The Structure and Function of Proline Recognition Domains

Ali Zarrinpar,[1,2] Roby P. Bhattacharyya,[1,2] Wendell A. Lim[2]*

One particularly abundant group of modular recognition domains consists of those that bind proline-rich motifs. Such modules, including the SH3, WW, and EVH1 domains, play a critical role in the assembly and regulation of many intracellular signaling complexes. These domains use strikingly similar molecular mechanisms of proline recognition. We discuss some of the potential biological advantages conferred by proline recognition, which may explain its widespread use in signaling.

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 (*1*) are the most common sequence motif in the *Drosophila* genome and the second most common in the *Caenorhabditis elegans* genome (*2*). 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).

Proline recognition domains are usually found in the context of larger multidomain signaling proteins. Their binding events often direct the assembly and targeting of protein complexes involved in cell growth (*3-5*), cytoskeletal rearrangements (*6*, *7*), transcription (*8*), postsynaptic signaling (*9*, *10*), and other key cellular processes (*11*). In addition, these interactions can play a regulatory role, often through autoinhibitory interactions that are alleviated by competing binding events (*12*).

Several recent reviews discuss individual proline recognition domains (*9*, *13*, *14*). This review 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 (*15*) and in many cases are thought to function as docking sites for signaling modules (*16*). 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. 1A): the unusual shape of its pyrrolidine ring, the conformational constraints on 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 that is frequently used in binding to signaling domains is their propensity to form a polyproline type II (PPII) helix. The PPII helix is an extended left-handed helical structure with three residues per turn and an

Organism	SH3	WW	EVH1	GYF
Saccharomyces cerevisiae	25	5	1	3
Caenorhabditis elegans	66	18	2	3
Drosophila melanogaster	90	27	5	2
Mus musculus	163	39	16	2
Homo sapiens	332	80	20	5

Table 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 (*90*), 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 (*91*). They are not included in the table because they lack certain key conserved residues, and the structure and function of these domains are unknown.

overall shape resembling a triangular prism (Fig. 1B) (*15*, *17*). A combination of steric and hydrogen-bonding properties of proline-rich motifs is thought to contribute to its preference for this unusual secondary structure (*15*, *17*). Two features of the PPII helix make it a useful recognition motif. First, in this structure both the side chains and the backbone carbonyls point out from the helical axis into solution at regular intervals (Fig. 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 car-

[1]Program in Biological Sciences, [2]Department of Cellular and Molecular Pharmacology and Department of Biochemistry and Biophysics, University of California San Francisco, 600 16th Street, San Francisco, CA 94143-2240, USA.

*Corresponding author. Telephone, 415-502-8080; fax, 415-514-4242; e-mail: wlim@itsa.ucsf.edu

bonyls can easily be "read" by interacting proteins (*18*). Second, because the backbone conformation in a PPII helix is already restricted, the entropy cost of binding is reduced (*16*, *19*). Nearly all of the domains described here bind their ligands in a PPII conformation. 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 twofold rotational pseudosymmetry: Side chains and backbone carbonyls are displayed with similar spacing in either of the two N- to C-terminal orientations (Fig. 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 orientational flexibility could play an important role in domain function. For example, one could imagine a complex in which binding in one orientation could be activating, whereas binding in the opposite orientation could be inhibitory. However, such an orientational switching 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 δ carbon on the substituted amide nitrogen (Fig. 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. 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 nonnatural N-substituted amino acids that do not resemble proline (*20*).

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% as compared to negligible amounts for the other amino acids (*15*). Moreover, the kinetic barrier for

C-substitution/N-substitution
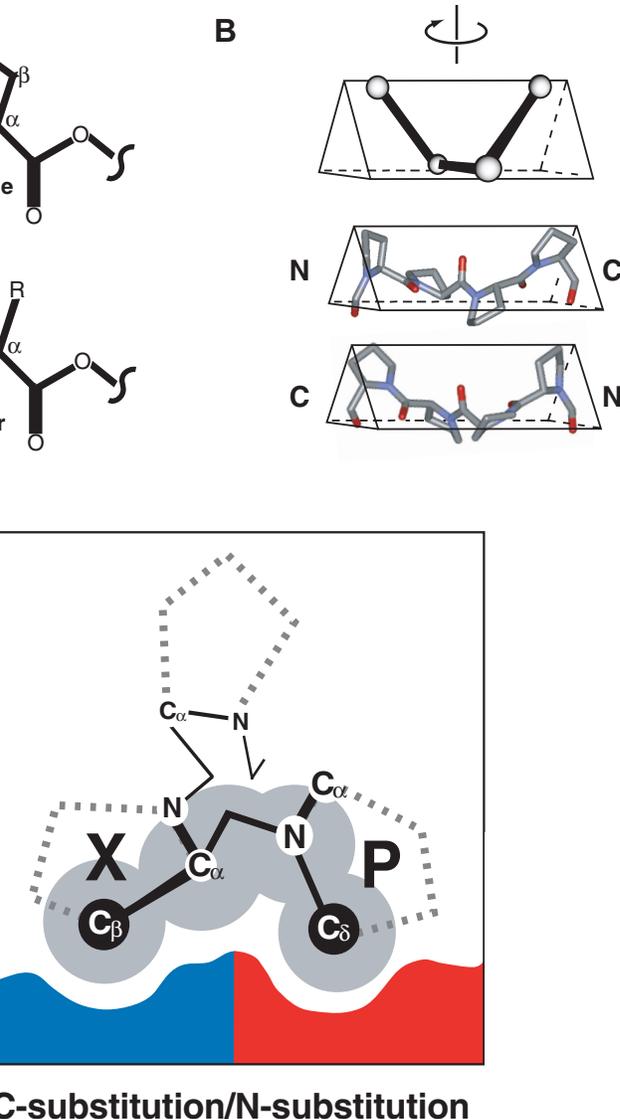
Fig. 1. Properties of proline and polyproline sequences. (**A**) Chemical structure of proline contrasted with that of other natural amino acids. Proline possesses a five-member ring fused onto the nitrogen, making it a secondary amine, whereas other amino acids have side chains that only branch off of the α carbon, leaving a primary amine. (**B**) Schematic and structural representation of a PPII helix. The helix has twofold pseudosymmetry: A rotation of 180° about a vertical axis leaves the proline rings and the carbonyl oxygens at approximately the same position. The Protein Data Bank (PDB) accession code for the poly-(l)-proline structure shown is 1CF0 (http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics&pdbId=1CF0&page=&pid=298021048527478). To view this structure in motion, see [Structures (http://stke.sciencemag.org/cgi/content/full/sigtrans;2003/179/re8/DC1)] (**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, whereas in the "P" position that requires N-substitution (red), the primary recognition element is the δ carbon that is unique to proline.

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 (*21*). 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.
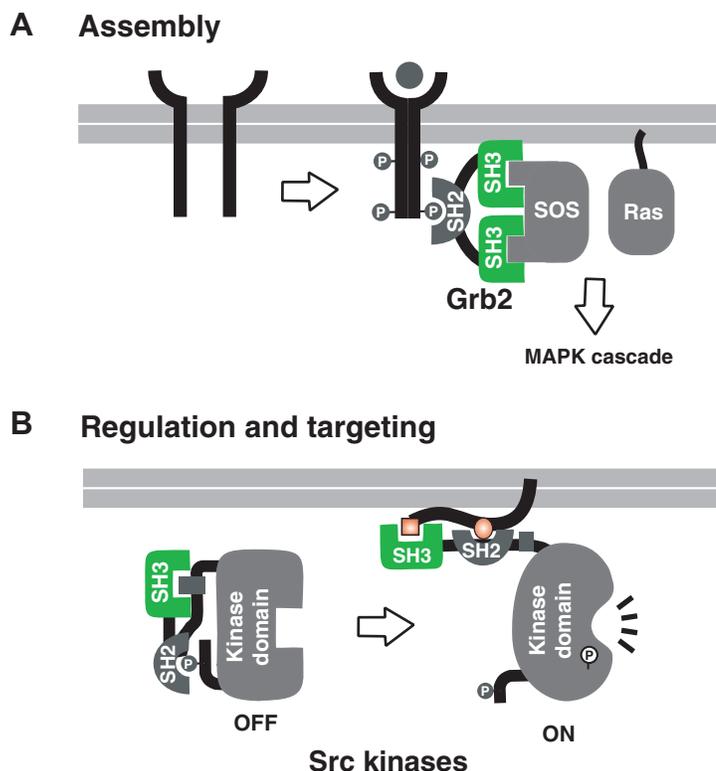
SH3 Domains

The first characterized and best understood example of the proline recognition modules is the SH3 domain (14). 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. 2A) (4). 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 (3, 5). Similar recruitment roles are played by SH3 domain-containing proteins in various other biological processes, including endocytosis (11) and cytoskeletal dynamics (22).

SH3 domains also play regulatory roles. An excellent example of this is the Src family of tyrosine kinases (Fig. 2B) (12, 23). 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 (24-26). 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,

A Assembly

Grb2

MAPK cascade

B Regulation and targeting

OFF   ON

Src kinases

Fig. 2. Functional roles of SH3 domains. (**A**) Assembly role of SH3 domains. Growth factor stimulation leads to the activation of receptor tyrosine kinases and to the phosphorylation of the receptor tail, of related adaptor proteins (not shown), or of 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 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.

which were initially viewed as static assembly elements, can function as dynamic switches by alternating binding partners (intra- versus intermolecular). SH3 interactions tend to be fairly weak, with typical dissociation constants (K_d's) in the μM range (14). 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 (27-30). 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 (*31-34*). Each of these two recognition motifs corresponds to the sequence preferences for a distinct orientation of binding. Efforts are under way to use the extensive peptide library data to generate algorithms to predict SH3 recognition (*30, 35, 36*).

Structures of SH3 domains both alone and in complex with ligand reveal their mechanism of recognition (Fig. 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 (*37, 38*). When bound, the proline-rich peptide ligand adopts a PPII helix conformation (*33, 34, 39*). Recognition of this structure is achieved by insertion of the ridges of the PPII helix into a 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. 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 (*20*). This mode of recognition also explains why SH3 domains can bind ligands in two possible orientations: A PPII ligand has twofold rotational pseudosymmetry, with respect to both the steric properties of the xP unit and the presentation of hydrogen-bonding groups (the backbone carbonyls) that are used in recognition (Fig. 1B).

Adjacent to the core recognition surface of SH3 domains are the more variable RT and n-Src loops (Fig. 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 (*28, 40*). 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 (*29, 41*). 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, used by a handful of SH3 domains, is the evolution of a noncanonical recognition
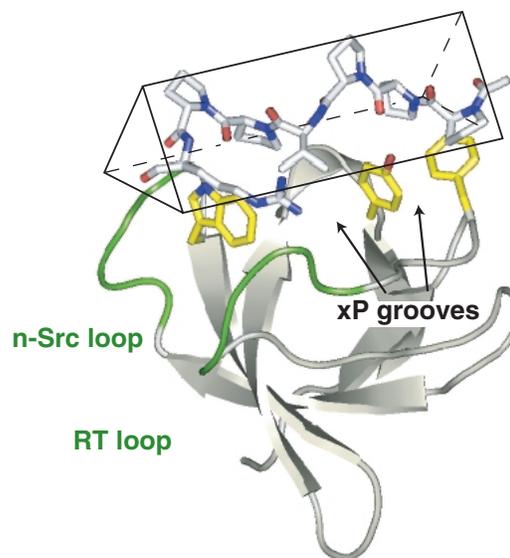
Sem5 SH3
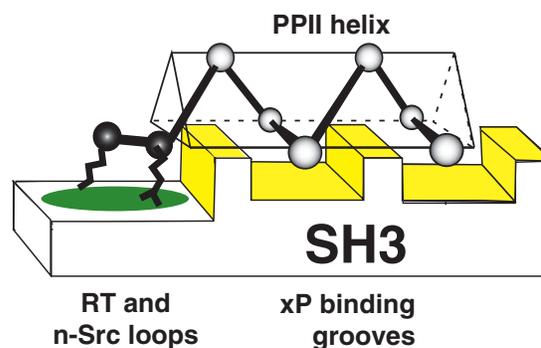
Fig. 3. Structure and binding mechanism of SH3 domains. The structure of the Sem5 SH3 domain in complex with a proline-rich ligand is shown. A cartoon of the proline-binding surface of these domains docked with a ligand, showing the general mechanism of recognition, is shown below. The core recognition 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 Protein Data Bank (PDB) accession code for this structure is 1SEM (http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics& pdbId=1SEM&page=&pid=1691048625252). To view this structure in motion, see [Structures (http://stke.sciencemag.org/cgi/content/full/ sigtrans;2003/179/re8/DC1)]

mechanism. Several SH3 domains recognize non-PxxP motifs. This is the case for the SH3 domains of Eps8, which recognizes PxxDY (*42*); Gads, which recognizes RxxK (*43*); and Fus1, which recognizes Arg-Ser-rich sequences (*41*). In most of these cases, it is unclear whether this recognition is mediated by the equivalent surface used by canonical SH3 domains to recognize PxxP ligands. 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 (*44*). This interaction may play a role in the assembly of signaling complexes at cell-cell junctions.

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 (*23*, *45*, *46*). 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 (*41*) or the above examples of Grb2 (*47*) and p47phox. Moreover, SH3 domains could function together with other modules such as SH2, PDZ (named after 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 (*48*). 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 (*49*). 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 (*50*).
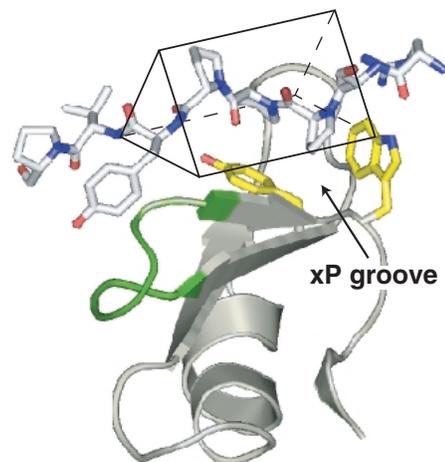
WW Domains

WW domains mediate protein-protein interactions in diverse processes (*13*). For example, the WW domains of the ubiquitin ligase Nedd4 bind to Na^+-channel subunits, thereby targeting ubiquitin-mediated down-regulation of channel activity (*51*). 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 (*8*). In addition, pre-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 (*52*, *53*).

WW domains can be divided into several classes based on recognition motifs (*54*). 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) (*52*, *55*); the FBP11 and FE65 WW domains prefer Pro-Pro-Leu-Pro (PPLP) (*56*); and the FBP21, FBP30, and Npw38 WW domains prefer Pro-Arg (P-R) repeats (*57*, *58*). 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 to their unphosphorylated counterparts. In contrast, interactions with PPxY motifs can be abolished by tyrosine phosphorylation (*59-61*).

The structures of WW domain-ligand complexes reveal a striking mechanistic similarity to those of SH3s and other proline recognition domains (Fig. 4) (*62*). Containing 35 to 45 residues, WW domains are highly compact binding domains, comprising an antiparallel three-stranded fold (*55*). 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 side chains against the ridges and grooves on the domain-binding surface (*52*, *60*). The aromatic groove in the WW domain also recognizes an xP

Dystrophin WW
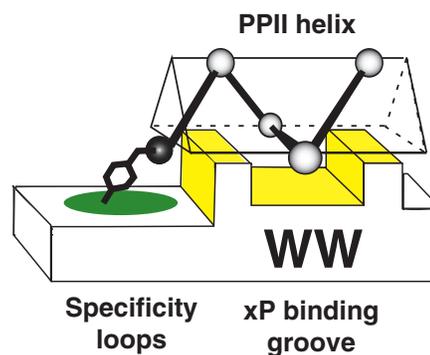
Fig. 4. Structure and binding mechanism of WW domains. The structure of the dystrophin WW domain in complex with a proline-rich ligand is shown. A cartoon of the proline-binding surface of these domains docked with a ligand, showing the general mechanism of recognition, is shown below. The core recognition 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 (http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics&pdbId=1EG4&page=&pid=2091048625292). To view this structure in motion, see [Structures (http://stke.sciencemag.org/cgi/content/full/sigtrans;2003/179/re8/DC1)]

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 as 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.
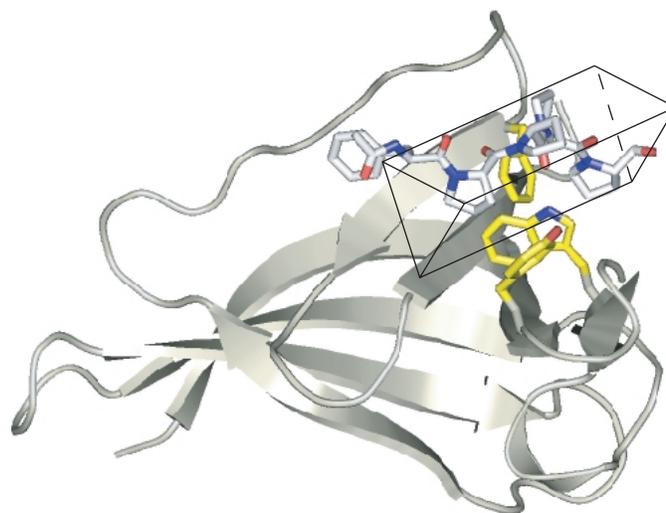
EVH1 Domains

A third class of polyproline-binding domains, EVH1 domains, also typically play a recruitment or targeting role, often in events involving actin cytoskeleton dynamics (*6*, *7*) and postsynaptic signaling (*9*). Examples of proteins with EVH1 domains that contribute to targeting of actin polymerization include the Wiskott-Aldrich syndrome protein (WASP) family (*63*) and the Ena/VASP family (*64*). Molecular targets of these domains include the cellular focal adhesion proteins vinculin and zyxin (*7*), as well as the ActA protein from *Listeria monocytogenes* (*65*). Recruitment of Ena/VASP proteins to ActA is thought to promote the motility of this intracellular pathogen. By hijacking the cellular actin polymerization machinery, *Listeria* can move from cell to cell through adjacent plasma membranes to evade immune detection (*66*, *67*). The Homer (also known as Vesl) family exemplifies EVH1 domain-containing proteins that are thought to direct targets to synaptic signaling complexes (*68*). The best characterized targets for this family are the cytosolic portion of group I metabotropic glutamate receptors (mGluRs), inositol 1,4,5-trisphosphate receptors (IP3Rs), ryanodine receptors (RyRs), and the Shank family of postsynaptic density proteins (*10*, *68*, *69*).

Individual EVH1 domains have distinct ligand recognition profiles (*9*). Most domains recognize relatively short peptides comprising a proline-rich core and a few flanking residues with K_d's in the low μM range. One group of EVH1 domains, including the Ena/VASP family of proteins, recognizes the consensus sequence FPxϕP (ϕ

is a hydrophobic residue), where x and ϕ are often prolines (*70*). Another group, including Homer, binds the distinct consensus PPxxF (*10*). More recently, a distinct class of EVH1 domains, including those from the WASP family member N-WASP and the nucleoporin RanBP1, has emerged that recognizes extended ligand sequences of up to 25 amino acids (*71-73*).

Structures of EVH1 domains from several proteins, some in complex with their proline-rich ligands, reveal that EVH1 ligands also bind in a PPII helical conformation and dock against an aromatic-rich binding surface on the domain (Fig. 5). However, the mechanism by which EVH1 domains recognize the PPII helix is different from that used by SH3 and WW domains.

Mena EVH1
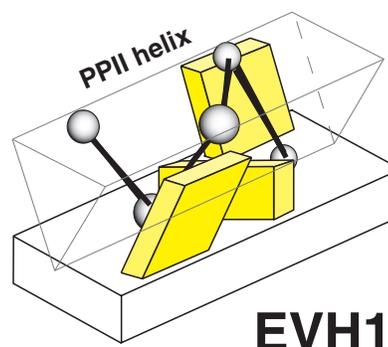
EVH1

Fig. 5. Structure and binding mechanism of EVH1 domains. A representative structure of the Mena EVH1 domain in complex with a peptide ligand is shown. Below is a schematic of the recognition mechanism showing the apex of the PPII helix fitting into an aromatic-rich wedge at the binding surface. Although a conserved set of aromatic residues (yellow) also contacts the PPII ligand, the manner in which the PPII helix docks against the domain surface differs from that observed in most other proline-binding domains discussed here. The PDB accession code for this structure is 1EVH (http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics&pdbId=1EVH&page=&pid=2741048625354).

The binding surface of the EVH1 domain is concave, allowing complementary packing of the apex of the PPII helix. This interface contrasts with the relatively flat recognition surface of SH3 and WW domains that allows complementary packing of the base of the PPII helix. Consequently, whereas SH3 and WW ligand recognition depends largely on side-chain contacts with the backbone amide substitution that is unique to proline, EVH1 ligand recognition seems to depend more on the conformation of the PPII helix. This difference in recognition mechanisms may explain differences in the precise ligand structure required for binding. Although a chemical dissection approach demonstrated that the "required" prolines in an SH3 or WW ligand could effectively be replaced by nonnatural N-substituted amino acids, such as sarcosine (20), this type of approach failed to reveal a similar trend for EVH1 domains (74). This observation is consistent with EVH1 domains recognizing a feature of the proline-rich core other than the repetitive presentation of a substituted amide nitrogen in an xP dipeptide motif.

EVH1 complex structures also demonstrate a critical role in recognition for motifs outside of the proline-rich core. The Phe residue that flanks the central prolines in EVH1 ligands docks into a hydrophobic pocket adjacent to the PPII recognition surface of the domain. Also, charged residues near the aromatic binding surface of some EVH1 domains interact with oppositely charged residues that flank the core proline residues in the ligand (75, 76). An additional role for flanking sequences in recognition is exemplified by the N-WASP EVH1 domain. Although ligands for this domain contain a short (~10-residue) proline-rich sequence, this motif alone is insufficient for detectable binding. This domain instead requires an extended ligand of about 25 amino acids. Structural and biochemical analysis revealed that the ligand wraps around the EVH1 domain, making energetically important contacts with extensive surfaces beyond the PPII binding interface (72). A similar extended interaction is seen between the EVH1 domain of the nucleoporin RanBP1 and its ligand, Ran (73).

EVH1 domains can recognize their ligands in either of two N- or C-terminal orientations (72). This flexibility is likely a consequence of the fact that, like SH3 and WW domains, EVH1 domains recognize their ligands in the pseudosymmetric PPII conformation (9, 72).

Other Proline-Rich Binding Domains
A number of other proline recognition domains deserve mention, though they may be less abundant within sequenced genomes and are less well characterized than those mentioned above. Among this group are Gly-Tyr-Phe (GYF) domains, profilin, and the ubiquitin E2 variant (UEV) domain from Tsg101. Despite their divergent sequences, these domains seem to share some molecular strategies for proline recognition.

The GYF domain was first identified in CD2BP2, a binding protein for the human transmembrane T cell adhesion molecule CD2. This interaction enhances interleukin-2 production upon T cell activation. These domains are named after a highly conserved Gly-Tyr-Phe motif. The domain in CD2BP2 recognizes tandem PPPPGHR sequences separated by seven amino acids in CD2 (77). Deletion of either PPPPGHR repeat in CD2 abolishes binding to the CD2BP2 GYF domain in a yeast two-hybrid assay (77), although a single repeat was sufficient for weakened but detectable binding in vitro (78). This behavior is consistent with two CD2BP2 molecules, each using its own GYF domain to bind the tandem recognition motif in a cooperative manner. Structural studies are also consistent with a 1:1 binding stoichiometry between the GYF domain and the

PPPPGHR motif.

A short peptide ligand binds the CD2BP2 GYF domain in a PPII helical conformation (Fig. 6) (78). The base of the helix docks against a relatively flat surface, as is observed in SH3 and WW domains. An xP dipeptide motif on one turn of the helix inserts into a hydrophobic, largely aromatic pocket on this surface. In addition, negatively charged residues near the binding pocket interact with Arg residues in the peptide that are required for efficient binding, and these electrostatic interactions outside of the PPII peptide core
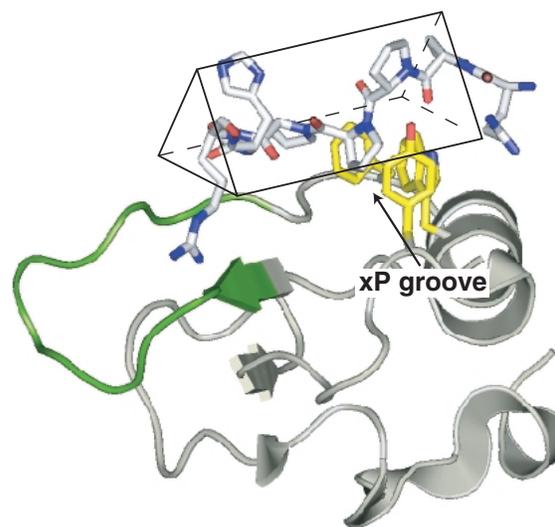
CD2BP2 GYF
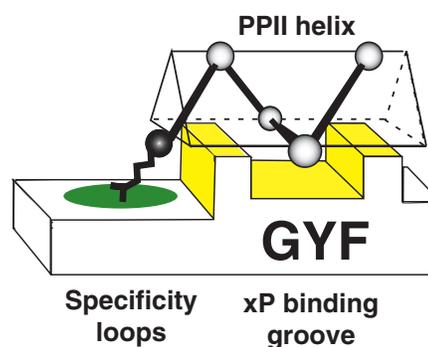
Fig. 6. Structure and binding mechanism of a GYF domain. The structure of the CD2BP2 GYF domain in complex with a proline-rich ligand is shown. A cartoon of the proline-binding surface of these domains docked with a ligand is shown below. The core recognition surface has one xP binding groove formed by aromatic amino acids (yellow) and adjacent, less conserved specificity pockets (green). The PDB accession code for this structure is 1L2Z (http://www.rcsb.org/ pdb/cgi/explore.cgi?job=graphics&pdbId=1L2Z&page=&pid=3141048625386). To view this structure in motion, see [Structures (http://stke.sciencemag.org/cgi/content/full/sigtrans;2003/179/re8/DC1)]

are implicated in determining the specificity of this GYF domain (*77*, *78*). There is also evidence for steric, nonelectrostatic specificity determinants in the loops around the proline-binding pocket in the CD2BP2 GYF domain. Whereas the aromatic residues involved in core PPII helix docking are conserved across all known GYF domains, the residues interacting with flanking residues differ among these domains, separating GYF domains into a number of apparently distinct subclasses (*78*). This variability supports the notion that in GYF domains, as in other proline recognition motifs, the residues flanking the core prolines contribute to specificity among members of a given domain family.

Profilin modulates actin dynamics by binding actin monomers and restricting their addition to one end of an actin filament, thus contributing to the polarization of actin polymerization (*79*, *80*). In addition, profilin binds proline-rich sequences on such targeting proteins as formins and focal adhesion proteins like VASP, localizing it to sites of extensive actin filament assembly (*7*, *81*). A crystal structure of profilin complexed with a polyproline decamer reveals extensive contacts between the ligand in a PPII helix conformation and five highly conserved, solvent-exposed aromatic residues on profilin (*82*). Mutagenic studies validate the importance of these residues in proline binding (*83*). Thus, profilin further exemplifies the intimate involvement of aromatic residues in PPII helix recognition.

UEV domains are found in the human protein Tsg101, which plays an important role in vesicular sorting and is coopted for the budding of HIV and Ebola virus. Virus budding involves the interaction between a 150-amino acid UEV domain in Tsg101 and a PTAP peptide motif in various structural virus proteins (with a K_d in the low µM range) (*84*, *85*). Several putative cellular targets of Tsg101 also contain related proline-rich motifs. These targets include hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs), which is also involved in Tsg101-dependent protein sorting, and plasma membrane proteins such as connexins 43 and 45, which are substrates of this sorting pathway (*86*). Tsg101 itself contains a PTAP sequence that binds its own UEV domain in vitro (*84*), suggesting that the UEV domain could play an autoregulatory role.

A structure of the UEV domain in complex with a PTAP-containing peptide reveals that the peptide adopts a PPII helical conformation over part of its length. The second xP dipeptide in this ligand is bound in a deep groove formed by the side chains of two key tyrosines from Tsg101 whose aromatic rings are positioned in a very similar manner to those in the proline-binding pockets of SH3 and WW domains (*84*).

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 such as phospho-

amino acids, as exemplified by SH2 and phosphotyrosine-binding (PTB) domains, (*87*) or carboxy termini, as exemplified by PDZ domains (*88*). 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 K_d's with values ranging from high nM to low µM. 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 (*89*). 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 interacts with an SH3 domain after stimulation (*78*).

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. 7). 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
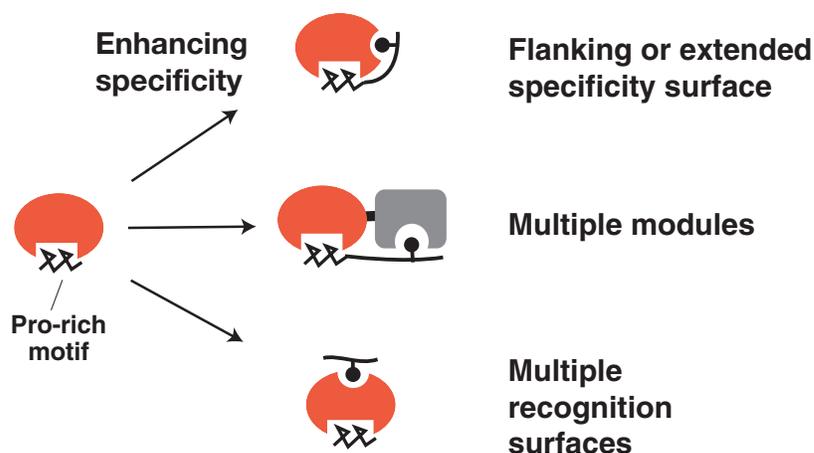
Fig. 7. Potential mechanisms for enhancing the specificity of proline-binding domains. One means of increasing specificity in proline-mediated interactions is by extending the interaction surface with the peptide to include residues beyond the proline-rich core. Another mechanism is to include a nearby 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.

work together to achieve specific recognition. 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.

References

1. See InterPro at http://www.ebi.ac.uk/interpro/.
2. G. M. Rubin, M. D. Yandell, J. R. Wortman, G. L. Gabor Miklos, C. R. Nelson, I. K. Hariharan, M. E. Fortini, P. W. Li, R. Apweiler, W. Fleischmann, J. M. Cherry, S. Henikoff, M. P. Skupski, S. Misra, M. Ashburner, E. Birney, M. S. Boguski, T. Brody, P. Brokstein, S. E. Celniker, S. A. Chervitz, D. Coates, A. Cravchik, A. Gabrielian, R. F. Galle, W. M. Gelbart, R. A. George, L. S. Goldstein, F. Gong, P. Guan, N. L. Harris, B. A. Hay, R. A. Hoskins, J. Li, Z. Li, R. O. Hynes, S. J. Jones, P. M. Kuehl, B. Lemaitre, J. T. Littleton, D. K. Morrison, C. Mungall, P. H. O'Farrell, O. K. Pickeral, C. Shue, L. B. Vosshall, J. Zhang, Q. Zhao, X. H. Zheng, S. Lewis, Comparative genomics of the eukaryotes. *Science* **287**, 2204-2215 (2000).
3. M. Rozakis-Adcock, R. Fernley, J. Wade, T. Pawson, D. Bowtell, The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Nature* **363**, 83-85 (1993).
4. E. J. Lowenstein, R. J. Daly, A. G. Batzer, W. Li, B. Margolis, R. Lammers, A. Ullrich, E. Y. Skolnik, D. Bar-Sagi, J. Schlessinger, The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* **70**, 431-442 (1992).
5. L. Buday, J. Downward, 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 (1993).
6. P. J. Renfranz, M. C. Beckerle, Doing (F/L)PPPPs: EVH1 domains and their proline-rich partners in cell polarity and migration. *Curr. Opin. Cell Biol.* **14**, 88-103 (2002).
7. M. R. Holt, A. Koffer, Cell motility: Proline-rich proteins promote protrusions. *Trends Cell Biol.* **11**, 38-46 (2001).
8. M. Sudol, K. Sliwa, T. Russo, Functions of WW domains in the nucleus. *FEBS Lett.* **490**, 190-195 (2001).
9. L. J. Ball, T. Jarchau, H. Oschkinat, U. Walter, EVH1 domains: Structure, function and interactions. *FEBS Lett.* **513**, 45-52 (2002).
10. J. C. Tu, B. Xiao, J. P. Yuan, A. A. Lanahan, K. Leoffert, M. Li, D. J. Linden, P. F. Worley, Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. *Neuron* **21**, 717-726 (1998).
11. P. S. McPherson, Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. *Cell Signal.* **11**, 229-238 (1999).
12. J. T. Nguyen, W. A. Lim, How Src exercises self-restraint. *Nat. Struct. Biol.* **4**, 256-260 (1997).
13. M. J. Macias, S. Wiesner, M. Sudol, WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.* **513**, 30-37 (2002).
14. B. J. Mayer, SH3 domains: Complexity in moderation. *J. Cell Sci.* **114**, 1253-1263 (2001).
15. M. W. MacArthur, J. M. Thornton, Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397-412 (1991).
16. B. K. Kay, M. P. Williamson, M. Sudol, The importance of being proline: The interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231-241 (2000).
17. M. P. Williamson, The structure and function of proline-rich regions in proteins. *Biochem. J.* **297** (part 2), 249-260 (1994).
18. G. Siligardi, A. F. Drake, The importance of extended conformations and, in particular, the PII conformation for the molecular recognition of peptides. *Biopolymers* **37**, 281-292 (1995).
19. E. C. Petrella, L. M. Machesky, D. A. Kaiser, T. D. Pollard, Structural requirements and thermodynamics of the interaction of proline peptides with profilin. *Biochemistry* **35**, 16535-16543 (1996).
20. J. T. Nguyen, C. W. Turck, F. E. Cohen, R. N. Zuckermann, W. A. Lim, Exploiting the basis of proline recognition by SH3 and WW domains: Design of N-substituted inhibitors. *Science* **282**, 2088-2092 (1998).
21. W. J. Wedemeyer, E. Welker, H. A. Scheraga, Proline *cis-trans* isomerization in protein folding. *Biochemistry* **41**, 14637-14644 (2002).
22. L. Buday, L. Wunderlich, P. Tamas, The Nck family of adapter proteins: Regulators of actin cytoskeleton. *Cell Signal.* **14**, 723-731 (2002).
23. I. Moarefi, M. LaFevre-Bernt, F. Sicheri, M. Huse, C. H. Lee, J. Kuriyan, W. T. Miller, Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. *Nature* **385**, 650-653 (1997).
24. D. Karathanassis, R. V. Stahelin, J. Bravo, O. Perisic, C. M. Pacold, W. Cho, R. L. Williams, 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 (2002).
25. F. Kuribayashi, H. Nunoi, K. Wakamatsu, S. Tsunawaki, K. Sato, T. Ito, H. Sumimoto, The adaptor protein p40(phox) as a positive regulator of the superoxide-producing phagocyte oxidase. *EMBO J.* **21**, 6312-6120 (2002).
26. H. Hiroaki, T. Ago, T. Ito, H. Sumimoto, D. Kohda, Solution structure of the PX domain, a target of the SH3 domain. *Nat. Struct. Biol.* **8**, 526-530 (2001).
27. R. J. Rickles, M. C. Botfield, Z. Weng, J. A. Taylor, O. M. Green, J. S. Brugge, M. J. Zoller, Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. *EMBO J.* **13**, 5598-5604 (1994).
28. S. Feng, C. Kasahara, R. J. Rickles, S. L. Schreiber, 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 (1995).
29. A. B. Sparks, J. E. Rider, N. G. Hoffman, D. M. Fowlkes, L. A. Quillam, B. K. Kay, 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 (1996).
30. G. Cesareni, S. Panni, G. Nardelli, L. Castagnoli, Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett.* **513**, 38-44 (2002).
31. S. Feng, J. K. Chen, H. Yu, J. A. Simon, S. L. Schreiber, Two binding orientations for peptides to the Src SH3 domain: Development of a general model for SH3-ligand interactions. *Science* **266**, 1241-1247 (1994).
32. W. A. Lim, F. M. Richards, Critical residues in an SH3 domain from Sem-5 suggest a mechanism for proline-rich peptide recognition. *Nat. Struct. Biol.* **1**, 221-225 (1994).
33. W. A. Lim, F. M. Richards, R. O. Fox, Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375-379 (1994).
34. H. Yu, J. K. Chen, S. Feng, D. C. Dalgarno, A. W. Brauer, S. L. Schreiber, Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-945 (1994).
35. A. M. Wollacott, J. R. Desjarlais, Virtual interaction profiles of proteins. *J. Mol. Biol.* **313**, 317-342 (2001).
36. B. Brannetti, A. Via, G. Cestra, G. Cesareni, M. Helmer-Citterich, SH3-SPOT: An algorithm to predict preferred ligands to different members of the SH3 gene family. *J. Mol. Biol.* **298**, 313-328 (2000).
37. A. Musacchio, M. Noble, R. Pauptit, R. Wierenga, M. Saraste, Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851-855 (1992).
38. R. Ren, B. J. Mayer, P. Cicchetti, D. Baltimore, Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157-1161 (1993).
39. H. Terasawa, D. Kohda, H. Hatanaka, S. Tsuchiya, K. Ogura, K. Nagata, S. Ishii, V. Mandiyan, A. Ullrich, J. Schlessinger, 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 (1994).
40. T. M. Kapoor, A. H. Andreotti, S. L. Schreiber, 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 (1998).
41. A. H. Tong, B. Drees, G. Nardelli, G. D. Bader, B. Brannetti, L. Castagnoli, M. Evangelista, S. Ferracuti, B. Nelson, S. Paoluzi, M. Quondam, A. Zucconi, C. W. Hogue, S. Fields, C. Boone, G. Cesareni, A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* **295**, 321-324 (2002).
42. A. M. Mongiovi, P. R. Romano, S. Panni, M. Mendoza, W. T. Wong, A. Musacchio, G. Cesareni, P. P. Di Fiore, A novel peptide-SH3 interaction. *EMBO J.* **18**, 5300-5309 (1999).
43. D. M. Berry, P. Nash, S. K. Liu, T. Pawson, C. J. McGlade, 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 (2002).
44. A. W. McGee, S. R. Dakoji, O. Olsen, D. S. Bredt, W. A. Lim, K. E. Prehoda, 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 (2001).
45. C. H. Lee, K. Saksela, U. A. Mirza, B. T. Chait, J. Kuriyan, Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* **85**, 931-942 (1996).
46. K. Saksela, G. Cheng, D. Baltimore, 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 (1995).
47. S. Yuzawa, M. Yokochi, H. Hatanaka, K. Ogura, M. Kataoka, K. Miura, V. Mandiyan, J. Schlessinger, F. Inagaki, Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. *J. Mol. Biol.* **306**, 527-537 (2001).
48. K. Kami, R. Takeya, H. Sumimoto, D. Kohda, Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. *EMBO J.* **21**, 4268-4276 (2002).
49. A. Douangamath, F. V. Filipp, A. T. Klein, P. Barnett, P. Zou, T. Voorn-Brouwer, M. C. Vega, O. M. Mayans, M. Sattler, B. Distel, M. Wilmanns, Topography for independent binding of alpha-helical and PPII-helical ligands to a peroxisomal SH3 domain. *Mol. Cell* **10**, 1007-1017 (2002).

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

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

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

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

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

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

56. M. T. Bedford, R. Reed, P. Leder, 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 (1998).

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

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

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

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

61. J. S. Lott, S. J. Coddington-Lawson, P. H. Teesdale-Spittle, F. J. McDonald, 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 (2002).

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

63. E. Caron, Regulation of Wiskott-Aldrich syndrome protein and related molecules. *Curr. Opin. Cell Biol.* **14**, 82-87 (2002).

64. M. Reinhard, T. Jarchau, U. Walter, Actin-based motility: Stop and go with Ena/VASP proteins. *Trends Biochem. Sci.* **26**, 243-249 (2001).

65. K. Niebuhr, F. Ebel, R. Frank, M. Reinhard, E. Domann, U. D. Carl, U. Walter, F. B. Gertler, J. Wehland, T. Chakraborty, A novel proline-rich motif present in ActA of Listeria monocytogenes and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *EMBO J.* **16**, 5433-5444 (1997).

66. D. A. Portnoy, V. Auerbuch, I. J. Glomski, The cell biology of *Listeria monocytogenes* infection: The intersection of bacterial pathogenesis and cell-mediated immunity. *J. Cell Biol.* **158**, 409-414 (2002).

67. D. Pantaloni, C. Le Clainche, M. F. Carlier, Mechanism of actin-based motility. *Science* **292**, 1502-1506 (2001).

68. P. R. Brakeman, A. A. Lanahan, R. O'Brien, K. Roche, C. A. Barnes, R. L. Huganir, P. F. Worley, Homer: A protein that selectively binds metabotropic glutamate receptors. *Nature* **386**, 284-288 (1997).

69. A. Kato, F. Ozawa, Y. Saitoh, K. Hirai, K. Inokuchi, vesl, a gene encoding VASP/Ena family related protein, is upregulated during seizure, long-term potentiation and synaptogenesis. *FEBS Lett.* **412**, 183-189 (1997).

70. L. J. Ball, R. Kuhne, B. Hoffmann, A. Hafner, P. Schmieder, R. Volkmer-Engert, M. Hof, M. Wahl, J. Schneider-Mergener, U. Walter, H. Oschkinat, T. Jarchau, Dual epitope recognition by the VASP EVH1 domain modulates polyproline ligand specificity and binding affinity. *EMBO J.* **19**, 4903-4914 (2000).

71. M. Zettl, M. Way, The WH1 and EVH1 domains of WASP and Ena/VASP family members bind distinct sequence motifs. *Curr. Biol.* **12**, 1617-1622 (2002).

72. B. F. Volkman, K. E. Prehoda, J. A. Scott, F. C. Peterson, W. A. Lim, Structure of the N-WASP EVH1 domain-WIP complex: Insight into the molecular basis of Wiskott-Aldrich Syndrome. *Cell* **111**, 565-576 (2002).

73. I. R. Vetter, C. Nowak, T. Nishimoto, J. Kuhlmann, A. Wittinghofer, Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: Implications for nuclear transport. *Nature* **398**, 39-46 (1999).

74. R. P. Bhattacharyya, W. A. Lim, unpublished results.

75. K. E. Prehoda, D. J. Lee, W. A. Lim, Structure of the enabled/VASP homology 1 domain-peptide complex: A key component in the spatial control of actin assembly. *Cell* **97**, 471-480 (1999).

76. U. D. Carl, M. Pollmann, E. Orr, F. B. Gertler, T. Chakraborty, J. Wehland, Aromatic and basic residues within the EVH1 domain of VASP specify its interaction with proline-rich ligands. *Curr. Biol.* **9**, 715-718 (1999).

77. K. Nishizawa, C. Freund, J. Li, G. Wagner, E. L. Reinherz, Identification of a proline-binding motif regulating CD2-triggered T lymphocyte activation. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14897-14902 (1998).

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

79. T. Finkel, J. A. Theriot, K. R. Dise, G. F. Tomaselli, P. J. Goldschmidt-Clermont, Dynamic actin structures stabilized by profilin. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1510-1514 (1994).

80. D. Pantaloni, M. F. Carlier, How profilin promotes actin filament assembly in the presence of thymosin beta 4. *Cell* **75**, 1007-1014 (1993).

81. M. Reinhard, K. Giehl, K. Abel, C. Haffner, T. Jarchau, V. Hoppe, B. M. Jockusch, U. Walter, The proline-rich focal adhesion and microfilament protein VASP is a ligand for profilins. *EMBO J.* **14**, 1583-1589 (1995).

82. N. M. Mahoney, P. A. Janmey, S. C. Almo, Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. *Nat. Struct. Biol.* **4**, 953-960 (1997).

83. C. Bjorkegren, M. Rozycki, C. E. Schutt, U. Lindberg, R. Karlsson, Mutagenesis of human profilin locates its poly(L-proline)-binding site to a hydrophobic patch of aromatic amino acids. *FEBS Lett.* **333**, 123-126 (1993).

84. O. Pornillos, S. L. Alam, D. R. Davis, W. I. Sundquist, Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nat. Struct. Biol.* **9**, 812-817 (2002).

85. O. Pornillos, S. L. Alam, R. L. Rich, D. G. Myszka, D. R. Davis, W. I. Sundquist, Structure and functional interactions of the Tsg101 UEV domain. *EMBO J.* **21**, 2397-2406 (2002).

86. J. E. Garrus, U. K. von Schwedler, O. W. Pornillos, S. G. Morham, K. H. Zavitz, H. E. Wang, D. A. Wettstein, K. M. Stray, M. Cote, R. L. Rich, D. G. Myszka, W. I. Sundquist, Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**, 55-65 (2001).

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

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

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

90. See Pfam at http://pfam.wustl.edu/.

91. J. C. Whisstock, A. M. Lesk, SH3 domains in prokaryotes. *Trends Biochem. Sci.* **24**, 132-133 (1999).