Orientation of Peptide Fragments from Sos Proteins Bound to the N-Terminal SH3 Domain of Grb2 Determined by NMR Spectroscopy

Michael Wittekind,* Claudio Mapelli, Bennett T. Farmer II, Ki-Ling Suen, Valentina Goldfarb, Jonglin Tsao, Thomas Lavoie, Mariano Barbacid, Chester A. Meyers, and Luciano Mueller

Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543-4000

Received June 20, 1994; Revised Manuscript Received September 14, 1994[⊗]

ABSTRACT: NMR spectroscopy has been used to characterize the protein−protein interactions between the mouse Grb2 (mGrb2) N-terminal SH3 domain complexed with a 15-residue peptide (SPLLPKLPPKTYKRE) corresponding to residues 1264−1278 of the mouse Sos-2 (mSos-2) protein. Intermolecular interactions between the peptide and $^{13}C-^{15}N$-labeled SH3 domain were identified in half-reverse-filtered 2D and 3D NOESY experiments. Assignments for the protons involved in interactions between the peptide and the SH3 domain were confirmed in a series of NOESY experiments using a set of peptides in which different leucine positions were fully deuterated. The peptide ligand-binding site of the mGrb2 N-terminal SH3 domain is defined by the side chains of specific aromatic residues (Tyr_7, Phe_9, Trp_{36}, Tyr_{52}) that form two hydrophobic subsites contacting the side chains of the peptide Leu_4 and Leu_7 residues. An adjacent negatively charged subsite on the SH3 surface is likely to interact with the side chain of a basic residue at peptide position 10 that we show to be involved in binding. The peptide-binding site of the SH3 is characterized by large perturbations of amide chemical shifts when the peptide is added to the SH3 domain. The mGrb2 N-terminal SH3 domain structure in the complex is well-defined (backbone RMSD of 0.56 ± 0.21 calculated over the backbone N, Cα, and C atoms of residues 1−54). The structure of the peptide in the complex is less well-defined but displays a distinct orientation. Relative to the SH3 domain binding site, the polypeptide backbone of the peptide in the complex is positioned in the opposite orientation compared to that recently described for the synthetic peptide (RKLPPRP)−PI3K SH3 domain complex [Yu et al. (1994) *Cell 76*, 933−945]. The mSos peptide−mGrb2 N-terminal SH3 complex has a dissociation constant of 54 μM as determined by isothermal titration calorimetry. Another mSos peptide (VPPPVPPRRR) exhibits tighter binding (3.5 μM dissociation constant) with the N-terminal Grb2 SH3 domain and appears to bind in a similar manner on the basis of NOESY experiments.

Src homology 3 (SH3)[1] domains are small (~60 residue) protein domains that are involved in cytoplasmic protein−protein interactions (Koch et al., 1991). One of the best understood proteins that utilizes SH3 domains to bind to a cellular target protein is the growth factor receptor-bound protein 2 (Grb2) (Lowenstein et al., 1992). Grb2 is a triple domain protein consisting of a single SH2 domain flanked by two SH3 domains. While the single SH2 domain of Grb2 binds to tyrosyl-phosphorylated intracellular domains of stimulated growth factor receptors, the dual SH3 domains of Grb2 have been shown to bind to segments of the proline-rich C-terminal domain of the guanine nucleotide exchange factor, son of sevenless (Sos) (Rogge et al., 1991; Egan et al., 1993; Rozakis-Adcock et al., 1993; Li et al., 1993). A proposed model postulates that, by forming a trimeric complex with the activated receptor and Sos, Grb2 positions the Sos exchange factor within physical proximity of the membrane-bound Ras protein, resulting in Ras activation, a critical step in the Ras signaling pathway (Chardin et al., 1993; Egan & Weinberg, 1993). Given the pivotal role of Grb2 in cellular signal-transduction processes and the existence of a characterized in vivo target protein, we are interested in determining the details of the ligand-binding modes exhibited by the Grb2 SH3 domains.

The crystal (Musacchio et al., 1992; Noble et al., 1993; Eck et al., 1994) and solution (Yu et al., 1992; Booker et al., 1993; Kohda et al., 1993; Koyama et al., 1993a,b; Yang et al., 1994) structures of several SH3 domains have established the SH3 global polypeptide folding pattern. A region of the SH3 domain surface rich in solvent-exposed conserved aromatic residues has been implicated in peptide ligand binding on the basis of perturbations of SH3 domain amide NMR chemical shifts observed when SH3-binding peptide ligands were added (Yu et al., 1992; Booker et al., 1993). The only structure of a SH3 domain−ligand complex reported to date (Yu et al., 1994) shows that the central portion of the synthetic peptide ligand RLP1 (RKLPPRPSK) bound to a SH3 domain of the phosphatidylinositol 3-kinase (PI3K) adopts a left-handed type II polyproline (PPII) helix (Sasisekharan, 1959; Adzhubei & Sternberg, 1993) backbone conformation. This conformation has a repeat of three residues per turn which positions the proline 4 and 7 residues of RLP1 in an optimal orientation for intercalation with the conserved aromatic residues Tyr_{14}, Trp_{55}, and Tyr_{73} in the PI3K SH3 domain ligand-binding site. Although not directly observed, specific electrostatic interactions were inferred on the basis of the results of mutagenesis experiments.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1994.

[1] Abbreviations: SH3 domain, Src homology 3 domain; Grb2, growth factor receptor-bound protein 2; mGrb2, mouse Grb2; Sos, "son-of-sevenless" GTP-exchange factor for Ras; N-SH3, N-terminal SH3 domain of mouse Grb2; C-SH3, C-terminal SH3 domain of mouse Grb2; PPII, left-handed type II polyproline; ITC, isothermal titration calorimetry; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; τ_m, mixing time; VDW, van der Waals; SE, sensitivity enhanced; GE, gradient enhanced; DTT, dithiothreitol; MS, mass spectrometry.

0006-2960/94/0433-13531$04.50/0

Simultaneously with the publication of the SH3 PI3K–RLP1 complex structure, a model was proposed for SH3–ligand recognition on the basis of data from an extensive set of mutagenesis experiments of the C-terminal SH3 domain of Sem-5 (Lim & Richards, 1994). This model also proposed that the peptide ligand binds with the polypeptide backbone, adopting a PPII helix. However, the proposed orientation of the peptide ligand in the SH3 ligand-binding site differs from that of RLP1. The model of Lim and Richards agrees with the suggestion of Schreiber and co-workers (Yu et al., 1994) that SH3 peptide ligands with basic C-terminal residues (termed class II ligands; Chen et al., 1993) might bind in an orientation opposite to that observed for the class I ligand RLP1.

During the course of our investigation directed at probing the binding of the murine Sos2 exchange factor to the Grb2 adaptor protein, we identified a 15-residue peptide (SPLLPKLPPKTYKRE, termed Sos-A) from the mSos-2 sequence (Bowtell et al., 1992) which binds to both the N- and the C-terminal SH3 domains of Grb2. In this paper, we provide experimental evidence that the Sos-A peptide binds to the N-terminal SH3 domain of Grb2 in an orientation opposite to that reported for RLP1 binding to the PI3K SH3 domain. The predominant intermolecular interactions observed involve the methyl groups of the Leu_4 and Leu_7 residues of the peptide with the aromatic side chains of the conserved Tyr_7, Phe_9, Trp_{36}, and Tyr_{52} residues of the N-terminal Grb2 SH3 domain. In addition, we demonstrate that a basic residue at position 10 of the Sos-A peptide is required for optimal binding. We also provide evidence that a peptide corresponding to mSos-1 residues 1152–1161 (VPPPVPPRRR, termed Sos-E) (Li et al., 1993; Chen, 1993), which exhibits 10-fold higher affinity for the Grb2 N-terminal SH3 domain, binds in a similar fashion.

METHODS

Overexpression of the Grb2 SH3 Domains. cDNAs encoding the mouse Grb2 N- and C-terminal SH3 domains (Suen et al., 1993) were cloned into the expression vector pGTX-2T (Pharmacia). ^{13}C–^{15}N-labeled GST–SH3 fusion proteins were expressed in minimal media (Constantine et al., 1993) using Escherichia coli strain BL21 as a host. The fusion proteins were purified from bacterial lysates by binding to a glutathione–agarose column in 50 mM Tris-HCl, pH 8.0, buffer and cleavage with human thrombin (Enzyme Research) on the column. The SH3 domains were further purified by chromatography on a Mono Q FPLC column (Pharmacia). The final samples were prepared by buffer exchange into 50 mM sodium phosphate, pH 6.0, and 10 mM perdeuterated DTT in 90% H_2O/10% D_2O or else 100% D_2O buffered to pD 6.0. Peptides were titrated into the SH3 preparations from concentrated peptide stock solutions in 99.9% D_2O. The final concentration of the Sos-A–N-SH3 complex was approximately 1.5 mM.

Peptide Synthesis. All peptides were synthesized as *N*-acetyl C-terminal amides using the FastMoc Fmoc/HBTU chemistry protocol of an Applied Biosystems Model 431A peptide synthesizer. The peptides were confirmed by amino acid composition and fast atom bombardment mass spectrometry. The position of the deuterated leucine residue in each labeled peptide was confirmed by MS/MS sequencing.

NMR Spectroscopy. All NMR experiments were performed on either a Varian Unity or UnityPlus 600-MHz spectrometer operating at 30.0 °C. Gradient-enhanced (GE) and sensitivity-enhanced (SE) 2D ^1H–^{15}N HSQC (GESE-HSQC) spectra were recorded as previously described (Kay et al., 1992). The other spectra used to obtain ^1H, ^{15}N, and ^{13}C resonance assignments will be described elsewhere (manuscript in preparation). Assignments for the ^1H_N, ^{15}N, $^1H_\alpha$, $^{13}C_\alpha$, $^1H_\beta$, and $^{13}C_\beta$ resonances of the Sos-A-bound N-SH3 domain were obtained with a semiautomated assignment algorithm (Friedrichs et al., 1994). Assignments for the side-chain ^1H and ^{13}C resonances of the Sos-A-bound N-SH3 domain were obtained manually using the data from the GE-HCCH-TOCSY experiment (Kay et al., 1993). Assignments for the amide ^1H_N and ^{15}N resonances of the unbound N-SH3 and the Sos-E-bound N-SH3 domains were obtained by titrating the uncomplexed ^{15}N-labeled SH3 domain with peptide and following the induced chemical shifts in the 2D ^1H–^{15}N HSQC spectra. Both of the peptide–SH3 domain complexes examined in this study are in fast exchange on the chemical shift NMR time scale.

The NOESY spectra used for identification of intramolecular N-SH3 distance restraints were the simultaneous 4D GESE-[^{15}N/^{13}C,^{15}N]-edited NOESY (τ_m = 80 ms) (Farmer & Mueller, 1994) and 3D [^{13}C]-separated GE-NOESY using ^{13}C/^1H coherence transfer echo for optimum H_2O suppression (τ_m = 80 ms) (B. T. Farmer II, unpublished results). To obtain intermolecular distance restraints, GE full- and half-reverse-filtered 2D ^1H–^1H NOE experiments (τ_m = 200 ms) were performed with suppression of both ^{15}N- and ^{13}C-bonded protons. These pulse sequences are similar to those published by Fesik and co-workers (Petros et al., 1993). A 3D version of the experiment was performed where the additional dimension encoded the ^{13}C frequencies of the carbons bound to the ^1H NOE donors. Details of the GE reverse-filtered NOESY and TOCSY pulse sequences will be published elsewhere (manuscript in preparation).

Isothermal Titration Calorimetry. ITC experiments were performed at 30 °C with a Microcal MCS isothermal titration calorimeter (Microcal, Northhampton, MA). The buffer used for all ITC studies was 25 mM sodium phosphate, pH 7.0, and 10 mM DTT with 0.01% sodium azide. Solutions of the SH3 domain (0.15 mM) in the sample cell were titrated with peptide solutions (peptide concentrations were in the range of 2.0–6.0 mM for the various peptide solutions). Experiments were performed as described (Wiseman et al., 1989). The data were fit using the ORGIN software supplied by the manufacturer. For all peptides, best fits to the data were obtained when the number of sites was between 0.8 and 1.0.

Structure Calculations. Intensities of assigned cross-peaks in the 3D ^{13}C-edited NOESY spectra and the 4D [^{15}N–^{15}N]- and [^{15}N–^{13}C]-edited NOESY spectra were quantitated and converted to distance restraints. Four distance restraint categories were defined, having upper bounds equal to very strong (2.5 Å), strong (3.0 Å), medium (4.0 Å), or weak (5.0 Å). The lower bound in all cases was 1.8 Å.

Amide protons protected from exchange with the solvent were identified in a qualitative manner by the lack of exchange cross-peak intensity at the H_2O ^1H resonance frequency in the F_1 dimension of a 3D gradient-enhanced ^{15}N-edited NOESY spectrum (τ_m = 80 ms). Restraints involving these amides, when hydrogen-bond acceptors could be corroborated by NOESY evidence, were included (1.8–2.2 Å for amide proton to carbonyl oxygen and 2.8–3.2 Å

Table 1: Binding Constants for Sos Peptide−Grb2 N-SH3 Interactions (ITC)[a]

Peptide sequence	Peptide name	mSOS sequence#	K_d	ΔG	ΔH	ΔS
S P L L P K L P P K T Y K R E	SOS-A	(2) 1264-1278	54	-5.9	-9.7	-12.5
S P L L P K L P P R T Y K R E	SOS-A[R_{10}]		48	-6.0	-9.1	-10.2
S P L L P K L P P A T Y K R E	SOS-A[A_{10}]		455	-4.6	-5.5	-2.9
V P P P V P P R R R	SOS-E	(1) 1152-1161	3.5	-7.6	-13.0	-17.9
I P P P L P P R K K	SOS-F	(2) 1110-1119	5.3	-7.3	-12.0	-15.5
P1 P2 P3						
ϕ/p p p ϕ p p r/k	*Class II peptide consensus sequence*					

[a] Peptides used in this study and thermodynamic data obtained using isothermal titration calorimetry. All ITC data were collected at 30 °C. The values are the result of a single titration experiment for each peptide−N-SH3 complex. The units for the reported values are as follows: K_d, μM; ΔG, kcal mol^{-1}; ΔH, kcal mol^{-1}; ΔS, cal mol^{-1} K^{-1}. The P1, P2, and P3 positions (see Discussion) and a class II SH3-binding peptide consensus sequence are designated below the table (p, proline; r, arginine; k, lysine; ϕ, hydrophobic residue).

for amide nitrogen to carbonyl oxygen). In addition, stereo assignments of $^1H_\beta$ methylenes were obtained for 15 residues, and χ_1 angle restraints corresponding to one of the preferred staggered conformations (±30°) were included in the calculations.

An initial crude N-SH3 model was built by replacing the side chains of the spectrin SH3 domain X-ray model (Musacchio et al., 1992; Brookhaven prerelease database entry 1SHG) with those of the mGrb2 N-terminal SH3 domain (residues 1−59; Suen et al., 1993). Additional residues corresponding to the N- and C-terminal amino acids present from the coding portion of the pGTX-2T cloning vector were added to the model in an extended conformation. The full sequence of the N-SH3 model (corresponding to the protein sample used in the study) was GSRRASVGS-MEAIAKYDFKATADDELSFKRGDILKVL-NEECDQNWYKAELNGKDGFIPKNYIEMK-PHPEFIVTD. The numbering of the 74-residue model starts at −8, with residue 1 corresponding to the first residue (Met_1) of the mGrb2 sequence (underlined). All of the structure manipulations were carried out using the INSIGHTII program (Biosym Technologies, San Diego, CA).

Calculations of the structure of the complex of the N-terminal Grb2 SH3 domain and the Sos-A peptide were begun using the crude SH3 domain model (described above) and a 15-residue Sos-A peptide model (in different conformations and orientations; see Results) placed in proximity of the SH3 peptide ligand-binding site as starting structures. Simulated annealing refinement was carried out with intra-N-SH3 restraints [824 NOE distance, 48 H-bond (corresponding to 24 H-bonds), and 15 χ_1 angle], 11 intermolecular distance restraints, and 4 intra-Sos-A distance restraints using the X-PLOR program (Brünger, 1992). A different numerical seed value was used for each structure to give a different set of initial velocities to the atoms to begin the dynamics phase of the structure calculation. The refinement protocol began with 2000 steps of Powell minimization followed by 10 ps of dynamics at 800 K with the multiplier for the VDW radius set to 1.0. Another 10 ps of dynamics was performed in which the VDW force constant was gradually increased from the initial value of 0.002 to 0.1 kcal mol^{-1} Å^{-4}, and the asymptote of the soft-square well potential was tilted from 0.1 to 1.0. The system was then cooled from 800 to 100 K over 2.8 ps, during which time the VDW force constant and the multiplier for the VDW radius were set to constant values of 4.0 kcal mol^{-1} Å^{-4} and 0.8, respectively. The procedure was completed with 1000 steps of Powell minimization.

RESULTS

In an effort to identify cellular proteins that bind to the mouse Grb2 protein (mGrb2), plaques from a λgt11 mouse cDNA library were probed with a ^{32}P-labeled GST−mGrb2 fusion protein. A high-affinity mGrb2 binding lacZ fusion protein was identified, and sequencing of the λgt11 phage DNA insert identified the mouse cDNA as that encoding the C-terminal 34 residues of the mouse Sos-2 protein (mSos-2 residues 1264−1297; Bowtell et al., 1992) (K.-L. Suen, unpublished results). This 34-residue peptide (SPLLPKLP-PKTYKRELSHPPLYRLPLLENAETPQ, peptide Sos-C) was synthesized, as were two other peptides corresponding to the sequences of each half of the 34-mer: Sos-A (see Table 1) and Sos-B (SHPPLYRLPLLENAETPQ). Each of these three peptides were titrated into 200 μM solutions of ^{15}N-labeled SH3 domains of Grb2, and the effects on the ^{1}H−^{15}N HSQC spectra of the SH3 domains were monitored. Peptides Sos-A and Sos-C induced large amide chemical shifts for a number of residues in both the N-terminal and C-terminal Grb2 SH3 domains (N-SH3 and C-SH3, respectively), with all peptide−SH3 combinations exhibiting fast exchange on the chemical shift NMR time scale. The magnitude of the N-SH3 amide ^{1}H and ^{15}N shifts induced by the addition of Sos-A is presented in the histogram shown in Figure 1.

NMR studies were initiated with a sample of the Sos-A−N-SH3 complex. For experiments in which the N-SH3 domain resonances were observed, the NMR sample contained a 1.2:1.0 ratio of peptide to N-SH3. This ratio was slightly higher than that needed to fully titrate the N-SH3 resonances to their bound positions (as monitored by ^{1}H−^{15}N HSQC spectra). For experiments where SOS-A resonances were assigned, a 1:1 ratio was used. Since the Sos-A−N-SH3 complex was shown to have a K_D of approximately 50 μM (see ITC results, Table 1), at 1.5 mM the percentage of free peptide and protein is less than 5%.

Complete ^{1}H, ^{13}C, and ^{15}N assignments were obtained for the N-SH3 domain in the complex. These assignments were

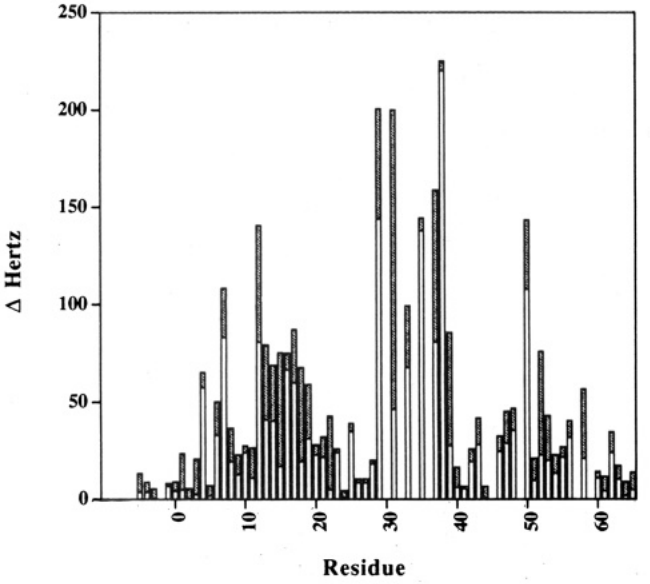

FIGURE 1: Histogram of ^1H and ^{15}N amide N-SH3 chemical shift perturbations caused by Sos-A peptide binding as determined by analysis of 2D $^1H-^{15}N$ HSQC spectra: filled bars, ^1H perturbations (Hz); open bars, ^{15}N perturbations (Hz).

used to identify intra-N-SH3 NOE peaks in the 3D and 4D NOESY spectra of the complex. In addition, stereo assignments for $^1H_\beta$ protons and H-bond restraints were obtained.

Intermolecular contacts between the Sos-A peptide and N-SH3 were identified by GE half-reverse-filtered 2D NOE experiments. Among the NOESY cross-peaks arising from intermolecular interactions, a number of NOEs were observed correlating resonances from the methyl region of the Sos-A peptide with aromatic side-chain resonances of the N-SH3 domain, as shown in Figure 2A. The identities of the N-SH3 ^1H resonances participating in these contacts were unambiguously confirmed by performing a 3D version of the half-reverse-filtered NOESY experiment in which the chemical shifts of the ^{13}C nuclei bonded to the ^1H NOE donors were recorded in the additional frequency dimension (data not shown).

Because the methyl region of the ^1H spectrum of the bound peptide suffered from overlap that complicated the assignment of these peptide resonances, we synthesized three versions of the Sos-A peptide, each with a different leucine position fully deuterated. By monitoring the disappearance of ^1H signals in the half-reverse-filtered 2D NOESY spectra of the various complexes, NOEs involving the resonances of the three leucine residues could be readily assigned. Intermolecular NOE cross-peaks between the methyl groups of Leu_4 and Leu_7 and aromatic protons of N-SH3 were observed while NOE cross-peaks involving the peptide Leu_3 methyl protons were not.

Three different calculations of 40 structures were performed, each differing in the starting conformation and orientation (relative to the starting SH3 domain model) of the peptide: all PPII helices with 0° orientation, all α-helical with 180° orientation, and all fully extended with 180° orientation. The 0° orientation roughly pairs the peptide Leu_4 methyls with the N-SH3 Tyr_7 side chain and the peptide Leu_7 methyls with the N-SH3 Trp_{36} side chain. The 180° orientation reverses this pairing. All of the resulting structures that exhibit no distance restraint violations greater than 0.5 Å adopt the 0° orientation of the peptide relative to

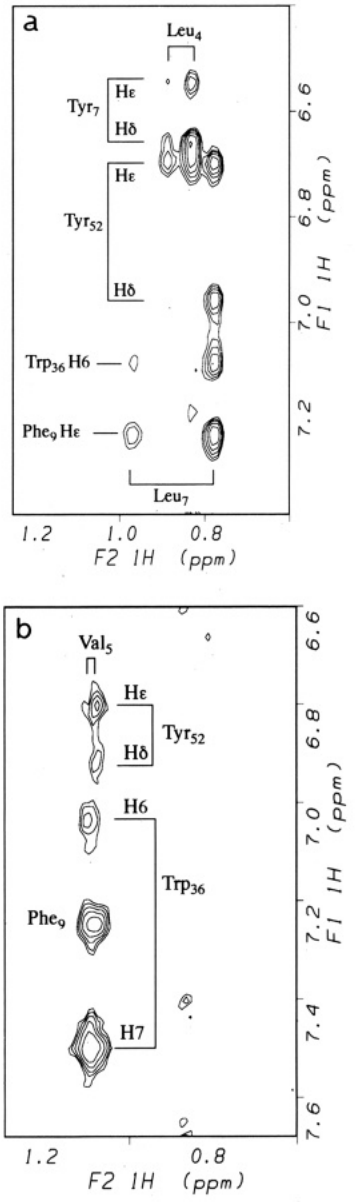

FIGURE 2: Two-dimensional GE half-reverse-filtered ^1H NOESY spectra (τ_m = 200 ms) of peptide−N-SH3 complexes: (A) Sos-A peptide−N-SH3 domain complex; (B) Sos-E peptide−N-SH3 domain complex. The spectra show NOE cross-peaks between aromatic protons of N-SH3 in F_1 and the methyl protons of the peptides in F_2. Assignments of the leucine methyl groups of the Sos-A peptide were confirmed by analysis of spectra from partially deuterated peptide−N-SH3 complexes (see text).

the N-SH3. This proves that the resulting structures are insensitive to the peptide's orientation and conformation in the starting structures and that the constraints alone are driving the peptide to its final orientation.

The final ensemble of 29 structures, shown in Figure 3, is comprised equally of structures originating from each of the three calculations described above. They are characterized by an average RMSD of 0.56 ± 0.21 Å for the backbone atoms (N, Cα, and CO) of N-SH3 residues 1−54 relative to those of the average structure. When the relatively disordered residues 27−35 are excluded from the calculation, this value improves to 0.24 ± 0.10 Å, indicating a very well-defined N-SH3 backbone conformation. When calculated over residues 1−54, the crude model used to begin the structure calculations has a backbone RMSD of 3.1 Å relative to the average structure, demonstrating that sufficient move-

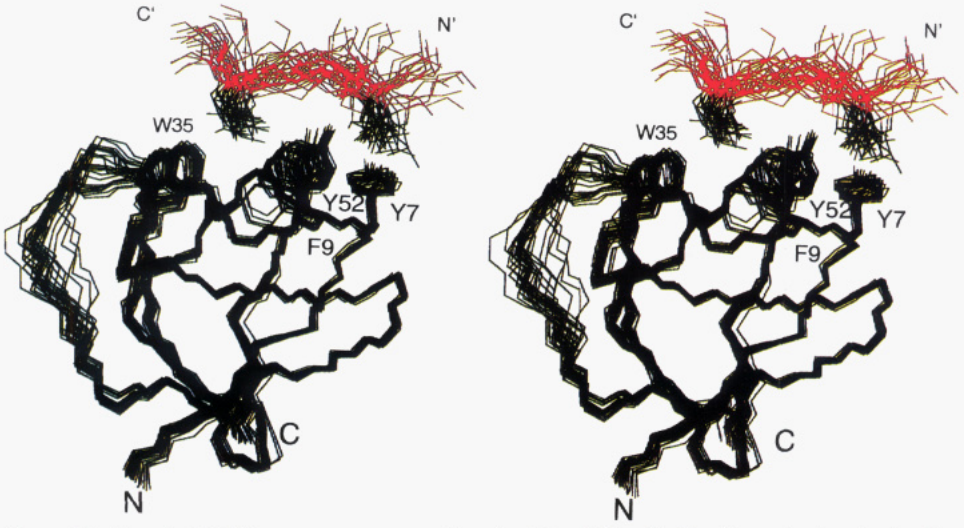

FIGURE 3: Superposition of the Sos-A−N-SH3 structures over residues 1−26 and 36−54. The figure depicts the backbone atoms of N-SH3 residues to −1 to 57 (black) and Sos-A peptide residues 3−8 (red) of the 14 lowest energy structures. Selected side chains of N-SH3 (labeled) and of Sos-A (residues Leu_4 and Leu_7) are shown in black. N, N′, C, and C′ denote the amino and carboxy termini of the displayed portions of N-SH3 and the Sos-A peptide, respectively. The unrestrained disordered backbone atoms at the N- and C-termini for both N-SH3 and the Sos-A peptide are not shown. Note: W35 should be W36 on the figure.

ment of atoms occurred during the simulated annealing procedure to provide for adequate sampling of conformational space. The set of 29 N-SH3 structures is characterized by good geometry and no distance restraint violations greater than 0.5 Å. There were no NOE distance restraints included for N-SH3 residues −8 to 0 and 55−65 and for Sos-A residues 1−3 and 9−15, and consequently their conformations are randomly distributed among the structures. Further evidence for adequate sampling is provided by the high backbone RMSD values of residues at the ends of the molecules: N-SH3 Gly_{-8}, 6.47; N-SH3 Asp_{64}, 7.91; Sos-A Ser_1, 6.32; Sos-A Glu_{15}, 14.76.

Other than some variation in the χ_1 angle of Phe_9, the N-SH3 side chains that have NOEs to the peptide (Tyr_7, Phe_9, Trp_{36}, and Tyr_{52}) are well-defined (Figure 3). The Tyr_7 and Tyr_{52} aromatic side chains form a site for the peptide Leu_4 methyl groups while the Phe_9, Trp_{36}, and Tyr_{52} side chains form a site for the peptide Leu_7 methyl groups. These interactions restrict the orientation of the peptide backbone relative to the SH3.

The thermodynamic parameters characterizing the binding of a series of Sos peptides to the N-SH3 domain of Grb2 were measured by ITC and are listed in Table 1. The strongest peptide−N-SH3 interaction is exhibited by the peptide Sos-E (VPPPVPPRRR, residues 1152−1161 of mouse Sos-1) complexed with N-SH3 (Li et al., 1993; Chen et al., 1993). The Sos-A peptide was found to bind more strongly to C-SH3 than it does to N-SH3 (data not shown).

As a test for the importance of electrostatic contributions to N-SH3−Sos-A peptide complex formation, Sos-A peptide analogs in which the lysine residue at position 10 was replaced by other amino acids were tested for their ability to bind to N-SH3 (Table 1). Binding was retained when arginine was used as a replacement for lysine (peptide Sos-A[R_{10}]), but it was reduced 10-fold when position 10 was substituted with alanine (peptide Sos-A[A_{10}]). This suggests a requirement for a basic amino acid residue at position 10 in the Sos-A peptide in order to retain full binding potential.

In order to test whether the relatively tight-binding Sos-E−N-SH3 complex is characterized by the same types of peptide−protein contacts observed for the Sos-A−N-SH3 complex, we collected GE full- and half-reverse-filtered 2D NOESY, 3D NOESY, and 2D TOCSY data for the Sos-E−N-SH3 complex, and resonances from a number of key residues of the bound Sos-E peptide and N-SH3 domain were assigned. Unambiguous NOEs were observed between the Sos-E Val_5 methyl groups and the aromatic side-chain protons of N-SH3 Phe_9, Trp_{36}, and Tyr_{52} (Figure 2B). From these results, it is clear that the methyl groups of Sos-E Val_5 occupy the same binding site in the Sos-E−N-SH3 complex as do the Sos-A Leu_7 methyl groups in the Sos-A−N-SH3 complex. In contrast, the Sos-E Val_1 methyl protons do not appear to strongly interact with aromatic protons of the SH3 domain. Although NOE cross-peaks between aromatic protons of N-SH3 Tyr_7 and peptide protons resonating near 2 ppm were observed (data not shown), further assignment work is in progress to confirm whether these signals are due to the Sos-E Pro_2 side-chain protons.

DISCUSSION

The structure of the peptide-bound N-SH3 domain of mGrb2 exhibits the same overall global folding pattern seen in the structures of other SH3 domains. This is not unexpected, given the overall high degree of homology shared by many SH3 domains (Koyama et al., 1993). The 29 final structures are well ordered over residues 1−26 and 36−54 while residues 27−35 show more disorder. This portion of the SH3 sequence shows the most variation among the family of SH3 sequences (Koyama et al., 1993) and is a region of the SH3 structure that appears to be inherently less well-defined than the rest of the domain, as has been observed with other SH3 solution structures (Yu et al., 1992; Booker et al., 1993; Kohda et al., 1993; Koyama et al., 1993a,b; Yang et al., 1994).

The solvent-exposed surface of the N-SH3 is predominantly comprised of polar residues, except for a portion of the peptide ligand-binding site. The surface of this part of the binding site is primarily composed of the side chains of aromatic residues, imparting an overall hydrophobic character to this portion of the domain surface. The hydrophobic part

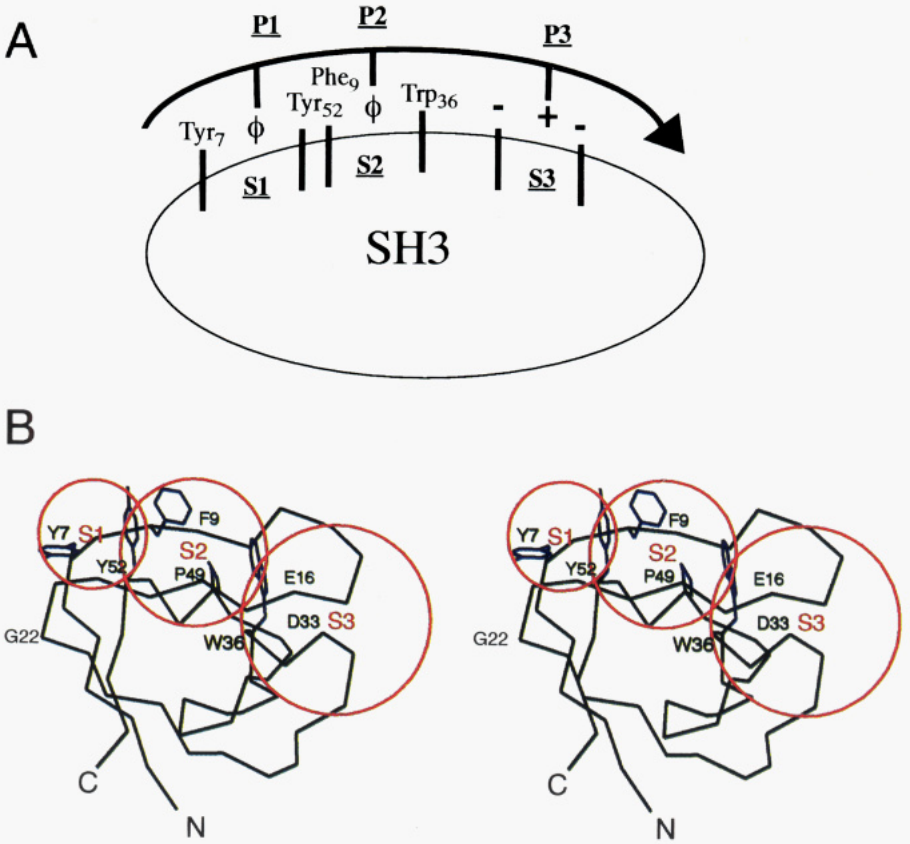

FIGURE 4: Diagram of the key interactions observed between the class II peptides and N-SH3. The hydrophobic subsites of the peptide ligand-binding site of the SH3 domain are labeled S1 and S2 while the negatively charged subsite adjacent to the S2 subsite is labeled S3. (A) The hydrophobic residues at peptide positions P1 and P2 contact the S1 and S2 ligand-binding subsites. The positively charged peptide residue at position P3 is likely to interact with the negatively charged S3 subsite on the SH3 domain (see Discussion). The symbol ϕ indicates a hydrophobic group while the arrow indicates the direction of the peptide backbone. B) Stereoview of the Cα trace of N-SH3 residues 1−54 (black) with the three peptide-binding subsites circled (red). The heavy atoms of the side chains involved in forming the S1 and S2 subsites are shown in blue. N and C denote the amino and carboxy termini of the of N-SH3.

of the peptide ligand-binding site can be divided into two smaller subsites, as outlined in Figure 4. These two indentations in the SH3 domain surface accept hydrophobic groups on the surface of the peptide structure in the peptide−protein complex. The main hydrophobic binding subsite (S2) is defined as the cleft formed by the aromatic side chains of N-SH3 residues Tyr_{52}, Phe_9, and Trp_{36}, with the side chain of Pro_{49} forming the floor of the binding cavity. The cavity formed at S2 is the deeper of the two hydrophobic subsites and accepts the methyl groups of the peptide Leu_7 residue in the Sos-A−N-SH3 complex and of the peptide Val_5 residue in the Sos-E−N-SH3 complex. We call the position occupied by the Leu_7 residue in the Sos-A peptide sequence the P2 position of class II peptides (Table 1). We note that, in the RLP1−PI3K SH3 complex (Yu et al., 1994), the S2 subsite is occupied by the methyl groups of the Leu_3 residue of the class I peptide, RLP1.

A second hydrophobic binding subsite (S1) is formed by the cavity made up of the aromatic rings of Tyr_{52} and Tyr_7, the planes of which are positioned nearly 90° relative to one another. The S1 subsite is shallower and interacts with the Leu_4 methyl groups in the Sos-A−N-SH3 complex. This residue in class II peptides is designated as the P1 position. As described below, the Sos-A Leu_4 methyl groups may be too bulky to optimally occupy the relatively shallow S1 subsite, and a smaller hydrophobic group at the P1 position may provide a better steric fit.

Immediately adjacent to the S2 subsite of N-SH3 is a large cavity in the domain surface, termed the S3 subsite (Figure 4), formed by the two N-SH3 domain loops that encompass the highly negatively charged D_{14}-D_{15}-E_{16} and E_{30}-E_{31}-C_{32}-D_{33} sequences. Residues of these two loops exhibit extensive ^{15}N and ^{1}H amide chemical shift changes when the peptide is added to the SH3 domain (Figure 1 and see Figure 5B,C). Glu_{16} is a highly conserved residue in SH3 domains, and the cognate negatively charged residues in the C-terminal SH3 domain of Sem-5 (Lim & Richards, 1994) and the PI3K SH3 domain (Yu et al., 1994) have been shown to be involved in peptide binding. In the Sos-A−N-SH3 complex, because of the P1/S1 and P2/S2 pairings, the S3 subsite is likely to be in proximity to the side chain of the lysine residue at position 10 of the Sos-A peptide. When Sos-A residue 10 is substituted by alanine, the ability to bind to N-SH3 is greatly diminished whereas binding is retained when arginine is used as the replacement (Table 1). This important basic residue position in class II SH3-binding peptides is termed P3 and is probably involved in electrostatic interactions with the negatively charged S3 subsite on the SH3 domain surface. It is yet to be established whether the Sos-A peptide Lys_{10} side chain makes a distinct salt bridge with a particular negatively charged residue of the SH3 domain or simply accelerates the rate of complex formation by nonspecific electrostatic interaction with the negatively charged region of the protein. The functional interchangeability of lysine and arginine at the Sos-A P3 position (Table 1) and at the positions adjacent to the P3 position in the Sos-E and Sos-F peptides (see below) suggests that exact side-chain topology of the basic residues may be unimportant and that the positive

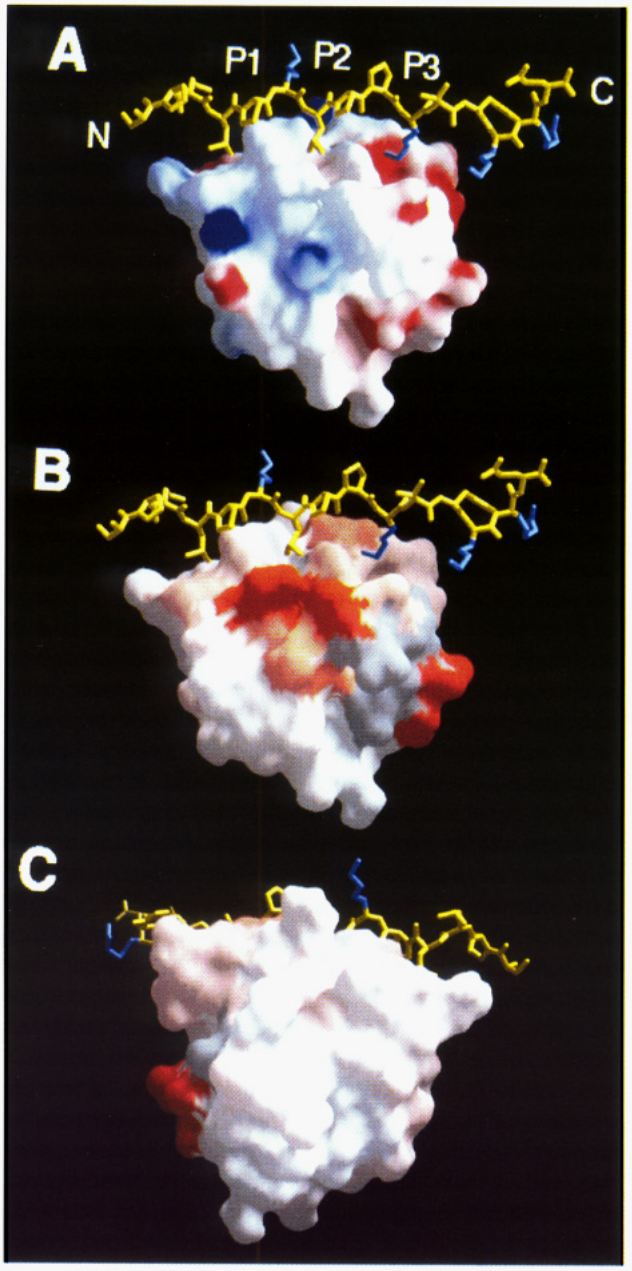

FIGURE 5: Hypothetical model of the N-SH3−Sos-A complex generated within X-PLOR by augmenting the experimentally derived restraints with angle restraints enforcing an idealized PPII helix backbone conformation for peptide residues 2−11 ($\phi = -75°$, $\psi = 145°$, $\pm 20°$) and distance restraints corresponding to a hydrogen bond between the Sos-A Lys_{10} and N-SH3 Glu_{16} side chains. In this model, there are no distance violations greater than 0.5 Å, and all angle restraints are satisfied. The N-SH3 domain is shown in color-coded surface representations. The Sos-A heavy atoms are depicted in the yellow rod representation with the side chains of Lys_6, Lys_{10}, Lys_{13}, and Arg_{14} colored blue. The images were generated with the GRASP program (Nicholls et al., 1991). (A) The electropositive (blue) and electronegative (red) portions of the charged surface of the N-SH3 domain are depicted, with the maximum values equal to ±10 kT. The P1, P2, and P3 designations correspond to the Sos-A peptide Leu_4, Leu_7, and Lys_{10} residues (see text). (B) The magnitudes of the summed ^1H and ^{15}N chemical shift perturbations for backbone and side-chain amide resonances of the N-SH3 residues are represented by the color intensity of the N-SH3 surface (deep orange, large perturbations; white, no perturbations). Residues for which no data could be obtained are colored gray. (C) Same as (B), but with the model rotated by 180°.

charge(s) may be the critical determinant for binding.

On the basis of the knowledge that both Leu_7 of Sos-A and Val_5 of Sos-E make contact with the S2 subsite of N-SH3, an alignment of the Sos peptide sequences can be abstracted to the class II SH3-binding peptide consensus sequence shown at the bottom of Table 1. Hydrophobic residues at the peptide P1 and P2 positions contact the S1 and S2 subsites of the SH3 domain while a basic residue is needed at the P3 position in order to interact with the acidic S3 subsite. On the basis of this limited set of peptide sequences, prolines appear to be favored at the residue positions separating the triple-spaced P1, P2, and P3 positions. The assignment of additional, low-intensity, intermolecular NOEs, now in progress for the N-SH3−Sos-E complex, may support a role for prolines at these peptide positions in contacting the SH3 domain.

The interdigitization of the methyl groups of the peptide residues Leu_4 and Leu_7 (peptide positions P1 and P2) with the well-defined aromatic side chains of N-SH3 residues that define the hydrophobic subsites S1 and S2 (Figures 3 and 4) provides an unambiguous peptide backbone orientation. This orientation is opposite to that observed for the RLP1 peptide, a class I ligand, bound to the PI3K domain (Yu et al., 1994). Because both the PI3K SH3 and N-SH3 domains share the same overall ligand-binding site topology (a cation-binding subsite adjacent to a hydrophobic subsite comprised of two subsites) on the common SH3 polypeptide fold, the difference in their binding modes can be largely attributed to the nature of the peptide ligands. For the class I and class II SH3-binding peptides, the S3 subsite of the SH3 appears to make a key interaction with a basic peptide residue: for the class I peptides, this residue is N-terminal to the hydrophobic contact residues, while for the class II peptides, the important basic residue(s) is (are) at the C-terminal end. The relative positioning of the residues responsible for the hydrophobic and electrostatic interactions along the peptide sequence appears to determine their SH3-bound orientations.

The lack of chemical shift dispersion in the ^1H spectrum of the N-SH3-bound Sos-A peptide precluded the complete assignment of all the Sos-A peptide resonances. As a result, while the orientation of the peptide backbone and the positioning of the Leu_4 and Leu_7 methyl groups are clearly established, the overall resolution of the peptide structure is not high enough to determine whether the backbone conformation of the SH3-bound peptide adopts a PPII left-handed helix, as has been observed in the RLP1−PI3K SH3 structure (Yu et al., 1994). The presentation of the triply spaced P1, P2, and P3 peptide groups responsible for specific interactions with the SH3 domain is consistent with the PPII helix backbone conformation of the bound peptide, due to the PPII helix periodicity of three residues per turn (Sasisekharan, 1959; Adzhubei & Sternberg, 1993). In NOESY spectra of the Sos-E−N-SH3 complex, sequential connectivities for the residues making up the central portion of the bound peptide support a PPII helix conformation (data not shown).

A hypothetical model of the complex was generated in which the Sos-A peptide backbone conformation was restricted to be close to an idealized PPII helix and the side-chain position of Sos-A residue Lys_{10} was restricted to be near the S3 subsite. The model is shown in Figure 5. The positioning of the triply spaced P1, P2, and P3 residues on one side of the peptide structure is evident in the model, due to the periodicity of three residues per turn dictated by the PPII helix conformation. The surface representation of the N-SH3 portion of the model gives an indication of the

cavities formed by the S1, S2, and S3 subsites on the N-SH3 surface. The S1 and S2 subsites share the aromatic side chain of N-SH3 Tyr_{52}, and its hydroxyl proton may form a hydrogen bond with the backbone carbonyl oxygen of the Pro_5 residue of the Sos-A peptide (W. Lau, personal communication).

In the model, note that once the P1/S1 and P2/S2 contacts are established, the P3 side chain can be easily positioned very close to the S3 subsite. The extreme electronegative potential of the surface of the S3 subsite near the P3 side chain is apparent (Figure 5A). When the amide chemical shift perturbations (Figure 1) are mapped to the N-SH3 surface, the highest levels of perturbations are localized to the surface area near the S3 subsite (Figure 5B,C). Because of the potential for favorable electrostatic interactions and the pronounced clustering of the amide chemical shift perturbations, we propose that the P3−S3 subsite interaction plays a key role in contributing to the binding energy and specificity of the class II peptide−N-SH3 complexes. In addition, because the SH3 folding topology appears to be relatively conserved among members of the SH3 family and the aromatic residues making up the S1 and S2 subsites are common to many SH3 domain sequences, the S3 subsite is probably important in defining the ligand-binding specificities of many SH3 domains.

Several features of the class II peptide ligands Sos-E and Sos-F suggest reasons why these peptides exhibit tighter binding to the Grb2 N-SH3 domain than does the Sos-A peptide (Table 1). First, the Sos-E and Sos-F peptides have two additional basic residues directly adjacent to the P3-position residue, and these positively charged residues are likely to provide additional electrostatic interactions with the negatively charged loops (D_{14}-D_{15}-E_{16} and E_{30}-E_{31}-C_{32}-D_{33}) of the S3 subsite of the Grb2 N-SH3 domain. Second, PPII helix formation should be favored by the high proline content of the Sos-E/F peptides, perhaps facilitating presentation of the P1, P2, and P3 residue side chains to the N-SH3 domain. Finally, since the binding cavity of the N-SH3 S1 subsite is relatively shallow, the less bulky Pro_2 residue of the Sos-E/F peptides may provide a better steric fit at the P1 position than the Leu_4 residue of Sos-A.

The calorimetry results provide information about the relative enthalpic and entropic contributions to the free energies of binding. The magnitudes of ΔH of binding for the Sos-E and Sos-F peptides are large enough to offset the losses in entropy these peptides accrue upon complexation with the N-SH3 domain. One interpretation of this result is that these two peptides make more energetically favorable contacts with N-SH3 than does the Sos-A peptide, and consequently their conformations become more ordered during the binding process.

The C-terminal domain of the Sos protein is established as one of the primary *in vivo* targets of the Grb2 SH3 domains (Rogge et al., 1991; Egan et al., 1993; Rozakis-Adcock et al., 1993; Li et al., 1993). The Sos-E peptide has been shown to interfere with the Grb2−Sos C-terminal domain interaction *in vitro* (Li et al., 1993), suggesting that the peptide is binding to the same ligand-binding site on the SH3 that the intact Sos C-terminal domain does. On the basis of the magnitude of the observed binding constants, it is possible that the Sos-E (F) sequence binds to N-SH3 *in vivo* while the Sos-A peptide sequence binds to C-SH3. While the dissociation constants of the individual Sos peptide−SH3 interactions are in the low micromolar range (Chen et al., 1993; this study), the combined effect of the two interactions would undoubtedly increase the avidity of the intact Grb2 protein's association with Sos. Segments of the Sos protein that comprise these two peptide sequences may be positioned within the C-terminal Sos structure such that their interactions with the two SH3 domains of Grb2 are optimal for binding.

The existence of at least two peptide ligand-binding modes for SH3 domains, differing in their relative orientations, raises the question of whether both of these binding modes are utilized by SH3 domains *in vivo*. In principle, a protein ligand sequence adopting a PPII helix could conceivably bind to a SH3 domain in either orientation. This is because the almost equidistant spacing between the S1, S2, and S3 subsites on the SH3 surface (Lim & Richards, 1994) causes the formation of a complementary set of linearly arranged "holes" for the "pegs" provided by the side chains of the triply spaced contact residues along the presentation side of the peptide ligand. As previously mentioned, since most SH3 domains share high sequence homology for the residues making up the S1 and S2 subsites, the orientation used by a particular ligand when binding to a SH3 domain is likely to be determined by the types of interactions it can make with the S3 subsite or some other portion of the SH3 surface flanking the S1−S2 subsites.

A second question arises as to whether additional binding modes are possible. The two established binding modes represent the most straightforward ways of achieving the necessary contacts between the relatively extended peptide structures and the linearly arranged S1, S2, and S3 subsites on the SH3 domain surfaces. However, as the family of known SH3 protein sequences is expanded, SH3 domains that share the same conserved residues for the domain core have been discovered that differ in their solvent-exposed surface residues. For example, the recently described structure of the GAP SH3 domain has leucine substitutions at the key SH3 ligand-binding site residues corresponding to aromatic residues Tyr_7 and Tyr_{52} of the Grb2 N-SH3 domain (Yang et al., 1994). Substitutions of this type must alter the nature of the SH3 S1 and S2 subsites, complicating model-building efforts based on the structures of the peptide−SH3 domain complexes described so far. It will be of interest to determine the structures of these nonstandard SH3 domains complexed to the ligands they bind in order to appreciate the full range of peptide ligand-binding modes exhibited by SH3 domains.

In conclusion, we have used NMR and ITC to characterize the binding of the N-terminal SH3 domain of Grb2 to a peptide derived from the C-terminal region of the Ras-activating Sos protein, a known physiologically relevant binding partner for Grb2 *in vivo*. The structures of the N-SH3 portion of the complex are of high resolution, providing information about the topology of the peptide binding site. The two hydrophobic subsites of the ligand-binding site of the N-SH3 domain are comprised of side chains of conserved aromatic residues that make contact with the methyl groups of the peptide Leu_4 and Leu_7 residues, unambiguously establishing that the Sos-A peptide binds to the N-terminal SH3 domain of Grb2 in an orientation opposite to that reported for a synthetic peptide bound to the PI3K SH3 domain. A third negatively charged subsite on the N-SH3 is likely to interact with a positively charged

residue at position 10 of the peptide, shown to be important for binding. The triple-spaced periodicity of the peptide residues interacting with the N-SH3 domain suggests that the peptide backbone may adopt a PPII helix conformation in the Sos-A−N-SH3 complex. The tighter binding Sos-E peptide (VPPPVPPRRR) appears to bind to the N-SH3 in a similar manner.

NOTE ADDED IN PROOF

X-ray structures of the 3BP2·Fyn and 3BP1·Abl (peptide·SH3) complexes have recently been published (Musacchio et al., 1994). Both structures exhibit the same peptide-binding orientation described for the RPL1·PI3K complex (Yu et al., 1994), which is opposite to the one described in this study.

ACKNOWLEDGMENT

The authors thank Paul Reiss, Clifford Klimas, Hann-Guang Chao, and Mike Bernatowicz for peptide microsequencing, amino acid analysis, and peptide synthesis at the early stages of this work. We thank Stan Krystek for help with the figures.

SUPPLEMENTARY MATERIAL AVAILABLE

Two tables giving intermolecular NOE distance restraints and structural statistics for the 29 final dynamical simulated annealing structures (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Adzhubei, A. A., & Sternberg, M. J. E. (1993) *J. Mol. Biol. 229*, 472−493.

Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., & Campbell, I. D. (1993) *Cell 73*, 813−822.

Bowtell, D., Fu, P., Simon, M., & Senior, P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6511−6515.

Brünger, A. T. (1992) *X-PLOR Version 3.1 User Manual*, Yale University, New Haven, CT.

Chardin, P., Camonis, J. H., Gale, N. W., Vanaelst, L., Schlessinger, J., Wigler, M. H., & Barsagi, D. (1993) *Science 260*, 1338−1343.

Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A., & Schreiber, S. L. (1993) *J. Am. Chem. Soc. 115*, 12591−12592.

Constantine, K. L., Goldfarb, V., Wittekind, M., Friedrichs, M. S., Anthony, J., Ng, S.-C., & Mueller, L. (1993) *J. Biomol. NMR 3*, 41−54.

Eck, M. J., Atwell, S. K., Shoelson, S. E., & Harrison, S. C. (1994) *Nature 386*, 764−769.

Egan, S. E., & Weinberg, R. A. (1993) *Nature 365*, 781−783.

Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., & Weinberg, R. A. (1993) *Nature 363*, 45−51.

Farmer, B. T., II, & Mueller, L. (1994) *J. Biomol. NMR 4*, 673−687.

Friedrichs, M. S., Mueller, L., & Wittekind, M. (1994) *J. Biomol. NMR 4*, 703−726.

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. 101B*, 333−337.

Koch, C. A., Anderson, D., Moran., M. F., Ellis, C., & Pawson, T. (1991) *Science 252*, 668−674.

Kohda, D., Hanataka, H., Odaka, M., Mandiyan, V., Ullrich, A., Sclessinger, J., & Inagaki, F. (1993) *Cell 72*, 953−960.

Koyama, S., Yu, H., Dalgarno, D. C., Shin, T. B., Zydowsky, L. D., & Schreiber, S. L. (1993a) *Cell 72*, 945−952.

Koyama, S., Yu, H., Dalgarno, D. C., Shin, T. B., Zydowsky, L. D., & Schreiber, S. L. (1993b) *FEBS Lett. 324*, 93−98.

Li, N., Batzer, A., Daly, R., Yajnik, V., Skolnik, E., Chardin, P., Bar-Sagi, D., Margolis, B., & Schlessinger, J. (1993) *Nature 363*, 85−87.

Lim, W., & Richards, F. M. (1994) *Struct. Biol. 1*, 221−225.

Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., & Schlessinger, J. (1992) *Cell 70*, 431−442.

Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., & Saraste, M. (1992) *Nature 359*, 851−855.

Musacchio, A., Saraste, M., & Wilmanns, M. (1994) *Nature Struct. Biol. 1*, 546−551.

Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 282−293.

Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A., & Wierenga, R. K. (1993) *EMBO J. 12*, 2617−2624.

Rogge, R. D., Karlovich, C. A., & Banerjee, U. (1991) *Cell 64*, 39−48.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., & Bowtell, D. (1993) *Nature 363*, 83−85.

Sasisekharan, V. (1959) *Acta Crystallogr. 12*, 897−909.

Suen, K.-L., Bustelo, X. R., Pawson, T., & Barbicid, M. (1993) *Mol. Cell. Biol. 13*, 5500−5512.

Wiseman, T., Williston, S., Brandts, J. F., & Lin, L-N. (1989) *Anal. Biochem. 179*, 131−137.

Yang, Y. S., Garbay, C., Duchesne, M., Cornille, F., Jullian, N., Fromage, N., Tocque, B., & Roques, B. P. (1994) *EMBO J. 13*, 1270−1279.

Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S., & Schreiber, S. L. (1992) *Science 258*, 1665−1668.

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