

© 2020 Nilmani Singh

REDEFINING THE SPECIFICITY OF PHOSPHOINOSITIDE-BINDING BY HUMAN
PH DOMAIN-CONTAINING PROTEINS

BY

NILMANI SINGH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Cell and Developmental Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Professor Jie Chen, Chair and Director of Research
Professor Supriya G. Prasanth
Professor Brian C. Freeman
Associate Professor Rutilio A. Fratti

ABSTRACT

Lipids and proteins are two major components of cells, which are crucial for structural integrity and maintenance of cellular homeostasis by integrating various extracellular and intracellular cues. Lipids are involved in the regulation of cellular functions and can interact with proteins and undergo diverse chemical modifications. The presence of distinct lipid species creates a unique membrane signature for various cellular organelles as well as acts as a localization signal for multiple proteins. Proteins have evolved to contain specific lipid-binding modules crucial for mediating protein-lipid interaction. Numerous human pathologies have been linked to perturbations in lipid signaling and lipid-protein interactions, thus underscoring the need to study protein-lipid interactions and associated cellular processes. In this dissertation, I have examined the binding characteristics of Pleckstrin Homology (PH) domain, a lipid-binding domain, with phosphoinositides (phosphatidylinositol phosphates or PIPs), an important class of regulatory phospholipids.

PH domains are presumed to bind PIPs, but specific interaction with and regulation by PIPs for most PH domain-containing protein remains unclear. Recapitulating biologically relevant protein-lipid interactions *in vitro* is challenging due to a myriad of membrane properties such as curvature, composition, membrane asymmetry, viscosity, lipid rafts etc. The traditional methods to investigate these interactions rely on purified proteins, isolated lipid-binding domains, or non-physiological lipid environments. These common techniques include lipid strips, liposome sedimentation, isothermal titration calorimetry (ITC), and surface plasmon resonance (SPR); however, despite their wide use, they are known to be artifact-prone because of their reliance on purified proteins, isolated domains, and/or non-physiological lipid presentation. To overcome those shortcomings, our laboratory has developed a single-molecule pull down (SiMPull) assay

to probe protein-lipid binding. This method, called lipid-SiMPull, has high specificity and sensitivity while utilizing lipid vesicles of desired composition and full-length proteins expressed in whole cell lysates without the need for purification.

In Chapter 2, utilizing lipid-SiMPull, I have examined the PIP binding for 67 full-length human PH domain-containing proteins. Approximately half of them were found to have affinity for PIPs with various specificity, the majority of which had not been previously reported. For 8 proteins, PI(3,4,5)P₃ (PIP₃) binding affinity was further validated *in vivo* by insulin stimulated PIP₃ production on the plasma membrane and subsequent translocation. In Chapter 3, I describe a probabilistic sequence comparison algorithm, called recursive functional classification (RFC), created using the lipid-SiMPull data that identified the amino acid determinants of PIP binding throughout the entire sequence of the PH domain. The results from lipid-SiMPull experiments and RFC algorithm suggest that about half of the 246 human PH domains should bind PIPs with some specificity.

Guanine nucleotide-exchange factors (GEFs) catalyze the GTP-bound active state of small G proteins. The Rho family of small G proteins are activated by a large family of proteins containing Dbl-homology (DH) domain, called RhoGEFs, which are known for conserved DH-PH domain structures. While the DH domain is responsible for GEF activity, the function of the PH domain in RhoGEFs is not well-understood. In Chapter 4, I have investigated the role of phosphoinositide binding by ARHGEF3, a RhoGEF, in its biological functions. In the lipid-SiMPull screening, ARHGEF3 was found to specifically bind PI(4,5)P₂ and PI(3,5)P₂ in Lipid-SiMPull screening through its PH domain. Of the two known biological function of ARHGEF3, lipid-binding regulates the GEF activity towards RhoA and but not the inhibition of mTORC2 kinase. This adds to the growing understanding of the structure-function relationship for RhoGEFs, especially the poorly-understood role of the conserved PH domains.

AKT is a serine/threonine kinase crucial for regulation of diverse cellular processes and a major drug target for various human pathologies. Binding of the AKT PH domain with PIP_3 or $\text{PI}(3,4)\text{P}_2$ is responsible for its membrane translocation and activation. Different models have been proposed to explain how lipid binding modulates the activity of AKT in cells. While the classical view holds that AKT dissociates from the membrane after activation, recent studies have found that membrane dissociation results in rapid dephosphorylation and deactivation of AKT. A more detailed understanding of the lipid-binding and associated intra-molecular conformational changes is crucial to describe the activation model of AKT. In Chapter 5, I have combined lipid-SiMPull and mutagenesis to address questions regarding the role of lipid-binding in the mechanism of AKT activation. The results suggest that the hydrophobic motif domain of AKT may play an important role in inhibiting lipid-binding and that future investigation is warranted to elucidate the interactions of hydrophobic motif domain with other domains in AKT.

*To my parents, Shri Arvind Kumar and Shrimati Manju Kumari,
whose dreams and sacrifices have paved my ways*

To my brother Lokesh, for his trust and strength

*To my wife Aditi, for making my life incredibly wonderful
for her unconditional love and support*

ACKNOWLEDGEMENTS

I am extremely fortunate to have worked under the guidance of an exceptional mentor Dr. Jie Chen, who poured her heart and soul in ensuring my success in her laboratory and outside of it. Working in the lab granted me an opportunity to witness her scientific intellect, attention to details and curiosity that helped me learn and grow as a scientist. Despite numerous setbacks and hurdles, her constant encouragement and support kept me going and on the right track.

During my PhD, I had an incredible opportunity to learn from many professors, who graciously spent their time listening to my ideas and providing feedbacks. I would like to express my sincere thanks to my thesis committee members, Professor Supriya Prasanth, Professor Brian Freeman and Professor Rutilio Fratti for their critiques and insightful advices on multiple projects.

I would like to thank Dr. Rachel Waldemer for teaching me the myogenesis related work and Dr. Edwin Arauz for training me in SiMPull. The PH domain project would not have been possible without the generous contributions of many colleagues and especially Adriana Reyes Ordonez. Without her organizational skills, painstaking and never-ending SiMPull assays, and intellectual insights about the project, my PhD would have been incomplete. I convey my sincere gratitude for the help and kindness of Adriana Reyes. I am also thankful to Michael Compagnone, for his help with expediting the PH domain project. I am also extremely grateful to members of the Chen lab: Dr. Christina Rosenberg, Dr. Nidhi Khanna, Dr. Kook Son, Dr. Jaesung You, Dongwook Kim, Chong Dai, Jesus Francisco Moreno Castillo, and Allison Ann Boss-Kennedy, who helped create a sense of belonging and productive comradery in the Chen lab.

Lastly, I want to thank all my family and friends, who have shown me incredible affection throughout my life. I am deeply indebted and grateful to my parents for their innumerable sacrifices and unconditional love and support. This adventurous journey, which began in a remote village of India reading science fiction under kerosene lamps, would not have been possible without the hard work and dreams from my parents. I am especially thankful to my wife, Aditi. She has been a consistent beacon of love, joy, happiness and hope in my world. I am fortunate to have shared my PhD journey with her. I am also incredibly thankful to my kind and loving in-laws whose help has been critical to tide us through challenging times. I am incredibly grateful to many of my friends, who have listened to me and lent their support during tough moments in my PhD journey.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION.....	1
1.1 Phosphoinositides.....	1
1.2 Lipid Binding Domains.....	3
1.3 Pleckstrin Homology (PH) Domains.....	4
1.4 Methods of Studying Protein-Lipid Interactions.....	5
1.5 Regulation of RhoGEF Activity by Lipid Binding.....	11
1.6 Allosteric Regulation of AKT Activity by Lipid Binding.....	13
1.7 Tables and Figures.....	15
1.8 References.....	26
CHAPTER 2. LIPID-BINDING SPECIFICITY OF FULL-LENGTH HUMAN PH DOMAIN-CONTAINING PROTEINS.....	30
2.1 Introduction.....	30
2.2 Materials and Methods.....	32
2.3 Results.....	35
2.4 Discussion.....	40
2.5 Figures and Tables.....	43
2.6 References.....	55
CHAPTER 3. PREDICTING PHOSPHOINOSITIDE BINDING BY PH DOMAINS.....	59
3.1 Introduction.....	59
3.2 Material and Methods.....	60
3.3 Results.....	62
3.4 Discussion.....	65
3.5 Figures and Table.....	67
3.6 References.....	75
CHAPTER 4. REGULATION OF ARHGEF3 BY LIPID BINDING.....	76
4.1 Introduction.....	76
4.2 Materials and Methods.....	77
4.3 Results.....	80
4.4 Discussion.....	83

4.5 Figures and Table.....	85
4.6 References.....	93
CHAPTER 5. REGULATION OF AKT-PIP ₃ INTERACTION BY INTER-DOMAIN INTERACTIONS.....	95
5.1 Introduction.....	95
5.2 Materials and Methods.....	96
5.3 Results.....	97
5.4 Discussion.....	99
5.5 Figures.....	101
5.6 References.....	106
CHAPTER 6. CONCLUSIONS AND PERSPECTIVES.....	108
6.1 References.....	112

CHAPTER 1: INTRODUCTION

1.1 Phosphoinositides

Lipids and proteins are two of the major building blocks of life. There are thousands of lipid species found in cells, which can be subdivided in two broad groups: a) bulk or structural lipids such as phosphatidylcholine (PC) and phosphatidylethanolamine (PE) and b) signaling lipids such as diacylglycerol (DAG) and phosphoinositides. Various cellular organelles have distinct steady-state lipid composition maintained by active transport of lipid species between organelle membranes, while endoplasmic reticulum (ER) acts as primary lipid synthesis hub. The functional paradigm of lipids can be categorized under three broad groups; as energy storage molecules, as means of cellular compartmentalization, and as signaling molecules actively regulating diverse cellular functions (1). Protein-lipid interactions are critical to the regulation of numerous signal transduction processes and cellular functions. Several human pathologies such as cancer, diabetes, and neurological disorders have been traced back to the dysregulation of lipid signaling or perturbations in lipid-protein interactions.

Phosphoinositides constitute a relatively minor fraction of cellular lipids, and yet play an outsized essential role in regulation of multiple cellular processes such as signaling, trafficking, cytoskeletal dynamics, migration and proliferation. Seven different phosphoinositide species can be generated through reversible phosphorylation at 3-, 4-, and 5- position of inositol ring of phosphoinositide (PI) (Figure 1.1). Concerted regulation by multiple kinases and phosphatases spatiotemporally maintains phosphoinositide pools in response to varied internal or external cues. Local phosphoinositide pools in cells can be drastically elevated or depleted by various cellular enzyme. In response to mitogenic

stimuli, PIP₃ and PI(3,4)P₂ are produced on plasma membrane by class I phosphoinositide-3 kinase (PI3K) followed by membrane recruitment and activation of AKT by lipid binding. Similarly, thrombin can stimulate PI(5)P production in platelets (2). While a complete map of organelle distribution of PIPs is not well-known due to their uneven distribution and lack of sensitive probes, total PIPs account for less than 1% of cellular phospholipids. PI(4)P and PI(4,5)P₂ are two most abundant PIPs in cells and are primarily enriched at the plasma membrane (3, 4). PI(4)P acts as the primary source for PI(4,5)P₂ synthesis on the plasma membrane. PI, the primary precursor of PI(4)P in cells, is synthesized at the endoplasmic reticulum (ER), but the organelle distribution and trafficking of PI is not well understood. Two recent studies failed to identify significant steady-state pools of PI at plasma membrane or ER. Rather, it was found to be enriched in Golgi complex, mitochondria, and peroxisomes (3, 5).

Cellular organelles are uniquely enriched in different PIPs (Figure 1.2), which helps recruit PIP specific proteins to different cellular membranes production at plasma membrane is crucial for AKT activation, which subsequently regulates diverse cellular processes downstream. Similarly, early endosomes and autophagosomes are characteristically enriched with PI(3)P, which plays a key functional role in these organelles (6). EEA1, a FYVE domain containing PI(3)P binding protein, is required for endosomal tethering and membrane fusion (7). Another mono-phosphorylated PIP, PI(4)P, is abundant in ER and *trans*-Golgi network. Numerous PI(4)P interacting proteins are recruited to these membranes and are essential for bud formation, vesicle extrusion, and vesicular trafficking. Recently a PI(4)P binding protein called, GOLPH3, was identified to be localized at *trans*-Golgi and budding vesicles. It forms a multi-protein complex with unconventional myosin, MYO18A, to promote vesicle budding by applying a tensile force to *trans*-Golgi (8).

Another important role of PIPs is generation of lipid second messengers. Hydrolysis of $\text{PI}(4,5)\text{P}_2$ by PLC can generate Inositol triphosphate (IP_3), and diacylglycerol (DAG), which can be subsequently acted upon by diacylglycerol kinases to produce phosphatidic acid (PA). All of the above mentioned second messengers have independent functional roles in regulation of multiple cellular process. Numerous studies have underlined the crucial role of protein-lipid interaction in maintaining cellular homeostasis. Further investigations will identify novel protein-lipid interactions and illuminate their role in various cellular functions.

1.2 Lipid Binding Domains

Proteins are modular structures made of one or many diverse units with specialized structural and functional objectives. Many proteins have evolved to contain specific lipid-binding domains (LBDs) that can bind their particular lipid moieties. Since lipids are found in organelle membranes, LBDs act as membrane tethering modules for proteins. Multiple LBDs have been identified including C1, C2, PH, PX, FYVE, BAR, ENTH, ANTH, and PROPPIN domain (9) (Table 1.1). These LBDs can be broadly characterized into two groups; domains with stereospecific recognition of particular lipid headgroups and domains that recognize physical properties of membranes such as curvature or charge. C1 domain of protein kinase C (PKC) was the first domain identified in 1980s to selectively bind lipid second messenger diacylglycerol (DAG) and phorbol esters (10, 11). Numerous studies over the last three decades have identified more selective lipid-binding domains such as PH, PX, FYVE, and PROPPIN. C2 domain was the second domain identified in PKC as a Ca^{2+} -dependent lipid binding domain. Majority of C2 domains are Ca^{2+} -dependent non-specific phospholipid binding domain preferring negatively charged lipids such as phosphatidylserine. Similarly, three domains ENTH,

ANTH, and BAR have little, if any, lipid-binding specificity but recognize membrane curvature as well as remodel membrane topology after binding (9).

Not only PH domains make the largest lipid-binding domain in many species (4), but there are also many structural homologs of PH domain, collectively termed as PH-like domains, which includes PTB, PDZ, GRAM, GLUE and FERM domain (9). Specific phosphoinositide binding sites have been identified in some PTB domains such as Dab1, IRS-1, and Shc1; however, the residues forming the PIP binding sites are not conserved across the PTB domains in the majority of proteins (12). Similarly, most of the other PH-like domains are primarily involved in protein-protein interaction, with only a minority known for lipid binding (9).

1.3 Pleckstrin Homology (PH) Domains

The largest family of putative lipid-binding domains is the PH domain, with over 250 members encoded by the human genome (13). PH domains are known for binding phosphoinositides and, according to the SMART database, they form the 11th most common domain in the human genome (14). Originally defined by its presence in the protein pleckstrin (15, 16), this domain of 100-120 amino acids has an invariable structure of seven-stranded β -barrel lined by a C-terminal α -helix (17). Some of the earliest characterized PH domains were found to have affinity for $\text{PI}(4,5)\text{P}_2$, such as those in pleckstrin, RasGAP, and GRK2 (18). Several amino acids in the variable loops of $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ strands and flanking regions can interact directly with the headgroup of PIPs (19). The X-ray crystallographic studies of the PH domain within PLC δ 1 showed it bound to the headgroup of $\text{PI}(4,5)\text{P}_2$ and provided the first model of interaction between amino acids and the phosphate groups of phosphoinositides (20). Direct and indirect hydrogen bonds

were identified between the amino acids within the $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ strands and flanking regions with 4- and 5-phosphate group of $\text{Ins}(1,4,5)\text{P}_3$ (Figure 1.3). Mutational analysis of the $\text{PLC}\delta 1$ PH domain identified many basic amino acids in the aforementioned region, which, when mutated to neutral or opposite charge, abolished $\text{PI}(4,5)\text{P}_2$ binding (21).

Many PH domains have been reported to bind with high selectivity to specific PIPs such as $\text{PLC}\delta 1$ to $\text{PI}(4,5)\text{P}_2$, or AKT1, BTK and GRP1 to PIP_3 (22-25). Indeed, this interaction is so specific that the PH domains of proteins such as AKT1, $\text{PLC}\delta 1$ have been commonly used as $\text{PI}(4,5)\text{P}_2$ and PIP_3 sensors in cells, respectively (26, 27). Despite these early examples of PH-PIP interactions, to date only a modest number of PH domain-containing proteins have been demonstrated to bind PIPs with specificity. A genome-wide analysis of the PH domains in *Saccharomyces cerevisiae* found that non-specific binding is common amongst PH domains, with only one PH domain, Num1p, displaying both high-affinity and high-specificity binding to $\text{PI}(4,5)\text{P}_2$ (28). Another study examined 130 mouse PH domains and found 20% of them to translocate to the plasma membrane in response to PIP_3 production in cells; however, most PIP_3 -responsive PH domains did not show specific binding to PIP_3 in lipid strips assay (29). Thus, it has been proposed that the specific phosphoinositide recognition may be the exception rather than the rule for the PH domains and that only 10% of all PH domains bind PIPs with high affinity and specificity (4).

1.4 Methods of Studying Protein-Lipid Interactions

1.4.1 Biochemical Methods

Many biochemical methods have been employed to study protein-lipid interaction including lipid strips, liposome sedimentation, surface plasmon resonance (SPR), isothermal calorimetry (ITC) etc. The Lipids Strips assay (also known as the dot blot

assay) is one of the most commonly used assays to identify the PIP binding partners for proteins. PIPs spotted on a cellulose membrane are incubated with purified proteins followed by primary and secondary antibodies to detect binding (30). Despite being a straightforward procedure, this method is hard to quantitate and often results in non-specific or poor binding (30, 31). Lipids spotted on a 2-D surface lose their 3-D membrane orientation as observed in biological membranes and thus may fail to capture biologically relevant interactions. A study of 130 mouse PH domains observed no correlation between PIP binding by lipid strips and *in vivo* localization for majority of the PH domains (29). Other biochemical methods such as liposome sedimentation, ITC, and SPR require relatively large amount of proteins and have low sensitivity for binding (31). In liposome centrifugation, the detection of binding occurs after the disruption of equilibrium and thus it may miss transient binding. Most of these methods require protein purification, which is labor intensive with no guarantee of success. Due to inherent drawbacks and hurdles of purifying large proteins, purification of the isolated PH domains are preferred over full-length proteins (29, 32, 33). Purified PH domains not only lack the post-translational modifications but also the intramolecular interactions of the full-length protein. Due to inherent practical limitations of most of these biochemical methods, further *in vivo* characterization and validation of results remain essential.

1.4.2 *In vivo* Methods

Many different approaches such as genetic, pharmacological, chemical-genetic and optogenetic are utilized to manipulate and study *in vivo* protein-lipid interaction. Lipid specific genetically-encoded fluorescent reporters are a versatile toolkit to probe the phosphoinositide dynamics in cell. Numerous *in vivo* phosphoinositide specific reporters have been designed based on the high-affinity and selective interaction between lipid-binding domains and corresponding PIPs (Table 1.2).

The first such probe employed to visualize PI(4,5)P₂ hydrolysis *in vivo* after PLC activation was the EGFP-tagged PH domain of PLCδ1 (27, 34). The repertoire of phosphoinositide-binding proteins has grown steadily over the last two decades, resulting in multiple *in vivo* lipid sensors (35, 36). Many such lipid biosensors are fluorescently tagged lipid-binding domains, such as PH, FYVE, and PX domain of different proteins. PIP₃ dynamics during different cellular processes such as chemotaxis, polarization, phagocytosis and more, has been probed using the PH domains of AKT1 or BTK (37). While the PH domain of AKT1 binds both PIP₃ and PI(3,4)P₂, the C-terminal PH domain of TAPP1 has been used as a PI(3,4)P₂ specific biosensor. A recent study used tandem trimer of the TAPP1 PH domain, which has high specificity and sensitivity for PI(3,4)P₂, to detect changes in cellular PI(3,4)P₂ not accessible to earlier probes (38). Using this improved ability to detect PI(3,4)P₂, the authors identified class I PI3K as the major contributor of PI(3,4)P₂ production on plasma membrane. Similarly, the PH domains of OSBP or FAPP1 have been traditionally used to visualize PI(4)P dynamics in cells. A new probe with improved specificity and avidity for PI(4)P was derived from P4M domain of SidM from bacterial pathogen *Legionella pneumophila* and identified new pools of PI(4)P in cellular membranes (39). PI(3)P is another phosphoinositide, whose *in vivo* dynamics has been studied using the FYVE domain of Hrs or EEA1 protein.

Many proposed lipid biosensors are prone to artifacts due to insufficient understanding of *in vivo* lipid binding characteristics. Recently, Hammond & Balla (40) proposed three criteria for comprehensive characterization of lipid biosensors; a) specificity for the targeted lipid, b) cellular localization of biosensor should be driven by the lipid, and c) the lipid alone should be sufficient for cellular localization. There are few biosensors satisfying all three criteria, thus it is crucial to acknowledge the limitations while interpreting the data from lipid biosensors. Exploration of protein-lipid binding with

biologically relevant methods will identify novel biosensors as well as characterize available ones.

Pharmacological methods can target different signaling pathways and enzymes involved in regulation of phosphoinositide levels in cells. Class I PI3K inhibitors such as wortmannin or LY294002 are frequently used to inhibit PIP₃ production on plasma membrane in response to mitogenic stimuli. Similarly, many different inhibitors have been described for different phosphoinositide kinases such as PI 4-kinase, PI(3)P 5-kinase and phosphatases such as PI 5-phosphatase (41). Many of these inhibitors have been involved in clinical trials as promising candidates for cancer treatment due to the roles played by PIPs in various cancer associated cellular functions.

Rapamycin inducible FRB (mTOR fragment that binds FKBP12) - FKBP12 (FK506 binding proteins 12) dimerization system is a routinely used chemical-genetic tool to perturb local concentration of PIPs (42, 43). This approach provides exquisite temporal and spatial control over manipulating phosphoinositide levels *in vivo*. In one such scheme (42), both FRB and FKBP12 constructs were tagged with a fluorescent probe and the FRB domain was anchored to the plasma membrane by a palmitoylation sequence. The FKBP12 fragment was tagged with catalytic fragment of type IV 5-phosphatase, which gets recruited to plasma membrane after rapamycin addition (Figure 1.4). The depletion of plasma membrane PI(4,5)P₂ can be monitored by live cell microscopy using EGFP-tagged PLCδ1 PH domain. A further modification of this approach fused two phosphatases to FKBP12 fragment and depleted plasma membrane PI(4,5)P₂ to PI(4)P, followed by further depletion of PI(4)P to PI (44). This FRB-FKBP dimerization system has also been utilized for the controlled induction of phosphoinositide synthesis. The role of PI(4,5)P₂ or PIP₃ in regulating the KCNQ ion channel was probed by tagging the FKBP12 construct with PI(4)P 5-kinase type I or iSH2 domain of p85, to induce plasma membrane synthesis of PI(4,5)P₂ or PIP₃, respectively (45). Similar approaches have been deployed to precisely

control phosphoinositide levels *in vivo* through optogenetic tools. Over the last decade, many different tools have been designed for spatiotemporal manipulation of different PIPs within the cell that has yielded critical insights into their roles in various cellular processes.

1.4.3 Lipid-SiMPull

Although methods investigating ensemble properties of biological systems are routinely used, recent developments in single-molecule fluorescence microscopy has made it possible to inspect the heterogeneity of molecular behaviors in real-time. One such method, Single-Molecule Pull-down (SiMPull), probes protein-protein interaction at the single-molecule level and was developed by Prof. T J Ha Laboratory (46). SiMPull uses a total internal reflection fluorescence (TIRF) microscopy based single-molecule visualization method to analyze subunits of a protein complex. The imaging surface is passivated with polyethylene glycol (PEG) to block non-specific interactions and simultaneously doped with biotinylated PEG to capture molecules of interest through biotin-streptavidin interaction. Immobilized biotinylated antibodies pull down the protein of interest and its interaction partners. SiMPull can identify the heterogeneity of molecular interactions in protein complexes, which are not revealed by ensemble measurements. Additionally, the stoichiometry of protein-complexes can be analyzed by fluorescence photobleaching analysis. Previous work employing SiMPull from our laboratory observed that mTORC2 exists as a homodimer in cells and dimerization is induced by high local concentration of multiple subunits (47).

Based on aforementioned SiMPull, a TIRF microscopy based single-molecule technique was developed by our laboratory in collaboration with Prof. T J Ha laboratory to probe lipid-protein interaction (48) (Figure 1.5). This method called, lipid-SiMPull, can detect specific protein-lipid interactions and determine kinetic and assembly properties.

The protein-lipid interaction is probed between small unilamellar vesicles (SUVs) tagged with lipophilic fluorescent dye and fluorescently tagged proteins in whole cell lysates. SUVs mimic the membrane orientation of lipids and thus provide a marked improvement over lipids spotted on 2-D surfaces. Compared to traditional methods of probing protein-lipid interactions, as discussed in chapter 1.4.1, lipid-SiMPull requires small amounts of samples and circumvents the need for protein purification thus preserving post-translational modifications. Since, there is no need for protein purification, full-length proteins can be directly assayed for lipid binding rather than using the purified lipid binding domain of the protein. The sensitivity and specificity of this assay has been demonstrated through highly selective pulldown of various lipid-binding domains such as the PH domain of AKT1 and PLC δ 1, tandem dimer of the FYVE domain of Hrs (48). Different LBDs bound selectively with SUVs containing specific PIPs such as the PH domain of AKT1 to PIP₃ and PLC δ 1 to PI(4,5)P₂.

Fluorescent time traces of binding can be analyzed to determine different assembly properties such as copy number, transient versus stable binding, and kinetic parameters such as dissociation constant. The AKT1 PH domain bound transiently with PIP₃, whereas the full-length AKT1 bound stably. Similarly, multiple copies of AKT1 PH domain bound PIP₃ vesicles compared to predominant monomeric binding of the full-length protein. Building on the evident merits of Lipid-SiMPull over traditional lipid-protein assay methods, in Chapter 2, I have investigated the phosphoinositide-binding specificity of 67 full-length PH domain proteins with all PIPs except PI(3,5)P₂. This screening has identified many novel PIP binding properties and serves as a call for a broad re-examination of lipid-binding specificities of the whole PH domain family.

1.5 Regulation of RhoGEF Activity by Lipid Binding

1.5.1 Rho Guanine Exchange Factors (RhoGEFs)

Small G proteins act as binary molecular switches integrating stimuli from various cell surface receptors to regulate diverse cellular processes such as cytoskeletal rearrangement, adhesion, migration, growth, differentiation, and vesicular transport. The family of small GTPases share a common structural fold of six β -sheets enclosed by five α -helices and are governed by a shared mechanism of switching between either GDP- or GTP-bound states. Two different group of proteins, guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs), catalyze the transition between the GDP- and GTP-bound states of small GTPases. The GEFs activate the small GTPases by accelerating the exchange of GTP for GDP, whereas GAPs inactivate the proteins by promoting GTP hydrolysis (Figure 1.6).

One of the major subfamilies of small GTPases is the Rho GTPase family with ~20 members identified in mammals to date (49). Rho GTPases regulate diverse cellular processes such as cytoskeleton rearrangement, cell-cycle, growth, and polarity. The first characterized GEF for Rho GTPases was a ~200 amino acid DH (Dbl homology) domain containing protein for Cdc42 . Thus far, ~70 RhoGEFs have been identified in human genome with a shared invariable domain structure of a ~100 amino acid PH domain adjacent to the DH domain (49). While the DH domain is required for its GEF activity in GTP loading on Rho GTPases, the role of the adjacent PH domain is not well-defined in the GEF activity. PH domains are commonly known for phosphoinositide binding, but in context of RhoGEFs, the contribution of lipid-binding towards their GEF activity for majority of RhoGEFs is not well understood.

1.5.2 Regulation of GEF Activity by Protein-Lipid Interactions

The conserved invariable arrangement of DH-PH domains suggests that the PH domain may play an important regulatory role in biological functions of RhoGEFs. Over the years, studies of a few RhoGEFs have identified multiple regulatory mechanisms through their PH domain. The PH domains in RhoGEFs have been proposed to act in multiple regulatory roles such as a membrane targeting module, or an allosteric regulator of GEF activity, or as a modulator for efficient activation through direct interaction with small GTPases (Figure 1.7).

The PH domain of a mouse RhoGEF, Lfc, was identified to be crucial for its transformation ability. PH domain deletion in Lfc abolished its transformation capacity, which was restored by a plasma membrane targeting site (50). Conversely, mutations in the PH domain of Dbs, a RhoGEF for Cdc42 and RhoA, which impaired lipid-binding, abrogated transformation of NIH3T3 cells without affecting cellular localization (51). It was proposed that lipid binding by PH domain reorients the DH and PH domain of Dbs and relieves inhibitory constraints for catalytic activity. Another RhoGEF, P-Rex-1, a Rac1 specific RhoGEF involved in neutrophil migration and metastasis, was reported to have phosphoinositide-regulated GEF activity. Structural and functional studies of P-Rex1 supports an allosteric mechanism of regulation of GEF activity by PIP₃ binding. PIP₃ binding of different PH domain mutants was strongly correlated with catalytic activity of P-Rex1, although impaired PIP₃ binding did not affect membrane localization (52). A similar effect is observed in Tiam-1, a Rac1 specific RhoGEF, where abolishing lipid-binding through mutations in PH domain did not affect *in vivo* localization but did decrease its GEF activity, suggesting yet another instance of allosteric regulation of GEF activity through its PH domain (53). The mutations in the Lbc family of RhoGEFs, which abolished the

interactions between the PH domain and active RhoA, decreased stimulated RhoA in cells suggesting a positive feedback mechanism to regulate GEF activity (54).

Notwithstanding the studies of some of the RhoGEFs, no general mechanism of regulation by the PH domain has been proposed and validated. Despite the conserved DH-PH structure, it appears that the PH domains do not share similar functional roles in different RhoGEFs and thus studying more RhoGEFs will yield a better insight into their structure and function. In Chapter 3, I will describe the role of lipid-binding towards regulating the GEF activity of ARHGEF3, a RhoGEF with classic DH-PH structure.

1.6 Allosteric Regulation of AKT Activity by Lipid Binding

Protein kinase B (PKB) / AKT is a serine/threonine kinase with a wide array of essential functions in cell proliferation, survival, differentiation, and metabolism. Growth factor stimulation of PI3K results in generation of PIP_3 and $\text{PI}(3,4)\text{P}_2$ at the plasma membrane, which recruits AKT through its PH domain. According to the classical view, AKT is phosphorylated and activated on plasma membrane followed by subsequent translocation and diffusion to cytoplasm or nucleus in order to phosphorylate its targets (55, 56) (Figure 1.8) Recent studies have proposed a contradicting model where AKT requires continued membrane association to maintain its active state (57, 58) (Figure 1.9)

In Chapter 4, I have employed the lipid-SiMPull assay to study lipid binding of AKT and understand its functional implication. Quantitative and kinetic measurements were made with wild-type and phosphor-mimetic or unphosphorylatable mutants of AKT in whole cell lysates for their binding to PIP_3 vesicles. Indeed, phosphorylation of AKT drastically decreased its affinity for the lipids. The potential role of phosphatidylserine (PS) in AKT membrane recruitment was also examined. Furthermore, distinct binding

behaviors were observed with full-length AKT and the PH domain alone, suggesting that other domains in AKT modulate its lipid interaction and thus membrane association. The deletion of hydrophobic motif increased lipid-binding compared to the wild-type. This suggests that hydrophobic motif may be involved in crucial intra-molecular interaction regulating AKT activity, as opposed to the generally accepted view that hydrophobic motif is disordered in inactive AKT. These results provide new insights into the molecular mechanism of AKT activation, and also highlight the unique advantages of the lipid-SiMPull method in probing protein-lipid interactions with full-length proteins in a near-physiological context.

1.7 Tables and Figures

Table 1.1: Common lipid-binding domains and their properties

Domain Name	Domain Size (amino acid)	3-D structure	Binding Specificity
ANTH	~280	α -helix solenoid	PIPs with low specificity
BAR	~250	α -helical coiled coil bundle dimer	acidic phospholipids, membrane curvature dependent
C1	~50	Zn ⁺² Finger	DAG, phorbol esters
C2	~130	8-stranded antiparallel β -sandwich	Non-specific Ca ²⁺ dependent binding to anionic lipids
ENTH	~150	α -helix solenoid	PI(4,5)P ₂
FYVE	~70-80	Zn ⁺² Finger	PI(3)P
PH	~100-120	β -barrel lined by α -helix on c-terminal	weak binding with limited specificity with PIPs, some highly affinity and specificity
PROPPIN	~500	7-blade β -propeller	PI(3,5)P ₂ and PI(3)P
PX	~130	3 β -sheets + 4 α -helices	most are PI(3)P specific

Table 1.2: Commonly used genetically encoded PIP biosensors

Phosphoinositide	Lipid Biosensor	Cellular Localization	References
PI(3,4,5,)P₃	PH-AKT	Plasma membrane	(26)
	PH-BTK	Plasma membrane	(59)
	PH-ARNO	Plasma membrane	(60)
PI(3,4)P₂	PH-TAPP1	Plasma membrane	(38, 61)
PI(4,5)P₂	PH-PLCδ1	Plasma membrane	(27)
PI(3)P	FYVE-Hrsx2	Early endosomes	(62, 63)
	FYVE-EEA1	Early endosomes	(64)
	PX-p40phox	Early endosomes	(65)
PI(4)P	FAPP1-PH	Golgi, Plasma membrane	(66, 67)
	OSBP-PH	Golgi, Plasma membrane	(66)
	P4M-SidMx2	Golgi, Plasma membrane, endosomes	(39)

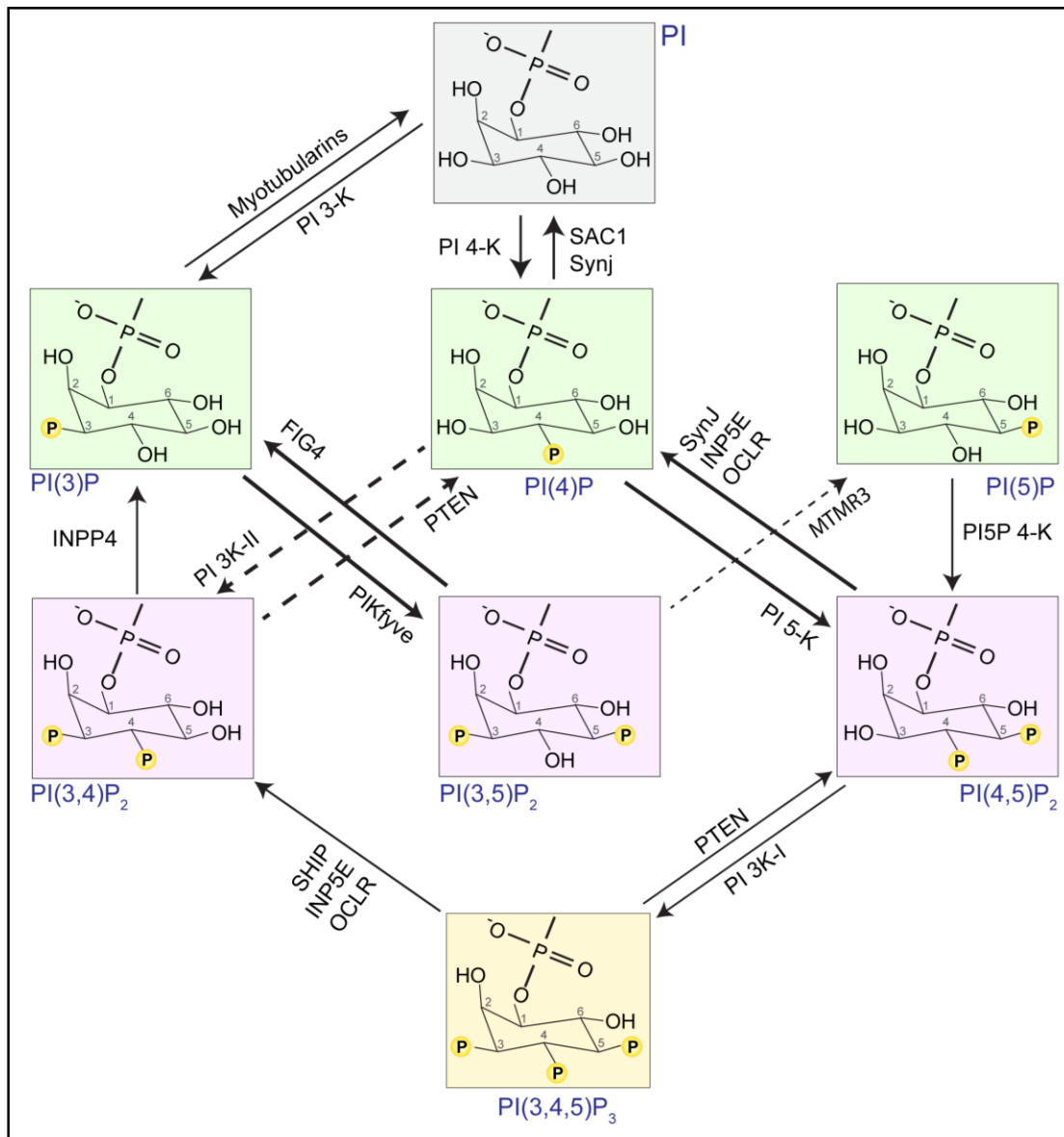


Figure 1.1 The phosphoinositide cycle in cells

Different lipid kinases and phosphatases can generate seven PIP species by attachment of phosphate group at 3-, 4- and 5- positions from the precursor, Phosphatidylinositol (PI). Previously reported kinases and phosphatases involved in PIP cycles are shown in figure.

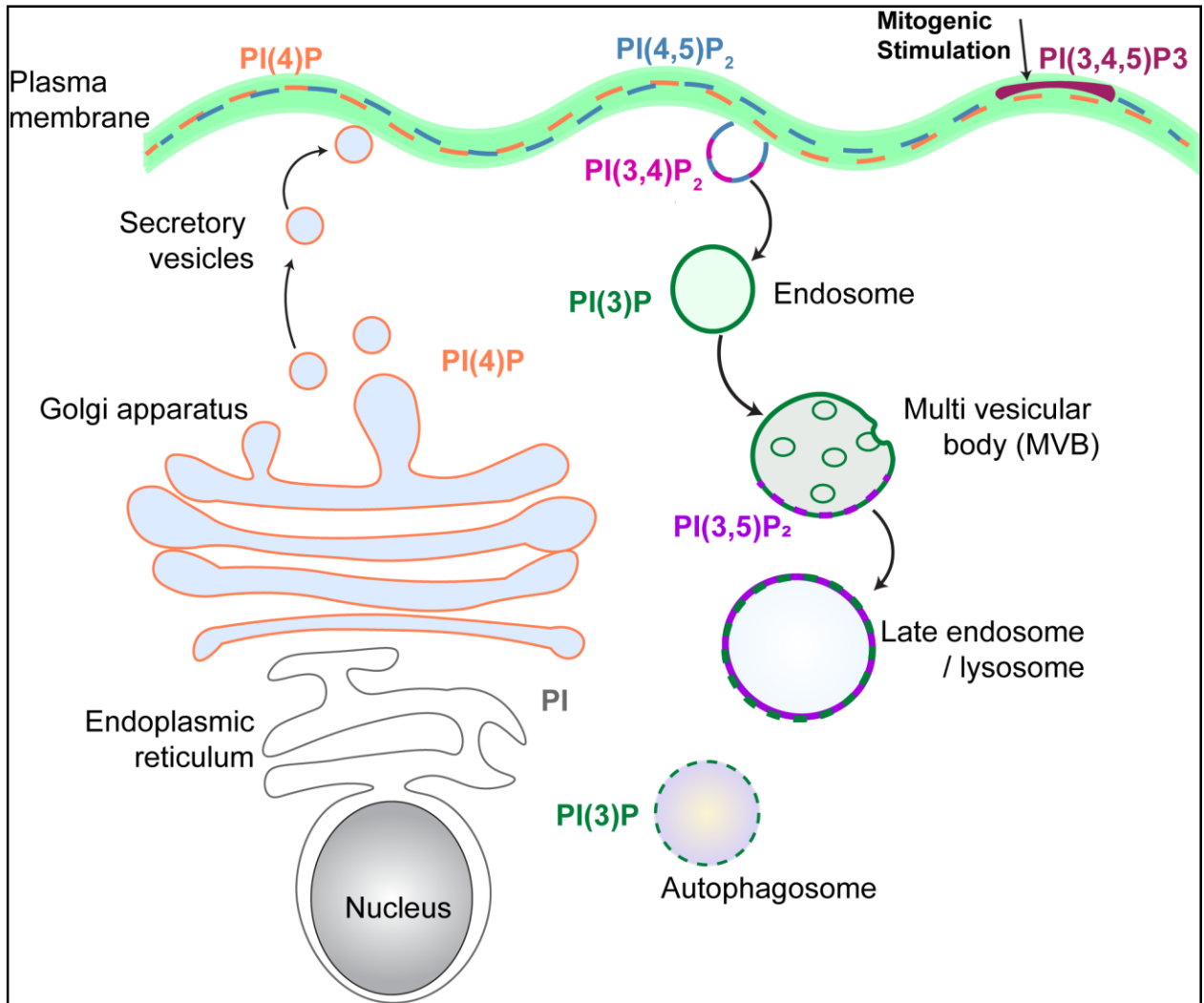


Figure 1.2 Intracellular localization of phosphoinositides

Different organelle membranes are uniquely enriched in specific PIPs, as shown in the figure above.

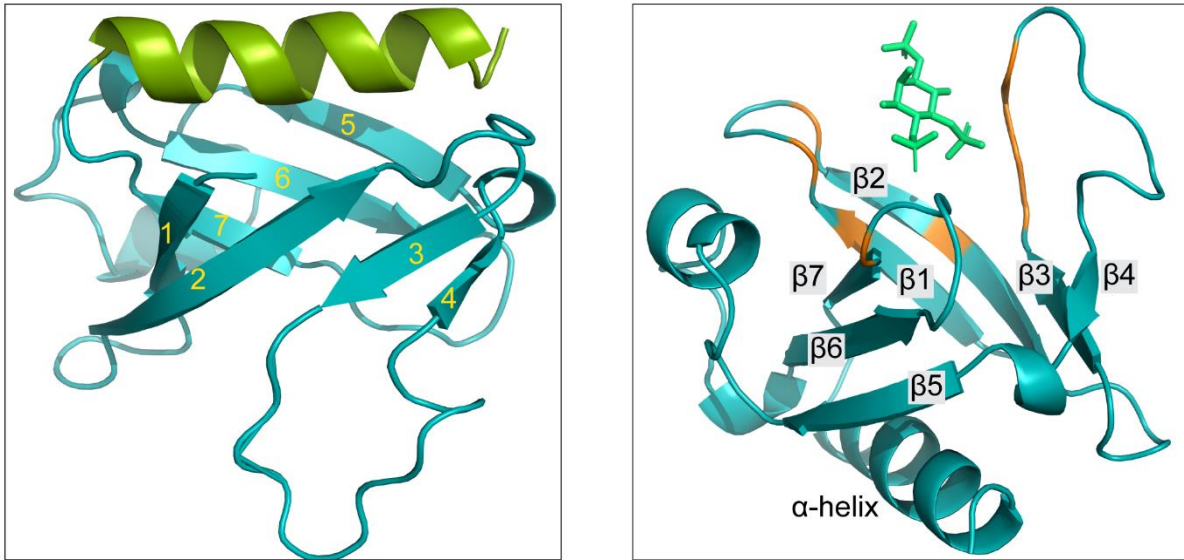


Figure 1.3 The crystal structure of PH domain of PLC δ 1 (PDB: 1MAI)

A) The β -barrel structure of the PH domain of PLC δ 1 with 7 β sheets (teal) c-terminal α -helix (lime green) **B)** a different orientation showing the amino acids located in the β 1- β 2 and β 3 - β 4 variable loop region (orange), which make direct hydrogen bonds with the headgroup of PI(4,5)P₂ (green).

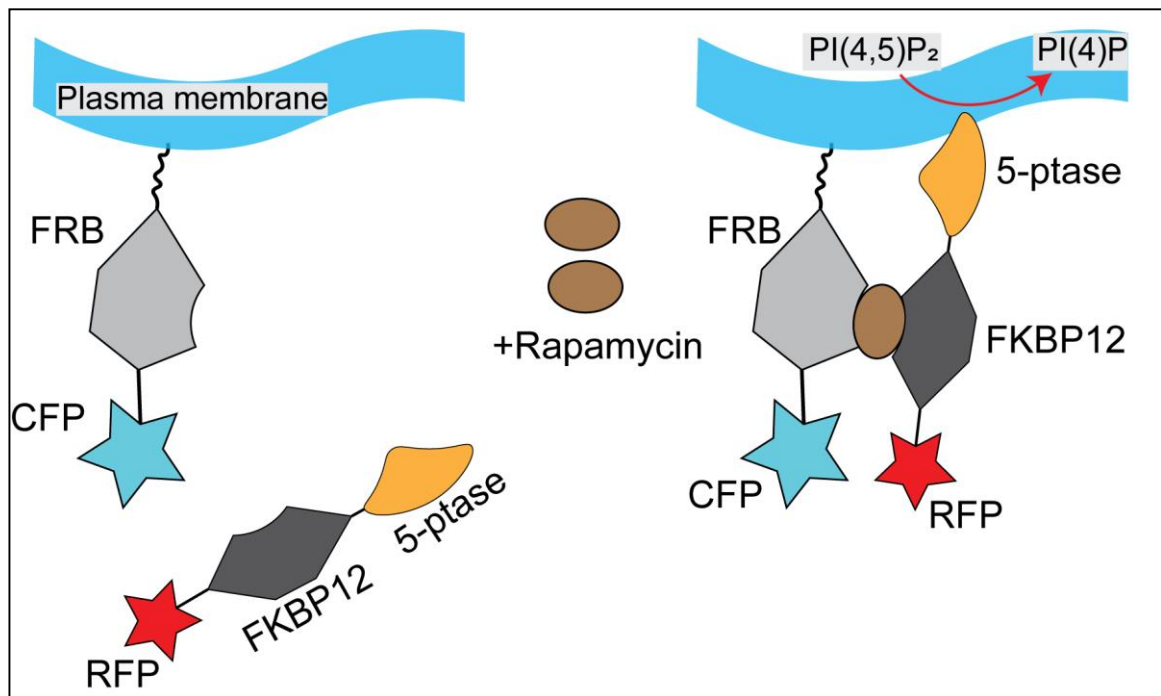


Figure 1.4 Schematics of FRB-FKBP12 dimerization system for depleting plasma membrane PI(4,5)P₂

FRB is tethered on the plasma membranes by a palmitoylation sequence and the FKBP12 fragment is present in cytoplasm containing the catalytic fragment of phosphoinositide 5-phosphatase. After rapamycin addition, FKBP12 binds FRB and gets recruited to plasma membrane with the 5-phosphatase, which subsequently dephosphorylates PI(4,5)P₂. Both FRB and FKBP12 fragments are labeled with a fluorescent tag.

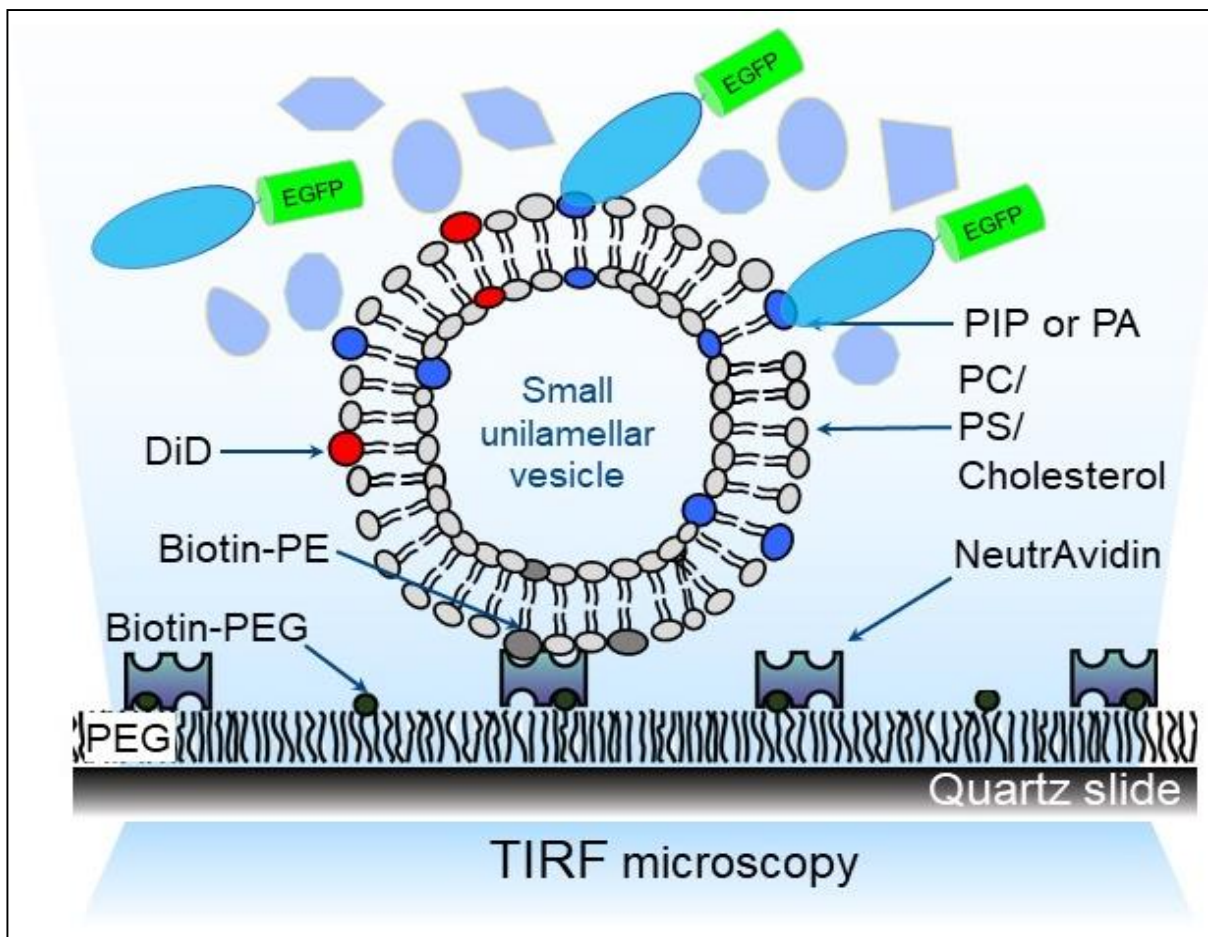


Figure 1.5 Schematic representation of the lipid-SiMPull assay

Quartz slides are passivated with PEG doped with biotinylated-PEG. Small unilamellar vesicles with desired composition containing a lipophilic dye are immobilized on surface through biotin-neutravidin interaction. Cells expressing fluorescently labeled proteins are lysed and lysates are flown through the imaging chamber. After a brief incubation, the binding with lipid-vesicles can be visualized with TIRF microscopy.

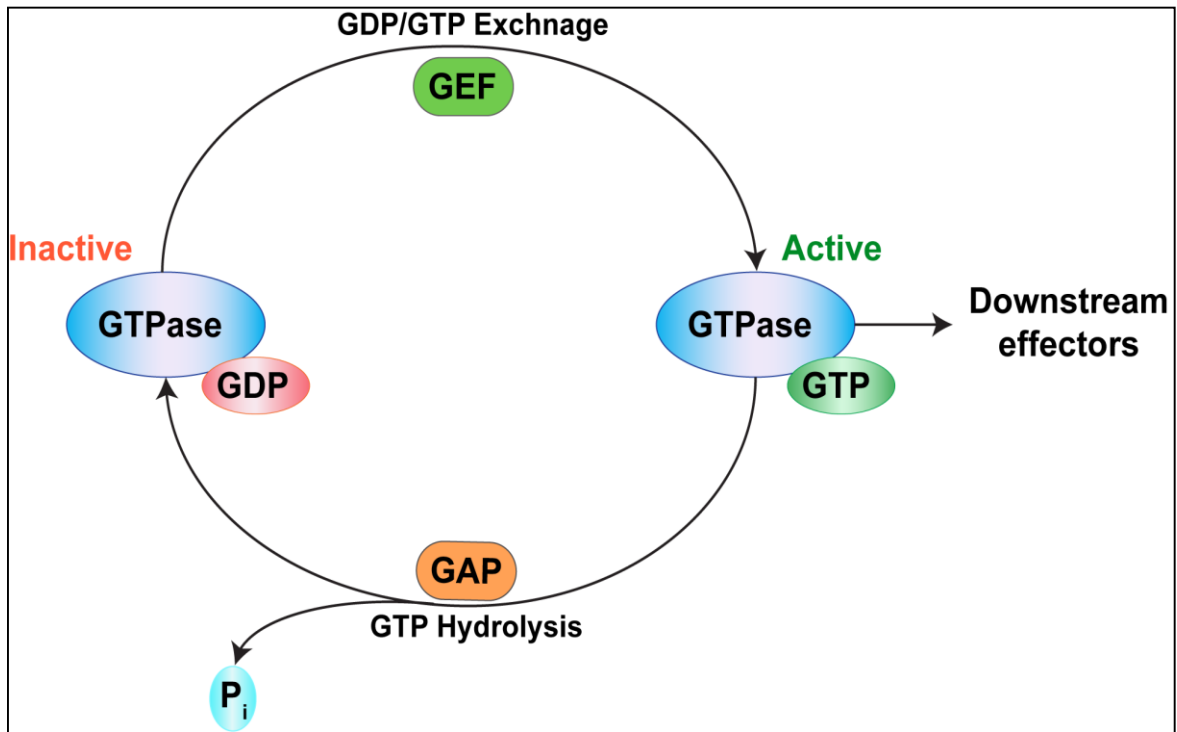


Figure 1.6 The GTPase cycle

GEFs activate small G proteins by catalyzing GTP loading of GTPases. GAPs promote the hydrolysis of GTP and thus deactivate the GTPases.

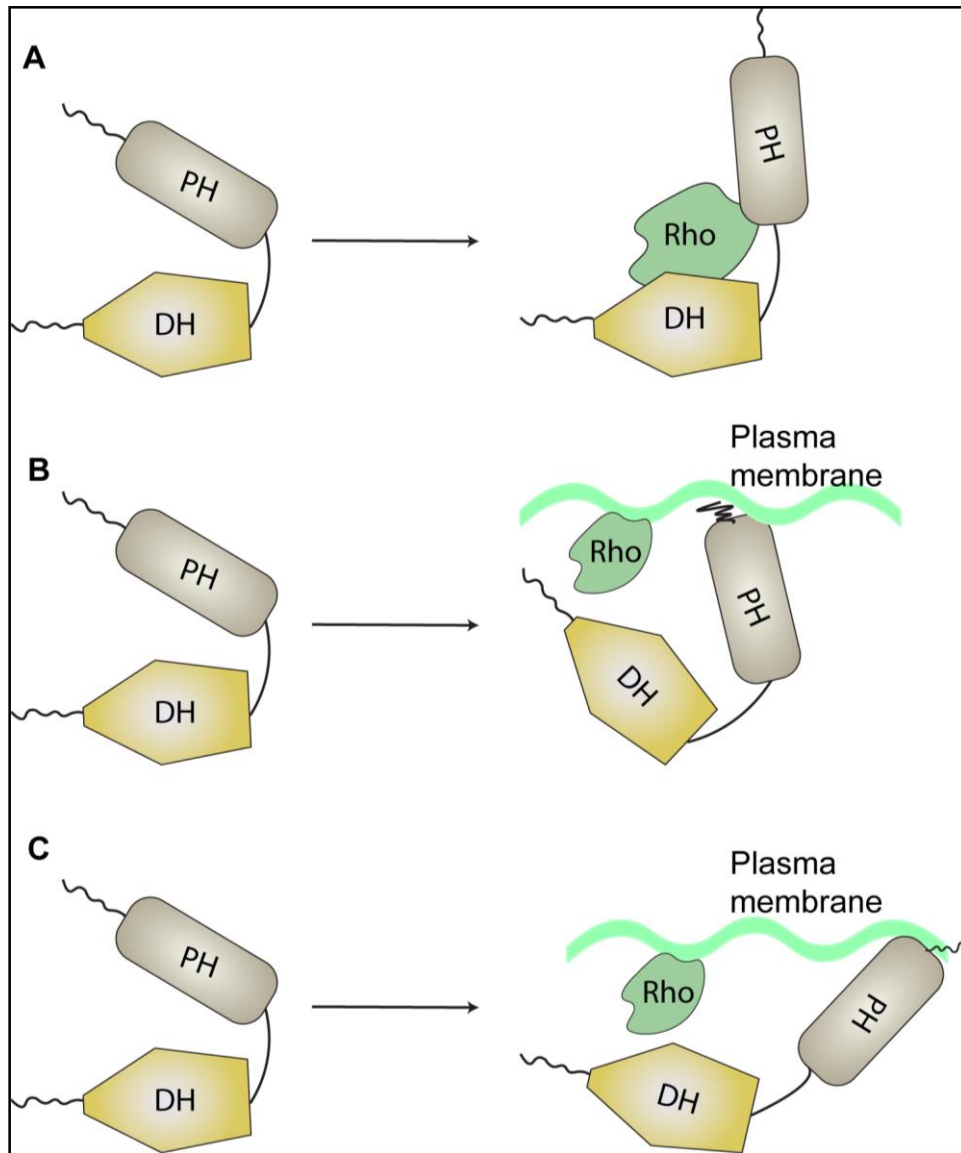


Figure 1.7 Diverse modes of GEF activity regulation in RhoGEFs by the PH domain

A) The PH domain can directly interact with the small G protein e.g. in Dbs-Cdc42 or LARG-RhoA, **B)** PH domain can act as the membrane targeting module for the RhoGEFs e.g. Lfc and, **C)** allosteric regulation of GEF activity by lipid binding of the PH domain e.g. P-Rex1.

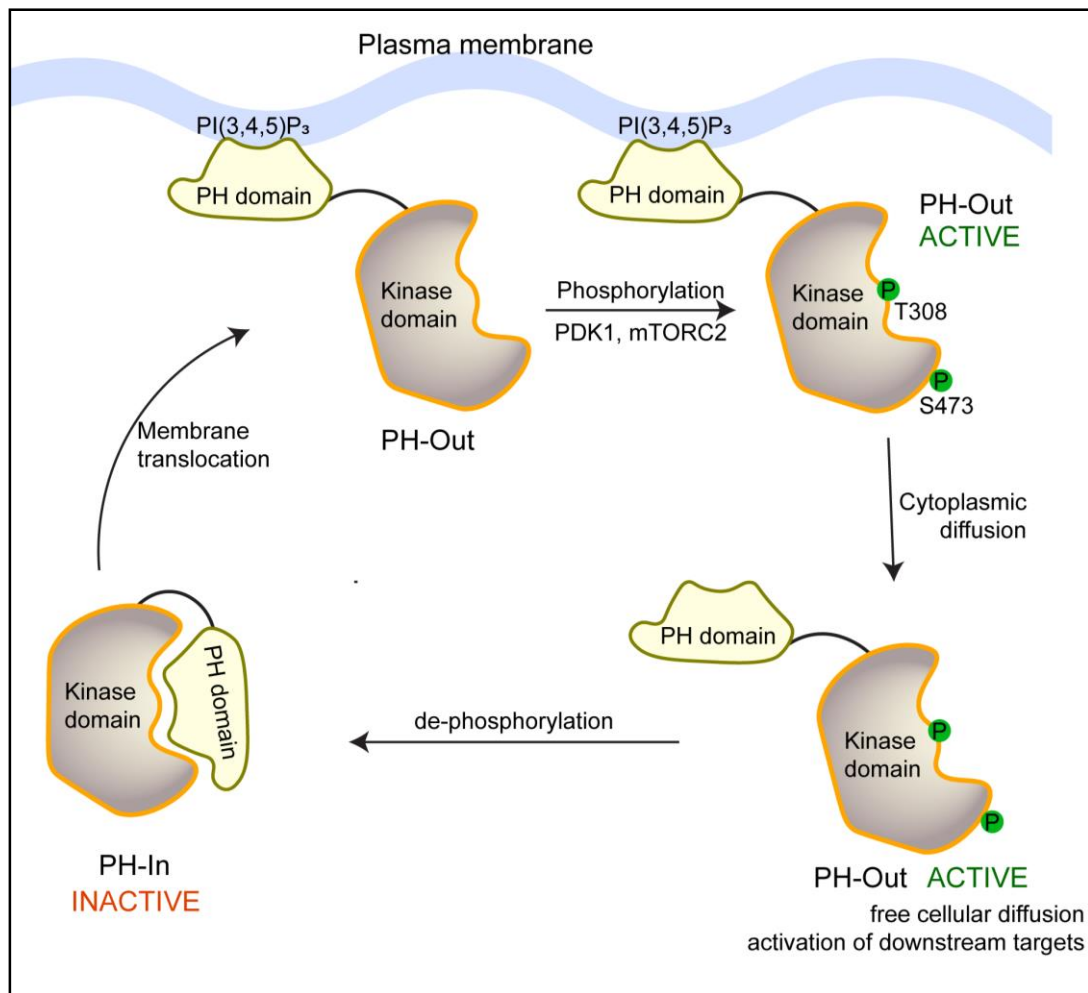


Figure 1.8 Classical mechanism of AKT activation

AKT is present in inhibited “PH-In” conformation in cytoplasm. It gets translocated to plasma membrane by binding PI(3,4,5)P₃ and releases inhibitory interaction between kinase domain and PH domain to adopt “PH-Out” conformation. Subsequent phosphorylation by PDK1 and mTORC2 activates AKT. After activation on plasma membrane, active AKT dissociates from the membrane and diffuses in cytoplasm in “PH-Out” active conformation, subsequently activating downstream targets.

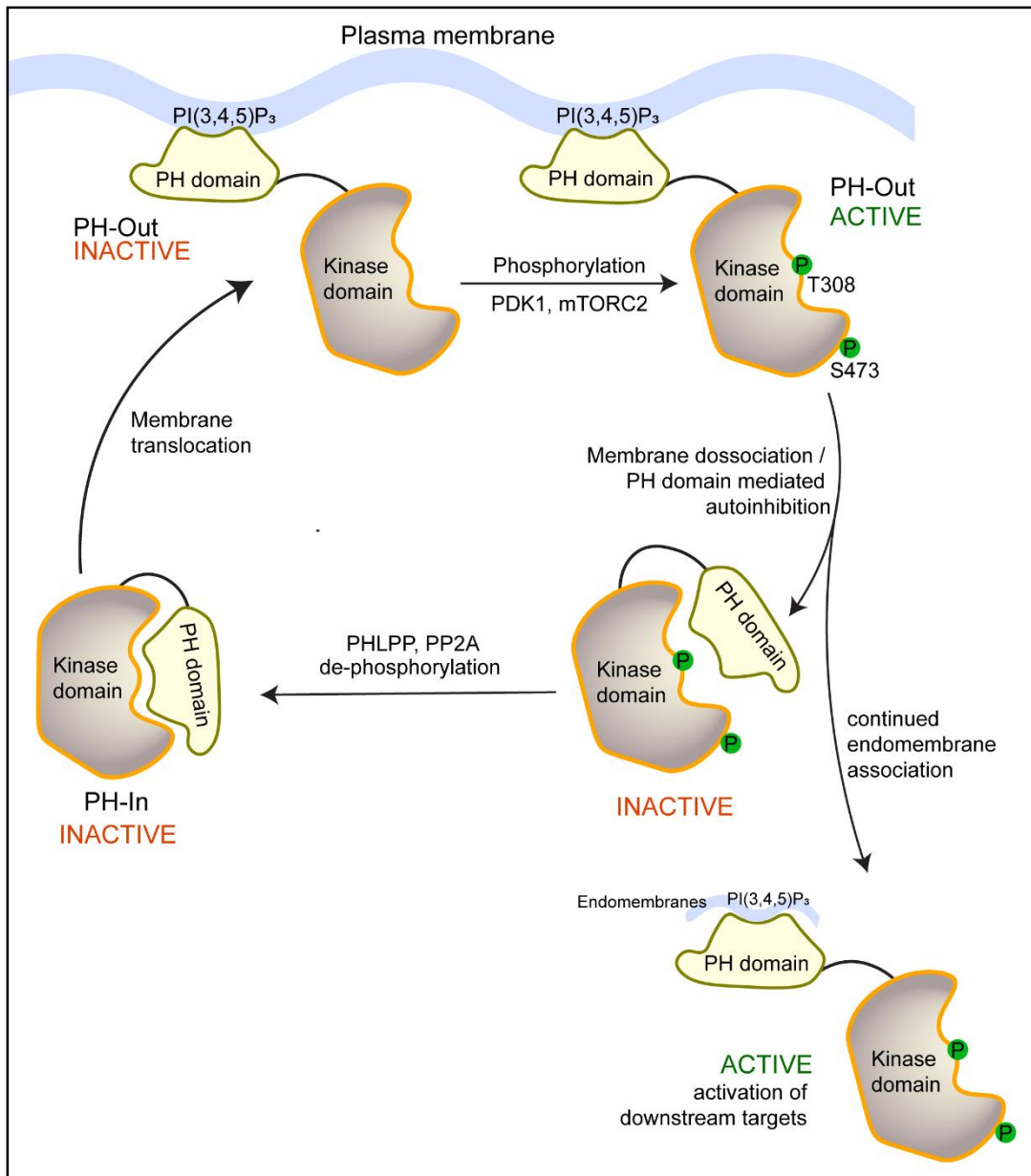


Figure 1.9 Continuous membrane engagement is required to maintain active state of AKT

After plasma membrane translocation, AKT adopts an “PH-Out” conformation and gets phosphorylated. Dissociation from membrane results in PH domain mediated inhibition of kinase activity of AKT and subsequent rapid dephosphorylating by cytoplasmic phosphatases. Continuous attachment with endomembranes is required to maintain prolonged active state of AKT in cytoplasm.

1.8 References

1. G. van Meer, D. R. Voelker, G. W. Feigenson, Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **9**, 112-124 (2008).
2. J. B. Morris, K. A. Hinchliffe, A. Ciruela, A. J. Letcher, R. F. Irvine, Thrombin stimulation of platelets causes an increase in phosphatidylinositol 5-phosphate revealed by mass assay. *FEBS Letters* **475**, 57-60 (2000).
3. J. P. Zewe, A. M. Miller, S. Sangappa, R. C. Wills, B. D. Goulden, G. R. V. Hammond, Probing the subcellular distribution of phosphatidylinositol reveals a surprising lack at the plasma membrane. *J Cell Biol* **219**, (2020).
4. M. A. Lemmon, Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99-111 (2008).
5. J. G. Pemberton, Y. J. Kim, J. Humpolickova, A. Eisenreichova, N. Sengupta, D. J. Toth, E. Boura, T. Balla, Defining the subcellular distribution and metabolic channeling of phosphatidylinositol. *J Cell Biol* **219**, (2020).
6. A. L. Marat, V. Haucke, Phosphatidylinositol 3-phosphates-at the interface between cell signalling and membrane traffic. *EMBO J* **35**, 561-579 (2016).
7. S. Christoforidis, H. M. McBride, R. D. Burgoyne, M. Zerial, The Rab5 effector EEA1 is a core component of endosome docking. *Nature* **397**, 621-625 (1999).
8. R. S. Kuna, S. J. Field, GOLPH3: a Golgi phosphatidylinositol(4)phosphate effector that directs vesicle trafficking and drives cancer. *J Lipid Res* **60**, 269-275 (2019).
9. J. G. Pemberton, T. Balla, in *Protein Reviews – Purinergic Receptors: Volume 20*, M. Z. Atassi, Ed. (Springer International Publishing, Cham, 2019), pp. 77-137.
10. N. A. Sharkey, K. L. Leach, P. M. Blumberg, Competitive inhibition by diacylglycerol of specific phorbol ester binding. *Proc Natl Acad Sci U S A* **81**, 607-610 (1984).
11. Y. Ono, T. Fujii, K. Igarashi, T. Kuno, C. Tanaka, U. Kikkawa, Y. Nishizuka, Phorbol ester binding to protein kinase C requires a cysteine-rich zinc-finger-like sequence. *Proc Natl Acad Sci U S A* **86**, 4868-4871 (1989).
12. J. P. DiNitto, D. G. Lambright, Membrane and juxtamembrane targeting by PH and PTB domains. *Biochim Biophys Acta* **1761**, 850-867 (2006).
13. M. Lenoir, I. Kufareva, R. Abagyan, M. Overduin, Membrane and Protein Interactions of the Pleckstrin Homology Domain Superfamily. *Membranes (Basel)* **5**, 646-663 (2015).
14. J. Schultz, F. Milpetz, P. Bork, C. P. Ponting, SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-5864 (1998).
15. R. J. Haslam, H. B. Koide, B. A. Hemmings, Pleckstrin domain homology. *Nature* **363**, 309-310 (1993).
16. B. J. Mayer, R. Ren, K. L. Clark, D. Baltimore, A putative modular domain present in diverse signaling proteins. *Cell* **73**, 629-630 (1993).
17. H. S. Yoon, P. J. Hajduk, A. M. Petros, E. T. Olejniczak, R. P. Meadows, S. W. Fesik, Solution structure of a pleckstrin-homology domain. *Nature* **369**, 672-675 (1994).
18. J. E. Harlan, P. J. Hajduk, H. S. Yoon, S. W. Fesik, Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature* **371**, 168-170 (1994).
19. T. Balla, Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093-2104 (2005).

20. K. M. Ferguson, M. A. Lemmon, J. Schlessinger, P. B. Sigler, Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain. *Cell* **83**, 1037-1046 (1995).
21. H. Yagisawa, K. Sakuma, H. F. Paterson, R. Cheung, V. Allen, H. Hirata, Y. Watanabe, M. Hirata, R. L. Williams, M. Katan, Replacements of single basic amino acids in the pleckstrin homology domain of phospholipase C-delta1 alter the ligand binding, phospholipase activity, and interaction with the plasma membrane. *J Biol Chem* **273**, 417-424 (1998).
22. M. A. Lemmon, K. M. Ferguson, R. O'Brien, P. B. Sigler, J. Schlessinger, Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. *Proc Natl Acad Sci U S A* **92**, 10472-10476 (1995).
23. S. R. James, C. P. Downes, R. Gigg, S. J. Grove, A. B. Holmes, D. R. Alessi, Specific binding of the Akt-1 protein kinase to phosphatidylinositol 3,4,5-trisphosphate without subsequent activation. *Biochem J* **315 (Pt 3)**, 709-713 (1996).
24. K. Salim, M. J. Bottomley, E. Querfurth, M. J. Zvelebil, I. Gout, R. Scaife, R. L. Margolis, R. Gigg, C. I. Smith, P. C. Driscoll, M. D. Waterfield, G. Panayotou, Distinct specificity in the recognition of phosphoinositides by the pleckstrin homology domains of dynamin and Bruton's tyrosine kinase. *The EMBO Journal* **15**, 6241-6250 (1996).
25. J. K. Klarlund, A. Guilherme, J. J. Holik, J. V. Virbasius, A. Chawla, M. P. Czech, Signaling by phosphoinositide-3,4,5-trisphosphate through proteins containing pleckstrin and Sec7 homology domains. *Science* **275**, 1927-1930 (1997).
26. S. J. Watton, J. Downward, Akt/PKB localisation and 3' phosphoinositide generation at sites of epithelial cell-matrix and cell-cell interaction. *Current Biology* **9**, 433-436 (1999).
27. P. Varnai, T. Balla, Visualization of phosphoinositides that bind pleckstrin homology domains: calcium- and agonist-induced dynamic changes and relationship to myo-[3H]inositol-labeled phosphoinositide pools. *J Cell Biol* **143**, 501-510 (1998).
28. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. DeWald, D. Murray, S. D. Emr, M. A. Lemmon, Genome-wide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol Cell* **13**, 677-688 (2004).
29. W. S. Park, W. D. Heo, J. H. Whalen, N. A. O'Rourke, H. M. Bryan, T. Meyer, M. N. Teruel, Comprehensive identification of PIP3-regulated PH domains from *C. elegans* to *H. sapiens* by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).
30. C. M. Shirey, J. L. Scott, R. V. Stahelin, Notes and tips for improving quality of lipid-protein overlay assays. *Anal Biochem* **516**, 9-12 (2017).
31. K. Narayan, M. A. Lemmon, Determining selectivity of phosphoinositide-binding domains. *Methods* **39**, 122-133 (2006).
32. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. DeWald, D. Murray, S. D. Emr, M. A. Lemmon, Genome-Wide Analysis of Membrane Targeting by *S. cerevisiae* Pleckstrin Homology Domains. *Molecular Cell* **13**, 677-688 (2004).
33. I. Vonkova, A. E. Saliba, S. Deghou, K. Anand, S. Ceschia, T. Doerks, A. Galih, K. G. Kugler, K. Maeda, V. Rybin, V. van Noort, J. Ellenberg, P. Bork, A. C. Gavin, Lipid Cooperativity as a General Membrane-Recruitment Principle for PH Domains. *Cell Rep* **12**, 1519-1530 (2015).
34. T. P. Stauffer, S. Ahn, T. Meyer, Receptor-induced transient reduction in plasma membrane PtdIns(4,5)P2 concentration monitored in living cells. *Current Biology* **8**, 343-346 (1998).

35. R. C. Wills, B. D. Goulden, G. R. V. Hammond, Genetically encoded lipid biosensors. *Mol Biol Cell* **29**, 1526-1532 (2018).
36. P. Varnai, G. Gulyas, D. J. Toth, M. Sohn, N. Sengupta, T. Balla, Quantifying lipid changes in various membrane compartments using lipid binding protein domains. *Cell Calcium* **64**, 72-82 (2017).
37. G. Halet, Imaging phosphoinositide dynamics using GFP-tagged protein domains. *Biol Cell* **97**, 501-518 (2005).
38. B. D. Goulden, J. Pacheco, A. Dull, J. P. Zewe, A. Deiters, G. R. V. Hammond, A high-avidity biosensor reveals plasma membrane PI(3,4)P2 is predominantly a class I PI3K signaling product. *J Cell Biol* **218**, 1066-1079 (2019).
39. G. R. Hammond, M. P. Machner, T. Balla, A novel probe for phosphatidylinositol 4-phosphate reveals multiple pools beyond the Golgi. *J Cell Biol* **205**, 113-126 (2014).
40. G. R. Hammond, T. Balla, Polyphosphoinositide binding domains: Key to inositol lipid biology. *Biochim Biophys Acta* **1851**, 746-758 (2015).
41. O. Idevall-Hagren, P. De Camilli, Detection and manipulation of phosphoinositides. *Biochim Biophys Acta* **1851**, 736-745 (2015).
42. P. Varnai, B. Thyagarajan, T. Rohacs, T. Balla, Rapidly inducible changes in phosphatidylinositol 4,5-bisphosphate levels influence multiple regulatory functions of the lipid in intact living cells. *J Cell Biol* **175**, 377-382 (2006).
43. J. Chen, X. F. Zheng, E. J. Brown, S. L. Schreiber, Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc Natl Acad Sci U S A* **92**, 4947-4951 (1995).
44. G. R. Hammond, M. J. Fischer, K. E. Anderson, J. Holdich, A. Koteci, T. Balla, R. F. Irvine, PI4P and PI(4,5)P2 are essential but independent lipid determinants of membrane identity. *Science* **337**, 727-730 (2012).
45. B. C. Suh, T. Inoue, T. Meyer, B. Hille, Rapid chemically induced changes of PtdIns(4,5)P2 gate KCNQ ion channels. *Science* **314**, 1454-1457 (2006).
46. A. Jain, R. Liu, B. Ramani, E. Arauz, Y. Ishitsuka, K. Ragunathan, J. Park, J. Chen, Y. K. Xiang, T. Ha, Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484-488 (2011).
47. A. Jain, E. Arauz, V. Aggarwal, N. Ikon, J. Chen, T. Ha, Stoichiometry and assembly of mTOR complexes revealed by single-molecule pulldown. *Proc Natl Acad Sci U S A* **111**, 17833-17838 (2014).
48. E. Arauz, V. Aggarwal, A. Jain, T. Ha, J. Chen, Single-Molecule Analysis of Lipid-Protein Interactions in Crude Cell Lysates. *Anal Chem* **88**, 4269-4276 (2016).
49. D. R. Cook, K. L. Rossman, C. J. Der, Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. *Oncogene* **33**, 4021-4035 (2014).
50. I. Whitehead, H. Kirk, C. Tognon, G. Trigo-Gonzalez, R. Kay, Expression cloning of lfc, a novel oncogene with structural similarities to guanine nucleotide exchange factors and to the regulatory region of protein kinase C. *J Biol Chem* **270**, 18388-18395 (1995).
51. K. L. Rossman, L. Cheng, G. M. Mahon, R. J. Rojas, J. T. Snyder, I. P. Whitehead, J. Sondek, Multifunctional roles for the PH domain of Dbs in regulating Rho GTPase activation. *J Biol Chem* **278**, 18393-18400 (2003).
52. J. N. Cash, E. M. Davis, J. J. G. Tesmer, Structural and Biochemical Characterization of the Catalytic Core of the Metastatic Factor P-Rex1 and Its Regulation by PtdIns(3,4,5)P3. *Structure* **24**, 730-740 (2016).

53. M. A. Baumeister, L. Martinu, K. L. Rossman, J. Sondek, M. A. Lemmon, M. M. Chou, Loss of phosphatidylinositol 3-phosphate binding by the C-terminal Tiam-1 pleckstrin homology domain prevents in vivo Rac1 activation without affecting membrane targeting. *J Biol Chem* **278**, 11457-11464 (2003).
54. F. Medina, A. M. Carter, O. Dada, S. Gutowski, J. Hadas, Z. Chen, P. C. Sternweis, Activated RhoA is a positive feedback regulator of the Lbc family of Rho guanine nucleotide exchange factor proteins. *J Biol Chem* **288**, 11325-11333 (2013).
55. V. Calleja, D. Alcor, M. Laguerre, J. Park, B. Vojnovic, B. A. Hemmings, J. Downward, P. J. Parker, B. Larijani, Intramolecular and intermolecular interactions of protein kinase B define its activation in vivo. *PLoS Biol* **5**, e95 (2007).
56. W. I. Wu, W. C. Voegtli, H. L. Sturgis, F. P. Dizon, G. P. Vigers, B. J. Brandhuber, Crystal structure of human AKT1 with an allosteric inhibitor reveals a new mode of kinase inhibition. *PLoS One* **5**, e12913 (2010).
57. M. Ebner, I. Lucic, T. A. Leonard, I. Yudushkin, PI(3,4,5)P3 Engagement Restricts Akt Activity to Cellular Membranes. *Mol Cell* **65**, 416-431 e416 (2017).
58. I. Lucic, M. K. Rathinaswamy, L. Truebestein, D. J. Hamelin, J. E. Burke, T. A. Leonard, Conformational sampling of membranes by Akt controls its activation and inactivation. *Proc Natl Acad Sci U S A* **115**, E3940-E3949 (2018).
59. P. Varnai, K. I. Rother, T. Balla, Phosphatidylinositol 3-kinase-dependent membrane association of the Bruton's tyrosine kinase pleckstrin homology domain visualized in single living cells. *J Biol Chem* **274**, 10983-10989 (1999).
60. K. Venkateswarlu, P. B. Oatey, J. M. Tavaré, P. J. Cullen, Insulin-dependent translocation of ARNO to the plasma membrane of adipocytes requires phosphatidylinositol 3-kinase. *Current Biology* **8**, 463-466 (1998).
61. W. A. Kimber, L. Trinkle-Mulcahy, P. C. Cheung, M. Deak, L. J. Marsden, A. Kieloch, S. Watt, R. T. Javier, A. Gray, C. P. Downes, J. M. Lucocq, D. R. Alessi, Evidence that the tandem-pleckstrin-homology-domain-containing protein TAPP1 interacts with Ptd(3,4)P2 and the multi-PDZ-domain-containing protein MUPP1 in vivo. *Biochem J* **361**, 525-536 (2002).
62. D. J. Gillooly, I. C. Morrow, M. Lindsay, R. Gould, N. J. Bryant, J. M. Gaullier, R. G. Parton, H. Stenmark, Localization of phosphatidylinositol 3-phosphate in yeast and mammalian cells. *EMBO J* **19**, 4577-4588 (2000).
63. D. J. Gillooly, A. Simonsen, H. Stenmark, Cellular functions of phosphatidylinositol 3-phosphate and FYVE domain proteins. *Biochem J* **355**, 249-258 (2001).
64. J. M. Gaullier, E. Ronning, D. J. Gillooly, H. Stenmark, Interaction of the EEA1 FYVE finger with phosphatidylinositol 3-phosphate and early endosomes. Role of conserved residues. *J Biol Chem* **275**, 24595-24600 (2000).
65. C. D. Ellson, S. Gobert-Gosse, K. E. Anderson, K. Davidson, H. Erdjument-Bromage, P. Tempst, J. W. Thuring, M. A. Cooper, Z. Y. Lim, A. B. Holmes, P. R. Gaffney, J. Coadwell, E. R. Chilvers, P. T. Hawkins, L. R. Stephens, PtdIns(3)P regulates the neutrophil oxidase complex by binding to the PX domain of p40(phox). *Nat Cell Biol* **3**, 679-682 (2001).
66. T. P. Levine, S. Munro, Targeting of Golgi-Specific Pleckstrin Homology Domains Involves Both PtdIns 4-Kinase-Dependent and -Independent Components. *Current Biology* **12**, 695-704 (2002).
67. A. Godi, A. Di Campli, A. Konstantakopoulos, G. Di Tullio, D. R. Alessi, G. S. Kular, T. Daniele, P. Marra, J. M. Lucocq, M. A. De Matteis, FAPPs control Golgi-to-cell-surface membrane traffic by binding to ARF and PtdIns(4)P. *Nat Cell Biol* **6**, 393-404 (2004).

CHAPTER 2. LIPID-BINDING SPECIFICITY OF FULL-LENGTH HUMAN PH DOMAIN-CONTAINING PROTEINS

2.1 Introduction

Specific phospholipid-protein interactions are critical to the regulation of many signal transduction processes and cellular functions. These interactions typically involve lipid-binding domains recognized by specific lipid species and/or physical properties of the membrane such as charge or curvature (1-4). The largest family of putative lipid-binding domains (LBDs) is the pleckstrin homology (PH) domain, with over 250 members encoded by the human genome (Pfam database). Originally defined by its presence in the protein Pleckstrin (5, 6), this domain of 100-120 amino acids has an invariable structure of seven-stranded β -sandwich lined by a C-terminal α -helix (7). PH domains are found in many different types of proteins that are involved in regulating diverse signaling pathways and functions. Some of the earliest characterized PH domains, including those in Pleckstrin, RasGAP, and GRK2, were found to have affinity for PI(4,5)P₂ (8). The PH domains in PLC δ and AKT1 bind with high selectivity to PI(4,5)P₂ and PI(3,4,5)P₃ (PIP₃), respectively (9, 10). Indeed, the interaction with PIP is so specific for those two PH domains that they have been commonly used as sensors to detect PI(4,5)P₂ and PIP₃ in cells (11, 12).

Despite those early examples of PH-PIP interactions, to date only a modest number of PH domain-containing proteins have been demonstrated to bind PIPs with specificity. A comprehensive analysis of *S. cerevisiae* PH domains revealed that only one of them had specific affinity for a particular PIP and the rest of them displayed little affinity or selectivity for PIPs (13). Another study examined a large number of mouse PH domains and found 20% of them to translocate to the plasma membrane (PM) in response to PIP₃

production in cells, but most of those PIP₃-responsive PH domains did not show specific binding to PIP₃ in lipid binding assays (14). Thus, it has been proposed that the specific PIP recognition by PLCδ-PH and AKT1-PH may be the exception rather than the rule for PH domains and that only a small percentage of all PH domains bind PIPs with high affinity and specificity (4). Protein partners have been identified for PH domains, which may cooperate with lipid-PH interaction or operate independently to regulate or mediate PH domain functions (1, 15).

It is important to note that our knowledge to date of the lipid binding properties of PH domain-containing proteins is largely derived from studies of isolated PH domains rather than full-length proteins, at least partly owing to the hurdle of purifying full-length proteins for lipid binding assays. Intra-molecular interaction is a universal mechanism for determining protein structure and activity, and this crucial determinant would be eliminated when a PH domain is taken out of the context of the protein. To circumvent the laborious process of purifying proteins and to preserve post-translational modifications of the proteins, we have developed a total internal reflection fluorescence (TIRF) microscopy-based single-molecule pulldown (SiMPull) assay to study protein-lipid interactions with whole cell lysates and lipid vesicles (16), herein referred to as lipid-SiMPull or SiMPull. The sensitivity and specificity of this assay have been demonstrated through highly selective pulldown of various LBDs, as well as full-length AKT1, from cell lysates by vesicles containing PIPs known to interact with those proteins (16). We have now applied the lipid-SiMPull assay to interrogate 67 human PH domain proteins for their binding to vesicles of various compositions. Our results suggest that PIP recognition by PH domain proteins is more prevalent than previously believed. We have also used the assay data to generate a recursive-learning algorithm, which reveals PH domain sequence determinants for PIP binding and predicts PIP binding for the entire family of human PH domains.

2.2 Materials and Methods

2.2.1 Cell Culture

HEK293 cells were maintained in high-glucose DMEM with 10% FBS, 2 mM L-glutamine, and penicillin/streptomycin at 37 °C in 5% CO₂. NIH3T3 cells were maintained in high-glucose DMEM with 10% newborn calf serum and 4 mM L-glutamine, and penicillin/streptomycin at 37 °C in 5% CO₂. For SiMPull experiments, HEK293 cells were transfected for 24 hours in 6-well plates using PolyFect® (3 µL/µg DNA) following manufacturer's recommendations. For immunofluorescence experiments, cells were plated in 12-well plate on poly-L-lysine coated coverslips and transfected with Lipofectamine™ 3000 following manufacturer's recommendations.

2.2.2 Cell Lysis and Western Blotting

For SiMPull assays, the cells were collected after 24-hr transfection in detergent-free buffer (40 mM HEPES, pH8.0, 150 mM NaCl, 10 mM β-glycerophosphate, 10 mM sodium pyrophosphate, 2 mM EDTA, 1x Sigma protease inhibitor cocktail). The cells were lysed by sonication for 3 seconds on ice followed by ultracentrifugation at 90,000xg in a TLA100.3 rotor for 1 hour at 4°C. For SiMPull, EGFP concentration was measured using a standard emission (ex488/em520) curve of pure recombinant EGFP, and each cell lysate was diluted in vesicle buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl) to yield 5 nM EGFP. For western blotting, cells were lysed in 1x SDS sample buffer or as described above and mixed at 1:1 with 2x SDS sample buffer, both containing β-mercaptoethanol at a final concentration of 5%, and heated for 5 min at 95 °C. Proteins were resolved by SDS-PAGE and transferred onto PVDF membrane. The membrane was incubated with primary and secondary antibodies following manufacturers' recommendations. HRP-conjugated

secondary antibody was reacted with West Pico PLUS Chemiluminescent Substrate, and the signal was detected using an Invitrogen iBright digital imager.

2.2.3 cDNA Cloning and Mutagenesis

Plasmids harboring cDNAs for human PH domain-containing proteins from human cDNA library hORFeome V5.1 were cloned as a pool into the pDest-eEGFP-N1 vector using Gateway Cloning, followed by identification of individual clones. The PH domains were cloned into the pEGFP-C1 vector using Gibson assembly. p40PhoxPX-EGFP was obtained from Addgene (#19010) (17). Point mutants for p40PhoxPX-EGFP were created using site-directed mutagenesis using QuikChange Lightning Site-Directed Mutagenesis Kit following the manufacturer's protocol. All cDNAs in the final plasmids were sequence-confirmed in their entirety.

2.2.4 Lipid Vesicle Preparation

All lipids were mixed in chloroform (0.166 μ mol total) and dried under nitrogen flow. The dried mixture was re-suspended in 100 μ L vesicle buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl) to a final concentration of 1.66 mM. After 30 min incubation at room temperature, vesicles were formed by water bath sonication (Laboratory Supplies, Hicksville, NY, model G112SPIT, 600 v, 80 kc, and 0.5 A) in 4 cycles of 4 min each. Small unilamellar vesicles were collected as the supernatant after ultracentrifugation at 194,398xg in a TLA100.3 rotor for 1 hr at 25 °C. Lipid vesicles were kept at 4° C for 10 days to be used in SiMPull assays.

2.2.5 Single-Molecule Pulldown Assay

Quartz slides were prepared as described in Jain et al. and Arauz et al. (16, 18). Briefly, the slides were thoroughly cleaned and passivated with PEG doped with 0.1-0.2% biotin-PEG. Neutravidin (200 μ g/mL) was incubated in chambers for 10 min, followed by

addition of biotinylated lipid-vesicles. Freshly prepared whole-cell lysate (80 μ L) was added by flowing into each slide chamber, replacing the vesicle solution. An inverted total internal reflection fluorescence (TIRF) microscope with Olympus 100x, NA1.47 lens and EMCCD camera (Andor iXon Ultra 897) was used to acquire single-molecule data at 10 frames/second. Diode pumped solid state lasers were used to excite EGFP at 488nm (Coherent) and DiD at 638nm (Cobalt). All SiMPull experiments were performed at room temperature. Average number of spots per imaging area ($\sim 1600 \mu\text{m}^2$) was calculated from 10 or more images.

2.2.6 SiMPull Image Analysis

TIRF image acquisitions were processed in IDL to generate image files, and the number of fluorescence spots were identified with point-spread function. MATLAB® was used to extract data from data files generated in IDL and import into Excel. The background number of EGFP spots was gathered prior to addition of lysate and subtracted from EGFP spots after lysate addition to generate data for each image. At least 10 SiMPull images ($1600 \mu\text{m}^2$ each) were analyzed to generate the average number of spots per imaging area for each assay. The IDL scripts used to process raw image files are publicly available from Taekjip Ha laboratory (<http://ha.med.jhmi.edu/resources/>).

2.2.7 Fluorescence Imaging

HEK293 or NIH3T3 cells on poly-L-lysine-coated glass coverslips were fixed in 3.7% paraformaldehyde at room temperature for 15 min and permeabilized with 0.1% Triton X-100 for 5 min. For visualizing stress fibers, Rhodamine phalloidin was diluted 1:3000 and incubated with 3% BSA/PBS at 4 °C for 60 min, followed by incubation with DAPI (1:2500) for 20 min at room temperature. For PM localization experiments, HEK293 cells were transfected and treated as described in Figure 2.5 legend, fixed in 3.7%

paraformaldehyde, and stained with DAPI for 20 min. A personal deconvolution microscope system (DeltaVision, Applied Precision) was used with a 100x or 60x NA 1.4 lens to capture fluorescence images. Deconvolution used an enhanced ratio iterative–constrained algorithm (19). The acquired images were processed in ImageJ. For quantification of PM-localized cells, 60-100 transfected cells were counted in each experiment.

2.3 Results

2.3.1 Lipid-SiMPull Assay for Human PH Domain Proteins

In order to investigate binding of PIPs by PH domain proteins using the lipid-SiMPull assay (16), we created cDNA constructs and transfected them in human embryonic kidney (HEK) 293 cells to express 67 full-length human PH domain proteins with EGFP fused at their C-termini. Those 67 proteins were randomly collected based on availability of intact cDNA clones from the hORFeome V5.1 library. Each cDNA in the final plasmid was sequence-validated in its entirety, and expression of the full-length fusion protein in HEK293 cells was confirmed by western analysis (see examples in Figure 2.1 A). Concentrations of EGFP-fusion proteins in the lysates were determined by fluorescence intensity and calculated using a standard emission (ex488/em520) curve generated with purified recombinant EGFP (Figure 2.1 B). Small unilamellar vesicles (SUVs) were made with 60 mol % phosphatidylcholine (PC), 15 mol % phosphatidylserine (PS), 20 mol % cholesterol, and 5 mol % of one of seven PIPs: PI(3)P, PI(4)P, PI(5)P, PI(3,4)P₂, PI(4,5)P₂, PI(3,5)P₂, and PI(3,4,5)P₃. Vesicles containing 5% phosphatidic acid (PA), a different type of signaling lipid than PIPs, were also included in our assays. As a negative control, vesicles were made to include an additional 5 mol % PC in place of PIP (referred to as PC or control vesicles hereafter). All vesicles also had 0.01 mol % biotinylated phosphatidylethanol (PE) to facilitate immobilization on imaging slides.

SiMPull assays were performed with the SUVs immobilized on slides and lysates containing 5 nM EGFP-fusion protein flowed into the slide chamber followed by TIRF imaging (Figure 2.2 A). Each of the 67 proteins was assayed against each of the 9 types of vesicles. The SUVs also contained a lipophilic red dye, DiD, to allow visualization of the immobilized vesicles on slides and confirmation of consistent vesicle density across assays. Pulldown of EGFP proteins was quantified by counting green fluorescent spots. Each preparation of SUV was validated by binding to a positive control in SiMPull, and only the experiments with validated SUVs were further processed. Each vesicle-protein pair was assayed using validated vesicles with at least three independent lysates. Based on characterization of many positive and negative controls with the SiMPull assay, we arrived at 100 EGFP spots per area of $1600 \mu\text{m}^2$ (after background subtraction) as the threshold for binding and the data were interpreted as a binary outcome – binding or no binding.

2.3.2 New PH Domain Protein-PIP Interactions are Revealed

The assay results for all proteins are summarized in Table 1. Of the 67 proteins examined, 4 proteins displayed promiscuous binding to lipids including PA and PC: CERT1, NET1, PLEKHO1 and SPATA13. Misfolding could not be ruled out for those proteins. Thirty-two proteins were found to bind PIPs with various selectivity, and none of them bound PA. Among the 31 proteins that did not bind lipids in our initial assays, some had been previously reported to bind PIPs as either PH domain alone or full-length protein. We considered the possibility that some of those proteins might require more than 5% PIP in the vesicle for effective binding, since PIPs in the cell membrane can cluster and their asymmetric distribution may lead to drastically higher local concentrations (20). Thus, we performed SiMPull assays with vesicles containing 20% PIP_3 or $\text{PI}(4,5)\text{P}_2$ for 10 proteins that had been reported to bind either PIP but not found to bind in our assays

(ARHGAP12, DOK1, FERMT2, GRB14, GRK2, OSBPL8, PLCD4, PLEK, SBF2, SKAP2).

Four of the proteins (ARHGAP12, PLCD4, SBF2, SKAP2) indeed bound 20% PIP₃ or PI(4,5)P₂.

In total, 36 of the 67 proteins (54%) were found to bind PIPs with some specificity (Table 1, Figure 2.2 B). The positive binding data are also illustrated in a binary table in Figure 2.2 C. Representative SiMPull assay images and quantification are shown in Figure 2.2 D for 7 proteins, each displaying specificity for a different PIP. Of the 36 proteins found to bind PIPs, our assay results confirmed reported PIP binding by 8 proteins (Group I in Figure 2.2 B & C), whereas 28 proteins (Group II) had been either not previously reported to bind any lipid or reported to display different lipid binding profiles than those found in our assays.

Of the 27 proteins found not to bind PIP in our assays, 11 had been reported to bind lipids (Group III) and 16 not reported (Group IV). It is possible that those 11 represent false negatives in our assay. Previously we showed that the assay detected LBD-lipid interactions that were reported to have K_d 's in the high nanomolar/sub-micromolar range (16). To estimate the threshold of detection under our assay conditions, we performed SiMPull with the PX domain of p40Phox, the WT, R60A, and K92A mutant of which bind PI(3)P with K_d of 5 μ M, 17.5 μ M, and >50 μ M, respectively (21). As shown in Figure 2.3, robust binding to PI(3)P, but not to PI(4,5)P₂, was detected for WT p40Phox-PX. Binding to the R60A mutant was also observed, at just above the cut-off of 100 EGFP spots, whereas the K92A mutant did not bind the lipid. Therefore, our assay under current conditions has a detection threshold of affinity in the 10-20 μ M range. It is unlikely that many physiologically relevant interactions would have been missed.

2.3.3 Phosphatidylserine is not Responsible for the Novel PH Domain Protein-PIP Interactions

Phospholipids may cooperate in the membrane to determine specific interaction with proteins (2, 3). The presence of sphingolipids and PS in the membrane has been found to contribute to enhanced affinity and/or specificity for PIPs by yeast PH domains (22). Since the SUVs in our assays contained 15 mol% PS (based on mammalian cell membrane composition), we wondered whether the newly discovered PIP-PH protein interactions were dependent on PS. To address this possibility, we re-assayed 23 of the new PIP-binding proteins (Group II) with SUVs containing only PC and a specific PIP, omitting PS and cholesterol. Remarkably, all 23 proteins bound their respective PIPs without PS in the vesicles (Figure 2.4 A). Representative SiMPull images are presented in Figure 2.4 B. Therefore, the presence of PS does not appear to account for the novel PIP-PH protein interactions emerged from our studies.

2.3.4 Plasma Membrane Targeting of PIP₃ Binding Proteins

One potential consequence of PIP binding is targeting of the protein to specific membranes in the cell. PIP₃ is acutely produced by mitogenic stimulation through activation of PI-3-kinase (PI3K) in the PM, resulting in recruitment of proteins to the PM through their binding to PIP₃. AKT1 is an example of such a protein. As a means to validate PIP binding in the cell, we set out to examine PM translocation of all the proteins found to bind PIP₃ but not PI(4,5)P₂ in our SiMPull assays (a total of 8). We excluded those that bind both PIP₃ and PI(4,5)P₂ because PI(4,5)P₂ is present at high concentrations in the PM and could potentially mask any effect of PIP₃. HEK293 cells transiently expressing EGFP fusion proteins were serum starved and then stimulated with insulin for 5 min, with or without treatment by wortmannin, a PI3K inhibitor. As shown in Figure 2.5, AKT1-PH as a positive control was found to translocate to PM in response to

acute insulin stimulation and this translocation was abolished by wortmannin treatment. BTK, FERMT3, GAB2, ITK, and SKAP1 displayed various degrees of PM translocation upon insulin stimulation in 50-75% of transfected cells, which were also abolished by wortmannin. PLCD4, which bound 20% PIP₃ in SiMPull assays, translocated to PM only in small patches (see image) and only in ~25% of transfected cells. Nevertheless, this insulin effect was unmistakable, as starved or wortmannin-treated cells never displayed any PM localization of PLCD4. Another protein that bound 20% PIP₃, SKAP2, was not found at the PM. ARHGEF39, shown to bind PIP₃ and PI(3,4)P₂ in our assays, displayed constitutive PM localization insensitive to wortmannin. It is possible that ARHGEF39 is targeted to the PM via lipid modification or interaction with a membrane protein. Taken together, the observed membrane recruitment is largely consistent with the protein-PIP₃ interactions we have discovered with the SiMPull assays.

2.3.5 Isolated PH Domains do not have the Same Lipid-Binding Properties as Their Full-Length Counterparts

To compare the behaviors of PH domains to their full-length protein counterparts, we subcloned 14 PH domains from 11 proteins (3 of them have two PH domains) to express as EGFP-fusions in HEK293 cells and performed SiMPull assays with the lysates. With the exception of AFAP1L1, which had been reported to bind PIP₃ but did not bind in SiMPull assays, those full-length proteins all displayed PIP binding in our assays. All the PH domains were first assayed against the negative control vesicles and PA vesicles, and no binding was observed for any. We then proceeded to assay the PH domains with selected PIP vesicles. The results are summarized in Table 2.2. Surprisingly, only one PH domain displayed the same PIP binding as its full-length counterpart: PH2 of PLEKHA1 for PI(3,4)P₂. The PH domains from the remaining 10 proteins did not behave like their

full-length counterparts: 9 PH domains did not bind any lipid, and 3 PH domains bound lipids differently than the full-length proteins.

One possibility is that lipid-binding domains other than PH could have been responsible for our observed PIP binding by full-length proteins. However, only 2 of those 10 proteins above contain other known lipid-binding domains: BAR domain in ACAP1 and FERM domain in FERMT3. A BAR-PH fragment of ACAP1 bound lipid vesicles non-specifically in our assays (data not shown). The FERM domain was not examined in our study but it had been reported to *not* contribute to PIP binding by FERMT3 (23). Although we cannot rule out the possibility of EGFP interfering with PH domain activity in the fusion proteins, our observations are consistent with the notion that a PH domain without the context of its full-length protein may not have the same lipid-binding properties. It is also important to note that given the diverse domain structures of the 34 proteins found to bind PIPs with PH domain being the only common domain among them, and the scarcity of other lipid-binding domains in those proteins, the PH domains most likely mediate the PIP binding by full-length proteins. Another possibility could be misfolding of the PH domains when expressed alone and/or interference by the EGFP tag, which is difficult to rule out. But misfolding alone cannot explain different binding behavior by 12 PH domain proteins, when EGFP tagged PH domains are routinely used in the field for probing protein-lipid interactions.

2.4 Discussion

Our analysis of a random collection of 67 human PH domain-containing proteins using the lipid-SiMPull assay reveals that 54% of them bind PIPs and that many of the specific interactions have not been reported before. The SiMPull assay offers a unique advantage of using full-length proteins expressed in the native environment (i.e., mammalian cells) without the need for purification. These proteins would be likely to

maintain their native structures and proper post-translational modifications. However, due to overexpression of the EGFP-fusion proteins it is unlikely that any endogenous proteins in the lysate would reach stoichiometric levels to contribute directly to the observed recombinant protein binding to lipid vesicles. Hence, the lipid-protein interactions observed in lipid-SiMPull assays are most likely intrinsic properties of the PH domain-containing proteins. Our experimental validation of the PIP binding predicted by a recursive learning algorithm based on PH domain sequences also suggests that the PH domain defines PIP binding by the full-length proteins.

We have determined the threshold of detection under the current lipid-SiMPull assay conditions to be in the K_d range of 10-20 μ M, which is sufficiently sensitive to detect most reported and/or physiologically relevant protein-PIP interactions. Although the possibility of interference by the EGFP tag has not been ruled out for those proteins that did not display any lipid binding in our assays, we do not consider it likely that many true interactions have been missed by our assay. Instead, we would like to suggest that non-physiological binding conditions (such as in the lipid strips assay) and/or isolation of the PH domains out of the context of full-length proteins may have been accountable for at least some of the discrepancies between our observations and those in the literature. Indeed, of those 11 proteins that were reported to bind PIPs but found not to bind in our assays (Group III), 10 were based on isolated PH domains and/or lipid strips assays.

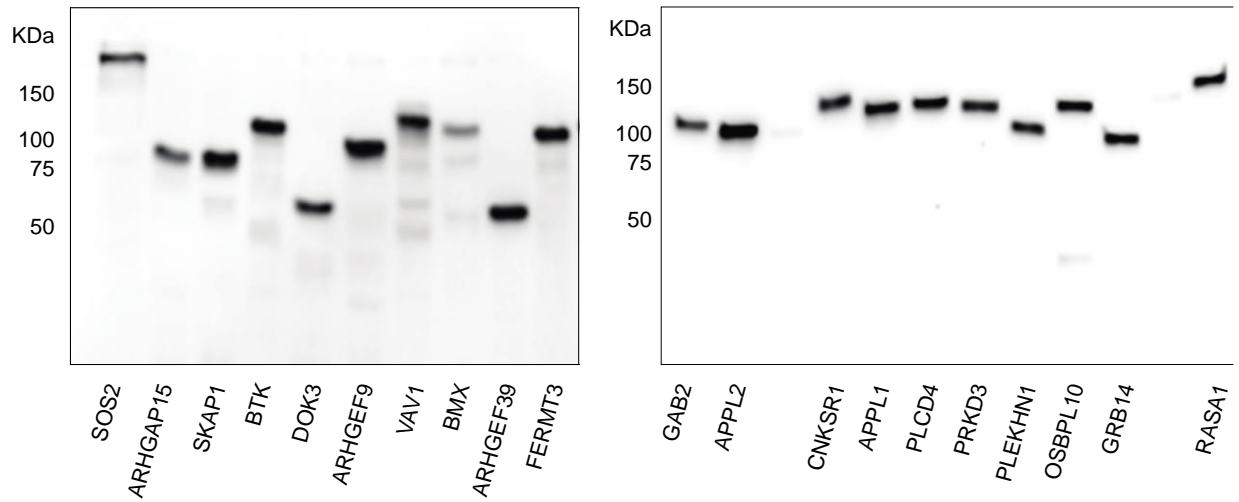
The PIP binding specificity discovered in our study can potentially lead to new understanding of the functions and mechanisms of the PH domain proteins. For instance, previously the PH domain of CNKSR1 (or CNK1) had been found to bind PIPs weakly and non-specifically (24), but in our assays the full-length CNKSR1 protein bound specifically to PI(4)P. Interestingly, CNKSR1 has been reported to promote PI(4,5)P₂ production from PI(4)P by phosphatidylinositol 4-phosphate 5-kinases (PIP5Ks) at the PM where insulin

receptor signaling takes place (25), although the mechanism is unclear. Based on our finding, it is conceivable that CNKSR1 is targeted to PI(4)P-enriched membrane in order to activate PIP5K or to recruit PIP5K to its substrate. Another example is EXOC8 (or Exo84), a component of the exocyst, which is an octameric protein complex involved in tethering secretory vesicles to the plasma membrane for fusion during exocytosis (26). A generally accepted view is that Sec3 and Exo70 in the exocyst bind PI(4,5)P₂ and target the entire octameric complex to the membrane. Our observation of PI(4,5)P₂ binding by EXOC8 suggests that EXOC8 may also contribute directly to the regulation of the exocyst by PI(4,5)P₂, and that further investigation could potentially lead to revision of the current model of exocyst assembly.

The vast majority of the 70 members of the Dbl family of RhoGEFs contain a PH domain immediately following the catalytic DH domain (27). Although those PH domains had been speculated to bind phospholipids and subsequently contribute to membrane localization and/or allosteric regulation of the RhoGEFs, very few of them have been reported to have marked affinity for specific PIPs. Thus, the idea of phospholipids regulating RhoGEFs through their PH domains remains controversial (28). Among the proteins we studied with the SiMPull assay, 14 are RhoGEFs belonging to the Dbl family. PIP binding was observed for 9 of them (ARHGEF3, ARHGEF5, ARHGEF7, ARHGEF9, ARHGEF16, ARHGEF39, MCF2L, NGEF, VAV1), all of which were either never reported to bind lipids or reported to bind lipids with different specificity than what we observed. For instance, the PH domain of ARHGEF3 was reported to bind phospholipids with little selectivity in lipid strips assays (14). We studied the functional relevance of lipid binding by ARHGEF3 and found it to be important for stress fibre formation. It will be discussed in detail in chapter 4. Follow-up studies on the 8 remaining RhoGEFs and investigation of lipid binding by other Dbl family members will also likely be illuminating.

2.5 Figures and Tables

A



B

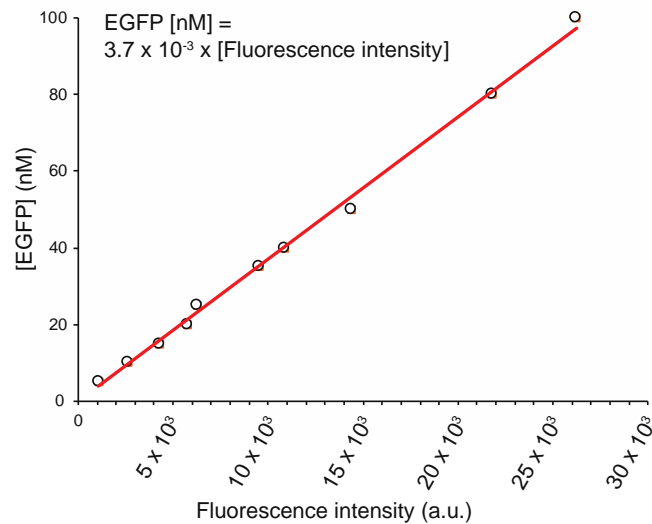


Figure 2.1

(A) Examples of EGFP-fusion proteins transiently expressed in HEK293 cells. Cell lysates were subjected to western blotting with an anti-GFP antibody.

(B) Fluorescence of pure recombinant EGFP (ex 488 / em 520) was measured at 10 different concentrations on a Spectramax GeminiXPS plate reader (Molecular Devices) to yield a linear standard curve.

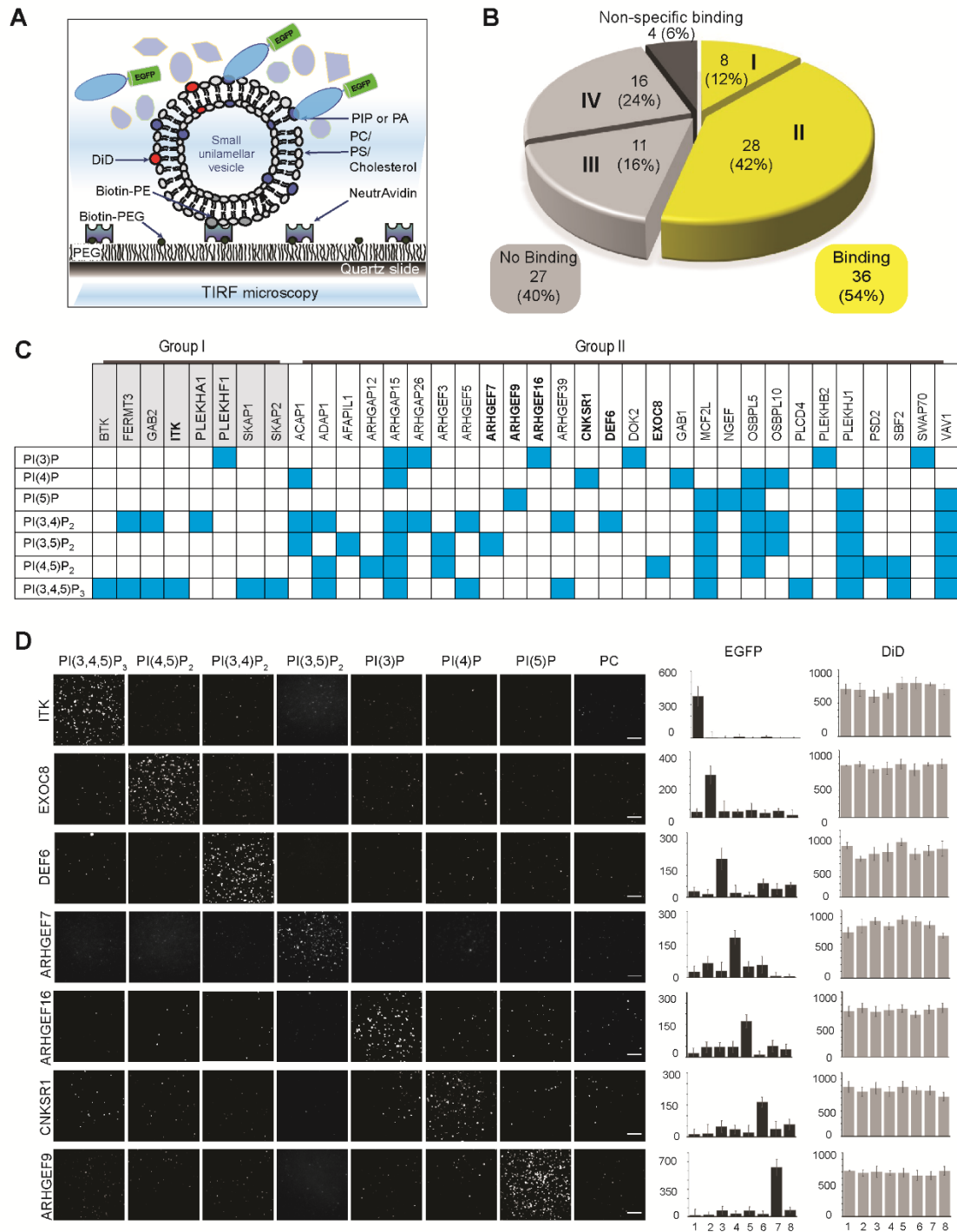


Figure 2.2 Specific PIP binding by full-length human PH domain proteins.

(A) Schematic representation of the lipid-SiMPull assay to detect lipid-protein interactions.

(B) A pie chart summarizing the results of PIP-binding assays for 67 full-length human PH domain proteins against 9 types of vesicles (see text for details). Group I: reported binding

Figure 2.2 (cont.)

confirmed. Group II: novel binding. Group III: binding reported but not found by SiMPull. Group IV: binding not reported nor found by SiMPull.

(C) A binary representation of results for the 36 proteins found to bind PIPs in SiMPull. The vesicles are identified by the unique PIP, but all contained PC, PS, cholesterol, DiD, and biotin-PE. Blue: binding; white: no binding. SiMPull data for the proteins in bold are shown in D.

(D) Representative TIRF images of 7 EGFP-fusions pulled down by vesicles containing 5% PIP, with PC as a negative control. EGFP and DiD spots were quantified from 10 or more image areas to yield the average number of spots per 1600 μm^2 for each assay, shown in the graphs on the right. Each assay was repeated with lysates from at least three independent transfections, and the data can be found in Supplementary Table 1. The results for 67 proteins are summarized in Table 1. Scale bars: 5 μm .

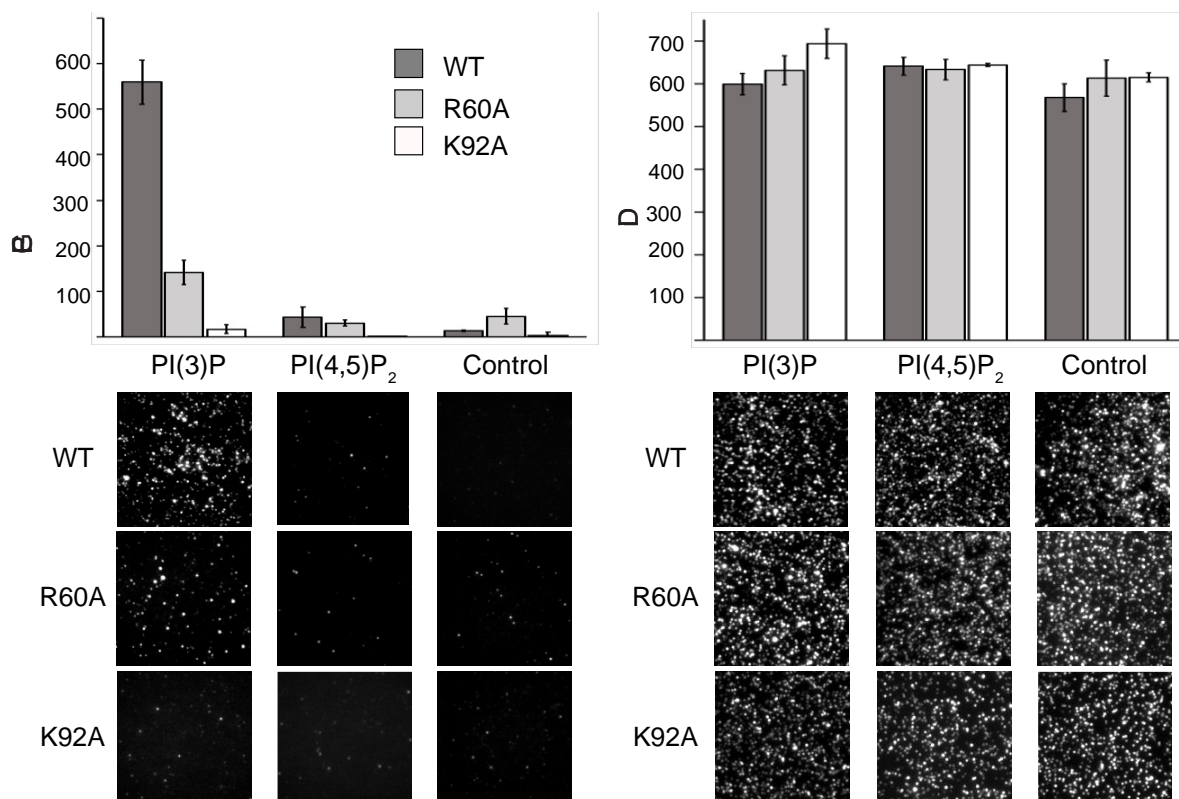


Figure 2.3 Loss of PI(3)P binding by point mutants of EGFP-p40PhoxPX

EGFP-p40PhoxPX and mutants were expressed in HEK293 cells, and cell lysates (5 nM EGFP-fusion) were subjected to SiMPull assay with PI(3)P, PI(4,5)P₂ and PC (control) vesicles. EGFP and DiD (vesicle) spots were quantified as described in Figure 1 legend. The average results of 3 independent experiments are shown with error bars representing SEM.

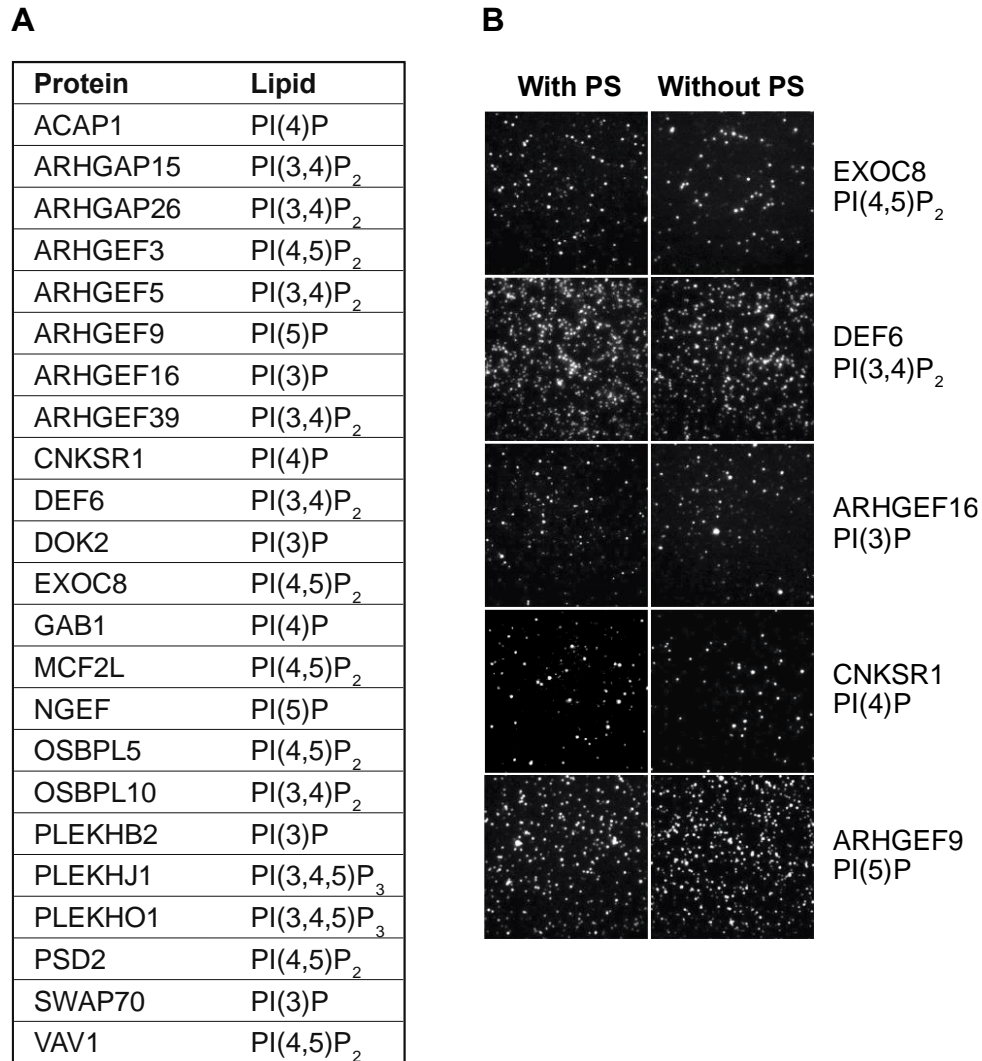


Figure 2.4 The presence of PS in lipid vesicles does not affect PIP binding by proteins.

(A) SiMPull was performed for 23 proteins against lipid vesicles containing the indicated PIPs, with or without PS and cholesterol. Binding was observed for all with both types of vesicles. For each condition at least two independent experiments were performed with similar results. Data were analyzed as described for Table 1 and Supplementary Table 1.

(B) Representative TIRF images of EGFP pull-down for 5 proteins that are also shown in Figure 1D, each by vesicles with or without PS.

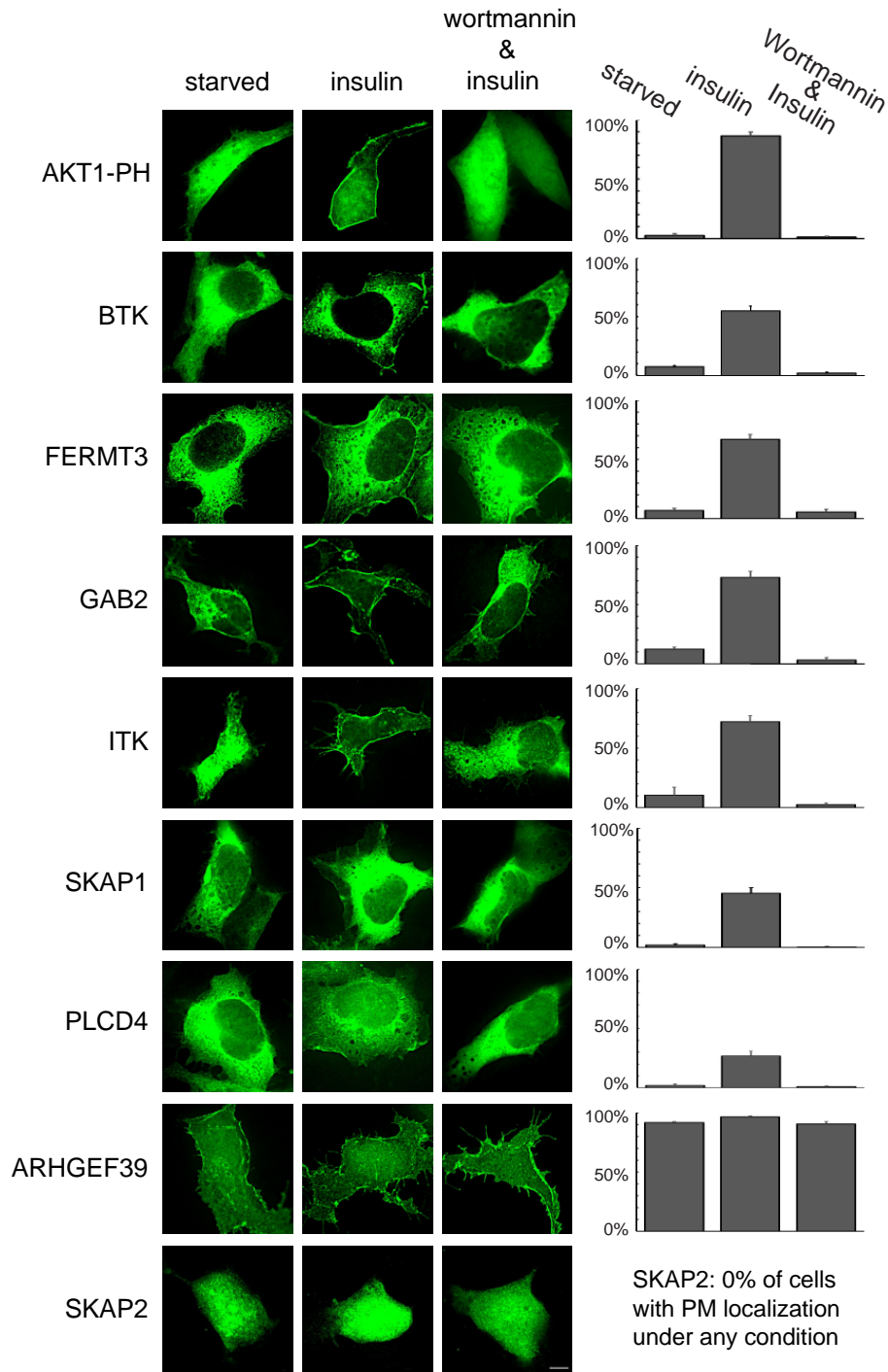


Figure 2.5 Membrane translocation of PIP₃-binding proteins.

HEK293 cells were transiently transfected with various EGFP-tagged proteins, followed by serum starvation overnight and then stimulation with 200 nM insulin for 5 min, with or without pretreatment by 100 nM wortmannin for 10 min. The cells were fixed for imaging. Representative images are shown. Scale bar: 5 μ m. The percentage of transfected cells with PM localization was scored, and the average results of 3 independent experiments are shown with error bars representing standard deviation.

Table 2.1

Group	Protein name	Entrez gene ID	Lipid bound in SiMPull	Reported lipid binding
I	BTK	695	PIP ₃	PIP ₃
	FERMT3 ^a	83706	PI(3,4)P ₂ , PIP ₃	PIP ₃
	GAB2 ^a	9846	PI(3,4)P ₂ , PIP ₃	PIP ₃
	ITK	3702	PIP ₃	PIP ₃
	PLEKHA1	59338	PI(3,4)P ₂	PI(3,4)P ₂
	PLEKHF1	79156	PI(3)P	PI(3)P
	SKAP1	8631	PIP ₃	PIP ₃
	SKAP2 ^{b,c}	8935	PIP ₃	PI(3,4)P ₂ , PIP ₃
II	ACAP1	9744	PI(4)P, PI(3,4)P ₂ , PI(3,5)P ₂	all PIPs
	ADAP1	11033	PI(3,4)P ₂ , PI(4,5)P ₂ , PIP ₃	PI(3,4)P ₂ , PIP ₃
	AFAP1L1	134265	PI(3,5)P ₂	PIP ₃
	ARHGAP12 ^b	94134	None/PI(4,5)P ₂	PI(4,5)P ₂ , PIP ₃
	ARHGAP15	9938	PI(3)P, PI(4)P, PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃	Not reported
	ARHGAP26	23092	PI(3)P, PI(3,4)P ₂	PI(4,5)P ₂
	ARHGEF3	50650	PI(4,5)P ₂ , PI(3,5)P ₂	Not reported
	ARHGEF5	7984	PI(3,4)P ₂ , PI(4,5)P ₂	PI(3)P, PI(4)P, PI(5)P
	ARHGEF7	8874	PI(3,5)P ₂	Not reported
	ARHGEF9	23229	PI(5)P	PI(3)P
	ARHGEF16	27237	PI(3)P	Not reported
	ARHGEF39	84904	PI(3,4)P ₂ , PIP ₃	Not reported
	CNKS1R	10256	PI(4)P	all PIPs
	DEF6	50619	PI(3,4)P ₂	PIP ₃
	DOK2	9046	PI(3)P	PI(3)P, PI(4)P, PI(5)P
	EXOC8	149371	PI(4,5)P ₂	Not reported
	GAB1	2549	PI(4)P	PIP ₃
	MCF2L	23263	PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃	PI(4,5)P ₂
	NGEF	25791	PI(5)P	Not reported
	OSBPL5	114879	PI(4)P, PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂	PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃
	OSBPL10	114884	PI(4)P, PI(3,4)P ₂ , PI(3,5)P ₂	PI(3)P, PI(4)P, PI(5)P, PA
	PLCD4 ^b	84812	None/PIP ₃	PI(4,5)P ₂ , PIP ₃
	PLEKHB2	55041	PI(3)P	PI(3,4)P ₂ , PIP ₃ , PS
	PLEKHJ1	55111	PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃	Not reported
	PSD2	84249	PI(4,5)P ₂	Not reported
	SBF2 ^b	81846	None/ PI(4,5)P ₂ , PIP ₃	PIP ₃
	SWAP70	23075	PI(3)P	PI(3)P, PI(4)P, PI(3,4)P ₂ , PIP ₃
	VAV1	7409	PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃	PI(4,5)P ₂ , PIP ₃

Table 2.1 (cont.)

Group	Protein name	Entrez gene ID	Lipid bound in SiMPull	Reported lipid binding
III	APPL1	26060	None	PI(3)P, PI(4)P, PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂
	APPL2	55198	None	PI(3)P, PI(4)P, PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂
	DOK1	1796	None	PI(5)P, PI(3,4)P ₂ , PI(4,5)P ₂ , PIP ₃
	FERMT2	10979	None	PI(3,5)P ₂ , PIP ₃
	GRB14	2888	None	all PIPs
	GRB7	2886	None	PI(3)P, PI(4)P, PI(5)P, PI(3,5)P ₂
	GRK2	156	None	PI(4,5)P ₂
	OSBPL8	114882	None	PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PIP ₃
	PLEK	5341	None	PI(4,5)P ₂
	PLEKHF2	79666	None	PI(3)P
	PLEKHN1	84069	None	PI(4)P, Cardiolipin, PS, PA
IV	AFAP1	60312	None	Not reported
	ARHGAP25	9938	None	Not reported
	BMX	660	None	Not reported
	DOK3	79930	None	Not reported
	DOK4	55715	None	Not reported
	FGD5	152273	None	Not reported
	KIF1B	23095	None	Not reported
	PHETA2	150368	None	Not reported
	PLCG2	5336	None	Not reported
	PLEKHO2	80301	None	Not reported
	PRKD3	23683	None	Not reported
	PSD4	23550	None	Not reported
	RASA1	5921	None	Not reported
	RASGRF1	5923	None	Not reported
	SOS2	6655	None	Not reported
	STAP1	26228	None	Not reported
Non-specific	CERT1	10087	all vesicles tested	PI(4)P, PI(4,5)P ₂
	NET1	10276	all vesicles tested	Not reported
	PLEKHO1	51177	PI(3)P, PI(5)P, PI(3,4)P ₂ , PI(4,5)P ₂ , PIP ₃ , PA	all PIPs
	SPATA13	221178	all vesicles tested	Not reported

Table 2.1 Summary of lipid-SiMPull assay results for 67 human PH domain-containing proteins.

Three independent experiments were performed for each protein; selected references are provided in Table 3. Based on our assay results and

Table 2.1 (cont.)

information from the literature, the 67 proteins are grouped as follows: **Group I**, proteins previously reported to bind PIP and binding confirmed by SiMPull; **Group II**, proteins found to bind PIP in SiMPull but not previously reported or previously reported to have different binding specificity; **Group III**, proteins previously reported to bind PIP but binding not observed in SiMPull; **Group IV**, proteins not reported to bind PIP and not found to bind PIP in SiMPull. ^a PI(3,4)P₂ had not been examined in the published work. ^b Binding in SiMPull was only observed with 20% PIP. ^c Binding to 20% PI(3,4)P₂ was not examine.

Table 2.2

PH Domain	Lipid binding by PH domain	Lipid binding by full-length protein
ACAP1-PH	PI(3,4)P ₂	PI(3,4)P ₂ , PI(4)P
ADAP1-PH1	None	PI(3,4,5)P ₃ , PI(4,5)P ₂ , PI(3,4)P ₂
ADAP1-PH2	None	PI(3,4,5)P ₃ , PI(4,5)P ₂ , PI(3,4)P ₂
ARHGEF3-PH	None	PI(4,5)P ₂ , PI(3,5)P ₂
ARHGEF5-PH	None	PI(4,5)P ₂ , PI(3,4)P ₂
ARHGEF9-PH	None	PI(5)P
DOK2-PH	PI(3,4,5)P ₃ , PI(3)P, PI(4)P	PI(3)P
FERMT3-PH	None	PI(3,4,5)P ₃ , PI(3,4)P ₂
PLEKHA1-PH1	None	PI(3,4)P ₂
PLEKHA1-PH2	PI(3,4)P ₂	PI(3,4)P ₂
SPATA13-PH	None	Bound all lipid vesicles
VAV1-PH	None	PI(3,4,5)P ₃ , PI(4,5)P ₂ , PI(3,4)P ₂ , PI(3,5)P ₂ , PI(5)P

Table 2.2 SiMPull assay results for selected PH domains. Experimental details and data processing were as described for Table 1 and Supplementary Table 1. At least three independent experiments were performed with similar outcome for each PH domain.

Table 2.3

Protein	ENTREZ GENE ID	Reported Lipid Binding by PH or protein	Reference
Proteins that bound lipids in SiMPull			
ACAP1	9744	all PIPs	(29)
ADAP1	11033	PI(3,4)P ₂ , PI(3,4,5)P ₃	(30)
AFAP1L1	134265	PI(3,4,5)P ₃	(31)
ARHGAP12 *	94134	PI(4,5)P ₂ , PI(3,4,5)P ₃	(32)
ARHGAP26	23092	PI(4,5)P ₂	(33)
ARHGEF5	7984	PI(3)P, PI(4)P, PI(5)P	(34)
ARHGEF9	23229	PI(3)P	(35)
BTK	695	PI(3,4,5)P ₃	(36)
CERT1	10087	PI(4)P, PI(4,5)P ₂	(37)
CNKSR1	10256	weak promiscuous binding for all PIPs.	(24)
DEF6	50619	PI(3,4,5)P ₃	(38)
DOK2	9046	PI(3)P, PI(4)P, PI(5)P	(39)
FERMT3	83706	PI(3,4,5)P ₃	(23)
GAB1	2549	PI(3,4,5)P ₃	(40)
GAB2	9846	PI(3,4,5)P ₃	(41)
ITK	3702	PI(3,4,5)P ₃	(42)
MCF2L	23263	PI(4,5)P ₂	(43)
OSBPL5	114879	PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PI(3,4,5)P ₃	(44)
OSBPL10	114884	PI(3)P, PI(4)P	(45)
PLCD4 *	84812	PI(4,5)P ₂ , PI(3,4,5)P ₃	(46)
PLEKHA1	59338	PI(3,4)P ₂	(47)
PLEKHB2	55041	PI(3,4)P ₂ , PI(3,4,5)P ₃	(47)
PLEKHF1	79156	PI(3)P	(48)
PLEKHO1	51177	promiscuous binding	(49)
SBF2 *	81846	PI(3,4,5)P ₃	(50)
SKAP1	8631	PI(3,4,5)P ₃	(51)
SKAP2 *	8935	PI(3,4)P ₂ , PI(3,4,5)P ₃	(52)
SWAP70	23075	PI(3,4,5)P ₃	(53)

Table 2.3 (cont.)

VAV1	7409	1) PI(4,5)P ₂ , PI(3,4,5)P ₃ 2) PI(3)P, PI(4)P, PI(5)P	(54, 55)
Proteins that did not bind lipids in SiMPull			
APPL1	26060	PI(3)P, PI(4)P, PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂	(56)
APPL2	55198	PI(3)P, PI(4)P, PI(5)P, PI(3,4)P ₂ , PI(3,5)P ₂	(56)
DOK1	1796	PI(5)P	(39)
FERMT2	10979	PI(3,5)P ₂ , PI(3,4,5)P ₃	(57)
GRB7	2886	PI(5)P, PI(3)P, PI(4)P, PI(3,5)P ₂	(58)
GRB14	2888	promiscuous binding	(59)
GRK2	156	PI(4,5)P ₂	(8)
OSBPL8	114882	PI(3,4)P ₂ , PI(3,5)P ₂ , PI(4,5)P ₂ , PI(3,4,5)P ₃	(44)
PLEK	5341	PI(4,5)P ₂	(8)
PLEKHF2	79666	PI(3)P	(60)
PLEKHN1	84069	PA, Cardiolipin, PS and PI(4)P	(61)

Table 2.3 References for reported lipid binding by proteins investigated in this study

In most cases the first report in the literature is cited here. Additional reports are cited when different conclusions are made. Some reports only examined the PH domains whereas others examined the full-length proteins.

* Bound 20% PIP vesicles in SiMPull assays

2.6 References

1. T. Balla, Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093-2104 (2005).
2. G. Di Paolo, P. De Camilli, Phosphoinositides in cell regulation and membrane dynamics. *Nature* **443**, 651-657 (2006).
3. K. Moravcevic, C. L. Oxley, M. A. Lemmon, Conditional peripheral membrane proteins: facing up to limited specificity. *Structure* **20**, 15-27 (2012).
4. M. A. Lemmon, Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99-111 (2008).
5. R. J. Haslam, H. B. Koide, B. A. Hemmings, Pleckstrin domain homology. *Nature* **363**, 309-310 (1993).
6. B. J. Mayer, R. Ren, K. L. Clark, D. Baltimore, A putative modular domain present in diverse signaling proteins. *Cell* **73**, 629-630 (1993).
7. H. S. Yoon, P. J. Hajduk, A. M. Petros, E. T. Olejniczak, R. P. Meadows, S. W. Fesik, Solution structure of a pleckstrin-homology domain. *Nature* **369**, 672-675 (1994).
8. J. E. Harlan, P. J. Hajduk, H. S. Yoon, S. W. Fesik, Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate. *Nature* **371**, 168-170 (1994).
9. M. A. Lemmon, K. M. Ferguson, R. O'Brien, P. B. Sigler, J. Schlessinger, Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. *Proc Natl Acad Sci U S A* **92**, 10472-10476 (1995).
10. S. R. James, C. P. Downes, R. Gigg, S. J. Grove, A. B. Holmes, D. R. Alessi, Specific binding of the Akt-1 protein kinase to phosphatidylinositol 3,4,5-trisphosphate without subsequent activation. *Biochem J* **315 (Pt 3)**, 709-713 (1996).
11. S. J. Watton, J. Downward, Akt/PKB localisation and 3' phosphoinositide generation at sites of epithelial cell-matrix and cell-cell interaction. *Current Biology* **9**, 433-436 (1999).
12. P. Varnai, T. Balla, Visualization of phosphoinositides that bind pleckstrin homology domains: calcium- and agonist-induced dynamic changes and relationship to myo-[3H]inositol-labeled phosphoinositide pools. *J Cell Biol* **143**, 501-510 (1998).
13. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. DeWald, D. Murray, S. D. Emr, M. A. Lemmon, Genome-wide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol Cell* **13**, 677-688 (2004).
14. W. S. Park, W. D. Heo, J. H. Whalen, N. A. O'Rourke, H. M. Bryan, T. Meyer, M. N. Teruel, Comprehensive identification of PIP3-regulated PH domains from *C. elegans* to *H. sapiens* by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).
15. M. A. Lemmon, Pleckstrin homology domains: not just for phosphoinositides. *Biochem Soc Trans* **32**, 707-711 (2004).
16. E. Arauz, V. Aggarwal, A. Jain, T. Ha, J. Chen, Single-Molecule Analysis of Lipid-Protein Interactions in Crude Cell Lysates. *Analytical chemistry* **88**, 4269-4276 (PMC4863138) (2016).
17. F. Kanai, H. Liu, S. J. Field, H. Akbary, T. Matsuo, G. E. Brown, L. C. Cantley, M. B. Yaffe, The PX domains of p47phox and p40phox bind to lipid products of PI(3)K. *Nat Cell Biol* **3**, 675-678 (2001).

18. A. Jain, R. Liu, B. Ramani, E. Arauz, Y. Ishitsuka, K. Ragunathan, J. Park, J. Chen, Y. K. Xiang, T. Ha, Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484-488. (PMC3103084) (2011).
19. D. A. Agard, Y. Hiraoka, P. Shaw, J. W. Sedat, Fluorescence microscopy in three dimensions. *Methods Cell Biol.* **30**, 353-377. (1989).
20. G. van den Bogaart, K. Meyenberg, H. J. Risselada, H. Amin, K. I. Willig, B. E. Hubrich, M. Dier, S. W. Hell, H. Grubmuller, U. Diederichsen, R. Jahn, Membrane protein sequestering by ionic protein-lipid interactions. *Nature* **479**, 552-555 (2011).
21. J. Bravo, D. Karathanassis, C. M. Pacold, M. E. Pacold, C. D. Ellson, K. E. Anderson, P. J. G. Butler, I. Lavenir, O. Perisic, P. T. Hawkins, L. Stephens, R. L. Williams, The Crystal Structure of the PX Domain from p40phox Bound to Phosphatidylinositol 3-Phosphate. *Molecular Cell* **8**, 829-839 (2001).
22. I. Vonkova, A. E. Saliba, S. Deghou, K. Anand, S. Ceschia, T. Doerks, A. Galih, K. G. Kugler, K. Maeda, V. Rybin, V. van Noort, J. Ellenberg, P. Bork, A. C. Gavin, Lipid Cooperativity as a General Membrane-Recruitment Principle for PH Domains. *Cell Rep* **12**, 1519-1530 (2015).
23. R. Hart, P. Stanley, P. Chakravarty, N. Hogg, The kindlin 3 pleckstrin homology domain has an essential role in lymphocyte function-associated antigen 1 (LFA-1) integrin-mediated B cell adhesion and migration. *J Biol Chem* **288**, 14852-14862 (2013).
24. A. B. Jaffe, P. Aspenstrom, A. Hall, Human CNK1 acts as a scaffold protein, linking Rho and Ras signal transduction pathways. *Mol Cell Biol* **24**, 1736-1746 (2004).
25. J. Lim, M. Zhou, T. D. Veenstra, D. K. Morrison, The CNK1 scaffold binds cytohesins and promotes insulin pathway signaling. *Genes Dev* **24**, 1496-1506 (2010).
26. T. F. Martin, PI(4,5)P(2)-binding effector proteins for vesicle exocytosis. *Biochim Biophys Acta* **1851**, 785-793 (2015).
27. D. R. Cook, K. L. Rossman, C. J. Der, Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. *Oncogene* **33**, 4021-4035 (2014).
28. K. L. Rossman, C. J. Der, J. Sondek, GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat Rev Mol Cell Biol* **6**, 167-180 (2005).
29. N. Shinozaki-Narikawa, T. Kodama, Y. Shibasaki, Cooperation of phosphoinositides and BAR domain proteins in endosomal tubulation. *Traffic* **7**, 1539-1550 (2006).
30. K. Venkateswarlu, P. J. Cullen, Molecular cloning and functional characterization of a human homologue of centaurin-alpha. *Biochem Biophys Res Commun* **262**, 237-244 (1999).
31. S. R. Tie, D. J. McCarthy, T. S. Kendrick, A. Louw, C. Le, J. Satiaputra, N. Kucera, M. Phillips, E. Ingley, Regulation of sarcoma cell migration, invasion and invadopodia formation by AFAP1L1 through a phosphotyrosine-dependent pathway. *Oncogene* **35**, 2098-2111 (2016).
32. D. J. Bae, J. Seo, S. Y. Kim, S. Y. Park, J. Do Yoo, J. H. Pyo, W. Cho, J. Y. Cho, S. Kim, I. S. Kim, ArhGAP12 plays dual roles in Stabilin-2 mediated efferocytosis: Regulates Rac1 basal activity and spatiotemporally turns off the Rac1 to orchestrate phagosome maturation. *Biochim Biophys Acta Mol Cell Res* **1866**, 1595-1607 (2019).

33. R. Lundmark, G. J. Doherty, M. T. Howes, K. Cortese, Y. Vallis, R. G. Parton, H. T. McMahon, The GTPase-activating protein GRAF1 regulates the CLIC/GEEC endocytic pathway. *Curr Biol* **18**, 1802-1808 (2008).
34. M. Kuroiwa, C. Oneyama, S. Nada, M. Okada, The guanine nucleotide exchange factor Arhgef5 plays crucial roles in Src-induced podosome formation. *J Cell Sci* **124**, 1726-1738 (2011).
35. V. M. Kalscheuer, L. Musante, C. Fang, K. Hoffmann, C. Fuchs, E. Carta, E. Deas, K. Venkateswarlu, C. Menzel, R. Ullmann, N. Tommerup, L. Dalpra, A. Tzschach, A. Selicorni, B. Luscher, H. H. Ropers, K. Harvey, R. J. Harvey, A balanced chromosomal translocation disrupting ARHGEF9 is associated with epilepsy, anxiety, aggression, and mental retardation. *Hum Mutat* **30**, 61-68 (2009).
36. K. Salim, M. J. Bottomley, E. Querfurth, M. J. Zvelebil, I. Gout, R. Scaife, R. L. Margolis, R. Gigg, C. I. Smith, P. C. Driscoll, M. D. Waterfield, G. Panayotou, Distinct specificity in the recognition of phosphoinositides by the pleckstrin homology domains of dynamin and Bruton's tyrosine kinase. *The EMBO Journal* **15**, 6241-6250 (1996).
37. T. P. Levine, S. Munro, Targeting of Golgi-Specific Pleckstrin Homology Domains Involves Both PtdIns 4-Kinase-Dependent and -Independent Components. *Current Biology* **12**, 695-704 (2002).
38. S. Gupta, J. C. Fanzo, C. Hu, D. Cox, S. Y. Jang, A. E. Lee, S. Greenberg, A. B. Pernis, T cell receptor engagement leads to the recruitment of IBP, a novel guanine nucleotide exchange factor, to the immunological synapse. *J Biol Chem* **278**, 43541-43549 (2003).
39. G. Guittard, A. Gerard, S. Dupuis-Coronas, H. Tronchere, E. Mortier, C. Favre, D. Olive, P. Zimmermann, B. Payrastre, J. A. Nunes, Cutting edge: Dok-1 and Dok-2 adaptor molecules are regulated by phosphatidylinositol 5-phosphate production in T cells. *J Immunol* **182**, 3974-3978 (2009).
40. C. R. Maroun, D. K. Moscatello, M. A. Naujokas, M. Holgado-Madruga, A. J. Wong, M. Park, A conserved inositol phospholipid binding site within the pleckstrin homology domain of the Gab1 docking protein is required for epithelial morphogenesis. *J Biol Chem* **274**, 31719-31726 (1999).
41. C. E. Edmead, B. C. Fox, C. Stace, N. Ktistakis, M. J. Welham, The pleckstrin homology domain of Gab-2 is required for optimal interleukin-3 signalsome-mediated responses. *Cell Signal* **18**, 1147-1155 (2006).
42. Y. H. Huang, J. A. Grasis, A. T. Miller, R. Xu, S. Soonthornvacharin, A. H. Andreotti, C. D. Tsoukas, M. P. Cooke, K. Sauer, Positive regulation of Itk PH domain function by soluble IP4. *Science* **316**, 886-889 (2007).
43. J. T. Snyder, K. L. Rossman, M. A. Baumeister, W. M. Pruitt, D. P. Siderovski, C. J. Der, M. A. Lemmon, J. Sondek, Quantitative analysis of the effect of phosphoinositide interactions on the function of Dbl family proteins. *J Biol Chem* **276**, 45868-45875 (2001).
44. R. Ghai, X. Du, H. Wang, J. Dong, C. Ferguson, A. J. Brown, R. G. Parton, J. W. Wu, H. Yang, ORP5 and ORP8 bind phosphatidylinositol-4, 5-bisphosphate (PtdIns(4,5)P₂) and regulate its level at the plasma membrane. *Nat Commun* **8**, 757 (2017).
45. E. Nissila, Y. Ohsaki, M. Weber-Boyvat, J. Perttila, E. Ikonen, V. M. Olkkonen, ORP10, a cholesterol binding protein associated with microtubules, regulates apolipoprotein B-100 secretion. *Biochim Biophys Acta* **1821**, 1472-1484 (2012).
46. K. Nagano, K. Fukami, T. Minagawa, Y. Watanabe, C. Ozaki, T. Takenawa, A novel phospholipase C delta4 (PLCdelta4) splice variant as a negative regulator of PLC. *J Biol Chem* **274**, 2872-2879 (1999).

47. S. Dowler, R. A. Currie, D. G. Campbell, M. Deak, G. Kular, C. P. Downes, D. R. Alessi, Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities. *Biochem J* **351**, 19-31 (2000).
48. W. J. Lin, C. Y. Yang, L. L. Li, Y. H. Yi, K. W. Chen, Y. C. Lin, C. C. Liu, C. H. Lin, Lysosomal targeting of phafin1 mediated by Rab7 induces autophagosome formation. *Biochem Biophys Res Commun* **417**, 35-42 (2012).
49. M. E. Olsten, D. A. Canton, C. Zhang, P. A. Walton, D. W. Litchfield, The Pleckstrin homology domain of CK2 interacting protein-1 is required for interactions and recruitment of protein kinase CK2 to the plasma membrane. *J Biol Chem* **279**, 42114-42127 (2004).
50. P. Berger, I. Berger, C. Schaffitzel, K. Tersar, B. Volkmer, U. Suter, Multi-level regulation of myotubularin-related protein-2 phosphatase activity by myotubularin-related protein-13/set-binding factor-2. *Hum Mol Genet* **15**, 569-579 (2006).
51. A. Witte, B. Meineke, J. Sticht, L. Philipsen, B. Kuropka, A. J. Muller, C. Freund, B. Schraven, S. Kliche, D120 and K152 within the PH Domain of T Cell Adapter SKAP55 Regulate Plasma Membrane Targeting of SKAP55 and LFA-1 Affinity Modulation in Human T Lymphocytes. *Mol Cell Biol* **37**, (2017).
52. K. D. Swanson, Y. Tang, D. F. Ceccarelli, F. Poy, J. P. Sliwa, B. G. Neel, M. J. Eck, The Skap-hom dimerization and PH domains comprise a 3'-phosphoinositide-gated molecular switch. *Mol Cell* **32**, 564-575 (2008).
53. M. Shinohara, Y. Terada, A. Iwamatsu, A. Shinohara, N. Mochizuki, M. Higuchi, Y. Gotoh, S. Ihara, S. Nagata, H. Itoh, Y. Fukui, R. Jessberger, SWAP-70 is a guanine-nucleotide-exchange factor that mediates signalling of membrane ruffling. *Nature* **416**, 759-763 (2002).
54. J. Han, K. Luby-Phelps, B. Das, X. Shu, Y. Xia, R. D. Mosteller, U. M. Krishna, J. R. Falck, M. A. White, D. Broek, Role of substrates and products of PI 3-kinase in regulating activation of Rac-related guanosine triphosphatases by Vav. *Science* **279**, 558-560 (1998).
55. S. Rodriguez-Fdez, C. Citterio, L. F. Lorenzo-Martin, J. Baltanas-Copado, C. Llorente-Gonzalez, S. Corbalan-Garcia, M. Vicente-Manzanares, X. R. Bustelo, Phosphatidylinositol Monophosphates Regulate Optimal Vav1 Signaling Output. *Cells* **8**, (2019).
56. H. J. Chial, R. Wu, C. V. Ustach, L. C. McPhail, W. C. Mobley, Y. Q. Chen, Membrane targeting by APPL1 and APPL2: dynamic scaffolds that oligomerize and bind phosphoinositides. *Traffic* **9**, 215-229 (2008).
57. H. Qu, Y. Tu, X. Shi, H. Larjava, M. A. Saleem, S. J. Shattil, K. Fukuda, J. Qin, M. Kretzler, C. Wu, Kindlin-2 regulates podocyte adhesion and fibronectin matrix deposition through interactions with phosphoinositides and integrins. *J Cell Sci* **124**, 879-891 (2011).
58. T. L. Shen, D. C. Han, J. L. Guan, Association of Grb7 with phosphoinositides and its role in the regulation of cell migration. *J Biol Chem* **277**, 29069-29077 (2002).
59. R. V. Rajala, M. D. Chan, A. Rajala, Lipid-protein interactions of growth factor receptor-bound protein 14 in insulin receptor signaling. *Biochemistry* **44**, 15461-15471 (2005).
60. W. J. Lin, C. Y. Yang, Y. C. Lin, M. C. Tsai, C. W. Yang, C. Y. Tung, P. Y. Ho, F. J. Kao, C. H. Lin, Phafin2 modulates the structure and function of endosomes by a Rab5-dependent mechanism. *Biochem Biophys Res Commun* **391**, 1043-1048 (2010).
61. E. Sano, S. Shono, K. Tashiro, H. Konishi, E. Yamauchi, H. Taniguchi, Novel tyrosine phosphorylated and cardiolipin-binding protein CLPABP functions as mitochondrial RNA granule. *Biochim Biophys Acta* **1783**, 1036-1047 (2008).

CHAPTER 3. PREDICTING PHOSPHOINOSITIDE BINDING BY PH DOMAINS

3.1 Introduction

Human proteome consists of over 250 PH domain containing proteins. Despite their identification as phosphoinositide (PIP) binding proteins almost 25 years ago (1, 2), the lipid binding properties for the majority of the family is not known. There are multiple challenges in the way to achieve a comprehensive understanding of lipid binding by the PH domain family. One of the major hurdles being that commonly used methods are often artifact prone such as lipid-strips and often rely on isolated PH domains instead of full-length proteins. Utilizing full-length proteins is often limited by feasibility of protein purification and associated functional characterization of purified proteins. This is compounded by a severe lack of tools for *in vivo* characterization of binding with different PIP moieties. Additionally, low sequence similarity among PH domain proteins is a major hurdle towards finding a consensus PIP binding motif.

Over the years, researcher have attempted to characterized the PIP binding of large number of PH domains (3-5), using either isolated PH domains or focusing on smaller number of PH domains found in Yeast. The structural and experimental observations suggest that the presence of a basic “KX_n(K/R)XR” motif in β 1- β 2 loop of PH domain plays the most prominent role in PIP binding (6). One of the early reports (7) proposed four major criteria to assess PIP₃ binding by the PH domains, the most important criteria being a β 1- β 2 loop consisting of 6 or more amino acids, with at least one lysine or arginine residue within the loop along with lysine or arginine residues at the +2 and +4 position from the beginning of this loop. Although, this criteria when applied to observed PIP₃ binding by mouse PH domains (4) did not successfully distinguish between

the binding and non-binding proteins. An unbiased probabilistic sequence comparison strategy to discern the binding and non-binding proteins was developed by Park et al., (4), which assumes that all amino acids in the PH domain contribute independently towards PIP₃ binding. The algorithm called recursive functional classification (RFC) was able to cleanly separate the binding and non-binding groups and correctly predict PIP₃ binding by the PH domains from different organisms. This analysis suggested that the entire PH domain sequence contributes towards the lipid binding properties.

While previous effort was limited to identifying PIP₃ binding PH domains, taking advantage of results from our lipid-SiMPull assay, we employed the RFC approach to predict PIP binding by the PH domain proteins in human proteome. Our RFC analysis predicts that almost half of the PH domains in human proteome should bind with PIPs with some specificity. Further, we randomly selected 20 PH domains predicted to bind PIPs based on cDNA availability and successful cloning and expression, and 19 of these successfully bound either PIP₃, PI(4,5)P₂ or PI(3,4)P₂.

3.2 Material and Methods

3.2.1 Cell Culture

See 2.2.1

3.2.2 Cell Lysis and Western Blotting

See 2.2.2

3.2.3 PH domain Sequence Alignment

The list of human PH domain-containing proteins was obtained from the Pfam database and individual PH domain sequences were acquired from UniProt. After initial

alignment of all sequences using PROMALS3D with default parameters, about 30 sequences were removed due to large insertions, and the remaining 246 PH domains were re-aligned. The amino acids aligned at 297 positions.

3.2.4 RFC Matrix Generation and PIP-binding Prediction

The recursive-learning strategy described by Park et al. (4) was followed, with the details specific to our study described here. MATLAB was used to create two 297 x 20 probability matrices for PIP-binding (P^B) and non-binding proteins (P^{NB}), respectively. For each position, we calculated the probability of individual amino acids in their respective groups. An amino acid was scored in the probability matrix if it was found in at least one PH domain in the group. The RFC matrix was then calculated by taking the log of the ratio of P^B to P^{NB} . For scoring a PH domain, the values in the RFC matrix for matching amino acids at all positions were summed. In Figure 3.3 B, the value at each amino acid position is the average of sum-square of values in the corresponding position. For determining the cutoff S_{RFC} value, the 246 PH domain sequences were randomly scrambled 10,000 times using MATLAB, followed by calculation of S_{RFC} value for each scramble. The distribution of mean of all scrambled sets were used to determine the cutoff for PIP binding.

3.2.5 Cloning of Full-Length PH Domain Proteins

Plasmids selected for validation of RFC predictions were obtained from DNASU plasmid library. The plasmids were cloned individually in BamHI/EcoRI digested pcDNA3-EGFP vector through Gibson assembly. Individual clones were confirmed by sequencing and western blotting against EGFP antibody.

3.2.6 Lipid Vesicle Preparation and Single-Molecule Pulldown Assay

See 2.2.4 and 2.2.5

3.3 Results

3.3.1 A Recursive-Learning Algorithm for Probabilistic Sequence Comparison Based on Simpull Data

Although our survey of 67 proteins covered only 1/4 of the PH domain-containing protein family in the human genome, we reasoned that our results could have predictive power for the rest of the family. Given the high likelihood that the PH domains mediate PIP binding of the full-length proteins, we compared the sequences of the PH domains of the 67 proteins we studied plus AKT1-PH as a positive control, excluding the 4 proteins that bound lipids non-specifically. We also excluded two proteins (ADAP1 and AFAP1L1) that each contained two PH domains and bound PIP as full-length proteins, because we did not know which PH domain contributed to the binding. This yielded a collection of 67 PH domains. Based on reported structural studies of the PH domain-PIP headgroup interaction, the loop between the first two β strands (β 1- β 2) with a basic sequence motif of “KX_n(K/R)XR” plays the most prominent role (6). However, we found the vast majority of the 67 PH domains to contain this motif, with no distinction between those that bound PIPs and those that did not (Figure 3.1). Therefore, we next considered the possibility that sequence features throughout the entire PH domain are determinants of PIP binding, as suggested by Park et al. for PIP₃-regulated PH domains (4).

To identify sequence signatures of PIP binding, we employed an unbiased probabilistic sequence comparison strategy developed by Park et al., called recursive functional classification (RFC) (4). This strategy relies on accurate sequence alignment of multiple PH domains, which can be a challenge given the low sequence similarity shared by PH domains despite the 3D structure conservation. We employed the PROMALS3D (PROfile Multiple Alignment with predicted Local Structures and 3D constraints) tool, which makes use of available 3D structural information to guide

multiple sequence alignments (8), to align the sequences of 246 PH domains from 222 human proteins in the Pfam and UniProt databases, including the 67 PH domains described above. We found a total of ~280 PH domains in the database, but excluded ~30 of them from the alignment due to their unusual sequence features that would have caused very large gaps in the alignment. Alignment of the 246 PH domains yielded 297 amino acid positions; the aligned sequences of 67 PH domains were used to generate the RFC matrix. Following the reported method (4) we created two probability matrices of 297 (positions) x 20 (amino acids): P^B for the PIP binding group (35 PH domains including AKT1-PH), and P^{NB} for the non-binding group (32 PH domains). Based on the assumption that each sequence position contributes independently to PIP binding, we created an RFC-matrix, with each element of the matrix calculated as $\log(P^B/P^{NB})$. In the RFC-matrix values > 0 represent positive correlation with PIP binding whereas values < 0 represent negative correlation. This RFC-matrix can be used to calculate the predicted PIP binding score (S_{RFC}) for any PH domain, as previously described (4).

3.3.2 Recursive Functional Classification Identifies Potential Amino Acid Determinants of PIP Binding Distributed Throughout the Entire PH Domain

Using the RFC-matrix created above, we first scored positions in the $\beta 1$ - $\beta 2$ loop and its flanking regions to calculate S_{RFC} . The resulting scores separated PIP-binding proteins from non-binding proteins without a large margin (Figure 3.2 A). When we expanded the positions to include $\beta 3$ - $\beta 4$ and $\beta 6$ - $\beta 7$ loops and their flanking regions, which are believed to also contribute to PH-PIP interactions (9), the separation of the two groups of proteins improved (Figure 3.2 A). However, when all positions of the PH domain were taken into calculation, the PIP-binding and non-binding groups were even more cleanly separated with markedly higher scores in both the positive and negative ranges (Figure 3.2 B). This outcome validates the idea that the entire PH domain may be involved in

determining whether a domain (or full-length protein) binds PIP. Interestingly, the relative contributions of various regions of the PH domain to PIP-binding vary drastically among those found to bind PIPs (Figure 3.2 B), suggesting that the structural basis of PH-PIP interactions may be diverse within this family of proteins.

We created a heat map (Figure 3.3 A) to represent the RFC-matrix for the 67 PH domains studied in SiMPull and aligned it to the sequence of AKT1-PH based on its crystal structure (PDB ID: 1UNP). Amino acids with high impact on lipid binding, either positively or negatively, were distributed throughout the PH domain. Following the reported method (4), we determined the overall contribution of each amino acid position to PIP binding, with 12 positions of the highest contribution to PIP binding marked (Figure 3.3 B). While some of those residues are located in the classically defined head group binding pocket, many are not. Other than those 12, additional positions spanning the entire PH domain are likely to also play important roles in PIP binding (Figure 3.3 B). It should be noted that this is a composite view, and that an individual PH domain would use only a subset of those residues for interaction with PIP (see Figure 3.2 B).

3.3.3 The RFC Algorithm Predicts 50% of Human PH Domain-Containing Proteins to Bind PIPs

Next, we applied the RFC algorithm obtained above to analyze the human PH domains not investigated by SiMPull. While in principle any protein with a positive S_{RFC} score can be considered to possibly bind PIP, we set out to establish a more stringent cut-off score for positive correlation with PIP binding by scrambling the sequences of all 246 PH domains and subjecting them to scoring by the same algorithm. Using mean + 3x SD, we arrived at 2.59 as the threshold S_{RFC} score for a PH domain to be considered likely to bind PIP (Figure 3.4). Of the 179 PH domains not tested by SiMPull, 86 (48%) scored

above the threshold (Figure 3.3 C / Table 3.1). Taking into consideration our assay data, the overall estimate is that 50% of the PH domains in the human genome bind PIPs.

To test the validity of the RFC algorithm, we performed SiMPull assays with predicted binders that have not been reported to bind any lipids to date. The availability of cDNA clones and confirmation of EGFP-fusion full-length protein expression led to the random selection of 20 proteins in that category (Figure 3.5). We first confirmed that none of the 20 proteins displayed nonspecific binding to PC. Because our goal was to validate PIP binding and the binding to any PIP would satisfy the criterion, we assayed the 20 proteins against one of the three PIPs frequently found to bind PH domains: PIP₃, PI(3,4)P₂, and PI(4,5)P₂. If binding was confirmed for one PIP, that protein was not assayed against other PIPs. As shown in Figure 3.5, 19 out of the 20 proteins were found to bind a PIP. The only exception was PSD3, which did not bind any of the three PIPs although it is still possible that it binds to a PIP not yet tested. These results validate the predictive power of our recursively learning approach and, furthermore, are consistent with the notion that the PH domain defines the observed PIP binding.

3.4 Discussion

A recursive-learning algorithm based on the results of the 67 proteins has predicted 48% of the remaining human PH domains in the family to bind PIPs with some specificity. The predictive power of this algorithm is confirmed by our finding that 19 out of 20 proteins predicted to bind PIP (but never reported to bind lipids) indeed display affinity for at least one PIP in SiMPull assays. Our work suggests that ~50% of all human PH domain proteins likely bind PIPs with specificity, in contrast to the previous estimate that only 10% of all PH domains bind PIPs with high affinity and specificity (10).

The recursive-learning algorithm developed by Park et al. (4) is an unbiased probabilistic sequence comparison strategy that is now facilitated by accurate multiple sequence alignment incorporating 3D structural information (8). The clean separation of PIP binding and nonbinding groups of PH domains by this algorithm and, more importantly, our validation of predicted novel PIP binding, attest to the power of the method. Our RFC analysis suggests that for some proteins PIP binding may involve regions of the PH domain outside of the previously known interacting site. Similar suggestions have been made by others for PIP₃ regulation (4) and organelle PIP binding (5) by PH domains. Interestingly, two most recent molecular dynamics simulation studies of the PH domain of GRP1 have revealed that there may be multiple PIP₃ binding sites – canonical and non-canonical – on this domain (11, 12). Future biochemical and structural studies to probe these putative interactions of novel modes will be informative.

3.5 Figures and Table

	$\beta 1$	$\beta 1$ - $\beta 2$ Variable loop	$\beta 2$		$\beta 1$	$\beta 1$ - $\beta 2$ Variable loop	$\beta 2$	
ACAP1	H L F K	R A S N	- - - - - A F K T	W S R R W	AFAP1-PH1	F L L R	K K R - - - - - F G Q W	T K L L C
ADAP1 PH1	L W K R	G R D -	- - - - - N G Q	F L S R K	APPL1	Y L N A	R N K - - - - - S S T W	D R Q F Y
ADAP1 PH2	Y M E K	T G P K	- - - - - Q T E G F	R K R W F	APPL2	Y L N I	R N K - - - - - T T T W	E R L Y F
AFAP1L1-PH1	F L L R	K K R -	- - - - - F G Q W	A K Q L T	ARHGAP25	L K K Q	R S I - - - - - V K N	W Q Q R Y
AFAP1L1-PH2	Y L N V	L V N -	- - - - - Q G W	K R R W C	BMX	L L L K	R S Q Q K K K M S	- - - - - P N N Y
AKT1	L H K R	G E Y -	- - - - - I K T	W R P P Y	DOK1	P L P L	Q S Q R F G	- - - - - T K R W
ARHGAP12	L L N V	T K I A E N G K	- - - - - K V R K N	W L S S W	DOK3	I L Y Q	Q H V K F G	- - - - - K K C W
ARHGAP15	Y L Q K	A K I A D G C K	- - - - - K L R K N	W S T S W	DOK4	Y V K M	K S R K -	- - - - - L G I Y
ARHGAP26	Y L Y V	Q E K R H -	- - - - - F G T S	W V K H Y	FERMT2	L K V F	K P K K L T	- - - - - L K G Y
ARHGEF3	E L K N	N R G -	- - - - - V	K L H V F	FGD5-PH2	Y L S R	C K R G -	- - - - - K R H W
ARHGEF5	E L T A	L E F S A S P G	- - - - - L R R K L N	T R P V H	GRB14	L E A K	E Q G -	- - - - - K K S
ARHGEF7	Q V L I	Q C A G -	- - - - - S E E K	N R K Y I	GRB7	F L Q L	R G S G -	- - - - - R K L W
ARHGEF9	E M A W	I Y Q P -	- - - - - Y G R	N Q Q R V	GRK2	S K M G	N P F -	- - - - - L T
ARHGEF16	E L F L	V E E T G L F -	- - - - - R K I A S	R P T C Y	KIF1B	Y L H F	K E P L -	- - - - - Y S N W
ARHGEF39	W L L V	V P P H -	- - - - - G E	P R P R M	OSBPL8	W L K I	R G T -	- - - - - L K S W
BTK	F L K R	S Q Q K K K T -	- - - - - S P L N	F K K R L	PHETA2	M G F L	- - - - -	- - - - -
CNKSR1	W L L L	R K A P G G F -	- - - - - M G P R	W R R R W	PLCG2	T V M T	V F S F R -	- - - - - K S T P
DEF6	Y L W K	R G H L -	- - - - - R R N	W A E R W	PLEK-PH1	L V K K	G S V -	- - - - - F N T W
DOK2	F L Y L	Q Q Q Q T -	- - - - - F G K K W R	R F G A S	PLEK-PH2	L L K Q	G H R R -	- - - - - R K N
EXOC8	D L V E	Y D A D H -	- - - - - M A Q	L Q R V H	PLEKHE2	V L T K	L C R -	- - - - - K K P
FERMT3	L R I F	R I P R R P R -	- - - - - K L T L K G Y	R Q H W V	PLEKHN1-PH1	R V Q L	R F Q H -	- - - - - S Q D V
GAB1	W L R K	S P P E K K L -	- - - - - K R Y A	W K R F W	PLEKHN1-PH2	R V K L	Q H L P -	- - - - - A Q E Q
GAB2	W L R K	S P P E K K L -	- - - - - R R Y A	W K R F W	PLEKHO2	W I K K	S S G G L -	- - - - - L G F W
ITK	Q L I K	K S Q Q K R R T -	- - - - - S P S N	F K V R F	PRKD3	V H Y T	S R -	- - - - - D N L R
MCF2L	S F S V	W T D E K R G H T K V K E L A R F K	- - - - - P M Q R H	F H E I Y	PSD4	I L A R	K M H Q D A D G K K	- - - - - K R G W
NGEF	E L Q Q	M S G P K T S R T -	- - - - - L R T K K L	W Q N R Y	RASA1	Y L L K	K G K G K -	- - - - - R W
OSBPL10	V L S K	Y T N L -	- - - - - L Q G	W Q N R Y	RASGRF1-PH1	Y L S K	R S S D -	- - - - - N T K W
OSBPL5	S L K I	R G T -	- - - - - L K S	W T K L W	SOS2	P L T R	I G -	- - - - - A K
PLCD4	P M R K	V R S -	- - - - - K S W	R K L Y I	STAP1	F L L I	K R S G -	- - - - - Y R E Y
PLEKHA1	C V K Q	G A V -	- - - - - M K N	W K R R Y				
PLEKHB2	L R Q S	T I -	- - - - - L K	W S K K N				
PLEKHF1	V L T K	E C R -	- - - - - K K	A K P R I				
PLEKHJ1	E L G M	R G P K -	- - - - - K G S V	L K R R L				
PSD2	V L T R	K T H A D M D G K R -	- - - - - T P R G R R G	W K K F Y				
SBF2	T L Y K	R G A L -	- - - - - L K G	W K P R W				
SKAP1	E K K S	K D H S F -	- - - - - F G S E W	Q K R W C				

(Bound Phosphoinositides in Lipid-SiMPull)

(Did not bind Phosphoinositides in Lipid-SiMPull)

(Bound Phosphoinositides in Lipid-SiMPull)

Figure 3.1 Both binding and non-binding proteins contain canonical PIP binding motif

Sequence alignment of PIP-binding (left) and non-binding (right) PH domains in the $\beta 1$ - $\beta 2$ region. Bold: PH domains with reported crystal structures. Red: residues conforming to the $KX_n(K/R)XR$ motif. Three PH domains (AFAP1-PH2, FGD5-PH1 and RASGRF1-PH2) were not shown in the alignment due to an unusually short $\beta 1$ sheet.

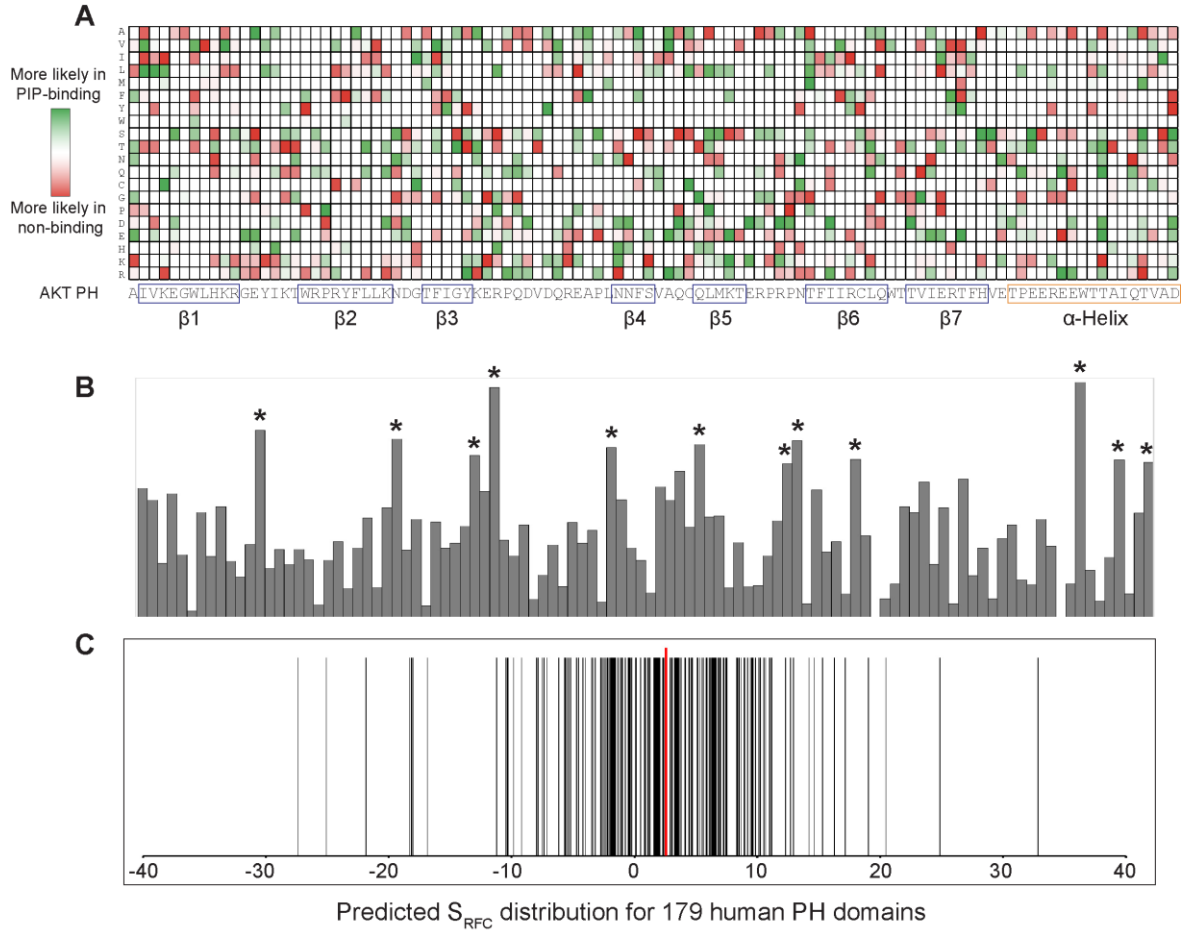


Figure 3.3 Recursive-learning algorithm identifies amino acid determinants of PIP-binding and predicts binding by human PH domains.

(A) A heat map showing the contribution of 20 amino acids at each position for the 67 PH domains. The amino acid positions are aligned to AKT1-PH domain based on its crystal structure (PDB ID: 1UNP).

(B) Contribution from different positions of PH domain towards PIP-binding was plotted from RFC-matrix. Twelve positions with the highest scores are marked by stars and aligned to the sequence in A.

(C) Prediction of PIP-binding for 179 human PH domains using the RFC-matrix. Each vertical black line represents one PH domain. The red line indicates the cutoff value (2.59) for PIP-binding.

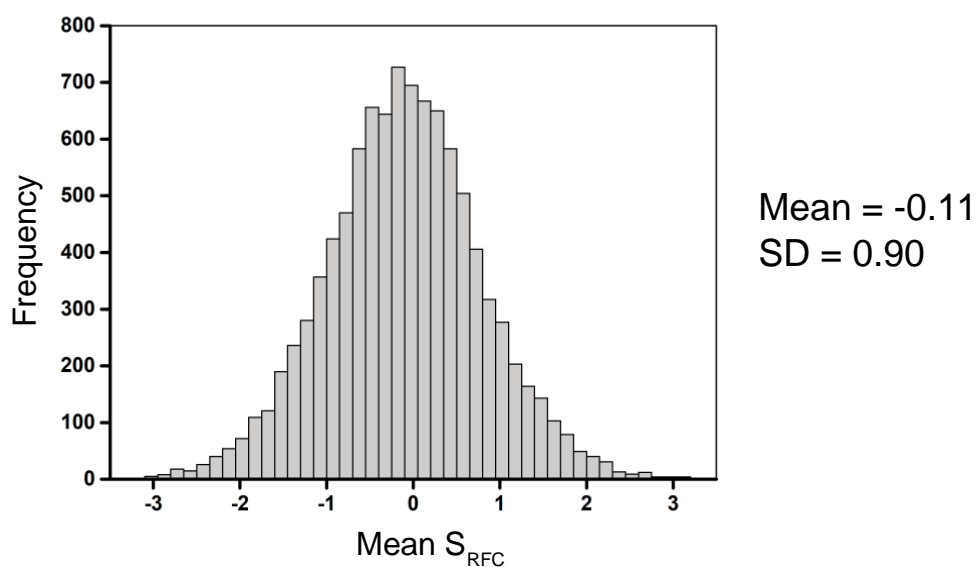


Figure 3.4 Determining the cutoff S_{RFC} value for PIP binding

The sequences of 246 human PH domains were scrambled 10,000 times, and the S_{RFC} for each scrambled sequence was calculated. The distribution of mean S_{RFC} is shown, with the overall mean and standard deviation indicated. The cutoff for PIP is 2.59 (mean + 3*SD).

	PC	PI(3,4,5)P ₃	PI(4,5)P ₂	PI(3,4)P ₂
ARHGAP27				
ARHGEF19				
ARHGEF28/ RGNEF				
DDEF2/ASAP2				
DDEFL1/ASAP3				
DNM1				
DNM3				
FGD1				
FGD2				
FGD3_PH1				
GAB3				
GAB4				
MCF2L2				
MPRIP_PH2				
OSBPL11				
PLCH2				
PLEKHG3				
PRKD1				
SH2B1				
PSD3				

	Does not Bind
	Binds PIP

Figure 3.5 Lipid-SiMPull assay results for 20 human PH domain-containing proteins predicted to bind PIPs.

20 proteins predicted to bind PIPs were assayed against PC, PI(3,4,5)P₃, PI(4,5)P₂, and PI(3,4)P₂ in lipid-SiMPull. 19 proteins bound one of the three PIP vesicles, only PSD3 did not display binding to the three PIP containing vesicles.

Table 3.1**A.**

Uniprot ID	Name	S _{RFC} score	Uniprot ID	Name	S _{RFC} score	Uniprot ID	Name	S _{RFC} score
Q96N96	SPATA13	32.83	Q9HD67	MYO10_PH2	8.94	O14492	SH2B2	5.49
Q9BRR9	ARHGAP9	24.85	Q8WZ64	CENTD1_PH2	8.64	Q05193	DNM1	5.32
Q9NR80	ARHGEF4	20.47	Q7Z6J4	FGD2_PH1	8.51	Q96P48	CENTD2_PH1	5.16
Q2WGN9	GAB4	19.05	Q01082	SPTBN1	8.42	Q8WZ64	CENTD1_PH4	5.11
Q7Z628	NET1	17.15	Q8WZ64	CENTD1_PH1	8.41	Q8TDY4	DDEFL1	4.73
Q8TCU6	PREX1	16.26	Q3KR16	PLEKHG6	8.32	P98174	FGD1_PH2	4.65
O95248	SBF1	15.31	P22059	OSBP	7.52	O75038	PLCH2	4.62
Q9ULL1	PLEKHG1	14.64	Q9BXB4	OSBPL11	7.43	Q9UN19	DAPP1	4.49
Q70Z35	DEPDC2	14.22	Q9Y243	AKT3	7.28	Q9ULH1	DDEF1	4.42
Q13424	SNTA1_PH2	12.93	P98174	FGD1__PH1	7.22	Q92974	ARHGEF2	4.14
Q58EX7	PLEKHG4	12.68	O43150	DDEF2	7.02	P31751	AKT2	4.12
O60890	OPHN1	12.29	Q9NYT0	PLEK2_PH2	6.90	Q8WWN8	CENTD3_PH1	3.79
Q8WWW8	GAB3	11.18	Q8N1W1	RGNEF	6.83	Q9Y5J5	PHLDA3	3.67
Q9Y5P4	CERT1	11.13	Q12802	AKAP13	6.79	Q96M96	FGD4_PH1	3.60
Q9HB20	PLEKHA3	11.04	Q9NYI0	PSD3	6.65	Q9NRF2	SH2B1	3.53
Q15283	RASA2	10.93	O14578	CIT	6.56	Q53GA4	PHLDA2	3.48
Q96DR7	SGEF	10.66	Q86SQ0	PHLDB2	6.56	Q8IW93	ARHGEF19	3.48
O94827	PLEKHG5	10.54	Q5JSP0	FGD3_PH1	6.51	P10911	MCF2	3.44
Q6ZUM4	ARHGAP27	10.25	Q8IVF5	TIAM2_PH1	6.49	Q9BYX2	TBC1D2	3.39
Q9UQ16	DNM3	10.13	Q9UPR0	PLCL2	6.46	Q7Z5H3	ARHGAP22	3.36
Q86YR7	MCF2L2	10.12	Q9UF11	PLEKHB1	6.40	Q15057	CENTB2	3.28
Q9H254	SPTBN4	9.88	Q86UU1	PHLDB1	6.32	O75962	TRIO_PH2	3.26
A1L390	PLEKHG3	9.85	Q9NZM3	ITSN2	6.27	Q7Z6J4	FGD2_PH2	3.08
Q8WZ64	CENTD1_PH3	9.85	Q9UQQ2	SH2B3	6.17	Q9BST9	RTKN	3.04
Q969R2	OSBP2	9.63	P51178	PLCD1	6.11	O94887	FARP2_PH2	3.02
Q9Y4F1	FARP1_PH2	9.61	Q96P48	CENTD2_PH2	6.10	Q96P50	CENTB5	2.92
Q96JA3	PLEKHA8	9.52	A5PKW4	PSD1	5.98			
Q9HB19	PLEKHA2_PH2	9.44	Q9NRC6	SPTBN5	5.87			
Q13425	SNTB2_PH2	9.18	Q9NQW6	ANLN	5.84			
Q14644	RASA3	8.96	Q15139	PRKD1	5.59			

Table 3.1 (cont.)**B.**

Uniprot ID	Name	S _{RFC} score	Uniprot ID	Name	S _{RFC} score	Uniprot ID	Name	S _{RFC} score
Q9Y2H5	PLEKHA6	2.40	Q5VT25	CDC42BPA	-0.41	Q8N4X5	AFAP1L2_PH2	-3.47
Q9HB19	PLEKHA2_PH1	2.35	Q8WWN8	CENTD3_PH3	-0.51	Q5JSL3	DOCK11	-3.96
Q9H4L5	OSBPL3	2.31	Q9HD67	MYO10_PH1	-0.77	Q07889	SOS1	-4.20
Q5JSP0	FGD3_PH2	2.07	Q6DT37	CDC42BPG	-0.98	Q6WCQ1	MPRIP_PH1	-4.22
Q9ULM0	PLEKHH1_PH1	2.00	Q9BZF2	OSBPL7	-1.01	Q86UW7	CADPS2	-4.56
Q6ZSZ5	ARHGEF18	1.98	P78314	SH3BP2	-1.16	O94887	FARP2_PH1	-4.71
Q96SU4	OSBPL9	1.96	Q9BZL6	PRKD2	-1.37	Q70E73	RAPH1	-5.19
Q8IVE3	PLEKHH2_PH2	1.93	Q15811	ITSN1	-1.55	P19174	PLCG1	-5.47
Q6WCQ1	MPRIP_PH2	1.91	P52735	VAV2	-1.61	O60346	PHLPP1	-5.60
Q8WZ64	CENTD1_PH5	1.89	Q9Y5S2	CDC42BPB	-1.67	Q9P227	ARHGAP23	-6.15
O15020	SPTBN2	1.87	Q9NYT0	PLEK2_PH1	-1.75	Q8WWN9	ICEF1	-7.11
Q13009	TIAM1_PH1	1.80	Q96P48	CENTD2_PH3	-1.76	Q9BZ29	DOCK9	-7.35
Q8WWN8	CENTD3_PH2	1.73	Q9Y4G2	PLEKHM1_PH1	-1.77	Q9UKW4	VAV3	-7.37
Q15111	PLCL1	1.69	Q96BY6	DOCK10	-1.77	O14827	RASGRF2_PH1	-7.51
P50570	DNM2	1.64	Q53GL0	PLEKHO1	-1.77	Q13393	PLD1	-7.83
P42680	TEC	1.60	O43374	RASA4_PH2	-1.81	Q7Z5R6	APBBLIP	-7.99
Q4KWH8	PLCH1	1.59	E7ERK1	RASA4_PH1	-1.81	P35626	ADRBK2	-9.19
O95294	RASAL1	1.55	Q9QX11	PSCD1	-1.85	Q5KSL6	DGKK	-9.84
Q9NPF8	CENTA2_PH2	1.25	O43739	CYTH3	-1.91	O14827	RASGRF2_PH2	-10.29
Q9ULM0	PLEKHH1_PH2	1.14	Q16760	DGKD	-1.94	Q8IZC4	PLEKHK1	-10.30
Q9Y4F1	FARP1_PH1	1.10	Q6ZVD8	PHLPP2	-1.96	Q8IZC4	RTKN2	-10.30
Q8IWE5	PLEKHM2	0.95	Q9NPF8	CENTA2_PH1	-2.07	Q9BQL6	C20ORF42	-10.33
Q9BZF3	OSBPL6	0.88	Q9ULU8	CADPS	-2.18	Q54A15	DTGCU2	-10.33
Q15052	ARHGEF6	0.77	Q8WXI2	CNKSR2	-2.23	Q12756	KIF1A	-10.47
Q86XP1	DGKH	0.48	O75962	TRIO_PH1	-2.40	Q5T5U3	ARHGAP21	-11.21
Q99418	CYTH2	0.14	Q9HAU0	PLEKHA5	-2.55	Q8N556	AFAP_PH2	-16.86
Q8IVE3	PLEKHH2_PH1	0.11	Q96M96	FGD4_PH2	-2.63	Q9P104	DOK5	-18.02
Q7Z736	PLEKHH3	0.08	Q9H2D6	TRIOBP	-2.74	Q8N4X5	AFAP1L2_PH1	-18.29
Q9Y4G2	PLEKHM1_PH2	-0.23	Q9H4M7	PLEKHA4	-2.76	Q13322	GRB10	-21.85
O14939	PLD2	-0.25	Q9UIA0	PSCD4	-3.21	Q6PKX4	DOK6	-25.08
Q8N264	ARHGAP24	-0.34	P35568	IRS1	-3.38	Q8N556	AFAP_PH1	-27.39

Table 3.1 Predicted S_{RFC} score for 179 PH domains

Table 3.1 (cont.)

(A) 86 PH domains that scored above the cutoff for binding (>2.59) and are predicted to bind PIPs with some specificity **(B)** 93 PH domains that scored below the cutoff for binding (>2.59)

3.6 References

1. B. J. Mayer, R. Ren, K. L. Clark, D. Baltimore, A putative modular domain present in diverse signaling proteins. *Cell* **73**, 629-630 (1993).
2. R. J. Haslam, H. B. Koide, B. A. Hemmings, Pleckstrin domain homology. *Nature* **363**, 309-310 (1993).
3. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. DeWald, D. Murray, S. D. Emr, M. A. Lemmon, Genome-wide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol Cell* **13**, 677-688 (2004).
4. W. S. Park, W. D. Heo, J. H. Whalen, N. A. O'Rourke, H. M. Bryan, T. Meyer, M. N. Teruel, Comprehensive identification of PIP3-regulated PH domains from *C. elegans* to *H. sapiens* by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).
5. I. Vonkova, A. E. Saliba, S. Deghou, K. Anand, S. Ceschia, T. Doerks, A. Galih, K. G. Kugler, K. Maeda, V. Rybin, V. van Noort, J. Ellenberg, P. Bork, A. C. Gavin, Lipid Cooperativity as a General Membrane-Recruitment Principle for PH Domains. *Cell Rep* **12**, 1519-1530 (2015).
6. K. Moravcevic, C. L. Oxley, M. A. Lemmon, Conditional peripheral membrane proteins: facing up to limited specificity. *Structure* **20**, 15-27 (2012).
7. S. J. Isakoff, T. Cardozo, J. Andreev, Z. Li, K. M. Ferguson, R. Abagyan, M. A. Lemmon, A. Aronheim, E. Y. Skolnik, Identification and analysis of PH domain-containing targets of phosphatidylinositol 3-kinase using a novel in vivo assay in yeast. *EMBO J* **17**, 5374-5387 (1998).
8. J. Pei, B. H. Kim, N. V. Grishin, PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* **36**, 2295-2300 (2008).
9. T. Balla, Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093-2104 (2005).
10. M. A. Lemmon, Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99-111 (2008).
11. S. Pant, E. Tajkhorshid, Microscopic Characterization of GRP1 PH Domain Interaction with Anionic Membranes. *J Comput Chem* **41**, 489-499 (2020).
12. E. Yamamoto, J. Domanski, F. B. Naughton, R. B. Best, A. C. Kalli, P. J. Stansfeld, M. S. P. Sansom, Multiple lipid binding sites determine the affinity of PH domains for phosphoinositide-containing membranes. *Sci Adv* **6**, eaay5736 (2020).

CHAPTER 4. REGULATION OF ARHGEF3 BY LIPID BINDING

4.1 Introduction

The RhoGEF family of proteins, consisting of ~70 members in the human genome, plays critical roles in multiple cellular processes (1, 2). The Dbl homology domain (DH) in RhoGEFs is responsible for facilitating the GDP to GTP catalysis in small GTPases, but the role of conserved PH domain adjacent to the DH domain is not well understood. As discussed in chapter 1, PH domains do not share the same functional significance in different RhoGEFs, and multiple mechanisms have been proposed for the PH domains in RhoGEFs. PH domains may be required for organelle targeting through interaction with membrane lipids, but they remain to be thoroughly explored to develop a comprehensive understanding. The PH domain may also act as an allosteric regulator of GEF activity or as a modulator for efficient activation through binding small GTPases (1, 3).

In the lipid-SiMPull screening described in chapter 2, we investigated PIP binding by 14 members of the RhoGEF family and found 9 of them to have affinity for specific PIPs. I set out to probe the functional relevance of the PH domain in one of the RhoGEFs, ARHGEF3 / XPLN (exchange factor found in platelets, leukemic, and neuronal tissues), with a classic DH-PH structure and containing a N-terminal region with no similarity to any known domains (Figure 4.1 A). ARHGEF3 is a GEF for RhoA and RhoB, and it has been reported to both activate the RhoA-ROCK kinase pathway to stimulate stress-fiber formation (4) and negatively regulate cell survival through inhibition of the mTORC2 kinase activity towards AKT (5). This inhibition of mTORC2 activation of AKT has been shown to be independent of its GEF activity with the N-terminal of ARHGEF3 being necessary and sufficient for the inhibition (5). Another recent study from our lab has found

that loss of ARHGEF3 in mice promotes skeletal muscle regeneration (6). The enhanced muscle regeneration was mediated by the GEF activity of ARHGEF3, but not the inhibition of mTORC2-Akt signaling, through activation of autophagy in regenerating muscle fibers. Despite the reported canonical and non-canonical functions of ARHGEF3, the role of the PH domain in these functions, if any, has not been investigated.

A previous study found that the PH domain of ARHGEF3 displayed little specificity towards PIPs in lipid strips assays (7); however, the full-length protein was not investigated for lipid-binding. The crystal structure of the PH domain of ARHGEF3 was found to have the classical PH domains structure formed by β -barrel lined by an α -helix at the C-terminus (8). In the SiMPull assays described in Chapter 2, we found the full-length ARHGEF3 to bind PI(4,5)P₂ and PI(3,5)P₂. Here, I set out to further investigate this novel lipid binding and to probe its functional relevance.

4.2 Materials and Methods

4.2.1 Cell Culture

See 2.2.1

4.2.2 Cell Lysis and Western Blotting

See 2.2.2

4.2.3 cDNA Cloning and Mutagenesis

The wild-type EGFP-ARHGFE3 construct was generated by cloning EGFP at N-terminal in pCMV-Myc-ARHGEF3 plasmid (Addgene #73365). The Point mutants for ARHGEF3 were created in the wild-type EGFP-ARHGFE3 plasmid using site-directed mutagenesis kit (QuikChange Lightning Site-Directed Mutagenesis Kit, Agilent #210518)

following the manufacturer's protocol. The list of primers used for mutagenesis can be found in Table 4.5.1. All cDNAs in the final plasmids were sequence-confirmed in their entirety followed by western blot against EGFP to confirm the expression.

4.2.4 Lipid Vesicle Preparation

See 2.2.4

4.2.5 Single-Molecule Pulldown Assay

See 2.2.5

4.2.6 Digitonin Permeabilization Assay

For digitonin permeabilization experiments, cells on coverslips were treated with 0.03% digitonin (20 mM HEPES, PH 7.5, 110 mM KOAc, 5 mM NaOAc, 2 mM Mg(OAc)₂, 1 mM EGTA) for 2.5 min, followed by fixation with 3.7% paraformaldehyde. Cells were stained with DAPI and mounted in Vectashield vibrance. Images were acquired using a personal deconvolution microscope system (DeltaVision, Applied Precision) at 100x magnification using 63x Objective lens. For PM localization ≥60 cells overexpressing either wild-type or K342E/K348E ARHGEF3 were scored.

4.2.7 Phalloidin Staining

NIH3T3 cells were plated on poly-L-lysine-coated glass coverslips. Cells were transfected with wild-type EGFP-ARHGEF3 or EGFP-ARHGEF3 (K342E/K348E) and fixed after 24 hours in 3.7% paraformaldehyde at room temperature for 15 min. After permeabilization with 0.1% Triton X-100 for 5 min, cells were incubated with Rhodamine phalloidin (Cytoskeleton Inc. #PHDR1) diluted (1:3000) in 3% BSA/PBS at 4° C for 60 min, followed by incubation with DAPI (1:2500) for 20 min at room temperature. The coverslips were mounted in Vectashield vibrance and visualized as described above.

4.2.8 Protein Purification

GST-ARHGEF3 wild-type and K342E/K348E proteins were expressed in *E. coli* from the pGEX-4T-1 vector and purified using glutathione Sepharose beads following the manufacturer's recommendations.

4.2.9 GTPase RhoA Activity Assays

Cellular RhoA activity was measured as previously described (9, 10). Briefly, HEK293 cells in 10-cm plate were transfected for 24 hrs and then lysed in buffer A (50 mM Tris (pH 7.4), 10mM MgCl₂, 500 mM NaCl, 1% (vol/vol) Triton X-100, 0.1%SDS, 0.5% sodium deoxycholate, and 1x Protease Inhibitors cocktail). The cell lysates were incubated for 1 hour at 4 °C with 60 µg of GST-rhotekin beads (Cytoskeleton Inc. #RT02), followed by washing with buffer B (50 mM Tris (pH 7.4), 10 mM MgCl₂, 150 mM NaCl, 1% (vol/vol) Triton X-100, and 1x Protease Inhibitors cocktail). The pulldown was analyzed by western blotting against anti-RhoA antibody (Cell Signaling Technology #2117).

To measure nucleotide exchange activity of RhoA in vitro, the mant-GTP exchange assay (Cytoskeleton Inc. #BK100), was performed following manufacturer's recommendation in 20 µL reaction volume in a 384-well plate using SpectraMax GeminiXPS Microplate reader. In short, purified RhoA (1 µM) was added to exchange buffer (40 mM Tris-HCl, pH 7.5, 100 mM NaCl, 20 mM MgCl₂, 4 µM mant-GTP) and fluorescence (ex360/em440) was recorded for 5 min before adding purified GST-ARHGEF3 or GST-ARHGEF3-K342E/K348E.

4.3 Results

4.3.1 Full-Length but not the PH Domain of ARHGEF3 Binds PIPs

We found the full-length ARHGEF3 protein to bind specifically to PI(4,5)P₂ and PI(3,5)P₂ in lipid-SiMPull assays (Figure 4.1 B). However, the PH domain alone fused with EGFP at either the N- or C-terminus did not bind PI(4,5)P₂ or PI(3,5)P₂ vesicles, neither did fragments of ARHGEF3 missing the N-terminal 125 amino acids (Δ N) or the PH domain (Δ PH) (Figure 4.1 C-D). Expression of the various ARHGEF3 proteins was confirmed by western blotting (Figure 4.2 A-B). Hence, it appears that the entire full-length ARHGEF3 may be necessary to confer lipid binding.

4.3.2 The PH Domain of ARHGEF3 is Necessary for PIP Binding

As another approach to assess whether the PH domain of ARHGEF3 was required for interaction with PIP, we considered structural information available in the literature. In reported PH domain-PIP headgroup interactions, positively charged amino acids in the β 1- β 2 and β 3- β 4 loops play critical roles (11). Applying the Patch Finder Plus software (12) to the reported ARHGEF3-PH crystal structure (8), we identified a positively charged patch containing K342, R345, K348, R376 and H427 to be potentially involved in the interaction between ARHGEF3 and PIP (Figure 4.3 A). These amino acids were mutated to alanine or acidic residues in full-length ARHGEF3 individually and in combination, and the mutant proteins (Figure 4.3 B) were subjected to SiMPull assays with PI(4,5)P₂ vesicles. Mutation of any one of the positively charged residues did not result in drastic reduction of PI(4,5)P₂ binding, but three double-mutants (K342E/K348E, R345D/R376D, R345D/H427D) and a triple-mutant (R345A/R376A/H427D) each abolished lipid binding by ARHGEF3 (Figure 4.3 B). Therefore, despite the requirement of the entire protein for

PI(4,5)P₂ binding, the PH domain of ARHGEF3 most likely contributes directly to the PIP-protein interaction.

Next, we investigated PI(3,5)P₂ binding by the ARHGEF3 mutants. As shown in Figure 4.3 C, the triple mutant and R345D/H427D abolished PI(3,5)P₂ binding. Interestingly, the K342E/K348E mutant retained PI(3,5)P₂ binding activity. These observations suggest that the PH domain of ARHGEF3 may interact with the two PIPs via structurally distinct mechanisms. Future investigation of these molecular interactions would be informative.

4.3.3 PI(4,5)P₂ Binding is Necessary for ARHGEF3 Plasma Membrane Targeting, Activation of RhoA, and Induction of Stress Fiber Formation

Because PI(4,5)P₂ is highly abundant in the plasma membrane (PM), we wondered if ARHGEF3 binding to PI(4,5)P₂ played a role in the protein's membrane recruitment. Recombinant EGFP-ARHGEF3 was found throughout the cell including the nucleus, and the nuclear localization is consistent with the presence of a nuclear localization sequence in ARHGEF3. We observed prominent presence of EGFP-ARHGEF3 in the PM of NIH3T3 cells and, importantly, the PI(4,5)P₂ binding-deficient mutant of ARHGEF3, K342E/K348E, displayed drastically reduced localization to the PM (Figure 4.4 A). This observation is consistent with ARHGEF3 targeting to the PM via its interaction with PI(4,5)P₂.

Next, we asked whether PI(4,5)P₂ binding is relevant to ARHGEF3's function in a cell. ARHGEF3 has two unrelated activities: as a GEF for RhoA/B (13) and as a GEF-independent inhibitor of mTORC2 phosphorylation of AKT (9). We examined these activities of ARHGEF3 in HEK293 cells. As shown in Figure 4.4 B, expression of wild-type ARHGEF3 suppressed AKT phosphorylation as expected, and expression of each lipid

binding-deficient mutant of ARHGEF3 (K342E/K348E, R345D/R376D, R345D/H427D, or R345A/R376A/H427D) was equally effective in inhibiting AKT, suggesting that PI(4,5)P₂ binding is not involved in ARHGEF3's function as an inhibitor of mTORC2 (nor is PI(3,5)P₂ binding). This is consistent with the previous observation that the N-terminal 125-amino acid fragment, devoid of the PH domain, is sufficient to inhibit mTORC2 (9). On the other hand, unlike the wild-type protein, the K342E/K348E mutant ARHGEF3 was ineffective in activating cellular RhoA activity upon overexpression, as measured by pulldown of active RhoA using an effector protein (Figure 4.4 C). Hence, lipid binding is apparently necessary for ARHGEF3's GEF activity toward RhoA in the cell.

We also performed in vitro GEF assays with ARHGEF3 and ARHGEF3-K342E/K348E proteins purified from bacteria, and found that the two proteins displayed near identical activities towards RhoA (Figure 4.4 D). We were not able to assess the effect of lipids on the GEF activity because we found addition of lipid vesicles to be incompatible with the GEF assay. Nevertheless, our observation confirmed that the K342E/K348E mutant retained proper folding and enzymatic activity. The simplest explanation for all our results combined is that PI(4,5)P₂ binding and/or PM targeting is necessary to achieve maximum GEF activity for ARHGEF3 in the cell.

Consistent with its role as a GEF for RhoA, ARHGEF3 has been reported to regulate actin cytoskeleton reorganization and assemble stress fibers in the cell (13). To probe a potential role of ARHGEF3-PI(4,5)P₂ interaction in this process, we expressed EGFP-ARHGEF3 in NIH3T3 cells and visualized actin cytoskeleton with phalloidin staining. As shown in Figure 4.4 E, expression of wild-type ARHGEF3, but not the K342E/K348E mutant, led to robust stress fiber formation, suggesting that PI(4,5)P₂ binding may be critical for ARHGEF3 regulation of actin cytoskeleton reassembly, a RhoA-dependent process. Taken together, our observations uncover a novel mechanism

of ARHGEF3 regulation and provide functional validation of a protein-lipid interaction identified in our lipid-SiMPull assays.

4.4 Discussion

We have identified specific lipid binding by full-length ARHGEF3, a RhoGEF, with PI(4,5)P₂ and PI(3,5)P₂, in contrast with previously reported non-specific binding by the PH domain of ARHGEF3 (7). The full-length protein of ARHGEF3 was required for its lipid binding. In-depth exploration of lipid-binding through site-directed mutagenesis in the full-length protein confirmed that the PH domain is indeed responsible for the lipid interaction. We identified 4 mutants with charge reversal of 2 or more positively charged amino acids in the variable loop of the PH domain, which resulted in a total loss of lipid-binding by the full-length ARHGEF3. Interestingly, the mutant K342E/K348E lost PI(4,5)P₂ binding while retaining its binding to PI(3,5)P₂, thus offering an ideal tool to probe functional relevance of PI(4,5)P₂ binding. The wild-type EGFP-ARHGEF3 was found to be localized on plasma membrane, in contrast to the K342E/K348E mutant that did not localize to plasma membrane. This observation suggests that ARHGEF3 binding to PI(4,5)P₂ is recapitulated in cells.

PI(3,5)P₂ is one of the phosphoinositides with least well-defined role in cellular physiology, which is produced in cells from PI(3)P by a PIKfyve lipid kinase complex (14, 15). This complex is comprised of a scaffold proteins Vac14, the lipid kinase PIKfyve, a PI(3)P 5-kinase, and the lipid phosphatase Fig4 (16). PI(3,5)P₂ was shown to be enriched in early and late endosome in mammalian cells (17) along with the PIKfyve lipid kinase complex (15). It is possible that ARHGEF3 is recruited to endosomes in cells through its interaction with PI(3,5)P₂. ARHGEF3 was reported to bind Rictor, a member of the mTORC2 complex, and specifically inhibit its kinase activity towards AKT (5). Recent studies have identified that mTORC2 is associated with various cellular compartments

including subpopulations of endosomes and lysosomes (18-20). This suggests that ARHGEF3 may regulate the mTORC2 activity in endosomal compartments through its interaction with PI(3,5)P₂. Future investigation is warranted to probe the role of PI(3,5)P₂ in ARHGEF3 function.

In agreement with the previous report (5) that N-terminal region of ARHGEF3 is necessary and sufficient for mTORC2 inhibition in cells, the lipid-binding deficient mutants, created by point mutations in the PH domain, didn't affect this function. Instead, we found that PI(4,5)P₂ binding is necessary for the GEF activity of ARHGEF3 towards RhoA in the cell and stress fiber formation. These results reveal a new regulation for ARHGEF3 and also offer an example of validation of the novel interactions identified in our lipid-SiMPull screen. Follow-up studies on the role of lipid-binding by the 8 remaining RhoGEFs, identified as PIP binders in lipid-SiMPull assay, will also likely be illuminating. These results also call for further studies of lipid-binding by the rest of the RhoGEF family members.

4.5 Figures and Table

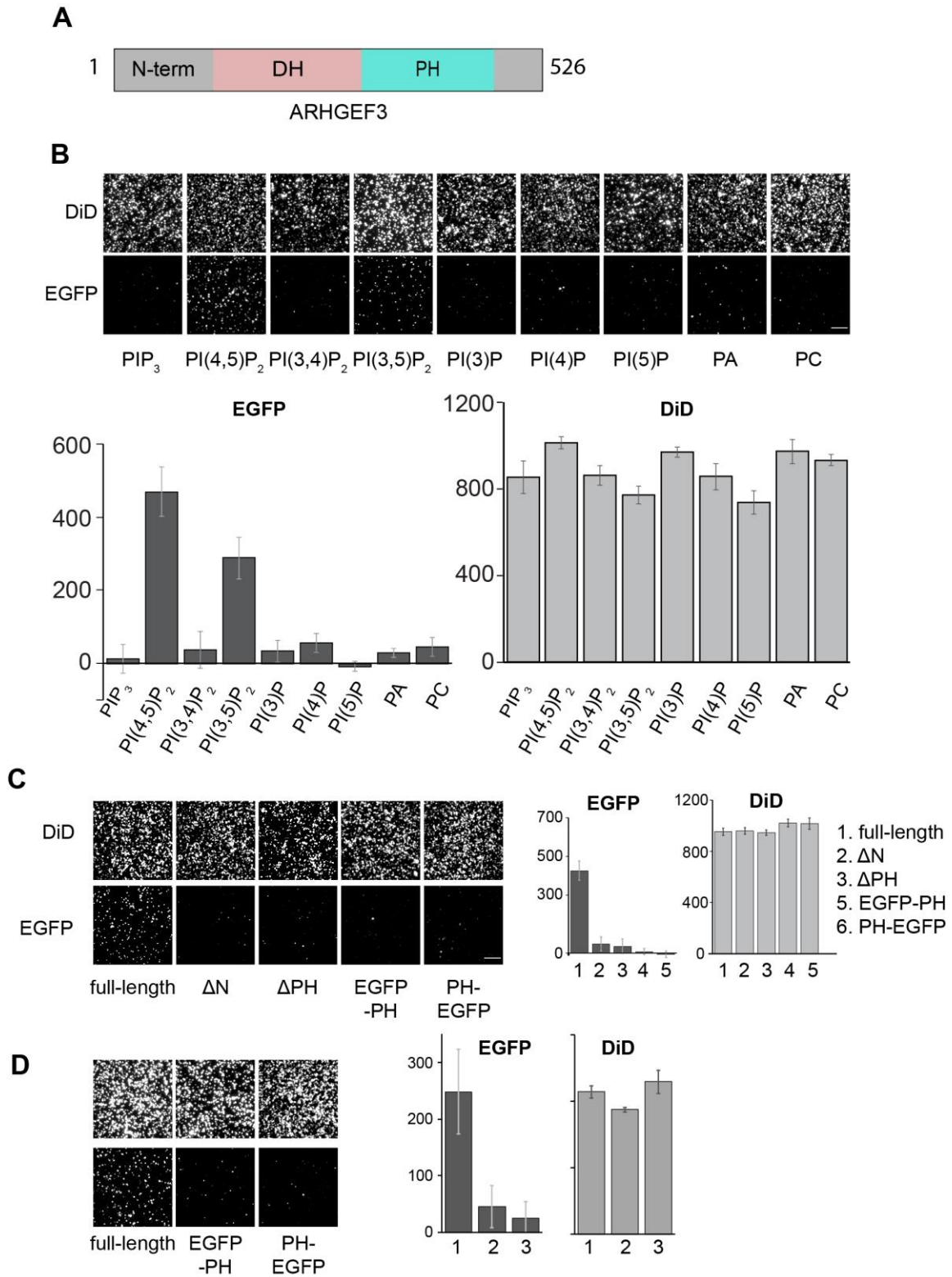


Figure 4.1 ARHGEF3 Binds PI(4,5)P₂ Through its PH Domain.

Figure 4.1 (cont.)

(A) Schematic diagram of ARHGEF3 domain structure.

(B) HEK293 cells were transiently transfected with EGFP-ARHGEF3, and the lysates were subjected to SiMPull assays against various types of lipid vesicles as indicated. Representative TIRF images are shown in upper panel. EGFP and DiD spots were quantified from 10 or more image areas to yield the average number of spots per 1600 μm^2 for each assay, shown in the lower panel. Each assay was performed with at least 3 independent transfections with similar results.

(C) Similar to B, various EGFP-tagged ARHGEF3 fragments were assayed for binding to vesicles containing 5% PI(4,5) P_2 . Representative TIRF images on the right and quantification in the left panel.

(D) Similar to C, various EGFP-tagged ARHGEF3 fragments were assayed for binding to vesicles containing 5% PI(3,5) P_2 . Representative TIRF images on the right and quantification in the left panel.

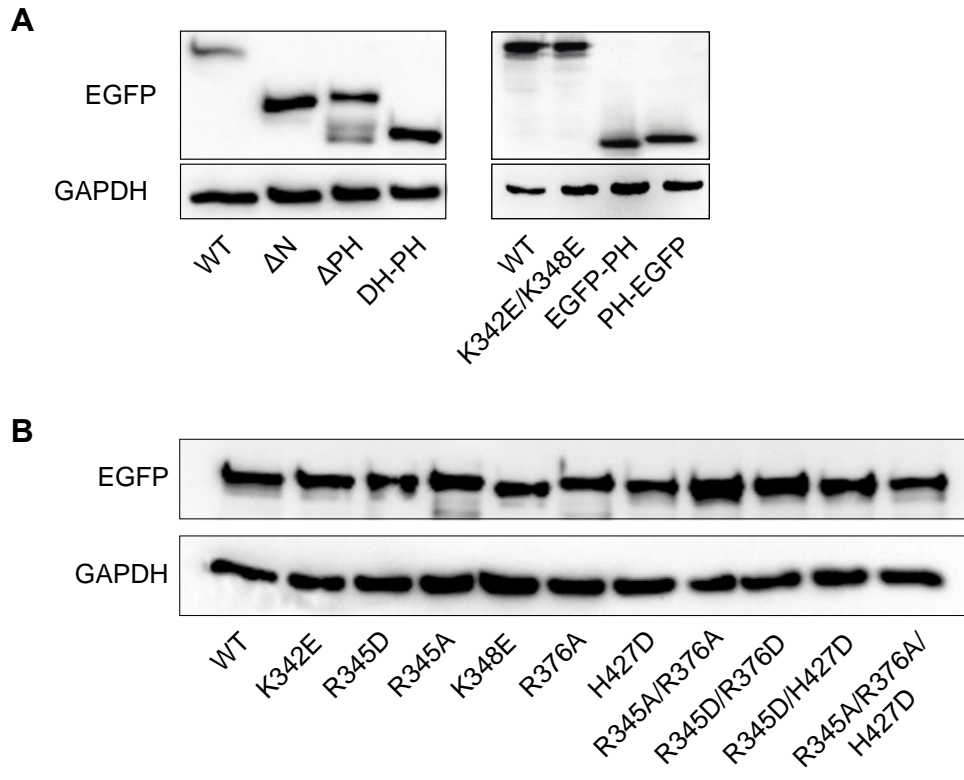


Figure 4.2 Expression of ARHGEF3 constructs in HEK293 cells

HEK293 cells were transiently transfected to express various truncations **(A)** or point mutants **(B)** of ARHGEF3 fused to EGFP. Cell lysates were analyzed by western blotting with an anti-GFP antibody, and anti-GAPDH blotting served as loading control.

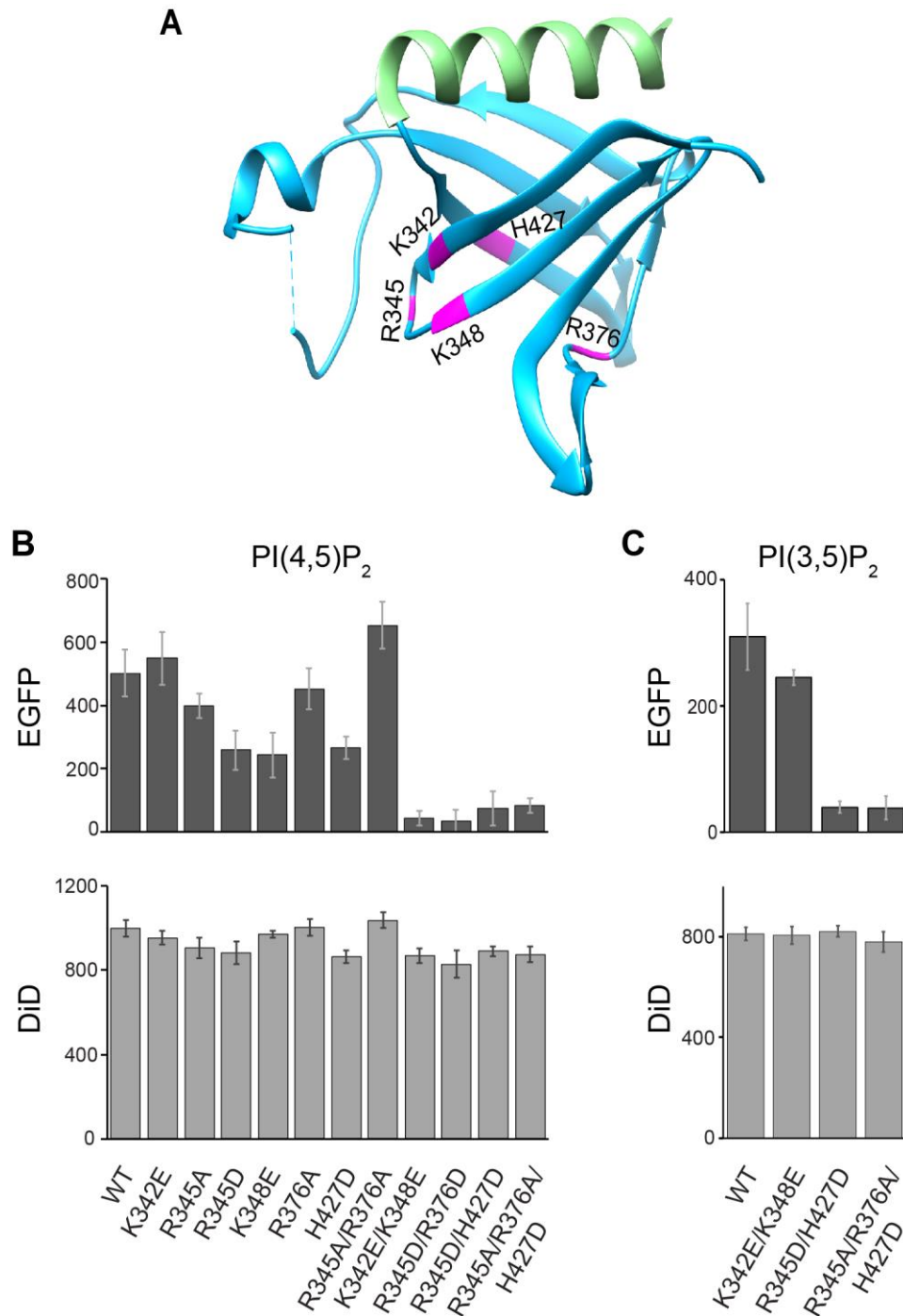


Figure 4.3 ARHGEF3 binds PI(4,5)P₂ through its PH domain.

(A) Ribbon representation of the structure of ARHGEF3 PH domain (PDB: 2Z0Q), with residues predicted to interact with PIP highlighted in magenta.

(B) HEK293 cells were transiently transfected with various mutants of EGFP-ARHGEF3 as indicated, and the lysates were subjected to SiMPull assays against 5% PI(4,5)P₂ lipid vesicles. EGFP and DiD spots were quantified as described above, quantification from

Figure 4.3 (cont.)

one experiment shown in upper panel (EGFP) and lower panel (DiD). Each assay was performed with at least 3 independent transfections with similar results.

(C) Similar to B, EGFP-tagged ARHGEF3 mutants were assayed for binding to 5% PI(3,5)P₂ vesicles. Each assay was repeated with lysates from at least 3 independent transfections.

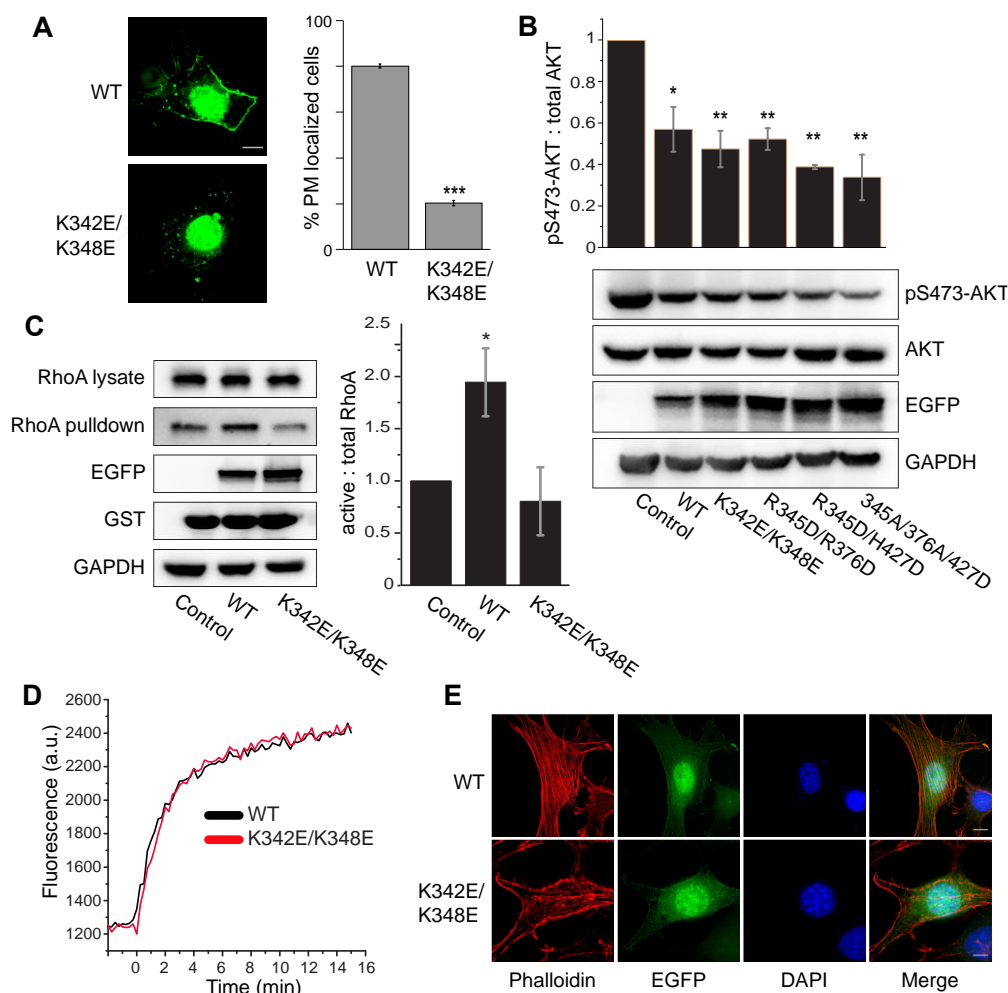


Figure 4.4 PI(4,5)P₂ binding is necessary for ARHGEF3 membrane targeting, cellular activity and function.

(A) NIH3T3 cells were transfected for 24 hrs with EGFP-tagged wild-type or lipid binding-deficient mutant K342E/K348E of ARHGEF3. The transfected cells were then treated with 0.03% digitonin for 2.5 min to remove cytosolic proteins before fixation. Percentage of transfected cells with PM localization was quantified from 3 independent experiments.

(B) HEK293 cells were transfected with wild-type or lipid binding-deficient mutants of ARHGEF3, serum-starved overnight, and then stimulated with 10% FBS for 30 min followed by cell lysis and western blotting. Western blot signals were quantified by densitometry to generate ratio of pAKT versus total AKT from 3 independent experiments.

(C) HEK293 cells were transfected for 24 hrs with empty vector (control), wildtype, or K342E/K348E mutant of ARHGEF3. Cell lysates were subjected to active RhoA pulldown assay using GST-rhotekin beads, and analyzed by western blotting. Western blot signals from 5 independent experiments were quantified by densitometry to generate the relative ratios of RhoA pulled down (active) and RhoA in lysates (total) shown in the graph.

Figure 4.4 (cont.)

(D) Purified wild-type and K342E/K348E mutant ARHGEF3 were subjected to in vitro RhoA guanine nucleotide exchange assay. Three independent experiments were performed with similar outcome, and representative results are shown.

(E) NIH3T3 cells were transfected for 24 hrs with wild-type or K342E/K348E mutant of ARHGEF3, followed by fixation and phalloidin/DAPI staining. Three independent experiments were performed with similar outcome, and representative images are shown.

Error bars represent SD. Student's t test was performed to compare WT and mutant in A, and each sample to the control in B & C. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Scale bars: 5 μm .

ARHGEF3 Mutagenesis Primers	
Primer Sequence	Name
5'-GTCATGGTGAAGTGGAGAACAATCGGGGCGTG -3'	ARHGEF3_K342E_Fwd
5'-CACGCCCCGATTGTTCTCCAGTTCACCATGAC -3'	ARHGEF3_K342E_Rev
5'-TGTTGTCATGGTGAAGTGAAGAACAATGCCGGCGT GAAACTGCA -3'	ARHGEF3_R345A_Fwd
5'-TGCAGTTTCACGCCGGCATTGTTCTTCAGTTCACCA TGACAACA -3'	ARHGEF3_R345A_Rev
5'-GAACTGAAGAACAATGACGGCGTGAACTGCATG -3'	ARHGEF3_R345D_Fwd
5'- CATGCAGTTTCACGCCGTCATTGTTCTTCAGTTC -3'	ARHGEF3_R345D_Rev
5'- GAACAATCGGGGCGTGGAAGTGCATGTTTTCCTG -3'	ARHGEF3_K348E_Fwd
5'-CAGGAAAACATGCAGTTCCACGCCCCGATTGTTC -3'	ARHGEF3_K348E_Rev
5'-GCTACCAGCTGTACGCTCAGCCAATCCCCG -3'	ARHGEF3_R376A_Fwd
5'-CGGGGATTGGCTGAGCGTACAGCTGGTAGC -3'	ARHGEF3_R376A_Rev
5'-GCTACCAGCTGTACGATCAGCCAATCCCCGTG -3'	ARHGEF3_R376D_Fwd
5'-CACGGGGATTGGCTGATCGTACAGCTGGTAGC -3'	ARHGEF3_R376D_Rev
5'-GGCTTGTAGCGAGTCGGTCTGACTTTGGG -3'	ARHGEF3_H427D_Fwd
5'-CCCAAAGTCAGACCGACTCGCTACAAGCC-3'	ARHGEF3_H427D_Rev
5'-CATGGTGAAGTGGAGAACAATCGGGGCGTGGAAGT GCATGTTTTCC -3'	ARHGEF3_K342E_K348E _Fwd
5'-GGAAAACATGCAGTTCCACGCCCCGATTGTTCTCCA GTTCAACCATG -3'	ARHGEF3_K342E_K348E _Rev

Table 4.1 List of PCR primers used for generating point mutations in wild-type EGFP-ARHGEF3 plasmid

4.6 References

1. K. L. Rossman, C. J. Der, J. Sondek, GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat Rev Mol Cell Biol* **6**, 167-180 (2005).
2. D. R. Cook, K. L. Rossman, C. J. Der, Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. *Oncogene* **33**, 4021-4035 (2014).
3. Y. Zheng, Dbl family guanine nucleotide exchange factors. *Trends Biochem Sci* **26**, 724-732 (2001).
4. W. T. Arthur, S. M. Ellerbroek, C. J. Der, K. Burrridge, K. Wennerberg, XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not RhoC. *J Biol Chem* **277**, 42964-42972 (2002).
5. N. Khanna, Y. Fang, M. S. Yoon, J. Chen, XPLN is an endogenous inhibitor of mTORC2. *Proc Natl Acad Sci U S A* **110**, 15979-15984 (2013).
6. J.-S. You, N. Singh, A. Reyes-Ordóñez, N. Khanna, Z. Bao, H. Zhao, J. Chen, ARHGEF3 regulates skeletal muscle regeneration and strength through autophagy. *bioRxiv*, (2020).
7. W. S. Park, W. D. Heo, J. H. Whalen, N. A. O'Rourke, H. M. Bryan, T. Meyer, M. N. Teruel, Comprehensive identification of PIP3-regulated PH domains from *C. elegans* to *H. sapiens* by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).
8. K. Murayama, M. Kato-Murayama, R. Akasaka, T. Terada, S. Yokoyama, M. Shirouzu, Structure of the Rho-specific guanine nucleotide-exchange factor Xpln. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**, 1455-1459 (2012).
9. N. Khanna, Y. Fang, M. S. Yoon, J. Chen, XPLN is an endogenous inhibitor of mTORC2. *Proc Natl Acad Sci U S A*. **110**, 15979-15984. (PMC3791717) (2013).
10. X. D. Ren, M. A. Schwartz, Determination of GTP loading on Rho. *Methods Enzymol* **325**, 264-272 (2000).
11. T. Balla, Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093-2104 (2005).
12. S. Shazman, G. Celniker, O. Haber, F. Glaser, Y. Mandel-Gutfreund, Patch Finder Plus (PFplus): a web server for extracting and displaying positive electrostatic patches on protein surfaces. *Nucleic Acids Res* **35**, W526-530 (2007).
13. W. T. Arthur, S. M. Ellerbroek, C. J. Der, K. Burrridge, K. Wennerberg, XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not RhoC. *J Biol Chem*. **277**, 42964-42972. (2002).
14. N. Jin, M. J. Lang, L. S. Weisman, Phosphatidylinositol 3,5-bisphosphate: regulation of cellular events in space and time. *Biochem Soc Trans* **44**, 177-184 (2016).
15. J. Hasegawa, B. S. Strunk, L. S. Weisman, PI5P and PI(3,5)P₂: Minor, but Essential Phosphoinositides. *Cell Struct Funct* **42**, 49-60 (2017).
16. J. A. Lees, P. Li, N. Kumar, L. S. Weisman, K. M. Reinisch, Insights into Lysosomal PI(3,5)P₂ Homeostasis from a Structural-Biochemical Analysis of the PIKfyve Lipid Kinase Complex. *Molecular Cell*, (2020).
17. S. Takatori, T. Tatematsu, J. Cheng, J. Matsumoto, T. Akano, T. Fujimoto, Phosphatidylinositol 3,5-Bisphosphate-Rich Membrane Domains in Endosomes and Lysosomes. *Traffic* **17**, 154-167 (2016).
18. R. Jia, J. S. Bonifacino, Lysosome Positioning Influences mTORC2 and AKT Signaling. *Mol Cell* **75**, 26-38 e23 (2019).

19. M. Ebner, B. Sinkovics, M. Szczygiel, D. W. Ribeiro, I. Yudushkin, Localization of mTORC2 activity inside cells. *J Cell Biol* **216**, 343-353 (2017).
20. W. Fu, M. N. Hall, Regulation of mTORC2 Signaling. *Genes (Basel)* **11**, (2020).

CHAPTER 5. REGULATION OF AKT-PIP₃ INTERACTION BY INTER-DOMAIN INTERACTIONS

5.1 Introduction

Since its discovery more than three decades ago (1), AKT has emerged as a major regulator of cellular homeostasis. AKT was found to be activated by growth factors through recruitment of PI3K by receptor tyrosine kinases to plasma membrane (2-4). At the plasma membrane, PI3K generates PI(3,4)P₂ and PI(3,4,5)P₃, which recruits AKT to through its PH domain. Perturbation in AKT signaling cascade has been associated with numerous pathophysiological conditions such as cancer, diabetes, cardiovascular disease. One of the lipid-binding mutants of AKT, E17K, was identified as having potential oncogenic activity due to increased membrane association (5). Many mutations have been identified in AKT that disrupt the interdomain interaction and subsequently modulating its lipid binding, kinase activity and oncogenicity. Disruption of potential PH-kinase domain interactions through L52R and D323H mutations in AKT resulted in upregulation of its oncogenic activity (6). Another report identified Trp80 on the PH domain that interacts with a cleft formed on kinase domain by Lys297, Glu298, and Glu314 and helps maintain the inactive conformation (7, 8). The disruption of PH-kinase domain interface, by D323A/D325A mutation in the kinase domain, resulted in a hyperphosphorylated AKT independent of its PIP₃ binding suggesting the loss of inhibited conformation in inactive AKT (9). Efficient interactions between the PH and Kinase domain was required for engagement with PIP₃ and PI(3,4)P₂ and subsequent allosteric activation (9). This study proposed an allosteric model of AKT activation where lipid binding of PH domain relieves autoinhibitory interactions, which is necessary for maintaining its inactive state. However, this was contradicted by another recent study that

did not find the evidence to support the above model, but discovered that the interaction between the C-tail and the PH-kinase domain linker is required for Ser473 phosphorylation (10).

The lipid-binding of AKT is a critical step for its plasma membrane recruitment and subsequent activation. Even though many studies have proposed different interdomain interaction models and their role in regulating AKT activation, the conformational mechanism of its activation still remains to be fully elucidated. Lipid-SiMPull provides a great advantage that protein-lipid interactions can be studied at single-molecule level. Taking advantage of strengths of single-molecule microscopy, I have studied the interaction between AKT and its binding partner PIP₃. While the majority of studies have been focused on unraveling the PH-kinase domain interactions, the role of the hydrophobic motif (HM) domain in PIP₃ binding and its role in AKT activation, if any, has not been investigated. My studies described in this chapter suggest regulatory roles for the HM domain.

5.2 Materials and Methods

5.2.1 Cell Culture

See 2.2.1

5.2.2 Cell Lysis and Western Blotting

See 2.2.2

5.2.3 cDNA Cloning and Mutagenesis

The AKT constructs were purchased from Addgene GFP-AKT (#39531), and GFP-AKT-PH (#39533). The Δ HM construct (aa 1-408) was cloned using Gibson assembly in EGFP-C1 vector. All constructs were verified by western blot and sequencing.

5.2.4 Lipid Vesicle Preparation

See 2.2.4

5.2.5 Single-Molecule Pulldown Assay

See 2.2.5

5.2.6 SiMPull Image Analysis

See 2.2.6

5.3 Results

5.3.1 The PH Domain of AKT Displays Different Lipid-Binding Behavior than Full-Length AKT

In lipid-SiMPull assay, stoichiometry of binding can be determined by acquiring movies of fluorophores and analyzing single-molecule traces to count the photobleaching steps (11-13). Next, the single-molecule traces can be analyzed for binding and re-binding events to determine stable or transient binding. Both of the above analysis require the events to happen within the lifetime of EGFP fluorescence in the setup. The average lifetime of recombinant EGFP in our system was 6.34 ± 0.25 sec (Figure 5.1 B).

As reported earlier (13), the PH domain of AKT displays multi-copy binding with PIP₃ lipid vesicles in lipid-SiMPull assay, whereas the full-length mostly displays monomeric binding. I further observed that the PH domain of the AKT had markedly

higher EGFP pulldown than the full-length AKT at same concentration of EGFP applied to PIP₃ lipid vesicles under similar conditions (Figure 5.1 C). This suggests that the isolated PH domain of AKT may have an exposed lipid-binding surface resulting in higher binding, whereas the lipid-binding of full-length AKT may be inhibited by tertiary structure of full-length protein. Next, I investigated the dynamics of copy number of PH domains bound to PIP₃ vesicles under different conditions. The copy number of PH domains bound to PIP₃ vesicles increased over time (Figure 5.1 D). When the surface was saturated with lipid vesicles, it had relatively higher monomeric binding after 10 minutes of incubation compared to optimal vesicle density (Figure 5.1 D). As mentioned above, the copy number of binding can only be determined if the binding events happen within the lifetime of EGFP. The full-length AKT was found to bind stably with PIP₃ vesicles, whereas the PH domain displayed transient binding, as reported earlier (13).

5.3.2 The HM domain regulates PIP₃ binding of AKT

Next, I investigated the role of HM of AKT towards regulating its lipid binding by comparing the binding with PIP₃ vesicles between the full-length and the HM deletion construct (Δ HM). The deletion of the HM resulted in increased EGFP pulldown with PIP₃ vesicles, similar to the PH domain (Figure 5.2 A). The AKT construct having deletion of HM also displayed increased rate of binding with a remarkable increase in EGFP pull down by PIP₃ (Figure 5.2 B). The binding with full-length AKT is low but kept going up at slow rate, while the binding with Δ HM increased quickly and plateaued around 40 minutes. These observations suggest that the HM of AKT may inhibit PIP₃ binding in the full-length protein and this inhibition may need to be released for its activity. The photobleaching step analysis of Δ HM binding with PIP₃ revealed mostly monomeric binding similar to the full-length protein (Figure 5.2 C), in contrast with the multi copy binding by the PH domain. While the PH domain mostly bound transiently, the Δ HM

construct bound stably similar to the full-length protein. As shown in Figure 5.3, the PH domain displayed transient binding with PIP₃ vesicles, whereas both the full-length and the Δ HM displayed stable binding.

5.4 Discussion

Based on reported crystal structures of AKT (10, 14), the lipid-binding site of the PH domain is buried inside the lobes of kinase domain and thus should inhibit direct PIP₃ interaction. This suggests that inactive AKT may exist in an equilibrium with most of the free AKT not bound to PIP₃, and a drastic increase of PIP₃ on plasma membrane such as by mitogenic stimulation (15, 16) may be necessary to shift the equilibrium towards binding. This further argues that the isolated PH domain can present an exposed lipid-binding surface that should result in better PIP₃ binding, which was not found in traditional lipid binding studies (17). But, in lipid-SiMPull, I have consistently observed higher PIP₃ binding by the isolated PH domain than the full-length protein. Another major difference observed has been the transient binding by the PH domain but stable binding of the full-length protein, as reported earlier by Arauz et al (13). This suggests that interdomain interaction may be necessary for maintaining stable and continuous association with PIP₃, which would allow for longer residence times. These observations may have been missed in assays such as SPR that use high density of lipid vesicle as compared to the low density allowing for single molecule resolution in SiMPull. Many commonly used assays to detect lipid-protein interactions such as lipid strips or liposome centrifugation would be unable to detect such differences in binding characteristics.

Another significant observation is that the deletion of HM (aa 408 – 480) in AKT results in apparent increase in EGFP pulldown against PIP₃ compared to the full-length

protein, while maintaining the stable interaction similar to the full-length construct. This suggests that the HM inhibits lipid binding in inactive AKT and assists in maintaining the inactive conformation with lipid-binding surface of the PH domain buried towards the kinase domain. Deletion of HM may result in significant conformational changes exposing the lipid-binding surface of the PH domain. While the HM may help in maintaining the inactive conformation, the interdomain engagement of PH–kinase domain may assist in stabilizing the PIP₃ binding. The Δ HM construct, which contains both the PH and the kinase domain, displayed stable binding similar to the full-length protein. The copy number of Δ HM construct was also similar to the full-length protein rather than the PH domain. In other words, like the PH domain the Δ HM construct has higher affinity for PIP₃ than the full-length protein, but unlike the PH domain it binds PIP₃ stably in a monomeric stoichiometry.

Our studies suggest that the lipid binding by AKT is regulated by various inter-domain interactions, which are yet to be comprehensively understood. The striking difference observed in the binding characteristics of the Δ HM construct warrants further investigation of the potential role of interdomain communication in modulating lipid binding by AKT. Future studies may include site-directed mutagenesis to disrupt domain-domain interactions in AKT followed by lipid-SiMPull assay, growth factor stimulated cellular translocation, and kinase activity assays.

5.5 Figures

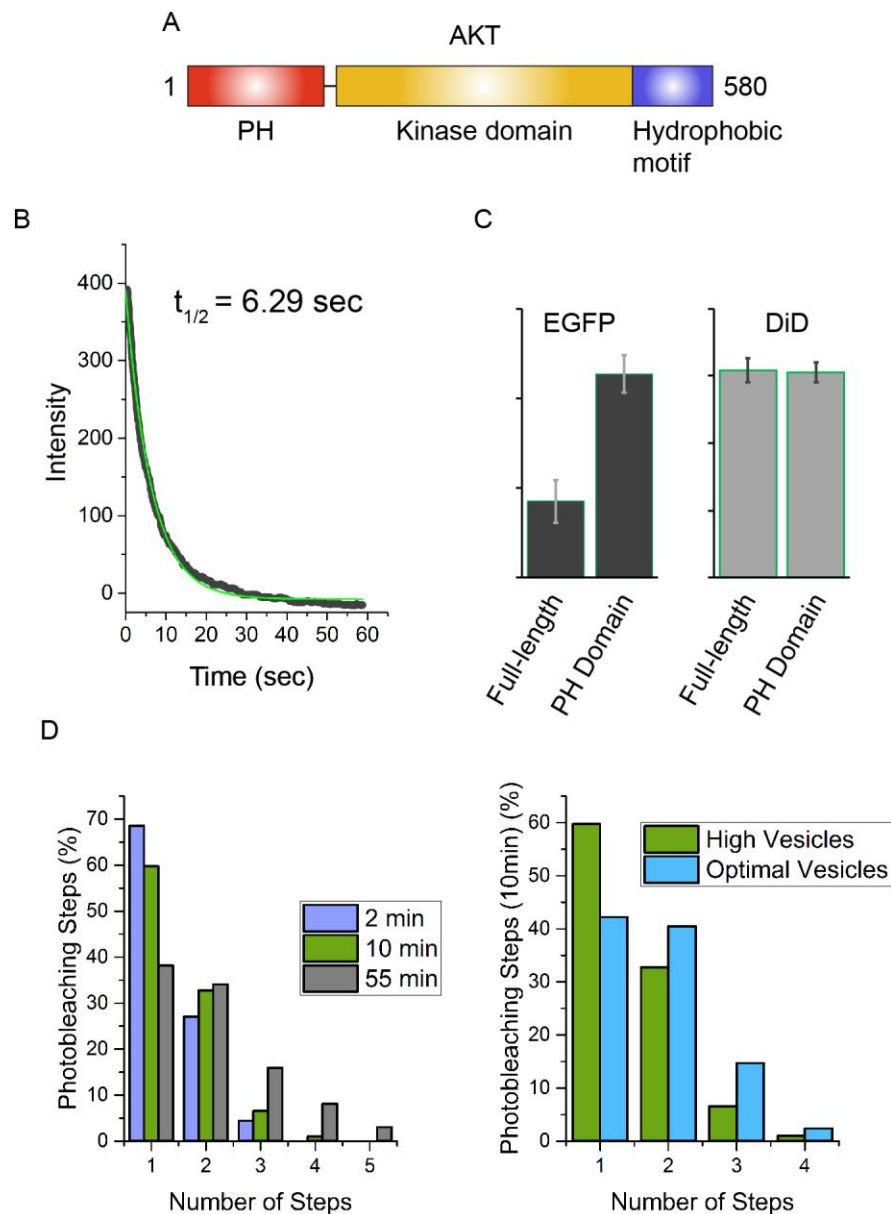


Figure 5.1 The PH domain of AKT displays different PIP₃ binding behavior than the full-length protein under similar conditions.

(A) the domain structure of AKT

(B) Recombinant EGFP was pulled down by biotinylated GFP antibody followed by acquiring movies for 60 seconds. A representative intensity decay plot is shown here. The average lifetime, $6.34 \pm 0.25 \text{ sec}$, was determined by quantifying > 600 single molecule traces in 4 movies. The black curve represents the original decay of EGFP intensity, the green curve is the exponential decay fit.

Figure 5.1 (cont.)

(C) HEK293 cells were transfected with EGFP-tagged full-length AKT and the PH domain of AKT (aa 1-147). Cell lysates were diluted to 5nM EGFP and added to 10% PIP₃ vesicles immobilized in slide chambers. The quantification for EGFP and DiD spots on surface is shown above.

(D) In the experiment of (B), photobleaching steps were calculated for the EGFP-tagged PH domain of AKT at 2 minutes, 10 minutes and 55 minutes of incubation with high density of PIP₃ vesicles (left panel). Photobleaching steps were also compared for two different PIP₃ vesicle densities on the surface (right panel)

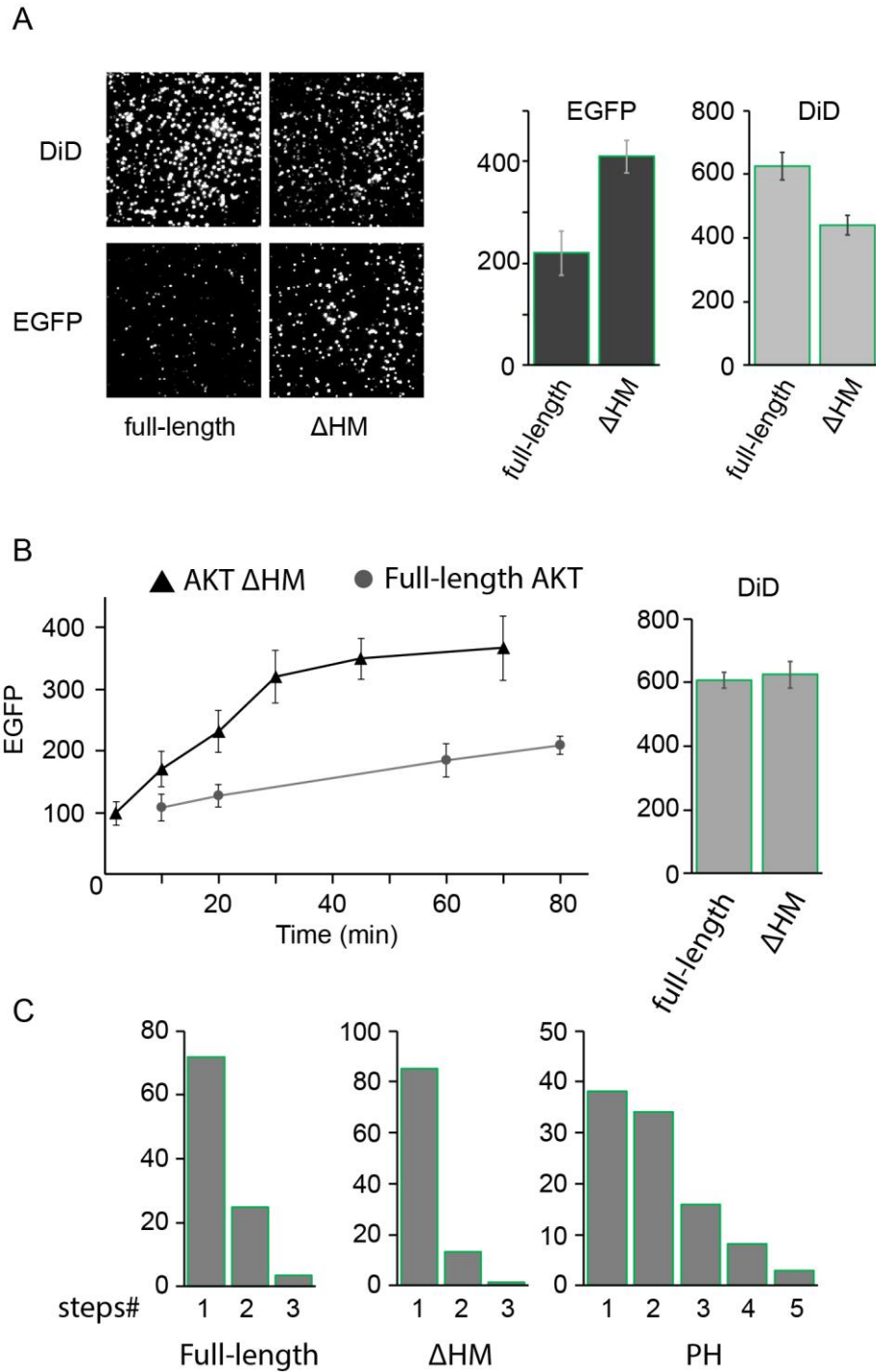


Figure 5.2 HM inhibits PIP₃ binding of full-length AKT

(A) HEK293 cells were transfected with EGFP-tagged full-length AKT and the HM deletion construct (Δ HM). Cell lysates were diluted to 5nM EGFP and added to 5% PIP₃ vesicles immobilized in slide chambers. Representative TIRF images (right panel) for EGFP

Figure 5.2 (cont.)

pulldown and PIP₃ lipid vesicles (DiD) and the quantification for EGFP and DiD spots is shown in the left panel from one experiment.

(B) Similar to (A), the binding of full-length and Δ HM constructs were quantified over the shown time period. The PIP₃ vesicles were quantified for both and shown in the left panel.

(C) Number of photobleaching steps were quantified for full-length AKT, the Δ HM construct and the PH domain binding with PIP₃. 500 or more single molecule traces were analyzed and are presented as percentage of total quantified traces.

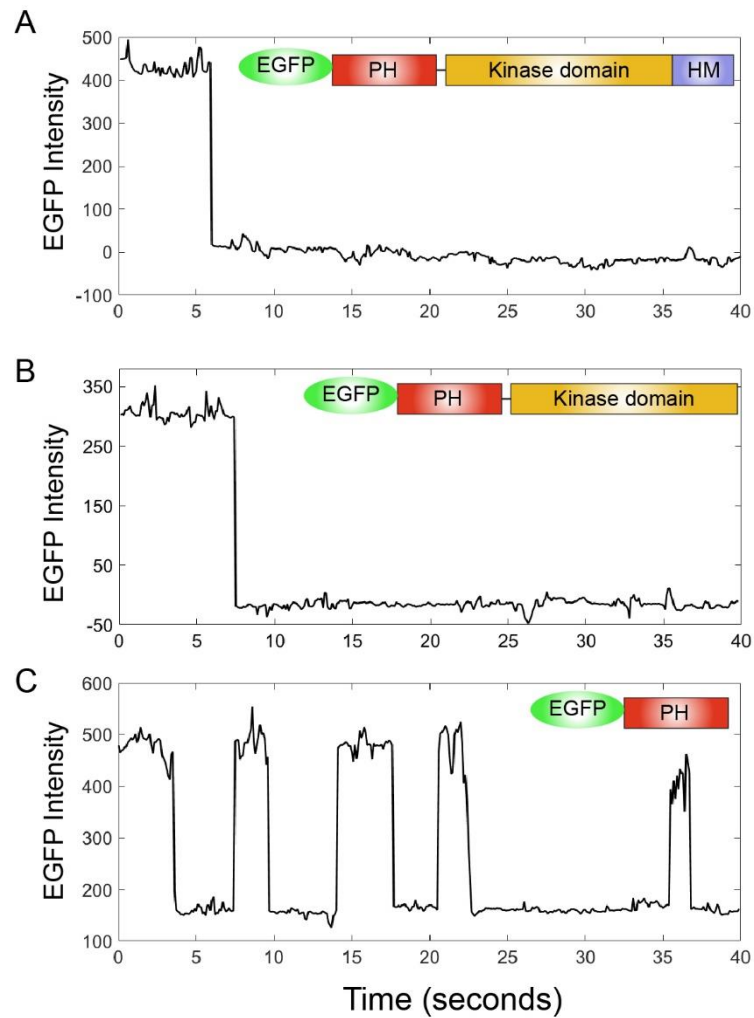


Figure 5.3 HM can stabilize binding with PIP_3

Single molecule intensity traces were acquired for 40 seconds and EGFP intensity Vs time was plotted for EGFP-tagged **(A)** full-length, **(B)** HM deletion, and **(C)** PH domain AKT. The domain structure for each construct is shown in each panel.

5.6 References

1. S. P. Staal, Molecular cloning of the akt oncogene and its human homologues AKT1 and AKT2: amplification of AKT1 in a primary human gastric adenocarcinoma. *Proc Natl Acad Sci U S A* **84**, 5034-5037 (1987).
2. T. F. Franke, S.-I. Yang, T. O. Chan, K. Datta, A. Kazlauskas, D. K. Morrison, D. R. Kaplan, P. N. Tsichlis, The protein kinase encoded by the Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase. *Cell* **81**, 727-736 (1995).
3. A. D. Kohn, K. S. Kovacina, R. A. Roth, Insulin stimulates the kinase activity of RAC-PK, a pleckstrin homology domain containing ser/thr kinase. *The EMBO Journal* **14**, 4288-4295 (1995).
4. B. M. Burgering, P. J. Coffey, Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. *Nature* **376**, 599-602 (1995).
5. J. D. Carpten, A. L. Faber, C. Horn, G. P. Donoho, S. L. Briggs, C. M. Robbins, G. Hostetter, S. Boguslawski, T. Y. Moses, S. Savage, M. Uhlik, A. Lin, J. Du, Y. W. Qian, D. J. Zeckner, G. Tucker-Kellogg, J. Touchman, K. Patel, S. Mousses, M. Bittner, R. Schevitz, M. H. Lai, K. L. Blanchard, J. E. Thomas, A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* **448**, 439-444 (2007).
6. C. Parikh, V. Janakiraman, W. I. Wu, C. K. Foo, N. M. Kljavin, S. Chaudhuri, E. Stawiski, B. Lee, J. Lin, H. Li, M. N. Lorenzo, W. Yuan, J. Guillory, M. Jackson, J. Rondon, Y. Franke, K. K. Bowman, M. Sagolla, J. Stinson, T. D. Wu, J. Wu, D. Stokoe, H. M. Stern, B. J. Brandhuber, K. Lin, N. J. Skelton, S. Seshagiri, Disruption of PH-kinase domain interactions leads to oncogenic activation of AKT in human cancers. *Proc Natl Acad Sci U S A* **109**, 19368-19373 (2012).
7. V. Calleja, D. Alcor, M. Laguerre, J. Park, B. Vojnovic, B. A. Hemmings, J. Downward, P. J. Parker, B. Larijani, Intramolecular and intermolecular interactions of protein kinase B define its activation in vivo. *PLoS Biol* **5**, e95 (2007).
8. V. Calleja, M. Laguerre, P. J. Parker, B. Larijani, Role of a novel PH-kinase domain interface in PKB/Akt regulation: structural mechanism for allosteric inhibition. *PLoS Biol* **7**, e17 (2009).
9. M. Ebner, I. Lucic, T. A. Leonard, I. Yudushkin, PI(3,4,5)P3 Engagement Restricts Akt Activity to Cellular Membranes. *Mol Cell* **65**, 416-431 e416 (2017).
10. N. Chu, A. L. Salguero, A. Z. Liu, Z. Chen, D. R. Dempsey, S. B. Ficarro, W. M. Alexander, J. A. Marto, Y. Li, L. M. Amzel, S. B. Gabelli, P. A. Cole, Akt Kinase Activation Mechanisms Revealed Using Protein Semisynthesis. *Cell* **174**, 897-907 e814 (2018).
11. A. Jain, R. Liu, B. Ramani, E. Arauz, Y. Ishitsuka, K. Ragunathan, J. Park, J. Chen, Y. K. Xiang, T. Ha, Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484-488 (2011).
12. A. Jain, E. Arauz, V. Aggarwal, N. Ikon, J. Chen, T. Ha, Stoichiometry and assembly of mTOR complexes revealed by single-molecule pulldown. *Proc Natl Acad Sci U S A* **111**, 17833-17838 (2014).
13. E. Arauz, V. Aggarwal, A. Jain, T. Ha, J. Chen, Single-Molecule Analysis of Lipid-Protein Interactions in Crude Cell Lysates. *Anal Chem* **88**, 4269-4276 (2016).
14. W. I. Wu, W. C. Voegtli, H. L. Sturgis, F. P. Dizon, G. P. Vigers, B. J. Brandhuber, Crystal structure of human AKT1 with an allosteric inhibitor reveals a new mode of kinase inhibition. *PLoS One* **5**, e12913 (2010).

15. M. A. Lemmon, Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99-111 (2008).
16. L. R. Stephens, T. R. Jackson, P. T. Hawkins, Agonist-stimulated synthesis of phosphatidylinositol(3,4,5)-trisphosphate. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1179**, 27-75 (1993).
17. D. Balzano, M. A. Fawal, J. V. Velazquez, C. M. Santiveri, J. Yang, J. Pastor, R. Campos-Olivas, N. Djouder, D. Lietha, Alternative Activation Mechanisms of Protein Kinase B Trigger Distinct Downstream Signaling Responses. *J Biol Chem* **290**, 24975-24985 (2015).

CHAPTER 6. CONCLUSIONS AND PERSPECTIVES

The lack of biochemical assays for large-scale studies of protein-lipid interactions that can capture physiologically relevant interactions has been a major challenge in understanding the functional consequences of lipid-protein interactions. The majority of the commonly used lipid binding assays are not amenable for high-throughput studies to faithfully recapitulate physiologically relevant interactions. However, over the last two decades, several studies have attempted large-scale studies by painstakingly studying interactions of individual pairs of PH domains and phosphoinositides (PIP) by a specific biochemical method of choice. Additionally, efforts have been made to identify a sequence signature that can predict lipid-binding by the PH domain proteins.

In one of such studies, 33 yeast PH domains were tested for their PIP binding specificity employing lipid strips assay, surface plasmon resonance (SPR), and membrane targeting in yeast (1). They found 67% PH domains to bind all PIPs with no clear PIP binding preference and only one PH domain displayed highly specific PIP recognition. Molecular modeling of PIP binding PH domains suggested a correlation between electrostatic properties of the ligand binding site and lipid binding. Structural and biochemical studies of PH domains have identified a basic sequence motif “KX_n(K/R)XR” in the β 1- β 2 loop of the PH domain as a determinant of PIP binding (2, 3).

Another study of 130 mouse PH domains employing live-cell imaging and mitogen induced plasma membrane (PM) translocation of YFP tagged proteins, identified 27 proteins that translocated to PM after mitogenic stimulation; however, the specific PIP₃ recognition was not recapitulated in lipid strips assays (4). The lack of correlation between PIP binding in lipid strips assay and PM translocation argues for the untrustworthiness of

the lipid strips assay, which is unfortunately a rather commonly used method. Using the amino acid sequence of PM translocating PH domains, Park et al. (4) devised a recursive functional classification (RFC) strategy that can probabilistically take into account the contribution of all amino acid in the PH domain towards lipid binding. This approach cleanly separated the binding and non-binding proteins and predicted the PIP₃ binding of PH domain based on their entire amino acid sequence. Another recent study used mass-spectrometry to identify 405 proteins in PIP interactome employing pulldown in cell lysates by PIPs attached to agarose beads (5). However, only 33 PH domain containing proteins bound PIPs and 18 of those only bound PIP₃. Based on sequence alignment of these proteins, the authors came up with an extended PIP₃ specific sequence motif, which predicted 36 PIP₃ binding PH domain proteins in human proteome. Despite these previous efforts, the PIP binding properties of the majority of the full-length PH domain proteins is not known and neither are there any known sequence motifs that can accurately predict PIP binding.

In my study, taking advantage of the strengths of the SiMPull assay (6-8), I have tackled several challenges associated with traditional lipid binding assays. We have expressed EGFP-tagged full-length proteins instead of PH domain in mammalian cells, thus eliminating the need for protein purification. Expression of full-length proteins in mammalian cells helps maintain the post-translational modification and preserves native milieu for proteins in cell lysate. Use of lipid vesicles in this assay provides a superior mimic of cellular membrane compared to the 2-D PVDF membranes spotted with lipids or agarose beads. Single-molecule studies further provide unique insights into the binding dynamics and kinetics such as transient / stable binding or binding stoichiometry. We have also established that the lipid-SiMPull assay can detect lipid-protein interactions with

dissociation constants as high as 10 – 20 μ M, which should capture the majority, if not all, physiologically relevant lipid-protein interactions.

We probed the phosphoinositide binding of 67 full-length PH domain from human proteome with all seven PIPs, phosphatidic acid and phosphatidylcholine as a negative control. We found that 36 proteins bound PIPs with some specificity, where we found novel binding partners for 28 of these proteins. Twenty-seven proteins were found to not bind any of the lipids, which includes 11 proteins with previously reported PIP binding. The previously reported lipid binding for 10/11 of these proteins were based on isolated PH domains and/or lipid strips assays. Our results further show that the presence of phosphatidylserine did not affect the specificity of binding, in contrast to an earlier report (9). Previously described lipid-binding motifs in the PH domain did not explain results from lipid-SiMPull assay, which prompted us to search for better methods for PIP binding prediction. We used the recursive functional classification strategy (4) to predict PIP binding from our lipid-SiMPull assay results. The RFC analysis predicted that almost half of the PH domain proteins in human proteome should bind with some specificity towards PIPs. The predictions were further validated by PIP binding of 19 proteins in lipid-SiMPull assays from a random selection of 20 proteins predicted to bind. In this study, we have experimentally confirmed specific lipid-binding for 54 full-length PH domain proteins out of the 87 proteins examined.

The combined list of PIP binding proteins from lipid-SiMPull assay and prediction by the RFC algorithm provides a unique picture of the PIP binding behavior of the largest lipid binding family found in human genome, in stark contrast to the common belief in the field that most PH domain proteins do not bind PIPs with specificity (3, 10). The distribution of PIPs in cells and their synthesis and turnover still remains unknown for many PIPs owing to the lack of specific high affinity probes for many PIPs in cells (11-13).

The PIP binding specificity discovered in this study can be utilized to develop specific genetically encoded lipid biosensors to study cellular localization and associated functions. The biological relevance of reported PIP binding in this study can be probed utilizing chemical-genetic tools such as the FRB-FKBP dimerization system (14). The results from lipid-SiMPull assay would prove to be a great resource for studying diverse cellular processes such as PIP metabolism, lipid dynamics, organelle localizations, vesicular trafficking and associated pathophysiologies.

6. 1 References

1. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. DeWald, D. Murray, S. D. Emr, M. A. Lemmon, Genome-wide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol Cell* **13**, 677-688 (2004).
2. T. Balla, Inositol-lipid binding motifs: signal integrators through protein-lipid and protein-protein interactions. *J Cell Sci* **118**, 2093-2104 (2005).
3. K. Moravcevic, C. L. Oxley, M. A. Lemmon, Conditional peripheral membrane proteins: facing up to limited specificity. *Structure* **20**, 15-27 (2012).
4. W. S. Park, W. D. Heo, J. H. Whalen, N. A. O'Rourke, H. M. Bryan, T. Meyer, M. N. Teruel, Comprehensive identification of PIP3-regulated PH domains from *C. elegans* to *H. sapiens* by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).
5. S. Jungmichel, K. B. Sylvestersen, C. Choudhary, S. Nguyen, M. Mann, M. L. Nielsen, Specificity and commonality of the phosphoinositide-binding proteome analyzed by quantitative mass spectrometry. *Cell Rep* **6**, 578-591 (2014).
6. E. Arauz, V. Aggarwal, A. Jain, T. Ha, J. Chen, Single-Molecule Analysis of Lipid-Protein Interactions in Crude Cell Lysates. *Anal Chem* **88**, 4269-4276 (2016).
7. A. Jain, R. Liu, B. Ramani, E. Arauz, Y. Ishitsuka, K. Rangunathan, J. Park, J. Chen, Y. K. Xiang, T. Ha, Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484-488 (2011).
8. A. Jain, E. Arauz, V. Aggarwal, N. Ikon, J. Chen, T. Ha, Stoichiometry and assembly of mTOR complexes revealed by single-molecule pulldown. *Proc Natl Acad Sci U S A* **111**, 17833-17838 (2014).
9. I. Vonkova, A. E. Saliba, S. Deghou, K. Anand, S. Ceschia, T. Doerks, A. Galih, K. G. Kugler, K. Maeda, V. Rybin, V. van Noort, J. Ellenberg, P. Bork, A. C. Gavin, Lipid Cooperativity as a General Membrane-Recruitment Principle for PH Domains. *Cell Rep* **12**, 1519-1530 (2015).
10. M. A. Lemmon, Membrane recognition by phospholipid-binding domains. *Nat Rev Mol Cell Biol* **9**, 99-111 (2008).
11. M. Maekawa, G. D. Fairn, Molecular probes to visualize the location, organization and dynamics of lipids. *J Cell Sci* **127**, 4801-4812 (2014).
12. G. R. Hammond, T. Balla, Polyphosphoinositide binding domains: Key to inositol lipid biology. *Biochim Biophys Acta* **1851**, 746-758 (2015).
13. R. C. Wills, B. D. Goulden, G. R. V. Hammond, Genetically encoded lipid biosensors. *Mol Biol Cell* **29**, 1526-1532 (2018).
14. P. Varnai, B. Thyagarajan, T. Rohacs, T. Balla, Rapidly inducible changes in phosphatidylinositol 4,5-bisphosphate levels influence multiple regulatory functions of the lipid in intact living cells. *J Cell Biol* **175**, 377-382 (2006).