A ‘three-pronged’ binding mechanism for the SAP/SH2D1A SH2 domain: structural basis and relevance to the XLP syndrome

Peter M. Hwang1, Chengjun Li2, Massimo Morra3, Jennifer Lillywhite2, D. Ranjith Muhandiram1,4, Frank Gertler5, Cox Terhorst3, Lewis E. Kay1,4,6, Tony Pawson7,7, Julie D. Forman-Kay1,8 and Shun-Cheng Li2,9

Departments of 1Biochemistry, 2Chemistry and 4Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5S 1A8, 3Department of Biochemistry, Faculty of Medicine and Dentistry, University of Western Ontario, London, Ontario N6A 5C1, 5Samuel Lunenfeld Research Institute, Mt Sinai Hospital, Toronto, Ontario MSG 1X5, 6Program in Structural Biology and Biochemistry, The Hospital for Sick Children, Toronto, Ontario MSG 1X8, Canada, 7Division of Immunology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215 and 8Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

9Corresponding author
e-mail: sli@uwo.ca

The SH2 domain protein SAP/SH2D1A, encoded by the X-linked lymphoproliferative (XLP) syndrome gene, associates with the hematopoietic cell surface receptor SLAM in a phosphorylation-independent manner. By screening a repertoire of synthetic peptides, the specificity of SAP/SH2D1A has been mapped and a consensus sequence motif for binding identified, T/S-x-x-x-x-I/I, where x represents any amino acid. Remarkably, this motif contains neither a Tyr nor a pTyr residue, a hallmark of conventional SH2 domain–ligand interactions. The structures of the protein, determined by NMR, in complex with two distinct peptides provide direct evidence in support of a ‘three-pronged’ binding mechanism for the SAP/SH2D1A SH2 domain in contrast to the ‘two-pronged’ binding for conventional SH2 domains. Differences in the structures of the two complexes suggest considerable flexibility in the SH2 domain, as further confirmed and characterized by hydrogen exchange studies. The structures also explain binding defects observed in disease-causing SAP/SH2D1A mutants and suggest that phosphorylation-independent interactions mediated by SAP/SH2D1A likely play an important role in the pathogenesis of XLP.

Keywords: immunodeficiency/SAP/SH2 domain/SH2D1A/XLP

Introduction

X-linked lymphoproliferative (XLP) syndrome, also known as Duncan’s disease, is a rare inherited immunodeficiency condition characterized by benign or malignant proliferation of lymphocytes, histiocytosis and alterations in serum immunoglobulin concentrations (Portillo et al., 1975, 1978). The gene altered in XLP has recently been identified (Coffey et al., 1998; Nichols et al., 1998; Sayos et al., 1998). It encodes a small protein of 128 amino acids (variably named SAP, SH2D1A and DSHP), consisting of a five amino acid N-terminal sequence, an SH2 domain and a 25 residue C-terminal tail. A number of SAP/SH2D1A mutations have been identified in XLP patients (Nichols et al., 1998; Sayos et al., 1998; Yin et al., 1999; Sumegi et al., 2000) and these fall into three broad categories: (i) micro/macro-deletions that lead to a complete or partial loss of the gene; (ii) mutations that interfere with mRNA transcription or splicing; and (iii) nonsense mutations that result in premature termination of protein synthesis or missense mutations that give rise to substitution of an amino acid conserved between the human and murine SAP/SH2D1A proteins (Morra et al., 2001a). The missense mutations are particularly interesting since, except for one case where a stop codon at the C-terminus of the protein is replaced by that for an Arg, all the other mutations occur within the boundaries of the SH2 domain of the protein and therefore directly implicate the SH2 domain in the pathogenesis of XLP.

Although mRNA analysis has strongly linked the SAP/SH2D1A gene to XLP, little is known about the molecular mechanism of the disease at the protein level. Since SAP/SH2D1A lacks an effector domain (e.g. a catalytic domain), it was suggested to function as an inhibitor of other SH2 domain-mediated interactions in T and natural killer (NK) cells so as to enhance or prolong the activities of these cells (Satterthwaite et al., 1998; Sayos et al., 1998). In line with this possibility, SAP/SH2D1A was found to bind to the cytoplasmic domain of the hematopoietic surface receptor known as the signaling lymphocyte activation molecule (SLAM), a glycosylated transmembrane protein also known as CD150 (Cocks et al., 1995). This interaction, mediated by the SAP/SH2D1A SH2 domain and a tyrosine phosphorylation site (Tyr281 in SLAM), was further shown to block the recruitment of an SH2 domain-containing tyrosine phosphatase Shp2 to the same site in SLAM (Sayos et al., 1998). Intriguingly, SAP/SH2D1A associated with SLAM regardless of the phosphorylation state of Tyr281, indicating that its SH2 domain possesses a novel binding specificity, which raises questions about the role of tyrosine phosphorylation in SAP/SH2D1A function. An analogous complex was also documented between SAP/SH2D1A and B24, an activating NK cell receptor that shares significant sequence identity with SLAM (Tangye et al., 1999, 2000). Moreover, engagement of B24 by SAP/SH2D1A prevented its association with Shp2. It is, therefore, likely that SAP/SH2D1A plays similar roles in regulating the activities of T cells (via SLAM) and NK cells (via B24). However, contrary to the above observations, binding of SAP/SH2D1A was recently shown to be necessary to
induce tyrosine phosphorylation of SLAM by facilitating the recruitment of the Src-related tyrosine kinase FynT, rather than displacing a phosphotyrosine phosphatase (PTP) that interferes with SLAM signal transduction (Latour et al., 2001).

Previously, we employed a series of truncated peptides derived from sequences flanking Tyr281 in SLAM to investigate the specificity of the SAP/SH2D1A SH2 domain (Li et al., 1999). Results demonstrated that this SH2 domain is not only capable of binding to a pTyr residue that is followed by a stretch of C-terminal amino acids in a manner akin to that of a conventional SH2 domain, but it is also capable of recognizing a pTyr residue preceded by an N-terminal fragment or an unphosphorylated Tyr residue embedded within an appropriate sequence context. Based on these observations, a ‘three-pronged’ plug that engages three binding sites’ model was put forward to account for the tripartite nature of the SAP/SH2D1A–SLAM–Tyr281 peptide interaction (Li et al., 1999). This model is consistent with the crystal structures of the SAP/SH2D1A SH2 domain in a complex with two SLAM–Tyr281 peptides, one containing a pTyr residue and the other Tyr, which demonstrated that, in addition to residues C-terminal to the pTyr/Tyr, those N-terminal are also involved in binding to the SH2 domain (Poy et al., 1999). Here, we present two solution structures of this protein: one in complex with a peptide that lacks the entire sequence C-terminal to pTyr281 and the other with an unphosphorylated SLAM peptide. These structures provide compelling evidence in support of a ‘three-pronged’ mechanism for peptide recognition by the SAP/SH2D1A SH2 domain, which is further supported by results obtained from a SPOTs (Frank and Doring, 1988; Blankenmeyer-Menge et al., 1990) peptide binding study demonstrating that SAP/SH2D1A binds selectively to peptides with a consensus sequence motif T/S-x-x-x-V/I, where x denotes any amino acid. The structures, together with amide hydrogen exchange rates measured by NMR spectroscopy, also provide a rationale for the reduced binding affinities for both the phosphorylated and unphosphorylated SLAM–Y281 peptides observed for disease-causing SAP/SH2D1A missense mutants.

Results

The SAP/SH2D1A SH2 domain binds to peptide motifs containing no tyrosine

In order to determine the specificity of the SAP/SH2D1A SH2 domain in a more comprehensive manner than previously reported (Li et al., 1999), we synthesized a set of undecamer peptides using the SPOTs method of simultaneous multiple peptide synthesis on a derivatized cellulose membrane (Frank and Doring, 1988; Blankenmeyer-Menge et al., 1990). These peptides are based on the sequence of the Tyr281 site in SLAM (KSLTlYAQVQK), which contains five amino acids each at the N- and C-terminus of the Tyr281 residue and is thus designated n-Y-c for simplicity. Specifically, each position in peptide n-Y-c was permuted to allow for a full representation of all 20 naturally occurring amino acids at that position, generating a total of $11 \times 20 = 220$ spots on a cellulose membrane. A series of truncated peptides with decreasing N- or C-terminal length was also synthesized, as was a group of Ala-scanning peptides derived from peptide n-pY-c, the phosphorylated version of peptide n-Y-c (Figure 1).

The peptide spots on the cellulose membrane were then screened for their ability to bind to purified glutathione S-transferase (GST)–SAP/SH2D1A. Results from the screen are displayed in Figure 1, where bright (fluorescent) spots indicate positive binding of the corresponding peptides to the protein and dark spots denote weak or negative binding. To facilitate analysis of the results, each residue in peptide n-Y-c was assigned a number according to its position relative to the central Tyr, which was given a value ‘0’ (Figure 1). As seen in Figure 1A, residues in the peptide do not play equal roles in binding to SAP/SH2D1A. While positions +4, +5, and −3 to −5, corresponding to the extreme N- and C-termini of the peptide, and position +2 could be substituted by almost any residue without adversely affecting binding, SAP/SH2D1A is highly selective for positions −2 and +3. Binding is facilitated by residues with a hydroxyl side chain, such as Thr and Ser at position −2, and residues with hydrophobic, β-branched side chains such as Val and Ile are favored at position +3. A certain degree of selectivity was also seen at positions −1 and +1. While position −1 demonstrated a proclivity for residues with large side chains, position +1 displayed a bias against bulky, hydrophobic residues. Interestingly, there is no apparent discrimination at position 0, corresponding to the central Tyr residue in peptide n-Y-c, suggesting that within the context of the full-length peptide, the Tyr residue is completely dispens-
able for binding. It should be pointed out that GST alone did not show detectable binding to any of the peptide spots (data not shown). Collectively, these data demonstrate that SAP/SH2D1A selectively binds to a peptide motif containing the consensus sequence T/S-x-x-x-x-V/I.

The importance of residues at positions –2 and +3 is also reflected by the N- and C-terminally truncated peptides. As shown in Figure 1B, residues at the –5 to –3 positions could be deleted without compromising the affinities of the resulting peptides. However, as soon as the peptide was shortened to exclude Thr at position –2, its affinity for SAP/SH2D1A was lost. Similarly, when the C-terminal truncation reached residue Val+3, the resulting peptide was no longer active (Figure 1C). The role of a pTyr residue in binding to SAP/SH2D1A was demonstrated by the Ala-scanning peptide series (Figure 1D). In the n-pY-c peptide background, residues at either position –2 or +3 appeared to be dispensable for binding since either could be replaced by an Ala residue without significantly compromising the affinity of the resulting peptides. Together, these observations have unambiguously identified three sites or ‘prongs’ in the SLAM-Tyr281 peptide that mediate high-affinity binding to SAP/SH2D1A. These prongs correspond to residues at positions –2 (Thr or Ser) and +3 (Val or Ile), and to a pTyr residue at position ‘0’. A combination of any two of the three prongs in a peptide is both necessary and sufficient for high-affinity SAP/SH2D1A binding.

**Figure 2.** Stereoview of a superposition of the 20 lowest energy structures of SAP/SH2D1A–SLAM peptide complexes (using backbone atoms of residues 6–56). The protein (residues 6–104) is shown in blue for the n-pY complex and in red for the n-Y-c complex. Peptide n-pY is depicted in green (only residues –4 to +1 relative to the pTyr are shown) and n-Y-c in magenta (only residues –4 to +5 are shown).

**Structural basis for the ‘three-pronged’ binding mode of the SAP/SH2D1A SH2 domain**

In order to understand the structural basis underlying the ‘three-pronged’ interaction between SAP/SH2D1A and its targets, we determined its three-dimensional (3D) solution structure in complex with two distinct peptides, n-pY (RKSLTIpYA) and n-Y-c (RKSLTIYAQVQK). Peptide n-pY was selected for structural analysis because it is representative of a novel mode of ligand recognition by an SH2 domain, unique to SAP/SH2D1A. The structure of the SAP/SH2D1A–n-pY complex offers an example of how an SH2 domain recognizes a peptide devoid of the typical contacts C-terminal to the phosphotyrosine. Peptide n-Y-c was chosen for comparative purposes, although a crystal structure of the SAP/SH2D1A SH2 domain in complex with an essentially identical peptide was recently reported (Poy et al., 1999).

The C-terminal 25 residues of SAP/SH2D1A were unstructured in both complexes. The 1H–15N HSQC cross-peaks corresponding to this region of the protein were very narrow and intense, indicative of rapid and unhindered tumbling in solution (Li et al., 1999). Moreover, no intermediate- or long-range nuclear Overhauser effects (NOEs) could be observed beyond residue 105. Consequently, only residues 1–107, corresponding essentially to the SH2 domain region of the protein (residues 6–102), were included in the structure calculations. Figure 2 shows an overlay of the 20 lowest energy structures for each complex calculated using ARIA (Nilges et al., 1998), with statistics of the structures given in Table I.

As shown in Figure 3A and B, the SAP/SH2D1A SH2 domain adopts a conventional SH2 domain fold characterized by a large central β-sheet flanked by two α-helices and a small β-sheet (Cohen et al., 1995). In both complexes, the peptides assume an extended conformation. Amino acids N-terminal to Tyr in peptide n-Y-c or pTyr in peptide n-pY make almost identical contacts with the protein in the two complexes (Figures 2 and 3). The N-terminal ‘prong’, Thr-2, fits into a well-defined binding site formed by Arg13, Glu17, Ile51 and Thr53 (Figure 4). Thr-2 forms a meshwork of interactions with these residues through both hydrogen bonds and hydrophobic interactions. The main-chain carbonyl group and the side-chain hydroxyl of Thr-2 are within hydrogen-bonding distance with the side chains of Thr53 and Glu17, although the latter hydrogen bond may be mediated by a water molecule (Poy et al., 1999). In addition, van der Waals contacts occur between the γ-methyl group of Ile51 and Cα of Thr-2, and between side chains of Arg13 and Thr-2 (Figure 4), as demonstrated by observed NOEs. The dual polar/hydrophobic nature of interactions occurring at position –2 explains the strong preference for Thr at this position seen in the SPOTs peptide binding study.

The middle ‘prong’, the phosphotyrosine, is present in peptide n-pY, but absent in n-Y-c. Even though the binding pocket occupied by the Tyr/pTyr side chain is virtually the same in both complexes, the unphosphorylated Tyr is more mobile. While numerous NOEs can be
Table I. Statistics for the 20 lowest energy SAP/SH2D1A–SLAM peptide complex structures

<table>
<thead>
<tr>
<th></th>
<th>SAP/SH2D1A–n-pY</th>
<th>SAP/SH2D1A–n-Y-c</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.m.s.d. from distance restraints (Å)</td>
<td>0.0083 ± 0.0001</td>
<td>0.0073 ± 0.0011</td>
</tr>
<tr>
<td>all (3157, 3950)</td>
<td>0.0080 ± 0.0002</td>
<td>0.0078 ± 0.0009</td>
</tr>
<tr>
<td>unambiguous (1930, 2252)</td>
<td>0.0082 ± 0.0003</td>
<td>0.0045 ± 0.0003</td>
</tr>
<tr>
<td>ambiguous (1177, 1653)</td>
<td>0.0104 ± 0.001</td>
<td>0.0198 ± 0.0002</td>
</tr>
<tr>
<td>hydrogen bonds (50, 52)</td>
<td>0.157 ± 0.019</td>
<td>0.163 ± 0.016</td>
</tr>
<tr>
<td>Deviations from idealized geometry</td>
<td>0.0011 ± 0.00002</td>
<td>0.0011 ± 0.00003</td>
</tr>
<tr>
<td>bonds (Å)</td>
<td>0.272 ± 0.001</td>
<td>0.273 ± 0.0002</td>
</tr>
<tr>
<td>angles (°)</td>
<td>0.188 ± 0.003</td>
<td>0.154 ± 0.004</td>
</tr>
<tr>
<td>Procheck Ramachandran map analysis</td>
<td>76.3%</td>
<td>79.2%</td>
</tr>
<tr>
<td>most favoured regions</td>
<td>22.7%</td>
<td>19.8%</td>
</tr>
<tr>
<td>additional allowed regions</td>
<td>1.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>generously allowed regions</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Energy (kcal/mol)</td>
<td>−148 ± 17</td>
<td>−276 ± 20</td>
</tr>
<tr>
<td>Atomic r.m.s.d. (Å) from mean structure</td>
<td>0.20 ± 0.05</td>
<td>0.29 ± 0.08</td>
</tr>
<tr>
<td>backbone</td>
<td>0.59 ± 0.14</td>
<td>0.70 ± 0.17</td>
</tr>
<tr>
<td>heavy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The number of restraints for the n-pY and n-Y-c peptide complex structures, respectively, are given in parentheses.

*Of these restraints, 392 of 1930 for the n-pY complex and 507 of 2252 for the n-Y-c complex are structurally redundant.

*Superposition involved protein residues 6–102 and peptide residues −4 to +1 relative to pTyr in peptide n-pY and −4 to +5 in peptide n-Y-c.

**Fig. 3.** Structures of the SAP/SH2D1A SH2 domain in complex with peptides n-Y-c (A and C) and n-pY (B and D). The protein is shown in ribbon representations in (A) and (B) with β-strands in cyan, α-helices in red and loops in gray, and in surface diagrams in (C) and (D) with blue indicating positive and red indicating negative electrostatic potentials. The bound peptides are in green. The secondary structural elements of the protein are labeled in (A) and (B), as are residues of the peptides important for binding. This figure was generated using MolMol (Koradi et al., 1996).

**Fig. 4.** Spatial orientation of residues in the SAP/SH2D1A SH2 domain (shown in cyan) involved in binding to the Thr-2 site (in gold) of the peptides. Protein residues shown are all within 4 Å of Thr-2 in both complexes. Oxygen atoms are colored red and nitrogen blue. Dashed lines indicate hydrogen-bonding interactions. Note that, except for the hydroxyls of the two Thr residues, hydrogen atoms are omitted for clarity.

observed between the pTyr ring protons of peptide n-pY and Arg13, Thr32, Val37, Val40, Cys42, Ile51, Thr53, Tyr54 and Arg55 of the protein, the Tyr residue in peptide n-Y-c shows a relatively limited number of NOE contacts with Arg13, Thr53 and Arg55. The increased rigidity of the pTyr ring is due mainly to hydrogen-bonding interactions between its phosphate group and the side chains of Arg32, Ser34, Ser36 and Arg55. The SPOTs peptide binding studies also demonstrate that the Tyr residue in peptide n-Y-c could be replaced by any of the remaining 19 amino acids without compromising its affinity for SAP/SH2D1A (Figure 1). These observations collectively suggest that the majority of the binding energy from the pTyr ‘prong’ is contributed by interactions to the phosphate group, not the aromatic ring.
The C-terminal ‘prong’, Val+3, is present in peptide n-Y-c, but absent in n-pY. It sits in a well-defined hydrophobic pocket lined by residues from the EF (Ala66, Glu67, Thr68, Ala69 and Lys74) and BG loops (Gly93 and Ile94). (The EF loop is located between strands E and F, while the BG loop is between helix B and strand G; Figure 3A and B). In the n-Y-c peptide complex, these loops adopt an open configuration in order to accommodate the C-terminal residues of the peptide (Figure 3A). In contrast, the two loops adopt a closed configuration in the n-pY complex (Figure 3B), a structure that bears a closer resemblance to the structure of the free protein determined by X-ray crystallography (Poy et al., 1999) than to that of the SAP/SH2D1A–SLAM-pY281 peptide complex (Table II). The switch from an open to a closed state buries a significant amount of surface area, which otherwise would be exposed to the solvent, and closes off the portion of the peptide binding groove that would otherwise interact with the C-terminal ‘prong’ of the peptide. This is seen most vividly in Figure 3C and D. While the two loops in the open state allow for the formation of an elongated binding groove in the n-Y-c complex (Figure 3C), the EF loop in the SAP/SH2D1A–peptide n-pY complex closes over the cleft and, in effect, creates a ridge on the protein surface, which interacts with the truncated C-terminus of peptide n-pY (Figure 3D). The use of flexible regions of a protein, such as loops, to modify the ligand-binding surface reflects the highly adaptive nature of the SAP/SH2D1A SH2 domain employing an ‘induced fit’ mechanism.

**Backbone amide accessibility probed by hydrogen exchange rates**

The most striking differences between the NMR structures of the two complexes are found in the C-terminal half of the protein, including the EF and BG loops. As seen in Figure 2, residues 6–56 superimpose very well. However, between residues 58 and 95, the structures of the two complexes diverge, suggesting considerable conformational flexibility. To characterize the dynamics of this region further, we conducted hydrogen exchange experiments to measure the accessibility of SAP/SH2D1A backbone amides to a solvent. Exchange of most amide hydrogens with a solvent requires a substantial opening of structure. Thus, amide exchange experiments measure larger amplitude structural fluctuations over much longer time scales than conventional NMR relaxation-based techniques. Amide exchange protection factors were calculated for both complexes and the results are shown in Figure 5A. In both complexes, residues 58–95 display lower protection factors than the rest of the SAP/SH2D1A molecule, confirming the dynamic nature of this region. The loops in this region are particularly mobile and even the secondary structure elements (strands βE and βF, and helix αB) are not well protected from the solvent. Thus, multiple conformations are possible even within a single complex. The NMR structure of the SAP/SH2D1A–peptide n-Y-c complex is virtually identical to its corresponding X-ray structure for the protein core (residues 6–56), with a backbone r.m.s.d. of 0.6 Å when the two structures are superimposed (Table II). However, the r.m.s.d. increases to 1.07 Å when residues 6–102 are superimposed, again suggesting considerable conformational variability between residues 58 and 95 even though both structures are in an ‘open’ configuration. The differences between the NMR structures and the previously determined X-ray structures (Table II) highlight the dynamic nature of this region, which have been confirmed and quantitated by hydrogen exchange measurements (Figure 5A). Amide and methyl group relaxation experiments probing fast time scale motion have also been performed, showing behavior in agreement with the hydrogen exchange results (P.J.Finerty Jr, D.R.Muhandiram and J.D.Forman-Kay, unpublished data).

The differences in protection factors between the SAP/SH2D1A–n-pY and n-Y-c structures reflect differences in binding between the two complexes. In the n-pY complex, the phosphotyrosine-binding BC loop of the protein displays much higher protection factors than in the n-Y-c complex (Figure 5). This loop appears to be quite mobile in the absence of a phosphate group, despite the many indirect hydrogen bonds observed in the crystal structure of the SAP/SH2D1A SH2 domain in complex with an unphosphorylated SLAM peptide (Poy et al., 1999). These observations support the notion that the unphosphorylated tyrosine in the peptide has a rather insignificant contribution to SAP/SH2D1A binding, as also suggested by the SPOTs assay and the limited number of NOEs observed. While the n-pY complex has higher protection factors in the BC loop, the n-Y-c complex displays higher protection factors for the entire C-terminal half of the protein (Figure 5B). This suggests that the binding of residues C-terminal to pTyr greatly reduces the conformational freedom present in this region, stabilizing the ‘open’ conformation of the EF and BG loops.

Another interesting observation from the hydrogen exchange measurements is that residues 96–102 are
relatively rigid (Figure 5A). It is surprising that residues 96–99 are well protected against exchange even though they are not part of any regular secondary structure element. As well, residues 100–102 form a peripheral, exposed β-strand. The high protection factors suggest that this region may be important for stabilizing the overall fold of the protein, which explains why mutations in this region compromise SAP/SH2D1A binding (see below).

**Binding defects of disease-causing SAP/SH2D1A mutants**

A variety of missense mutations have been identified in XLP patients. Except for the mutation VTail, which changes a stop codon to an Arg and results in an addition of 11 amino acids to the C-terminus of SAP/SH2D1A (Sayos et al., 1998), all mutations are sequestered within the boundaries of the protein’s SH2 domain. The fact that replacement of a single amino acid causes disease points to the pivotal importance of maintaining the integrity of the protein and suggests that the specificity of the SH2 domain of SAP/SH2D1A is fine-tuned for its function. Since SAP/SH2D1A binds to both phosphorylated and non-phosphorylated peptides, it is possible that the mutated proteins are defective in binding to either or both. To explore these possibilities, we generated 10 single amino acid substitution mutants according to missense mutations found in XLP patients and measured their affinities for both the phosphorylated and non-phosphorylated SLAM-Y281 peptides by fluorescence polarization. As summarized in Table III, all mutants except T53I exhibited reduced affinities for the fluorescein-labeled SLAM-pY281 peptide and all showed decreased binding to the SLAM-Y281 peptide. In addition, for the majority of mutants, binding to the non-phosphorylated SLAM peptide appeared to be more profoundly perturbed (Table III). Again, this is in accordance with the ‘three-pronged’ model, e.g. mutation at a single site of the protein has a less detrimental effect to binding by a peptide with three ‘prongs’ (e.g. SLAM-pY281) than one with only two (e.g. SLAM-Y281).

The 10 mutants can be divided into two groups according to their discrete binding behavior. Group I mutants, including Y7C, S28R, C42W, P101L and VTail, displayed partially reduced affinities for both peptides compared with wild-type SAP/SH2D1A. The $K_d$ values of the corresponding complexes of these mutants with either peptide were found to be $<5 \mu M$ (Table III). Group II mutants, including R32Q, T53I, T68I, Q99P and V102G,
exhibited appreciable affinities for the SLAM-pY281 peptide (with corresponding $K_d$ values $\leq 5$ μM), but were incapable of binding to the SLAM-Tyr281 peptide (with corresponding $K_d$ values $>10$ μM). Figure 6 shows binding curves observed for Y7C, the representative of the group I mutants, and R32Q, the representative of the group II mutants, in comparison with those for the wild-type SAP/SH2D1A protein.

Understandably, the mutation of residues directly involved in peptide recognition had a greater impact on binding than did that of residues outside of the peptide binding cleft (Figures 3 and 4). Thus, mutants R32Q, C42W, T53I and T681I all involve residues interacting with the 'prongs' of the peptide and all display markedly reduced affinity for one or both peptides. Interestingly, mutation of Thr53, a residue involved in binding to the N-terminal 'prong' Thr-2 of the peptides (Figure 4), drastically affected the affinity of mutant T53I for the SLAM-Y281 peptide, while leaving its affinity for the SLAM-pY281 peptide intact. Again, this is readily explained by the 'three-pronged' binding mechanism of SAP/SH2D1A. The presence of three 'prongs' in the SLAM-pY281 peptide makes it possible to leave out one 'prong' (e.g. Thr-2) in binding and yet maintain a high affinity for SAP/SH2D1A. This is certainly not the case for the SLAM-Y281 peptide, where only two prongs (e.g. Thr-2 and Val+3) are present. Since mutation of Thr53 is associated with XLP, this observation underscores the importance of phosphorylation-independent interaction in SAP/SH2D1A function and in the pathogenesis of XLP.

The only mutant that displays a greater affinity for the SLAM-Tyr281 peptide than for the phosphorylated peptide is C42W. Since Cys42 forms part of the binding pocket for pTyr, its replacement by a bulky Trp residue likely creates a steric clash with the phosphate group of pTyr in the SLAM-pY281 peptide. The same mutation, however, may have a less detrimental effect on binding to the less bulky Tyr residue in the SLAM-Y281 peptide.

Mutants Y7C, S28R, Q99P, P101L, V102G and VTail do not involve residues at the peptide–protein interface, and their reduced affinities for both peptides are likely to be attributable to protein folding and stability. Mutant VTail is extremely unstable both in vitro and in vivo (Morra et al., 2001b). Surface area calculations reveal that the side chains of residues Tyr7, Gln99 and Pro101 are largely buried in both complexes, while those of Ser28 and Val102 are completely inaccessible to solvent (data not shown). Although residues Gln99, Pro101 and Val102 are not at the core of the protein, they may nonetheless be involved in stabilizing the overall fold of the protein by interacting with the protein’s central β-sheet. Consistent with this, solvent exchange experiments also showed that the amide hydrogens of these residues are well protected (Figure 5).

**Discussion**

SH2 domains are a group of structurally conserved protein modules of ~100 amino acids, which, in general, bind selectively to phosphotyrosine-containing sequences.
of the SH2 domain for a given physiological ligand. This may be the result of steric clashes, loss of favorable interactions or changes in protein dynamics. The observation that all mutants examined in the present study displayed markedly reduced affinities for the non-phosphorylated SLAM peptide suggests that phosphorylation-independent interactions are important for proper functioning of SAP/SH2D1A and likely also for the pathogenesis of XLP. In agreement with this hypothesis, a disease-causing mutation, T33I, results in drastically reduced affinity for the unphosphorylated SLAM-Y281 peptide while retaining full capacity to bind the phosphorylated peptide.

The considerable dynamics observed in SAP/SH2D1A contribute to its ability to accommodate non-optimal sequences. Structural flexibility can improve the stability of various complexes by modulating the binding surface and by maintaining conformational entropy in the absence of favorable interactions. The flexibility of the two segments that regulate the peptide binding cleft, the BG and EF loops, allows for considerable structural plasticity, enabling the extreme versatility of binding specificity observed in this modular domain.

Materials and methods

Synthesis of peptide SPO Ts on derivatized cellulose membranes

A set of undecamer SLAM-Tyr281 peptide analogs was assembled on a derivatized cellulose membrane using Auto-Spot Robot ASP 222 (Abimed, CA) according to standard solid-phase peptide synthesis chemistry. The SPOTs peptide sheet was moistened sequentially with ethanol and water prior to screening using either purified GST–SAP/SH2D1A or GST alone. Specifically, the peptide sheet was washed three times with Tris-buffered saline-T (TBS-T) buffer containing 20 mM Tris–HCl, 140 mM NaCl and 0.1% (v/v) Triton X-100 pH 7.6, and blocked with 5% skimmed milk in TBS-T for 1 h at room temperature (RT). GST–SAP/SH2D1A (~1.0 µM) or GST protein was added directly into the blocking solution and the peptide sheet was incubated at RT for an additional 1 h. The sheet was then washed three times, 5 min each, with TBS-T and once with TBS prior to addition of anti-GST antibodies. After incubation of 30 min at RT, the cellulose sheet was washed three times with TBS and developed using the ECF western blotting kit (AmershamPharmacia Biotech) following the manufacturer’s recommendations and documented using a Fluor-S Multi-Imager (Bio-Rad). For reprobing, the peptide sheet was stripped by treating it sequentially with buffer A [containing 8.0 M urea, 1% (w/v) SDS and 0.5% (v/v) β-mercaptoethanol] and buffer B [containing 10% (v/v) acetic acid and 50% (v/v) ethanol], followed by several washes with deionized water.

NMR spectroscopy and structure calculation

Full-length SAP/SH2D1A (128 residues) cloned into a pET3a vector was expressed in Escherichia coli strain BL21(DE3) and purified as described earlier (Li et al., 1999) with 1H or 13C labeling as required. All NMR samples contained 1.0–1.5 mM protein, 20 mM sodium phosphate pH 6.0 and 100 mM NaCl, in H2O or D2O. Unlabeled peptides (n-PY, RSK1TIPyA amide; and n-Y-C, RKS1TYA/QQQK) were synthesized and purified as reported previously, and titrated into protein samples to form 1:1 peptide–protein complexes as monitored using HSQC spectra.

NMR experiments were carried out at 30°C on either a Varian Inova 500 or 600 MHz spectrometer equipped with z-axis pulsed-field gradients. Data were processed with NMRPipe (Delaglio et al., 1995) and analyzed with NMRView (Johnson et al., 1994). Backbone chemical shift assignments were obtained by analyzing CBCA(CO)NH (Grzesiek and Bax, 1992), HNCACB (Witek and Muller, 1993), HNCO (Kay et al., 1994) and (Hβ)CBCA(CO)C(A)Hα (Kay, 1993) spectra. Initial side-chain assignments were extracted from 3D (H(CC)(CO)NH-TOSCY and H(CC)(CO)NH-TOSCY experiments (Montelione et al., 1992; Grzesiek et al., 1993), as well as 2D (Hβ)(C)β(Cy3)H6 and (Hβ)(C)β(Cy3)CτHe experiments (Yamazaki et al., 1993) for aromatic
residues. 3D HCCH-TOCSY (Kay et al., 1993) and CN-NOESY (Pascal et al., 1994) experiments were used to further refine side-chain assignments. Stereo-specific methyl assignments for Val and Leu residues were obtained using a 10% 13C-labeled sample, according to the method of Neri et al. (1989). Stereo-specific methylene assignments and 1H dihedral angles were derived using a combination of a short mixing time CN-NOESY (50 ms) and 1H HNHB and HNC(CC)HB experiments (Grzesiek et al., 1992). Peptide 1H chemical shift assignments were obtained using 2D double-filtered TOCSY and NOESY (Ikura and Bax, 1992) experiments.

Distance restraints for structural calculations were obtained from CN-NOESY (50 ms and 150 ms) data and a single-filtered NOESY (200 ms; Zwanenburg et al., 1997) for peptide-protein restraints. Note that a 200 ms mixing time was only used in the case of the intermolecular NOE experiment, which required a longer mixing time to build up appreciable intensity. In the case of the SH2 domain structure determination, the shortest mixing time for observing significant NOEs was used. Hydrogen-bonding partners were assigned based on slowly exchanging backbone amides. Dihedral angle (φ and ψ) restraints were calculated from chemical shifts using TALOS (Cornilescu et al., 1999). Manually assigned NOE restraints, dihedral angle restraints and hydrogen-bonding restraints were used as input to calculate 100 structures for each complex using CNS v.1.0 (Brünger et al., 1998) simulated annealing protocols. The 10 lowest energy structures were used as input structures for the automated NOE assignment program ARRA (Nilges et al., 1998) along with stereo-specific assignments, dihedral angle restraints, hydrogen-bonding restraints and peak lists from NOESY spectra (no manually assigned distance restraints were used as input for these calculations). A total of 100 final structures were calculated (without water) for each complex and the 20 lowest energy structures were retained. Coordinates of the energy minimized average structure for each complex have been deposited in the Protein Data Bank, accessible under code 1K6A for the n-pY complex and 1K7A for the n-Y-c complex.

Measurement of backbone amide proton exchange rates by NMR

Slow water–amide proton exchange rates were measured by lyophilizing an H2O protein sample and then redissolving it in D2O. 1H–15N HSQC spectra were then recorded at various time points and the decay of amide proton signals due to D2O exchange was measured. A number of amide protons displayed no changes in peak intensity even after 6 weeks, indicating extremely slow exchange with D2O. The fastest exchange rate measurable using this technique was that of T15 from the SAP/SH2D1A–n-pY complex, which had a time constant of 174 s. Amides that exchanged more rapidly than this disappeared too quickly to allow the acquisition of a decay curve. More rapid exchange rates were observed using a CLEANE-PM pulse sequence, which selectively transfers magnetization from H2O to amide groups for detection (Hwang et al., 1998). The slowest exchange rate that could be measured by this technique belonged to the backbone amide of D26 from the SAP/SH2D1A–n-Y-c complex, which had an exchange time constant of 5 s. The most rapid exchange rate was seen for the amide of R78, with a time constant of 0.029 s. The signal for the R78 amide is noticeably broadened in the conventional 1H–15N HSQC spectrum. Hence, exchange rates spanning nine orders of magnitude could be measured. Protection factors were calculated using predicted random coil exchange rates, as described by Englander and co-workers (Bai et al., 1993).

Generation of SAP/SH2D1A mutants and measurement of affinities for SLAM peptides

SAP/SH2D1A mutants were generated by PCR using primers incorporating the point mutations found in XLP patients and a template bearing wild-type cDNA in a pfGE4-IT2 vector. All mutations were confirmed by DNA sequencing. Expression and purification of the mutant proteins were carried out according to established procedures (Li et al., 1999). The affinities of each mutant for the phosphorylated and non-phosphorylated SLAM-Tyr281 peptides were measured by fluorescence polarization performed on a Beacon 2000 fluorescence polarization system (Panvera Co.; Li et al., 1997).

Acknowledgements

We thank Qin Liu for assistance in the preparation of Figure 3. This work was supported by grants from the National Cancer Institute of Canada (NCIC) to S.-C.L., J.D.F.-K. and T.P. P.M.H. is the recipient of a scholarship from the Canadian Institute of Health Research. S.-C.L. is a Research Scientist of NCIC with funds made available by the Cancer Society of Canada. T.P. and L.E.K. are international scholars of the Howard Hughes Research Institutes.

References

Li,S.-C., Songyang,Z., Vincent,S.J.F., Zwahlen,C., Wiley,S., Cantley,L.,


Received September 3, 2001; revised November 14, 2001; accepted December 5, 2001.