The Specificity of the N-Terminal SH2 Domain of SHP-2 Is Modified by a Single Point Mutation†

Gregory Huyer‡,§,|| and Chidambaram Ramachandran*,‡

Department of Biochemistry and Molecular Biology, Merck Frosst Centre for Therapeutic Research, P.O. Box 1005, Pointe-Claire-Dorval, Québec H9R 4P8, Canada, and Department of Chemistry and Biochemistry, Concordia University, 1455 de Maisonneuve Boulevard West, Montréal, Québec H3G 1M8, Canada

Received September 15, 1997; Revised Manuscript Received December 8, 1997

ABSTRACT: SH2 domains are small protein domains of ∼100 amino acids that bind to phosphotyrosine (pY) in the context of a specific sequence surrounding the target pY. In general, the residues C-terminal to the pY of the binding target are considered most important for defining the binding specificity, and in particular the pY + 1 and pY + 3 residues (i.e., the first and third amino acids C-terminal to the pY). However, our previous studies with the SH2 domains of the protein tyrosine phosphatase SHP-2 [Huyer, G., Li, Z. M., Adam, M., Huckle, W. R., and Ramachandran, C. (1995) *Biochemistry 34*, 1040−1049] indicated important interactions with the pY − 2 residue as well. In the SH2 domains of SHP-2, the highly conserved αA2 Arg is replaced by Gly. A comparison of the published crystal structures of the Src SH2 domain and the N-terminal SH2 domain of SHP-2 complexed with high-affinity peptides suggested that the αA2 Gly of SHP-2 creates a gap which is filled by the side chain of the pY − 2 residue of the bound peptide. It was predicted that replacing this Gly with Arg would alter or eliminate the involvement of the pY − 2 residue in binding. The αA2 Gly → Arg mutant was constructed, and indeed, this mutant no longer required residues N-terminal to the target pY for high-affinity binding, making its specificity more like that of other SH2 domains. The αA2 Gly is clearly involved in directing the unusual requirement for the pY − 2 residue in the binding sequence of this SH2 domain, which has important implications for its in vivo targeting and specificity.

SH2 domains (for Src homology 2)[1] are discrete protein domains of ∼100 amino acids that bind to phosphotyrosine (pY) in the context of a specific sequence surrounding the pY (*1*−*4*). SH2 domains have been identified in as many as 100 proteins, including two mammalian protein tyrosine phosphatases (PTPs), SHP-1 and SHP-2. These PTPs are approximately 58% identical and have two tandem SH2 domains in the N-terminal portion and a single C-terminal catalytic domain. SHP-1 and SHP-2 play important roles in regulating many signaling pathways, including hematopoietic cell signal transduction and growth factor signaling (*5*, *6*).

The involvement of SH2 domain-containing proteins in a particular cell signaling pathway is directed by their SH2 domains; thus, knowledge of SH2 domain binding specificity is important for predicting and identifying their in vivo targets. For most SH2 domains, interactions with residues C-terminal to the target pY, and in particular the pY + 1 and pY + 3 residues (i.e., the first and third positions C-terminal to pY), are responsible for defining specificity. Using affinity selection of peptides from degenerate libraries, Songyang et al. (*7*) defined the binding sequence specificities of a number of SH2 domains which fell into two main classes: (i) pY-hydrophilic-hydrophilic-hydrophobic and (ii) pY-hydrophobic-X-hydrophobic, where X indicates any amino acid. Site-directed mutagenesis has demonstrated the involvement of particular residues of SH2 domains in defining binding specificity (*8*−*10*). In some cases, a single amino acid change in the SH2 domain led to dramatic changes in binding sequence specificity.

We have previously shown that the SH2 domains of SHP-2 have an unusual specificity in that residues N-terminal to the target pY, in particular the pY − 2 residue, influence binding (*11*). Recent studies with SHP-1 also revealed a similar involvement of N-terminal residues (*12*). The crystal structure of the N-terminal SH2 domain of SHP-2 bound to a peptide corresponding to its in vivo target (*13*) suggests a structural reason for this pY − 2 interaction. Almost all SH2 domains have a highly conserved Arg (or occasionally Lys) residue at the αA2 position; however, in the SH2 domains of SHP-2 (and SHP-1), this position is occupied by Gly. As a result, a gap is created in the SH2 domain that appears to be filled by the pY − 2 side chain of the target peptide and is a likely reason for the requirement for N-terminal residues. Thus, changing this Gly to Arg was predicted to eliminate this pY − 2 interaction. This change was introduced, along with five other αA2 substitutions which molecular modeling suggested may have a similar

† G.H. was supported by a MRC/PMAC Health Program Studentship and a Stanley G. French Graduate Fellowship from Concordia University.

* To whom correspondence should be addressed. Phone:(514) 428-8543, Fax:(514) 695-0693, e-mail:chidambaram_ramachandran@merck.com

‡ Merck Frosst Centre for Therapeutic Research.

§ Concordia University.

|| Present address: T Cell Laboratory, The Babraham Institute, Babraham, Cambridge CB2 4AT, U.K.

[1] Abbreviations: SH2, Src homology 2; pY, phosphotyrosine; PDGFR, platelet-derived growth factor receptor *β*; GST, glutathione *S*-transferase; GT, glutathione; PBS, phosphate-buffered saline; DTT, dithiothreitol; TBS, Tris-buffered saline; HBS, HEPES-buffered saline; PTP, protein tyrosine phosphatase.

S0006-2960(97)02291-5 CCC: $15.00 © 1998 American Chemical Society
Published on Web 02/12/1998

effect. The Gly → Arg substitution resulted in an SH2 domain with peptide binding affinity comparable to that of the wild type. As predicted, the requirement for residues N-terminal to the target pY was eliminated in the αA2 Arg mutant, resulting in a specificity more typical of other SH2 domains. The αA2 Gly is clearly involved in the unusual specificity of the N-terminal SH2 domain of SHP-2 and reveals an additional mode of SH2 domain binding that involves residues N-terminal to the target pY.

EXPERIMENTAL PROCEDURES

Materials. Phosphotyrosine peptides were obtained from California Peptide Research (Napa, CA), and PCR primers were synthesized by Research Genetics. Anti-phosphotyrosine antibody was from Upstate Biotechnology Inc., and GT-Sepharose was from Pharmacia. All other chemicals were of reagent grade from Sigma.

Cloning and Mutagenesis of the N-Terminal SH2 Domain of SHP-2. The N-terminal SH2 domain of SHP-2 was cloned by PCR from a full-length clone of SHP-2 in the vector pBluescript (*14*). The PCR primer used at the 5′ end was GGGGGATCCATGACATCGCGGAGATGGTTT and at the 3′ end was AGGAATTCTATGCACAGTTCAGAGGATATTT. The primers were designed to amplify residues 1−105 of SHP-2 with the addition of a *Bam*HI site at the 5′ end and a stop codon and *Eco*RI site at the 3′ end. The amplification conditions were 30 s at 95 °C, 30 s at 52 °C, and 30 s at 72 °C, for 20 cycles. The PCR products were purified using a PCR purification kit (QIAgen), digested with *Bam*HI and *Eco*RI, gel purified, and ligated into pBluescript SK+. Sequencing ensured that no errors had been introduced by PCR. The SH2 domain insert was subsequently cut out from pBluescript by digestion with *Bam*HI and *Eco*RI, gel purified, and ligated in-frame into the GST fusion vector pGEX-2T (Pharmacia) that had been digested with the same enzymes.

The αA2 substitutions were created using a PCR mutagenesis technique (*15*). To introduce the change at the αA2 position (the 13th codon), 5′ primers of 58−60 bases were used that included the 5′ start and *Bam*HI site plus 10−12 bases after the mutated codon; the same 3′ primer from the original subcloning of the SH2 domain (see above) was used. The SH2 domain cloned into pGEX-2T was used as the template DNA. The PCR conditions were 40 s at 95 °C, 30 s at 55 °C, and 30 s at 72 °C, for 20 cycles, and the PCR products were purified, digested, and ligated into pGEX-2T as described above. Sequencing verified that the correct mutations had been made.

Expression and Purification of GST−SH2 Constructs. *Escherichia coli* DH5α cells containing the GST−SH2 constructs were grown at 37 °C in 700 mL of LB with 100 μg/mL ampicillin to an A_{600} of 0.6−0.8. Cultures were chilled on ice to ∼27 °C, and expression was induced by the addition of 50 μM isopropyl thio-β-D-galactopyranoside (IPTG) with incubation overnight at 27 °C. All subsequent steps were performed at 4 °C. Cells were harvested by centrifugation at 4000*g* for 10 min, resuspended in 15 mL of lysis buffer [PBS containing 5 mM DTT, 1% Triton X-100, and Complete Protease Inhibitor Cocktail (Boehringer Mannheim)], and lysed by sonication. Lysates were cleared by centrifugation at 23000*g* for 15 min, and the fusion proteins were purified by incubating with 2 mL of GT-Sepharose beads for 1 h with end-over-end mixing. The beads were washed extensively with lysis buffer and then with TBS [50 mM Tris (pH 8.0) and 150 mM NaCl] containing 5 mM DTT. Beads were stored at −20 °C after the addition of an equal volume of 100% glycerol to a 50% suspension of the beads in TBS/DTT. The fusion proteins were >95% pure as estimated by SDS−PAGE.

To elute the GST−SH2 fusion proteins from the beads, 200 μL of beads was washed with TBS containing 5 mM DTT and 1 mM EDTA. Then 500 μL of the same buffer containing 10 mM GT was added and the mixture incubated on ice for 15 min with occasional mixing. The beads were pelleted by centrifugation, and the supernatant was removed. Typically, 2.5−3 mg of fusion protein was recovered from 200 μL of beads. The fusion proteins had the expected mass as determined by electrospray ionization mass spectrometry.

BIAcore Assay for SH2 Domain Binding. All experiments were carried out in HBS buffer [10 mM HEPES (pH 7.4), 0.15 M NaCl, 3 mM EDTA, and 0.005% Surfactant P20]. The PDGFR 1009 peptide [DTSSVL(pY)TAVQPN] was biotinylated as described (*11*) and immobilized on a streptavidin-coated sensor chip (Pharmacia) by injecting 10 μL of an 8.3 ng/mL peptide (flow rate of 5 μL/min). For all binding experiments, a flow rate of 10 μL/min was used, and samples were also injected over a flow cell with no immobilized peptide to provide a blank sensogram that was subtracted from the binding sensogram. Control experiments involving binding of anti-pY monoclonal antibody (injected at 10 μg/mL) reproducibly resulted in a 400−450 change in resonance units. The sensor surface was regenerated at the end of each experiment with 100 mM HCl which did not disrupt the high-affinity biotin−streptavidin interaction but did remove the bound proteins.

RESULTS

Modeling of SH2 Domain−Peptide Interactions. A unique feature of the SH2 domains of SHP-2 and SHP-1 is that Gly occupies the αA2 position instead of a normally highly conserved Arg (Figure 1). In the crystal structures of the Src (*16*, *17*) and Lck (*18*) SH2 domains complexed with peptides, an amino nitrogen of the αA2 Arg side chain makes an amino−aromatic interaction with the pY ring; in addition, the αA2 Arg forms several hydrogen bonds with the phosphate of the pY and with the pY − 1 residue. In the structure of the N-terminal SH2 domain of SHP-2 complexed with peptide (*13*), the replacement of this Arg by Gly is not compensated by any structural rearrangement of the SH2 domain relative to Src or Lck. However, while the phenyl ring of the pY residue is bound in a very similar conformation in SHP-2 and Src, the phosphate is rotated ∼180° relative to that in Src. This rotation results in additional interactions with the βB5 Arg residue, and overall, the same number of hydrogen bonds are formed as compared to those in the Src SH2 domain (*13*).

An additional consequence of the αA2 Gly of SHP-2 is that a gap is created in the SH2 domain that appears in the crystal structure to be filled by the side chain of the pY − 2 Val of the PDGFR peptide, helping to cover the phenyl ring of the pY (Figure 2). We rationalized that replacing this Gly with Arg would eliminate the involvement of the pY − 2 residue in the SH2 binding specificity. In addition to the Arg mutant, five other substitutions were made: Lys,

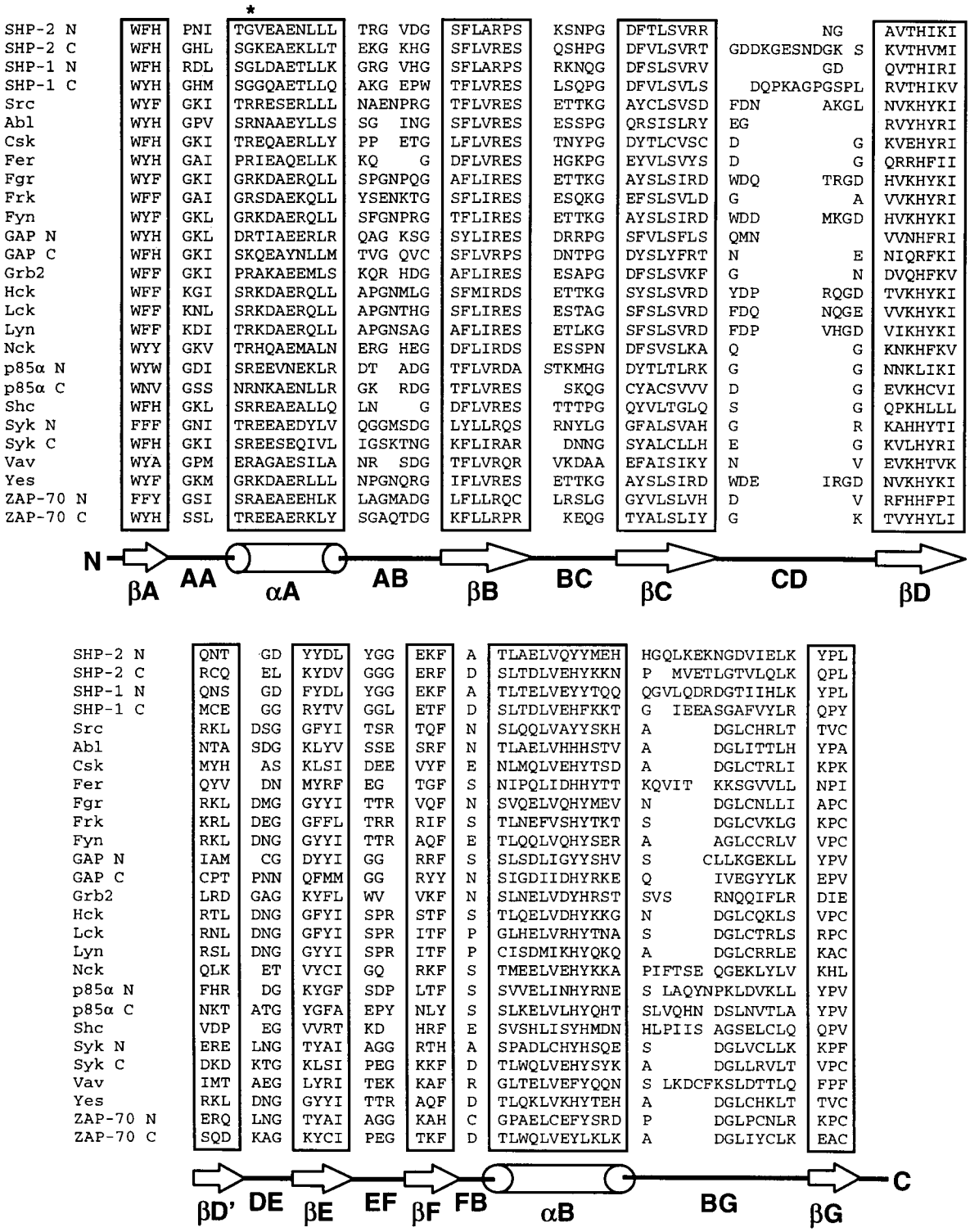

FIGURE 1: Alignment of human SH2 domain sequences. A subset of human SH2 domain-containing proteins was chosen from the Prosite SH2 domain profile, and their SH2 domains were aligned using GCG PileUp. The boxes show the boundaries of secondary structure elements as defined by Eck et al. (*18*) and are indicated schematically at the bottom along with the notation for these elements. The αA2 residue is marked by an asterisk. For proteins with two SH2 domains (e.g., SHP-2), the SH2 domains are indicated as N or C, for N-terminal or C-terminal.

another basic residue that is occasionally observed at this position in other SH2 domains; His, the only other basic residue, for completeness; Gln, as it appeared by modeling to be able to reach far enough to interact favorably with the pY; Met, to occupy the space created by Gly at αA2; and Ser, because modeling suggested that it would eliminate the need for a bound water molecule that hydrogen bonds with the peptide.

Binding Studies with Mutant SH2 Domains. The ability of the SH2 domains to bind to the PDGFR 1009 peptide was assessed by surface plasmon resonance, using a BIAcore system. The SH2 domains were eluted from the Sepharose beads by competition with glutathione, diluted to 500 nM in HBS buffer, and injected over immobilized PDGFR 1009 peptide on a sensor chip. As is apparent from the overlay sensograms (Figure 3), the wild-type SH2 domain caused

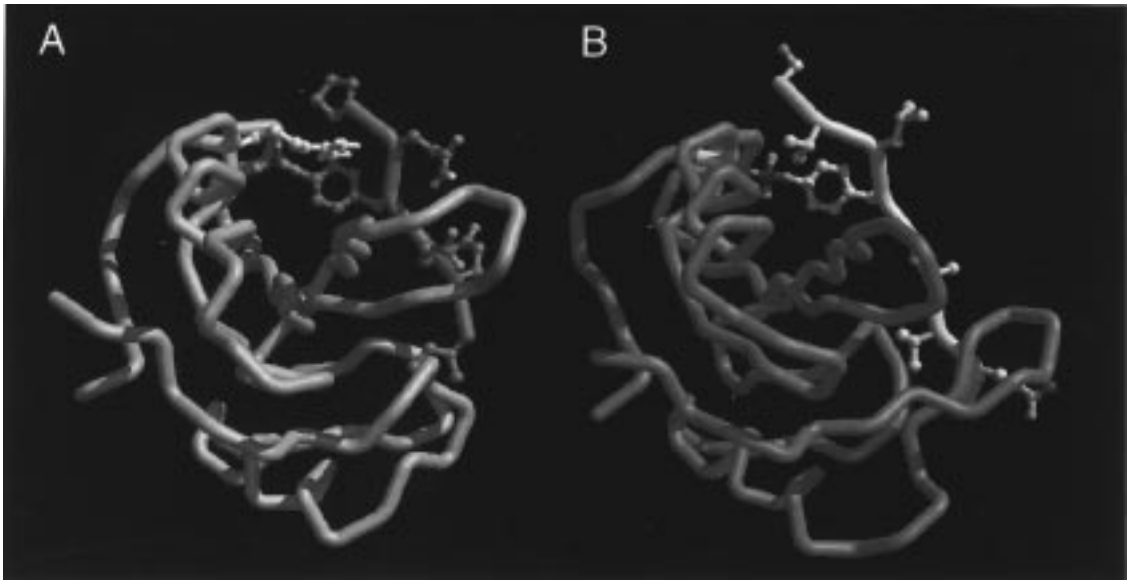

FIGURE 2: Comparison of the αA2 residue−pY peptide interactions of Src and SHP-2: (A) Src SH2 domain (orange) with bound PQ-(pY)EEI peptide (green) and (B) N-terminal SH2 domain of SHP-2 (cyan) with bound SVL(pY)TAVQPNE peptide (yellow). In both SH2 domain structures, the αA2 residue is shown in white. The orientations were matched by overlaying the structures, with an rms deviation of 0.72 Å between 40 residues in conserved structural elements. Note the gap created by the replacement of the αA2 Arg with Gly in SHP-2, which is filled by the Val side chain of the pY − 2 residue. The images were produced using the MidasPlus software system from the Computer Graphics Laboratory of the University of California at San Francisco (*37*). Coordinates were obtained from the Brookhaven Protein Data Bank (Src, ID code 1SPS; and SHP-2, ID code 1AYA).

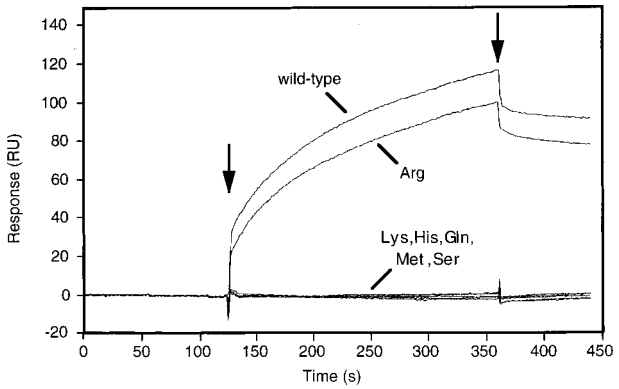

FIGURE 3: Overlay sensograms for the binding of wild-type and αA2 substituted SH2 domains to immobilized PDGFR 1009 peptide. The various SH2 domains (indicated by the αA2 residue or wild-type) were diluted to 500 nM in HBS buffer and injected over the sensor chip surface with immobilized PDGFR 1009 peptide, and the background (i.e., injection over the sensor surface without immobilized peptide) was subtracted: first arrow, injection of SH2 domain; and second arrow, wash with HBS buffer.

the largest change in resonance units upon binding, followed by the αA2 Arg mutant. The other five mutants displayed no detectable binding to the peptide under similar conditions. From the sensograms, the kinetic binding constants were calculated for the wild-type and αA2 Arg SH2 domains and are summarized in Table 1. The dissociation constants for the two SH2 domains are essentially identical, with an affinity of ∼40 nM.

Binding Sequence Specificities of Wild-Type and αA2 Arg SH2 Domains. We previously demonstrated (*11*) that the minimum sequence surrounding Tyr_{1009} of the PDGFR for high-affinity binding to the SH2 domains of SHP-2 was VL-(pY)TAV. To determine if the αA2 Arg mutation changed the binding specificity of the SH2 domain, a number of peptides were tested for their ability to compete in the BIAcore assay. In referring to the peptides, the 13mer

Table 1: Association and Dissociation Rate Constants for the Interaction of Wild-Type and αA2 Arg SH2 Domains with PDGFR 1009 Peptide

SH2 domain	$k_a \pm$ SE ($\times 10^{-4}$ M^{-1} s^{-1})[a]	$k_d \pm$ SE ($\times 10^4$ s^{-1})[a]	$K_d \pm$ SE (nM)[b]
wild-type	2.10 ± 0.42	8.87 ± 1.17	42.9 ± 5.7
αA2 Arg	2.19 ± 0.77	7.66 ± 0.29	38.3 ± 12.9

[a] Kinetic parameters for binding were calculated from the sensograms in Figure 3 and at three other SH2 domain concentrations, after blank subtraction using the BIAevaluation software package. The values represent the average and standard deviation of the four determinations. [b] The K_d was calculated for each of the four SH2 domain concentrations from the ratio of the k_d over the k_a, and the values were averaged.

PDGFR 1009 peptide immobilized on the sensor chip [DTSSVL(pY)TAVQPN] is designated as "1−13" and the numbering of all truncated peptides is based on this numbering. First, the concentration of free 1−13 peptide that reduced the binding by 50% (i.e., the IC_{50}) was determined, which was ∼1.5 μM for both the wild-type and αA2 Arg SH2 domains (Table 2). The various peptide competitors were tested at this concentration. If they competed as well as the 1−13 peptide, then binding would similarly be reduced by 50%; if they were poorer competitors, little or no reduction in binding would be observed.

The specificity for the pY − 2 residue was examined by testing a series of peptides based on the minimum high-affinity sequence [VL(pY)TAV, "5−10"] with substitutions at the pY − 2 position. The pY − 2 residue was an important determinant of binding to the wild-type SH2 domain. Peptides with Val, Ile, and Leu at this position competed strongly; Met, Phe, and Thr were reasonable competitors, and Ser and Ala were poor competitors (Figure 4A). For the αA2 Arg SH2 domain, all peptides competed equally well (Figure 4B), indicating that the side chain at this position was unimportant. The IC_{50}'s of a subset of the

Table 2: IC_{50} Values for Peptide Competitors with Wild-type and αA2 Arg SH2 Domains

peptide competitor	sequence[a]	IC_{50} (wild-type) ± S.E. (μM)[b]	IC_{50} (αA2 Arg) ± S.E. (μM)[b]
1-13	DTSSVL(pY)TAVQPN	1.3 ± 0.1	1.2 ± 0.1
5-10	VL(pY)TAV	0.56 ± 0.04	2.5 ± 0.3
5-10 Ala5	AL(pY)TAV	11.3 ± 0.7	1.5 ± 0.2
5-10 Thr5	TL(pY)TAV	1.7 ± 0.5	1.8 ± 0.2
5-10 Ser5	SL(pY)TAV	19.2 ± 1.5	2.2 ± 0.3
5-10 Leu5	LL(pY)TAV	1.0 ± 0.1	2.7 ± 0.1
5-10 Ala10	VL(pY)TAA	5.3 ± 0.6	24.8 ± 1.5
5-14	VL(pY)TAVQPNE	0.90 ± 0.03	2.5 ± 0.4
6-14	L(pY)TAVQPNE	24.3 ± 2.6	5.4 ± 0.7
7-14	Ac-(pY)TAVQPNE	35.1 ± 2.0	2.2 ± 0.2

[a] All peptides have carboxyamide at the C terminus and a free amino group at the N terminus, except the 7−14 peptide which is N-acetylated (indicated by Ac). [b] Values were determined from BIAcore binding assay, on the basis of the amount of peptide required to reduce the maximal SH2 domain binding to the immobilized 1−13 peptide in the absence of peptide competitors by 50%. The IC_{50} values were determined from a graphical fit as described previously (*11*).

peptides were determined (Table 2), further emphasizing the difference between the wild-type and αA2 Arg SH2 domains. For the wild-type SH2 domain, the IC_{50}'s varied over a 35-fold range, while for the αA2 Arg SH2 domain, less than a 2-fold range of IC_{50}'s was observed. The specificity for the pY + 3 residue was unaffected by the αA2 Arg substitution; replacing the pY + 3 residue with Ala increased the IC_{50} 10-fold relative to that with Val for both the wild-type and αA2 Arg SH2 domains (Table 2).

To explore further the peptide binding determinants, truncated peptides were tested for their ability to compete for binding. The results with the wild-type SH2 domain were as expected. Peptides that contained the minimum 5−10 sequence [VL(pY)TAV] competed as well as the 1−13 peptide, while truncations that extended into this sequence (i.e., the 1−9, 6−14, and 7−14 peptides) did not compete (Figure 5A). With the αA2 Arg SH2 domain, the same result was observed with C-terminal truncated peptides. The 1−10 peptide competed well, while the 1−9 peptide was a poor competitor (Figure 5B). However, substantial differences were observed with the N-terminal truncated peptides. Removal of N-terminal residues had little effect on the ability of the peptides to compete for binding (Figure 5B). The differences are clearly apparent by comparing the IC_{50}'s of the N-terminal truncated peptides (Table 2). Removal of the N-terminal residues increased the IC_{50} 40-fold for the wild-type SH2 domain but had no effect on the IC_{50} with the αA2 Arg SH2 domain. The higher IC_{50} of the 6−14 peptide is presumably due to the fact that it is not N-acetylated, since nonacetylated 7−14 peptide was a poorer competitor than N-acetylated 7−14 peptide (data not shown).

Comparison of in Vivo Binding Target Sequences. The involvement of the pY − 2 residue in directing high-affinity binding to the SH2 domains of SHP-1 has also been demonstrated (*12*), which by analogy to SHP-2 is likely a consequence of their having Gly residues at the αA2 position (Figure 1). A number of in vivo binding targets have been identified for SHP-2 and SHP-1 and the sites of association mapped to particular pY residues (Table 3). A comparison of the sequences surrounding the target pY residues further

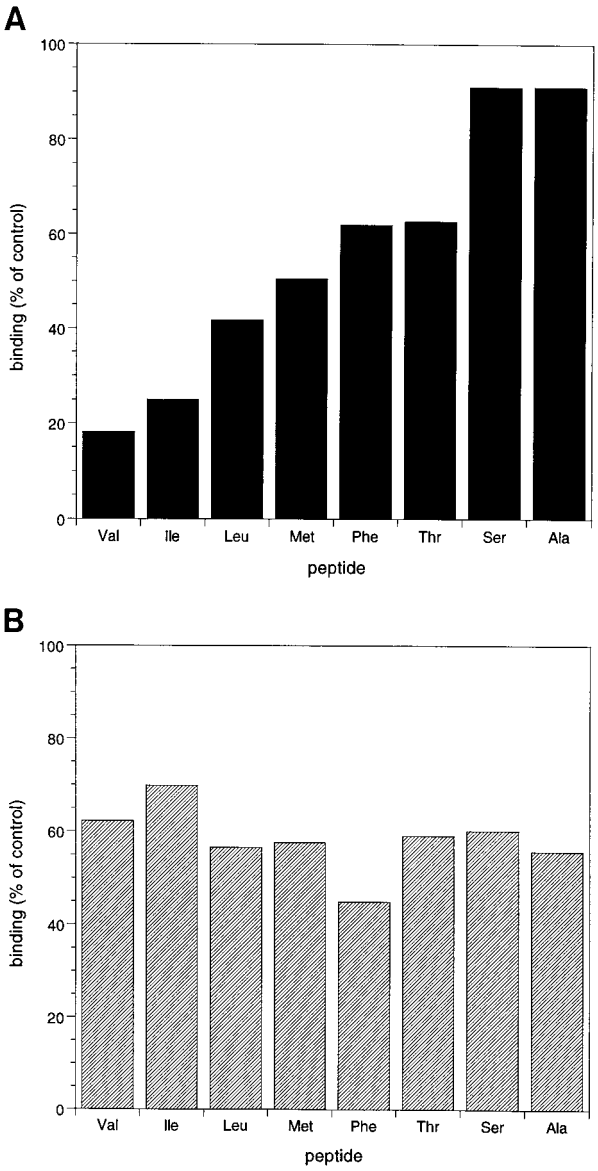

FIGURE 4: Effect of the pY − 2 residue in peptide binding competitions. Wild-type (A) or αA2 Arg mutant (B) SH2 domains (500 nM each) were incubated with the indicated peptides at 1.5 μM in HBS buffer. The increase in resonance units at steady state was determined, and the background (i.e., injection over the sensor surface without immobilized peptide) was subtracted. Binding is expressed as a percentage of the SH2 domain binding in the absence of competitor to the 1−13 peptide immobilized on the sensor chip. All peptides have the sequence XL(pY)TAV, with the residue at the pY − 2 position (X) indicated.

supports the importance of the pY − 2 residue, as all of these sequences have a hydrophobic residue at this position. The consensus binding sequence for these SH2 domains can thus be extended to hydrophobic-X-(pY)-hydrophobic-X-hydrophobic, with a preference for Val, Leu, and Ile at the pY − 2 and pY + 3 positions.

SHP-1 has also been shown to associate with c-Kit (*29*), but the actual pY residue to which the SH2 domains bind has not been determined. On the basis of this extended consensus binding sequence, a likely site for SHP-1 association with c-Kit is at Tyr_{730} (VSYVVP), assuming this Tyr represents an in vivo phosphorylation site. Similarly, SHP-2 associates with the platelet endothelial cell adhesion molecule-1 (PECAM-1) at an unknown site (*30*). Of the five

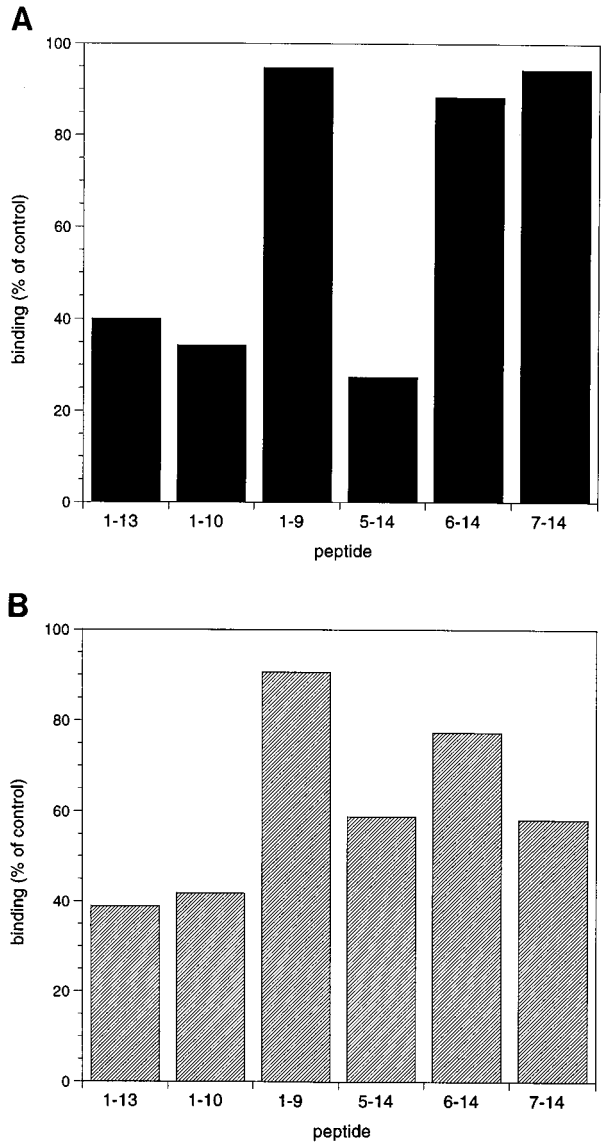

FIGURE 5: Effect of peptide truncations in peptide binding competitions. Wild-type (A) or αA2 Arg mutant (B) SH2 domains (500 nM each) were incubated with the indicated peptides at 1.5 μM in HBS buffer. Binding is expressed as described for Figure 4.

intracellular Tyr residues, only the sequence surrounding Tyr_{663} (VQYTEV) conforms to the high-affinity binding sequence (*31*) and represents the likely site of association if it is phosphorylated in vivo. Interestingly, there is also a Tyr at position 686 which may provide a binding target for the C-terminal SH2 domain of SHP-2.[2]

DISCUSSION

The PTPs SHP-2 and SHP-1 play essential roles in several tyrosine phosphorylation signaling pathways, as shown by the number of in vivo binding targets that have been identified (Table 3). Furthermore, mice carrying mutations in these PTPs display severe developmental defects (*32*, *33*). The SH2 domains of these PTPs direct their association with the appropriate targets to form specific, regulated signaling complexes, which may serve to bring SHP-1 and SHP-2 close to their substrates. Also, the SH2 domains modulate the catalytic activity of these PTPs (*6*). Thus, knowledge of the SH2 domain binding specificity is clearly important in determining the interaction partners of SHP-1 and SHP-2 and placing them in the appropriate pathways as well as in understanding their regulation.

Table 3: Sequences of Identified in Vivo Binding Sites of SHP-2 and SHP-1

Enzyme	*In vivo* binding target[a]	Binding sequence[b]	Reference
SHP-2	IRS-1	SLN(pY$_{1172}$)IDLDLV	*20*
SHP-2	PDGFR	SVL(pY$_{1009}$)TAVQPN	*21*
SHP-2	BIT	DIT(pY$_{436}$)ADLNLP	*22*
SHP-2	BIT	TLT(pY$_{477}$)ADLDMV	*22*
SHP-2	IR	HIP(pY$_{1322}$)THMNGG	*23*
SHP-2	EpoR[c]	SFE(pY$_{427}$)TILDPS	*24*
SHP-2	IL-3 ßc	SLE(pY$_{612}$)LCLPAG	*25*
SHP-1	CD22	TVS(pY$_{783}$)AILRFP	*26*
SHP-1	CD22	SIH(pY$_{843}$)SELVQF	*26*
SHP-1	CD22	DVD(pY$_{863}$)VTLKH	*26*
SHP-1	EpoR[c]	HLK(pY$_{455}$)LYLVVS	*27*
SHP-1	IL-3 ßc	SLE(pY$_{612}$)LCLPAG	*25*

[a] IRS-1, insulin receptor substrate 1; BIT, brain immunoglobulin-like molecule with tyrosine-based activation motifs; IR, insulin receptor; IL-3 βc, interleukin 3 receptor βc subunit; EpoR, erythropoietin receptor. [b] The number indicates the location of the pY residue in the target protein sequence. The pY − 2, pY + 1, and pY + 3 residues are highlighted in bold. [c] The amino acid sequence numbering for EpoR is taken from Yi et al. (*28*).

The SH2 domains of SHP-2 and SHP-1 have an unusual requirement for residues N-terminal to the target pY for high-affinity binding (*11*, *12*). We have shown here that, for the N-terminal SH2 domain of SHP-2, this requirement arises at least in part from the αA2 Gly residue in place of a highly conserved Arg. The wild-type and αA2 Arg mutant SH2 domains have distinct binding specificities as demonstrated by the peptide binding competitions (Figures 4 and 5 and Table 2). In particular, interactions between the SH2 domain and residues N-terminal to the target pY were eliminated in the αA2 Arg mutant. The αA2 Arg mutant showed no selectivity toward a variety of pY − 2 substituted peptides as compared to wild type (Figure 4B vs Figure 4A, Table 2). Furthermore, removal of all residues N-terminal to the pY had no appreciable effect on binding to the αA2 Arg SH2 domain (Table 2, compare 5−14 and 7−14 peptides), while affinity for the wild-type SH2 domain was decreased 40-fold.

Changing the αA2 Gly of the N-terminal SH2 domain of SHP-2 to Arg had little effect on binding to the PDGFR 1009 peptide; in fact, the binding kinetic constants were almost identical for the wild-type and αA2 Arg SH2 domains (Table 1). The αA2 position has previously been shown to be relatively tolerant to changes, as mutations at this position in the N-terminal SH2 domain of Gap had only minor effects on binding (*19*). However, in this study, only the αA2 Arg-substituted SH2 domain was competent for high-affinity binding; substitution with five other residues at this position resulted in weaker binding that was not detectable in the BIAcore assay (Figure 3). Because of the αA2 Gly, there is a major change in the pY binding as the phosphate of the pY residue is rotated ∼180° relative to the Src and Lck complexes. It is likely that, in the Gly → Arg mutant, the pY phosphate binds in an orientation similar to that for the

[2] While this manuscript was being reviewed, it was demonstrated that Tyr_{663} and Tyr_{686} of PECAM-1 do in fact bind to the N- and C-terminal SH2 domains, respectively, of SHP-2 in vivo (*38*).

Src and Lck SH2 domains, although a structural determination is required to confirm this. Thus, while the interaction with the pY − 2 residue is eliminated, compensatory amino−aromatic interactions between the αA2 Arg and the phenyl ring of the pY may be created like that observed in the Src and Lck SH2 domains, accounting for the similar binding affinities of the wild-type and αA2 Arg SH2 domains. Such interactions would not be formed in the other SH2 domain mutants made here, consistent with these SH2 domains not binding efficiently to the PDGFR 1009 peptide. These data suggest that the αA2 Arg is directly involved in maintaining a high-affinity interaction and is responsible for the change in specificity.

The role of the αA2 Gly in creating the pY − 2 interaction reinforces the involvement of N-terminal residues in the binding of the SH2 domains of SHP-2 (*11*). Similarly, the SH2 domains of SHP-1, the other mammalian SH2 domain-containing PTP, have Gly at the αA2 position (Figure 1), and an interaction with the pY − 2 residue has been demonstrated (*12*). An αA2 Gly is also found in the SH2 domain of Ctk (*34*) [also known as Matk (*35*) and Hyl (*36*)], a member of the Csk family of protein tyrosine kinases. While there is no information on the binding specificity of the Ctk SH2 domain, the presence of the αA2 Gly suggests by analogy to SHP-2 and SHP-1 that residues N-terminal to the target pY are involved. Thus, these αA2 Gly-containing SH2 domains may define another class with specificity for residues both N- and C-terminal to the target pY. From the results presented in this study, the consensus binding sequence for high-affinity binding appears to be hydrophobic-X-(pY)-hydrophobic-X-hydrophobic, with a preference for aliphatic amino acids at the pY − 2 and pY + 3 positions. Indeed, almost every in vivo binding target for the SH2 domains of SHP-2 and SHP-1 has Val, Leu, or Ile at the pY − 2 position (Table 3), further supporting these observations. This binding mode has important implications for identifying in vivo binding targets for these critical PTPs in cell signaling and reinforces the fact that SH2 binding specificity is defined by more than the three residues C-terminal to the target pY.

ACKNOWLEDGMENT

We thank M. Abramovitz for assistance with the mutagenesis strategies and techniques, C. Bayly for molecular modeling work and many stimulating and helpful discussions, L. Stolz for assistance with the BIAcore analysis, and A. Filosa for mass spectrometry. We also thank M. Gresser for his support and A. English for critical reading of the manuscript.

REFERENCES

1. Cohen, G. B., Ren, R., and Baltimore, D. (1995) *Cell 80,* 237−248.
2. Pawson, T. (1995) *Nature 373,* 573−580.
3. Schaffhausen, B. (1995) *Biochim. Biophys. Acta 1242,* 61−75.
4. Songyang, Z., and Cantley, L. C. (1995) *Trends Biochem. Sci. 20,* 470−475.
5. Neel, B. G., and Tonks, N. K. (1997) *Curr. Opin. Cell Biol. 9,* 193−204.
6. Frearson, J. A., and Alexander, D. R. (1997) *BioEssays 19,* 417−427.
7. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) *Cell 72,* 767−778.
8. von Bonin, A., Wienands, J., Manning, U., Zuber, J. F., and Baumann, G. (1994) *J. Biol. Chem. 269,* 33035−33041.
9. Songyang, Z., Gish, G. D., Mbamalu, G., Pawson, T., and Cantley, L. C. (1995) *J. Biol. Chem. 270,* 26029−26032.
10. Marengere, L. E. M., Songyang, Z., Gish, G. D., Schaller, M. D., Parsons, J. T., Stern, M. J., Cantley, L. C., and Pawson, T. (1994) *Nature 369,* 502−505.
11. Huyer, G., Li, Z. M., Adam, M., Huckle, W. R., and Ramachandran, C. (1995) *Biochemistry 34,* 1040−1049.
12. Burshtyn, D. N., Yang, W., Yi, T., and Long, E. O. (1997) *J. Biol. Chem. 272,* 13066−13072.
13. Lee, C.-H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E., and Kuriyan, J. (1994) *Structure 2,* 423−438.
14. Bastien, L., Ramachandran, C., Liu, S., and Adam, M. (1993) *Biochem. Biophys. Res. Commun. 196,* 124−133.
15. Horton, R. M., Cai, Z., Ho, S. N., and Pease, L. R. (1990) *BioTechniques 8,* 528−535.
16. Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. D., Silverman, L., and Kuriyan, J. (1992) *Nature 358,* 646−653.
17. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., and Kuriyan, J. (1993) *Cell 72,* 779−790.
18. Eck, M. J., Shoelson, S. E., and Harrison, S. C. (1993) *Nature 362,* 87−91.
19. Marengere, L. E. M., and Pawson, T. (1992) *J. Biol. Chem. 267,* 22779−22786.
20. Sun, X. J., Crimmins, D. L., Myers, M. G., Jr., Miralpeix, M., and White, M. F. (1993) *Mol. Cell. Biol. 13,* 7418−7428.
21. Lechleider, R. J., Sugimoto, S., Bennett, A. M., Kashishian, A. S., Cooper, J. A., Shoelson, S. E., Walsh, C. T., and Neel, B. G. (1993) *J. Biol. Chem. 268,* 21478−21481.
22. Ohnishi, H., Kubota, M., Ohtake, A., Sato, K., and Sano, S. (1996) *J. Biol. Chem. 271,* 25569−25574.
23. Staubs, P. A., Reichart, D. R., Saltiel, A. R., Milarski, K. L., Maegawa, H., Berhanu, P., Olefsky, J. M., and Seely, B. L. (1994) *J. Biol. Chem. 269,* 27186−27192.
24. Tauchi, T., Damen, J. E., Toyama, K., Feng, G.-S., Broxmeyer, H. E., and Krystal, G. (1996) *Blood 87,* 4495−4501.
25. Bone, H., Dechert, U., Jirik, F., Schrader, J. W., and Welham, M. J. (1997) *J. Biol. Chem. 272,* 14470−14476.
26. Doody, G. M., Justement, L. B., Delibrias, C. C., Matthews, R. J., Lin, J., Thomas, M. L., and Fearon, D. T. (1995) *Science 269,* 242−244.
27. Klingmuller, U., Lorenz, U., Cantley, L. C., Neel, B. G., and Lodish, H. F. (1995) *Cell 80,* 729−738.
28. Yi, T., Zhang, J., Miura, O., and Ihle, J. N. (1995) *Blood 85,* 87−95.
29. Yi, T., and Ihle, J. N. (1993) *Mol. Cell. Biol. 13,* 3350−3358.
30. Masuda, M., Osawa, M., Shigematsu, H., Harada, N., and Fujiwara, K. (1997) *FEBS Lett. 408,* 331−336.
31. Newman, P. J., Berndt, M. C., Gorski, J., White, G. C., Lyman, S., Paddock, C., and Muller, W. A. (1990) *Science 247,* 1219−1222.
32. Saxton, T. M., Henkemeyer, M., Gasca, S., Shen, R., Rossi, D. J., Shalaby, F., Feng, G. S., and Pawson, T. (1997) *EMBO J. 16,* 2352−2364.
33. Tsui, H. W., Siminovitch, K. A., de Souza, L., and Tsui, F. W. L. (1993) *Nat. Genet. 4,* 124−129.
34. Klages, S., Adam, D., Class, K., Fargnoli, J., Bolen, J. B., and Penhallow, R. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91,* 2597−2601.
35. Bennet, B. D., Cowley, S., Jiang, S., London, R., Deng, B., Grabarek, J., Groopman, J. E., Goeddel, D. V., and Avarham, H. (1994) *J. Biol. Chem. 269,* 1068−1074.
36. Sakano, S., Iwama, A., Inazawa, J., Ariyama, T., Ohno, M., and Suda, T. (1994) *Oncogene 9,* 1155−1161.
37. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6,* 13−27.
38. Jackson, D. E., Kupcho, K. R., and Newman, P. J. (1997) *J. Biol. Chem. 272,* 24868−24875.

BI9722913