Nuclear Magnetic Resonance Solution Structure of the Growth Factor Receptor-Bound Protein 2 Src Homology 2 Domain

Kevin H. Thornton,[‡] W. Tom Mueller,[§] Patrick McConnell,[§] Guochang Zhu,[||] Alan R. Saltiel,[||] and Venkataraman Thanabal*,[‡]

Departments of Chemistry, Biotechnology, and Signal Transduction, Parke-Davis Pharmaceutical Research, Division of Warner-Lambert Company, 2800 Plymouth Road, Ann Arbor, Michigan

Received November 3, 1995; Revised Manuscript Received June 20, 1996[⊗]

ABSTRACT: A family of NMR solution structures of the growth factor receptor-bound protein 2 (Grb2) SH2 domain has been determined by heteronuclear multidimensional NMR. Proton, nitrogen, and carbon chemical shift assignments have been made for the SH2 domain of Grb2. Assignments were made from a combination of homonuclear two-dimensional and ^{15}N- and ^{13}C-edited three-dimensional spectra at pH 6.2 and 298 K. Structure-induced proton and carbon secondary shifts were calculated and used to facilitate the spectral assignment process. NOE, scalar coupling, secondary chemical shift, and amide proton exchange data were used to characterize the secondary structural elements and hydrogen-bonding network in the Grb2 SH2 domain. The three-dimensional structure of the Grb2 SH2 domain was calculated using 1112 restraints obtained from NOE, coupling constant, and amide proton exchange data. The rmsd for the 24 calculated structures to the mean structure of the Grb2 SH2 domain was 0.75 Å for backbone and 1.28 Å for all heavy atoms. The three-dimensional fold of the Grb2 SH2 domain is similar to that observed for other SH2 domains and consists of two α-helical segments and eight β-strands, six strands that make up two contiguous antiparallel β-sheets, and two strands that form two short parallel β-sheets. The structure of the phosphotyrosine binding pocket of Grb2 is similar to that observed for other SH2 domains. The hydrophobic binding pocket of Grb2 is similar to that observed for Src with the exception that tryptophan 121 of Grb2 occupies part of the pY+3 binding pocket. Structural implications for the Grb2 SH2 domain selectivity at the pY+2 and pY+3 sites are discussed.

Src homology 2 (SH2[1]) domains are noncatalytic regions of proteins shown to play an important role in cellular signaling pathways [for reviews, see Aaronson (1991), Schlessinger and Ullrich (1992), Fantl et al. (1993), Fry et al. (1993), Kuriyan and Cowburn (1993), Margolis and Skolnik (1994), Downward (1994), Cowburn (1995), and Chardin et al. (1995)]. SH2 domains have been found in many cytoplasmic signaling proteins and have been shown to bind to phosphorylated growth factor receptors or other proteins phosphorylated in response to extracellular signals. SH2 domains are approximately 100 amino acid residues in length and are often found in conjunction with a 50-amino acid domain known as a Src homology 3 (SH3) domain. There are two subfamilies of SH2 domain-containing proteins. The first family includes proteins that have catalytic or regulatory activity, and the second family consists of proteins that function only as adapter proteins (Fry et al., 1993).

Growth factor receptor-bound protein 2 (Grb2) (Lowenstein et al., 1992) is a mammalian homologue of Drk (*Drosophila*) and Sem-5 (*Caenorhabditis elegans*) that functions as an adapter protein (Fry et al., 1993; Margolis & Skolnik, 1994; Downward, 1994). It is a 217-amino acid residue protein that has no catalytic domain but consists of one SH2 domain flanked by two SH3 domains. Grb2 is of interest due to its important role in the Ras signaling pathway. As part of this pathway, Grb2 activates Ras by recruiting a nucleotide exchange factor, Sos (son of sevenless), to the plasma membrane where Ras is located (Buday & Downward, 1993). The nucleotide exchange factor, Sos, complexes with Grb2 through its SH3 domains. Grb2 is involved in many regulatory pathways by binding to receptors directly (EGFR) or indirectly (through proteins such as Shc, FAK, Syp, and IRS-1) (Rozakis-Adcock et al., 1992; Skolnik et al., 1993; Buday & Downward, 1993; Pronk et al., 1994; Schlaepfer et al., 1994).

The SH2 domain of Grb2 recognizes the motif pYXNZ (where pY is a phosphorylated tyrosine and X and Z are hydrophilic and hydrophobic residues, respectively) on the receptor binding sequence (Songyang et al., 1994). Currently, there is no structural information to explain the specificity of the Grb2 SH2 domain at the pY+1, pY+2, and pY+3 pockets. However, experiments in which the SH2 domain of Grb2 was screened against peptide libraries show

* To whom correspondence should be addressed.
[‡] Department of Chemistry.
[§] Department of Biotechnology.
[||] Department of Signal Transduction.
[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1996.
[1] Abbreviations: Grb2, growth factor receptor-bound protein 2; SH2, Src homology 2; SH3, Src homology 3; Sos, son of sevenless; GST, glutathione *S*-transferase; EGFR, epidermal growth factor receptor; DQF-COSY, two-dimensional double-quantum filtered correlation spectroscopy; TQF-COSY, two-dimensional triple-quantum filtered correlation spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; TPPI, time-proportional phase incrementation; TOCSY, two-dimensional total correlation spectroscopy; SCUBA, stimulated cross-peaks under bleached α-proton; FID, free induction decay; PMSF, phenylmethanesulfonyl fluoride; TSP, sodium 3-(trimethylsilyl)propionate.

S0006-2960(95)02615-8 CCC: $12.00

a strong preference for asparagine at the pY+2 pocket (Songyang et al., 1994). This preference for asparagine at the pY+2 site of the Grb2 SH2 domain is unique from other SH2 domains. Mutation experiments on the Src SH2 domain suggest that the residue at the first position in the loop between β-strands βE and βF (referred to as position EF1) plays an important role in pY+2 and pY+3 selectivity (Marengere et al., 1994). Marengere et al. (1994) showed that mutation of the threonine at position EF1 to tryptophan resulted in the SH2 domain of Src selecting a Grb2-like peptide that contains asparagine and proline at the pY+2 and pY+3 sites, respectively. These experiments also showed that mutation of threonine at position EF1 to tryptophan did not influence pY+1 selectivity and suggest that other residues of the SH2 domain influence selectivity at this site.

NMR (Goudreau et al., 1994; Kohda et al., 1994; Terasawa et al., 1994) and X-ray (Guruprasad et al., 1995) structural studies of the Grb2 protein have focused primarily on the SH3 domains. Recently, a low-resolution X-ray structure of full length Grb2 was reported (Maignan et al., 1995); however, no details of its structure have been published. Although the general structure of the Grb2 SH2 domain is expected to be similar to that of other SH2 domains (Overduin et al., 1992; Booker et al., 1992; Waksman et al., 1992, 1993; Eck et al., 1993; Lee, 1994; Pascal, 1994; Hensmann et al., 1994; Xu et al., 1995; Zhou et al., 1995; Narula et al., 1995; Metzler et al., 1996), it is the local structure that gives the SH2 domain of Grb2 its unique specificity and molecular recognition at the binding site. Since the motif recognized by the Grb2 SH2 domain differs from that of other SH2 domains, it is of interest to understand the structural features that lead to its unique specificity. Design of molecules that inhibit binding of Grb2 to its target receptor should interrupt the Ras signaling pathway and may promise therapeutic leads for certain diseases, like cancer. Structural information on the molecular interactions of the Grb2 SH2 domain and its inhibitors will provide a valuable tool for modeling and design of better drug candidates. Here we report the first NMR assignment and solution structure of the Grb2 SH2 domain.[2]

EXPERIMENTAL PROCEDURES

Expression and Purification of the Grb2 SH2 Domain. JM101 or BL21(DE3) *Escherichia coli* was transformed with a pGEX-KT (Hakes & Dixon, 1992) vector encoding the Grb2 SH2 domain residues 60−158. Fermentation was carried out in a 2 L fermenter at 37 °C under pH control. Isotope-labeled cultures were grown in M9 minimal media containing 1 g/L [^{15}N]ammonium chloride (99%) (Isotech) and 10 g/L glucose (Sigma, cell culture) or 4 g/L uniformly ^{13}C-labeled glucose (Isotech, 99%). All fermentations were supplemented with vitamins and trace elements (Weber et al., 1992). The cells were lysed in phosphate-buffered saline (pH 7.2) that contained 1% Triton-X, 1% β-mercaptoethanol, 0.1 mM PMSF, 10 μg/mL aprotinin, and 10 μg/mL leupeptin. Glutathione *S*-transferase−Grb2 SH2 fusion protein (GST−Grb2) was affinity purified on glutathione−agarose beads. The purified fusion protein was subjected to digestion with bovine thrombin for 24 h to give GST and Grb2 SH2 domain. The SH2 domain of Grb2 was purified from the digest mixture on a Superdex 75 (Pharmacia) column equilibrated in 150 mM NaCl and 50 mM sodium phosphate (pH 7.2). The Grb2 SH2 domain was characterized by electrospray mass spectrometry and N-terminal sequencing. The observed mass (12 364.2 ± 3.6 Da) agrees with the theoretical mass (12 364 Da) based upon sequence. The purified protein contains 107 residues, 99 of which are derived from full length Grb2 (residues W60−P158) and eight residues (two N-terminal and six C-terminal) that come from the pGEX vector. Numbering of the Grb2 SH2 domain residues is with respect to the full length protein.

NMR Sample Preparation. The solution containing the Grb2 SH2 domain was concentrated and then buffer exchanged into 50 mM phosphate and 100 mM NaCl (pH 6.2) using an Amicon concentrator equipped with a YM3 membrane. Typical NMR sample concentrations were 1.5 mM, based on BioRad protein assay. NMR samples in D_2O were made by exchanging H_2O samples into D_2O (100 mM NaCl and 50 mM phosphate at pH 6.2) in an Amicon concentrator.

NMR Spectroscopy. All NMR data were acquired on a Bruker AMX 500 or AMX 600 MHz spectrometer at 25 °C equipped with triple-resonance hardware. Some experiments were run on an instrument equipped with triple-axis gradient hardware. Unless stated otherwise, water suppression was accomplished either by on-resonance low-power continuous wave saturation at the water resonance or by off-resonance saturation using a DANTE type sequence (Kay et al., 1989) during the 1.2 s acquisition delay. All NMR data were processed and analyzed with Felix software (Biosym) on an Iris workstation (Silicon Graphics Inc.). Time domain convolution (Marion et al., 1989c) was used to remove the H_2O signal during processing of all nongradient experiments recorded in H_2O. Prior to Fourier transform, a sine bell-squared apodization function was applied in t_1, t_2, and t_3. All spectra were referenced to water at 4.75 ppm relative to TSP at 25 °C and pH 6.2. Carbon and nitrogen dimensions were referenced indirectly to TSP using the ratio of gammas (Hγ/Cγ or Hγ/Nγ) (Bax & Subramanian, 1986).

Two-Dimensional Homonuclear NMR Experiments. All homonuclear two-dimensional NMR data of the unlabeled Grb2 SH2 domain were collected using time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983). Homonuclear DQF-COSY (Piantini et al., 1982; Shaka & Freeman, 1983; Rance et al., 1983) and TQF-COSY (Piantini et al., 1982; Shaka & Freeman, 1983; Rance et al., 1983) spectra were acquired in D_2O at 500 MHz. TOCSY (Braunschweiler & Ernst, 1983) and clean-TOCSY (Cavanagh & Rance, 1992) experiments were run in D_2O and H_2O at 500 MHz using a DIPSI-2 pulse sequence. Each t_1 increment of the TOCSY experiments was acquired with three mixing times (30, 45, and 60 ms for TOCSY and 22, 40, and 50 ms for clean-TOCSY) that were coadded together to yield a single data set. NOESY (Jeener et al., 1979; Kumar et al., 1980) experiments were run at 600 MHz (H_2O) and 500 MHz (D_2O). The SCUBA technique (Brown et al., 1988) was used in TOCSY and NOESY experiments to recover resonances lost due to saturation of H_2O.

[2] A preliminary account of this work was presented at the Frontiers of NMR in Molecular Biology−IV Keystone Symposium, April 3−9, 1995, Keystone, CO [Thornton, K. H. & Thanabal, V. (1995) Solution Structure of Src Homology-2 Domain of Growth Factor Receptor Bound Protein 2 (Grb2). *J. Cell. Biochem.* (supplement 21B), 38 (abstract D2-163)].

Two-Dimensional Double-Resonance NMR Experiments. Two-dimensional $^1H-^{13}C$ HMQC (Morris & Freeman, 1979; Bodenhausen & Ruben, 1980) (500 MHz) and $^1H-^{15}N$ HMQC (600 MHz) spectra were collected with sweep widths in the indirect dimension of 8928 and 3125 Hz, respectively. Phase sensitive data were collected using time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983) in the indirect dimension. Amide proton exchange was followed by a series of $^1H-^{15}N$ HMQC experiments run over a 24 h period. Spectra were collected with 128 increments and 32 scans at each increment.

A HMQC-J (Kay & Bax, 1990) spectra was recorded on a uniformly ^{15}N-labeled Grb2 SH2 domain with 288 scans and 320 complex points in t_1 and the transmitter placed at 8.48 ppm. Complex linear prediction was used to extend the FID in the ^{15}N dimension (Olejniczak & Eaton, 1990; Zhu & Bax, 1990). A series of two-dimensional constant time HMQC-J experiments (Kuboniwa et al., 1994) were also run on a ^{15}N-labeled Grb2 SH2 domain. A set of six experiments was acquired with mixing times of 45, 60, 80, 100, 120, and 140 ms. The data were processed and analyzed as described by Kuboniwa et al. (1994).

Three-Dimensional Double-Resonance NMR Experiments. All three-dimensional experiments were acquired using the States-TPPI (Marion et al., 1989b) method in t_1 and t_2 to obtain phase sensitive data. The initial t_2 delay (^{15}N or ^{13}C evolution period) in folded spectra was calculated so that the zero- and first-order phase corrections in ω_2 were 90 and $-180°$, respectively (Bax et al., 1991). During processing, the FID in the indirect dimension was extended by complex linear prediction (Olejniczak & Eaton, 1990; Zhu & Bax, 1990). Broad-band decoupling during acquisition was accomplished by use of a garp or waltz16 decoupling field.

Three-dimensional ^{15}N-edited TOCSY-HSQC and NOESY-HSQC (Fesik & Zuiderweg, 1988; Marion et al., 1989a; Zuiderweg & Fesik, 1989) spectra were acquired at 600 MHz on a uniformly ^{15}N-labeled sample of Grb2 SH2 domain in H_2O. Sweep widths for both experiments were 7812 Hz (^1H, F_1), 1667 Hz (^{15}N, F_2), and 11 905 Hz (^1H, F_3). The TOCSY-HSQC spectrum was collected with 84, 25, and 1024 complex points in t_1, t_2, and t_3, respectively, and with 32 scans per increment. The first 16 scans of the TOCSY-HSQC had a mixing time of 35 ms and the second 16 scans a mixing time of 60 ms. The 150 ms NOESY-HSQC spectrum was collected with 110, 25, and 1024 complex points in t_1, t_2, and t_3, respectively, and with 32 scans per increment. A fifth-order polynomial baseline correction function was applied to both spectra in ω_3 during processing. Both data sets were zero filled to give a final matrix size of $512 \times 128 \times 512$ in ω_1, ω_2, and ω_3, respectively.

A 65 ms ^{15}N-edited NOESY-HSQC spectrum was acquired at 500 MHz with 32 scans per increment. Water suppression was accomplished by using a watergate gradient sequence (Sklenar et al., 1993). Sweep widths were 5405 Hz (^1H, F_1), 1082 Hz (^{15}N, F_2), and 3597 Hz (^1H, F_3). The data were folded in F_2. A total of 128, 64, and 256 complex points were acquired in t_1, t_2 and t_3, respectively. The transmitter was placed at 8.23, 4.75, and 119.7 ppm in ω_1, ω_2, and ω_3, respectively.

A 14.8 ms HCCH-TOCSY (Kay et al., 1990; Bax et al., 1990) spectrum was acquired on a sample that was uniformly labeled with ^{15}N and ^{13}C. The data were acquired at 500 MHz with 110, 31, and 1024 complex points in t_1, t_2, and t_3, respectively, and with 16 scans per increment. The spectrum was collected with sweep widths of 3846 Hz (^1H, F_1), 2666 Hz (^{15}N, F_2) and 7575 Hz (^1H, F_3). The transmitter was placed at 3.00, 43.0, and 3.00 ppm in t_1, t_2 and t_3, respectively. Carbonyl decoupling was accomplished by off-resonance decoupling (Kay et al., 1989).

A three-dimensional ^{13}C-edited TOCSY-HSQC (Ikura et al., 1990) spectrum was collected at 500 MHz on a uniformly ^{15}N and ^{13}C-labeled sample. Sweep widths were 7576 Hz (^1H, F_1), 3846 Hz (^{15}N, F_2), and 2667 Hz (^1H, F_3). The spectrum was collected with 16 scans per increment and 220, 32, and 1024 complex points in t_1, t_2 and t_3, respectively. The data were zero filled to give a final matrix size of $512 \times 128 \times 512$ in ω_1, ω_2 and ω_3, respectively. A 65 and 150 ms three-dimensional ^{13}C NOESY-HMQC (Ikura et al., 1990) spectrum was collected at 500 and 600 MHz, respectively, with 16 scans per increment on a uniformly ^{15}N- and ^{13}C-labeled Grb2 SH2 domain sample in D_2O. The 65 ms three-dimensional ^{13}C NOESY-HMQC spectrum was collected with 110, 50, and 512 complex points and sweep widths of 4545 Hz (^1H, F_1), 5814 Hz (^{13}C, F_2) and 7575 Hz (^1H, F_3). The 150 ms data were collected with 84, 60, and 512 complex points and sweep widths of 5555 Hz (^1H, F_1), 6993 Hz (^{13}C, F_2) and 7412 Hz (^1H, F_3). The data were zero filled to give a final matrix size of $512 \times 256 \times 512$ in ω_1, ω_2 and ω_3, respectively.

Assignment of Grb2 SH2 Domain. Amino acid spin system assignments were accomplished using a combination of DQF-COSY, TOCSY, ^{15}N-edited TOCSY-HSQC, ^{13}C-edited TOCSY-HMQC, and HCCH-TOCSY experiments. Sequential assignments were made by the standard methodology using NOE connectivities of the type $d_{\alpha N}(i,i+1)$, $d_{NN}(i,i+1)$, and $d_{\beta N}(i,i+1)$ in the ^{15}N-edited NOESY-HSQC and TOCSY-HSQC data (Wüthrich, 1986). Medium and long range NOE assignments were made from ^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY- HMQC spectra.

Distance and Dihedral Restraints. Distance restraints were generated from 65 ms nuclear Overhauser enhancement data by converting NOE volumes to a continuum of distances and then adjusting the distances to account for error in the data. Cross-peaks from 65 ms NOE spectra were peak picked using Felix (Biosym). NOE volumes from ^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY-HMQC spectra were converted to a continuous range of distances using the procedure described by Xu et al. (1995). This was accomplished by assigning the average of the 10 strongest and the 10 weakest NOE volumes to 1.8 and 5.0 Å, respectively, in the ^{15}N-edited NOESY-HSQC or ^{13}C-edited NOESY-HMQC data sets. These two calibration points were then used to calculate distances from the ^{15}N-edited NOESY-HSQC or ^{13}C-edited NOESY-HMQC NOE volumes (Xu et al., 1995). Distances were obtained from the 65 ms two-dimensional NOESY data by setting the distance corresponding to the volume of the $H\alpha_i-H\alpha_j$ NOE of F83−V99 to 2.2 Å and then using this calibration point to calculate distances from the measured volumes.

Upper limit NOE distance restraints were defined as the calculated distance plus 10% (Xu et al., 1995) and given a force constant of 40 kcal mol^{-1} $Å^{-2}$. All lower bound distance restraints were set to 1.8 Å. Restraints with distances less than 2.1 Å were set to 2.1 Å to avoid excessive restraint. All restraints to non-stereospecifically assigned

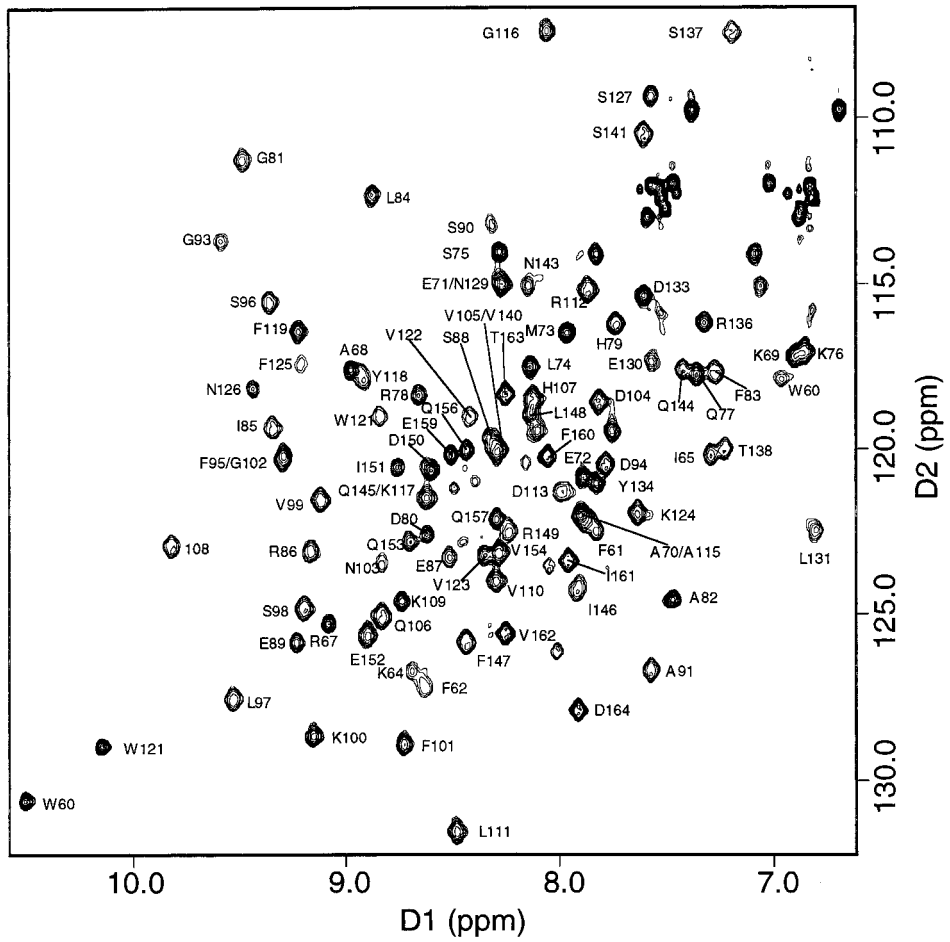

FIGURE 1: $^{1}H-^{15}N$ HSQC spectrum of the Grb2 SH2 domain. All amide backbone protons and tryptophan indole NH protons are labeled. Glutamine and asparagine amide side chain protons are not labeled. Minor unlabeled peaks are due to C-terminal proteolysis of the Grb2 SH2 domain (see discussion on protein stability).

methyl and methylene protons were made to a center-averaged pseudoatom and given the appropriate corrections (Wüthrich, 1983).

Torsion angle, ϕ, restraints for residues with $^{3}J_{HN-H\alpha} > 8.0$ Hz were given as $-120 \pm 40°$, and residues with $^{3}J_{HN-H\alpha} < 6.0$ Hz were given as $-50 \pm 40°$. In the α-helical regions of the protein, a ψ angle restraint of $-60 \pm 50°$ was used. All torsion angle restraints were given a force constant of 200 kcal mol^{-1} rad^{2}. Hydrogen bond restraints were included when a hydrogen bond occurred in a region of regular secondary structure and was associated with a slowly exchanging amide proton. Each hydrogen bond is characterized by two distance restraints: 1.8 Å $\leq d_{HN-O} \leq$ 2.3 Å and 2.5 Å $\leq d_{N-O} \leq$ 3.3 Å.

Structure Calculations. Grb2 SH2 domain structures were calculated using the Biosym programs DGII and Discover interfaced to InsightII for visualization. Initial structures were calculated using distance geometry (DGII) by first using a triangle inequality strategy for bound smoothing followed by embedding in four-dimensional space and majorization (Havel, 1991). The embedded structures were next optimized using a simulated annealing protocol (Havel, 1991) and minimized with 250 steps of conjugate gradient minimization. Structures with a DGII error function value of less than 1 were minimized without restraints in Discover (Biosym) using 500 steps of steepest descent minimization and 1000 steps of conjugate gradient minimization. This was followed by another 1000 steps of conjugate gradient minimization with restraints. Minimization and dynamics were carried out using a CVFF force field (Dauber-Osguthorpe, 1988) that included a potential energy term in the form of a skewed biharmonic function for distance and dihedral angle restraints. Lennard-Jones terms were included for nonbonded interactions. Electrostatic terms were not included in the calculation. Minimization was followed by 20 ps of molecular dynamics at 298 K, during which the force constants for the NMR restraint potential energy term were reduced to 75% of their original value. After molecular dynamics, the structures were again minimized using the previous three-step minimization with the exception that force constants for the potential energy term were at 75% (30 kcal mol^{-1} Å^{-2} and 150 kcal mol^{-1} rad^{2}) of their original value during the final 1000 steps of conjugate gradient minimization.

RESULTS AND DISCUSSION

Protein Stability. Electrospray mass spectrometry and N-terminal protein sequencing results on the Grb2 SH2 domain showed the mass and primary amino acid sequence to match that expected on the basis of the Grb2 amino acid sequence. However, within a few hours of the preparation of a NMR sample of the Grb2 SH2 domain, the solution became cloudy due to precipitation of the protein. The precipitation process was observed to be temperature dependent, being faster at 37 than at 15 °C. After a few days, the precipitation process stopped and the concentration of

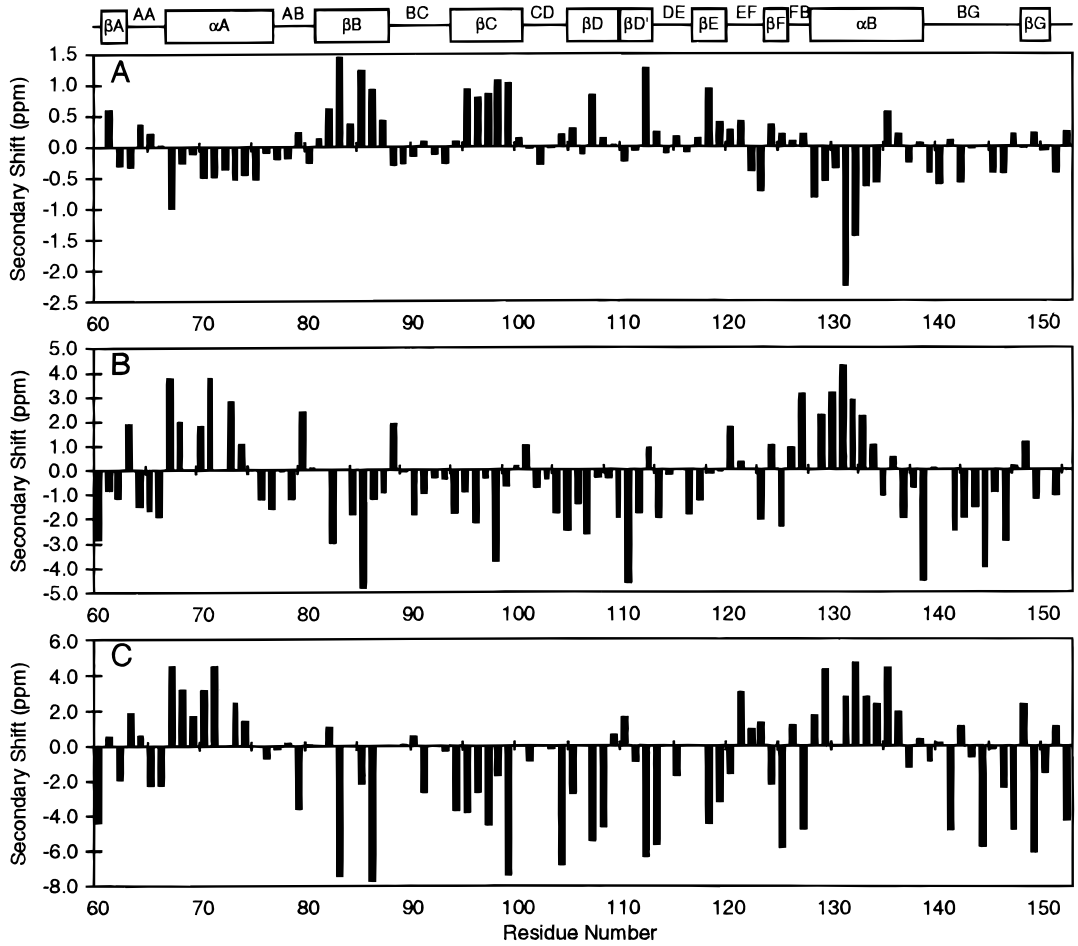

FIGURE 2: Secondary chemical shifts of Grb2 SH2 Hα (A), Cα (B), and Cα−Cβ (C) resonances. Secondary chemical shifts were calculated by subtracting the random coil Hα (Wütrich, 1986), Cα, and Cβ chemical shifts (Thanabal et al., 1994) from the respective Hα, Cα, and Cβ observed chemical shifts. Cα and Cβ chemical shifts of residues following proline were corrected for proline chemical shift effects (Wishart et al., 1995). (C) Difference of Cα and Cβ secondary chemical shifts. The observed secondary structural elements of the Grb2 SH2 domain are shown at the top of the figure.

the Grb2 SH2 domain remained constant. The precipitation process appears to be a function of concentration. Conditions were not found to stop the precipitation problem entirely. Light scattering experiments (Protein Solutions, Inc., data not shown) as well as proton line width measurements suggest that the Grb2 SH2 domain is forming aggregates in solution at NMR sample concentrations (1−1.5 mM). The light scattering results suggest that the protein is not a dimer but is a mixture of aggregates.

In addition to the problem with precipitation, electrospray mass spectrometry and sequencing data have shown that up to six residues in the C-terminal tail of the Grb2 SH2 domain were being removed through proteolysis. The proteolysis, however, did not seem to be the cause of the precipitation. Both the protein remaining in solution and the protein that had precipitated out had the same degree of proteolysis on the C-terminal tail. Further purification of the Grb2 SH2 domain by ion exchange chromatography did not prevent the protease activity. The proteolysis appeared to slow significantly after the last five residues were removed. At this stage, the Grb2 SH2 domain was stable for several weeks. Since the residues lost due to proteolysis were not part of the natural Grb2 SH2 domain sequence, it was felt that this was not a serious problem. From mass spectrometry and $^{1}H-^{15}N$ HMQC NMR data, the protein appeared to be structurally intact except for the last five or six residues. These problems did not appear to affect the structural integrity of the Grb2 SH2 domain. Since the structure of the Grb2 SH2 domain still appeared to be intact, efforts to solve its structure were continued.

Assignments. The strategy used for making the sequential NMR assignments of the SH2 domain of Grb2 involved identifying amide proton to side chain proton intraresidue correlations using DQF-COSY, TOCSY, and three-dimensional ^{15}N-edited TOCSY-HSQC and then using sequential NOE correlations to connect the spin systems. Although the chemical shift dispersion was good in $^{1}H-^{15}N$ HSQC spectra (see Figure 1), the spin system assignments proved to be difficult due to poor magnetization transfer from the amide protons to their respective side chain protons in TOCSY and three-dimensional ^{15}N-edited TOCSY. This poor magnetization transfer is due to the short T_2 relaxation times (manifested by proton line widths of approximately 25 Hz) in the Grb2 SH2 domain in solution. Aggregation of the Grb2 SH2 domain is most likely the cause of the short T_2 values. To optimize the observed magnetization transfer, TOCSY experiments were acquired with multiple mixing times that were coadded together. Even in these experiments, coherence was rarely observed to transfer from the amide to β-protons or beyond. Often, coherence from the amide protons would transfer no further than to the α-proton. The only residues to show good scalar coupling properties were residues E152−D164 as these residues comprise the C-terminal flexible tail. In order to assign the amino acid side

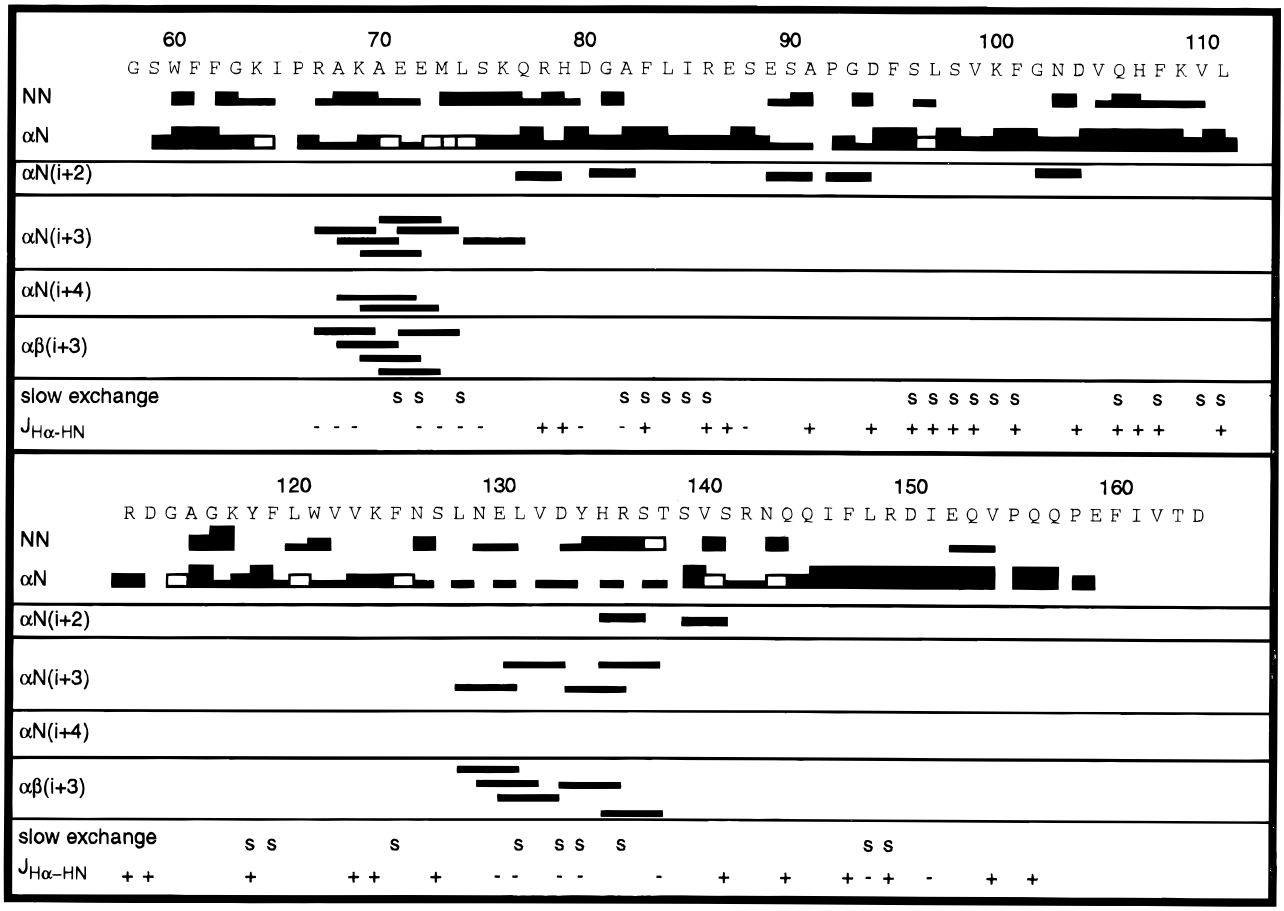

FIGURE 3: Summary of sequential and medium range NOEs, amide proton exchange, and $^3J_{HN-H\alpha}$ coupling constant data. The amino acid sequence for the SH2 domain of Grb2 is shown (numbering is with respect to the full length Grb2). Displayed in the figure are the αN and NN sequential NOEs as well as the αN(i+2), αN(i+3), αβ(i+3), and αN(i+4) medium range NOEs. The relative intensity of sequential NOE data is shown. Empty boxes indicate sequential NOEs that were ambiguous due to overlap in NOESY spectra. All slowly exchanging amide protons are labeled with an s in the figure. Coupling constant data are represented with a + for $^3J_{HN-H\alpha}$ greater than 8 Hz and a − for $^3J_{HN-H\alpha}$ less than 6 Hz.

chain protons, ^{13}C TOCSY-HMQC and ^{13}C-HCCH-TOCSY data were used to connect amino acid side chain protons to their α- and β-protons. In spite of the difficulty in connecting amide protons to their respective side chain protons, assignments were made for 98 of the 103 backbone amide protons of Grb2 SH2 domain (Figure 1).

After identification of residues according to spin system type, sequential nuclear Overhauser enhancements of the type $d_{\alpha N}(i,i+1)$, $d_{NN}(i,i+1)$, and $d_{\beta N}(i,i+1)$ in three-dimensional ^{15}N-edited NOESY-HSQC data were used to connect spin systems. The sequential assignment process was facilitated by the use of gradients as amide protons for residues L120, R142, and N143 were only observed in HSQC and NOESY-HSQC spectra that used gradient water suppression. Sequential assignment of Grb2 was initiated by first identifying the unique amino acid pairs, S90−A91, V122−V123, and V162−T163, and then extending the assignments from these residues. Sequential NOE correlations between proline Hδ and the Hα of the previous residue were observed for all four prolines. This indicates that the four prolines in the SH2 domain of Grb2 are in the *trans* conformation. The sequential assignments in the two helical regions of the Grb2 SH2 domain were difficult to make due to degeneracy in the α-proton chemical shifts. Sequential assignments in these regions relied heavily on NN and βN sequential NOE connectivities. The assignments in the α-helical regions of the protein were confirmed by observation of the expected αN(i+3), αβ(i+3), and αN(i+4) medium range NOEs.

Amino acid side chain resonances were assigned by use of ^{13}C-edited TOCSY-HMQC and HCCH-TOCSY experiments. These experiments allowed the assignment of most amino acid side chain proton resonances. The aromatic ring protons of most phenylalanine residues have not been assigned due to chemical shift degeneracy in both proton and carbon dimensions. Amide protons for residues S59, G114, L128, V132, and S139 have also not been identified. The reason that these resonances were not observed is not clear but could be due to overlap with other amide protons or chemical exchange. The ^{1}H, ^{15}N, and ^{13}C chemical shift assignments for all assigned resonances are available as supporting information.

During the assignment of the Grb2 SH2 domain, information was obtained about its secondary structural elements. Proton and carbon chemical shifts obtained from the ^{13}C-edited three-dimensional NOESY-HMQC spectra of the Grb2 SH2 domain were analyzed for structure-induced secondary shifts. Secondary chemical shifts of Hα, Cα, and Cβ have been shown to correlate well with the observed secondary structure found in proteins (Spera & Bax, 1991; Wishart et al., 1991; de Dois et al., 1993) and peptides (Lee et al., 1992; Reily et al., 1992) (Figure 2). An enhanced secondary shifts can be observed by plotting the difference of Cα and Cβ secondary shift (Figure 2C), since secondary shifts of Cβ are opposite in sign from that of Cα. The secondary shift

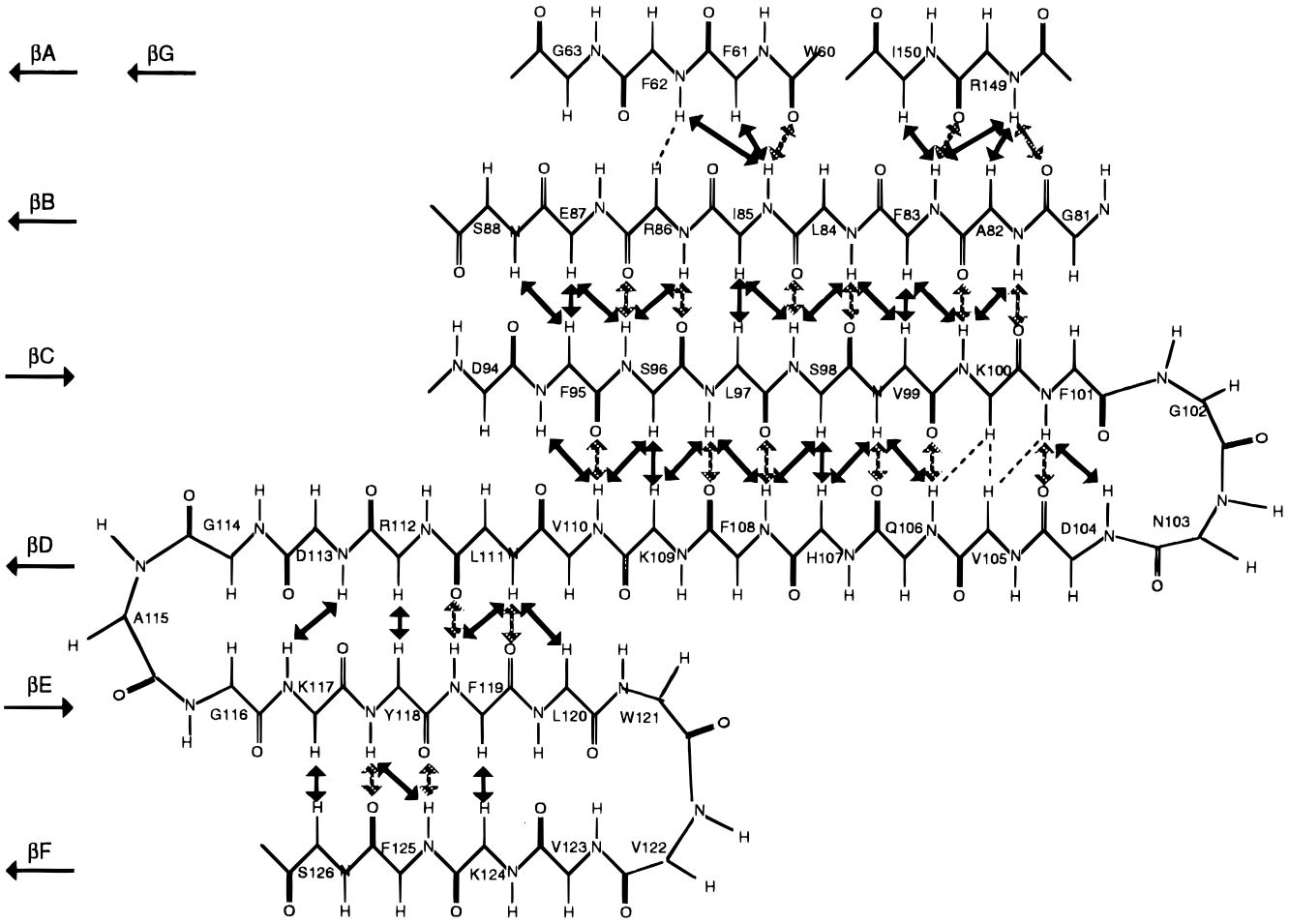

FIGURE 4: β-sheet structure of the Grb2 SH2 domain as determined from NOE experiments and amide proton exchange. Solid arrows indicate observed interstrand NOEs, and hashed arrows represent hydrogen bonds determined from slowly exchanging amide protons. Dashed lines indicate areas where NOEs could not be assigned due to overlap in spectra. Amino acid residues are numbered with respect to full length Grb2. Strands βA−βG are labeled in the figure above the arrows that indicate the parallel or antiparallel nature of the β-sheet structure of the SH2 domain of Grb2.

information obtained for the SH2 domain of Grb2 was used to assist in the assignment of spin systems in regions of regular secondary structure. On the basis of the secondary chemical shifts, as shown in Figure 2, the Grb2 SH2 domain is predicted to have two helical segments, R67−Q77 and L128−H135 (NOE data suggest that this helix extends to T138), as well as the following β-strands: A82−E87 (strand βB), F95−K100 (strand βC), Q106−R112 (strand βD), Y118−F119 (strand βE), K124−F125 (strand βF), and R149−D150 (strand βG).

Analysis of two- and three-dimensional NOE data has provided additional evidence for the structural features in the Grb2 SH2 domain. Assignment of several short range interresidue NOEs ($2 < |i - j| < 5$) (Figure 3) including d_{NN} connectivities suggest that residues R67−Q77 (αA_1−αA_{11}) and L128−T138 (αB_1−αB_{11}) are α-helical (Figure 3). Furthermore, several interstrand NOEs of the type $H\alpha_i$−$H\alpha_j$, $H\alpha_i$−HN_j, and HN_i−HN_j were observed. The observed NOEs are consistent with two adjacent three-stranded antiparallel β-sheets that are comprised of strands βB, βC, and βD and strands βD′, βE, and βF as shown in Figure 4. In addition, the data suggest the presence of two short parallel β-sheets formed by strands βB−βA and βB−βG. This core structure is consistent with that observed for other SH2 domains. The absence of medium and long range NOEs in the segment Q153−D164 along with sharp line widths suggests that these residues are predominantly random coil.

Hydrogen Bonds. Through a series of HMQC spectra acquired immediately after exchanging of the Grb2 SH2 domain into D_2O, 27 slowly exchanging amide protons were observed after 4 h (see Figure 3). Of the 27 slowly exchanging amide protons, 25 were assigned to amide protons involved in regular secondary structure. Amide protons for residues V110, L111, and L131 show very slow exchange with no visible signs of exchange even after 24 h. On the other hand, amide protons for E71, E72, L74, A82, F101, and R136 were present after 4 h but fully exchanged within 24 h. All other slowly exchanging amide protons were still present after 24 h, although partially exchanged (Figure 3). These exchange data agree with the relative rate of exchange observed in our 65 ms gradient three-dimensional ^{15}N-edited NOESY-HSQC. The ω_2−ω_3 two-dimensional plane at the ω_1 water frequency shows HN−^{15}N cross-peaks corresponding to rapidly exchanging amide protons. The relative intensity of these exchange cross-peaks corresponds to the amide proton exchange rate. Amide protons that show large exchange peaks in the gradient three-dimensional ^{15}N-edited NOESY-HSQC are the same ones that exchange rapidly in the D_2O exchange experiments. The amide proton exchange data are also consistent with the observed α-helical and β-sheet regions of the Grb2 SH2 domain. Most slowly exchanging amide protons were assigned to residues in these regions of regular secondary structure as shown in Figures 3 and 4. The hydroxyl proton

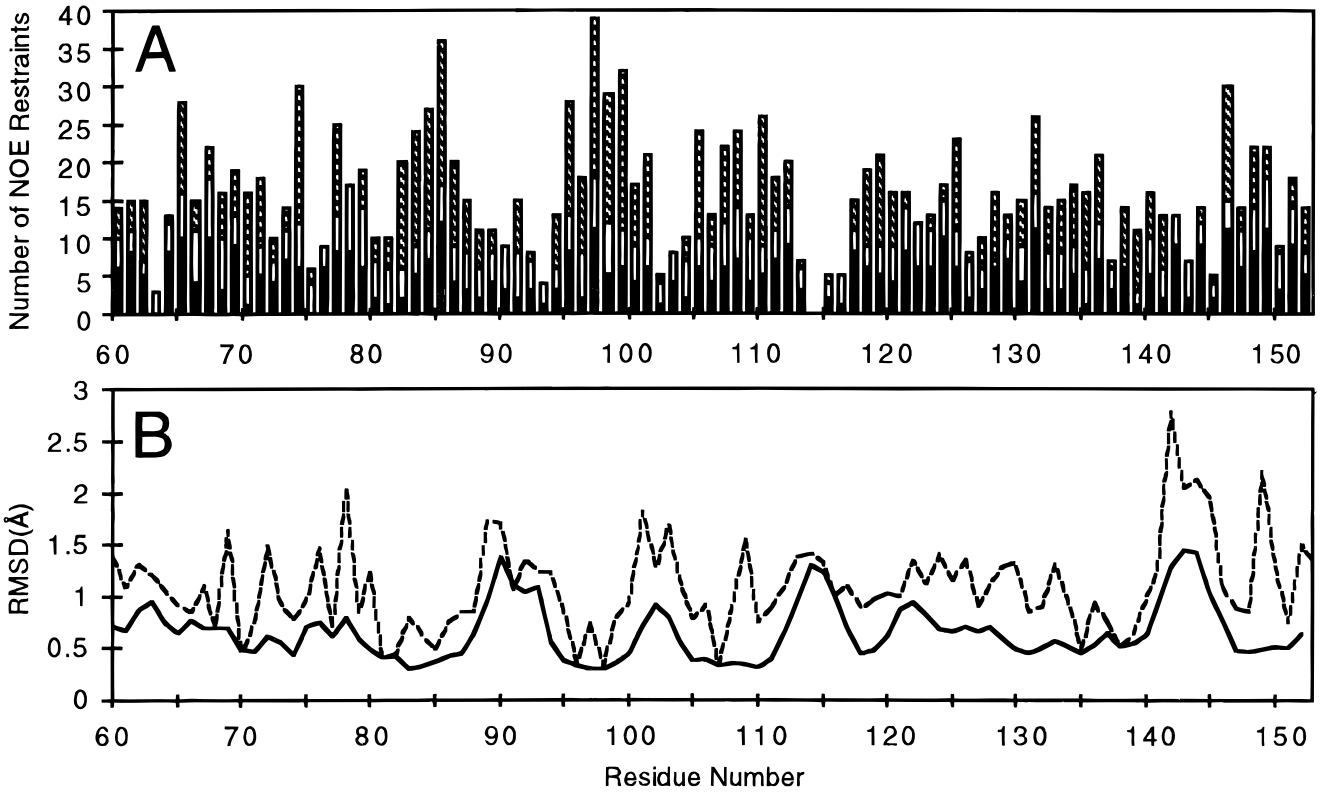

FIGURE 5: (A) Histogram showing the NMR distance restraints for residues 60−152 used in calculation of the Grb2 SH2 domain structure. Three different types of restraints are shown in the figure: black, intraresidue; white, sequential; and hashed, medium and long range ($2 < |i - j|$). The bar height represents the total number of restraints. (B) Plot showing the rmsd for each residue with respect to the average calculated structure of the Grb2 SH2 domain. The solid and dashed lines represent the rmsd with respect to the average structure for each residue of all backbone (N, Cα, and C′) and all heavy atoms, respectively.

of residue S139 may also be slowly exchanging as it is observed in presaturated NOESY spectra recorded in water.

Coupling Constants. Coupling constants of the type $^3J_{HN-H\alpha}$ were extracted from two different HMQC-J experiments. CT-HMQC-J (Kuboniwa et al., 1994) experiments suffered from poor sensitivity due to the large proton line widths of the Grb2 SH2 domain which makes scalar coupling dependent coherence transfer inefficient. The HMQC-J (Kay & Bax, 1990) experiment had better sensitivity. However, the large ^{15}N line widths made it difficult to measure accurate coupling constants. The $^3J_{HN-H\alpha}$ coupling constants were estimated only qualitatively as $J < 6$ Hz, $6 < J < 8$ Hz, or $J > 8$ Hz.

Restraints. NOE distance restraints were obtained from three 65 ms NOESY experiments: two-dimensional NOESY in D_2O (26 restraints), gradient ^{15}N-edited NOESY-HSQC in H_2O (538 restraints), and ^{13}C-edited NOESY-HMQC in D_2O (422 restraints). Cross-peaks from these NOE experiments were picked and converted to distance restraints as described in Experimental Procedures. A total of 986 distance restraints, including 475 intraresidue, 210 sequential, 85 medium range, and 216 long range, were obtained after making pseudoatom corrections to restraints involving nonstereospecifically assigned protons. The number of restraints for each residue is shown in Figure 5A. These include restraints to stereospecifically assigned β-methylene protons of 14 residues. In addition to the NOE-derived distance restraints, 43 ϕ, 19 ψ, and 14 χ_1 dihedral restraints and 25 hydrogen bond restraints (two restraints for each hydrogen bond) were included in the calculations.

Calculations. The structure calculations were carried out starting from a linear peptide built in InsightII that included residues G58−E152. Residues Q153−D164 were omitted from the calculations since no medium or long range NOEs were observed in this region of the protein. Initial structures were calculated using unambiguous NOE distance restraints only. These structures were then analyzed for restraint violations and the inevitable bad restraints removed. The resulting structures from these initial calculations were used to assist in assignment of additional NOE cross-peaks. The dihedral restraints were added next, and the structures were recalculated and analyzed for restraint violations. Finally, hydrogen bond restraints were added, and the resulting structures were analyzed further for restraint violations. The final set of DGII calculations produced 83 structures, 62 of which were selected for further refinement on the basis of their DGII error function being less than 1. Structures were refined with minimization and dynamics routine as described in Experimental Procedures. From the 62 minimized structures, 24 were selected using a threshold value for E_{NOE} of 13 kcal/mol. The stereoview of 24 energy-minimized Grb2 SH2 domain structures is shown in Figure 6.

Quality of NMR Structures. Statistics for the 24 calculated structures of the Grb2 SH2 domain (Figure 6) are given in Table 1. After molecular dynamics and minimization, the structures had no NOE distance violations greater than 0.26 Å with most errors less than 0.1 Å. The final structures converged with a rmsd of 0.75 Å for all backbone atoms and 1.28 Å to the average structure for all atoms. Another measure of the quality of the structures is a large negative Lennard-Jones energy potential. The calculated structures of the Grb2 SH2 domain have an average Lennard-Jones energy of -417 ± 6 kcal/mol and suggest good nonbonded atom contacts. The rmsd for backbone or all non-hydrogen

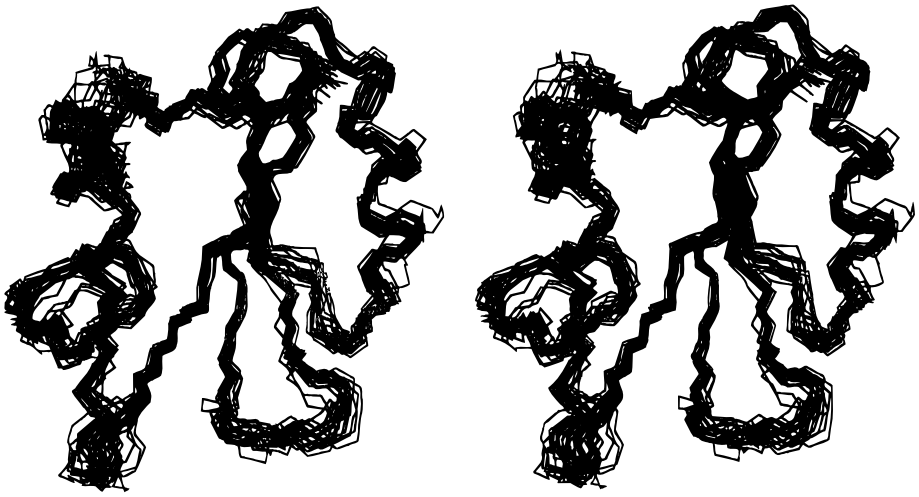

FIGURE 6: Stereoview of 24 overlaid NMR-derived structures of the Grb2 SH2 domain. Residues 60−152 are shown. Structures were superimposed by doing a least squares fit to the average N, Cα, and C′ coordinates of residues 60−152.

Table 1: Statistics for 24 Minimized Structures of the Grb2 SH2 Domain

Average Number of Violations in the 986 Experimental NOE Distance Restraints[a]	
>0.0 and <0.1 Å	90 ± 1.7
>0.1 and <0.2 Å	8.7 ± 1.7
>0.2 and <0.26 Å	0.3 ± 0.5
>0.26 Å	0.0
average maximum violation[b]	0.17 ± 0.03 Å
average sum of violations[c]	4.2 ± 0.2 Å
Average Number of Violations in the 76 Experimental Dihedral Restraints[a]	
>0.0 and <2.2°	9 ± 2
>2.2°	0.0

rmsd (Å) to the Mean Structure[d]	backbone	all heavy atoms
residues 60−152	0.75 ± 0.14	1.28 ± 0.14
all secondary structure[e]	0.51 ± 0.14	1.06 ± 0.12
β-strands βB, βC, and βD	0.23 ± 0.04	0.76 ± 0.1
helices αA and αB	0.42 ± 0.14	0.98 ± 0.14

[a] Average number of violations across an ensemble of 24 calculated structures. [b] Average of the largest violation for each of the 24 calculated structures. [c] Average sum of the violations for each of the 24 converged Grb2 SH2 domain structures. [d] The rmsd was calculated with respect to the mean of the 24 structures of the Grb2 SH2 domain. [e] Residues used in the calculation of the rmsd for regions of secondary structure include F61−F62, R67−Q77, A82−E87, F95−K100, V105−R112, Y118−F119, K124−F125, L128−T138, and R149−D150.

atoms with respect to the mean structure is shown for each residue in Figure 5B. The loop and termini regions are not as well-defined as the core regions of the Grb2 SH2 domain (Figure 5B). If only core regions of the protein are overlaid, the average rmsd is 0.51 and 1.06 Å for backbone and all heavy atoms, respectively. Another method for assessing the quality of a family of NMR-derived structures is based on the angular order parameter (Hyberts et al., 1992). In well-defined regions of a protein, the angular order parameter should be 1. On the other hand, highly disordered regions of a protein are expected to show a random distribution of angles and would have an angular order parameter of 0. The average angular order parameters for residues 60−152 of the Grb2 SH2 domain are shown in Figure 7 for ϕ, ψ, and χ_1 angles. Residues in the Grb2 SH2 domain with low angular order parameters are found primarily in loop regions (Figure 7). The loop regions of the Grb2 SH2 domain are not well-characterized due to few NOE correlations being observed in these regions. These regions are expected to be more flexible, although this cannot be confirmed from the current data. Characterization of the motional properties of the Grb2 SH2 domain would require information on its relaxation properties. Angular order parameters for both the ϕ and ψ angles of residues 60−152 have mean values of 0.91 and suggest a good convergence for these torsion angles.

Dihedral torsion angle violations were less than 2.2° in all structures. As shown in Figure 8 most ϕ and ψ angles fall in the energetically allowed regions of ϕ,ψ space in a Ramachandran map. When residues with angular order parameters of less than 0.9 are excluded, there are virtually no ϕ and ψ angles observed in disfavored regions of ϕ,ψ space. G81 is clustered in the positive ϕ angle region of Figure 8B.

Description of the Structure of the Grb2 SH2 Domain. The topology of the calculated Grb2 SH2 domain structures is similar to that observed for other SH2 domains. In describing the secondary structure of the Grb2 SH2 domain, we have used the notation put forth by Eck et al. (1993). As shown in Figure 9, it consists of a short β-strand (βA) followed by an α-helix (αA). Helix αA is followed by a large three-stranded antiparallel β-sheet (strands βB−βD). The end of strand βD (βD′) forms a smaller antiparallel β-sheet consisting of strands βD′−βF. A second helix (αB) follows βF and lies on a face of the β-sheet opposite that of helix αA. This results in the core β-sheet (strands βB−βD) being flanked by the two helices αA and αB. Following a large loop, BG, the structure ends with another short β-strand βG that forms a parallel sheet with strand βA. Residues that are part of the regions of regular secondary structure are 61−62 (βA), 67−77 (αA), 82−87 (βB), 95−100 (βC), 105−110 (βD), 111−112 (βD′), 118−119 (βE), 124−125 (βF), 128−138 (αB), and 149−150 (βG). Residues 77−80 (AB), 88−91 (BC), and 101−104 (CD) form tight β-turns. Residues 153−164 (residues 159−164 are derived from an expression vector) were observed to be random coil.

The phosphotyrosine binding site of Grb2 is similar to that observed for other SH2 domains (Overduin et al., 1992; Booker et al., 1992; Waksman et al., 1992, 1993; Eck et al., 1993; Lee, 1994; Pascal et al., 1994; Hensmann et al., 1994; Xu et al., 1995; Zhou et al., 1995; Narula et al., 1995; Metzler

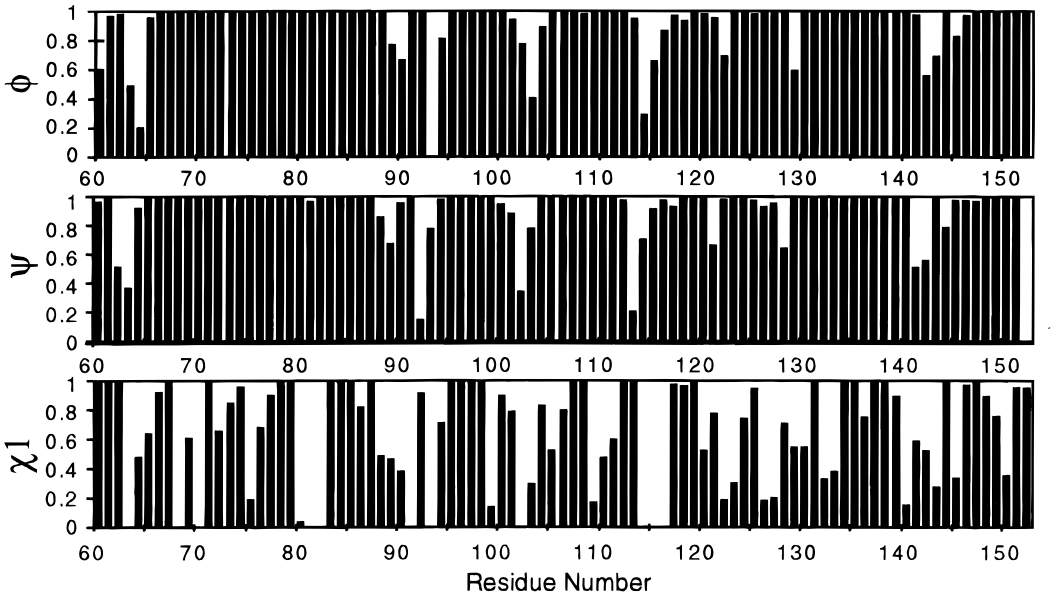

FIGURE 7: Plot of the angular order parameter for the ϕ, ψ, and χ_1 torsion angles of the 24 calculated Grb2 SH2 domain structures as a function of residue number.

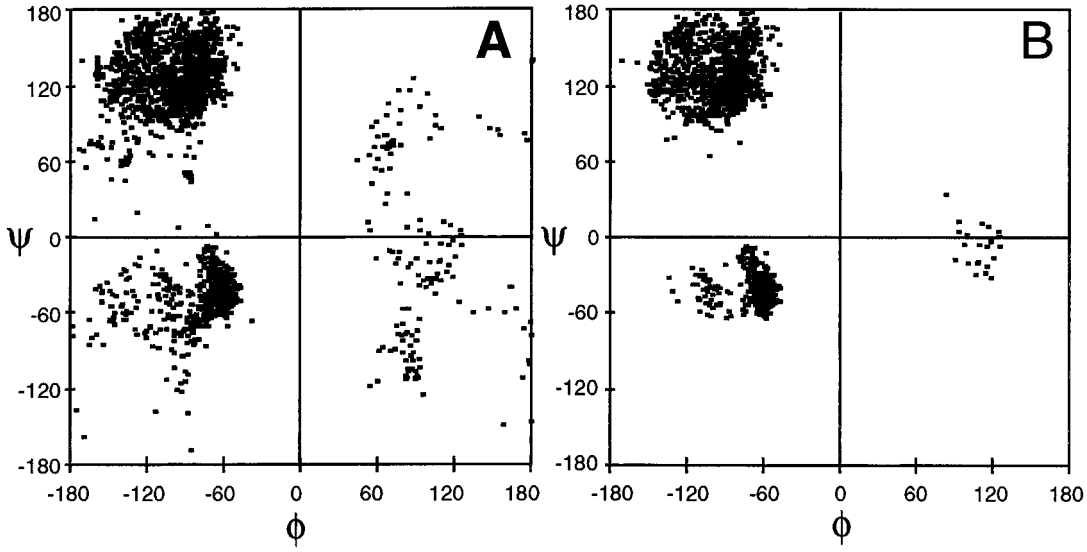

FIGURE 8: Ramachandran plot showing the ϕ and ψ angles of the 24 calculated Grb2 SH2 structures. Plot A shows a ϕ,ψ map of all ϕ and ψ angles. Plot B shows a ϕ,ψ map for ϕ and ψ angles with an angular order parameter greater than 0.9.

et al., 1996). The phosphotyrosine pocket is formed by residues from helix αA (R67) and strands βB (R86), βC (S96), and βD (H107). Residues at these positions are generally well-conserved in the known SH2 domains (Lee et al., 1994). Residues from the BC loop are also predicted to form part of the phosphotyrosine binding pocket. However, in the uncomplexed Grb2 SH2 domain, the BC loop (consisting of residues 88−94) is not in the immediate vicinity of the phosphotyrosine pocket. It is possible that upon complexation with a phosphopeptide the BC loop may move closer to the phosphotyrosine binding pocket and make contact with the phosphopeptide. This hingelike movement of the BC loop has been observed in the case of the lck SH2 domain (Mikol et al., 1995).

The loops BG and EF along with strand βD surround the hydrophobic binding pocket of the SH2 domain of Grb2 (Figure 9). This results in residues F108 (FβD5), L120 (LβE4), W121 (WEF1), and V140 (VBG4) lining the expected pY+3 binding pocket of the Grb2 SH2 domain. Residues at positions βD5, βE4, and BG4 generally show conservative mutations in the known SH2 domains, whereas residues at position EF1 are not conserved (Lee et al., 1994). Grb2 and its homologues sem-5 and Drk are unique among the known SH2 domains in that they are the only SH2 domains to have tryptophan at position EF1 (Lee et al., 1994). In the calculated Grb2 SH2 domain structures, W121 is located in the expected pY+3 binding site as shown in Figure 10A. Preliminary data from our studies of the Grb2 SH2 domain complexed with a target peptide are consistent with the placement of W121 in the hydrophobic binding pocket of Grb2. $^{1}H-^{15}N$ HSQC data (data not shown) show that both the W121 amide and indole protons experience significant chemical shift perturbations upon formation of the Grb2−peptide complex. This suggests that W121 is involved in the Grb2 SH2 domain binding to its target peptide.

Comparison of Phosphotyrosine and Hydrophobic Binding Pockets of Grb2 and Src SH2 Domains. We have compared the structure of the Grb2 SH2 domain with that of Src (Waksman et al., 1993). The overall structures of the Grb2 and Src SH2 (Waksman et al., 1993) domains are quite

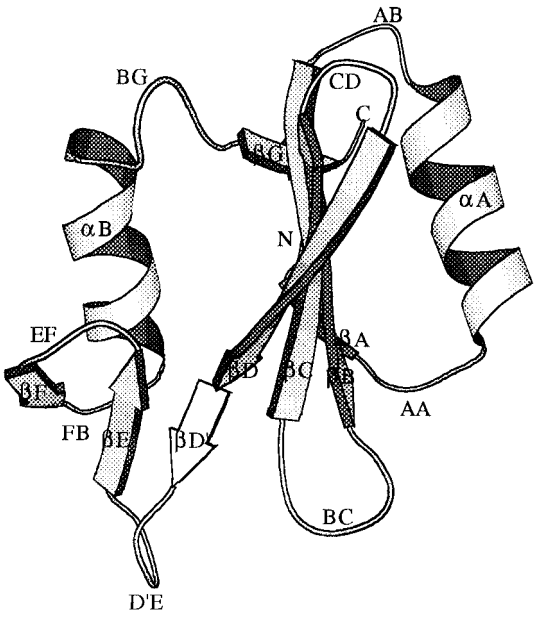

FIGURE 9: Backbone and secondary structure ribbon diagram of the Grb2 SH2 domain. Residues 60−152 are shown. The figure was made using the program Molscript (Kraulis, 1991). The regular secondary structure elements and loop regions are labeled.

similar with the core backbone residues showing a rmsd of 0.86 Å with respect to each other. The phosphotyrosine binding sites of Grb2 and Src are quite similar as would be expected. Residues expected to make contact with phosphotyrosine [R67 (RαA2), R86 (Rβ B5), S96 (Sβ C3), and H107 (Hβ D4)] are in the same positions as their counterparts in Src [R12 (RαA2), R32 (Rβ B5), C42 (Cβ C3), and H58 (Hβ D4)] [Src residues numbered according to Waksman et al. (1993)]. Residues RαA2, Rβ B5, and Hβ D4 are highly conserved among the various SH2 domains.

Although the structure of the phosphotyrosine pocket is quite similar in Grb2 and Src, it is the structure of the hydrophobic pocket that is responsible for differences in binding specificity and molecular recognition. As can be seen in Figure 10, the spacial orientations of residues at positions βD5, βE4, αB8, and BG4 in the Src and Grb2 hydrophobic binding pockets are similar. There are some differences observed in the orientation of the EF and BG loops of Src and Grb2 due to the insertion of extra amino acid residues within the Src EF and BG loops. These extra residues in the Src EF and BG loops do not appear to result in significant differences in the spatial orientation of residues βE4 and BG4 in Grb2 and Src. The primary difference between Grb2 and Src is the placement of tryptophan (colored black) in the hydrophobic binding pocket of Grb2 relative to threonine (colored black) in Src at position EF1. The tryptophan at position EF1 in Grb2 (Figure 10A) takes up considerabely more space than does a threonine in Src (Figure 10B) and results in Grb2 having a smaller pY+3 binding pocket than Src. The placement of tryptophan at the EF1 position could explain the preference of Grb2 for residues that are somewhat smaller than that observed for the Src SH2 domain in the pY+3 pocket (Src prefers isoleucine in the pY+3 pocket while Grb2 prefers valine and proline).

W121 could play an important role in pY+3 selectivity in the Grb2 SH2 domain. Mutation experiments on Src have shown that the residue at position EF1 plays a role in pY+2 and pY+3 selectivity (Marengere et al., 1994). It has been shown that mutation of the residue at position EF1 in the Src SH2 domain from threonine to tryptophan causes Src to have Grb2 SH2-like specificity (Marengere et al., 1994). The location of tryptophan in the pY+3 binding pocket of Grb2 should result in the selection of a smaller residue at the pY+3 site and/or require that the pY+3 residue bind in a different orientation or location within the pY+3 site as compared to that observed for the Src SH2 domain.

The side chain of the residue at position βD3 in SH2 domains plays an important role in the selection of amino acid type at the pY+1 position of the target peptide. In the case of Src, a positively charged residue at position βD3 (lysine) selects for a negatively charged residue (glutamate) at the pY+1 site (Waksman, 1993; Xu et al., 1995). In contrast, peptide screening experiments (Songyang et al., 1994) have shown that Grb2 prefers neutral residues such as glutamine, tyrosine, or valine in the pY+1 position. This

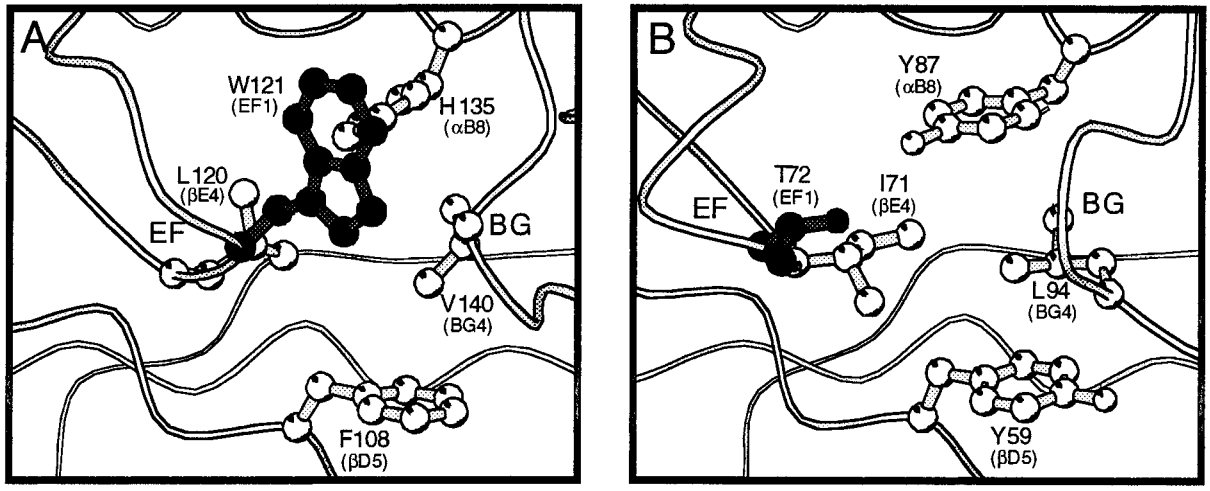

FIGURE 10: Comparison of the Grb2 and SH2 domain hydrophobic binding pockets. Residues expected to influence binding in the hydrophobic binding pocket are labeled. (A) One of the 24 Grb2 SH2 domain structures. The shown structure was selected on the basis of its having the lowest rmsd (0.32 Å) for the regular secondary structure element backbone atoms with respect to the average Grb2 SH2 domain structure. (B) SH2 domain X-ray structure (Waksman et al., 1993). The backbones of the SH2 domains are drawn as white strings. Residues that interact with the pY+3 residue are shown in ball-and-stick mode. Hydrogen atoms were omitted for clarity. The figure was generated with Molscript (Kraulis, 1991).

preference of the Grb2 SH2 domain for a neutral residue at pY+1 is most likely due to the presence of a glutamine at βD3.

Mutation studies on Src (Marengere et al., 1994) have linked the placement of tryptophan at EF1 to a preference for asparagine at the pY+2 position (Marengere et al., 1994). On the basis of the calculated structures of the apo Grb2 SH2 domain, the EF1 tryptophan would not be expected to make contact with the pY+2 residue. In the case of wild type Src, the side chain of the pY+2 residue does not make close contact with the protein (Waksman et al., 1993). However, the high selectivity of Grb2 and mutant Src (threonine mutated to tryptophan at EF1) for asparagine suggests that the pY+2 side chain may be making contact with amino acids in the protein. It is possible that steric interactions in the Grb2 pY+3 pocket, caused by the EF1 tryptophan, may result in the pY+2 residue orienting itself differently than in the case of wild type Src. This may result in the pY+2 asparagine side chain making contact with the protein. Structural studies of the complex of the Grb2 SH2 domain with its target sequence peptide will provide details of these molecular interactions, and such studies are now in progress.

CONCLUSIONS

Chemical shift assignments have been made for the Grb2 SH2 domain ^{1}H, ^{15}N, and ^{13}C resonances. Chemical shift information has been obtained for the apo Grb2 SH2 domain that will provide a basis for the future study of Grb2 SH2 domain−inhibitor complexes. The structure of the Grb2 SH2 domain has been calculated using data obtained from homonuclear and heteronuclear multidimensional NMR experiments. These structures have been calculated with 1112 experimentally obtained restraints. The calculated structures satisfied the experimental restraints and display low restraint violations and a low rmsd. Angular order parameters for ϕ (0.91) and ψ (0.91) torsion angles also indicate a good convergence of the calculated Grb2 SH2 domain structure. The regions of regular secondary structure of the Grb2 SH2 domain were observed to be similar to that observed for other SH2 domains (Overduin et al., 1992; Booker et al., 1992; Waksman et al., 1992, 1993; Eck et al., 1993; Lee, 1994; Pascal, 1994; Hensmann et al., 1994; Xu et al., 1995; Zhou et al., 1995; Narula et al., 1995; Metzler et al., 1996). The hydrophobic binding pocket of Grb2 has been shown to be similar to that observed for Src except for the placement of tryptophan 121 in the pY+3 binding pocket. The NMR assignment and structural information obtained in the present study will lay the foundation for future studies on the molecular interactions of the Grb2 SH2 domain complexed with various inhibitors.

ACKNOWLEDGMENT

The authors thank Tracy Stevenson and Drs. Joseph Loo and Peifeng Hu for mass spectral data and Margie Whitton for amino acid sequencing of the Grb2 SH2 domain. We thank Dr. Michael D. Reily for helpful discussions. The authors also thank one of the reviewers of this paper for helpful suggestions.

NOTE ADDED IN PROOF

During the review of this article, the chemical shift assignments and secondary structure of the Grb2 SH2 domain were reported (Wang et al., 1996)

SUPPORTING INFORMATION AVAILABLE

A table of the ^{1}H, ^{15}N, and ^{13}C NMR assignments for the SH2 domain of Grb2 (9 pages). Ordering information is given on any current masthead page.

REFERENCES

Aaronson, S. A. (1991) *Science 254*, 1146−1153.
Bax, A., & Subramanian, S. (1986) *J. Magn. Reson. 67*, 565−569.
Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., & Kay, L. E. (1990) *J. Magn. Reson. 87*, 620−627.
Bax, A., Ikura, M., Kay, L. E., & Zhu, G. (1991) *J. Magn. Reson. 91*, 174−178.
Bodenhausen, G., & Ruben, D. L. (1980) *Chem. Phys. Lett. 69*, 185−188.
Booker, G. W., Breeze, A. L., Downing, A. K., Panayotou, G., Gout, I., Waterfield, M. D., & Campbell, I. D. (1992) *Nature 358*, 684−687.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166−169.
Buday, L., & Downward, J. (1993) *Cell 73*, 611−620.
Cavanagh, J., & Rance, M. (1992) *J. Magn. Reson. 96*, 670−678.
Chardin, P., Cussac, D., Maignan, S., & Ducruix, A. (1995) *FEBS Lett. 369*, 47−51.
Cowburn, D. (1995) *Structure 3*, 429−430.
Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., & Hagler, A. T. (1988) *Proteins: Struct., Funct., & Genet. 4*, 31−47.
de Dois, A. C., Pearson, J. G., & Oldfield, E. (1993) *Science 260*, 1491−1496.
Downward, J. (1994) *EMBO J. 338*, 113−117.
Eck, M. J., Shoelson, S. E., & Harrison, S. C. (1993) *Nature 362*, 87−91.
Fantl, W. J., Johnson, D. E., & Williams, L. T. (1993) *Annu. Rev. Biochem. 62*, 453−481.
Fesik, S. W., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588−593.
Fry, F. J., Panayotou, G., Booker, G. W., & Waterfield, M. D. (1993) *Protein Sci. 2*, 1785−1797.
Goudreau, N., Cornille, F., Duchesne, M., Parker, F., Tocque, B., Garbay, C., & Roques, B. P. (1994) *Structure 2*, 1029−1040.
Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T. L., Gout, I., & Waterfield, M. D. (1995) *J. Mol. Biol. 248*, 856−866.
Hakes, D. J., & Dixon, J. E. (1992) *Anal. Biochem. 202*, 293−298.
Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43−78.
Hensmann, M., Booker, G. W., Panayotou, G., Boyd, J., Linacre, J., Waterfield, M., & Cambell, I. D. (1994) *Protein Sci. 3*, 1020−1030.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.
Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204−209.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.
Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72−84.
Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888−889.
Kohda, D., Terasawa, H., Ichikawa, S., Ogura, K., Hatanaka, H., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1994) *Structure 2*, 1029−1040.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kuboniwa, H., Grzesiek, S., Delaglio, F., & Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.

Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
Kuriyan, J., & Cowburn, D. (1993) *Curr. Opin. Struct. Biol. 3*, 828−831.
Lee, C. H., Kominos, D., Jacques, S., Margolis, B., Schlessinger, S. E., Shoelson, S. E., & Kuriyan, J. (1994) *Structure 2*, 423−438.
Lee, M. S., Palmer, A. G., & Wright, P. E. (1992) *J. Biomol. NMR, 2*, 307−322).
Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., Bar-Sagi, D., & Schlessinger, J. (1992) *Cell 70*, 431−442.
Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., & Ducruix, A. (1995) *Science 268*, 291−293.
Marengere, L. E. M., Songyang, Z., Gish, G. D., Schaller, M., Parsons, J. T., Stern, M. J., Cantley, L. C., & Pawson, T. (1994) *Nature 369*, 502−505.
Margolis, B., & Skolnik, E. (1994) *J. Am. Soc. Nephrol. 5*, 1288−1299.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
Marion, D., Driscol, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A., & Clore, G. M. (1989a) *Biochemistry 28*, 6150−6156.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989b) *J. Magn. Reson. 85*, 393−399.
Marion, D., Ikura, M., & Bax, A. (1989c) *J. Magn. Reson. 84*, 425−430.
Metzler, W. J., Leiting, B., Pryor, K., Mueller, L., & Farmer, B. T. (1996) *Biochemistry 35*, 6201−6211.
Mikol, V., Baumann, G., Keller, T. H., Manning, U., & Zurini, M. G. (1995) *J. Mol. Biol. 246*, 344−355.
Morris, G. A., & Freeman, R. (1979) *J. Am. Chem. Soc. 101*, 760−762.
Narula, S. S., Yuan, R. W., Adams, S. E., Green, O. M., Green, J., Philips, T. B., Zydowsky, L. D., Botfield, M. C., Hatada, M., Laird, E. R., Zoller, M. J., Karas, J. L., & Dalgarno, D. C. (1995) *Structure 3*, 1061−1073.
Olejniczak, E. T., & Eaton, H. L. (1990) *J. Magn. Reson. 87*, 628−632.
Overduin, M., Mayer, B., Rios, C. B., Baltimore, D., & Cowburn, D. (1992) *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.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
Pronk, G. J., de Vries-Smits, A. M. M., Buday, L., Downward, J., Maassen, J. A., Medema, R. H., & Bos, J. L. (1994) *Mol. Cell. Biol. 14*, 1575−1581.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
Reily, M. D., Thanabal, V., & Omecinsky, D. O. (1992) *J. Am. Chem. Soc. 114*, 6251−6252.
Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R., Li, W., Batzer, A., Thomas, S., Brugge, J., Pelicci, P. G., Schlessinger, J., & Pawson, T. (1992) *Nature 260*, 689−692.
Schlaepfer, D. D., Hanks, S. K., Hunter, T., & van der Geer, P. (1994) *Nature 372*, 786−791.
Schlessinger, J., & Ullrich, A. (1992) *Neuron 9*, 383−391.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169−173.
Sklenar, V., Piotto, M., Leppik, R., & Saudek, V. (1993) *J. Magn. Reson. 102*, 241−245.
Skolnik, E. Y., Batzer, A., Li, N., Lowenstein, L. E., Mohammadi, M., Margolis, B., & Schlessing, J. (1993) *Science 260*, 1953−1955.
Songyang, Z., Shoelson, S. E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X. R., Barbacid, M., Hanafusa, H., Yi, T., Ren, R., Baltimore, D., Ratnofsky, S., Feldman, R. A., & Cantley, L. C. (1994) *Mol. Cell. Biol. 14*, 2777−2785.
Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490−5492.
Terasawa, H., Kohda, D., Hatanaka, H., Tsuchiya, S., Ogura, K., Nagata, K., Ishii, S., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1994) *Nat. Struct. Biol. 1*, 891−898.
Thanabal, V., Omecinsky, D. O., Reily, M. D., & Cody, W. L. (1994) *J. Biomol. NMR 4*, 47−59.
Waksman, G., Kominos, D., Robertson, S. R., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H. M., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L., & Kuriyan, J. (1992) *Nature 358*, 646−653.
Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., & Kuriyan, J. (1993) *Cell 72*, 779−790.
Wang, Y. S., Frederick, A. F., Senior, M. M., Lyons, B. A., Black, S., Kirschmeier, P., Perkins, L. M., & Wilson, O. (1996) *J. Biomol. NMR 7*, 89−98.
Weber, D. J., Gittis, A. G., Mullen, G. P., Abeygunawardana, C., Lattman, E. E., & Mildvan, A. S. (1992) *Proteins: Struct., Funct., & Genet. 13*, 275−287.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., & Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Wüthrich, K., Billetter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
Xu, R. X., Word, M., Davis, D. G., Rink, M. J., Willard, D. H., & Gampe, R. T. (1995) *Biochemistry 34*, 2107−2121.
Zhou, M. M., Meadows, R. P., Logan, T. M., Yoon, H. S., Wade, W. S., Ravichandran, K. S., Burakoff, S. J., & Fesik, S. W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7784−7788.
Zhu, G., & Bax, A. (1990) *J. Magn. Reson. 90*, 405−410.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387−2391.

BI952615S