Volume 324, number 1, 87–92 FEBS 12554 June 1993

^1H and ^{15}N assignments and secondary structure of the Src SH3 domain

Hongtao Yu, Michael K. Rosen and Stuart L. Schreiber

Department of Chemistry, Harvard University, Cambridge, MA 02138, USA

Received 15 April 1993

The ^1H and ^{15}N sequential assignments of the Src SH3 domain have been determined through a combination of 2D and 3D Nuclear Magnetic Resonance (NMR) methods. The secondary structure of the protein has been identified based on long-range NOE patterns. The SH3 domain of Src consists largely of six β-strands that form two anti-parallel β-sheets.

Nuclear magnetic resonance; Secondary structure; SH3 domain; Src

1. INTRODUCTION

Src Homology 3 (SH3) domains consisting of 55 to 75 residues are found in many cytoplasmic signaling molecules [1,2]. SH3 domains of the Src family of protein tyrosine kinases regulate the catalytic functions, the substrate specificities and the transforming potentials of the kinases [3–6]. The existence of SH3 domains in many cytoskeletal proteins suggests that they play important roles in mediating membrane–cytoskeletal interactions [2]. In addition, recent evidence suggests that SH3 domains may link tyrosine kinase and Ras signaling pathways. A protein (3BP-1) that binds to the Abl and Src SH3 domains in vitro is similar in sequence to the guanosine triphosphatase (GTPase) activation domain of several GTPase activating proteins (GAPs) [7]. Further, SH3-containing adaptor proteins such as GRB2/SEM5 may interact with guanine nucleotide exchange factors (GEFs) [1,8,9].

The X-ray structure of the spectrin SH3 domain and solution structure of the Src SH3 domain have been reported previously [10,11]. The three dimensional structure of the Src SH3 was determined by simulated annealing, and the tertiary fold of this domain is shown in Fig. 1. We now report the ^1H and ^{15}N assignments and secondary structure analysis of the Src SH3 domain. A set of 2D and 3D homonuclear and heteronuclear experiments were performed for the purpose of sequential assignment. The secondary structure of the c-Src SH3 domain is derived from the NOE data and $^3J_{NH\text{-}C\alpha H}$ vicinal coupling constants obtained by a variety of homonuclear and heteronuclear methods.

Correspondence address S.L. Schreiber, Department of Chemistry, Harvard University, Cambridge, MA 02138, USA.

2. MATERIALS AND METHODS

Recombinant Src SH3 domain was obtained as a glutathione S-transferase fusion protein by overexpression in *E. coli* strain XA90 of the SH3 cDNA contained in the plasmid pGEX-2T (Pharmacia). The fusion protein was purified on a glutathione-agarose affinity column and cleaved by digestion with thrombin. The cleaved mixture was purified by gel filtration chromatography. The sequence of the protein obtained by this procedure is shown in Fig. 1. The N-terminal four residues (GSHM) are derived from the fusion construct. The NMR sample contained 3.5 mM protein in 100 mM phosphate (pH 6.0) and 100 mM NaCl. D_2O samples were prepared by lyophilizing the protein from H_2O and dissolving it in 99.9% D_2O buffer to the desired concentration. Uniformly ^{15}N-labeled Src SH3 was produced by growing the bacteria in M9 minimal media containing $^{15}NH_4Cl$ (Isotech) as the sole nitrogen source.

All NMR data were recorded at 298K using either Bruker AMX600 or AM500 instruments. For the 2D experiments, quadrature detection in the t_1 dimension was achieved through Time Proportional Phase Incrementation (TPPI), and normally 450 to 512 t_1 increments were recorded. Two-dimensional homonuclear double quantum filtered (DQF)-COSY, double quantum (DQ, τ = 32 ms), Relay-COSY (τ = 32 ms), total correlation spectroscopy (TOCSY, τ = 30 ms or 70 ms), and nuclear Overhauser effect spectroscopy (NOESY, τ = 50 ms and 150 ms) were taken in both H_2O and D_2O using standard pulse sequences. For the 3D ^{15}N TOCSY-HMQC [12] and NOESY-HMQC [13] experiments, the States-TPPI method was used to achieve quadrature detection in the indirectly detected dimensions. In both cases, $512 \times 128 \times 32$ complex data points were recorded with folding in the ^{15}N dimension to increase digital resolution. The NMR data were processed with the software package FELIX with appropriate apodization, base line correction and zero-filling [14].

3. RESULTS AND DISCUSSION

3.1. *Spin system identification*

The sequential assignment was performed in two stages: spin system identification and through-space sequential connection of the spin systems [15,16]. The preliminary spin system database was constructed by peak-picking from either the 3D TOCSY-HMQC experiment or the 2D TOCSY experiments. The magnetization transfer in TOCSY experiments was sufficient to provide most of the expected crosspeaks between amide

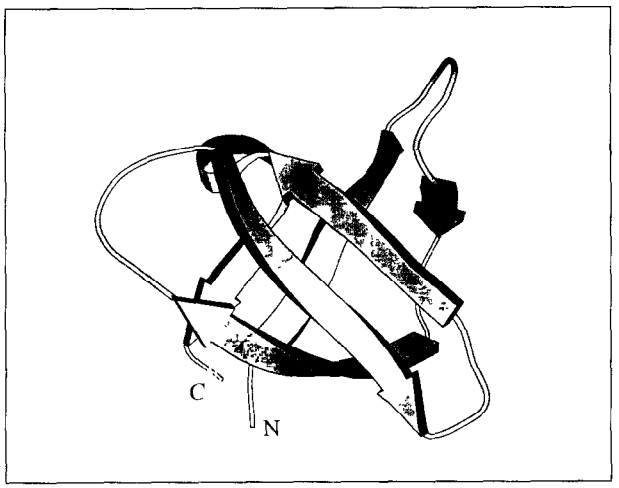

Fig. 1. Schematic drawing of the tertiary fold of the Src SH3 domain. The N and C termini are labeled 'N' and 'C', respectively. The diagram was generated using the program MOLSCRIPT [18].

NHs and their corresponding sidechain protons. The subsequent grouping of these cross-peaks based on their amide NH and ^{15}N chemical shifts gave rise to the spin systems, which were then tentatively assigned to amino acid types based on sidechain proton chemical shifts and through-bond coupling patterns.

The alanine and threonine assignments were confirmed by observation of strong crosspeaks between methyl protons and $C_\alpha H/C_\beta H$, respectively, in the DQF-COSY spectrum. Threonines were distinguished from the alanines by the existence of the strong $C_\gamma H_3/C_\alpha H$ crosspeaks in the relay-COSY spectrum. All of the seven glycines showed remote crosspeaks at ω_2 = NH and $\omega_1 = C_{\alpha 1}H + C_{\alpha 2}H$ in a DQ spectrum recorded in H_2O. The glycine assignments were also confirmed by their high field amide ^{15}N chemical shifts. By analyzing the patterns of the strong methyl coupling peaks in the D_2O DQF-COSY, DQ and Relay-COSY spectra, all the methyl resonances belonging to valine, isoleucine and leucine residues were tentatively assigned and were correlated to their amide proton chemical shifts in the TOCSY experiments. Other non-characteristic spin systems were divided into three subgroups: three-spin residues (Asp, Asn, Ser, Phe, Tyr, Trp, and His), five-spin residues (Glu, Gln, and Met), and long residues (Arg, Pro and Lys). The ring protons of the aromatic residues showed NOEs to their corresponding $C_\beta Hs$. The sidechain amide NH of the asparagines and glutamines also showed NOE connections to their own $C_\beta Hs$ and $C_\gamma Hs$. The sidechain NH of arginine showed the same cross-peaks as the backbone NH in the TOCSY experiments. These NOESY or TOCSY correlations allowed us to assign the four tyrosines, two arginines, two phenylalanines, two tryptophans, three asparagines and two glutamines to their individual residue types. Tentative assignments of the serine residues were also made based on their slightly lower field $C_\beta H$ chemical shifts. The remaining spin systems, including the prolines, were only identified at the sequential assignment stage.

3.2. *Sequential assignments*

The sequential assignment of this protein was largely based on $C_\alpha H_i$–NH_{i+1}, NH_i–NH_{i+1}, and $C_\beta H_i$–NH_{i+1} NOE connectivities. Long stretches of peptide segments could easily be traced out with the 3D NOESY-HMQC starting from unique residues. Fig. 2 shows the NOE connectivities from Asp-41 to Ser-47. Fig. 3 shows all the connections used to make the sequential assignment. The sequential connections were greatly facilitated by the completeness of the spin system identification. The sequential assignment in the β-sheet region was also confirmed by interstrand long-range NOE connectivities. Stereospecific assignments of some methylene $C_\beta Hs$ and valine methyl protons were based on coupling constants and cross-peak magnitudes in a DQF-COSY spectrum, and on intra-residue NOEs from amide NH and $C_\alpha H$ [17]. The ^{15}N and ^1H chemical shifts of the Src SH3 domain are shown in Table I.

3.3. *Secondary structure analysis*

The main secondary structural elements of the Src SH3 domain are two three-stranded anti-parallel β sheets. One β sheet consists of strands S1, S2 and S6,

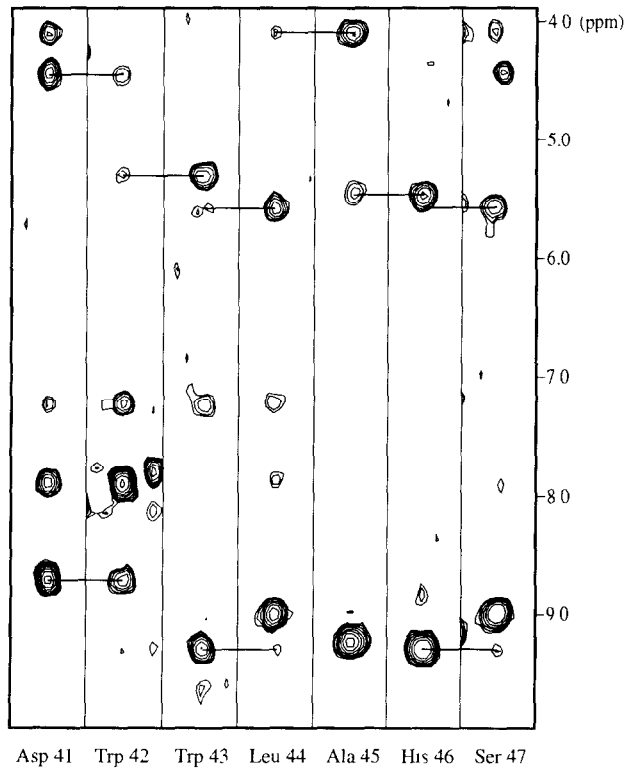

Fig. 2. Three-dimensional NOESY-HMQC strip plots for the region from Asp-41 to Ser-47. Each strip was taken from the 3D spectrum at the ^{15}N chemical shift of an assigned ^{15}N resonance. Solid lines indicate the $d_{\alpha N}$ and d_{NN} connectivities through this region.

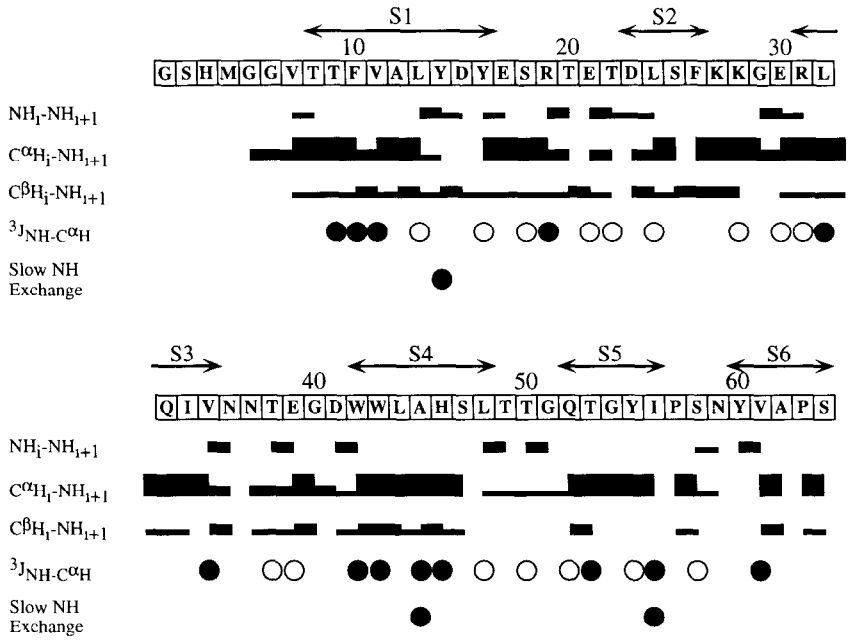

Fig. 3. Summary of sequential NOEs, $^{3}J_{NH\text{-}C\alpha H}$ coupling constants and slowly exchanging amide protons The magnitudes of $^{3}J_{NH\text{-}C\alpha H}$ coupling constants were estimated from the intensities of the NH/C_αH crosspeaks in a DQF-COSY spectrum. Slowly exchanging amide protons were identified from a 2D ^{1}H-^{15}N HSQC spectrum recorded 3 h after dissolving the lyophilized protein in D_2O buffer The widths of the bars indicate the intensities of the corresponding NOEs (strong, medium or weak). Small $^{3}J_{NH\text{-}C\alpha H}$ coupling constants are shown as open circles, while large $^{3}J_{NH\text{-}C\alpha H}$ coupling constants are shown as closed circles. Slowly exchanging amide protons are also represented by closed circles. Above the amino acid sequence, the proposed secondary structural elements are shown; S stands for β-strand.

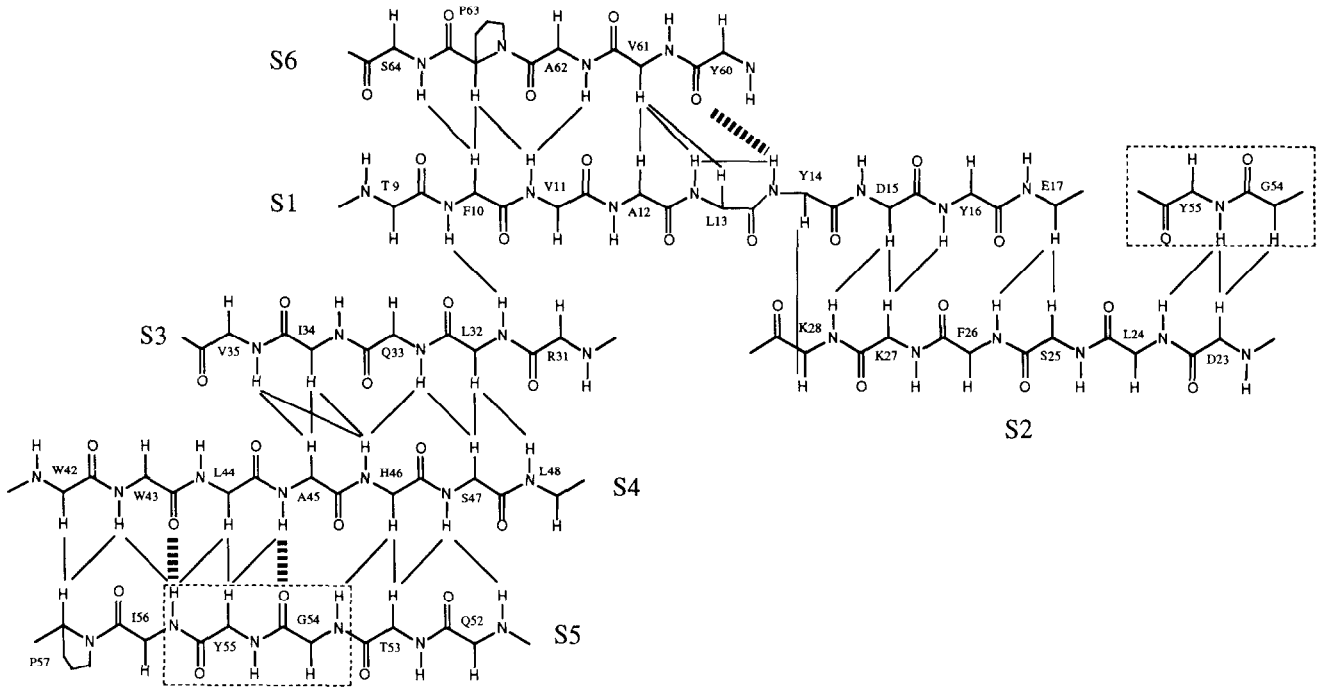

Fig. 4. A schematic diagram of the inter-strand NOEs in the β-sheet regions of the Src SH3 domain. Solid lines indicate inter-strand NOEs and dashed lines show deduced hydrogen bonds from slowly exchanging amide protons and long-range NOEs. The region from Gly-54 to Tyr-55 in strand S5 which is enclosed by a dashed square forms an antiparallel β-sheet with the strand S4 and is associated with the strand S2 in parallel manner.

Table 1: ^{1}H and ^{15}N Chemical Shifts[a] of Recombinant Src SH3 Domain

Residue	amide ^{15}N	amide ^{1}H	$C^{\alpha}H$	$C^{\beta}H$ proS	$C^{\beta}H$ proR	other
Gly 1						
Ser 2						
His 3	129 7	8.07	4.53	2.87, 2.70		$C^{\delta 2}H$ 8.25; $C^{\varepsilon 1}H$ 7.23
Met 4		8.56	4.08	2.20, 2.07		$C^{\gamma}H$ 2.65, 2.58
Gly 5	111.7	8.38	4.60, 4.60			
Gly 6	110.3	8.38	4.11, 4.04			
Val 7	120.8	8.10	4.35	2.15		$C^{\gamma}H$ 1.00, 0.93
Thr 8	120.8	8.57	4.60	4.35		$C^{\gamma}H$ 1.22
Thr 9	121.1	8.45	5.11	4.13		$C^{\gamma}H$ 1.28
Phe 10	128.2	9.49	5.10	*2.96	2.85	$C^{\delta}H$ 7.11; $C^{\varepsilon}H$ 7 48; $C^{\zeta}H$ 7.41
Val 11	120.7	9.64	5.19	1.91		$C^{\gamma}H$ 0.95, 0.93
Ala 12	128.4	8.66	4.73	1.83		
Leu 13	127.6	9 58	4.09	*0.69	1.06	$C^{\gamma}H$ 1.49; $C^{\delta}H$ 0.78
Tyr 14	113 7	7 28	5.00	*3.43	2.58	$C^{\delta}H$ 6.74; $C^{\varepsilon}H$ 7.28
Asp 15	119.3	8.40	4.89	2.90, 2.76		
Tyr 16	122.9	8.75	4.63	2.42, 2.42		$C^{\delta}H$ 7.14; $C^{\varepsilon}H$ 6.76
Glu 17	129 6	7.56	4.40	1.83, 1.75		$C^{\gamma}H$ 2.20, 2.20
Ser 18	119.7	8.27	4.21	3.96, 3.76		
Arg 19	126.4	9.81	4.60	*2.06	1.88	$C^{\gamma}H$ 1.85, 1.78; $C^{\delta}H$ 3.31; $N^{\varepsilon}H$ 7.32; N^{ε} 85 8
Thr 20	115 4	8.54	4.86	4.40		$C^{\gamma}H$ 1.19
Glu 21	121.6	8 99	4.46	2.30, 2.25		$C^{\gamma}H$ 2.51
Thr 22	107.6	7.88	4.73	4.73		$C^{\gamma}H$ 1.35
Asp 23	124.8	8.03	5.58	3.29, 3.09		
Leu 24	125 7	8.77	4 72	*1.07	1.92	$C^{\gamma}H$ 1.76; $C^{\delta}H$ 0.95, 0.78
Ser 25	118.9	8.11	5.00	*4.01	4.17	
Phe 26	118.8	8.75	5.23	*3.26	3.03	$C^{\delta}H$ 7.15; $C^{\varepsilon}H$ 7.53; $C^{\zeta}H$ 6.94
Lys 27	121.5	8.81	5.18	2.08, 1 95		$C^{\gamma}H$ 1.63
Lys 28	122 7	9.33	3 57	*1.80	1.68	$C^{\gamma}H$ 1.30; $C^{\varepsilon}H$ 3.13
Gly 29	116.3	9.10	4.55, 3.62			
Glu 30	125.0	8 48	4.27	2.58, 2.25		$C^{\gamma}H$ 2 77, 2.43
Arg 31	123.4	8.20	5.19	*1.78	1.97	$C^{\gamma}H$ 1.90, 1 67; $C^{\delta}H$ 3.30; $N^{\varepsilon}H$ 7.31; N^{ε} 86 5
Leu 32	125.3	9.32	5.17	*1.21	1.43	$C^{\gamma}H$ 1.51, $C^{\delta}H$ 0.69, 0.28
Gln 33	123.2	9.25	5.13	2.21, 2.17		$C^{\gamma}H$ 2.49, 2.39, $N^{\varepsilon 2}H$ 7.58, 6.96; $N^{\varepsilon 2}$ 113.0
Ile 34	127.6	9 20	4.40	2.01		$C^{\gamma 2}H$ 0.54; $C^{\gamma 1}H$ 1.55, 0.92; $C^{\delta}H$ 0.59
Val 35	128.8	8.85	4.08	1.79		*$C^{\gamma}HS$ 0.87; $C^{\gamma}HR$ 0.96
Asn 36	117.8	7.86	4.92	2.98, 2.91		$N^{\delta 2}H$ 7.84, 7 18; $N^{\delta 2}$ 114.9
Asn 37	119.7	8.48	4.60	2.05, 1.72		$N^{\delta 2}H$ 6.74, 6.52; $N^{\delta 2}$ 111.4
Thr 38	113.0	8.03	4.04	4.32		$C^{\gamma}H$ 1.14
Glu 39	123.2	8.62	4.70	*2.42	2 31	$C^{\gamma}H$ 2.49, 2.24
Gly 40	110.1	8.50	4.42, 4.13			
Asp 41	120.2	8 69	4.51	2.64, 2 59		
Trp 42	122.4	7.90	5.32	3.07, 3 07		$C^{\delta 1}H$ 7.25; $N^{\varepsilon 1}H$ 10.08; $N^{\varepsilon 1}$ 130.5; $C^{\varepsilon 3}H$ 7.31; $C^{\eta}H$ 7.41; $C^{\zeta 2}H$ 7.52; $C^{\zeta 3}H$ 6.92
Trp 43	125.9	9.28	5.60	*3.10	2.99	$C^{\delta 1}H$ 7.25; $N^{\varepsilon 1}H$ 9.76; $N^{\varepsilon 1}$ 130.6, $C^{\varepsilon 3}H$ 6.90; $C^{\eta}H$ 7.18; $C^{\zeta 2}H$ 7 62; $C^{\zeta 3}H$ 5.99

(continued on page 91)

(Table 1 continued)

Table 1: ^{1}H and ^{15}N Chemical Shifts[a] of Recombinant Src SH3 Domain

Residue	amide ^{15}N	amide ^{1}H	$C^{\alpha}H$	$C^{\beta}H$ proS	$C^{\beta}H$ proR	other
Leu 44	126.8	8.99	4.15	1.23, 1.03		$C^{\gamma}H$ 1.14; $C^{\delta}H$ 0.76, 0.27
Ala 45	132.7	9.25	5.48	1.37		
His 46	119.1	9.29	5.60	*3.17	2.96	$C^{\delta 2}H$ 8.31; $C^{\varepsilon 1}H$ 6.95
Ser 47	121.8	8.99	4.82	*4.15	4.07	
Leu 48	131.3	8.99	4.39	1.88		$C^{\gamma}H$ 1.80; $C^{\delta}H$ 1.03, 0.79
Thr 49	117.2	8.51	4.41	4.10		$C^{\gamma}H$ 1.44
Thr 50	110.3	8.42	4.60	4.60		$C^{\gamma}H$ 1.43
Gly 51	112.4	7.91	4.37, 3.93			
Gln 52	120.7	7.61	4.63	*2.30	2.03	$C^{\gamma}H$ 2.53; $N^{\varepsilon 2}H$ 7.59, 7.06; $N^{\varepsilon 2}$ 113.5
Thr 53	118.8	8.72	5.80	4.09		$C^{\gamma}H$ 1.24
Gly 54	114.2	9.23	4.13, 4.13			
Tyr 55	121.4	9.11	5.46	*3.15	3.11	$C^{\delta}H$ 7.17; $C^{\varepsilon}H$ 6.99
Ile 56	114.3	9.60	5.23	1.67		$C^{\gamma 2}H$ 1.13; $C^{\gamma 1}H$ 1.44, 1.11; $C^{\delta}H$ 0.48
Pro 57			3.75	1.53, 1.16		$C^{\gamma}H$ 0.89, 0.74; $C^{\delta}H$ 2.96, 2.90
Ser 58	121.1	7.80	2.66	*2.11	1.76	
Asn 59	116.7	8.13	4.76	3.08, 2.81		$N^{\delta 2}H$ 7.59, 6.79; $N^{\delta 2}$ 113.3
Tyr 60	122.7	8.04	4.85	3.74, 3.49		$C^{\delta}H$ 7.26; $C^{\varepsilon}H$ 7.08
Val 61	110.4	7.28	5.53	2.24		*$C^{\gamma}HS$ 0.56; $C^{\gamma}HR$ 0.73
Ala 62	124.2	8.92	5.21	1.53		
Pro 63			3.69	2.03, 1.91		$C^{\gamma}H$ 2.18, 1.30; $C^{\delta}H$ 3.80
Ser 64	123.4	7.93	4.23	3.83, 3.65		

[a] Chemical shifts are reported for the Src SH3 domain at pH 6.0 and 298K in buffer containing 50 mM sodium phosphate and 150 mM sodium chloride. Proton chemical shifts are referenced to H_2O at 4.86 ppm. Nitrogen chemical shifts are referenced to external $^{15}NH_4Cl$ (5M) in 1M aqueous HCl at 26.1 ppm.

* indicates stereospecific assignments.

while the other sheet is composed of strands S3, S4 and S5. The observed interstrand NOEs are shown in Fig. 4. The Src SH3 domain contains only three slowly exchanging amide protons even though 35 of the 64 residues are involved in β strands. In contrast to Src SH3, PI3K SH3 has 38 slowly exchanging amide protons at the same pH (see accompanying manuscript). The absence of many expected slowly exchanging amide protons suggests that the structure of the Src SH3 domain might possess somewhat unusual dynamic properties that could facilitate proton exchange under our experimental conditions. Therefore Src SH3, together with PI3K SH3, could provide us with an interesting model system to study protein folding pathways, protein dynamics and protein stabilities in solution. Another interesting feature of the Src SH3 domain is that a peptide segment (Thr-22 to Leu-24) in β strand S2 associates with a portion of strand S5 (residue 54 to 55) in parallel manner, as evidenced by the strong NOE cross-peak between the amide NH of Tyr-55 and the $C_{\alpha}H$ of Asp-23.

Comparison of our secondary structure analysis reported herein and of the three dimensional fold reported elsewhere [11] (Fig. 1) of the Src SH3 domain with the X-ray structure of the spectrin SH3 domain shows a significant difference between the two structures with regard to the interstrand connection between β-strands S1 and S3 [10]. The structure of the spectrin SH3 domain was described as a continuous five-stranded antiparallel β-sheet. However, we were unable to observe many of the interstrand NOEs (e.g. $C_{\alpha}H_i$–$C_{\alpha}H_j$, $C_{\alpha}H_i$–NH_j, etc.) between S1 and S3 in our NMR studies. This may not reflect a genuine difference between the two structures. The C_{α}Hs of both of the residue pairs that would be expected to show defining interstrand NOEs are degenerate; the C_{α}Hs of Val-11 and Arg-31 are at

5.19 ppm, and those of Thr-9 and Gln-33 are at approximately 5.12 ppm. Thus, interstrand $C_\alpha H_i$–NH_j NOEs from the amide NHs of Phe-10 and Leu-32 could not be distinguished from sequential $C_\alpha H_i$–NH_{i+1} NOEs from these protons. These difficulties, along with the absence of slowly exchanging amide protons, make it impossible to unambiguously confirm the association of S1 and S3 using our current data. A ^{13}C dispersed NOESY spectrum of suitably labeled material could help to resolve this discrepancy.

Although the sequences of SH3 domains are highly divergent, they seem to have the same overall molecular structure. The differences found among these domains may contribute to their ligand specificities in mediating the cellular interactions of different signaling molecules.

Acknowledgements We thank C. Seidel-Dugan and J.S. Brugge (ARIAD Pharmaceuticals, Inc.) for providing the Src SH3 domain cDNA; D.C Dalgarno for assisting in the acquisition of several NMR data sets, and Bruker Instruments for providing instrument time This research was supported by fellowships from the American Chemical Society and Merck Sharp & Dohme Research Laboratories (to M.K.R.) and grant from the National Institute of General Medical Sciences (GM44993, to S.L S.).

REFERENCES

[1] Pawson, T. and Gish, G.D (1992) Cell 71, 359–362.
[2] Musacchio, A., Gibson, T., Lehto, V.P. and Saraste, M. (1992) FEBS Lett. 307, 55–61.
[3] Potts, W.M., Reynolds, A.B., Lansing, T.J. and Parsons, J.T. (1988) Oncogene Res 3, 343–355.
[4] Nemuth, S.P., Fox, L.C., DeMarco, M. and Brugge, J S. (1989) Mol. Cell. Biol. 9, 1109–1119.
[5] Seidel-Dugan, C., Meyer, B.E., Thomas, S.M. and Brugge, J.S. (1992) Mol. Cell. Biol. 12, 1835–1845.
[6] Kanner, S B., Reynolds, A.B., Wang, H.-C.R., Vines, R R. and Parsons, J.T (1991) EMBO J. 10, 1689–1698.
[7] Cicchetti, P., Mayer, B.J., Thiel, G. and Baltimore, D. (1992) Science 257, 803–806.
[8] Williams, L.T (1992) Curr Biol 2, 601–603.
[9] Lowenstein, E J., Daly, R.J., Datzer, A.G., Li, W., Margolis, B., Lammers, R , Ullrich, A.L., Skolnik, F.Y , Bar-Sag, H. and Schlessinger, J. (1992) Cell 70, 431–442.
[10] Musacchio, A., Noble, M., Pauptit, R., Wierenga, R and Saraste, M. (1992) Nature 359, 851–855.
[11] Yu, H., Rosen, M.K , Shin, T B., Seidel-Dugan, C , Brugge, J.S and Schreiber, S.L. (1992) Science 258, 1665–1668
[12] Marion, D , Driscoll, P.C , Kay, L.E , Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) Biochemistry 28, 6150–6156.
[13] Zuiderweg, E R.P. and Fesik, S.W. (1989) Biochemistry 28, 2387–2391.
[14] Hare, D. (1991) FELIX manual (Version 2.0).
[15] Chazin, W.J., Rance, M. and Wright, P E. (1988) J. Mol. Biol. 202, 603–622.
[16] Chazin, W.J. and Wright, P.E. (1988) J. Mol Biol. 202, 623–636.
[17] Wagner, G., Braun, W., Havel, T F , Schaumann, T., Go, N. and Wuthrich, K. (1987) J. Mol. Biol. 196, 611–639.
[18] Kraulis, P J (1991) J. Appl Cryst. 24, 946–950.