REVIEW ARTICLE

Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction

Shawn S.-C. Li¹

Department of Biochemistry, Schulich School of Medicine, University of Western Ontario, London, Ontario, Canada N6A 5C1

Protein–protein interactions occurring via the recognition of short peptide sequences by modular interaction domains play a central role in the assembly of signalling protein complexes and larger protein networks that regulate cellular behaviour. In addition to spatial and temporal factors, the specificity of signal transduction is intimately associated with the specificity of many co-operative, pairwise binding events upon which various pathways are built. Although protein interaction domains are usually identified via the recognition code, the consensus sequence motif, to which they selectively bind, they are highly versatile and play diverse roles in the cell. For example, a given interaction domain can bind to multiple sequences that exhibit no apparent identity, and, on the other hand, domains of the same class or different classes may favour a given consensus motif. This promiscuity in ligand selection is typified by the SH3 (Src homology 3) domain and several other interaction modules that commonly recognize

proline-rich sequences. Furthermore, interaction domains are highly adaptable, a property that is essential for the evolution of novel pathways and modulation of signalling dynamics. The ability of certain interaction domains to perform multiple tasks, however, poses a challenge for the cell to control signalling specificity when cross-talk between pathways is undesired. Extensive structural and biochemical analysis of many interaction domains in recent years has started to shed light on the molecular basis underlying specific compared with diverse binding events that are mediated by interaction domains and the role affinity plays in affecting domain specificity and regulating cellular signal transduction.

Key words: interaction domain, proline-recognition domain, proline-rich motif, signal transduction, Src homology 3 domain (SH3 domain), WW domain.

INTRODUCTION

Modular interaction domains are conserved regions in proteins that are specialized in mediating interactions of proteins with one another or with other biomolecules such as lipids and nucleic acids [1]. A subgroup of interaction domains, namely protein or peptide interaction domains on which this review is focused, are devoted to promoting associations between proteins in the cell by binding modified amino acids and/or defined sequence motifs. These interaction domains are generally 30–200 amino acids in size and share certain characteristics. First, members of a given class of domain are related in both sequence and structure, and the three-dimensional fold of a domain is usually retained when isolated from the native protein. This latter property renders interaction domains particularly amenable to biochemical and biophysical manipulations. Secondly, protein interaction domains are often identified through consensus sequence motifs, or the so-called protein-recognition codes, to which they selectively bind [2]. In most cases, domain-mediated protein–protein interactions can be reconstructed *in vitro* using isolated domains and synthetic peptides containing these motifs [3]. Thirdly, interaction domains of the same or different classes are often linked in tandem in a single polypeptide chain. The combinatorial and repetitive use of interaction domains is the main driving force behind the

sophisticated mammalian signalling networks [1,4]. Fourthly, these domains are widespread in eukaryotic cells. There has been a drastic expansion in the recognition of interaction domains in metazoans compared with yeast, in accordance with the more complex signal transduction systems that accompany multicellularity [5–7].

Protein interaction domains participate in and regulate almost all essential cellular functions, including cell growth, differentiation, motility, polarity and apoptosis [1,4,8,9]. Establishing the principles that govern ligand recognition by these domains is critical to the understanding of their diverse functions in the cell. Much of our current knowledge on the specificity of protein interaction domains has been gleaned through the use of purified domains to screen peptide libraries that have been synthesized chemically or prepared using phage display [10–13]. The groups of Songyang and Cantley have pioneered the use of the oriented peptide library approach for the determination of binding specificity for a group of SH2 domains [13,14]; this method was subsequently extended to identify consensus binding motifs for other protein interaction domains and to map the preferred substrates for protein kinases and phosphatases [15,16]. More recently, peptide libraries displayed in an array format on a solid support have been used for charting domain specificity [17]. Collectively, these studies have greatly expanded our knowledge

Abbreviations used: CAP, cytoskeleton-associated protein; CD2BP2, CD2-binding protein 2; CDR, complementarity-determining region; Csk, C-terminal Src kinase; EVH1, Enabled/VASP (vasodilator-stimulated protein) homology; Fyb, Fyn-binding protein; HOG, high-osmolarity glycerol; MAPK, mitogen-activated protein kinase; PB1, phox (phagocyte oxidase) and Bem1; PEP, proline-enriched phosphatase; PH, pleckstrin homology; phox, phagocyte oxidase; PPII, polyproline type II; PRD, proline-recognition domain; PTB, phosphotyrosine-binding; PX, phox homology; SAP, signalling-lymphocytic-activation-molecule-associated protein; SH, Src homology; SH3-C, C-terminal SH3 domain; SLAM, signalling lymphocytic activation molecule; SLAP130, SH2-domain-containing-leucocyte-protein-of-76-kDa-associated protein; SLP-76, SH2-domain-containing leucocyte protein of 76 kDa; STAM, signal transducing adaptor molecule; UEV, ubiquitin E2 variant.

¹ email sli@uwo.ca

on the ligand preferences for a number of interaction domains. For instance, the SH2 (Src homology 2) domain, a prototypical interaction module that is known to bind phosphotyrosine-containing sequences [4], has been classified into four main groups based on distinct selections of three to six amino acids C-terminal to the phosphotyrosine residue [13,14]. By comparison, SH3 domains generally favour peptides that bear a PxxP core motif (where x represents any amino acid). As for an SH2 domain, selectivity of a given SH3 domain is defined by residues that flank this core motif [10]. Consensus motifs recognized by interaction domains have been used to predict potential physiological partners for proteins that contain these domains [18].

Despite the delineation of consensus motifs, mounting evidence suggests that members of a given domain class can recognize distinct peptides that are not apparently related and possess multiple modes of binding. Complexity may also arise from recognition of a consensus sequence or structural scaffold by different domain classes. Numerous structural and biochemical studies have been devoted to uncovering the molecular basis of domain specificity and diversity since the initial identification of the SH2 domain nearly 20 years ago [19]. As it is not practical to survey the entire literature on the topic, this review will revisit some recent studies on the versatility and novel binding properties for a group of interaction domains that commonly recognize proline. Emphasis will be given to the SH3 domain, a group of small interaction modules (~60 amino acids) that are ubiquitous in eukaryotes.

PROLINE-RICH SEQUENCES AND PRDs (PROLINE-RECOGNITION DOMAINS)

Proline-rich sequences are widely distributed in distinct proteomes from prokaryotes to eukaryotes [7,20]. For example, *Drosophila* is estimated to harbour 579 proline-rich regions, making them the most abundant sequence pattern in its proteome [7]. Together with their binding proteins, proline-rich sequences play an indispensable role in mediating a multitude of protein-protein interactions that are essential for a host of cellular processes [20,21]. Why are proline-rich motifs favoured in a cell? The answer to this intriguing question appears to lie within proline itself. First, for a peptide sequence to function in a binary interaction, it has to been exposed to the solvent and be accessible to the binding partner. Of the 20 naturally occurring amino acids, proline may be best suited for such a role. It is a well-known breaker of regular secondary structures such as α -helices and β -sheets that are essential for protein folding and topology. Consequently, proline-containing sequences are often found on the surface of a protein, as opposed to being buried within the core [22]. Secondly, the closure of the side chain of proline in a five-member ring restricts one of its dihedral angles, Φ , at approx. -60° . This severely restrains the types of conformation that proline and proline-rich sequences can adopt. The most common structure formed by two or more proline residues in a row is PPII (polyproline type II), a left-handed helix with three residues per turn. This structure is more relaxed than an ideal α -helix that has a pitch of 3.6 residues. Thirdly, the PPII conformation can arise automatically from a stretch of proline residues of sufficient length [23]. It is believed that restraints in the side chain of proline and, in some cases, the pre-formed structure would significantly reduce the entropic cost associated with binding of a proline-rich sequence. Fourthly, since both the side chains and the backbone carbonyls in a PPII structure are projected outwards from the axis of the helix, they are poised to interact with another molecule. Moreover, the lack of an amide proton in proline to participate in intramolecular hydrogen bonding frees its

carbonyl group for intermolecular interactions [24]. Finally, the PPII structure is stable and resistant, to a large extent, to amino acid substitutions. Therefore various combinations of non-proline and proline residues can be incorporated into a peptide sequence without compromising the integrity of the PPII structural frame (see below). This unique property of the PPII helix might have played an important part in the evolution of modular domains that bind to proline-rich sequences.

To make use of the abundant proline-rich sequences found in a cell, an array of PRDs have been developed in metazoans. Comparative genome analysis suggest that proline-rich sequences and PRDs may have co-evolved in Nature [7]. As for proline-rich sequences, PRDs, as a superfamily, are the most abundant protein interaction modules found in metazoan proteomes [21]. Modular domains currently known to recognize proline-rich sequences include the SH3 domain [25], the WW domain (named after two highly conserved tryptophan residues) [26,27], the EVH1 [Enabled/VASP (vasodilator-stimulated protein) homology] domain [28–30], the GYF domain (named after the presence of a Gly-Tyr-Phe triad) [31], profilin [32,33], UEV (ubiquitin E2 variant) [34] and the CAP (cytoskeleton-associated protein)-Gly domain [35]. These domains range in size from 30–35 residues for the WW domain to approx. 150 residues for UEV. SH3 and WW are the most abundant PRDs in vertebrates, with an estimated 409 and 125 copies respectively in the human proteome alone [21]. They have been the topic of numerous studies and have been reviewed extensively in the literature (e.g. [24–28]). Other PRDs are less abundant, reflecting, perhaps, their more dedicated functions. The EVH1 domain is found in a number of proteins that are implicated in actin cytoskeletal remodelling and post-synaptic signalling [22,28]. The GYF domain was originally identified in CD2BP2, a protein that binds to and regulates the function of the T-cell adhesion molecule CD2 [31]. Profilin promotes exchange of actin-bound ADP for ATP, thereby facilitating the formation of actin filaments [22]. The UEV domain was identified in Tsg101, a human protein that is recruited by structural proteins of HIV and Ebola to facilitate virus budding [34]. The CAP-Gly domain, found in a group of cytoskeleton-associated proteins, was recently added to the arsenal of PRDs [35]. CYLD, originally identified as a human familial cylindromatosis tumour suppressor, contains three CAP-Gly domains, of which the third was recently shown to bind weakly to a proline-rich sequence from NEMO (nuclear factor- κ B essential modifier)/IKK γ (inhibitory κ B kinase γ). It is likely that the CAP-Gly domain represents a variant of the SH3 domain, since the two domains share an essentially identical fold. Nevertheless, the binding site on the CAP-Gly domain lies outside the conserved ligand-binding surface for an SH3 domain [35].

Different PRD families prefer different consensus motifs or canonical sequences (Table 1) [2]. A prominent feature of most binding motifs identified for PRDs to date is the presence of one or more proline residue that form the ligand core. While these core proline residues are required for the formation of the ligand structure (i.e. the PPII helix) and play important roles in binding, selectivity of a ligand for a PRD is believed to be conferred by residues that flank the core. Some PRDs, such as profilin and the GYF domain, are dedicated to bind to a defined sequence pattern. Others, such as the SH3, WW and EVH1 domains, are far less selective. Based on ligand propensities determined using NMR and peptide arrays, Otte and co-workers [36,37] categorized 42 WW domains into six distinct groups (Table 1). This division, however, is not absolute. A number of WW domains, such as those found in HYBP, WAC (WW-domain-containing adaptor with coiled-coil) and Fe65, are capable of binding in a rather indiscriminate fashion to multiple types of ligands [26]. Along similar lines, members of the SH3 domain family bind to a

Table 1 Specificity and affinity of PRDs

For target sequences, upper-case letters indicate strong selection, lower-case letters denote moderate selection. Φ , hydrophobic residue; *po*, phosphate; x, any amino acid. K_d values listed represent consensus values from the literature, and are not meant to be inclusive. NA, data not available.

Domain	Target sequence or motif	Affinity (K_d , μ M)	Ligand structure	Reference(s)
SH3	(R/K)xxPxxP (class I)	1–200	PPII	[10,11]
	PxxPx(R/K) (class II)	1–200	PPII	[10,11]
	RxxK (class III)	0.1–30	3^{10} helix	[66–72]
	RKxxYxxY	20–60	NA	[59]
	PxxDY	NA	NA	[58]
	(R/K/G)XXPPGX(R/K)	10–200	PPII	[43]
	R/K-rich	10–100	N/A	[74,76]
	WQTDfEKLKE	NA	α -Helix	[61]
	RPSADLILNRCSESTRKCLASA	2–13	α -Helix	[102]
	PWTDQFEKLEKEVAEN...	NA	α -Helix	[60]
	PX domain	50	PPII*	[87]
	SH2 domain	1–5	Tertiary	[91,93]
	LIM domain	3000	Tertiary	[99]
WW	PPx(Y/poY)	1–500	PPII	[27,36,50]
	PPLPp	1–500	PPII	[27,36,50]
	(p/ Φ)PPPP	5–100	PPII	[36,37,106]
	(p/ Φ)P(p/g)PPpR	>10	NA	[36,106]
	(p/ Φ)PPRgpPp	>10	NA	[36,106]
	(poS/poT)P	15–150	PPII	[36,52]
EVH1	FPx Φ P (class I)	1–500	PPII	[28,29]
	PPxxF (class II)	1–500	PPII	[28,107]
	DLPPPEPYVQ... (25-mer)	NA	PPII	[30]
GYF	(R/K/G)XXPPGX(R/K)	10–200	PPII	[43,108]
Profilin	Poly(L-proline)	1–100?	PPII	[32,33]
UEV	PTAP	3	PPII	[34]
CAP-Gly	SQRRSPPEPPDF	>100	PPII?	[35]

large repertoire of peptide and protein ligands (Table 1). Since all members in a domain family adopt the same fold, questions arise as to how different peptides may be accommodated by a given domain class.

PPII: A UNIQUE STRUCTURAL SCAFFOLD RECOGNIZED BY ALL PRDs

The availability of structural information for numerous PRD–peptide complexes has made it possible to extract the general features of PRD–ligand interactions (Figure 1). In this regard, it is remarkable that the basis of ligand recognition established for the SH3 domain more than 10 years ago [38–41] has proved to be a general mechanism that is shared by all PRDs. For instance, the majority of SH3 domains characterized to date recognize the class I or II peptides that share a PxxP core element and assume the PPII conformation when bound to an SH3 domain [11]. Similarly, proline-rich peptides recognized by other PRD families all adopt the PPII helical structure in the bound states, despite large variations in peptide sequences and domain folds (Figure 1). In other words, the PPII helix is a common structural scaffold that is recognized by all PRDs. This observation is important since it implies that different PRD families have all evolved through exploitation of the PPII helix. It is unlikely, however, that all PRDs are descendants of a common ancestor because of the drastic differences in both the amino acid sequence and the three-dimensional fold between domain classes. This assertion is in agreement with the finding that, while some folds are dedicated to binding proline-rich sequences, others have been found in distinct types of interaction. For example, the EVH1 domain has essentially the same topology as that of a PTB (phosphotyrosine-

binding) or PH (pleckstrin homology) domain that recognizes phosphotyrosine-containing sequences or phosphatidylinositides respectively [29]. In fact, an EVH1 domain is more related in sequence to a PTB or PH domain than to any PRD [29].

It is interesting to speculate why Nature has targeted the PPII scaffold for the evolution of binding domains. To a certain degree, domain specificity, which plays a central part in signal transduction, is an oxymoron. On the one hand, a domain class that bound indiscriminately to many types of ligands would be of little value in regulating defined cellular processes. On the other hand, a domain class that was dedicated to binding one or a few selected sequences and hence was incapable of adapting to various binding events occurring in the cell would not find widespread use in a proteome, or be favoured by evolution. The PPII structure represents a good compromise between these two extreme scenarios. Since the PPII helix is only found in sequences that are rich in proline residues, and proline is unique in that it is the only N-substituted amino acid in Nature [42], a basal level of specificity is already built into a proline-rich sequence [24]. Nevertheless, the PPII scaffold, which can tolerate multiple amino acid insertions and/or substitutions, provides an excellent framework on which to explore the side-chain chemistry of non-proline residues in a peptide in order to maximize its affinity and selectivity for a PRD. Furthermore, because of the inherent 2-fold rotational pseudosymmetry of the PPII helix, a PRD can, in principle, engage a proline-rich sequence in one of the two opposing orientations [24,39]. This unique mode of bi-directional binding, adapted by most PRDs, enlarges further their ligand pools [24].

Besides similarities in ligand structure and binding modes, the binding surfaces for various PRDs, with the exception of the UEV domain, are all enriched in aromatic residues such as tryptophan, tyrosine and phenylalanine. Taking the SH3 domain as an example, its binding surface comprises three discrete patches: two hydrophobic grooves lined mainly by aromatic residues to accommodate the xP dipeptides (x represents, in most cases, a hydrophobic amino acid) in the PPII helix and a specificity pocket formed by residues primarily from the RT and n-Src loops (Figure 2) [37–44]. Similar xP grooves are found in the WW and GYF domains [24]. In the profilin–L-Pro¹⁰ complex, a slightly distorted PPII helix makes intimate contact with a patch of five highly conserved aromatic residues [32,33]. In the interesting case of the CD2BP2 GYF–peptide complex, the centre of the ligand-binding face is defined entirely by seven aromatic residues [43]. Aromatic residues such as tryptophan, tyrosine and phenylalanine are frequently found at protein–protein recognition sites [44]. These residues possess some unique properties that render them particularly favourable at the binding site of a PRD. The bulky side chain of an aromatic residue ensures large van der Waals contacts with the ligand. Their planar structure and near-parallel disposition at the binding site creates a shallow groove that is ideally suited for accommodating the xP unit that form the base of a PPII helix. The topology of the xP grooves in a PRD is therefore markedly different from the deep cavity used for binding the phosphotyrosine residue by an SH2 domain [45,46]. Apart from the EVH1 domain, which utilizes a concave surface to engage the apex, instead of the base, of a PPII helix [24,29,30], the binding grooves on most PRDs are shallow and have few distinguishing features [47].

In summary, despite drastic differences in three-dimensional structures for the six classes of PRDs depicted in Figure 1 and the diverse sequences that they recognize, PRDs employ conserved mechanisms in binding to their preferred ligands. This may be the main cause for cross-reactivity within and between classes of PRDs.

Figure 1 Structures of six PRDs in complex with their cognate ligands

All six PRDs recognize a common structural scaffold: the PPII helix. A PPII helix formed by a poly-L-Pro⁷ peptide is depicted in the centre to show the characteristics of the structure. Domains are portrayed in ribbon representations, while the bound peptides are shown in heavy atom traces. Structures were generated using MOLMOL [109]. The Protein Data Bank (PDB) accession codes for the structures are: 1QWE (SH3), 1EVH (EVH1), 1K9R (WW), 1L2Z (GYF), 1CJF (Profilin) and 1M4Q (UEV).

PROMISCUITY AND VERSATILITY OF PRDs

Large-scale mapping of protein–protein interactions mediated by the SH3 [48,49] and WW domains [50] have given us a glimpse of the complex nature of domain–ligand interactions. One important finding from these proteomic studies is that a given SH3 or WW domain could potentially interact with a few to several dozens of different peptide ligands [48–50]. Although it is generally believed that only a fraction of the interactions mapped in these types of *in vitro* studies would actually take place *in vivo*, the highly promiscuous binding patterns displayed by a number of SH3 and WW domains imply that these domains possess intrinsic ability to recognize a diverse set of ligands.

While the detailed molecular mechanisms for the observed promiscuity may differ from one class of PRD to another, the conserved binding mode shared by all PRDs may have played a central part in bestowing broad specificity to a given PRD. The shallow binding surface of a PRD, although ideal for accommodating the PPII structural frame, lacks intricate features that allow it to distinguish subtle differences between two proline-rich sequences [47]. This is contrary to the combining site of an antibody where the CDRs (complementarity-determining regions) from the light and heavy chains come together to form a binding surface that is rich in pockets and clefts [51]. The variability of the CDRs also dictates that the topology of the combining surface varies from one antibody to another. In comparison, the ligand-binding surfaces of different members within a given PRD family and, in some cases, between two PRD classes are remarkably similar, resulting in cross-binding within and between PRD classes. In addition, the interface of a typical

PRD–peptide complex is significantly smaller than that of a typical antibody–antigen complex. Therefore a PRD-mediated interaction is usually much weaker than one between an antigen and its antibody. Since the difference in K_d values between various PRD–peptide interactions spans only two orders of magnitude, most PRD-mediated interactions are of low specificity [47]. In the case of the WW domain, this low affinity was shown to originate from sub-optimal interactions between the ligand and the domain [52,53]. Structural analysis demonstrated that multiple contact sites were involved in the complex formation, of which no single site played an essential role. Accordingly, the affinity and specificity of a WW domain arises from the collective contribution of sub-optimal interactions. This mode of ligand recognition is also seen in other PRDs such as the SH3 and EVH1 domains. Thus a proline-rich sequence that is slightly different from the canonical ligand of a PRD may still be capable of binding to the PRD.

Accumulating evidence suggests that PRDs are remarkably versatile. The specificity pocket of a PRD is usually composed of residues from loops that connect regular secondary structures (Figure 2). Since loops are generally variable in sequence and flexible in structure, they play important roles in modulating the specificity of a domain [38]. Moreover, regions outside the conserved binding surface of a PRD can participate in ligand binding, which often leads to increased affinity and/or novel specificity. For instance, the Tsg101 UEV domain, which functions in both HIV budding and vacuolar protein sorting, is capable of binding to both a Pro-Thr/Ser-Ala-Pro peptide from the HIV-1 p6 proteins and to ubiquitin. Interestingly, these two unrelated ligands engage separate surfaces of the UEV domain and can

Figure 2 Structure of the Crk SH3-N domain in complex with a high-affinity peptide from C3G [62]

(A) The structure shown is based on PDB accession code 1CKB. The SH3 domain is depicted in ribbons with secondary-structural elements shown in different colours and labelled. The bound peptide, PPPALPPKKR, is shown in blue with side chains. Interface residues on the SH3 domain are shown in pink, for aromatic residues, and in light blue, for non-aromatic residues. The locations of the xP grooves and specificity pocket on the SH3 domain are identified by broken arrows. (B) A schematic representation of the same structure to highlight the characteristics of the ligand-binding surface on the SH3 domain such as the enrichment of aromatic residues. The same colouring scheme is used in (A) and (B) for purposes of comparison.

thus bind independently of each other [54]. PRDs are also highly adaptable. Phage library screens demonstrated that substitution of two or three residues within an SH3 domain was sufficient to alter its specificity [55]. In the case of the Fyn SH3 domain, mutation of a single tyrosine residue (Tyr¹²³) to isoleucine renders it incompetent in binding the class I ligands, but still capable of recognizing class II ligands [56]. This adaptability might have played an important part in the development of novel signalling properties and/or in the expansion of ligand space for PRDs. Tong et al. [48] mapped the ligand space of 20 yeast SH3 domains using a strategy that combined peptide phage display with two-hybrid screening. Although the majority of SH3 domains examined in this study selected for the class I and/or class II consensus, unusual motifs were identified. Of note, the Bem-1 SH3 domain bound to a PPxVxPY consensus, whereas the Fus1 SH3 domain preferred the RxxR(s/t)(s/t)SI motif. It is possible that the human proteome, which contains hundreds of SH3 domains, may have evolved more elaborate mechanisms for the recognition of a broader spectrum of peptide and protein ligands.

RECOGNITION BEYOND THE PxxP MOTIF BY SH3 DOMAINS: THE IMPORTANCE OF BEING POSITIVE

Although SH3 domains have been the focus of numerous studies since their discovery more than 15 years ago [57], our understanding of the mechanisms of ligand recognition and signalling by this important family of PRD is far from complete. A growing body of literature suggests that SH3 may possess the most diverse specificity among interaction domains. In addition to recognizing the class I and II peptides that contain a PxxP core element, a number of SH3 domains have been shown to bind peptide sequences that lack such a motif. For example, the SH3 domains from the tyrosine kinase substrate Eps8 and related proteins bind selectively to the PxxDY motif [58]. The SH3 domains of Fyn and the Fyn-binding protein, Fyb or SLAP130 [SLP-76 (SH2-domain-containing leucocyte protein of 76 kDa)-associated protein], engages a site in SKAP55 (Src kinase associated protein of 55 kDa) bearing a consensus sequence RKxxYxxY that is devoid of proline residues [59]. The SH3 domain of the yeast protein Pex13p is capable of binding, using distinct surfaces, two proteins of the peroxisomal import pathway that are related in neither sequence nor structure. Whereas it interacts with Pex14p through a PPII helix in a conventional manner, it binds independently to an α -helix formed by a novel sequence motif WxxxFxxLE present in Pex5p through a site that is removed from the conventional surface [60,61]. Chemical shift perturbations monitored by NMR spectroscopy revealed that the Pex13p SH3 used separate sites located at opposite faces of the domain for binding Pex14p and Pex5p respectively. That the Pex13p SH3 domain is capable of interacting simultaneously with two ligands in a non-competitive manner [60] suggests that the functions of SH3 domains are not limited to mediating stoichiometric binary protein-protein interactions [61].

Positively charged residues such as arginine and lysine have been known to play an important part in binding an SH3 domain, not only by providing additional binding energy through electrostatic interactions with residues in the specificity pocket, but also by orienting the ligand with respect to the binding groove on the SH3 domain (Figure 2) [38,41,62]. The importance of these basic residues is underscored in ligands identified for a subgroup of SH3 domains. The SH3 domain of amphiphysin binds to dynamin through a sequence that is enriched in both proline and basic residues. Combinatorial peptide library screening indicated that this domain recognized a novel consensus PxR PxR(H)R(H) [63]. The SH3 domains of STAM (signal transducing adaptor molecule), EAST (epidermal growth factor receptor-associated protein with SH3 and tyrosine-based activation motif domains) and Hbp (Hrs-binding protein), a family of proteins that are involved in cytokine-mediated signalling and receptor-mediated endocytosis and exocytosis [64], select for a consensus motif Px(V/I)(D/N)RxxKP contained in AMSH (associated molecule with the SH3 domain of STAM), an endosome-associated ubiquitin isopeptidase and UBPY (ubiquitin isopeptidase Y), a deubiquitinating enzyme [65]. Similar sequences bearing a conserved (R/K)xx(K/R) core motif have been found in several other proteins that include the scaffolding proteins Gab1 and Gab2, the B-cell signalling protein BLNK (B-cell linker protein), SLAP130/Fyb, the haematopoietic progenitor kinase HPK1 [66] and the T-cell docking/adaptor protein SLP-76 [67]. *In vitro* and *in vivo* experiments demonstrated that the (R/K)xx(K/R) motif in these proteins mediate their respective interactions with the C-terminal SH3 (SH3-C) domains of Grb2 and/or Gads [67–69]. Since the (R/K)xx(K/R) motif is found in a large number of proteins, it is perhaps appropriate to designate it as the class III consensus recognized by the SH3 domain.

Figure 3 Novel mode of peptide recognition for the Gads SH3-C domain

The structure of the Gads SH3 domain (left) is depicted in surface representation to show the three distinct binding pockets for the SLP-76 peptide (in green). Areas of positive and negative charges are shown in blue and red respectively. Residue Glu²⁷⁵ in the RT loop of the SH3 domain, which encloses the second pocket, is labelled in white. For comparison, the structure of the c-Src SH3 domain in complex with a PxxP-containing peptide [38] is shown on the right. Note the differences in characteristics and size for the second and third pockets on the two SH3 domains.

What is the structural basis for the specific recognition of an (R/K)xx(K/R) motif by an SH3 domain? Binding of Gads to SLP-76 serves an important physiological function by coupling the latter to membrane-associated LAT (linker for activated T-cells) upon T-cell receptor activation [70]. The unusually high affinity ($K_d = 0.24 \mu\text{M}$) of the Gads SH3-C domain for an SLP-76 peptide, APSIDRSTKPA, allowed for the determination of the solution structure of the complex by NMR [68,71]. In contrast with proline-rich ligands that form the PPII helix, the SLP-76 peptide adopts an extended conformation at the N-terminal region, followed by one turn of a right-handed 3^{10} helix at the RSTK (Arg-Ser-Thr-Lys) motif (Figure 3). This unique structure of the RSTK motif predisposes the side chains of the arginine and lysine residues to engage a pair of glutamate residues in the RT loop of the SH3 domain [68]. Although the peptide occupies a conserved surface on the Gads SH3-C domain, the xP grooves are significantly different from those found in a conventional SH3–ligand interaction. Compared with the Src SH3 domain, the second xP groove in Gads SH3-C is enclosed on one side by the side chain of a glutamate residue (Glu²⁷⁵) due to a conformational change in the RT loop (Figure 3). Consequently, this groove is narrower and deeper than the corresponding xP groove on the Src SH3 domain and selects for an aliphatic isoleucine residue instead of the usual xP dipeptide unit. This selectivity of the Gads SH3-C domain is, however, not absolute. A peptide containing both a PxxP and an RxxK motif derived from HPK1 was recently shown to bind the Gads SH3-C domain in a manner that combines the classical mode of a PPII helix binding with the atypical recognition of a 3^{10} helix. Nonetheless, the interaction of the Gads SH3-C domain with the HPK1 peptide is approx. one order of magnitude weaker than with the SLP-76 peptide [66], suggesting that the PPII scaffold is not ideally suited for the ligand-binding surface of the Gads SH3-C domain. In another instance, the STAM SH3 domain bound with different affinities to both an RxxK-containing peptide PMVNRENK ($K_d = 27 \mu\text{M}$) and a class I ligand ($K_d = 74 \mu\text{M}$) [72].

What is the importance of basic residues relative to proline residues in SH3 binding? Owing to the weak to moderate affinity between a class I or II ligand and an SH3 domain, mutating either

a proline residue in the PxxP core or the flanking arginine/lysine residue essentially abolishes binding. In the case of the Gads SH3-C domain, however, positively charged residues appear to play a dominant role, since replacing either the arginine or lysine residue with alanine in the RSTK motif eliminated binding, whereas substituting alanine for a proline residue outside this motif had a less detrimental effect [68]. Can positively charged residues alone confer sufficient affinity for an SH3 domain under physiological circumstances? Recent studies on several SH3 domains suggest that this scenario can indeed take place in biologically relevant interactions. Bin1/M-amphiphysin-II, a protein that is involved in transverse tubule biogenesis in striated muscle, contains an N-terminal BAR (Bin/amphiphysin/Rvs) domain, a basic region called Exon 10 and SH3-C domain [73]. Exon 10 bears a basic motif (K/R)xxxxKx(K/R)(K/R) that is required for binding to phosphoinositides, an interaction that serves to localize Bin1 to the membrane. Interestingly, when Bin1 is not associated with the membrane, Exon 10 is engaged by its own SH3 domain via an intramolecular interaction [74]. This, mediated exclusively by clusters of arginine and lysine residues in the basic motif, blocks binding of the SH3 domain to its physiological target dynamin [74]. Thus the Bin1 SH3 domain displays dual specificity for two distinct types of ligands: a basic motif in Exon 10 and a canonical PxxP motif from dynamin. Intriguingly, unlike the Pex13p SH3 domain–ligand interactions, these two unrelated motifs engage an overlapping surface in the Bin1 SH3 domain, and are therefore mutually exclusive. This creates a binary switch for Bin1 function so that binding of its SH3 domain to dynamin is permitted in the presence of phosphoinositides, but inhibited in the absence of the lipid [74,75].

The role of basic residues in promoting and/or mediating SH3 domain binding is not limited to the above examples. In a small-scale proteomic study using peptide arrays synthesized on nitrocellulose membranes, the potential binding sites in SLP-76 were mapped for a group of 15 SH3 domains [76]. The proline-rich region in SLP-76 encompasses 250 amino acids and is rich in both proline and basic amino acids. Numerous peptides were identified as potential SH3-binding sites from this study, of which a significant proportion lacked the PxxP motif. Instead,

Figure 4 Diverse roles of SH3 domains in the phox proteins

A model for the translocation of the regulatory phox complex to the catalytic core of the NADPH oxidase through co-operative interactions mediated by SH3 and PX domains, leading to the assembly and activation of the holoenzyme (not shown in whole in the diagram). The p40–p47–p67^{phox} regulatory complex resides in the cytosol in the absence of a stimulating signal. The complex is kept together through high-affinity interactions involving SH3 and PB1 domains, as shown. The tandem SH3 domains of p47^{phox} were shown to interact with its polybasic region through a novel mode of ligand recognition (see text). This unique interaction renders p47^{phox} translocation-incompetent. Following a stimulus, p47^{phox} is phosphorylated on multiple serine residues in the polybasic region. Phosphorylation releases the auto-inhibition in p47^{phox}, allowing it to associate with p22^{phox} through its SH3 domains. The PX domains in p47^{phox} and p40^{phox}, capable of binding phosphoinositides, are believed to play an important role in the translocation of the regulatory phox complex by interacting with the membrane. aa, amino acids.

these sequences are characterized by the presence of multiple basic residues. A number of SH3 domains, including those from PLC- γ 1 (phospholipase C- γ 1), Grb2 and Nck, exhibited weak to moderate binding (i.e. a K_d of 10–100 μ M) to a selection of peptides that were rich in arginine and/or lysine residues [76].

VERSATILITY OF THE SH3 DOMAIN: A LESSON FROM THE PHOX (PHAGOCYTE OXIDASE) PROTEINS

The versatile nature of SH3 domains in ligand recognition and in regulating the formation and dissolution of protein complexes is illustrated by a group of phox proteins that contain multiple SH3 domains and binding motifs (Figure 4). The superoxide-producing NADPH oxidase, which plays a critical role in host defence against microbial infection, is a membrane-bound enzyme composed of a catalytic core formed by p22^{phox} and gp91^{phox} and cytosolic regulators. The regulators include the small GTPase Rac and three phox proteins named p44^{phox}, p67^{phox} and p40^{phox} respectively [77]. The p40^{phox} protein contains an N-terminal PX (phox homology) domain, a central SH3 domain and a C-terminal PB1 (phox and Bem1) domain. p47^{phox} contains, from the N- to the C-terminus, a PX domain, two SH3 domains, a polybasic region and a proline-rich region. p67^{phox} is also a multidomain protein containing, among others, two SH3 domains that flank a PB1 domain (Figure 4). These three proteins form a ternary complex mediated by two intermolecular interactions: heterodimerization of the p40^{phox} and p67^{phox} PB1 domains and a high-affinity binding of the p67^{phox} SH3_B domain to a 32-residue proline-rich sequence from p47^{phox} [78]. Activation of the NADPH oxidase requires the translocation of the p40–p47–p67^{phox} complex from the cytoplasm to the membrane, a process mediated by binding of the p47^{phox} SH3 domains to a PxxP motif present at the C-terminal tail of p22^{phox} (Figure 4). In resting cells, however, this is prevented by an intramolecular interaction between the two SH3 domains and the polybasic region of p47^{phox} that keeps the protein

in an auto-inhibited conformation [77,79,80]. The X-ray structure of auto-inhibited p47^{phox} determined recently by Groemping et al. [80] revealed how this unique interaction took place. The two SH3 domains of p47^{phox} were shown to interact with its polybasic region through a novel mode of ligand recognition (see text). This unique interaction renders p47^{phox} translocation-incompetent. Following a stimulus, p47^{phox} is phosphorylated on multiple serine residues in the polybasic region. Phosphorylation releases the auto-inhibition in p47^{phox}, allowing it to associate with p22^{phox} through its SH3 domains. The PX domains in p47^{phox} and p40^{phox}, capable of binding phosphoinositides, are believed to play an important role in the translocation of the regulatory phox complex by interacting with the membrane. aa, amino acids.

The X-ray structure of auto-inhibited p47^{phox} determined recently by Groemping et al. [80] revealed how this unique interaction took place. The two SH3 domains of p47^{phox} were shown to interact with its polybasic region through a novel mode of ligand recognition (see text). This unique interaction renders p47^{phox} translocation-incompetent. Following a stimulus, p47^{phox} is phosphorylated on multiple serine residues in the polybasic region. Phosphorylation releases the auto-inhibition in p47^{phox}, allowing it to associate with p22^{phox} through its SH3 domains. The PX domains in p47^{phox} and p40^{phox}, capable of binding phosphoinositides, are believed to play an important role in the translocation of the regulatory phox complex by interacting with the membrane. aa, amino acids.

in an auto-inhibited conformation [77,79,80]. The X-ray structure of auto-inhibited p47^{phox} determined recently by Groemping et al. [80] revealed how this unique interaction took place. The two SH3 domains of p47^{phox} were shown to interact with its polybasic region through a novel mode of ligand recognition (see text). This unique interaction renders p47^{phox} translocation-incompetent. Following a stimulus, p47^{phox} is phosphorylated on multiple serine residues in the polybasic region. Phosphorylation releases the auto-inhibition in p47^{phox}, allowing it to associate with p22^{phox} through its SH3 domains. The PX domains in p47^{phox} and p40^{phox}, capable of binding phosphoinositides, are believed to play an important role in the translocation of the regulatory phox complex by interacting with the membrane. aa, amino acids.

in an auto-inhibited conformation [77,79,80]. The X-ray structure of auto-inhibited p47^{phox} determined recently by Groemping et al. [80] revealed how this unique interaction took place. The two SH3 domains of p47^{phox} were shown to interact with its polybasic region through a novel mode of ligand recognition (see text). This unique interaction renders p47^{phox} translocation-incompetent. Following a stimulus, p47^{phox} is phosphorylated on multiple serine residues in the polybasic region. Phosphorylation releases the auto-inhibition in p47^{phox}, allowing it to associate with p22^{phox} through its SH3 domains. The PX domains in p47^{phox} and p40^{phox}, capable of binding phosphoinositides, are believed to play an important role in the translocation of the regulatory phox complex by interacting with the membrane. aa, amino acids.

The auto-inhibited conformation of p47^{phox} is released upon phosphorylation of a group of serine residues in the polybasic region, presumably by protein kinase C, following a stimulating signal [81]. Activated p47^{phox} regains ability to bind the C-terminal tail of p22^{phox} through its SH3 domains. It is believed that the

phosphate group in phosphoserine disrupted the electrostatic interaction network between the positively charged residues in the basic region and the acidic residues in the tandem SH3 domains [80]. Intriguingly, progressive substitution of serine in the polybasic region by glutamate residues to mimic phosphorylation drastically increased the affinity of p47^{phox} for a proline-rich peptide derived from C-terminus of p22^{phox} [80]. The fact that phosphorylation coupled to cell activation is employed here to regulate SH3 domain–ligand interaction is fascinating, since it suggests that a phosphorylation signal can be relayed not only by canonical phosphate-binding modules such as SH2, PTB and FHA (forkhead-associated) [1], but also by interaction domains that normally bind to unmodified peptides. In this regard, it is worth mentioning that interactions mediated by WW domains, in particular those involving the PPPY and (pS/pT)P motifs, can also be modulated by tyrosine and serine/threonine phosphorylation [36].

It is not clear whether one or both SH3 domains of p47^{phox} are required for binding the p20^{phox} PxxP motif *in vivo*. Sumimoto et al. [82] showed that the N-terminal SH3 domain (SH3_A) of p47^{phox} alone bound with high affinity ($K_d = 0.34 \mu\text{M}$) to the C-terminal tail of p22^{phox}. Notwithstanding this observation, recent studies indicated that the tandem SH3 domains had a higher affinity for p22^{phox} [80], suggesting that the superSH3 domain may be the functional unit responsible for the association of p47^{phox} with p22^{phox}. Furthermore, recruitment of p47^{phox}, and hence the regulatory complex, to the membrane may be facilitated by the PX domains in p47^{phox} and p40^{phox}, both of which have been shown to bind phosphoinositides specifically [83–85]. A mutant of p47^{phox} with the PX domain deleted failed to migrate to the membrane upon cell stimulation [86]. Interestingly, the lipid-binding surface of the PX domain is masked by an unusual intramolecular interaction with the C-terminal SH3 domain (SH3_B) in the auto-inhibited form of p47^{phox} [86,87]. NMR spectroscopic analysis demonstrated that a proline-rich segment between α -helices 1 and 2 of the PX domain assumed a twisted PPII conformation and was recognized by the p47^{phox} SH3_B domain [87]. Taken together, these studies suggest that the p47^{phox} SH3 domains are capable of participating in a variety of interactions that involve either a classical PxxP motif, a novel basic sequence or a lipid-binding module. Although the mechanism of the NADPH oxidase activation is not yet completely understood [78], it is remarkable that these interactions take place in a highly controlled, orderly fashion in the cell in response to extracellular cues.

RECOGNITION VIA TERTIARY CONTACTS

Binding of the p47^{phox} SH3 domain to a PPII helix presented by a folded PX domain is reminiscent of an interaction between the Fyn SH3 domain and the HIV-1 Nef protein [88]. In the latter case, a PPII helix formed by a short PxxP-containing fragment derived from a highly conserved region of Nef makes contact with the Fyn SH3 domain in a manner akin to that seen between an SH3 domain and a class II peptide. Interestingly, a 12-residue peptide encompassing the Nef motif bound with a 300-fold reduced affinity to a related SH3 domain from Hck than did the intact Nef protein, suggesting that tertiary interactions may contribute to the Nef–SH3 interaction. This was indeed shown to be the case. In addition to binding mediated by the PPII helix, an isoleucine residue from the RT loop (named after its important arginine and threonine residues) of the Fyn SH3 domain inserts into a well-defined hydrophobic pocket formed by two α -helices of Nef [88]. This unique interaction not only provides a hydrophobic anchor for the Fyn SH3 domain on Nef, but also significantly increases the interface area between the two proteins [88]. This high-affinity

interaction may also be facilitated by the pre-formed PPII helix in the context of the folded Nef protein. Since the PPII structure moulded in a folded protein undergoes little structural changes before and after binding, this type of interaction is entropically more favourable than that between a domain and an isolated peptide.

In contrast with the examples detailed above, binding of an SH3 domain to another domain or protein does not necessarily involve specific motifs. This is illustrated in the regulation of the guanine nucleotide-exchange factor Vav by Grb2 through heterodimerization of its C-terminal SH3 domain with the N-terminal SH3 domain of Vav. Since the interaction is mediated by a complementary interface between the two SH3 domains rather than a continuous sequence motif [89], it can only be rationalized in terms of tertiary contacts. As shape and chemical complementarity underlies all specific molecular recognition, tertiary contacts, in a broad sense, may be a more appropriate gauge of ligand specificity for an interaction domain than is a simplified consensus motif.

Domain–ligand association mediated by tertiary contacts can play important roles in signal transduction, as has been shown for the tyrosine kinase Fyn and its substrates. In T-cells, Fyn is recruited to the immunoreceptor SLAM (signalling lymphocytic activation molecule) by SAP (SLAM-associated protein), a protein that is deleted or mutated in patients with the X-linked lymphoproliferative syndrome. Here, the SAP SH2 domain plays an adaptor's role by binding simultaneously to the Fyn SH3 domain and to Tyr²⁸¹ of SLAM in a fashion that is independent of its phosphorylation [90–93]. As shown in Figure 5(A), the formation of a ternary complex of Fyn–SAP–SLAM not only serves to recruit Fyn to SLAM, but also activates the kinase function of Fyn by releasing an inhibitory intramolecular interaction between the SH3 domain and the SH2 kinase linker of Fyn. This allows for the phosphorylation of multiple tyrosine residues within the cytoplasmic tail of SLAM and the propagation of an activating signal to downstream molecules (Figure 5A). The X-ray structure of a (Fyn SH3)–(SAP SH2)–(SLAM Tyr²⁸¹) peptide complex determined by Eck and colleagues [93] uncovered some novel features of domain–ligand interaction. In particular, the Fyn SH3 and SAP SH2 domains associated with each other through a surface–surface interaction that involved no conventional binding motifs for either the SH3 domain or the SH2 domain. Using residues primarily from the RT loop and the β B, β C and β D strands, the Fyn SH3 domain engaged a continuous surface of the SAP SH2 domain formed by residues that were far apart in the primary structure [93] (Figure 5B). Moreover, the SH3 domain-binding surface on the SAP SH2 domain was located on a face opposite to where the binding groove for the SLAM peptide resided. Therefore, by using separate surfaces to engage two different ligands, the single SH2 domain of SAP fulfils the function of a multidomain adaptor in a highly economical fashion.

Regulation of signalling events through the exploitation of distinct surfaces on an interaction domain as illustrated above is also seen in Crk [94]. Phosphorylation of residue Tyr²²¹ in Crk by the Abl (Abelson) kinase initiates an intramolecular interaction with its own SH2 domain. Interestingly, binding of Crk SH2 domain to the pTyr²²¹ site induces a conformational change in the SH2 domain that promotes its association with the Abl SH3 domain. NMR analysis demonstrated that the Abl SH3 domain recognized a proline-rich motif present in the β D– β E loop of the Crk SH2 domain. Since this loop is located on a separate face to that of the canonical pTyr²²¹ peptide-binding site, the Crk SH2 is capable of simultaneously binding the Abl SH3 domain and the pTyr²²¹ peptide in a fashion similar to the SAP SH2 domain-mediated interactions in SLAM signalling. It should be noted that the interaction between the Abl SH3 domain and the highly

Figure 5 SAP, an adaptor composed of a single SH2 domain

(A) Signal transduction through the SLAM receptor is regulated by SAP. Fyn is kept in an auto-inhibited state in the absence of SAP. Recruitment of Fyn to the SLAM receptor by SAP leads to activation of the kinase and subsequent phosphorylation of SLAM on multiple tyrosine residues. Phosphorylated tyrosine residues serve as docking sites for the SH2 domain of SHIP (SH2-containing inositol phosphatase). Binding of SHIP to SLAM initiates a signalling cascade that ultimately leads to modulation of IFN- γ production in T cells. **(B)** Structure of a ternary complex of (Fyn SH3)–(SAP SH2)–SLAM peptide. The Fyn SH3 domain, in green, binds to the SAP SH2 domain, in cyan, via surface–surface association. The SAP SH2 domain engages the SLAM peptide (shown in magenta) on the opposite surface using a groove conserved for SH2 domains. The structure shown is Protein Data Bank accession code 1M27.

mobile β D– β E loop in the Crk SH2 domain involved a much smaller interface area than did the Fyn SH3–SAP SH2 complex, consistent with the functional flexibility for the former interaction [94].

AFFINITY AND SPECIFICITY IN LIGAND RECOGNITION BY SH3 DOMAINS

Because interaction domains play a central role in the assembly of multiprotein complexes, their specificity and affinity for a particular target or targets often determine the efficiency and out-

come of cellular signalling [89]. Zarrinpar et al. [24] suggest that there are two mechanisms of interaction specificity: domain-mediated specificity that is embedded within individual domain–ligand pairs, and contextual specificity in which factors such as cellular context, co-operative events and subcellular localization play an important role in the promotion or facilitation of the unique interactions. The importance of cellular compartmentalization in dictating specificity of signal transduction is illustrated in the following example. In T-cells, the CD2-binding protein CD2BP2 engages, through its GYF domain, tandem proline-rich motifs that are present in the cytoplasmic tail of CD2. Interestingly, the

Fyn SH3 domain binds to the same motifs *in vitro* as does the CD2BP2 GYF domain [43]. These two potentially competing interactions, however, appear to be segregated in T-cells. While CD2BP2 associates with CD2 in the detergent-soluble membrane fraction, Fyn is detected only in lipid rafts to which CD2 is translocated upon its ectodomain clustering [43].

The importance of domain-mediated specificity was nicely demonstrated in a series of experiments performed on components of the yeast HOG (high-osmolarity glycerol) pathway. The osmosensor Sho1 is a membrane protein that contains an SH3 domain in its cytoplasmic region. Activation of the osmolarity-specific MAPK (mitogen-activated protein kinase) Hog1 is mediated by the scaffolding protein Pbs2, a MAPKK (MAPK kinase) that interacts with both Ste11 and Sho1 [95]. Binding of Pbs2 to Sho1 involves a proline-rich motif in the former and the SH3 domain of the latter. Intriguingly, a Pbs2 peptide encompassing the proline-rich motif was shown to bind the Sho1 SH3 domain with absolute selectivity so that it did not display significant affinity for any of the remaining SH3 domains in the yeast proteome *in vitro* or *in vivo* [96]. Because the yeast proteome contains only 28 SH3 domains, the specificity of the Pbs2 peptide was believed to be conferred by 'negative selection' in a manner whereby other SH3 domains were disfavoured for binding Pbs2 in comparison with the Sho1 SH3 domain. Interestingly, this absolute selectivity of the Pbs2 peptide for the Sho1 SH3 domain was lost when certain residues flanking the PxxP motif were mutated [96]. It remains to be seen whether this type of negative selection seen for the Sho1 SH3 domain is employed by the human proteome to evolve domain specificity. In the highly modular human proteome, where hundreds of SH3 domains coexist, other mechanisms of specificity enhancement, such as subcellular sequestration, scaffolding and co-operative interactions, may come into play [9,96].

Despite the highly specific interaction between the Sho1 SH3 domain and the Pbs2 proline-rich motif, this interaction could be functionally replaced *in vivo* by other modular interactions of sufficient affinity, with or without the involvement of an SH3 domain. Using a series of Sho1 SH3 domain mutants with decreased affinities for the Pbs2 peptide, Marles et al. [97] showed that a strong correlation existed between *in vitro* binding affinity of an SH3 domain mutant and the ability of the corresponding Sho1 mutant to confer osmoresistance in yeast. In a separate study, Park et al. [98] showed that the native interaction pair of Sho1 and Pbs2 could be functionally replaced by a PDZ-domain-mediated interaction, demonstrating the extremely modular nature of signalling pathways in yeast.

Typically, interactions mediated by PRDs have dissociation constants in the 1–100 μ M range. This type of weak interaction may be advantageous under certain circumstances. In cases where the formation and dissolution of signalling complexes occur rapidly in the cell, weak interactions, especially those with high on and off rates, would favour the dynamic flux of macromolecular assemblies. Weak interactions mediated by interaction domains have been shown to play important regulatory roles in certain processes. For example, the rapid turnover of focal adhesions during integrin signalling is thought to be triggered by a transient association between NCK2 and PINCH1. This focal adhesion complex is formed via an extremely weak interaction between the third SH3 domain of NCK2 and the fourth LIM domain of PINCH1 ($K_d \sim 3$ mM) [99]. NMR analysis demonstrated that this ultraweak interaction was maintained by an extremely small and polar interface between the two interacting domains. It is worth noting that the interface of the complex involves an area outside the conserved ligand-binding surface of the SH3 domain [99].

The presence of multiple weak binding sites in a protein may confer tight binding. This notion is appealing given that proline-

rich regions are widespread in a mammalian proteome and can span tens to hundreds of amino acids in a single polypeptide chain. It is common that multiple binding sites coexist for one or more PRDs. These sites may co-operate in binding to PRDs. For instance, although the CD2BP2 GYF domain binds with high affinity to the tandem proline-rich motifs in CD2, it binds to a peptide bearing a single copy of the motif with a K_d of approx. 200 μ M [43]. NMR analysis suggests that both the short and long CD2 peptides engage the same set of residues in the GYF domain. It is likely that the tandem proline-rich motifs act co-operatively to sequester the GYF domain to the vicinity of the peptide and thereby effectively increasing its local concentration.

Despite the advantage of weak interactions in certain cellular contexts, it is reasonable to assume that greater affinity would usually result in greater specificity for a binding event. Under this consideration, the lower affinity and high promiscuity displayed by some PRDs present a dilemma for specificity control in the cell. To solve this dilemma, SH3 domains have evolved a number of mechanisms to improve affinity and specificity. One such mechanism is to utilize non-PxxP motifs to enhance specificity. This is exemplified by the specific recognition of a RSTK motif in SLP-76 by the Gads SH3-C domain. The significance of this interaction in T-cell receptor signalling is demonstrated in the observation that mutation or deletion of the RSTK motif in SLP-76 completely abrogated Gads binding [68]. Furthermore, expression of a short peptide containing the Gads-binding site in T-cells blocked Gads–SLP-76 association and interrupted SLP-76 trafficking [100].

Affinity and specificity can also be enhanced by residues outside the conventional PxxP motif. In haematopoietic cells, the activity of the c-Src tyrosine kinase is negatively regulated by Csk (C-terminal Src kinase) and PEP (proline-enriched phosphatase). PEP is localized to activated Src via a specific interaction with the Csk SH3 domain. A 25-residue PEP peptide containing a PxxP motif binds to the Csk SH3 domain with high affinity. However, specificity of the PEP peptide resides in a pair of isoleucine and valine residues more than ten residues removed from the PxxP motif [101]. These residues are part of a 3^{10} helix formed by the C-terminal fragment of the PEP peptide that engages the SH3 domain through extensive hydrophobic interactions outside the PPII-binding groove (Figure 6).

Enhancing affinity through extended binding surfaces is another mechanism that is exploited by a number of SH3 domains including that from p67^{phox}. Binding of p67^{phox} to p47^{phox} is mediated by the C-terminal SH3 domain of p67^{phox} and a 32-residue fragment from the C-terminal tail of p47^{phox} (Figure 6). This interaction, with a K_d value of 24 nM, represents one of the strongest physiological binding events that is mediated by an SH3 domain [102]. Structural analysis on the p67^{phox} SH3 domain in complex with a 32-residue p47^{phox} peptide revealed some interesting features regarding this high-affinity, high-specificity interaction [102]. The peptide appears to use a 'two-prong' mechanism in binding the SH3 domain. While the N-terminal part of the peptide assumes a typical PPII helix and docks to the conventional binding site on the SH3 domain, the C-terminal 20 residues of the peptide adopt a unique helix–turn–helix structure and make extensive contacts with residues outside the PPII-binding groove (Figure 6). The importance of these C-terminal residues was confirmed in the observation that a peptide containing only the N-terminal proline-rich sequence of the p47^{phox} peptide bound to the p67^{phox} SH3 domain with a 1000-fold reduced affinity than did the 32-residue peptide [102,103].

Because modular interaction domains often occur in tandem in regulatory proteins, affinity, and hence specificity, can often be augmented through co-operative binding involving multiple

Figure 6 Interactions outside the PxxP motif augments affinity and specificity for SH3 domain

Shown on the left is the structure of the Csk SH3 domain–PEP peptide complex while on the right is the structure of the p67^{phox} SH3-C domain in complex with a 32-residue peptide derived from p47^{phox}. SH3 domains in both complexes are depicted in cyan (with loops in grey). The structures of the bound peptides are labelled as follows: 3¹⁰ and α -helices in red, PPII in magenta, and others in yellow.

domains of the same or different kinds [1,4,8,24]. Enhanced specificity can also be achieved by increased local concentrations of two interacting partners as a result of subcellular localization or scaffolding. Moreover, specific intramolecular interactions can occur between a domain and its binding sequence when they coexist in a single polypeptide chain. Such is the case for the autoinhibitory interaction between the SH3 domain and the SH2-kinase linker in Src family kinases [104].

PERSPECTIVES

It is estimated that 1000–5000 distinct stable globular protein folds exist in Nature [105]. This number agrees well with the 2000 or so domain families that are encoded by the human genome [5]. Given that the human genome encodes an estimated 30 000 or so proteins [5,6], it can be expected that some modular domains, including many protein interaction modules, are capable of interacting with multiple partners and performing diverse functions. The promiscuous nature and the versatile functions displayed by the SH3 and other PRDs suggest that the cellular role of some interaction domains may not be limited to mediate binary protein–protein associations. Understanding the molecular basis that underpins specific compared with diverse binding by PRDs and other interaction domains would provide invaluable insights into the organization and regulation of protein networks that are mediated by these domains.

Diversity in specificity may impart novel functions to an interaction domain and provide a mechanism for enhancing plasticity and dynamics of signal transduction. At the same time, the promiscuous characteristics of some interaction domains make it difficult to predict protein–protein interactions based on consensus motifs alone. These motifs have been used widely to predict potential protein–protein interactions by homology-based sequence search [18]. Most of the motifs used in these predictive schemes have been derived from screening oriented peptide libraries. Since peptide libraries used in these screening methods invariably contain a core motif that is recognized by the class of interaction domains under investigation, peptide sequences that do not conform to the predetermined core motif may be overlooked. Furthermore, as has been noted above, many domain–ligand interactions occur in the context of folded proteins instead of through binding to a short peptide sequence. These interactions

may not be readily predicted using current bioinformatic approaches. It can be envisaged that a strategy that combines proteomics and genomics methods, including systematic yeast two-hybrid screening, MS-aided protein complex identification and computer-aided pattern search, structure prediction and modelling, would ultimately afford a map of the global ligand space for interaction domains and protein networks that are mediated by them.

I thank Dr David Litchfield and Dr Tony Pawson for critical reading of the manuscript before publication. Work from my laboratory was supported by funds from the Cancer Research Society Inc. and from the Canadian Cancer Society. I am a scientist of the National Cancer Institute of Canada.

REFERENCES

- Pawson, T. and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445–452
- Sudol, M. (1998) From Src homology domains to other signaling modules: proposal of the 'protein recognition code'. *Oncogene* **17**, 1469–1474
- Songyang, Z. (1999) Recognition and regulation of primary-sequence motifs by signaling modular domains. *Prog. Biophys. Mol. Biol.* **71**, 359–372
- Pawson, T. (2004) Specificity in signal transduction: from phosphotyrosine–SH2 domain interactions to complex cellular systems. *Cell* **116**, 191–203
- Morgan, M. J. (2001) Initial sequencing and analysis of the human genome. *Nature (London)* **409**, 860–921
- International Human, Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature (London)* **431**, 931–945
- 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. et al. (2000) Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215
- Pawson, T. and Nash, P. (2000) Protein–protein interactions define specificity in signal transduction. *Genes Dev.* **14**, 1027–1047
- Pawson, T. and Scott, J. D. (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science* **278**, 2075–2080
- Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quilliam, L. A. and Kay, B. K. (1996) Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLC γ , Crk, and Grb2. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1540–1544
- Sparks, A. B., Rider, J. E. and Kay, B. K. (1998) Mapping the specificity of SH3 domains with phage-displayed random-peptide libraries. *Methods Mol. Biol.* **84**, 87–103
- Songyang, Z. (2001) Peptide library screening for determination of SH2 or phosphotyrosine-binding domain sequences. *Methods Enzymol.* **332**, 183–195
- Songyang, Z., Shoelson, S. E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X. R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T. et al. (1994) Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell Biol.* **14**, 2777–2785

- 14 Songyang, Z. (1995) SH2 domain specificity determination using oriented phosphopeptide library. *Methods Enzymol.* **254**, 523–535
- 15 Turk, B. E. and Cantley, L. C. (2003) Peptide libraries: at the crossroads of proteomics and bioinformatics. *Curr. Opin. Chem. Biol.* **7**, 84–90
- 16 Songyang, Z. and Cantley, L. C. (1995) Recognition and specificity in protein tyrosine kinase-mediated signalling. *Trends Biochem. Sci.* **20**, 470–475
- 17 Rodriguez, M., Li, S. S.-C., Harper, J. W. and Songyang, Z. (2004) An oriented peptide array library (OPAL) strategy to study protein–protein interactions. *J. Biol. Chem.* **279**, 8802–8807
- 18 Obenauer, J. C., Cantley, L. C. and Yaffe, M. B. (2003) Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31**, 3635–3641
- 19 Sadowski, I., Stone, J. C. and Pawson, T. (1986) A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. *Mol. Cell. Biol.* **6**, 4396–4408
- 20 Chandra, B. R., Gowthaman, R., Akhouri, R. R., Gupta, D. and Sharma, A. (2004) Distribution of proline-rich (PxxP) motifs in distinct proteomes: functional and therapeutic implications for malaria and tuberculosis. *Protein Eng. Des. Sel.* **17**, 175–182
- 21 Castagnoli, L., Costantini, A., Dall'Armi, C., Gonfloni, S., Montecchi-Palazzi, L., Panni, S., Paoluzi, S., Santonico, E. and Cesareni, G. (2004) Selectivity and promiscuity in the interaction network mediated by protein recognition modules. *FEBS Lett.* **567**, 74–79
- 22 Holt, M. R. and Koffer, A. (2001) Cell motility: proline-rich proteins promote protrusions. *Trends Cell Biol.* **11**, 38–46
- 23 Schweimer, K., Hoffmann, S., Bauer, F., Friedrich, U., Kardinal, C., Feller, S. M., Biesinger, B. and Sticht, H. (2002) Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Lck. *Biochemistry* **41**, 5120–5130
- 24 Zarrinpar, A., Bhattacharyya, R. P. and Lim, W. A. (2003) The structure and function of proline recognition domains. *Science STKE* 2003, RE8
- 25 Mayer, B. (2001) SH3 domains: complexity in moderation. *J. Cell Sci.* **114**, 1253–1263
- 26 Kato, Y., Nagata, K., Takahashi, M., Lian, L., Herrero, J. J., Sudol, M. and Tanokura, M. (2004) Common mechanism of ligand recognition by group II/III WW domains: redefining their functional classification. *J. Biol. Chem.* **279**, 31833–31841
- 27 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
- 28 Ball, L. J., Jarchau, T., Oschkinat, H. and Walter, U. (2002) EVH1 domains: structure, function and interactions. *FEBS Lett.* **513**, 45–52
- 29 Prehoda, K. E., Lee, D. and Lim, W. A. (1999) Structure of the enabled/VASP homology 1 domain–peptide complex: a key component in the spatial control of actin assembly. *Cell* **97**, 471–480
- 30 Volkman, B. F., Prehoda, K. E., Scott, J. A., Peterson, F. C. and Lim, W. A. (2002) Structure of the N-WASP EVH1 domain–WIP complex: insight into the molecular basis of Wiskott–Aldrich syndrome. *Cell* **111**, 565–576
- 31 Freund, C., Dotsch, V., Nishizawa, K., Reinherz, E. L. and Wagner, G. (1999) The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. *Nat. Struct. Biol.* **6**, 656–660
- 32 Mahoney, N. M., Janmey, P. A. and Almo, S. C. (1997) Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. *Nat. Struct. Biol.* **4**, 953–960
- 33 Mahoney, N. M., Rozwarski, D. A., Fedorov, E., Fedorov, A. A. and Almo, S. C. (1999) Profilin binds proline-rich ligands in two distinct amide backbone orientations. *Nat. Struct. Biol.* **6**, 666–671
- 34 Pornillos, O., Alam, S. L., Davis, D. R. and Sundquist, W. I. (2002) Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nat. Struct. Biol.* **9**, 812–817
- 35 Saito, K., Kigawa, T., Koshida, S., Sato, K., Matsuo, Y., Sakamoto, A., Takagi, T., Shirouzu, M., Yabuki, T. and Nunokawa, E. (2004) The CAP-Gly domain of CYLD associates with the proline-rich sequence in NEMO/IKK γ . *Structure* **12**, 1719–1728
- 36 Otte, L., Wiedemann, U., Schlegel, B., Pires, J. R., Beyermann, M., Schmieder, P., Krause, G., Volkmer-Engert, R., Schneider-Mergener, J. and Oschkinat, H. (2003) WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Sci.* **12**, 491–500
- 37 Schleinkofer, K., Wiedemann, U., Otte, L., Wang, T., Krause, G., Oschkinat, H. and Wade, R. C. (2004) Comparative structural and energetic analysis of WW domain–peptide interactions. *J. Mol. Biol.* **344**, 865–881
- 38 Feng, S., Kasahara, C., Rickles, R. and Schreiber, S. (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
- 39 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 (London)* **372**, 375–379
- 40 Musacchio, A., Noble, M., Paupit, R., Wierenga, R. and Saraste, M. (1992) Crystal structure of a Src-homology 3 (SH3) domain. *Nature (London)* **359**, 851–855
- 41 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
- 42 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
- 43 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
- 44 Chakrabarti, P. and Janin, J. (2002) Dissecting protein–protein recognition sites. *Proteins* **47**, 334–343
- 45 Bradshaw, J. M. and Waksman, G. (2002) Molecular recognition by SH2 domains. *Adv. Protein Chem.* **61**, 161–210
- 46 Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. and Kuriyan, J. (1993) Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790
- 47 Ladbury, J. E. and Arold, S. S. (2000) Searching for specificity in SH domains. *Chem. Biol.* **7**, R3–R8
- 48 Tong, A. H. Y., 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
- 49 Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R. and Cesareni, G. (2004) Protein interaction networks by proteome peptide scanning. *PLoS Biol.* **2**, e14
- 50 Hu, H., Columbus, J., Zhang, Y., Wu, D., Lian, L., Yang, S., Goodwin, J., Luczak, C., Carter, M., Chen, L. et al. (2004) A map of WW domain family interactions. *Proteomics* **4**, 643–655
- 51 Davies, D. R., Padlan, E. A. and Sheriff, S. (1990) Antibody–antigen complexes. *Annu. Rev. Biochem.* **59**, 439–473
- 52 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
- 53 Huang, X., Poy, F., Zhang, R., Joachimaki, A., Sudol, M. and Eck, M. J. (2000) Structure of a WW domain containing fragment of dystrophin in complex with β -dystroglycan. *Nat. Struct. Biol.* **7**, 634–638
- 54 Sundquist, W. I., Schubert, H. L., Kelly, B. N., Hill, G. C., Holton, J. M. and Hill, C. P. (2004) Ubiquitin recognition by the human TSG101 protein. *Mol. Cell* **13**, 783–789
- 55 Panni, S., Dente, L. and Cesareni, G. (2002) *In vitro* evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *J. Biol. Chem.* **277**, 21666–21674
- 56 Fernandez-Ballester, G., Blanes-Mira, C. and Serrano, L. (2004) The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains. *J. Mol. Biol.* **335**, 619–629
- 57 Mayer, B. J., Hamaguchi, M. and Hanafusa, H. (1988) A novel viral oncogene with structural similarity to phospholipase C. *Nature (London)* **332**, 272–275
- 58 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
- 59 Kang, H., Freund, C., Duke-Cohan, J. S., Musacchio, A., Wagner, G. and Rudd, C. E. (2000) SH3 domain recognition of a proline-independent tyrosine-based RKxxYxxY motif in immune cell adaptor SKAP55. *EMBO J.* **19**, 2889–2899
- 60 Barnett, P., Bottger, G., Klein, A. T. J., Tabak, H. F. and Distel, B. (2000) The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. *EMBO J.* **19**, 6382–6391
- 61 Douangamath, A., Filipp, F. V., Klein, A. T. J., 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 α -helical and PPII-helical ligands to a peroxisomal SH3 domain. *Mol. Cell* **10**, 1007–1017
- 62 Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H. and Kuriyan, J. (1995) Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* **3**, 215–226
- 63 Grabs, D., Slepnev, V. I., Songyang, Z., David, C., Lynch, M., Cantley, L. C. and De Camilli, P. (1997) The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. *J. Biol. Chem.* **272**, 13419–13425

- 64 Lohi, O. and Lehto, V.-P. (2001) STAM/EAST/Hbp adapter proteins – integrators of signalling pathways. *FEBS Lett.* **508**, 287–290
- 65 Kato, M., Miyazawa, K. and Kitamura, N. (2000) A deubiquitinating enzyme UBPY interacts with the Src homology 3 domain of Hrs-binding protein via a novel binding motif PX(V/I)(D/N)RXKK. *J. Biol. Chem.* **275**, 37481–37487
- 66 Lewitzky, M., Harkiolaki, M., Domart, M.-C., Jones, E. Y. and Feller, S. M. (2004) Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. *J. Biol. Chem.* **279**, 28724–28732
- 67 Berry, D. M., Nash, P., Liu, S. K.-W., 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
- 68 Liu, Q., Berry, D., Nash, P., Pawson, T., McGlade, C. J. and Li, S. S. (2003) Structural basis for specific binding of the Gads SH3 domain to an RxxK motif-containing SLP-76 peptide: a novel mode of peptide recognition. *Mol. Cell.* **11**, 471–481
- 69 Lewitzky, M., Kardinal, C., Gehring, N. H., Schmidt, E. K., Konkol, B., Eulitz, M., Birchmeier, W., Schaeper, U. and Feller, S. M. (2001) The C-terminal SH3 domain of the adapter protein Grb2 binds with high affinity to sequences in Gab1 and SLP-76 which lack the SH3-typical P-x-x-P core motif. *Oncogene* **20**, 1052–1062
- 70 Liu, S. K., Berry, D. M. and McGlade, C. J. (2001) The role of Gads in hematopoietic cell signalling. *Oncogene* **20**, 6284–6290
- 71 Harkiolaki, M., Lewitzky, M., Gilbert, R. J., Jones, E. Y., Bourette, R. P., Mouchiroud, G., Sondermann, H., Moarefi, I. and Feller, S. M. (2003) Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76. *EMBO J.* **22**, 2571–2582
- 72 Kaneko, T., Kumasaka, T., Ganbe, T., Sato, T., Miyazawa, K., Kitamura, N. and Tanaka, N. (2003) Structural insight into modest binding of a non-PXXP ligand to the signal transducing adaptor molecule-2 Src homology 3 domain. *J. Biol. Chem.* **278**, 48162–48168
- 73 Wigge, P. and McMahon, H. T. (1998) The amphiphysin family of proteins and their role in endocytosis at the synapse. *Trends Neurosci.* **21**, 339–344
- 74 Kojima, C., Hashimoto, A., Yabuta, I., Hirose, M., Hashimoto, S., Kanaho, Y., Sumimoto, H., Ikegami, T. and Sabe, H. (2004) Regulation of Bin1 SH3 domain binding by phosphoinositides. *EMBO J.* **23**, 4413–4422
- 75 Lee, E., Marcucci, M., Daniell, L., Pyper, M., Weisz, O. A., Ochoa, G.-C., Farsad, K., Wenk, M. R. and De Camilli, P. (2002) Amphiphysin 2 (Bin1) and T-tubule biogenesis in muscle. *Science* **297**, 1193–1196
- 76 Jia, C. Y., Nie, J., Wu, C., Li, C. and Li, S. S. (2005) Novel SH3 domain-binding motifs identified from proteomic screen of a Pro-rich region. *Mol. Cell. Proteomics*, in the press
- 77 Takeya, R. and Sumimoto, H. (2003) Molecular mechanism for activation of superoxide-producing NADPH oxidases. *Mol. Cells* **16**, 271–277
- 78 Groemping, Y. and Rittinger, K. (2005) Activation and assembly of the NADPH oxidase: a structural perspective. *Biochem. J.* **386**, 401–416
- 79 Yuzawa, S., Ogura, K., Horiuchi, M., Suzuki, N. N., Fujioka, Y., Kataoka, M., Sumimoto, H. and Inagaki, F. (2004) Solution structure of the tandem Src homology 3 domains of p47^{phox} in an autoinhibited form. *J. Biol. Chem.* **279**, 29752–29760
- 80 Groemping, Y., Lapouge, K., Smerdon, S. J. and Rittinger, K. (2003) Molecular basis of phosphorylation-induced activation of the NADPH oxidase. *Cell* **113**, 343–355
- 81 Fontayne, A., Dang, P. M., Gougerot-Pocidallo, M. A. and El-Benna, J. (2002) Phosphorylation of p47^{phox} sites by PKC α , β , δ , and ζ : effect on binding to p22^{phox} and on NADPH oxidase activation. *Biochemistry* **41**, 7743–7750
- 82 Sumimoto, H., Hata, K., Mizuki, K., Ito, T., Kage, Y., Sakaki, Y., Fukumaki, Y., Nakamura, M. and Takeshige, K. (1996) Assembly and activation of the phagocyte NADPH oxidase. *J. Biol. Chem.* **271**, 22152–22158
- 83 Bravo, J., Karathanassis, D., Pacold, C. M., Pacold, M. E., Ellison, C. D., Anderson, K. E., Butler, P. J. G., Lavenir, I., Perisic, O. and Hawkins, P. T. (2001) The crystal structure of the PX domain from p40^{phox} bound to phosphatidylinositol 3-phosphate. *Mol. Cell* **8**, 829–839
- 84 Ellison, C. D., Gobert-Grosse, S., Anderson, K. E., Davidson, K., Erdjument-Bromage, H., Tempst, P., Thuring, J. W., Cooper, M. A., Lim, Z. Y., Holmes, A. B. et al. (2001) PtdIns(3)P regulates the neutrophil oxidase complex by binding to the PX domain of p40^{phox}. *Nat. Cell Biol.* **3**, 679–682
- 85 Kanai, F., Liu, H., Field, S. J., Akbary, H., Matsuo, T., Brown, G. E., Cantley, L. C. and Yaffe, M. B. (2001) The PX domains of p47^{phox} and p40^{phox} bind to lipid products of PI(3)K. *Nat. Cell Biol.* **3**, 675–678
- 86 Ago, T., Kuribayashi, F., Hiroaki, H., Takeya, R., Ito, T., Kohda, D. and Sumimoto, H. (2003) Phosphorylation of p47^{phox} directs phox homology domain from SH3 domain toward phosphoinositides, leading to phagocyte NADPH oxidase activation. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4474–4479
- 87 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
- 88 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
- 89 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
- 90 Li, C., Josef, C., Jia, C. Y. H., Han, V. K. M. and Li, S. S.-C. (2003) Dual functional roles for the X-linked lymphoproliferative syndrome gene product SAP/SH2D1A in signaling through the signaling lymphocyte activation molecule (SLAM) family of immune receptors. *J. Biol. Chem.* **278**, 3852–3859
- 91 Li, C., Josef, C., Jia, C. Y., Gkouras, T., Han, V. K. and Shun-Cheng Li, S. (2003) Disease-causing SAP mutants are defective in ligand binding and protein folding. *Biochemistry* **42**, 14885–14892
- 92 Latour, S., Roncagalli, R., Chen, R., Bakinowski, M., Shi, X., Schwartzberg, P. L., Davidson, D. and Veillette, A. (2003) Binding of SAP SH2 domain to FynT SH3 domain reveals a novel mechanism of receptor signalling in immune regulation. *Nat. Cell Biol.* **5**, 149–154
- 93 Chan, B., Lanyi, A., Song, H. K., Griesbach, J., Simarro-Grande, M., Poy, F., Howie, D., Sumegi, J., Terhorst, C. and Eck, M. J. (2003) SAP couples Fyn to SLAM immune receptors. *Nat. Cell Biol.* **5**, 155–160
- 94 Donaldson, L. W., Gish, G., Pawson, T., Kay, L. E. and Forman-Kay, J. D. (2002) Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14053–14058
- 95 Zarrinpar, A., Bhattacharyya, R. P., Nittler, M. P. and Lim, W. A. (2004) Sho1 and Pbs2 act as coscaffolds linking components in the yeast high osmolarity MAP kinase pathway. *Mol. Cell* **14**, 825–832
- 96 Zarrinpar, A., Park, S.-H. and Lim, W. A. (2003) Optimization of specificity in a cellular protein interaction network by negative selection. *Nature (London)* **426**, 676–680
- 97 Marles, J. A., Dahe, S., Haynes, J., Andrews, B. J. and Davidson, A. R. (2004) Protein–protein interaction affinity plays a crucial role in controlling the Sho1p-mediated signal transduction pathway in yeast. *Mol. Cell* **14**, 813–823
- 98 Park, S.-H., Zarrinpar, A. and Lim, W. A. (2003) Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061–1064
- 99 Vaynberg, J., Fukuda, T., Chen, K., Vinogradova, O., Velyvis, A., Tu, Y., Ng, L., Wu, C. and Qin, J. (2005) Structure of an ultraweak protein–protein complex and its crucial role in regulation of cell morphology and motility. *Mol. Cell* **17**, 513–523
- 100 Singer, A. L., Bunnell, S. C., Obstfeld, A. E., Jordan, M. S., Wu, J. N., Myung, P. S., Samelson, L. E. and Koretzky, G. A. (2004) Roles of the proline-rich domain in SLP-76 subcellular localization and T cell function. *J. Biol. Chem.* **279**, 15481–15490
- 101 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
- 102 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
- 103 Dutta, K., Shi, H., Cruz-Chu, E. R., Kami, K. and Ghose, R. (2004) Dynamic influences on a high-affinity, high-specificity interaction involving the C-terminal SH3 domain of p67^{phox}. *Biochemistry* **43**, 8094–8106
- 104 Lim, W. A. (2002) The modular logic of signaling proteins: building allosteric switches from simple binding domains. *Curr. Opin. Struct. Biol.* **12**, 61–68
- 105 Burley, S. K. (2000) An overview of structural genomics. *Nat. Struct. Biol.* **7**, 932–934
- 106 Bedford, M. T., Sarbassova, D., Xu, J., Leder, P. and Yaffe, M. B. (2000) A novel Pro-Arg motif recognized by WW domains. *J. Biol. Chem.* **275**, 10359–10369
- 107 Beneken, J., Tu, J., Xiao, B., Nuriya, M., Yuan, J. P., Worley, P. F. and Leahy, D. J. (2000) Structure of the Homer EVH1 domain–peptide complex reveals a new twist in polyproline recognition. *Neuron* **26**, 143–154
- 108 Kofler, M., Heuer, K., Zech, T. and Freund, C. (2004) Recognition sequences for the GYF domain reveal a possible spliceosomal function of CD2BP2. *J. Biol. Chem.* **279**, 28292–28297
- 109 Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51–55