Title
Function and specificity of SH3 domains in Saccharomyces cerevisiae

Author
Zarrinpar, Ali,

Publication Date
2003

Peer reviewed|Thesis/dissertation

Function and Specificity of SH3 Domains in *Saccharomyces cerevisiae*

by

Ali Zarrinpar

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

To my parents.

It is the least I can do.

ACKNOWLEDGEMENTS:

No useful work happens without the contributions of many, and this is no exception. This work is the direct result of the creative insights and the careful analysis of Wendell Lim. What is less visible is the care and consideration he expended in teaching how to carry out good experiments, how to attain creative insights, and how to perform careful analyses. Much gratitude and appreciation must also go to his partners in doing the above, lab members past and present who have encouraged, enabled, inquired, and ridiculed, mostly for the betterment of the projects.

I wish I could have thanked Ira Herskowitz more for all his guidance, help, inspiration, and intuition all of which he gave in his very unassuming manner. Many thanks to Sean O'Rourke, not only for his yeast strains, knowledge base, and establishing much of the background, but also for his cheering on.

Ideas and criticism were not the only things doled out by my thesis committee. Within that context and without, Erin O'Shea and Keith Yamamoto, each in their own right and in their own way taught me very much about the practice of and the life in science.

What remains of my sanity I owe to my family and friends who dealt with my serious miscalculations of when my lab work would be finished and the tardiness that always ensued: Annie Countryman for her indomitable verve and endless kindness, my MSTP classmates Greg Ku, Aylin Rodan, and Michael Rothenberg for being there (for a long time), Markus Hägele for being funny, and Connie You for just not buying it sometimes.

I would also like to acknowledge the Herbert W. Boyer PIBS Tetrad and

the Medical Science Training programs at UCSF for their generous support.

Portions of the text and figures presented in this thesis are reproduced with permission from material published previously. Chapter 1, "Introduction: Role of SH3 Domains in the Assembly of Signaling Complexes" was published in *Science STKE* **2003**, RE8 (2003). Roby P. Bhattacharyya was a co-author on that paper and wrote the sections on the EVH1 domain and other non-SH3, non-WW proline rich peptide binding domains. Chapter 2, "Optimization of Specificity Within a Protein Interaction Network Through System-Wide Negative Selection", has been submitted for publication at *Nature*. This paper owes much to Sang-Hyun Park for his scientific contributions. Chapter 3, "Redundant scaffolds Sho1 and Pbs2 direct activity and specificity in the yeast osmoregulatory MAPK pathway," benefited from the scientific and technical contributions of Roby P. Bhattacharyya and M. Paige Nittler. With the exception of those items listed above, the work presented in this thesis was performed by its author, Ali Zarrinpar, under the supervision of Dr. Wendell Lim.

Function and Specificity of SH3 Domains in *Saccharomyces cerevisiae*

by

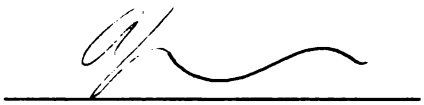
Ali Zarrinpar

ABSTRACT

Specific protein-protein interactions are essential to cellular signaling, but they are achieved is not yet apparent. Many protein-protein interactions are mediated by binding domains. Because there are many domains of the same family within a single proteome, it is generally thought that isolated domains cannot specify unique, biologically relevant interactions. Instead, specificity is hypothesized to be encoded in the context in which the domain is presented (flanking domains, co-localization, etc.). We show that a proline-rich motif from the yeast protein Pbs2 recognizes its biologically relevant partner, the Src Homology 3 (SH3) domain from Sho1, with near absolute specificity--none of the other 26 yeast SH3 domains cross-reacts with the Pbs2 ligand, *in vivo* or *in vitro*. This high level of specificity, however, is not observed among a set of non-yeast SH3 domains, suggesting that the interaction has been optimized through negative selection against cross-reactivity. Thus, in this case, negative selection provides a mechanism to optimize interaction specificity in a network of highly overlapping recognition domains.

On another level is the problem of achieving signaling specificity when the same protein is involved in multiple pathways. This is the case with the kinase Ste11, which acts in three different signaling mitogen activated protein kinase (MAPK) cascades in *Saccharomyces cerevisiae*. Despite this fact, each pathway maintains specificity and limits cross-activation of the other pathways through the effects of scaffolding proteins. MAPK kinase Pbs2 has been proposed to act as a scaffold in the osmo-response pathway because it interacts with Sho1, Ste11, and Hog1. We show that the membrane protein Sho1 also functions as a scaffold and that it determines the flow of the pathway by interacting with Ste20, Ste11, Pbs2, and

Hog1. Direct interaction with Sho1 is necessary for the osmo-responsive activation of Ste11. The interaction of Sho1 with Pbs2 is required to direct the activity of Ste11 towards the HOG pathway upon osmoshock. Multiple interactions lead to the formation of complexes including Sho1, Ste20, and Ste11 in one case and Sho1, Pbs2, and Hog1 in another, hinting at a step-wise assembly of the osmolarity MAPK module by Sho1.



Ali Zarrinpar
Doctoral Candidate

Wendell A. Lim, Ph.D.
Thesis Advisor

TABLE OF CONTENTS

LIST OF TABLES:

LIST OF FIGURES:

Chapter 1

Introduction:

The Structure and Function of Proline Recognition Domains

Ali Zarrinpar, Roby P. Bhattacharyya, Wendell A. Lim

Introduction

Domains that bind proline-rich motifs are critical to the assembly of many intracellular signaling complexes and pathways. The importance of proline-rich motifs in biology is highlighted by the finding that "proline-rich regions" (InterPro) are the most common sequence motif in the *Drosophila* genome and the second-most common in the *Caenorhabditis elegans* genome (Rubin et al., 2000). The number of defined protein domains that recognize proline-rich motifs has expanded considerably in recent years to include such common motifs as Src Homology 3 (SH3), WW (named for a conserved Trp-Trp motif), and Enabled/VASP Homology (EVH1, also known as WASP Homology 1 or WH1) domains, as well as other proline-binding domains. The number of domains in an organism roughly corresponds to its perceived complexity (Table 1.1).

Proline recognition domains are usually found in the context of larger multi-domain signaling proteins. Their binding events often direct the assembly and targeting of protein complexes involved in cell growth (Buday and Downward, 1993; Lowenstein et al., 1992; Rozakis-Adcock et al., 1993), cytoskeletal rearrangements (Holt and Koffer, 2001; Renfranz and Beckerle, 2002), transcription (Sudol et al., 2001), postsynaptic signaling (Ball et al., 2002; Tu et al., 1998), and other key cellular processes (McPherson, 1999). In addition, these interactions can play a regulatory role, often through autoinhibitory interactions that are alleviated by competing binding events (Nguyen and Lim, 1997).

More detail about the individual proline recognition domains can be found in several recent reviews (Ball et al., 2002; Macias et al., 2002; Mayer, 2001).

Organism	SH3	WW	EVH1	GYF
S. cerevisiae	25	5	1	3
C. elegans	66	18	2	3
D. melanogaster	90	27	5	2
M. musculus	163	39	16	2
H. sapiens	332	80	20	5

Table 1.1. Abundance of proline recognition domains. The number of proteins with proline recognition domains in some commonly studied eukaryotic organisms, as found in the Pfam homology database, is shown. Those listed in the table are meant only to reflect the relative abundance in each proteome; different numbers are obtained from other domain identification databases. SH3-like domains are found in some prokaryotes. They are not included in the table because they lack certain key conserved residues, and the structure and function of these domains are unknown.

This chapter aims to compare the biological role and the molecular mechanisms of these domains and to address the implications of having multiple domains with similar ligand specificities within a single cell.

Properties of Proline and Polyproline Sequences

Repetitive proline-rich sequences are found in many proteins (MacArthur and Thornton, 1991) and in many cases are thought to function as docking sites for signaling modules (Kay et al., 2000). Why might proline be singled out for recognition by so many key protein-protein interaction modules? Several features of proline distinguish it from the other 19 naturally occurring amino acids (Fig. 1.1A): the unusual shape of its pyrrolidine ring, the conformational constraints upon its dihedral angles imposed by this cyclic side chain, its resulting secondary structural preferences, its substituted amide nitrogen, and the relative stability of the *cis* isomer in a peptide bond. Each recognition domain exploits some combination of these distinctive features of proline in order to achieve specific binding to proline-rich regions.

One feature of proline-rich motifs frequently utilized in binding to signaling domains is their propensity to form a polyproline type II (PP II) helix. The PPII helix is an extended left-handed helical structure with three residues per turn and an overall shape resembling a triangular prism (Fig. 1.1B) (MacArthur and Thornton, 1991; Williamson, 1994). A combination of steric and hydrogen-bonding properties of proline-rich motifs is thought to contribute to its preference for this unusual secondary structure (MacArthur and Thornton, 1991; Williamson, 1994). Two features of the PPII helix make it a useful recognition motif. First, in this structure both the side chains and the backbone carbonyls

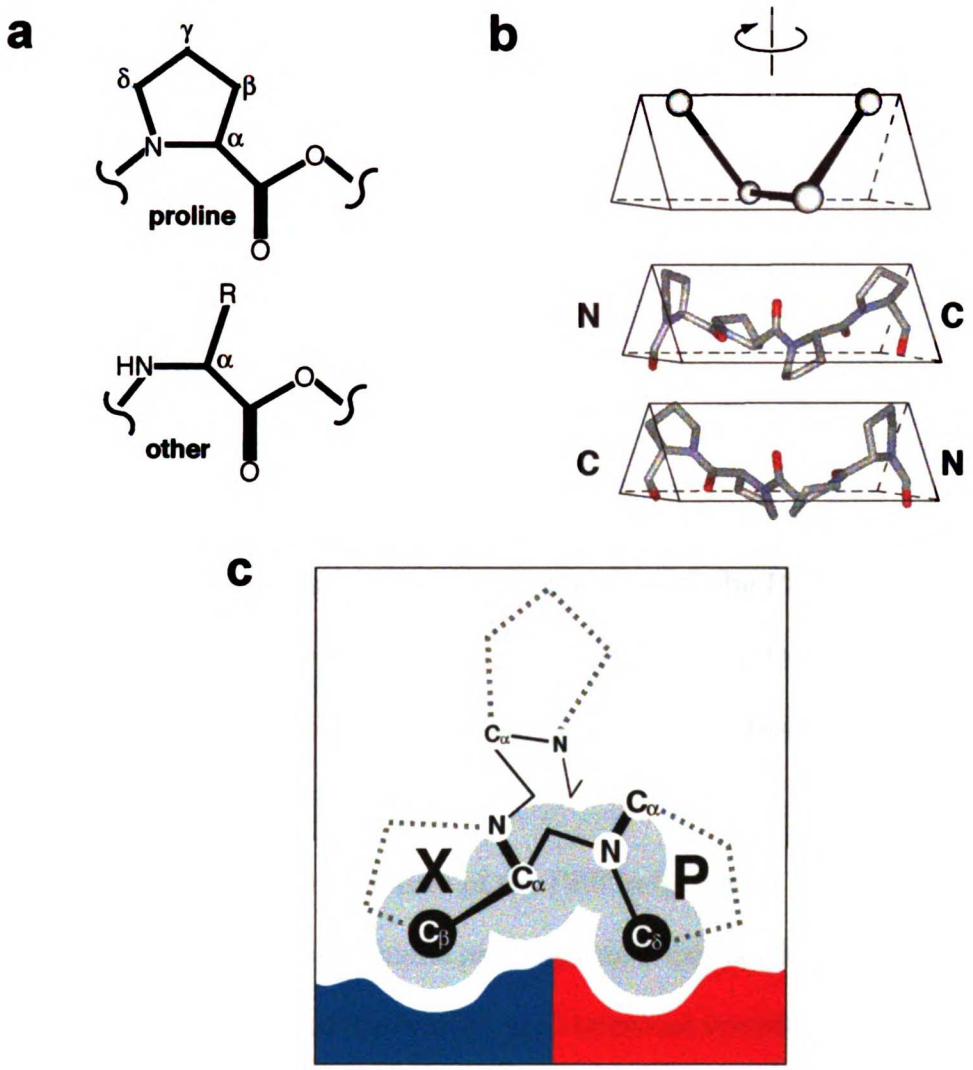


Fig. 1.1. Properties of proline and polyproline sequences. (A) Chemical structure of proline contrasted with other natural amino acids. Proline possesses a five-member ring fused onto the nitrogen making it a secondary amine, whereas other amino acids have sidechains that only branch off the α-carbon leaving a primary amine. (B) Schematic and structural representation of a polyproline II (PPII) helix. The helix has two-fold pseudosymmetry: A rotation of 180 degrees about a vertical axis leaves the proline rings and the carbonyl oxygens at approximately the same position. The PDB accession code for the poly-(l)-proline structure shown is 1CF0. (C) A view down the axis of the PPII helix highlighting the position of the carbons in the xP dipeptide. In the "x" position that requires C-substitution (blue), the primary recognition element is the β-carbon, while in the "P" position that requires N-substitution (red), the primary recognition element is the δ-carbon that is unique to proline.

point out from the helical axis into solution at regular intervals (Fig. 1.1B). The lack of intramolecular hydrogen bonds in the PPII structure, due largely to the absence of a backbone hydrogen-bond donor on proline, leaves these carbonyls free to participate in intermolecular hydrogen bonds. Thus, both side chains and carbonyls can easily be "read" by interacting proteins (Siligardi and Drake, 1995). Second, because the backbone conformation in a PPII helix is already restricted, the entropic cost of binding is reduced (Kay et al., 2000; Petrella et al., 1996). Nearly all of the domains described here bind their ligands in a PPII conformation. Interestingly, many of the interactions with the PPII helical ligand involve aromatic residues. The planar structure of aromatic side chains appears to be highly complementary to the ridges and grooves presented on the PPII helix surface.

One interesting structural feature of the PPII helix is that it has two-fold rotational pseudosymmetry: side chains and backbone carbonyls are displayed with similar spacing in either of the two N- to C-terminal orientations (Fig. 1.1B). This feature may explain why many proline-binding domains are observed to bind ligands in two possible orientations, a property unique among characterized peptide recognition modules. In principle, this flexibility could play an important role in domain function. For example, binding in one orientation could be activating, whereas binding in the other orientation could be inhibitory. However, this role has not been demonstrated.

Another unique property of proline is that it is the only naturally occurring *N*-substituted amino acid. Proteins that recognize the d-carbon on the substituted amide nitrogen (Fig. 1.1A) within the context of the otherwise

standard peptide backbone can select precisely for proline at a given position without making extended contacts with the rest of the side chain (Fig. 1.1C). Thus, sequence-specific recognition can be achieved without requiring a particularly high affinity interaction. Interactions that are specific and low-affinity can be quite useful in intracellular signaling environments where rapidly reversible interactions may be required. Among proline-binding domains, this phenomenon has been best characterized for SH3 domains, in which required prolines can be replaced without a significant loss in binding affinity by a number of non-natural *N*-substituted amino acids that do not resemble proline (Nguyen et al., 1998).

Proline also stands out from other natural amino acids in its ability to exist stably as a *cis* isomer about the peptide bond. In an unfolded chain, proline residues adopt the *cis* conformation with a probability of ~20% compared to negligible amounts for the other amino acids (MacArthur and Thornton, 1991). Moreover, the kinetic barrier for *cis-trans* isomerization is higher for proline than for the other amino acids and is even the rate-limiting step in the folding of certain proteins (Wedemeyer et al., 2002). In principle, recognition of *cis* proline moieties could be a useful way of achieving regulation, potentially even with some degree of kinetic control. However, none of the major proline recognition modules discussed here are known to exploit recogntion of *cis* isomers. Still, the intriguing possibility remains that *cis-trans* isomerization could provide a mechanism to modulate such recognition events.

Thus, many chemical properties of proline distinguish it from the other 19 naturally occurring amino acids, and proline recognition domains exploit several

of these properties. If a recognition event involves a property of proline that is sufficiently distinct among the natural set of 20 amino acids, the interaction does not have to be of particularly high affinity to be selective. The benefits of weak, but specific, interactions in intracellular signaling pathways may help explain the abundance of proline-based recognition motifs.

SH3 Domains

The first characterized and best understood example of the proline recognition modules is the SH3 domain (Mayer, 2001). SH3 domains comprise about 60 residues and typically play an assembly or regulatory function. An assembly role is exemplified by the adaptor protein Grb2, which is involved in the p21 Ras-dependent growth factor signaling pathway (Fig. 1.2A) (Lowenstein et al., 1992). Grb2 has a single Src Homology 2 (SH2) domain, which recognizes phosphotyrosine motifs, flanked by two SH3 domains. Upon growth factor stimulation, receptor tyrosine kinase activation results in autophosphorylation and phosphorylation of other membrane-associated proteins. These phosphorylation events create docking sites for the Grb2 SH2 domain, thereby resulting in membrane recruitment of Grb2. The Grb2 SH3 domains bind to proline-rich motifs in the protein SOS, a guanine nucleotide exchange factor for Ras, ultimately recruiting SOS to the membrane. Because Ras is myristoylated and membrane localized, this colocalization with SOS promotes guanosine triphosphate (GTP) loading of Ras. The resultant stimulation of Ras activates a mitogen-activated protein kinase (MAPK) cascade, leading to cell growth and differentiation (Buday and Downward, 1993; Rozakis-Adcock et al., 1993). Similar recruitment roles are played by SH3 domain-containing proteins in

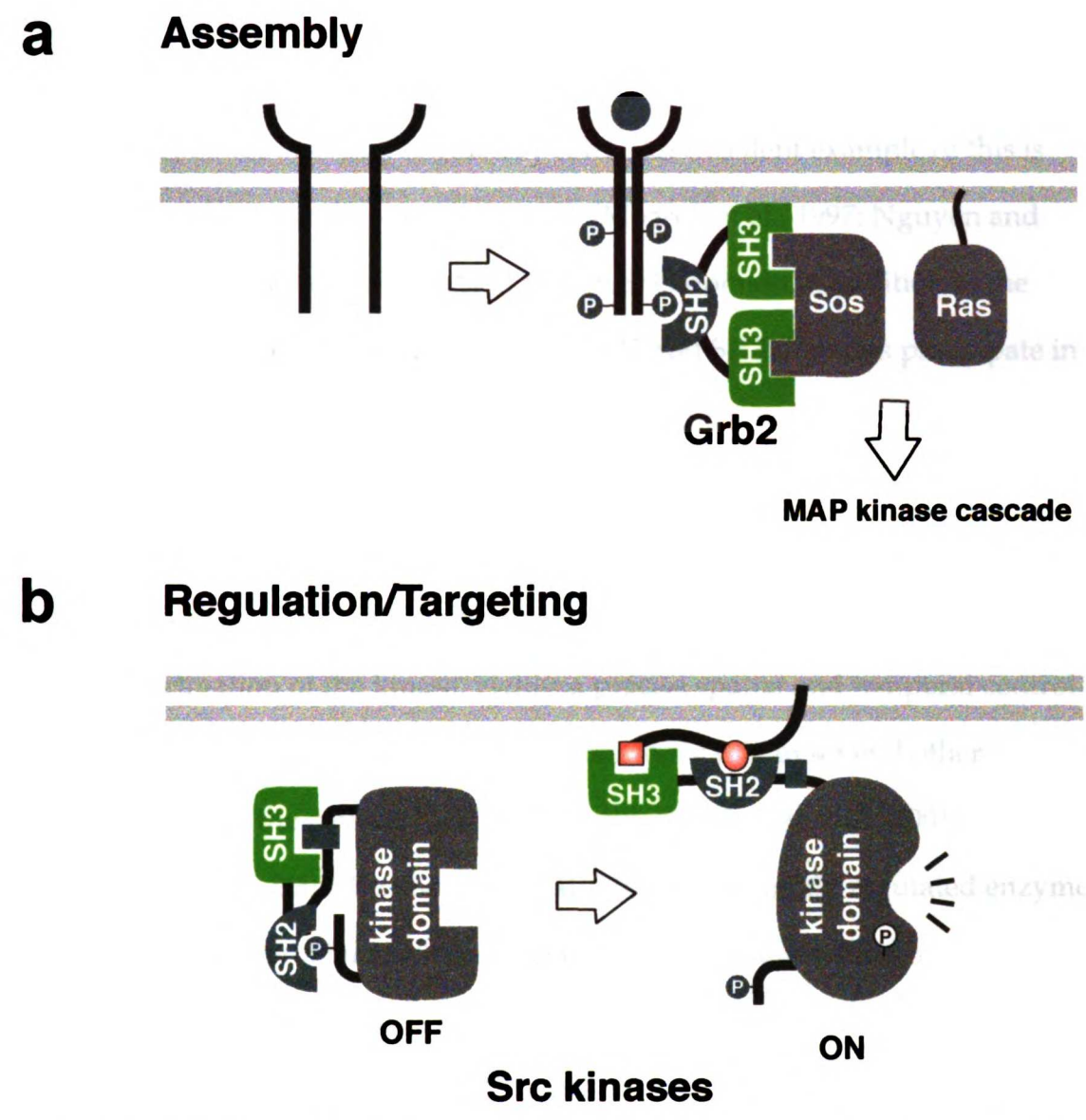


Fig. 1.2. Functional roles of SH3 domains. (A) Assembly role of SH3 domains. Growth factor stimulation leads to the activation of receptor tyrosine kinases and phosphorylation of the receptor tail, related adaptor proteins (not shown), or both. The resultant phosphotyrosines form docking sites for the adaptor protein Grb2 (through its SH2 domain). The Grb2 SH3 domains bind proline-rich motifs in SOS, the guanine nucleotide exchange factor for Ras, recruiting SOS to the membrane and colocalizing it with Ras. The resultant stimulation of Ras activates a mitogen-activated protein kinase (MAPK) cascade, leading to cell growth and differentiation. (B) Regulatory role of SH3 domains. Intramolecular interactions of the SH2 and SH3 domains of Src kinases hold their kinase domains in an inactive conformation. These autoinhibitory interactions can be disrupted by external SH2 and SH3 ligands, yielding spatial and temporal control of kinase activation.

various other biological processes, including endocytosis (McPherson, 1999) and cytoskeletal dynamics (Buday et al., 2002).

SH3 domains also play regulatory roles. An excellent example of this is the Src family of tyrosine kinases (Fig. 1.2B) (Moarefi et al., 1997; Nguyen and Lim, 1997). Src kinases contain an SH2 and an SH3 domain in addition to the kinase domain. Under basal conditions, the SH2 and SH3 domains participate in intramolecular interactions that hold the kinase domain in an inactive conformation. Binding to external SH2 and SH3 ligands can disrupt these autoinhibitory interactions, thereby yielding activation. An important feature of such a regulatory role is that targeting by the SH2 and SH3 domains is directly coupled to activation of the kinase, yielding precise spatial and temporal control. SH3 domains appear to play a similar autoinhibitory role in several other systems, including the neutrophil NADPH oxidase (Hiroaki et al., 2001; Karathanassis et al., 2002; Kuribayashi et al., 2002). This tightly regulated enzyme produces the antimicrobial reactive oxygen species only upon proper stimulation. Activation involves the assembly and membrane localization of the SH3-containing proteins $p40^{phox}$, $p47^{phox}$, and $p67^{phox}$.

Such regulatory mechanisms reveal how SH3 domains, which were initially viewed as static assembly elements, can function as dynamic switches by alternating binding partners (intra- versus intermolecular). Interestingly, SH3 interactions tend to be fairly weak, with typical dissociation constants (K_ds) in the mM range (Mayer, 2001). Such weak affinities may be essential for this kind of reversible switching mechanism.

Much effort has been dedicated to understanding the ligand preferences of SH3 domains (Cesareni et al., 2002; Feng et al., 1995; Rickles et al., 1994; Sparks et al., 1996). In vitro peptide selection studies revealed that the majority of SH3 domains require the conserved consensus motif PxxP for recognition. In individual SH3 domains, however, this core PxxP motif is flanked by different specificity elements. For example, a large group of SH3 domains recognize the PxxP core flanked by the basic residues R or K. However, early studies were confounded by the observation that two classes of such ligand motifs emerged--K/RxxPxxP and PxxPxK/R (where K or R are required flanking residues and x is any amino acid). This confusion was clarified by structural studies that revealed that SH3 domains could use a single recognition surface to bind ligands in two possible N- to C-terminal orientations (Feng et al., 1994; Lim and Richards, 1994; Lim et al., 1994; Yu et al., 1994). Each of these two recognition motifs corresponds to the sequence preferences for a distinct orientation of binding. Efforts are underway to utilize the extensive peptide library data to generate algorithms to predict SH3 recognition (Brannetti et al., 2000; Cesareni et al., 2002; Wollacott and Desjarlais, 2001).

Structures of SH3 domains both alone and in complex with ligand reveal their mechanism of recognition (Fig. 1.3). The SH3 fold consists of two antiparallel β sheets at right angles to one another. Within this fold are two variable loops, referred to as the RT and the n-Src loops (Musacchio et al., 1992; Ren et al., 1993). When bound, the proline-rich peptide ligand adopts a PPII helix conformation (Lim et al., 1994; Terasawa et al., 1994; Yu et al., 1994). Recognition of this structure is achieved by insertion of the ridges of the PPII helix into a

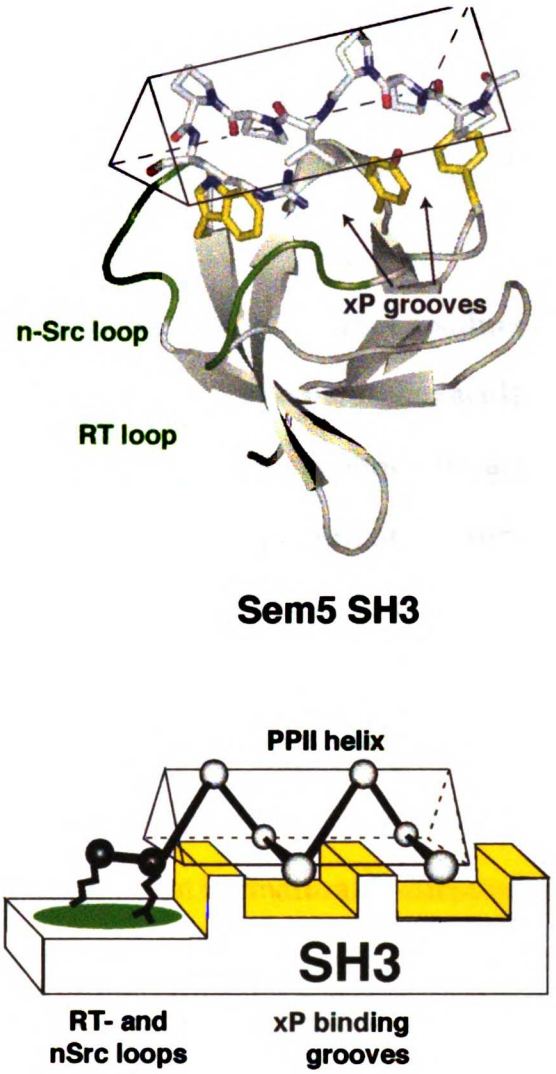


Fig. 1.3. Structure and binding mechanism of SH3 domains. Structure of the Sem5 SH3 domain in complex with a proline-rich ligand. A cartoon of the proline-binding surface of these domains docked with a ligand, showing the general mechanism of recognition, is shown below. The core recognition surface has two xP binding grooves formed by aromatic amino acids, shown in yellow, and the adjacent, less conserved specificity pockets are designated in green. The PDB accession code for this structure is 1SEM.

complementary pair of grooves on the SH3 surface. These surface grooves are defined by a series of nearly parallel, well-conserved aromatic residues. In addition, hydrogen-bonding donors are well positioned to recognize ligand backbone carbonyl moieties.

Each groove actually recognizes a pair of residues of the sequence xP (where x is a variable, usually hydrophobic, amino acid). This mode of recognition explains the requirement for prolines. Because the xP dipeptide unit has the unique backbone substitution pattern of a C-substituted residue followed by an N-substituted residue, it forms a relatively continuous ridge that can pack efficiently into the aromatic grooves on the SH3 surface (Fig. 1.1C). Because this mechanism relies only on the N-substitution of proline and not the entire proline ring, it allows recognition to be highly selective without being of high affinity. Moreover, it has been shown that nonnatural *N*-substituted groups can be used to make synthetic SH3 inhibitors (Nguyen et al., 1998). This mode of recognition also explains why SH3 domains can bind ligands in two possible orientations--a PPII ligand has two-fold rotational pseudosymmetry, both with respect to the steric properties of the xP unit and presentation of hydrogen-bonding groups (the backbone carbonyls) that are used in recognition (Fig. 1.1B).

Adjacent to the core recognition surface of SH3 domains are the more variable RT and n-Src loops (Fig. 1.3). In many cases, residues in these loops are observed to make numerous unique interactions with key residues in the ligand that flank the PxxP core. Thus, in general, these loops can be considered to form a flanking specificity pocket. The specificity provided by these pockets has been explored through both phage display techniques and combinatorial synthetic

strategies (Feng et al., 1995; Kapoor et al., 1998). These studies show that there is sufficient variability in these pockets to allow for some differential binding among SH3 family members.

Despite having distinct specificity pockets, many SH3 domains appear to have highly overlapping recognition profiles. For example, a large majority of SH3 domains recognize R/KxxPxxP or PxxPxR/K motifs (Sparks et al., 1996; Tong et al., 2002). Thus, an unanswered question is how specificity within SH3 domain-mediated interaction networks is achieved, especially in cells and organisms with many SH3 domains. One solution, utilized by a handful of SH3 domains, is the evolution of a noncanonical recognition mechanism. Several SH3 domains recognize non-PxxP motifs. This is the case for the SH3 domains of Eps8, which recognizes PxxDY (Mongiovi et al., 1999); Gads, which recognizes RxxK (Berry et al., 2002); and Fus1, which recognizes Arg-Ser-rich sequences (Tong et al., 2002). In most of these cases, it is unclear whether this novel recognition is mediated by the equivalent surface used by canonical SH3 domains to recognize PxxP ligands. The only structurally characterized domain from the list above, Eps8, defines a sub-family of SH3 domains that are domain swapped dimers. Another class of unusual SH3 domains is found in membrane-associated guanylate kinases (MAGUKs). MAGUK SH3 domains do not appear to bind PxxP motifs, but instead can associate with an adjacent guanylate kinase domain in an intra- or intermolecular fashion (McGee et al., 2001). This interaction may play a role in the assembly of signaling complexes at cell-cell junctions. One more interesting example of noncanonical recognition is the

interaction between the N-terminal SH3 domain of Vav and the C-terminal SH3 domain of Grb2 (Nishida et al., 2001).

Several other mechanisms may contribute to enhancing specificity in SH3 domain-mediated interactions. There may be tertiary structure elements involved in recognition, as is the case for the recognition of the human immunodeficiency virus (HIV) protein Nef by the SH3 domains of Src family kinases Hck, Fyn, and Lyn (Lee et al., 1996; Moarefi et al., 1997; Saksela et al., 1995). Nef presents a canonical PPII core in the context of a folded structure. Thus, there are additional interactions between other parts of Nef with unique elements in the RT loops of these SH3 domains.

Specificity and affinity enhancements may also come from combinatorial recognition by multiple recognition domains working in concert. There are many examples of proteins containing multiple SH3 domains, such as the yeast proteins Bem1 and Sla1 (Tong et al., 2002) or the above examples of Grb2 (Yuzawa et al., 2001) and $p47^{phox}$. Moreover, SH3 domains could function together with other modules such as SH2, PDZ (named after signaling proteins PSD-95, Dlg, and ZO-1), or EVH1 domains that are often found in the same proteins or complexes.

Additionally, some SH3 domains participate in multiple interactions (Kami et al., 2002). For example, the SH3 domain from the yeast protein Pex13 has two binding surfaces: a canonical surface that binds a PxxP ligand from Pex14 and a second surface that binds a nonproline motif in Pex5 (Douangamath et al., 2002). This set of distinct interactions achieved by the Pex13 SH3 domain is thought to reinforce the assembly of the specific trimeric complex. Several other

SH3 domains also appear to have binding surfaces distinct from their proline binding interface (Nishida et al., 2001).

What has not been explored until now has been the effect of limiting the number of potential interacting proteins. Not only can this be done through cellular compartmentalization and transcriptional regulation but also through manipulating the sequences of the proteome in general. As we show in the following chapters, the isolated proline-rich motif from the yeast protein Pbs2 recognizes its biologically relevant partner, the Src Homology 3 (SH3) domain from Sho1, with near absolute specificity--none of the other 26 yeast SH3 domains cross-reacts with the Pbs2 ligand, *in vivo* or *in vitro*. This high level of specificity, however, is not observed among a set of non-yeast SH3 domains, suggesting that the Pbs2 ligand motif has been optimized through negative selection against cross-reactivity with competing domains within the yeast proteome. System-wide negative selection is a further mechanism to optimize interaction specificity among a network of highly overlapping recognition domains.

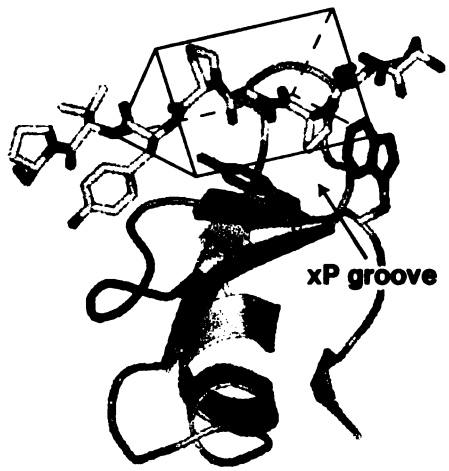
WW Domains

WW domains mediate protein-protein interactions in diverse processes (Macias et al., 2002). For example, the WW domains of the ubiquitin ligase Nedd4 bind to Na^+-channel subunits, thereby targeting ubiquitin-mediated down-regulation of channel activity (Farr et al., 2000). A mutation in the recognition motif on the Na^+-channel subunit, as occurs in the human disease Liddle's syndrome, increases the number of Na^+ channels in the membrane,

leading to increased blood pressure. WW domains are found in several ubiquitin ligases that bind to other targets (Sudol et al., 2001). In addition, pre-messenger RNA (mRNA) splicing involves an interaction between the WW domains in the splicing factor PRP40 and a proline-rich region in the branchpoint-binding protein BBP. Another example of a biologically important role of WW domains is the organization of the dystrophin-syntrophin-β-dystroglycan complex (Huang et al., 2000; Ilsley et al., 2002).

WW domains can be divided into several classes based on recognition motifs (Sudol and Hunter, 2000). All recognize proline-containing motifs that are distinct from, though overlapping with, SH3 domains. For example, the WW domains from the Yes-associated protein YAP65 and dystrophin prefer the motif Pro-Pro-X-Tyr (PPxY) (Huang et al., 2000; Macias et al., 1996); the FBP11 and FE65 WW domains prefer Pro-Pro-Leu-Pro (PPLP) (Bedford et al., 1998); and the FBP21, FBP30, and Npw38 WW domains prefer Pro-Arg (P-R) repeats (Bedford et al., 2000a; Bedford et al., 2000b). Interestingly, phosphorylation can play an important negative or positive regulatory role in WW domain recognition. For example, the WW domains of the mitotic peptidyl prolyl isomerase (PPIase) Pin1 and the ubiquitin ligase Nedd4 bind specifically to phospho-Ser/Thr-Pro motifs, but not their unphosphorylated counterparts. In contrast, interactions with PPxY motifs can be abolished by tyrosine phosphorylation (Lott et al., 2002; Lu et al., 1999; Verdecia et al., 2000).

The structures of WW domain-ligand complexes reveal a striking mechanistic similarity to those of SH3s and other proline recognition domains (Fig. 1.4) (Zarrinpar and Lim, 2000). Containing 35 to 45 residues, WW domains



dystrophin WW

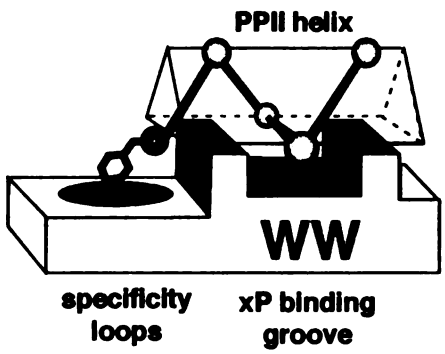


Fig. 1.4. Structure and binding mechanism of WW domains. Structure of the dystrophin WW domain in complex with a proline-rich ligand. A cartoon of the proline binding surface of these domains docked with a ligand, showing the general mechanism of recognition, is shown below. The core recognition surface has one xP binding groove formed by aromatic amino acids (yellow) and adjacent, less conserved specificity pockets (green). The PDB accession code for this structure is 1EG4.

are highly compact binding domains, comprising an antiparallel three-stranded fold (Macias et al., 1996). Like SH3 domains, their binding surfaces are composed of a series of nearly parallel aromatic residues. Correspondingly, their ligands adopt PPII helices that position the proline sidechains against the ridges and grooves on the domain binding surface (Huang et al., 2000; Verdecia et al., 2000). The aromatic groove in the WW domain also recognizes an xP pair in the ligand core. A consequence of this common mode of proline recognition is that WW domains, like SH3 domains, can recognize their ligands in two opposite orientations. WW domains differ from SH3 domains in that they typically have only one xP binding groove compared to two adjacent xP binding grooves found in SH3 domains. Thus, a shorter proline-rich core is required for WW domain recognition.

How then, outside the requirement for the xP core, do WW domains achieve specific recognition of their ligands? Like SH3 domains, WW domains use variable loops and neighboring domains to enhance specificity. The WW domain fold has two variable loops that are adjacent to the aromatic xP-binding groove. These loops are observed to participate in interactions with key specificity elements including the required phospho-Ser residues within the proline-rich motif bound by Pin1 or the nonphosphorylated Tyr residue within the PPxY motif bound by the dystrophin WW domain. This mechanism of specificity is conceptually similar to that used by the n-Src and RT loops of SH3 domains.

Multiple cooperative interactions with neighboring domains can also contribute to specificity in WW domain-mediated recognition. The interaction of

dystroglycan with dystrophin requires both the WW domain and an adjacent helical EF hand-like domain (EF domains are calcium-binding domains). The two domains form a contiguous recognition surface where approximately half of the dystroglycan peptide ligand contacts only the EF domain. The structure of Pin1 in complex with a phosphopeptide also shows significant contacts between the ligand and the adjacent PPIase domain.

Conclusions

The domains discussed here recognize proline-containing motifs by focusing on unique chemical properties of proline and proline-rich sequences. These recognition mechanisms take advantage of the fact that proline is chemically distinct from the other 19 natural amino acids. Thus, these domains are similar to other recognition domains used in signaling, which often focus on a highly distinct recognition anchor like phosphoamino acids, as exemplified by SH2 and phosphotyrosine binding (PTB) domains, (Yaffe, 2002) or carboxy-termini, as exemplified by PDZ domains (Harris and Lim, 2001). Such features may simply stand out within the chemical milieu of the cell.

An advantage of focusing on such distinct chemical features is that such interactions can be discriminatory without resorting to extremely high affinities. The domains discussed here all tend to have dissociation constants ranging from high nM to low mM. Signaling pathways are often dynamic; they must be activated and inactivated quickly, and their interactions often involve domains switching between multiple interaction partners. Thus, these interactions cannot be so tight as to inhibit the dynamic nature of cellular processes.

Why are there so many proline-recognition domains? This abundance may be a simple result of the proliferation of a successful solution to the problem of protein recognition. Having more domain types presumably allows the evolution of more complex signaling networks. Further, having a suite of domains that recognize similar or overlapping motifs may provide additional modes of interaction regulation (Sudol, 1996). If domains from distinct family members recognize a single motif, the competition between these alternative partners could, in principle, act as a regulatory switch. Relatively little is known about the functional intersection between different domain families in vivo. However, in one case, T cell activation appears to promote this type of domain interaction swap: a receptor proline-rich motif that initially interacts with a GYF domain, after stimulation interacts with an SH3 domain (Freund et al., 2002).

The number of proline-binding domains, however, exacerbates the problem of selectivity: how are incorrect interactions avoided? Most domains discussed here have multiple mechanisms for recognizing ligands with higher specificity (Fig. 1.5). Almost all have specificity pockets flanking surfaces used to recognize a proline-rich core. A few have multiple binding sites on a single domain, which may facilitate more specific, cooperative assembly. In some cases it is clear that multiple domains work together to achieve specific recognition. Additionally, proteome-wide negative selection plays a role in generating specific binding partners. The molecular mechanisms by which multiple domains cooperate to achieve biologically specific functions remains one of the major questions concerning these and other recognition modules.

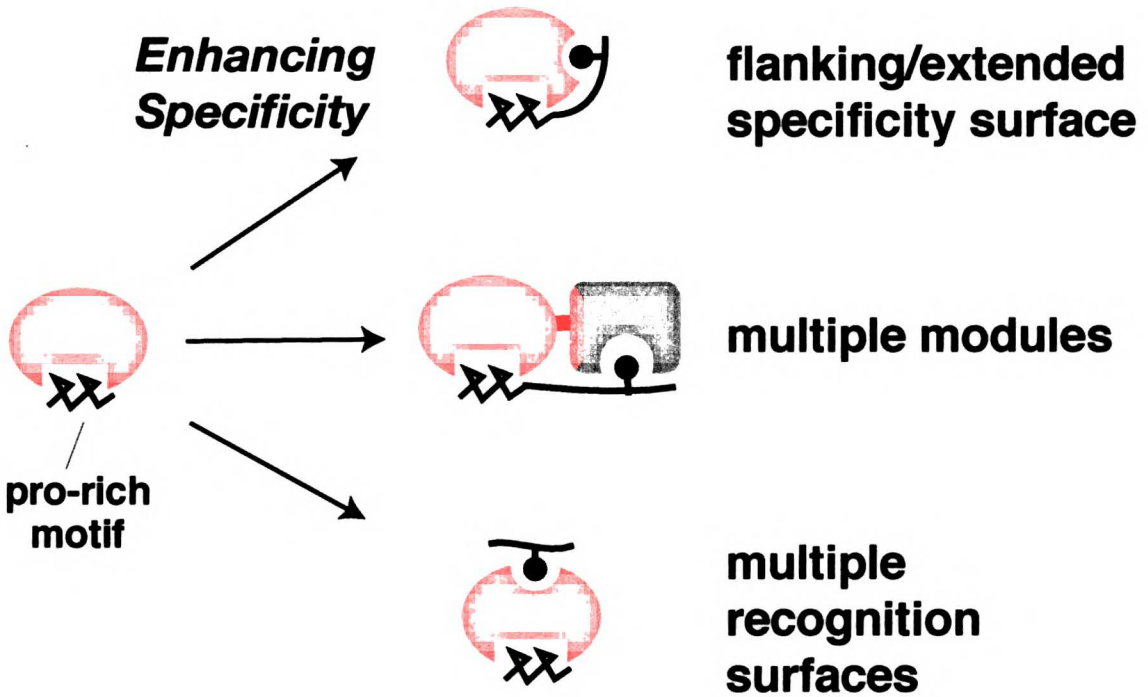


Fig. 1.5. Potential mechanisms for enhancing specificity of proline binding domains. One means of increasing specificity in proline-mediated interactions is by extending the interaction surface with the peptide to include residues beyond the proline-rich core. Another mechanism is to include a nearby sequence on the ligand that interacts with another binding module in the same complex as the proline recognition module. A third mechanism adds a separate recognition surface onto the proline recognition domain that recognizes a distinct peptide.

Chapter 2

Optimization of Specificity Within a Protein Interaction Network Through System-Wide Negative Selection

Ali Zarrinpar, Sang-Hyun Park, and Wendell A. Lim

ABSTRACT

Modular protein interaction domains function as key links in cell signaling networks. Because there are many domains of the same family within a single proteome, it is generally thought that isolated domains lack sufficient information to independently specify unique, biologically relevant interactions. Instead, specificity may be encoded in the context in which the domain is presented (flanking domains, co-localization, etc.). Here we show that the isolated proline-rich motif from the yeast protein Pbs2 recognizes its biologically relevant partner, the Src Homology 3 (SH3) domain from Sho1, with near absolute specificity--none of the other 26 yeast SH3 domains cross-reacts with the Pbs2 ligand, *in vivo* or *in vitro*. This high level of specificity, however, is not observed among a set of non-yeast SH3 domains, suggesting that the Pbs2 ligand motif has been optimized through negative selection against cross-reactivity with competing domains within the yeast proteome. System-wide negative selection provides a subtle but powerful mechanism to optimize interaction specificity among a network of highly overlapping recognition domains.

INTRODUCTION

Protein interaction domains appear repeatedly within the same organism(Pawson and Nash, 2003). The yeast proteome, for example, contains 27 SH3 domains. While duplicating such modular domains can lead to new protein connections, it also creates a problem of specificity: how do domains avoid cross-reactivity? One model postulates that domains have diverged sufficiently and have distinct recognition profiles (Fig. 2.1a). However, extensive peptide library studies have shown that the recognition profiles of most SH3 domains are highly overlapping(Cesareni et al., 2002; Kay et al., 2000; Sparks et al., 1996). Despite a few examples of SH3 domains with unusual recognition profiles(Berry et al., 2002; Ghose et al., 2001; Nishida et al., 2001), the majority bind canonical peptides with core proline-rich motifs flanked by basic residues on either the N- or C-terminus (e.g. R/KxxPxxP or xPxxPxR/K)(Feng et al., 1994; Lim et al., 1994). Thus, it is generally thought that most SH3 domain-peptide pairs do not, by themselves, contain sufficient information to determine unique, biologically relevant interactions(Ladbury and Arold, 2000; Mayer, 2001). Instead, it is postulated that specificity is encoded largely through the context in which the domain-ligand partners are presented, including cooperativity with other interaction domains or subcellular co-localization (Fig. 2.1a).

To examine the specificity of SH3 domains and to better understand how such domains are used to assemble intracellular interaction networks, we have examined a physiologically relevant SH3-ligand pair from yeast and probed whether the wild-type domain can be interchanged with alternative SH3

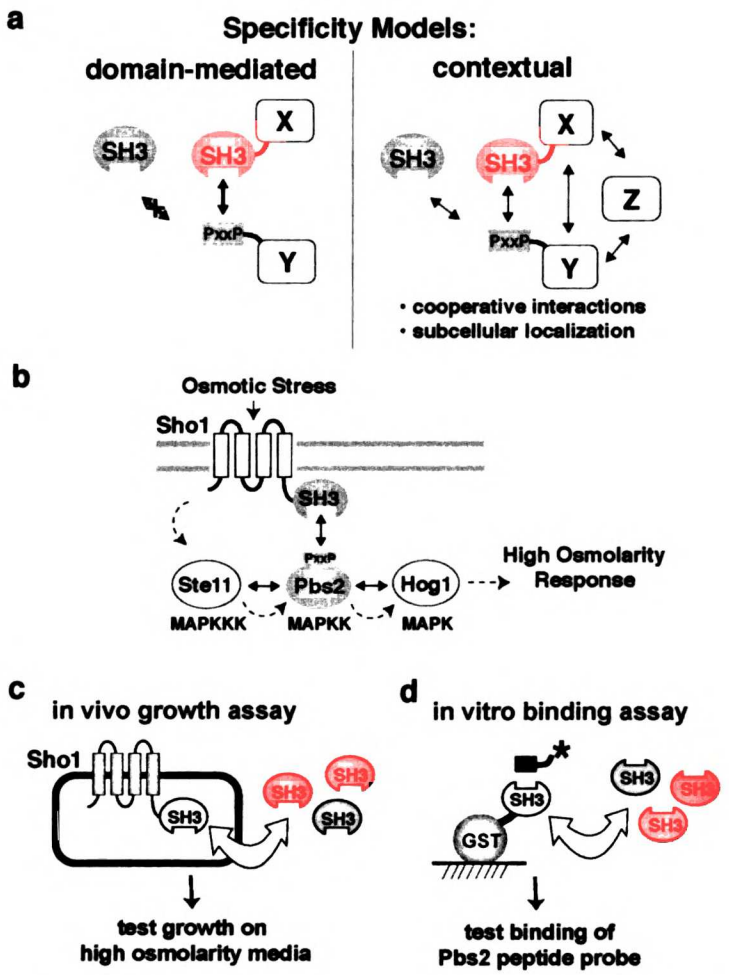


Figure 2.1. Yeast high-osmolarity pathway as a system for studying SH3 network specificity. General models for specificity in domain interaction networks: a, In domain-mediated specificity, individual domain-ligand pairs contain enough information to independently specify a unique interaction. In contextual or distributed specificity, individual domain-ligand pairs lack sufficient information to encode a unique interaction. Other factors (cooperative interactions, subcellular colocalization, etc.) are required for specificity. b, In the yeast high-osmolarity MAPK pathway, the SH3 domain of Sho1 interacts with a PxxP motif in Pbs2. Pbs2 also interacts with other proteins in the pathway (solid arrows— physical interactions; dashed arrows — activating interactions). A second branch of the osmoresponse pathway,{Posas, 1997 #29} involving the two-component sensor protein Sln1, has been omitted for simplicity. This branch does not require Sho1 or Ste11 and all studies here were performed with strains deficient in this branch (ssk2Δ and ssk22Δ). c, Growth assay to test SH3 domain functional interchangeability in vivo. Sho1 chimeras bearing swapped SH3 domains are tested for rescue of osmoresistance in a sho1Δ strain. d, Array binding assay to test SH3 domain interchangeability in vitro. A set of GST-SH3 fusions arrayed on nitrocellulose can be probed for binding to a His-tagged proline-rich peptide.

domains. The fraction of alternative domains that cannot functionally replace the original can be taken as a measure of the interaction information content (inverse degree of entropy in the system(Schneider, 2000)). If individual domains carry little specificity information, then many SH3 domains should be able to functionally replace a native SH3 domain.

The interaction of the SH3 domain from the *Saccharomyces cerevisiae* osmosensor protein Sho1 with a proline-rich motif from the kinase Pbs2 (Fig. 2.1b) is an ideal model for studying specificity. First, it is one of the few SH3 domain interactions that, through genetic studies, has unequivocally been shown to be biologically relevant: it is essential for signaling in one branch of the yeast high osmolarity stress response pathway(Posas and Saito, 1997). Second, peptide library screens show that the Sho1 SH3 domain falls into the canonical SH3 recognition class (Fig. 2.2)(Cesareni et al., 2002; Tong et al., 2002). Finally, there are excellent methods to assay domain function and specificity both *in vivo* and *in vitro*.

Results and Discussion

To probe SH3 domain information content *in vivo*, we generated Sho1 constructs in which the wild-type domain was replaced by alternative SH3 domains (Fig. 2.3, 2.4) and assayed their ability to rescue growth of a Sho1 deletion strain on high osmolarity media (Fig. 2.1c). To probe SH3 domain information content *in vitro*, we generated spatially defined arrays of SH3 domains fused to glutathione–S-transferase (GST) and assayed these for binding to a His-tagged Pbs2 ligand (Fig. 2.1d).

in vitro Ligand Profile	Yeast SH3 Domains
$^{R}/_{K}$xxPxxP (class I)	Abp1, Boi1, Myo3, Myo5, Nbp2, Pex13, Rvs167, Sho1, Sla1-3, YFR024c, YGR136, YHL002w, YHR114w-1, YHR114w-2, YJL020c, YPR154, Ysc84
xPxxPx$^{R}/_{K}$ (class II)	Bem1-1, Boi1, Boi2, Myo3, Myo5, Pex13, Rvs167, Sho1, YFR024c, YGR136, YHL002w, YJL020c, YPR154, Ysc84
unusual motifs	Bem1-1, Boi1, Boi2, Fus1, Myo3, Myo5, YHL002w
no identified peptides	Bem1-2, Cdc25, Hof1, Sla1-1, Sla1-2,YAR014c, YDL117w

Figure 2.2 Binding profiles of all yeast SH3 domains from phage display experiments. SH3 domains that fall into distinct classes based on binding profile are shown. Class I and II peptides bind SH3 domains in two opposite orientations. The peptide from Pbs2 fits the class I profile. Data are from Cesareni, G. et al. Can we infer peptide recognition specificity mediated by SH3 domains?, *FEBS Lett* **513**, 38-44 (2002), and Tong, A. H. et al. A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* **295**, 321-4 (2002).
CONCLUSION: The general ligand profile of the Sho1 SH3 domain overlaps that of many other yeast SH3 domains.

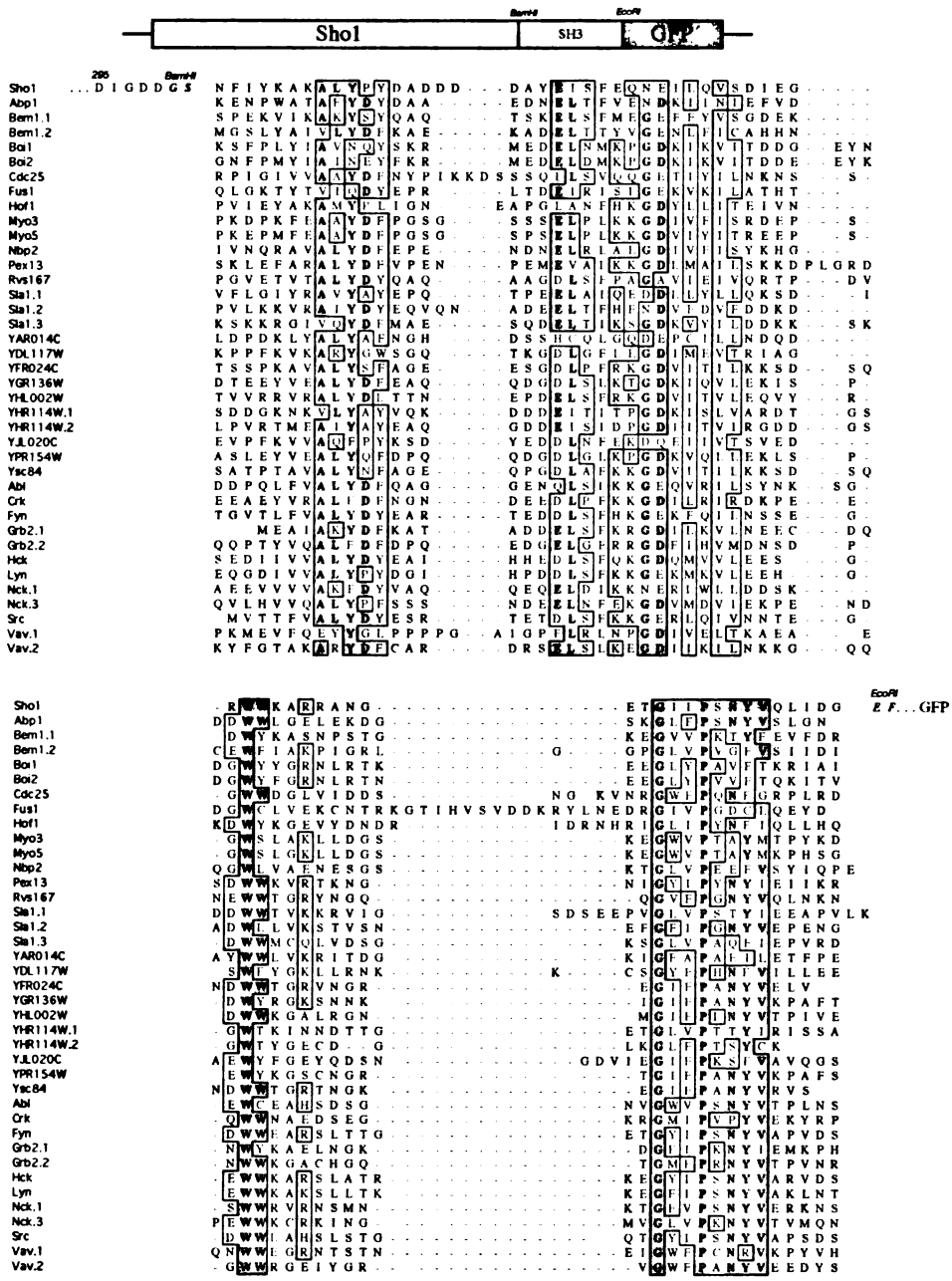


Figure 2.3 Construction and alignment of the Sho1 SH3 domain replacements. Alternative SH3 domains were cloned into Sho1 using engineered BamHI and EcoRI restirction sites flanking the SH3 boundaries. These sites each introduced two amino acid insertions, as indicated below. All constructs also contained a C-terminal fusion to green fluorescent protein (GFP), in order to follow protein localization. These insertions and the GFP fusion had no effect on the function Sho1, and all mutants are compared to a "wild-type" construct containing the native Sho1 SH3 domain in this context. Boundaries for domains were determined based on alingnment in ClustalW. Sdc25 contains an incomplete SH3-like region and was not considered to be a true SH3 domain here.

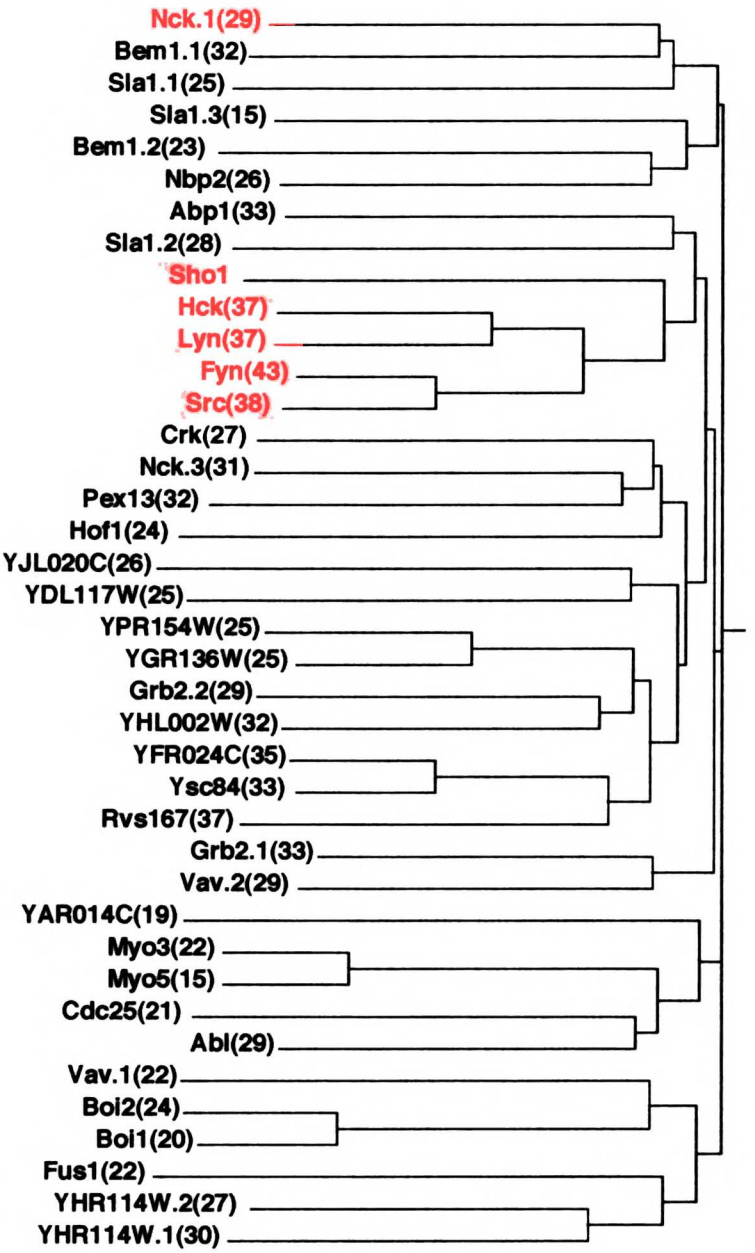


Figure 2.4 Sequence dendogram of all the SH3 domains used in this study. SH3 domains that reconstitute osmo-response and bind the proline-rich region of Pbs2 are highlighted in yellow. Percent identity indicated in parentheses. Dendogram and percent identities were generated using ClustalW.

CONCLUSION: There is no simple relationship between SH3 domain sequence and ability to replace Sho1 SH3 domain. Although a number of functional metazoan SH3 domains are clustered near Sho1 (Hck, Lyn, Fyn, Src), several are far more distant (Nck.1, Grb2.1). More than 30% of the yeast SH3 domains are closer in sequence to Sho1 than Nck.1, yet none of these can replace the native Sho1 domain.

Of twelve metazoan SH3 domains tested, six were able to reconstitute osmoresistance when swapped into Sho1 (Fig. 2.5a). The same six domains showed binding to the Pbs2 ligand on *in vitro* SH3 domain arrays (Fig. 2.5b). Binding results were corroborated by quantitative fluorescence-based solution binding assays (Fig. 2.5c). There was a good correlation between binding affinity and ability to rescue function (Fig. 2.6). Binding to the Pbs2 peptide with a K_d of ≤ 40 μM (wt K_d =1.3 μM) appeared to be sufficient to restore detectable pathway function *in vivo*. These results are consistent with low information content within the individual SH3 domains: the Pbs2 ligand motif is promiscuously recognized by this set of domains and the SH3 domains show a relatively high degree of functional interchangeability.

In contrast, a much higher level of specificity was observed when similar assays are performed with the set of 27 yeast SH3 domains. None of the 26 alternative SH3 domains could reconstitute osmoresistance (Fig. 2.7a). This lack of function was not due to changes in protein expression or localization (see Fig. 2.8). Moreover, in the *in vitro* array binding assays, none of the twenty-three alternative domains tested (3 of the SH3 domains were insoluble) showed detectable binding to the Pbs2 peptide (Fig. 2.7b). This lack of binding was confirmed by quantitative solution binding assays (Fig. 2.7c). Thus, within the context of the *S. cerevisiae* SH3 domain network, the Sho1 domain appears to have high information content.

These results suggest that the isolated SH3 domain-ligand pair contains sufficient information to encode interaction specificity among the yeast set of SH3 domains. This model is supported by several other observations. A non-functional Sho1-construct bearing a swapped SH3 domain can be complemented

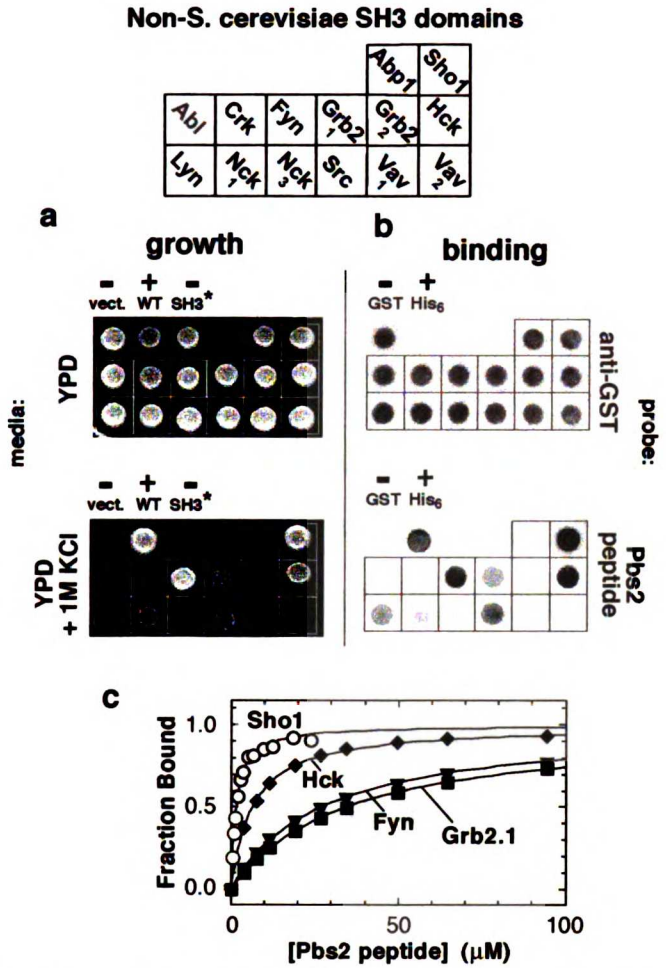
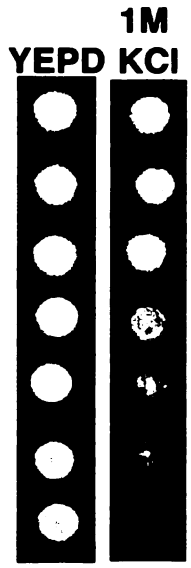


Figure 2.5. The Sho1 SH3 domain can be replaced by SH3 domains from other organisms. **a**, Osmolarity growth assays probe function of Sho1 bearing swapped SH3 domains in vivo (strain: sho1Δ, ssk2Δ, and ssk22Δ). Strains were plated on YPD media with (bottom) or without (top) 1 M KCl. Growth requires Sho1 function. Positive control is a strain transformed with wild-type Sho1 (WT); negative controls are transformed with either vector alone or Sho1 bearing non-binding mutation (SH3* - W338F). Key indicating arrangement of different SH3 constructs is shown at top; where given, small numbers indicate domain identity from N- to C- terminus in multidomain proteins. **b**, SH3 arrays assess binding of alternative SH3 domains in vitro. Arrays of GST fusions of SH3 domains on nitrocellulose were probed with the Pbs2 proline-rich motif fused to a His-tagged protein (subsequently detected by probing with anti-His antibody), or anti-GST antibody (spotting control). Positive control is a His-tagged protein directly spotted on the filter; negative control was GST alone. Arrangement of SH3 fusions is same as given above. **c**, Solution fluorescence binding assays of Pbs2 peptide binding to alternative SH3 domains. Quantitative binding affinities correlate well with intensities observed in array assays. The dissociation constants of the Pbs2 peptide for the SH3 domains are: Sho1 (wt) 1.3μM; Hck 7μM; Fyn 28μM; Grb2.1 35μM.



Domain	Ligand	Kd
Sho1 SH3 (wt)	VNKPLPPLPV (WT)	1.3±0.1 μM
	*R*SKPLPPLPV	1.0±0.2 μM
	VN*R*PLPPLPV	2.2±0.2 μM
	*R*SKPLP*LTPN*	8.3±1.3 μM
	*RSRA*LPPLPV	13±2.3 μM
	*KSRV*LPPLPV	30±8 μM
	VNKPL*A*PL*A*V	>50 μM

Figure 2.6 Osmoresistance correlates with SH3-peptide affinity Testing osmoresistance of Pbs2 variants bearing mutations in the Sho1 SH3 interaction motif (strain: Δssk2/22, Δpbs2). Dissociation constants for these peptides were meaure in vitro using the assay described in Methods.
CONCLUSIONS: Osmoresistance correlates with interaction affinity. Kd < 30 μM appears to be the approximate threshold for detectable osmoresistance.

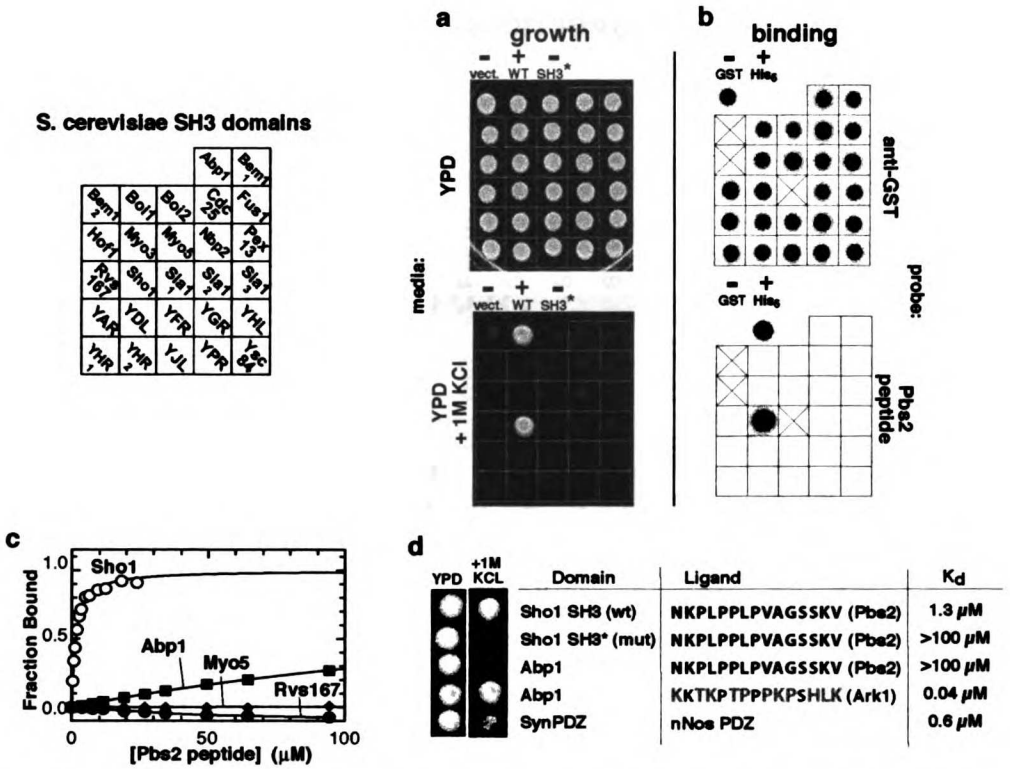
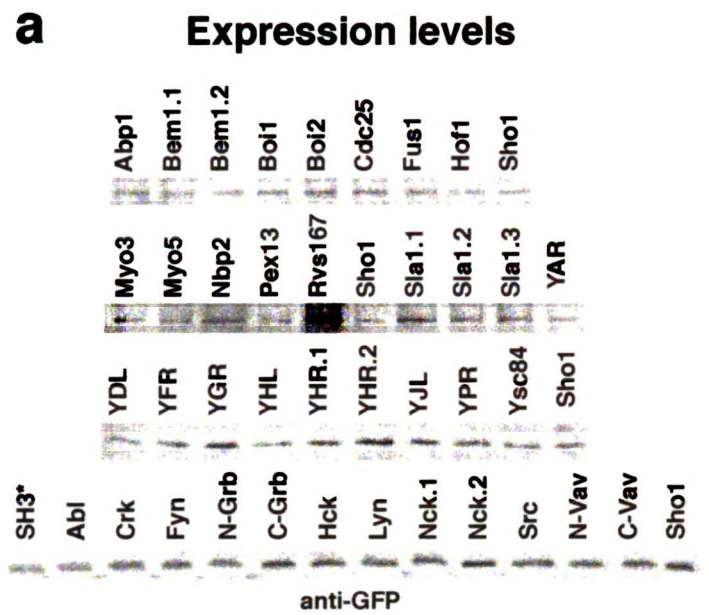


Figure 2.7. The Sho1 SH3 domain cannot be functionally replaced, in vivo or in vitro, by any other S. cerevisiae SH3 domain. **a**, Osmolarity growth assays of Sho1 bearing swapped SH3 domains in vivo (strain: sho1Δ, ssk2Δ, and ssk22Δ). Strains were plated on YPD media with (bottom) or without (top) 1 M KCl. Positive control is a strain transformed with wild-type Sho1 (WT); negative controls are transformed with either vector alone or Sho1 bearing non-binding mutation (SH3* - W338F). Key indicating arrangement of different SH3 constructs is shown on the left; where given, small numbers indicate domain identity from N- to C- terminus in multidomain proteins. **b**, SH3 arrays assess binding of alternative SH3 domains in vitro. Arrays of GST fusions of SH3 domains on nitrocellulose were probed with the Pbs2 proline-rich motif fused to a His-tagged protein and then with an anti-His antibody, or anti-GST antibody (spotting control). Positive control is a His-tagged protein directly spotted on the filter; negative control was GST alone. Arrangement of SH3 fusions is same as given above. Array positions with "X" indicate those SH3 domain fusions that are insoluble and therefore were not included in the arrays. **c**, Solution fluorescence binding assays of Pbs2 peptide binding to alternative SH3 domains. The dissociation constants of the Pbs2 peptide for the SH3 domains are: Sho1 (wt) 1.3μM; Abp1, Rvs167, and Myo5 >70μM. **d**, Compensatory changes in Pbs2 can rescue osmoresistance of non-functional Sho1 SH3 chimeras. Mutation of the Pbs2 proline-rich region to a sequence that binds Abp1 rescues the function of the Abp1 SH3 domain-swapped chimera. A heterologous interaction pair, a PDZ/PDZ heterodimer from the proteins syntrophin (syn PDZ) and neuronal nitric oxide synthase (nNOS PDZ), can also reconstitute osmo-resistance.



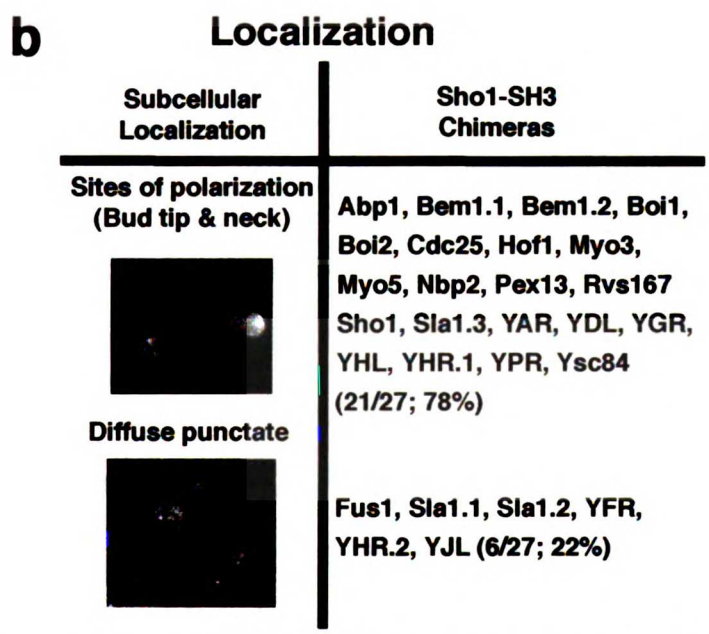


Figure 2.8 Chimeric (swapped SH3) Sho1 constructs do not change in expression levels and subcellular localization a, Anti-GFP western blots of cell lysates from yeast strains expressing chimeric Sho1 constructs show that the expression levels of these chimeras do not vary significantly. **b**, Fluorescence micrographs of the chimeric Sho1 constructs in yeast show that the vast majority of them stillshow wild-type localization (cell membrane; sites of polarization). Only a few chimeras show altered localization Sample micrographs of the two classes of subcellular localization are shown.
CONCLUSION: Changes in the expression levels or of the subcellular localization of the chimeric proteins do not account for observation that Sho1 SH3 domain cannot be replaced by other yeast SH3 domains.

by compensatory changes in the Pbs2 peptide motif (Figs. 2.7d, 2.9): the yeast Abp1 SH3 domain regains function when combined with Pbs2 bearing an Abp1 binding peptide(Fazi et al., 2002). Moreover, the native interaction pair can be functionally replaced by a completely heterologous PDZ domain/ligand pair(Park et al., 2003) (Fig. 2.7d). Thus diverse interactions appear capable of functionally replacing the wild-type SH3 domain/ligand pair, as long as the interaction is of sufficient affinity. These data indicate that other yeast SH3 domains cannot be functionally swapped into Sho1 because they simply do not cross-react with Pbs2.

Why does the Sho1 SH3-Pbs2 ligand interaction show such a high level of specificity within the set of yeast SH3 domains, but not within a set of non-yeast SH3 domains? There is no simple explanation based on sequence clustering of the two SH3 domain sets (Fig. 2.3, 2.4). Instead, an attractive model is that the specificity observed among yeast SH3 domains results not only from positive selection of the Pbs2 ligand for interaction with Sho1, but also from negative selection against binding to competing SH3 domains from *the same organism* (Fig. 2.10). If the recognition profile of the Sho1 SH3 domain overlaps with those of many other SH3 domains, both from yeast and other species, then most random ligands that bind Sho1 will show high levels of cross-reactivity (Fig. 2.10a). However, if the Pbs2 motif were specifically selected to minimize cross-reaction with other yeast SH3 domains (Fig. 2.10b), then high specificity would be observed only within the yeast domain set and not within the non-yeast domain set, i.e. only domains within the same proteome would be targets for negative selection. In summary, this model suggests that as interaction domains proliferate over the course of evolution, specificity can be enhanced by the

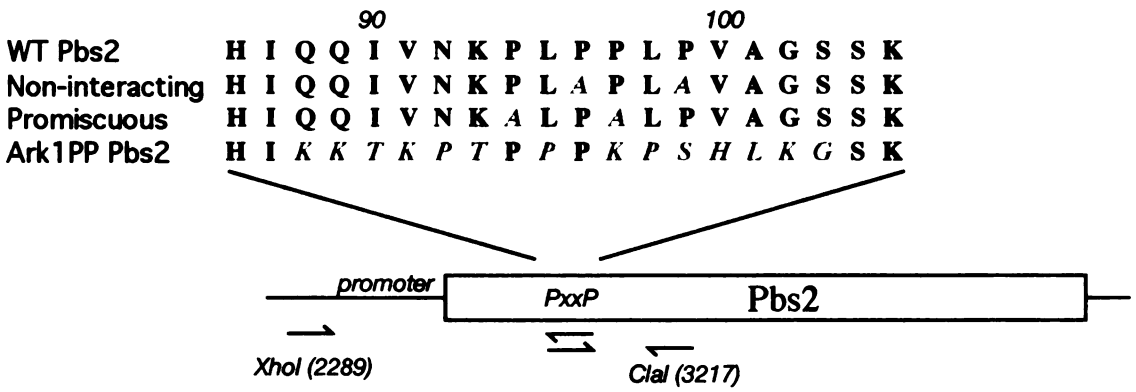


Figure 2.9 Construction of Pbs2 polyproline motif mutations. The Pbs2 proline-rich region mutants were made via two-step PCR. The PCR product of the second round was cut by XhoI and ClaI and ligated into Pbs2 (plasmid pRS304 Pbs2-GFP). The resultant mutants were integrated as a single copy into the yeast genome (at the TRP1 locus). Mutations indicated in Fig. S4 were also made in this manner.

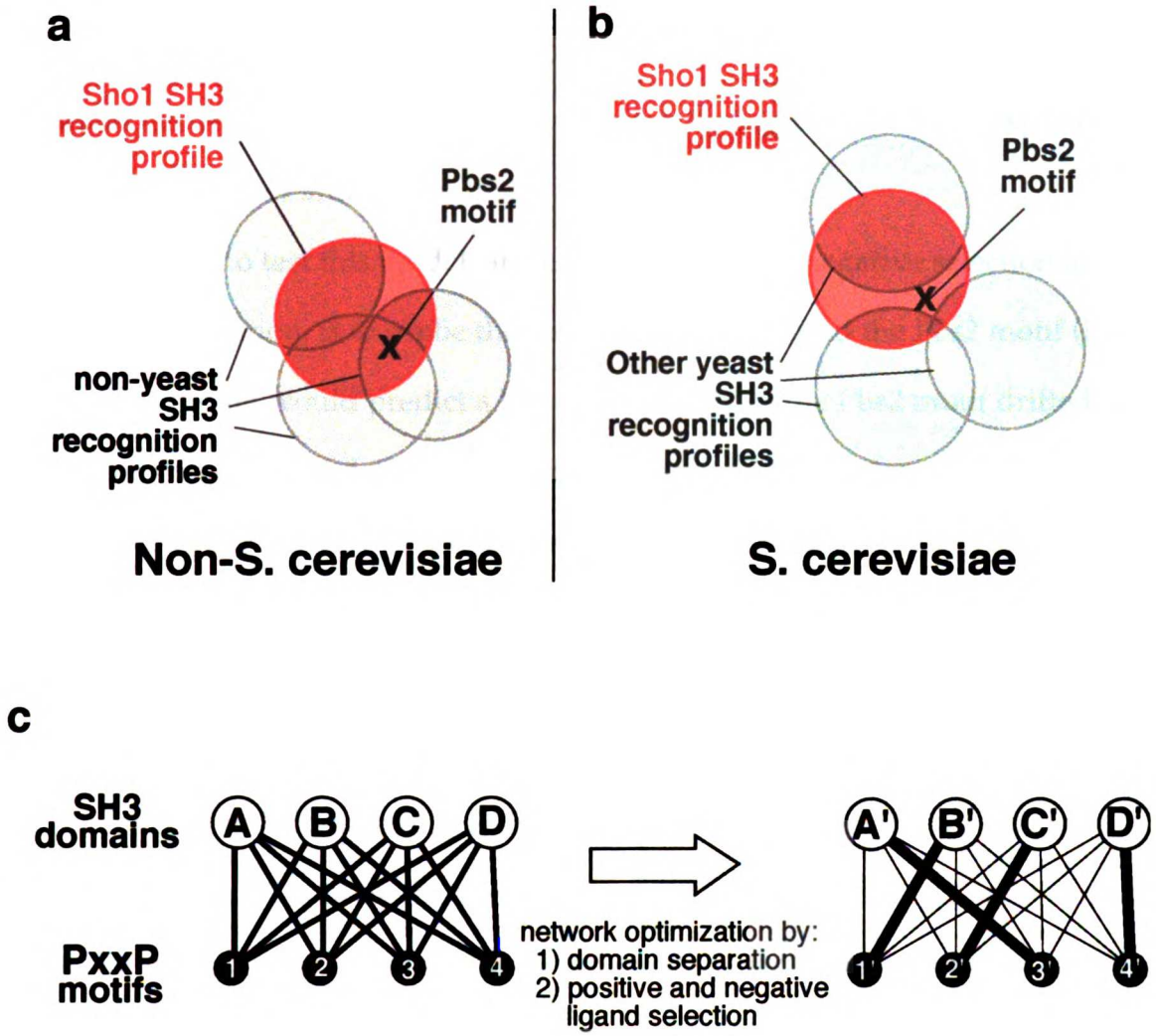


Figure 2.10. Model: role of proteome-wide negative selection in interaction network specificity. a, The Pbs2 peptide is a canonical PxxP motif that falls within the recognition space of a number of SH3 domains, and therefore shows cross-reactivity with many non-yeast SH3 domains. **b**, However, negative selection against cross-reactivity with other natural competitor domains (i.e. other yeast domains) could drive the Pbs2 motif into a sequence space niche only compatible with the Sho1 SH3 domain. **c**, Two forces can optimize network-wide specificity after domain proliferation: domain diversification and niche exploitation through positive and negative selection.

combination of two distinct mechanisms: increased divergence in the domain recognition profiles and pruning of cross-reactivity by negative selection (Fig. 2.10c). Binding interactions may diverge and be rendered orthogonal through evolution much in the same way species diverge through evolution to exploit ecological niches(Orr and Smith, 1998).

One way to test this model, and the importance of negative selection in network optimization, is to probe the sequence space around the Pbs2 motif (Fig. 2.11a). This model would predict a loss of specificity as the Pbs2 motif drifted away from this optimized point in sequence space. To this end, we made a library of 19 of the possible 47 single base pair missense mutations of the Pbs2 motif (leaving the core prolines unchanged) (Figs. 2.11b and 2.12).

Specificity and affinity of this point mutant library was assayed using the yeast SH3 arrays (Fig. 2.11b). Intensity of the Sho1 spot was used as an index of affinity for the Sho1 domain (Fig. 2.13). As an index of specificity, we divided intensity of the Sho1 spot by the average intensity of the remaining 23 non-Sho1 spots. Some mutations increase affinity, others decrease affinity, but they all yield an increase in cross-reactivity with other yeast SH3 domains (Fig. 2.11c). Based on this mutant analysis, several residues in the ligand appear to play a more significant role than others in determining specificity. However, it is difficult to precisely rationalize these effects based on structural comparisons (Fig. 2.14).

This analysis indicates that the wild-type Pbs2 motif is not optimized for affinity for the Sho1 SH3 domain, but it does appear to be optimized for specificity. In fact, by combining two promiscuous point mutations (P94A, P97A), we were able to construct a Pbs2 motif double mutant that bound to the

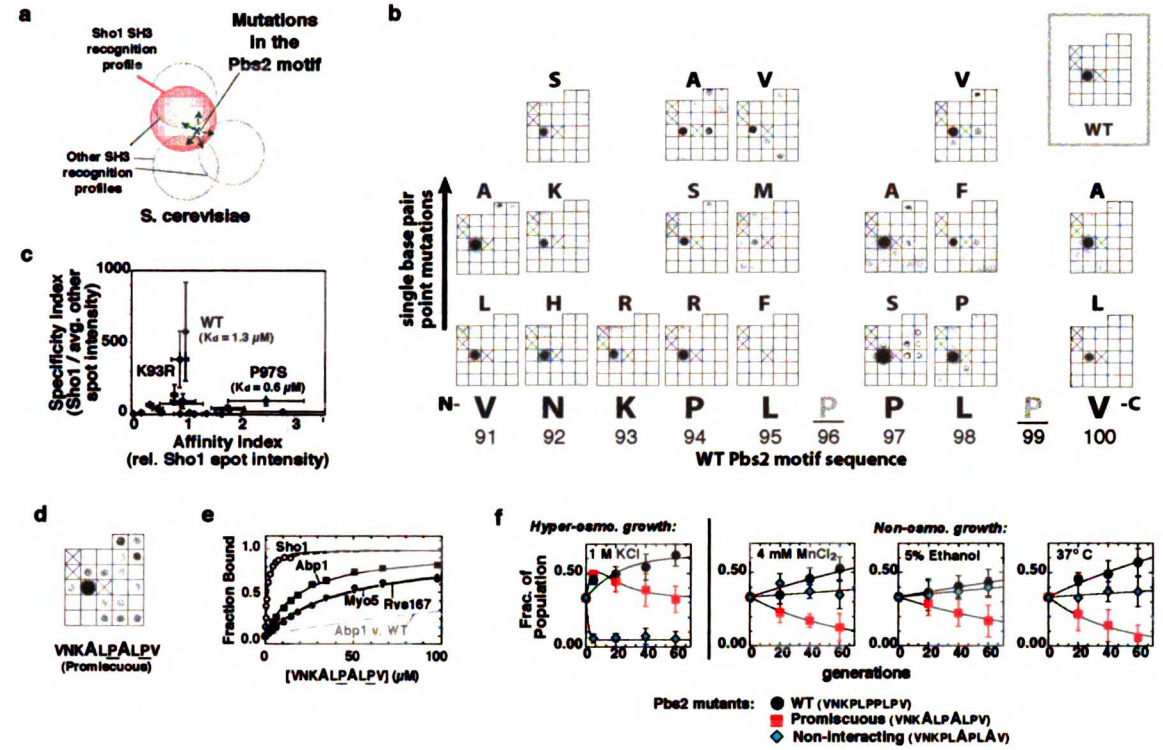


Figure 2.11. The Pbs2 proline-rich motif is optimized to minimize cross-reactivity with other yeast SH3 domains. a, To test specificity optimization of the Pbs2 peptide, we probed the effects of mutational drift. b, We examined the yeast SH3 array binding profiles of 19 of the 47 possible missense point mutants of the Pbs2 peptide. Wild-type sequence is shown at the bottom (core PxxP prolines are underlined). Each array above was probed with ligand bearing the indicated point mutation at this position. Array probed with wild-type ligand (WT) is shown for reference. Exposures were calibrated using His-tagged standards on each array (removed for clarity). All mutations show increased cross-reactivity with other yeast SH3 domains. c, Quantitation of arrays in (b) shows that the wild-type Pbs2 ligand has the highest specificity, although it does not have the highest affinity among the mutant set (affinity ~ Sho1 spot intensity; specificity ~ ratio of the Sho1 spot intensity to the average intensity of all other competing spots). Data for wild-type and other key mutant ligands are labeled (Kd measured by fluorescence given in parentheses). d, Combining the P94A and P97A mutations yields a more promiscuous peptide as assayed by SH3 domain array binding. (e) The promiscuity of double mutant was confirmed by in vitro fluorescence binding measurements; mutation improves affinity to Abp1, Myo5, and Rvs167 domains, but maintains similar affinity to the Sho1 domain. Binding curves of the wild-type peptide to the Sho1 SH3 (dashed line) and to the Abp1 SH3 (grey line) are shown for comparison. f, Increasing domain-ligand promiscuity leads to fitness defects under several non-hyperosmotic conditions. Fitness in competitive cultures was measured by assaying the fraction of a growing population composed of cells bearing a specific Pbs2 variant (starting with equal fractions). Cells with promiscuous Pbs2 (P94A/P97A) are outcompeted by cells with wild-type Pbs2 under conditions indicated. Cells with non-interacting Pbs2 (P96A/P99A) were included in the growths and did not display fitness defects under these conditions (except hyperosmotic), indicating that defects were not caused by increased osmosensitivity.

WT sequence

V	**N**	**K**	**P**	**L**	**P**	**P**	**L**	**P**	**V**

Amino acids accessible by single base pair change (Dayhoff score):

F(1)	Y(4)	**R(48)**	**S(27)**	**M(21)**	-	**S(27)**	**F(17)**	-	**L(30)**
L(30)	I(4)	M(9)	L(4)	S(3)		L(4)	**P(4)**		**A(37)**
I(66)	**K(32)**	N(32)	T(7)	**F(17)**		T(7)	I(25)		I(66)
A(37)	**H(23)**	Q(15)	Q(9)	**V(30)**		H(5)	H(4)		G(10)
D(2)	T(17)	T(20)	**A(35)**	W(1)		**A(35)**	**V(30)**		E(4)
G(10)	D(53)	E(10)	**R(7)**			R(7)	R(2)		
	S(43)								

Of the 47 possible missense point mutations, 19 (bold) were selected to give a representative sampling. Mutations were selected with Dayhoff scores ≥17 and ≤50 so that they would be sufficiently different to be meaningful but not too different so as not to bind. Exceptions (underlined) were made in the following cases:

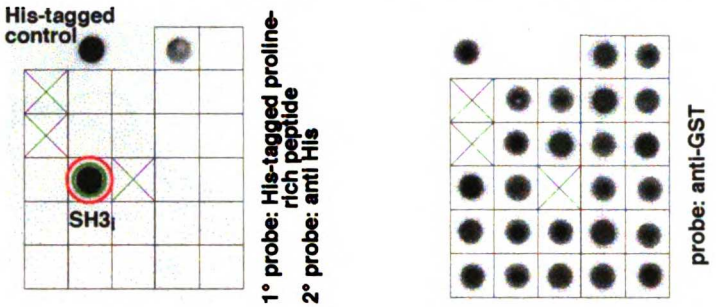
- no two substitutions have a score ≥50 for each other (s.a. I/V and S/T); of these, the higher score substitution was selected;
- K93 substituted only with R because of requirement for basic residue in that position;
- P94 also substituted with R because phage display data from Cesareni *et al* show that the Sho1 SH3 domain can bind peptides with an R in that position;
- L98 also substituted with P to explore effect of expanding the polyproline repeat.

Figure 2.12. Sampling single base pair missense mutations in the Pbs2 proline-rich peptide

Supporting Reference: Dayhoff, M.O. (1978) *Atlas of Protein Sequence and Structure*, **5**, suppl. 3

a

I. probe replica SH3 arrays with experimental and control probes



II. measure spot intensity (Alpha Innotech scanner) and correct

raw intensity= intensity within green circle

background = average pixel intensity in surrounding (red) ring x area of green circle

background corrected intensity = raw Intensity - background

$$\text{CORRECTED SPOT INTENSITY} = \frac{\text{background corrected intensity}}{\text{(background corrected intensity of His-tagged control spot)} \times \text{(intensity of spot on replica array probed directly with anti-GST)}}$$

exposure variability correction term *spotting variability correction term*

b

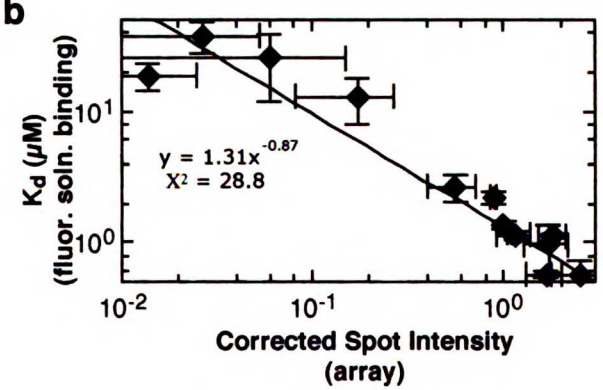


Figure 2.13. Analysis of peptide binding to SH3 domain arrays. a, Scheme for quantitation of arrays (see methods for array generation) **b,** comparison of array measurements with *in vitro* binding measurements (fluorescence) demonstrate a good correlation to the equation:

$$\log K_d = k \log I + c$$

where K_d is the dissociation constant, I is the spot intensity and k and c are constants.

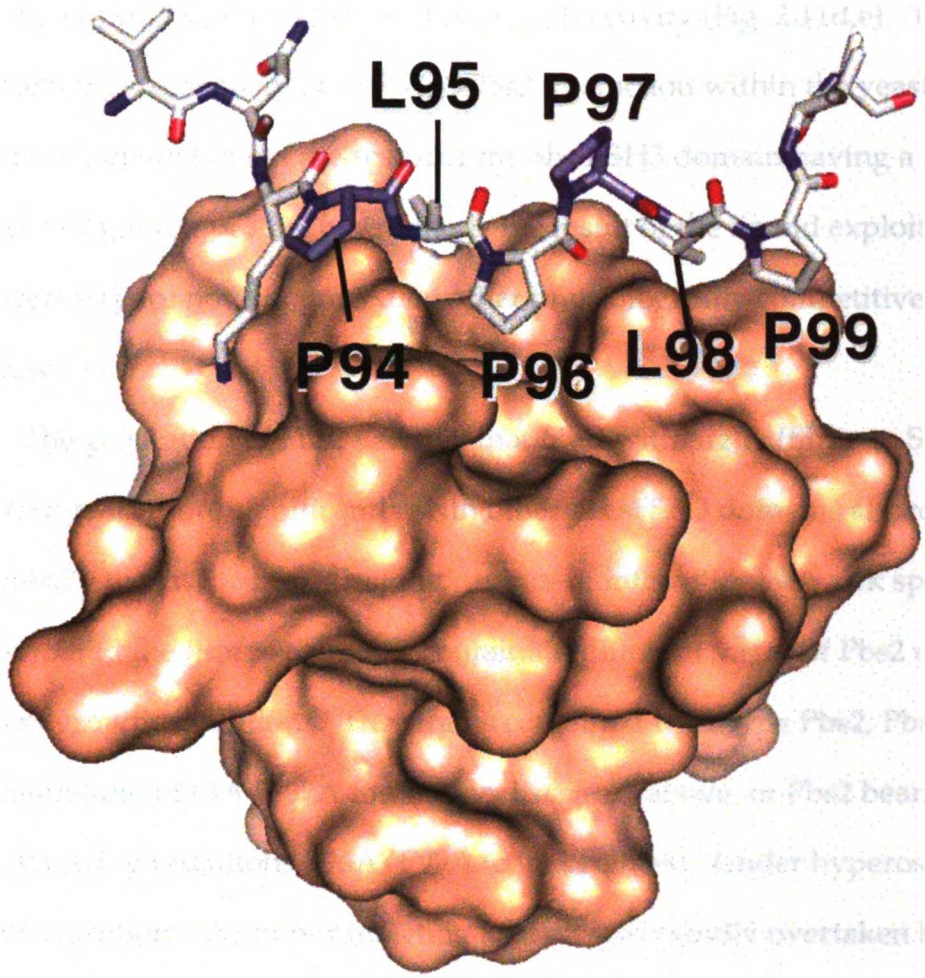


Figure 2.14. Location of mutated Pbs2 peptide residues in an SH3-peptide complex. Peptide residues where mutation most significantly alters SH3 cross-reactivity are P94, L95, P97, and L98 (P96 and P99 are the core prolines of the conserved PxxP motif). A model for how these residues might contact the Sho1 SH3 domains was generated starting with the structure of the Fyn SH3 domain (seq. ident. with Sho1 SH3: 43%) bound to a peptide from PI3 kinase (PDB accession 1AZG). The peptide sequence was mutated to that of Pbs2 using the PyMOL graphics system.

CONCLUSIONS: L95 and L98 are expected to contact the SH3 domain surface and could exploit unique surface properties to alter cross-reactivity. However, P94 and P97 lie at apical positions on the ligand that are not expected to contact the SH3 surface. It is therefore unclear how these residues alter cross-reactivity. It is possible that conformational properties at these apical positions subtly alter the presentation of residues to the SH3 surface.

Sho1 SH3 domain with a slightly higher affinity than the wild-type Pbs2 motif, but with a dramatically higher level of cross-reactivity (Fig. 2.11d,e). Thus the extremely high specificity of the Sho1-Pbs2 interaction within the yeast SH3 interaction network is not the result of the Sho1 SH3 domain having a highly distinct recognition profile, but rather is the result of the ligand exploiting niches in sequence space not recognized by other physiologically competitive SH3 domains.

The generation of Pbs2 variants that bind with high affinity to Sho1 but also cross-react more significantly with other yeast SH3 domains affords us the opportunity to test the biological importance of interaction network specificity. We compared the fitness of strains containing different forms of Pbs2 under different growth conditions. The strains contained wild-type Pbs2, Pbs2 bearing the promiscuous P94A/P97A mutations discussed above, or Pbs2 bearing the non-interacting mutations P96A/P99A (core prolines). Under hyperosmotic growth conditions, the non-interacting mutant was rapidly overtaken by the other strains, as expected since this mutant is osmosensitive. The promiscuous mutant strain, in contrast, shows growth under hyperosmotic conditions that is competitive with the wild-type strain. However, under some conditions, such as growth in minimal media at 37°C, the promiscuous mutant strain is overtaken by both the wild-type strain and the non-interacting mutant strain (Fig. 2.11f). Thus, the promiscuous mutant strain appears to have a fitness defect under these conditions that is not due to a defect in the osmolarity response pathway, suggesting that promiscuous interactions may lead to small but possibly evolutionarily important disadvantages.

The generality of the use of negative selection for specificity enhancement is difficult to probe because so few biologically verified SH3-ligand pairs in yeast have been clearly identified. Nonetheless, we examined two of the better characterized yeast SH3 domains, those from Abp1 and Pex13 (Fig. 2.15). A putative ligand for the Abp1 SH3 domain, a peptide from Ark1(Fazi et al., 2002), was observed to bind the Abp1 SH3 domain with minimal cross-reactivity against other yeast SH3 domains. On the other hand, a proline-rich peptide from Pex14 is found to cross-react with 7 other yeast SH3 domains in addition to Pex13, its native partner(Barnett et al., 2000) (Fig. 2.15). This promiscuity, however, is consistent with previous findings that a functional interaction of Pex13 and Pex14 is dependent on the interaction of both of these proteins with a third protein Pex5(Bottger et al., 2000), a case of multipartner cooperativity in recognition. Moreover, cellular localization studies show that Pex13 is the only SH3 domain-containing protein in peroxisomes. Pex14 also localizes to the peroxisome independent of the Pex13 SH3 domain(Girzalsky et al., 1999). In contrast, Sho1 and Pbs2 both overlap in subcellular localization with up to sixteen other SH3 domain-containing proteins (personal communication, E.K. O'Shea and SGD). Thus, because of subcellular colocalization and cooperative interactions, the Pex13-Pex14 interaction pair may not have had the same selective pressure to achieve the level of discrimination observed for Sho1-Pbs2. It is also possible that in some cases SH3 promiscuity may be required function(Sudol, 1998). These results show how negative selection is only one of several possible mechanisms used to enhance interaction specificity.

In conclusion, negative domain-ligand selection can play a powerful role in optimizing protein interaction network specificity. Negative selection as a

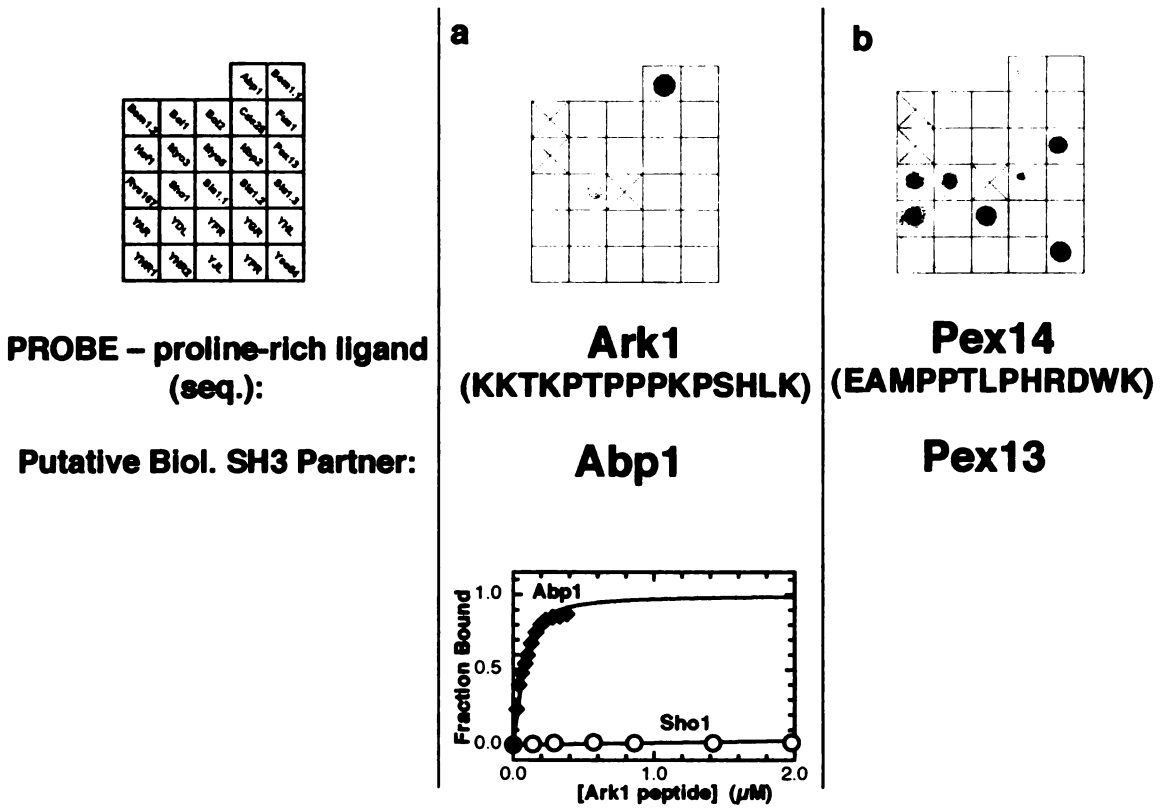


Figure 2.15. Analysis of cross-reaction for other putative physiologically relevant yeast SH3 domain/ligand pairs. **a**, The Ark1 peptide binds with high specificity to its putative biological partner, the Abp1 SH3 domain (K_d = 40 nM). The closest cross-reacting SH3 domain is Sho1, which binds ~1000 weaker (K_d ≥ 40 μM). **b**, The Pex14 peptide binds a number of yeast SH3 domains with similar affinities, including its biological partner Pex13. However, as discussed in the text, the Pex14-Pex13 interaction probably uses other mechanisms to enhance specificity.

driving force in specificity has previously been recognized in immunology (Palmer, 2003). Negative selection is likely to play a key role in the construction of many biological networks, ranging from protein signaling networks(Yaffe et al., 2001) to DNA-binding/transcriptional networks(Newman and Keating, 2003). The importance of negative selection suggests that in order to map cellular interaction networks, it will be critical not only to search for potential ligands with optimized affinity, but also to characterize cross-reactivity of these ligands with relevant sets of competing receptors. In the case of higher eukaryotes, in which only a fraction of a genome is expressed in each cell type,(Jiang et al., 2001; Kim et al., 2001; Reinke et al., 2000) accurate interaction mapping may require characterization of cell-type specific domain expression profiles in order to delineate physiologically competitive domain sets.

ACKNOWLEDGEMENTS

We thank A. Davidson and the Davidson Lab, I. Herskowitz, E. O'Shea, A. Sali, K. Yamamoto, R. Zuckerman, P. Chien, K. Chak, R. Bhattacharyya, J. Dueber, N. Sallee, B. Yeh, and members of the Lim Lab for assistance and discussion.

Supported by grants from the NIH and the Packard Foundation.

MATERIALS AND METHODS

Constructs and Strains. Yeast strains were grown in YPD medium or synthetic dropout media (for maintaining plasmids) at 30° C, unless otherwise indicated. All yeast strains were derived from the *ssk2Δ, ssk22Δ* mutant of the W303 strain background (*trp1 leu2 ura3 his3 ADE2 can1*). Gene disruptions were confirmed by phenotypic analysis and/or PCR reactions with gene-specific primers. Sho1 chimeras were constructed as shown in Figure 2.3 and expressed from a CEN/ARS plasmid (pRS316) driven by the native Sho1 promoter (strains: *ssk2Δ, ssk22Δ, sho1Δ* or *ssk2Δ, ssk22Δ, pbs2Δ, sho1Δ*). Pbs2 mutants were constructed in pRS304 or pRS306 bearing the Pbs2 promoter and gene (Fig. 2.9) and integrated as a single copy into the genome (strain: *ssk2Δ, ssk22Δ, pbs2Δ*, or *ssk2Δ, ssk22Δ, pbs2Δ, sho1Δ*).

Protein Expression and Purification. Sho1, Pbs2, and all the yeast SH3 domains were cloned by PCR from *Saccharomyces cerevisiae* genomic DNA. The other SH3 domains were cloned from appropriate cDNA libraries. His6-tagged Pbs2 peptides, fused to the N-terminal domain of lambda repressor (res. 1-99), were constructed as described by Maxwell and Davidson(Maxwell and Davidson, 1998). *Escherichia coli* strain TG1 was used for cloning and propagating the plasmids, strain BL21 (DE3) RIL for the expression of recombinant proteins. To express proteins, cultures were grown to an OD_{600} of 0.6-0.8 at 20 °C and induced with 1 mM IPTG for 3-6 h. Cells were harvested by centrifugation, resuspended in PBS (50 mM sodium phosphate, 100 mM NaCl, pH 7.4), and frozen at -80°C. Subsequently, cell suspensions were thawed and lysed using a Branson model 250 sonifier. Lysates were cleared by centrifugation at 20,000*g*. The His6-fusions

were bound to Ni^{2+}-NTA resin (Qiagen) at 4 °C, washed three times with PBS containing 20 mM imidazole, eluted with PBS containing 250 mM imidazole, and dialyzed three times into 100 mM NaCl, 10 mM HEPES (pH 8.0). GST fusions were bound to glutathione agarose at 4 °C, washed three times with PBS, eluted with PBS containing 10mM reduced glutathione, and dialyzed three times into 100 mM NaCl, 10 mM HEPES, 2 mM DTT (pH 8.0). Protein concentration was measured by UV absorbance using calculated extinction coefficients and individual aliquots were stored at -80 °C.

Hyper-osmotic Plate Growth Assay. 10^3 cells were spotted onto YPD plates with or without 1M KCl. Plates were incubated at 30° C for 3 to 5 days.

Peptide Synthesis. Peptides (acetylated and amidated) were synthesized on an ABI 381 synthesizer using Fmoc chemistry, and were purified on a Vydac 25 cm x 2.2 cm, 10 μm C18 reverse phase column (gradient of 0 to 90% acetonitrile in 0.1% TFA). Molecular mass was verified to within 0.5 Da by electrospray mass spectrometry, and final stocks were made in water. Concentration was verified by quantitative amino acid analysis.

SH3 Domain Array Binding Assay. 100μL each of 0.1 μM solutions of purified GST-SH3 fusion proteins in TBST were spotted in array format onto pre-wetted nitrocellulose membrane using a Dot-Blot apparatus. Array membranes were blocked in 3% milk/1% BSA in TBST for 1 h at RT, and then probed with 6 mL of TBST containing a His-tagged fusion protein containing the proline-rich peptide of interest (50μM) for 4-16 h at 4°C. The membrane was washed four times in

TBST, and reprobed with a horseradish peroxidase-conjugated anti-His antibody (1:2000 dilution, Santa Cruz Biotech.) for 1 h at 4°C. Finally, the blot was developed with an ECL system and quantitated on an AlphaInnotech CCD camera and analytical software. To control for variation in antibody levels and development exposure, standards of a His-tagged protein (100 μL of 100 nM and 10 nM solutions) were directly spotted onto the membrane (reference spots are not shown in figures for clarity).

Spot intensities were quantitated as described in Figure 2.13a. The raw value for each spot was taken to be the intensity inside a circle entirely enclosing the spot. The background for each spot (calculated based on the intensity within a outer ring surrounding the spot) was subtracted. Variations in spotting were corrected by dividing by the intensity of the spot on a replica array probed only with anti-GST (i.e. measuring total GST-SH3 protein in spot). Variations in exposure were corrected by dividing by the intensities of the reference His-tag spots described above. The corrected intensities for each spot are given relative to the intensity for the Sho1 SH3 domain spot probed with wild-type peptide. The semi-quantitative nature of this assay was validated by comparing spot intensities from the SH3 domain arrays to *in vitro* measured dissociation constants (Fig. 2.13b). The two sets of measurements show a correlation fit to the equation: $\log K_d = k \log I + c$ (K_d is the dissociation constant, I is the spot intensity, k and c are constants).

Measurement of Binding Affinities. Affinities for binding to SH3 domains were measured by following the increase in domain Trp fluorescence upon titration of ligand into a 1 cm x 1 cm stirred-cell cuvette containing a 1300 μL

solution of SH3 domain at a fixed concentration of 0.01 – 0.5 μM (always less than one-fourth the K_d)(Lim and Richards, 1994). The ligand stock concentration was typically between 0.1-2 mM. Data were fit to the following equation by nonlinear least-squares analysis using the program ProFit 5.6.3 (Quantum Soft) where y is the fluorescence reading, x is the concentration of ligand, K_d is the dissociation constant of the SH3 domain and peptide, F_o is the initial fluorescence value (fraction bound = 0), and F_{max} is the fluorescence value at saturation (fraction bound = 1).

$$y = F_0 + \frac{(F_{max} - F_0)\frac{x}{K_d}}{1 + \frac{x}{K_d}}$$

Competition Growths. Starter cultures of the three strains (wt, promiscuous Pbs2, non-interacting Pbs) were grown independently to OD_{600} = 0.5. Equal amounts of each were combined into one tube. An aliquot was removed from this tube as a standard against which all subsequent samples were measured. Cells were diluted 1:100 into various media and incubated at the appropriate temperature until OD_{600} = 0.5, whereupon they were diluted 1:100 (approximately 1-2 times each day). Samples were removed at various timepoints and lysed by incubation with Zymolyase and boiling. The lysates were subjected to PCR and the PCR product was sequenced according to standard protocol provided by Applied Biosystems and run on an ABI Prism® 3700 DNA Analyzer with DNA Sequencing Analysis Software™ Version 3.6.1 (Applied Biosystems, Foster City, CA).

Mutant frequencies within the culture were estimated according to the sequencing-based protocol developed by Kwok and Duan(Kwok and Duan,

2003). Briefly, we normalize the sequencing traces of the pooled DNA samples and the reference by identifying a base of the same type and of similar height to the reference allele in the reference sample from 20-base window around the polymorphic site. The peak heights of the reference allele and the normalizing base for the reference and the pools are measured and the allele frequency is estimated as = $c(P_{pool}/N_{pool})/(P_{ref}/N_{ref})$, where c is 0.333 in a mixture of three competing mutants, P is the peak height of the base at polymorphic site and N is the peak height of the normalizing base. The frequency of the wild-type is estimated by subtracting the mutant frequencies of the others from 1. Data were fit to multiple exponential equations that account for changes in growth of the mutant of interest and changes in growth of competitors.

Chapter 3

Redundant scaffolds Sho1 and Pbs2 direct activity and specificity in the yeast osmoregulatory MAPK pathway interactions.

Ali Zarrinpar, Roby P. Bhattacharyya, M. Paige Nittler, and Wendell A. Lim

ABSTRACT

The kinase Ste11 acts in three different mitogen activated protein kinase (MAPK) cascades in *Saccharomyces cerevisiae*. Nevertheless, each pathway maintains specificity and limits cross-activation of other pathways. Scaffolding interactions of Pbs2 and Ste5 are thought to dictate this specificity by directing the activity of upstream components to the appropriate targets. Here we show that the membrane protein Sho1 also functions as a scaffold and that it determines the flow of the pathway by interacting with Ste20, Ste11, Pbs2, and Hog1. Osmo-responsive activation of Ste11 requires direct interaction with Sho1. The interaction of Sho1 with Pbs2 is required to direct the activity of Ste11 towards the HOG pathway upon osmoshock. Multiple interactions lead to the formation of complexes including Sho1, Ste20, and Ste11 in one case and Sho1, Pbs2, and Hog1 in another, suggesting a Sho1-mediated assembly of the yeast osmolarity MAPK module.

INTRODUCTION

Considering the variety of environments to which cells are exposed, it is imperative for cells to respond rapidly and specifically to each stimulus. Conceptually, the simplest means of achieving specificity is for all the individual protein interactions to be specific. Some members of signaling cascades do seem to interact only with their physiological partners. This specificity can stem from either the interaction of individual domains or that of a combination of domains. For example, targets of MAPKs have specific consensus binding sequences that are separate from their consensus phosphorylation sequences (Pawson and Nash, 2003). But this mechanism fails to explain how some signals are routed through only one pathway when their associated kinases can function in multiple pathways. Another way to aid the maintenance of specificity is cross-inhibition of one pathway by a competing one, as is the case for the JNK and the ERK pathways (Shen et al., 2003). A third mechanism, exemplified by Ste5 in the yeast pheromone (or mating) response pathway, utilizes scaffolds, proteins which interact with multiple members of a pathway, to channel signal to a specific output (Elion, 2001).

Perhaps the most thoroughly studied cases of the specific transmission of information in eukaryotic signal transduction are the highly conserved MAPK pathways (Whitmarsh and Davis, 1999). These pathways exemplify many of the still unresolved issues in signaling. One of the MAPK cascades in the budding yeast *Saccharomyces cerevisiae* regulates the response to increases in environmental osmolarity by eliciting among other things an increase in intracellular concentrations of glycerol. This high osmolarity/glycerol (HOG) pathway is activated by at least two branches that converge upon the MAPK

kinase (MAPKK) Pbs2 and the MAPK Hog1(O'Rourke and Herskowitz, 1998). The sensor histidine kinase Sln1 initiates one branch and an integral membrane SH3 domain containing protein Sho1 defines the other well-characterized branch (Fig. 3.1). In the Sho1 branch, an increase in osmolarity results in the activation of the MAPKK kinase (MAPKKK) Ste11p, which in turn phosphorylates MAPKK Pbs2, thus activating the MAPK Hog1.

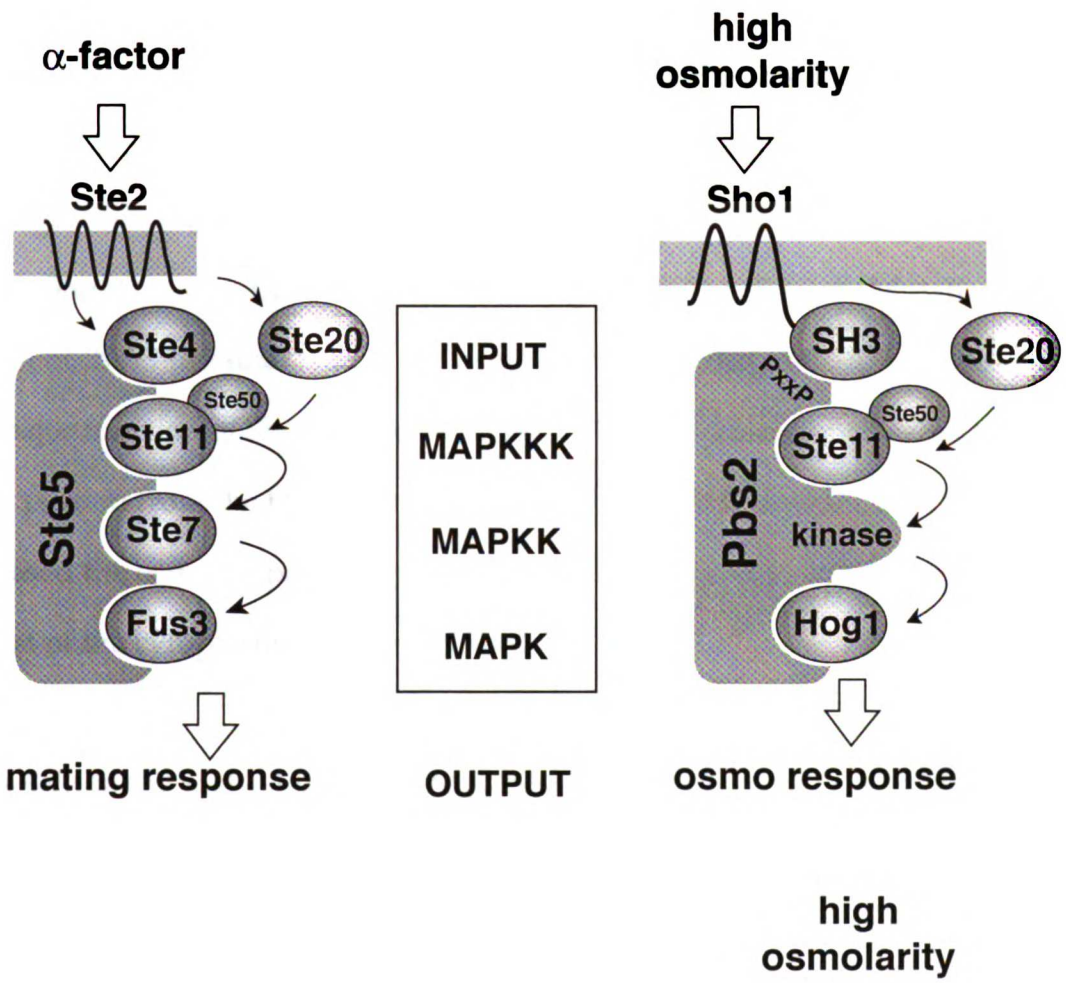
What remains puzzling is how specificity is maintained despite the overlapping set of proteins involved in multiple MAPK cascades. Sho1, Cdc42, Ste20, Ste50, and Ste11 are all involved in the pseudohyphal growth pathway, and the last four proteins are also members of the mating signal transduction cascade. Despite this high degree of overlap, there is very little aberrant signaling or cross-talk, a property attributed at least partly to the scaffolding roles of Ste5 and Pbs2. While there is evidence that Pbs2 and presumably its associated kinases are recruited to Sho1 upon osmoshock, the role of Sho1 and the events upstream of Pbs2 are still unknown. To address these issue we examined more closely the role of Sho1 in the HOG pathway.

RESULTS

Optimal osmo-signal requires a functional SH3 domain in Sho1.

The osmosensor Sho1 consists of four transmembrane segments in the amino-terminus and an approximately 220 amino acid cytoplasmic tail ending in a Src homology 3 (SH3) domain. Other than its SH3 domain, Sho1 has very little similarity to any proteins other than its direct homologs. To explore its function and mechanism in the osmo-response pathway, we tested a series of mutants of

Figure 3.1. Current model of scaffolds in yeast MAPK cascades. Many proteins are shared.



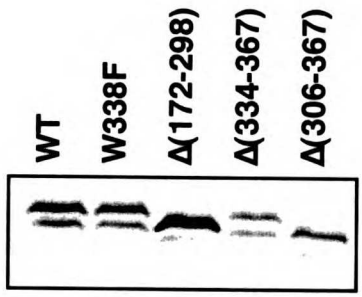
Sho1 for their ability to reconstitute growth on high-osmolarity media. Since the Sho1 pathway merges and is redundant with the Sln1 pathway starting at the MAPKK Pbs2, all our strains bore deletions of Ssk2 and Ssk22 to remove any input from that pathway. Most of the Sho1 mutants we tested were expressed at comparable levels in yeast as assayed by Western blot and the ones that included the transmembrane regions all localized to points in the membrane (Fig. 3.2). Previous work (Raitt et al., 2000; Reiser et al., 2000), as well as mutations we made in the N-terminus and the three loops between the transmembrane regions (M.P.N. unpublished results), suggested that most of the activity of Sho1 stems from the C-terminal tail. We found that mutations in the SH3 domain predicted to abrogate binding to proline-rich peptides also destroyed the ability of the mutant Sho1 to reconstitute high-osmolarity growth (Fig. 3.3a-c). Mutations in the rest of the C-terminal tail left the osmo-response mostly intact, indicating that one major function of Sho1 is to interact with Pbs2. A deletion removing the entire portion from the transmembrane region to the SH3 domain (Δ172-298) decreased the osmo-response significantly, indicating a more subtle function for the rest of the cytoplasmic portion of Sho1.

Cross-talk to the mating pathway requires Sho1.

To test further the function of Sho1 in activating the HOG MAPK pathway, we relied on another previously developed assay. O'Rourke and Herskowitz(O'Rourke and Herskowitz, 1998) showed that in the absence of Pbs2 or Hog1 (or their kinase activities) osmoshock aberrantly elicits a mating response in a Sho1 dependent manner. Sho1's being required for cross-talk indicates that it is necessary for the activation of Ste11. We used this property of

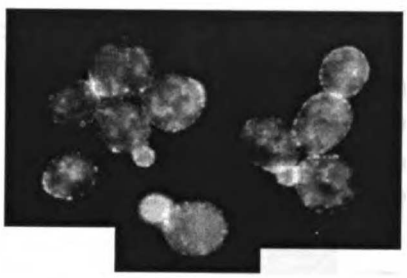
Figure 3.2.

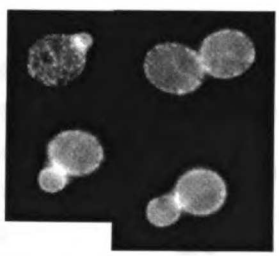
a **Expression of Sho1 Mutants**



b **Localization of Sho1 Mutants**


WT


Δ(172-298)


Δ(306-367)

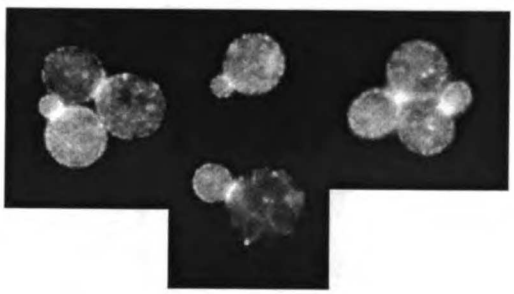
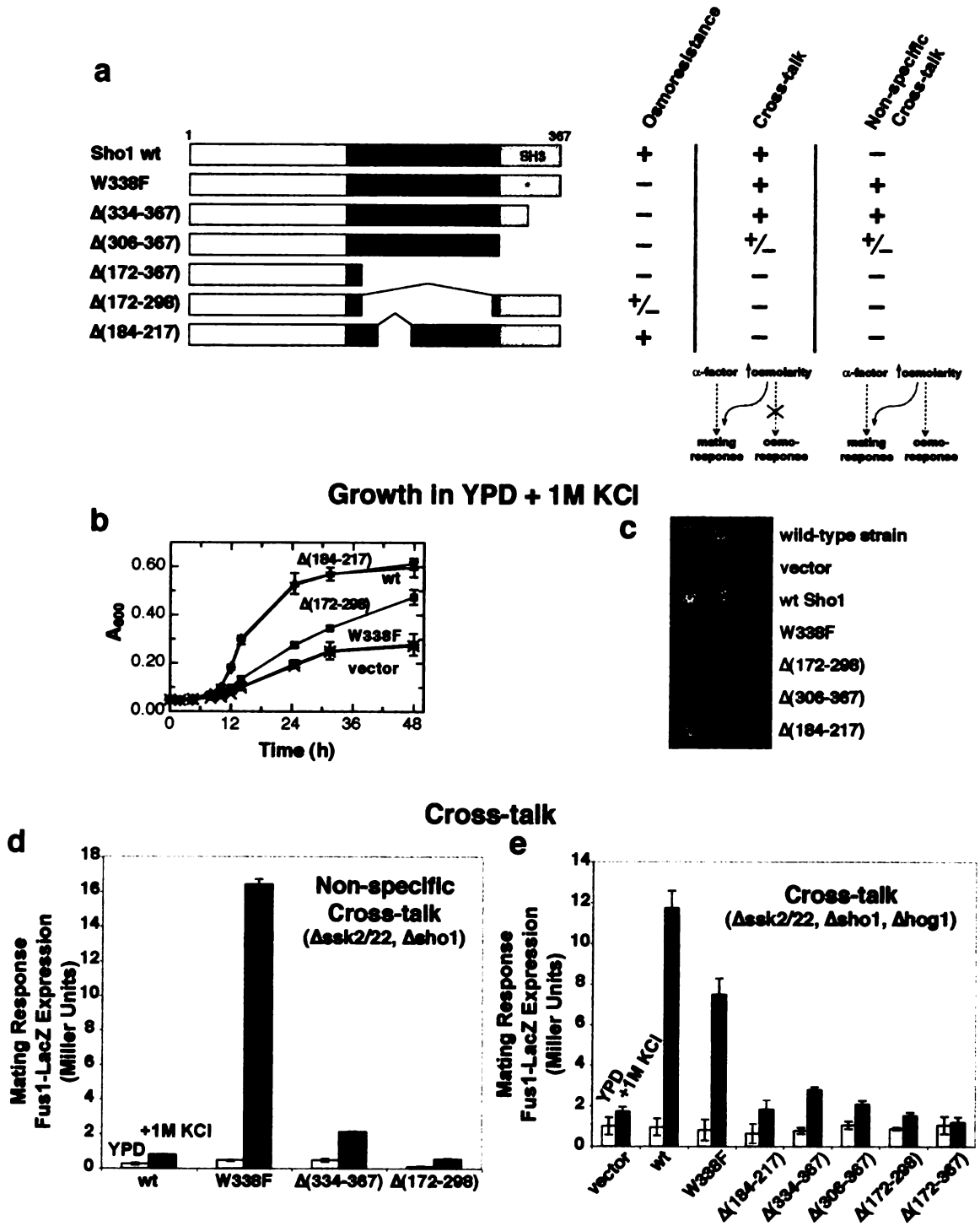

W338F

Figure 3.3. The SH3 domain of Sho1 is required for osmoresistance; the intervening region is required for cross-talk.

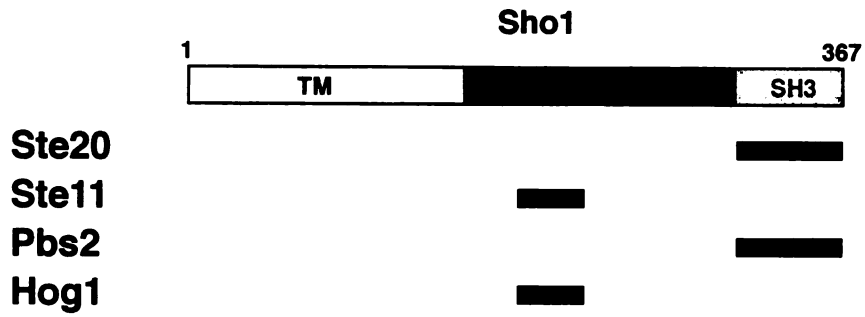


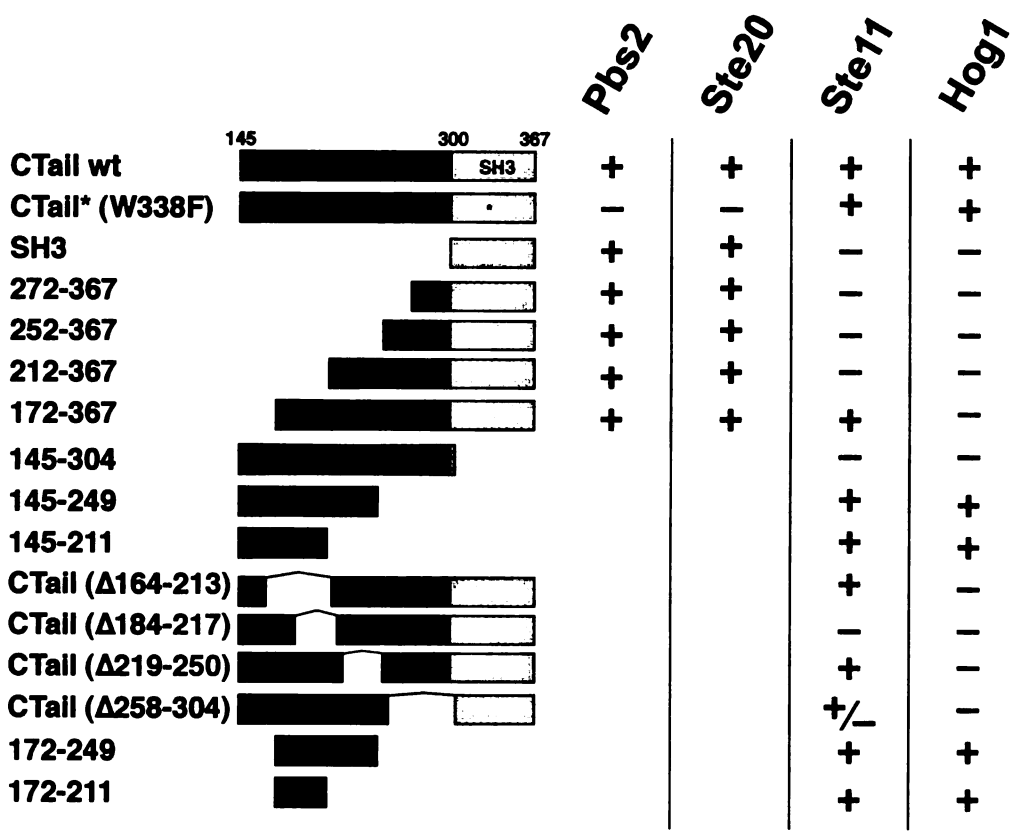
the MAPK pathway to looks for mutations in Sho1 that lead to cross-talk (Fig. 3.3d).

Sho1-W338F, a tryptophan to phenylalanine mutant that does not interact with Pbs2 (Figs. 3.4), also aberrantly activates the mating pathway, mimicking deletions of Pbs2 or of Hog1. The ability of Sho1 to activate the mating pathway without interacting with Pbs2 suggests that Sho1's interactions are not limited to the SH3-peptide interaction, but rather that they extend to other proteins, specifically ones that also act in the mating pathway. Mutants with deletions of the SH3 domain also activate the mating pathway, though they not as effectively as the single point mutation. In fact, the more the SH3 domain is truncated, the lower the level of activation of mating pathway in cross-talk. Furthermore, the aberrant activation of the mating pathway is not observed in strains with an active Sln1 pathway (data not shown), indicating that Hog1 activation regardless of input is sufficient to downregulate cross-talk to the mating pathway.

We then tested various mutants of Sho1 to see which regions were required for activating the mating pathway in cross-talk in Δpbs2 or Δhog1 strains (Fig. 3.3e). This activity seems limited to two regions of Sho1. The SH3 domain, though not its full structure, and a small region in the N-terminal third of the cytoplasmic tail (184-217) appear to be important in activating the mating response. For Sho1 to be competent to cross-talk, the Sho1 SH3 needs to remain intact, though not necessarily capable of binding Pbs2. This suggests the presence of another, non-canonical, proline-rich peptide-independent SH3 function. Interactions between Sho1 and Pbs2 do not affect cross-talk efficiency, because mating response activation levels are the same in hog1Δ and pbs2Δ strains, as well as in wild-type and the mutant of Sho1 that cannot bind Pbs2. In

Figure 3.4. Sho1 binds Ste20, Ste11, Pbs2, and Hog1.





other words, Sho1's ability to activate Ste11 is independent of Sho1's ability to interact with Pbs2.

Sho1 interacts with Ste20, Ste11, Pbs2, and Hog1.

The cross-talk data above suggest interactions between Sho1 and proteins other than Pbs2. To address this issue we conducted binding studies using glutathione S-transferase (GST) fusions of the cytoplasmic tail of Sho1 (CTail, residues 145-367) and some of its fragments. We found Sho1 to interact with many members of the yeast HOG pathway (Fig. 3.4). First, we corroborated previous work (Posas and Saito, 1997) that showed the SH3 domain of Sho1 to be necessary and sufficient for interaction with Pbs2 (Fig. 3.5a) and, because the proteins were both expressed in *E. coli*, the interaction is independent of any other yeast proteins. We found that Sho1 also interacts directly with Ste20 in an SH3 domain dependent manner (Fig. 3.5b). A small portion of the Sho1 CTail (residues 172-211) is sufficient to interact both with the N-terminal 200 residues of Ste11 expressed in bacteria and with full-length Ste11 expressed in a sho1Δ pbs2Δ yeast strain (Figs. 3.5c-e). Deletion of this region of Sho1 greatly attentuates the binding to Ste11. This mutant was also unable to activate the mating pathway in cross-talk (Fig. 3.3e). We also found that a small region in Sho1 (172-211) is also sufficient for direct interaction with the MAPK Hog1 (Fig. 3.6) both in yeast lysates (Δsho1, Δpbs2) and as a purified bacterially expressed protein.

Figure 3.5. Pbs2, Ste20, and Ste11 binding data.

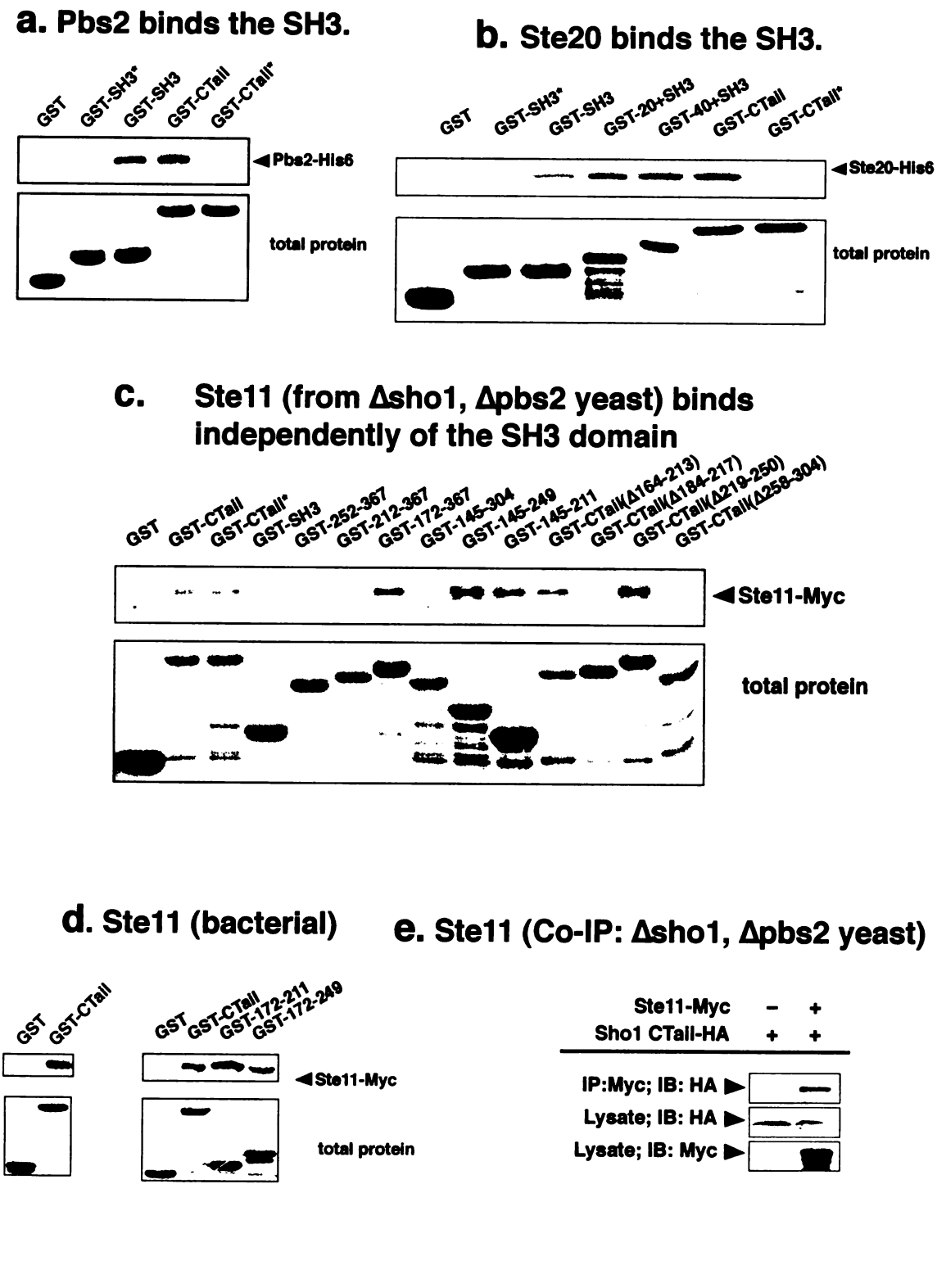
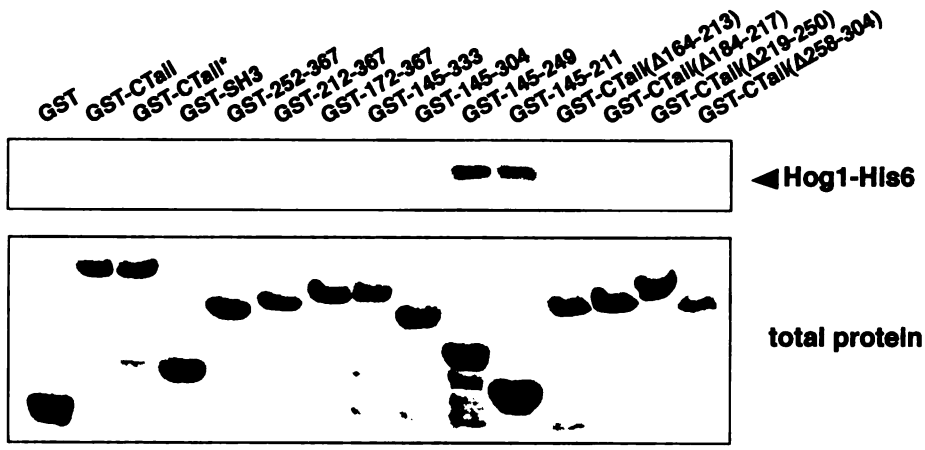
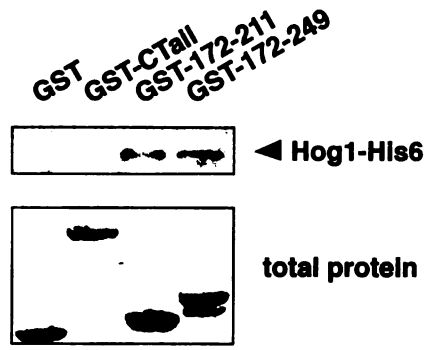


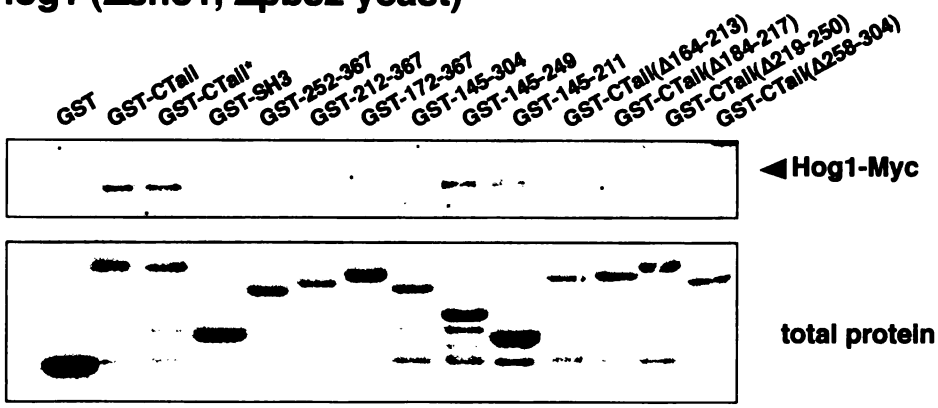
Figure 3.6. Hog1 binding data.

a. Hog1 (bacterial) binding is SH3 domain independent and overlaps with the region important for Ste11 binding.





b. Hog1 (Δsho1, Δpbs2 yeast)



Pbs2 N-terminal region interacts with Ste11 and Hog1

We next examined regions of Pbs2 required for binding Ste11 and Hog1. It has already been shown that Pbs2 interacts with the upstream and the downstream kinases expressed in yeast(Posas and Saito, 1997). We mapped these binding interactions using small fragments from the N-terminal (non-kinase) portion of Pbs2. Residues 51-113 of Pbs2 were sufficient for direct binding to Ste11 (Fig. 3.7a). This data is in good agreement with previous work by the Saito lab who show by in vivo assays that this region was important for Sho1 mediated osmo-response(Tatebayashi et al., 2003). An overlapping region (residues 2-162) also appears to be important in the direct interaction between Pbs2 and Hog1. These overlapping binding regions hint at the possibility of Ste11 and Hog1 competing for the same binding sites on Pbs2. The same may be true in the interactions of Sho1 with Ste11 and Hog1.

The binding of the kinases to Sho1 affects Sho1's affinity for the other kinases.

The number of interactions of Sho1 with the other members of the HOG pathway, as well as their overlapping regions of recognition on Sho1 hints at the possibility of cooperative or competitive modes of binding. We tested for the effects of kinases had on the binding affinities of each other and found an intriguing array of binding interactions. The binding of Ste20 to Sho1 increases the binding of Ste11 and vice versa (Fig. 3.8a). Previous work has shown that Ste20 phosphorylates and thus activates Ste11(Drogen et al., 2000), though binding interactions between the two kinases have not yet been demonstrated. The binding of Pbs2 to Sho1 increases binding of Ste11 to Sho1 (Fig. 3.8b). Finally, the binding of Pbs2 (and surprisingly Ste20) greatly increases the binding

Figure 3.7. Overlapping regions in Pbs2 N-terminus bind Ste11 and Hog1.

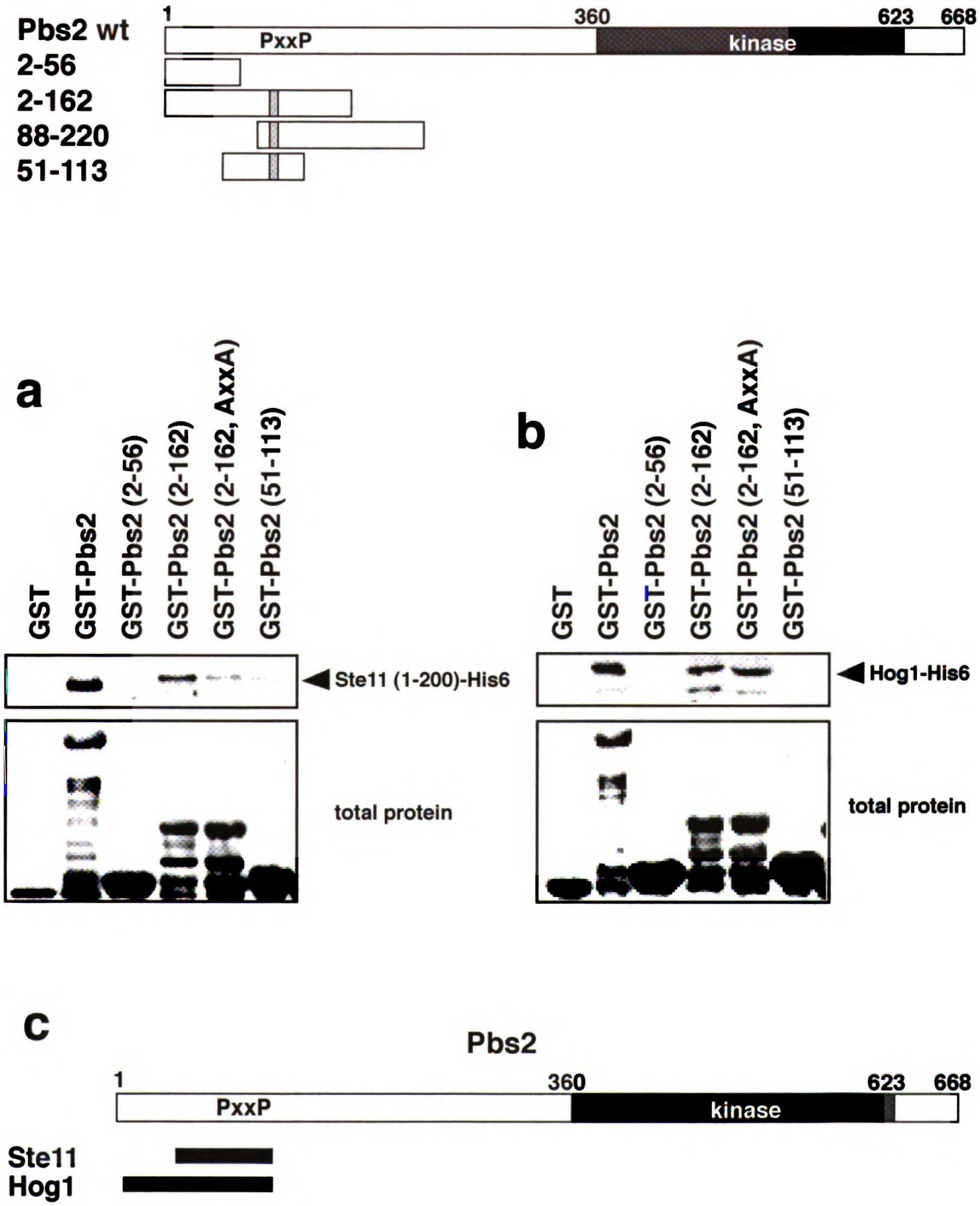
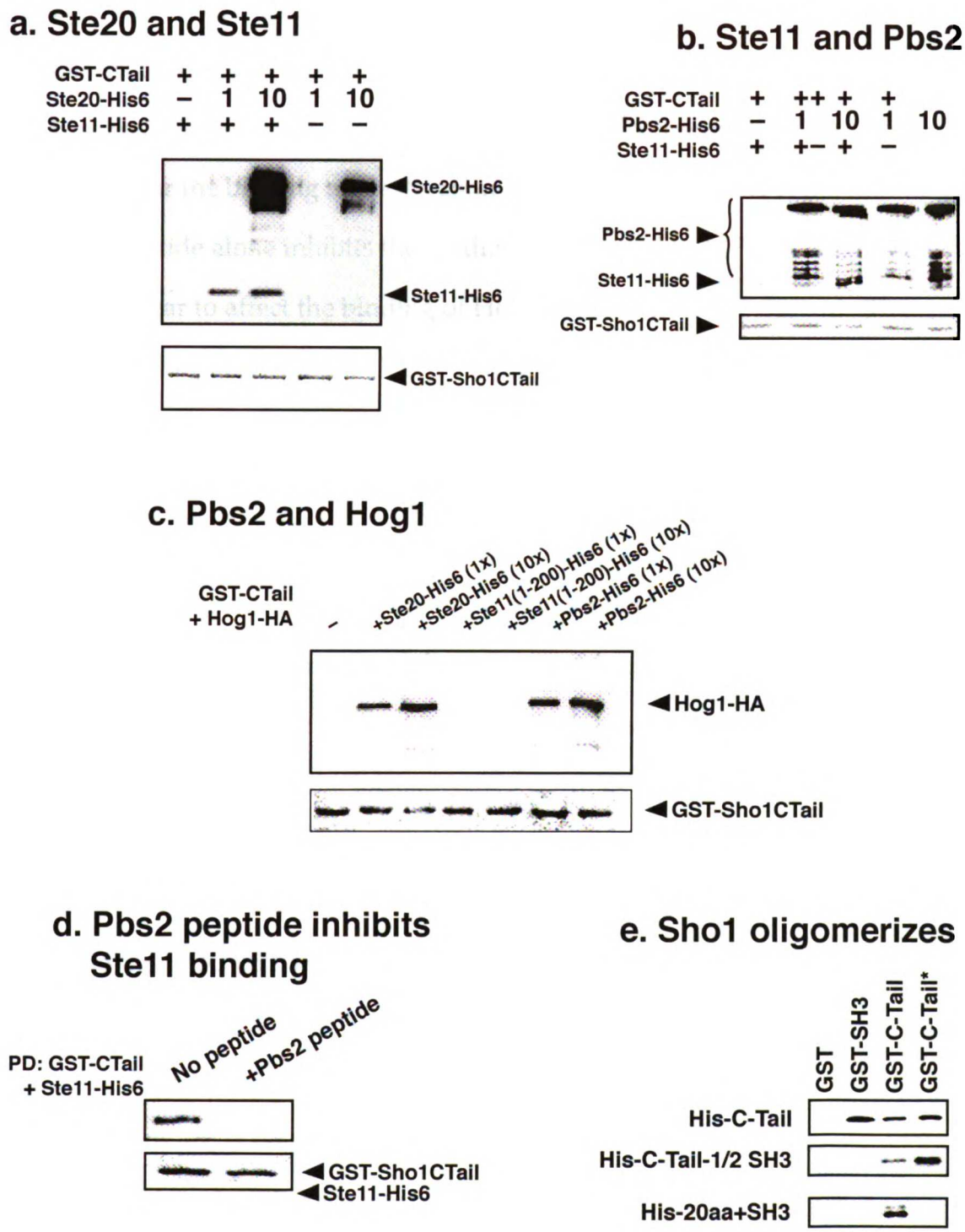


Figure 3.8. Kinases affect each other's binding to Sho1.



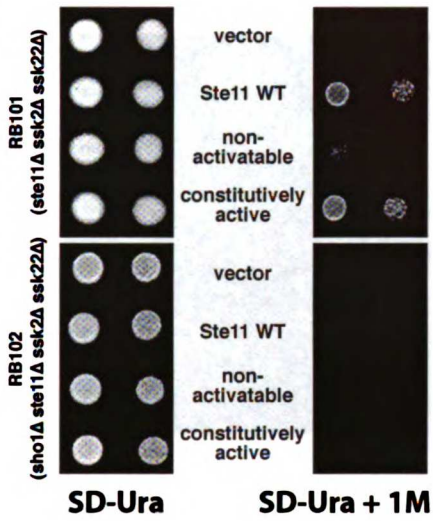
of Hog1 to Sho1 (Fig. 3.8c). This, in addition to high affinity for Hog1 of Sho1 mutants with a truncation of the C-terminal 120 amino acids (Fig. 3.6a), could mean that the binding of Pbs2 to the SH3 domain relieves an auto-inhibitory interaction in Sho1, thus allowing for increased affinity for Hog1. It is also possible that the observed increase in affinity could be simply due to a bridging effect by Pbs2 linking Sho1 and Hog1. Suggestive of a competitive, step-wise mechanism for the binding and activation of the kinases on Sho1 is the finding that Pbs2 peptide alone inhibits the binding of Ste11 to Sho1 (Fig. 3.8d), though it does not appear to affect the binding of Hog1 to Sho1 (data not shown). The finding that Sho1 oligomerizes (Fig. 3.8e) provides another complication in the multiple overlapping interactions.

The complex is necessary for optimal signaling.

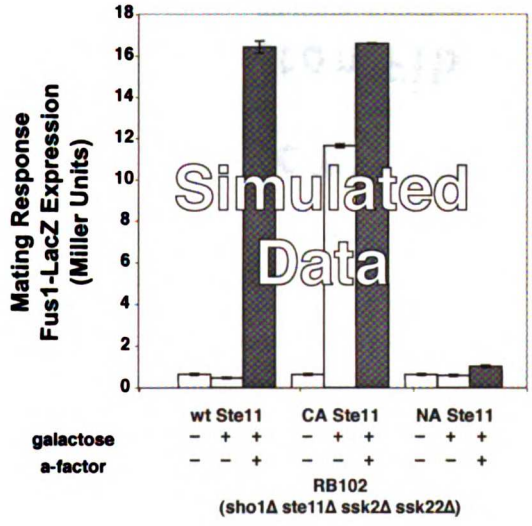
Why are there so many interactions in this pathway? The number of interactions is not surprising given the number of interactions seen in the mating pathway both between the scaffold Ste5 and the kinases and among the kinases themselves. Nonetheless, the formation of a complex appears to be required for optimal signaling. Most importantly, the interaction of Sho1 with both Ste11 and Pbs2 is required for osmo-signaling, both to activate Ste11 and to direct its activity to the HOG pathway. This is based on experiments showing that constitutively active Ste11 still requires the presence of Sho1 for osmo-resistance (Fig. 3.9a). Further evidence of the role of Sho1 in the formation of an active signaling complex, is that Hog1-GFP coalesces into discrete points, similar to the localization of Sho1-GFP, before being transported into the nucleus (Fig. 3.10). We still need to show that the constitutively active Ste11 is active in the mating

Figure 3.9.

a. Sho1 is required for the transmission of signal from activated Ste11 to osmo-pathway.

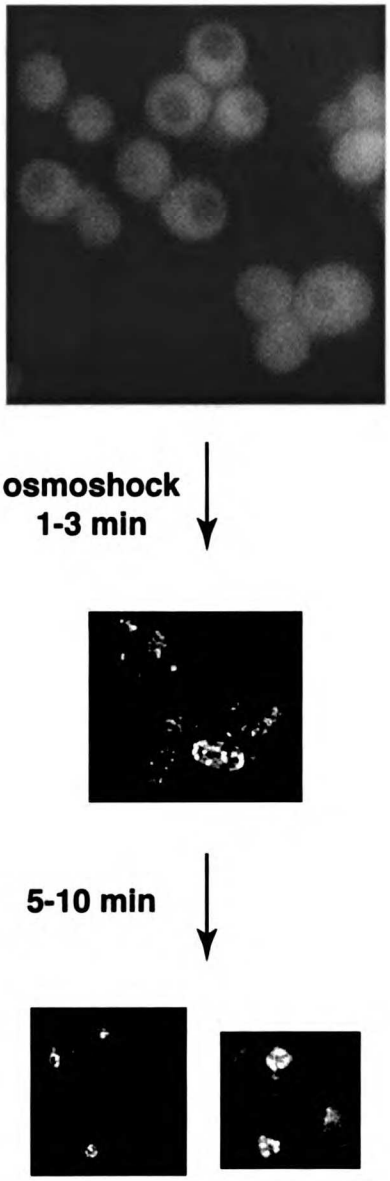


b. contol for activity of Ste11



c. correlation of growth with pathway activity

Figure 3.10. Hog1 localizes to points before going to the nucleus.



pathway and also to show that pathway activation measured by Hog1 phosphorylation correlates with the growth phenotype. Additional evidence for a scaffolding role of Sho1 could be provided by *in vitro* phosphorylation studies (Fig. 3.11a). Can Sho1 enhance the activities of the kinases on their downstream targets in the absence of the scaffolding role of Pbs2? Another unanswered question involves Hog1's ability to suppress cross-talk through its kinase activity. The target of this activity could be Sho1, the most upstream member of the pathway. In that case, we should be able to find mutant Sho1 proteins that are resistant to downregulation, presumably non-phosphorylatable, and thus cross-talk to the other pathways (Fig. 3.11b).

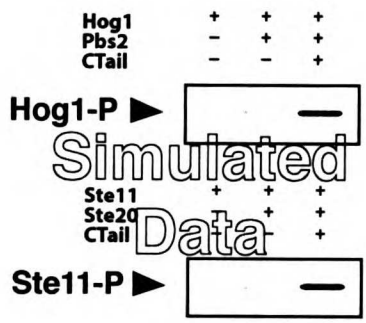
Discussion

As evidenced by the data above, the HOG pathway overlaps significantly with other yeast MAPK cascades. Nevertheless it maintains tight control over the flow of information by mechanisms that are still not well understood. One proposed mechanism involves scaffold proteins that can interact with multiple members of pathways simultaneously, thus limiting the number of possible interactions. In the yeast osmoregulatory pathway, Pbs2 is thought to act as such an element. We have shown above that Sho1 also interacts with multiple members of the pathway, independent of Pbs2, and is essential to the fidelity and activity of the pathway.

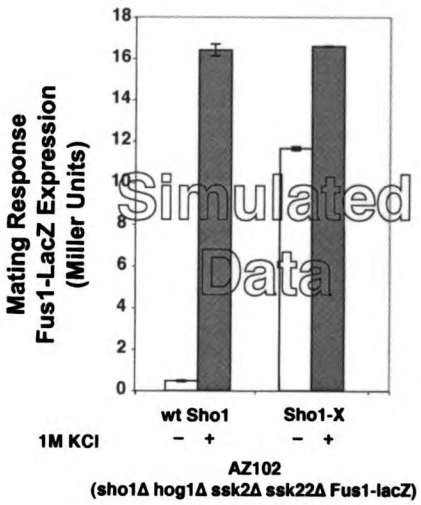
Sho1 does not simply fit this mold. It does not appear to tether the various kinases together in a large complex; rather its overlapping regions of recognition, and its possibly cooperative or competitive interactions suggest a more active role in the step-wise assembly and activation of a signaling complex.

Figure 3.11. Future directions.

a. Can Sho1 enhance in vitro phosphorylation in the osmo-pathway?



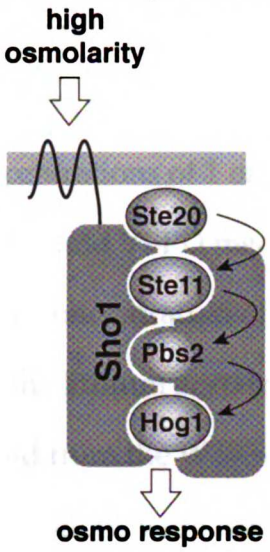
b. Can mutations in Sho1 render it resistant to downregulation, and thus (susceptible to cross-talk)?



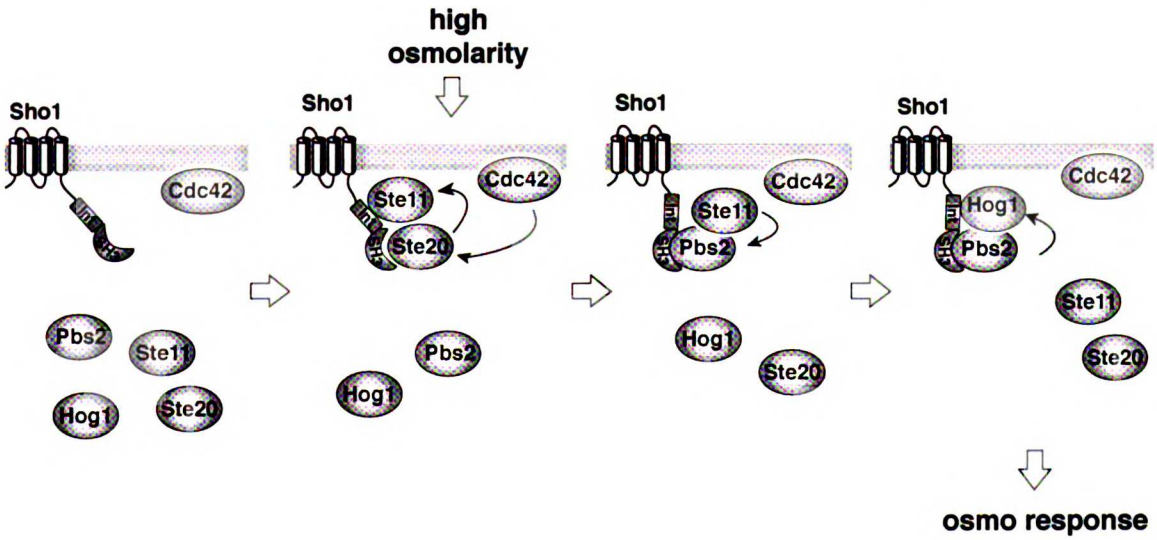
Possible models for the mechanism of activation of Sho1 include a simple tethering mechanism (Fig. 3.12a) and a more involved and active partner in the activation and exchange of kinases (Fig. 3.12b). Both these models imply the presence of scaffold-mediated complexes that reinforce the set of interactions necessary for the maintenance of specificity and signal strength. It also allows for the possibility that a signaling cascade can be broken up into discreet, sequential, and swappable steps, each of which is regulated by a scaffold. These steps could then be regulated as modules. For example, different modules could function in different tissues or at different stages of development. The modular architecture also allows for rapid generation of new pathways.

Figure 3.12.

Model A: Sho1 binds all the other proteins and holds them together as an oligomer



Model B: Sho1 binds the other proteins in discrete complexes



MATERIALS AND METHODS

Strains, media, and genetic techniques

Yeast strains were grown in YEPD medium (1% yeast extract, 2% Bacto-Peptone, 2% glucose) at 30°C. Synthetic complete medium (Rose et al. 1990) was used for maintaining plasmids and selecting gene replacements. D-sorbitol and NaCl (Sigma) were used at final concentrations of 1 or 1.2 M as indicated. For a-factor treatments, cells were grown in liquid YEPD medium, and 0.5 mg/ml a-factor in 0.01 M HCl was added to a final concentration of 0.005 mg/ml. Yeast transformations were done by the lithium acetate procedure (Schiestl and Gietz 1989). Yeast strains were derived from the W303 strain background (*trp1 leu2 ura3 his3 ADE2 can1*). Gene disruptions were confirmed by phenotypic analysis and/or PCR reactions with gene-specific primers.

Escherichia coli strain TG1 was used for cloning and propagating plasmids and strain BL21 (DE3) RIL for the expression of recombinant proteins. To express proteins, cultures were grown to an OD_{600} of 0.6-0.8 at 20 °C and induced with 1 mM IPTG for 3-6 h. Cells were harvested by centrifugation, resuspended in PBS (50 mM sodium phosphate, 100 mM NaCl, pH 7.4), and frozen at -80°C. Subsequently, cell suspensions were thawed and lysed using a Branson model 250 sonifier. Lysates were cleared by centrifugation at 20,000*g*. The His6-fusions were bound to Ni^{2+}-NTA resin (Qiagen) at 4 °C, washed three times with PBS containing 20 mM imidazole, eluted with PBS containing 250 mM imidazole, and dialyzed three times into 100 mM NaCl, 10 mM HEPES (pH 8.0). GST fusions were bound to glutathione agarose at 4 °C, washed three times with PBS, eluted with PBS containing 10mM reduced glutathione, and dialyzed three times into

100 mM NaCl, 10 mM HEPES, 2 mM DTT (pH 8.0). Protein concentration was measured by UV absorbance using calculated extinction coefficients and individual aliquots were stored at -80 °C.

Hyper-osmotic Plate Growth Assay. 10^3 cells were spotted onto YPD plates with or without 1M KCl. Plates were incubated at 30° C for 3 to 5 days.

b-*Galactosidase Assay.* LacZ expression was measured as described previously (Stern et al. 1984), except that log-phase cells were treated for 5 hrs as indicated by diluting into fresh medium, medium containing a-factor, or medium containing 1 M KCl prior to harvesting.

Microscopy

Green fluorescent protein (GFP) tagged proteins and cell morphology were visualized by using a Nikon Microphot-SA microscope with a 100 objective lens and a Princeton Instruments cooled charge-coupled device camera (RTE/CCD-1300-V).

Coimmunoprecipitation experiments, Immunoblots.

Protein extracts and immunoblots were carried out as described previously (Park et al., 2003).

Chapter 4

Discussion and Future Directions

Signal transduction is based primarily on the interaction of macromolecules with one another. For these processes to occur efficiently, they must be of high fidelity. Examples of specific, macromolecular complex assembly were known in a number of biological processes, from DNA replication, transcription, and translation, to tyrosine kinase and G-protein coupled receptor signaling, prior to the work described above. Our findings are two-fold. First, we have shown that through proteome-wide negative selection, binding partners can evolve to achieve specific interaction. Second, we have shown that in the case of proteins with multiple physiological partners, scaffold proteins can organize specific pair-wise interactions into functional signaling complexes that direct the activity of one protein onto its relevant target. The various mechanisms described in the previous chapters, i.e. interactions requiring multiple partners and selection for specificity only within the milieu of the organism of origin, are a subset of the possible means of achieving specificity. Much more work remains to address some of the questions raised by these models.

In the case of proteome-dependent domain-mediated specificity, the question of why there are so many proline-recognition domains still hasn't been explored adequately. This abundance may be a simple result of the proliferation of a successful solution to the problem of protein recognition. Having more domain types presumably allows the evolution of more complex signaling networks. Further, having a suite of domains that recognize similar or overlapping motifs may provide additional modes of interaction regulation (Sudol, 1996). If domains from distinct family members recognize a single motif,

the competition between these alternative partners could, in principle, act as a regulatory switch. Relatively little is known about the functional intersection between different domain families in vivo. In one known case, T cell activation appears to promote this type of domain interaction swap: a receptor proline-rich motif that initially interacts with a GYF domain, after stimulation interacts with an SH3 domain (Freund et al., 2002). More studies may reveal more such examples.

More studies are also needed in a number of other areas of domain-ligand specificity. With the SH3 domain arrays in hand, it would be interesting to study a number of known proline-rich peptides with or without known binding partners. With which domains do these peptides interact? Are these interactions specific? Are all physiological interactions specific or are there physiological interactions that are required to be non-specific (or at least have multiple partners)? A distinct advantage of having modular binding domains is the relative ease with which such interactions can be generated.

Our study also examined the evolution of only one pair of interacting proteins. We showed that the peptide was selected for specific interaction. Though more complex, it would be interesting to see if the binding domains themselves were under selective pressure of any kind. A more nuanced approach would be required in this case. A structurally informed or peptide-library approach to deciphering the binding profiles of all the SH3 domains in a genome could lead to the discovery of a more or less orthogonal set of SH3 domains, each of which can have a specific peptide ligand. What is the largest

number of mutually orthogonal SH3 domains and can we develop a corresponding set of specific peptides for them?

Further studies are also warranted on other genomes and other domains. Do other modular binding domains, such as SH2, PDZ, EVH1, and WW domains, also have binding partners that are as specific for them as the Pbs2 peptide is for the Sho1 SH3 domain? Do the SH3 domain networks in other yeast, such as *Candida albicans* also display such specificity? Is the Pbs2 peptide no longer specific when probed against the library of SH3 domains from closely related species such as *C. albicans* or *Kluyveromyces lactis*?

Many questions also remain unanswered in the case of the scaffolding interactions of Sho1 and Pbs2. Is there a requirement for the cooperative interactions *in vivo*? In other words, are the binding sites for Hog1 or Ste11 on both Pbs2 and Sho1 required for function? Does deletion of that region or ablation of the binding interaction have some effect on function, efficacy, or specificity? Further quantitation of the cooperative effects of the proteins on one another is an important step in understanding the mechanism of osmo-responsive MAPK signaling. What is the proper order of addition of the proteins to the scaffold? Also unexplored is the effect of phosphorylation on the binding affinities within this complex. Signaling complexes are, by nature, transient and their assembly needs to be regulated. Phosphorylation is one mechanism of control over the binding events and signal intensity and endurance.

Overall, the roles that negative domain-ligand selection and scaffold-mediated protein complex assembly play in optimizing protein interaction network specificity are only beginning to be understood. The mechanism have a

key role in the construction of many biological networks, ranging from protein signaling networks to DNA-binding/transcriptional networks. Questions about the transient assembly of active signaling complexes abound in all areas of biology. Both negative selection and scaffold-mediated assembly are likely to affect the construction of many biological networks and subsequently how these networks are used to achieve specific and efficient signaling. Their importance is only accentuated by the vast amounts of raw data generated by the push to sequence whole genomes. The decoding of these large datasets depends partly on vast computational power and partly on the discovery of simple rules by which genomes and proteomes have been evolutionarily organized.

APPENDIX A

METHODS FOR SH3 ARRAYS

Protein Expression and Purification. All the yeast SH3 domains were cloned by PCR from *Saccharomyces cerevisiae* genomic DNA. The other SH3 domains were cloned from appropriate cDNA libraries. His6-tagged Pbs2 peptides, fused to the N-terminal domain of lambda repressor (res. 1-99), were constructed as described by Maxwell and Davidson(Maxwell and Davidson, 1998). *Escherichia coli* strain TG1 was used for cloning and propagating the plasmids, strain BL21 (DE3) RIL for the expression of recombinant proteins. To express proteins, cultures were grown to an OD_{600} of 0.6-0.8 at 20 °C and induced with 1 mM IPTG for 3-6 h. Cells were harvested by centrifugation, resuspended in PBS (50 mM sodium phosphate, 100 mM NaCl, pH 7.4), and frozen at -80°C. Subsequently, cell suspensions were thawed and lysed using a Branson model 250 sonifier. Lysates were cleared by centrifugation at 20,000*g*. The His6-fusions were bound to Ni^{2+}-NTA resin (Qiagen) at 4 °C, washed three times with PBS containing 20 mM imidazole, eluted with PBS containing 250 mM imidazole, and dialyzed three times into 100 mM NaCl, 10 mM HEPES (pH 8.0). Alternatively, cell suspensions with the His6-tagged Pbs2 peptides were harvested by centrifugation, resuspended in guanidine lysis buffer (PBS + 6M guanidine HCl + 20 mM imidazole), incubated for 1 hour, and sonicated briefly (30 seconds) to shear the DNA. After clearing the lysates by centrifugation, the His6-fusions were incubated with the nickel resin for 15 minutes at 4 °C, and the resin was washed three times with the lysis buffer. The proteins were then eluted and dialyzed as above. GST fusions were bound to glutathione agarose at 4 °C, washed three times with PBS, eluted with PBS containing 10mM reduced glutathione, and dialyzed three times into 100 mM NaCl, 10 mM HEPES, 2 mM DTT (pH 8.0). Protein concentration was measured by UV absorbance using

calculated extinction coefficients from ProtParam (us.expasy.org/tools/protparam.html) and individual aliquots were stored at -80 °C.

Generating the GST-SH3 Array. To make the arrays, the Dot-Blot apparatus was set up with a pre-wetted nitrocellulose membrane placed on top of a pre-wetted Whatman paper (both 8 x 12 cm). 100μL each of 100 nM solutions of purified GST-SH3 fusion proteins in TBST were added to the wells and incubated for 15 minutes at RT. To control for variation in antibody levels and development exposure, standards of a GST and His-tagged protein (100 μL of 100 nM and 10 nM solutions of empty pETARA) were directly spotted onto the membrane. Vacuum was applied slowly such that it took approximately one minute to draw the solutions were drawn through the membrane. The nitrocellulose membranes were then immediately placed in blocking buffer (3% milk/1% BSA in TBST) and shaken for 1 h at RT.

Probing and developing the GST-SH3 Array. Array membranes were probed with 6 mL of blocking buffer containing a His-tagged fusion protein containing the proline-rich peptide of interest at a concentration of 50μM for 4-16 h at 4°C. The membranes were washed four times in TBST (two times quickly and vigorously and two times for 5 minutes each) and reprobed with a horseradish peroxidase-conjugated anti-His antibody (1:2000 dilution, Santa Cruz Biotech.) for 1 h at 4 °C. Finally, the blot was developed with an ECL system (Pierce SuperSignal West Pico) and quantitated on an AlphaInnotech CCD camera and

analytical software.

Analyzing the GST-SH3 Array. Spot intensities were quantitated as described in Figure 2.13a. The raw value for each spot was taken to be the intensity inside a circle entirely enclosing the spot. The background for each spot (calculated based on the total intensity within a inner circle surrounding the spot subtracted from the total intensity within an outer circle 50% larger than the inner circle divided by the area of this ring) was subtracted. Variations in spotting were corrected by dividing by the intensity of the spot on a replica array probed only with anti-GST (i.e. measuring total GST-SH3 protein in spot). Variations in exposure were corrected by dividing by the intensities of the reference His-tag spots described above. The corrected intensities for each spot are given relative to the intensity for the Sho1 SH3 domain spot probed with wild-type peptide. The semi-quantitative nature of this assay was validated by comparing spot intensities from the SH3 domain arrays to *in vitro* measured dissociation constants (Fig. 2.13b). The two sets of measurements show a correlation fit to the equation: $\log K_d = k \log I + c$ (K_d is the dissociation constant, I is the spot intensity, k and c are constants).

APPENDIX B

REFERENCES

Saccharomyces Genome Database, http://www.yeastgenome.org/.

Ball, L. J., Jarchau, T., Oschkinat, H., and Walter, U. (2002). EVH1 domains: structure, function and interactions. FEBS Lett *513*, 45-52.

Barnett, P., Bottger, G., Klein, A. T., Tabak, H. F., and Distel, B. (2000). The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. Embo J *19*, 6382-6391.

Bedford, M. T., Frankel, A., Yaffe, M. B., Clarke, S., Leder, P., and Richard, S. (2000a). Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. J Biol Chem *275*, 16030-16036.

Bedford, M. T., Reed, R., and Leder, P. (1998). WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. Proc Natl Acad Sci U S A *95*, 10602-10607.

Bedford, M. T., Sarbassova, D., Xu, J., Leder, P., and Yaffe, M. B. (2000b). A novel pro-Arg motif recognized by WW domains. J Biol Chem *275*, 10359-10369.

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.

Bottger, G., Barnett, P., Klein, A. T., Kragt, A., Tabak, H. F., and Distel, B. (2000). Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner. Mol Biol Cell *11*, 3963-3976.

Brannetti, B., Via, A., Cestra, G., Cesareni, G., and Helmer-Citterich, M. (2000). SH3-SPOT: an algorithm to predict preferred ligands to different members of the SH3 gene family. J Mol Biol *298*, 313-328.

Buday, L., and Downward, J. (1993). Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. Cell *73*, 611-620.

Buday, L., Wunderlich, L., and Tamas, P. (2002). The Nck family of adapter proteins: regulators of actin cytoskeleton. Cell Signal *14*, 723-731.

Cesareni, G., Panni, S., Nardelli, G., and Castagnoli, L. (2002). Can we infer peptide recognition specificity mediated by SH3 domains? FEBS Lett *513*, 38-44.

Douangamath, A., Filipp, F. V., Klein, A. T., Barnett, P., Zou, P., Voorn-Brouwer, T., Vega, M. C., Mayans, O. M., Sattler, M., Distel, B., and Wilmanns, M. (2002). Topography for independent binding of alpha-helical and PPII-helical ligands to a peroxisomal SH3 domain. Mol Cell *10*, 1007-1017.

Drogen, F., O'Rourke, S. M., Stucke, V. M., Jaquenoud, M., Neiman, A. M., and Peter, M. (2000). Phosphorylation of the MEKK Ste11p by the PAK-like kinase Ste20p is required for MAP kinase signaling in vivo. Curr Biol *10*, 630-639.

Elion, E. A. (2001). The Ste5p scaffold. J Cell Sci *114*, 3967-3978.

Farr, T. J., Coddington-Lawson, S. J., Snyder, P. M., and McDonald, F. J. (2000). Human Nedd4 interacts with the human epithelial Na^+-channel: WW3 but not WW1 binds to Na^+-channel subunits. Biochem J *345 Pt 3*, 503-509.

Fazi, B., Cope, M. J., Douangamath, A., Ferracuti, S., Schirwitz, K., Zucconi, A., Drubin, D. G., Wilmanns, M., Cesareni, G., and Castagnoli, L. (2002). Unusual binding properties of the SH3 domain of the yeast actin-binding protein Abp1: structural and functional analysis. J Biol Chem *277*, 5290-5298.

Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. Science *266*, 1241-1247.

Feng, S., Kasahara, C., Rickles, R. J., and Schreiber, S. L. (1995). Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. Proc Natl Acad Sci U S A *92*, 12408-12415.

Freund, C., Kuhne, R., Yang, H., Park, S., Reinherz, E. L., and Wagner, G. (2002). Dynamic interaction of CD2 with the GYF and the SH3 domain of compartmentalized effector molecules. Embo J *21*, 5985-5995.

Ghose, R., Shekhtman, A., Goger, M. J., Ji, H., and Cowburn, D. (2001). A novel, specific interaction involving the Csk SH3 domain and its natural ligand. Nat Struct Biol *8*, 998-1004.

Girzalsky, W., Rehling, P., Stein, K., Kipper, J., Blank, L., Kunau, W. H., and Erdmann, R. (1999). Involvement of Pex13p in Pex14p localization and peroxisomal targeting signal 2-dependent protein import into peroxisomes. J Cell Biol *144*, 1151-1162.

Harris, B. Z., and Lim, W. A. (2001). Mechanism and role of PDZ domains in signaling complex assembly. J Cell Sci *114*, 3219-3231.

Hiroaki, H., Ago, T., Ito, T., Sumimoto, H., and Kohda, D. (2001). Solution structure of the PX domain, a target of the SH3 domain. Nat Struct Biol *8*, 526-530.

Holt, M. R., and Koffer, A. (2001). Cell motility: proline-rich proteins promote protrusions. Trends Cell Biol *11*, 38-46.

Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. Nat Struct Biol *7*, 634-638.

Ilsley, J. L., Sudol, M., and Winder, S. J. (2002). The WW domain: linking cell signalling to the membrane cytoskeleton. Cell Signal *14*, 183-189.

InterPro http://www.ebi.ac.uk/interpro/.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., and Kim, S. K. (2001). Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. Proc Natl Acad Sci U S A *98*, 218-223.

Kami, K., Takeya, R., Sumimoto, H., and Kohda, D. (2002). Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. Embo J *21*, 4268-4276.

Kapoor, T. M., Andreotti, A. H., and Schreiber, S. L. (1998). Exploring the specificity pockets of two homologous SH3 domains using structure-based, split-pool synthesis and affinity-based selection. J Am Chem Soc *120*, 23-29.

Karathanassis, D., Stahelin, R. V., Bravo, J., Perisic, O., Pacold, C. M., Cho, W., and Williams, R. L. (2002). Binding of the PX domain of p47(phox) to phosphatidylinositol 3,4-bisphosphate and phosphatidic acid is masked by an intramolecular interaction. Embo J *21*, 5057-5068.

Kay, B. K., Williamson, M. P., and Sudol, M. (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. Faseb J *14*, 231-241.

Kim, S. K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J. M., Eizinger, A., Wylie, B. N., and Davidson, G. S. (2001). A gene expression map for Caenorhabditis elegans. Science *293*, 2087-2092.

Kuribayashi, F., Nunoi, H., Wakamatsu, K., Tsunawaki, S., Sato, K., Ito, T., and Sumimoto, H. (2002). The adaptor protein p40(phox) as a positive regulator of the superoxide-producing phagocyte oxidase. Embo J *21*, 6312-6320.

Kwok, P. Y., and Duan, S. (2003). SNP discovery by direct DNA sequencing. Methods Mol Biol *212*, 71-84.

Ladbury, J. E., and Arold, S. (2000). Searching for specificity in SH domains. Chem Biol *7*, R3-8.

Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J. (1996). Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell *85*, 931-942.

Lim, W. A., and Richards, F. M. (1994). Critical residues in an SH3 domain from Sem-5 suggest a mechanism for proline-rich peptide recognition. Nat Struct Biol *1*, 221-225.

Lim, W. A., Richards, F. M., and Fox, R. O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature *372*, 375-379.

Lott, J. S., Coddington-Lawson, S. J., Teesdale-Spittle, P. H., and McDonald, F. J. (2002). A single WW domain is the predominant mediator of the interaction between the human ubiquitin-protein ligase Nedd4 and the human epithelial sodium channel. Biochem J *361*, 481-488.

Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., Bar-Sagi, D., and Schlessinger, J. (1992). The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. Cell *70*, 431-442.

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.
MacArthur, M. W., and Thornton, J. M. (1991). Influence of proline residues on protein conformation. J Mol Biol *218*, 397-412.

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.

Macias, M. J., Wiesner, S., and Sudol, M. (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. FEBS Lett *513*, 30-37.

Maxwell, K. L., and Davidson, A. R. (1998). Mutagenesis of a buried polar interaction in an SH3 domain: sequence conservation provides the best prediction of stability effects. Biochemistry *37*, 16172-16182.

Mayer, B. J. (2001). SH3 domains: complexity in moderation. J Cell Sci *114*, 1253-1263.

McGee, A. W., Dakoji, S. R., Olsen, O., Bredt, D. S., Lim, W. A., and Prehoda, K. E. (2001). Structure of the SH3-guanylate kinase module from PSD-95 suggests a mechanism for regulated assembly of MAGUK scaffolding proteins. Mol Cell *8*, 1291-1301.

McPherson, P. S. (1999). Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. Cell Signal *11*, 229-238.

Moarefi, I., LaFevre-Bernt, M., Sicheri, F., Huse, M., Lee, C. H., Kuriyan, J., and Miller, W. T. (1997). Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. Nature *385*, 650-653.

Mongiovi, A. M., Romano, P. R., Panni, S., Mendoza, M., Wong, W. T., Musacchio, A., Cesareni, G., and Di Fiore, P. P. (1999). A novel peptide-SH3 interaction. Embo J *18*, 5300-5309.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., and Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. Nature *359*, 851-855.

Newman, J. R., and Keating, A. E. (2003). Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science *300*, 2097-2101.

Nguyen, J. T., and Lim, W. A. (1997). How Src exercises self-restraint. Nat Struct Biol *4*, 256-260.

Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N., and Lim, W. A. (1998). Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. Science *282*, 2088-2092.

Nishida, M., Nagata, K., Hachimori, Y., Horiuchi, M., Ogura, K., Mandiyan, V., Schlessinger, J., and Inagaki, F. (2001). Novel recognition mode between Vav and Grb2 SH3 domains. Embo J *20*, 2995-3007.

O'Rourke, S. M., and Herskowitz, I. (1998). The Hog1 MAPK prevents cross talk between the HOG and pheromone response MAPK pathways in Saccharomyces cerevisiae. Genes Dev *12*, 2874-2886.

Orr, M. R., and Smith, T. B. (1998). Ecology and speciation. Trends Ecol Evol *13*, 502-506.

Palmer, E. (2003). Negative selection--clearing out the bad apples from the T-cell repertoire. Nat Rev Immunol *3*, 383-391.

Park, S. H., Zarrinpar, A., and Lim, W. A. (2003). Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. Science *299*, 1061-1064.

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

Petrella, E. C., Machesky, L. M., Kaiser, D. A., and Pollard, T. D. (1996). Structural requirements and thermodynamics of the interaction of proline peptides with profilin. Biochemistry *35*, 16535-16543.

Posas, F., and Saito, H. (1997). Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: scaffold role of Pbs2p MAPKK. Science *276*, 1702-1705.

Raitt, D. C., Posas, F., and Saito, H. (2000). Yeast Cdc42 GTPase and Ste20 PAK-like kinase regulate Sho1-dependent activation of the Hog1 MAPK pathway. Embo J *19*, 4623-4631.

Reinke, V., Smith, H. E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S. J., Davis, E. B., Scherer, S., Ward, S., and Kim, S. K. (2000). A global profile of germline gene expression in C. elegans. Mol Cell *6*, 605-616.

Reiser, V., Salah, S. M., and Ammerer, G. (2000). Polarized localization of yeast Pbs2 depends on osmostress, the membrane protein Sho1 and Cdc42. Nat Cell Biol *2*, 620-627.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science *259*, 1157-1161.

Renfranz, P. J., and Beckerle, M. C. (2002). Doing (F/L)PPPPs: EVH1 domains and their proline-rich partners in cell polarity and migration. Curr Opin Cell Biol *14*, 88-103.

Rickles, R. J., Botfield, M. C., Weng, Z., Taylor, J. A., Green, O. M., Brugge, J. S., and Zoller, M. J. (1994). Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. Embo J *13*, 5598-5604.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., and Bowtell, D. (1993). The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature *363*, 83-85.

Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., *et al.* (2000). Comparative genomics of the eukaryotes. Science *287*, 2204-2215.

Saksela, K., Cheng, G., and Baltimore, D. (1995). Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. Embo J *14*, 484-491.

Schneider, T. D. (2000). Evolution of biological information. Nucleic Acids Res *28*, 2794-2799.

Shen, Y. H., Godlewski, J., Zhu, J., Sathyanarayana, P., Leaner, V., Birrer, M. J., Rana, A., and Tzivion, G. (2003). Cross-talk between JNK/SAPK and ERK/MAPK pathways: sustained activation of JNK blocks ERK activation by mitogenic factors. J Biol Chem *278*, 26715-26721.

Siligardi, G., and Drake, A. F. (1995). The importance of extended conformations and, in particular, the PII conformation for the molecular recognition of peptides. Biopolymers *37*, 281-292.

Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quillam, L. A., and Kay, B. K. (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. Proc Natl Acad Sci U S A *93*, 1540-1544.

Sudol, M. (1996). The WW module competes with the SH3 domain? Trends Biochem Sci *21*, 161-163.

Sudol, M. (1998). From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'. Oncogene *17*, 1469-1474.

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

Sudol, M., Sliwa, K., and Russo, T. (2001). Functions of WW domains in the nucleus. FEBS Lett *490*, 190-195.

Tatebayashi, K., Takekawa, M., and Saito, H. (2003). A docking site determining specificity of Pbs2 MAPKK for Ssk2/Ssk22 MAPKKKs in the yeast HOG pathway. Embo J *22*, 3624-3634.

Terasawa, H., Kohda, D., Hatanaka, H., Tsuchiya, S., Ogura, K., Nagata, K., Ishii, S., Mandiyan, V., Ullrich, A., Schlessinger, J., and et al. (1994). Structure of the N-terminal SH3 domain of GRB2 complexed with a peptide from the guanine nucleotide releasing factor Sos. Nat Struct Biol *1*, 891-897.

Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., *et al.* (2002). A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science *295*, 321-324.

Tu, J. C., Xiao, B., Yuan, J. P., Lanahan, A. A., Leoffert, K., Li, M., Linden, D. J., and Worley, P. F. (1998). Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. Neuron *21*, 717-726.

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.

Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002). Proline *cis-trans* isomerization and protein folding. Biochemistry *41*, 14637-14644.

Whitmarsh, A. J., and Davis, R. J. (1999). Signal transduction by MAP kinases: regulation by phosphorylation-dependent switches. Sci STKE *1999*, PE1.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. Biochem J *297 (Pt 2)*, 249-260.

Wollacott, A. M., and Desjarlais, J. R. (2001). Virtual interaction profiles of proteins. J Mol Biol *313*, 317-342.

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

Yaffe, M. B., Leparc, G. G., Lai, J., Obata, T., Volinia, S., and Cantley, L. C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. Nat Biotechnol *19*, 348-353.

Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. Cell *76*, 933-945.

Yuzawa, S., Yokochi, M., Hatanaka, H., Ogura, K., Kataoka, M., Miura, K., Mandiyan, V., Schlessinger, J., and Inagaki, F. (2001). Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. J Mol Biol *306*, 527-537.

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