JBC Papers in Press. Published on July 20, 2006 as Manuscript M604135200
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M604135200

STRUCTURAL BASIS OF ROBO PROLINE-RICH MOTIF RECOGNITION BY THE SRGAP1 SH3 DOMAIN IN THE SLIT-ROBO SIGNALING PATHWAY

Xiaofeng Li[1,2], Yushu Chen[1,2], Yiwei Liu[1], Jia Gao[1], Feng Gao[1], Mark Bartlam[1,2], Jane Y. Wu[2,3‡] & Zihe Rao[1,2‡]

[1] "Tsinghua-IBP Joint Research Group for Structural Biology", Tsinghua University, Beijing 100084, China. [2] National Laboratory of Biomacromolecules, Institute of Biophysics (IBP), Chinese Academy of Sciences, Beijing 100101, China. [3]Department of Neurology; Center of Genetic Medicine, Lurie Cancer Center; Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA.

Running title: Structure of the srGAP1 SH3 domain.

‡ Address correspondence to Jane Y. Wu, Tel: 312-503-0684, Fax: 312-503-5603, Email: jane-wu@northwestern.edu; or Zihe Rao, Tel: 86-10-62771493, Fax: 86-10-62773145, E-mail: raozh@xtal.tsinghua.edu.cn.

The srGAPs are important components in the intracellular pathway mediating Slit-Robo signaling in axon guidance and cell migration. We report the first crystal structure of the srGAP1 SH3 domain at 1.8 Å resolution. The unusual side chain conformation of the conserved Phe13 in the P1 pocket renders the ligand binding pocket shallow and narrow, which contributes towards the low binding affinity. Moreover, the opposing electrostatic charge and the hydrophobic properties of the P3 specificity pocket are consistent with the observed binding characteristics of the srGAP1 SH3 domain to its ligand. Surface plasmon resonance experiments indicate that the srGAP1 SH3 domain interacts with its natural ligand in a C to N orientation. The srGAP1 SH3 domain can bind to both the CC2 and CC3 motifs *in vitro*. The N-terminal two acidic residues in the CC3 motif recognition site are necessary for srGAP1 SH3 domain binding. A longer CC3 peptide (CC3-FL) binds with greater affinity than its shorter counterpart, suggesting that the residues surrounding the proline-rich core are important for protein-peptide interactions. Our study reveals previously unknown properties of the srGAP-Robo interaction. Our data provide a structural basis for the srGAP-Robo interaction, consistent with the role of the Robo intracellular domain in interacting with other downstream signaling molecules and mediating versatile and dynamic responses to axon guidance and cell migration cues.

In the developing central nervous system, the axon must navigate through a complex terrain consisting of various cell types, distinct neuronal processes and many extracellular matrix molecules. Within this complex terrain are guidance cues that direct the growth cone, the motile tip of the axon, to its target. Actin assembly is a key process that controls the growth and steering of the axon growth cones (1,2). The Rho family of small GTPases, which includes Rho, Rac and Cdc42, has important roles in regulating actin cytoskeletal dynamics, and has been implicated in axon guidance and cell migration (3-7). Rho proteins have low intrinsic GTPase activities which can be stimulated by GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs). Rho GAPs and Rho GEFs are themselves regulated by several

extracellular and intracellular signaling pathways (8-10).

A role for Rho GTPase-activating proteins in repulsion mediated by Roundabout (Robo) has been demonstrated in Slit signaling in neuronal migration. Robo is a cell surface receptor that is responsible for the repulsive effect of Slit. Wong and colleagues used a two-hybrid screen to identify a novel family of GAPs that interact with the intracellular domain of rat Robo1 (11) . Three members of this family were identified and named slit-robo (sr) GAP1, GAP2 and GAP3, corresponding to KIAA1304, KIAA0456 and KIAA0411, respectively. The srGAPs contain a RhoGAP domain, an SH3 domain and a Fes/CIP4 homology (FCH) domain. The proline-rich conserved cytoplasmic motif 3 (CC3) motif in Robo binds directly to the SH3 domain of the srGAP subfamily of Rho GAPs (11). In cultured mammalian cells, srGAP1 can bind to and decrease the level of active Cdc42 and RhoA, but not Rac1. Extracellular application of Slit to primary neurons increases the intracellular binding of Robo and srGAP1, and inhibits Cdc42 activity in a Robo- and srGAP-dependent manner, which emphasizes the importance of the CC3 motifs in Robo1 and the SH3 domain of srGAPs in Slit-Robo signaling pathway (11).

The intracellular region of *Drosophila* Robo contains four identifiable conserved motifs, designated CC0, CC1, CC2 and CC3 (12,13), which may interact with several downstream effectors independently or simultaneously. The proline-rich CC3 motif contains the conventional PXXP sequence generally recognized by SH3 domains. However the proline-rich sequence of the CC3 motif, ^{1476}TYTDDLPPPPVPPPAIKSP^{1493} (12), may be a mixture of class I and class II proline-rich peptides (see below). It is therefore interesting to characterize the binding orientation and exact sites in the CC3 motif for the srGAP1 SH3 domain.

Found in a wide variety of intracellular proteins in living cells, the SH3 domain is a ubiquitous protein-interaction module that contributes to many aspects of complex signaling networks (14). The SH3 peptide recognition surface includes a hydrophobic cleft that is flanked on one side by variable loops (RT and n-Src loops) that contribute to the recognition specificity and determine ligand register and orientation (14,15). The SH3 domain recognizes a proline-rich sequence of its ligand protein with a PXXP motif, which is further classified into +XφPXφP (class I) and φPXφPX+ (class II) (where φ and + are usually a hydrophobic residue and an arginine residue, respectively) (16-19). Each motif adopts a left-handed type II polyproline (PPII) helix (20), known as a collagen chain conformation, and fits the ligand-binding site on the SH3 surface. An SH3 domain may interact with several ligand proteins *in vivo*, due not only to the low ligand recognition specificity of the domain but also to the weak interactions, with K_d (dissociation constant) values of 1–102 μM in most cases (14,16,21). This results in the formation of a complicated network consisting of many SH3-containing proteins and their interaction partners (or ligand proteins).

To better understand the molecular basis of the polyproline-SH3 interactions within the Robo-srGAP complex, we report the first three-dimensional structure of the srGAP1 SH3 module at 1.8 Å resolution. This structure reveals the relevant protein-protein interaction surface and provides a rationale for the weak binding

properties exhibited by srGAPs. We have also investigated the recognition site in the intracellular domain of Robo recognized by the srGAP1 SH3 domain using BIAcore experiments.

EXPERIMENTAL PROCEDURES

Protein expression and purification

The DNA fragments encoding the srGAP SH3 domains (srGAP1 SH3, srGAP2 SH3, srGAP3 SH3) from Ala9 to Val60 were amplified by PCR using the full-length srGAP sequences (KIAA1304, KIAA0456, KIAA0411) as templates. The PCR fragments were cloned into pGEX 6p-1 (GE Healthcare). The GST-SH3 proteins were overexpressed in the *Escherichia coli* BL21(DE3) strain then purified by GST-glutathione affinity chromatography, cleaved with PreScission Protease. The recombinant srGAPs SH3 domains were stored in 10 mM Tris, pH 8.0, 100mM NaCl and concentrated to 15 mg/ml.

Crystallization

Crystals were grown by the hanging-drop vapor diffusion technique by mixing equal 1 µl volumes of protein and reservoir solutions. The best diffracting crystals of srGAP1 SH3 domain were obtained at 18 °C from a protein solution at 15 mg/ml in 10 mM Tris, pH 8.0, 100 mM NaCl, and a reservoir solution containing 1.5 M Lithium sulfate, pH 7.5. Cubic-shaped crystals grew within one week. Crystals were soaked in 100% paraffin oil (Hampton Research) as a cryoprotectant prior to X-ray diffraction experiments.

Data collection, structure determination, and refinement

Initial crystals of the srGAP1 SH3 domain diffracted to 2.5 Å in-house under cryo-conditions with a Rigaku MM-007 rotating Cu-Kα anode X-ray source and a Mar345 detector. The space group is body-centered cubic (I23) and different kinds of crystal packing were observed in one drop. Later, data from another primitive cubic crystal (space group P23) diffracting to 1.8 Å were collected in-house. Data were reduced and scaled by the program HKL2000 (22). Data collection statistics are summarized in Table 1.

The structure of the srGAP1 SH3 domain was solved by molecular replacement using the program CNS (23) with the α-spectrin SH3 domain D48G Mutant (PDB code: 1BK2) (24) as a search model. The phase problem was solved using the 2.5 Å diffraction data in space group I23, and refinement of the srGAP1 SH3 domain model was performed using CNS with the 1.8 Å diffraction data in space group P23. The asymmetric unit contains two srGAP1 SH3 domains with a V_m of 2.1 $Å^3$ Da^{-1}. The srGAP1 SH3 domain structure was refined by iterative cycles of manual corrections with O (25) and energy minimization or simulated annealing followed by B-factor refinement using CNS. The final 1.8 Å resolution structure of the srGAP1 SH3 domain, consisting of residues 6-60, has an R_{work} of 20.8% and R_{free} of 25.7%. Refinement statistics are summarized in Table 1. Atomic coordinates have been deposited into the Protein Data Bank with the accession number 2GNC.

Peptide synthesis and surface plasmon resonance (BIAcore)

All peptides used for BIAcore experiments were synthesized by Beijing Scilight Biotechnology except the CC3-I13

and CC3-II14 peptides, which were synthesized by GL Biochem (Shanghai) Ltd. A C-terminal GGK extension was added to several peptides for coupling to the CM5 chip via the lysine side chain amino group, but ultimately this scheme was substituted by SH3 domain binding to the chip for our experiments. The peptides were in freeze-dried powder and were dissolved in sterile water and stored at -80 °C.

Surface plasmon resonance (SPR) experiments were carried out on a Pharmacia Biosensor BIAcore 3000 instrument according to the manufacturer's instructions. Three srGAP SH3 domains (srGAP1 SH3, srGAP2 SH3, srGAP3 SH3) were covalently coupled to different flow cells of the CM5 chip respectively, binding to the dextran matrix of a CM5 sensorchip (Pharmacia Bio-sensor). Immobilizations were performed in 10 mM sodium acetate buffer (pH 5.0). A reference surface was generated simultaneously under the same conditions but without protein injection and used as a blank to correct for instrumental and buffer effects. The amount of protein bound to the sensor chip was monitored by the change in refractive index. Binding experiments were performed at 25 °C in HBS running buffer (20mM Hepes, pH 7.4, 1mM DTT, 150mM NaCl, 0.005% (v/v) Tween20) with a flow-rate of 60 µl/min. After each binding experiment, the sensor chip was regenerated by a 90 second pulse of 90 µl of HBS buffer at a flow rate of 60 µl/min, a treatment that did not produce significant changes in baseline and response after each run. Thus, estimation of kinetic parameters was realized by repetitive injections of a series of peptides at variable concentrations at 60 µl/min flow rate to the immobilized srGAP SH3 domains. Response curves were prepared by extracting the signal generated using the control flow cell. Analysis of experimental data and kinetic parameters were performed with BIAevalution software v4.1 (Pharmacia Biasensor).

RESULTS

Crystal structure of the srGAP1 SH3 domain at 1.8 Å resolution

We have solved the crystal structure of the srGAP1 SH3 domain, encompassing the region from residues Ala377 to Val428 of the full-length srGAP1 protein, at 1.8 Å resolution. For convenience, residues are subsequently renumbered from Ala9 to Val60 throughout the manuscript. The general topology of the srGAP1 SH3 domain is similar to that of other SH3 domains, consisting of five β-strands arranged as two orthogonal β-sheets and forming a compact anti-parallel β-barrel (16,17). As shown in Figure 1A, one of the sheets is formed by β-strands βa, βe, and the first half of βb, while the other is formed by β-strands βc, βd, and the second half of βb. A kink in β-strand b allows it to participate in both β-sheets.

Strands βa and βb are connected by a long hairpin loop (RT-loop), which flips onto the top of the β-barrel and is approximately orthogonal to the central axis of the barrel with a left-handed twist. The conformation is stabilized by hydrophobic interactions involving residues Tyr15, Leu23, Phe25 and the upper part of the β-barrel hydrophobic core. The loop is further stabilized by extensive intra-loop hydrogen bonds, as well as those between the loop and rest of the SH3 domain. The majority of residues in the RT-loop are extremely well defined in the electron density map, including hydrophilic residues

such as Asp14 and Glu22.

There are two classic β-bulges found in the βb strand, which endows this long strand with greater flexibility than in other SH3 domain structures (Figure 2). The defining features of the β-bulge in the SH3 domain are hydrogen bonds between the amide nitrogens of both Tyr34 and His35 and the carbonyl group of Lys43(26). This is a highly conserved β-bulge in the middle of the second β-strand of the SH3 domain which appears to provide a necessary kink in this strand, thus enabling it to hydrogen bond to both sheets comprising the fold. The other β-bulge is adjacent to the n-Src loop, in which the amide nitrogens of Ala37 and Ser38 both hydrogen bond to the carbonyl of the Trp41, and the NH of Trp41 is bonded to the carbonyl of Ser38. The right-handed helical angles generally observed at the Tyr34 and Ala37 amino acids are characteristic of position 1 in the classical β-bulge.

Interestingly, there are five β-turns in the overall srGAP1 SH3 domain structure, which includes almost all types of β-bend (27) (Figure 1A). All of them are located in the loop regions of the srGAP1 SH3 domain. Four residues on the tip of the RT loop (Ser19, Ala20, Arg21, and Glu22) form a type I β-turn. At the end of the RT loop, four residues (Lys26, Lys27, Gly28, and Ala29) form a type II β-turn N-terminal to strand βb. The whole n-Src loop (Ser38, Glu39, Asp40 and Trp41) is just a classic type I β-turn. The distal loop (His46 to Ile49) linking strands βc and βd is a classic type II β-turn. Finally, in the center of the ligand recognition groove, the four residues (Pro54, His55, Gln56 and Tyr57) separating strands βd and βe are in a 3_{10}-helical conformation.

There are two molecules in one asymmetric unit of the crystal structure. The two chains of the srGAP1 SH3 domain structure can be superimposed with an overall RMSD of 0.58 Å for all C_α atoms. The structures of the two srGAP1 SH3 domains are essentially identical, including side-chain orientations, with a few notable exceptions (Figure 3). The largest deviation is at position 13. In molecule B, Phe13 stacks against Tyr57 on the same side of the binding groove. This stacking of Phe13 against the conserved Tyr57 on the same side of the aromatic ligand binding groove is the same general conformation observed in almost all other SH3 structures to date. However, in molecule A, the conformation of the phenyl group of Phe13 is directed away from Tyr57 and is located on the opposite side of the binding groove where it lies close to Phe8 of a symmetry-related molecule A. This suggests that even conserved residues in the SH3 fold involved in ligand binding have plasticity and can adapt to different conditions. Another possibility may be that the conformational change of Phe13 in molecule A is a result of crystal packing.

The ligand binding region of the srGAP1 SH3 domain

In the crystal structure of the srGAP1 SH3 domain, the ligand binding area is a hydrophobic patch on the surface surrounded by several charged residues. As shown in Figure 1B, the side chains of highly conserved residues such as Phe13, Tyr15, Glu22, Trp41, Pro54 and Tyr57, Leu52, Ala37 generate three general clefts labeled pocket 1 (P1), pocket 2 (P2) and pocket 3 (P3) and an additional fourth pocket labeled pocket 4 (P4) (Figure 1B). P1 is composed of two hydrophobic residues with Phe13 on one side and Tyr57 on the other. This pocket is the most

anomalous one in the SH3 ligand binding groove with a very shallow bottom and almost no surrounding side face. P2 is constructed with Pro54 on the bottom and Tyr54, Tyr15, Trp41, Gln56 and Asp40 backbones at the sides. It can be further divided into two subsites, one formed by residues Tyr15, Trp41, Pro54 and Tyr57 and the other formed by the side chains of Asp40, Trp41 and Gln56. Among the three general clefts, P2 is the largest and deepest.

P3 is formed by the long alkyl side chain of Glu22 and the indole ring of Trp41 at one side, the long alkyl side chain of Arg21 on the other side, the side chain of Ser19 at the top and Leu52 on the bottom. Generally, the P3 pocket is large, contributing to the ligand binding affinity and specificity among the SH3 domain family. In the srGAP1 SH3 domain, there is an unusual basic arginine residue on the side wall of P3 adjacent to the conserved negative residue Glu22 (Figure 1B). Arg21 is in the i+2 position of the classic I β-turn on the tip of the RT-loop, and the dihedral angle makes its side chain point out of the pocket with its long alkyl side chain constituting one side wall of the P3 pocket. Moreover, the conserved acidic residue Glu22 is fixed by two strong hydrogen bonds to Thr19 and Tyr15, with its long alkyl side chain constituting the other side wall of the P3 pocket. Thus, the so-called specificity pocket P3 is particularly unsuitable for binding to the charged residues in the ligand. On the contrary, it is inclined to be a hydrophobic pocket together with the adjacent hydrophobic pocket P4, accommodating the aromatic residues of the binding ligand (Figure 5).

The three clefts (P1, P2, P3) are arranged linearly from left to right and are approximately parallel to the direction of the hairpin loop. The fourth hydrophobic pocket termed P4 is aligned vertical to the preceding groove containing the P1, P2 and P3 pockets. It is made up of three hydrophobic residues with Trp41 on one side, Leu52 on the other and Ala37 on the distal side.

Comparison with other SH3 domains reveals the unique binding scaffold of the srGAP1 SH3 domain

We compared the srGAP1 SH3 domain with the crystal structures of three other SH3 domains: the α-spectrin SH3 domain D48G mutant (PDB code: 1BK2) (24), the c-Crk N terminal SH3 domain (PDB code: 1CKA) (28), and the Abl tyrosine kinase SH3 domain (PDB code: 1ABO) (29). The RMS deviations between the C_α atoms of these structures and the srGAP1 SH3 domain are 0.98 Å, 1.3 Å, and 2.1 Å respectively. Superposition of these structures reveals the unique binding surface presented by the srGAP1 SH3 domain. In the conventional PXXP binding site, most of the conformations of the conserved hydrophobic residues in these structures resemble each other except for Phe13 and Tyr57 in the P1 pocket (Figure 3). The phenyl ring of Phe13 (Tyr70 in 1ABO) in the srGAP1 SH3 domain is directed sidewards and upwards from the bottom of the P1 pocket relative to the other SH3 domain. The aromatic side chain of Tyr57 in the srGAP1 SH3 domain is deflected to one side from the center of the P1 pocket compared to the other SH3 domain. These differences make the P1 pocket of the srGAP1 SH3 domain shallower and smaller than the general groove surrounding the first "XP" dipeptide in other SH3 domains (19).

Spectrin involves in a multi-protein scaffold attached to diverse cellular

membranes and functions in actin dynamics and focal adhesions, partially through the central region hosting an SH3 domain(30,31). As shown by the sequence alignment in Figure 4, the α-spectrin SH3 domain D48G mutant was among those with highest sequence similarity to the srGAP1 SH3 domain, and was therefore used as a search model for the molecular replacement solution of the structure. In general, there is a conserved class I β-turn in the tip of the RT-loop which includes a cluster of polar residues forming the specificity pocket P3. It is interesting to note that the srGAP SH3 domain has oppositely charged residues in the i+2 and i+3 positions similar to the α-spectrin SH3 domain. These two residues have almost the same conformations in the two structures (Figure 3).

In the structure of the Crk-N SH3 complex with a high affinity peptide from C3G, a lysine in the C3G peptide is tightly coordinated by three acidic residues in the class I β-turn in the tip of the RT-loop of Crk-N SH3 domain (Figure 5). The conformations of these three acidic residues are similar to those of the srGAP1 SH3 domain (Figure 3). However, in the srGAP1 SH3 domain, a basic residue Arg21 replaces Glu149, while Ser19 replaces Asp147 of the Crk-N SH3 domain. The opposing electrostatic residues lining the P3 specificity pocket make it unsuitable for tight binding to ligand as in the other high affinity binding SH3 domains. This may offer an explanation for the differences in their ligand binding specificities and affinities as discussed later.

The SH3 domain of Abl kinase was selected for comparison as it binds to the same proline-rich CC3 motif of Robo as the srGAP1 SH3 domain(32,33). In the complex of the Abl SH3 domain complex with synthetic peptide 3BP1, the four residues in the tip of the RT-loop form a type II β-turn in the P3 specificity pocket. Asp77 (Ala20 in srGAP1 SH3) points towards the center of the pocket to form a hydrogen bond with Met4 of its binding peptide, compensating for the function of the conserved acidic residue in the bottom of P3 (29) (Figure 3, 5). In the srGAP1 SH3 domain, the classical type I β-turn directs the Arg21 residue (Asn78 in Abl SH3) into the pocket and positions it adjacent to the conserved Asp22 in the side face of the grooves. So, we speculate that the srGAP1 SH3 domain will not employ the conserved polar residue (D22) in one side of the specificity pocket to form a hydrogen bond with the binding peptides, which is similar to the Abl SH3 recognition mode.

The free energy of binding between proteins includes an electrostatic and a hydrophobic component. Displaying the electrostatic potential and hydrophobicity on the surface of a protein can provide information concerning the nature of its interactions with other proteins. In Figure 5, the electrostatic and hydrophobic surface properties of the srGAP1 SH3 domain are compared with those of the Abl and Crk-N SH3 domains. The most striking aspects of the srGAP1 SH3 domain are the narrow and shallow P1 pocket recognizing the sequence –PXXP–, together with the opposing electrostatic potential and hydrophobic characteristics of the P3 specificity pocket.

Specificity of the recognition site on the Robo1 intracellular domain

The SH3 domains of srGAP proteins interact with Robo predominantly through the proline-rich CC3 motif in the receptor's cytoplasmic domain. Activation of Robo leads to the srGAP dependent

down-regulation of the small GTPase Cdc42 (7,11). The srGAP family has three members, called srGAP1, srGAP2 and srGAP3, which have different expression profiles and may have distinct functions in diverse tissues (11). The consensus sequence of the full length CC3 motif, ^{1476}TYTDDLPPPPVPPPAIKSP^{1493} (12), could be interpreted either as an extended class II peptide, containing an extra negatively charged residue at the amino side, or an extended class I peptide, containing an extra positively charged residue at the C terminus. Therefore, we expressed the three srGAP SH3 domains and synthesized a series of peptides corresponding to the conserved cytoplasmic motif 3 (CC3). The peptides were analyzed using surface plasmon resonance (BIAcore) to gain a better understanding of the specificity and strength of the interactions between the srGAPs and the precise recognition sequence in the Robo1 CC3 motif. The names and the sequences of the synthesized peptides are summarized in Table 2. Binding analyses results indicate that the SH3 domains of different srGAP interact with Robo peptides with slightly different affinities.

Peptides derived from the full length Robo1 CC3 motif (CC3-FL) bound to the immobilized srGAP1-SH3 surface with an equilibrium dissociation constant (K_d) of 4.8 μM, which is the highest affinity for SH3-peptide interactions (Figure 1A, 1B. Table 2). We could not detect any significant sensorgram response for cluster of the hydrophobic residues in the center of the CC3 motif (CC3-10) binding to the SH3 domains of srGAPs, indicating that those charged residues flanking the central hydrophobic cluster are critical for the interaction of srGAP SH3 to the Robo CC3 motif. Adding two aspartic acids to the N-terminus of the hydrophobic cluster (CC3-I13) had a profound effect on the ability of the peptides to interact with the srGAP SH3 domains, indicating that those two acidic residues are critical for the interaction, and that ligands recognized by the srGAP SH3 domain might be class I peptides. Adding the basic residues to the C-terminus of the hydrophobic cluster (CC3-II14), however, resulted in no detectable binding to the srGAP1 SH3 domains, suggesting that srGAP SH3 domains may not recognize class II sequences in the CC3 motif. In contrast, the peptide CC3 which included flanking residues at both ends of the hydrophobic cluster resulted in weak binding similar to CC3-I13. As a result, these data together with the CC3-FL experiment indicate that srGAP SH3 domains may recognize CC3 motifs of the class I peptide type, requiring the hydrophobic cluster and the two N-terminal two acidic residues for weak binding affinity. The full length CC3 motif binding sequence of srGAP1 has the highest affinity for Robo1-CC3.

There are two proline-rich sequences in the intracellular region of Robo1. The CC2 motif is a proline-rich sequence (^{1184}DLLPPPPAHPPPHSN^{1198}) containing the consensus binding site (LPPPP) for the EVH1 domain of the *Drosophila* Enabled protein (12,32). However, it also matches the general consensus binding site (PXXP) of the general SH3 domain. Interestingly, we found that peptides derived from the Robo1 CC2 motif (CC2) can bind to the immobilized srGAP SH3 domains, with affinities comparable to the 15aa CC3 peptides. This affinity is much lower than that for the CC3-FL peptide (Figure 1C, 1D). We also examined the CC3-FL peptide

binding to the immobilized βPIX-SH3 domain (34,35) and found no detectable binding response (data not shown), thus revealing the specificity of the Robo1 CC3 motif for srGAP SH3 domains.

DISCUSSION

Now that a large number of extracellular guidance molecules and their receptors have been identified, the focus has turned to examining the intracellular signaling mechanisms that transduce the signals at the cell surface into changes in growth cone dynamics and cell motility (6,7,36). One of the emerging themes from recent studies is the major importance of the cytoplasmic domains of guidance receptors in signaling and the diverse mechanisms utilised by the cytoplasmic domains to regulate axon guidance. The srGAP1 SH3 domain directly interacts with the intracellular domain of Robo and mediates the repulsive response of Slit, a well studied guidance cue.

The high resolution structure of the srGAP1 SH3 domain is the first X-ray crystal structure of the intracellular signaling module in the Slit-Robo-srGAP pathway. The elaborate structure of the srGAP1 SH3 domain provides a structural basis for the srGAP-Robo interaction, an important step in mediating the repulsive Slit signal in axon guidance and neuronal migration. The P1 pocket is relatively shallow and narrow. This unusual feature in the SH3 binding groove provides an explanation for the lack of binding to the hydrophobic core sequence in the CC3 motif and for the weak binding affinity to the CC3-I13 and CC3 peptides. The opposing electrostatic charge and the hydrophobic properties of the P3 specificity pocket are consistent with our BIAcore results showing the relatively low affinity and weak binding of the srGAP1 SH3 domain. This characteristic endows the srGAP1 SH3 domain with the ability to adapt to variable conditions and to modulate the corresponding downstream pathways. It is conceivable that several polar residues in the solvent-exposed surface of srGAP1 SH3 domain may contribute to the potential binding site for the flanking residues of the CC3 motif hydrophobic core. For example, Lys26, Lys27 and Lys12 comprise a cluster of basic residues in the C-terminal of the binding slot. There is a cluster of acidic residues in the n-Src loop adjacent to the P4 pocket. The interaction between the flanking residues of the CC3 motif hydrophobic core and the SH3 domain may enhance the binding of the Robo intracellular domain to srGAP1. These binding properties and structural features provide a structural mechanism allowing Robo cytoplasmic repeats to interact with other downstream signaling molecules.

There are several potential binding sites in the Robo intracellular domain, including CC0, CC1, CC2 and CC3, that can directly transduce the extracellular axon guidance cues. There are also many effectors downstream of the intracellular domain of Robo such as Abl (32), Ena (32), srGAP, Vilse (37) and Dock (38). CC2 matches the consensus binding site for the EVH1 domain of the *Drosophila* Enabled protein and CC3 is a poly-proline stretch (32,39). Abl functions to antagonize Robo signaling, likely through a mechanism involving direct phosphorylation of the Robo receptor on the CC0 and CC1 motifs, but Abl may also bind via its SH3 domain to the CC3 motif in Robo (32,33). Dock can directly bind to the cytoplasmic domain of Robo, but is dependent on the SH3 domains of Dock and

the CC2 and CC3 motifs in Robo (38). Furthermore, Vilse/CrGAP, a new discovered RhoGTPase, was also reported to bind the intracellular CC2 Robo domain through its WW domain and may play an analogous role to srGAP in locally down-regulating actin polymerization (37,40). Thus, there is a complex network involving Robo and its downstream effectors. In different neuronal cells and under different cellular conditions, Slit-binding may stimulate the intracellular conformational changes of Robo, resulting in the exposure of different binding sites to the matrix. As a consequence, Robo would bind to different effectors or combinations of effectors, eventually resulting in different responses in axon guidance and cell motility. These different downstream effectors may interact with the intracellular domains of Robo in different fashions, resulting in a dynamic spatial and temporal regulation in GTPase effectors. This complex network may be vital for adapting to different cellular conditions and for coordinating precise responses in axon pathfinding and cell movement.

REFERENCES

1. Bear, J. E., Krause, M., and Gertler, F. B. (2001) *Curr Opin Cell Biol* **13**(2), 158-166

2. Luo, L. (2002) *Annu Rev Cell Dev Biol* **18**, 601-635

3. Bishop, A. L., and Hall, A. (2000) *Biochem J* **348 Pt 2**, 241-255

4. Dickson, B. J. (2001) *Curr Opin Neurobiol* **11**(1), 103-110

5. Luo, L. (2000) *Nat Rev Neurosci* **1**(3), 173-180

6. Patel, B. N., and Van Vactor, D. L. (2002) *Curr Opin Cell Biol* **14**(2), 221-229

7. Rao, Y., Wong, K., Ward, M., Jurgensen, C., and Wu, J. Y. (2002) *Genes Dev* **16**(23), 2973-2984

8. Lamarche, N., and Hall, A. (1994) *Trends Genet* **10**(12), 436-440

9. Moon, S. Y., and Zheng, Y. (2003) *Trends Cell Biol* **13**(1), 13-22

10. Schmidt, A., and Hall, A. (2002) *Genes Dev* **16**(13), 1587-1609

11. Wong, K., Ren, X. R., Huang, Y. Z., Xie, Y., Liu, G., Saito, H., Tang, H., Wen, L., Brady-Kalnay, S. M., Mei, L., Wu, J. Y., Xiong, W. C., and Rao, Y. (2001) *Cell* **107**(2), 209-221

12. Kidd, T., Brose, K., Mitchell, K. J., Fetter, R. D., Tessier-Lavigne, M., Goodman, C. S., and Tear, G. (1998) *Cell* **92**(2), 205-215

13. Zallen, J. A., Yi, B. A., and Bargmann, C. I. (1998) *Cell* **92**(2), 217-227

14. Mayer, B. J. (2001) *J Cell Sci* **114**(Pt 7), 1253-1263

15. Dalgarno, D. C., Botfield, M. C., and Rickles, R. J. (1997) *Biopolymers* **43**(5), 383-400

16. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) *Science* **266**(5188), 1241-1247

17. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., and Schreiber, S. L. (1994) *Cell* **76**(5), 933-945

18. Feng, S., Kasahara, C., Rickles, R. J., and Schreiber, S. L. (1995) *Proc Natl Acad Sci U S A* **92**(26), 12408-12415

19. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature* **372**(6504), 375-379

20. Adzhubei, A. A., and Sternberg, M. J. (1993) *J Mol Biol* **229**(2), 472-493

21. Larson, S. M., and Davidson, A. R. (2000) *Protein Sci* **9**(11), 2170-2180

22. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology. Macromol. Crystallog. A* **276** 307-326

23. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**, 905-921

24. Martinez, J. C., Pisabarro, M. T., and Serrano, L. (1998) *Nat Struct Biol* **5**(8), 721-729

25. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr A* **47 (Pt 2)**, 110-119

26. Chan, A. W., Hutchinson, E. G., Harris, D., and Thornton, J. M. (1993) *Protein Sci* **2**(10), 1574-1590

27. Shin, I., Ting, A. Y., and Schultz, P. G. (1997) *J. Am. Chem. Soc.* **119**(51), 12667 -12668

28. Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H., and Kuriyan, J. (1995) *Structure* **3**(2), 215-226

29. Musacchio, A., Saraste, M., and Wilmanns, M. (1994) *Nat Struct Biol* **1**(8), 546-551

30. Rotter, B., Bournier, O., Nicolas, G., Dhermy, D., and Lecomte, M. C. (2005) *Biochem J* **388**(Pt 2), 631-638

31. Merilainen, J., Palovuori, R., Sormunen, R., Wasenius, V. M., and Lehto, V. P. (1993) *J Cell Sci* **105 (Pt 3)**, 647-654

32. Bashaw, G. J., Kidd, T., Murray, D., Pawson, T., and Goodman, C. S. (2000) *Cell* **101**(7), 703-715

33. Hsouna, A., Kim, Y. S., and VanBerkum, M. F. (2003) *J Neurobiol* **57**(1), 15-30

34. Manser, E., Loo, T. H., Koh, C. G., Zhao, Z. S., Chen, X. Q., Tan, L., Tan, I., Leung, T., and Lim, L. (1998) *Mol Cell* **1**(2), 183-192

35. Li, X., Liu, X., Sun, F., Gao, J., Zhou, H., Gao, G. F., Bartlam, M., and Rao, Z. (2006) *Biochem Biophys Res Commun* **339**(1), 407-414

36. Garbe, D., and Bashaw, G. (2004) *Crit Rev Biochem Mol Biol* **39**(5-6), 319-341

37. Lundstrom, A., Gallio, M., Englund, C., Steneberg, P., Hemphala, J., Aspenstrom, P., Keleman, K., Falileeva, L., Dickson, B. J., and Samakovlis, C. (2004) *Genes Dev* **18**(17), 2161-2171

38. Fan, X., Labrador, J. P., Hing, H., and Bashaw, G. J. (2003) *Neuron* **40**(1), 113-127

39. Yu, T. W., Hao, J. C., Lim, W., Tessier-Lavigne, M., and Bargmann, C. I. (2002) *Nat Neurosci* **5**(11), 1147-1154

40. Hu, H., Li, M., Labrador, J. P., McEwen, J., Lai, E. C., Goodman, C. S., and Bashaw, G. J. (2005) *Proc Natl Acad Sci U S A* **102**(12), 4613-4618

41. Esnouf, R. M. (1997) *J Mol Graph Model* **15**(2), 132-134, 112-133

42. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res* **22**(22), 4673-4680

43. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15**(4), 305-308

44. Laskowski, R. A. a. M., M.W.; Moss, D.S.; Thornton, J.M. . (1993) *Journal of Applied Crystallography* **26**(Part 2), 283-291

FOOTNOTES

* We wish to thank Yuanyuan Chen for her help with the BIAcore experiments and Dr. Fei Sun for his valuable discussion about the paper. ZR was supported by Project 973 of the Ministry of Science and Technology (Grant number 2004CB520801); JYW was supported by NIH (Grant number NIH RO1 CA114197). Coordinates of the srGAP1 SH3 domain structure have been deposited in the Protein Data Bank with accession number 2GNC. (http://www.rcsb.org/pdb/).

FIGURE LEGENDS

Figure 1: Crystal structure and ligand binding surface of the srGAP1 SH3 domain. A: Ribbon representation of the monomer structure of the srGAP1 SH3 domain. B: The transparent ligand binding pockets P1, P2, P3 and P4 are mapped onto the solvent accessible surface. The residues involved in ligand recognition are labeled. This figure was drawn using PyMOL (http://www.pymol.org/), Molscript, Bobscript and Raster 3D (41).

Figure 2: Stereo view of the two classic β-bulges found in the βb strand, which endows this long strand with greater flexibility than in other SH3 domain structures. Hydrogen bonds are denoted by dashed lines.

Figure 3: Stereo view showing a comparison of the srGAP1 SH3 domain with other SH3 domain structures. The residues forming the ligand binding pockets are shown in ball-and-stick representations. Molecule A of the srGAP1 SH3 domain is colored blue and molecule B is colored spring green, the α-spectrin SH3 domain (PDB ID: 1BK2) (24) is colored gold, the Crk-N SH3 domain (PDB ID: 1CKA) (28) is colored brown and the Abl tyrosine kinase SH3 domain (PDB ID: 1ABO) (29) is colored magenta. Residues are labeled corresponding to their positions in the srGAP1 SH3 domain.

Figure 4: Sequence alignment and structural fold of srGAP1 SH3 domain with other SH3 domains. βa-βe represent β-strands; T represents a β-turn. The four specific loops, including a 310 helix, are labeled. Identical residues are highlighted in red and the most conserved residues are highlighted in yellow. The green box indicates the conserved β-turn in the tip of RT loop which constitutes half of the Pocket3. Sequences were aligned using CLUSTALW (42) and the figure was produced by ESPript (43).

Figure 5: Surface representations of the SH3 domains from Crk-N complexed with C3G peptide, Abl complex with 3BP1, and srGAP1. The views on the left depict surface accessible hydrophobic regions colored in yellow. On the right are representations of electrostatic potential showing the peptide binding surface of the SH3 domains with the n-Src loops pointing toward the top of the page, in which positive electrostatic potential is colored in blue, negative electrostatic potential is colored in red and hydrophobic surface is colored white. Peptides are in ball-and-stick representation. Coordinates for the SH3 domains of

Crk-N and Abl were obtained from the Protein Data Bank. In the surface of the SH3 domain of srGAP1, the red circles indicate the four pockets in the ligand binding groove. The figure was produced by CCP4mg (http://www.ysbl.york.ac.uk/~ccp4mg/) and PyMOL (http://www.pymol.org/).

Figure 6: BIAcore analysis of the binding specificity of the srGAP1 SH3 domain to the intracellular domain of Robo1. (A) represents the sensorgram response profile with different concentrations of the CC3-FL peptide binding to the immobilized srGAP1 SH3 domain. (B) shows the steady curve of CC3-FL peptide binding to the srGAP1 SH3 domain. (C) and (D) show a comparison of the binding affinity of the synthesized peptides at the same concentration of 600 μM binding to the srGAP1 SH3 domain *in vitro*.

TABLES

Table 1: Data collection and refinement statistics

Crystallographic data collection and refinement statistics		
Data collection		
Space group	P23	I23
Unit Cell		
a (Å)	70.3	69.8
b (Å)	70.3	69.8
c (Å)	70.3	69.8
α,β,γ (°)	α=90, β=90, γ=90	α=90, β=90, γ=90
Resolution (Å)	50-1.8 (1.87-1.81)	50-2.5 (2.59-2.5)
Completeness (%)	93 (99.8)	99.1 (96.6)
Reflections		
No. of measured reflections	67116	10695
No. of unique reflections	11938	2063
Redundancy	3.3 (3.3)	5.2(4.0)
[a] R_{merge} (%)	13.1 (30.4)	4.9 (33.2)
I/σ(I)	11.6 (5.3)	30.6 (3.2)
Refinement statistics		
Resolution (Å)	1.8	
[b] R-factor (%)		
Working set	20.8	
Test set	25.7	
R.m.s deviation		
Bonds (Å)	0.117	
Angles (°)	1.603	
[c]Ramachandran plot (100%)		
Most favored	97.9	
Allowed	2.1	
Generously allowed	0	
Disallowed	0	

[a] $R_{merge}= \Sigma_h\Sigma_l \mid I_{ih} - \langle I_h \rangle \mid / \Sigma_h\Sigma_I \langle I_h \rangle$, *where* $\langle I_h \rangle$ *is the mean of the observations* I_{ih} *of reflection .*

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

[c] *Ramachandran plots were generated using program PROCHECK (44)*

Table 2: Dissociation constants of various peptides derived from the intracellular domains of Robo1 binding to the SH3 domains of srGAPs, from BIAcore experiments. Residues underlined represent the additional linker. Ac at the N-terminal of the peptides indicates the addition of an acetyl group and the -NH2 shows the imido C-terminal. 1SH3, 2SH3 and 3SH3 correspond to the srGAP1 SH3, srGAP2 SH3 and srGAP3 SH3 domain, respectively. ND means that we could not detect any binding response.

Peptide	Amino acid sequence of peptides	Number of residues	Molecular weight (Dalton)	Dissociation constant (K_d) (μM)		
				1SH2	2SH2	3SH3
CC3FL	Ac-TYTDDLPPPPVPPPAIKSPGGK-NH2	22aa	2284.2	4.8	4.9	5.6
CC3	DDLPPPPVPPPAIKS	15aa	1539.7	876	925	107
CC3-I13	Ac-DDLPPPPVPPGGK-CONH2	13aa	1285.4	160	137	149
CC3-II14	Ac-PPPVPPPAIKSGGK-CONH2	14aa	1341.6	ND	ND	ND
CC3-10	PPPPVPPPAI	10aa	981.2	ND	ND	ND
CC2	DLLPPPPAHPPPHSN	15aa	1585.7	682	667	902

FIGURES

Figure 1

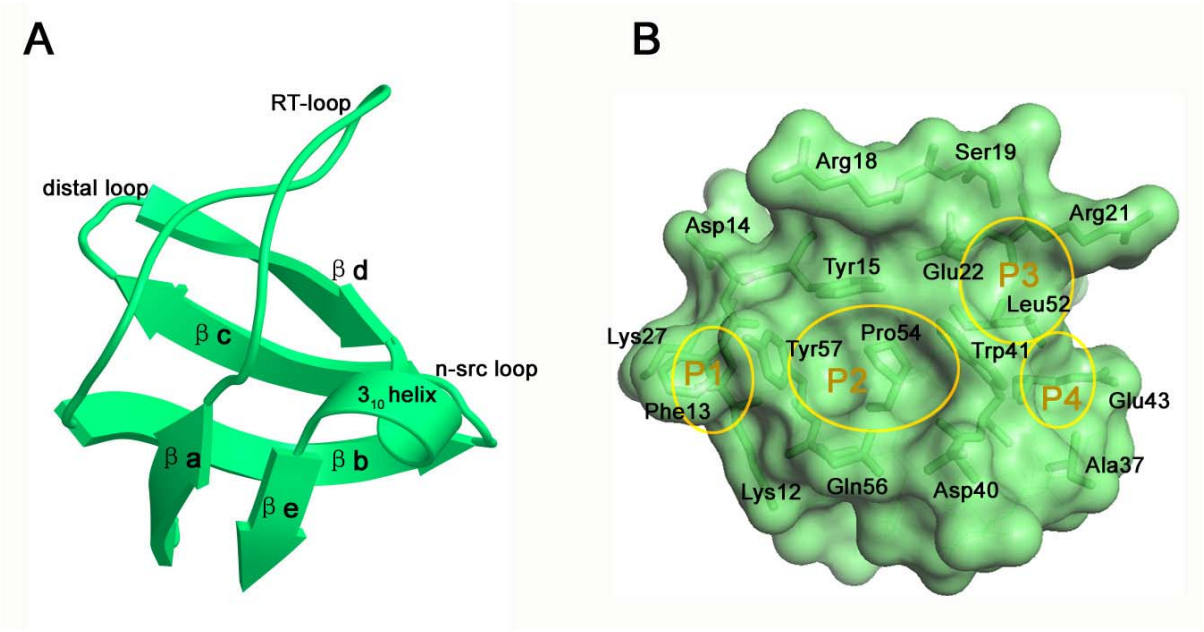

Figure 2

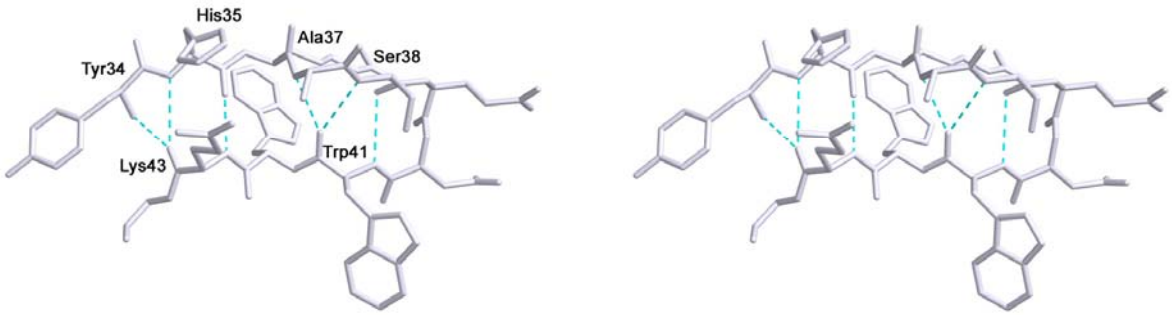

Figure 3

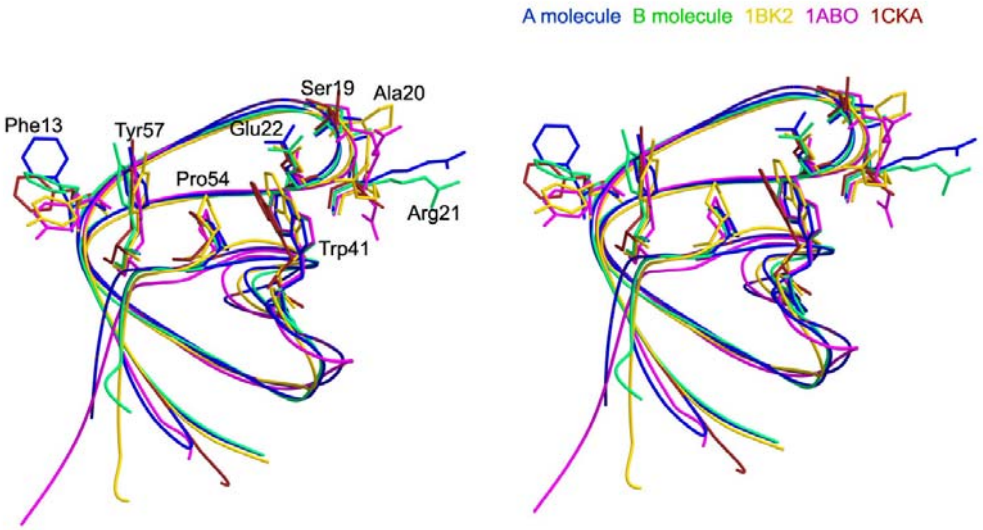

Figure 4

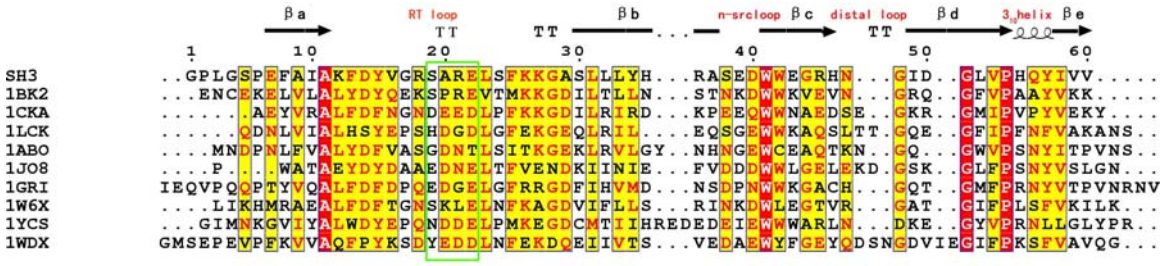

Figure 5

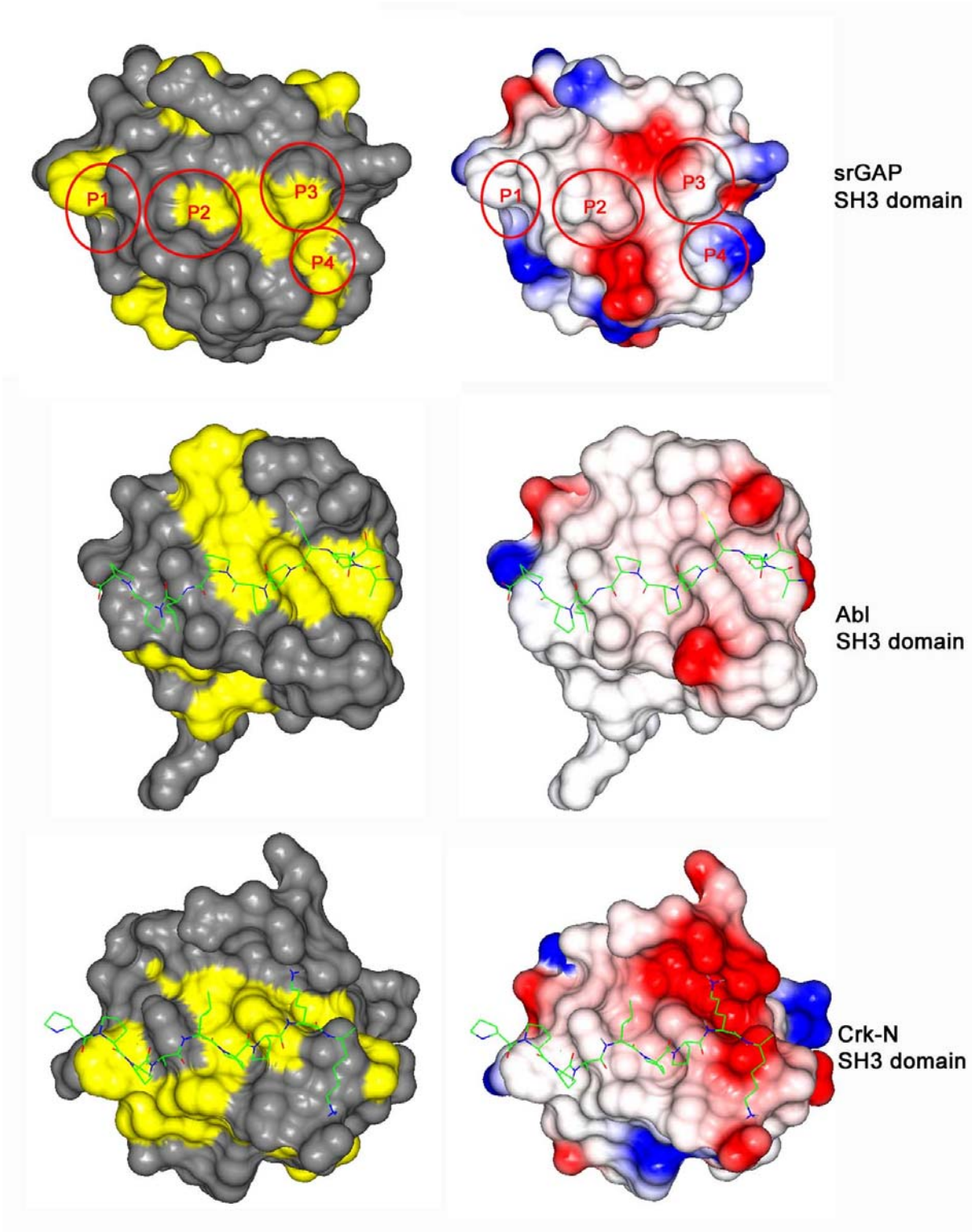

Figure 6