Directed Evolution of the Forkhead-associated Domain to Generate Anti-Phosphospecific Reagents

BY

KRITIKA PERSHAD
B.S., Aurora's Degree College, Osmania University, India, 2002
M.S., Osmania University, India, 2004

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biological Sciences
in the Graduate College of the University of Illinois at Chicago, 2012

Chicago, Illinois

Defense committee:

David E. Stone, Chair
Brian K. Kay, Advisor
Teresa V. Orenic
Hua Jin
Bernard D. Santarsiero, Medicinal Chemistry and Pharmacognosy

This thesis is dedicated to my husband, Aasheesh and to my parents for their constant support and encouragement, without whom this journey would have been impossible.

ACKNOWLEDGEMENTS

Dr. Brian Kay, my PhD advisor is a wonderful person. I am very lucky to be his student. I cannot thank him enough for everything he has done for me in the past six years. He always encouraged me, gave exciting research ideas and the freedom to plan and execute experiments. He has provided me with extraordinary opportunities of attending various conferences, giving research talks and collaborating with other labs. Brian has always challenged me to do my best and prove myself. I shall remember his words 'No Ventures No Gains'. I look up to him and hope that someday I will be at least half as accomplished as him.

I thank all my committee members, Dr. David Stone, Dr. Teresa Orenic, Dr. Hua Jin, Dr. Bernard Santarsiero and my previous committee members Dr. Susan Liebman, Dr. Don Morrison and Dr. Athar Chisti who have always given me expert scientific advice. Their suggestions have brought a lot of value to my research. The committee's expertise has helped me think about and answer questions which otherwise would have not been addressed. I thank Dr. Teresa Orenic for training me to use her microscope, Dr. Hua Jin for sharing her cell lines and her graduate student Livana Soetedjo for teaching me cell-staining and Dr. Jennifer Schmidt for allowing me to use her cell culture facility.

ACKNOWLEDGEMENTS (continued)

The Kay lab has been a wonderful place to work. The last six years has been a great learning experience for me personally & professionally. My friends in the Kay lab, Renhua, Gosia, Mike and Sujatha and the former lab members, Zengping and Sang have always been a great source of help, encouragement, support, suggestions and positive criticism. I thank all of them for making the lab a place I look forward to going each day. I also want to thank Karolina, an exchange student from Germany, who helped me with my research for two semesters.

Many thanks to all the awesome undergraduate students we had over the years who made our day to day research life easier.

I want to thank Dr. John McCafferty at the University of Cambridge, Cambridge, UK and his lab members for all their help and suggestions.

My family and friends in Chicago have helped me out on numerous occasions and I could share my despairs and delights with them. I want to thank everyone for their positive impact on me.

My many thanks to Omar, Corinna and Beth for answering all my non-scientific questions and organizing the everyday activities in LMB. The smooth functioning of LMB is a result of their hard work and dedication.

ACKNOWLEDGEMENTS (continued)

Lastly, I want to thank my husband, Aasheesh and my parents for their support and for putting up with my demands and needs each and every day. They play a major role in making this day come true in my life.

TABLE OF CONTENTS

CHAPTER PAGE

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

FIGURE PAGE

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS

pT	Phosphothreonine
pS	Phosphoserine
pY	Phosphotyrosine
pS/T	Phosphoserine/threonine
FHA	Forkhead-associated domain
GST	Glutathione S-transferase
PCR	Polymerase chain reaction
ELISA	Enzyme-linked immunosorbent assay
ITC	Isothermal calorimetry
k_d	Dissociation constant
IMAC	Immobilized metal affinity chromatography
HRP	Horse radish peroxidase
PBS	Phosphate buffered saline
PBST	PBS with 0.1% Tween 20
FTS	Fluorescent based thermal shift
LB	Luria Bertani
WT	Wild-type

LIST OF ABBREVIATIONS (continued)

PEG	Polyethylene glycol
NaCl	Sodium chloride
CB	Carbenicillin
Kan	Kanamycin
EGF	Epidermal growth factor
E. coli	*Escherichia coli*
S. cerevisiae	*Saccharomyces cerevisiae*
scFv	Single-chain fragments of variable region
Fab	Antigen binding Fragment
PBD	Phosphopeptide-binding domain
FN3	Fibronectin type III domain
PBD	Polobox-domain
PVDF	Polyvinylidene difluoride
DMEM	Dulbecco’s Modified Eagle Medium
BRCT	BRCA1 carboxy-terminal domains
BRCA1	Breast cancer susceptibility gene 1
WW	Tryptophan-tryptophan
SH2	Src homology 2
SH3	Src homology 3

LIST OF ABBREVIATIONS (continued)

COMP	Cartilage oligomeric matrix protein
IgG	Immunoglobulin
Fc	Fragment crystallizable
ERK	Extracellular signal-regulated kinases
MAPK	Mitogen-activated protein kinase
MEK	Mitogen-activated protein kinase kinase
HEK293	Human Embryonic Kidney 293 cells
ATF2	Activating transcription factor 2
siRNA	small interfering RNA
Cdc	cell division control

SUMMARY

Phosphorylation is a very important post-translational modification that has various functions in the cell, such as regulating the catalytic activity of proteins, generating binding sites for phosphopeptide-binding domains as well as disrupting protein-protein interactions, targeting proteins for phosphorylation-dependent degradation, and translocating proteins to their proper subcellular location.

With the development of sophisticated phosphopeptide enrichment and mass spectrometry techniques, thousands of phosphorylation sites have been identified in the phosphoproteome. Immunizing animals with synthetic phosphopeptides has been previously the standard route chosen for generating antibodies that recognize phosphorylated residues in proteins. While antibodies are valuable tools for monitoring protein phosphorylation and studying signaling events in cells, generating them through immunization of animals with phosphopeptides is expensive and time consuming. An attractive alternative is to use recombinant methods to generate antibodies or affinity reagents in less time. To explore such a strategy, we chose to engineer a naturally occurring phosphothreonine-binding domain, namely the Forkhead-associated domain (FHA1), from *Saccharomyces cerevisiae* Rad53 protein, generate variants of

this domain by oligonucleotide-directed mutagenesis, and isolate variants that bind specifically to phosphopeptides of interest.

To this end, we have displayed functional variants of the FHA1 domain on the surface of bacteriophage M13, improved its thermal stability, identified residues in the FHA1 domain that are important for recognizing its cognate pT-peptide, generated two phage-displayed libraries of FHA1 variants, screened a total of 11 different pT-containing peptides from which we successfully isolated anti-phosphospecific reagents to 7 phospho peptides. Our success rate for isolating phospho specific reagents from these libraries is about 60%.

For the first time it has been shown that the specificity of a phosphopeptide-binding domain can be engineered to recognize new phosphopeptides for which binding was not previously detected. The phage-displayed library of FHA1 variants is a useful resource to isolate anti-phosphospecific reagents, *in vitro,* by phage-display. These reagents are useful in binding their cognate phosphopeptides in ELISA. As these reagents have micromolar binding dissociation constants, we will use *in vitro* affinity maturation to improve the affinity of these reagents to broaden their applicability for use as reagents in Western blotting and pull-down experiments. Our research serves as a platform for exploiting other phosphopeptide-binding domains as scaffolds for generating

recombinant phosphospecific affinity reagents. These reagents will be an attractive alternative for antibodies, because they are renewable, have excellent expression in *E. coli* (20-25 mg/L), can be generated in a short time (<2 weeks), and are cost effective compared to immunizing animals.

CHAPTER 1

INTRODUCTION

Parts of the text have been published

Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display, *Protein Eng Des Sel* 23, 279-288.

Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

1. INTRODUCTION

1.1 Protein phosphorylation

In response to a stimulus, various signaling events are orchestrated within the cell during which the concerted action of protein kinases, protein phosphatases, and protein interaction domains are responsible for translating this signal into a specific biological response. The importance of phosphorylation events in the cell is revealed by the facts that nearly 1/3 of all the proteins in the cell are phosphorylated at any given time (*1, 2*) and 2.5% of the human genome encodes for protein kinases, protein phosphatases, and protein interaction domains (*3*). Of the various post-translational modifications that occur in a cell, such as methylation of arginine (Arg) or lysine (Lys) residues, acetylation, ubiquitylation or sumoylation of Lys residues, hydroxylation of proline residues (*4*) , the most abundant is phosphorylation of hydroxyl groups of serine (Ser), threonine (Thr) or tyrosine (Tyr) residues (*5*). About 76% of the phosphosites are pSer, 20% are pThr and 4% are pTyr (*6*).

Various protein isoforms, with distinct biological activity, are generated through phosphorylation of the same protein at different sites (*4*). Protein phosphorylation has various functions in the cell (*7*), such as 1) making proteins functionally active or inactive, 2) generating binding sites for phosphopeptide-binding domains, such as Forkhead-associated (FHA) domains that recognize

pT-containing motifs, 14-3-3 proteins, BRCA1 carboxy-terminal (BRCT) domains and Polo-boxes that bind to pT/S containing-peptides and SH2, or phosphotyrosine-binding domains that recognize pY-containing motifs (*8, 9*) (Fig. 1A), 3) regulating the activity of catalytic proteins by either activating them or by maintaining them in an inactive conformation (e.g., in Src family kinases), the interaction of the SH2 and SH3 domains with their cognate peptide-binding motifs (pY-containing motif and PXXP motif) restrain the Src kinase in an inactive conformation (*10*) (Fig. 1B), 4) targeting proteins for degradation (e.g., phosphorylated cyclin-E creates a binding site for the WD40X8 repeat domain of an F-box protein which then targets cyclin-E for protein degradation by the E3 ligase complex) (*11, 12*) (Fig. 1C). Similarly, binding of WD40X8 repeat of an F-box containing protein (Cdc4) to its multiply phosphorylated substrate (Sic1), a cyclin dependent kinase inhibitor, targets Sic1 for polyubiquitylation and degradation (*13, 14*), and 5) disrupting protein-protein interactions (e.g., phosphorylation of a specific Ser residue on histone 3 inhibits the interaction between the tri-methylated Lys residue of histone 3 and the Chromodomain of a heterochromatin protein-1, leading to changes in gene expression) (*15*) (Fig. 1D). In another example, when a Ser residue in the C-terminal peptide motif of the NR2B subunit of NMDA receptor is phosphorylated, it blocks the interaction of two PDZ domain containing proteins (PSD-95 and SAP102) with the receptor, which is a mechanism to regulate protein trafficking and cell surface expression

of the receptor (*16*), and 6) translocating proteins to their proper subcellular location. In addition to phosphorylation of proteins, lipids are often phosphorylated (e.g., phosphoinositides and diacylglycerol) and recognized by phospholipid-binding domains.

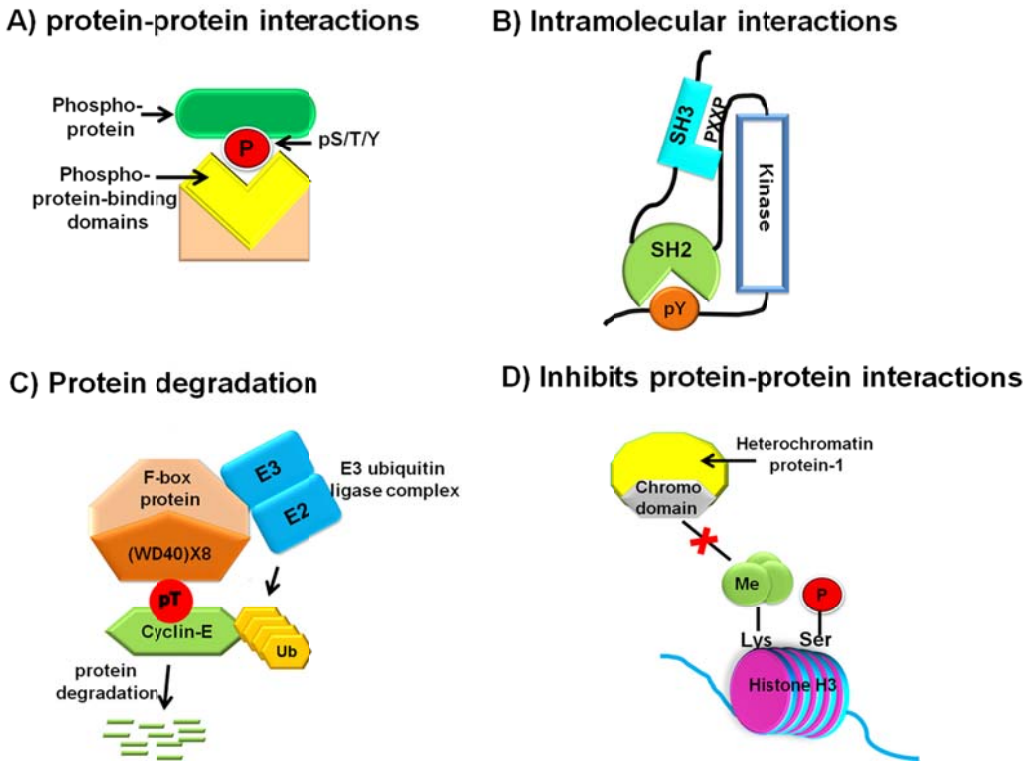


Figure 1. Functional consequences of protein phosphorylation. A) Protein-protein interactions**.** Phosphorylation on Ser, Thr or Tyr residues in proteins creates binding sites for various pS/T/Y-binding domains. B) Intramolecular interactions. In Src tyrosine kinases, phosphorylation promotes intramolecular interaction between the SH2 domain and a pY-containing motif and simultaneously the SH3 domain interacts with its proline-rich peptide ligand (PXXP), and both these intramolecular interactions maintain the protein kinase in an inactive conformation. C) Protein degradation. The WD40X8 repeat domain of an F-box protein interacts with phosphorylated cyclin-E and targets it for polyubiquitylations by the E3 ubiquitin ligase complex and subsequent degradation by the proteasome complex**.** D) Inhibits protein-protein interactions. Phosphorylation of a Ser residue in histone H3 inhibits the interaction between the chromodomain of heterochromatin protein-1 (a domain that recognized methylated residues in proteins) and the methylated Lys residue. This interaction is restored when the phosphate group is removed.

Protein phosphorylation regulates various processes such as gene expression, cytoskeletal rearrangements, cell cycle progression, DNA repair, and apoptosis (*8*). The human genome encodes for about 500 protein kinases and a third of that number for protein phosphatases (*17*). Defective expression of kinases and phosphatases or mutations in phosphopeptide-binding domains is the cause of a number of human diseases (*18*). Over the years, there has been great interest in unraveling the phosphoproteome by constantly improving the phosphopeptide enrichment and detection technologies (*19, 20*). According to a bioinformatics analysis, available at http://www.phosphonet.ca, more than 650,000 potential phosphorylation sites have been predicted in the human proteome, of which 10,000 sites have already been identified by mass spectrometry (*21*).

1.2 Protein interaction domains

Protein interaction domains play an indispensible role in signaling events with their diverse ligand recognition ranging from recognizing specific peptide sequences, to lipids to DNA, and some of them are dedicated for the recognition of post-translational modifications such as acetylation, methylation or phosphorylation (Table I). Protein interaction domains are mostly non-catalytic domains that are conserved in structure with varying levels of sequence similarity (e.g., FHA domains have 20-30% sequence similarity (*22*), and the SH2 and SH3

domains have 20-90% sequence identity (*23*)). Protein domains are present in almost 70% of the proteins in the human proteome (*24*), and at least 155 domains, which are involved in various cellular signaling pathways, have been identified in the SMART database (http://smart.embl-heidelberg.de/).

Table I. Protein interaction domains with diverse ligand specificities.

Protein domains	Ligands	References
SH3	Proline-rich peptides (PXXP, KKPP, PXXDY motifs) RXXK motif, surface of ubiquitin	(*25-27*)
WW domains	Proline-rich sequences (PPXY, PPLP, PPR) pS-Pro and pT-Pro motifs, and pSPXpSP	(*28-32*)
PDZ	C-termini of proteins, phospholipids	(*33, 34*)
SH2	pY-containing peptides- pY-hydrophobic-X-hydrophobic and pY-hydrophilic-hydrophilic-I/P	(*35-37*)
PTB	pY-containing peptides (N-P-X-pY) and non-phosphorylated peptides (N-X-X-Y/F)	(*38-42*)
14-3-3	pS/T-containing peptides-R(S/ϕ)X(pS/pT)XP, RX(ϕ/S)X(pS/pT)XP and SWpTX, non-phosphorylated peptides	(*43-47*)
BRCT	pS/T-containing peptides with a stronger preference for pS-containing peptides- pSXXϕ motif	(*48, 49*)
FHA	pT-containing peptides-pTXX(I/L/D), pTXXpTXX	(*50-53*)
WD40 repeats	pS/T-containing motifs-(L/I)-(L/I/P)-(pS/pT)-P and DpSGXXpS	(*13, 14, 54*)
Bromo domains	Acetylated lysine residues	(*55*)
Chromo	Methylated lysine residues	(*56*)
PH, ENTH, PX	Phosphoinositides	(*57-59*)

SH3 = Src homology 3 domain, WW = named after the two conserved Tryptophan residues, PDZ = PSD-95/Dlg/ZO-1, SH2 = Src homology 2 domain, PTB = phospho-tyrosine binding domain, 14-3-3 = pSer/Thr-binding proteins, BRCT = breast cancer susceptibility gene, C-terminal domain, FHA = Forkhead-associated domain, PH = pleckstrin homology domain, ENTH = epsin N-terminal homology domain, PX = phagocyte NADPH oxidase (phox) complex, known as Phox-homology domain

Even though protein interaction domains are components of a larger protein, they are independently folding protein modules that remain functional when expressed as discrete proteins. This has allowed defining the binding specificity and identifying binding partners of various interaction domains such as PDZ domains, SH2 and phosphotyrosine-binding (PTB) domains, SH3 domains, FHA domains, etc., by screening combinatorial peptide libraries, protein or peptide arrays and mass spectrometry (*60*). A single protein can contain multiple domains; for example, the Rad53 Ser/Thr protein kinase contains two FHA domains, two SQ-TQ cluster domains and a kinase domain, the human MDC1 and NBS1 proteins have two BRCT domains and one FHA domain, the 53BP1 had one Tudor domain and two BRCT domains) (http://pfam.sanger.ac.uk) (Fig. 2). Cooperative interactions of multiple domains with their ligands is one of the mechanisms that generates a specific signaling event in the cell, in addition to other factors such as subcellular localization of the protein and its ligands, expression levels under different stimulation or stress conditions and negative feedback regulation that also contribute to generating a specific output signal in response to an external or internal stimulus.

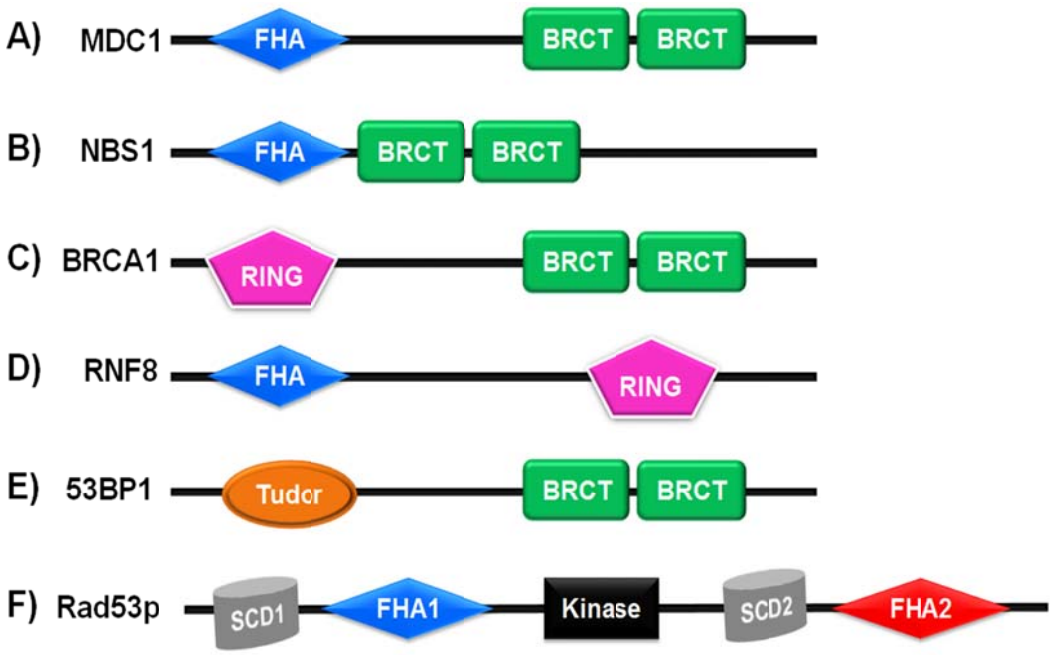


Figure 2. Multiple domains present in a single protein. Shown here are examples of six proteins involved in DNA damage repair. Each protein comprises of a unique combination of protein domains which have different ligand recognition specificities or catalytic activity. The Forkhead-associated (FHA) and BRCA1 C-terminal (BRCT) domains are pS/T-binding domains. The SQ-TQ cluster domains (SCD) are protein regions that are concentrated with Ser and Thr residues and are phosphorylated at multiple sites by upstream kinases such as Mec1 and Tel1 in yeast and ATM and ATR like kinases in humans. The RING finger domain is a protein interaction domain which is known to have E3 ubiquitin-protein ligase activity. Tudor domain is present in many RNA and DNA-binding proteins and may play a role in binding to nucleic acids. The figure is adapted from (*61*).

1.3 Phosphopeptide-binding domains

Phosphopeptide-binding domains (PBDs) are divided into two categories, 1) pSer/Thr-binding domains (BRCT domains, WW domains, FHA domains, 14-3-3 proteins and Polo-box domains) and 2) pTyr-binding domains (SH2 domains and phosphotyrosine-binding (PTB) domains). There are various mechanisms by which PBDs recognize their cognate phosphorylated peptide ligands, such as, 1) two PBDs present in tandem can bind to two phosphopeptides simultaneously (e.g., the two SH2 domains in ZAP-70 protein tyrosine kinase binds two pY-containing peptides) (*62*), 2) a single PBD can bind to a doubly phosphorylated peptide due to the presence of two phosphopeptide-binding pockets (e.g., the FHA domain of Dun1 protein kinase binds a dual phosphorylated pT-containing peptide) (*53*), 3) two PBDs present in tandem can be involved in binding to the same phosphopeptide (e.g., BRCT domains) (*49*), 4) a PBD can form non-covalent dimers and interact with two phosphoepitopes simultaneously (e.g., 14-3-3 proteins form dimers that interact with two pT/S-containing peptides) (*63*), and 5) a single PBD can bind to multiple phosphorylation sites on a protein (e.g., WD40 repeat domain of Cdc4 recognizes its target only when it is phosphorylated on six or more pT/S residues) (*64*). The most commonly occurring phosphopeptide-binding domains that are found in a variety of proteins are discussed below.

1.3.1 Src homology 2 domains

SH2 domains were first discovered in the 1980s as non-catalytic domains that mediate protein-protein interactions (*65, 66*). SH2 domains were first identified as conserved protein regions present in protein tyrosine kinases belonging to the Src family (*65*). SH2 domains constitute the largest class of pTyr-binding domains, with 120 SH2 domains encoded by the human genome (*67, 68*). SH2 domains are structural components of a wide variety of kinases, phosphatases, adaptors, and transcription factors, and play important roles in cell signaling (*69, 70*). SH2 domains are composed of ~100 residues that fold into anti-parallel β sheets, flanked by two α-helices. Like many other protein interaction modules, they fold independently into functional modules, with their N- and C-termini on the opposite side to the pTyr-binding surface. Even though SH2 domains are structurally similar, they bind to different pTyr-containing peptide motifs. The core-binding motif is typically 3-6 residues, positioned C-terminally to the pTyr residue. The optimal binding motifs for various SH2 domains have been identified by screening combinatorial peptide (*36*) and oriented peptide array libraries (*37*). This information is valuable for the identification of putative binding partners for SH2 domain containing proteins of unknown function.

1.3.2 14-3-3 proteins

14-3-3 proteins are pS/T-binding proteins (*44, 71*), which are found exclusively in eukaryotic proteins (*72*). While performing DEAE-cellulose chromatography, 14-3-3 proteins were eluted as fraction no. 14 and in starch-gel electrophoresis they were located at position 3.3, and therefore they were named 14-3-3 proteins (*73*). There are 7 isotypes of 14-3-3 proteins in mammals (β, α, e, η, σ, t, and ς), which form homodimers or heterodimers with other members within the 14-3-3 family. Each monomer, consisting of nine α-helices, binds to one phosphopeptide, thus there are binding sites for two phosphopeptides in the 14-3-3 dimer (*74*) (Fig. 3A). Either addition of a competing pS-containing peptide or incubation with a phosphatase will dissociate 14-3-3 zeta protein from Raf protein, indicating a phosphoserine specific interaction (*75*). Through screening of combinatorial phosphoserine peptide library, the optimal binding sequences for 14-3-3 proteins have been determined, such as RSX(**pS/pT)**XP, RXXX(**pS/pT)**XP or SW(**pT)**X, where X is any amino acid (*46, 47*). The main function of 14-3-3 proteins is in DNA repair, where they maintain Cyclin-Cdk1 in a phosphorylated and inactive state, by interacting with phosphorylated Cdc25 family of phosphatases and sequestering them in the cytoplasm (*76*). More than 100 signaling proteins have been identified as binding ligands for 14-3-3 proteins, which include kinases, phosphatases, and membrane receptors (http://pfam.sanger.ac.uk/family/14-3-3). Unlike other phosphopeptide binding

domains, such as FHA domains, BRCT domains, Polo-box domains, SH2 domains, etc., which are structural components of a larger protein, 14-3-3 proteins are proteins that can form homo- or hetero-dimers with their other isoforms.

1.3.3 Tryptophan-tryptophan domains

WW domains are small protein modules of about 40 amino acids found in many signaling proteins, proline isomerases, ubiquitin ligases, and cytoskeletal proteins (http://smart.embl-heidelberg.de/). WW domains contain two conserved Tryptophan (W) residues that are usually 20-23 residues apart in the primary sequence, hence the name WW domains. WW domains are composed of three anti-parallel β-strands (*77*) (Fig. 3B), and their ligand specificity ranges from recognizing proline-rich motifs to phosphopeptides with the general motif (pSer/Thr)-Pro (*29, 30*). The WW domain from proline isomerase Pin1 binds to substrates of Pin1 only after the (pSer/Thr)-Pro bonds have been isomerized by the proline isomerase from cis to trans configuration (*78*). Solving the crystal structure of a dual phosphorylated peptide (Y(pS)PT(pS)PS) from RNA polymerase II binding to the WW domain from proline isomerase Pin1 revealed that only one phosphate group contributes to binding, and that the major binding determinant in the Pin1 WW domain is an Arg residue in the β1-β2 loop (*31*). This Arg residue is not present in most of the WW domains, which explains why

only a few selected WW domains are involved in phosphospecific interactions. Therefore, WW domains from different proteins can recognize both phosphorylated and non-phosphorylated motifs with the requirement of a Pro residue in the binding motif.

1.3.4 WD40 repeat domain

WD40 repeats are found in F-box containing proteins that are components of the E3 ubiquitin ligases. The WD40 repeats bind to their substrates in a phosphorylation dependent manner (pSer/Thr-containing motifs), initiating polyubiquitylation and degradation of their targets (*79-81*). Coupling phosphorylation with polyubiquitylation for protein degradation is commonly seen in the degradation of cyclins (*11, 12*) and cyclin dependent kinase (Cdk) inhibitors, during cell cyclin progression (*13, 14*). WD40 repeats form a seven or eight bladed β-propeller structure with each blade usually consisting of four β-strands (*13, 82*). The crystal structure of a pT-containing peptide from cyclin E has been solved in complex with the WD40X8 repeat domain of F-box protein Cdc4 (*13*), and is shown in Fig. 3C. The pT-peptide occupies the pore formed in between the eight blades and interacts with the WD40 domain by electrostatic and hydrophobic interactions. The optimal binding motif for the WD40 repeat domain of Cdc4, (L/I)-(L/I/P)-pS/pT-P has been determined by screening peptide arrays (*13*).

1.3.5 Carboxy-terminal domains of BRCA1 protein

BRCT (BRCA1 C-terminal) domains are found in many proteins involved in DNA repair (*83*). BRCT domains bind both to pS and pT-containing peptides, and tighter binding is observed for pS-containing peptides (*48, 84*). The BRCT domains are often present as tandem repeats, and the BRCT domains interact with each other to form a phosphopeptide-binding grove/cleft which accommodates a single pS/T- containing peptide (*85*). Each BRCT domain contains a β-sheet composed of four β-strands which are flanked by three α-helices (*86*) (Fig. 3D). The optimal binding motif for BRCT domains is pSXXϕ (where ϕ is an aromatic amino acid) (*48, 49*). BRCT domains bind both to pS and pT-containing peptides (*48*). Most of the breast cancer causing mutations in the BRCA1 protein are found in its BRCT domains (*87*).

1.3.6 Polo-box domain

Polo-like kinases are composed of catalytic N-terminal Ser/Thr kinase domain and C-terminal polo box domain that is composed of two polo boxes, PB1 and PB2 that are structurally similar but have only 20-30% sequence identity (*88*). Six β-strands from each polo-box are packed together into a β-sandwich, which is flanked by three α-helices (*89*) (Fig. 3E). Human Polo-like kinase 1 (Plk1) regulates the different stages of mitosis and cytokinesis during cell division

(*90*), and abnormal function or over expression causes cancer. Currently, a number of cancer treatments are aimed at inhibiting Plk (*91*).

1.3.7 Forkhead-associated domain

The FHA domain was first discovered by Hofmann and Bucher in 1995, in the Forkhead family of transcription factors, and hence the name, Forkhead-associated domain (*92*). FHA domains are found in a wide variety of proteins, such as protein kinases, protein phosphatases, kinesins, transcription factors, glycoproteins, etc. that are involved in various biological processes such as signal transduction, protein transport, transcription, DNA repair, protein degradation, etc. (http://smart.embl-heidelberg.de/) (*93, 94*). The FHA domains are phosphothreonine-binding domains that are composed of 100-180 residues, and are present in prokaryotic as well as eukaryotic proteins. They are composed of 11 or 12 β-strands, which are connected by loops to form a β-sandwich (Fig. 3F). In the SMART (Simple Modular Architecture Research Tool) database (http://smart.embl-heidelberg.de/) there are 5457 FHA domains in 5145 proteins, of which 2476 proteins are from bacteria, 2635 from eukaryotes and the remaining are found in archaea bacteria. In *Saccharomyces cerevisiae,* there are 13 FHA domain containing proteins, and in humans 65 proteins contain FHA domains.

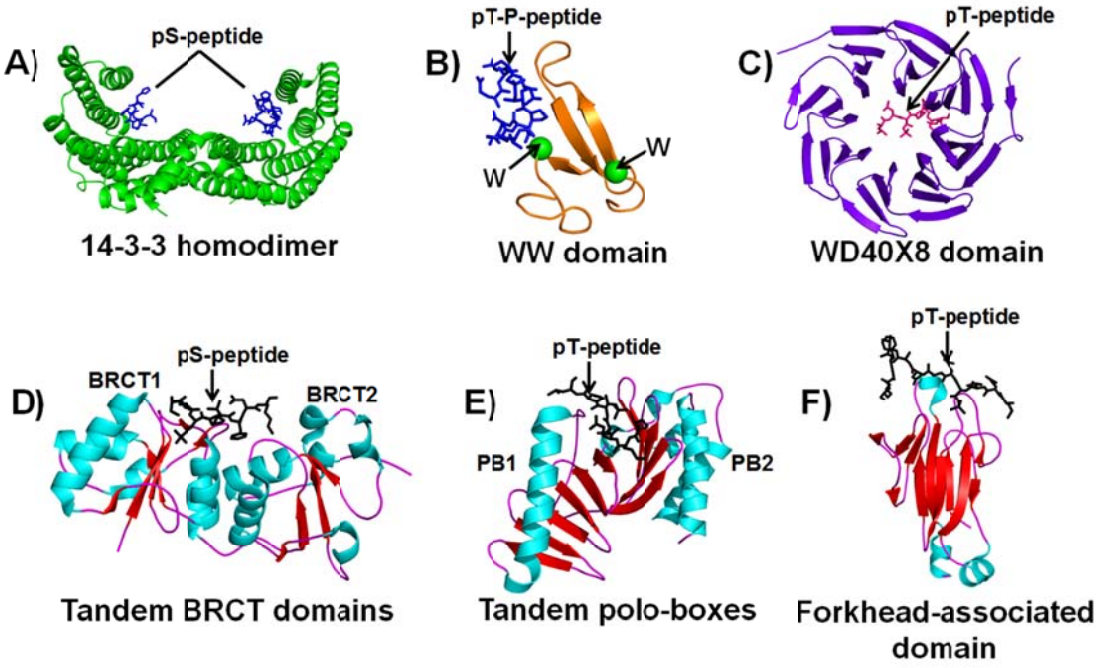


Figure 3. Structures of phosphoserine/threonine-binding domains. Structures are generated using PyMol software (www.pymol.org) and their Protein Data Bank (PDB) coordinates are noted. A) 14-3-3 protein sigma homodimer binding to two same pS peptides (PDB code: 1YWT) (*95*). B) WW domain bound to a pT-P peptide, the Tryptophan (W) residues are shown as green spheres (PDB code: 1I8G) (*96*).C) WD40X8 repeat domain bound to a pT-containing peptide (PDB code: 1NEX) (*13*). D) Tandem BRCT domains (BRCT 1 and BRCT2) bound to a single pS-containing peptide (PDB code: 1T15) (*86*). E) Tandem polo-boxes (PB1 and PB2) bound to a single pT-containing peptide (PDB code: 1UMW) (*88*). E) Forkhead-associated domain complexed with a pT-containing peptide (PDB code: 1G6G) (*22*).

1.4 Forkhead-associated domain 1 of *Saccharomyces cerevisiae* Rad53 protein

In the budding yeast, *S. cerevisiae*, damage to DNA leads to a cascade of phosphorylation events that stop cell cycle progression and allow for repair of DNA. Rad53 protein kinase plays a large role in the DNA repair process in budding yeast (*97*). Rad53 protein contains two FHA domains, an N-terminal FHA1 domain and a C-terminal FHA2 domain flanking a central S/T kinase domain (Figure 1.4). Upon DNA damage, Rad53 protein interacts with phosphorylated Rad9 protein (*98*), and from mutational analysis it was learned that the FHA2 domain is responsible for the interaction of Rad53p with phosphorylated Rad9p (*99*). Later, it was found out by Durocher *et al.* (*100*) that the Rad53p FHA1 domain reacts with Rad9p as well. The interaction between the FHA1 domain of Rad53 protein and phosphorylated Rad9 protein could be disrupted by a pT-containing peptide from p53 protein which competes for binding to the FHA1 domain in a phosphorylation dependent manner. The non-phosphorylated peptide did not compete for binding to the FHA1 domain and the Rad53-Rad9 protein complex remained intact (*100*). This was the first time that it was demonstrated that FHA domains directly recognize phosphorylated epitopes, much like the SH2 domains that recognize phosphotyrosine residues.

1.5 Structure of Forkhead-associated domains

FHA domains, like other protein interaction domains, are independently folding modules, which can be overexpressed and purified from *Escherichia coli (22).* The three-dimensional structure of the first FHA domain to be solved was the FHA2 domain of Rad53p (*101*). Since then, the structure of 23 (http://pfam.sanger.ac.uk/) FHA domains from various species has been solved by NMR or X-ray crystallography. The crystal structures of FHA1 and FHA2 domains of Rad53 protein complexed to their cognate pT and pY-containing peptides from Rad9 protein have also been solved (*22, 50, 102, 103*). While the FHA domains do not share much sequence similarity (20-30%), they are structurally similar and are composed of 11 or 12 β-strands connected by loops to form a β-sandwich (Fig. 4A-C). The loops vary in length, sequence and structure and determine the ligand binding specificity. For instance, the Rad53 FHA1 domain and human Chk2 FHA domain have helical insertions in the β2-β3 and β4-β5 loops respectively (*22*). Due to this helical insertion, the β4-β5 loop of Chk2 FHA domain is longer and has a different conformation, which results in a different binding specificity (*51*). From the structural information of various FHA domains complexed with their cognate pT peptides (*22, 51, 53, 102*) (http://pfam.sanger.ac.uk/), (www.pymol.org), it is evident that residues from four loops β3-β4, β4-β5, β6-β7 and β10-β11 interact with the pT peptide. Residues from the β3-β4 and β6-β7 loops are mainly involved in making contacts with the

phosphate group, and the γ-methyl group of threonine which confers specificity for binding to pT versus pS peptides. The residues in the β4-β5 and β10-β11 loop also interact with residues N- and C- terminal of the pT residue and contribute to the binding specificity. Only six residues are conserved among FHA domains from various organisms (G69, R70, S85, H88, N107, and N112), some of which are important for binding to the phosphopeptide (R70, S85, N107) and the others are required for structural stability (G69 and H88) (Fig. 4D) (*22*). FHA domains specifically recognize pT-containing peptides and not those in which the pT is replaced by either pS or Asp (*94*); this binding specificity is presumably due to the additional contacts made by the pT γ-methyl group with various residues (N107, S82, R83, L84) in the FHA1 domain (*22, 104*).

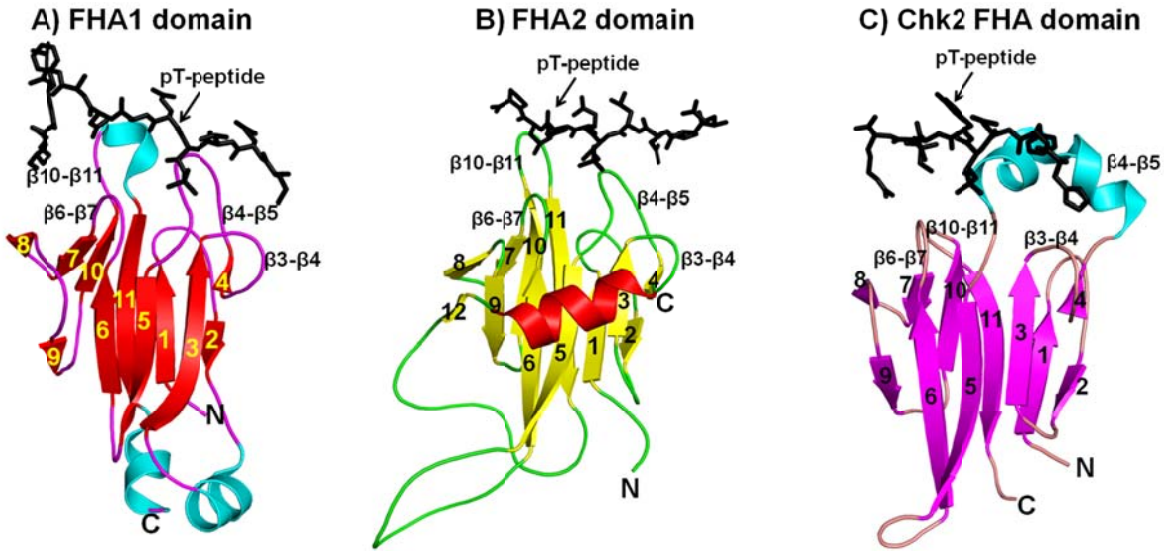
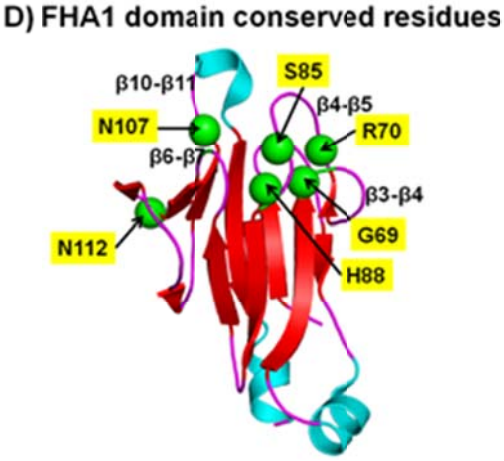


Figure 4. Structures of FHA domains. FHA domains are composed of 11 or 12 β-strands which are connected by loops to form a β-sandwich structure. The β-strands are numbered from 1 through 12. The N- and C-terminus are present on one side of the structure opposite to the pT-peptide binding site. The β3-β4, β4-β5, β6-β7 and β10-β11 loops that are involved in binding to the pT-peptide ligand are shown. The pT-peptide is shown as sticks and the FHA domains are shown in a cartoon representation using PyMol software (www.pymol.org) and their Protein Data Bank (PDB) coordinates are noted. A) FHA1 domain of *S. cerevisiae* Rad53p (*22*) (PDB code: 1G6G). B) FHA2 domain of *S. cerevisiae* Rad53p (*102*) (PDB code: 1K2N). C) FHA domain of human Chk2 protein kinase (*51*) (PDB code: 1GXC).



D) Six residues that are conserved among FHA domains from prokaryotes and eukaryotes are shown as green spheres. These residues either play a role in pT-peptide recognition or are important for the structural stability of FHA domains.

1.6 Ligand specificity of Forkhead-associated domains

From screening combinatorial phosphopeptide libraries, it was demonstrated that the pT (+3) position (i.e., the residue at the +3 position C-terminus of the pT moiety) was a major determinant of binding specificity, and that all the FHA domains prefer a specific residue at the pT (+3) position; this was called the pT (+3) rule (Table II). These screens revealed that the FHA1 and FHA2 domains preferentially recognized **pT**XXD and **pT**XX(I/L) motifs (*22*) (Fig. 5). Structural studies and site directed mutagenesis have revealed that residue R83 (located in the β4-β5 loop) from the FHA1 domain makes ionic interactions

with the pT (+3) Asp (D) and is a major determinant of the binding specificity (*22*). It is interesting to note that the Chk2 FHA domain does not have a preference for phosphopeptides with Asp in the pT+3 position, in spite of the presence of two Arg residues within the β4-β5 loop. It has been observed that the ChK2 FHA domain has an additional 8 residues in the β4-β5 loop, which forces this loop away from the pT+3 residue of the phosphopeptide, and results in preference for a hydrophobic residue in the pT(+3) position. This +3 residue fits into a hydrophobic pocket formed from the residues within β4-β5, β6-β7 and β10-β11 (*51*). The residues in the β10-β11 loop of the FHA1 domain have also been shown to be important determinants of binding specificity. Mutating two residues in the FHA1 domain (G133I/G135D) changes its specificity from binding its cognate peptide with Asp in the pT (+3) position to binding to a pT-containing peptide with a hydrophobic residue in the pT (+3), i.e., the binding specificity of the FHA1 domain is switched to be more like FHA2 (*105*). Another interesting observation was that the human Ki67 FHA domain did not recognize short phosphopeptides, but could bind to longer versions of the same phosphopeptide (44 residues), suggesting that some interactions require an extended binding surface to form stable FHA-phosphopeptide complexes (*106*). The FHA domain from yeast Dun1 protein binds to a doubly phosphorylated peptide due to the presence of two phosphopeptide-binding pocking in the FHA domain structure (*53*). These observations suggest that the FHA1 domain can bind to a variety of

phosphopeptides with different sequences, making it a suitable scaffold for library construction.

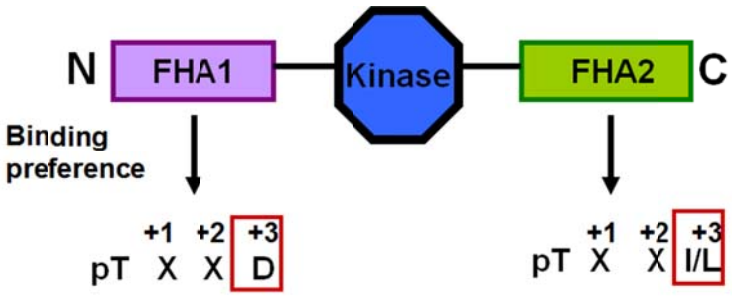


Figure 5. The architecture of the Rad53 protein. Rad53p has two phosphopeptide-binding domains, the N-terminal FHA1 domain and C-terminal FHA2 domain flanking a central Ser/Thr kinase domain. The FHA1 domain has a preference for peptide sequences with Asp (D) in the pT (+3) position, whereas the FHA2 domain prefers hydrophobic residues, such as Ile (I) or Leu (L) in the pT (+3) position.

Table II. Ligand specificity of FHA domains.

Organism	FHA domain	pT(+3) specificity	Reference
Saccharomyces cerevisiae	Rad53p FHA1	Asp	(*22, 50*)
Saccharomyces cerevisiae	Rad53p FHA2	Ile/Leu	(*22, 52*)
Saccharomyces pombe	Cds1 FHA	Asp	(*22*)
Arabidopsis thaliana	KAPP FHA	Ser, Ala	(*22*)
Homo sapiens	Chk2 FHA	Ile	(*51*)
Homo sapiens	RNF8 FHA	Tyr, Phe	(*107*)

1.7 Antibodies against phosphoepitopes

Sophisticated phosphopeptide enrichment and mass spectrometric techniques have led to the identification of more than 100,000 phosphorylation sites in the human proteome. To study phosphorylation dependent signaling events in the cells or for the purpose of phosphoprotein pull-down from cell lysates for analysis by mass-spectrometry, the most commonly used method to generate anti-pY and anti-pS/T antibodies has been to immunize animals, such as mice and rabbits, with synthetic phosphopeptides (*108*). The resulting antibodies, once purified from serum, are invaluable for studying the phosphorylation of proteins in cells, and for unraveling biologically important signal transduction pathways. However, with the existence of thousands of

phosphorylation sites in the human proteome, immunizing animals with a specific phosphopeptide sequence is time consuming, expensive and laborious.

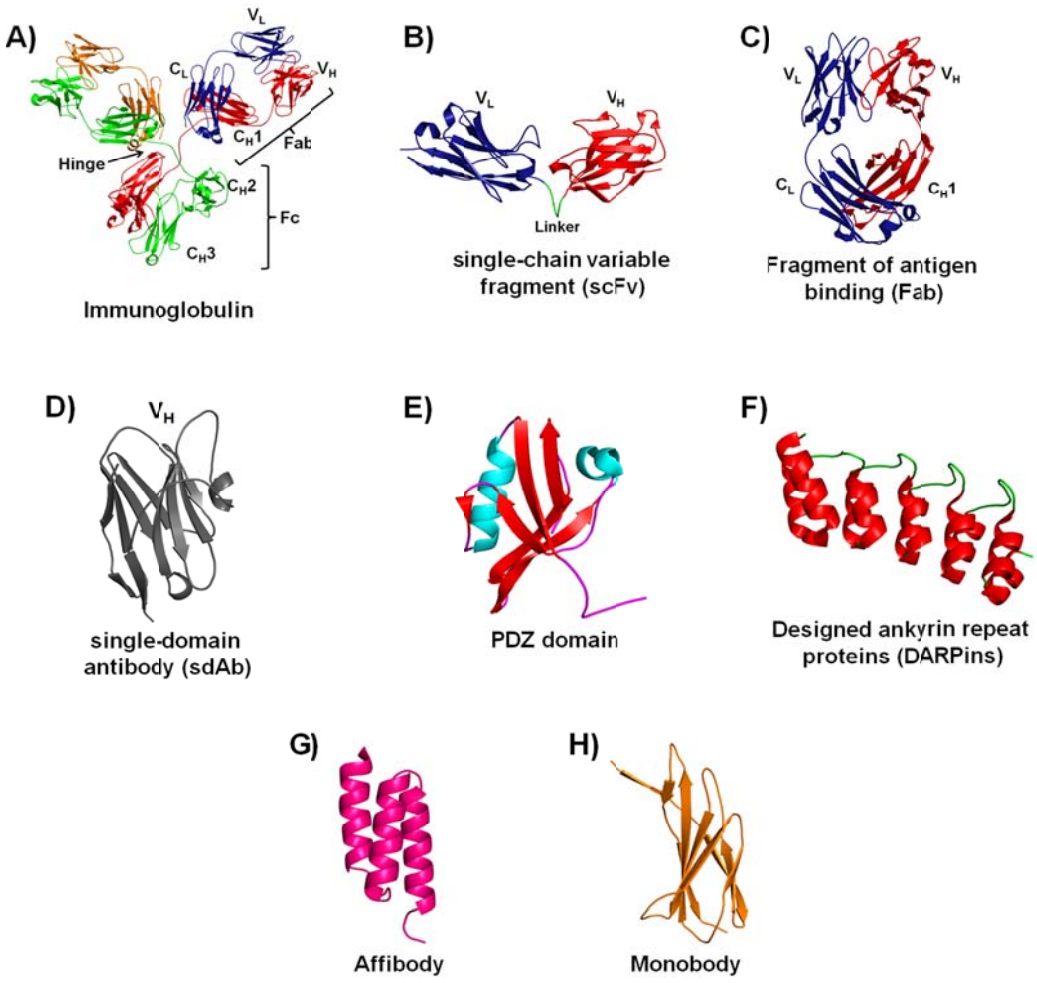
An ideal alternative would be to bypass the immunization step and produce antibodies in less time by using recombinant techniques. My proposal is to utilize the Forkhead-associated (FHA) domain 1, a naturally occurring pT-peptide binding domain, from the *S. cerevisiae* Rad53 protein, as a scaffold for generating new anti-phosphopeptide binding specificities. My goal is to generate new affinity reagents to various pT-containing peptides through affinity selection of a phage library displaying FHA1 domain variants which have been randomized at selected positions by oligonucleotide-directed mutagenesis (*109, 110*). The resulting FHA1 affinity reagents will be evaluated in enzyme linked immunosorbent assays (ELISA), Western blots, and cell-staining experiments. The advantages of using the FHA domain as a scaffold is that the affinity reagents derived from it are monoclonal and can be overexpressed in *E. coli.* In the future, other naturally occurring pS/pT and pY- binding domains may be used as scaffolds from which additional affinity reagents can be generated.

1.8 Phage-display

Antibodies generated by immunizing animals are polyclonal and, unfortunately, are not a renewable source of an affinity reagent. Consequently, hybridomas (*111-113*) were developed, which can be stored frozen and grown up

in large quantities for secreting monoclonal antibodies. However, with the existence of thousands of proteins in the human proteome, these methods for generating antibodies by animal immunization or hybridoma production are expensive, time consuming and laborious. An attractive alternative is to generate 'antibody-like reagents' or 'recombinant affinity reagents' through one of several display technologies. In 1985, it was discovered for the first time that a protein can be displayed on the surface of bacteriophage (*114*). Since then, libraries of single-chain variable fragments (scFvs) (*115, 116*), single domain antibodies (sdAbs) (*117*), Fragments of antigen binding (Fab) (*118-120*), bivalent Fab ($F(ab)_2$) (*121, 122*), Fibronectin type III domain (FN3) (*123, 124*), combinatorial peptides (*125-127*), PDZ domains (*128, 129*), designed ankyrin repeat proteins (DARPins) (*130-132*), and Affibodies (*133, 134*) have been displayed on the surface of bacteriophage M13 (Fig. 6). From large (i.e., 10^9 to 10^{10}) libraries of variants of a particular scaffold, one can affinity select individual clones that bind selectively and with nanomolar to micromolar affinities for particular targets.

Figure 6. Structures of recombinant affinity reagents that have been displayed on bacteriophage. Structures are generated using PyMol software and their Protein Data Bank (PDB) coordinates are noted. A) IgG contains two heavy chain and two light chains. The light chain contains one variable region (V_L) and one constant region (V_H). The heavy chain comprises of one variable region (C_L) and three constant regions (C_H1, C_H2 and C_H3). The V_H and V_L together constitute the antigen binding site (*135*) PDB:1IGY. B) The V_H and V_L of an antibody are joined by a flexible linker to form the single-chain fragment of variable region (scFv) which retains antigen-binding properties (*136*) PDB: 1MOE. C) Fragments of antigen binding (Fab) contain the antigen binding V_H and V_L and two constant regions (C_L and C_H1) (*137*) PDB: 1N8Z. D) single-domain antibodies (sdAbs) contain only the V_H which is sufficient for antigen recognition (*138*) PDB: 1MEL. E) PDZ domains are 90 residue repeats that consist of six β-strands and two α-helices (*139*) PDB: 2IWO. F) Designed ankyrin repeat proteins (DARPins) consist of 4-6 repeats, with each repeat comprising of a β-turn, followed by two α-helices which are connected by a loop to the next repeat (*140*) PDB: 1MJO. G) Affibody is made of 3 helices and it is a variant of domain *Z*, which is an Fc recognizing domain in *Staphylococcus aureus* protein G (*141*) PDB: 1LP1. H) Monobody is derived from the 10^{th} repeat of the fibronectin type III domain (*142*) PDB: 3K2M. Not drawn to scale.


A)
V_L
C_L
V_H
C_H1
Fab
Hinge
C_H2
Fc
C_H3
Immunoglobulin
B)
V_L
V_H
Linker
single-chain variable fragment (scFv)
C)
V_L
V_H
C_L
C_H1
Fragment of antigen binding (Fab)
D)
V_H
single-domain antibody (sdAb)
E)
PDZ domain
F)
Designed ankyrin repeat proteins (DARPins)
G)
Affibody
H)
Monobody

Display on the surface of bacteriophage M13 viral particles is the most popular phage-display system. Other bacteriophage systems, such as λ (*143*) and T7 (*144*) have also been used for display purposes. The M13 bacteriophage consists of a single stranded DNA molecule that is packaged by five coat proteins of the phage (pIII, pVI, pVII, pVIII, and pIX) (*63*). After assembly in the bacterial periplasm, the infectious phage particles are secreted from the bacterial cells in a non-lytic mode (*145*). Recombinant proteins are most commonly displayed on the phage surface as N-terminal fusions to pIII or pVIII, with the display efficiency largely dependent on the route of transport of the recombinant fusion proteins to the bacterial periplasm. To improve functional phage-display of a variety of proteins, three pathways have been used for transporting proteins to the bacterial periplasm: the general secretory (SEC) pathway (*84, 146*), which transports unfolded proteins to the bacterial periplasm, signal recognition particle (SRP) pathway (*74, 147, 148*), which allows proteins to be transported as they are being translated and the twin-arginine translocation (Tat) pathway (*149, 150*), which transports only folded proteins. Each of these pathways offers unique advantages and increases the plethora of proteins amenable for functional phage-display.

There are 3-5 copies of pIII at one end of the phage, and the intact pIII is essential for infectivity of the bacteriophage particles. The entire surface of the bacteriophage is covered with ~2700 copies of the major coat protein pVIII.

Keeping in mind the large number of pVIII molecules, proteins or peptides fused to them will be displayed in more numbers on the phage, whereas fusion to pIII leads to a maximum display of 5 copies of a protein per phage particle. However, the display number depends on the type of display vector used. Two types of display vectors are commonly used to display proteins on the bacteriophage surface: 1) phagemid vector (plasmids with phage packaging signal, *E. coli* and phage origins of replication, and gene III or VIII coding sequence), and 2) phage vector (modified complete phage genome). The protein to be displayed is cloned in-frame with the coding sequence of the phage coat protein. In a phage vector system (*125, 151*), the phage genome encodes for all the proteins essential for packing the phage particles, therefore, all the copies of the phage coat protein will display the protein of interest. As this system allows multivalent display (Fig. 7A), it facilitates the isolation of even weak binding reagents. In one study, it was shown that more binding clones are isolated in fewer rounds of selection with phage libraries compared to phagemid libraries, due to the avidity effect (i.e., multivalent display) (*151*). However, there is a limitation on the size of the peptides or proteins that can be displayed in a multivalent fashion by phage particles. For instance, to be displayed on all the copies of pVIII, the peptides have to be ≤ 9 residues (*152, 153*). Larger size proteins (such as scFvs, ~25 kDa) have been displayed in a multivalent fashion as pIII fusions, and due to their multivalent display have a greater apparent affinity resulting in isolation of even

weak binding clones (*154*). On the other hand, a phagemid vector (*155, 156*) encodes for only one of the coat protein that serves as a fusion partner to the protein of interest, with all the other phage proteins supplied by a helper phage (M13KO7 or VCS-M13). Competition between the helper phage wild-type coat protein and the recombinant coat protein from the phagemid vector for getting incorporated into the phage particle results in only 1-3% of the phage particles displaying the recombinant protein and the rest will be the wild-type M13 bacteriophage. Due to monovalent display, high affinity binders can be isolated (*23, 157-159*) (Fig. 7B).

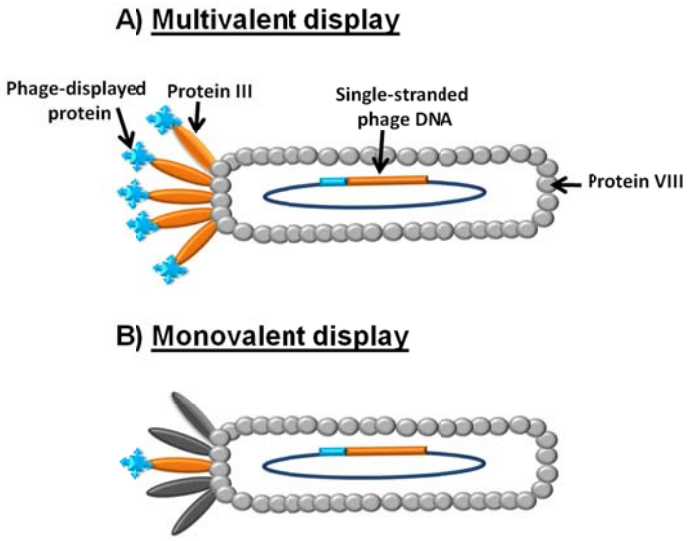


Figure 7. Multivalent vs. monovalent display. A) The protein of interest is displayed on all the five copies of protein III when a phage vector is used. B) One or zero copies of the fusion protein are displayed on the bacteriophage surface, when a phagemid vector is used and the system requires a helper phage to allow phage propagation.

Phage-displayed libraries have been screened to isolate affinity reagents against a wide variety of targets, such as peptides (*160*), protein-protein interaction domains (*23, 161*), cell surface antigens (*162*), membrane receptors (*66*), disease biomarkers (*163, 164*), etc. The process of screening large phage-displayed libraries containing ~10^9-10^{10} clones to isolate variants that specifically bind to a target protein is called 'biopanning' or 'affinity selection' (*165, 166*), which is depicted in Fig. 8. In this process, phage particles displaying binding peptides, proteins, or antibodies for a given target are selectively recovered, amplified and enriched with each subsequent round of affinity selection. In general, two to three rounds of affinity selection are performed followed by a phage enzyme-linked immunosorbent assay (ELISA) to identify phage clones specifically binding to their target. The phage genome encodes for the protein displayed on its surface, which makes it straightforward to obtain the DNA sequence of the binding phage clones with relative ease.

Once affinity reagents are confirmed to bind to their target, the second step is to transfer inserts from the binding clones into expression vectors. For instance, these reagents can be overexpressed in *Escherichia coli, Pichia pastoris*, or Human embryonic kidney 293 (HEK-293). They can be fused to various enzymes, such as cutinase for immobilizing the reagents in a specific orientation (*167, 168*), or alkaline phosphatase as a direct probe to use in ELISA, western blotting, and immunohistochemistry (*157, 169-171*). Moreover, a free

cysteine residue can be engineered in the protein so that it can be chemically linked to solvatochromatic dyes, to be used as biosensors for monitoring the activation of proteins in the cells (*172*).

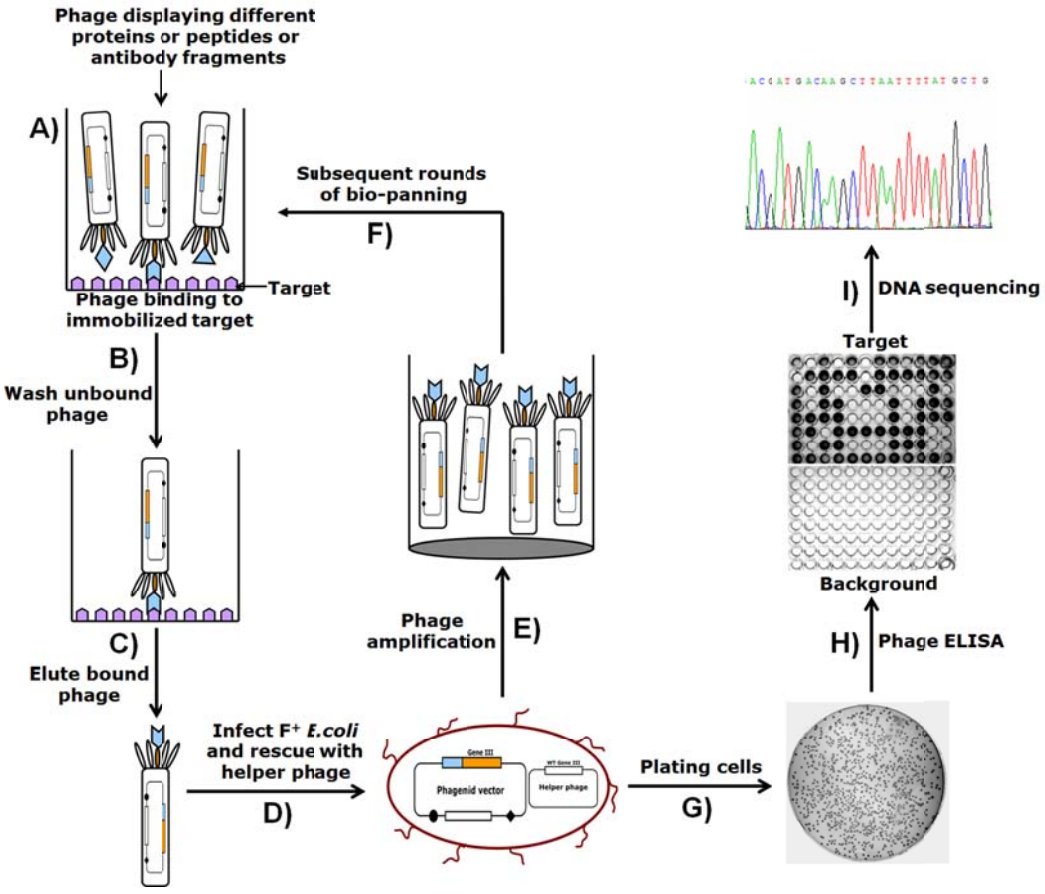


Figure 8. Overview of the affinity selection process. A) The phage library is first incubated with the immobilized target protein. B) The unbound phage are removed by washing in the presence of detergent C) followed by elution of the bound phage under low pH conditions. D) Bacterial cells with F-pilus (F^+) are infected *en masse* with the eluted phage, and E) amplified overnight. F) The secreted phage particles are subsequently used for the second round of biopanning. G) After two or three rounds of affinity selection, the infected cells are diluted and plated. H) Phage are amplified from 96 individual colonies and the phage clones are analyzed by phage ELISA, for binding to the target protein (target) and the negative control (background). I) DNA is isolated from the binding phage and sequenced.

1.9 Advantages of recombinant affinity reagents

Over the years, phage display technology (*63, 173*) has become increasingly popular as an alternative for generating 'affinity reagents' to numerous proteins *in vitro.* Compared to generating antibodies by animal immunizations, which takes up to few months, using recombinant techniques, monoclonal affinity reagents can be generated in 2-3 weeks by screening large libraries (~10^9-10^{10}) of variants against the protein of interest by a process known as 'biopanning' or 'affinity selection'. These reagents can be overexpressed in *E. coli*, *P. pastoris*, or mammalian cells, and offer a renewable source of the affinity reagents. The DNA sequence of the binding clone is easily available, and if needed, can be artificially synthesized with codon optimization for improved expression in prokaryotic or eukaryotic hosts. One of the major strengths of recombinant antibody technology is that the affinity and specificity of the binding clones can be affinity matured *in vitro* by directed evolution to obtain reagents with an affinity comparable to or even greater than antibodies derived after animal immunizations (*174, 175*). The affinity reagents can also be dimerized or multimerized to improve their apparent affinity through avidity (*176, 177*). By using stringent selection conditions, and appropriate selective pressures, reagents with improved thermal stability, expression, affinity (*178, 179*), and aggregation resistance (*180*) have been isolated. Recombinant antibody technology has made it possible to initiate proteome scale projects for isolating

reagents to hundreds or even thousands of proteins (*181-183*), which is practically impossible to do with animal immunizations.

1.10 Conclusion

To this end, we have engineered the FHA1 domain to display functional variants of this domain on the surface of bacteriophage M13, improved its thermal stability, identified residues in the FHA1 domain that are important for recognizing its cognate pT-peptide, generated two phage-displayed libraries of FHA1 variants, screened a total of 11 different pT-containing peptides, and successfully isolated anti-phosphospecific reagents to seven phosphopeptides. Our success rate for isolating phosphospecific reagents from these libraries is 60%. For the first time it has been shown that the specificity of a phosphopeptide-binding domain can be engineered to recognize new phosphopeptides for which binding was not previously detected. We have achieved our first goal of using the FHA1 domain as a scaffold and demonstrated the applicability of these reagents in ELISA. As these reagents have micromolar binding dissociation constants, we will use *in vitro* affinity maturation to improve the affinity of these reagents to broaden their applicability for use as reagents in Western blotting and pull-down experiments. This research opens up the prospect of engineering other pS/T- binding domains to generate a resource of

useful candidates from which reagents with desired properties can be isolated by affinity selection from a phage-displayed library of variants.

1.11 References

1. Ahn, N. G., and Resing, K. A. (2001) Toward the phosphoproteome, *Nat Biotechnol 19*, 317-318.

2. Knight, Z. A., Schilling, B., Row, R. H., Kenski, D. M., Gibson, B. W., and Shokat, K. M. (2003) Phosphospecific proteolysis for mapping sites of protein phosphorylation, *Nat Biotechnol 21*, 1047-1054.

3. Yu, J. W., and Lemmon, M. A. (2003) Genome-wide analysis of signaling domain function, *Curr Opin Chem Biol 7*, 103-109.

4. Yang, X. J. (2005) Multisite protein modification and intramolecular signaling, *Oncogene 24*, 1653-1662.

5. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications, *Nat Biotechnol 21*, 255-261.

6. Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., and Pandey, A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome, *Trends Biotechnol 20*, 261-268.

7. Seet, B. T., Dikic, I., Zhou, M. M., and Pawson, T. (2006) Reading protein modifications with interaction domains, *Nat Rev Mol Cell Biol 7*, 473-483.

8. Yaffe, M. B., and Elia, A. E. (2001) Phosphoserine/threonine-binding domains, *Curr Opin Cell Biol 13*, 131-138.

9. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains, *Science 300*, 445-452.

10. Brown, M. T., and Cooper, J. A. (1996) Regulation, substrates and functions of src, *Biochim Biophys Acta 1287*, 121-149.

11. Ye, X., Nalepa, G., Welcker, M., Kessler, B. M., Spooner, E., Qin, J., Elledge, S. J., Clurman, B. E., and Harper, J. W. (2004) Recognition of phosphodegron motifs in human cyclin E by the SCF(Fbw7) ubiquitin ligase, *J Biol Chem 279*, 50110-50119.

12. Strohmaier, H., Spruck, C. H., Kaiser, P., Won, K. A., Sangfelt, O., and Reed, S. I. (2001) Human F-box protein hCdc4 targets cyclin E for

proteolysis and is mutated in a breast cancer cell line, *Nature 413*, 316-322.

13. Orlicky, S., Tang, X., Willems, A., Tyers, M., and Sicheri, F. (2003) Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase, *Cell 112*, 243-256.

14. Nash, P., Tang, X., Orlicky, S., Chen, Q., Gertler, F. B., Mendenhall, M. D., Sicheri, F., Pawson, T., and Tyers, M. (2001) Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication, *Nature 414*, 514-521.

15. Hirota, T., Lipp, J. J., Toh, B. H., and Peters, J. M. (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin, *Nature 438*, 1176-1180.

16. Chung, H. J., Huang, Y. H., Lau, L. F., and Huganir, R. L. (2004) Regulation of the NMDA receptor complex and trafficking by activity-dependent phosphorylation of the NR2B subunit PDZ ligand, *J Neurosci 24*, 10248-10259.

17. Yaffe, M. B., and Smerdon, S. J. (2001) PhosphoSerine/threonine binding domains: you can't pSERious?, *Structure 9*, R33-38.

18. Cohen, P. (2001) The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture, *Eur J Biochem 268*, 5001-5010.

19. Paradela, A., and Albar, J. P. (2008) Advances in the analysis of protein phosphorylation, *J Proteome Res 7*, 1809-1818.

20. Rosenqvist, H., Ye, J., and Jensen, O. N. (2011) Analytical strategies in mass spectrometry-based phosphoproteomics, *Methods Mol Biol 753*, 183-213.

21. Kalume, D. E., Molina, H., and Pandey, A. (2003) Tackling the phosphoproteome: tools and strategies, *Curr Opin Chem Biol 7*, 64-69.

22. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms, *Mol Cell 6*, 1169-1182.

23. Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display, *Protein Eng Des Sel 23*, 279-288.

24. Pawson, T., and Warner, N. (2007) Oncogenic re-wiring of cellular signaling pathways, *Oncogene 26*, 1268-1275.

25. Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) The structure and function of proline recognition domains, *Sci STKE 2003*, RE8.

26. Berry, D. M., Nash, P., Liu, S. K., Pawson, T., and McGlade, C. J. (2002) A high-affinity Arg-X-X-Lys SH3 binding motif confers specificity for the interaction between Gads and SLP-76 in T cell signaling, *Curr Biol 12*, 1336-1341.

27. Stamenova, S. D., French, M. E., He, Y., Francis, S. A., Kramer, Z. B., and Hicke, L. (2007) Ubiquitin binds to and regulates a subset of SH3 domains, *Mol Cell 25*, 273-284.

28. Sudol, M., and Hunter, T. (2000) NeW wrinkles for an old domain, *Cell 103*, 1001-1004.

29. Zarrinpar, A., and Lim, W. A. (2000) Converging on proline: the mechanism of WW domain peptide recognition, *Nat Struct Biol 7*, 611-613.

30. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules, *Science 283*, 1325-1328.

31. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000) Structural basis for phosphoserine-proline recognition by group IV WW domains, *Nat Struct Biol 7*, 639-643.

32. Otte, L., Wiedemann, U., Schlegel, B., Pires, J. R., Beyermann, M., Schmieder, P., Krause, G., Volkmer-Engert, R., Schneider-Mergener, J., and Oschkinat, H. (2003) WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains, *Protein Sci 12*, 491-500.

33. Doyle, D. A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996) Crystal structures of a complexed and peptide-free membrane

protein-binding domain: molecular basis of peptide recognition by PDZ, *Cell 85*, 1067-1076.

34. Kornau, H. C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95, *Science 269*, 1737-1740.

35. Kessels, H. W., Ward, A. C., and Schumacher, T. N. (2002) Specificity and affinity motifs for Grb2 SH2-ligand interactions, *Proc Natl Acad Sci U S A 99*, 8524-8529.

36. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., and et al. (1993) SH2 domains recognize specific phosphopeptide sequences, *Cell 72*, 767-778.

37. Huang, H., Li, L., Wu, C., Schibli, D., Colwill, K., Ma, S., Li, C., Roy, P., Ho, K., Songyang, Z., Pawson, T., Gao, Y., and Li, S. S. (2008) Defining the specificity space of the human SRC homology 2 domain, *Mol Cell Proteomics 7*, 768-784.

38. Zhou, M. M., Ravichandran, K. S., Olejniczak, E. F., Petros, A. M., Meadows, R. P., Sattler, M., Harlan, J. E., Wade, W. S., Burakoff, S. J., and Fesik, S. W. (1995) Structure and ligand recognition of the phosphotyrosine binding domain of Shc, *Nature 378*, 584-592.

39. van der Geer, P., Wiley, S., Gish, G. D., Lai, V. K., Stephens, R., White, M. F., Kaplan, D., and Pawson, T. (1996) Identification of residues that control specific binding of the Shc phosphotyrosine-binding domain to phosphotyrosine sites, *Proc Natl Acad Sci U S A 93*, 963-968.

40. Trub, T., Choi, W. E., Wolf, G., Ottinger, E., Chen, Y., Weiss, M., and Shoelson, S. E. (1995) Specificity of the PTB domain of Shc for beta turn-forming pentapeptide motifs amino-terminal to phosphotyrosine, *J Biol Chem 270*, 18205-18208.

41. Zhang, Z., Lee, C. H., Mandiyan, V., Borg, J. P., Margolis, B., Schlessinger, J., and Kuriyan, J. (1997) Sequence-specific recognition of the internalization motif of the Alzheimer's amyloid precursor protein by the X11 PTB domain, *EMBO J 16*, 6141-6150.

42. Dhalluin, C., Yan, K. S., Plotnikova, O., Lee, K. W., Zeng, L., Kuti, M., Mujtaba, S., Goldfarb, M. P., and Zhou, M. M. (2000) Structural basis of SNT PTB domain interactions with distinct neurotrophic receptors, *Mol Cell 6*, 921-929.

43. Rittinger, K., Budman, J., Xu, J., Volinia, S., Cantley, L. C., Smerdon, S. J., Gamblin, S. J., and Yaffe, M. B. (1999) Structural analysis of 14-3-3 phosphopeptide complexes identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding, *Mol Cell 4*, 153-166.

44. Muslin, A. J., Tanner, J. W., Allen, P. M., and Shaw, A. S. (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine, *Cell 84*, 889-897.

45. Petosa, C., Masters, S. C., Bankston, L. A., Pohl, J., Wang, B., Fu, H., and Liddington, R. C. (1998) 14-3-3zeta binds a phosphorylated Raf peptide and an unphosphorylated peptide via its conserved amphipathic groove, *J Biol Chem 273*, 16305-16310.

46. Yaffe, M. B. (2002) How do 14-3-3 proteins work?-- Gatekeeper phosphorylation and the molecular anvil hypothesis, *FEBS Lett 513*, 53-57.

47. Coblitz, B., Shikano, S., Wu, M., Gabelli, S. B., Cockrell, L. M., Spieker, M., Hanyu, Y., Fu, H., Amzel, L. M., and Li, M. (2005) C-terminal recognition by 14-3-3 proteins for surface expression of membrane receptors, *J Biol Chem 280*, 36263-36272.

48. Manke, I. A., Lowery, D. M., Nguyen, A., and Yaffe, M. B. (2003) BRCT repeats as phosphopeptide-binding modules involved in protein targeting, *Science 302*, 636-639.

49. Yu, X., Chini, C. C., He, M., Mer, G., and Chen, J. (2003) The BRCT domain is a phospho-protein binding domain, *Science 302*, 639-642.

50. Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I. J., Pei, D., and Tsai, M. D. (2000) Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9, *J Mol Biol 304*, 941-951.

51. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and

functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2, *Mol Cell 9*, 1045-1054.

52. Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M. D. (2000) II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides, *J Mol Biol 302*, 927-940.

53. Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E. S., Wu, M. R., Su, M. I., Heierhorst, J., and Tsai, M. D. (2008) Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade, *Mol Cell 30*, 767-778.

54. Winston, J. T., Strack, P., Beer-Romero, P., Chu, C. Y., Elledge, S. J., and Harper, J. W. (1999) The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro, *Genes Dev 13*, 270-283.

55. Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p, *EMBO J 19*, 6141-6149.

56. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain, *Nature 410*, 120-124.

57. Lemmon, M. A., Ferguson, K. M., O'Brien, R., Sigler, P. B., and Schlessinger, J. (1995) Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain, *Proc Natl Acad Sci U S A 92*, 10472-10476.

58. Ellson, C. D., Andrews, S., Stephens, L. R., and Hawkins, P. T. (2002) The PX domain: a new phosphoinositide-binding module, *J Cell Sci 115*, 1099-1105.

59. Itoh, T., Koshiba, S., Kigawa, T., Kikuchi, A., Yokoyama, S., and Takenawa, T. (2001) Role of the ENTH domain in phosphatidylinositol-4,5-bisphosphate binding and endocytosis, *Science 291*, 1047-1051.

60. Liu, B. A., Engelmann, B. W., and Nash, P. D. (2012) High-throughput analysis of peptide-binding modules, *Proteomics 12*, 1527-1546.

61. Mohammad, D. H., and Yaffe, M. B. (2009) 14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response, *DNA Repair (Amst) 8*, 1009-1017.

62. Wilson, D. S., Keefe, A. D., and Szostak, J. W. (2001) The use of mRNA display to select high-affinity protein-binding peptides, *Proc Natl Acad Sci U S A 98*, 3750-3755.

63. Kehoe, J. W., and Kay, B. K. (2005) Filamentous phage display in the new millennium, *Chemical reviews 105*, 4056-4072.

64. van Meerten, D., Olsthoorn, R. C., van Duin, J., and Verhaert, R. M. (2001) Peptide display on live MS2 phage: restrictions at the RNA genome level, *J Gen Virol 82*, 1797-1805.

65. Sadowski, I., Stone, J. C., and Pawson, T. (1986) A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps, *Mol Cell Biol 6*, 4396-4408.

66. Martinez-Torrecuadrada, J., Cifuentes, G., Lopez-Serra, P., Saenz, P., Martinez, A., and Casal, J. I. (2005) Targeting the extracellular domain of fibroblast growth factor receptor 3 with human single-chain Fv antibodies inhibits bladder carcinoma cell line proliferation, *Clin Cancer Res 11*, 6280-6290.

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

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

69. Moran, M. F., Koch, C. A., Anderson, D., Ellis, C., England, L., Martin, G. S., and Pawson, T. (1990) Src homology region 2 domains direct protein-protein interactions in signal transduction, *Proceedings of the National Academy of Sciences of the United States of America 87*, 8622-8626.

70. Pawson, T., Gish, G. D., and Nash, P. (2001) SH2 domains, interaction modules and cellular wiring, *Trends in cell biology 11*, 504-511.

71. Yaffe, M. B., Rittinger, K., Volinia, S., Caron, P. R., Aitken, A., Leffers, H., Gamblin, S. J., Smerdon, S. J., and Cantley, L. C. (1997) The structural basis for 14-3-3:phosphopeptide binding specificity, *Cell 91*, 961-971.

72. Bridges, D., and Moorhead, G. B. (2005) 14-3-3 proteins: a number of functions for a numbered protein, *Sci STKE 2005*, re10.

73. Moore, B. E., and Perez, V. J. (1967) Physiological and Biochemical Aspects of Nervous Integration (ed. Carlson, F. D.)

(Prentice–Hall, Englewood Cliffs), 343–359.

74. Schierle, C. F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. (2003) The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway, *J Bacteriol 185*, 5706-5713.

75. Tzivion, G., Luo, Z., and Avruch, J. (1998) A dimeric 14-3-3 protein is an essential cofactor for Raf kinase activity, *Nature 394*, 88-92.

76. Mils, V., Baldin, V., Goubin, F., Pinta, I., Papin, C., Waye, M., Eychene, A., and Ducommun, B. (2000) Specific interaction between 14-3-3 isoforms and the human CDC25B phosphatase, *Oncogene 19*, 1257-1265.

77. Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide, *Nature 382*, 646-649.

78. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism, *Science 278*, 1957-1960.

79. Craig, K. L., and Tyers, M. (1999) The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction, *Prog Biophys Mol Biol 72*, 299-328.

80. Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., and Harper, J. W. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex, *Cell 91*, 209-219.

81. Willems, A. R., Goh, T., Taylor, L., Chernushevich, I., Shevchenko, A., and Tyers, M. (1999) SCF ubiquitin protein ligases and phosphorylation-dependent proteolysis, *Philos Trans R Soc Lond B Biol Sci 354*, 1533-1550.

82. Fulop, V., and Jones, D. T. (1999) Beta propellers: structural rigidity and functional diversity, *Curr Opin Struct Biol 9*, 715-721.

83. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., and Koonin, E. V. (1997) A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins, *FASEB J 11*, 68-76.

84. den Blaauwen, T., and Driessen, A. J. (1996) Sec-dependent preprotein translocation in bacteria, *Arch Microbiol 165*, 1-8.

85. Shiozaki, E. N., Gu, L., Yan, N., and Shi, Y. (2004) Structure of the BRCT repeats of BRCA1 bound to a BACH1 phosphopeptide: implications for signaling, *Mol Cell 14*, 405-412.

86. Clapperton, J. A., Manke, I. A., Lowery, D. M., Ho, T., Haire, L. F., Yaffe, M. B., and Smerdon, S. J. (2004) Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer, *Nat Struct Mol Biol 11*, 512-518.

87. Castilla, L. H., Couch, F. J., Erdos, M. R., Hoskins, K. F., Calzone, K., Garber, J. E., Boyd, J., Lubin, M. B., Deshano, M. L., Brody, L. C., and et al. (1994) Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer, *Nat Genet 8*, 387-391.

88. Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain, *Cell 115*, 83-95.

89. Cheng, K. Y., Lowe, E. D., Sinclair, J., Nigg, E. A., and Johnson, L. N. (2003) The crystal structure of the human polo-like kinase-1 polo box domain and its phospho-peptide complex, *Embo J 22*, 5757-5768.

90. Barr, F. A., Sillje, H. H., and Nigg, E. A. (2004) Polo-like kinases and the orchestration of cell division, *Nat Rev Mol Cell Biol 5*, 429-440.

91. Strebhardt, K., and Ullrich, A. (2006) Targeting polo-like kinase 1 for cancer therapy, *Nat Rev Cancer 6*, 321-330.

92. Hofmann, K., and Bucher, P. (1995) The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors, *Trends Biochem Sci 20*, 347-349.

93. Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) The FHA domain mediates phosphoprotein interactions, *J Cell Sci 113 Pt 23*, 4143-4149.

94. Durocher, D., and Jackson, S. P. (2002) The FHA domain, *FEBS Lett 513*, 58-66.

95. Wilker, E. W., Grant, R. A., Artim, S. C., and Yaffe, M. B. (2005) A structural basis for 14-3-3sigma functional specificity, *J Biol Chem 280*, 18891-18898.

96. Wintjens, R., Wieruszeski, J. M., Drobecq, H., Rousselot-Pailley, P., Buee, L., Lippens, G., and Landrieu, I. (2001) 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides, *J Biol Chem 276*, 25150-25156.

97. Zhou, B. B., and Elledge, S. J. (2000) The DNA damage response: putting checkpoints in perspective, *Nature 408*, 433-439.

98. Vialard, J. E., Gilbert, C. S., Green, C. M., and Lowndes, N. F. (1998) The budding yeast Rad9 checkpoint protein is subjected to Mec1/Tel1-dependent hyperphosphorylation and interacts with Rad53 after DNA damage, *EMBO J 17*, 5679-5688.

99. Sun, Z., Hsiao, J., Fay, D. S., and Stern, D. F. (1998) Rad53 FHA domain associated with phosphorylated Rad9 in the DNA damage checkpoint, *Science 281*, 272-274.

100. Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999) The FHA domain is a modular phosphopeptide recognition motif, *Mol Cell 4*, 387-394.

101. Liao, H., Byeon, I. J., and Tsai, M. D. (1999) Structure and function of a new phosphopeptide-binding domain containing the FHA2 of Rad53, *J Mol Biol 294*, 1041-1049.

102. Byeon, I. J., Yongkiettrakul, S., and Tsai, M. D. (2001) Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes, *J Mol Biol 314*, 577-588.

103. Yuan, C., Yongkiettrakul, S., Byeon, I. J., Zhou, S., and Tsai, M. D. (2001) Solution structures of two FHA1-phosphothreonine peptide complexes provide insight into the structural basis of the ligand specificity of FHA1 from yeast Rad53, *J Mol Biol 314*, 563-575.

104. Zhou, M. M. (2000) Phosphothreonine recognition comes into focus, *Nat Struct Biol 7*, 1085-1087.

105. Yongkiettrakul, S., Byeon, I. J., and Tsai, M. D. (2004) The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues, *Biochemistry 43*, 3862-3869.

106. Li, H., Byeon, I. J., Ju, Y., and Tsai, M. D. (2004) Structure of human Ki67 FHA domain and its binding to a phosphoprotein fragment from hNIFK reveal unique recognition sites and new views to the structural basis of FHA domain functions, *J Mol Biol 335*, 371-381.

107. Huen, M. S., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M. B., and Chen, J. (2007) RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly, *Cell 131*, 901-914.

108. Sun, T., Campbell, M., Gordon, W., and Arlinghaus, R. B. (2001) Preparation and application of antibodies to phosphoamino acid sequences, *Biopolymers 60*, 61-75.

109. Karatan, E., Merguerian, M., Han, Z., Scholle, M. D., Koide, S., and Kay, B. K. (2004) Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain, *Chem Biol 11*, 835-844.

110. Fellouse, F. A., Li, B., Compaan, D. M., Peden, A. A., Hymowitz, S. G., and Sidhu, S. S. (2005) Molecular recognition by a binary code, *J Mol Biol 348*, 1153-1162.

111. Zhang, C. (2012) Hybridoma technology for the generation of monoclonal antibodies, *Methods Mol Biol 901*, 117-135.

112. Kohler, G., and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity, *Nature 256*, 495-497.

113. Sovova, V., Cerna, H., Dostalova, V., and Hlozanek, I. (1980) Use of the hybridoma technique for antibody production to group-specific antigens of avian oncornaviruses, *Folia Biol (Praha) 26*, 366-368.

114. Smith, G. P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface, *Science 228*, 1315-1317.

115. Sheets, M. D., Amersdorfer, P., Finnern, R., Sargent, P., Lindquist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J. C., and Marks, J. D. (1998) Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens, *Proc Natl Acad Sci U S A 95*, 6157-6162.

116. Bliss, J. M., Sullivan, M. A., Malone, J., and Haidaris, C. G. (2003) Differentiation of Candida albicans and Candida dubliniensis by using recombinant human antibody single-chain variable fragments specific for hyphae, *Journal of clinical microbiology 41*, 1152-1160.

117. Even-Desrumeaux, K., Fourquet, P., Secq, V., Baty, D., and Chames, P. (2012) Single-domain antibodies: a versatile and rich source of binders for breast cancer diagnostic approaches, *Mol Biosyst 8*, 2385-2394.

118. Rothe, C., Urlinger, S., Lohning, C., Prassler, J., Stark, Y., Jager, U., Hubner, B., Bardroff, M., Pradel, I., Boss, M., Bittlingmaier, R., Bataa, T., Frisch, C., Brocks, B., Honegger, A., and Urban, M. (2008) The human combinatorial antibody library HuCAL GOLD combines diversification of all six CDRs according to the natural immune system with a novel display method for efficient selection of high-affinity antibodies, *J Mol Biol 376*, 1182-1200.

119. Lu, D., Jimenez, X., Zhang, H., Bohlen, P., Witte, L., and Zhu, Z. (2002) Selection of high affinity human neutralizing antibodies to VEGFR2 from a large antibody phage display library for antiangiogenesis therapy, *Int J Cancer 97*, 393-399.

120. Rauchenberger, R., Borges, E., Thomassen-Wolf, E., Rom, E., Adar, R., Yaniv, Y., Malka, M., Chumakov, I., Kotzer, S., Resnitzky, D., Knappik, A., Reiffert, S., Prassler, J., Jury, K., Waldherr, D., Bauer, S., Kretzschmar, T., Yayon, A., and Rothe, C. (2003) Human combinatorial Fab library yielding specific and functional antibodies against the human fibroblast growth factor receptor 3, *J Biol Chem 278*, 38194-38205.

121. Lee, C. V., Sidhu, S. S., and Fuh, G. (2004) Bivalent antibody phage display mimics natural immunoglobulin, *J Immunol Methods 284*, 119-132.

122. Tso, J. Y., Wang, S. L., Levin, W., and Hakimi, J. (1995) Preparation of a bispecific F(ab')2 targeted to the human IL-2 receptor, *J Hematother 4*, 389-394.

123. Koide, A., Bailey, C. W., Huang, X., and Koide, S. (1998) The fibronectin type III domain as a scaffold for novel binding proteins, *J Mol Biol 284*, 1141-1151.

124. Koide, S., Koide, A., and Lipovsek, D. (2012) Target-binding proteins based on the 10th human fibronectin type III domain ((1)(0)Fn3), *Methods in enzymology 503*, 135-156.

125. Scholle, M. D., Kehoe, J. W., and Kay, B. K. (2005) Efficient construction of a large collection of phage-displayed combinatorial peptide libraries, *Comb Chem High Throughput Screen 8*, 545-551.

126. Scholle, M. D., Banach, B. S., Hamdan, S. M., Richardson, C. C., and Kay, B. K. (2008) Peptide ligands specific to the oxidized form of Escherichia coli thioredoxin, *Biochim Biophys Acta 1784*, 1735-1741.

127. Kriplani, U., and Kay, B. K. (2005) Selecting peptides for use in nanoscale materials using phage-displayed combinatorial peptide libraries, *Curr Opin Biotechnol 16*, 470-475.

128. Ernst, A., Sazinsky, S. L., Hui, S., Currell, B., Dharsee, M., Seshagiri, S., Bader, G. D., and Sidhu, S. S. (2009) Rapid evolution of functional complexity in a domain family, *Science signaling 2*, ra50.

129. Ernst, A., Gfeller, D., Kan, Z., Seshagiri, S., Kim, P. M., Bader, G. D., and Sidhu, S. S. (2010) Coevolution of PDZ domain-ligand interactions analyzed by high-throughput phage display and deep sequencing, *Mol Biosyst 6*, 1782-1790.

130. Huber, T., Steiner, D., Rothlisberger, D., and Pluckthun, A. (2007) In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: The Na(+)-citrate symporter CitS as an example, *Journal of structural biology 159*, 206-221.

131. Stefan, N., Martin-Killias, P., Wyss-Stoeckle, S., Honegger, A., Zangemeister-Wittke, U., and Pluckthun, A. (2011) DARPins recognizing the tumor-associated antigen EpCAM selected by phage and ribosome display and engineered for multivalency, *J Mol Biol 413*, 826-843.

132. Steiner, D., Forrer, P., and Pluckthun, A. (2008) Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display, *J Mol Biol 382*, 1211-1227.

133. Lindborg, M., Cortez, E., Hoiden-Guthenberg, I., Gunneriusson, E., von Hage, E., Syud, F., Morrison, M., Abrahmsen, L., Herne, N., Pietras, K., and Frejd, F. Y. (2011) Engineered high-affinity affibody molecules targeting platelet-derived growth factor receptor beta in vivo, *J Mol Biol 407*, 298-315.

134. Friedman, M., Nordberg, E., Hoiden-Guthenberg, I., Brismar, H., Adams, G. P., Nilsson, F. Y., Carlsson, J., and Stahl, S. (2007) Phage display selection of Affibody molecules with specific binding to the extracellular domain of the epidermal growth factor receptor, *Protein Eng Des Sel 20*, 189-199.

135. Harris, L. J., Skaletsky, E., and McPherson, A. (1998) Crystallographic structure of an intact IgG1 monoclonal antibody, *J Mol Biol 275*, 861-872.

136. Carmichael, J. A., Power, B. E., Garrett, T. P., Yazaki, P. J., Shively, J. E., Raubischek, A. A., Wu, A. M., and Hudson, P. J. (2003) The crystal structure of an anti-CEA scFv diabody assembled from T84.66 scFvs in V(L)-to-V(H) orientation: implications for diabody flexibility, *J Mol Biol 326*, 341-351.

137. Cho, H. S., Mason, K., Ramyar, K. X., Stanley, A. M., Gabelli, S. B., Denney, D. W., Jr., and Leahy, D. J. (2003) Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab, *Nature 421*, 756-760.

138. Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. (1996) Crystal structure of a

camel single-domain VH antibody fragment in complex with lysozyme, *Nat Struct Biol 3*, 803-811.

139. Elkins, J. M., Papagrigoriou, E., Berridge, G., Yang, X., Phillips, C., Gileadi, C., Savitsky, P., and Doyle, D. A. (2007) Structure of PICK1 and other PDZ domains obtained with the help of self-binding C-terminal extensions, *Protein Sci 16*, 683-694.

140. Kohl, A., Binz, H. K., Forrer, P., Stumpp, M. T., Pluckthun, A., and Grutter, M. G. (2003) Designed to be stable: crystal structure of a consensus ankyrin repeat protein, *Proc Natl Acad Sci U S A 100*, 1700-1705.

141. Hogbom, M., Eklund, M., Nygren, P. A., and Nordlund, P. (2003) Structural basis for recognition by an in vitro evolved affibody, *Proc Natl Acad Sci U S A 100*, 3191-3196.

142. Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K. L., Barkinge, J., Jones, R. B., Koide, A., Superti-Furga, G., and Koide, S. (2010) A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain, *Nat Struct Mol Biol 17*, 519-527.

143. Gupta, A., Onda, M., Pastan, I., Adhya, S., and Chaudhary, V. K. (2003) High-density functional display of proteins on bacteriophage lambda, *J Mol Biol 334*, 241-254.

144. Danner, S., and Belasco, J. G. (2001) T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries, *Proc Natl Acad Sci U S A 98*, 12954-12959.

145. Feng, J. N., Russel, M., and Model, P. (1997) A permeabilized cell system that assembles filamentous bacteriophage, *Proc Natl Acad Sci U S A 94*, 4068-4073.

146. Fekkes, P., and Driessen, A. J. (1999) Protein targeting to the bacterial cytoplasmic membrane, *Microbiol Mol Biol Rev 63*, 161-173.

147. Steiner, D., Forrer, P., Stumpp, M. T., and Pluckthun, A. (2006) Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display, *Nat Biotechnol 24*, 823-831.

148. Valent, Q. A. (2001) Signal recognition particle mediated protein targeting in Escherichia coli, *Antonie Van Leeuwenhoek 79*, 17-31.

149. Paschke, M., and Hohne, W. (2005) A twin-arginine translocation (Tat)-mediated phage display system, *Gene 350*, 79-88.

150. Speck, J., Arndt, K. M., and Muller, K. M. (2011) Efficient phage display of intracellularly folded proteins mediated by the TAT pathway, *Protein Eng Des Sel 24*, 473-484.

151. O'Connell, D., Becerril, B., Roy-Burman, A., Daws, M., and Marks, J. D. (2002) Phage versus phagemid libraries for generation of human monoclonal antibodies, *J Mol Biol 321*, 49-56.

152. Iannolo, G., Minenkova, O., Petruzzelli, R., and Cesareni, G. (1995) Modifying filamentous phage capsid: limits in the size of the major capsid protein, *J Mol Biol 248*, 835-844.

153. Malik, P., Terry, T. D., Gowda, L. R., Langara, A., Petukhov, S. A., Symmons, M. F., Welsh, L. C., Marvin, D. A., and Perham, R. N. (1996) Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage, *J Mol Biol 260*, 9-21.

154. MacKenzie, R., and To, R. (1998) The role of valency in the selection of anti-carbohydrate single-chain Fvs from phage display libraries, *J Immunol Methods 220*, 39-49.

155. Pershad, K., Sullivan, M. A., and Kay, B. K. (2011) Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats, *Anal Biochem 412*, 210-216.

156. Zhao, Q., Chan, Y. W., Lee, S. S., and Cheung, W. T. (2009) One-step expression and purification of single-chain variable antibody fragment using an improved hexahistidine tag phagemid vector, *Protein Expr Purif 68*, 190-195.

157. Gonzalez-Techera, A., Umpierrez-Failache, M., Cardozo, S., Obal, G., Pritsch, O., Last, J. A., Gee, S. J., Hammock, B. D., and Gonzalez-Sapienza, G. (2008) High-throughput method for ranking the affinity of peptide ligands selected from phage display libraries, *Bioconjugate chemistry 19*, 993-1000.

158. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J., and Johnson, K.

S. (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library, *Nat Biotechnol 14*, 309-314.

159. Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S., Clarke, K. F., Conquer, J. S., Crofts, A. M., Crowther, S. R., Dyson, M. R., Flack, G., Griffin, G. J., Hooks, Y., Howat, W. J., Kolb-Kokocinski, A., Kunze, S., Martin, C. D., Maslen, G. L., Mitchell, J. N., O'Sullivan, M., Perera, R. L., Roake, W., Shadbolt, S. P., Vincent, K. J., Warford, A., Wilson, W. E., Xie, J., Young, J. L., and McCafferty, J. (2007) Application of phage display to high throughput antibody generation and characterization, *Genome biology 8*, R254.

160. Memic, A., Volgina, V. V., Gussin, H. A., Pepperberg, D. R., and Kay, B. K. (2011) Generation of recombinant guinea pig antibody fragments to the human GABAC receptor, *J Immunol Methods 368*, 36-44.

161. Huang, R., Fang, P., and Kay, B. K. (2012) Isolation of monobodies that bind specifically to the SH3 domain of the Fyn tyrosine protein kinase, *N Biotechnol 29*, 526-533.

162. Zhang, J. L., Gou, J. J., Zhang, Z. Y., Jing, Y. X., Zhang, L., Guo, R., Yan, P., Cheng, N. L., Niu, B., and Xie, J. (2006) Screening and evaluation of human single-chain fragment variable antibody against hepatitis B virus surface antigen, *Hepatobiliary Pancreat Dis Int 5*, 237-241.

163. Delcommenne, M., and Klingemann, H. G. (2012) Detection and characterization of syndecan-1-associated heparan sulfate 6-O-sulfated motifs overexpressed in multiple myeloma cells using single chain antibody variable fragments, *Hum Antibodies 21*, 29-40.

164. Abdolalizadeh, J., Nouri, M., Zolbanin, J. M., Baradaran, B., Barzegari, A., and Omidi, Y. (2012) Downstream characterization of anti-TNF-alpha single chain variable fragment antibodies, *Hum Antibodies 21*, 41-48.

165. Kay, B. K., Kasanov, J., and Yamabhai, M. (2001) Screening phage-displayed combinatorial peptide libraries, *Methods (San Diego, Calif 24*, 240-246.

166. Turunen, L., Takkinen, K., Soderlund, H., and Pulli, T. (2009) Automated panning and screening procedure on microplates for antibody generation from phage display libraries, *J Biomol Screen 14*, 282-293.

167. Hodneland, C. D., Lee, Y. S., Min, D. H., and Mrksich, M. (2002) Selective immobilization of proteins to self-assembled monolayers presenting active site-directed capture ligands, *Proc Natl Acad Sci U S A 99*, 5048-5052.

168. Kwon, Y., Han, Z., Karatan, E., Mrksich, M., and Kay, B. K. (2004) Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers, *Analytical chemistry 76*, 5713-5720.

169. Martin, C. D., Rojas, G., Mitchell, J. N., Vincent, K. J., Wu, J., McCafferty, J., and Schofield, D. J. (2006) A simple vector system to improve performance and utilisation of recombinant antibodies, *BMC Biotechnol 6*, 46.

170. Muller, B. H., Chevrier, D., Boulain, J. C., and Guesdon, J. L. (1999) Recombinant single-chain Fv antibody fragment-alkaline phosphatase conjugate for one-step immunodetection in molecular hybridization, *J Immunol Methods 227*, 177-185.

171. Wright, R. M., Dudas, D., Gavin, B., Dottavio, D., Hexham, J. M., and Lake, P. (2001) A high-capacity alkaline phosphatase reporter system for the rapid analysis of specificity and relative affinity of peptides from phage-display libraries, *J Immunol Methods 253*, 223-232.

172. Gulyani, A., Vitriol, E., Allen, R., Wu, J., Gremyachinskiy, D., Lewis, S., Dewar, B., Graves, L. M., Kay, B. K., Kuhlman, B., Elston, T., and Hahn, K. M. (2012) A biosensor generated via high-throughput screening quantifies cell edge Src dynamics, *Nat Chem Biol 8*, 737.

173. Schmitz, U., Versmold, A., Kaufmann, P., and Frank, H. G. (2000) Phage display: a molecular tool for the generation of antibodies--a review, *Placenta 21 Suppl A*, S106-112.

174. Dyson, M. R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B. K., Pawson, T., and McCafferty, J. (2011) Mapping protein interactions by combining antibody affinity maturation and mass spectrometry, *Anal Biochem 417*, 25-35.

175. Razai, A., Garcia-Rodriguez, C., Lou, J., Geren, I. N., Forsyth, C. M., Robles, Y., Tsai, R., Smith, T. J., Smith, L. A., Siegel, R. W., Feldhaus, M., and Marks, J. D. (2005) Molecular evolution of antibody affinity for sensitive detection of botulinum neurotoxin type A, *J Mol Biol 351*, 158-169.

176. de Kruif, J., and Logtenberg, T. (1996) Leucine zipper dimerized bivalent and bispecific scFv antibodies from a semi-synthetic antibody phage display library, *J Biol Chem 271*, 7630-7634.

177. Duan, J., Wu, J., Valencia, C. A., and Liu, R. (2007) Fibronectin type III domain based monobody with high avidity, *Biochemistry 46*, 12656-12664.

178. Jung, S., Honegger, A., and Pluckthun, A. (1999) Selection for improved protein stability by phage display, *J Mol Biol 294*, 163-180.

179. Pershad, K., and Kay, B. (Accepted for publication) Generating Thermally Stable Variants of Protein Domains through Phage-display *Methods Journal*.

180. Famm, K., Hansen, L., Christ, D., and Winter, G. (2008) Thermodynamically stable aggregation-resistant antibody domains through directed evolution, *J Mol Biol 376*, 926-931.

181. Colwill, K., and Graslund, S. (2011) A roadmap to generate renewable protein binders to the human proteome, *Nat Methods 8*, 551-558.

182. Taussig, M. J., Stoevesandt, O., Borrebaeck, C. A., Bradbury, A. R., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T. J., Gloriam, D., Gold, L., Herberg, F. W., Hermjakob, H., Hoheisel, J. D., Joos, T. O., Kallioniemi, O., Koegl, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P. A., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D. J., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nat Methods 4*, 13-17.

183. Blow, N. (2007) Antibodies: The generation game, *Nature 447*, 741-744.

CHAPTER 2

ENGINEERING THE FORKHEAD-ASSOCIATED DOMAIN

FOR FUNCTIONAL PHAGE-DISPLAY

This work has been published

Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

2.1 Abstract

Conventionally antibodies to monitor protein phosphorylation were generated by immunizing animals with synthetic phosphopeptides. With an estimated 100,000 phosphorylation sites in the human proteome, this method is time consuming, expensive and laborious. An ideal alternative would be to produce affinity reagents using recombinant proteins displayed on the surface of bacteriophage M13. To explore such a strategy, we attempted to display on the surface of bacteriophage M13, the N-terminal Forkhead-associated domain (FHA1) of *Saccharomyces cerevisiae* Rad53 protein, which is a naturally occurring phosphothreonine (pT)-binding domain, and found it to be non-functional, presumably due to misfolding in the bacterial periplasm. To overcome this limitation, a library of FHA1 variants was constructed by mutagenic PCR and functional variants were isolated after three rounds of affinity selection with its pT peptide ligand. The FHA1D2 variant, which showed the highest binding signal, had a single residue substitution (S34F) that rescued binding to the cognate pT peptide. Further evaluation revealed that to facilitate proper folding of the FHA1 variant in the bacterial periplasm, where phage particles are assembled in *E. coli*, position 34 has to be occupied with a hydrophobic amino acid (with the exceptions of Trp, Cys and Pro). The FHA1D2 variant bound to the pT peptide with a dissociation constant similar to that of the wild-type FHA1 domain (~1 μM). By fluorescence-based thermal shift assay, it was determined that the thermal stability of the FHA1D2 variant was lower than that of the wild-type FHA1 domain. This functional, phage-displayed FHA1D2 variant can be a potential scaffold for generating anti-phosphospecific affinity reagents.

2.2 Introduction

Phage-display is a powerful technology for isolating affinity reagents that bind selectively to their target proteins with high affinity by performing two or three rounds of biopanning (*1, 2*). Phage-displayed libraries have been constructed from different scaffold proteins, such as single-chain variable fragments (scFv) (*3, 4*) , Fragments of antigen binding (Fab) (*5*), Fibronectin type III domain (FN3) or monobodies (*6*), combinatorial peptides (*7*), PDZ domains (*8*), designed ankyrin repeat proteins (DARPins) (*9*), cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) (*10*), Affibodies (*11*) and anticalins (*12*). The phage library composed of billions of variants is displayed, most commonly, as fusions to the coat protein III or VIII of the bacteriophage particle. Using phage-display technology, different kinds of affinity reagents have been isolated against various target proteins, such as, single-chain variable fragments (scFv) to mammalian proteins (*13*) and 20 human Src homology 2 (SH2) domains (*14*), monobodies to Src homology 3 (SH3) domains (*15*), Fragments of antigen binding to HER2 (*16*), designed anykyrin repeat proteins (DARPins) to epidermal growth factor receptor (*17*), etc. As the phage particles are assembled in the bacterial periplasm, all the proteins that have to be displayed on the surface of bacteriophage M13 are transported to the bacterial periplasm via their N-terminal signal sequence where they get incorporated into the phage as fusions to the coat protein during the assembly process.

Generating good quality and specific antibodies to phosphoproteins is still a challenge. In cell signaling, many proteins get phosphorylated on serine, threonine, and

tyrosine residues by protein kinases (*18*). Protein phosphorylation is a very important post-translational modification, which is responsible for regulating the activity of proteins, translocating them to their proper subcellular location and facilitating the formation of multisubunit protein complexes for transducing signals to downstream effectors.

The most commonly used anti-phosphotyrosine and anti-phosphoserine/threonine antibodies available in the market as of today (for instance 4G10 from Millipore, PY20 from Oncogene Research Products, and 7F12, IC8, IE11 from Alexis Biochemicals) have been generated by immunizing animals, such as mice and rabbits, with synthetic phosphopeptides (*19, 20*). An attractive alternative for generating affinity reagents that recognize phosphoepitopes is to use recombinant proteins and isolate new anti-phosphopeptide binding specificities from a library of variants of a phosphopeptide-binding domain. To explore such a strategy, we chose to engineer a naturally occurring phosphothreonine (pT)-binding domain, namely the Forkhead-associated domain (FHA1), from Saccharomyces cerevisiae Rad53 protein.

Among the pS/T-binding domains, the FHA domains are unique in that they recognize only pT-containing peptides and do not show binding to either non-phosphorylated or pS-containing peptides (*21, 22*). The optimal binding motifs for various FHA domains, from *S. cerevisiae*, *Schizosaccharomyces pombe*, *Arabidopsis thaliana* and *Mycobacterium tuberculosis* were determined by using oriented phosphopeptide libraries that contain a fixed pT residue, which is flanked by four degenerate residues on either side of it (*22*). From these screens, the pT +3 residue

was found to be one of the major determinants of binding specificity. For example, the N-terminal FHA1 domain from *S. cerevisiae* Rad53 protein kinase prefers Asp at the pT +3 position (*23*), the C-terminal FHA2 domain from the same protein prefers Leu/Iso at the pT +3 position (*24*) and Met/Leu/Phe at the pY +3 position (*25*), and the FHA domain of the human Chk2 DNA damage check point kinase prefers Leu/Iso at the pT +3 position (*26*). From alanine-scanning experiments of the pT peptide, it was determined that only the pT (+3) residue contributed significantly to binding to the FHA1 domain (*21*). Interestingly, two non-conserved residues in the FHA1 domain (G133 and G135) also contribute to the pT (+3) residue specificity (*27*). The specificity of FHA domains ranges from recognizing singly or doubly phosphorylated sequences (*28*) to binding to an extended binding surface (*29*). The tightest FHA domain:pT peptide interaction (K_d), reported to date is 100 nM (*30*).

As the FHA1 domain was found to be non-functional when displayed on the surface of bacteriophage M13, functional variants were isolated by affinity selecting against the cognate pT peptide (Rad9-pT from *S. cerevisiae* Rad9 protein) from a library of FHA1 variants generated by error-prone PCR. A hydrophobic residue at position 34 in the β1-strand was discovered to be essential for phage-display of a functional FHA1 domain. We characterized the binding affinity, specificity, and thermal stability of the functional phage-displayed FHA1D2 variant. The advantages of the FHA domain as a scaffold are that 1) they can be overexpressed in *Escherichia coli*, 2) no animal immunizations are required, 3) it is a more economical and faster way of generating affinity reagents than animal immunization, and 4) the affinity reagents are monoclonal

and renewable. In the future, these anti-phosphospecific reagents can be evaluated for various applications, such as detection reagents for ELISA, western blotting and cell staining, and as biosensors for detecting protein phosphorylation in live cells.

2.3 Materials and methods

2.3.1 Gene synthesis and subcloning into phage and expression vectors

A variant of the FHA1 domain from the *S. cerevisiae* Rad53 protein, named the 3C-3S variant, which has three cysteine residues mutated to serine (C34S, C38S and C154S), was commercially synthesized (Blue Heron Biotechnology, Bothell, WA) with codons optimized for expression in *E. coli* and the DNA was provided after subcloning into the a derivative of the pUC 119 plasmid. The 3C-3S coding sequence was amplified by polymerase chain reaction (PCR), using primers FHA1-NcoI-Fw and FHA1-NotI-Rv and the AccuPrimeTM *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA), for creating flanking *Nco* I/*Not* I sites for subcloning into the phagemid vector (pKP600, Fig. 1) in-frame with the gene III coding sequence. The phagemid vector (*31*) used is a modified version of the pKP300 vector except that it has a DsbA signal sequence and lacks the alkaline phosphatase coding sequence. Following the protocol for Kunkel mutagenesis (*32*), the WT FHA1 domain (containing four cysteine residues at positions 34, 38, 74 and 154) was generated using the 3C-3S coding sequence as the template and two oligonucleotides; the first oligonucleotide (KM-S34C+S38C-FHA1) mutated S34 and S38 to cysteine, and the second oligonucleotide (KM-S154C-FHA1) converted position S154 serine to cysteine. Another FHA1 variant, 4C-4S (all four cysteines mutated to

serines) was generated from the 3C-3S variant using one oligonucleotide, KM-C74S-FHA1, which mutated position 74 cysteine to serine. All the phagemid vectors were sequenced using the primer DsbA-Fw. For generating glutathione-S-transferase (GST) fusions of the FHA1 domains for cytoplasmic expression, their coding sequence was amplified by PCR creating *Bam* HI/*Eco* RI flanking sites using the primers-FHA1-BamH1-Fw and FHA1-EcoRI-Rv and AccuPrimeTM *Pfx* DNA polymerase. The pGEX-2T GST fusion vector (GE Healthcare, Piscataway, NJ) was cleaved with the same two restriction endonucleases and the FHA1 domains were subcloned in-frame with the GST coding sequence. The final construct was sequenced using the primer Seq-pGEX-Fw. (All restriction enzymes were purchased from New England BioLabs, Ipswich, MA).

For expression on a large scale, the FHA domains were subcloned into a modified version of the pET29b expression vector (gift from Dr. Brian Kuhlman, University of North Carolina) in-frame with a C-terminal six-histidine tag for protein purification, by immobilized metal affinity chromatography (IMAC). The FHA1 domain coding sequences were amplified by PCR using the AccuPrimeTM *Pfx* DNA polymerase and FHA1-NdeI-Fw and FHA1-XhoI-Rv primers, which created flanking *Nde* I/*Xho* I recognition sites for subcloning into the pET29b expression vector (for cytoplasmic expression). All constructs were DNA sequenced. All the primers were ordered from Integrated DNA Technologies and their sequences are listed in Table I.

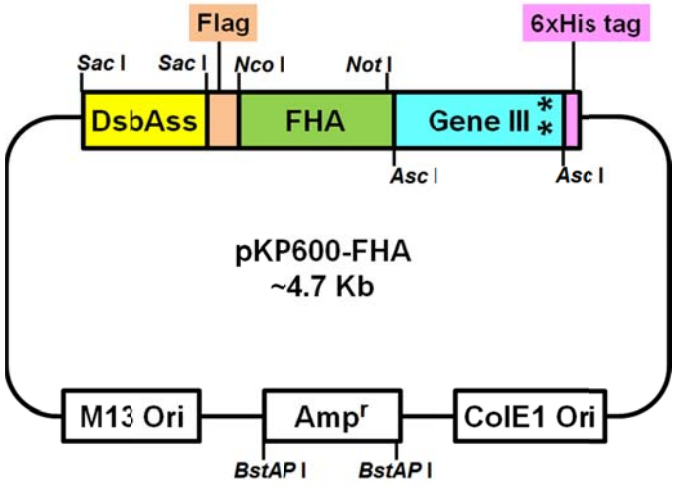


Figure 1. Vector map of pKP600 phagemid vector. FHA1 variants, cloned in-frame with the gene III coding sequence, are displayed on the surface of bacteriophage M13 viral particles. The DsbA signal sequence (Dsbss) directs the recombinant protein into the periplasm of *E. coli*. All FHA1 variants display a Flag epitope (DYKDDDDKL) at their N-termini. Gene III is truncated (~630 bp) and has opal and ochre stop codons at the end of the coding sequence, which are represented by two asterisks. They serve to prevent the expression of the six-histidine tag at the C-terminus in the phage-display format, which gets fused in-frame with the coding sequence of the FHA1 variant, after dropping out the Gene III coding region with restriction digestion with *Asc* I and vector re-ligation.

Table I. List of primers used for construction of phage-display and expression constructs.

Primer	Sequence (5′-3′)
FHA1-NcoI-Fw	TCCAGCCCATGGCGATGGAAAATATTACACAACCA
FHA1-NotI-Rv	CGAGTCTAGATGCGGCCGCGGTA
KM-S34C+S38C-FHA1	GAATTTGACCAGTTGTGCAGATTACGCGGCATACGATGTTTTCGC
KM-S154C-FHA1	GACTTTATTTTGTTCCAGGCATTGTTTGAATTTATCG
KM-C74S-FHA1	ACCTAAGTGATAATCAGAGGCTGGGTTACGTCCA
DsbA-Fw	CGCTGGCTGGTTTAGTTTTAGCGT
FHA1-BamH1-Fw	ATCATCGGATCCATGGAAAATATTACACAACCA
FHA1-EcoRI-Rv	GTAGATGAATTCGGTATTTTTAAGATTTGAACGGATACG
Seq-pGEX-Fw	CATGGCCTTTGCAGGGCTGGCAAG
FHA1-NdeI-Fw	TAGCTACATATGACCATGGCGATGGAAAATAT

2.3.2 Protein and phage enzyme-linked immunosorbent assay (ELISA)

To amplify the phage particles displaying the recombinant FHA1 variants, TG1 bacterial cells (5 mL; Stratagene, La Jolla, CA) harboring the phagemid DNA were infected at mid-log (OD_{600nm}= 0.5-0.6) with M13K07 helper phage (New England BioLabs) at a multiplicity of infection (MOI) of 20 for 1 h at 37°C at 150 rpm. Infected cells were centrifuged, and the pellet was resuspended in fresh Luria Bertani medium (LB: 10 g Tryptone, 5 g Yeast extract, and 10 g NaCl per liter) supplemented with 50 µg/mL Carbenicillin (CB) and 50 µg/mL Kanamycin (Kan), and phage were amplified overnight at 30°C at 250 rpm.

To set up the ELISA, biotinylated peptides (100 µL, 5 µg/mL) were immobilized on Nunc MaxiSorp flat-bottom 96 well plates (Thermo Fisher Scientific, Pittsburgh, PA) via

NeutrAvidin™ Biotin binding protein (100 µL, 10 µg/mL; Thermo Fisher Scientific) and blocked with 2% skim milk in Phosphate Buffered Saline (PBS: 0.14 M Sodium Chloride, 0.003 M Potassium Chloride, 0.002 M Potassium Phosphate, and 0.01 M Sodium Phosphate). After the wells were washed with PBS, the wells were filled with phage particles (100 µL of culture supernatant, diluted 1:2 with PBS containing 0.1% Tween 20; PBST) for 1 h and washed three times with PBST. The binding phage particles were detected using anti-M13 antibody conjugated to Horseradish Peroxidase (HRP; GE Healthcare), diluted 1:5000 with PBST. After washing away the unbound antibody, the chromogenic substrate, 2,2'-Azinobis (3-ethylbenzothiazoline-6-Sulfonic Acid) diammonium salt (Thermo Fisher Scientific), supplemented with hydrogen peroxide, was added (100 µL per well), and the absorbance of the green colored complex was measured at 405 nm on a POLARstar OPTIMA microtiter plate reader (BMG Labtech, Cary, NC).

For overexpression of the FHA1-GST fusion protein, BL21 DE3 cells (10 mL; Stratagene) harboring the expression vector was grown overnight at 30°C using the Overnight Express™ Autoinduction System 1 (Novagen, Madison, WI). The next day, cells were lysed using BugBuster® 10X Protein Extraction Reagent (Novagen), following the manufacturer's instructions. Cell lysate (100 µL diluted 1:5 with PBST) was incubated with the biotinylated peptides as described above and detected using anti-GST antibody conjugated to HRP (diluted 1:10,000 with PBST; GE Healthcare).

2.3.3 Peptides

Peptides were synthesized at the Research Resource Center, University of Illinois at Chicago and were >90% purity. All the peptides, except for the one used for isothermal titration calorimetry, were biotinylated at their N-terminus, and amidated at their C-terminus. The peptide used for ITC is SLEVpTEADATFYAKK (*22*). It was purified by high performance liquid chromatography (HPLC) and does not contain a linker and N- or C- terminal modifications. Many The peptides contain a tripeptide linker, SGS, between the N-terminal biotin and the amino acid sequence, and some required the addition of two or three lysine residues at the C-terminus to increase their solubility. The peptide sequences used in this study are listed in Table II.

Table II. Peptides used in this study.

Peptide name	Sequence	Reference
Rad9-pT	SGS-SLEVpTEADATFYAKK	(*22*)
Rad9-T	SGS-SLEVTEADATFAKK	-
Rad9-pS	SGS-SLEVpSEADATFAKK	-
Rad9-S	SGS-SLEVSEADATFAKK	-
Rad9-pY	SGS-SLEVpYEADATFAKK	-
Rad9-Y	SGS-SLEVYEADATFAKK	-
Plk1-pT	SGS-AGPMQSpTPLNGAKK	(*33*)
BRCT-pS	SGS-AYDIpSQVFPFAKKK	(*34*)

2.3.4 Fluorescence-based thermal shift (FTS) assay

The FTS assay was performed on a MxPro-Mx3005P instrument (Stratagene), following a published protocol (*35*). This real-time assay uses the SYBR® Green to monitor the thermal stability of a protein, as the dye fluorescence is low in the presence of a folded proteins and increases when it binds to the denatured form of the protein (Fig. 2). The default FRROX filter set was used with an excitation wavelength of 492 nm and emission wavelength of 610 nm. The proteins (wild-type FHA1 domain and its FHA1D2 variant) were mixed with SYPRO® Orange protein gel stain (Invitrogen; 5000X concentration in DMSO) and added to wells containing either the pT-containing peptide or PBS. The final reaction volume was 20 µL, in which the proteins were at 1 µM final concentration, the SYPRO® Orange dye was at 5X final concentration and the pT-containing peptide was at 50 µM or 250 µM final concentration. The assay was performed in duplicate wells of 96 well white PCR plate (Bio-Rad Laboratories, Hercules, CA) which was covered with optically clear Microseal® 'B' Film (Bio-Rad) and heated incrementally from 25°C to 95°C. The melting curve was obtained with fluorescence (R) values plotted on Y-axis and increasing temperature (°C) on X-axis. The mid-point of the curve is considered the inflection point of the melting curve.

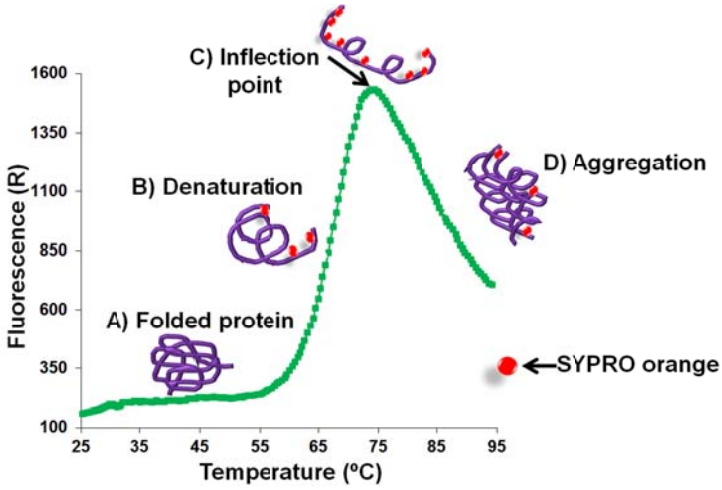


Figure 2. Fluorescence-Based Thermal Shift (FTS) Assay. A) When the protein is folded, the fluorescence of SYPRO orange is quenched in the aqueous environment. B) When the protein starts to denature, SYPRO orange dye interacts with the hydrophobic core of the protein and its fluorescence increases. C) Infliction point is referred to the temperature at which the dye exhibits maximum fluorescence. D) Beyond the infliction point temperature, there is a drop in fluorescence due to protein aggregation at a higher temperature. Figure adapted from (*35*).

2.3.5 Construction of phage-displayed FHAD2 library by random mutagenesis

To generate the FHA1 library, from which functional phage-displayed variants can be isolated by affinity selection, mutagenic PCR (*36*) was performed (primers: MP-FHA1-Fw and MP-FHA1-Rv), to amplify the coding sequence of the 3C-3S variant. The PCR product (~0.34 µg) was digested with *Nco* I and *Not* I, and subcloned into a phagemid vector (pKP600, ~1 µg), generating in-frame fusions with the gene III coding sequence at the C-terminus and a FLAG epitope tag at the N-terminus. The recombinant DNA was concentrated using a phenol:chloroform:isoamyl alcohol mixture

(Sigma-Aldrich Corp, St. Louis, MO), and electroporated into TG1 bacterial cells. The cells were allowed to recover for 40 minutes, with shaking at 250 rpm at 37°C, of which various dilutions (10 µL and 100 µL of 10^{-1} and 10^{-2}) were plated on 10 cm LB/CB agar plates. The remaining cells were plated on three 15 cm LB/CB (50 µg/mL) agar plates. The next day, colonies were counted on the titration plates and the library diversity was determined to be $2x10^4$ clones. The bacterial lawn on 15 cm plates was scraped, 30 mL of LB/CB media was inoculated with ~$1x10^8$ cells and grown to mid-log phase (OD_{600nm} = 0.5) followed by infection with M13KO7 helper phage (MOI=20) for 1 h at 37°C at low speed (150 rpm). Infected cells were collected by centrifugation, resuspended in 30 mL of fresh LB/CB/Kan (50 µg/mL) media and incubated overnight at 30°C, with shaking at 250 rpm. The following day, phage particles were precipitated using 1/5 volume of 24% polyethylene glycol (PEG) and 3 M NaCl, and the phage pellet was resuspended in PBS (1 mL), and stored in aliquots at -80°C with 16% final glycerol concentration. Similarly, another mutagenic library (FHA1D2 library) was constructed, with a final diversity of $6x10^7$ variants, using the pKP600 vector with the D2 variant coding sequence as the starting template.

2.3.6 Affinity selection for isolating functional, phage-displayed FHA1 variants

To isolate functional, phage-displayed FHA1 variants, three rounds of affinity selection were performed with its cognate pT peptide (SLEVpTEADATFYAKK) and the FHA1 library. All the selection steps were performed at room temperature. The biotinylated peptide (200 µL, 10 µg/mL) was immobilized on Nunc polystyrene tube (Thermo Fisher Scientific) via NeutrAvidin™ Biotin binding protein (200 µL, 20 µg/mL;

Thermo Fisher Scientific) and blocked with 2% skim milk in PBS. Then, $1x10^{10}$ phage particles from the library (*13*) were incubated in the tube with the blocked target for 1 h, followed by six washes with PBST and six washes with PBS. Phage particles bound to the target were eluted using 100 mM glycine-HCl (100 µL; pH 2.0), neutralized with 2 M Tris-base (6 µL; pH 10) and used to infect 800 µL of TG1 cells at mid-log phase (OD_{600nm}=0.5) for 40 min at 37ºC. The cells were plated after infection, scraped the next day; phage was amplified and precipitated as described above. The second and third rounds of affinity selection were conducted in the same manner, except that only 1/2 the volume of the eluted phage were used to infect bacterial cells after round 2 and 1/4 of the volume was used in round 3. After the third round of affinity selection, 96 individual clones were propagated as phage, followed by phage-ELISA to identify functional clones that recognize the pT peptide ligand. Positive binding clones were sequenced, and further specificity tests were performed.

2.3.7 Mutating phenylalanine at position 34 to 18 other amino acids in the FHA1D2 coding sequence

This experiment was performed in two steps: first, the phenylalanine (F) at position 34 was mutated to TAA stop codon, and second, the TAA stop codon was mutated one at a time to the 18 other L-amino acids. Following the Kunkel mutagenesis protocol (*32*), the pKP600 phagemid vector containing the FHA1D2 coding sequence was converted to single-stranded uracilated phagemid DNA. To mutate F34 to TAA, the oligonucleotide F-TAA (5’ phosphorylated; 5’- TGTTGAGATTACGCGTTATACGATGTTTTCGC -3’) was annealed to the single-stranded template at a molar ratio of 20:1 (oligonucleotide:

template), extended using T7 DNA polymerase and the covalently closed circular DNA was sealed by T4 DNA ligase (both enzymes were purchased from New England BioLabs). The covalently closed DNA was transformed into TG1 electrocompetent cells, the cells were recovered by shaking at 250 rpm at 37°C for 40 min and plated on 10 cm LB/CB agar plates. The next day, colony PCR was performed on 20 colonies using the using primers MP-FHA1-Fw (TGCTAGCGCCATGGCGATGGAAAATA) and MP-FHA1-Rv (5'- TCGACTGCGGCCGCGGTATTTTTA -3') and GoTaq® Flexi DNA polymerase (Promega Corporation, Madison, WI). The PCR products were cut with *Mlu* I restriction enzyme for 3 h at 37°C and run on a 1% agarose gel. The positive clones (that have the *Mlu* I restriction enzyme site) were sequenced at the Research Resource Center at UIC. To mutate TAA at position 34 to 18 other amino acids, the pKP600-FHA1D2 (F34TAA) phagemid DNA was used as the starting template. Eighteen oligonucleotides (sequences shown in Table III) were designed to anneal to the same single-stranded template, extended, ligated and transformed in TG1 electrocompetent cells as described above. For each mutation, 8 individual colonies were inoculated into 0.1 mL of LB/CB media in 96 deep well 2 mL polypropylene plates (Thermo Fisher Scientific), grown to mid-log phase (OD_{600nm} = 0.5), followed by infection with M13KO7 helper phage (MOI=20) for 1 h at 37°C at low speed (150 rpm). Infected cells were collected by centrifugation, resuspended in 0.5 mL of fresh LB/CB/Kan media and incubated overnight at 30ºC with shaking at 250 rpm. Next day, the cells were pelleted by centrifugation and the phage supernatant was used for phage ELISA. The Nunc MaxiSorp flat-bottom 96 well plates (Thermo Fisher Scientific) were coated with

monoclonal anti-Flag® M2 antibody (Sigma-Aldrich, 100 µL of 1 µg/mL) for 1 h, blocked with 2% skim milk in PBS (200 µL/well) and incubated with phage supernatant (100 µL) diluted 1:2 with PBST. Phage binding was detected by incubating with anti-M13 antibody conjugated to HRP (GE Healthcare, diluted 1:5000 with PBST, 100 µL/well) and the binding signal was measured at 405 nm on POLARstar OPTIMA microtiter plate reader (BMG Labtech), after adding ABTS (chromogenic substrate for HRP, 100 µL/well, Thermo Fisher Scientific).

Table III. Oligonucleotides used to mutate position 34 of FHA1.

Oligonucleotides	Sequence (5′-3′)
F-Leu	TGTTGAGATTACGCGCAGTACGATGTTTTCGC
F-Ile	TGTTGAGATTACGCGAATTACGATGTTTTCGC
F-Met	TGTTGAGATTACGCGCATTACGATGTTTTCGC
F-Val	TGTTGAGATTACGCGAACTACGATGTTTTCGC
F-Pro	TGTTGAGATTACGCGCGGTACGATGTTTTCGC
F-Thr	TGTTGAGATTACGCGGGTTACGATGTTTTCGC
F-Ala	TGTTGAGATTACGCGTGCTACGATGTTTTCGC
F-Tyr	TGTTGAGATTACGCGATATACGATGTTTTCGC
F-His	TGTTGAGATTACGCGGTGTACGATGTTTTCGC
F-Asn	TGTTGAGATTACGCGGTTTACGATGTTTTCGC
F-Lys	TGTTGAGATTACGCGTTTTACGATGTTTTCGC
F-Asp	TGTTGAGATTACGCGATCTACGATGTTTTCGC
F-Glu	TGTTGAGATTACGCGTTCTACGATGTTTTCGC
F-Cys	TGTTGAGATTACGCGACATACGATGTTTTCGC
F-Trp	TGTTGAGATTACGCGCCATACGATGTTTTCGC
F-Arg	TGTTGAGATTACGCGACGTACGATGTTTTCGC
F-Gly	TGTTGAGATTACGCGGCCTACGATGTTTTCGC
F-Glu	TGTTGAGATTACGCGCTGTACGATGTTTTCGC

2.4 Results and discussion

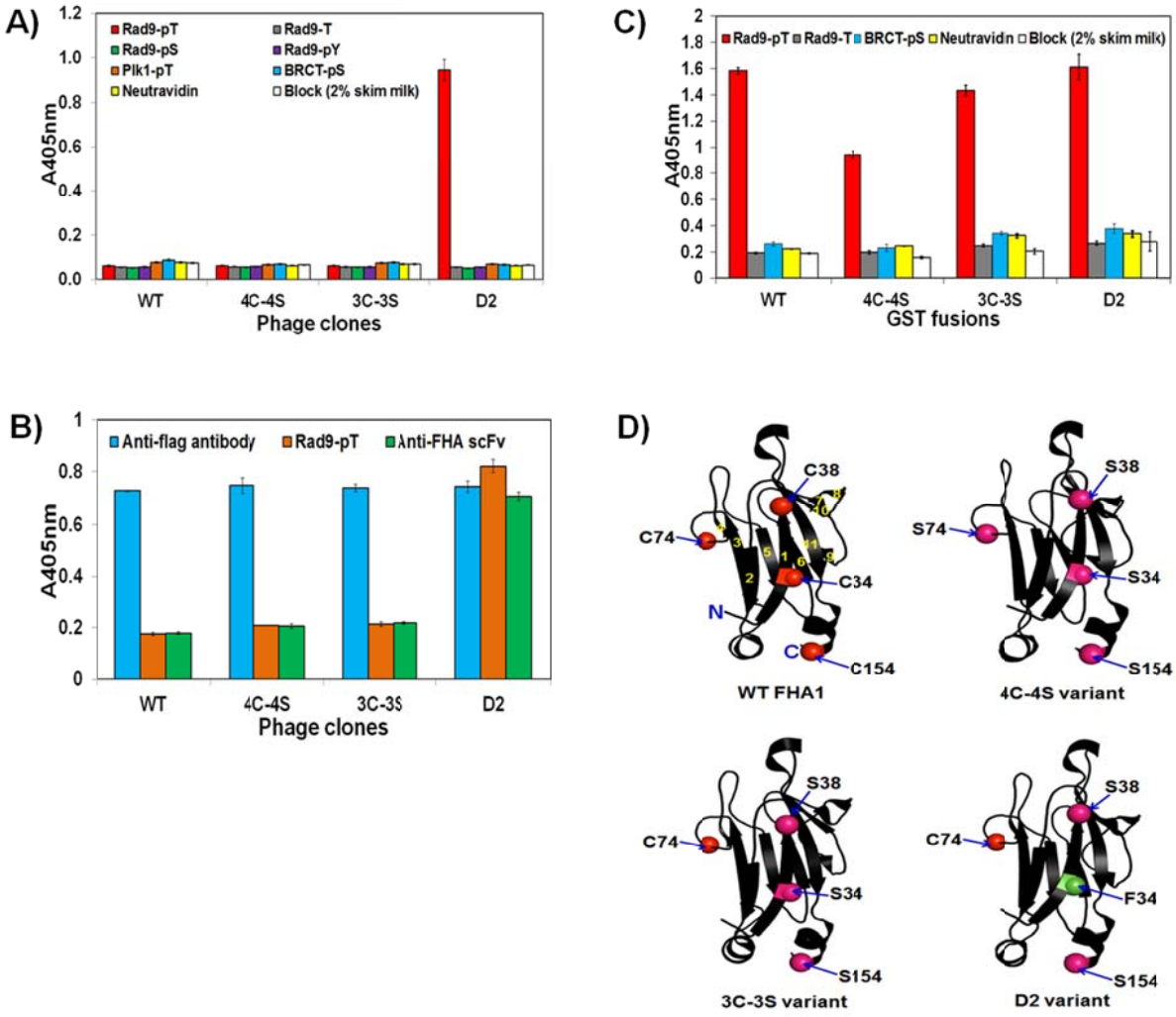
2.4.1 The phage-displayed wild-type FHA1 domain is non-functional

When the wild-type (WT) FHA1 domain was displayed as an N-terminal fusion to capsid protein III of M13 bacteriophage, binding to its cognate pT peptide (Rad9-pT: SLEVpTEADATFYAKK) was undetectable. We considered the possibility that loss of binding by the phage-displayed domain was a consequence of disulfide bonds forming incorrectly, upon folding in the oxidizing environment of the periplasm, during viral morphogenesis. The FHA1 domain has four cysteine residues (C34, C38, C74 and C154), but they do not participate in disulfide bond formation, as evident in its three-dimensional structure (*22*). To rule out the likelihood that the cysteines were interfering with phage-display of a functional form of the domain, three or all four cysteines were mutated to serines, in the 3C-3S and 4C-4S variants, respectively. Nevertheless, both of these variants still remained non-functional when examined by phage ELISA (Fig. 3A). The D2 variant was functional when displayed on the surface of bacteriophage M13. Affinity selection for isolating the D2 variant is described below.

As we could detect the presence of the Flag-epitope (with an anti-Flag antibody) at the N-terminus of the phage-displayed WT FHA1 domain and its 3C-3S and 4C-4S variants, we hypothesized that while the FHA1 domains were being displayed, they lacked the proper conformation for ligand recognition. To confirm our hypothesis, we generated a single-chain antibody fragment (scFv) that recognizes the FHA1 domain when it is folded, but not denatured (isolation of the conformation specific scFv is described in the Appendix B). It was observed that the scFv bound only to the D2

variant, and not to the WT or the other two variants, which confirms that they are misfolded in the bacterial periplasm, and consequently were non-functional when displayed on the surface of bacteriophage (Fig. 3B). This was a surprise given that the WT FHA1 domain, and its 3C-3S and 4C-4S variants, when expressed in the bacterial cytoplasm as Glutathione S-transferase (GST) fusions could bind their cognate peptide ligand. The D2 variant is functional, both as a GST fusion and when displayed on the surface of bacteriophage M13 (Fig. 3C). The three-dimensional structure of the WT FHA1 domain (*22*) represented in cartoon fashion with PyMol software (http://www.pymol.org) is shown in Fig. 3D, along with the hypothetical structures of the 3C-3S, 4C-4S and D2 variants.

Figure 3. ELISA for detecting interaction between the FHA1 domain and the immobilized pT peptide ligand. A) Phage ELISA demonstrates that the WT FHA1 domain and two of its variants, 3C-3S and 4C-4S, are non-functional and do not bind to the cognate pT peptide (Rad9-pT), when displayed on the surface of bacteriophage M13 as protein III fusions, whereas the D2 variant is functionally active when phage displayed. Binding of phage particles was detected using anti-M13 antibody conjugated to HRP. B) Phage ELISA shows that only the D2 variant, which was selected for functional display (i.e., binding to the Rad9-pT peptide ligand), is recognized by the anti-FHA scFv, an antibody that recognizes a conformational epitope in the FHA1 domain. The ELISA values were normalized to the detection of the Flag epitope by anti-flag antibody. Binding of phage particles was detected using anti-M13 antibody conjugated to HRP. C) GST fusions of the WT FHA1 domain and three of its variants, 3C-3S, 4C-4S, and D2, bind specifically to the cognate pT peptide (Rad9-pT) captured on plastic microtiter plate wells via Neutravidin. Binding was detected using anti-GST antibody conjugated to HRP. D) The structures of the WT FHA1 domain (*22*) (PDB code: 1G6G) and three variants- 4C-4S, 3C-3S and D2 prepared using PYMOL (http://www.pymol.org/) are shown as a cartoon representation with the cysteine (C), serine (S), and phenylalanine (F) residues shown as red, pink and green spheres respectively. The β-strands are labeled from 1 to 11 in the WT FHA1 domain structure and the N- and C- termini are noted.


A)
Rad9-pT
Rad9-T
Rad9-pS
Rad9-pY
Plk1-pT
BRCT-pS
Neutravidin
Block (2% skim milk)
A405nm
WT
4C-4S
3C-3S
D2
Phage clones
C)
Rad9-pT Rad9-T BRCT-pS Neutravidin Block (2% skim milk)
A405nm
WT
4C-4S
3C-3S
D2
GST fusions
B)
Anti-flag antibody
Rad9-pT
Anti-FHA scFv
A405nm
WT
4C-4S
3C-3S
D2
Phage clones
D)
C38
C74
C34
N
C
C154
WT FHA1
S38
S74
S34
S154
4C-4S variant
S38
C74
S34
S154
3C-3S variant
S38
C74
F34
S154
D2 variant

To investigate whether the WT FHA1 domain could be properly phage-displayed with overexpression of chaperone proteins, we co-expressed five chaperones (dsbA, dsbC, fkpA, skp and surA) that have been shown to improve folding and yields of scFv antibody fragments (*37*). Co-expression of the chaperone proteins, however, did not restore the activity of the phage-displayed WT FHA1 domain (Fig. 4A).

Another possibility we investigated is whether a different signal sequence for protein secretion would overcome the inability to display a functional form of the FHA1 domain. In our original effort, the WT FHA1 was transported to the bacterial periplasm via the DsbA signal sequence (MKKIWLALAGLVLAFSASAE), which transports proteins to the bacterial periplasm as they are being translated (*38*). When we tried a different signal sequence, TorA (MNNNDLFQASRRRFLAQLGGLTVAGMLGPS LLTPRRATAA QAA), which transports only fully folded proteins to the periplasm (*39-41*), the activity of the phage-displayed WT FHA1 domain was only partially restored showing a weak signal even after 50-fold concentration of the phage particles, indicating a very low level of display. The D2 variant that was functional while using either of the signal sequences was used as a positive control; however, the phage particles needed to be concentrated ~100 times when using the TorA signal sequence to give the same level of ELISA signal as the DsbA signal sequence (Fig. 4B). Therefore, with the DsbA signal sequence, the phage-display efficiency is good, however, the protein is not functional and with the TorA signal sequence, weak activity is observed, but the display efficiency is highly compromised.

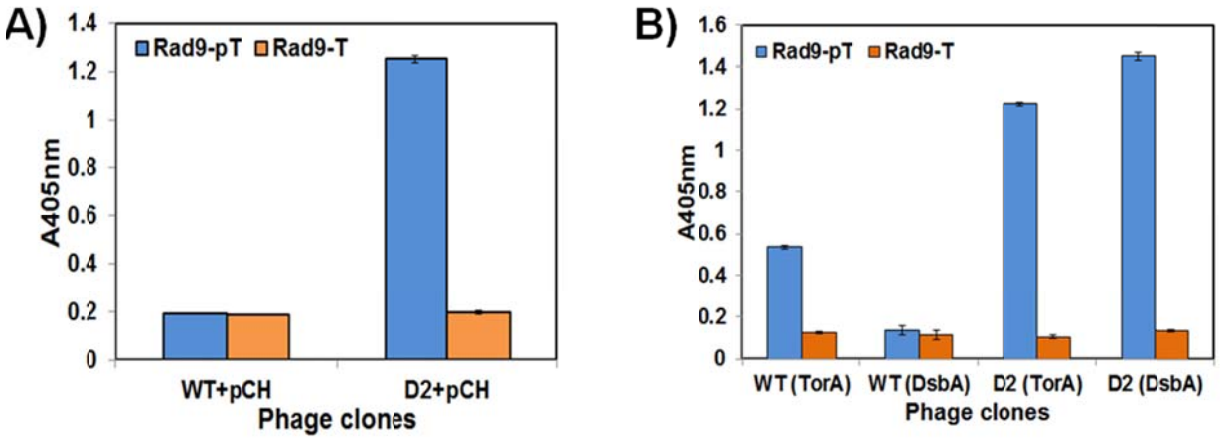


Figure 4. Effect of chaperones and TorA signal sequence on functional phage-display. A) Five protein folding chaperones were expressed from the pCH vector (*37*) and transported to the bacterial periplasm, along with the phage-displayed FHA1 domains to facilitate their folding. The WT domain remained non-functional and did not demonstrate any binding to the cognate pT peptide ligand (Rad9-pT), whereas the D2 variant was functional, with or without the co-expression of the five chaperones. B) FHA domains were transported to the bacterial periplasm with the DsbA and TorA signal sequences. The WT domain was non-functional when transported via the DsbA signal sequence, however, when TorA signal sequence was used, binding to the cognate pT peptide (Rad9-pT) was weakly detected. The D2 variant that was functional while using either of the signal sequences was used as a positive control.

2.4.2 Isolating functional phage-displayed FHA1 variants through directed evolution

A mutagenic library comprising of $2x10^4$ variants was constructed by mutagenic PCR (*36*), using the 3C-3S version of the domain as the starting template. (We were interested in leaving one cysteine in the domain, which was away from the binding surface, so in future experiments it could be derivatized through maleimide coupling chemistry for immobilization to resin). Mutagenic PCR was performed to amplify and generate mutations randomly across the coding region of the 3C-3S variant. This method (*36*) generates an error rate of 0.66% per position and the estimated mutants in the library are ~4% wild-type, 12%, 20%, 22%, 18%, and 12%, with 1 to 5 mutations, respectively, and 12% with 6 or more mutations. From sequencing 30 clones chosen at random from the library, it was observed that the numerical distribution of mutations in the library matched the predictions. Affinity selection of the library of variants with the cognate pT peptide, captured on plastic wells via a NeutrAvidin^{TM} coating, yielded six clones that specifically bound to the phosphorylated, but not the non-phosphorylated, peptide ligand (Fig. 5). Mutations observed in the six binding FHA1 variants are listed in Table IV. The D2 variant, which is the tightest binder, carried one mutation (S34F), suggesting that this position, in the β1 strand, facilitates proper folding of the phage-displayed FHA1 domain.

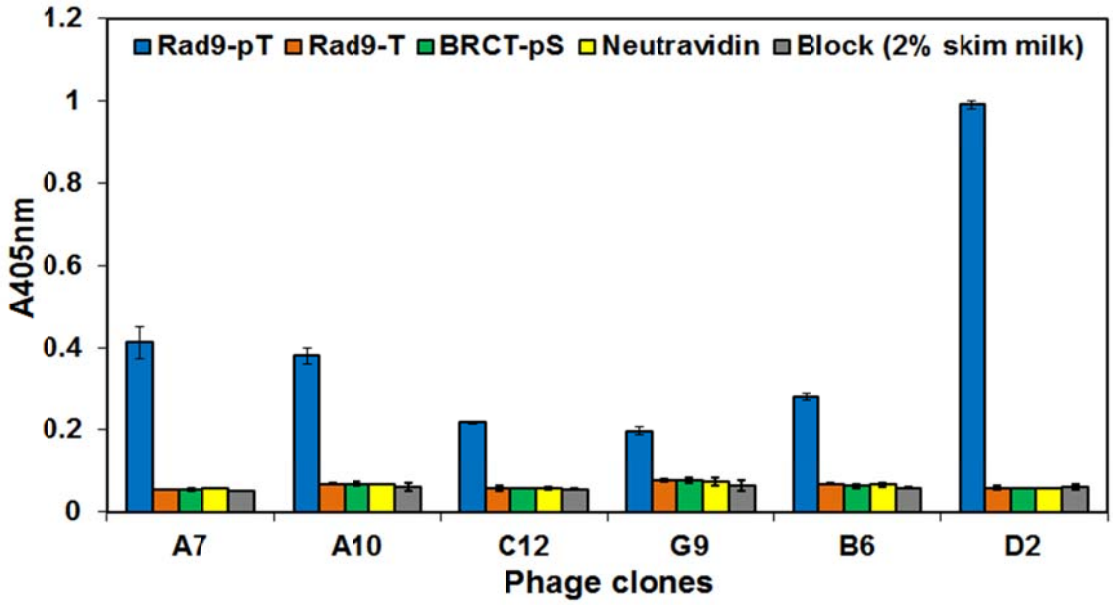


Figure 5. Phage ELISA of FHA1 domain variants selected to bind to a phosphopeptide. Three rounds of affinity selection against the pT peptide (Rad9-pT: SGS-SLEVpTEADATFYAKK), using a phage-displayed library of FHA1 variants, yielded six functional variants that specifically bound only to the cognate pT peptide (Rad9-pT) and not to any of the negative controls tested.

Table IV. Mutations observed in the functional phage-displayed FHA1 variants isolated after three rounds of affinity selection.

FHA variants[a]	Mutations	Position[b]
A7	N121Y	β8-β9 loop
A10	Q25H	N-terminus
	N121Y	β8-β9 loop
	S142R	β11 strand
C12	W66R	β3 strand
G9	Q13R	N-terminus
	I104V	β6-β7 loop
	N121Y	β8-β9 loop
	R164S	C-terminus
B6	S34A	β1-strand
D2	S34F	β1-strand

[a] FHA1 variants that specifically bind to the Rad9-pT peptide when phage-displayed.
[b] β = beta-strand. Loop is the region between two β-strands. The N-terminal mutation is present before the β1-strand and the C-terminal mutation is after the β11-strand.

2.4.3 Hydrophobic residues are preferred at position 34 for functional phage-display

To survey what amino acids at position 34 permit proper folding for functional phage-display, this position was mutated one at a time 34 in the D2 coding sequence by Kunkel mutagenesis to the other 18 L-amino acids (*32*). Phage ELISA showed that the FHA1 domain bound to the Rad9-pT peptide only when a hydrophobic amino acid (F, A, M, I, L, Y, and V) was present at position 34, with a few exceptions (W, C, and P). On the other hand, hydrophilic residues at this position rendered the FHA1 domain non-

functional with respect to phage-display (Fig. 6). Of all the 20 amino acids, phenylalanine (i.e., the D2 variant) at position 34 that gave the highest level of binding when phage-displayed.

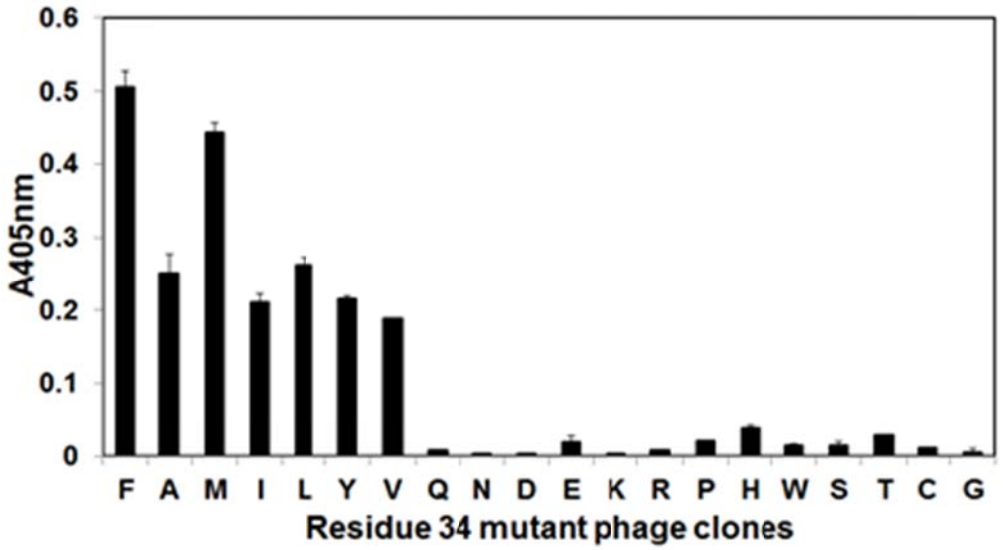


Figure 6. Hydrophobic residues are preferred at position 34 for activity of phage-displayed FHA1 variants. Through Kunkel mutagenesis, position 34 in the β1 strand of the D2 FHA1 domain was replaced one at a time with each amino acid and binding to the immobilized Rad9-pT peptide was monitored by phage ELISA. Phage binding was detected using anti-M13 antibody conjugated to HRP. The phage-displayed FHA1 variants that contain a hydrophobic residue at position 34 (with exceptions of W, C, and P) retain binding to the Rad9-pT. Presence of polar or hydrophilic amino acids at this position fails to rescue binding of the phage to the Rad9-pT peptide. All the variants display an N-terminal Flag peptide and the amount of phage particles used for each variant in this experiment has been normalized for this epitope.

2.4.4 Comparing the thermal stability of the wild-type FHA1 domain with the FHA1D2 variant

The WT FHA1 domain and the D2 variant were mixed with SYPRO orange dye with or without the Rad9-pT-containing peptide and heated from 25°C to 95°C in a real-time PCR experiment known as the Fluorescence-based thermal shift (FTS) assay (*35*). In this assay, initially, the fluorescence of SYPRO orange dye is quenched in the aqueous environment of a folded protein, but as the protein starts to unfold upon heating, the dye interacts with the hydrophobic core of the protein and its fluorescence increases. Therefore, an increase in fluorescence of the dye is directly proportional to protein unfolding, until a temperature (referred to as the inflection point here) is reached at which the dye fluorescence decreases due to aggregation and precipitation of the protein. It was observed that the inflection point of the melting curve (mid-point) of D2 variant was ~ 5°C lower than that of the WT FHA1 domain (Fig. 7). Even though, the S34F mutation, generated a functional phage-displayed FHA1 variant; the thermal stability of the original domain was compromised.

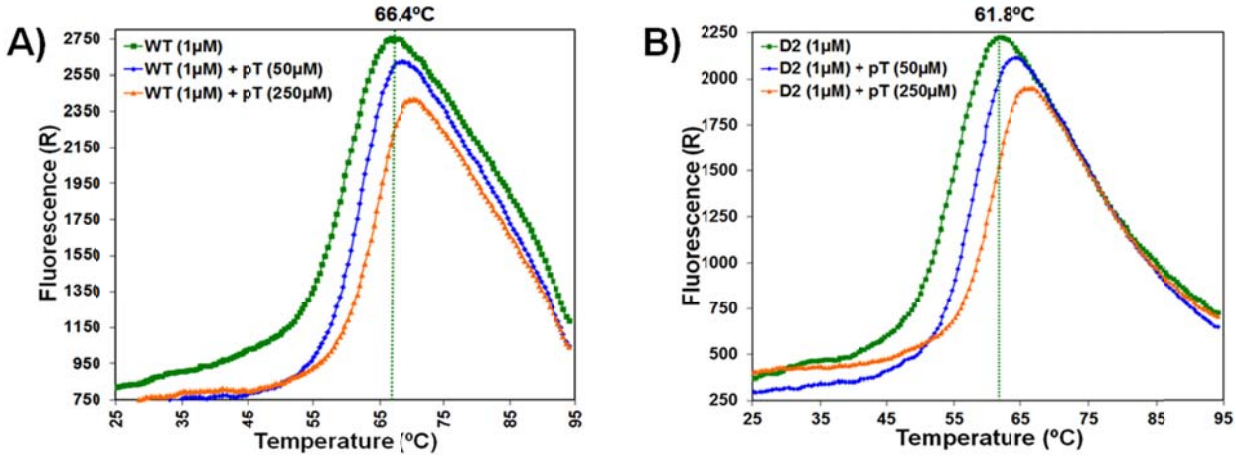


Figure 7. Determining the thermal stability of the WT FHA1 domain and the FHA1D2 variant by FTS Assay. The protein domains were mixed with SYPRO orange dye, with or without the Rad9-pT-containing peptide and heated from 25ºC to 95ºC on an Mx3000P Real-time thermal cycler instrument. The inflection point of the melting curve of the (A) WT FHA1 domain and the (B) D2 variant, without the Rad9-pT peptide (green curves), were determined to be 66.4°C and 61.8°C, respectively. In the presence of 50 μM and 250 μM of the Rad9-pT peptide, increase in inflection points of (1.7°C and 3.8°C) and (2.6°C and 4.6°C) were observed for the WT FHA1 domain and D2 variant, respectively.

2.4.5 Determining the dissociation equilibrium constants (K_d) by isothermal calorimetry (ITC)

The affinity of the D2 variant and the WT FHA1 domain to the pT peptide (SLEVpTEADATFYAKK) was determined by ITC (Fig. 8). The sample cell contained the purified FHA1 domains (35 μM) and the injection syringe was filled with the pT peptide (350 μM). The D2 variant bound with a K_d of ~0.97 μM, which was similar to the K_d (~1 μM) of WT FHA1 domain. The FHA1 domains bound to the pT peptide with a

stoichiometry of 1:1, and fit a single-site binding model. The previously reported K_d for WT FHA1 domain by ITC is 0.53 μM (*22*).

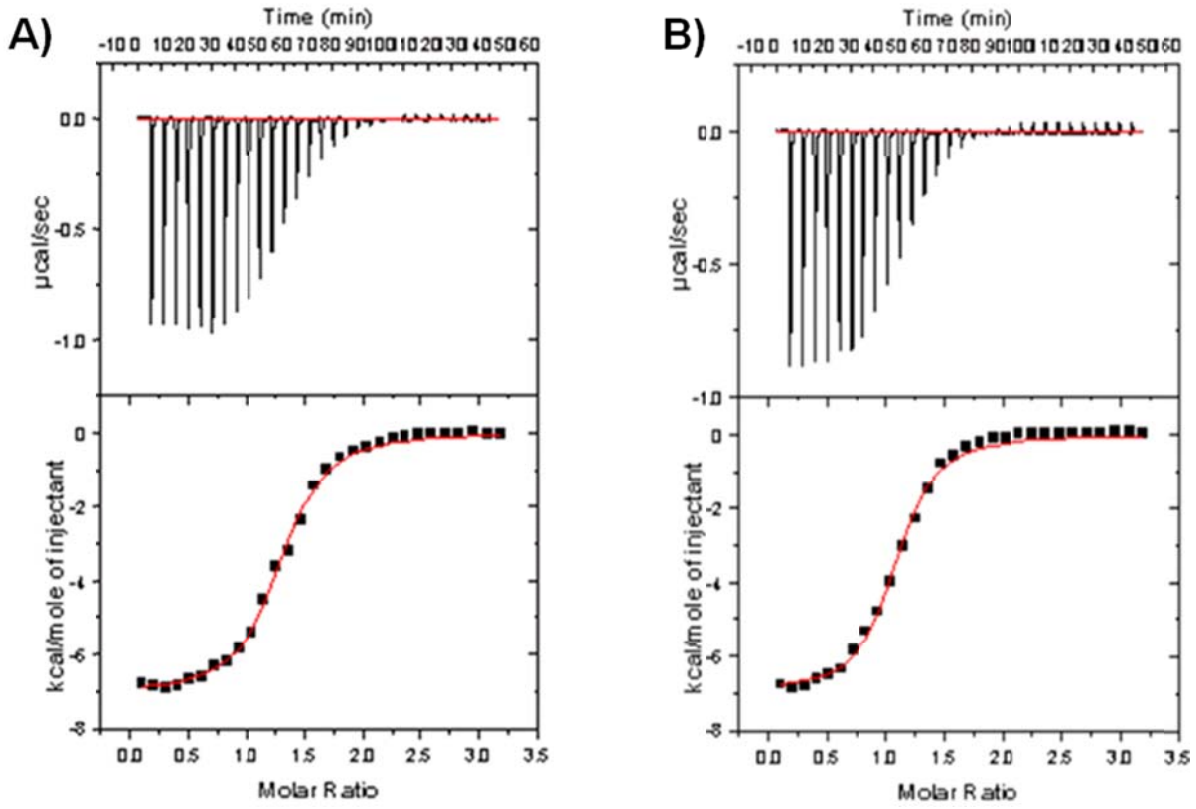


Figure 8. Determining the K_d in solution by ITC. A fixed amount (10 μL) of the pT peptide (SLEVpTEADATFYAKK; 350 μM) was injected from the syringe into the sample cell containing the FHA1 domain (35 μM). The K_d values of the (A) WT FHA1 domain and the (B) D2 variant were determined to be 1 ± 0.17 μM and 0.97 ± 0.15 μM respectively. A stoichiometry of 1:1 was observed for both the FHA1 domains with their pT peptide complexes. Each measurement was repeated three times.

2.5 Conclusions

The phage-displayed proteins fold in the oxidizing environment of the bacterial periplasm; therefore, the presence of cysteine residues can cause protein misfolding resulting in a non-functional displayed protein. Since, this was a bottleneck that we observed; we mutated all the cysteines to serines (except one for future conjugation) and then identified mutations in this mono-cysteine scaffold that facilitated functional phage-display. We used directed evolution to engineer the FHA1 domain of yeast Rad53 protein and generate variants of this domain that are functional when displayed on the surface of bacteriophage M13. We further characterized one variant, FHA1D2 that had the highest binding signal to the cognate pT-containing peptide and determined its thermal stability by using a real time PCR based assay. In this endeavor, we found out that the FHA1D2 variant was thermally less stable than the wild-type FHA1 domain. It is always preferred to use a thermal stable protein as a starting scaffold because it offers various advantages, such as greater stability at room temperature, more resistance to proteolysis and is soluble at high concentrations. Therefore, our next step was to generate thermal stable variants of the FHA1D2 domain which is described in detail in Chapter 3.

2.6 References

1. Kay, B. K., Kasanov, J., and Yamabhai, M. (2001) Screening phage-displayed combinatorial peptide libraries, *Methods (San Diego, Calif 24*, 240-246.

2. Turunen, L., Takkinen, K., Soderlund, H., and Pulli, T. (2009) Automated panning and screening procedure on microplates for antibody generation from phage display libraries, *J Biomol Screen 14*, 282-293.

3. Sheets, M. D., Amersdorfer, P., Finnern, R., Sargent, P., Lindquist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J. C., and Marks, J. D. (1998) Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens, *Proc Natl Acad Sci U S A 95*, 6157-6162.

4. Bliss, J. M., Sullivan, M. A., Malone, J., and Haidaris, C. G. (2003) Differentiation of Candida albicans and Candida dubliniensis by using recombinant human antibody single-chain variable fragments specific for hyphae, *Journal of clinical microbiology 41*, 1152-1160.

5. Rothe, C., Urlinger, S., Lohning, C., Prassler, J., Stark, Y., Jager, U., Hubner, B., Bardroff, M., Pradel, I., Boss, M., Bittlingmaier, R., Bataa, T., Frisch, C., Brocks, B., Honegger, A., and Urban, M. (2008) The human combinatorial antibody library HuCAL GOLD combines diversification of all six CDRs according to the natural immune system with a novel display method for efficient selection of high-affinity antibodies, *J Mol Biol 376*, 1182-1200.

6. Koide, A., Bailey, C. W., Huang, X., and Koide, S. (1998) The fibronectin type III domain as a scaffold for novel binding proteins, *J Mol Biol 284*, 1141-1151.

7. Scholle, M. D., Kehoe, J. W., and Kay, B. K. (2005) Efficient construction of a large collection of phage-displayed combinatorial peptide libraries, *Comb Chem High Throughput Screen 8*, 545-551.

8. Ernst, A., Sazinsky, S. L., Hui, S., Currell, B., Dharsee, M., Seshagiri, S., Bader, G. D., and Sidhu, S. S. (2009) Rapid evolution of functional complexity in a domain family, *Science signaling 2*, ra50.

9. Huber, T., Steiner, D., Rothlisberger, D., and Pluckthun, A. (2007) In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: The Na(+)-citrate symporter CitS as an example, *Journal of structural biology 159*, 206-221.

10. Hufton, S. E., van Neer, N., van den Beuken, T., Desmet, J., Sablon, E., and Hoogenboom, H. R. (2000) Development and application of cytotoxic T

lymphocyte-associated antigen 4 as a protein scaffold for the generation of novel binding ligands, *FEBS Lett 475*, 225-231.

11. Nilsson, F. Y., and Tolmachev, V. (2007) Affibody molecules: new protein domains for molecular imaging and targeted tumor therapy, *Current opinion in drug discovery & development 10*, 167-175.

12. Skerra, A. (2008) Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities, *The FEBS journal 275*, 2677-2683.

13. Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S., Clarke, K. F., Conquer, J. S., Crofts, A. M., Crowther, S. R., Dyson, M. R., Flack, G., Griffin, G. J., Hooks, Y., Howat, W. J., Kolb-Kokocinski, A., Kunze, S., Martin, C. D., Maslen, G. L., Mitchell, J. N., O'Sullivan, M., Perera, R. L., Roake, W., Shadbolt, S. P., Vincent, K. J., Warford, A., Wilson, W. E., Xie, J., Young, J. L., and McCafferty, J. (2007) Application of phage display to high throughput antibody generation and characterization, *Genome biology 8*, R254.

14. Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display, *Protein Eng Des Sel 23*, 279-288.

15. Karatan, E., Merguerian, M., Han, Z., Scholle, M. D., Koide, S., and Kay, B. K. (2004) Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain, *Chem Biol 11*, 835-844.

16. Fisher, R. D., Ultsch, M., Lingel, A., Schaefer, G., Shao, L., Birtalan, S., Sidhu, S. S., and Eigenbrot, C. (2010) Structure of the complex between HER2 and an antibody paratope formed by side chains from tryptophan and serine, *J Mol Biol 402*, 217-229.

17. Boersma, Y. L., Chao, G., Steiner, D., Wittrup, K. D., and Pluckthun, A. (2011) Bispecific designed ankyrin repeat proteins (DARPins) targeting epidermal growth factor receptor inhibit A431 cell proliferation and receptor recycling, *J Biol Chem 286*, 41273-41285.

18. Cohen, P. (2001) The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture, *Eur J Biochem 268*, 5001-5010.

19. Bangalore, L., Tanner, A. J., Laudano, A. P., and Stern, D. F. (1992) Antiserum raised against a synthetic phosphotyrosine-containing peptide selectively recognizes p185neu/erbB-2 and the epidermal growth factor receptor, *Proc Natl Acad Sci U S A 89*, 11637-11641.

20. Sun, T., Campbell, M., Gordon, W., and Arlinghaus, R. B. (2001) Preparation and application of antibodies to phosphoamino acid sequences, *Biopolymers 60*, 61-75.

21. Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999) The FHA domain is a modular phosphopeptide recognition motif, *Mol Cell 4*, 387-394.

22. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms, *Mol Cell 6*, 1169-1182.

23. Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I. J., Pei, D., and Tsai, M. D. (2000) Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9, *J Mol Biol 304*, 941-951.

24. Byeon, I. J., Yongkiettrakul, S., and Tsai, M. D. (2001) Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes, *J Mol Biol 314*, 577-588.

25. Wang, P., Byeon, I. J., Liao, H., Beebe, K. D., Yongkiettrakul, S., Pei, D., and Tsai, M. D. (2000) II. Structure and specificity of the interaction between the FHA2 domain of Rad53 and phosphotyrosyl peptides, *J Mol Biol 302*, 927-940.

26. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2, *Mol Cell 9*, 1045-1054.

27. Yongkiettrakul, S., Byeon, I. J., and Tsai, M. D. (2004) The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues, *Biochemistry 43*, 3862-3869.

28. Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E. S., Wu, M. R., Su, M. I., Heierhorst, J., and Tsai, M. D. (2008) Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade, *Mol Cell 30*, 767-778.

29. Li, H., Byeon, I. J., Ju, Y., and Tsai, M. D. (2004) Structure of human Ki67 FHA domain and its binding to a phosphoprotein fragment from hNIFK reveal unique recognition sites and new views to the structural basis of FHA domain functions, *J Mol Biol 335*, 371-381.

30. Pennell, S., Westcott, S., Ortiz-Lombardia, M., Patel, D., Li, J., Nott, T. J., Mohammed, D., Buxton, R. S., Yaffe, M. B., Verma, C., and Smerdon, S. J. (2010) Structural and functional analysis of phosphothreonine-dependent FHA domain interactions, *Structure 18*, 1587-1595.

31. Pershad, K., Sullivan, M. A., and Kay, B. K. (2011) Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats, *Anal Biochem 412*, 210-216.

32. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A. (2000) Phage display for selection of novel binding peptides, *Methods in enzymology 328*, 333-363.

33. Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain, *Cell 115*, 83-95.

34. Williams, R. S., Lee, M. S., Hau, D. D., and Glover, J. N. (2004) Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1, *Nat Struct Mol Biol 11*, 519-525.

35. Giuliani, S. E., Frank, A. M., and Collart, F. R. (2008) Functional assignment of solute-binding proteins of ABC transporters using a fluorescence-based thermal shift assay, *Biochemistry 47*, 13974-13984.

36. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR, *PCR methods and applications 3*, S136-140.

37. Schaefer, J. V., and Pluckthun, A. (2010) Improving Expression of scFv Fragments by Co-expression of Periplasmic Chaperones, in *Antibody engineering* (Kontermann, R., and Du¨bel, s., Eds.), pp 345-361, Springer-Verlag Berlin Heidelberg.

38. Schierle, C. F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. (2003) The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway, *J Bacteriol 185*, 5706-5713.

39. Paschke, M., and Hohne, W. (2005) A twin-arginine translocation (Tat)-mediated phage display system, *Gene 350*, 79-88.

40. Wu, L. F., Ize, B., Chanal, A., Quentin, Y., and Fichant, G. (2000) Bacterial twin-arginine signal peptide-dependent protein translocation pathway: evolution and mechanism, *J Mol Microbiol Biotechnol 2*, 179-189.

41. Speck, J., Arndt, K. M., and Muller, K. M. (2011) Efficient phage display of intracellularly folded proteins mediated by the TAT pathway, *Protein Eng Des Sel 24*, 473-484.

CHAPTER 3

GENERATING THERMAL STABLE VARIANTS OF THE FORKHEAD-ASSOCIATED DOMAIN AND IDENTIFYING RESIDUES CRITICAL FOR PHOSPHOTHREONINE PEPTIDE RECOGNITION

Part of this work have been accepted for publication

Pershad, K., and Kay, B. Generating Thermally Stable Variants of Protein Domains through Phage-display, *Methods Journal* (*In vitro* Protein Selection issue), with targeted publication date in February/March 2013.

Part of this work has been published

Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

3.1 Abstract

Advancements in protein evolution strategies and computational biology have made it possible to generate proteins with useful and desired characteristics such as improved protein stability, higher affinity and specificity for their target, longer half-life, and greater expression in *Escherichia coli*. One route to identifying mutations that yield domains that remain folded, and active, at elevated temperatures is through the use of directed evolution. In a previous chapter of this thesis, I identified a functional, phage-displayed variant (FHA1D2) of the eukaryotic Forkhead-associated (FHA1) domain. However, the D2 variant was unfortunately less thermal stable than the wild-type FHA1 domain. By applying high temperature as a selective pressure during biopanning, I isolated thermal stable variants from a phage-displayed library of FHA1D2 variants generated by mutagenic PCR. The FHA1G2 variant showed ~ 13ºC increase in thermal stability, compared to the starting variant, as measured in a fluorescence-based thermal shift assay. Interestingly, such variants are expressed at high levels in *E. coli* (30-60 mg/L shake flask) and remain soluble in solution at higher concentrations for longer periods of time than the wild-type form of the domain. Subsequent alanine scanning of a thermal stable variant (FHA1G2) led to the identification of 10 residues (L78, R83, L84, S105, T106, G133, V134, G135, V136, and D139) that play an important role in interacting with the pT peptide. We propose that randomization of these residues within this thermally stable scaffold offers great potential for generating new anti-phosphopeptide binding specificities.

3.2 Introduction

In protein design research, it is possible to engineer a variety of 'new traits' into proteins, such as high affinity and specificity for their target protein, improved protein stability, longer half-life, and greater expression in *E. coli*. This can be achieved either by site-directed mutagenesis of specific residues or by random mutagenesis and then selecting for variants with desired properties and applying selective pressure. Using random or site-directed mutagenesis, proteins variants can be selected with higher thermal stability and expression (*1*), slower off rates (*2, 3*), and higher affinity for their target (*4-6*).

As phage particles are intrinsically very stable, one can subject them to conditions that disrupt the three-dimensional structure and binding of a displayed protein, without affecting the integrity or infectivity of the phage particle. For example, bacteriophage M13 particles can be heated to high temperatures (60ºC), or incubated in 10 M urea, 50 mM dithiothreitol, 7 M guanidine chloride, acidic (i.e., pH 2.0) or basic (i.e., pH 10.0) solutions (*7*), certain organic solvents (*8*), proteases (i.e., trypsin, chymotrypsin, endoproteinase Glu-C, proteinase K), without loss of infectivity . Thus, if one has a wild-type form of a protein that loses its three-dimensional structure or binding function under a particular condition, one can create a mutagenized library and select for folded, stable proteins by applying that particular condition as a selective pressure during affinity selection.

Phage-displayed library of variants can be subjected to selective pressures, such as heat, protease treatment, or protein denaturants, and if the phage displaying protein

variants remain folded and active, they can be isolated, propagated and enriched with each subsequent round of selection. Proteins can be inserted between the N2 and CT domains (Fig. 1) of the minor coat protein (protein III) of the phage without disturbing the infectivity of the recombinant phage particle. For example, one can insert the 110 amino acid barnase protein (*7*), a bacterial ribonuclease T1 (*9*), β-lactamase (*9*), and cold shock protein B (*9*) at this site, expose the recombinant phage particles to trypsin, and cleavage in the "guest' proteins would render the phage particles non-infective. In a mock library selection experiment (*10*), with two barnase mutants of differing stability, the barnase mutant that was trypsin resistant at a certain temperature could be enriched 10^4-fold from the other barnase variant. Phage bearing a more stable form of ribonuclease T1 could also be enriched via trypsin cleavage from a mixture of phage particles bearing stable and less stable variants or ribonuclease T1 (*9*). Furthermore, a library of ribonuclease T1 proteins, mutated at three non-interacting sites on the surface of the protein, could be screened for protease resistance (i.e., only resistant phage particles would be infective and propagated). When the conformational stability of the selected proteins was examined by thermal denaturation, all the mutants showed increased stability. In a similar approach, display of hexahistidine-tagged ubiquitin at the N-terminus of pIII, led to capture of particles to surfaces coated with nickel-nitrilotriacetic acid chelates. When the phage particles were exposed to chymotrypsin, only ubiquitin variants that carried hydrophobic residues at the core could be recovered (*11*). Well-folded proteins remained bound to the support in the presence of chymotrypsin, whereas misfolded ones were eluted away (*12*).

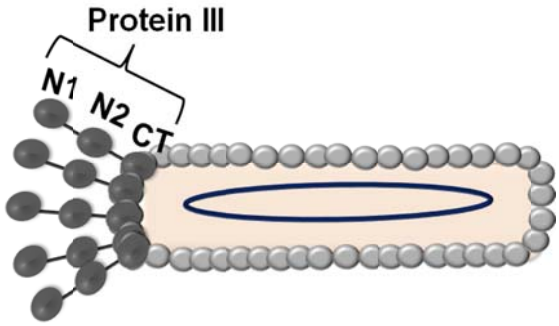


Figure 1. Domains in the minor coat protein (protein III) of bacteriophage M13. Protein III is composed of three domains, two N-terminal domains (N1 and N2), and the C-terminal (CT) domain. During phage infection, the N2 domain first binds the F-pilus of an F^+ bacterial cell (*13*) and then the N2 domain binds to the bacterial membrane receptor TolA (*14*). Both the N-terminal domains are essential for phage infection of bacterial cells (*15*).

The same approach has been used to select for protein folding as it relates to binding a peptide ligand. Eight contiguous residues, which comprise one-third of the entire β-sheet in the WW domain (34 amino acids), were randomized in a phage-display construct, and were selected for binding to its peptide ligand, GTPPPPYTVG, which was biotinylated and captured by immobilized streptavidin (*12*). Affinity selections were performed in the presence of proteinase K, as a selective pressure in favor of mutated WW domains that folded properly for binding. Finally, the method of using proteolytic selection for sorting a phage-display library according to the resistance of the mutants towards proteolysis has been applied to examine the natural variability between the

related ribonucleases, barnase and binase, from *Bacillus amyloliquefaciens* and *B. intermedius*, respectively, and to select for stabilized variants of barnase (*10*).

Heat has also been used as a selective agent for phage-display. Solutions of bacteriophage M13 can be heated at 65ºC for 15 minutes, without loss of infectivity; however, whether the displayed protein or domain retains its structure and function, will depend on the intrinsic thermal stability of the protein or domain, independent of protein III or the viral particle. In one example, variants of the *B. subtilis* endoxylanase (XynA) protein were displayed on phage, and affinity selected against an immobilized endoxylanase inhibitor, each time preceded by an incubation step at an elevated temperature (*16*). Analysis of binders isolated after three rounds of selection led to the identification of mutations that conferred enhanced thermal stability; each single mutation increased the half-inactivation temperature by 2-3ºC over that of the wild-type enzyme. Interestingly, none of these mutations was identified in earlier engineering studies. A few examples of proteins created by this method that are resistant to aggregation upon heating are single-domain human antibody variable domains that bind to protein A (*17*) or chicken egg lysozyme (*18*).

Finally, it should be noted that instead of using random mutagenesis and phage selection to identify protein variants with thermal stability, it is also possible to use a computational approach (*19, 20*) when the three-dimensional structure is known. In such case, certain amino acid residues, which are predicted to cause instability in protein folding, due to their large range of movement, are replaced with smaller ones that permit

Thermal stability has a number of practical advantages for proteins. First, one can use a protein at elevated temperatures, and it still remains well-folded and active. Second, more of the protein will be in the folded state at room temperature. Third, thermally stable proteins have been observed to accumulate to a higher degree in *E. coli*. Fourth, often thermally stable proteins remain more soluble at high concentrations, because they adopt a stable structure that is less prone to aggregation, which enhances the possibility of studying the kinetics of protein-protein interactions or performing three-dimensional structural determination by NMR or X-ray crystallography. Stable proteins serve as excellent scaffolds for generating affinity reagents to target proteins of interest. Proteins with these superior properties can further be used for a variety of purposes, such as biosensors, assay reagents, or for fabrication of novel materials.

3.3 Materials and methods

3.3.1 Constructing and amplification of a phage-displayed library of FHA1D2 variants

To generate a phage-displayed library of FHA1D2 variants, from which thermal stable variants can be isolated by affinity selection, mutagenic PCR (*21*) was performed using primers: MP-FHA1-Fw (5'- TGCTAGCGCCATGGCGATGGAAAATA -3') and MP-FHA1-Rv (5'- TCGACTGCGGCCGCGGTATTTTTA -3'), to amplify the coding sequence of the FHA1D2 variant. The PCR product was digested with *Nco* I/*Not* I restriction enzymes, and subcloned into a phagemid vector (pKP600, ~4 µg) at a molar ratio of 3:1 (insert: vector), generating in-frame fusions with the gene III coding sequence at the C-terminus and a FLAG epitope tag at the N-terminus. The ligated DNA was purified using phenol:chloroform:isoamyl alcohol mixture (Sigma-Aldrich Corp, St. Louis, MO), and concentrated by ethanol precipitation. TG1 cells were transformed with the recombinant DNA by performing 10 transformations, and the cells were recovered for 40 minutes, with shaking at 250 rpm at 37°C, of which various dilutions (10 µL and 100 µL of 10^{-2}, 10^{-3} and 10^{-4}) were plated on 10 cm LB/CB agar plates. The remaining cells were plated on six 15 cm LB/CB agar plates and all the plates were incubated overnight at 30°C. The next day, colonies were counted on the titration plates and the library diversity was determined to be 6.5×10^{7} clones. The bacterial lawn on the 15 cm plates was scraped with a total of 6 mL of freezing media (LB/CB/16% glycerol) and the library cells were stored at -80°C in single use aliquots. In order to determine the error rate of

mutagenic PCR, 60 clones were sequenced, of which 8% were wild type, and 16%, 20%, 23%, 15%, 10%, 2%, 5%, 2% clones had one through eight mutations, respectively. LB/CB media (20 mL) was inoculated with ~10^9 library cells and grown to mid-log (OD_{600nm} = 0.5) followed by infection with M13KO7 helper phage (MOI=20) for 1 h at 37°C at low speed (150 rpm). Infected cells were collected by centrifugation, resuspended in 100 mL of fresh LB/CB/Kan media and incubated overnight at 30ºC with shaking at 250 rpm. The following day, phage particles were precipitated using 1/5 volume of 24% polyethylene glycol (PEG) and 3 M NaCl, and the phage pellet was resuspended in PBS (1 mL), and stored in aliquots at -80°C with 16% final glycerol concentration. The schematic representation of library construction is shown in Fig. 2. All the restriction enzymes and the M13KO7 helper phage were purchased from New England BioLabs (Ipswich, MA). The final concentration of carbenicillin (CB) and kanamycin (kan) antibiotics is 50 µg/mL.

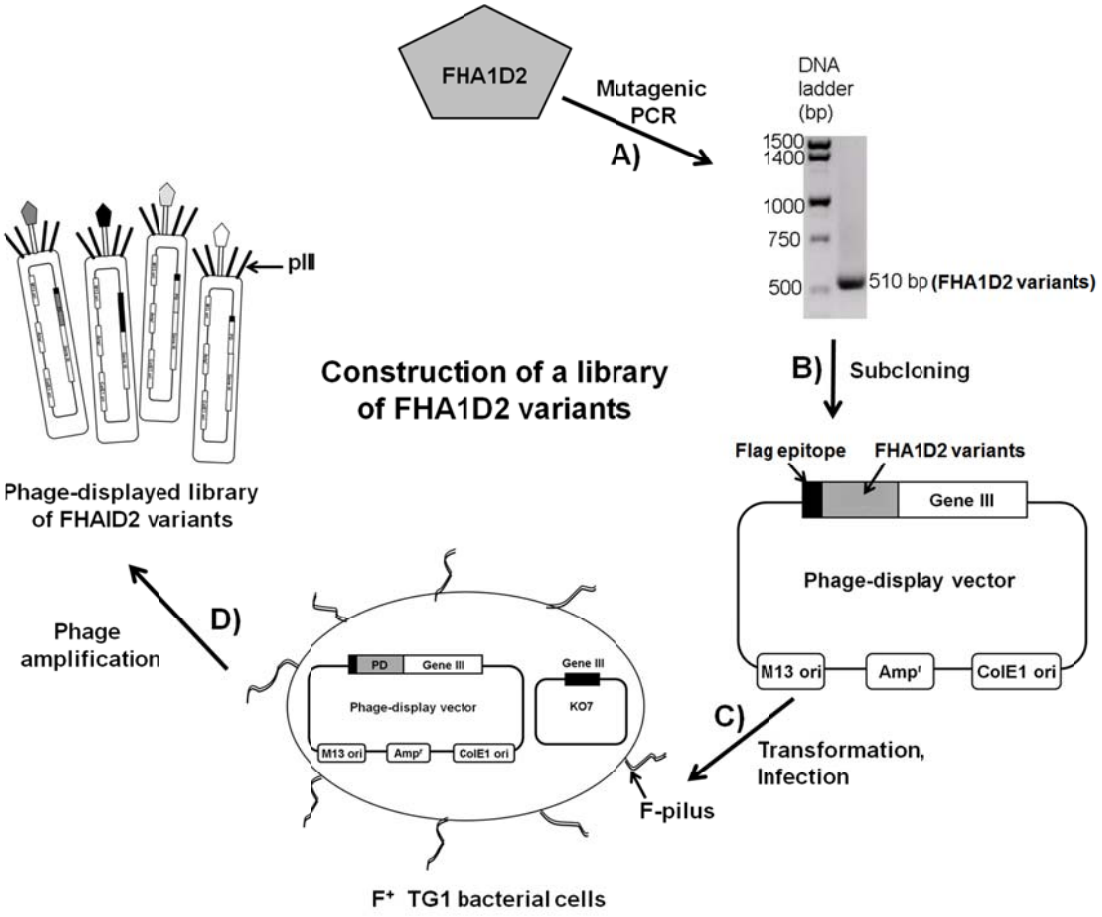


Figure 2. Constructing a phage-displayed library of the FHA1D2 variants. A) The coding sequence of the FHA1D2 domain was amplified by mutagenic PCR. B) The insert DNA pool was subcloned in a phage-display vector in-frame with the gene III coding sequence. C) The library DNA was transformed into TG1 strain of *E. coli*, and super-infected with the helper phage, M13KO7, for propagation of secreted viral particles. D) Phage particles were amplified overnight to produce a library of FHA1D2 ariants displayed as N-terminal fusions with the minor capsid protein III (pIII).

3.3.2 Effect of high temperature on eliminating thermally less stable clones

In order to confirm that the FHA1D2 library consists of thermal stable variants and that elevated temperatures are effective in eliminating thermally unstable phage clones, the phage library was heated at 25ºC, 40°C, 50°C or 60°C for 3 h, followed by cooling to room temperature. The heat-treated and untreated libraries (150 µL) were incubated with the cognate phosphothreonine (pT)-containing peptide for FHA1D2 (Rad9-pT; biotin conjugated, 2 µg) immobilized on Nunc^{TM} MaxiSorp polystyrene tube (Thermo Fisher Scientific, Pittsburgh, PA) via NeutrAvidin (2 µg; Thermo Scientific) and blocked with 2% skim milk in 1x PBS. After washing away the unbound/non-specific phage particles, the binding phage were eluted with 150 µL of 100 mM glycine-HCl (pH 2), and neutralized using 9 µL of 2 M Tris-base (pH 10). TG1 *E. coli* cells (800 µL) at mid-log (OD_{600nm} = 0.5) were infected with the entire volume of the eluted phage for 40 minutes at 37ºC. Infected cells (10 µL) from each treatment were plated on 10 cm LB/CB agar plates and incubated overnight at 30ºC.

3.3.3 Determining the appropriate temperature for affinity selection

TG1 bacterial cells (5 mL) harboring the pKP600-FHA1D2 phagemid vector were grown to mid-log (OD_{600nm} = 0.5) and infected with M13KO7 helper phage (MOI=20) for 1 h at 37°C at low speed (150 rpm). Infected cells were collected by centrifugation, resuspended in 30 mL of fresh LB/CB/Kan media and incubated overnight at 30ºC with shaking at 250 rpm. The following day, phage particles displaying the FHA1D2 domain were precipitated using PEG/NaCl mixture, and the phage pellet was resuspended in

PBS (1 mL). To determine the phage concentration, mid-log TG1 cells (200 µL) were infected with 10 µL of various phage dilutions (10^{-8}, 10^{-9} and 10^{-10}) for 40 min at 37°C, plated on 10 cm LB/CB plates and incubated overnight at 37°C. The next day colonies were counted and the phage titer was determined to be $3.7X10^{12}$ cfu/mL. Phage particles displaying the FHA1D2 domain were either incubated at room temperature or heated at various elevated temperatures (30ºC, 40ºC, 50ºC, 60ºC, 70ºC, and 95ºC) for 3 h and allowed to cool. The biotinylated pT-containing peptide (Rad9-pT; 500 ng/well) was immobilized on Nunc™ microtiter plate (Fischer Scientific) wells via NeutrAvidin (Thermo Scientific; 1 µg/well), and blocked with 2% skim milk in 1x PBS. The treated phage particles were incubated with the target pT-peptide ($5x10^{10}$ phage/well) and the binding phage were detected using anti-M13 antibody conjugated to Horseradish Peroxidase (HRP) (GE Healthcare, Piscataway, NJ; diluted 1:5000 with PBST).

3.3.4 Affinity selection for isolating thermally stable FHA1D2 variants

Three rounds of affinity selection were performed applying high temperature as the selective pressure to eliminate thermally unstable variants of the protein domain. The affinity selection process is depicted in Fig. 3. During the first round of selection, the phage library was pre-heated before the selection process at 40°C for 3 h and cooled to room temperature to denature and inactivate variants that are unstable at or below this temperature. The treated phage library (150 µL) was incubated with the Rad9-pT peptide (2 µg) immobilized on Nunc™ MaxiSorp polystyrene tube (Fisher Scientific) via NeutrAvidin (2 µg; Thermo Scientific) and blocked with 2% skim milk in 1x PBS. After

washing away the unbound/non-specific phage particles (Fig. 4C), the binding phage were eluted with 150 µL of 100 mM glycine-HCl (pH 2), and neutralized using 9 µL of 2 M Tris-base (pH 10). TG1 *E. coli* cells (800 µL) at mid-log (OD_{600nm} = 0.5) were infected with the entire volume of the eluted phage for 40 minutes at 37ºC and the cells were plated on a 15 cm LB/CB plate and incubated overnight at 30ºC. The cells harboring the genomes of phage particles recovered from round 1, were scrapped off the plate using freezing media (2 mL), ~10^9 cells were grown to mid-log (OD_{600nm}=0.5-0.6) and infected with M13KO7 helper phage (MOI=20; New England BioLabs) for phage propagation. Phage particles were amplified overnight at 30ºC and the enriched phage library was concentrated 20 fold by precipitation with PEG/NaCl, and carried forward for the second round of selection. The next two rounds of affinity selection were more stringent and the phage library was heated at 50°C for 3 h, cooled, and then incubated with the Rad9-pT peptide. The experimental procedure for rounds 2 and 3 is the same as round 1, except that the volume of eluted phage used to infect TG1 cells is decreased with each subsequent round (i.e., half and one quarter of the total eluted phage was used for infection in rounds 2 and 3, respectively). After the third round of affinity selection, 10 µL and 100 µL of 10^{-2} and 10^{-4} dilutions were plated on LB/CB agar plates (10 cm). The following day, 96 individual clones were propagated as phage, followed by phage-ELISA to identify variants that were more thermal stable compared to the FHA1D2 variant.

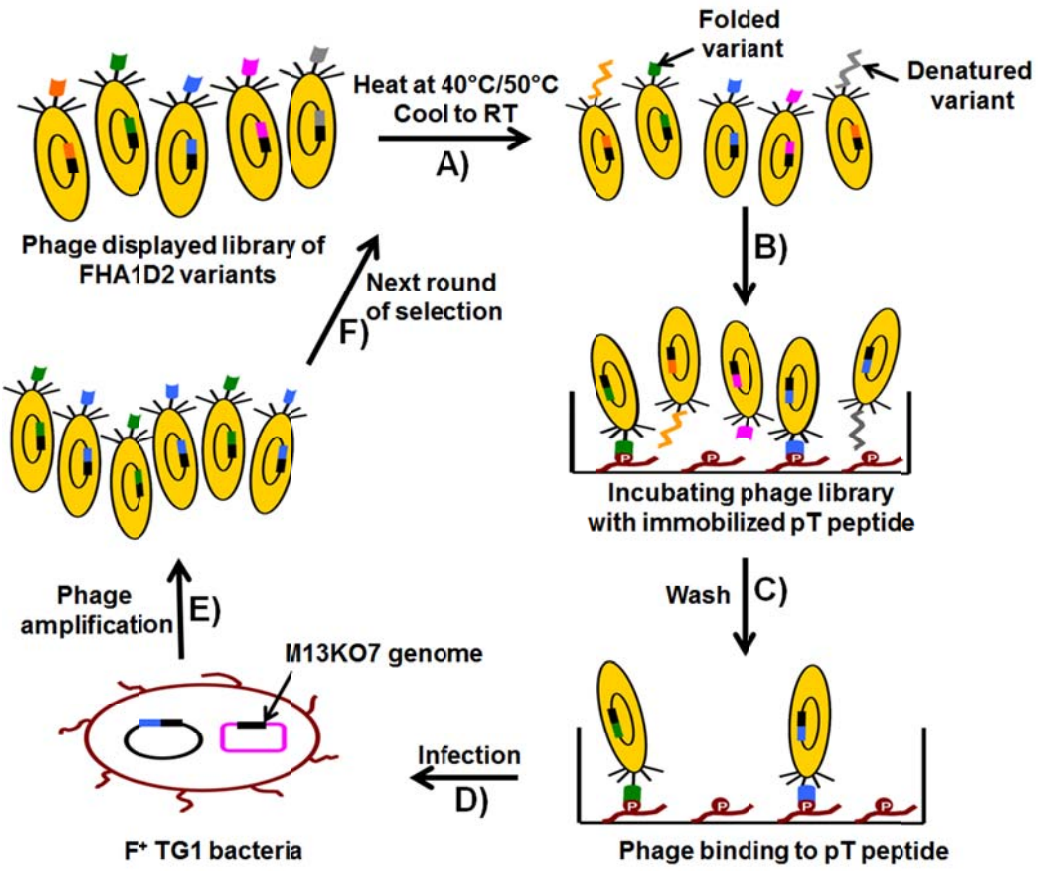


Figure 3. Affinity selection process for isolating thermal stable variants. A) The phage library was heated at 40°C for 3 h and cooled to room temperature (RT). B) The heat-treated library was incubated with the Rad9-pT peptide immobilized on the surface of a polystyrene tube. C) The unbound phage were washed using a detergent solution. D) The target bound phage were eluted with acid and used to infect TG1 cells, which are grown to mid-log and then infected with M13KO7 helper phage. E) Infected cells were grown overnight to allow for phage propagation which will be used for the next round of selection. During the subsequence rounds of selected the phage library was heated at 50°C for 3 h.

3.3.5 Testing output clones by phage-ELISA

Phage particles produced from 96 clones isolated after round 3, were heated at 50ºC for 3 h and cooled to room temperature. The biotinylated target peptide (Rad9-pT) and the non-phosphorylated form of the same peptide (Rad9-T) were immobilized on two Nunc MaxiSorp flat-bottom 96 well plates (Thermo Fisher Scientific; 500 ng/well) via NeutrAvidin™ Biotin binding protein (Thermo Fisher Scientific; 1 µg/well) and blocked with 2% skim milk in PBS. The phage supernatants (diluted 1:1 with PBST) were then incubated with peptides coated on both the plates. Binding phage clones were detected using anti-M13 antibody conjugated to HRP (GE Healthcare). The green, colored product of the enzymatic reaction in the microtiter plate wells was measured at 405 nm using the FLUOstar OPTIMA plate reader (BMG Labtech, Cary, NC). All the ELISA steps were performed at room temperature with 1 h incubations. The positive clones were sequenced (Research Resource Center, UIC) and the clones with unique DNA sequences (A1, A12, G2, G7 and H7) were retested in phage-ELISA to confirm binding to the target pT-containing peptide after heating at 50°C.

3.3.6 Expression and purification of FHA1D2 variants by immobilized metal affinity chromatography

The Overnight Express™ Autoinduction System 1 (Novagen, Madison, WI; 200 mL) was inoculated with 6 mL (3% of the final culture volume) of BL21 DE3 cells (Stratagene, La Jolla, CA) harboring the expression vector (pET29b-FHA1 variant) and grown for 24 h at 30°C/300 rpm. The cells were lysed on ice using 1x BugBuster

supplemented with Benzonase® Nuclease (0.125 units/µL; Novagen) and complete EDTA-free protease inhibitor cocktail (1X concentration; Roche Applied Science, Indianapolis, IN). BugBuster® 10x Protein Extraction Reagent (Novagen) was diluted to 1x BugBuster used cell lysis buffer (50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM Imidazole, 20 mM Beta-mercapto ethanol). Cell lysis was carried out for 30 min at room temperature on a shaking platform, followed by centrifugation at 8,000 rpm for 10 min, at 4°C, to pellet the cell debris. Ni-NTA agarose (Qiagen, Valencia, CA) was equilibrated with 5 volumes of lysis buffer, and then incubated with the clarified cell lysate for 1 h at room temperature on a revolver. The unbound protein was removed by washing with a total of 80 mL of wash buffer 1 (50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM Imidazole, 0.5% Tween 20) and 50 mL of wash buffer 2 (50 mM NaH_2PO_4, 0.3 M NaCl, 20 mM Imidazole, 0.5% Tween 20). The FHA1 variants were eluted with 2 mL of elution buffer (50 mM NaH_2PO_4, 0.3 M NaCl, 250 mM Imidazole) for 1 h on a shaking platform. The purified proteins were desalted using Zeba™ Spin Desalting Columns (5 mL; Thermo Scientific), glycerol was added (16% final) and stored at -80°C.

3.3.7 Fluorescence-based thermal shift (FTS) assay

The FTS assay was performed on a MxPro-Mx3005P instrument (Stratagene), following a published protocol (*22*). It is a real-time assay using the SYBR® Green (with dissociation curve) to measure the increase in fluorescence, as the protein unfolds upon heating and binds the dye, making it fluorescent. The default FRROX filter set was used with an excitation wavelength of 492 nm and emission wavelength of 610 nm. FHA1

variants (2 μM and 8 μM, 2X of final concentration) were mixed with SYPRO® Orange protein gel stain (Invitrogen; Carlsbad, CA; 5000X concentration in DMSO) to give a final dye concentration of 10X. In the protein only wells, the dye+protein mixture (10 μL) was diluted 1:1 with PBS. In the other wells, 10 μL of either the pT peptide or its non-phosphorylated form (2X the final concentration) was added to 10 μL of dye+protein mixture. The final reaction volume was 20 μL, with final concentrations of 1 μM or 4 μM FHA1 domains, 50 μM or 250 μM of pT and non-phosphorylated peptides, and 5X final concentration of the SYPRO Orange dye. The assay was performed in duplicates in 96 well white PCR plates (Bio-Rad) covered with optically clear Microseal® 'B' Film (Bio-Rad Laboratories, Hercules, CA) and heated from 25°C to 95°C. The melting curve was obtained with fluorescence (R) values plotted on Y-axis and increasing temperature (°C) on X-axis. The mid-point of the curve is considered the melting temperature inflection point.

3.3.8 Alanine-scanning of the most thermal stable variant (FHA1G2)

Following the Kunkel mutagenesis protocol (Fig. 4) (*23*), 24 amino acids from three loops (11 from β4-β5 loop, 5 from β6-β7 and 8 from β10-β11) in the FHA1G2 variant were mutated to alanine (gcg codon) one at a time. The 24 oligonucleotides used for alanine-scanning are listed in Table I. The resulting covalently closed DNA was transformed into TG1 electrocompetent cells (1 ng of DNA into 40 μL of TG1 cells, pulsed at 1800 volts in 0.1 cm cuvettes) and recovered by adding 960 μL of SOC media for 1 h at 37°C/250rpm. The cells (20 μL) were then plated on 10 cm LB/carbenicillin

agar plates and incubated overnight at 30°C. DNA was purified from a total of 72 clones (3 from each alanine-scanning) using the Wizard® Plus SV Minipreps DNA Purification System (Promega Corporation, Madison, WI) and sequenced at the Research Resource Center at UIC to identify the mutated clones.

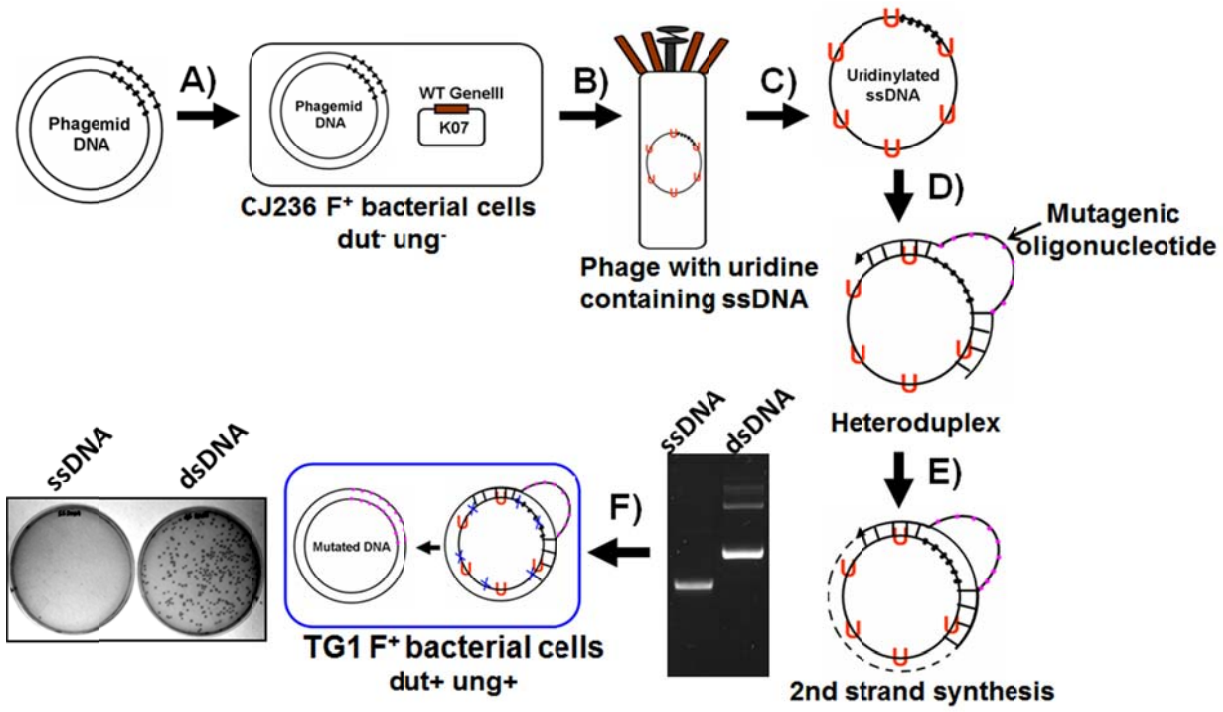


Figure 4. Kunkel mutagenesis. A) The phagemid DNA is transformed into F-pilus containing (F^+) CJ236 bacterial cells that lack the enzymes dUTPase and uracil deglycosidase (dut^- ung^-). B) Then the cells are grown to mid-log, infected with the M13KO7 helper phage, and propagated overnight, in the presence of uridine, to generate phage particles with UTP (U) incorporated into their single-stranded phagemid DNA (ssDNA). C) The uracil containing ssDNA is isolated. D) The oligonucleotide harboring the desired mutations is annealed to the ssDNA template to form a heteroduplex. E) Second strand synthesis is initiated to form the double-stranded DNA (dsDNA) by the action of T7 DNA polymerase and T4 DNA ligase. F) The ds DNA is electroporated into F^+ TG1 bacterial cells (dut^+ ung^+), which destroy the uridine containing wild-type DNA.

Table I. List of primers used in alanine-scanning.

Primer	Sequence (5′-3′)
KM-G2-L78A	GCGGCTAATGTTACCCGCGTGATAATCGCAGGC
KM-G2-G79A	GCGGCTAATGTTCGCTAAGTGATAATCGC
KM-G2-N80A	AGATAAGCGGCTAATCGCACCTAAGTGATAATC
KM-G2-I81A	ATTAGATAAGCGGCTCGCGTTACCTAAGTGATA
KM-G2-S82A	GTGTTTATTAGATAAGCGCGCAATGTTACCTAAGTGAT
KM-G2-R83A	GAAAGTGTTTATTAGATAACGCGCTAATGTTACCTAAGTGA
KM-G2-L84A	GTGTTTATTAGACGCGCGGCTAATGTTACCT
KM-G2-S85A	GATTTGAAAGTGTTTATTCGCTAAGCGGCTAATGTTAC
KM-G2-N86A	GAGGATTTGAAAGTGTTTCGCAGATAAGCGGCTAATGT
KM-G2-K87A	CAGGAGGATTTGAAAGTGCGCATTAGATAAGCGGCTAAT
KM-G2-H88A	GCCCAGGAGGATTTGAAACGCTTTATTAGATAAGCGGCT
KM-G2-G133A	ATCGCTTTCTACACCTACCGCTACCGTAATTTCGTCGC
KM-G2-V134A	ATATCGCTTTCTACACCCGCGCCTACCGTAATTTCGTC
KM-G2-G135A	TATCGCTTTCTACCGCTACGCCTACCGTA
KM-G2-V136A	ATATCGCTTTCCGCACCTACGCCTACCG
KM-G2-E137A	GCTTAAAATATCGCTCGCTACACCTACGCCTAC
KM-G2-S138A	GCTTAAAATATCCGCTTCTACACCTACGCC
KM-G2-D139A	GACTAAGCTTAAAATCGCGCTTTCTACACCTAC
KM-G2-I140A	ATGACTAAGCTTAACGCATCGCTTTCTACACCTA
KM-G2-I104A	GCCATGTACCATTTGTTGACGCGTCGTTGAGTAATAAGT
KM-G2-S105A	GAGCCATGTACCATTTGTCGCGATGTCGTTGAGTAATA
KM-G2-T106A	TGAGCCATGTACCATTCGCTGAGATGTCGTTGAGTA
KM-G2-N107A	CGTTGAGCCATGTACCCGCTGTTGAGATGTCGTTGAG
KM-G2-G108A	TGACCGTTGAGCCATGTCGCATTTGTTGAGATGTCGT

3.4 Results and discussion

3.4.1 Thermal stability profile of the FHA1D2 variant

In order to isolate variants that are more thermal stable than the FHA1D2 variant, it is important to determine the thermal stability profile of the domain. Phage particles displaying the FHA1D2 variant were either incubated at room temperature or heated at various, elevated temperatures prior to incubation with the pT-containing peptide (Rad9-pT). Phage ELISA showed that the FHA1D2 variant was functional when heated at 30°C and 37°C, but there was a severe drop in binding (by ~60%) upon heating at 50°C (Fig. 5). Thus, the hypothesis is that performing affinity selection at this temperature will eliminate variants that are less stable than 50°C. This effort is based on prior research demonstrating that the thermal stability, expression and affinity of a single-chain variable Fragment (scFv) could be improved by employing high temperature and denaturants during affinity selection (*1*). We have extended this approach to a phosphopeptide-binding domain using mutagenic PCR to construct the library and high temperature as selective pressure during biopanning.

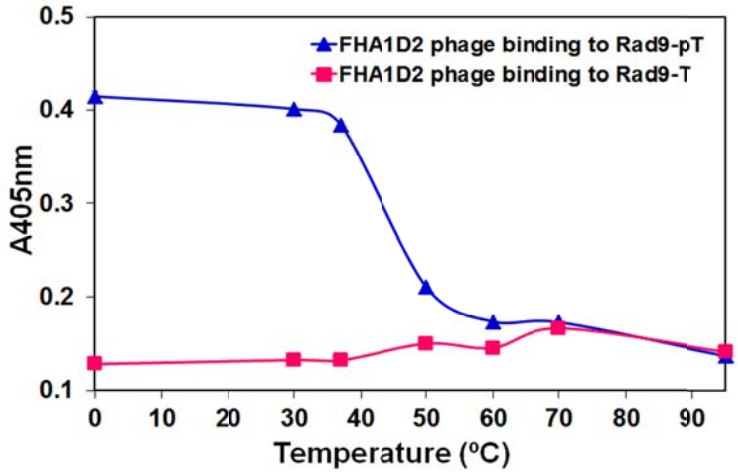


Figure 5. Thermal stability of bacteriophage particles displaying the FHA1D2 variant. Phage particles displaying the FHA1D2 variant were heated at various temperatures (30ºC, 37ºC, 50ºC, 60ºC, 70ºC, 80ºC and 95ºC) for 3 h and allowed to cool to room temperature. The treated phage particles were then incubated with the cognate pT peptide (Rad-pT; triangles) and the negative control peptide (Rad9-T; squares) immobilized on plastic via Neutravidin. Phage binding was detected using an anti-M13 antibody conjugated to HRP.

3.4.2 Using high temperature as a selective pressure during affinity selection

The higher the temperature applied before the selection process, the greater is the reduction in the number of colonies (Fig. 6A), confirming that high temperature is an effective negative, selective pressure on the displayed protein domain. Incubation at room temperature (RT) and at 40°C yielded almost the same number of binding clones after the first round of selection; therefore, we took forward the selection done at 40°C to perform the second and third round of selection. It is a good idea not to perform a very stringent selection in the first round of selection to facilitate the isolation of maximum

number of binding clones. The stringency can be increased in the subsequent rounds of selection to isolate binding clones with the desired properties. The second and third rounds of selection were carried out by heating the library at 50°C prior to the selection process to eliminate variants less stable than 50°C and to isolate and propagate the more thermal stable pool of variants. A quantitative assessment of the number of colonies isolated after each heat treatment can easily be determined by plating (Fig. 6B). This also shows that the mutagenized FHA1D2 library contains variants that have accumulated favorable mutations (by mutagenic PCR) which give them more thermal stability than the original FHA1D2 variant. This preliminary experiment provides the proof that this library is a good resource of thermal stable variants which can be selectively isolated from the less stable ones by performing affinity selection using high temperature as the selective pressure.

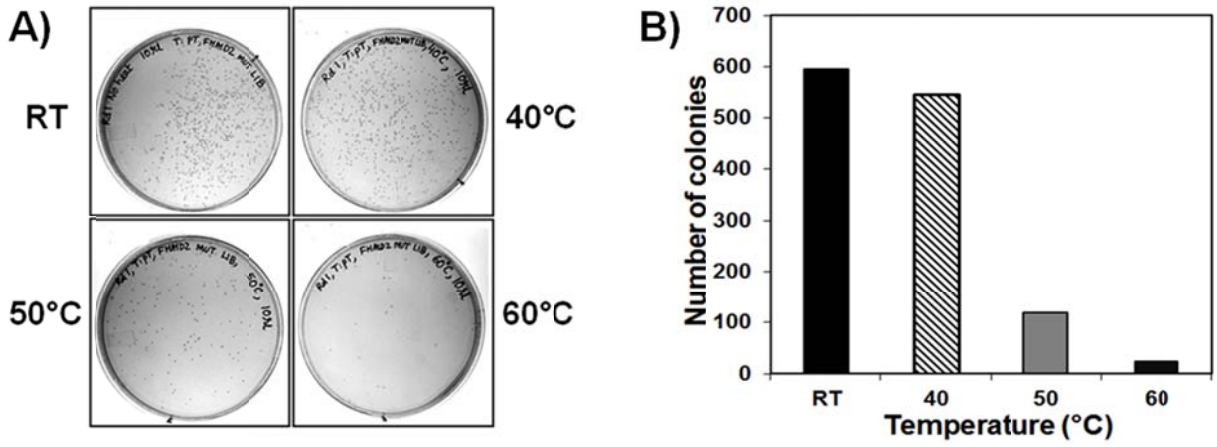


Figure 6. High temperature works as a selective pressure. A) The phage-displayed library of FHA1D2 variants was untreated (incubated at room temperature (RT)) or heated at 40°C, 50°C or 60°C prior to affinity selection. The four libraries were incubated with the Rad9-pT peptide and the phage clones bound to the target peptide after the first round of selection were plated after infecting bacterial cells. B) A quantitative assessment of the number of binding clones isolated after the first round of selection with or without heat treatment of the phage library.

3.4.3 Affinity selection with the mutagenized FHA1D2 library

Since the FHA1D2 variant was ~5°C thermally less stable compared to the wild-type FHA1 domain, to improve its thermal stability, a library of FHA1D2 variants generated by error-prone PCR was displayed on the surface of the phage, heated at 40°C/50°C prior to incubation with the pT peptide ligand to favor the isolation of variants that remained folded and functional above 50°C. After 3 rounds of affinity selection, 96

individual colonies were propagated as phage, heated at 50°C for 3 h and tested for binding to the Rad9-pT peptide in a phage-ELISA. Microtiter plate wells that generated signal on the target plate (coated with Rad9-pT), but not on the background plate (coated with Rad9-T peptide), suggest that those phage clones represent potential thermal stable binders for the target (Fig. 7A). The ratio of the specific signal, over background, ranged between 3 to 20 times for various positive clones. A similar approach was employed to improve the thermal stability, expression and affinity of a recombinant antibody fragment (*1*). Five positive clones (A1, A12, H7, G2, and G7), with unique sequences, were reconfirmed by phage-ELISA to bind specifically to the Rad9-pT peptide and not to the non-phosphorylated form of the same peptide, before and after heat treatment (Fig. 7B).The mutations observed in each one of the thermally stable mutants are listed in Table II. Mutations were observed in 4 out of 11 β-strands, and in 4 loops of which two loops (β3-β4 and β4-β5) are involved in interaction with the pT peptide ligand and the other two loops (β1-β2 and β8-β9) do not interact with the pT peptide ligand. Mutations were also observed at the N-terminus of the FHA domain, before the β1-strand, indicating that this region may be critical for structural stability and the mutation may be causing better packaging and folding of the domain, henceforth resulting in thermal stability.

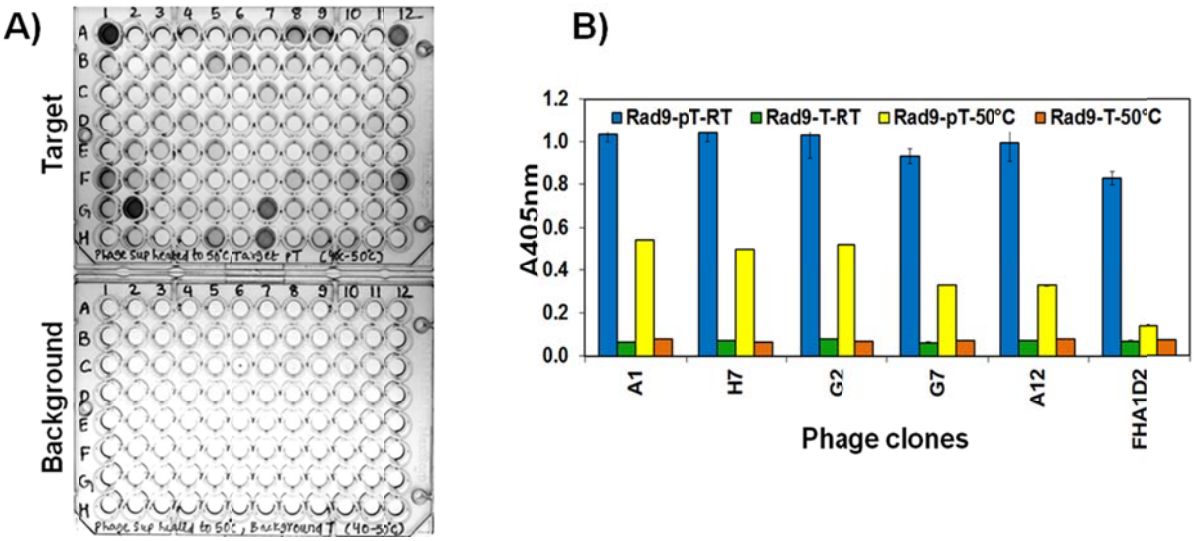


Figure 7. Affinity selection for isolating thermal stable variants. A) After three rounds of affinity selection, 96 phage clones were tested by phage-ELISA, after heating at 50°C, for binding to the cognate phosphorylated peptide (Rad9-pT) and to the non-phosphorylated form of the same peptide (Rad9-T). B) Five thermal stable phage clones with different sequences were reconfirmed by phage-ELISA to bind to the Rad9-pT peptide and not to the non-phosphorylated peptide. The starting variant, FHA1D2, was used to compare its binding signal with the thermal stable variants.

Table II. Mutations observed in the thermal stable FHA1D2 variants.

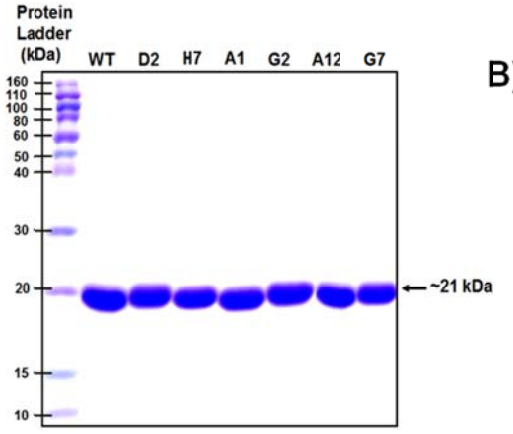
FHA1 variants[a]	No. of mutations	Mutations	Location[b]
A1	2	N121Y L141I	β8-β9 loop β11 strand
G2	3	T15A L48F N121Y	N-terminus β2 strand β8-β9 loop
H7	5	Q16R T39P K59E C74S G94D	N-terminus β1-β2 loop β2-β3 loop β3-β4 loop β5-β6 loop
G7	5	S11T A14V S82R N121Y L141I	N-terminus N-terminus β4-β5 loop β8-β9 loop β11 strand
A12	4	C74S G94D E129V N158I	β3-β4 loop β5-β6 loop β10 strand C-terminus

[a] FHA1 variants that are thermally more stable than the starting D2 variant.
[b] β = beta-strand. Loop is the region between two β-strands. The N-terminal mutation is present before the β1-strand and the C-terminal mutation is after the β11-strand.

3.4.4 SDS-PAGE analysis and yields of purified FHA1 variants

Coomassie staining of the purified FHA1 variants (4 µg) run on a 15% SDS-PAGE gel revealed protein bands of the correct size (~21 kDa) with high yields per liter listed in Fig. 8. It was observed that the most thermal stable protein (FHA1G2) had the highest yield when expressed and purified from *E. coli*. This protein was also more stable when stored at 4°C compared to the FHA1D2 variant.

A)



B)

FHA1 variants	Yield (mg/L)
WT	36.8
D2	34.8
H7	36.8
A12	36
G7	44
A1	55.7
G2	63.4

Figure 8. SDS-PAGE analysis of purified FHA1 variants. A) The FHA1 variants were purified by IMAC via their C-terminal six-histidine tag. Purified FHA variants (4 µg) were resolved on a 15% SDS-PAGE gel, followed by staining with Coomassie Brilliant Blue. The molecular weights corresponding to protein standards are shown in kilodaltons (kDa). The FHA1 domains are ~21 kDa in size. B) The yields per liter of culture are between 35 and 63 mg/L.

3.4.5 Fluorescence-based thermal shift assay as a measure of thermal stability

The purified FHA1 variants were mixed with SYPRO orange dye with or without the peptide and heated from 25°C to 95°C in a real-time PCR experiment known as the Fluorescence-based thermal shift (FTS) assay (*22*). In this assay, initially, the fluorescence of SYPRO orange dye is quenched in the aqueous environment of a folded protein, but as the protein starts to unfold upon heating, the dye interacts with the hydrophobic core of the protein and its fluorescence increases. Therefore, an increase in fluorescence of the dye is directly proportional to protein unfolding, until a temperature (referred here as infliction point) is reached at which the dye fluorescence decreases due to aggregation and precipitation of the protein.

The melting curves of the thermal stable FHA1D2 variants, obtained by fluorescence-based thermal shift assay, at 1 μM concentration are shown, alone (green curve) and in the presence of 50 μm (blue curve) and 250 μm (orange curve) of the pT peptide (Fig. 9). No shift in the infliction point was observed in the presence of the non-phosphorylated form of the same peptide (data not shown). The infliction points of the thermally stable variants (along with wild-type FHA1 domain and FHA1D2 variant) at 1 μM and 4 μM concentration, along with the shift in infliction point in the presence of two different concentrations (50 μM and 250 μM) of the pT peptide, are listed in Table III. Using this method, the thermal stability of the wild-type FHA1 domain was increased by ~8°C. Thus, using high temperature as a selective pressure, variants with favorable mutations that improved the thermal stability were isolated from a phage-displayed mutagenized library using the thermal less stable variant as a starting scaffold.

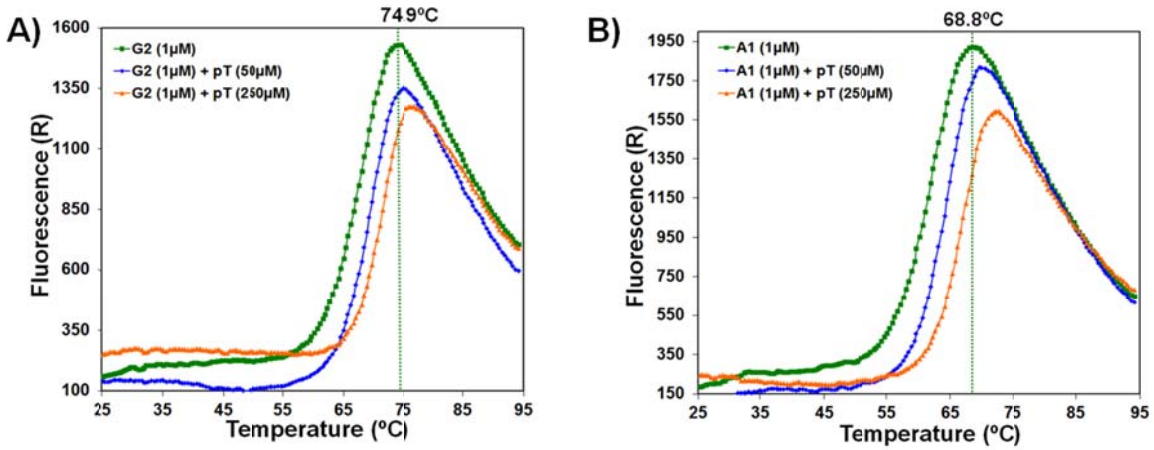


Figure 9. Determining the inflection points of FHA1 variants by FTS Assay. A-B) Shown here are the melting temperature curves of two FHA1 variants (G2 and A1) that were determined to be more thermal stable than the wild-type FHA1 domain. The variants were mixed with SYPRO orange dye, with or without the Rad9-pT peptide ligand, and heated from 25ºC to 95ºC on an Mx3000P Real-time thermal cycler instrument. The melting curve for the variants (without any pT-peptide) is shown as a green curve with the infliction points noted as the midpoint of the curve. The melting curves in the presence of 50 μM and 250 μM of the pT-peptide are shown as blue and orange curves respectively.

Table III. The infliction points for various FHA1 variants along with the shift in infliction point in the presence of the cognate Rad9-pT peptide.

FHA1 variants	Protein concentration (μM)	Infliction point (°C)[a]	Shift in infliction point (°C) in the presence of pT peptide[b]	
			50 μM	250 μM
WT	1	66.7	1.7	3.8
	4	66.4	1.95	3.55
D2	1	61.6	2.6	4.6
	4	61.8	1.85	4.45
A12	1	66	1.25	2.0
	4	64	1.0	2.4
H7	1	65.2	1.25	3.0
	4	65.1	1.35	3.65
G7	1	67	4.0	6.5
	4	65.7	3.5	6.3
A1	1	68.5	1.2	3.95
	4	68.8	1.2	4.5
G2	1	73.75	1.2	2.75
	4	74.9	0.5	2.5

[a] The temperature corresponding to the mid-point of the melting curve/thermal denaturation curve is the infliction point.
[b] An increase in the infliction point temperature is observed when the FHA1 variants are incubated with the pT peptide and heated from 25°C to 95°C. This increase corresponds to the shift in infliction point.

3.4.6 Alanine-scanning to identify residues in the FHA1G2 variant important for interaction with the pT peptide ligand

From previous structural studies on the FHA1-pT complex from various species, it has been shown that residues from four loops (i.e., β3-β4, β4-β5, β6-β7 and β10-β11) contact the pT peptide (*24-26*). Interestingly, the β4-β5 loop varies in sequence,

structure and length among FHA domains (*27*), which may account for different specificities among FHA domains: for instance, the β4-β5 loop in FHA1 domain comprises of 11 amino acid residues, and on the other hand, the β4-β5 loop in ChK2 FHA domain is longer and contains 19 residues with a helical insertion in the loop (*28*). The length and structure of this loop determines its positioning either close to or away from the pT +3 residue, which is an important determinant of binding specificity, therefore, resulting in specific binding to phosphopeptides with either charged (for FHA1 domain) or hydrophobic residues (for ChK2 FHA domain) in the pT +3 position. Mutating residues in the β10-β11 loop has been shown to alter the binding specificity of the FHA1 domain to be more like FHA2 (*29*) and the residues in this loop may play an important role in the binding of the FHA1 domain to a pT peptide (from Mdt1 protein) containing a hydrophobic residue at the pT +3 position (*27*). Residues from β6-β7 loop are known to be responsible for conferring preference for binding to pT- and not pS-containing peptides (*26*).

To establish which residues in the G2 variant are important for interaction with the pT peptide (SLEVpTEADATFYAKK), alanine-scanning of each of the residues from the three loops (11 from β4-β5 loop, 5 from β6-β7 and 8 from β10-β11) was performed and binding of the mutants to the cognate pT-peptide was tested by phage-ELISA. Five mutants (L78A, R83A, L84A, S85A and H88A) from the β4-β5 loop, four mutants (S105A, T106A, N107A and G108A) from the β6-β7 loop and five mutants (G133A, V134A, G135A, V136A and D139A) from the β10-β11 loop had reduced (marked with arrows), or no binding to the pT peptide when mutated to alanine indicating that the

interaction of these residues with the pT-peptide is important for binding (Fig. 10).

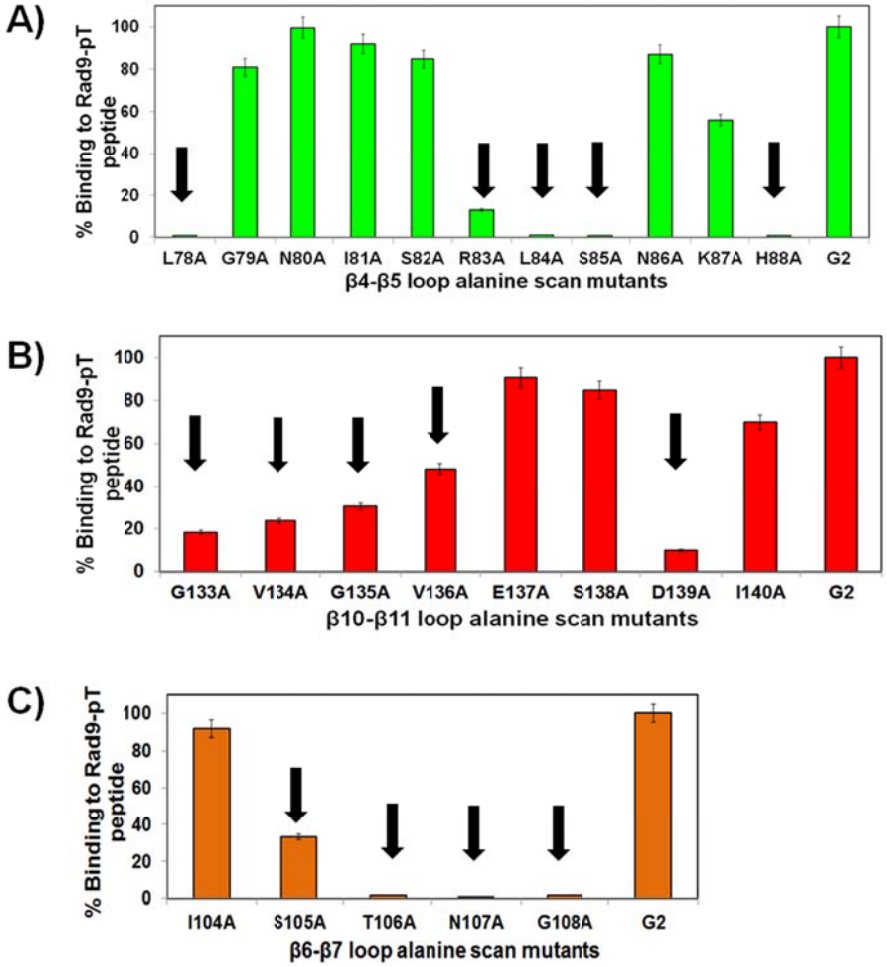
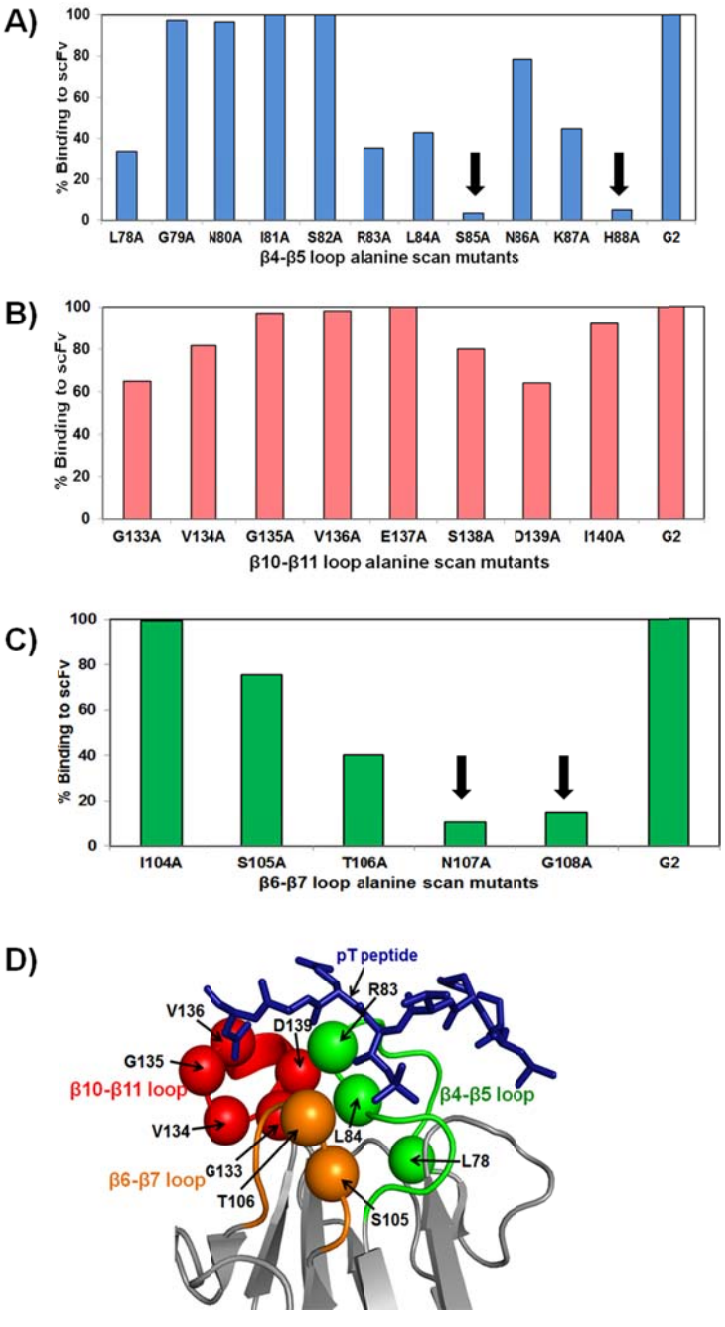


Figure 10. Alanine-scanning of residues from three loops in the FHA1G2 variant. Residues 78-88 in the β4-β5 loop, 104-108 in the β6-β7 loop, and 133-139 in the β10-β11 loop were mutated to alanine one at a time and the mutants were tested for binding to the Rad9-pT peptide by phage-ELISA. Binding was detected using an anti-M13 antibody conjugated to HRP. Phage-ELISA for A) β4-β5 loop alanine-scan mutants, B) β10-β11 loop alanine-scan mutants, and C) β6-β7 loop alanine-scan mutants. Fourteen residues that show reduced or no binding to the Rad9-pT when mutated to alanine are indicated by arrows. The G2 variant was the original variant from which the alanine-scan variants were generated.

Four out of these 14 mutations (S85A, H88A, N107A and G108A) destroyed folding or denatured the FHA1 domain, as determined with a recombinant antibody fragment that recognized folded, but not denatured, FHA1 domain. Therefore, 10 residues from the three loops, 3 from β4-β5 (green spheres), 5 from β10-β11 loop (red spheres) and 2 from β6-β7 (orange spheres) were identified to be critical for pT peptide recognition (Fig. 11), and not for folding of the FHA1 domain. These residues are good candidates for oligonucleotide-directed mutagenesis and generation of a phage-displayed library of FHA1 variants that can be screened against various pT peptides. According to our alanine-scanning experiments, the G133A and G135A mutants showed reduced binding to the pT peptide (SLEVpTEADATFAKK), indicating that these two positions in the FHA1 domain may play an important role in interaction with the pT peptide, as previously described (*29*).

Figure 11. Detecting folded alanine-scan mutants. The phage-displayed alanine-scan point mutants were tested for binding to an antibody fragment that recognizes only folded FHA1 variants. A) Out of the 11 alanine-scan mutants from the β4-β5 loop, the conformation of two of them (S85A and H88A) was disrupted. B) All the 8 alanine-scan mutants from the β10-β11 loop remained folded. C) Of the five alanine-scan mutants from the β6-β7 loop, the folding of two of them (N107A and G108A) was disrupted. Binding of all the variants was normalized for the display of the Flag epitope. The binding signal of the G2 mutant for the Rad9-pT peptide was taken as 100% and the percentage binding of all the twenty-four alanine-scan mutants was calculated accordingly. D) Residues from the β4-β5 loop (green spheres), β6-β7 loop (orange spheres), and β10-β11 loop (red spheres) which are important for interaction with the Rad9-pT peptide, but not for folding of the FHA1 domain, are shown.


A)
% Binding to scFv
L78A G79A N80A I81A S82A R83A L84A S85A N86A K87A H88A G2
β4-β5 loop alanine scan mutants
B)
% Binding to scFv
G133A V134A G135A V136A E137A S138A D139A I140A G2
β10-β11 loop alanine scan mutants
C)
% Binding to scFv
I104A S105A T106A N107A G108A G2
β6-β7 loop alanine scan mutants
D)
pT peptide
R83
V136
D139
G135
β10-β11 loop
β4-β5 loop
V134
L84
L78
β6-β7 loop
G133
T106
S105

A library of FHA1 variants, generated by mutating residues important for interaction with the pT peptide, will be a useful resource for isolating affinity reagents with new anti-phosphopeptide binding specificities different from that of the original FHA1 domain scaffold. Previously, it has been shown that variants of the Erbin Post synaptic density protein - Disc large tumor suppressor - Zonula occludens-1 protein (PDZ) domain, generated by mutating ten residues known to be important for interaction with peptide ligands from structural studies, had different binding specificities compared to the wild-type PDZ domain (*30*). Ernst et al. have demonstrated that it is possible to change the specificity of a protein interaction domain so that it binds a ligand for which an interaction has not been previously detected.

3.5 Conclusions

We identified mutations in the FHA1 variants that improved their thermal stability upto 8°C, compared to the wild-type FHA1 domain. With the goal of using the most thermal stable variant, FHA1G2, as a scaffold protein for generating new anti-phosphopeptide binding specificities, we identified 8 residues in the FHA1 domain, by alanine-scanning, that were important for interaction with its pT-containing peptide ligand. Next, to test our hypothesis to use a phosphopeptide-binding domain as a scaffold for generating affinity reagents to phosphorylated epitopes, we constructed two different phage-displayed libraries of FHA1 variants and screened them against various pT-containing peptides. This work is described in Chapter 4.

3.6 References

1. Jung, S., Honegger, A., and Pluckthun, A. (1999) Selection for improved protein stability by phage display, *J Mol Biol 294*, 163-180.

2. Hawkins, R. E., Russell, S. J., and Winter, G. (1992) Selection of phage antibodies by binding affinity. Mimicking affinity maturation, *J Mol Biol 226*, 889-896.

3. Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Pluckthun, A. (2001) Tailoring in vitro evolution for protein affinity or stability, *Proc Natl Acad Sci U S A 98*, 75-80.

4. Zahnd, C., Spinelli, S., Luginbuhl, B., Amstutz, P., Cambillau, C., and Pluckthun, A. (2004) Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scFv) with low picomolar affinity, *J Biol Chem 279*, 18870-18877.

5. Zahnd, C., Wyler, E., Schwenk, J. M., Steiner, D., Lawrence, M. C., McKern, N. M., Pecorari, F., Ward, C. W., Joos, T. O., and Pluckthun, A. (2007) A designed ankyrin repeat protein evolved to picomolar affinity to Her2, *J Mol Biol 369*, 1015-1028.

6. Razai, A., Garcia-Rodriguez, C., Lou, J., Geren, I. N., Forsyth, C. M., Robles, Y., Tsai, R., Smith, T. J., Smith, L. A., Siegel, R. W., Feldhaus, M., and Marks, J. D. (2005) Molecular evolution of antibody affinity for sensitive detection of botulinum neurotoxin type A, *J Mol Biol 351*, 158-169.

7. Kristensen, P., and Winter, G. (1998) Proteolytic selection for protein folding using filamentous bacteriophages, *Fold Des 3*, 321-328.

8. Olofsson, L., Ankarloo, J., Andersson, P. O., and Nicholls, I. A. (2001) Filamentous bacteriophage stability in non-aqueous media, *Chem Biol 8*, 661-671.

9. Sieber, V., Pluckthun, A., and Schmid, F. X. (1998) Selecting proteins with improved stability by a phage-based method, *Nat Biotechnol 16*, 955-960.

10. Pedersen, J. S., Otzen, D. E., and Kristensen, P. (2002) Directed evolution of barnase stability using proteolytic selection, *J Mol Biol 323*, 115-123.

11. Finucane, M. D., Tuna, M., Lees, J. H., and Woolfson, D. N. (1999) Core-directed protein design. I. An experimental method for selecting stable proteins from combinatorial libraries, *Biochemistry 38*, 11604-11612.

12. Dalby, P. A., Hoess, R. H., and DeGrado, W. F. (2000) Evolution of binding affinity in a WW domain probed by phage display, *Protein Sci 9*, 2366-2376.

13. Deng, L. W., Malik, P., and Perham, R. N. (1999) Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of Escherichia coli, *Virology 253*, 271-277.

14. Martin, A., and Schmid, F. X. (2003) The folding mechanism of a two-domain protein: folding kinetics and domain docking of the gene-3 protein of phage fd, *J Mol Biol 329*, 599-610.

15. Kramer, R. A., Cox, F., van der Horst, M., van der Oudenrijn, S., Res, P. C., Bia, J., Logtenberg, T., and de Kruif, J. (2003) A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein, *Nucleic acids research 31*, e59.

16. Belien, T., Verjans, P., Courtin, C. M., and Delcour, J. A. (2008) Phage display based identification of novel stabilizing mutations in glycosyl hydrolase family 11 B. subtilis endoxylanase XynA, *Biochemical and biophysical research communications 368*, 74-80.

17. Jespers, L., Schon, O., Famm, K., and Winter, G. (2004) Aggregation-resistant domain antibodies selected on phage by heat denaturation, *Nat Biotechnol 22*, 1161-1165.

18. Jespers, L., Schon, O., James, L. C., Veprintsev, D., and Winter, G. (2004) Crystal structure of HEL4, a soluble, refoldable human V(H) single domain with a germ-line scaffold, *J Mol Biol 337*, 893-903.

19. Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., and Baker, D. (2003) Design of a novel globular protein fold with atomic-level accuracy, *Science 302*, 1364-1368.

20. Bondugula, R., Wallqvist, A., and Lee, M. S. (2011) Can computationally designed protein sequences improve secondary structure prediction?, *Protein Eng Des Sel 24*, 455-461.

21. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR, *PCR methods and applications 3*, S136-140.

22. Giuliani, S. E., Frank, A. M., and Collart, F. R. (2008) Functional assignment of solute-binding proteins of ABC transporters using a fluorescence-based thermal shift assay, *Biochemistry 47*, 13974-13984.

23. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A. (2000) Phage display for selection of novel binding peptides, *Methods in enzymology 328*, 333-363.

24. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms, *Mol Cell 6*, 1169-1182.

25. Pennell, S., Westcott, S., Ortiz-Lombardia, M., Patel, D., Li, J., Nott, T. J., Mohammed, D., Buxton, R. S., Yaffe, M. B., Verma, C., and Smerdon, S. J. (2010) Structural and functional analysis of phosphothreonine-dependent FHA domain interactions, *Structure 18*, 1587-1595.

26. Mahajan, A., Yuan, C., Lee, H., Chen, E. S., Wu, P. Y., and Tsai, M. D. (2008) Structure and function of the phosphothreonine-specific FHA domain, *Science signaling 1*, re12.

27. Mahajan, A., Yuan, C., Pike, B. L., Heierhorst, J., Chang, C. F., and Tsai, M. D. (2005) FHA domain-ligand interactions: importance of integrating chemical and biological approaches, *J Am Chem Soc 127*, 14572-14573.

28. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2, *Mol Cell 9*, 1045-1054.

29. Yongkiettrakul, S., Byeon, I. J., and Tsai, M. D. (2004) The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues, *Biochemistry 43*, 3862-3869.

30. Ernst, A., Sazinsky, S. L., Hui, S., Currell, B., Dharsee, M., Seshagiri, S., Bader, G. D., and Sidhu, S. S. (2009) Rapid evolution of functional complexity in a domain family, *Science signaling 2*, ra50.

CHAPTER 4

DIRECTED EVOLUTION OF THE FORKHEAD-ASSOCIATED DOMAIN TO GENERATE ANTI-PHOSPHOSPECIFIC REAGENTS BY PHAGE-DISPLAY

Part of this research has been published

Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

4.1 Abstract

Previously, we engineered the Forkhead-associated (FHA) domain 1 from the *S. cerevisiae* Rad53 protein to generate functional and thermal stable variants and identified residues in the FHA1 domain that are important for interaction with the pT peptide ligand. Using a thermal stable variant (FHA1G2) as a scaffold, we constructed two phage-displayed libraries of FHA1 variants, randomizing 8 or 10 residues by Kunkel mutagenesis, and affinity selected for variants that bound selectively to seven different pT-containing peptides. These reagents are renewable and have high protein yields (~20-25 mg/L), when expressed in *Escherichia coli*. Thus, we have changed the specificity of the FHA1 domain and demonstrated that engineering this phosphopeptide-binding domain is an attractive avenue for generating new anti-phosphopeptide binding specificities *in vitro* by phage-display. For use in applications such as western blotting, the affinity of the reagents needs to be improved by directed evolution. Cell staining of paraformaldehyde-fixed NIH/3T3 cells with an anti-ERK1/2 affinity reagent, A5, yielded punctate staining in the nucleus. Curiously, staining occurred in cells with or without stimulation by epidermal growth factor (EGF), and persisted even after treatment of the cells with MEK1/2 inhibitor. Thus, the identity of the protein the A5 reagent is binding to in cells is unknown and further experiments will be necessary to identify its cellular target.

4.2 Introduction

A cascade of signaling events, involving many protein-protein interactions, are initiated within the cell in response to external or internal stimuli. The signal is translocated to downstream effectors by the reversible action of protein kinases, phosphatases and phosphopeptide-binding domains. The human genome encodes for ~500 protein kinases and a third of that number for protein phosphatases (*1*), and defective expression of either group of proteins can cause human disease (*2*). Radioisotopic labeling studies have shown that a third of the total proteins in the cell are phosphorylated at any given time (*3*). Phosphorylation of serine/threonine residues on proteins can lead not only to conformational changes in proteins but also create binding sites for phosphopeptide-binding domains, which play a critical role in the formation of multiprotein signaling complexes for relaying the signal to downstream signaling proteins (*4*). Similar to the recognition of phosphotyrosine (pY) residues by Src Homology 2 (SH2) domains and phosphotyrosine binding (PTB) domains (*5, 6*), phosphorylation of serine/threonine residues creates binding sites for various proteins, such as the 14-3-3 proteins (*7, 8*), the tryptophan-tryptophan (WW) domain of Pin1 (*9*), the FHA domain found in prokaryotic and eukaryotic signaling proteins (*10-12*) , and WD40 repeats of F-box proteins (*13*). These phosphoprotein-binding domains play a critical role in the formation of signaling complexes, translocating proteins to their proper subcellular location and

modifying the activity of catalytic proteins. Protein interaction domains regulate various activities in the cell, such as gene expression, cytoskeletal rearrangements, cell cycle progression, DNA repair and apoptosis (*14, 15*).

With the development of sophisticated phosphopeptide enrichment and mass spectrometry techniques (*16, 17*), thousands of phosphorylation sites have been identified in the phosphoproteome (*18, 19*). However, producing good quality antibodies to each phosphorylated peptide is a challenge because of the flexibility of the peptide ligands, and small number of contacts between the peptide ligand and the antibody. Immunizing animals with synthetic phosphopeptides has been previously the standard route chosen for generating antibodies that recognize phosphorylated residues in proteins (*20, 21*). Such antibodies are valuable tools for studying the phosphorylation of proteins upon cellular stimulation, for instance by epidermal growth factor (EGF) or insulin, and for unraveling biologically important signal transduction pathways. With the existence of thousands of phosphorylation sites in the human proteome, the typical method for generating anti-phosphopeptide antibodies requires immunization with a specific phosphopeptide, making this process time consuming, expensive, and laborious.

An ideal alternative would be to bypass the immunization step and use recombinant methods to generate recombinant antibodies or affinity reagents in less time. For example, antibody fragments and various engineered proteins (*22*)

have been exploited as scaffolds for generating useful affinity reagents. In order to generate affinity reagents against phosphopeptides, an ancillary route would be to use a naturally occurring phosphopeptide-binding domain as a scaffold, generate variants of this domain by Kunkel mutagenesis (*23*) and isolate variants that bind specifically to phosphopeptides of interest.

For this study we have chosen to use the naturally occurring pT peptide-binding domain, the FHA1 domain from the *S. cerevisiae* Rad53 protein. From our previous work (*24, 25*), we isolated functional and thermal stable phage-displayed variants of the FHA1 domain. Through alanine-scanning, we identified residues critical in the FHA1 domain for binding to the pT peptide ligand, and constructed two phage-displayed libraries of FHA1 variants, with library diversities of $3X10^9$ and $1X10^{10}$ members. After two rounds of biopanning, we isolated variants that specifically recognize seven different pT peptide sequences from jun-B, jun-D, c-Myc and Activating transcription factor 2 (ATF2) transcription factors, Mitogen-activated protein kinase 1 (MAPK1)/ERK2, and Mitogen-activated protein kinase 3 (MAPK3)/ERK1. By this approach, we optimized this domain for use as a scaffold from which novel anti-phosphospecific affinity reagents can be generated and evaluated in ELISA, Western blots, and cell staining.

4.3 Materials and methods

4.3.1 Construction of phage-displayed libraries of FHA1 variants by oligonucleotide-directed mutagenesis

4.3.1.1 FHA1G2 library

The pKP700 vector, with the FHA1G2 coding region subcloned into it, was used as the starting template for constructing site-directed libraries of FHA1 variants. Eight residues (L78, R83, L84 from the β4-β5 loop and G133, V134, G135, V136, D139 from the β10-β11), in the FHA1G2 coding sequence were randomized, based on my alanine-scanning results. Following the Kunkel mutagenesis protocol (*23*), two oligonucleotides- β4-β5 Lib (5'-CCCAGGAGGA TTTGAAAGTGTTTATTAGAMNNMNNMNNAATGTTACCTAAGTGATAATCGC AGGC -3', 5’ phosphorylated) and β10-β11 Lib (5'-GTTTGAATTTATCGTTAAT AAAAATGACTAAGCTTAAAATMNNMNNMNNMNNMNNMNNMNNTACCGT AATT TCGTCGC CTTGA-3'; 5’ phosphorylated) with NNK codons (N = A, G, C, or T; K = T or G and M is the complement of K; M = A or C) at these 8 positions were annealed to the single-stranded uracilated phagemid DNA (pKP700) at a molar ratio of 1:5 (single-stranded DNA:oligonucleotides), extended using T7 DNA polymerase and the covalently closed circular DNA was sealed by T4 DNA ligase (both from New England BioLabs, Ipswich, MA). A total of 15 transformations were performed with electrocompetent TG1 bacterial cells

(Lucigen Corporation, Middleton, WI). After recovery, the cells were pooled, plated on three 15 cm 2xYT/CB agar plates, and incubated overnight at 30°C. The lawn of colonies were scraped with a total of 6 mL of freezing media (2xYT/CB/16% glycerol) media and the library cells were stored at -80°C.

4.3.1.2 G2-XmaI library

A second library (G2-XmaI) was constructed, using two oligonucleotides: Phos β4-β5 Lib (/5Phos/CCCAGGAGGATTTGAAAGTGTTTATTAGAMNNMNNMNNA ATGTTACCTAAGTGATAATCGCAGGC -3') and Phos β10-β11 Lib (5Phos-GTTTGAATTTATCGTTAATAAAAATGACTAAGCTTAAAATMNNMNNMNNMNN MNNMNNMNNTACCGTAATTTCGTCGCCTTGA-3'). The two oligonucleotides randomized the FHA1D2 variant at a total of 10 positions (S82, R83, L84 from the β4-β5 loop, and G133, V134, G135, V136, E137, S138, D139 from the β10-β11) with NNK codons. S82 was randomized in this library, because it was shown to contact the pT peptide (*26*), whereas L78 was excluded because preliminary results revealed that many binding clones retained Leu at this position. E137 and S138 were randomized to facilitate efficient annealing of the oligonucleotide to the template DNA, however, according to alanine-scanning results, these residues were not important for binding to the pT peptide. The pKP700 phagemid DNA used for constructing this library has two *Xma* I restriction sites, one in the β4-β5 loop and the second one in the β10-β11 loop.

They were introduced into the FHA1 coding region by Kunkel mutagenesis, with the primers pKP700-G2-XmaI#1 (5'-CCAGGAGGATTTGAAAGTGTTTATTAGA CCCGGGGCTAATGTTACCTAAGTGATAATCG CAG -3') and pKP700-G2-XmaI#2 (5'-CGTTAATAAAAATGACTAAGCTTAAAATATCGCTTTCCCCGGG TACGCCTACCGTAATTTCGTCG-3'). The G2-XmaI library consisted of four sublibraries, each with a diversity of ~5-7x10^9 members. Each sublibrary was constructed by performing 25 transformations of the double-stranded DNA into electrocompetent TG1 cells (total of ~2.5 x10^9 transformants). The recovered cells were pooled from the electroporations, grown to an OD_{600nm}=1.0 (10^9 cells/mL), and phagemid DNA was purified from 3x10^{11} cells using the PureLink™ HiPure Plasmid Filter Maxiprep Kit (Invitrogen, Carlsbad, CA). The covalently, closed circular DNA (10 µg) was cleaved with *Xma* I (5 units/ µg of DNA) for 16 h at 37°C. DNA from the *Xma* I digest was purified using one QIAquick® PCR Purification Kit (QIAGEN Sciences, Valencia, CA), and 10 transformations were performed to return the DNA into electrocompetent TG1 cells (total of ~5-7x10^9 transformants). The transformed cells were plated onto ten 15 cm 2xYT/CB agar plates, and the next day, colonies were scraped with a total of 40 mL of freezing media and the library stored at -80°C.

4.3.2 Amplification of the FHA1G2 and G2-XmaI libraries

For amplifying the library phage, 2xYT/CB media was inoculated with sufficient number of scraped cells (corresponding to 10-fold coverage of the library diversity), grown to mid-log phase ($4x10^8$ cells/mL), and infected with the trypsin-cleavable helper phage (TM13KO7, 10^{10} pfu/mL) for 1 h at 37°C at 150 rpm. Infected cells, recovered after centrifugation, were resuspended in fresh 2xYT/CB/Kan medium (10 times the initial volume) and phage were amplified for 18-19 h at 30°C, with 250 rpm shaking. Phage were concentrated 100-fold by PEG/NaCl precipitation, passed through 0.45 µm filters, glycerol added to a final concentration of 16%, and the phage library was stored at -80°C. For the four sublibraries, phage was amplified from each library separately and pooled before performing affinity selections with various pT-containing peptides.

4.3.3 Affinity selection against various phosphothreonine peptides

Dynabeads® MyOne™ Streptavidin T1 magnetic beads (Invitrogen Dynal AS, 100 µL) were incubated with the biotinylated pT peptide (1.2 µg; 1.5 µM concentration) for 30 min. All the selection steps were performed at room temperature. The unbound target was removed and the beads were blocked for 1 h with blocking buffer (2% skim milk in PBS with 1 µM free biotin; 1 mL). The phage library ($3x10^{12}$ phage) was incubated for 15 min with equal volume of 4% skim milk in PBS, and then added to the blocked beads. Washing the beads

three times with PBST and twice with PBS minimized non-specific binding of phage particles to the beads. Phage bound to the target were eluted using TPCK treated trypsin (Sigma-Aldrich Corp, St. Louis, MO; 400 µL at 100 µg/mL concentration) and used to infect 800 µL of TG1 cells at mid-log growth phase (OD_{600nm}=0.5) for 40 min at 37ºC. The cells were then plated on one 15 cm 2xYT/CB agar plate and the colonies were scraped the next day with 8 mL of freezing media. For amplifying the phage for the second round of selection, ~10^8-10^9 cells were inoculated into 40 mL of 2xYT/CB media, grown to mid-log, and 5 mL was infected with trypsin cleavable helper phage (TM13KO7; 10^{10} pfu/mL). The infected cells, after centrifugation, were resuspended in 30 mL of 2xYT/CB/Kan media, phage were amplified overnight at 30°C at 250 rpm, and precipitated with PEG/NaCl mixture. The second round of affinity selection was conducted in the same manner, except that beads were washed more (five times with PBST and three times with PBS) before eluting the bound phage. TG1 cells at mid-log were infected with eluted phage, and 10 µL and 100 µL of 10^{-2} and 10^{-4} dilutions were plated on 2xYT/CB agar plates (10 cm). After the second round of affinity selection, 96 individual clones were propagated as phage particles, followed by a phage ELISA to identify clones that recognize the pT peptide ligand. The DNA of positive binding clones was sequenced at the Research Resource Center (RRC) of University of Illinois at Chicago (UIC).

4.3.4 Peptides

Peptides were synthesized at the UIC Research Resource Center, and were >90% pure. All the peptides are biotinylated at their N-terminus, and amidated at their C-terminus. Some of the peptides required the addition of a tyrosine (Y) residue at their C-terminus, which permits determining the peptide's concentration by measuring the solution's optical absorbance with 280 nm wavelength light. Two or three lysine (K) residues were included at the C-terminus of some of the peptides to increase their solubility. All the peptide sequences (except Rad9-pT: SGS-SLEVpTEADATFYAKK (*26*) are obtained from http://www.phosida.com/ (posttranslational modification database) website, and their sequences are listed in Table I.

Table I. Peptide sequences used in this study obtained from http://www.phosida.com/ website.

Accession#	Protein	Phosphosite	Sequence
IPI00003479	Mitogen-activated protein kinase 1	MAPK1-pT(185) MAPK1/3-pTpY or ERK1/2-pTpY (185,187)	HDHTGFLpTEYVATKK HTGFLpTEpYVATRW
IPI00018195	Mitogen-activated protein kinase 3	MAPK3-pT(197) MAPK3-pY(203)	ADPEHDHpTGFLTEYKKK HTGFLTEpYVATRWYR
IPI00013439	Transcription factor jun-B	JunB-pT(255) JunB-3P(251,255, 259)	EARSRDApTPPVSPYKK EARpSRDApTPPVpSPYKK
IPI00234446	Activating transcription factor 2	ATF2-pTpT(69,71) ATF2-pT(69) ATF2-pT(71)	IVADQpTPpTPTRFLKY IVADQpTPTPTRFLKY IVADQTPpTPTRFLKY
IPI00289547	Transcription factor jun-D	JunD-pT(245)	ALKDEPQpTVPDVPYKKK
IPI00796046	Transcription factor Myc	Myc-pT(58) Myc-pTpS(58,62)	KKFELLPpTPPLSPSY KKFELLPpTPPLpSPSY
IPI00165135	Src homology 2 domain containing transforming protein 1	Shc1-pT(35)	GSFVNKPpTRGWLHKK
IPI00018274	Isoform 1 of epidermal growth factor receptor precursor	EGFR-pT(993) EGFR-pY1(998) EGFR-pY2(1092)	RMHLPSPpTDSNFYRA SPTDSNFpYRALMDKK TFLPVPEpYINQSVKK
IPI00017305	Ribosomal protein S6 kinase alpha-1	RSKA1-pT(359)	DTEFTSRpTPKDSPYKK
IPI00003783	Dual specificity mitogen-activated protein kinase 2	MAP2K2-pT(394)	LRLNQPGpTPTRTAYKK

4.3.5 Subcloning FHA affinity reagents into an expression vector

AviTag is a 15 amino acid peptide sequence (GLNDIFEAQ**K**IEWHE) to which a single biotin molecule is conjugated at the lysine (K) residue by the biotin ligase (BirA) enzyme of *E. coli (27-29).* Two oligonucleotides, pET29b-Avitag-Fw (5'- TCGAGTTCGTGCCATTCAATTTTCTGAGCTTCAAAAATATCATTCAGGCC TGC-3') and pET29b-Avitag-Rv (5'-GGCCGCAGGCCTGAATGATATTTTTGAA GCTCAGAAAATTGAAT GGCACGAAC-3'), were annealed to create a double-stranded DNA insert with *Not* I/*Xho* I sticky ends. The pET29b vector, cut with the corresponding *Not* I/*Xho* I restriction enzymes was ligated with the insert overnight at 16°C using T4 DNA Ligase (New England BioLabs). In the final construct, the FHA coding region is in frame with the AviTag and a C-terminal six-histidine tag.

4.3.6 Expression, purification and SDS-PAGE of purified FHA affinity reagents

BL21 DE3 cells (200 mL; Stratagene, La Jolla, CA) harboring the expression vector were grown for 24 h at 30°C/300 rpm in the Overnight Express™ Autoinduction System 1 (Novagen, Madison, WI). The next day, cells were lysed on ice using 1X BugBuster supplemented with Benzonase® Nuclease (0.125 units/µL; Novagen) and complete EDTA-free protease inhibitor cocktail (1X concentration; Roche Applied Science, Indianapolis, IN). BugBuster® 10X

Protein Extraction Reagent (Novagen) was diluted to 1X BugBuster in cell lysis buffer (50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM Imidazole, 20 mM beta-mercaptoethanol). Cell lysis was carried out for 30 min at room temperature on a shaking platform, followed by centrifugation at 8,000 rpm for 10 min (at 4°C), to pellet the cellular debris. Ni-NTA agarose (Qiagen) was equilibrated with 5 volumes of lysis buffer, and then incubated with the clarified cell lysate for 1 h at room temperature on a rotating platform. The unbound protein was removed by washing with a total of 80 mL of wash buffer 1 (50 mM NaH_2PO_4, 0.3 M NaCl, 10 mM Imidazole, 0.5% Tween 20) and 50 mL of wash buffer 2 (50 mM NaH_2PO_4, 0.3 M NaCl, 20 mM Imidazole, 0.5% Tween 20). The FHA-AviTag fusions were eluted off the resin with 2 mL of elution buffer (50 mM NaH_2PO_4, 0.3 M NaCl, 250 mM Imidazole) for 1 h on a shaking platform. The purified proteins were biotinylated using BirA enzyme (Avidity, LLC, Aurora, CO), following the manufacturer's instructions.

4.3.7 Cell culture

The mouse embryonic fibroblast cell line (NIH/3T3) was a gift of Dr. Hua Jin (University of Illinois at Chicago). The cells were cultured in Dulbecco's minimal essential medium (DMEM), which was supplemented with 2 mM L-Glutamine (200 Mm; 100X, Gibco Life Technologies, Grand Island, NY) and 10% heat-inactivated Newborn Calf serum (Gibco Life Technologies). In some cases, cells

were serum starved for 24 h in DMEM media with 20 mM HEPES buffer solution (1M; Gibco Life Technologies). Penicillin (50 units/mL) and Streptomycin (50 µg/mL) antibiotics (100X; Gibco Life Technologies) were added to all media

4.3.8 Cell staining

NIH/3T3 ($8x10^4$) were cultured in complete media on 35 mm tissue culture dish (Becton Dickinson Labware, Franklin Lakes, NJ) containing 3 or 4 poly-L-Lysine coated coverslips (12mm round, No. 1 German glass, BD BioCoat™ Cellware, Batavia, IL) for 18-20 h. The cells were serum starved for 24 h at 37°C, stimulated with human Epidermal Growth Factor (EGF; 25 ng/mL; 2 mL/plate; Sigma-Aldrich) for 10 min at 37°C. In some cases, the cells were treated with 10 µM MEK1/2 inhibitor (U0126, Cell Signaling Technology, Danvers, MA) and incubated for 2 h at 37°C. Later the inhibitor was aspirated out and the cells were stimulated with EGF (25 ng/mL; 2 mL/plate) added to the serum starvation media along with MEK1/2 inhibitor (10 µM) and incubated for 10 at 37°C. After washing once time with PBS without calcium and magnesium, the cells were fixed for 10 min at room temperature with 4% formaldehyde solution (Thermo Scientific, Pittsburgh, PA; 16% Formaldehyde (w/v) ampules, Methanol-free). After washing the fixed cells three times with PBS, non-specific binding sites on the coverslips were blocked for 40 min at room temperature with 3% Bovine serum albumin (Sigma-Aldrich) and 0.1% Triton® X-100 (Bio-Rad Laboratories, Hercules, CA) in

PBS supplemented with 10% normal goat serum (Cell Signaling Technology). The blocking solution was aspirated off the coverslips using a Pasteur pipette connected to vacuum, and the cells were incubated with the primary antibody or FHA affinity reagent for 1 h at room temperature or overnight at 4°C. The cells were washed 5 times (5 min each wash) with PBST (PBS - 0.1% Tween 20), followed by incubation with a secondary antibody for 30 min at room temperature. The cells were washed 5 times (5 min each wash) with PBST, and mounted on Gold Seal Microscope slides (Thermo Scientific) using VECTASHIELD® Hard Set™ Mounting Medium with DAPI (Vector Laboratories, Inc.). The cells were visualized using an epifluorescence microscope at University of Illinois at Chicago.

4.3.9 Western blotting of mammalian cell lysates

NIH/3T3 cells were cultured in complete media in 100 mm cell culture dishes (Corning Incorporated, Corning, NY) until 70-80% confluent. The cells were serum starved for 24 h at 37°C, and stimulated with human Epidermal Growth Factor (EGF; 25 ng/mL; 10 mL/plate, Sigma-Aldrich) for 10 min at 37°C. In instances where the cells were treated with MEK1/2 inhibitor (U0126, Cell Signaling Technology), the inhibitor was added to the cells (10 μM, 10 mL/plate) and incubated for 2 h at 37°C, before stimulation with EGF. The inhibitor was aspirated out and the cells were stimulated with EGF (25 ng/mL), along with

MEK1/2 inhibitor (10 μM), and incubated for 10 min at 37°C. Immediately following treatment, the cells were washed once with ice cold PBS, lysis buffer (0.5 mL/plate), was added and the plates were incubated on ice for 10 min (lysis buffer recipe was from (*30*)). Using an ice-cold cell scrapper, the cells were scrapped off the plates and incubated on ice for an additional 10 min. The cells were centrifuged at 14,000 rpm 4°C to remove the cellular debris. The protein concentration in the clarified cell lysate was determined by measuring the absorbance at 280 nm. The cell lysate was stored in single-use aliquots at -80C. The cell lysate (50 μg) was electrophoretically resolved in a 10% Mini-Protean® TGX™ Precast Gel (Bio-Rad Laboratories) and transferred to an Immobilon®FL membrane (Millipore Corporation, Billerica, MA). Non-specific binding sites on the membrane were blocked by incubation for 2 h with Odyssey blocking buffer (*LI-COR*® Biosciences, Lincoln, NE), incubated with the primary antibody for 2.5 h, washed 3 times with PBST (2 min each wash), and then incubated with the secondary antibody for 30 min. After washing 3 times with PBST (2 min each wash), the membrane was rinsed once with PBS and bands were visualized on the Odyssey® Fc imaging system. For detecting phospho-ERK1/2, the positive control used was Anti-phospho-ERK1 [pThr^{202}/pTyr^{204}] & ERK2 [pThr^{185}/pTyr^{187}] (MAPK), affinity-purified rabbit antibody (Sigma-Aldrich), 0.5 μg/mL final concentration). The secondary antibody used was Goat anti-Rabbit IRDye® 680 (*LI-COR*® Biosciences; 0.1 μg/mL). For detecting phospho-ERK1/2 proteins, the

membrane was probed with 1 μM of the biotinylated FHA affinity reagent, A5. The secondary reagent was Streptavidin IRDye® 700DX conjugated (ROCKLAND, Gilbertsville, PA; 1:10,000 dilution).

4.4 Results and discussion

4.4.1 Characterization of the two libraries of FHA1G2 variants

A library of FHA1G2 variants, generated by mutating residues important for interaction with the pT-peptide, is anticipated to be a useful source of affinity reagents with new anti-phosphopeptide binding specificities. Previously, it has been shown that variants of Erbin PDZ domain, generated by mutating ten residues known to be important for interaction with peptide ligands from structural studies, had different binding specificities compared to the wild-type PDZ domain (*31*). Ernst *et al*. have demonstrated that it is possible to change the specificity of a protein interaction domain such that it binds to ligands for which interaction was not previously detected.

To test the feasibility of generating novel phosphopeptide-binding specificities, we constructed two different phage-displayed libraries of FHA1G2 variants. The first library (FHA1G2 library) had 8 residues randomized (from the β4-β5 and β10-β11 loops combined), which were found to be important to bind to the pT peptide ligand from our alanine-scanning results. Sequencing the DNA of 30 individual clones, randomly chosen from the FHA1G2 library, revealed that

61% clones had both loops mutated, 17% had mutations in one loop, 11% were wild-type sequences, and 11% contained stop codons. Clones with stop codons will not get displayed on the surface of phage, so the phage-displayed library used for selection will not have any representation of such phage. The library's size was determined to be $5x10^9$, of which $3x10^9$ clones have mutations in both loops.

The G2-XmaI libraries were constructed using the pKP700 phagemid vector (a modified version of the pKP600 phagemid vector, in which the truncated gene III was replaced with the full-length gene III) containing two *Xma* I recognition sites (one in β4-β5 loop and the other in β10-β11 loop) in the FHA1G2 coding sequence. Colony PCR, followed by *Xma* I digestion of 93 colonies, randomly chosen from the four G2-XmaI I subibraries, revealed that 99% of the clones had both the loops mutated (Fig. 1). Thus, the efficiency for library construction increased from ~60% to 99%. In this method, we have combined two strategies for efficiently cleaving the wild-type phagemid vector; one is to use the intrinsic ability of TG1 cells to cleave the uracil-containing wild-type DNA (50-60% efficient, based on our experience) and the second is eliminating the remaining wild-type DNA by cleavage with *Xma* I.

Another well-established strategy to prevent phage-display of the wild-type protein is to incorporate stop codons in the coding sequence of the original template or the gene III coding sequence. The recombinant protein is displayed

only when the stop codon is replaced by sequence of the annealed oligonucleotide. This method has been used successfully for the construction of phage-displayed combinatorial peptide libraries which resulted in 100% recombinant phage (*32*).

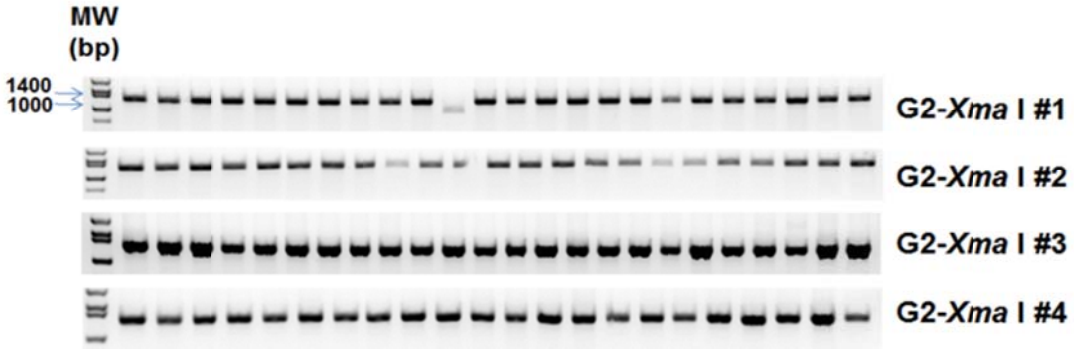


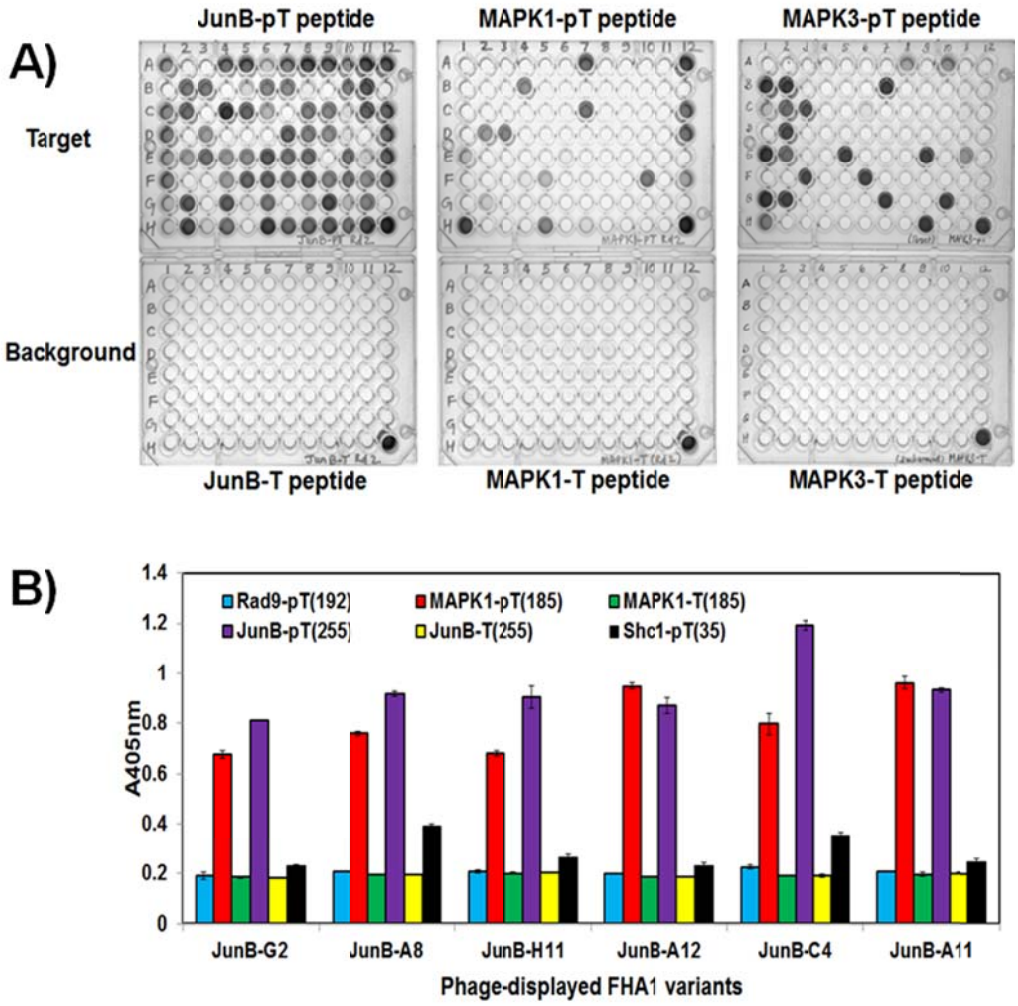
Figure 1. Determining the percentage of recombinants in the G2-*Xma* I libraries. Colony PCR was performed on a total of 93 colonies, from the four libraries combined, followed by digestion with *Xma* I restriction enzyme. The cut PCR products were run on a 1% agarose gel. The expected PCR product size of the recombinants with mutations in both the loops is 1273 bp. The molecular weight (MW) marker was electrophoresed in the far left-hand lane; the molecular weights are noted in base pairs.

4.4.2 Isolating FHA1 variants with novel binding specificities

4.4.2.1 Affinity reagents isolated from the FHA1G2 library against MAPK1-pT (185), MAPK3-pT (197) and JunB-pT (255) peptides

Using the FHA1G2 library, with 8 residues randomized, two rounds of affinity selection were performed against the MAPK1-pT (185) peptide, MAPK3-pT (197) and JunB-pT (255) peptide. Ninety-five individual colonies were propagated as phage particles and tested by phage ELISA for binding to the cognate pT peptide (target) and non-phosphorylated form of the same peptide (background) (Fig. 2A). From examining the DNA sequences of 12 binding clones from each selection experiment, 2, 6 and 5 clones with unique sequences were identified for the affinity selection experiments with the MAPK1-pT (185), JunB-pT (255) and MAPK3-pT (197) peptides, respectively. The specificity of the two MAPK1-pT (185) clones (H1 and B4), six JunB-pT (255) clones (A8, A11, A12, C4, G2 and H11), and five MAPK3-pT (197) clones (B1, B2, E9, G1, G2) was further tested by phage ELISA. All of the JunB clones were observed to be cross-reactive with the MAPK1-pT (185) peptide and *vice versa* (Fig. 2B-C), and a possible explanation for this is that both these pT peptides have the same pT (+3) residue (in this case Val), which has been shown from previous studies (*12, 26*) to be a major determinant of binding. The specificity phage-ELISA of five MAPK3 clones revealed that two clones (B1 and G1) bound specifically only to

the cognate MAPK3-pT (197) peptide and not to other unrelated phosphopeptides (Fig. 2D). Three other clones (B2, E9 and G2) showed weak cross-reactivity with JunB-pT peptide; cross-reactivity may be a consequence that both peptides contain a hydrophobic amino acid in the pT (+3) position (Leu and Val for MAPK3-pT and JunB-pT peptides, respectively).



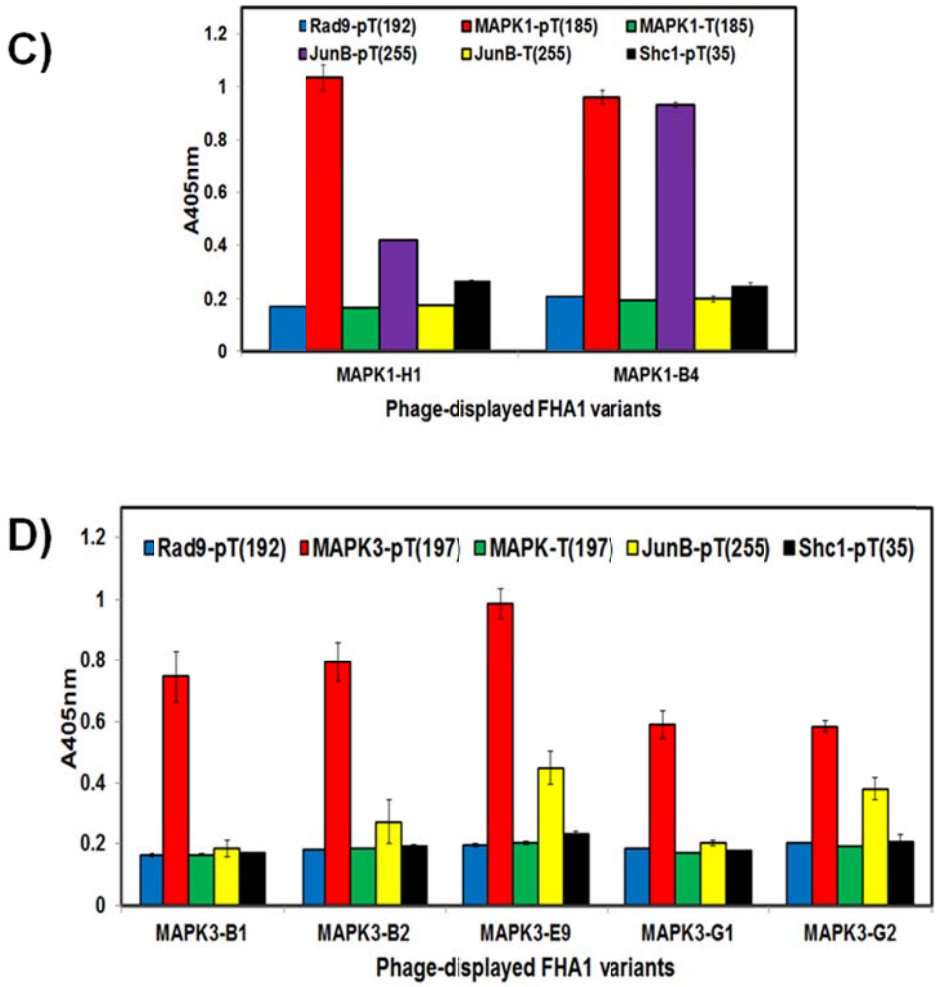


Figure 2. Phage ELISA. A) The target plate was coated with the cognate pT peptide and the background plate was coated with the non-phosphorylated form of the same peptide. The last microtiter plate well (H12) was coated with a control pT peptide and incubated with its specific binding phage (positive control). Binding was detected using anti-M13/HRP conjugated antibody. Positive signal only on the target plate represents potential binding clones. The binding of FHA1 affinity reagents isolated against B) JunB-pT peptide, C) MAPK1-pT peptide, and D) MAPK3-pT peptide were tested for binding to the non-phosphorylated form of the same peptide, the Rad9-pT peptide (cognate peptide for wild-type FHA1 domain) and few other pT peptide sequences.

We purified the MAPK3-B1 clone and further tested its binding specificity on a large number of phosphopeptides. This protein ELISA revealed that the B1 clone bound its cognate pT peptide ligand with, little or no cross-reaction with 10 other phosphopeptide sequences and the non-phosphorylated form of peptide ligand (Fig. 3). The only observable cross-reactivity was with the Myc-pT peptide, which was again not surprising because both phosphopeptides contain leucine in the pT (+3) position. The FHA1 domain has been previously observed (*33*) to cross-react with various pT peptides that contain the same pT (+3) residue, with dissociation constants ranging from 0.36 to 71 μM, suggesting that residues on either side of the pT moiety contribute to and fine-tune the binding specificity of the FHA1 domain.

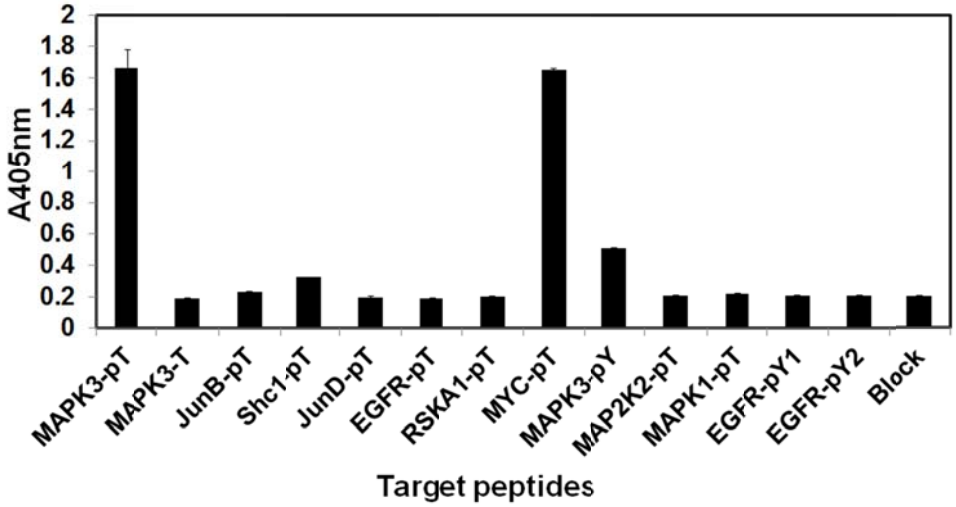


Figure 3. Characterizing the specificity of an anti-MAPK3 affinity reagent. The binding specificity of an anti-MAPK3 affinity reagent (B1) was monitored with 11 different phosphopeptide sequences, along with its cognate pT peptide (MAPK3-pT) and the non-phosphorylated form of the same peptide (MAPK3-T).

In conclusion, the specificity of the FHA1 domain can be engineered to bind to new pT-containing peptides for which binding was not previously detected. Binding of the FHA variants to peptides is phosphorylation dependent, as no binding was observed for the non-phosphorylated form of peptides. While specific binding clones were isolated, cross-reactivity was observed whenever the peptides shared the same pT (+3) residue. A strategy that is commonly used to eliminate cross reactive clones from the library is a deselection approach, which is depicted in Figure 4.

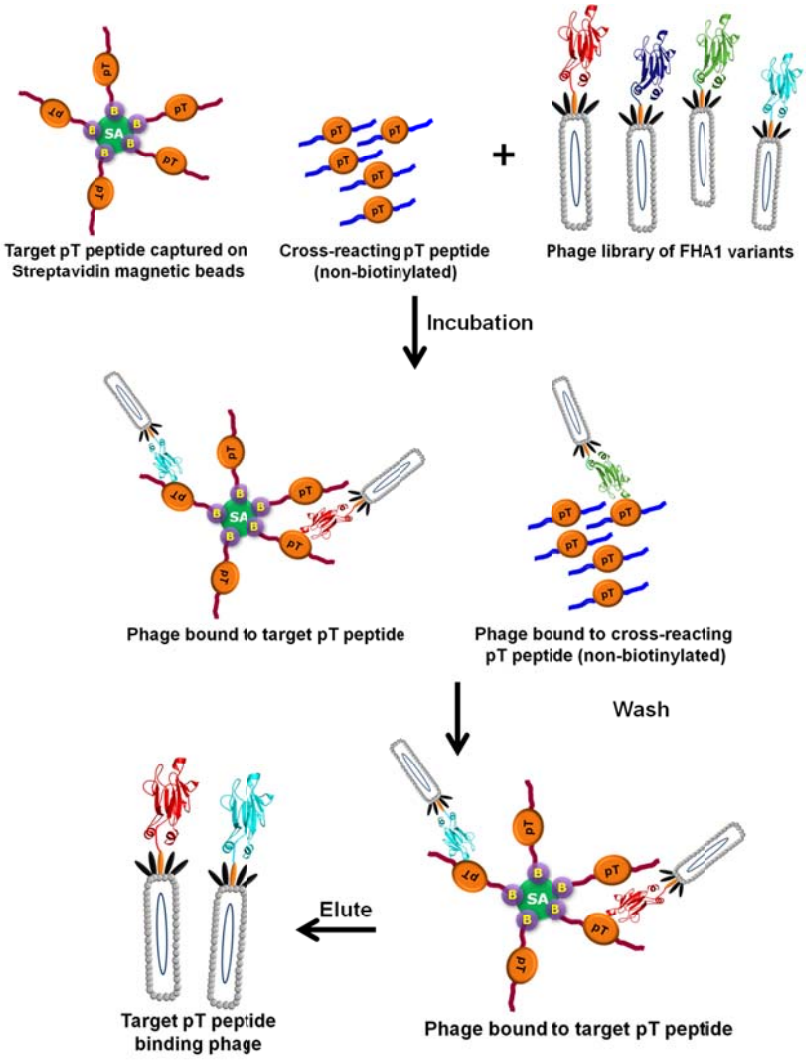


Figure 4. Deselection strategy for eliminating cross-reactive clones. The phage library displaying FHA1 variants is incubated with the target pT peptide (biotinylated and captured on streptavidin magnetic beads) along with the pT peptide (non-biotinylated) for which cross-reactivity was observed. Depending on their affinity, the library phage will bind to one or the other pT peptide. Washing will eliminate the phage bound to the non-biotinylated cross-reactive pT peptide. Target pT peptide bound phage are captured on streptavidin magnetic beads which are then eluted by acid or trypsin and propagated to amplify the binding clones which will be used for subsequent rounds of affinity selection.

4.4.2.2 Affinity reagents isolated from the FHA1G2 library against a dual phosphorylated peptide from MAPK1/3 or ERK1/2 Ser/Thr protein kinase

During the ERK-MAPK pathway, in response to an external stimulus, Ser/Thr protein kinases, MEK1 and MEK2 activate MAPK3/ERK1 (p44) and MAPK1/ERK2 (p42) protein kinases by phosphorylating both Thr (185) and Tyr (187) residues present adjacent to each other in the activation loop of the kinase (*34*). The activated ERK1 and ERK2 proteins in turn activate numerous cytoplasmic and nuclear substrates involved in various signaling events (*35*). A dual phosphorylated peptide (ERK1/2-pTpY (185,187)) containing these two phosphosites was used as a target for performing affinity selection. After two rounds of affinity selection, binding clones were isolated and upon sequencing 12 clones, 3 clones with unique sequences (A5, A9 and C5) were identified. All the three clones showed specific binding for their cognate pT peptide (shown as light blue histograms), with minimal or no cross-reactivity against seven other phosphopeptides from the ERK-MAPK pathway (Fig. 5A). Thr (185) and Tyr (187) in the activation loop are represented as red and blue spheres respectively in the crystal structure of ERK2 (*36*) (Fig. 5B). Binding was observed primarily to the dual phosphorylated peptide with minimal or no binding detected for the two mono-phosphorylated peptides (Fig. 5C).

The FHA domain of Dun1 protein from *S. cerevisiae* binds well to doubly phosphorylated peptides. This tight binding observed for Dun1 FHA domain with the dual phosphorylated peptides is attributed to the presence of two phosphopeptide binding sites in its three dimensional structure (*37*). It has never been shown before that an FHA1 variant can recognize a doubly phosphorylated peptide. This screening result demonstrates that the phage-displayed library of FHA1 variants contains members that can specifically recognize doubly phosphorylated epitopes in addition to those that bind to singly phosphorylated epitopes.

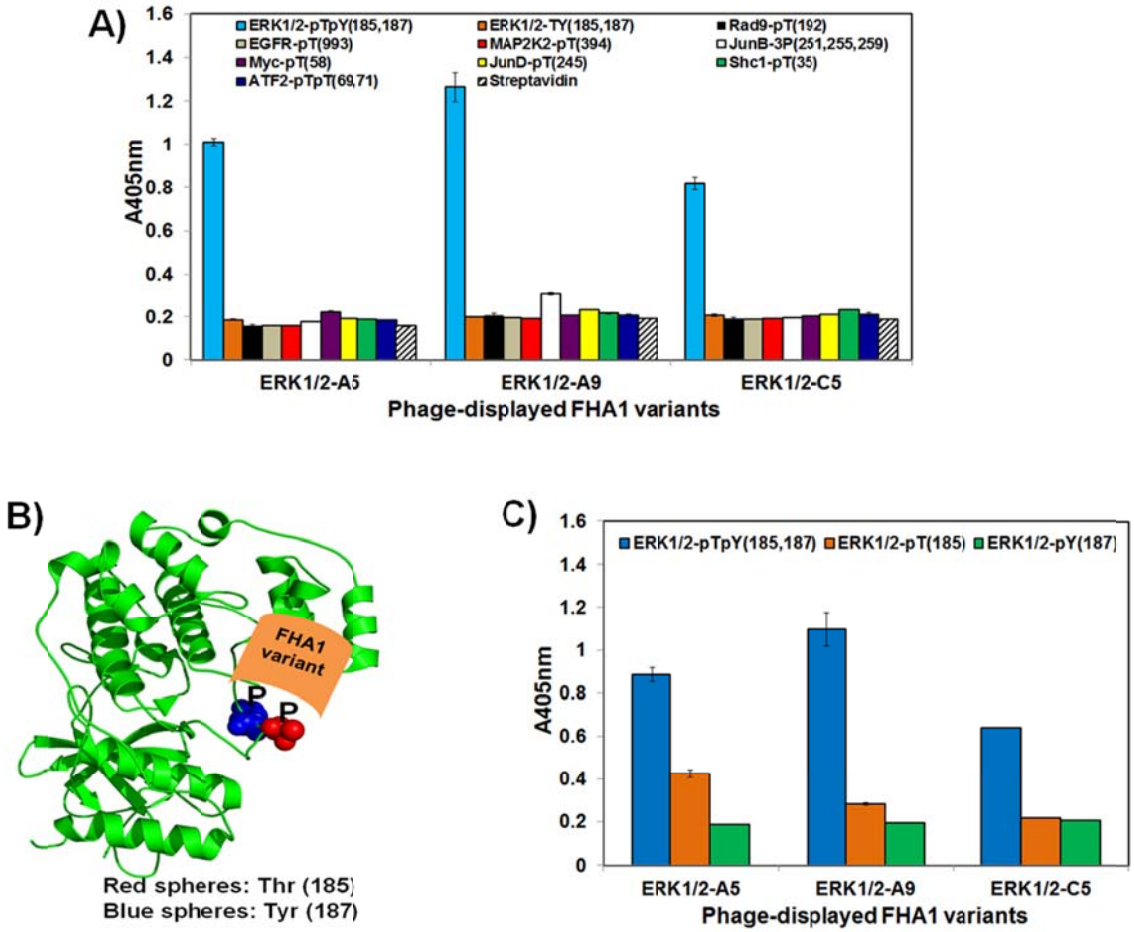


Figure 5. Specificity test for the anti-MAPK affinity reagents. A) The binding specificity of three FHA affinity reagents isolated against a dual phosphorylated peptide from MAPK was tested for binding to the non-phosphorylated form of the same peptide, the Rad9-pT peptide (cognate peptide for wild-type FHA1 domain) and seven other phosphopeptides from the ERK-MAPK pathway. B) The crystal structure of ERK2 protein (Protein data base code-2FYS (*36*), showing Thr (185) and Tyr (187) in the activation loop, which get phosphorylated by MEK1/2 kinase upon stimulation (*34*). FHA1 affinity reagents were isolated against a dual phosphorylated peptide, harboring these two phosphosites. C) The three anti-MAPK FHA1 affinity reagents were tested for binding to the dual phosphorylated peptide, as well as to the two singly-phosphorylated peptides. Phage binding was detected using anti-M13/HRP conjugated antibody.

The nucleotide and amino acid sequences of the 8 positions, which were randomized during library construction, are shown in Table II. Mutations at two positions, R83 and L84, in the FHA1G2 domain changed its specificity from recognizing its cognate Rad9-pT peptide (with Asp in the pT (+3) position) to a new pT peptide with Val in the pT (+3) position. Structural studies and site directed mutagenesis experiments have revealed that residue R83 (located in the β4-β5 loop) from the FHA1 domain makes ionic interactions with the pT (+3) Asp and is a major determinant of the binding specificity (*26*). Interestingly, the FHA2 domain, with Arg at residue 617, does not bind to the FHA1 pT-peptide ligand, but instead binds to a peptide that has a hydrophobic residue in the pT (+3) position. From structural studies, it is clear that the Arg617 in the FHA2 domain is involved in an intramolecular interaction with Asp683, making it unavailable to bind to the Asp in the pT-peptide (*38, 39*). Therefore, in our affinity selection experiments, mutation of the R83 residue resulted in loss of recognition of the original peptide sequence. Among the selected clones, 75% of them retained the original L78 residue, which means that it might not contribute to the binding specificity, therefore, while constructing the second library (G2-XmaI), this residue was not selected, instead S82 residue was included as it was shown from structural studies to interact with the pT-peptide (*26*). The residues selected for randomization in the β10-β11 loop were mutated in about 60% of the clones, indicating that this loop contributes to the binding specificity.

Table II. The sequences of the FHA1 affinity reagents isolated from the FHA1G2 library.

Clone ID	Nucleotide sequence (β4-β5 loop) (78-84)	Amino acid sequence (β4-β5 loop) (78-84)	Nucleotide sequence (β10-β11 loop) (133-139)	Amino acid sequence (β10-β11 loop) (133-139)
FHA1G2	TTAGGTAACATTAGCCGCTTA	L G N I S R L	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
MAPK1-B4	TTAGGTAACATTAGCCATTAT	L G N I S H Y	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
MAPK1-H1	TTAGGTAACATTAGCACGCAT	L G N I S T H	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
MAPK3-B1	TTGGGTAACATTAGCTATATT	L G N I S Y I	CGTCTGGATGCGGAAAGCCAT	R L D A E S H
MAPK3-B2	TTGGGTAACATTAGCTATATT	L G N I S Y I	AGGTTGGGGGAGGAAAGCTCT	R L G E E S S
MAPK3-G2	TTGGGTAACATTAGCTATTTG	L G N I S Y L	CTTCCTTCTCTTGAAAGCGAT	L P S L E S D
MAPK3-E9	TTTGGTAACATTAGCGAGTAT	F G N I S E Y	CCGGGTCGGAGGGAAAGCGAT	P G R R E S D
MAPK3-G1	TTAGGTAACATTAGCACGATT	L G N I S T I	CATACTTCTGATGAAAGCAAT	H T S D E S N
JunB-A8	TTAGGTAACATTAGCTATTAT	L G N I S Y Y	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
JunB-A11	TTAGGTAACATTAGCCATTAT	L G N I S H Y	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
JunB-C4	TTGGGTAACATTAGCTATTTG	L G N I S Y L	CTTCCTTCTCTTGAAAGCGAT	L P S L E S D
JunB-A12	TTAGGTAACATTAGCCATTTT	L G N I S H F	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
JunB-G2	TTAGGTAACATTAGCTATTTT	L G N I S Y F	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
JunB-H11	TTAGGTAACATTAGCTTTTAT	L G N I S F Y	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
ERK1/2-A5	ATGGGTAACATTAGCAGTATT	M G N I S S I	TTGACTGGGCGTGAAAGCACT	L T G R E S T
ERK1/2-C5	CACGGTAACATTAGCATGATT	H G N I S M I	CGTGCTAGGACGGAAAGCCGT	R A R T E S R
ERK1/2-A9	ATGGGTAACATTAGCATGATT	M G N I S M I	CGGCAGGGGCGGGAAAGCCGG	R Q G R E S R

The sequences of all the FHA affinity reagents are compared to the sequence of the FHA1G2 variant, which is the starting scaffold used for library construction. The 8 residues randomized during library construction are show in blue in the FHA1G2 sequence. If these residues differ from G2, they are shown in red in the binding clones, whereas residues that were selected for randomization during library construction are shown in black.

4.4.2.3 Affinity reagents isolated from the G2-XmaI library against JunD-pT (245), Myc (Myc-pT (58) and ATF2-pTpT (69, 71) peptides.

Using the G2-XmaI library, in which 10 residues were randomized in the FHA1 domain, I performed rounds of affinity selection against a doubly phosphorylated peptide (ATF2-pTpT) (69, 71), from the Activating transcription factor 2, and singly phosphorylated peptides from transcription factors jun-D (JunD-pT (245)) and Myc (Myc-pT (58)). Ninety-five individual clones were evaluated by phage ELISA for binding to the cognate pT peptide and the non-phosphorylated form of the same peptide. The potential binders were sequenced and the clones with unique sequences were further characterized for specificity.

ATF2 transcription factor is phosphorylated by Jun and p38 mitogen-activated protein kinases and high levels of phosphorylated ATF2 transcription factor are observed in tumor cells (*40, 41*). For the ATF2 transcription factor, two unique binding clones, A3 and B12, were isolated. When they were tested against an array of other phosphopeptides (Fig. 6A), they showed the strongest binding to their cognate pT peptide ligand. No binding was detected for the non-phosphorylated form of the same peptide, or the Rad9-pT peptide (cognate pT peptide for the wild-type FHA1 domain). In the phage ELISA, the signal above background was only 2-fold in phage, which suggests that its affinity (K_d) for the target phosphopeptide is low. Presumably, its affinity can be improved by affinity maturation, in which error-prone-PCR of the coding sequence of the binding

clone generates a secondary library from which variants with higher affinity can be selected. Clone B12 binds equally well to the doubly phosphorylated peptide (shown as blue histograms), as well as to the peptide phosphorylated at T69 (shown as orange histograms), while A3 shows better binding to the ATF2-pT(69) mono phosphorylated peptide (Fig. 6B). Interestingly, even though a dual phosphorylated peptide was used for affinity selection, the pT(69) residue in the peptide contributes the bulk of the peptide's recognition by the selected FHA domain.

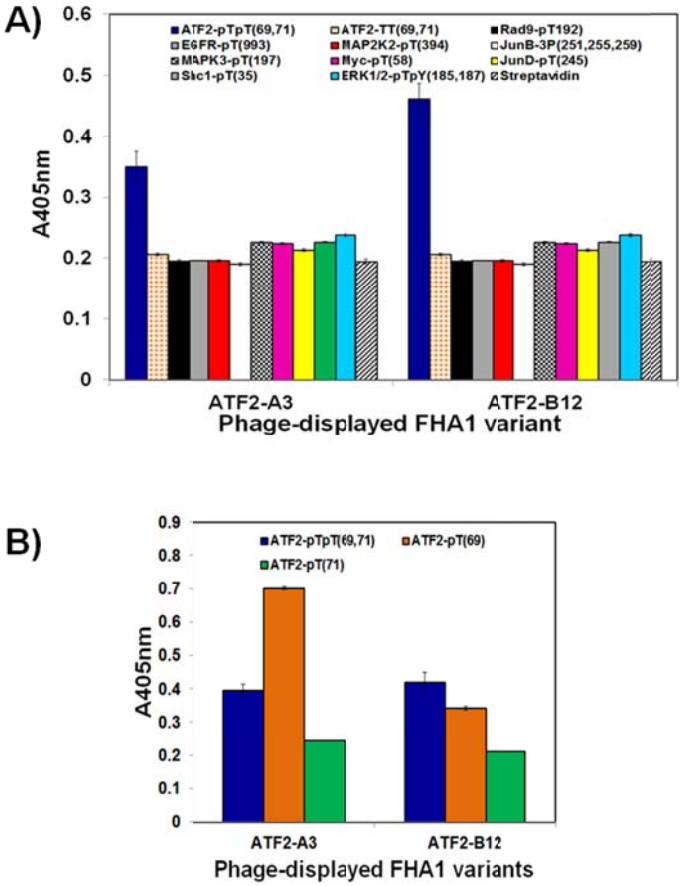


Figure 6. Specificity test for the anti-ATF2 affinity reagent. A) The binding specificity of two anti-ATF2 reagents was tested for binding to the non-phosphorylated form of the same peptide, the Rad9-pT peptide (cognate peptide for wild-type FHA1 domain) and eight other phosphopeptides from the ERK-MAPK pathway. B) The anti-ATF2 reagents were tested for binding to the dual phosphorylated peptide, as well as to the two singly phosphorylated peptides. Phage binding was detected using anti-M13/HRP conjugated antibody.

To isolate an affinity reagent for phosphorylated Myc, I performed affinity selection against a peptide phosphorylated at T58 (Myc-pT (58)). One unique binding clone, A4, was isolated, which in phage ELISA exhibited very specific binding to the cognate peptide (shown as blue histogram; Fig. 7A). Similarly, one affinity reagent was isolated for JunD-pT peptide (Fig. 7B); however, it showed cross-reactivity with the original Rad9-pT peptide, which can again be rationalized from the fact that both these peptides have Asp in the pT (+3) position, and this residue is known to determine the binding specificity for various different FHA domains. As anticipated, the FHA1G2 variant, which is the starting scaffold for library construction, also cross-reacted with the JunD-pT peptide. It is possible that the observed cross-reactivity is due to the formation of ionic interactions between the positively charged Arg residues located in the loop regions of the FHA domain and the negatively charged Asp in the pT (+3) position of the peptide. It has previously been shown, from screening combinatorial peptide libraries, that the FHA1 domain from Rad53 protein can bind to different pT peptide sequences from the Rad9 protein, all of which have Asp in the pT(+3) position (*33*). However, the binding affinity varies for different peptides because residues on either site of the pT peptide also contribute to binding to the FHA1 domain. The nucleotide and amino acid sequences of all the affinity reagents, at the 10 randomized positions, are shown in Table III.

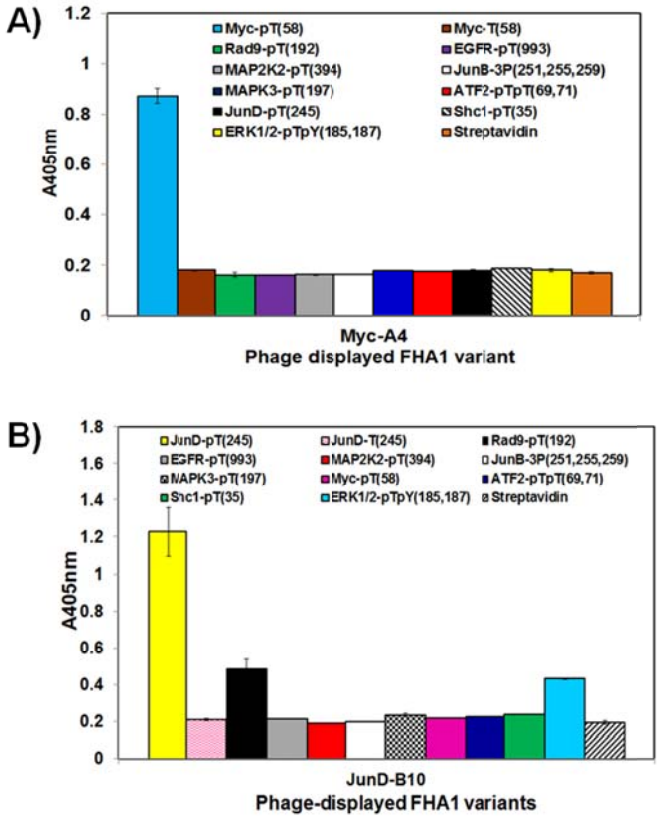


Figure 7. Specificity of the anti-Myc and anti-JunD affinity reagents. A) The binding specificity of anti-Myc reagent was tested for binding to the non-phosphorylated form of the same peptide, the Rad9-pT peptide (cognate peptide for wild-type FHA1 domain) and eight other phosphopeptides from the ERK-MAPK pathway. B) The binding specificity of one FHA reagent isolated against a phosphorylated peptide from JunD transcription factor was tested for binding to the non-phosphorylated form of the same peptide, the Rad9-pT peptide (cognate peptide for wild-type FHA1 domain) and eight other phosphopeptides from the ERK-MAPK pathway. Phage binding was detected using anti-M13/HRP conjugated antibody.

Table III. The sequences of the FHA1 affinity reagents isolated from the G2-XmaI library.

Clone Name	Nucleotide sequence (β4-β5 loop) (82-84)	Amino acid sequence (β4-β5 loop) (82-84)	Nucleotide sequence (β10-β11 loop) (133-139)	Amino acid sequence (β10-β11 loop) (133-139)
FHA1G2	AGCCGCTTA	S R L	GGCGTAGGTGTAGAAAGCGAT	G V G V E S D
Myc-A4	CTTCCTGTG	L P V	CGGACTGATCCTACGGGTACG	R T D P T G T
ATF2-A3	AGCTGGGTT	S W V	CAGCCTGCTAGGGAAAGCAAT	Q P A R E S N
ATF2-B12	CCGTGGGTG	P W V	GAGCCGGGGACGCGTAGGCGG	E P G T R R R
JunD-B12	CCTACGGTG	P T V	GGGCGGCGTTCTGGTGCTCCT	G R R S G A P

The sequences of all the FHA affinity reagents are compared to the sequence of the FHA1G2 variant, which is the starting scaffold used for library construction. The 10 residues randomized during library construction are show in blue in the FHA1G2 sequence. If these residues differ from G2, they are shown in red in the binding clones.

4.4.3 Determining peptide residues important for FHA recognition

To map the residues in the FHA1 domain that are most important in recognizing the pT-peptide, I performed alanine scanning of the MAPK3-pT peptide ligand (ADPEHDH**pT**GFLTEYKKK). My results revealed that the pT (-1, +2 and +3) residues are critical for binding to the FHA1 domain, as replacing these residues led to >90% loss of binding (Fig. 8). Intriguingly, there was 60-70% loss of binding when pT (+1, +4, -2 or -3) residues were replaced with alanine, indicating that these residues also contribute to binding. Mutating the pT (-4) residue had no effect on binding, suggesting that this residue is not

important. Additionally, no binding was detected for either the non-phosphorylated form of the same peptide or for a pT-containing peptide in which the residues (-4 to +4) were replaced with alanine.

From previous studies (*42*) it has been shown that the pT (+3) residue is a key determinant for binding specificity for FHA domains. Therefore, a pT peptide was synthesized in which the +3 residue was fixed and residues from -4 to +4 were replaced with alanine. To our surprise there was only 15% residual binding, which suggests that the other residues surrounding the pT also contribute to the binding specificity in addition to the pT (+3) residue. Together, these results indicate that the pT (-3 to +4) residues form the binding epitope for the FHA1 domain. My findings corroborate a published study (*26*), in which it was demonstrated that the pT-peptide residues (-3 to +2) are important for binding in addition to the pT-moiety and the pT (+3) residue. When the (-3 to +2) residues in the pT-containing peptide were mutated, there was a > 160-fold decrease in the binding affinity of the FHA1 domain for its cognate-pT peptide. Similarly mutating these peptide residues in the cognate peptide for FHA2 domain completely abolished its binding to its cognate pT-peptide. This supports my observation that, in addition to the pT and the +3 residue, residues N- and C-terminal to the pT moiety also contribute to binding.

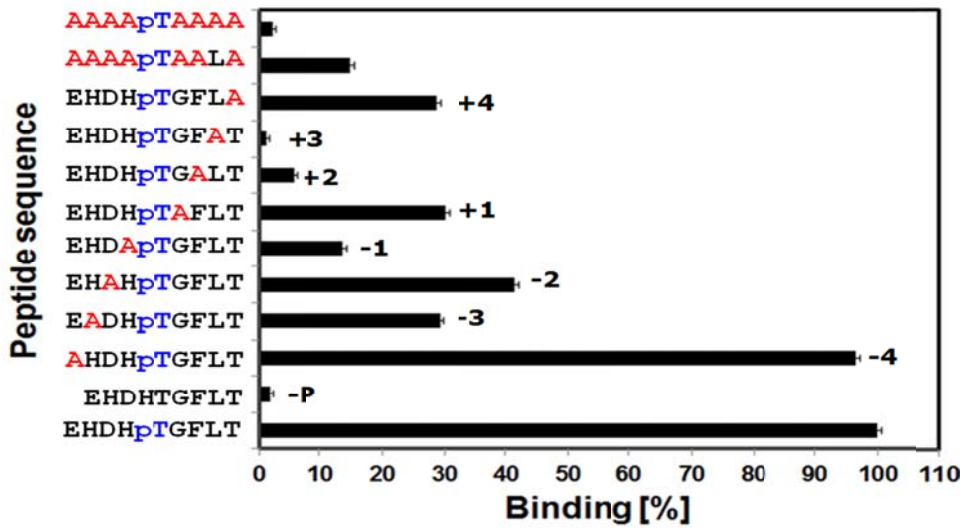


Figure 8. Characterizing the specificity of an anti-MAPK3 affinity reagent, and determining the peptide residues important for interaction. Binding of the anti-MAPK3 affinity reagent (B1) to its cognate MAPK3-pT peptide is set to 100% and the binding to various alanine-scanned peptide variants is compared to it.

4.4.4 SDS-PAGE analysis of purified FHA affinity reagents

The purified FHA-AviTag fusions were resolved in a 10% SDS-PAGE gel, followed by visualization with Coomassie Brilliant Blue (Fig. 9). The FHA-AviTag fusions were the correct size (~22 kDa). The yield of the FHA-AviTag fusion proteins is ~20-25 mg/L (culture flask).

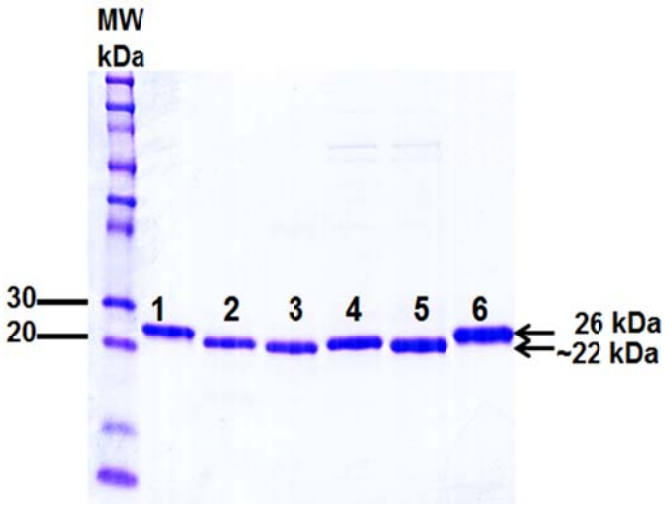


Figure 9. SDS-PAGE analysis of AviTag fusions of FHA variants. 1 µg of the AviTag fusions of FHA1 variants (~22 kDa) were run on a 10% SDS-PAGE gel and visualized after staining with Coomassie Blue stain. Protein loading order- Molecular weight marker (kDa), Lane 1: GST (1 µg, 26 kDa), Lane 2: anti-phospho-Myc FHA reagent (A4), Lane 3: anti-phospho-ERK1/2 FHA reagent (A5), Lane 4: the original scaffold protein (FHA1G2), Lane 5: point mutant of the FHA1G2 domain (D139A), Lane 6: GST (3 µg, 60 kDa).

4.4.5 Using FHA affinity reagents as probes in Western blotting

An ideal format to test the specificity of the affinity reagent is Western blotting of a total cell lysate. Cell lysates were prepared both before and after stimulating NIH/3T3 cells with EGF or after incubation with MEK1/2 inhibitor (U0126), followed by EGF stimulation. U1026 is a non-competitive inhibitor of MEK1 and MEK2 protein kinases (*43*), which are responsible for activating ERK1

and ERK2 protein kinases by phosphorylating both the Thr202/Tyr204 and Thr185/Tyr187 residues in the activation loop of the kinase domains of ERK1 and ERK2 kinases, respectively (*44*). Cell lysates (50 µg), from all the treatments were resolved in an SDS-PAGE gel and the proteins were transferred to a nitrocellulose membrane. An anti-phospho-ERK1/2 IgG (positive control) revealed specific staining of ERK1 (44 kDa) and ERK2 (42 kDa) proteins, only when the cells were stimulated with EGF (Fig. 10, panel b). Unfortunately, I did not observe any binding with the anti-phospho-ERK1/2 FHA reagent A5 with any cellular protein, regardless of the cellular treatment (Fig. 10, panel a).

I suspect that the negative result is due to the fact that A5 reagent has a modest affinity to its peptide ligand. From preliminary experiments with the BIND reader, I determined its K_d to be ~2 µM. With a micromolar K_d value, the reagent-target complex will have a short half-life (i.e., seconds) and the affinity reagent will be quickly lost during the various washing steps of Western blotting. In contrast, affinity reagents with low nM K_d work well in pull-down experiments and western blotting (*45*).

Instead of improving the intrinsic affinity of the A5 reagent through affinity maturation, I decided to generate bivalent forms of the A5 affinity reagent by fusing it to the Fc portion of an IgG to improve its apparent affinity through avidity. Bivalent and multivalent reagents have been shown to improve the apparent K_d and demonstrate wide applicability in recognizing the cognate target

in various biochemical and cell-based applications (*46-48*). The FHA-Fc proteins were overexpressed in HEK293F cells and purified by immobilized metal affinity chromatography. Regrettably, the protein yields were very low (10-15 times lower than the monovalent FHA reagents expressed in *E. coli*) and the ELISA signal for binding to the pT peptide was even weaker than before (data not shown). Since the original FHA scaffold was codon optimized for expression in *E. coli*, it is possible that the protein is poorly expressed in HEK-293 cells, folding improperly, or is glycosylated. For this reason, I plan to fuse the FHA reagents either to a leucine zipper dimerization domain (*48, 49*) or a cartilage oligomeric matrix protein (COMP) pentamerization domain (*46, 50*), overexpress the constructs in bacteria, and repeat the Western blotting experiments.

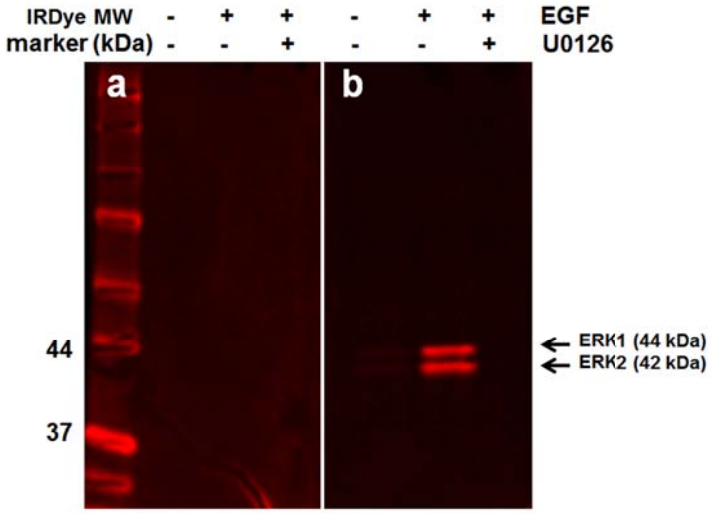


Figure 10. Western blotting of NIH/3T3 cell lysates. NIH/3T3 cell lysates (50 µg) from three treatments (before and after EGF stimulation and EGF stimulation after treatment with U0126), were loaded on an SDS-PAGE gel, transferred to a nitrocellulose membrane, and probed with anti-phospho ERK1/2 FHA reagent A5 (biotin conjugated) and detected using Streptavidin IRDye® 700DX conjugated (panel a) or with an IgG that recognizes the same dual phosphorylated ERK1/2 phosphoepitope and detected using Goat anti-Rabbit IRDye® 680 (panel b).

4.4.6 Cell staining of endogenous phospho-ERK1 and ERK2 proteins

NIH/3T3 cells were cultured on poly-L-Lysine coated coverslips for 20 h in complete medium and then serum starved for 24 h. The cells were fixed with 4% formaldehyde, either before or after stimulation with EGF or after treatment with MEK1/2 inhibitor (U0126) followed by EGF stimulation. The dual phosphorylated ERK1/2 proteins were detected using a commercially available IgG developed in

rabbit (as a positive control) and a FHA affinity reagent, A5, generated against the dual phosphorylated ERK1/2 peptide.

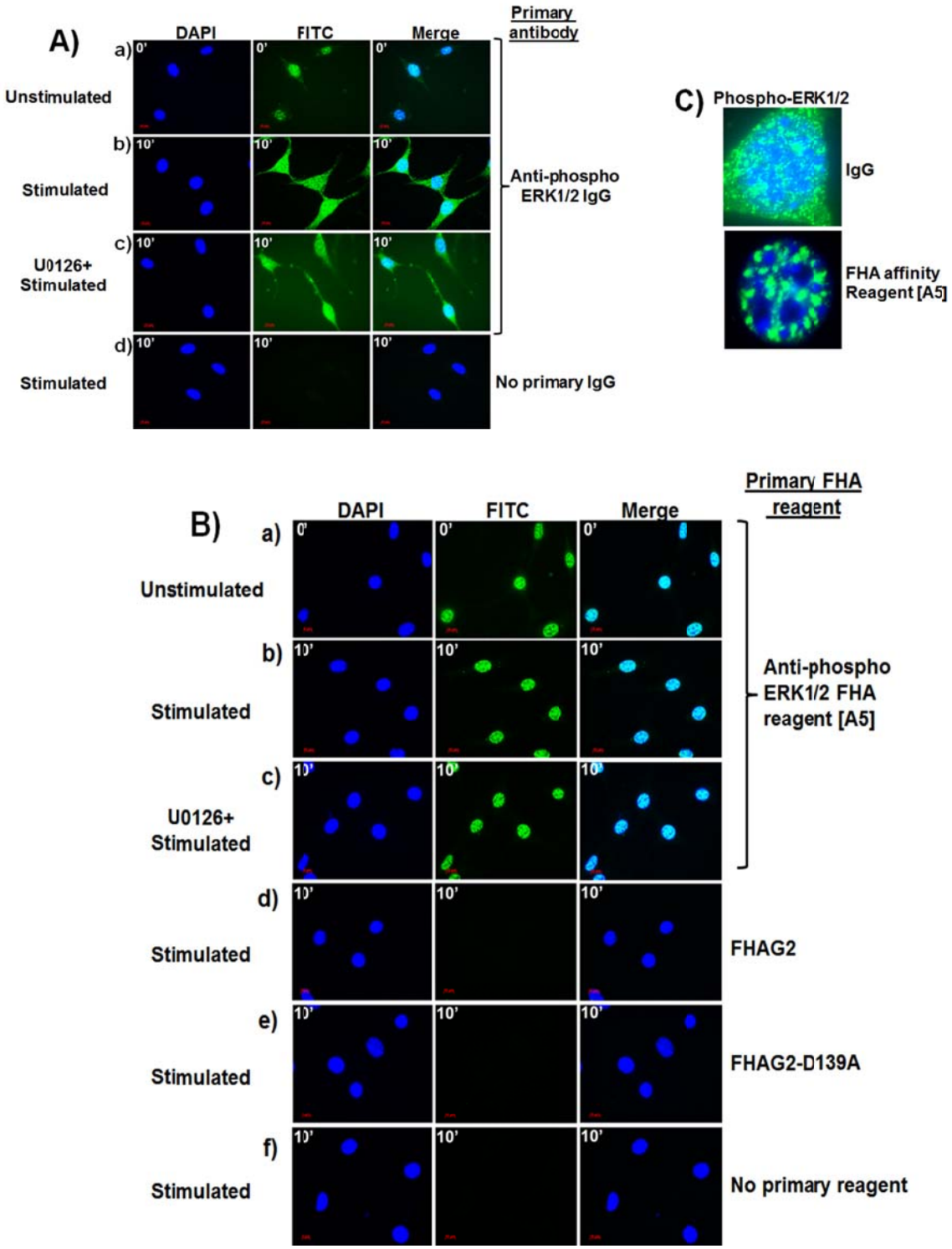
Cell-staining with anti-phospho ERK1/2 IgG

Staining with the IgG (Fig. 11A) showed an increased level of phospho-ERK1/2 proteins upon stimulation with EGF (panel b) compared to unstimulated cells (panel a). The level of phosphorylation was similar to unstimulated cells when the cells were treated with MEK1/2 inhibitor (U0126) for 2 h prior to stimulation with EGF (panel c). U0126 is a non-competitive inhibitor of MEK1/2 Ser/Thr protein kinase (*43*), which is the upstream kinase responsible for phosphorylating ERK1/2 proteins (*44*). Therefore, when MEK1/2 kinases are inactivated, they do not phosphorylate ERK1/2 protein kinases and the cell staining signal should be reduced to background. In panel d, cells were staining only with the Alexa Fluor 488 labeled secondary antibody (negative control). Prior research has shown that phospho-ERK1/2 proteins are almost exclusively localized to the nucleus after stimulation of Swiss 3T3 cells with EGF (*51*). Dual phosphorylation is essential for activation of ERK1 and ERK2 and for its translocation into the nucleus (*52*).

Cell-staining with anti-phospho FHA affinity reagent [A5]

Staining endogenous phospho-ERK1/2 proteins with an FHA affinity reagent A5 (Fig. 11B) showed nuclear staining without or with EGF stimulation (panels a and b), and this staining pattern did not change when the cells were treated with the U0126 inhibitor (panel c). Non-specific staining was not observed, when the cells were stained with the original FHA1G2 domain, which was used as a scaffold for library construction (panel d) or with a point mutant (D139A) of the FHA1G2 domain (panels e), indicating that the FHA reagent A5 is recognizing a protein that cannot be detected with the original starting scaffold. We have shown that we have changed the binding specificity of the original FHA1 variant.

Figure 11. Staining dual phosphorylated ERK1/2 proteins. A) Cell-staining with anti-phospho ERK1/2 IgG. NIH/3T3 cells were stained with anti-phospho ERK1/2 IgG a) before EGF stimulation (0'), b) after stimulating with EGF for 10 min (10'), and c) after pre-treating the cells with MEK1/2 inhibitor (U0126) following stimulation with EGF for 10 min. d) Cells were not stained with the primary IgG. The secondary antibody used for detection is Alexa Fluor® 488 goat anti-rabbit IgG. B) Cell-staining with anti-phospho FHA affinity reagent [A5]. NIH/3T3 cells were stained with FHA reagent [A5] a) before EGF stimulation (0'), b) after stimulating with EGF for 10 min (10'), and c) after pre-treating the cells with MEK1/2 inhibitor (U0126) following stimulation with EGF for 10 min. Cells were stimulated with EGF for 10 min and stained with d) FHA1G2 variant, which was the original scaffold used for library construction and e) a point mutant (D139A) of the FHA1G2 variant that was unable to recognize its cognate pT-peptide. f) Cells were not stained with any FHA reagent. All the FHA reagents are biotinylated and the secondary antibody used for detection is Streptavidin, Alexa Fluor® 488 conjugate. C) Comparing the punctate nuclear staining patterns obtained using the anti-phospho ERK1/2 IgG and the FHA affinity reagent A5.


A)
DAPI
FITC
Merge
Primary antibody
a)
Unstimulated
b)
Stimulated
c)
U0126+ Stimulated
d)
Stimulated
0'
10'
Anti-phospho ERK1/2 IgG
No primary IgG
C)
Phospho-ERK1/2
IgG
FHA affinity Reagent [A5]
B)
DAPI
FITC
Merge
Primary FHA reagent
a)
Unstimulated
b)
Stimulated
c)
U0126+ Stimulated
d)
Stimulated
e)
Stimulated
f)
Stimulated
0'
10'
Anti-phospho ERK1/2 FHA reagent [A5]
FHAG2
FHAG2-D139A
No primary reagent

Punctate staining was observed in the nucleus while staining with both the anti-phospho ERK1/2 IgG and the FHA affinity reagent A5 (Fig. 11C), but no staining was observed in the cytoplasm. To confirm my cell staining result, two different cell lines were used, NIH/3T3 and mIMCD3, two different growth hormones were used as stimulants (EGF and PDGF), different fixation methods, such as using 100% methanol, 4% formaldehyde and a combination of both, were used, and different formats of the FHA affinity reagents (such as GFP fusions and biotinylated versions) were tested. All methods gave the same staining pattern, indicating that it is a reproducible staining result. Since staining was not observed in the cytoplasm, and the two FHA proteins used as negative controls (FHAG2 and its D139A point mutant) did not stain the cells, either in the cytoplasm or in the nucleus, we believe that it is specific nuclear staining. However, the identity of the protein(s) responsible for binding the FHA domain in the cell is unknown.

One approach to prove the identity of the protein is to perform cell staining in an ERK1/2 gene knock out cell line or use siRNA to down-regulate the expression of either MEK1/2 or ERK1/2 proteins (*53, 54*). An alternate method is to perform a co-staining experiment using an IgG that binds to ERK1/2 protein and the A5 affinity reagent, each labeled with a fluorescent probe having a different emission wavelength. These approaches will answer whether the reagent is staining ERK1/2 or not. If it is ERK1/2, the second question is to figure

out if the reagent is phosphospecific. For this purpose, we propose to generate a mutant of the A5 affinity reagent in which Arg70, which is important for pT recognition (*26*) is mutated to Ala. If this R70A mutant does not stain the nucleus, then we can say with confidence that the reagent is phosphospecific. On the other hand, if both the A5 reagent and its R70A mutant show similar staining patterns, then we know that the persistent staining may be due to binding with the non-phosphorylated form of ERK1/2.

However, if the reagent shows nuclear staining in an ERK1/2 knockout cell line, then it is cross-reacting with some other protein. How do we identify this protein? We recommend two strategies: 1) Using peptide or protein arrays and 2) Pull-down experiment and mass spectrometry. To identify if the A5 reagent is cross-reacting with any of the closely related members in the ERK-MAPK family, an array can be probed to identify the cross-reacting proteins (Human phospho-MAPK array kit, R & D systems, Inc., McKinley Place NE, MN), similar to the work done previously to confirm the specificity of a monobody to Fyn SH3 domain by probing an array spotted with 150 human SH3 domains (*55*) and to characterize the specificity of scFvs specific for four SH2 domains by probing an array spotted with 432 proteins (*56*). On the other hand, phosphopeptide arrays with hundreds of phosphopeptides of known sequences can be custom synthesized on a cellulose membrane and probed with the A5 affinity reagent, as has been previously done to define the binding specificity and motifs of SH2

domains (*57*), BRCT domains (*58*) and Polo-box domain (*59*). Using a second strategy, the proteins can be pulled down from the nuclear extracts of mammalian cell lysates using the A5 reagent conjugated to magnetic beads and the proteins can be identified by mass spectrometry. It should be noted that pull-down experiments also depend upon the accessibility of the epitope, the concentration of the endogenous phosphorylated proteins, and on the affinity of the recombinant reagent. Recombinant affinity reagents usually do not bind as tightly compared to antibodies obtained from animal immunizations, which have undergone *in vivo* affinity maturating resulting in picomolar binding. Recombinant affinity reagents are amenable to *in vitro* directed evolution and various reagents have been affinity matured *in vitro* to obtain K_d values similar to or even better than antibodies isolated by animal immunizations (*60, 61*) which is described in more detail in the Conclusions chapter.

Since the FHA domain is a novel scaffold, we are in the process of improving the affinity reagents derived from this protein domain to make them widely applicable for use in various formats. We have proved the most important proposal of this project that it is possible to generate new anti-phosphopeptide binding specificities using the FHA1 domain as a scaffold. To this end, we have been able to show specific binding of these affinity reagents to their cognate phosphopeptides in phage ELISA. However, we were not able to show specific staining of the reagents in cell staining and Western blotting. For such

applications, directed evolution of these reagents is essential to increase their affinity and specificity for their cognate targets.

4.5 Conclusion

In conclusion, we have engineered the specificity of the FHA1 domain to recognize various phosphothreonine peptides, for which binding was not previously detected. The phage-displayed library of FHA1 variants is a useful resource to isolate renewable anti-phosphospecific reagents, *in vitro,* by phage-display. All the FHA1 variants bind to their cognate pT peptide, with little or no binding to the original Rad9-pT peptide. All the interactions are pT dependent and no binding was detected for the non-phosphorylated forms of any of the peptides in phage ELISA. Our success rate with isolating affinity reagents for various pT containing peptides from these phage-displayed libraries of FHA1 variants is ~60%. For the first time, a pT-binding domain has been engineered for displaying its functional variant on the surface of bacteriophage M13 and using it as a potential scaffold for generating affinity reagents to various pT peptides. This strategy serves as a platform for exploiting other phosphopeptide-binding domains as scaffolds for generating recombinant phosphospecific affinity reagents. These reagents will be an attractive alternative for antibodies, because they are renewable, have excellent expression in *E. coli* (20-25 mg/L), can be

generated in a short time (<2 weeks), and are cost effective compared to immunizing animals. In the next chapter, we discuss strategies for the directed evolution of affinity reagents and some other phosphopeptide-binding domains that can be used as potential scaffolds in the future.

4.6 References

1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of the human genome, *Science 298*, 1912-1934.

2. Cohen, P. (2001) The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture, *Eur J Biochem 268*, 5001-5010.

3. Sefton, B. M., and Shenolikar, S. (2001) Overview of protein phosphorylation, *Curr Protoc Protein Sci Chapter 13*, Unit13 11.

4. Yaffe, M. B., and Elia, A. E. (2001) Phosphoserine/threonine-binding domains, *Curr Opin Cell Biol 13*, 131-138.

5. Schlessinger, J., and Lemmon, M. A. (2003) SH2 and PTB domains in tyrosine kinase signaling, *Sci STKE 2003*, RE12.

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

7. Muslin, A. J., Tanner, J. W., Allen, P. M., and Shaw, A. S. (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine, *Cell 84*, 889-897.

8. Bridges, D., and Moorhead, G. B. (2005) 14-3-3 proteins: a number of functions for a numbered protein, *Sci STKE 2005*, re10.

9. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) Function of WW domains as phosphoserine- or phosphothreonine-binding modules, *Science 283*, 1325-1328.

10. Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) The FHA domain mediates phosphoprotein interactions, *J Cell Sci 113 Pt 23*, 4143-4149.

11. Tsai, M. D. (2002) FHA: a signal transduction domain with diverse specificity and function, *Structure 10*, 887-888.

12. Durocher, D., and Jackson, S. P. (2002) The FHA domain, *FEBS Lett 513*, 58-66.

13. Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., and Harper, J. W. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex, *Cell 91*, 209-219.

14. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains, *Science 300*, 445-452.

15. Mohammad, D. H., and Yaffe, M. B. (2009) 14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response, *DNA Repair (Amst) 8*, 1009-1017.

16. Paradela, A., and Albar, J. P. (2008) Advances in the analysis of protein phosphorylation, *J Proteome Res 7*, 1809-1818.

17. Rosenqvist, H., Ye, J., and Jensen, O. N. (2011) Analytical strategies in mass spectrometry-based phosphoproteomics, *Methods Mol Biol 753*, 183-213.

18. Kalume, D. E., Molina, H., and Pandey, A. (2003) Tackling the phosphoproteome: tools and strategies, *Curr Opin Chem Biol 7*, 64-69.

19. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell 127*, 635-648.

20. Sun, T., Campbell, M., Gordon, W., and Arlinghaus, R. B. (2001) Preparation and application of antibodies to phosphoamino acid sequences, *Biopolymers 60*, 61-75.

21. Bangalore, L., Tanner, A. J., Laudano, A. P., and Stern, D. F. (1992) Antiserum raised against a synthetic phosphotyrosine-containing peptide selectively recognizes p185neu/erbB-2 and the epidermal growth factor receptor, *Proc Natl Acad Sci U S A 89*, 11637-11641.

22. Gebauer, M., and Skerra, A. (2009) Engineered protein scaffolds as next-generation antibody therapeutics, *Curr Opin Chem Biol 13*, 245-255.

23. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A. (2000) Phage display for selection of novel binding peptides, *Methods in enzymology 328*, 333-363.

24. Pershad, K., and Kay, B. (Accepted for publication) Generating Thermally Stable Variants of Protein Domains through Phage-display *Methods Journal*.

25. Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

26. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms, *Mol Cell 6*, 1169-1182.

27. Howarth, M., and Ting, A. Y. (2008) Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin, *Nat Protoc 3*, 534-545.

28. O'Callaghan C, A., Byford, M. F., Wyer, J. R., Willcox, B. E., Jakobsen, B. K., McMichael, A. J., and Bell, J. I. (1999) BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation, *Anal Biochem 266*, 9-15.

29. Kay, B. K., Thai, S., and Volgina, V. V. (2009) High-throughput biotinylation of proteins, *Methods Mol Biol 498*, 185-196.

30. Mackeigan, J. P., Murphy, L. O., Dimitri, C. A., and Blenis, J. (2005) Graded mitogen-activated protein kinase activity precedes switch-like c-Fos induction in mammalian cells, *Mol Cell Biol 25*, 4676-4682.

31. Ernst, A., Sazinsky, S. L., Hui, S., Currell, B., Dharsee, M., Seshagiri, S., Bader, G. D., and Sidhu, S. S. (2009) Rapid evolution of functional complexity in a domain family, *Science signaling 2*, ra50.

32. Scholle, M. D., Kehoe, J. W., and Kay, B. K. (2005) Efficient construction of a large collection of phage-displayed combinatorial peptide libraries, *Comb Chem High Throughput Screen 8*, 545-551.

33. Liao, H., Yuan, C., Su, M. I., Yongkiettrakul, S., Qin, D., Li, H., Byeon, I. J., Pei, D., and Tsai, M. D. (2000) Structure of the FHA1 domain of yeast Rad53 and identification of binding sites for both FHA1 and its target protein Rad9, *J Mol Biol 304*, 941-951.

34. Raman, M., Chen, W., and Cobb, M. H. (2007) Differential regulation and properties of MAPKs, *Oncogene 26*, 3100-3112.

35. Yoon, S., and Seger, R. (2006) The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions, *Growth Factors 24*, 21-44.

36. Liu, S., Sun, J. P., Zhou, B., and Zhang, Z. Y. (2006) Structural basis of docking interactions between ERK2 and MAP kinase phosphatase 3, *Proc Natl Acad Sci U S A 103*, 5326-5331.

37. Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E. S., Wu, M. R., Su, M. I., Heierhorst, J., and Tsai, M. D. (2008) Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade, *Mol Cell 30*, 767-778.

38. Yongkiettrakul, S., Byeon, I. J., and Tsai, M. D. (2004) The ligand specificity of yeast Rad53 FHA domains at the +3 position is determined by nonconserved residues, *Biochemistry 43*, 3862-3869.

39. Byeon, I. J., Yongkiettrakul, S., and Tsai, M. D. (2001) Solution structure of the yeast Rad53 FHA2 complexed with a phosphothreonine peptide pTXXL: comparison with the structures of FHA2-pYXL and FHA1-pTXXD complexes, *J Mol Biol 314*, 577-588.

40. Zoumpourlis, V., Papassava, P., Linardopoulos, S., Gillespie, D., Balmain, A., and Pintzas, A. (2000) High levels of phosphorylated c-Jun, Fra-1, Fra-2 and ATF-2 proteins correlate with malignant phenotypes in the multistage mouse skin carcinogenesis model, *Oncogene 19*, 4011-4021.

41. Gozdecka, M., and Breitwieser, W. (2012) The roles of ATF2 (activating transcription factor 2) in tumorigenesis, *Biochem Soc Trans 40*, 230-234.

42. Durocher, D., Henckel, J., Fersht, A. R., and Jackson, S. P. (1999) The FHA domain is a modular phosphopeptide recognition motif, *Mol Cell 4*, 387-394.

43. Favata, M. F., Horiuchi, K. Y., Manos, E. J., Daulerio, A. J., Stradley, D. A., Feeser, W. S., Van Dyk, D. E., Pitts, W. J., Earl, R. A., Hobbs, F., Copeland, R. A., Magolda, R. L., Scherle, P. A., and Trzaskos, J. M. (1998) Identification of a novel inhibitor of mitogen-activated protein kinase kinase, *J Biol Chem 273*, 18623-18632.

44. Crews, C. M., Alessandrini, A., and Erikson, R. L. (1992) The primary structure of MEK, a protein kinase that phosphorylates the ERK gene product, *Science 258*, 478-480.

45. Dyson, M. R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B. K., Pawson, T., and McCafferty, J. (2011) Mapping protein interactions by combining antibody affinity maturation and mass spectrometry, *Anal Biochem 417*, 25-35.

46. Duan, J., Wu, J., Valencia, C. A., and Liu, R. (2007) Fibronectin type III domain based monobody with high avidity, *Biochemistry 46*, 12656-12664.

47. Powers, D. B., Amersdorfer, P., Poul, M., Nielsen, U. B., Shalaby, M. R., Adams, G. P., Weiner, L. M., and Marks, J. D. (2001) Expression of single-chain Fv-Fc fusions in Pichia pastoris, *J Immunol Methods 251*, 123-135.

48. de Kruif, J., and Logtenberg, T. (1996) Leucine zipper dimerized bivalent and bispecific scFv antibodies from a semi-synthetic antibody phage display library, *J Biol Chem 271*, 7630-7634.

49. Weber-Bornhauser, S., Eggenberger, J., Jelesarov, I., Bernard, A., Berger, C., and Bosshard, H. R. (1998) Thermodynamics and kinetics of the reaction of a single-chain antibody fragment (scFv) with the leucine zipper domain of transcription factor GCN4, *Biochemistry 37*, 13011-13020.

50. Terskikh, A. V., Le Doussal, J. M., Crameri, R., Fisch, I., Mach, J. P., and Kajava, A. V. (1997) "Peptabody": a new type of high avidity binding protein, *Proc Natl Acad Sci U S A 94*, 1663-1668.

51. Murphy, L. O., Smith, S., Chen, R. H., Fingar, D. C., and Blenis, J. (2002) Molecular interpretation of ERK signal duration by immediate early gene products, *Nat Cell Biol 4*, 556-564.

52. Khokhlatchev, A. V., Canagarajah, B., Wilsbacher, J., Robinson, M., Atkinson, M., Goldsmith, E., and Cobb, M. H. (1998) Phosphorylation of the MAP kinase ERK2 promotes its homodimerization and nuclear translocation, *Cell 93*, 605-615.

53. Gailhouste, L., Ezan, F., Bessard, A., Fremin, C., Rageul, J., Langouet, S., and Baffet, G. (2010) RNAi-mediated MEK1 knock-down prevents ERK1/2

activation and abolishes human hepatocarcinoma growth in vitro and in vivo, *Int J Cancer 126*, 1367-1377.

54. Bessard, A., Fremin, C., Ezan, F., Fautrel, A., Gailhouste, L., and Baffet, G. (2008) RNAi-mediated ERK2 knockdown inhibits growth of tumor cells in vitro and in vivo, *Oncogene 27*, 5315-5325.

55. Huang, R., Fang, P., and Kay, B. K. (2012) Isolation of monobodies that bind specifically to the SH3 domain of the Fyn tyrosine protein kinase, *N Biotechnol 29*, 526-533.

56. Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display, *Protein Eng Des Sel 23*, 279-288.

57. Huang, H., Li, L., Wu, C., Schibli, D., Colwill, K., Ma, S., Li, C., Roy, P., Ho, K., Songyang, Z., Pawson, T., Gao, Y., and Li, S. S. (2008) Defining the specificity space of the human SRC homology 2 domain, *Mol Cell Proteomics 7*, 768-784.

58. Rodriguez, M., Yu, X., Chen, J., and Songyang, Z. (2003) Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains, *J Biol Chem 278*, 52914-52918.

59. Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain, *Cell 115*, 83-95.

60. Schier, R., McCall, A., Adams, G. P., Marshall, K. W., Merritt, H., Yim, M., Crawford, R. S., Weiner, L. M., Marks, C., and Marks, J. D. (1996) Isolation of picomolar affinity anti-c-erbB-2 single-chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody binding site, *J Mol Biol 263*, 551-567.

61. Hufton, S. E. (2012) Affinity maturation and functional dissection of a humanised anti-RAGE monoclonal antibody by ribosome display, *Methods Mol Biol 805*, 403-422.

CHAPTER 5

CONCLUSIONS

5. CONCLUSIONS

5.1 Evaluating the specificity of affinity reagents

Recombinant techniques have made it possible to generate affinity reagents to hundreds or thousands of proteins in the human proteome (*1-4*). However, an equally important challenge is to validate these reagents for binding specificity and provide a pipeline of well characterized affinity reagents that are useful to the scientific community for using as reagents in various biochemical and cell-based applications. The Human Protein Atlas (HPA) project currently collects immunohistochemistry, immunofluorescence, Western blot, and protein array data about a large variety of monoclonal, polyclonal, and recombinant antibodies that recognize human proteins (www.proteinatlas.org) (*5-8*).

One very useful method for evaluating the specificity of an affinity reagent is to probe large numbers of proteins that have been spotted into arrays. While a reagent might seem specific when screened against a few proteins on the laboratory bench, it needs to be tested against a much large set of potential targets. For example, when a number of monoclonal and polyclonal antibodies, which had been generated to specific yeast proteins, were used to probe an array of ~5000 yeast proteins, cross-reactivity was observed (*9*). The array results were confirmed in Western blots: many of the cross-reactive proteins in the arrays were confirmed on the blots, but some of them were not, either due to

the low protein concentration in the cell lysate or due to the loss of conformational epitopes during Western blotting.

The recent availability of protein and antibody arrays has made it possible to validate the affinity reagents for cross-reactivity in a more realistic scenario. Using chips spotted with all the proteins encoded in an organism's genome is a very useful resource for mapping protein-protein interactions (*10*). The specificity of a monobody binding to the Src homology 3 (SH3) domain of Fyn tyrosine kinase was validated by probing an array onto which 150 human SH3 domains were spotted, which comprises of half the total number of SH3 domains in the human proteome (*11*). In an international effort, 398 affinity reagents (polyclonal or monoclonal antibodies or recombinant single-chain variable fragments), which had been generated against various human Src homology 2 (SH2) domains, were validated for specificity by probing 432 proteins with each one (*12*). Of these, 90 scFvs were provided by Dr. John McCafferty at University of Cambridge, with whom we collaborated to generate a panel of scFvs to 20 human SH2 domains (*13*). Of the 398 affinity reagents tested for specificity, about 50% showed specific binding for their cognate target and 10% were highly cross-reactive. In parallel, Dr. John McCafferty's group tested the specificity of scFvs isolated against four SH2 domains from NCK1, SYKN, RASA1_C and NCK1 proteins, by using the same method (*13*) and observed that all the scFvs recognized primarily their cognate antigens with little/no cross-reactivity to

unrelated proteins. Similar to monoclonal and polyclonal antibodies generated by animal immunizations, recombinant antibody fragments generated by phage display technology show specific recognition of their cognate targets and offer an attractive alternative to generating affinity reagents *in vitro* by recombinant techniques.

Another useful technique to confirm the specificity of affinity reagents is the use of cell lines in which the gene of interest is knocked out by genetic engineering (*14, 15*) or the expression of the protein is down-regulated with small interfering RNAs (siRNAs) (*16, 17*). To confirm that an antibody is binding to its cognate target, such engineered cells should yield no signal in immuno-fluorescence experiments or Western blotting experiments, if the affinity reagent is specific. A high-throughput immunofluorescence assay, in which images are captured with a confocal microscope, was used to validate the specificity of 75 polyclonal antibodies, which had been generated against 65 proteins that were silenced using 130 siRNAs (*18*). Using this approach, it was observed that 80% of the antibodies stained the correct part of the cell and showed a decrease in fluorescence intensity when cells were transfected with siRNA. The cross-reactive antibodies staining other parts of the cell in addition to their cognate target showed a persistent staining even in siRNA transfected cells. Similar approaches have also been used to validate the specificity of antibodies for specific targets (*19, 20*).

5.2 *In vitro* affinity maturation

While it is generally desired that an affinity reagent binds selectively to its cognate target, on occasion, the desired affinity of a binder is inadequate for a particular biochemical application. This is generally a consequence due to the short half-life of the antigen bound to its affinity reagent. For example, assuming a k_{on} rate of 10^6 $M^{-1}s^{-1}$, complexes with dissociation rate constants of 1 μM, 100 nM, and 1 nM will have half-lives of 1 sec, 10 sec, and 16.7 min, respectively (*21*). Thus, for pull-down experiments (followed by Western blotting or mass spectrometry), in which there can be several lengthy wash steps to reduce non-specific binding of thousands of abundant cellular proteins, it is essential that the affinity reagents have the tightest binding possible.

Antibody fragments and other affinity reagents isolated from naive (non-immunized) or synthetic libraries without the involvement of any affinity maturation usually have dissociation constants (K_ds) in the high nanomolar to micromolar range (*22*). Over the years, it has been a common practice to generate secondary libraries by random mutagenesis of the original binding clone, followed by selecting for variants with desired properties, as depicted in Fig. 1. Different types of affinity reagents isolated from either phage, ribosome or yeast displayed libraries have been further affinity matured *in vitro* by using random mutagenesis or site-directed mutagenesis to isolate high affinity binders with affinities comparable to antibodies isolated after animal immunizations (*23,*

24). In many respects, directed evolution mimics the *in vivo* immune response of B-cells for generating affinity-matured antibodies. The opportunity to affinity mature recombinant affinity reagents *in vitro* renders these display technologies as powerful alternatives to generating antibodies by animal immunization.

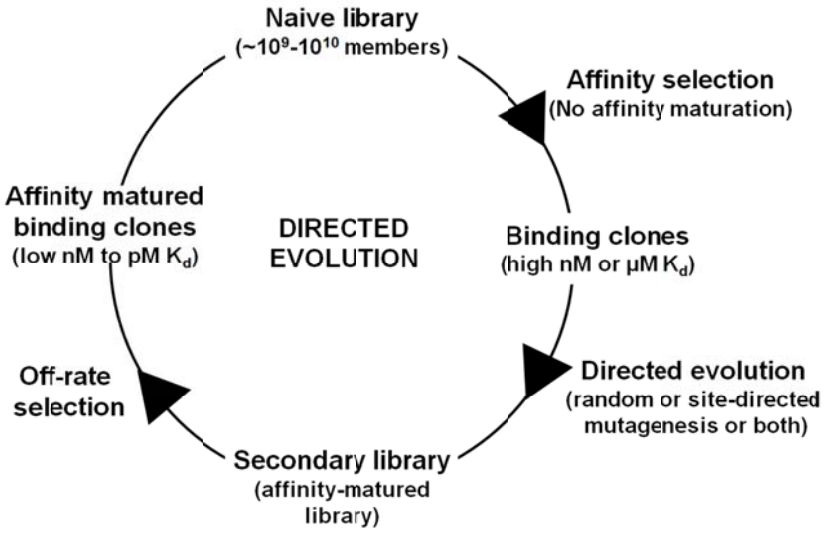


Figure 1. Directed evolution of affinity reagents *in vitro* by random mutagenesis. Affinity reagents specific for a target protein are isolated after 2-3 rounds of affinity selection from a library containing billions of variants. To improve the affinity of the best binding clones, they are subjected to random or site-directed mutagenesis, and variants are selected from the secondary libraries that bind tighter by performing an off-rate selection.

Two methods of random mutagenesis are most commonly employed to improve the affinity of binding clones: 1) chain shuffling and 2) error-prone PCR. While performing chain shuffling (*25, 26*), the variable region of heavy chain (V_H) of a single binding clone is allowed to recombine with a library of light chain variable fragments (V_L), or vice versa, and new V_H-V_L pairs are tested for variants with improved properties. This tactic works well, as I used it to improve the affinity of scFvs binding to 8 different Src Homology 2 (SH2) domains (*27*). In this study, we observed that for each target, a greater number of binding clones were isolated from the chain shuffled libraries that gave a high signal in phage ELISA. The anti-SHC1 scFvs with low nM K_ds values were able to immunoprecipitate SHC1 from lysates of breast cancer epithelial cells, and these reagents worked equally well in Western blotting. It was a very important observation that scFvs having low nM k_ds values of <60 nM worked successfully in immunoprecipitation experiments. This, however, is also dependent on the accessibility of the epitope and the endogenous protein concentration. Using this approach, the affinity of scFvs binding to human epidermal growth factor receptor 2, a hapten (2-phenyloxazol-5-one), and Hepatitis B virus coat protein were improved 5-fold, 6-fold and 300-fold, respectively by shuffling their V_H with a repertoire of V_L and *vice versa*, and screening for improved affinity (*28-30*). In addition to using chain shuffling to improve the affinity of scFvs, it has been possible to improve the levels of expression and thermal stability of scFvs (*31*). The expression and

thermal stability of three human G-protein coupled receptors have also improved by directed evolution (*32*).

When working with protein scaffolds, one can use error-prone PCR to generate a secondary library from which variants with desired properties are affinity selected under stringent conditions. In error-prone PCR (*33, 34*), mutations are randomly generated across the entire coding sequence; this method was used to improve the affinity of various affinity reagents, such as, an FN3 monobody binding to Fyn tyrosine kinase (*11*), designed ankyrin repeat proteins specific for cell-surface antigens and receptors (*35, 36*), and single-chain variable fragments isolated against tumor antigens, neurotoxins and peptides (*37-39*). Randomly mutating the complementarity determining regions of antibody fragments binding to cancer-associated antigen and a growth factor led to an affinity improvement of 20 to 100-fold (*40, 41*). By combining error-prone PCR and chain shuffling or by directed mutagenesis along with random mutagenesis, tight binding clones have been isolated (*37, 38, 42*).

Ribosome display technology includes an affinity maturation step during the affinity selection procedure, which involves an error-prone PCR on the entire pool of the selected population after each round of affinity selection. This favors the isolation of high affinity binders with picomolar or low nanomolar dissociation constants after 3-5 rounds of affinity selection (*43-45*). Affinity matured antibodies

have been shown to be more sensitive in recognizing their cognate targets, in various formats.

In addition to improving the affinity, error-prone PCR was used to improve the phage-display efficiency of Fragment of antigen binding (Fab) (*46*), *E. coli* adenylate cyclase and *Bordetella pertussis* (*47*), and to enhance the thermal stability of the Forkhead-associated domain (*48*). We have shown that by error-prone PCR in the Forkhead-associated domain coding sequence, we were able to isolate variants that were functional when displayed on the surface of bacteriophage M13 and variants that were more thermal stable compared to the wild-type domain (*48*).

In conclusion, random mutagenesis has been used to improve the affinity, specificity, thermal stability, expression, functional phage-display and phage-display efficiencies of different types of affinity reagents. In the future, to generate high affinity reagents to various phosphopeptides, we plan to make the overall selection process more stringent by introducing random mutations in the pool of clones after the second round of selection, by error-prone PCR, followed by a stringent third round of selection with off-rate selection to isolate high affinity binding clones. The final affinity reagents can be fused to a multimerization domain, to generate multivalent reagents with even higher apparent affinity. We predict that by using these strategies, we will be able to isolate affinity reagents with low nanomolar binding dissociation constants and these reagents can be

used for various applications because of their slower off-rate. All the reagents will first be tested in Western blotting and the ones that look specific can be further tested for other applications.

5.3 Improving affinity by generating multimeric affinity reagents

If a target is present at a local high concentration, or exists as a multimer in solution, then it is possible to produce a more effective affinity reagent by generating multivalent forms of the affinity reagents. Such an improvement is based on the phenomenon of chelation or avidity. There are many examples in nature of avidity playing an important role in biological processes, such as the pentavalent immunoglobulin IgM that binds to the repetitive cell surface antigens of bacteria and viruses (*49*). Similarly, due to avidity, the affinity of complement factor C1q is increased 100-fold when it binds to a number of IgG molecules present in an immune complex (*50*).

Various strategies have been adopted to generate bivalent, trivalent or multivalent affinity reagents. One common approach is to fuse them to domains that self dimerize or form multimers. For instance, bivalent affinity reagents are generated by fusion via a flexible linker to the Fragment Crystallization (Fc) portion of an IgG (Fig. 2A) (*51, 52*) or to the leucine zipper dimerization domain of Fos or Jun proteins or yeast GCN4 transcription factor (Fig. 2B) (*53, 54*). Pentameric fibronectin type III domain (FN3) based monobody was generated by

fusing it via a 45 amino acid linker to the pentamerization domain of cartilage oligomeric matrix protein (COMP) (Fig. 2C). The pentameric monobody had an increased binding affinity of ~460-fold, compared to the monomeric monobody, and detected its cognate target at much lower concentrations in ELISAs and dot blots. The pentameric reagent showed brighter cell staining of the cognate cell surface antigen at concentrations 200-fold lower than the monomer (*55*). Similarly, a peptide-COMP fusion had a K_d value of 1 nM, compared to the original peptide with a very weak K_d value of 200 μM, to its cell surface target. Moreover, the pentameric peptide was more effective in activating cells than that used for a monoclonal antibody to the same cell surface, indicating that the functional affinity of the pentameric peptide was greater than the affinity of the monoclonal IgG (*56*). Pentamers of single-domain antibodies have also been generated by fusing them to the verotoxin B-subunit of *E. coli* (*57*).

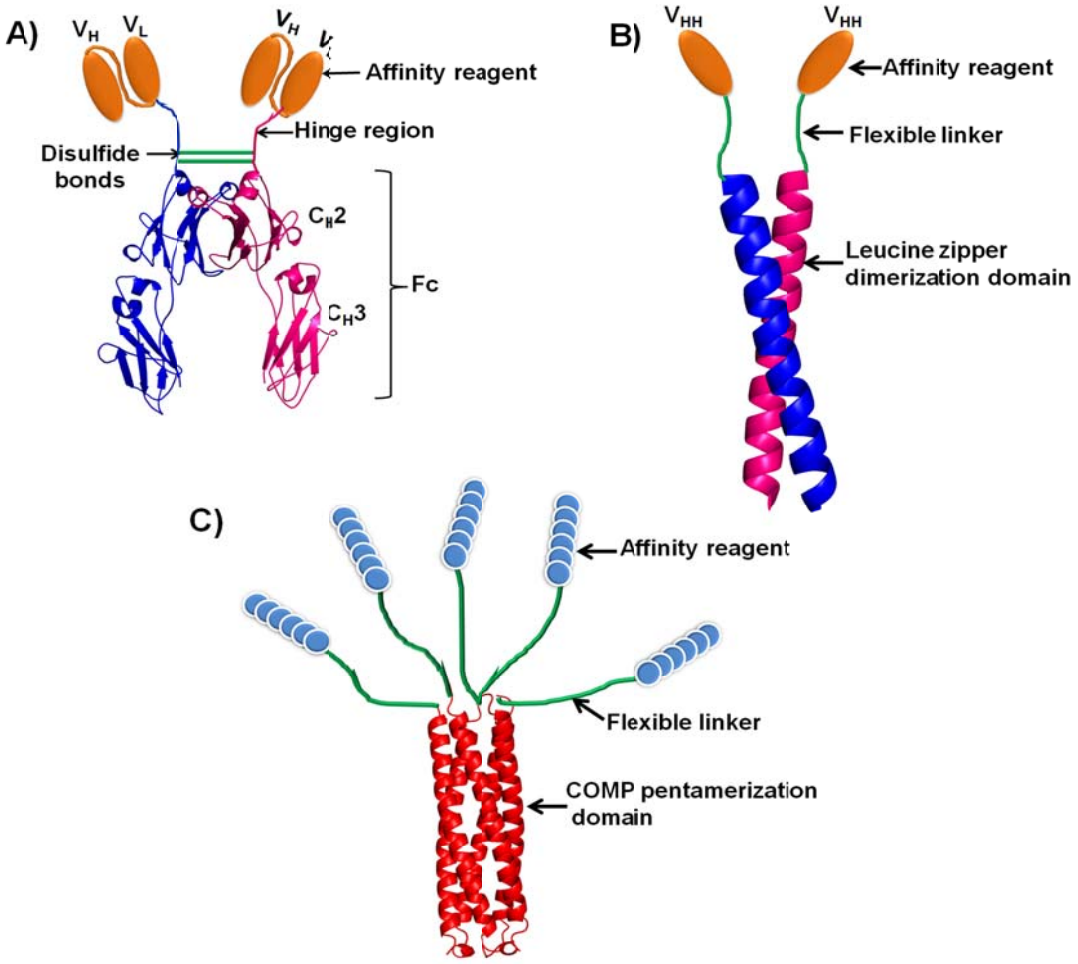


Figure 2. Dimerization and multimerization domains. A) The monomeric affinity reagent is fused to the Fc portion of an IgG molecule and the two polypeptides are connected by disulfide bonds to generate a bivalent affinity reagent. B) Leucine zippers are structural motifs found in transcription factors such as Fos, Jun and GCN4. They are rich in Leu residues (every 7th position is a Leu) and the two halves of a Leucine zipper zip together with strong hydrophobic interactions between the leucines in each half, dimerizing affinity reagents linked to them. C) COMP protein has a domain that forms homopentamers. Affinity reagents fused to this pentamerization domain via a long flexible linker produces pentavalent affinity reagents.

Table I summarizes a more detailed list of the various dimerization and oligomerization domains that have been commonly used to increase the valency of affinity reagents. From various studies, it has been shown that an increase in valency of the affinity reagents is associated with an increase in affinity by avidity effect and the multivalent reagents specifically detect their cognate targets in various applications such as western blotting (*51*), fluorescence activated cell sorting (FACS) assay (*52*), enzyme-linked immunosorbent assay (ELISA), flow cytometry, cell staining and immunohistochemistry (*53, 58*), and usually do did not require further chemical modification or fixation of the cells, which are necessary while using monovalent affinity reagents owing to their short half-life.

In conclusion, multivalent affinity reagents will work well when the target is present at a high density, such as in ELISA, western blotting and recognizing cell surface antigens that are present in a repetitive fashion. However, if this is not the case, then affinity maturation of the binding clone has to be performed followed by affinity selection under stringent conditions to isolate high affinity reagents that can then be used for various applications in which the target concentration is low.

Table I. List of affinity reagents and their fusion partners that promote dimerization or multimerization.

Affinity reagent	Fusion partner	Valency of the affinity reagent	Reference
scFv, affibody	Fc portion of an IgG	Bivalent	(*51, 52, 59, 60*)
scFv	Different linker lengths between the V_H and V_L 3-4 amino acids 1-2 amino acids 0 amino acids	Bivalent (Diabodies) Trimeric (Triabodies) Tetramers (Tetrabodies)	(*61-65*)
Di-diabody	C_H3 of an IgG	Bivalent diabodies	(*66*)
scFv, $F(ab)_2$, sdAb, DARPins	Leucine zippers	Bivalent	(*35, 53, 54, 67-71*)
FN3, peptide	COMP pentamerization domain	Pentamer	(*55, 56*)
sdAb	two sdAbs linked via a 29 amino acid peptide	Bivalent	(*72*)
sdAb	*E. coli* verotoxin B-subunit	Pentamer	(*57, 73, 74*)
Fab, scFv, FN3, peptides	Alkaline phosphatase	Bivalent	(*68, 75, 76*)

scFv, single-chain variable fragment; FN3, Fibronectin type III domain; Fab, fragment of antigen binding; DARPins, designed ankyrin repeat proteins; sdAb, single domain antibody.

5.4 Potential scaffolds for generating anti-phosphospecific reagents

In my thesis, I engineered new binding specificities from a naturally occurring pT-peptide binding domain, the N-terminal Forkhead-associated domain (FHA1) from *S. cerevisiae* Rad53 protein. My goal was to generate antibody-like 'affinity reagents' to various phosphoepitopes phage-display technology in place of immunizing animals. As a proof-of-principle, we selected the FHA1 domain, for four reasons: 1) the FHA1 domain specifically binds pT-containing peptides, 2) its binding is phosphorylation dependent interaction (*77-79*), 3) FHA1 domains can be overexpressed in *E. coli* (*77*), and 4) three-dimensional structures of the FHA1 domain have been solved with a pT-containing peptide bound to it (*77, 80*). These structural data provide valuable information regarding the residues in the FHA1 domain that are important for interaction with the phosphate group, the phosphopeptide backbone, and the side chain residues.

Based on published structural studies, supplemented with my alanine scanning data of 24 amino acid residues in the FHA1 domain, I constructed two phage-displayed libraries with diversities of $3x10^9$ and 10^{10} members. If the library design is optimal, I postulated that I should be able to isolate unique binding sequences from these libraries. In fact, I observed a 50% success rate of generating FHA1 domain variants that recognized different pT-peptides, which is very gratifying.

For downstream applications, the success of an affinity reagent is dependent on specificity and affinity for its target. At the moment, it is very challenging to generate highly specific affinity reagents to various phosphoepitopes because of the flexibility of the pT-peptide and the limited number of contacts between the pT-peptide and the affinity reagent. In practice, it is possible that an affinity reagent that works well in ELISA, may not work for western blotting, cell staining, immunoprecipitation, or FACS, because of difference in accessibility of a phosphoepitope in a native or denatured forms of the phosphorylated protein, and it is present in at a low concentration in a cell. Therefore, we think that certain FHA variants may need to be optimized, several FHA1 domain variants should be used for each target, or that other FHA scaffolds should be a source of affinity reagents.

It would be interesting to test two other FHA domains, the human Chk2 FHA domain and *S. cerevisiae* Dun1 FHA domain, as scaffolds for generating affinity reagents to phosphoepitopes. The human Chk2 FHA domain has a different ligand preference compared to the Rad53 FHA1 domain; it shows a strong preference for a hydrophobic amino acids at the pT (+3) position, in contrast to the Rad53 FHA1 domain that strongly prefers a negatively charged amino acid at the pT (+3) position. Chk2 FHA domain also contains a longer β4-β5 loop with an eight amino acid helical insertion (*81*). By constructing phage-

displayed libraries of Chk2 FHA variants containing β4-β5 loops with varying lengths, the range of ligands recognized can be potentially increased.

Another characteristic of the FHA domain of Dun1 protein from *S. cerevisiae*, which can be exploited with its use as a scaffold, is that it binds well to doubly phosphorylated peptides. The dissociation equilibrium constant (K_d) of Dun1 FHA domain for a diphosphorylated peptide (NI**pT**QP**pT**QQST), corresponding to the Rad53 protein, is ~0.3 μM. This tight binding observed for Dun1 FHA domain with the dual phosphorylated peptides is attributed to the presence of two phosphopeptide binding sites in its three dimensional structure (*82*). This structural information will be valuable for constructing a library of Dun1 variants that potentially can bind dual phosphorylated epitopes. Such dual-phosphorylated epitopes are present in many naturally occurring proteins, such as mitogen-activated protein kinase, transcription factors such as Myc and ATF2 (*83*).

While I have demonstrated the ability of generating affinity reagents that recognize phosphothreonine-containing peptides, it would be of great interest and value to the scientific community to generate recombinant affinity reagents to phosphoserine- and phosphotyrosine-containing sequences as well. To explore the ability to generate affinity reagents to pS-containing peptides, the BRCA1 carboxyl-terminal (BRCT) domains and Polo-box domain can be used as potential scaffolds. The structures of BRCT domains from different proteins have

been solved (*34, 84-86*). The BRCT domains of Breast Cancer Gene 1 (BRCA1) are present as tandem repeats and both the BRCT domains have been shown to interact with the pS-containing peptide from BACH protein (*87*), very much like the V_H and V_L domains of antibodies form the antigen binding pocket. The interaction of BRCT domains with BACH1 is phosphorylation dependent, as demonstrated by various methods, such as mutating the pS to alanine, treating the cell lysate (of cells transfected with a BACH1 expressing construct) with lambda protein phosphatase prior to pull-down, or including a competing pS-peptide sequence from BACH1, all of which disrupted the interaction between recombinant BRCT domains and BACH1 protein (*87*). Similar to FHA domains, BRCT domains also have diverse binding specificities, with a strong preference for a specific amino acid at the pS (+3) position (*88*); for instance, the BRCT domains of BRCA1, MDC1, and Rad9 prefer Phe, Tyr, and Ile, respectively. Interaction was not detected between BRCT domains and their target proteins or phosphopeptides in which the pS (+3) residue was mutated to alanine (*87, 88*). A good starting point would be the BRCT domains of BRCA1, as its ligand specificity and its crystal structure in complex with a pS peptide from BACH protein has been solved at a very high resolution (*89*). This structural information, combined with alanine-scanning of residues in the binding surface of the BRCT domains, will provide useful insights for designing a library of variants of the BRCT domain, from which reagents to pS-containing peptides can be isolated.

As two BRCT domains form the binding surface, we predict that the binding affinity will be higher than the binding of a single phosphopeptide-binding domain to its phosphorylated peptide. In an attempt to identify the interacting partners for the BRCT domains of BRCA1, it was shown by Western blotting and immunoprecipitation of nuclear cell lysates of HeLa cells that the BRCT domains bind a 130 kDa protein (*90*). Thus, the BRCT domains already have a tight binding for their target proteins and will serve as a good scaffold for isolating affinity reagents that will be useful in applications such as Western blotting and pull down experiments.

Another phosphopeptide-binding domain that can be a potential scaffold is the Polo-box domain (PBD). The name, PBD, is given to the C-terminal region of a class of Ser/Thr kinases known as Polo-like kinases (Plks), which play an important role in various stages of cell cycle progression. Similar to the BRCT domains, the PBD contains two tandem polo boxes, which together bind to pS/T sequences in various interacting proteins. The interaction of PBD with its substrates is phosphorylation dependent. Interaction is not observed with non-phosphorylated peptides and a PDB mutant (H538A, K540A) in which residues important for phosphate recognition are mutated to alanine (*85*). The optimal binding motifs of PBDs were determined for five proteins by screening combinatorial phosphopeptide libraries and the K_d of Plk1 PBD with its optimal motif was determined to be 280 nM by isothermal titration calorimetry (*91*).

Crystal structures of the Plk1 PBD, in complex with various phosphopeptides, have been solved (*85, 92*). The PBDs have a strong preference for Ser in the pT (-1) position and replacing it with any of the 19 amino acids disrupted binding of the Plk1 PBD (*91*). Within the domain, the Trp414 residue interacts with the -1 residue of the peptide and it will be essential to select this residue for randomization during library construction such that the library members are not biased towards binding to phosphopeptide with Ser in the -1 position. As both the Polo boxes interact with the phosphopeptide, we anticipate that affinity reagents isolated from a library of variants derived from this domain will bind strongly to their cognate phosphopeptides. The availability of the structural information should be very beneficial in library design.

Both the PBD and BRCT domains have six cysteine residues that are not involved in disulfide bond formation. While it may be advantageous to replace the cysteine residues with other amino acids, there is a possibility that the domain will no longer be functionally active when displayed on the surface of bacteriophage M13 particles as I discovered with the FHA1 domain. If this is the case, I suggest that error-prone PCR be used to isolate variants that permit functional display and to include thermal selection in the same process to select functional as well as thermal stable variants which will be suitable as scaffolds for library construction.

We propose a strategy for the selection process (Fig. 3), which will incorporate *in vitro* affinity maturation during the selection process, such that the variants isolated after 3 rounds of selection will have a tight binding for their cognate target peptides. This process of combining ‘affinity selection’ and ‘directed evolution’ will yield high affinity binding clones that will be ready for use as reagents in cell staining and Western blotting. This will be a time saving process, as we will not have to affinity mature each and every binding clone separately.

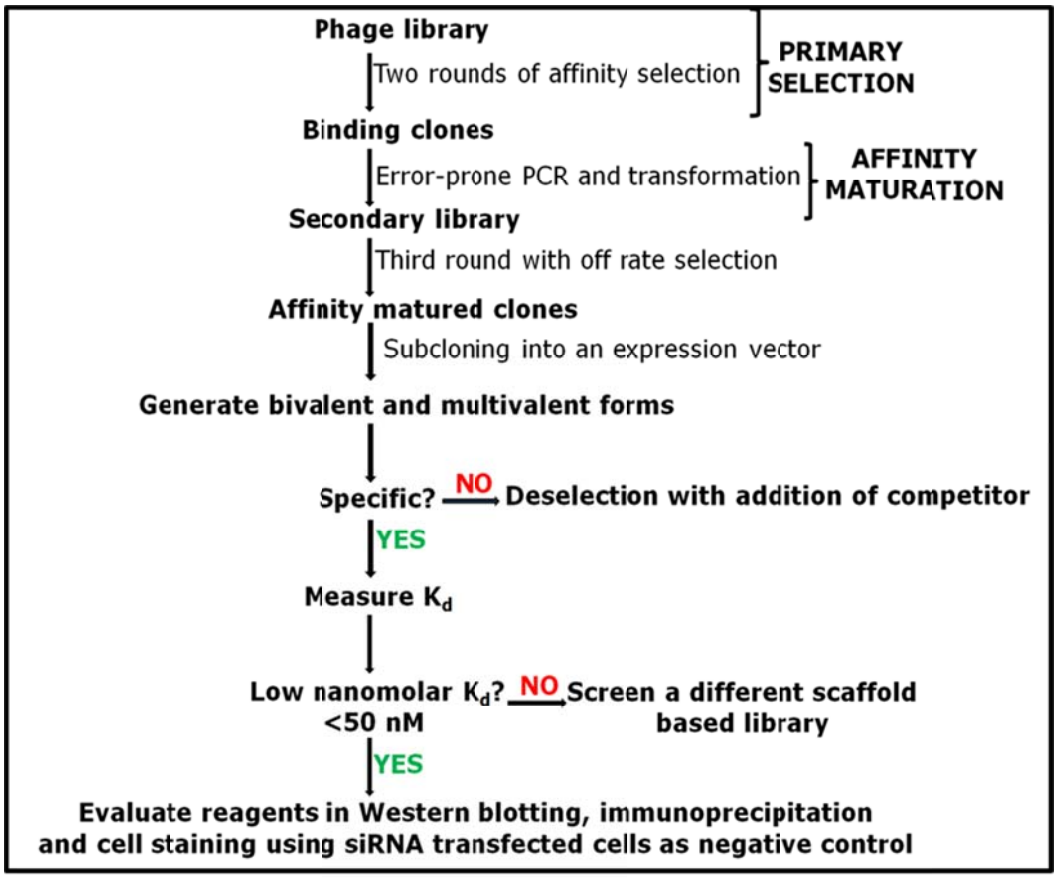


Figure 3. Strategy for isolating high affinity binding clones by combining affinity selection with affinity maturation. After performing two rounds of affinity selection against phosphopeptides of interest, mutagenic PCR will be performed on the coding sequence of all the binding clones to generate an affinity matured library. This library will be used in the third round of selection by incorporating off rate selection to isolate high affinity binding clones. After the third round, the pool of binding clones will be subcloned into an expression vector in-frame with the coding sequence of a dimerization or pentamerization domain. The multimerized proteins will be tested for binding to an array of other phosphopeptides and their affinity will be determined. If the reagents are not specific, then the phage library will be incubated with the competitor/cross-reactive peptide to eliminate them from the library and this process is termed as deselection. If the FHA scaffold is not optimal for isolating reagents to certain phosphopeptides, other scaffolds are attractive alternatives that can be tested.

Using a phosphopeptide-binding domain as a scaffold protein for generating anti-phosphopeptide reagents is an innovative idea and we have succeeded in our first goal in being able to change the specificity of FHA1 domain to bind to different pT-containing peptides. We have shown the applicability of these reagents in recognizing their target phosphopeptides in ELISA. Even though we have not been able to successfully use these reagents as probes in Western blotting and cell staining, we believe that by using the strategy described above, we will be able to isolate superior quality binding clones. Our reagents, like antibodies, will not be useful in all applications. We predict that some will be useful in detecting the phosphoprotein in assays where the protein is denatured and some will be able to recognize the protein in its native conformation. It will be very interesting to construct and screen phage-displayed libraries of other naturally occurring phosphopeptide-binding domains to develop affinity reagents that have different properties than the FHA domain. The power of *in vitro* directed evolution has made it possible to generate affinity reagents, which bind with high affinity and specificity to various target proteins.

We are in the process of optimizing a protocol for isolating affinity reagents that bind with high affinity and specificity to various phosphoepitopes. We will evaluate our reagents for various applications, such as western blotting, cell staining, immunoprecipitation, and immunohistochemistry. Since our reagents can be produced in large amounts by overexpression in *E. coli,* and their DNA

sequence is known, they can be readily made available to the scientific community who wish to use them to monitor phosphorylation of particular proteins in cells. We plan to collaborate with other labs in the future, and obtain three-dimensional structures of the domain bound to their peptide ligands through X-ray diffraction, as well as convert the reagents into biosensors that can monitor protein phosphorylation inside living cells. Our long-term goal is to develop a technology for generating reagents to large numbers of phosphorylated epitopes in a cost-effective and efficient manner.

5.5 References

1. Taussig, M. J., Stoevesandt, O., Borrebaeck, C. A., Bradbury, A. R., Cahill, D., Cambillau, C., de Daruvar, A., Dubel, S., Eichler, J., Frank, R., Gibson, T. J., Gloriam, D., Gold, L., Herberg, F. W., Hermjakob, H., Hoheisel, J. D., Joos, T. O., Kallioniemi, O., Koegl, M., Konthur, Z., Korn, B., Kremmer, E., Krobitsch, S., Landegren, U., van der Maarel, S., McCafferty, J., Muyldermans, S., Nygren, P. A., Palcy, S., Pluckthun, A., Polic, B., Przybylski, M., Saviranta, P., Sawyer, A., Sherman, D. J., Skerra, A., Templin, M., Ueffing, M., and Uhlen, M. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome, *Nat Methods 4*, 13-17.

2. Blow, N. (2007) Antibodies: The generation game, *Nature 447*, 741-744.

3. Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S., Clarke, K. F., Conquer, J. S., Crofts, A. M., Crowther, S. R., Dyson, M. R., Flack, G., Griffin, G. J., Hooks, Y., Howat, W. J., Kolb-Kokocinski, A., Kunze, S., Martin, C. D., Maslen, G. L., Mitchell, J. N., O'Sullivan, M., Perera, R. L., Roake, W., Shadbolt, S. P., Vincent, K. J., Warford, A., Wilson, W. E., Xie, J., Young, J. L., and McCafferty, J. (2007) Application of phage display to high throughput antibody generation and characterization, *Genome biology 8*, R254.

4. Haab, B. B., Paulovich, A. G., Anderson, N. L., Clark, A. M., Downing, G. J., Hermjakob, H., Labaer, J., and Uhlen, M. (2006) A reagent resource to identify proteins and peptides of interest for the cancer community: a workshop report, *Mol Cell Proteomics 5*, 1996-2007.

5. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Bjorling, L., and Ponten, F. (2010) Towards a knowledge-based Human Protein Atlas, *Nat Biotechnol 28*, 1248-1250.

6. Barbe, L., Lundberg, E., Oksvold, P., Stenius, A., Lewin, E., Bjorling, E., Asplund, A., Ponten, F., Brismar, H., Uhlen, M., and Andersson-Svahn, H. (2008) Toward a confocal subcellular atlas of the human proteome, *Mol Cell Proteomics 7*, 499-508.

7. Berglund, L., Bjorling, E., Oksvold, P., Fagerberg, L., Asplund, A., Szigyarto, C. A., Persson, A., Ottosson, J., Wernerus, H., Nilsson, P., Lundberg, E., Sivertsson, A., Navani, S., Wester, K., Kampf, C., Hober, S.,

Ponten, F., and Uhlen, M. (2008) A genecentric Human Protein Atlas for expression profiles based on antibodies, *Mol Cell Proteomics 7*, 2019-2027.

8. Uhlen, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergstrom, K., Brumer, H., Cerjan, D., Ekstrom, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Bjorklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Skollermo, A., Steen, J., Stenvall, M., Sterky, F., Stromberg, S., Sundberg, M., Tegel, H., Tourle, S., Wahlund, E., Walden, A., Wan, J., Wernerus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Ponten, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics, *Mol Cell Proteomics 4*, 1920-1932.

9. Michaud, G. A., Salcius, M., Zhou, F., Bangham, R., Bonin, J., Guo, H., Snyder, M., Predki, P. F., and Schweitzer, B. I. (2003) Analyzing antibody specificity with whole proteome microarrays, *Nat Biotechnol 21*, 1509-1512.

10. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips, *Science 293*, 2101-2105.

11. Huang, R., Fang, P., and Kay, B. K. (2012) Isolation of monobodies that bind specifically to the SH3 domain of the Fyn tyrosine protein kinase, *N Biotechnol 29*, 526-533.

12. Sjoberg, R., Sundberg, M., Gundberg, A., Sivertsson, A., Schwenk, J. M., Uhlen, M., and Nilsson, P. (2012) Validation of affinity reagents using antigen microarrays, *N Biotechnol 29*, 555-563.

13. Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to

20 human SH2 domains by phage display, *Protein Eng Des Sel 23*, 279-288.

14. Santiago, Y., Chan, E., Liu, P. Q., Orlando, S., Zhang, L., Urnov, F. D., Holmes, M. C., Guschin, D., Waite, A., Miller, J. C., Rebar, E. J., Gregory, P. D., Klug, A., and Collingwood, T. N. (2008) Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases, *Proc Natl Acad Sci U S A 105*, 5809-5814.

15. Wu, J., Kandavelou, K., and Chandrasegaran, S. (2007) Custom-designed zinc finger nucleases: what is next?, *Cell Mol Life Sci 64*, 2933-2944.

16. Filipowicz, W., Jaskiewicz, L., Kolb, F. A., and Pillai, R. S. (2005) Post-transcriptional gene silencing by siRNAs and miRNAs, *Curr Opin Struct Biol 15*, 331-341.

17. Gilmore, I. R., Fox, S. P., Hollins, A. J., and Akhtar, S. (2006) Delivery strategies for siRNA-mediated gene silencing, *Curr Drug Deliv 3*, 147-145.

18. Stadler, C., Hjelmare, M., Neumann, B., Jonasson, K., Pepperkok, R., Uhlen, M., and Lundberg, E. (2012) Systematic validation of antibody binding and protein subcellular localization using siRNA and confocal microscopy, *J Proteomics 75*, 2236-2251.

19. Mannsperger, H. A., Uhlmann, S., Schmidt, C., Wiemann, S., Sahin, O., and Korf, U. (2010) RNAi-based validation of antibodies for reverse phase protein arrays, *Proteome Sci 8*, 69.

20. Moore, K. G., Speckmann, W., and Herzig, R. P. (2007) The use of siRNA to validate immunofluorescence studies, *Methods Mol Biol 356*, 245-251.

21. Corzo, J. (2006) Time, the forgotten dimension of ligand binding teaching, *Biochem Mol Biol Educ 34*, 413-416.

22. Yau, K. Y., Dubuc, G., Li, S., Hirama, T., Mackenzie, C. R., Jermutus, L., Hall, J. C., and Tanha, J. (2005) Affinity maturation of a V(H)H by mutational hotspot randomization, *J Immunol Methods 297*, 213-224.

23. Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L., Kontermann, R. E., Jones, P. T., Low, N. M., Allison, T. J., and et al. (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires, *EMBO J 13*, 3245-3260.

24. Schier, R., McCall, A., Adams, G. P., Marshall, K. W., Merritt, H., Yim, M., Crawford, R. S., Weiner, L. M., Marks, C., and Marks, J. D. (1996) Isolation of picomolar affinity anti-c-erbB-2 single-chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody binding site, *J Mol Biol 263*, 551-567.

25. Stemmer, W. P. (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution, *Proc Natl Acad Sci U S A 91*, 10747-10751.

26. Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling, *Nature 370*, 389-391.

27. Dyson, M. R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B. K., Pawson, T., and McCafferty, J. (2011) Mapping protein interactions by combining antibody affinity maturation and mass spectrometry, *Anal Biochem 417*, 25-35.

28. Schier, R., Bye, J., Apell, G., McCall, A., Adams, G. P., Malmqvist, M., Weiner, L. M., and Marks, J. D. (1996) Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection, *J Mol Biol 255*, 28-43.

29. Marks, J. D., Griffiths, A. D., Malmqvist, M., Clackson, T. P., Bye, J. M., and Winter, G. (1992) By-passing immunization: building high affinity human antibodies by chain shuffling, *Bio/technology (Nature Publishing Company) 10*, 779-783.

30. Park, S. G., Lee, J. S., Je, E. Y., Kim, I. J., Chung, J. H., and Choi, I. H. (2000) Affinity maturation of natural antibody using a chain shuffling technique and the expression of recombinant antibodies in Escherichia coli, *Biochemical and biophysical research communications 275*, 553-557.

31. Jung, S., Honegger, A., and Pluckthun, A. (1999) Selection for improved protein stability by phage display, *J Mol Biol 294*, 163-180.

32. Dodevski, I., and Pluckthun, A. (2011) Evolution of three human GPCRs for higher expression and stability, *J Mol Biol 408*, 599-615.

33. Cadwell, R. C., and Joyce, G. F. (1994) Mutagenic PCR, *PCR methods and applications 3*, S136-140.

34. Krishnan, V. V., Thornton, K. H., Thelen, M. P., and Cosman, M. (2001) Solution structure and backbone dynamics of the human DNA ligase IIIalpha BRCT domain, *Biochemistry 40*, 13158-13166.

35. Stefan, N., Martin-Killias, P., Wyss-Stoeckle, S., Honegger, A., Zangemeister-Wittke, U., and Pluckthun, A. (2011) DARPins recognizing the tumor-associated antigen EpCAM selected by phage and ribosome display and engineered for multivalency, *J Mol Biol 413*, 826-843.

36. Zahnd, C., Wyler, E., Schwenk, J. M., Steiner, D., Lawrence, M. C., McKern, N. M., Pecorari, F., Ward, C. W., Joos, T. O., and Pluckthun, A. (2007) A designed ankyrin repeat protein evolved to picomolar affinity to Her2, *J Mol Biol 369*, 1015-1028.

37. Razai, A., Garcia-Rodriguez, C., Lou, J., Geren, I. N., Forsyth, C. M., Robles, Y., Tsai, R., Smith, T. J., Smith, L. A., Siegel, R. W., Feldhaus, M., and Marks, J. D. (2005) Molecular evolution of antibody affinity for sensitive detection of botulinum neurotoxin type A, *J Mol Biol 351*, 158-169.

38. Zahnd, C., Spinelli, S., Luginbuhl, B., Amstutz, P., Cambillau, C., and Pluckthun, A. (2004) Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scFv) with low picomolar affinity, *J Biol Chem 279*, 18870-18877.

39. Graff, C. P., Chester, K., Begent, R., and Wittrup, K. D. (2004) Directed evolution of an anti-carcinoembryonic antigen scFv with a 4-day monovalent dissociation half-time at 37 degrees C, *Protein Eng Des Sel 17*, 293-304.

40. Yoon, S. O., Lee, T. S., Kim, S. J., Jang, M. H., Kang, Y. J., Park, J. H., Kim, K. S., Lee, H. S., Ryu, C. J., Gonzales, N. R., Kashmiri, S. V., Lim, S. M., Choi, C. W., and Hong, H. J. (2006) Construction, affinity maturation, and biological characterization of an anti-tumor-associated glycoprotein-72 humanized antibody, *J Biol Chem 281*, 6985-6992.

41. Chen, Y., Wiesmann, C., Fuh, G., Li, B., Christinger, H. W., McKay, P., de Vos, A. M., and Lowman, H. B. (1999) Selection and analysis of an optimized anti-VEGF antibody: crystal structure of an affinity-matured Fab in complex with antigen, *J Mol Biol 293*, 865-881.

42. Wu, H., Beuerlein, G., Nie, Y., Smith, H., Lee, B. A., Hensler, M., Huse, W. D., and Watkins, J. D. (1998) Stepwise in vitro affinity maturation of Vitaxin, an alphav beta3-specific humanized mAb, *Proc Natl Acad Sci U S A 95*, 6037-6042.

43. Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H. R., and Pluckthun, A. (1998) Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries, *Proc Natl Acad Sci U S A 95*, 14130-14135.

44. Hanes, J., Schaffitzel, C., Knappik, A., and Pluckthun, A. (2000) Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display, *Nat Biotechnol 18*, 1287-1292.

45. Zhao, X. L., Chen, W. Q., Yang, Z. H., Li, J. M., Zhang, S. J., and Tian, L. F. (2009) Selection and affinity maturation of human antibodies against rabies virus from a scFv gene library using ribosome display, *J Biotechnol 144*, 253-258.

46. Huovinen, T., Sanmark, H., Yla-Pelto, J., Vehniainen, M., and Lamminmaki, U. (2010) Oligovalent Fab display on M13 phage improved by directed evolution, *Mol Biotechnol 44*, 221-231.

47. Strobel, H., Ladant, D., and Jestin, J. L. (2003) Efficient display of two enzymes on filamentous phage using an improved signal sequence, *Mol Biotechnol 24*, 1-10.

48. Pershad, K., and Kay, B. (Accepted for publication) Generating Thermally Stable Variants of Protein Domains through Phage-display *Methods Journal*.

49. Roitt, I. M. (1991) *Essential Immunology*, 7th ed., Oxford/Blackwell, London.

50. Male, D., Champion, B. & Cooke, A. (1987) *Advanced Immunology*, Gower Medical, London.

51. Ronnmark, J., Hansson, M., Nguyen, T., Uhlen, M., Robert, A., Stahl, S., and Nygren, P. A. (2002) Construction and characterization of affibody-Fc chimeras produced in Escherichia coli, *J Immunol Methods 261*, 199-211.

52. Powers, D. B., Amersdorfer, P., Poul, M., Nielsen, U. B., Shalaby, M. R., Adams, G. P., Weiner, L. M., and Marks, J. D. (2001) Expression of

single-chain Fv-Fc fusions in Pichia pastoris, *J Immunol Methods 251*, 123-135.

53. de Kruif, J., and Logtenberg, T. (1996) Leucine zipper dimerized bivalent and bispecific scFv antibodies from a semi-synthetic antibody phage display library, *J Biol Chem 271*, 7630-7634.

54. Weber-Bornhauser, S., Eggenberger, J., Jelesarov, I., Bernard, A., Berger, C., and Bosshard, H. R. (1998) Thermodynamics and kinetics of the reaction of a single-chain antibody fragment (scFv) with the leucine zipper domain of transcription factor GCN4, *Biochemistry 37*, 13011-13020.

55. Duan, J., Wu, J., Valencia, C. A., and Liu, R. (2007) Fibronectin type III domain based monobody with high avidity, *Biochemistry 46*, 12656-12664.

56. Terskikh, A. V., Le Doussal, J. M., Crameri, R., Fisch, I., Mach, J. P., and Kajava, A. V. (1997) "Peptabody": a new type of high avidity binding protein, *Proc Natl Acad Sci U S A 94*, 1663-1668.

57. Zhang, J., Tanha, J., Hirama, T., Khieu, N. H., To, R., Tong-Sevinc, H., Stone, E., Brisson, J. R., and MacKenzie, C. R. (2004) Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents, *J Mol Biol 335*, 49-56.

58. Ravn, P., Danielczyk, A., Jensen, K. B., Kristensen, P., Christensen, P. A., Larsen, M., Karsten, U., and Goletz, S. (2004) Multivalent scFv display of phagemid repertoires for the selection of carbohydrate-specific antibodies and its application to the Thomsen-Friedenreich antigen, *J Mol Biol 343*, 985-996.

59. Kato, T., Sato, K., Suzuki, S., Sasakawa, H., Kurokawa, M., Nishioka, K., and Yamamoto, K. (1995) Mammalian expression of single chain variable region fragments dimerized by Fc regions, *Mol Biol Rep 21*, 141-146.

60. Hong, J. W., Cho, W. D., Hong, K. P., Kim, S. S., Son, S. M., Yun, S. J., Lee, H. C., Yoon, S. S., and Song, H. G. (2012) Generation of 1E8 Single Chain Fv-Fc Construct Against Human CD59, *Immune Netw 12*, 33-39.

61. Todorovska, A., Roovers, R. C., Dolezal, O., Kortt, A. A., Hoogenboom, H. R., and Hudson, P. J. (2001) Design and application of diabodies,

triabodies and tetrabodies for cancer targeting, *J Immunol Methods 248*, 47-66.

62. Atwell, J. L., Breheney, K. A., Lawrence, L. J., McCoy, A. J., Kortt, A. A., and Hudson, P. J. (1999) scFv multimers of the anti-neuraminidase antibody NC10: length of the linker between VH and VL domains dictates precisely the transition between diabodies and triabodies, *Protein Eng 12*, 597-604.

63. Kortt, A. A., Dolezal, O., Power, B. E., and Hudson, P. J. (2001) Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting, *Biomol Eng 18*, 95-108.

64. Takemura, S., Asano, R., Tsumoto, K., Ebara, S., Sakurai, N., Katayose, Y., Kodama, H., Yoshida, H., Suzuki, M., Imai, K., Matsuno, S., Kudo, T., and Kumagai, I. (2000) Construction of a diabody (small recombinant bispecific antibody) using a refolding system, *Protein Eng 13*, 583-588.

65. MacKenzie, R., and To, R. (1998) The role of valency in the selection of anti-carbohydrate single-chain Fvs from phage display libraries, *J Immunol Methods 220*, 39-49.

66. Lu, D., Jimenez, X., Zhang, H., Atkins, A., Brennan, L., Balderes, P., Bohlen, P., Witte, L., and Zhu, Z. (2003) Di-diabody: a novel tetravalent bispecific antibody molecule by design, *J Immunol Methods 279*, 219-232.

67. Pack, P., and Pluckthun, A. (1992) Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in Escherichia coli, *Biochemistry 31*, 1579-1584.

68. Kostelny, S. A., Cole, M. S., and Tso, J. Y. (1992) Formation of a bispecific antibody by the use of leucine zippers, *J Immunol 148*, 1547-1553.

69. Lee, C. V., Sidhu, S. S., and Fuh, G. (2004) Bivalent antibody phage display mimics natural immunoglobulin, *J Immunol Methods 284*, 119-132.

70. Fraile, S., Munoz, A., de Lorenzo, V., and Fernandez, L. A. (2004) Secretion of proteins with dimerization capacity by the haemolysin type I transport system of Escherichia coli, *Mol Microbiol 53*, 1109-1121.

71. Cocca, B. A., Seal, S. N., and Radic, M. Z. (1999) Tandem affinity tags for the purification of bivalent anti-DNA single-chain Fv expressed in Escherichia coli, *Protein Expr Purif 17*, 290-298.

72. Els Conrath, K., Lauwereys, M., Wyns, L., and Muyldermans, S. (2001) Camel single-domain antibodies as modular building units in bispecific and bivalent antibody constructs, *J Biol Chem 276*, 7346-7350.

73. Zhang, J., Li, Q., Nguyen, T. D., Tremblay, T. L., Stone, E., To, R., Kelly, J., and Roger MacKenzie, C. (2004) A pentavalent single-domain antibody approach to tumor antigen discovery and the development of novel proteomics reagents, *J Mol Biol 341*, 161-169.

74. Tanha, J., Dubuc, G., Hirama, T., Narang, S. A., and MacKenzie, C. R. (2002) Selection by phage display of llama conventional V(H) fragments with heavy chain antibody V(H)H properties, *J Immunol Methods 263*, 97-109.

75. Pershad, K., Sullivan, M. A., and Kay, B. K. (2011) Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats, *Anal Biochem 412*, 210-216.

76. Han, Z., Karatan, E., Scholle, M. D., McCafferty, J., and Kay, B. K. (2004) Accelerated screening of phage-display output with alkaline phosphatase fusions, *Comb Chem High Throughput Screen 7*, 55-62.

77. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., and Yaffe, M. B. (2000) The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms, *Mol Cell 6*, 1169-1182.

78. Mahajan, A., Yuan, C., Lee, H., Chen, E. S., Wu, P. Y., and Tsai, M. D. (2008) Structure and function of the phosphothreonine-specific FHA domain, *Science signaling 1*, re12.

79. Li, J., Lee, G. I., Van Doren, S. R., and Walker, J. C. (2000) The FHA domain mediates phosphoprotein interactions, *J Cell Sci 113 Pt 23*, 4143-4149.

80. Yuan, C., Yongkiettrakul, S., Byeon, I. J., Zhou, S., and Tsai, M. D. (2001) Solution structures of two FHA1-phosphothreonine peptide complexes

provide insight into the structural basis of the ligand specificity of FHA1 from yeast Rad53, *J Mol Biol 314*, 563-575.

81. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P., and Smerdon, S. J. (2002) Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2, *Mol Cell 9*, 1045-1054.

82. Lee, H., Yuan, C., Hammet, A., Mahajan, A., Chen, E. S., Wu, M. R., Su, M. I., Heierhorst, J., and Tsai, M. D. (2008) Diphosphothreonine-specific interaction between an SQ/TQ cluster and an FHA domain in the Rad53-Dun1 kinase cascade, *Mol Cell 30*, 767-778.

83. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell 127*, 635-648.

84. Zhang, X., Morera, S., Bates, P. A., Whitehead, P. C., Coffer, A. I., Hainbucher, K., Nash, R. A., Sternberg, M. J., Lindahl, T., and Freemont, P. S. (1998) Structure of an XRCC1 BRCT domain: a new protein-protein interaction module, *EMBO J 17*, 6404-6411.

85. Joo, W. S., Jeffrey, P. D., Cantor, S. B., Finnin, M. S., Livingston, D. M., and Pavletich, N. P. (2002) Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brca1 BRCT structure, *Genes Dev 16*, 583-593.

86. Williams, R. S., Green, R., and Glover, J. N. (2001) Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1, *Nat Struct Biol 8*, 838-842.

87. Yu, X., Chini, C. C., He, M., Mer, G., and Chen, J. (2003) The BRCT domain is a phospho-protein binding domain, *Science 302*, 639-642.

88. Rodriguez, M., Yu, X., Chen, J., and Songyang, Z. (2003) Phosphopeptide binding specificities of BRCA1 COOH-terminal (BRCT) domains, *J Biol Chem 278*, 52914-52918.

89. Clapperton, J. A., Manke, I. A., Lowery, D. M., Ho, T., Haire, L. F., Yaffe, M. B., and Smerdon, S. J. (2004) Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer, *Nat Struct Mol Biol 11*, 512-518.

90. Cantor, S. B., Bell, D. W., Ganesan, S., Kass, E. M., Drapkin, R., Grossman, S., Wahrer, D. C., Sgroi, D. C., Lane, W. S., Haber, D. A., and Livingston, D. M. (2001) BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function, *Cell 105*, 149-160.

91. Elia, A. E., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J., and Yaffe, M. B. (2003) The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain, *Cell 115*, 83-95.

92. Cheng, K. Y., Lowe, E. D., Sinclair, J., Nigg, E. A., and Johnson, L. N. (2003) The crystal structure of the human polo-like kinase-1 polo box domain and its phospho-peptide complex, *Embo J 22*, 5757-5768.

APPENDICES

Part of the work in Appendix A has been published

Pershad, K., Sullivan, M. A., and Kay, B. K. (2011) Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats, *Anal Biochem 412*, 210-216.

APPENDIX A

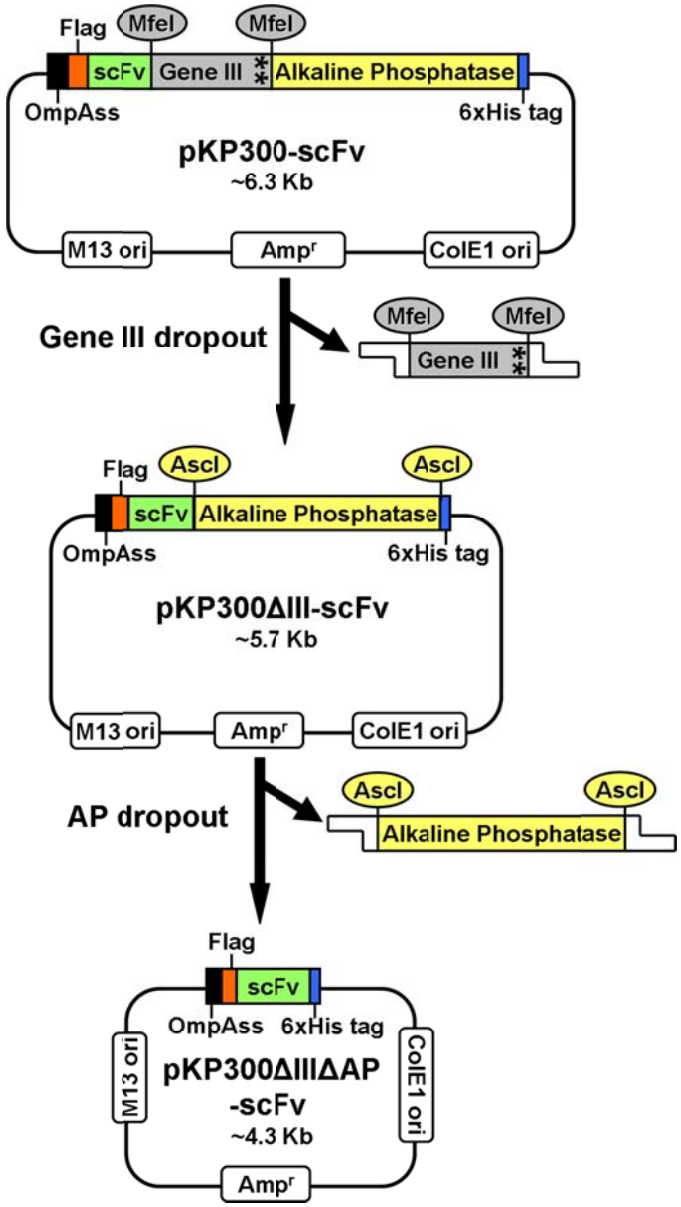
Plasmid vectors

pKP300, a Double-dropout Phagemid Vector

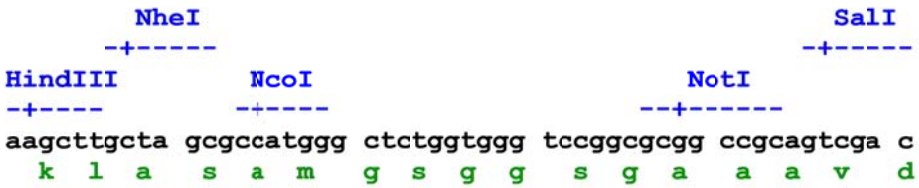
The scFv coding sequence is cloned in frame with the gene III coding region of the pKP300 vector, allowing display of the scFv on the surface of M13 bacteriophage particles. The OmpA signal sequence (MKKTAIAIAVALAGFAT VA) directs the recombinant protein into the periplasm of *E. coli*. All proteins display a Flag peptide epitope (DYKDDDDKL) at their N-termini. Gene III is truncated (~630 bp), and carries opal and ochre stop codons, which are represented by two asterisks, at the C-terminus of the coding regions of truncated gene III.

After the affinity selection process, the gene III coding sequence is dropped out by digestion with the restriction endonuclease, *Mfe* I, followed by re-ligation of the vector. This results in an in frame fusion of the alkaline phosphatase (AP) coding region to the C-terminus of the scFv. Subsequent digestion of the recombinant plasmid with *Asc* I leads to loss of the AP coding region, with re-ligation fusing the coding region of the scFv in frame with a C-terminal 6xHis tag. The proteins can be purified by immobilized metal affinity chromatography (IMAC) via the C-terminal six-histidine tag. There is a TAA stop codon after the six-histidine tag. Protein expression is under the control of PhoA (endogenous

alkaline phosphatase) promoter. The sizes of the three vector DNAs including the scFv coding sequences are represented in kilobases (kb).



Multiple cloning site (MCS)



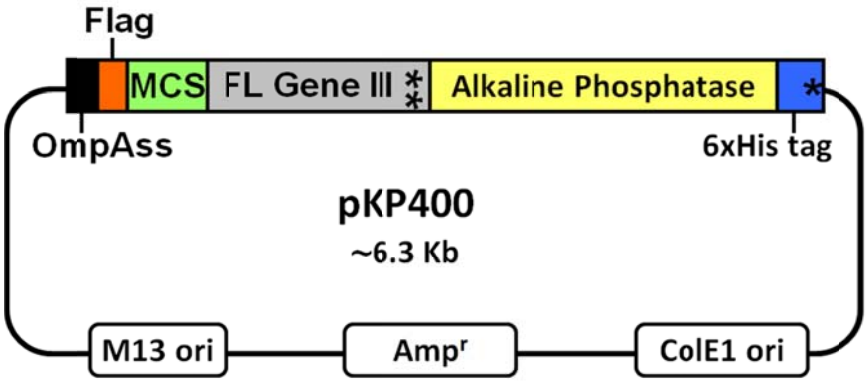
pKP400 phagemid vector

The pKP400 vector is similar to the pKP300 vector, except that it has full-length version of gene III (1272 bps), instead of a truncated one. The multiple cloning site (MCS) is similar to the pKP300 vector. The pKP400 vector is 6.3 kilobases in size.

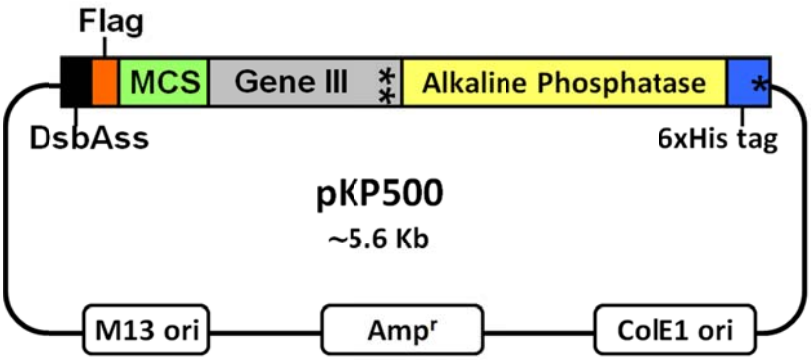
When a library of inserts is fused in frame to the gene III coding sequence in a phagemid vector, it leads to display of zero or one copy (i.e., monovalent display) of the recombinant protein on the surface of bacteriophage M13, with four or five copies of protein III provided by the helper phage (i.e., M13K07). Typically, the displayed protein is present on only 1-3% of phage particles (*1*) .

During affinity selection, phage particles that only display wild-type protein III can contribute to the background of non-binding phage. To eliminate such

phage, the helper phage has been engineered to contain a trypsin cleavable site (PAGLSEGSTIEGRGAHE) between the N2 and CT domains (*2, 3*) of protein III. If such a helper phage is used to package recombinant phage particles, during affinity selection, trypsin can be used to convert particles that display only wild-type protein III phages non-infective. This is a very effective approach of eliminating wild-type phage during each round of selection. However, for this system to work, the library of inserts has to be fused in frame to a full-length form of gene III, so that the recombinant phages can be specifically eluted with trypsin, and remain infective. We have used this system to isolate scFvs to human SH2 domains (*4*) and to isolate FHA variants to phosphopeptides (*5*) .



pKP500 phagemid vector



Differences compared to pKP300 phagemid vector:

1. Signal sequence: DsbA (MKKIWLALAGLVLAFSASA). The DsbA signal sequence is flanked on either side with a *Sac* I restriction enzyme site, which can be used for convenient switching to any other signal sequence.
2. The alkaline phosphatase coding sequence is flanked by *Sfi* I sites, which allows directional cloning of other genes, in this case the AviTag and Halotag.
3. Vector sizes:

    pKP500, with AP coding sequence, 5.6 kb

    pKP500-Avitag, 4.3 kb

    pKP500-Halotag, 5.2 kb

4. AviTag is a 15 amino acid peptide sequence (GLNDIFEAQ**K**IEWHE) to which a single biotin molecule is conjugated at the lysine (K) residue by the biotin ligase (BirA) enzyme of *E. coli (6-8).*
5. DNA containing the Halotag coding sequence was purchased from Promega. A range of Halotag ligands that make a covalent bond with the halotag protein are available as fusions to affinity tags or fluorescent probes that are then useful in the purification or detection of Halotag fusions (*9, 10*).
6. The β-lactamase gene has *BstAP* I sites on either side to facilitate directional cloning of different antibiotic resistance genes.
7. This phagemid vector offers easy subcloning of a different signal sequence or a fusion partner or the antibiotic resistance gene, due to the presence of unique restriction sites flanking these regions.

APPENDIX B

Generating conformation specific antibody fragments

When a residue is mutated in a protein (or domain), sometimes there is a loss of binding activity. There are two possible explanations for the loss of binding: a) the residue is important for the function of the protein or b) the residue is critical for the structural stability of the protein and mutating it disrupts the three dimensional structure of the protein. An antibody that recognizes a protein conformation will be a useful resource to address this question.

To map which residues in the FHA1 domain contribute to peptide ligand binding, I generated 24 alanine-scan mutants, and tested binding to the cognate pT peptide by phage ELISA. Several point mutants lost binding, but it was unclear if the loss of binding/reduced binding indicated that the mutated residue was important for binding to the cognate pT peptide or that it was important for the structural stability of the FHA1G2 variant. To discriminate between these two possible explanations, I decided to generate an affinity reagent that specifically recognizes the folded FHA1G2 protein and does not bind to the denatured form of the domain.

Materials and methods

Affinity selection for isolating an antibody fragment that recognizes conformational epitopes on the FHAG2 variant

Two rounds of affinity selection were performed against the FHA1G2 variant to identify antibody fragments from a phage-displayed scFv library (*11*) that recognized specifically the folded form of the FHA1G2 variant. All the selection steps were performed at room temperature, with 1 h incubation between each step (Fig. 1). The FHA1G2 protein (200 µL of 20 µg/mL; total 4 µg) was immobilized on Nunc polystyrene tube (Thermo Fisher Scientific), and then blocked with 2% skim milk in PBS (5 mL). The FHA1G2 protein (20 µL of 100 µg/mL; total 2 µg) was denatured by heating for 1 h at 95°C. The phage-displayed scFv library (*11*) (200 µL, diluted 1:1 with PBS) was mixed with denatured FHA1G2 protein (20 µL of 100 µg/mL; total 2 µg) and incubated for 30 min. Then 1/10 volume of 20% skim milk (22 µL; final is 2% skim milk) was added and incubation was carried out for 10 additional min. This pre-treated library (~240 µL) was added to the blocked FHA1G2 protein and after 1 h incubation, the non-bound phage particles were eliminated by six washes with PBST (PBS with 0.1% Tween 20) and six washes with PBS. A double elution was done to elute all the phage bound to the target. First, the bound phage were eluted for 15 min with TPCK treated trypsin (Sigma-Aldrich; 200 µL at 100 µg/mL

concentration), and used to infect 800 µL of TG1 cells, grown at mid-log phase (OD_{600nm}=0.5), for 40 min at 37ºC. Next, the tube was rinsed once with PBS and 200 µL of TG1 cells at mid-log phase (OD_{600nm}=0.5) were directly added to the tube to capture phage not eluted efficiently by trypsin treatment. Infection was carried out for 40 min at 37ºC. All the infected TG1 cells were pooled, plated onto one 15 cm LB/carbenicillin agar plate; colonies were scraped the next day with 2 mL of LB/CB/16% glycerol media. Scraped cells (10 µL; ~10^8 cells) were inoculated into 5 mL of LB/CB media, grown to mid-log phase, infected with M13KO7 helper phage (MOI=20) and amplified overnight at 30°C/250 rpm in 10 mL of LB/CB/Kan medium. The next day, phage particles were precipitated with PEG/NaCl, and the pellet resuspended in PBS. The second round of affinity selection was conducted in the same manner, except that only 1/2 the volume of the eluted phage was used to infect bacterial cells. Titration plates were prepared by plating 10 µL and 100 µL of 10^{-2} and 10^{-3} dilutions on 10 cm LB/CB agar plates and incubated overnight at 30°C.

Seventy-two individual colonies were propagated in 96 well deep well plates, and the phage supernatant was used for phage ELISA. Nunc MaxiSorp flat-bottom 96 well plates (Thermo Fisher Scientific) were coated with FHA1G2 protein (100 µL of 5 µg/mL, total of 0.5 µg) for 1 h, blocked with 2% skim milk in PBS (200 µL/well) and incubated with phage supernatant (100 µL) diluted 1:1 with PBST. Phage binding was detected by incubating with anti-M13 antibody

conjugated to HRP (GE Healthcare, diluted 1:5000 with PBST, 100 µL/well) and the binding signal was measured at 405 nm on POLARstar OPTIMA microtiter plate reader (BMG Labtech), after adding (2',2'-Azino-Bis 3-Ethylbenzothiazoline-6-Sulfonic Acid; Thermo Fisher Scientific, 100 µL/well).

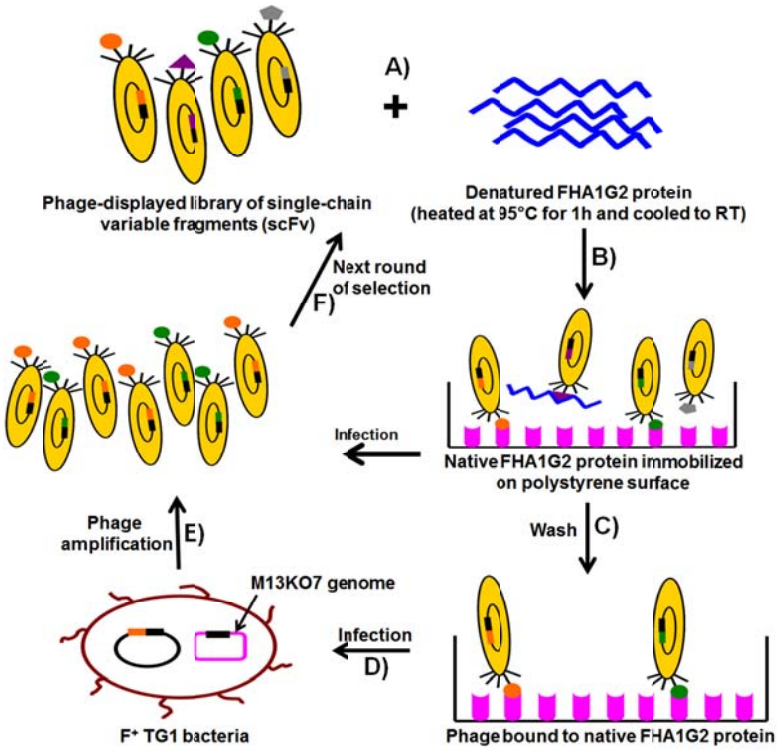


Figure 1. Affinity selection for isolating scFvs that bind native, but not the denatured form, of the FHA1G2 domain. A) The phage library was first incubated with the denatured FHA1G2 protein such that the scFvs that have a greater affinity for the denatured protein will bind to it and remain bound during the subsequent steps of the selection process. B) Then the library was incubated with the FHA1G2 protein in its native conformation that has been immobilized on the surface of a polystyrene tube by passive adsorption. C) The unbound phage, including the ones bound to the denatured protein were washed using a detergent solution. D) The target bound phage were eluted with trypsin and used to infect TG1 cells, which were then grown to mid-log phase and infected with M13KO7 helper phage. E) Infected cells were grown overnight to allow for phage propagation, which is used for the second round of selection.

Phage-ELISA to confirm specific binding of scFvs to folded FHA1G2 protein

FHA1G2 protein was chemically biotinylated using a biotinylation kit (Pierce). The biotinylated protein (100 µL of 60 µg/mL; total 6 µg) was heated at 95°C for 1 h, cooled, centrifuged to remove the precipitated protein pellet and the clarified protein was used for the phage-ELISA. The biotinylated FHA1G2 protein (100 µL of 5 µg/mL; total 0.5 µg) and the denatured FHA1G2 protein (100 µL of 60 µg/mL; total 6 µg) were immobilized on Nunc MaxiSorp flat-bottom 96 well plates (Thermo Fisher Scientific), in duplicates, via NeutrAvidin™ (100 µL of 5 µg/mL for folded FHA1G2 and 100 µL of 40 µg/mL for denatured FHA1G2). The proteins were blocked with 2% skim milk in PBS (200 µL per well), and incubated with the phage supernatant (100 µL of phage diluted 1:1 with PBST). The binding phage were detected using anti-M13 antibody conjugated to HRP (GE Healthcare) diluted 1:5000 with PBST. After washing away the unbound antibody, the chromogenic substrate for HRP; ABTS (Thermo Fisher Scientific), supplemented with hydrogen peroxide was added (100 µL per well), and the absorbance of the green colored complex was measured at 405 nm on POLARstar OPTIMA microtiter plate reader (BMG Labtech). Similarly, the folded and denatured proteins were detected with anti-His antibody (Sigma-Aldrich; diluted 1:3000 with PBST) after blocking to confirm that similar amounts of both the proteins were immobilized on the microtiter plate wells.

Results

Evaluating conformation specific anti-FHA1G2 scFvs by ELISA

After two rounds of affinity selection against the FHA1G2 protein, two scFvs with different sequences (D12 and A5) were isolated that showed specific binding to the folded FHA1G2 protein and not to the denatured from of the protein (Fig. 2A). Equal amounts of the folded and denatured proteins were immobilized on the microtiter plate wells for the phage-ELISA (Fig. 2B). This experiment demonstrates that the deselection strategy used during affinity selection is effective in eliminating clones that recognize linear epitopes. While we have not mapped the binding epitope of these scFvs, our experiment confirms that the scFvs do not bind linear epitopes in the denatured protein and recognize a conformational epitope that is lost after heating the protein.

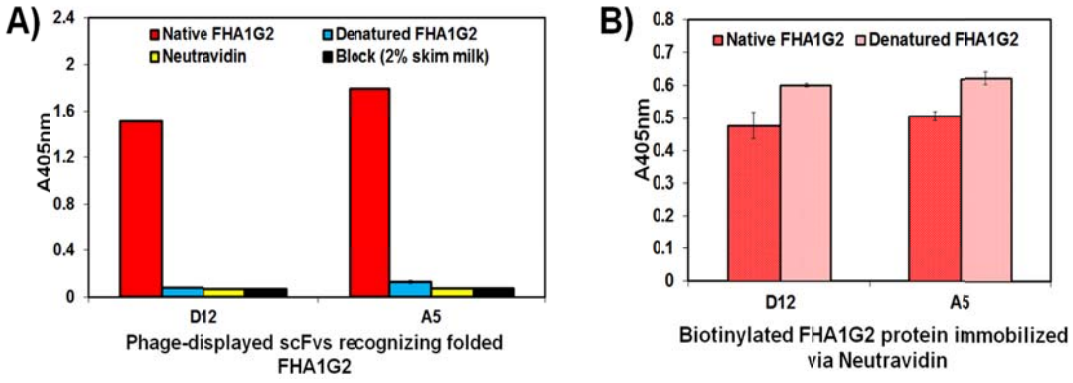


Figure 2. Isolating conformation specific scFvs for the FHA1G2 variant. A) Three phage-displayed scFvs with different sequences were tested for binding to FHA1G2 protein in its folded state and when it is denatured by heating. Binding was also tested for two other negative controls. The folded and denatured FHA1G2 proteins are biotinylated and captured on microtiter plate wells pre-coated with NeutrAvidin. Phage binding was detected using anti-M13 antibody conjugated to HRP. B) NeutrAvidin captured folded and denatured FHA1G2 proteins were detected using an anti-His antibody, which confirms that equal amount of both the proteins are captured on the microtiter plate wells.

The scFv A5 was subcloned into the pKP300 expression vector, between the *Nco* I and *Not* I restriction sites, and purified by immobilized metal affinity chromatography via the C-terminal 6XHis tag, following the protocol described in (*12*). All the 24 alanine-scan mutants were tested for binding to the scFv-A5 to determine which one of them remained folded after the specific mutation (*5*). This is described in Chapter 3 (Page 130).

Previously, conformation specific Fab antibody fragments were isolated for two different conformations of a protease, Caspase, by deselecting for one conformation or the other during the affinity selection process (*13*). ScFvs were isolated from yeast displayed libraries that can distinguish between Calmodulin protein in its free state and when bound to calcium (*14*).

Nucleotide and amino acid sequences of the three conformation specific scFvs

The Flag epitope (DYKDDDDKL) at the N-terminus of the scFv primary structure is highlighted yellow. The scFv contains the variable region of heavy chain (V_H) in between *Nco* I and *Xho* I recognition sites and the variable region of light chain (V_L) in between *Nhe* I and *Not* I recognition sites, joined by a 15 amino acid flexible linker ($Gly_4SerGly_4SerGly_3AlaSer$; highlighted blue), sites. There is a 6XHis tag at the C-terminus (highlighted grey).

A5 scFv

```
GACTACAAGGACGACGATGACAAGCTTGCTAGCGCCATGGCCCAGGTGCAGCTGGTGGAG
 D  Y  K  D  D  D  D  K  L  A  S  A  M  A  Q  V  Q  L  V  E
TCCGGAGCAGAGGTGAAAAAGCCCGGGGAGTCTCTGAGGATCTCCTGTAAGGGTTCTGGA
 S  G  A  E  V  K  K  P  G  E  S  L  R  I  S  C  K  G  S  G
TACAGCTTTACCAGCTACTGGATCGGCTGGGTGCGCCAGATGCCCGGGAAAGGCCTGGAG
 Y  S  F  T  S  Y  W  I  G  W  V  R  Q  M  P  G  K  G  L  E
TGGATGGGGATCATCTATCCTGGTGACTCTGATACCAGATACAGCCCGTCCTTCCAAGGC
 W  M  G  I  I  Y  P  G  D  S  D  T  R  Y  S  P  S  F  Q  G
CAGGTCACCATCTCAGCCGACAAGTCCATCAGCACCGCCTACCTGCAGTGGAGCAGCCTG
 Q  V  T  I  S  A  D  K  S  I  S  T  A  Y  L  Q  W  S  S  L
AAGGCCTCGGACACCGCCATGTATTACTGTGCGAGGATCGCATCTGGGATGGAACTTTAT
 K  A  S  D  T  A  M  Y  Y  C  A  R  I  A  S  G  M  E  L  Y
TCTTTTGATATCTGGGGCCAAGGGACAATGGTCACCGTCTCTTCACTCGAGGGTGGAGGC
 S  F  D  I  W  G  Q  G  T  M  V  T  V  S  S  L  E  G  G  G
GGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGCTAGCGACATCCAGATGACCCAGTCTCCA
 G  S  G  G  G  G  S  G  G  G  A  S  D  I  Q  M  T  Q  S  P
TCCTCCCTGTCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCGGGCAAGTCAGGGC
 S  S  L  S  A  S  V  G  D  R  V  T  I  T  C  R  A  S  Q  G
ATTAGCAATTATTTAGCCTGGTATCAGCAGAAACCAGGGAAAGCCCCTAAGCGCCTGATC
 I  S  N  Y  L  A  W  Y  Q  Q  K  P  G  K  A  P  K  R  L  I
TATGCTGCATCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGCGGCAGTGGATCTGGG
 Y  A  A  S  S  L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G
ACAGAATTCACTCTCACCATCAGCAGCCTGCAGCCTGAAGATTTTGCAACTTACTACTGT
 T  E  F  T  L  T  I  S  S  L  Q  P  E  D  F  A  T  Y  Y  C
CAACAGAGTTACAGTACCCCGCTCACTTTCGGCGGAGGGACCAAGCTGGAGATCAAACGT
 Q  Q  S  Y  S  T  P  L  T  F  G  G  G  T  K  L  E  I  K  R
GCGGCCGCAGTCGACGGGCGCGCCGCTTCCCATCACCATCACCATCAC
 A  A  A  V  D  G  R  A  A  S  H  H  H  H  H  H
```

D12 scFv

```
GACTACAAGGACGACGATGACAAGCTTGCTAGCGCCATGGCCGAGGTGCAGCTGGTGCAG
 D  Y  K  D  D  D  D  K  L  A  S  A  M  A  E  V  Q  L  V  Q
TCTGGAGCAGAGGTGAAAAAGCCGGGGGAGTCTCTGAAGATCTCCTGTAAGGGTTCTGGA
 S  G  A  E  V  K  K  P  G  E  S  L  K  I  S  C  K  G  S  G
TACAGCTTTACCAGCTACTGGATCGGCTGGGTGCGCCAGATGCCCGGGAAAGGCCTGGAG
 Y  S  F  T  S  Y  W  I  G  W  V  R  Q  M  P  G  K  G  L  E
TGGATGGGGATCATCTATCCTGGTGACTCTGATACCAGATACAGCCCGTCCTTCCAAGGC
 W  M  G  I  I  Y  P  G  D  S  D  T  R  Y  S  P  S  F  Q  G
CAGGTCACCATCTCAGCCGACAAGTCCATCAGCACCGCCTACCTGCAGTGGAGCAGCCTG
 Q  V  T  I  S  A  D  K  S  I  S  T  A  Y  L  Q  W  S  S  L
AAGGCCTCGGACACCGCCATGTATTACTGTGCGAGATCCAAGGGGCTCGGTAGATTCGAT
 K  A  S  D  T  A  M  Y  Y  C  A  R  S  K  G  L  G  R  F  D
CTCTGGGGCCGTGGCACCCTGGTCACCGTCTCCTCACTCGAGGGTGGAGGCGGTTCAGGC
 L  W  G  R  G  T  L  V  T  V  S  S  L  E  G  G  G  G  S  G
GGAGGTGGCTCTGGCGGTGGCGCTAGCGATATTGTGATGACGCAGTCTCCATCCTCCCTG
 G  G  G  S  G  G  G  A  S  D  I  V  M  T  Q  S  P  S  S  L
TCTGCATCTGTAGGAGACAGAGTCACCATCACTTGCCGGGCAAGTCAGAGCATTAGCACC
 S  A  S  V  G  D  R  V  T  I  T  C  R  A  S  Q  S  I  S  T
TATTTAAATTGGTATCAGCAGAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCA
 Y  L  N  W  Y  Q  Q  K  P  G  K  A  P  K  L  L  I  Y  A  A
TCCAGTTTGCAAAGTGGGGTCCCATCAAGGTTCAGTGGCAGTGGATCTGGGACAGATTTC
 S  S  L  Q  S  G  V  P  S  R  F  S  G  S  G  S  G  T  D  F
ACTCTCACCATCAGCAGTCTGCAACCTGAAGATTTTGCAACTTACTACTGTCAACAGAGT
 T  L  T  I  S  S  L  Q  P  E  D  F  A  T  Y  Y  C  Q  Q  S
TACAGTACCCCCCTCACTTTCGGCGGAGGGACCAAAGTGGATATCAAACGTGCGGCCGCA
 Y  S  T  P  L  T  F  G  G  G  T  K  V  D  I  K  R  A  A  A
GTCGACGGGCGCGCCGCTTCCCATCACCATCACCATCAC
 V  D  G  R  A  A  S  H  H  H  H  H  H
```

APPENDIX C

Nucleotide and amino acid sequences of wild-type FHA1 domain and its variants

Wild-type FHA1 domain

```
ATGGAAAATATTACACAACCAACCCAACAATCCACCCAAGCAACTCAACGTTTTTTAATC
 M  E  N  I  T  Q  P  T  Q  Q  S  T  Q  A  T  Q  R  F  L  I
GAAAAATTTTCTCAAGAACAAATCGGCGAAAACATCGTATGCCGCGTAATCTGCACAACT
 E  K  F  S  Q  E  Q  I  G  E  N  I  V  C  R  V  I  C  T  T
GGTCAAATTCCGATCCGCGATCTCAGTGCAGATATCTCTCAGGTCTTAAAAGAAAAACGT
 G  Q  I  P  I  R  D  L  S  A  D  I  S  Q  V  L  K  E  K  R
AGCATTAAAAAAGTATGGACCTTTGGACGTAACCCAGCCTGCGATTATCACTTAGGTAAC
 S  I  K  K  V  W  T  F  G  R  N  P  A  C  D  Y  H  L  G  N
ATTAGCCGCTTATCTAATAAACACTTTCAAATCCTCCTGGGCGAAGACGGTAACTTATTA
 I  S  R  L  S  N  K  H  F  Q  I  L  L  G  E  D  G  N  L  L
CTCAACGACATCTCAACAAATGGTACATGGCTCAACGGTCAAAAAGTAGAAAAAAATAGC
 L  N  D  I  S  T  N  G  T  W  L  N  G  Q  K  V  E  K  N  S
AATCAATTACTCTCTCAAGGCGACGAAATTACGGTAGGCGTAGGTGTAGAAAGCGATATT
 N  Q  L  L  S  Q  G  D  E  I  T  V  G  V  G  V  E  S  D  I
TTAAGTTTAGTCATTTTTATTAACGATAAATTCAAACAATGCCTGGAACAAAATAAAGTC
 L  S  L  V  I  F  I  N  D  K  F  K  Q  C  L  E  Q  N  K  V
GATCGTATCCGTTCAAATCTTAAAAATACC
 D  R  I  R  S  N  L  K  N  T
```

The four cysteines, C34, C38, C74 and C154 are highlighted green

4C-4S variant

```
ATGGAAAATATTACACAACCAACCCAACAATCCACCCAAGCAACTCAACGTTTTTTAATC
 M  E  N  I  T  Q  P  T  Q  Q  S  T  Q  A  T  Q  R  F  L  I
GAAAAATTTTCTCAAGAACAAATCGGCGAAAACATCGTATCTCGCGTAATCTCAACAACT
 E  K  F  S  Q  E  Q  I  G  E  N  I  V  S  R  V  I  S  T  T
GGTCAAATTCCGATCCGCGATCTCAGTGCAGATATCTCTCAGGTCTTAAAAGAAAAACGT
 G  Q  I  P  I  R  D  L  S  A  D  I  S  Q  V  L  K  E  K  R
AGCATTAAAAAAGTATGGACCTTTGGACGTAACCCAGCCTCTGATTATCACTTAGGTAAC
 S  I  K  K  V  W  T  F  G  R  N  P  A  S  D  Y  H  L  G  N
ATTAGCCGCTTATCTAATAAACACTTTCAAATCCTCCTGGGCGAAGACGGTAACTTATTA
 I  S  R  L  S  N  K  H  F  Q  I  L  L  G  E  D  G  N  L  L
CTCAACGACATCTCAACAAATGGTACATGGCTCAACGGTCAAAAAGTAGAAAAAAATAGC
 L  N  D  I  S  T  N  G  T  W  L  N  G  Q  K  V  E  K  N  S
AATCAATTACTCTCTCAAGGCGACGAAATTACGGTAGGCGTAGGTGTAGAAAGCGATATT
 N  Q  L  L  S  Q  G  D  E  I  T  V  G  V  G  V  E  S  D  I
TTAAGTTTAGTCATTTTTATTAACGATAAATTCAAACAATCTCTGGAACAAAATAAAGTC
 L  S  L  V  I  F  I  N  D  K  F  K  Q  S  L  E  Q  N  K  V
GATCGTATCCGTTCAAATCTTAAAAATACC
 D  R  I  R  S  N  L  K  N  T
```

The four serines, S34, S38, S74 and S154 are highlighted blue

3C-3S variant

```
ATGGAAAATATTACACAACCAACCCAACAATCCACCCAAGCAACTCAACGTTTTTTAATC
 M  E  N  I  T  Q  P  T  Q  Q  S  T  Q  A  T  Q  R  F  L  I
GAAAAATTTTCTCAAGAACAAATCGGCGAAAACATCGTATCTCGCGTAATCTCAACAACT
 E  K  F  S  Q  E  Q  I  G  E  N  I  V  S  R  V  I  S  T  T
GGTCAAATTCCGATCCGCGATCTCAGTGCAGATATCTCTCAGGTCTTAAAAGAAAAACGT
 G  Q  I  P  I  R  D  L  S  A  D  I  S  Q  V  L  K  E  K  R
AGCATTAAAAAAGTATGGACCTTTGGACGTAACCCAGCCTGCGATTATCACTTAGGTAAC
 S  I  K  K  V  W  T  F  G  R  N  P  A  C  D  Y  H  L  G  N
ATTAGCCGCTTATCTAATAAACACTTTCAAATCCTCCTGGGCGAAGACGGTAACTTATTA
 I  S  R  L  S  N  K  H  F  Q  I  L  L  G  E  D  G  N  L  L
CTCAACGACATCTCAACAAATGGTACATGGCTCAACGGTCAAAAAGTAGAAAAAAATAGC
 L  N  D  I  S  T  N  G  T  W  L  N  G  Q  K  V  E  K  N  S
AATCAATTACTCTCTCAAGGCGACGAAATTACGGTAGGCGTAGGTGTAGAAAGCGATATT
 N  Q  L  L  S  Q  G  D  E  I  T  V  G  V  G  V  E  S  D  I
TTAAGTTTAGTCATTTTTATTAACGATAAATTCAAACAATCTCTGGAACAAAATAAAGTC
 L  S  L  V  I  F  I  N  D  K  F  K  Q  S  L  E  Q  N  K  V
GATCGTATCCGTTCAAATCTTAAAAATACC
 D  R  I  R  S  N  L  K  N  T
```

Compared to the wild-type FHA1 domain, the 3C-3S variant has three cysteine residues mutated to serines. C74 was not mutated to serine.

FHA1D2 variant

```
ATGGAAAATATTACACAACCAACCCAACAATCCACCCAAGCAACTCAACGTTTTTTAATC
 M  E  N  I  T  Q  P  T  Q  Q  S  T  Q  A  T  Q  R  F  L  I
GAAAAATTTTCTCAAGAACAAATCGGCGAAAACATCGTATTTCGCGTAATCTCAACAACT
 E  K  F  S  Q  E  Q  I  G  E  N  I  V  F  R  V  I  S  T  T
GGTCAAATTCCGATCCGCGATCTCAGTGCAGATATCTCTCAGGTCTTAAAAGAAAAACGT
 G  Q  I  P  I  R  D  L  S  A  D  I  S  Q  V  L  K  E  K  R
AGCATTAAAAAAGTATGGACCTTTGGACGTAACCCAGCCTGCGATTATCACTTAGGTAAC
 S  I  K  K  V  W  T  F  G  R  N  P  A  C  D  Y  H  L  G  N
ATTAGCCGCTTATCTAATAAACACTTTCAAATCCTCCTGGGCGAAGACGGTAACTTATTA
 I  S  R  L  S  N  K  H  F  Q  I  L  L  G  E  D  G  N  L  L
CTCAACGACATCTCAACAAATGGTACATGGCTCAACGGTCAAAAAGTAGAAAAAAATAGC
 L  N  D  I  S  T  N  G  T  W  L  N  G  Q  K  V  E  K  N  S
AACCAATTACTCTCTCAAGGCGACGAAATTACGGTAGGCGTAGGTGTAGAAAGCGATATT
 N  Q  L  L  S  Q  G  D  E  I  T  V  G  V  G  V  E  S  D  I
TTAAGTTTAGTCATTTTTATTAACGATAAATTCAAACAATCTCTGGAACAAAATAAAGTC
 L  S  L  V  I  F  I  N  D  K  F  K  Q  S  L  E  Q  N  K  V
GATCGTATCCGTTCAAATCTTAAAAATACC
 D  R  I  R  S  N  L  K  N  T
```

FHA1D2 was isolated from a phage-displayed library of FHA1 variants constructed by error-prone PCR using the 3C-3S variant as a starting scaffold. The 3C-3S variant with one amino acid substitution (S34F, highlighted in yellow), isolated after three rounds of affinity selection was named as FHA1D2.

FHA1G2 variant

```
ATGGAAAATATTACACAACCAACCCAACAATCCACCCAAGCAGCTCAACGTTTCTTAATC
 M  E  N  I  T  Q  P  T  Q  Q  S  T  Q  A  A  Q  R  F  L  I
GAAAAATTTTCTCAAGAACAAATCGGCGAAAACATCGTATTTCGCGTAATCTCAACAACT
 E  K  F  S  Q  E  Q  I  G  E  N  I  V  F  R  V  I  S  T  T
GGTCAAATTCCGATCCGCGATTTCAGTGCAGATATCTCTCAGGTCTTAAAAGAAAAACGT
 G  Q  I  P  I  R  D  F  S  A  D  I  S  Q  V  L  K  E  K  R
AGCATTAAAAAAGTATGGACCTTTGGACGCAACCCAGCCTGCGATTATCACTTAGGTAAC
 S  I  K  K  V  W  T  F  G  R  N  P  A  C  D  Y  H  L  G  N
ATTAGCCGCTTATCTAATAAACACTTTCAAATCCTCCTGGGCGAAGACGGTAACTTATTA
 I  S  R  L  S  N  K  H  F  Q  I  L  L  G  E  D  G  N  L  L
CTCAACGACATCTCAACAAATGGTACATGGCTCAACGGTCAAAAAGTAGAAAAAAATAGC
 L  N  D  I  S  T  N  G  T  W  L  N  G  Q  K  V  E  K  N  S
TACCAATTACTCTCTCAAGGCGACGAAATTACGGTAGGCGTAGGTGTAGAAAGCGATATT
 Y  Q  L  L  S  Q  G  D  E  I  T  V  G  V  G  V  E  S  D  I
TTAAGCTTAGTCATTTTTATTAACGATAAATTCAAACAATCTCTGGAACAAAATAAAGTC
 L  S  L  V  I  F  I  N  D  K  F  K  Q  S  L  E  Q  N  K  V
GATCGTATCCGTTCAAATCTTAAAAATACC
 D  R  I  R  S  N  L  K  N  T
```

FHA1G2 variant is the most thermal stable variant isolated from a phage-displayed library of FHA1D2 variants using high temperature as a selective pressure during affinity selection. Compared to FHA1D2, it has three additional mutations-T15A, L48F and N121Y (highlighted grey).

References

1. Qi, H., Lu, H., Qiu, H. J., Petrenko, V., and Liu, A. (2012) Phagemid vectors for phage display: properties, characteristics and construction, *J Mol Biol 417*, 129-143.

2. Goletz, S., Christensen, P. A., Kristensen, P., Blohm, D., Tomlinson, I., Winter, G., and Karsten, U. (2002) Selection of large diversities of antiidiotypic antibody fragments by phage display, *J Mol Biol 315*, 1087-1097.

3. Kristensen, P., and Winter, G. (1998) Proteolytic selection for protein folding using filamentous bacteriophages, *Fold Des 3*, 321-328.

4. Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K., and McCafferty, J. (2010) Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display, *Protein Eng Des Sel 23*, 279-288.

5. Pershad, K., Wypisniak, K., and Kay, B. K. (2012) Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display, *J Mol Biol*, DOI: 10.1016/j.jmb.2012.1009.1006.

6. Howarth, M., and Ting, A. Y. (2008) Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin, *Nat Protoc 3*, 534-545.

7. O'Callaghan C, A., Byford, M. F., Wyer, J. R., Willcox, B. E., Jakobsen, B. K., McMichael, A. J., and Bell, J. I. (1999) BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation, *Anal Biochem 266*, 9-15.

8. Kay, B. K., Thai, S., and Volgina, V. V. (2009) High-throughput biotinylation of proteins, *Methods Mol Biol 498*, 185-196.

9. Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., and Wood, K. V. (2008) HaloTag: a novel

protein labeling technology for cell imaging and protein analysis, *ACS chemical biology 3*, 373-382.

10. Los, G. V., and Wood, K. (2007) The HaloTag: a novel technology for cell imaging and protein analysis, *Methods Mol Biol 356*, 195-208.

11. Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S., Clarke, K. F., Conquer, J. S., Crofts, A. M., Crowther, S. R., Dyson, M. R., Flack, G., Griffin, G. J., Hooks, Y., Howat, W. J., Kolb-Kokocinski, A., Kunze, S., Martin, C. D., Maslen, G. L., Mitchell, J. N., O'Sullivan, M., Perera, R. L., Roake, W., Shadbolt, S. P., Vincent, K. J., Warford, A., Wilson, W. E., Xie, J., Young, J. L., and McCafferty, J. (2007) Application of phage display to high throughput antibody generation and characterization, *Genome biology 8*, R254.

12. Pershad, K., Sullivan, M. A., and Kay, B. K. (2011) Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats, *Anal Biochem 412*, 210-216.

13. Gao, J., Sidhu, S. S., and Wells, J. A. (2009) Two-state selection of conformation-specific antibodies, *Proc Natl Acad Sci U S A 106*, 3071-3076.

14. Weaver-Feldhaus, J. M., Miller, K. D., Feldhaus, M. J., and Siegel, R. W. (2005) Directed evolution for the development of conformation-specific affinity reagents using yeast display, *Protein Eng Des Sel 18*, 527-536.

VITA

NAME:	Kritika Pershad
EDUCATION:	B.S., Biochemistry, Microbiology and Chemistry, Osmania University, Hyderabad, India, 2002
	M.S., Biochemistry, Osmania University, Hyderabad, India, 2004
	Ph.D., Molecular Biology, University of Illinois at Chicago, Chicago, Illinois, 2012
TEACHING EXPERIENCE:	Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois: Cell Biology Laboratory, 2009
	Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois: Microbiology Laboratory, 2006-2007
	Department of Biochemistry, Aurora's Degree College, Hyderabad, India, 2004-2005
HONORS:	Dean's Scholar, University of Illinois at Chicago, Chicago, Illinois, 2011-2012
	Research Achievement, University of Illinois at Chicago, Chicago, Illinois, 2010, 2011
	Best Poster, Gordon Research Conference, Chemistry and Biology of Peptides, Ventura, California, 2010
	Excellence in Teaching Microbiology Laboratory, University of Illinois at Chicago, Chicago, Illinois, 2008
	Certificate of Excellence, Osmania University, Hyderabad, India, 1999-2002
ABSTRACTS	Pershad K., Sullivan M.A., Kay B.K.: A Vector Facilitating the

(POSTERS): Downstream Use of Phage Display Antibodies. Fourth Annual Protein Engineering Summit (PEGS), Cambridge Healthtech Institute, Boston, Massachusetts, 2008

Pershad K., Kay B.K.: The Forkhead-associated domain as a scaffold for generating new phosphopeptide binding specificities. Gordon Research Conference, Chemistry and Biology of Peptides, Ventura, California, 2010

Pershad K., Kay B.K.: Engineering the Forkhead-associated domain for phage display and improved thermal stability. Chicago Biomedical Consortium, Eighth Annual Symposium, Protein Folding and Misfolding in Health and Disease, Chicago, Illinois, 2010

ABSTRACTS (TALKS): Pershad K: The Forkhead-associated domain as a scaffold for generating new phosphopeptide binding specificities, FEBS Workshop, Protein Modules and Networks in Health and Disease, Seefeld in Tirol, Austria, 2009

Pershad K., Kay B.K.: Monitoring Signal Transduction Events with Designer Affinity Reagents. Invited to give a talk at Donelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada, 2012

PUBLICATIONS: Pershad, K., Wypisniak, K. & Kay, B. K. (2012). Directed evolution of the forkhead-associated domain to generate anti-phosphospecific reagents by phage-display. *J Mol Biol*, DOI: 10.1016/j.jmb.2012.09.006.

Pershad K., Kay B.K. Generating thermally stable variants of protein domains through phage-display. *Methods J*. In vitro Protein Selection issue (accepted for publication, the publication date of this issue is February/March 2013).

Pershad, K., Sullivan, M. A. & Kay, B. K. (2011). Drop-out phagemid vector for switching from phage displayed affinity reagents to expression formats. *Anal Biochem* 412, 210-6.

Huang, R., Pershad, K., Kokoszka, M., and Kay, B.K. Phage-Displayed Combinatorial Peptides. Amino Acids, Peptides and Proteins in Organic Chemistry, Vol.4 – Protection Reactions, Medicinal Chemistry, Combinatorial Synthesis. Copyright 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, ISBN: 978-3-527-32103-2.

Dyson, M. R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B. K., Pawson, T. & McCafferty, J. (2011). Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *Anal Biochem* 417, 25-35.

Colwill K; Renewable Protein Binder Working Group, & Graslund, S. (2011). A roadmap to generate renewable protein binders to the human proteome. *Nat Methods* 8, 551-8.

Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., Schofield, D. J., Dyson, M. R., Pawson, T., Kay, B. K. & McCafferty, J. (2010). Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display. *Protein Eng Des Sel* 23, 279-88.