The Three-Dimensional Solution Structure of the SH2 Domain from p55^{blk} Kinase

William J. Metzler,*,‡ Barbara Leiting,§,|| Kelly Pryor,§,|| Luciano Mueller,‡ and Bennett T. Farmer II‡

Departments of Macromolecular NMR and Macromolecular Biochemistry, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08543-4000

Received January 22, 1996; Revised Manuscript Received March 14, 1996[⊗]

ABSTRACT: Signal transduction in B cells is mediated, in part, by the interaction of the cytoplasmic components of the antigen receptor complex and various members of the *src* family tyrosine kinases. Key to this process appears to be the interaction of the tyrosine kinase SH2 domains with the tyrosine-phosphorylated cytoplasmic domain of Ig-α, a disulfide-bonded heterodimeric (with Ig-β or Ig-γ) transmembrane protein that noncovalently associates with the antigen receptor immunoglobin chains. In addition to binding to the phosphorylated cytoplasmic domains of Ig-α and Ig-β, *blk* and *fyn*(T), two members of the *src* family kinases, have been shown to bind overlapping but distinct sets of phosphoproteins [Malek & Desiderio (1993) *J. Biol. Chem. 268*, 22557−22565]. A comparison of their three-dimensional structures may elucidate the apparently subtle differences required for phosphoprotein discrimination. To begin characterizing the *blk*/*fyn*/phosphosphoprotein interactions, we have determined the three-dimensional solution structure of the SH2 domain of *blk* kinase by nuclear magnetic resonance (NMR) spectroscopy. ^{1}H, ^{13}C, and ^{15}N resonances of the SH2 domain of *blk* kinase were assigned by analysis of multidimensional, double- and triple-resonance NMR experiments. Twenty structures of the *blk* SH2 domain were refined with the program X-PLOR using a total of 2080 experimentally derived conformational restraints. The structures converged to a root-mean-squared (rms) distance deviation of 0.51 and 0.95 Å for the backbone atoms and for the non-hydrogen atoms, respectively. The *blk* SH2 domain adopts the prototypical SH2 fold. Structurally, *blk* SH2 is most similar to the crystal structure of the v-*src* SH2 domain [Waksman et al. (1993) *Nature 358*, 646−653] and superimposes on the crystal structure with an rmsd of 1.52 Å for the backbone atoms. The largest deviations occur in the four loops interconnecting β-strands A−E, which are the least well-defined regions in the NMR structure. Exclusion of these loops lowers this rmsd to 0.82 Å. The conformation of the BC loop in the *blk* SH2 domain is similar to the open conformation in the apo *lck* SH2 domain, suggesting that, like the *lck* SH2 domain, the *blk* SH2 domain may have a gated phosphopeptide binding site. Finally, it is proposed that the amino acid substitution of Lys 88 (*blk*) for Glu [*fyn*(T)] is important for the observed differences in specificity between *blk* and *fyn*(T) SH2 domains.

Src homology-2 (SH2)[1] domains are globular protein modules of ∼100 amino acids that are found in a wide variety of intracellular signal-transduction proteins (Koch et al., 1991; Pawson, 1995; Pawson & Gish, 1992). SH2 domains mediate specific protein−protein interactions (Koch et al., 1991) through binding to specific phosphotyrosine (pY)-containing peptide sequences in their target ligands (Moran et al., 1990). The structures of various SH2 domains, both free in solution and complexed with pY-peptides, have been determined both by X-ray crystallography (Eck et al., 1993; Mikol et al., 1995; Waksman et al., 1992, 1993) and by solution-state NMR spectroscopy (Booker et al., 1992; Narula et al., 1995; Overduin et al., 1992a,b; Pascal et al., 1994; Xu et al., 1995). These structural studies have shown that all SH2 domains adopt a similar fold: a central β-sheet flanked on opposite sides by α-helices. The pY-peptide binding site on *src*-family SH2 domains is composed essentially of two binding pockets, one for the pY moiety and a second for the third residue C-terminal to pY (pY+3). The sole requirement for low-affinity binding appears to be the presence of the phosphotyrosine moiety on the peptide; in fact, isolated SH2 domains can directly bind to phosphotyrosine (Lemmon & Ladbury, 1994) and to phosphotyrosine coupled to agarose beads (Mayer et al., 1991). Discrimination between low- and high-affinity peptide binding is dependent upon occupancy of the second binding pocket by the pY+3 residue (Eck et al., 1993; Waksman et al., 1993). Studies with short peptides have indicated that a minimum of five residues can effectively block SH2-mediated binding to activated receptors (Fantl et al., 1992; Gilmer et al., 1994) and that the binding affinity of SH2 domains for the phosphopeptides is not increased with peptides having more than 12 residues (Piccione et al., 1993). Phosphopeptide binding can cause conformational changes in the SH2 domain, particularly in the BC and EF loops (Mikol et al., 1995; Waksman et al., 1993).

p55^{blk} kinase, or *blk*, is one of several members in the *src* family of kinases, including *lyn*, the *fyn*(T) isoform of *fyn*, and *lck*, that are expressed in B cells and are activated by engagement of the B cell antigen receptor (Burkhardt et al.,

* To whom correspondence should be addressed.
‡ Department of Macromolecular NMR.
§ Department of Macromolecular Biochemistry.
|| Current address: Merck & Company, RY50105, P.O. Box 200, Rahway, NJ 07065.

⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1996.

[1] Abbreviations: NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; 3D, three dimensional; 4D, four dimensional; rms, root mean squared; SH2, src homology-2; pY, phosphotyrosine.

S0006-2960(96)00157-2 CCC: $12.00

1991). Both *blk* and *fyn* have been observed to bind to the cytoplasmic domains of Ig-α and Ig-*β*, disulfide-linked transmembrane proteins that noncovalently associate with the B cell antigen receptor (Burkhardt et al., 1991; Campbell & Sefton, 1992; Law et al., 1993; Yamanashi et al., 1991). Despite extensive sequence homology, differences in the primary structure of these *src* kinases have suggested that these proteins interact with distinct sets of proteins and consequently serve distinct signaling functions. This suggestion has been confirmed, in part, by binding studies in which specific SH2 binding patterns were exchanged between various SH2 domains (Malek & Desiderio, 1993; Marengere & Pawson, 1994; von Bonin et al., 1994). For example, the binding properties of the *blk* SH2 domain were transferred to the *fyn*(T) SH2 domain in chimeras that modified as few as 14 residues of *fyn*(T) (Malek & Desiderio, 1993). Similarly, the phosphoprotein binding patterns of the *src* SH2 domain were transferred to the Sem-5 SH2 domain by altering only a single residue in the *src* SH2 domain (Marengere & Pawson, 1994). These binding studies would indicate that the determinants for SH2-specific binding lie in the structural details dictated by the intrinsic amino acids of the particular SH2 domain.

Given the cell-type restricted expression of some SH2 domains, the critical role of tyrosine phosphorylation in intracellular signaling pathways, the specificity of the binding interactions, and the highly defined binding site, SH2 domains are an attractive medium for intervention in immunological responsiveness and a number of human diseases, including a wide variety of carcinomas. Agents that target SH2 domains, however, must be able not only to bind with high affinity but also to bind with selectivity to the SH2 domain of interest. While the determinants that segregate SH2/phosphopeptide interactions have been discerned, those that provide specificity *within* subclasses of SH2 domains are less well understood. We have initiated a project to compare the structures of the SH2 domains of *blk* and *fyn*(T). A comparison of their three-dimensional structures may elucidate the apparently subtle differences required for phosphoprotein discrimination. Here, we report the three-dimensional structure of the SH2 domain of *blk*. The *blk* SH2 domain structure is found to be most similar to that determined for the v-*src* SH2 domain by X-ray crystallography (Waksman et al., 1993), with primary differences residing in loop regions. The BC loop of the *blk* SH2 domain appears to lie in a conformation analogous to the "open" conformation of the *lck* SH2 domain, leading to the suggestion that *blk* may also have a gated phosphopeptide binding site. Finally, it is proposed that the amino acid substitution of Lys 88 (*blk*) for Glu [*fyn*(T)] is important for the observed differences in specificity between *blk* and *fyn*(T) SH2 domains.

EXPERIMENTAL PROCEDURES

Sample Preparation. The recombinant *blk* SH2 domain was prepared by subcloning the cDNA-encoding amino acid residues 107−218 from $p55^{blk}$ kinase into the pGEX-2T expression vector (gift from J. B. Bolen, BMSPRI). *Escherichia coli* BL21 cells were transformed with pGEX-2T and grown at 37 °C in minimal medium to an OD_{600} between 0.8 and 1.0. The cells were pelleted (3200*g*, room temperature), resuspended in fresh minimal medium, and allowed to recover for 10 min. Expression was then induced at 37 °C with 0.5 mM IPTG. The cells were harvested after a 3−4 h expression phase, at which time the final OD_{600} was ∼1.3. The cell medium was chilled to 4 °C, and the cells were pelleted (3500*g*, 4 °C), resuspended in TNE buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM EDTA), pelleted again (4000*g*, 4 °C), and frozen in liquid nitrogen. Cell pellets from 1−1.5 Liter cell cultures were resuspended in ice cold lysis buffer [20 mM sodium phosphate, pH 6.0, 10 mM DTT, 2 mM EDTA, 1 mM Pefabloc (Boehringer), 1 mM benzamidine (Sigma)] to a final volume of 30−35 mL. This suspension was transferred into a chilled French press and lysed once between 1200 and 1500 psi. The volume was increased to 50 mL with lysis buffer, and the supernatant was separated from the cell debris by ultracentrifugation at 35 000 rpm for 35 min in a Beckman TI45 rotor at 4 °C. The lysis supernatant was diluted 3-fold with column buffer (20 mM sodium phosphate, pH 6.0, 1 mM DTT, 2 mM EDTA) and loaded onto a preequilibrated glutathione−agarose (Sigma G-4510) column (25−30 mL). The column was washed with at least 10 column volumes of column buffer and then eluted with 0.1 M NaOH. The NaOH eluate was neutralized by adding Tris-HCl, pH 7.4, to a final concentration of 50 mM and HCl to obtain a pH value between 8.0 and 9.0. Under these conditions, the fusion protein begins to precipitate, so thrombin digestion was initiated as quickly as possible. NaCl was added to a concentration of 150 mM, and bovine thrombin (Amour Pharmaceutical Co.) was added at 7 units/mg of fusion protein. After being allowed to stand 30−40 min at room temperature, the thrombin digestion was quenched with 1 mM benzamidine. The digestion mixture was diluted 8-fold with lysis buffer that did not contain Pefabloc or benzamidine and loaded again onto a glutathione-agarose column. The flow-through was collected and adjusted to pH 8.0 with NaOH. It was then loaded onto an SP-Sepharose column (Pharmacia, fast flow 17-0729-01) previously equilibrated with 20 mM phosphate buffer, pH 8.0, and 1 mM DTT. A salt gradient of 0−500 mM NaCl was then applied to the column, with the *blk* SH2 domain eluting at ∼200 mM NaCl. The fractions containing *blk* SH2 domain were pooled, dialyzed against water for 1 day (MW cutoff ∼3500), and concentrated by using centriprep and Centricon filters. Typical yields for a 1 L growth were ∼30 mg of purified *blk* SH2 domain. The protein identity was confirmed by mass spectroscopy; its ability to bind phosphotyrosine-containing peptides was demonstrated with isothermal calorimetry (data not shown).

NMR Spectroscopy. NMR experiments were collected at 20 °C on a single 1.4 mM sample of ^{13}C,^{15}N-labeled *blk* SH2 domain in 10 mM phosphate buffer, pH 6.3. All pulse sequences used in the NMR resonance assignment and structure determination of *blk* SH2 domain are listed in Table 1. In general, the pulse sequences were modified to include gradients for enhanced solvent suppression, ^{15}N-coherence selection with enhanced sensitivity (Kay et al., 1992), and/or reduced phase cycling. A detailed description of all NMR experiments and data processing are provided as Supporting Information.

Sequence-Specific Resonance Assignments. The residues of the *blk* SH2 domain have been numbered to facilitate alignment with previously published sequences of SH2 domains (see Figure 4). In this regard, the conserved tryptophan is residue 5, and the sequence of the protein

Table 1: Pulse Sequences for *blk* SH2 Domain Resonance Assignment and Structure Calculation[a]

experiment	acquisition time (h)	references
resonance assignments		
3D CBCA(CO)NH	24	Grzesiek & Bax, 1992
3D HNCACB	24	Yamazaki et al., 1994
3D HBHA(CO)NH	34	Grzesiek & Bax, 1993a
3D HBHANH	34	(unpublished results)
3D HNCO	12	Kay et al., 1994)
2D ^{1}H−^{13}C CT-HSQC^{Met}	3	Farmer et al., 1995
3D HCCH-TOCSY	36	Kay et al., 1993
2D CBHD	6	Yamazaki et al., 1993
2D CBHE	12	Yamazaki et al., 1993
distance restraints		
4D CC-NOESY	142	Vuister & Bax, 1993
4D CN/NN-NOESY	133	Farmer & Mueller, 1994
torsion restraints		
3D HNHB	54	Archer et al., 1991
3D HN(CO)HB	64	Grzesiek & Bax, 1992
2D HSQC-J	25	Billeter et al., 1992
{^{15}N} SE-difference CT-HSQC	7	Bax et al., 1992
{^{13}CO} SE-difference CT-HSQC	7	Vuister et al., 1993
hydrogen bond restraints		
2D HSQC-exchange	4 26 days (total)	Grzesiek & Bax, 1993b

[a] A detailed description of all NMR experiments and data processing are provided as Supporting Information.

therefore extends from residue −9 through 105.

Sequential resonance assignments were made through analysis of the experiments listed in Table 1. The assignment strategy was similar to that described previously (Constantine et al., 1994; Friedrichs et al., 1994; Metzler et al., 1993) and is only briefly described here. Cross-peaks detected in the 3D HNCACB, CBCA(CO)NH, HBHA(CO)NH, and HBHANH experiments were organized into spin systems consisting of the ^{1}HN[i], ^{15}N[i], ^{13}Cα[i],^{13}Cβ[i], ^{1}Hα[i], ^{1}Hβ-[i],^{13}Cα[$i-1$],^{13}Cβ[i-1], ^{1}Hα[$i-1$], and ^{1}Hβ[$i-1$] resonances. Spin systems were linked by matching the ^{13}Cα[i],^{13}Cβ[i], ^{1}Hα[i], and ^{1}Hβ[i] chemical shifts of an HN/N pair with the ^{13}Cα[$i-1$], ^{13}Cβ[$i-1$], ^{1}Hα[$i-1$] and ^{1}Hβ[$i-1$] chemical shifts of another HN/N pair. In this way, numerous stretches of sequentially aligned spin systems were generated. To obtain sequence specific assignments, "marker" spin systems were identified. These markers included threonine and serine residues, which were easily identified by their significantly downfield-shifted Cβ resonances; alanine residues, which were identified by the characteristic upfield chemical shift of their Cβ resonances; and glycine residues, which were identified by the characteristic chemical shift of their Cα resonances and the absence of a Cβ resonance. The identities of the remaining spin systems were restricted to a subset of possible residue types based on their Cα/Cβ chemical shifts and the probabilities that these chemical shifts indicate a particular amino acid type. At this point, it became straightforward to assign specific residues to each of the stretches of connected spin systems, and backbone resonance assignments were obtained for all residues. Sequential connectivity were typically redundant with the only breaks occurring at proline residues. These gaps were bridged with NOESY data; in addition, NOESY data were used to confirm all sequential linkages. Sequential carbonyl ^{13}C resonances were assigned with the 3D HNCO spectra.

Aliphatic side chain ^{1}H and ^{13}C resonances were assigned primarily by analyzing the 3D HCCH-TOCSY spectra. Methionine ϵ-methyl ^{1}H and ^{13}C resonances were identified with the 2D ^{1}H−^{13}C CT -HSQC^{MET} spectrum and subsequently assigned with the 4D ^{13}C/^{13}C-separated NOESY. Sidechain ^{1}H and ^{15}N for glutamines and asparagines were assigned by identifying NOEs between the side chain amide protons and Hγ and Hβ protons in the 4D ^{13}C/^{15}N-separated NOESY. Cross-peaks between these resonances were also observed in the HNCACB and CBCA(CO)NH spectra.

Aromatic side chain ^{1}H and ^{13}C resonances of Tyr and Phe were assigned by analyzing the CBHD, CBHE, and aromatic HCCH-TOCSY spectra. Aromatic side chain ^{1}H and ^{13}C resonances of Trp were assigned via the 4D ^{13}C/^{13}C-separated and ^{13}C/^{15}N-separated NOESY spectra. Assignments were confirmed by detection of NOEs from the aromatic side chain to the intraresidue ^{1}Hβs and ^{1}Hαs. Tryptophan indole ^{1}H$_N$ and ^{15}N resonances were assigned by detection of NOEs to the other ring protons in the 4D ^{13}C/^{15}N-edited and ^{13}C/^{13}C-edited NOESY spectra.

The side chain assignments confirmed the previously described sequential backbone assignments. This has provided support for the use of chemical shifts to group spin systems into subclasses of amino acid types.

The ^{1}H, ^{15}N, and ^{13}C assignments have been submitted as Supporting Information.

Derivation of Structural Restraints. Three types of restraints for the structure calculations were initially derived from the NMR data: interproton distance restraints, torsion angle restraints, and hydrogen bond restraints. After the initial rounds of structure refinement, additional restraints based on the experimental aliphatic ^{1}H chemical shifts (Kuszewski et al., 1995a), ^{13}Cα and ^{13}Cβ chemical shifts (Kuszewski et al., 1995b), and $^3J_{HNHA}$ (Garrett et al., 1994) were included. A list of all experimental restraints is available from the authors.

A total of 1454 physically significant distance restraints were derived from the cross-peak intensities of the 4D ^{13}C/^{15}N-separated and 4D ^{13}C/^{13}C-separated NOESY spectra acquired with a mixing time of 80 ms. These restraints included 583 intraresidue restraints, 306 sequential restraints, 173 medium-range restraints ($2 \leq |i - j| \leq 4$), and 392 long-range restraints ($|i - j| > 4$). Each interresidue NOE was converted into an interproton distance by normalizing its integrated cross-peak volume against a calibrated volume. For NOEs arising from aromatic protons, intraresidue Trp ring NOEs were used as calibrations. For the remaining protons, backbone−backbone NOEs within the identified β-sheet regions and α-helices were used with reference to the typical distance ranges for these interactions (Wüthrich, 1986). For example, the average volume of Hα−Hα NOEs between β-strands was used to establish a calibration standard for very strong NOEs in the aliphatic 4D ^{13}C/^{13}C-separated NOESY spectrum. The distance for this interaction is 2.3 Å in regular antiparallel β-sheets (Wüthrich, 1986). Restraint ranges were generated by setting the lower bound equal to the sum of the van der Waals radii (1.8 Å) and by setting the upper bound equal to 1.2 times the calculated target distance.

A total of 130 torsion angle restraints for either the backbone torsion angle Φ or side chain torsion angle χ_1 were derived from measured J values. Restraints on Φ were derived from a gradient-enhanced (GE) version of the HSQC-J-modulated [^{15}N,^{1}H]-COSY experiment recorded with variable delays of 10, 20, 40, 60, 80, 100, 120, and 140 ms.

Observed *J* values were corrected for the effect of relaxation due to $T_{1sel}(^1H\alpha)$ (Kuboniwa et al., 1994). In general, this correction led to an ~1 Hz increase in $^3J_{HNH\alpha}$. Torsion angle restraint ranges for Φ were set to $-120° \pm 40°$ (for $^3J_{HNH\alpha} \geq 8.5$ Hz), $-120° \pm 50°$ (for 8.5 Hz > $^3J_{HNH\alpha} \geq 8.0$ Hz), and $-60° \pm 40°$ (for $^3J_{HNH\alpha} < 6.0$ Hz and when the NOE data indicated the presence of a helix). Stereospecific Hβ and Val γ-methyl assignments and restraints on χ_1 were based on the values of J_{NH^β} and $J_{C'H^\beta}$ obtained from the GE 3D HNHB and HN(CO)HB experiments, respectively, and on the values of J_{NC^γ} and $J_{C'C^\gamma}$ obtained from the 2D $\{^{15}N\}$ and $\{^{13}C\}$ spin-echo difference, constant-time HSQC experiments. In cases where the data indicated one predominant χ_1 rotamer (~60°, ~−60°, or ~180°), the dihedral angle was restrained to be within 30° of the appropriate staggered conformation. In some cases, it was possible to exclude only one of the three favored rotamers. In total, 54 Φ restraints and 76 χ_1 restraints were included in the structure refinement, along with stereospecific assignments for 36 Hβ pairs and 3 Val γ-methyl pairs. For a restraint to be included in the structure calculations, the derived value of Φ or χ_1 must have been consistent with the both the coupling constant and NOE data.

Amide hydrogen-exchange rates were estimated from the buildup of magnetization in a series of 2D water-exchange HSQC spectra collected with delay times of 4, 8, 16, 32, 64, and 128 ms. Hydrogen bond restraints were included only if the amide hydrogen was shown to have a reduced exchange rate and the initial structure calculations provided unambiguous identification of the hydrogen bond acceptor. In total, 70 restraints defining 35 hydrogen bonds were included; for each hydrogen bond, the HN−O and the N−O distances were restrained to be 2.0 ± 0.4 and 3.0 ± 0.4 Å, respectively.

The last stage of structure refinement included 80 restraints for $^3J_{HNHA}$, 102 restraints based on $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, and 244 restraints based on nonexchangeable ^1H chemical shifts. For the ^1H chemical shifts arising from methylene groups, restraints were included only for those hydrogens that yielded nondegenerate, stereo assigned resonances. These restraints were incorporated as previously described by Clore and co-workers (Garrett et al., 1994; Kuszewski et al., 1995a,b).

Structure Calculations. All calculations were performed on an SGI Power ONYX computer with the program X-PLOR version 3.1 (Brünger, 1992) that had been modified to include chemical shift and *J*-refinement (Garrett et al., 1994; Kuszewski et al., 1995a,b). Structures were generated using a combined distance geometry/restrained dynamic simulated annealing approach (Nilges et al., 1988). In addition to the experimental restraints, functions restraining bond lengths and bond angles to their idealized values were included, as was a soft van der Waals repulsion term. The standard CHARMm22 parameter set was modified by setting all bond length, bond angle, and improper angle force constants to 500 in order to maintain proper stereochemistry. Electrostatics, Leonard-Jones potentials, and empirical dihedral angle functions were excluded. The force constants on the NOE restraints and torsion angle restraints were set to 50 kcal mol^{-1} Å^{-2} and 50 kcal mol^{-1} rad^{-2}, respectively.

Structure refinement proceeded in stages, with the only difference between stages being the number of experimental restraints included in the calculations. During each stage, 50 starting structures were generated with distance geometry by sub-embedding the Cα, Hα, N, HN, C′, Cβ, and Cγ atoms. These structures were subsequently regularized and refined. Refinement included 4000 steps of dynamic simulated annealing at 600 K with the van der Waals radii set to 0.002 times their standard value in the CHARMm22 parameter set; 8000 steps of dynamic simulated annealing during which the temperature was reduced from 600 to 200 K, the vdW radii were increased to 0.75 their standard value, and the vdW repulsive term was increased from 0.003 to 4.0; and 1000 steps of dynamic simulated annealing at 200 K with the vdW scaled by 0.75 and the vdW repulsive term set to 4.0. After structure refinement was complete, the structures were energy minimized for 500 steps with the standard CHARMm22 parameter set, electrostatics, and Leonard-Jones potentials.

The initial stage of structure calculations included only unambiguous restraints, i.e., those NOE restraints that could be defined uniquely by the chemical shift information and those torsion angle restraints derived from measured *J*-values greater than 8.5 Hz. NOE restraints assignable to equivalent and/or nonstereospecifically assigned groups were included as $(\Sigma r^{-6})^{-1/6}$ effective distances (Brünger, 1992; Constantine et al., 1992). Before each subsequent stage of refinement, structures were analyzed to resolve ambiguities in NOE cross-peak assignments, to identify hydrogen bond acceptors of slowly exchanging amide hydrogens, and to restrict the range of torsion angles to their appropriate region of Φ,Ψ space (Kline et al., 1990; Metzler et al., 1992). These additional restraints were then used to further refine the preliminary structures, and the process was repeated. During later stages of refinement, structures were analyzed to identify additional stereo assignments, particularly for the leucine δ-methyl groups. If all pertinent NOEs from all spectra were consistent with only a single stereo assignment for the entire ensemble of structures, the stereo assignment was made. In this way, the δ methyls of 8 of 12 leucines and the α protons for 4 of 8 glycine residues were stereo assigned. The stereo assignments of these groups led to additional restraints as each $(\Sigma r^{-6})^{-1/6}$ effective distance restraint was then expanded to two tighter restraints.

Analysis of the solution conformational ensemble during the later stages of refinement has indicated that the first 11 N-terminal residues and last 4 C-terminal residues are poorly defined by the NMR data. Consistent with this result is the observation of intense, sharp resonances for these residues and *J*-values that suggest conformational averaging. In addition, no long-range NOEs are observed to these residues. As these residues are extraneous to the primary structure of the consensus SH2 domain, no experimental restraints to these residues have been included during the final structure calculations.

After the extraction of all possible restraints, the structures were further refined directly against the measured values $^3J_{HNHA}$ (Garrett et al., 1994), the $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts (Kuszewski et al., 1995b), and nonexchangeable ^1H chemical shifts (Kuszewski et al., 1995a). These structures were then minimized with the standard CHARMm22 parameter set to ensure that they were in a local minimum with respect to all potential terms applied. At this point, Leonard-Jones and electrostatic potential terms were included in addition to bond lengths, bond angles, impropers, and experimental restraints. The final structures were analyzed

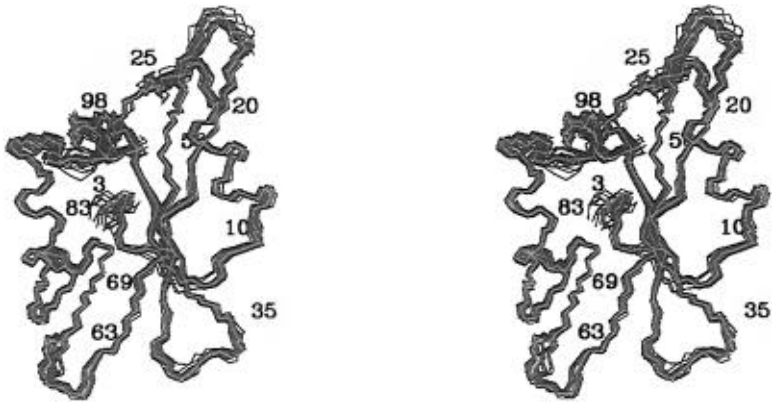

FIGURE 1: Stereoview of the 20 structures representing the ensemble of NMR solution conformations of the *blk* SH2 domain (black). The structures were superimposed for best fit of the backbone (N, Cα, C′) atoms of residues 3−101 to the average structure (red). In this and all following figures, the first 11 N-terminal residues (−9 through 2) and last 4 C-terminal residues (102−105) are not shown.

for best fit to the experimental NMR restraints and the overall energy. As the energy did not vary significantly among the structures, the 20 structures with the lowest residual experimental restraint violations were chosen to represent the ensemble of solution conformations for the *blk* SH2 domain.

Molecular graphics analysis was performed within the program INSIGHT II (Biosym Technologies). The CHARMm force field parameter set was generated by the program QUANTA (Molecular Simulations Inc.). Coordinates for the 20 refined structures, both before and after chemical shift and $^3J_{HNHA}$ refinement, have been deposited in the Brookhaven Protein Data Bank. The ID codes for these entries are 1blj and 1blk, respectively.

RESULTS

Distance Restraints Derived from 4D NOESY Spectra. The NOE restraints used in the structure calculations of the *blk* SH2 domain were derived solely from 4D NOESY spectra. Some concerns have been raised over the interpretation of interproton distance restraints derived solely from 4D NOESY data, due primarily to the differential effects that coherence transfers and relaxation might have on the observed NOE intensities. The overall signal-to-noise ratio of the 4D NOESY spectra relative to their 3D counterparts has also been an issue in this regard. We therefore compared the intensities of numerous NOE cross-peaks in 4D NOESY spectra with the intensities of corresponding NOE cross-peaks in 3D NOESY spectra. The relative intensities within the sets of cross-peaks were found to be essentially indistinguishable. Moreover, no cross-peaks were observed in the 3D NOESY spectra that were not also observed in the 4D NOESY spectra; in fact, due to the additional resolution provided by the fourth dimension, a larger number of resolved cross-peaks were identified in the 4D NOESY spectra. This observation should generally hold under the following conditions: (1) the 4D data collection time is twice that of the 3D; and (2) the protein MW is ≤ 15 000−20 000. For proteins whose MW is >20 000, the extra INEPT transfer in the 4D may measurably degrade the signal-to-noise ratio. By using the 4D NOESY spectra, more unambiguous cross-peak assignments can be made, which both facilitates the initial structure calculations and subsequently reduces the dependence on a structure-based assignment strategy for the remaining ambiguous NOEs.

Experimental Restraint Satisfaction. Twenty structures were selected to represent the ensemble of solution conformations for the *blk* SH2 domain (Figure 1). These structures were refined with the program X-PLOR and were selected on the basis of their rms difference from the experimental restraints and their deviations from idealized stereochemistry. Statistical values for the final ensemble of structures are listed in Table 2. Before refinement against $^3J_{HNHA}$ and the ^{1}H, ^{13}Cα, and ^{13}Cβ chemical shifts, all of the individual structures were consistent with the experimental NOE restraints; only a single NOE violation was greater than 0.3 Å, and relatively few were larger than 0.2 Å: the largest violation for any individual structure was 0.31 Å. Similarly, the structures satisfied all Φ and χ_1 torsion angle restraints quite well, with the largest individual violation being 2.4°. Refinement against $^3J_{HNHA}$ and the ^{1}H, ^{13}Cα, and ^{13}Cβ chemical shifts resulted in an increase in the residual NOE violations, a decrease in residual torsion angle restraint violations, and a minimal effect on deviations from idealized geometry (Table 2). Although the rms NOE violation increased from 0.027 to 0.044 Å, the NOE restraint satisfaction remained good. In the structures refined against $^3J_{HNHA}$ and the ^{1}H, ^{13}Cα, and ^{13}Cβ chemical shifts, on average only one restraint was violated by more than 0.4 Å per structure, with the largest single violation being 0.47 Å. All of these violations involved NOEs for which at least one of the partners was a methyl group in the core of the protein and for which there were obvious pathways leading to spin diffusion.

It has previously been shown that inclusion of $^3J_{HNH\alpha}$, the ^{1}H chemical shift, and the ^{13}Cα and ^{13}Cβ chemical shifts during refinement improves the accuracy of the experimentally determined NMR structures relative to a high resolution crystal structure (Garrett et al., 1994; Kuszewski et al., 1995a,b; Le et al., 1995; Pearson et al., 1995). Nevertheless, because the refinement against $^3J_{HNH\alpha}$, ^{1}H chemical shifts, and ^{13}Cα and ^{13}Cβ chemical shifts led to small increases in the violation of NOE restraints, we investigated whether including these restraints might have a detrimental effect on the quality of the refined structures of the *blk* SH2 domain. We found that the residual violations of the experimental restraints were independent of the order in which the $^3J_{HNH\alpha}$ and chemical shift restraints were introduced when these restraints were introduced successively rather than all at once. Moreover, successive introduction of one class of these restraints always led to a reduction in the overall violations of all previously introduced classes, except for the NOE restraints. For example, refinement against $^3J_{HNH\alpha}$ reduced

Table 2: Structural Statistics for the *blk* SH2 Domain

	ensemble A[a]	ensemble B[a]
rmsd (N, Cα, C′, residues 3−101) (Å)	0.60	0.51
rmsd (non-hydrogens, residues 3−101) (Å)	1.06	0.95
deviations from experimental restraints[b]		
rmsd of NOE restraints (Å)	0.027 ± 0.001	0.043 ± 0.004
largest NOE violation[c] (Å)	0.31	0.47
no. of NOE violations >0.4 (Å)	0	0.95 ± 1.1
no. of NOE violations >0.3 Å	0.1 ± 0.3	4.5 ± 1.9
no. of NOE violations >0.2 Å	2.2 ± 1.4	19.7 ± 5.6
rmsd of torsion restraints		
all (Φ and χ_1)	0.33 ± 0.04	0.11 ± 0.08
Φ	0.12	0.02
χ_1	0.41	0.14
largest torsion angle violation[c]	2.4	2.4
rmsd of $^3J_{HNHA}$ (Hz)	1.76	0.47
rmsd of chemical shifts (ppm)		
$^{13}C\alpha$	1.45	1.10
$^{13}C\beta$	1.50	1.21
^1H (all)	0.44	0.30
^1H (Hα)		0.29
^1H (methyl)		0.19
^1H (other)		0.32
deviations from idealized covalent geometry		
rmsd of bond lengths (Å)	0.013 ± 02e−4	0.015 ± 6e−4
rmsd of bond angles (deg)	3.16 ± 0.03	3.54 ± 0.11
rmsd of improper angles (deg)	1.56 ± 0.08	2.1 ± 0.1
energies[d]		
E(NOE)	59 ± 5	157 ± 25
E(cdih)	2.1 ± 0.6	0.34 ± 0.44
E(L-J)	−978 ± 20	−759 ± 45
E(bond)	90 ± 2	127 ± 10
E(angle)	420 ± 11	509 ± 30
E(improper)	39 ± 3	63 ± 6
$E(^3J_{HNHA})$	499 ± 54	36 ± 8
$E(^{13}C$ shift)	674 ± 75	397 ± 22
$E(^1H$ shift)	347 ± 26	162 ± 9

[a] Ensembles A and B refer to the ensemble of *blk* SH2 structures refined without and with $^3J_{HNHA}$ (Garrett et al., 1994), $^{13}C\alpha$ and $^{13}C\beta$ chemical shift (Kuszewski et al., 1995b), and ^1H chemical shift (Kuszewski et al., 1995a) restraints, respectively. [b] All deviations are represented as the average ± 1 standard deviation and are calculated for the ensemble of 20 *blk* SH2 domain structures. [c] The largest individual NOE or torsion angle restraint violation in any of the 20 refined structures is listed. [d] Energies are calculated within X-PLOR and are reported in units of kcal mol^{-1}. *E*(NOE) and *E*(cdih) were calculated with a square-well quadratic potential with force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively. *E*(L-J) is the Leonard-Jones van der Waals energy calculated with the X-PLOR empirical energy function and CHARMm22 parameter set.

the rms deviation of these restraints from 1.76 to 0.61 Hz and reduced the rms deviation of Φ torsion angle restraints from 0.12° to 0.04°. Additional refinement against $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts further reduced the rms deviation of the $^3J_{HNH\alpha}$ restraints from 0.61 to 0.51 Hz. Refinement against ^1H chemical shifts reduced the rms deviation of the $^3J_{HNH\alpha}$ restraints still further, to 0.47 Hz, and reduced the rms deviation of χ_1 torsion angle restraints from 0.37° to 0.14°. We note that improvements on the order of 0.1 Hz in the rms deviation of $^3J_{HNH\alpha}$ restraints are less than our precision in the measurements of $^3J_{HNH\alpha}$ (∼0.4 Hz). Improvements of this magnitude may therefore be insignificant. However, such trends were observed for all combinations of refinement.

Refinement against $^3J_{HNHA}$ and ^1H, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts yielded significant improvement in the satisfaction of these restraints (Table 2), with only a small perturbation to the overall structure of the protein. The energy-minimized average structures of the ensembles refined with and without $^3J_{HNHA}$ and ^1H, $^{13}C\alpha$, and $^{13}C\beta$ chemical shift restraints superimpose with an rmsd of 0.73 Å for the backbone atoms and an rmsd of 0.87 Å for all non-hydrogen atoms. The most significant deviations between the structures lie in the region where the NMR structures using NOEs only are least well defined. The precision of the structure determination improved from 0.60 to 0.51 Å for the ensemble upon $^3J_{HNHA}$ and ^1H, $^{13}C\alpha$, and $^{13}C\beta$ chemical shift refinement. As the improved precision most likely results directly from the incorporation of additional experimental restraints, we have chosen this ensemble of structures for further analysis and discussion. We note, however, that neither the results nor conclusions are affected by this choice.

Analysis of the backbone dihedral angles Φ and Ψ indicate that most lie in the energetically favored regions of conformational space as defined by the program PROCHECK (Laskowski et al., 1993). Furthermore, the values for the Leonard-Jones energy are large and negative, indicating that the structures are well folded with no or minimal bad nonbonded contacts.

Analysis of the Solution Conformational Ensemble of the blk SH2 Domain. The secondary structural elements for the *blk* SH2 domain were determined by analysis of the Φ,Ψ angles and hydrogen bonding patterns of the refined structures. The ends of the secondary structural elements were defined by analysis of the deviation of ^1H, $^{13}C'$, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts from random coil values (Supporting Information), in conjunction with the NOE data. Examination of the peptide chain starting at its N-terminus leads to the following secondary structure: βA (residues W5−F7); αA (R12−L19); βB (S28−E33); βC (A40−I48); βD (E53−

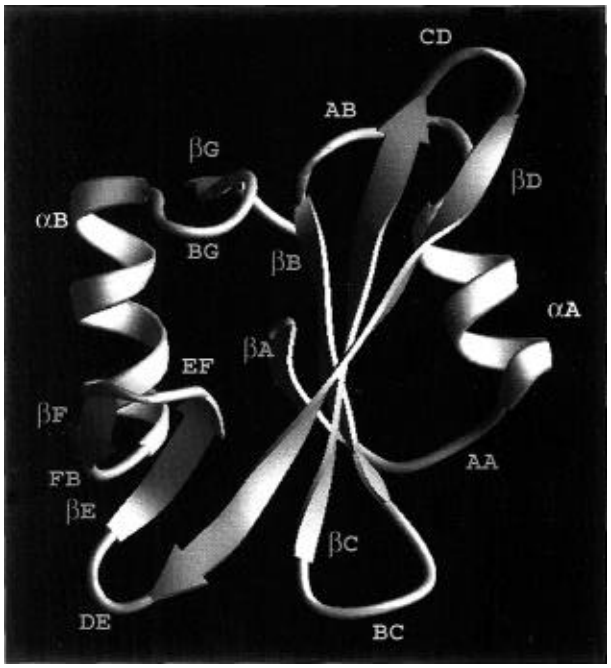

FIGURE 2: Ribbon representation depicting the three-dimensional structure of the *blk* SH2 domain. The sequence is numbered such that the conserved tryptophan is residue 5, and the secondary structure elements are labeled according to the nomenclature of Eck et al. (1993). The *blk* SH2 domain contains the following secondary structure: βA (residues W5−F7); αA (R12−L19); βB (S28−E33); βC (A40−I48); βD (E53−L63); βE (G67−I70); βF (I74−F76); αB (L79−K89); βG (T97−L98).

L63); βE (G67−I70); βF (I74−F76); αB (L79−K89); βG (T97−L98).

The *blk* SH2 domain folds into a single compact globular domain. The structure determined here for the *blk* SH2 domain (Figure 2) is very similar to previously determined structures of *src* family SH2 domains (Eck et al., 1993, 1994; Gilmer et al., 1994; Lee et al., 1994; Maignan et al., 1995; Mikol et al., 1995; Narula et al., 1995; Overduin et al., 1992b; Waksman et al., 1992, 1993; Xu et al., 1995). A central three-stranded antiparallel β-sheet (strands βB, βC, and βD) makes up the core of the protein, with an additional parallel β-strand βA extending the βB edge of the sheet and parallel β-strand βG extending the βD edge of the sheet. A second smaller antiparallel sheet (βE and βF) adjoins the C-terminal end of βD. Two α-helices (αA and αB) lie on opposite faces of the central sheet.

As previously described, the first eleven N-terminal residues and the last 4 C-terminal residues are conformationally averaged (see Experimental Procedures) and are excluded from the following discussion. Otherwise, the structural ensemble is quite well defined by the NMR data. The average rms distance deviation for the backbone N, Cα, and C′ atoms from the average structure is 0.51 Å. Figure 3a shows a plot of the rms distances of the backbone atoms from the average structure vs primary structure. Several regions of the structure exhibit relatively higher conformational variability, namely, residues 34−39, 50−53, and 64−67. These residues correspond to the loop regions between β-strands B and C, C and D, and D and E of the β-sheets and generally correspond to the more solvent-exposed regions of the structure (Figure 3c). The higher rmsd for residues 64−67 arises from the existence of two discrete conformers for the DE loop in the NMR ensemble. The two conformers differ primarily in the Φ and Ψ dihedral angles of glycines 65 and 66. Low order parameters were calculated for the analogous residues (Q65, T66) in the SH2 domain of PLC-γ1 on the basis of ^{15}N and ^{1}H relaxation measurements (Farrow et al., 1994). Low order parameters were also calculated for the PLC-γ1 residues analogous to *blk* residues 35−39 (Farrow et al., 1994); thus these residues are likely to possess enhanced mobility on the nanosecond to picosecond time scales. Also consistent with this observation are the above average temperature factors for the loop regions in the crystal structure of the v-*src* SH2 domain (Waksman et al., 1993). If residues 35−39, 49−52, and 64−66 are excluded from the calculation, the average backbone atom rmsd to the mean structure is 0.42 Å. Many of the side chain residues are also well defined by the NMR data. The average rms distance deviations from the average structure are 0.95 Å for all non-hydrogen atoms and 0.83 Å for all non-hydrogen atoms excluding those in the aformentioned dynamic residues.

DISCUSSION

Phosphate Ion/blk Interactions. Isolated SH2 domains can directly bind to free phosphotyrosine (Lemmon & Ladbury, 1994) and to phosphotyrosine coupled to agarose beads (Mayer et al., 1991). In addition, inorganic phosphate binds to the SH2 domains in a fashion similar to that of the phosphate group of pY (Pascal et al., 1995; Waksman et al., 1993). The structure of the *blk* SH2 domain reported here was determined in the presence of 10 mM phosphate; therefore, it may not be entirely appropriate to refer to it as the "uncomplexed" *blk* SH2 domain. However, given the abundance of various forms of phosphate (P_i, PP_i, ATP, etc.) in the cell, the structure in the presence of phosphate *may* more adequately represent the uncomplexed *blk* SH2 domain than one solved in the presence of imidazole or acetate ion.

Recent analysis of the NMR relaxation, hydrogen exchange, and chemical shifts of arginine guanidino groups has led Pascal et al. (1995) to suggest that the phosphate ion interacts primarily with the arginines in the phosphotyrosine binding pocket. To examine whether this interaction might also be occurring in *blk* SH2, 2D NH_2- and NH-selective HSQC experiments and a 2D NH2/NH-H(N)CZ experiment (Farmer & Venters, 1996) were acquired in an effort to detect and assign any of the Arg guanidino ^{1}Hη/^{15}Nη resonances. A single ^{1}Hη /^{15}Nη resonance was observed for the six Arg residues of *blk* SH2, suggesting that one of the Arg residues is conformationally restrained. In addition, the Hϵ and Nϵ resonances of Arg 32 of *blk* SH2 have anomalous chemical shifts. Although the spectroscopically observed Nη and Hη resonances have not yet been explicitly assigned, these results are consistent with Arg 32 of *blk* SH2 forming hydrogen bonds with phosphate ion in the pY binding pocket, analogous to the results for R37 in PLCγ_1 SH2 domain (Pascal et al., 1995).

Spectroscopic Anomalies Arising from the βC Strand of blk. Anomalous behavior was detected for several residues in the βC strand of the *blk* SH2 domain. These include the absence of resonances corresponding to ^{1}Hα/^{13}Cα of Thr 49 and backbone amide ^{1}H/^{15}N of Ile 48 and Thr 50. While the failure to detect the backbone amide protons could be explained by rapid exchange with the H_2O solvent protons, such a process cannot account for the absence of the Cα,

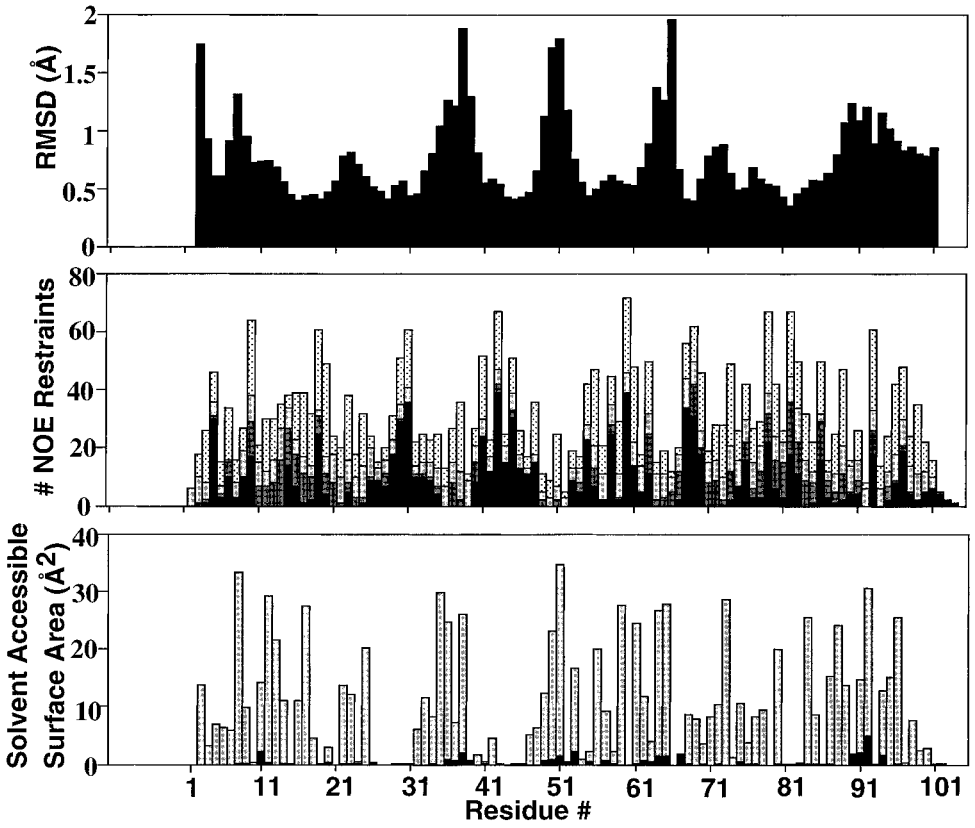

FIGURE 3: (a, top) Plot of the rmsd of the backbone (N, Cα, C′) atoms vs the primary structure of the *blk* SH2 domain. The structures were superimposed for best fit of the backbone (N, Cα, C′) atoms of residues 3−101 to the average structure; rmsds were then calculated for each residue. (b, middle) Stacked bar plot showing the NOE restraint density per residue. For each residue, long-range restraints are indicated by solid bars, and medium-range, sequential, and intraresidue restraints are indicated by increasingly less dense shaded bars. Both residues defined by the NOE interaction contain the information on each interresidue restraint; thus, each interresidue restraint is counted twice in this figure. (c, bottom) Stacked bar plot showing the solvent-acessible surface area of the backbone (solid bars) and side chain (hatched bars) atoms for each residue of the *blk* SH2 domain.

Hα resonances of Thr 49. Instead, it is more likely that these resonances are not detected because of chemical exchange arising from conformational averaging on the appropriate time scale. This inference is supported by the observation of multiple resonances for the γCH_3 of Ile 48, indicating localized conformational heterogeneity, and is borne out structurally in the NMR ensemble by these residues having above average rmsds upon superposition of the backbone atoms of *blk* SH2 (Figure 3a). While the cause of the exchange broadening of these resonances is not currently known, one may speculate that it is the result of weak dimerization between *blk* SH2 monomers. The above set of atoms lies at the C-terminal end of β-strand C, with this edge of the strand being exposed to solvent in the monomeric structure of *blk* SH2 determined here. Model building indicates that dimerization is conformationally plausible and could proceed by the formation of an intermolecular β-sheet involving the C-terminal end of β-strand C from two monomers. We note here that dimer interfaces are frequently composed of an intermolecular β-sheet, as in, for example, P22 Arc repressor (Raumann et al., 1994), RANTES (Chung et al., 1995), and apo-CRABPI (Thompson et al., 1995). Farrow et al. (1994) have previously shown that the uncomplexed form of the phospholipase Cγ1 SH2 domain is in a monomer−dimer equilibrium. In their system, the dimerization is removed upon complexation of phosphopeptide to the SH2 domain.

Comparison with the Structures of Other SH2 Domains. Sequence alignments and secondary structures are shown for all structures whose coordinates were available in the Brookhaven Protein Data Bank at the time of manuscript preparation (Figure 4). The structure of *blk* SH2 is most similar to that of *src* SH2, as determined by superposition of the backbone atoms in the secondary structure elements. Figure 5a also shows the superposition of the backbone atoms of *blk* and *src* SH2 domains. The rmsd of all structurally defined backbone atoms between the two proteins is 1.52 Å and drops to only 0.82 Å if the AB, BC, CD, and DE loops are excluded. The per residue backbone rmsd is plotted against primary structure in Figure 5b. In many regions of the sequence, the structure of *src* SH2 is as similar to *blk* SH2 as the NMR ensemble is to itself (comparison of the thin and thick lines in Figure 5b). The three large peaks in this plot indicate those regions where *blk* SH2 and *src* SH2 are most dissimilar. The peak at residue 51 reflects the structural differences arising from the one-residue insertion in *src* SH2 relative to *blk* SH2. The peak near residue 65 corresponds to the DE loop of *blk* SH2 in which two solution-state conformations were detected. Interestingly, one of the two conformations observed in the NMR structures for the DE loop is very similar to that seen in the crystal structure of *src* SH2. The dashed line in Figure 5b indicates the rms between this particular conformation and the crystal structure. The most striking difference between the crystal structure of *src* SH2 and the NMR structure of *blk* SH2 is the position of the BC loop (*blk* SH2 residues 34−39). Although only minor structural differences are observed between the crystal structures of apo *src* SH2 and phospopeptide-complexed *src*

```
          1        10        20        30        40        50        60        70        80        90       100       110       120
          ....:....+....:....+....:.... +....:....+....:.. ..+....:.... +....:....+ ....:....+....:....+....:....+....:....+....:....+
blk       GSVAPVETLEVEKWFFRTISRKDAERQL-LAPMNKAGSFLIRESESN-KGAFSLSVKDIT-TQGEVVKHYKI-RSLDNGGYYISPRITFPTLQALVQHYSKKGDGLCQKLTLPCVNLA----------
lck       ASMTGGQQMGRGSWFFKNLSRKDAERQL-LAPGNTHGSFLIRESEST-AGSFSLSVRDFDQNQGEVVKHYKI-RNLDNGGFYISPRITFPGLHDLVRHYTNASDGLCTRLSRPC----------
src       ----------AEEWYFGKITRRESERLL-LNPENPRGTFLVRESETT-KGAYCLSVSDFDNAKGLNVKHYKI-RKLDSGGFYITSRTQFSSLQQLVAYYSKHADGLCHRLTNVCPTS----------
abl       ---GSGNSLEKHSWYHGPVSRNAAEYLL-SSGIN--GSFLVRESESS-PGQRSISLR-YEG----RVYHYRI-NTASDGKLYVSSESRFNTAELVHHHSTVADGLITTLHYPAPKR----------
pi3       -------LQDA-EWYWGDISREEVNEKL-RDTAD--GTFLVRDASTKMHGDYTLTLR-KGG----NNKLIKIFHRDGKYGFSDP--LTFNSVVELINHYRNE-----SLAQY-NPKLDVKLLYPVSKY----
plc       ---GSPGIHESKEWYHASLTRAQAEHMLMRVPRD--GAFLVRKRNE--PNSYAISFR-AEG----KIKHCRV-QQEGQTVMLGN--SEFDSLVDLISYYEKHPLYRKMKLRYPINEENSSDNDYIIPLPDPK
                  βA        αA                βB        βC                βD        βE  βF        αB                βG
```

FIGURE 4: Sequence alignment of SH2 domains whose coordinates have been deposited in the Brookhaven protein Data Bank. Bars below the sequences indicate residues determined to be in secondary structural elements.

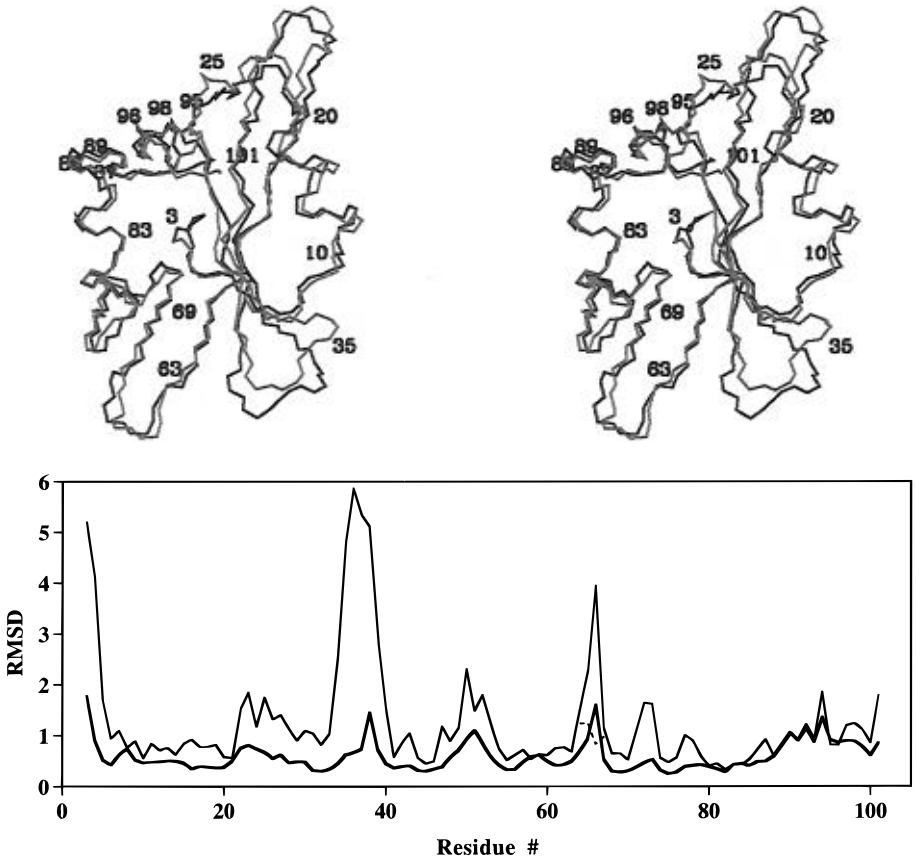

FIGURE 5: Comparison of the structure of the *blk* SH2 domain to the structure of the uncomplexed v-*src* SH2 domain. (a, top) Stereoview of the *blk* (black) and v-*src* (red) SH2 domain structures superimposed for best fit of the backbone (N, Cα, C′) atoms of residues in ordered secondary structures (see Figure 4). (b, bottom) Plot of the rmsds of the backbone (N, Cα, C′) atoms of the *blk* SH2 domain structural ensemble against primary structure. The structures were superimposed for best fit of the backbone atoms of residues 3−101 for each structure in the ensemble either to the average structure of the *blk* SH2 domain (thick line) or to the structure of the uncomplexed v-*src* SH2 domain (thin line). The rmsds for the two solution-state conformations of residues 64−67 are indicated by the thin solid and thin dashed lines. Rmsds were calculated and averaged for each residue after the superposition.

SH2 (Waksman et al., 1993), large conformational changes observed in the BC loop of *lck* SH2 upon peptide binding have led to the suggestion of a gated peptide binding site for the *lck* SH2 domain (Mikol et al., 1995). The conformation of the BC loop in *blk* SH2 is analogous to the open conformation in uncomplexed *lck* SH2, and it is likely that *blk* SH2 also contains a gated peptide binding site. This conclusion is consistent with NMR binding studies of high-affinity peptides to *blk* SH2. These studies indicate a perturbation in the chemical shifts of ^{15}N and $^{1}H_N$ atoms in the BC loop (data not shown).

blk/fyn SH2 Domain Specificities. Although SH2 domains share the common property of binding phosphotyrosine-containing peptides, they vary with respect to the phosphopeptide binding specificity. Three classes of SH2 domains have been identified on the basis of the ligands they bind (Songyang et al., 1993): class I SH2 domains, which include those of the *src* family SH2 domains, preferentially bind the motif pY−hydrophilic−hydrophilic−Ile/Pro; class II SH2 domains, which include phosphatidylinositol 3-kinase and phospholipase C-γ SH2 domains, preferentially bind the motif pY−hydrophobic−X−hydrophobic; and class III SH2 domains, which include several SH2 domain-containing protein tyrosine phosphatases (e.g., SH-PTP2), preferentially bind a motif in which the pY+1 site requires a β-branched residue, the pY+3 site requires a hydrophobic residue, and the pY−2 site is also critical, although its precise determinants are not yet clear (Huyer et al., 1995). Within class I

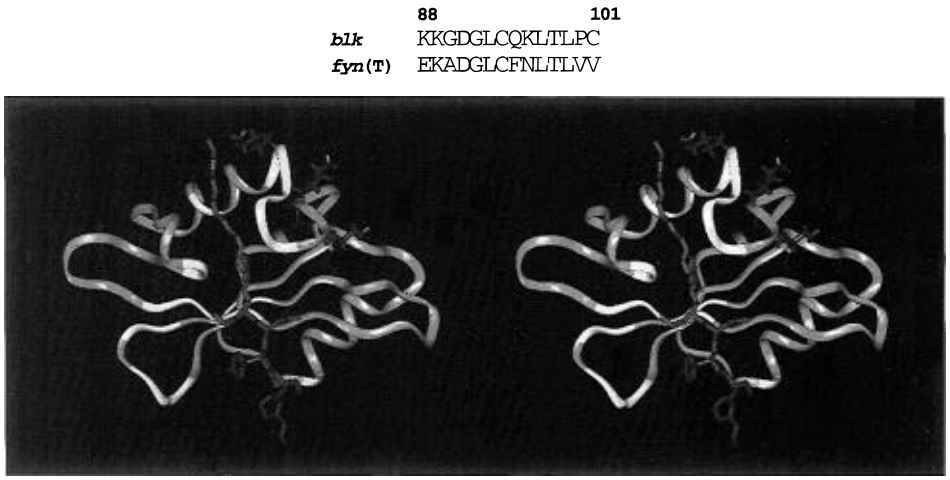

FIGURE 6: (a, top) Alignment of the 14 residues of *blk* and *fyn*(T) SH2 domains that are sufficient to transfer phosphoprotein specificity between *blk* and *fyn*(T) in chimeric proteins (Malek & Desiderio, 1993). (b, bottom) The structure of the *blk* SH2 domain is colored to highlight the potential discriminants between *blk*/*fyn*(T) SH2 specificities. Residues of the *blk* SH2 domain (green) whose amide $^{1}H/^{15}N$ resonances are perturbed in chemical shift by pY-peptide binding (data not shown) are highlighted in cyan. The side chain residues that are nonconserved between *blk* and *fyn*(T) SH2 domains (see text) are colored red. The pY-peptide is colored magenta and is built into the binding site in an orientation analogous to that found for the v-*src* SH2 domain. The peptide is extended at its C-terminus past the conserved Ile occupying the hydrophobic binding pocket.

SH2 domains, additional specificity has been identified. Class I SH2 domains from *src* family protein−tyrosine kinases have been shown to bind distinct sets of phosphoproteins from B lymphocytes under conditions of high stringency, i.e., relatively low SH2 concentrations (Malek & Desiderio, 1993). Furthermore, a subset of the *blk* SH2 specificity for phosphoproteins could be transferred to *fyn*(T) SH2 in chimeric SH2 proteins derived from *blk* and *fyn*(T) sequences. In particular, replacing the C-terminal 14 residues of *fyn*(T) SH2 with the corresponding sequence from *blk* was sufficient to confer *blk*-like binding properties to *fyn*(T) (Malek & Desiderio, 1993). These 14 residues correspond to the αB, BG, and βG regions of the SH2 structure (Figures 2 and 6) and help form the hydrophobic binding pocket in class I SH2 domains. Of these residues, eight are identically conserved between *blk* and *fyn*(T) SH2 domains, including the two residues (G95, L96) that have been shown to contact the phosphopeptide in the *src* SH2 structure. The determinants of differential specificity between *blk* and *fyn*(T) SH2 domains are likely to lie in nonconserved residues whose contact with phosphopeptide has yet to be demonstrated. The reason that these contacts have not yet been observed may be because previous structural studies on pY−peptide−SH2 complexes have been typically carried out with the minimal peptide sequence required for high-affinity binding (five residues) and because such peptides were therefore of insufficient length to make these additional contacts. We have performed binding studies between the *blk* SH2 domain and peptides of various lengths and have found that the addition of residues C-terminal to the pY+3 residue can result in an approximately 10-fold increase in affinity (manuscript in preparation). Similar results have been observed for peptides binding to the class II SH2 domain from phosphatidylinositol 3-kinase p85; in those studies, however, the additional residues C-terminal to the pY+3 residue did not appear to confer specificity but rather were necessary for the higher affinity (Piccione et al., 1993). By extending a phosphopeptide at its C-terminal end in the orientation that has been observed in the structures of class I SH2−peptide complexes (Figure 6b), additional contacts between SH2 domain and pY-peptide are possible. We note that the binding of a pY-peptide to the *blk* SH2 domain perturbs the chemical shifts of the amide $^{1}H/^{15}N$ resonances of several residues in this region, providing indirect support for the placement of the peptide residues in this location. Lys 88 of *blk* SH2 is a nonconserved residue [Glu in *fyn*(T)] that lies poised to interact either directly with a pY-peptide or indirectly through electrostatic effects on Lys 89. Thus we suggest that Lys 88 may be an important determinant for the observed difference in specificity between *blk* and *fyn*(T).

ACKNOWLEDGMENT

The authors gratefully acknowledge Drs. Joseph Bolen and Anne Burkhardt for providing the cDNA clone of *blk* SH2 domain and for stimulating interest in this project. We also thank Dr. Julie Forman-Kay for a preprint of her manuscript.

SUPPORTING INFORMATION AVAILABLE

A description of NMR experiments and data processing, a table listing chemical shifts for the *blk* SH2 domain, a histogram showing the deviations of the experimental chemical shifts from their random coil values, and a figure showing the pulse sequence for the 4D $^{13}C/^{13}C$-separated NOESY experiment (28 pages). Ordering information is given on any current masthead page.

REFERENCES

Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636−641.

Bax, A., Max, D., & Zax, D. (1992) *J. Am. Chem. Soc. 114*, 6923−6925.

Billeter, M., Neri, D., Otting, G., Qian, Y., & Wuthrich, K. (1992) *J. Biomol. NMR 2*, 257−274.

Booker, G. W., Breeze, A. L., Downing, A. K., Panayotou, G., Gout, I., Waterfield, M. D., & Campbell, I. D. (1992) *Nature 358*, 684−687.

Brünger, A. T. (1992) *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.

Burkhardt, A., Brunswick, M., Bolen, J., & Mond, J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7410−7414.

Campbell, M. A., & Sefton, B. M. (1992) *Mol. Cell. Biol. 12*, 2315−2321.
Chung, C., Cooke, R., Proudfoot, A., & Wells, T. (1995) *Biochemistry 34*, 9307−9314.
Constantine, K., Madrid, M., Banyai, L., Trexler, M., Patthy, L., & Llinas, M. (1992) *J. Mol. Biol. 223*, 281−298.
Constantine, K. L., Colson, K. L., Wittekind, M., Friedrichs, M. S., Zein, N., Tuttle, J., Langley, D. R., Leet, J. E., Schroeder, D. R., Lam, K. S., Farmer, B. T., II, Metzler, W. J., Bruccoleri, R. E., & Mueller, L. (1994) *Biochemistry 33*, 11438−11452.
Eck, M. J., Shoelson, S. E., & Harrison, S. C. (1993) *Nature 362*, 87−91.
Eck, M. J., Atwell, S. K., Shoelson, S. E., & Harrison, S. C. (1994) *Nature 368*, 764−769.
Fantl, W. J., Escobedo, J. A., Martin, G. A., Turck, C. W., del Rosario, M., McCormick, F., & Williams, L. T. (1992) *Cell 69*, 413−423.
Farmer, B. T., II, & Mueller, L. (1994) *J. Biomol. NMR 4*, 673−687.
Farmer, B., II, & Venters, R. (1996) *J. Biomol. NMR 7*, 59−71.
Farmer, B., II, Lavoie, T., Mueller, L., & Metzler, W. (1995) *J. Magn. Reson., Ser. B 107*, 197−200.
Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman Kay, J. D., & Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
Friedrichs, M. S., Mueller, L., & Wittekind, M. (1994) *J. Biomol. NMR 4*, 703−726.
Garrett, D., Kuszewski, J., Hancock, T., Lodi, P., Vuister, G., Gronenborn, A., & Clore, G. (1994) *J. Magn. Reson., Ser. B 104*, 99−103.
Gilmer, T., Rodriguez, M., Jordan, S., Crosby, R., Alligood, K., Green, M., Kimery, M., Wagner, C., Kinder, D., Charifson, P., et al. (1994) *J. Biol. Chem. 269*, 31711−31719.
Grzesiek, S., & Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291−6293.
Grzesiek, S., & Bax, A. (1993a) *J. Biomol. NMR 3*, 185−204.
Grzesiek, S., & Bax, A. (1993b) *J. Biomol. NMR 3*, 627−638.
Huyer, G., Li, Z. M., Adam, M., Huckle, W. R., & Ramachandran, C. (1995) *Biochemistry 34*, 1040−1049.
Kay, L. E., Keifer, P., & Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., & Forman-Kay, J. D. (1993) *J. Magn. Reson., Ser. B 101*, 333−337.
Kay, L. E., Xu, G. Y., & Yamazaki, T. (1994) *J. Magn. Reson., Ser. A 109*, 129−133.
Kline, T. P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805−7813.
Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., & Pawson, T. (1991) *Science 252*, 668−674.
Kuboniwa, H., Grzesiek, S., Delaglio, F., & Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.
Kuszewski, J., Gronenborn, A., & Clore, G. (1995a) *J. Magn. Reson., Ser. B 107*, 293−297.
Kuszewski, J., Qin, J., Gronenborn, A., & Clore, G. (1995b) *J. Magn. Reson., Ser. B 106*, 92−96.
Laskowski, R. A., McArthur, M., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Law, D. A., Chan, V. W. F., Datta, S. K., & DeFranco, A. L. (1993) *Curr. Biol. 3*, 645−657.
Le, H.-b., Pearson, J. G., de Dios, A. C., & Oldfield, E. (1995) *J. Am. Chem. Soc. 117*, 3800−3807.
Lee, C. H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, J., Shoelson, S. E., & Kuriyan, J. (1994) *Structure 2*, 423−438.
Lemmon, M. A., & Ladbury, J. E. (1994) *Biochemistry 33*, 5070−5076.
Maignan, S., Guilloteau, J. P., Fromage, N., Arnoux, B., Becquart, J., & Ducruix, A. (1995) *Science 268*, 291−293.
Malek, S. N., & Desiderio, S. (1993) *J. Biol. Chem. 268*, 22557−22565.
Marengere, L. E., & Pawson, T. (1994) *J. Cell Sci., Suppl. 18*, 97−104.
Mayer, B., Jackson, P., & Baltimore, F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 627−631.
Metzler, W. J., Valentine, K., Roebber, M., Friedrichs, M. S., Marsh, D. G., & Mueller, L. (1992) *Biochemistry 31*, 5117−5127.
Metzler, W. J., Constantine, K. L., Friedrichs, M. S., Bell, A. J., Ernst, E. G., Lavoie, T. B., & Mueller, L. (1993) *Biochemistry 32*, 13818−13829.
Mikol, V., Baumann, G., Keller, T. H., Manning, U., & Zurini, M. G. (1995) *J. Mol. Biol. 246*, 344−355.
Moran, M. F., Koch, C. A., Anderson, D., Ellis, C., England, L., Martin, G. S., & Pawson, T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8622−8626.
Narula, S., Yuan, R., Adams, S., Green, O., Green, J., Philips, T., Zydowsky, L., Botfield, M., Hatada, M., Laird, E., Zoller, M., Karas, J., & Dalgarno, D. (1995) *Structure 3*, 1061−1073.
Nilges, M., Clore, G., & Gronenborn, A. (1988) *FEBS Lett. 229*, 317−324.
Overduin, M., Mayer, B., Rios, C. B., Baltimore, D., & Cowburn, D. (1992a) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11673−11677.
Overduin, M., Rios, C. B., Mayer, B. J., Baltimore, D., & Cowburn, D. (1992b) *Cell 70*, 697−704.
Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelson, S. E., Pawson, T., Kay, L. E., & Forman Kay, J. D. (1994) *Cell 77*, 461−472.
Pascal, S., Yamazaki, T., Singer, A., Kay, L., & Forman-Kay, J. (1995) *Biochemistry 34*, 11353−11362.
Pawson, T. (1995) *Nature 373*, 573−580.
Pawson, T., & Gish, G. D. (1992) *Cell 71*, 359−362.
Pearson, J. G., Wang, J.-F., Markley, J. L., Le, H.-b., & Oldfield, E. (1995) *J. Am. Chem. Soc. 117*, 8823−8829.
Piccione, E., Case, R. D., Domchek, S. M., Hu, P., Chaudhuri, M., Backer, J. M., Schlessinger, J., & Shoelson, S. E. (1993) *Biochemistry 32*, 3197−3202.
Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994) *Nature 367*, 754−757.
Songyang, Z., Shoelson, S., CHaudhuri, M., Gish, G., Pawson, T., Haser, W., King, F., Roberts, T., Ratnofsky, S., Lechleider, R., Neel, B., Birge, R., Fajardo, J., Chou, M., Hanafusa, J., Schaffhausen, B., & Cantley, L. (1993) *Cell 72*, 767−778.
Thompson, J., Bratt, J., & Banaszak, L. (1995) *J. Mol. Biol. 252*, 443−446.
von Bonin, A., Wienands, J., Manning, U., Zuber, J. F., & Baumann, G. (1994) *J. Biol. Chem. 269*, 33035−33041.
Vuister, G. W., & Bax, A. (1993) *J. Magn. Reson., Ser. B 101*, 210−213.
Vuister, G. W., Wang, A. C., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 5334−5335.
Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., et al. (1992) *Nature 358*, 646−653.
Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., & Kuriyan, J. (1993) *Cell 72*, 779−790.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Xu, R. X., Word, J. M., Davis, D. G., Rink, M. J., Willard, D. H., Jr., & Gampe, R. T., Jr. (1995) *Biochemistry 34*, 2107−2121.
Yamanashi, Y., Katiuchi, T., Mizuguchi, J., Yamamoto, T., & Totoshima, K. (1991) *Science 251*, 192−194.
Yamazaki, T., Forman-Kay, J. D., & Kay, L. E. (1993) *J. Am. Chem. Soc. 115*, 11054−11055.
Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., & Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11655−11666.

BI960157X