Biol. Chem., Vol. 391, pp. 33–42, January 2010 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2010.003

Structural studies of the phosphatidylinositol 3-kinase (PI3K) SH3 domain in complex with a peptide ligand: role of the anchor residue in ligand binding

Renu Batra-Safferling[1,*], Joachim Granzin[1], Susanne Mödder[2], Silke Hoffmann[2] and Dieter Willbold[2,3,*]

[1] Institut für Strukturbiologie und Biophysik ISB-2, Forschungszentrum Jülich, D-52425 Jülich, Germany
[2] Institut für Strukturbiologie und Biophysik ISB-3, Forschungszentrum Jülich, D-52425 Jülich, Germany
[3] Institut für Physikalische Biologie, Heinrich-Heine-Universität, D-40225 Düsseldorf, Germany

* Corresponding authors
e-mail: r.batra-safferling@fz-juelich.de; d.willbold@fz-juelich.de

Abstract

Src homology 3 (SH3) domains are mediators of protein-protein interactions. They comprise approximately 60 amino acid residues and are found in many intracellular signaling proteins. Here, we present the crystal structure of the SH3 domain from phosphatidylinositol 3-kinase (PI3K) in complex with the 12-residue proline-rich peptide PD1R (HSKRPLPPLPSL). The crystal structure of the PI3K SH3-PD1R complex at a resolution of 1.7 Å reveals type I ligand orientation of the bound peptide with an extended conformation where the central portion forms a left-handed type II polyproline (PPII) helix. The overall structure of the SH3 domain shows minimal changes on ligand binding. In addition, we also attempted crystallization with another peptide ligand (PD1) where the residue at anchor position P_{-3} is a tyrosine. The crystals obtained did not contain the PD1 ligand; instead, the ligand binding site is partially occupied by residues Arg18 and Trp55 from the symmetry-related PI3K SH3 molecule. Considering these crystal structures of PI3K SH3 together with published reports, we provide a comparative analysis of protein-ligand interactions that has helped us identify the individual residues which play an important role in defining target specificity.

Keywords: phosphatidylinositol 3-kinase; polyproline helix; protein-ligand interaction; Src homology 3; π stacking.

Introduction

Src homology 3 (SH3) domains are found in several proteins involved in signal transduction pathways, cell polarization, and membrane-cytoskeleton interactions (Koch et al., 1991; Musacchio et al., 1992; Pawson and Gish, 1992; Pawson and Schlessinger, 1993). SH3 domains are generally small protein domains composed of 55–70 amino acid residues. The role of SH3 domains is to mediate protein-protein interactions, for example, via binding to proline-rich motifs in the target proteins, enabling the formation of multimeric signaling complexes (Koch et al., 1991). These functions have been extensively studied in the family of protein tyrosine kinases, critical components of signaling pathways which catalyze phosphorylation of tyrosine residues on protein substrates, a key step in cellular signaling (Koch et al., 1991; Hanks and Hunter, 1995; Stout et al., 2004). One prototype family of tyrosine kinases is the Src-type tyrosine kinase family. Members of this family consist of a unique N-terminus followed by a SH3 domain, a SH2 domain, and a C-terminal tyrosine kinase domain. The SH3 domain is known to play a role in localization of the protein, as well as in substrate recognition (Kaplan et al., 1994; Weng et al., 1994; Faux and Scott, 1996). The focus of our research are various SH3 domains that are known to negatively regulate kinase activity because mutations in this domain often result in constitutively active kinases (Franz et al., 1989; Jackson and Baltimore, 1989; Superti-Furga et al., 1993; Rodrigues et al., 1997). Molecular level studies of the tyrosine kinases and their subdomains are particularly important owing to their implications in the treatment of cancer, because kinases are frequently overexpressed and/or hyperactive in cancer cells.

Three-dimensional structures of several SH3 domains as well as SH3 domains in complex with physiologically relevant or engineered ligands have been deposited in the Protein Data Bank (PDB). SH3 domains recognize unique proline-rich peptides bearing the core sequence PxxP (where x denotes any amino acid), flanked by other residues specific to the type of SH3 domain (Feng et al., 1994; Yu et al., 1994; Kuriyan and Cowburn, 1997; Tran et al., 2005). The xP-dipeptide part of the ligand peptide core sequence PxxP contacts the hydrophobic pockets while a common basic residue at position P_{-3} (anchor position) contacts a highly conserved acidic residue within the SH3 compass pocket and determines the binding orientation of the ligand (Feng et al., 1994). Dependent on the binding orientation, which can be type I or type II, positions P_0 and P_3 or P_2 and P_{-1}, respectively, are occupied by proline residues (for clarity on positions, see Figure 1A). The binding constants of these peptides are rather low, in the milli- to micromolar range, and frequently the interactions are not specific.

Previously, we identified several artificial peptide ligands by screening a phage-displayed peptide library against the

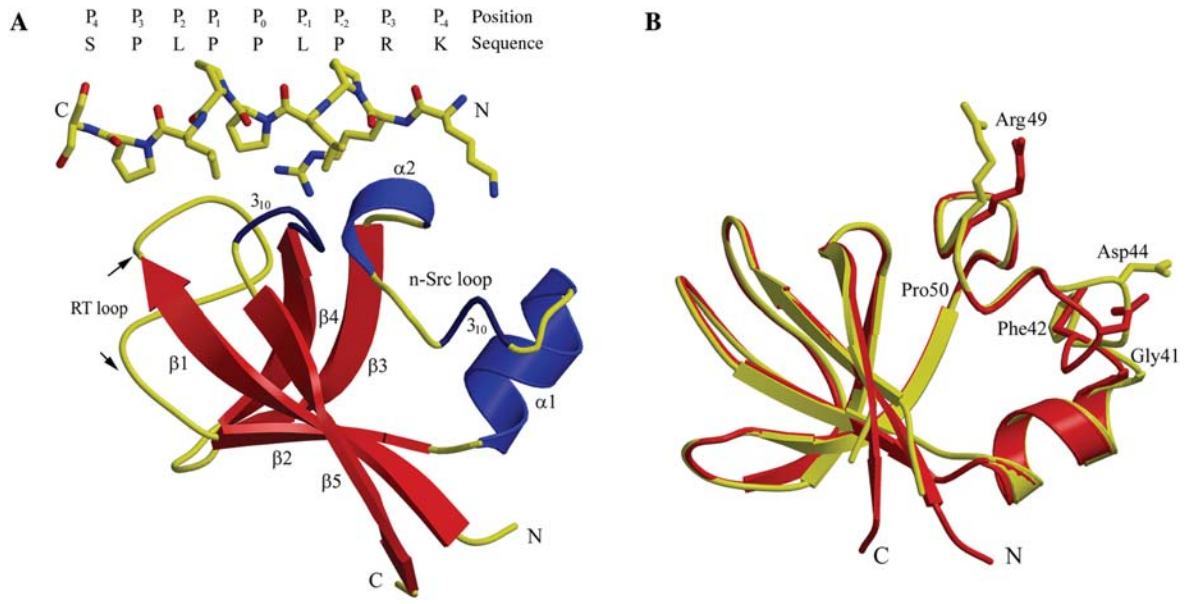

Figure 1 Crystal structure of the PI3K SH3 domain in complex with ligand peptide PD1R.
(A) PI3K SH3 domain is shown as ribbon structure where β strands (β1 4–11, β2 28–33, β3 54–60, β4 64–70, β5 74–80), loops (12–27, 42–49, 61–63, 70–73), and α helices (α1 34–41, α2 50–53) are colored red, yellow, and blue, respectively. Two 3_{10} helices (45–47, 70–73) are shown within the loops in dark blue. Residues 3–11 of the ligand peptide (with side chains) are shown as stick model and colored by element: carbon, yellow; nitrogen, blue; oxygen, red. The peptide interacts in a class I orientation, where the central part adopts a left-handed type II polyproline (PPII) helix. The amino acid sequence of the peptide PD1R is ^{1}HSKRPLPPLPSL^{12}. (B) Superposition of crystal structures of the PI3K SH3 domain bound to PD1R (red) and the free PI3K SH3 domain (yellow, PDB ID 1PHT). Note the residues in n-Src loop that shows conformational differences; side chains of selected residues are shown in stick mode.

lymphocyte specific kinase (Lck) SH3 domain (Tran et al., 2005). Most of the selected peptides follow a Class I (+XXPXXP) consensus (where + is a basic residue, X is any amino acid, and P is the conserved residue Pro), although none of these contains the very common positively charged amino acid at position P_{-3} (Feng et al., 1994; Musacchio, 2002). The binding affinities of the dominating peptide named PD1 (HSKYPLPPLPSL) to SH3 domains from Lck, Hck, Fyn, Src, and PI3K vary in a broad range between 0.2 μM and 120 μM (Schmidt et al., 2007; Stangler et al., 2007). The tightest complex formed is between PD1 and Hck SH3, and the weakest between PD1 and the SH3 domain from PI3K. We have reported the solution structure of the high affinity complex between Hck SH3 and the PD1 peptide, where the P_{-3} anchor position is occupied by a Tyr residue instead of the common basic residue, and which revealed novel interaction modes (Schmidt et al., 2007). A PD1 variant named PD1R (HSKRPLPPLPSL), which carries the common Arg anchor instead of P_{-3} tyrosine, binds to the SH3 domains of Lck, Hck, Src, Fyn, and PI3K kinases with binding constants from 0.5 μM to 40 μM (Schmidt et al., 2007; unpublished results). Compared with the peptide PD1, affinity of PD1R varies in a smaller range. The tightest PD1R complex is formed with Hck SH3, the weakest with the SH3 domain of PI3K.

Here, we present the crystal structure of the SH3 domain from phosphatidylinositol 3-kinase (PI3K) in complex with the peptide PD1R at 1.7 Å resolution. PI3K belongs to the lipid kinase family and is a heterodimer consisting of a regulatory subunit p85 that contains one SH3 and two SH2 domains, and a catalytic subunit p110 that contains the lipid kinase domain (Otsu et al., 1991; Skolnik et al., 1991; Hiles et al., 1992). When activated by growth factor stimulation, the catalytic subunit of PI3K phosphorylates phosphatidylinositols at the D3 position of the inositol ring. PI3K plays important roles in a diverse number of cellular functions including cell growth, proliferation, differentiation, motility, and vesicle transport (Fruman et al., 1998; Katso et al., 2001; Knight et al., 2006). Structural investigation of the PI3K SH3 domain in complex with peptide PD1R is expected to first elucidate the molecular interactions involved in binding, and second to shed light on the issue of affinity and specificity of ligand binding to SH3 domains in general.

Both X-ray and solution structures of the PI3K SH3 domain have been determined previously (Booker et al., 1993; Koyama et al., 1993; Yu et al., 1994; Liang et al., 1996). However, no structure of a ligand-bound complex of the PI3K SH3 domain is available in the PDB database. With a focus on the crystal structure of PI3K SH3 domain in complex with PD1R, we provide a comparative analysis of protein-ligand interactions that will enhance our understanding of the molecular determinants of binding specificity of these versatile adaptor modules.

Results and discussion

Overall structure

The crystal structure of the SH3 domain from human PI3K kinase (residues 1–83) in complex with the PD1R peptide (^{1}HSKRPLPPLPSL^{12}) (Figure 1A) has been determined at 1.7 Å resolution (Table 1). The final atomic model contains two chains, a SH3 domain with residues 2–82, a peptide chain with residues 3–11, and 69 water molecules. The other residues of the SH3 domain are disordered, and those of the peptide show weak electron density. According to Ramachandran plots generated with PROCHECK (CCP4), the model exhibits good geometry (Table 1), with none of the residues in the disallowed region of the Ramachandran plot.

The overall structure of the PI3K SH3 follows the typical β-barrel fold observed for the SH3 domains (Booker et al., 1993; Koyama et al., 1993; Yu et al., 1994; Liang et al., 1996; Larson and Davidson, 2000). The root mean square (rms) distance among 81 equivalent Cα atoms in ligand-bound (this work) and free PI3K SH3 (PDB ID 1PHT) structures is 0.63 Å. As described in previous publications, the structure consists of two orthogonal β-sheets containing five antiparallel strands (4–11, 28–33, 54–60, 64–70, 74–80), two helices (34–41, 50–53), and four loops (12–27, 42–49, 61–63, 70–73) containing two short 3_{10} helices (45–47,

Table 1 Data collection and refinement statistics.

	PI3K SH3-PD1R	PI3K SH3
Data collection		
Space group	$P3_121$	$P2_1$
Cell dimensions		
a (Å)	60.37	49.33
b (Å)	60.37	61.05
c (Å)	40.11	61.01
α, β, γ (°)	90, 90, 120	90, 111.4, 90
Resolution (Å)	1.7	3.0
Beamline	ESRF ID14-1	ESRF ID14-1
Detector	ADSC 210	ADSC 210
Wavelength (Å)	0.934	0.934
Unique reflections	11076	6390
Average redundancy	6.7	3.3
Completeness (%)	99.7 (96.5)[a]	94.0 (95.2)[a]
R_{merge} (%)	4.7 (24.3)[a]	11.3 (50.9)[a]
Wilson B-factor	18.12	54.05
I/σ (I)	26.5 (5.5)[a]	10.3 (2.6)[a]
Refinement		
Resolution	34.75–1.7	45.9–3.0
R_{work} (%)	17.9	24.0
R_{free} (%) (test set size, %)	21.8 (4.75)[b]	29.9 (4.56)[b]
Residue range protein	2–82	A, B, C, D 4–80 (ncs restrained)
Residue range peptide	3–11	–
Number of atoms		
Protein	644	2489 (4 mol/asu)
Peptide	70	–
Water	69	–
Average B-factor		
Protein	19.86	34.99
Peptide	31.91	–
Water	33.97	–
rms deviations		
Bonds (Å)	0.008	0.005
Angles (Å)	1.254	0.799

[a]Numbers in parentheses indicate the value for the outermost resolution shell.
[b]Values in parentheses refer to the percentage of (reserved free R) reflections that were not included in the refinement.
Common residue range (only Cα atoms):
RMSD: for PI3K SH3; molA to molB: 0.170 Å, molA to molC: 0.204 Å, molA to molD: 0.169 Å;
RMSD: from PI3K SH3-PD1R to PI3K SH3-molA: 0.755 Å;
RMSD: from PI3K SH3-molA to 1PHT: 0.457 Å;
RMSD: from PI3K SH3-PD1R to 1PHT: 0.633 Å.

70–73) in loop 2 and loop 4, respectively (Figure 1A). The RT loop located between strands β1 and β2 contains important residues which are generally involved in ligand binding. Strands β2 and β3 are connected by the second loop called n-Src loop. In the case of PI3K SH3, the n-Src loop contains a unique insertion of 15 amino acid residues (for explanation of nomenclature of the loops, see Musacchio, 2002). In contrast to the free PI3K SH3 crystal structure (PDB ID 1PHT), residues in the n-Src loop are well defined in the complex structure. The main difference is seen in the first part of the loop (residues 41–44), which is most probably as a result of the presence of ligand and/or in addition, owing to differences in crystal packing causing rearrangements in this loop located on the surface (Figure 1B). Residues in this loop make a close contact with the ligand. Moreover, rearrangement of side chains of residues Gln46 and Glu52 in this loop lead to an additional salt bridge providing stability to the conformation of the loop. Loop 3 connecting β3 and β4; and loop 4 connecting β4 and β5 are short, have well-defined electron densities and are similar to the previously reported crystal structure of the free PI3K SH3 domain (PDB ID 1PHT).

Protein-ligand binding site

SH3 ligands are pseudosymmetrical and can bind in two opposite orientations, type I orientation with the consensus sequence +pXPpP (class I ligands, where + is a basic residue; P is conserved residue Pro; p is Pro preferred), and type II orientation with the consensus sequence XPpXPp+ (class II ligands) (Feng et al., 1994; Lim et al., 1994; Yu et al., 1994; Kuriyan and Cowburn, 1997). In the two XP-dipeptide parts, Pro residues are main contributors to van der Waals interactions within the complex. The basic residue at position P_{-3} (for nomenclature, see Figure 2A) is the anchor position that interacts with a conserved acidic residue in the SH3 domain and is thought to determine the binding orientation of the peptide (Kuriyan and Cowburn, 1997).

Electron densities for residues 3–11 of the bound peptide PD1R are well defined in our structure (Figure 2A). The

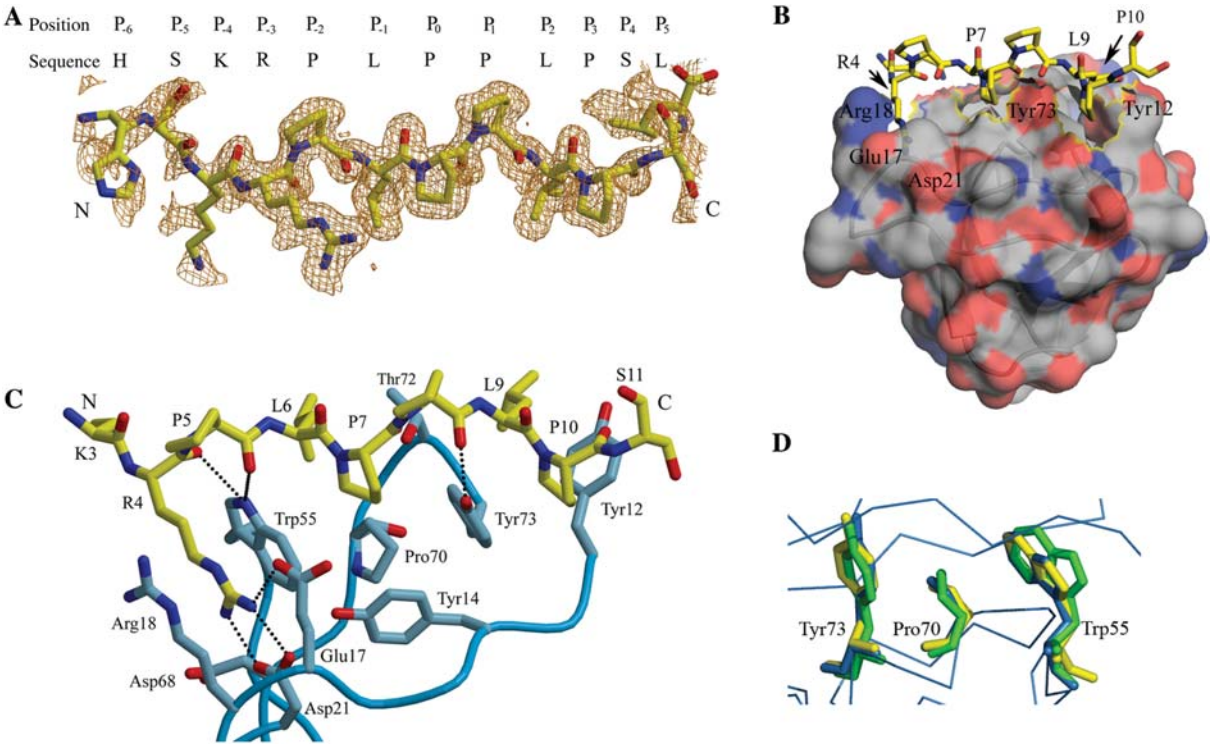

Figure 2 Protein-ligand interactions in the PI3K SH3-PD1R complex.
(A) The $2F_o$-F_c electron density map of the ligand PD1R contoured at 1.3 σ. The density of the core segment of the peptide (residues R4 to P10) is well defined. (B) Molecular surface of PI3K SH3 with the bound PD1R peptide showing binding pockets – for residues R4, P7, and P10. For clarity, the amino acid designation used for the protein and the peptide is in three-letter and one-letter code, respectively. (C) Intermolecular interactions at the ligand binding site. Both residues from the bound PD1R peptide and the side chains of protein residues (light blue) playing a crucial role in ligand binding are shown as stick models. The Cα trace of the SH3 domain is shown in blue. Only the hydrogen bonds between the SH3 domain and its ligand are shown as dotted lines. (D) Superimposition of ligand-bound SH3 structures showing the conformations of conserved residues Trp55, Pro70, and Tyr73 (numbering according to PI3K SH3 sequence). The Abl SH3 domain in complex with a type I ligand (PDB code 1ABO, colored green) and the c-Crk SH3 domain in complex to a type II ligand (PDB code CKB, colored yellow) were superimposed on the PI3K SH3 domain in complex with peptide PD1R (colored blue). The conserved residues Trp and Pro are known to adopt the SH3-I and SH3-II orientation in the case of Abl and c-Crk, respectively. In the PI3K SH3-PD1R complex, they adopt the SH3-II orientation, classifying PD1R as a class I′ ligand.

bound peptide adopts a left-handed type II polyproline (PPII) helix conformation and is seen in the type I orientation. The residues with their side chains facing the protein are Arg4, Pro7, and Pro10, where Arg4 is at the anchor position P_{-3}, and residues Pro7 and Pro10 correspond to positions P_0 and P_3, critical for ligand binding. Pro7 sits in the pocket formed by the residues Tyr14, Trp55, Pro70, and Tyr73 of the SH3 domain (Figure 2B,C). Pro10 and the side chain of residue Leu9 are located in a common pocket, interacting with residues Tyr12 and Tyr73 of the SH3 domain. Residues Lys3 and Pro5 have no contact to the protein. The side chain of Leu6 is close to the protein surface but shows no specific interaction with the SH3 domain. Arg4 lies sandwiched between residues Arg18 and Trp55 of the protein, and is surrounded by acidic residues Glu17, Asp21, and Asp68, where it forms salt bridges with Glu17 and Asp21 located on the RT loop (Figure 2C).

The salt bridge between anchor residue Arg4 of the ligand and negatively charged residue Asp21 is critical for ligand binding and ligand specificity. In a previous study, Yu et al. reported the solution structure of the PI3K SH3 domain complexed to another class I peptide ligand called RLP1 (RKLPPRPSK) containing Arg at position P_{-3} (Yu et al., 1994). The authors did not observe the salt bridges mentioned above in the structure, but their importance could be confirmed by mutagenesis experiments where the mutation of residue Asp21 to Asn led to a 50-fold decrease in binding affinity of the ligand. Interestingly, the acidic nature of position 21 is highly conserved among all SH3 domains where it is occupied by either Asp or Glu, with the single exception of Abl SH3 domain that contains Thr at this position. Located at the far end of the RT Loop, residue Asp21 is involved in highly conserved interactions within the loop. For example, the side chain of Asp21 is hydrogen bonded to the amide nitrogen of Arg18, and to the side chain of Tyr14. Such interactions are common to several SH3 domains, highlighting the importance of conserved residues in maintaining the optimal conformation of RT loop for peptide binding (Larson and Davidson, 2000). Mutation of these residues will thus perturb the interactions which in turn will have a significant effect on ligand affinity or specificity.

The main chain carbonyl group of residue Pro8 of ligand PD1R forms a hydrogen bond to the hydroxyl group of Tyr73 of the protein, and the Pro side chain is otherwise turned away from the protein molecule. Residue Leu9 shows a loose stacking with the Tyr12 of the protein. Despite the absence of interaction with the SH3 domain, the first and last residues in the peptide ligand (Lys3 and Ser11) show electron density (Figure 2A).

PD1R peptide binds to the SH3 domains as a class I′ ligand

The PD1R peptide contains the core consensus LPX(L/A)P with Leu at positions P_{-1} and P_2 as found in SH3 class I′ ligands (Tran et al., 2005; Schmidt et al., 2007). Classification of ligands in class I′ is based on the orientation of a conserved Trp in the ligand binding site of the SH3 domain, which determines the type of ligand (I, I′ or II) binding to the domain (Fernandez-Ballester et al., 2004). In general, class I′ ligands bind to the SH3 domain in type I ligand orientation, but the PPII helix conformation shows higher similarity to the class II ligands, with the two Leu residues that are typical for class I′ ligands occupying the same binding pocket positions as usually the Pro residues in the type II orientation. A superimposition of the PI3K SH3-PD1R complex structure with different SH3 domains bound to class I and class II ligands reveals that in our PI3K SH3 structure the conformation of the conserved residues Trp55 and Pro70 is similar to that in class II ligand complexes (Figure 2D). Thus, PD1R binds to PI3K SH3 as a class I′ ligand. Previously, we have reported the binding of ligand PD1 to the Hck SH3 domain as a class I′ ligand (Schmidt et al., 2007). Furthermore, we compared the Chi torsion angles for the conserved Trp and Pro residues from SH3 domains bound to class I, class II, and PD1R peptides. Interestingly, we find that the Chi angles of the Pro residue (Pro70 in our structure) show a significant difference between class I and class II ligands. A typical Pro residue can have three rotamers with Chi angles 1 and 2 as -14° and 30° for rotamer 1; 29° and -28° for rotamer 2; and 0° and 0° for rotamer 3. When bound to class I ligands, the Pro residue is in rotamer 1 conformation while in the presence of class II ligands it has the typical torsion angles of rotamer 2. In the presence of peptide PD1R, the Pro residue is in rotamer 2 conformation as expected for a class I′ ligand. The torsion angles of the Trp residue did not show significant variations in class I, class II, and class I′ ligands.

We then compared our PI3K SH3-PD1R structure with the available structures of SH3 domains bound to class I′ ligands. Interestingly, the structures available in the literature for SH3 domains in complex with class I′ ligands are from NMR studies in solution. Superimposition of the PI3K SH3-PD1R structure on eight different class I′ ligand-bound SH3 domains reveal similar conformations of the conserved Trp residue (Trp55 in our structure). The other conserved residue Pro (Pro70 in our structure) adopts the rotamer 2 conformation in all cases. Our complex structure is the first crystal structure in complex with a class I′ ligand.

A novel self-interaction mode of SH3 domains via the ligand binding site

Next, we attempted to co-crystallize the PI3K SH3 domain with the originally selected peptide PD1, carrying a Tyr at position P_{-3} (^{1}HSKYPLPPLPSL^{12}). The binding affinity of PD1 for the PI3K SH3 domain, as measured by surface plas-

Table 2 Affinities of the peptide ligands to the PI3K SH3 domain.

	Sequence	K_d value (μM)
PD1R	HSKRPLPPLPSL	40
PD1	HSKYPLPPLPSL	120

mon resonance analysis, was three times weaker than for PD1R (Table 2). Theoretically, the K_d value in the micromolar range and the presence of two-fold molar excess of the peptide ligand during crystallization (2 mM peptide and 1 mM protein) should predominantly result in a ligand-bound population. However, in the case of the PD1 peptide, the protein crystallized as free PI3K SH3 domain and not as PI3K SH3-PD1 complex.

The crystals obtained of free PI3K SH3 domain belong to space group $P2_1$, diffracted to 3 Å resolution and the structure was determined by molecular replacement using the X-ray structure of the PI3K SH3 domain (PDB code 1PHT; Liang et al., 1996) (Table 1). The structure is similar to those determined previously with an rms distance of 0.46 Å between 81 equivalent Cα atoms of the two free structures, and 0.75 Å between our free and the PD1R complex structure. No electron density could be observed for the peptide ligand PD1, and part of the binding site of the PI3K SH3 domain is occupied by residues Arg18 and Trp55 of a symmetry-related molecule (by operation of x, y+1/2, z, shown in light coral in Figure 3A). As seen in Figure 3A, residues Arg18 and Trp55 of the protein form a stack with the residues Arg18′ and Trp55′ of the symmetry related molecule such that the Arg side chains are sandwiched between the indole moieties of the Trp residues. The Trp-Arg-Arg interactions seen here are similar to those observed between residues Arg18′ and Trp55′ of the protein and the anchor residue Arg4 of the ligand peptide PD1R in the PI3K SH3-PD1R complex (Figure 3B).

Residue Trp55 is strictly conserved among SH3 domains. It participates in the formation of the hydrophobic pocket where ligand binding takes place. In contrast, structure-based alignments show that residue Arg18 is not conserved and much structural variation is seen at this position upon ligand binding (Larson and Davidson, 2000). Such non-conserved positions important for ligand binding are predicted to play a role in defining specificity for the ligand. Here, we show that the ligand PD1, which is PD1R with a Tyr at the anchor position instead of Arg, interacts with PI3K SH3 in solution (Table 2), but in the crystal the PD1 ligand is not present. It is plausible that the Trp-Arg-Arg stack formation (as seen in the PD1R complex structure) is entropically a favorable event. The stabilizing role of Trp-Arg-mediated π-π stacking has been recognized in crystal structures for several receptors (de Vos et al., 1992; Livnah et al., 1996). Because the Arg residue is absent at the anchor position in the PD1 ligand, the protein prefers residue Arg18 of the symmetry related protein molecule instead, allowing for crystal packing where the ligand is omitted. Residue Tyr when modeled in place of the Arg in the PI3K SH3-PD1R complex structure does not clash with the SH3 domain, because the space is sufficient in the binding site. However, the site favors a positively charged residue as shown in the case of PI3K SH3-PD1R and Hck SH3-PD1 complexes, where the residues Arg4 and Lys3 are present, respectively. Owing to a shorter side chain of Tyr, the stacking with Trp55 and Arg18 residues in the SH3 domain is not ideal and will be less favored than with Arg. In addition, instead of the conserved salt bridge between anchor residue Arg (peptide) with Asp21 (SH3 domain), the only possibility with residue Tyr (peptide) and Asp21 is a hydrogen bond.

In a previous report, the solution structure of the PI3K SH3 domain was determined in a peptide complex using NMR spectroscopy where the peptide RLP1 (RKLPPRPSK) binds to the SH3 domain with a K_d of 9.1 μM (Yu et al., 1994), which is approximately 10-fold and 4-fold stronger

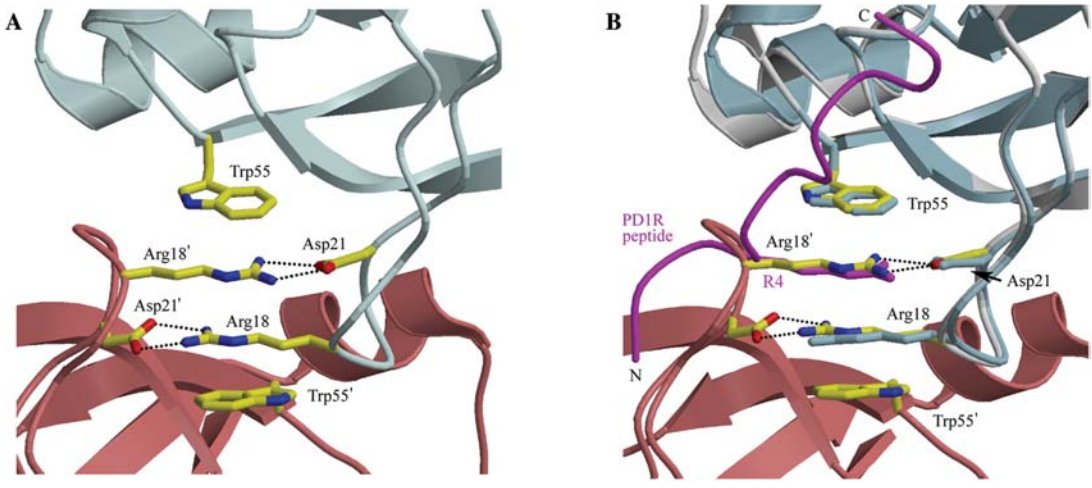

Figure 3 Ligand binding site in the crystal structure of unliganded PI3K SH3 domain at pH 5.5.
(A) Residues Arg18 and Trp55 are stacking with the same residues (indicated with the prime symbol) from the symmetry related molecule in the crystal. Ribbon structures of PI3K SH3 domain and the symmetry related molecule are shown in light blue and in coral, respectively. The salt bridge between Arg18 (usually contributed by the basic anchor residue in the ligand) and Asp21′ is shown as a dotted line. (B) Superimposition of unliganded PI3K SH3 domain as shown in (A) and the crystal structure of the PI3K SH3-PD1R complex (dark blue ribbon structure where the backbone of bound peptide PD1R is shown in magenta). Note the exact overlap of residue Arg4 of the PD1R ligand (R4 in magenta) and the Arg18′ of the symmetry related molecule.

than PD1 and PD1R, respectively. However, co-crystallization in the presence of the same ligand at a 1:1 molar ratio resulted in a free PI3K SH3 domain crystal structure (Liang et al., 1996). In this case, the binding site in the crystal structure is occupied by residues from two symmetry related molecules. These are Arg18 from one molecule occupying the anchor residue position, and the four C-terminal residues ^{82}Ile-Ser-Pro-Pro^{85} of another molecule where the diproline takes the place of the Pro-Pro motif of the class I ligand. Although the authors emphasize the significance of the C-terminus interaction owing to the presence of two prolines, these residues are absent in our own free structure which terminates at position Ser83. The polyproline binding site is thus unoccupied in our structure and the only close interaction seen is with the Arg18 of the symmetry related molecule via Trp-Arg-Arg stacking (Figure 3).

To understand which parameters exactly determine whether SH3-SH3 dimers or certain SH3 ligand complexes will crystallize more readily is complicated to predict and is certainly not solely dependent on the K_d value of the respective SH3-ligand complex but also on the absolute concentrations of the components, as well as solvent conditions, e.g., pH and salt concentrations.

Binding modes of PD1R and PD1 peptides – comparison of PI3K SH3-PD1R and Hck SH3-PD1 complexes

We compared the PI3K SH3-PD1R complex crystal structure with the previously reported NMR solution structure of hematopoietic cell kinase (Hck) SH3 domain in complex with PD1 (Figure 4A) (Schmidt et al., 2007). Good agreement is observed for the conformation of residues 5–10 of the ligand peptides, but terminal residues 1–4 and 11–12 show significant differences, with residues 10–12 being poorly defined in both structures. The residues of peptide PD1R adopt a regular left-handed PPII conformation throughout, whereas the N-terminal residues 1–4 of PD1 in the Hck SH3-PD1 structure bind to the protein surface in a S-shaped manner. Here, the lysine residue at position P_{-4} substitutes the function of the anchor residue (usually found at position P_{-3}) (Figure 4B). The Cα-Cα distance between Lys3 of PD1 and Arg4 of PD1R is 0.74 Å and Cγ-Cγ is 0.92 Å. In the Hck SH3-PD1 structure, the side chain of residue Lys3 is not well defined because the intermolecular NOEs between the side-chain protons of Lys3 and protons of Hck SH3 residues are absent (Schmidt et al., 2007). However, among 20 calculated conformers, some yielded the side chain of Lys3 to be within the compass pocket. The final position of the Lys3 side chain, as shown in Figure 4B, was based on the structure calculations employing additional constraints defining a salt bridge between Lys3 and Asp95 (of the HckSH3 domain) (Schmidt et al., 2007). Residues 6–10 occupy the conventional PxxP motif site where Pro7 and Pro10 occupy the binding pockets and adopt left-handed PPII helix conformation. The kink at Pro5 is a result of the shifted anchor residue (Figure 4B). N- and C-terminal residues of PD1R do not show interactions with the protein and are floating 'free' in solvent, whereas in PD1 they generate several hydrophobic and electrostatic interactions giving the peptide a compact conformation on the protein surface. Structural alignment of both SH3 domains shows that the core residues

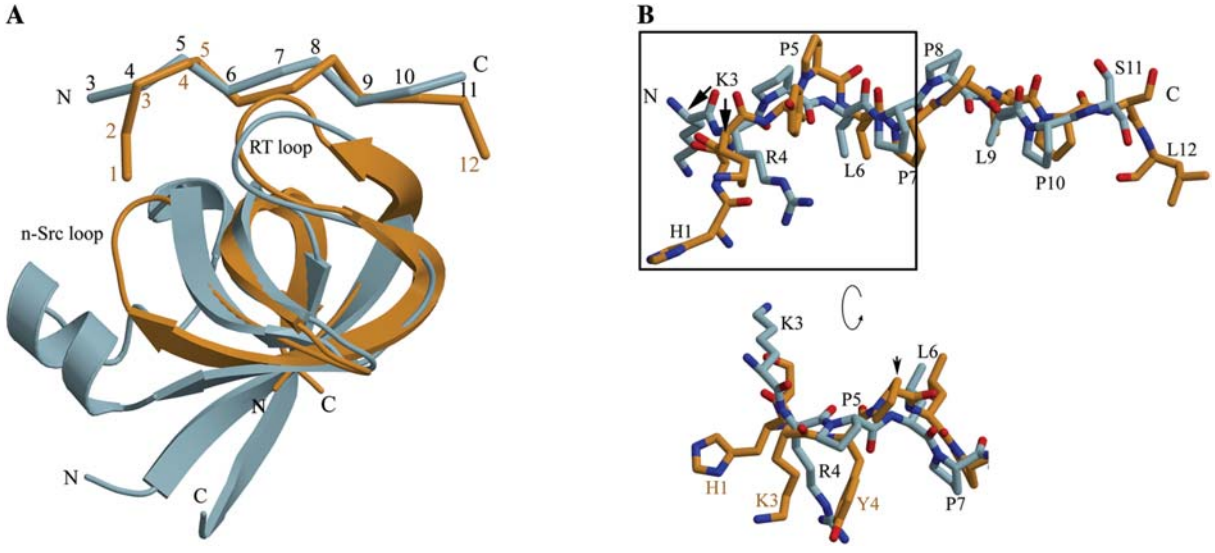

Figure 4 Superposition of the crystal structure of the PI3K SH3-PD1R complex (blue) and the solution structure of the Hck SH3-PD1 complex (orange).

Only the coordinates of the SH3 domains were the basis for the overlay with the UCSF Chimera package using the Smith-Waterman algorithm (weightings used: 70% sequence and 30% secondary structure). (A) Ribbon representation of both SH3 domains and bound peptides. (B) Two views (lower view is rotated by ∼45° along the x-axis and ∼20° along the y-axis relative to the upper view) on the stick mode representation of the bound peptides PD1R and PD1. Note residue Lys3 in PD1 taking the anchor position (the Cα-Cα distance between K3 of PD1 and R4 of PD1R is 0.74 Å) and the kink at position Pro5. Unlike the N- and the C-terminus, the core region of both peptides overlays well from position Leu6 to Pro10.

involved in ligand binding are well conserved. The N-terminal residues of PD1 show interaction with residues of the n-Src loop of the Hck SH3 domain. The residues of this loop show the least degree of sequence conservation among SH3 domains (Larson and Davidson, 2000). Such additional interactions, as seen in the case of the Hck SH3-PD1 structure, could be the reason for a many-fold higher affinity of PD1 peptide to Hck SH3 (K_d: 0.2 μM) (Schmidt et al., 2007) compared with PI3K SH3. In the case of the PI3K SH3 domain, the n-Src loop is 15 residues longer and does not show interactions to the N-terminal residues of the ligand PD1R (K_d: 40 μM). This suggests that residues of this loop play an important role in affinity selection of ligands. Future studies on the role of non-conserved residues of the RT loop and the n-Src loop will provide additional insights on how SH3 domains select their respective ligands.

Biological relevance

Protein-ligand interactions characterize and govern the current state and the fate of a living cell. It is important to realize that regulatory networks within the cell are largely based on transient (low affinity) interactions between proteins. Thus, high and low affinity interactions are important *in vivo*. The availability of potential ligands and the specificities of each protein determine the extent to which complexes will form. As an important consequence of this, not only the absolute affinity of a certain protein-ligand pair governs the concentration of its complex but the affinity in relation to other available ligands and the absolute concentrations of both protein and ligands have to be considered.

The majority of knowledge on protein-ligand interactions is based on high affinity complexes. This is simply because of the fact that it is easier to obtain co-crystals of proteins with high affinity ligands, and NOE-based NMR structures rely on the average lifetime of a protein-ligand contact, which is long enough in most cases only for high affinity complexes. The complex structure of the PI3K SH3 domain with the low affinity ligand PD1R reported herein is a rare case of a 3D structure showing a low affinity complex. It reveals that basically only the anchor residue and the PXXP motif engage in direct contacts with the SH3 domain. This could lead to the conclusion that increased affinity is achieved only by additional contacts of residues other than the anchor residue and the PXXP motif, as seen in the Hck SH3-PD1 complex structure. The important role of the anchor residue in peptide-SH3 recognition is underlined by the observation that even formation of a PI3K SH3 dimer, which does not occupy the PXXP binding sites, is favored over the formation and crystallization of the SH3-PD1 peptide complex, which is lacking the optimal anchor residue.

Concluding remarks

To summarize, both ligands PD1R and PD1 bind to the PI3K SH3 in solution in the micromolar range. Crystallization trials of both complexes yielded co-crystals of PI3K SH3 only with PD1R, but not with PD1. The ligand binding site in the latter case is partially occupied by residue Arg18 from a symmetry related SH3 molecule. The 12 amino acid residues long peptides PD1R and PD1 are identical except for position P_{-3} where Arg and Tyr are present, respectively. Our investigation of the two crystal structures presented here suggest that the residues Arg18 and Trp55 in the ligand binding site of the PI3K SH3 domain prefer Arg over Tyr at the anchor position. The study thus helped us to identify residues which play an important role in defining target specificity, a subject of great interest in the area of protein chemistry.

Materials and methods

Protein expression and purification

The PI3K SH3 cDNA was expressed as a glutathione-S-transferase (GST) fusion protein, as described previously (Tran et al., 2005). Briefly, the DNA encoding the SH3 domain of PI3K (residues 1–83) was amplified from plasmid IMAGp998J21580 (imaGenes GmbH, Berlin, Germany) by PCR and cloned into the vector pGEX-6P-2 (GE Healthcare, Freiburg, Germany) using the restriction sites BamHI and XhoI. DNA sequencing of the final construct confirmed 100% identity with human PI3K SH3 (accession code: P27986). The GST-PI3K SH3 fusion protein was produced in *Escherichia coli* strain BL21 transformed with the vector pGEX-PI3K SH3. The cells were grown at 37°C until they reached an $A_{600\ nm}$ of 0.6. At this stage, induction was performed by the addition of 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and the cells were grown further for 4 h at 37°C before the harvest. GST-PI3K SH3 fusion protein was purified using affinity chromatography on glutathione sepharose. The N-terminal GST tag was removed from the PI3K SH3 protein by proteolysis using PreScission protease (GE Healthcare) according to the manufacturer's protocol. The resulting protein mixture was then loaded on the gel filtration column Superdex-75 prep grade to obtain the purified PI3K SH3 protein used for crystallization studies.

Peptide PD1R was purchased from JPT Peptide Technologies GmbH (Berlin, Germany). Identity of the peptide was confirmed by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS).

Determination of dissociation constants (K_d) by surface plasmon resonance (SPR) analysis

The dissociation constants for the complexes between the SH3 domain from PI3K and peptides PD1 or PD1R were measured using a BIAcore X instrument (GE Healthcare). The PI3K SH3 domain was immobilized on a CM5 carboxymethyldextran biosensor chip (GE Healthcare) via amine coupling. Immobilization of the SH3 domain (in 10 mM sodium acetate, pH 4.5) was performed at 25°C at a flow rate of 5 μl/min in HBS-EP running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% P20 surfactant). In each case a second channel on the same chip was treated only with the sodium acetate buffer as a negative control. The binding curves of peptide PD1 or PD1R at various concentrations to the SH3 domain were determined at 25°C with a flow rate of 5 μl/min. The K_d values were calculated using the Biacore Analysis Software (BIAevaluation, GE Healthcare) by fitting the mean maximum response values from the binding curves to a simple 1:1 ligand receptor binding model.

Crystallization and data collection

The purified protein was concentrated to 1 mM in 10 mM Tris buffer (pH 8.0) and mixed with a 2-fold molar excess of the peptide in 10 mM Tris buffer (pH 8.0). The protein-peptide mixture was incubated for 12 h at 23°C prior to crystallization setup. Crystals were grown at 17°C using the sitting drop vapor diffusion method by combining 1 µl of the protein-peptide mixture and 1 µl of the reservoir solution (for the PD1R complex: 100 mM CAPS buffer, pH 10.5, 2 M ammonium sulfate and 0.2 M lithium sulfate; for PD1 complex: 100 mM Na-citrate buffer, pH 5.5, 0.5 M ammonium sulfate and 1 M lithium sulfate). After several weeks, single crystals appeared that were flash frozen in liquid nitrogen for data collection at 100 K. Prior to cryocooling, the crystals were soaked in reservoir solution containing up to 25% (v/v) glycerol for a few minutes.

Native data were recorded at beamline ID14-1 of the ESRF (Grenoble, France) tuned to a wavelength of 0.934 Å on a Quantum ADSC-210 detector. Data processing including reflections up to 1.7 Å resolution was carried out using MOSFLM (Leslie, 1992) and SCALA, which is part of the CCP4 software suite (http://www.ccp4.ac.uk).

Structure determination and refinement

Crystals of the PI3K SH3-PD1R complex belonged to space group $P3_121$. The structure was determined by molecular replacement using MOLREP (CCP4) with a single native dataset. The search model was created from the crystal structure of PI3K SH3 (PDB code 1PHT) (Liang et al., 1996). Crystals were found to contain one molecule per asymmetric unit, corresponding to a Matthews coefficient of 2.3 $Å^3$/Da and a solvent content of 46.3%. The model was improved by several cycles of rigid body refinement and in a second stage by positional, B-factor and TLS (Translation/Libration/Screw) refinement using the PHENIX package (Adams et al., 2002). Separate TLS groups were chosen for the protein and the peptide chain. For manual rebuilding the program O was used (Jones et al., 1991). For statistics on data collection and refinement refer to Table 1. The final model comprises amino acids 2–82 of the protein molecule (chain A) and amino acids 3–11 of the peptide (chain B). The structure of PI3K SH3 in the unbound state was determined essentially using the same strategy. In this case, the crystals belonged to space group $P2_1$ and contained four molecules per asymmetric unit (Table 1).

Figure preparation

Unless otherwise indicated, Figures were generated with MOLSCRIPT (Kraulis, 1991) and RASTER3D (Merritt and Bacon, 1997) using secondary structure assignments as given by the DSSP program (Kabsch and Sander, 1983). The surface representation in Figure 2 was prepared with PyMOL (DeLano, 2008).

The superposition in Figure 4 was performed with the UCSF Chimera package (Meng et al., 2006) using a Smith-Waterman algorithm (weightings used: 70% sequence and 30% secondary structure).

Database accession codes

The atomic coordinates and structure factors for free PI3K SH3 and the complex with PD1R (codes 3I5R and 3I5S, respectively) have been deposited in the PDB (http://rutgers.rcsb.org/pdb).

Acknowledgments

The authors wish to thank Georg Büldt for continuous generous support. Furthermore, technical assistance by Esther Jonas and the beamline scientists at the ESRF (Grenoble, France) is acknowledged. We also thank Oliver Weiergräber for critical reading of the manuscript. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 575, B11) to D.W.

References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Cryst. *D58*, 1948–1954.

Booker, G.W., Gout, I., Downing, A.K., Driscoll, P.C., Boyd, J., Waterfield, M.D., and Campbell, I.D. (1993). Solution structure and ligand-binding site of the SH3 domain of the p85 α subunit of phosphatidylinositol 3-kinase. Cell *73*, 813–822.

DeLano, W.L. (2008). The PyMOL Molecular Graphics System (Palo Alto, CA, USA: DeLano Scientific LLC), http://www.pymol.org.

de Vos, A.M., Ultsch, M., and Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science *255*, 306–312.

Faux, M.C. and Scott, J.D. (1996). More on target with protein phosphorylation: conferring specificity by location. Trends Biochem. Sci. *21*, 312–315.

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.

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.

Franz, W.M., Berger, P., and Wang, J.Y. (1989). Deletion of an N-terminal regulatory domain of the c-abl tyrosine kinase activates its oncogenic potential. EMBO J. *8*, 137–147.

Fruman, D.A., Meyers, R.E., and Cantley, L.C. (1998). Phosphoinositide kinases. Annu. Rev. Biochem. *67*, 481–507.

Hanks, S.K. and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *9*, 576–596.

Hiles, I.D., Otsu, M., Volinia, S., Fry, M.J., Gout, I., Dhand, R., Panayotou, G., Ruiz-Larrea, F., Thompson, A., Totty, N.F., et al. (1992). Phosphatidylinositol 3-kinase: structure and expression of the 110 kd catalytic subunit. Cell *70*, 419–429.

Jackson, P. and Baltimore, D. (1989). N-terminal mutations activate the leukemogenic potential of the myristoylated form of c-abl. EMBO J. *8*, 449–456.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110–119.

Kabsch, W. and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.

Kaplan, K.B., Bibbins, K.B., Swedlow, J.R., Arnaud, M., Morgan D.O., and Varmus, H.E. (1994). Association of the amino-terminal half of c-Src with focal adhesions alters their properties

and is regulated by phosphorylation of tyrosine 527. EMBO J. *13*, 4745–4756.

Katso, R., Okkenhaug, K., Ahmadi, K., White, S., Timms, J., and Waterfield, M.D. (2001). Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. Annu. Rev. Cell Dev. Biol. *17*, 615–675.

Knight, Z.A., Gonzalez, B., Feldman, M.E., Zunder, E.R., Goldenberg, D.D., Williams, O., Loewith, R., Stokoe, D., Balla, A., Toth, B., et al. (2006). A pharmacological map of the PI3-K family defines a role for p110α in insulin signaling. Cell *125*, 733–747.

Koch, C.A., Anderson, D., Moran, M.F., Ellis, C., and Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science *252*, 668–674.

Koyama, S., Yu, H., Dalgarno, D.C., Shin, T.B., Zydowsky, L.D., and Schreiber, S.L. (1993). Structure of the PI3K SH3 domain and analysis of the SH3 family. Cell *72*, 945–952.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946–950.

Kuriyan, J. and Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling. Annu. Rev. Biophys. Biomol. Struct. *26*, 259–288.

Larson, S.M. and Davidson, A.R. (2000). The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. Protein Sci. *9*, 2170–2180.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. Jnt. CCP4/ESF-EAMCB Newslett. Protein Crystallogr. *26*.

Liang, J., Chen, J.K., Schreiber, S.T., and Clardy, J. (1996). Crystal structure of P13K SH3 domain at 2 Å resolution. J. Mol. Biol. *257*, 632–643.

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

Livnah, O., Stura, E.A., Johnson, D.L., Middleton, S.A., Mulcahy, L.S., Wrighton, N.C., Dower, W.J., Jolliffe, L.K., and Wilson, I.A. (1996). Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 Å. Science *273*, 464–471.

Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., and Ferrin, T.E. (2006). Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinformatics *7*, 339.

Merritt, E.A. and Bacon, D.J. (1997). Raster3D photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Musacchio, A. (2002). How SH3 domains recognize proline. Adv. Protein Chem. *61*, 211–268.

Musacchio, A., Gibson, T., Lehto, V.P., and Saraste, M. (1992). SH3: an abundant protein domain in search of a function. FEBS Lett. *307*, 55–61.

Otsu, M., Hiles, I., Gout, I., Fry, M.J., Ruiz-Larrea, F., Panayotou, G., Thompson, A., Dhand, R., Hsuan, J., Totty, N., et al. (1991). Characterization of two 85 kD proteins that associate with receptor tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase. Cell *65*, 91–104.

Pawson, T. and Gish, G.D. (1992). SH2 and SH3 domains: from structure to function. Cell *71*, 359–362.

Pawson, T. and Schlessinger, J. (1993). SH2 and SH3 domains. Curr. Biol. *3*, 434–442.

Rodrigues, G.A., Park, M., and Schlessinger, J. (1997). Activation of the JNK pathway is essential for transformation by the Met oncogene. EMBO J. *16*, 2634–2645.

Schmidt, H., Hoffmann, S., Tran, T., Stoldt, M., Stangler, T., Wiesehan, K., and Willbold, D. (2007). Solution structure of a Hck SH3 domain ligand complex reveals novel interaction modes. J. Mol. Biol. *365*, 1517–1532.

Skolnik, E.Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A., and Schlessinger, J. (1991). Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell *65*, 83–90.

Stangler, T., Tran, T., Hoffmann, S., Schmidt, H., Jonas, E., and Willbold, D. (2007). Competitive displacement of full-length HIV-1 Nef from the Hck SH3 domain by a high-affinity artificial peptide. Biol. Chem. *388*, 611–615.

Stout, T.J., Foster, P.G., and Matthews, D.J. (2004). High-throughput structural biology in drug discovery: protein kinases. Curr. Pharm. Des. *10*, 1069–1082.

Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S.A., and Draetta, G. (1993). Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. EMBO J. *12*, 2625–2634.

Tran, T., Hoffmann, S., Wiesehan, K., Jonas, E., Luge, C., Aladag, A., and Willbold, D. (2005). Insights into human Lck SH3 domain binding specificity: different binding modes of artificial and native ligands. Biochemistry *44*, 15042–15052.

Weng, Z., Thomas, S.M., Rickles, R.J., Taylor, J.A., Brauer, A.W., Seidel-Dugan, C., Michael, W.M., Dreyfuss, G., and Brugge, J.S. (1994). Identification of Src, Fyn, and Lyn SH3-binding proteins: implications for a function of SH3 domains. Mol. Cell. Biol. *14*, 4509–4521.

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

Received July 28, 2009; accepted October 5, 2009