Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 339 (2006) 407–414
BBRC
www.elsevier.com/locate/ybbrc

# Crystal structure of the N-terminal SH3 domain of mouse βPIX, p21-activated kinase-interacting exchange factor

Xiaofeng Li [a], Xueqi Liu [a,b], Fei Sun [a], Jia Gao [a], Hongwei Zhou [a], George F. Gao [c], Mark Bartlam [a,b], Zihe Rao [a,b,*]

[a] *Laboratory of Structural Biology, Tsinghua University, Beijing 100084, China*
[b] *National Laboratory of Biomacromolecules, Institutes of Biophysics, Chinese Academy of Sciences, Beijing 100101, China*
[c] *Center for Molecular Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, China*

Received 20 October 2005
Available online 14 November 2005

**Abstract**

The mouse βPIX–SH3 domain, residues 8–63 of P21-activated kinase interacting exchange factor, has been characterized by X-ray diffraction. Crystals belonging to space group $P3_221$ diffracted to 2.0 Å and the structure was phased by the single-wavelength anomalous diffraction method. The domain is a compact β-barrel with an overall conformation similar to the general SH3 structure. The X-ray structure shows mouse βPIX–SH3 domain binding the way in which the βPIX characteristic amino acids do so for an unconventional ligand binding surface. This arrangement provides a rationale for the unusual ligand recognition motif exhibited by mouse βPIX–SH3 domain. Comparison with another SH3/peptide complex shows that the recognition mode of the mouse βPIX–SH3 domain should be very similar to the RXXK ligand binding mode. The unique large and planar hydrophobic pocket may contribute to the promiscuity of βPIX–SH3 domain resulting in its multiple biological functions.


*Keywords:* βPIX; SH3 domain; Crystal structure; Ligand; Binding pocket

Pak-interacting exchange factor (PIX) was first cloned as p85SPR (SH3 domain-containing proline-rich protein) [1] and then identified as a putative guanine nucleotide exchange factor (GEF) for Rac/Cdc42 by several groups. Recently, the PIX protein has been identified as a new PAK (p21-activated protein kinase) binding partner [2]. The PIX family consists of two members, αPIX and βPIX, which are different gene products. Interestingly, mouse βPIX appears to have multiple splicing variants, designated βPIX-a, -b, -c, and -d [3]. The variants of βPIX isoforms exhibit different tissue distributions. βPIX-a is ubiquitously expressed in all the tissues examined, whereas other isoforms are expressed mainly in the central nervous system. βPIX exhibits a GEF activity toward both Rac1 and Cdc42 in vitro and in vivo [2]. It has also been revealed that p85 βPIX is phosphorylated in response to βFGF, and its upstream pathway is defined by the components Ras/Raf-1/MEK/ERK/PAK2 [4]. PAK2-dependent phosphorylation of p85 βPIX is a signal for targeting of the p85 βPIX-PAK2 complex to the lamellipodia at growth cones, where PAK2 regulates reorganization of the actin cytoskeleton for βFGF-induced neurite extension in PC12 cells [4]. The biological function and regulation of βPIX seem to be more complex than originally thought. βPIX forms a complex with various proteins: PAK [2], Rac/Cdc42 [5,6] GIT [7,8], and Shank [9]. Through these interactions, βPIX is involved in cell adhesion and migration, neurite outgrowth [4,10], and synapse formation [11].

βPIX proteins consist of several protein modules that could contribute to protein–protein and protein–lipid interactions, thus enabling modulation of multiple signaling pathways. All isoforms share SH3, Dbl homology (DH),

* Corresponding author. Fax: +86 10 6277 3145.
*E-mail address:* raozh@xtal.tsinghua.edu.cn (Z. Rao).

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.10.212

and pleckstrin homology (PH) domains at their N terminal. The other PIX domains, a GIT1-binding (GB) domain and a leucine zipper domain, are tandemly arranged at the C terminus [2]. The leucine zipper domain at the C terminus was recently shown to mediate the homodimerization of PIX, which is essential for several features of cytoskeletal reorganization, such as membrane ruffles [12]. αPIX has an additional calponin homology domain at the N terminus compared with βPIX. The SH3 domain is the most conserved domain from comparison of rat, mouse, and human α- and βPIX and putative *Drosophila* and *Caenorhabditis elegans* PIX sequences.

For the typical SH3 domain, the majority of known ligands contain a conserved PxxP motif preceded or followed by a positive residue such as Arg, which adopts a polyproline II (PPII) helix conformation upon binding and can bind in two opposite orientations, class I ligand (RxxPxxP) in the 'plus' orientation and class II ligand (PxxPxR) in the opposite 'minus' orientation [13–15]. It has, however, become apparent that some SH3 domains recognize motifs that do not contain the PXXP motif [16–18]. PX(V/ I)(D/N)RXXKP derived from a deubiquitinating enzyme UBPY is a binding partner of the SH3 domain of mouse signal transducing adaptor molecule-2(STAM2) [19,20]. Furthermore, Gads (a Grb2-related protein) SH3(C) interacts with the novel motif, PX(V/I)(D/N)RXXKP, of SLP-76 very strongly [21]. The solved structure of Gads SH3(C) in complex with the novel motif revealed that, in addition to hydrophobic interactions with the (Val/Ile) residue of the motif, electrostatic interactions with the RXXK region enable this strong binding [22]. As previously revealed, the typical ligand sequence is absent in the αPAK ligand and the residues of αPAK182–203 (DATPPPVIAPRPEHTKSIYTRS) are the binding motif for βPIX–SH3 domain [2], which contains no PxxP motif characteristic of conventional SH3-binding sequences. Although there have been numerous sequences lacking the classical PxxP motif were reported to bind to SH3 domains, none of them have any similarity to this protein.

In this paper, we report the crystal structure of the SH3 domain from mouse βPIX at high resolution and reveal an unconventional ligand-binding mode based on structural analysis. The NMR structure of the human βPIX–SH3 domain complex with a peptide derived from αPAK was recently determined [23]. In this paper, we characterize the elaborate high resolution three-dimensional structure of mouse βPIX–SH3 domain using X-ray crystallography. Based on the X-ray structure and comparison with the those determined SH3 domain structures, we reveal that, in its unusual pocket, the βPIX–SH3 domain has a unique Trp residue with its indole ring extending parallel to the ligand binding surface, which may explain the promiscuity of this domain implying the multiple function in the signal transduction pathway. The unconventional charged binding pocket accommodating the unique second part of the non-PXXP ligand, which may have a similar recognition mode to RXXK, is also discussed.

## Materials and methods

*Crystallization and data collection of the βPIX SH3 domain.* The SH3 domain was cloned into the pGEX 6p-1 vector and expressed in *Escherichia coli* strain BL21. After purification, the protein was concentrated to 10 mg/ml. Crystallization trials were set up at 18 °C using the vapor diffusion method, with the drops consisting of a 1:1 mixture of 1 mM protein solution and reservoir solution (0.1 M Hepes–Na and 1.45 M lithium sulfate, pH 7.0). Crystals suitable for diffraction were obtained after one week. The first native data set at 2.5 Å resolution of PIX SH3 domain was collected at 100 K in-house using a Mar345 detector and Rigaku RU2000 rotating Cu-Kα anode X-ray source, with 100% paraffin oil used as cryoprotectant during data collection. Data were reduced and scaled by the program HKL2000 [24]. The crystals contain two molecules per asymmetric unit. The crystal had unit cell dimensions $a = b = 89.253$ Å, $c = 40.313$ Å with space group $P3_121$ or $P3_221$ (Table 1).

*Heavy-atom derivative search and phase refinement of the βPIX SH3 domain.* Extensive attempts to use molecular replacement with either the X-ray structure of Grb2 SH3 domain (1GCQ) [25] or the crystal structure of the Sem5 SH3 domain (1SEM) [15] as search models failed to yield any solution. To overcome the phase problem, heavy-atom derivatives were searched by soaking crystals in mother liquor (1.6–1.8 M lithium sulfate, 1.0 M Hepes–Na, pH 7.0 or pH 7.5) containing various concentrations of heavy-atom compounds. Soaking times varied from 15 min to 1 day. The soaked crystals were used to collect diffraction data, and each data set was analyzed using the programs CNS [26] and SOLVE [27] trying MIR (multiple isomorphous replacement) and SAD (single anomalous diffraction) protocols. After extensive trials with many heavy-atom compounds, the final derivative was from 1 mM $HgCl_2$ soaking for 16 h, which could diffract to 2.0 Å. The phases were determined by the single-wavelength anomalous diffraction technique using program SOLVE. Six independent heavy atom sites with high occupancy were found yielding an overall figure of merit (FOM) of 0.28 after calculation of initial SAD phases to 2.0 Å. The phases were further improved by density modification and an

Table 1
Crystallographic data collection and refinement statistics

| *Data collection* | |
|---|---|
| Space group | $P3_221$ |
| Unit cell | |
| a (Å) | 88.3 |
| b (Å) | 88.3 |
| c (Å) | 40.5 |
| αβγ (°) | $\alpha = 90$, $\beta = 90$, $\gamma = 120$ |
| Resolution (Å) | 50–2.0 (2.08–2.01) |
| Completeness (%) | 98.5 (99.8) |
| Reflections | |
| No. of measured reflections | 123,226 |
| No. of unique reflections | 12,181 |
| Redundancy | 10.1 (8.9) |
| [a] $R_{merge}$ (%) | 10.3 (58.2) |
| $I/\sigma(I)$ | 19.7 (2.8) |
| *Refinement statistics* | |
| Resolution (Å) | 2.0 |
| [b] *R*-factor (%) | |
| Working set | 23.24 |
| Test set | 25.61 |
| Rms deviation | |
| Bonds (Å) | 0.007 |
| Angles (°) | 1.20 |

[a] $R_{merge} = \sum_h \sum_i |I_{ih} - \langle I_h \rangle| / \sum_h \sum_I \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of the observations $I_{ih}$ of reflection.

[b] $R_{work} = \sum(||F_{obs}| - F_{calc}||)/\sum|F_{obs}|$; $R_{free}$ is the *R*-factor for a subset (10%) of reflections that was selected prior refinement calculations and not included in the refinement.

initial experimental electron density map was calculated by RESOLVE [28]. The space group ambiguity ($P3_121$ or $P3_221$) was resolved by calculating two maps. Only the map calculated in space group $P3_221$ had clear molecular boundaries and idenitifiable secondary structure, indicating it as the correct space group.

*Model building and structure refinement.* A preliminary discontinuous polyalanine model was built automatically by RESOLVE. The high quality electron density allowed tracing of the main-chain atoms of most residues with the exception of residue 1 of molecule A and residues 1–4 of molecule B. Side chains were built into the sequence for most residues based on the RESOLVE map, and manual adjustment using O [29] quickly reduced the *R*-factor to 35 % while maintaining reasonable stereochemistry and rigid-body refinement. A 10% random data set was excluded from further refinement at this stage so that an $R_{free}$ could be used to monitor for the further refinement. Positional refinement and energy minimization were alternated with re-examination of newly constructed $2F_o - F_c$ or $F_o - F_c$ maps in O for model adjustment. Restrained individual temperature factor refinement was included at 2.0 Å resolution and, in the final stages of refinement, 111 ordered water molecules were assigned to peaks higher than 2.5σ in the $F_o - F_c$ difference electron density map. The quality of the refined PIX SH3 domain can be assessed from the statistics given in Table 1. The final model of PIX SH3 domain, consisting of residues 2–61 in molecule A and residues 5–61 in molecule B, has an $R_{work}$ of 23.2% and an $R_{free}$ of 25.39%. Atomic coordinates have been deposited into the Protein Data Bank with the accession number 2ESW.

## Results and discussion

### *Overall structure of the SH3 domain*

The structure of the βPIX–SH3 domain shares the typical SH3 domain fold with other SH3 domain structures determined to date [13,14]. It consists of five β-strands (residues 8–11, 30–36, 41–46, 49–54, and 58–60) and a $3_{10}$-helical segment from residues 55 to 57, arranged into two orthogonal anti-parallel sheets. As shown in Fig. 3, the overall conformation resembles a compact anti-parallel β-barrel with a Greek key motif. One of the sheets is formed by βa, βe and the first half of βb, while the other is formed by βc, βd, and the second part of βb. There are two molecules constructing a close dimer in one asymmetric unit.

Several hydrophobic residues from these β-strands, namely Val9, Ala11, Leu23, Phe25, Ile31, Val33, Trp42, Gly44, Phe53, and Val58, pack against each other to form the hydrophobic core of the domain which is critical for protein stabilization [30]. These hydrophobic core residues are predominantly inaccessible to solvent due to their close packing in the interior of the protein. Furthermore, these residues are highly conserved and are mostly occupied by hydrophobic residues (Fig. 1).

Strands βa and βb are separated by the long RT loop consisting of 18 residues (Lys12 to Asp29), which has a long hairpin-like structure. The RT loop flips on the β-barrel and is approximately orthogonal to the central axis of the barrel with a left-handed twist. The loop is stabilized by extensive intra-loop hydrogen bonds as well as those between the loop and rest of the protein (Table 2). The conformation of the loop is further stabilized by hydrophobic interactions involving residues Phe13 and Phe15, and the upper part of the β-barrel hydrophobic core. Conformational rigidity of the loop should play a key role in the function of the SH3 domain. Indeed, several of the most highly conserved residues important for the binding of the peptide ligand are located on the RT loop.

The shorter n-Src loop consisting of four residues (Glu37 to Gly40) is located between strands βb and βc. It is stabilized in both subunits by several hydrogen bonds.

Table 2
Intra-loop hydrogen bonds involving residues on the RT loop in the crystal structure of mouse βPIX SH3 domain

| Donor | Distance (Å) | | |
|---|---|---|---|
| | Acceptor | Chain A | Chain B |
| NZ-K27 | O-K12 | 3.33 | — |
| N-K27 | OD1-D14 | 2.96 | 2.76 |
| N-F15 | O-F25 | 2.78 | 2.88 |
| N-F25 | O-F15 | 2.94 | 2.83 |
| N-Q17 | OG-S24 | 3.11 | — |
| N-T18 | OE1-E22 | 2.69 | 2.94 |
| N-N19 | OE1-E22 | 3.21 | 3.54 |
| N-E22 | O-N19 | 3.24 | 3.09 |
| N-D21 | OE1-E20 | 3.16 | — |
| N-D29 | O-S26 | 2.95 | 2.93 |
| N-S26 | OD2-D29 | 2.70 | 2.75 |
| *Loop-protein core hydrogen bonds* | | | |
| N-W52 | O-D21 | 3.0 | 3.14 |
| N-L23 | O-W52 | 2.71 | 2.71 |
| NH2-R49 | O-S24 | 2.49 | — |

A hydrogen bond is formed when the distance between the donor and the acceptor atoms is less than 3.4 Å.

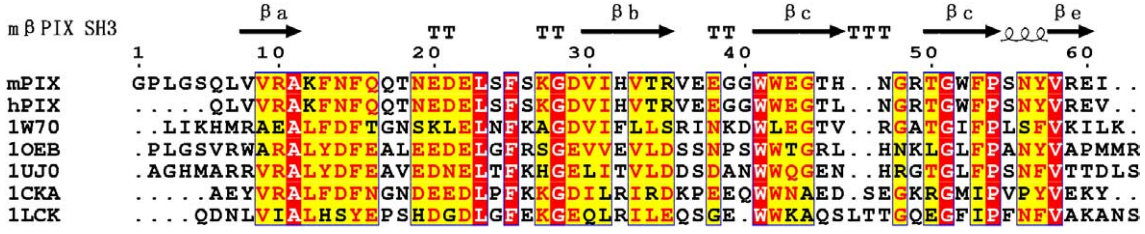


Fig. 1. Sequence alignment of mouse βPIX–SH3 domain with human βPIX–SH3 domain and the other SH3 domain revealed that it has the common conserved residues corresponding to peptide recognition and structural fold. βa–βe represent β-strands. T represents a β-turn. The helix represents a $3_{10}$ helix. Identical residues are highlighted in red and the most conserved residues are highlighted in yellow. Sequences were aligned using CLUSTALW [35] and the figure was produced by ESPript [36]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

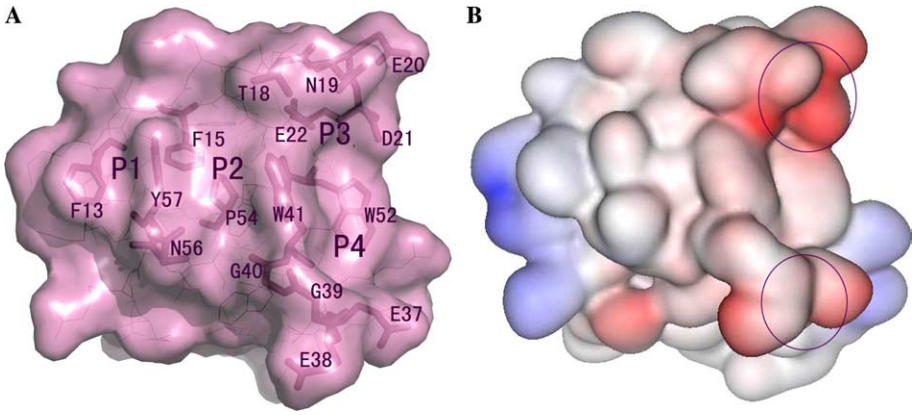


Fig. 2. The solvent accessible surface of the crystal structure of βPIX SH3 involving ligand recognition. (A) The ligand binding pocket P1, P2, P3, and P4. The residues involving ligand recognition are showed as cleavated shape. (B) The electronic-potential surface of mouse βPIX SH3. The upper circle indicates the specific negative electronic P3 recognizing the conserved basic residue in the ligand. The lower circle indicates the unique negative electronic surface at the skirt of the additional P4. The figures were produced by PyMOL. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

For example, the carbonyl oxygen of Gly40 makes a hydrogen bond to the nitrogen of Ser55, while Glu37 can be stabilized via its nitrogen atom by the carbonyl oxygen of Trp41. The n-Src loop is more independent of the core region of SH3 domain and has fewer interactions with the rest of the domain than the RT loop. Some residues are close to the ligand binding area, while the side chains of charged residues such as Glu37 and Glu38 point to the solvent and create a clustered region of negative charge.

Interestingly, there are five β-turns in the overall structure, which includes almost all types of β-bend (Fig. 3). All of them are located in the loop region of the mouse βPIX SH3 domain. Four residues on the tip of the RT loop (Asn19, Glu20, Asp21, and Glu22) form a type I β-turn, which directs the cluster of three acidic residues into the ligand-binding pocket. The four residues N-terminal to βb (Ser26, Lys27, Gly28, and Asp29) form a type II β-turn. The whole n-Src loop (Glu37, Gly38, and Gly39) is just a classic type II β-turn. The distal loop (Asn47 to Gly48) links strands βc and βd to create an unusual type I β-turn. Finally, the three residues (Ser55, Asn56, and Tyr57) separating strands βd and βe are in a $3_{10}$-helical conformation.

Interaction between the two molecules in the dimer may help to fix the loop and immobilize those hydrophilic residues in the asymmetric dimer. Several hydrophobic residues in the interface of the dimer from both molecules pack together via hydrophobic interactions, including Val8, Val30, and Ile61 in molecular A which face the side chains of Trp 42 and Val36 in molecular B. There are four tight hydrogen bonds between the two molecules. Atom NH1 of Arg10 in molecular A makes hydrogen bonds to the OE1 and OE2 atoms of Glu60 in molecular B with

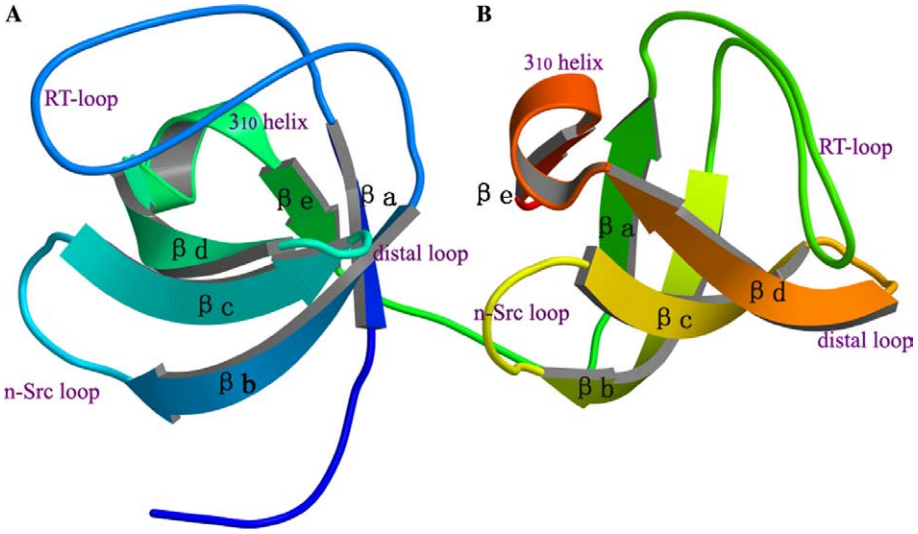


Fig. 3. Crystal structure of the dimer of mouse βPIX–SH3 domain, marked as A and B, respectively. There are two β-sheets composing of βa, βb, βc, βd, βe, forming a compact β-barrel in one monomer. One of the sheets is formed by βa, βe and the first half of βb, while the other one is constructed by βc, βd and the second part of βb. Those four specific loops, RT-loop, n-Src loop, distal loop, and $3_{10}$-helix, are also labeled painting with purple. This figure and the following figures were drawn using Molscript, Bobscript, and Raster 3D [37]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

respective distances of 2.6 and 2.9 Å. The other NH atom of Arg10 makes hydrogen bonds to the oxygen atoms of Ser55 and Val58 in molecular B with respective distances of 2.3 and 2.8 Å (Fig. 5).

As the crystal is a heavy atom derivative, six Hg ions are found in one asymmetric unit. It is interesting to note that all Hg ions bind at the outer surface of the dimer and form contacts with the symmetry-related molecules, thus contributing to increased compactness of the crystal packing. Two His46 residues in the dimer are the main metal binding residues in co-operation with the two Trp41 residues and one Arg49 residue in molecular A. Metal co-ordination sites in proteins can play important functional modulation roles. $Zn^{2+}$ is usually chelated by the Nε2 nitrogen atom of a histidine residue, which is somewhat similar to the binding mode observed in this structure. This suggests that His46 which is absent in human βPIX may be a potential metal binding site for important modulatory ions, such as calcium and zinc, in the mouse βPIX functional pathway. From this, we hypothesize that mouse βPIX–SH3 may have potential metal coordination sites that are involved in regulation of the signaling pathway.

The NMR structure of human βPIX–SH3/PAK peptide complex has been determined recently [23]. Overall, the conformation of the mouse βPIX–SH3 domain is similar to the NMR structure of the βPIX–SH3 complex, with the largest differences in the RT-loop, n-Src loop, and the distal loop (Fig. 4A). The rms deviation between the $C_\alpha$ atoms of the representative model of the human βPIX–SH3 domain and the crystal structure of mouse SH3 domain is 1.8 Å from residues 6 to residue 61.

Conformational differences for residues involved in ligand binding are also observed, especially those located in the P3 and P4 pockets. As shown in Fig. 4B, the side chain of Glu20 at the edge of the specificity pocket points into the ligand binding pocket, while accordingly Glu22 in the NMR structure points out of the pocket and makes no contact with the ligand. This acidic residue together with Asp21 (corresponding to Asp23 in the NMR structure) forms a large negatively charged surface. Thr18 and Glu22 are much closer to each other and may form an ionic interaction with the basic side chain of the conserved Arg in the ligand. In the NMR structure, the amine group of Glu39 (corresponding to Glu37 in the crystal structure) makes a salt bridge with Lys198 of the ligand [23]. Glu37 in the crystal structure makes a very different torsion angle in the side chain and the tip of the chain pointing out of the pocket. This acidic residue, together with Glu38, forms a negative patch at the edge of the hydrophobic fourth pocket and may make that salt bridge to enhance the specificity of the ligand binding.

In the distal loop, the higher temperature factor of the crystal structure in this region and the large rms derivation between the crystal structure and NMR structure reveal that these residues (Asn47, Gly48, and Arg49) are extremely flexible.

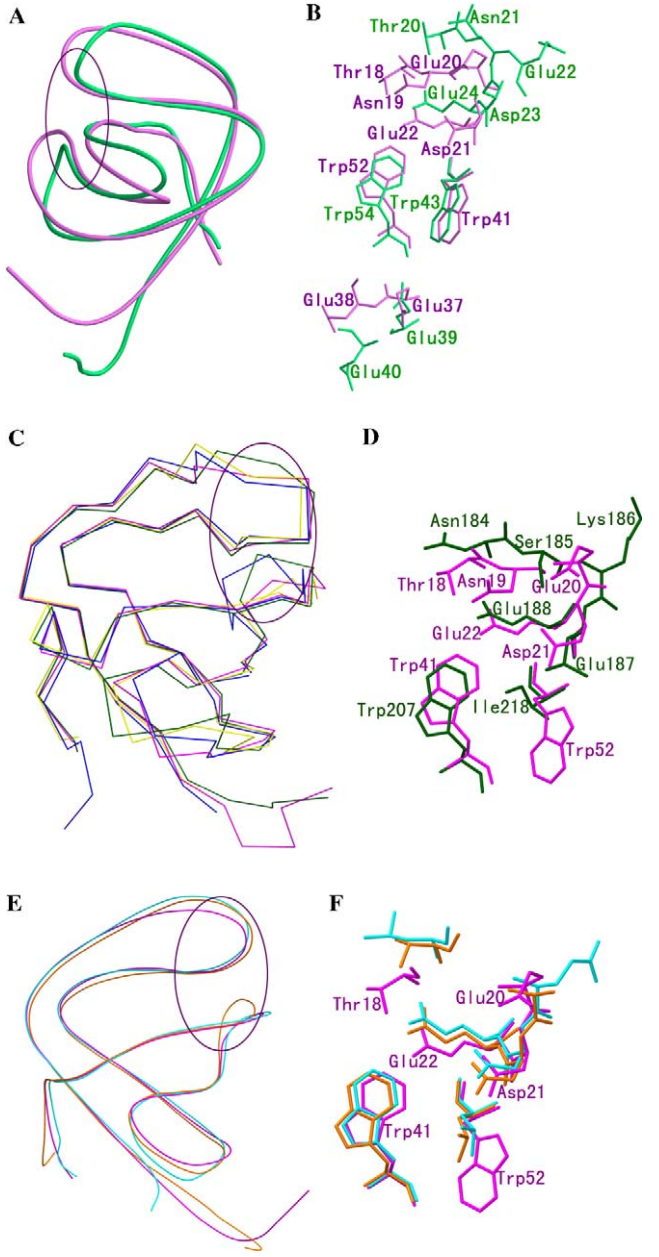


Fig. 4. Comparison with the other SH3 domains and the NMR structure. All residues involved in ligand binding are shown as ball-and-stick. In all comparisons, mouse βPIX SH3 domain is colored magenta. (A,B) The superposition of the crystal structure of mouse βPIX SH3 domain to the NMR structure of human βPIX–SH3. The NMR structure coordinates are taken from the first model of PDB entry 1ZSG and shown in green. (C,D) The superposition of the mouse βPIX–SH3 domain on P40phox SH3 (1W70) [31], Crk SH3N (1CKA) [32], and P56-Lck (1LCK)[33]. The P40phox SH3 domain is shown in green. Crk SH3N is shown in yellow. P56-Lck SH3 domain is shown in blue. D is a zoomed view the superposition of the specificity ligand binding pocket shown by the circle marked on C. (E,F) The superposition of the mouse βPIX SH3 domain on the Gads SH3 domain (1OEB) [22] and STAM2 SH3 domain (1UJ0) [20]. The Gads SH3 domain is colored orange and the STAM2 SH3 domain is colored cyan. E is a zoomed view of the superposition of the specificity ligand binding pocket shown by the circle marked on F. The conserved acidic residues, Asp21 and Glu22, with similar conformation suggest that they might have a similar unconventional RXXK ligand recognition mode in pocket3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

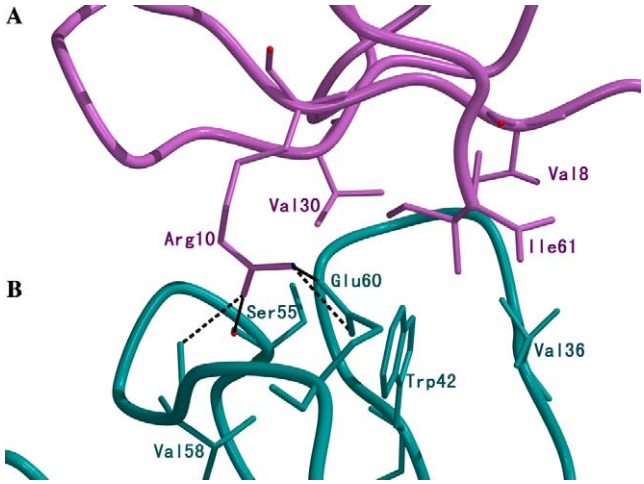


Fig. 5. The interaction between the two monomers in the dimer of mouse βPIX–SH3 domain. Hydrogen bonds are denoted by dashed lines. Molecule A is shown in magenta and molecule B is shown in cyan. The residues are labeled with the same color with the monomer. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

*The ligand binding pocket*

In the crystal structure of the mouse βPIX–SH3 domain, the ligand binding region is a hydrophobic patch on the surface surrounded by several charged residues. As shown in Fig. 2A, the side chains of highly conserved residues, such as Phe13, Phe15, Phe25, Trp41, Phe54, and Tyr57, form three well-defined clefts labeled Pocket 1 (or P1) Pocket 2 (or P2) and Pocket 3 (or P3). These three clefts are arranged linearly from left to right and lie approximately parallel to the direction of the hairpin loop. It is interesting to note that the βPIX–SH3 domain has an atypical Pocket 4 (or P4) flanking P3 and lying approximately at right angles to the conventional PXXP binding cleft. P1 is composed of Phe13 on one side and Tyr57 on the other. P2 is constructed by Pro54 on the bottom and Tyr57, Phe15, Trp41, Asn56, and Gly40 side chains on the sides. It can be further divided into two subsites, one formed by residues Tyr57, Phe15, Asn56, and Pro54 and the other formed by Trp41, Gly40, and Pro54. P3 is formed by Thr18, Asn19, Glu20, Asp21, Glu22, and Trp41. This so-called specificity pocket is formed by a cluster of negatively charged residues and represents a potential binding side for the second part of the ligand, which is rich in positively charged residues. The unusual P4 pocket is formed by the hydrophobic residues Thr41, Thr52, and Gly39 and the acidic residue Glu37 in the type II β-turn. It is notable that the ligand binding interface of the βPIX–SH3 domain is remarkably large, which may be the basis for the high affinity or the high specificity of the protein-ligand interaction. In the n-Src loop, the type II β-turn pulls the main chain atoms of Glu38 and Gly39 away from the center of P4, which confers it with a larger space to accommodate the longer recognition motif of the ligand (Fig. 2A).

The free energy of binding between proteins includes an electrostatic and a hydrophobic component in the interaction surface. Displaying the electrostatic potential and hydrophobicity at the surface of a protein can provide information concerning the nature of its interactions with other proteins. In the βPIX–SH3 domain, a larger patch of negative electrostatic potential is located in the Pocket 3 region of the peptide binding surface, while Pocket 4 is larger and much flatter compared with other SH3 domain structures. Significantly, there is another negative electrostatic potential region on the distal skirt of the Pocket 4, which might reinforce the binding affinity of the long ligand covering these pockets (Fig. 2B). These properties of the interface form the bedrock of the unconventional ligand recognition specificity of the βPIX–SH3 domain.

*Comparison with other SH3 domains*

The crystal and solution structures of several SH3 domains have established the general SH3 folding and polypeptide binding pattern. Although the sequence identities between these SH3 domains are relatively low (Fig. 1), they all share an overall topology of two small β-sheets packed against each other at approximately right angles, a C-terminal $3_{10}$ helix and three variable loops (a long hairpin RT loop, n-Src loop, and distal loop). The characteristic tertiary fold, namely the β-barrel, results in many of the conserved residues adopting nearly identical side-chain conformations. More interestingly, all of the highly conserved residues (Phe13, Phe15, Trp41, Pro54, and Tyr57) are involved in forming the solvent-exposed hydrophobic ligand-binding surface. Therefore, construction of the peptide-binding pockets for various SH3 domains are highly similar: two hydrophobic clefts (P1 and P2) are composed of side-chains of highly conserved aromatic residues flanked by the specificity pocket (P3) formed by charged residues. In general, the affinity of SH3 domains for their peptide ligands is quite low and selectivity is generally marginal, resulting in a promiscuous ligand binding mode.

*Comparison with similar structures*

The numerous entries within the Protein Data Bank for SH3 or SH3-like domains allow extensive comparisons with the βPIX–SH3 domain. Adjacent structural neighbors were identified from the DALI server and analyzed further using superposition. All homologous structures displayed rms deviations of 0.8–1.1 Å with the βPIX–SH3 domain structure. The structures that most closely resemble the βPIX–SH3 domain are P40phox SH3 (1W70) [31], Crk SH3N (1CKA) [32], and P56-Lck (1LCK) [33]. Of these structures, P40phox SH3 (1W70) and Crk SH3N (1CKA) were in complex with peptides. In both complex structures the peptides, although representing different motifs suited

for their respective binding partners, are bound in the 'minus' orientation. The binding ligands both have a positive residue in the same specificity pocket, namely arginine in the P40phox SH3/P47phox complex and lysine in the Crk SH3N/C3G peptide complex. Here we compare the βPIX–SH3 domain to its closest neighbor, P40phox SH3. As shown in Figs. 4C and D, the conserved residue Glu22, residue Thr18 (instead of Asn184) and Asn19 (replacing Ser185) form the conserved specificity pocket. While located at the edge of this pocket, the side chains of Glu20 (Lys186) and Asp21 (Leu187) point into the center of the pocket to make a much larger electronegative ligand binding surface. Significantly, Trp41 replaces Ile218 in the corresponding position and the indole ring extends smoothly to the surface of the pocket to form a broader and flatter hydrophobic binding surface. Recently it has been shown that the Cbl protein, with no obvious homology to the PAK ligand, can also bind to the βPIX–SH3 domain [34]. This kind of large and planar hydrophobic solvent-exposed surface may account for the binding promiscuity of the βPIX–SH3 domain.

*Comparison with other SH3 structures with unconventional RXXK recognition mode*

Comparison between those SH3 domains binding a long peptide with higher specificity or affinity via four binding pockets on the SH3 surface reveals the characteristic properties of this kind of recognition mode. The Gads SH3 domain (1OEB) [22] and STAM2 SH3 domain (1UJ0) [20] both have the P4 pocket flanking the specificity pocket to recognize the unique RXXK motif in the ligand, thus explaining their higher affinity and specificity for the ligands. Both of the ligands are atypical non-PXXP and form a helical conformation different from the conventional PPII helix. Gads SH3 and STAM2 SH3 can be superimposed onto the structure of the βPIX SH3 domain, with rms deviation values for all $C_\alpha$ atoms of 1.4 and 1.3 Å, respectively (Fig. 4E). We compare the crystal structure of βPIX–SH3 with that of Gads SH3. The rms deviation of the $C_\alpha$ atoms between Gads SH3 domain and crystal structure of the βPIX–SH3 domain is 1.4 Å from residue 6 to residue 61. As shown in Figs. 4E and F, superposition of the two structures reveals that the βPIX–SH3 domain possibly has the same RXXK ligand recognition mode. In the conventional PXXP binding site, the conformations of the conserved hydrophobic residues in the two structures resemble each other. In the P3 pocket, the conserved Glu20 residue (Glu15 in 1OEB), Asp21 (Asp16), and Glu22 (Glu17) assemble together to form a large, highly negative electrostatic potential surface defining the interface for the basic residues in the binding ligand such as residues Arg7 and Lys10 in the Gads SH3/SLP76 complex. The effect of the electrostatic potential surface can be made by a conformational change of Glu20, resulting in its involvement in ligand interactions. Thr18 (replacing Leu13 in the Gads SH3 structure) can potentially enhance the interaction with the basic residue of the ligand via hydrogen bonding. In the P4 pocket, the conserved Trp41 has a similar conformation to Trp36 in Gads SH3, while Trp52 (replacing Leu47) and its outward facing side chain form a large and planar hydrophobic binding pocket. The differences in this region may endow βPIX–SH3 with the specificity to accommodate more hydrophobic residues in a longer ligand, with a lower specificity or affinity compared to the Gads SH3 domain.

SH3–ligand interactions are usually weak, with ordinary $K_d$ of ~10 μM. Gads SH3 domain recognizes the novel RXXK motif from SLP-76 with a tight and specific association ($K_d = 0.12$ μM) [22], which is the highest affinity measured in solution for an SH3 binding peptide of this length. However, the interaction between the STAM2 SH3 domain and the UBPY-derived peptide is much weaker, revealing that this novel motif does not necessarily function to enhance binding to the SH3 domain [20]. This result is consistent with the βPIX–SH3/PAK peptide interaction. More recently, Mott and co-workers reported that the free βPIX–SH3 binds to the PAK183-204 derived ligand with an average $K_d = 7.5$ μM, which was much lower than before [2]. The study by Mott and co-workers also revealed that the βPIX–SH3 domain may recognize the novel motif of PAK with a moderate affinity as the other SH3-ligand affinity. This property can be approved by the analysis based on the atom resolution crystal structure of βPIX–SH3 domain.

## Conclusions

The high resolution crystal structure of the mouse βPIX–SH3 domain provides a picture of this important SH3 domain at an atomic level. The unusual conformation of the side chain of Trp41 in the P4 pocket makes a broad and planar hydrophobic binding pocket flanking the specificity pocket which can account for the promiscuity of this domain. Comparison with the NMR structure and with other SH3–peptide complexes reveals that in the P3 pocket, the recognition mode of the βPIX–SH3 domain is very similar to the unconventional RXXK mode. The clear three-dimensional structure of the βPIX–SH3 domain reported here should provide a structural basis for SH3-directed drug discovery for seeking effective SH3 domain inhibitors.

## Acknowledgments

This work was supported by Project "973" (Grant No. G1999075602), Project "863" (Grant No. 2002BA711A12) and the NSFC (Grant No. 30221003).

## References

[1] W.K. Oh, J.C. Yoo, D. Jo, Y.H. Song, M.G. Kim, D. Park, Biochem. Biophys. Res. Commun. 235 (1997) 794–798.
[2] E. Manser et al., Mol. Cell 1 (1998) 183–192.

[3] T. Kim, D. Park, Mol. Cells 11 (2001) 89–94.
[4] E.Y. Shin et al., J. Biol. Chem. 277 (2002) 44417–44430.
[5] B. Aghazadeh, K. Zhu, T.J. Kubiseski, G.A. Liu, T. Pawson, Y. Zheng, M.K. Rosen, Nat. Struct. Biol. 5 (1998) 107–1098.
[6] E.Y. Shin et al., J. Biol. Chem. 279 (2004) 1994–2004.
[7] Z.S. Zhao, E. Manser, T.H. Loo, L. Lim, Mol. Cell. Biol. 20 (2000) 6354–6363.
[8] H. Zhang, D.J. Webb, H. Asmussen, S. Niu, A.F. Horwitz, J. Neurosci. 25 (2005) 3379–3388.
[9] E. Park et al., J. Biol. Chem. 278 (2003) 19220–19229.
[10] C. Albertinazzi, L. Za, S. Paris, I. de Curtis, Mol. Biol. Cell 14 (2003) 1295–1307.
[11] D. Parnas, A.P. Haghighi, R.D. Fetter, S.W. Kim, C.S. Goodman, Neuron 32 (2001) 415–424.
[12] S. Kim, S.H. Lee, D. Park, J. Biol. Chem. 276 (2001) 10581–10584.
[13] H. Yu, J.K. Chen, S. Feng, D.C. Dalgarno, A.W. Brauer, S.L. Schreiber, Cell 76 (1994) 933–945.
[14] S. Feng, J.K. Chen, H. Yu, J.A. Simon, S.L. Schreiber, Science 266 (1994) 1241–1247.
[15] W.A. Lim, F.M. Richards, R.O. Fox, Nature 372 (1994) 375–379.
[16] G. Bottger, P. Barnett, A.T. Klein, A. Kragt, H.F. Tabak, B. Distel, Mol. Biol. Cell 11 (2000) 3963–3976.
[17] K. Kami, R. Takeya, H. Sumimoto, D. Kohda, EMBO J. 21 (2002) 4268–4276.
[18] J.R. Pires, X. Hong, C. Brockmann, R. Volkmer-Engert, J. Schneider-Mergener, H. Oschkinat, R. Erdmann, J. Mol. Biol. 326 (2003) 1427–1435.
[19] M. Kato, K. Miyazawa, N. Kitamura, J. Biol. Chem. 275 (2000) 37481–37487.
[20] T. Kaneko, T. Kumasaka, T. Ganbe, T. Sato, K. Miyazawa, N. Kitamura, N. Tanaka, J. Biol. Chem. 278 (2003) 48162–48168.
[21] M. Lewitzky et al., Oncogene 20 (2001) 1052–1062.
[22] M. Harkiolaki et al., EMBO J. 22 (2003) 2571–2582.
[23] H.R. Mott, D. Nietlispach, K.A. Evetts, D. Owen, Biochemistry 44 (2005) 10977–10983.
[24] W., O.Z.M. (1997) Methods in Enzymology. Macromol. Crystallog. A 276 307–326.
[25] M. Nishida, K. Nagata, Y. Hachimori, M. Horiuchi, K. Ogura, V. Mandiyan, J. Schlessinger, F. Inagaki, EMBO J. 20 (2001) 2995–3007.
[26] A.T. Brunger et al., Acta Crystallogr. D Biol. Crystallogr. 54 (Pt. 5) (1998) 905–921.
[27] T.C. Terwilliger, J. Berendzen, Acta Crystallogr. D Biol. Crystallogr. 55 (Pt. 4) (1999) 849–861.
[28] T.C. Terwilliger, Acta Crystallogr. D Biol. Crystallogr. 56 (Pt. 8) (2000) 965–972.
[29] T.A. Jones, J.Y. Zou, S.W. Cowan, Kjeldgaard, Acta Crystallogr. A 47 (2) (1991) 110–119.
[30] K.A. Dill, Biochemistry 29 (1990) 7133–7155.
[31] C. Massenet, S. Chenavas, C. Cohen-Addad, M.C. Dagher, G. Brandolin, E. Pebay-Peyroula, F. Fieschi, J. Biol. Chem. 280 (2005) 13752–13761.
[32] X. Wu, B. Knudsen, S.M. Feller, J. Zheng, A. Sali, D. Cowburn, H. Hanafusa, J. Kuriyan, Structure 3 (1995) 215–226.
[33] M.J. Eck, S.K. Atwell, S.E. Shoelson, S.C. Harrison, Nature 368 (1994) 764–769.
[34] J.A. Flanders, Q. Feng, S. Bagrodia, M.T. Laux, A. Singavarapu, R.A. Cerione, FEBS Lett. 550 (2003) 119–123.
[35] J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res. 22 (1994) 4673–4680.
[36] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, Bioinformatics 15 (1999) 305–308.
[37] R.M. Esnouf, J. Mol. Graph. Model 15 (1997) 132–134, 112–113.