Molecular Cell, Vol. 4, 555–561, October, 1999, Copyright ©1999 by Cell Press

Crystal Structures of the XLP Protein SAP Reveal a Class of SH2 Domains with Extended, Phosphotyrosine-Independent Sequence Recognition

Florence Poy,[1] Michael B. Yaffe,[3,4] Joan Sayos,[7] Kumkum Saxena,[1,6] Massimo Morra,[7] Janos Sumegi,[2] Lewis C. Cantley,[3,5] Cox Terhorst,[7] and Michael J. Eck[1,6,8]

[1]Department of Cancer Biology
Dana-Farber Cancer Institute
44 Binney Street
Boston, Massachusetts 02115
[2]Department of Pathology and Microbiology
University of Nebraska Medical Center
Omaha, Nebraska 68198-5660
[3]Division of Signal Transduction
[4]Departments of Medicine and Surgery
Beth Israel Deaconess Medical Center
Boston, Massachusetts 02215
[5]Department of Cell Biology
[6]Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
[7]Division of Immunology
Beth Israel Deaconess Medical Center
Harvard Medical School
Boston, Massachusetts 02215

Summary

SAP, the product of the gene mutated in X-linked lymphoproliferative syndrome (XLP), consists of a single SH2 domain that has been shown to bind the cytoplasmic tail of the lymphocyte coreceptor SLAM. Here we describe structures that show that SAP binds phosphorylated and nonphosphorylated SLAM peptides in a similar mode, with the tyrosine or phosphotyrosine residue inserted into the phosphotyrosine-binding pocket. We find that specific interactions with residues N-terminal to the tyrosine, in addition to more characteristic C-terminal interactions, stabilize the complexes. A phosphopeptide library screen and analysis of mutations identified in XLP patients confirm that these extended interactions are required for SAP function. Further, we show that SAP and the similar protein EAT-2 recognize the sequence motif TIpYXX(V/I).

Introduction

X-linked lymphoproliferative syndrome (XLP), or Duncan's disease, is an inherited syndrome characterized by extreme susceptibility to Epstein-Barr Virus (EBV) infection (Purtilo et al., 1975; Seemayer et al., 1995). Affected boys usually die of fulminant infectious mononucleosis or of EBV-induced lymphomas. The genetic basis of this disease has recently been elucidated (Coffey et al., 1998; Nichols et al., 1998; Sayos et al., 1998), but the function of the XLP gene product and the molecular physiology of the disease remain unclear. The *XLP* gene encodes SAP (also called SH2D1A), a small protein of 128 residues that consists of just a single SH2 domain with a short C-terminal extension (Coffey et al., 1998; Nichols et al., 1998; Sayos et al., 1998). SAP has been shown to bind the cytoplasmic tail of SLAM (Sayos et al., 1998), a self-ligand glycoprotein found on the surface of activated B and T cells (Cocks et al., 1995; Aversa et al., 1997). Surprisingly, SAP binds a site encompassing tyrosine 281 in SLAM irrespective of the phosphorylation state of this site (Sayos et al., 1998).

SAP is unusual among SH2 domains for its ability to bind a nonphosphorylated partner protein. All previously characterized SH2 domains bind with high affinity only to specific tyrosine-phosphorylated sequences; specificity is conferred by recognition of 3–5 residues just C-terminal to the phosphotyrosine (Kuriyan and Cowburn, 1997). The single domain architecture of SAP is also unusual; SH2 domains are typically found in multidomain signaling enzymes or in adaptor proteins, where they mediate protein–protein interactions and regulate associated catalytic subunits (Pawson, 1995). The lack of additional domains in SAP suggests that it is unlikely to function in either of these roles. It could, however, function as a signaling inhibitor, by blocking or regulating binding of other signaling proteins to particular docking sites. Indeed, SAP has been shown to block recruitment of the SHP-2 phosphatase to the tail of SLAM (Sayos et al., 1998).

In order to better understand the function of this atypical SH2 protein, we have studied its binding specificity and determined its three-dimensional structure, alone and in complex with SLAM peptides. Both phosphorylated and nonphosphorylated SLAM peptides bind in a similar mode, with Tyr-281 or phosphoTyr-281 inserted into the classical phosphotyrosine-binding pocket. Unexpectedly, we observe that specific interactions with amino acid residues N-terminal to Tyr-281, in addition to more characteristic C-terminal interactions, stabilize the complexes. Missense mutations identified in XLP patients cluster along the length of the SLAM-binding site. The structures, together with data obtained by screening a degenerate phosphopeptide library, show that SAP recognizes the sequence motif TIpYXX(V/I), which is found in the cytoplasmic domains of a restricted number of T, B, and NK cell surface receptors. Surprisingly, the SAP-like protein EAT-2 (Thompson et al., 1996) recognizes the same motif. These data support the notion that SAP and EAT-2 are endogenous inhibitors that regulate recruitment of SH2 signaling proteins to specific docking sites (Sayos et al., 1998).

Results

Structure of SAP and SAP/SLAM Complexes

We have determined atomic resolution structures of SAP (residues 1–104) alone or complexed with 14 residue peptides encompassing Tyr-281 in SLAM (Sayos et al., 1998). We crystallized SAP in complex with both phosphorylated (pY281) and nonphosphorylated (Y281)

[8]To whom correspondence should be addressed (e-mail: eck@red.dfci.harvard.edu).

Table 1. Data Collection, Phasing, and Refinement Statistics

	SAP/Y281	SAP/Y281 MeHg Derivative	SAP/pY281	SAP (Unliganded)
Data Collection Statistics				
Resolution (Å)	22–1.1	22–1.5	20–1.8	22–1.4
Space group	$P2_1$	$P2_1$	C2	P1
Unit cell (Å)	a = 25.9, b = 42.9, c = 44.9 $\beta = 98.3°$	a = 25.9, b = 42.5, c = 44.9 $\beta = 97.1°$	a = 100.2, b = 49.0, c = 55.6 $\beta = 94.2°$	a = 32.48, b = 62.3, c = 65.1 $\alpha = 65.7° \beta = 78.7° \gamma = 89.7°$
Molecules/a.s.u.	1	1	2	4
R_{sym} (%)	4.3	7.1	7.6	6.2
Reflections (total/unique)	134,878/34,081	46,063/14,014	113,987/23,946	274,082/83,110
Completeness (%)	86.0	89.6	94.8	92.9
Phasing Statistics				
R_{iso} (%)		53.6		
Phasing Power (centric/acentric)		1.83/2.75		
FOM		0.57		
Number of sites		2		
SHELXL Refinement Statistics				
Resolution range (Å)	8–1.1		10–1.8	10–1.4
R_{cryst}/R_{free} (all data, %)	12.7/16.5		17.2/24.0	14.7/19.5
R_{cryst}/R_{free} (F > 4Sig(F),%)	11.5/15.3		15.3/22.3	13.9/18.7
Rmsd bond lengths/angles	0.017 Å/0.032 Å		0.007 Å/.025 Å	0.011 Å/0.029 Å
Number of data[a] parameters/ restraints	32,715/10,326/12,194		22,677/8827/7370	78,862/33,861/41,338
Number of nonhydrogen atoms	1,147		2,203	3,772
Solvent atoms	236		413	577

[a] This is the number of unique reflections actually used in refinement and does not include the ~5% of reflections omitted for calculation of R_{free}.

SLAM peptides (see Table 1 and Experimental Procedures). SAP has the characteristic SH2 fold (Kuriyan and Cowburn, 1997), which includes a central β sheet with α helices packed against either side (Figure 1). The phosphorylated and nonphosphorylated SLAM peptides bind in the same manner across the surface of the domain (Figures 1, 2A, and 2B). Classical SH2 domains bind phosphopeptides in a "two-pronged" fashion; the phosphorylated tyrosine residue binds in a pocket on one side of the central sheet, and the 3–5 residues C-terminal to it bind in a pocket or groove on the opposite side (Kuriyan and Cowburn, 1997). Both SLAM peptides retain these general binding features, with Tyr-281 or phosphoTyr-281 extending into the phosphotyrosine-binding pocket. However, SAP forms unexpected additional interactions with SLAM residues N-terminal to Tyr-281. The SLAM peptides make a parallel β sheet interaction with the βD strand of the domain, and the side chains of residues at positions pY −3 and pY −1 (we refer to residues in both peptides by their position relative to Tyr-281) intercalate with hydrophobic residues in strand βD (Figures 2A and 3). Additionally, Thr-279 (at position pY −2 in the SLAM peptide) hydrogen bonds with Glu-17 and a buried water molecule (Figures 1A, 2D, and 3E). Significant N-terminal specificity has not previously been demonstrated in SH2/phosphopeptide recognition. Here, these N-terminal interactions likely provide the additional binding energy that allows recognition of the nonphosphorylated SLAM sequence.

In both complexes, Tyr-281 is positioned similarly in the phosphotyrosine-binding pocket, and the phosphotyrosine-binding BC loop adopts the same conformation (Figures 2B and 2C). The phosphorylated Tyr-281 is coordinated in a manner not unlike that observed in other SH2 complexes; the universally conserved arginine 32 forms the expected bidentate hydrogen bonds with phosphate oxygens (Figures 1A and 2D). In the unphosphorylated complex, ordered water molecules replace the interactions of the missing phosphate oxygens (Figure 2E). Arg-32 hydrogen bonds with one of these waters and with the hydroxyl of the unphosphorylated tyrosine. Additionally, Arg-32 coordinates two buried water molecules, one of which also hydrogen bonds with Thr (pY −2) in the bound peptide (Figures 2D and 2E). Thus, binding of the nonphosphorylated peptide does not create an energetically unfavorable solvent-excluding environment for this positively charged residue.

The five residues following Tyr-281 are coordinated by interactions with the EF and BG loops and the βD strand of the central sheet. These C-terminal interactions are similar to those seen in other SH2/phosphopeptide complexes. The backbone conformation of the +1 to +3 residues is identical to that observed in high-affinity Src family SH2 complexes, and Val-284 (in the pY +3 position) inserts into a hydrophobic cleft in a manner analogous to that of an isoleucine at this position in Src family SH2 complexes (Eck et al., 1993; Waksman et al., 1993). In the absence of a bound peptide, the EF and BG loops come together to close this hydrophobic cleft (Figure 2C).

Missense Mutations in XLP Patients

Examination of missense mutations identified in XLP patients (Coffey et al., 1998; J. Sumegi, unpublished data) suggests that the full complement of binding interactions we observe is critical for SAP function. These point mutations cluster in three regions: in and around the pTyr-binding pocket, in the Val (+3) pocket, and on the

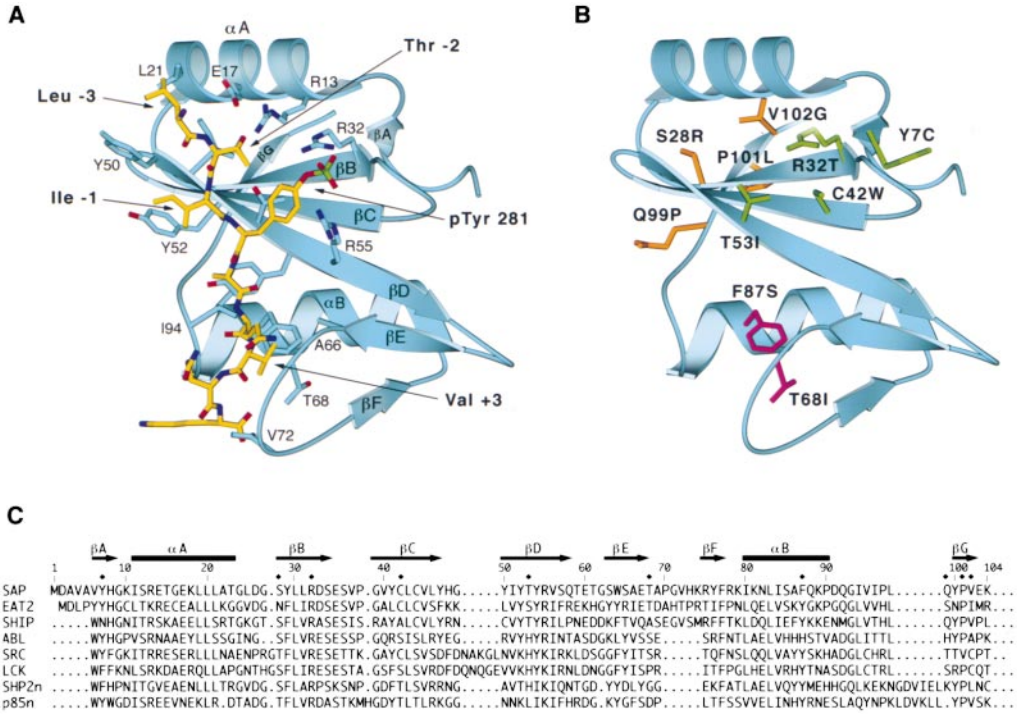

Figure 1. Structure of SAP and the Location of Missense Mutations Identified in XLP Patients

(A) Ribbon diagram showing the SAP/SLAM pY281 complex. The bound phosphopeptide is shown in a stick representation (yellow). Selected SAP residues that form the binding site are shown in blue. Elements of secondary structure are labeled using the standard SH2 domain nomenclature (Eck et al., 1993). Note that the pY −3 to pY −1 residues of the peptide make a parallel β sheet interaction with strand βD; the side chains of these peptide residues make hydrophobic contacts with Tyr-50, Ile-51, and Tyr-52 in strand βD, and with Leu-21. Thr (pY −2) in the peptide hydrogen bonds with Glu-17 and with a buried water molecule. The phosphotyrosine is coordinated in a manner similar to that observed in the N-terminal domain of SHP-2, and as in SHP-2, the phosphate group is rotated "above" the plane of the phosphotyrosine ring. Interestingly, arginine 13 (at position αA2), which is conserved in almost all SH2 domains and usually contributes to phosphotyrosine coordination, does not participate in phosphate binding in the SAP complex. Instead, arginine 55 (βD6) hydrogen bonds with the phosphate group. C-terminal to phosphotyrosine, Val(pY +3) binds in a mostly hydrophobic cleft.

(B) Point mutations identified in XLP patients cluster along the peptide-binding site and at the back of the domain. Mutations that would be expected to directly disrupt the phosphotyrosine-binding pocket are shown in green, and those that would disrupt C-terminal interactions in magenta. The remaining mutations (gold) are remote from the peptide-binding surface and may destabilize the folded protein (see text).

(C) Structure-based sequence comparisons of human SAP, murine EAT-2, and other SH2 domains. Elements of secondary structure are indicated above the alignment. Numbering corresponds to human SAP. The black diamonds indicate the mutations illustrated in (B).

"back" of the domain, largely in strand βG (Figure 1B). Some mutations in the first and second categories are likely to grossly disrupt peptide binding (including Arg-32 → Thr and Cys-42 → Trp), but others would be expected to have a more modest effect. For example, modeling of the Thr-53 → Ile substitution suggests that it would disrupt interactions with the N-terminal residues of the SLAM peptide, but not with the phosphotyrosine and C-terminal residues. Consistent with this hypothesis, an isoleucine or leucine is found at this position in other SH2 domains that exhibit the classical two-pronged binding mode. In the pY +3 binding pocket, the Thr-68 → Ile substitution (Coffey et al., 1998) may be compatible with recognition of a valine but might decrease binding affinity by stabilizing the "closed" conformation of the +3 pocket that we observe in the unliganded SAP structure. We predict that most of the substitutions in the third category decrease protein stability (rather than disrupt an additional, as yet unknown, function of the domain). For example, the Val-102 → Gly and Gln-99 → Pro mutations likely disrupt hydrophobic and hydrogen bonding interactions, respectively, that stabilize the βG strand.

SAP and EAT-2 Bind the Motif T-I-Y-X-X-(V/I)

To estimate the level of selectivity conferred by interactions of SAP with N- and C-terminal residues in the SLAM motif, we analyzed binding of a GST–SAP fusion protein to a degenerate phosphopeptide library (Songyang et al., 1993). The library contained four degenerate positions preceding and four degenerate positions following an orienting phosphotyrosine residue. In sharp distinction to all other SH2 domains (Songyang et al., 1993; Cantley and Songyang, 1994), SAP showed significant selectivity both N- and C-terminal to the phosphotyrosine, selecting Thr in the pY −2 position, Ile in the pY −1 position, and Val in the pY +3 position (Figure 4A). Based upon these results, and the interactions we observe in the structure, we deduce the binding motif T-I-Y-X-X-(V/I). This motif occurs twice in SLAM and is also found in the cytoplasmic tails of other cell surface proteins with similarity to SLAM, including 2B4, Ly9, and

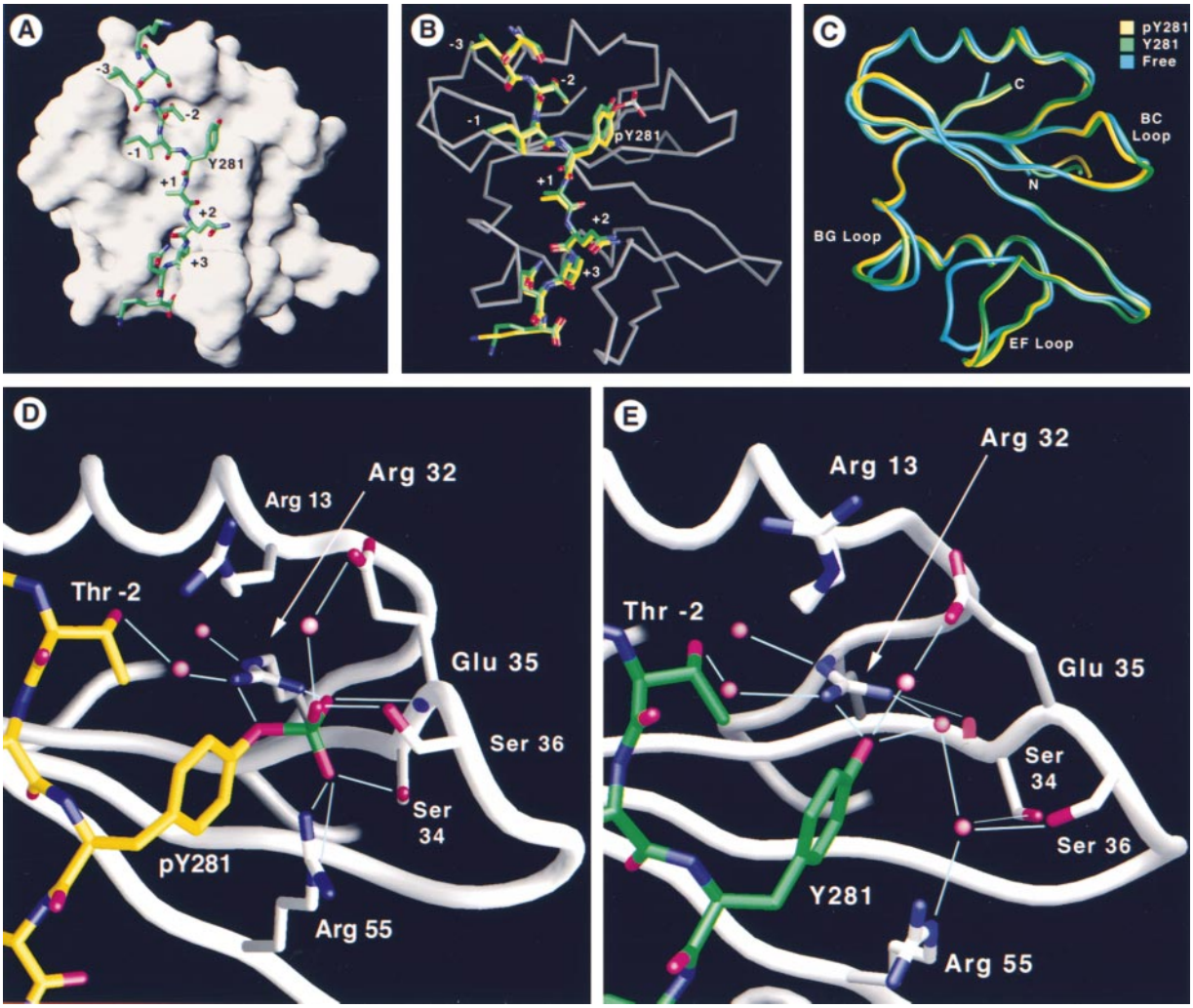

Figure 2. Structure and Comparisons of the SAP/SLAM Y281 Complex

(A) Surface representation of the SAP domain with the bound nonphosphorylated peptide shown in green. Hydrophobic residues at the −1 and −3 positions of the peptide intercalate with hydrophobic and aromatic residues on the surface of the domain (see also [D] and Figure 1A). C-terminal to phosphotyrosine, Val+3 is buried in a mostly hydrophobic groove.

(B) Superposition of the phosphorylated and nonphosphorylated peptides shows that they adopt an essentially identical conformation. An alpha-carbon trace of the domain is shown in gray.

(C) Superposition of the unliganded domain (blue) and the phosphopeptide (yellow) and nonphosphorylated peptide complexes (green). In the absence of bound peptide, the EF and BG loops fold inward to close the hydrophobic +3 binding groove. The conformation of the phosphotyrosine-binding pocket is essentially the same in all structures. In the unliganded structure, a sulfate ion occupies the position of the phosphate group in the phosphopeptide complex.

(D) Detail of the phosphotyrosine-binding pocket in the SLAM/pY281 complex. Red spheres represent ordered water molecules. The pY281 peptide is shown in yellow. Thin cyan lines indicate potential hydrogen bonds. Note that Arg-13 is poorly ordered and does not participate in phosphotyrosine coordination.

(E) Detail of the phosphotyrosine-binding pocket in the nonphosphorylated SLAM/Y281 complex. Red spheres represent ordered water molecules. The Y281 peptide is shown in green. Thin cyan lines indicate potential hydrogen bonds. Note that Arg-32 organizes an extensive network of hydrogen bonds in spite of the lack of phosphorylation of Tyr-281.

CD84. The extracellular domains of these proteins are homologous to SLAM, and their genes map with the *SLAM* gene to the same area of chromosome 1 in human and mouse (Brown et al., 1998). Perhaps more importantly, the ligand for 2B4, which is expressed on the surface of $CD8^+$ T cells and NK cells, is CD48 (Brown et al., 1998; Latchman et al., 1998), another membrane protein that maps to chromosome 1 and whose ectodomain is homologous to that of SLAM. CD48 is expressed as a PI-linked protein on activated B cells and was originally identified as the major B cell protein induced after EBV infection (Fisher and Thorley-Lawson, 1991). SAP may therefore regulate both SLAM/SLAM and CD48/2B4 signaling during infectious mononucleosis.

The unique selectivity of the SAP SH2 domain for residues N-terminal to the phosphotyrosine is readily understood in the context of the crystal structures: the constellation of residues that interacts with the pY −1

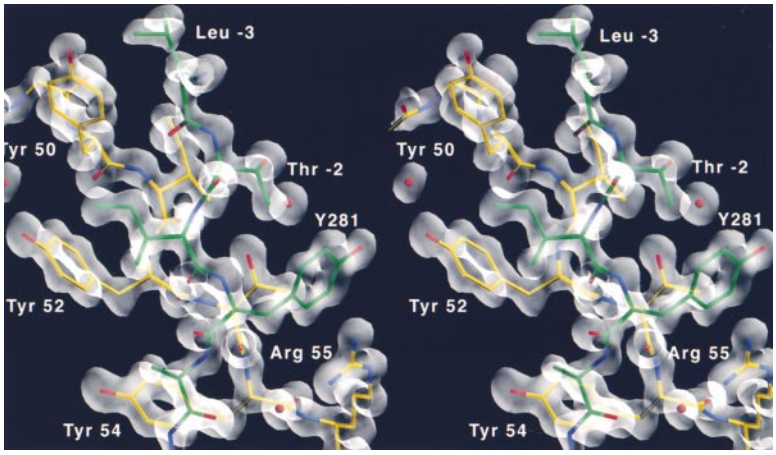

Figure 3. Stereo View of a Representative Portion of the Electron Density Map for the SAP/SLAM Y281 Complex

The transparent surface represents a $2F_o - F_c$ electron density map, contoured at 1.8 σ, and calculated with data extending to 1.15 Å resolution (see Experimental Procedures). Note the parallel β sheet interaction between the βD strand of the SH2 domain (yellow) and the N-terminal portion of the peptide (green). A grossly similar β sheet interaction was noted in the structure of a p85 SH2 domain complexed with a c-Kit receptor phosphopeptide, but this interaction appeared to be stabilized primarily by crystal-packing contacts (Nolte et al., 1996).

to pY −3 residues in the SLAM peptide are not present in most SH2 domains. In particular, the presence of three hydrophobic or aromatic residues followed by a serine or threonine in the first four positions of strand βD is unusual ("YIYT" in SAP, Figure 1C). However, this pattern is present in the SH2 domains of EAT-2 and SHIP (Lioubin et al., 1996), two SH2 domains most similar in sequence to SAP. EAT-2 was identified by virtue of its upregulation in Ewing's sarcoma cells, and like SAP, it consists solely of an SH2 domain with a short C-terminal tail (Thompson et al., 1996). Peptide library screening with EAT-2 reveals a pattern of residue selectivity nearly identical to that observed with the SAP SH2 domain (Figure 4B). These results define SAP and EAT-2 as members of a novel class of SH2 domains whose binding specificity extends on both sides of the phosphotyrosine residue. Furthermore, this extended recognition allows ligand binding to occur in a phosphorylation-independent manner. Both SAP and EAT-2 have been found to bind to nonphosphorylated motifs (Sayos et al., 1998; unpublished data). In fact, SAP binds nonphosphorylated SLAM with an affinity comparable to that measured for other SH2 domains with their cognate phosphopeptides. The Src SH2 domain, for example, recognizes a high-affinity phosphopeptide ligand with a dissociation constant of 550 nM, and the p85N SH2 domain of PI-3 kinase binds the pY751 site in the PDGF receptor with a dissociation constant of 470 nM (Ladbury et al., 1995). Using fluorescence polarization, we measure a binding constant of 650 nM for SAP with a nonphosphorylated SLAM Y281 peptide (data not shown). We observe approximately 5-fold tighter binding to the phosphorylated SLAM peptide (∼120 nM; data not shown).

Discussion

Although binding of nonphosphorylated sequences by SAP is unusual among SH2 domains, the phenomenon has been observed in phosphotyrosine-binding (PTB) domains (Zhang et al., 1997). Like SAP, PTB domains recognize longer sequences than classical SH2 domains. However, in PTB domains the arginine residues that coordinate phosphotyrosine differ among domains and are in some cases absent, suggesting that some

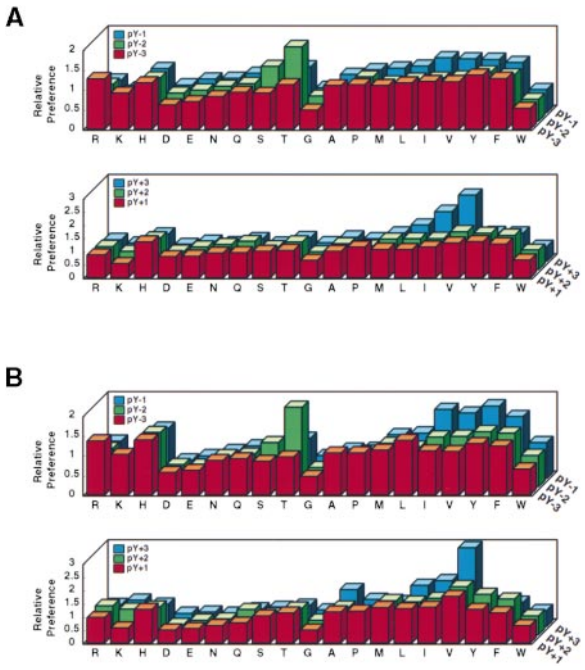

Figure 4. The SAP and EAT-2 Proteins Selectively Bind Phosphopeptides with the Motif TIpYXX(V/I)

(A) SAP selects for threonine in the pY −2 position (upper panel, green bars), and for valine or isoleucine in the pY +3 position (lower panel, blue bars). SAP also shows modest selection for isoleucine and other hydrophobic residues in the pY −1 position (upper panel, blue bars).

(B) EAT-2 selects for threonine in the pY −2 position (upper panel, green bars), for isoleucine, valine, or aromatic residues at pY −1 (upper panel, blue bars), and for valine in the pY +3 position (lower panel, blue bars). In (A) and (B), the histograms represent the abundance of each amino acid residue in the pool of specifically bound peptides, relative to its abundance in the starting degenerate library (see Experimental Procedures). Amino acid residues are indicated using the one-letter code. Sequence positions N-terminal to the orienting phosphotyrosine are presented in the upper panels of (A) and (B) and are labeled pY −3 (red bars), pY −2 (green bars), and pY −1 (blue bars). Positions C-terminal to the orienting phosphotyrosine are shown in the lower panels and are labeled pY +1 (red bars), pY +2 (green bars), and pY +3 (blue bars). The pY −4 and pY +4 positions are not shown, as no significant selectivity was observed.

PTB domains fulfill functions unrelated to phosphotyrosine signaling. The X11 PTB domain, for example, lacks these arginines and binds a nonphosphorylated segment of the β-amyloid precursor protein (Zhang et al., 1997). In contrast, SAP retains the ability to bind tightly to phosphorylated SLAM sequences.

What physiological purpose might be served by recognition of both phosphorylated and nonphosphorylated sequences? The structural studies reported here and the protein interactions described previously (Sayos et al., 1998) suggest two, not mutually exclusive, functions for the extended interactions between SAP or EAT-2 and their respective partners. First, binding with nonphosphorylated sequences could serve to block phosphorylation of these sites. Second, tight and highly specific binding to phosphorylated sequences could prevent downstream signaling molecules from binding to SH2 docking sites (Sayos et al., 1998). Taken together, these observations are therefore consistent with a model in which SAP and EAT-2 are natural inhibitors or regulators of the physiological role of a small family of receptors on the surface of T, B, and NK cells.

Experimental Procedures

Expression, Purification, and Crystallization

Human SAP (residues 1–104) was expressed in *E. coli* strain BL21(DE3) using the T7 expression plasmid pRSET (Invitrogen). Cultures (4× 1 liter) were grown at 25°C, induced overnight with IPTG, and harvested by centrifugation. Cell pellets were lysed by sonication in 50 mM Tris (pH 8.0), 150 mM NaCl, 10 mM DTT, and 1 mM PMSF. The SAP protein was purified to homogeneity from clarified lysates using cation exchange chromatography (S-Sepharose Fast-Flow, Pharmacia) followed by phosphotyrosine-affinity chromatography (Eck et al., 1993). The purified protein was concentrated to 20 mg/ml in storage buffer (20 mM HEPES [pH 7.5], 150 mM NaCl, and 10 mM DTT).

All SAP crystals were grown in hanging drops at 22°C by combining 2.5 μl of protein in storage buffer with 2.5 μl of a well solution. The unliganded crystals were grown with a well solution containing 1.64 M ammonium sulfate, 100 mM sodium citrate (pH 5.6), and 10 mM DTT. For cocrystallization with the nonphosphorylated SLAM Y281 peptide (VEKKSLTIYAQVQK), a 1.5-fold molar excess of peptide was added to the SAP protein in storage buffer. The Y281 complex was crystallized over a well solution containing 30% PEG 8000, 20% glycerol, 100 mM HEPES, (pH 8.0), and 10 mM DTT. The phosphorylated SLAM pY281 peptide (VEKKSLTIpYAQVQK) complex was obtained with 3-fold molar excess of peptide over a well solution containing 30% PEG 8000, 20% glycerol, 100 mM sodium citrate (pH 5.6), and 10 mM DTT. All crystals were transferred to stabilizing solutions containing the components of their respective crystallization buffers and at least 20% glycerol. The mercury derivative was prepared by soaking a SAP/Y281 crystal overnight in stabilizing solution containing 0.8 mM methyl mercury nitrate and no DTT. All diffraction data were recorded at −165°C. Unit cell constants and other structure determination statistics for all crystal forms are given in Table 1. Diffraction data for the unliganded SAP crystals and the SAP/Y281 complex were recorded using the Princeton 2k CCD detector on the A1 beam line at CHESS. The SAP/pY281 data were recorded using a Mar Research image plate detector mounted on a Rigaku rotating anode source with mirror optics. All diffraction data were integrated and scaled using the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997).

Structure Determinations

The structure of the SAP/Y281 complex was determined by isomorphous replacement using a single mercury derivative (Table 1). Heavy atom sites were located with difference Patterson methods using the CCP4 (Collaborative Computational Project Number 4, 1994) program package, and phases were calculated using MLPHARE with data extending to 1.5 Å resolution. Structure factor phases were improved with solvent flattening and histogram matching using DM (Collaborative Computational Project Number 4, 1994) and then extended to the limit of the native data set (1.15 Å) using the program ARP/wARP (Version 5; Lamzin and Wilson, 1997). A nearly complete (~90%) atomic model was constructed in a totally automated manner using ARP/wARP (Lamzin and Wilson, 1997); the model was manually completed using the molecular graphics program O (Jones and Kjeldgaard, 1997) and refined using ARP/wARP (Lamzin and Wilson, 1997) and REFMAC (Collaborative Computational Project Number 4, 1994) to a crystallographic R value of 19.8%. The model was further refined to an R value of 11.5% (12.7% for all data) with SHELX97 (Table 1). The final SHELXL refinement included "riding" hydrogen atoms, a diffuse solvent correction, and restrained anisotropic thermal parameters for all atoms. The final model included residues 1–104 of SAP, residues 276–286 of the SLAM peptide, and 236 water molecules. Nine SAP residues were modeled in two conformations.

The SAP/Y281 structure, sans peptide, was used as a search model to determine the structure of the SAP/pY281 complex by molecular replacement using the program AmoRe (Navaza, 1992). Unambiguous rotation and translation solutions were were readily obtained for the two molecules in the asymmetric unit of the SAP/pY281 complex. Electron density maps calculated after an initial round of refitting and simulated-annealing refinement revealed very clear density for the bound peptide. At this stage, residues 276–286 of the phosphopeptide were constructed, and solvent structure was built using ARP. This model was refined to an R value of 19.9% using XPLOR (Brunger, 1992) and was further refined with SHELXL to an R value of 17.2%. Noncrystallographic symmetry restraints were applied initially but were released in the final stages of refinement. The final SHELXL refinement included an overall anisotropic scale correction and diffuse solvent correction. Refinement statistics are presented in Table 1.

The unliganded SAP structure was determined by molecular replacement with AmoRe (Navaza, 1992). Unambiguous rotation and translation solutions were obtained for each of the four molecules in the asymmetric unit using the SAP/Y281 structure, with the peptide removed, as a search model. The structure was manually rebuilt using O (Jones and Kjeldgaard, 1997), and solvent structure was built and refined using ARP/wARP and REFMAC. The unliganded structure was further refined using SHELX-97 (Sheldrick and Schneider, 1997). In the final refinement, noncrystallographic symmetry restraints were released, all atoms were made anisotropic, and a diffuse solvent correction was applied. The final R value was 14.7% for all data (Table 1).

Peptide library screen

The phosphopeptide library methodology has been described previously (Songyang et al., 1993). The library used here contained $\sim 1.7 \times 10^{10}$ distinct peptides of general sequence MAXXXXpYXXXXAKKK, where X denotes all natural amino acids except cysteine, and pY is the orienting phosphotyrosine residue. The degenerate library (0.6 mg) was incubated with bead-immobilized GST-SAP(1-128) or GST-EAT-2(1-132) fusion proteins. The beads were washed twice with PBS plus 0.5% NP-40 and then twice with PBS. Specifically bound peptides were eluted with 200 μl of 30% acetic acid, dried in a speed-vac apparatus, resuspended in water, and sequenced as described (Songyang et al., 1993).

Illustrations

Figure 1 was prepared with Setor (Evans, 1993), Figure 2 with GRASP (Nicholls et al., 1991), and Figure 3 with O (Jones and Kjeldgaard, 1997).

Acknowledgments

We thank C. Dahl for synthesis and purification of the SLAM peptides, and X. Huang and W. Meng for assistance with data collection and crystallographic analysis. M. J. E. is a recipient of a Burroughs-Wellcome Career Award in the Biomedical Sciences, and a member of the Harvard-Armenise Center for Structural Biology. This work

was supported in part by NIH grants to M. B. Y., C. T., and L. C. C. Diffraction data were recorded at the Cornell High Energy Synchrotron Source (CHESS), which is supported by grants from the NSF and NIH.

Received May 28, 1999; revised August 2, 1999.

References

Aversa, G., Carballido, J., Punnonen, J., Chang, C.C., Hauser, T., Cocks, B.G., and De Vries, J.E. (1997). SLAM and its role in T cell activation and Th cell responses. Immunol. Cell Biol. *75*, 202–205.

Brown, M.H., Boles, K., Anton van der Merwe, P., Kumar, V., Mathew, P.A., and Barclay, A.N. (1998). 2B4, the natural killer and T cell immunoglobulin superfamily surface protein, is a ligand for CD48. J. Exp. Med. *188*, 2083–2090.

Brunger, A. (1992). X-PLOR Version 3.0: a System for Crystallography and NMR (New Haven, CT: Yale University Press).

Cantley, L.C., and Songyang, Z. (1994). Specificity in recognition of phosphopeptides by src-homology 2 domains. J. Cell Sci. Suppl. *18*, 121–126.

Cocks, B.G., Chang, C.C., Carballido, J.M., Yssel, H., de Vries, J.E., and Aversa, G. (1995). A novel receptor involved in T-cell activation. Nature *376*, 260–263.

Coffey, A.J., Brooksbank, R.A., Brandau, O., Oohashi, T., Howell, G.R., Bye, J.M., Cahn, A.P., Durham, J., Heath, P., Wray, P., et al. (1998). Host response to EBV infection in X-linked lymphoproliferative disease results from mutations in an SH2-domain encoding gene. Nat. Genet. *20*, 129–135.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–776.

Eck, M.J., Shoelson, S.E., and Harrison, S.C. (1993). Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56^{lck}. Nature *362*, 87–91.

Evans, S.V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. J. Mol. Graph. *11*, 134–138.

Fisher, R.C., and Thorley-Lawson, D.A. (1991). Characterization of the Epstein-Barr virus-inducible gene encoding the human leukocyte adhesion and activation antigen BLAST-1 (CD48). Mol. Cell. Biol. *11*, 1614–1623.

Jones, T.A., and Kjeldgaard, M. (1997). Electron-density map interpretation. Methods Enzymol. *277*, 173–208.

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

Ladbury, J.E., Lemmon, M.A., Zhou, M., Green, J., Botfield, M.C., and Schlessinger, J. (1995). Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. Proc. Natl. Acad. Sci USA *92*, 3199–3203.

Lamzin, V.S., and Wilson, K.S. (1997). Automated refinement for protein crystallography. Methods Enzymol. *277*, 269–305.

Latchman, Y., McKay, P.F., and Reiser, H. (1998). Identification of the 2B4 molecule as a counter-receptor for CD48. J. Immunol. *161*, 5809–5812.

Lioubin, M.N., Algate, P.A., Tsai, S., Carlberg, K., Aebersold, A., and Rohrschneider, L.R. (1996). p150Ship, a signal transduction molecule with inositol polyphosphate-5-phosphatase activity. Genes Dev. *10*, 1084–1095.

Navaza, J. (1992). AMoRe: a new package for molecular replacement. In Molecular Replacement: Proceedings of the CCP4 Study Weekend, E.J. Dodson, S. Gover, and W. Wolf, eds. (Daresbury, UK: SERC), pp. 87–90.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Nichols, K.E., Harkin, D.P., Levitz, S., Krainer, M., Kolquist, K.A., Genovese, C., Bernard, A., Ferguson, M., Zuo, L., Snyder, E., et al. (1998). Inactivating mutations in an SH2 domain-encoding gene in X-linked lymphoproliferative syndrome. Proc. Natl. Acad. Sci. USA *95*, 13765–13770.

Nolte, R.T., Eck, M.J., Schlessinger, J., Shoelson, S.E., and Harrison, S.C. (1996). Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. Nat. Struct. Biol. *3*, 364–374.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, C.W. Carter and R. M. Sweet, eds. (San Diego, CA: Academic Press), pp. 307–326.

Pawson, T. (1995). Protein modules and signaling networks. Nature *373*, 573–580.

Purtilo, D.T., Cassel, C.K., Yang, J.P., and Harper, R. (1975). X-linked recessive progressive combined variable immunodeficiency (Duncan's disease). Lancet *1*, 935–940.

Sayos, J., Wu, C., Morra, M., Wang, N., Zhang, X., Allen, D., van Schaik, S., Notarangelo, L., Geha, R., Roncarolo, M.G., et al. (1998). The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM. Nature *395*, 462–469.

Seemayer, T.A., Gross, T.G., Egeler, R.M., Pirruccello, S.J., Davis, J.R., Kelly, C.M., Okano, M., Lanyi, A., and Sumegi, J. (1995). X-linked lymphoproliferative disease: twenty-five years after the discovery. Pediatr. Res. *38*, 471–478.

Sheldrick, G.M., and Schneider, T.R. (1997). Shelxl: high resolution refinement. Methods Enzymol. *277*, 319–343.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., et al. (1993). SH2 domains recognize specific phosphopeptide sequences. Cell *72*, 767–778.

Thompson, A.D., Braun, B.S., Arvand, A., Stewart, S.D., May, W.A., Chen, E., Korenberg, J., and Denny, C. (1996). EAT-2 is a novel SH2 domain containing protein that is up regulated by Ewing's sarcoma EWS/FLI1 fusion gene. Oncogene *13*, 2649–2658.

Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D., and Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell *72*, 779–790.

Zhang, Z., Lee, C.H., Mandiyan, V., Borg, J.P., Margolis, B., Schlessinger, J., and Kuriyan, J. (1997). Sequence-specific recognition of the internalization motif of the Alzheimer's amyloid precursor protein by the X11 PTB domain. EMBO J. *16*, 6141–6150.

Protein Data Bank Accession Numbers

Atomic coordinates have been deposited with the Protein Data Bank (Brookhaven National Laboratory) under accession numbers 1D1Z, 1D4T, and 1D4W.