Journal of Biomolecular NMR, 8 (1996) 105–122
ESCOM

J-Bio NMR 377

Determination of the solution structure of the SH3 domain of human p56 Lck tyrosine kinase

Hidekazu Hiroaki*, Werner Klaus** and Hans Senn

Pharmaceutical Research, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

Received 26 April 1996
Accepted 27 June 1996

Keywords: SH3; Protein structure; Tyrosine kinase; p56 Lck

Summary

The solution structure of the SH3 domain of human p56 Lck tyrosine kinase (Lck-SH3) has been determined by multidimensional heteronuclear NMR spectroscopy. The structure was calculated from a total of 935 experimental restraints comprising 785 distance restraints derived from 1017 assigned NOE cross peaks and 150 dihedral angle restraints derived from 160 vicinal coupling constants. A novel combination of the constant-time ^{1}H-^{13}C NMR correlation experiment recorded with various delays of the constant-time refocusing delays and a fractionally ^{13}C-labelled sample was exploited for the stereospecific assignment of prochiral methyl groups. Additionally, 28 restraints of 14 identified hydrogen bonds were included. A family of 25 conformers was selected to characterize the solution structure. The average root-mean-square deviations of the backbone atoms (N, C^{α}, C', O) among the 25 conformers is 0.42 Å for residues 7 to 63. The N- and C-terminal residues, 1 to 6 and 64 to 81, are disordered, while the well-converged residues 7 to 63 correspond to the conserved sequences of other SH3 domains. The topology of the SH3 structure comprises five anti-parallel β-strands arranged to form two perpendicular β-sheets, which are concave and twisted in the middle part. The overall secondary structure and the backbone conformation of the core β-strands are almost identical to the X-ray structure of the fragment containing the SH2–SH3 domains of p56 Lck [Eck et al. (1994) *Nature*, **368**, 764–769]. The X-ray structure of the SH3 domain in the tandem SH2–SH3 fragment is spatially included within the ensemble of the 25 NMR conformers, except for the segment of residues 14 to 18, which makes intermolecular contacts with an adjacent SH2 molecule and the phosphopeptide ligand in the crystal lattice. Local structural differences from other known SH3 domains are also observed, the most prominent of which is the absence in Lck-SH3 of the two additional short β-strands in the regions Ser^{15} to Glu^{17} and Gly^{25} to Glu^{27} flanking the so-called 'RT-Src' loop. This loop (residues Glu^{17} to Leu^{24}), together with the 'n-Src' loop (residues Gln^{37} to Ser^{46}) confines the ligand interaction site which is formed by a shallow patch of hydrophobic amino acids (His^{14}, Tyr^{16}, Trp^{41}, Phe^{54} and Phe^{59}). Both loops are flexible and belong to the most mobile regions of the protein, which is assessed by the heteronuclear ^{15}N,^{1}H-NOE values characterizing the degree of internal backbone motions. The aromatic residues of the ligand binding site are arranged such that they form three pockets for interactions with the polyproline ligand.

Introduction

The human p56 Lck tyrosine kinase is a T-lymphocyte-specific member of the Src family and regulates early events of the signal-transduction process leading to T-cell activation and proliferation (Veillette et al., 1988; Ostergaard et al., 1989; Rudd et al., 1989; Veillette et al., 1989; Klausner and Samelson, 1991; Sefton, 1991). This signal transduction is a highly specific event, which includes specific molecular recognition processes in which SH2 and SH3 domains play a key role. Members of the Src family of tyrosine kinases each contain unique N-terminal sequences followed by well-conserved domains, the regulatory Src homology domains, SH2 and SH3, and the

*Present address: Department of Medicinal Chemistry, Nippon Roche Research Center, Kamakura, Japan.
**To whom correspondence should be addressed.

Abbreviations: CT, constant time; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; SH2, Src homology domain 2; SH3, Src homology domain 3.

TABLE 1
OVERVIEW OF THE ISOTOPIC COMPOSITION OF SH3 AND THE SOLVENT USED WITH THE DIFFERENT SAMPLES IN THE NMR EXPERIMENTS

Pulse sequence	Solvent	Purpose	Reference
Sample 1: unlabelled			
2D clean-TOCSY (mixing time 40 ms)	H_2O	spin-system identification	Griesinger et al., 1988
2D NOESY (mixing time 100 ms)	H_2O	sequential assignment	Anil Kumar et al., 1980
2D DQF-COSY	H_2O	spin-system identification	Rance et al., 1983
2D DQF-COSY	D_2O	spin-system identification	
Sample 2: 99% ^{15}N-labelled			
2D ^{15}N-{^{1}H}-NOE	H_2O	protein backbone dynamics	Grzesiek and Bax, 1993
2D HNHA	H_2O	coupling constant $^3J_{HNH\alpha}$	Kuboniwa et al., 1994
3D HNHB	H_2O	coupling constant $^3J_{NH\beta}$	Archer et al., 1991; Bax et al., 1994
Sample 3: 99% ^{15}N-, 12% ^{13}C-labelled ('fractionally')			
2D ^{15}N HSQC	H_2O	assignment of N-H	Grzesiek and Bax, 1993
2D ^{15}N HSQC (H-D exchange)	D_2O	solvent accessibility	Grzesiek and Bax, 1993
2D ^{13}C CT-HSQC	D_2O	stereospecific assignment of Val-/Leu-methyl groups	Vuister and Bax, 1992; Neri et al., 1989; Senn et al., 1989
3D ^{15}N TOCSY-HSQC (mixing time 40 ms)	H_2O	spin-system identification, assignment of side-chain protons	Marion et al., 1989a; Driscoll et al., 1990
3D ^{15}N NOESY-HSQC (mixing time 100 ms)	H_2O	distance information	Marion et al., 1989a,b
Sample 4: 99% ^{15}N-, 99% ^{13}C-labelled			
2D ^{13}C CT-HSQC	D_2O	assignment of C-H	Vuister and Bax, 1992
3D ^{13}C HCCH-TOCSY (mixing time 15.2 ms)	D_2O	assignment of side-chain protons	Bax et al., 1990
3D ^{13}C CT-HSQC-NOESY (mixing time 100 ms)	D_2O	distance information	Ikura et al., 1990
3D CBCANH	H_2O	sequential assignment	Grzesiek and Bax, 1992a
3D CBCA(CO)NH	H_2O	sequential assignment	Grzesiek and Bax, 1992b
3D HN(CO)HB	H_2O	coupling constant $^3J_{C\beta H\beta}$	Grzesiek et al., 1992; Bax et al., 1994

Compilation of NMR experiments performed with the unlabelled and the three differently labelled samples of the SH3 domain of p56 Lck. All measurements were done at a temperature of 303 K and a pH of 6.5.

C-terminal catalytic tyrosine kinase domain (Pawson, 1995).

SH2 and SH3 domains are not restricted to particular types of signal-transduction proteins. They occur also in lipid kinases, protein phosphatases, phospholipases, Ras-controlling proteins, and even in transcription factors, although they are rarely found in receptors. They are a part of non-enzymatic adapter proteins which solely act to aggregate other proteins. Furthermore, SH3 domains are seen in cytoskeleton proteins, where they may mediate the action of signal transduction on cellular architecture and cell movement. While SH2 domains have a tight binding affinity to their specific target sequences containing phosphotyrosine residues (Songyang et al., 1993), SH3 domains associate preferentially with proline-rich peptides which contain a PXXP motif for high-affinity (micromolar) binding (Feng et al., 1994; Alexandropoulos et al., 1995; Cohen et al., 1995).

Several reports on the determination of the three-dimensional structure for various SH3 protein modules have been published so far. They include the X-ray analyses of the SH3 domains of chicken α-spectrin (Musacchio et al., 1992), human Fyn (Noble et al., 1993), human Csk (Borchert et al., 1994), human Abl (Musacchio et al., 1994), Grb2-N (Guruprasad et al., 1995) and the peptide complexes of human Fyn and Abl (Musacchio et al., 1994), Sem-5-C (Lim et al., 1994) and c-Crk-N (Wu et al., 1995). NMR studies are reported for chicken Src (Yu et al., 1992), human PLC-γ (Kohda et al., 1993), GAP (Yang et al., 1994) and the P85 subunits of human (Koyama et al., 1993; Yu et al., 1994) and bovine (Booker et al., 1993) PI3-kinases, Grb2-C (Kohda et al., 1994) and Grb2-N (Wittekind et al., 1994).

Recently, the regulatory fragment of human p56 Lck containing the tandem SH2–SH3 domains was crystallized and the structure solved with and without the phosphorylated peptide that corresponds to the carboxyl tail of Lck (Eck et al., 1994). From this work, structural evidence was derived in support of functional observations that SH2 and SH3 domains may co-operate to regulate signal transduction.

One of the goals of this work was to gain a detailed insight into the three-dimensional structure of the isolated SH3 domain of p56 Lck in solution and to compare it with the same protein module of the intact tandem SH2–SH3 domain structure.

Materials and Methods

Sample preparation

The SH3-Lck protein used in this study is a construct comprising 81 amino acids. The residues 4 to 71 correspond to the amino acid residues 59 to 126 of the intact human p56 Lck sequence (Veillette et al., 1987), whereas the N- and C-terminal residues Met^{1} to Ile^{3} and Gly^{72} to Ser^{81} originate from the expression vector.

SH3-Lck was expressed at high levels using the *Escherichia coli* strain M15/pREP4 which carried the plasmid pDS56/RBSII-2. The thawed cells from a 2-liter culture in minimal medium were suspended in 300 ml of 50 mM Tris/HCl (pH 7.0) containing 1 mM Na_2EDTA, 10% (w/v) saccharose, 1 mM 2-mercaptoethanol, 5 mg $MgCl_2$, 33 mg/l DNase I, 2 mM phenylmethylsulfonyl fluoride. The cells were broken by passing twice through a pressure homogenizer (RANNIE Model Mini-Lab, Type 8.3.H) at 150 bar and 800 bar, respectively. The cell extract was centrifuged at 40 000 × g for 30 min and the supernatant was concentrated in a 400-ml stirring cell to a volume of approximately 15 ml by ultrafiltration using a membrane with 3 kDa MW cutoff. SH3-Lck was isolated and purified in two chromatographic steps (FPLC, Pharmacia, Uppsala, Sweden). The protein solution was first loaded onto a Superdex 75 prep-grade column (2.6 × 100 cm) equilibrated in 25 mM Tris/HCl, pH 7.6, 0.02% NaN_3 and eluted at 3 ml/min. The SH3-containing fractions were pooled and loaded on a Resource Q anion-exchange column equilibrated with the same buffer. The protein was eluted with a nonlinear gradient from 0 to 1 M NaCl. The final yield was 11 to 13 mg of highly pure protein per 2 liter of cell culture in minimal medium.

The characterization of the isolated protein in terms of structural identity and purity was carried out using electrospray mass spectrometry (ES/MS), iso-electric focusing (IEF) and SDS gel electrophoresis, and reverse-phase high-performance liquid chromatography (HPLC). The molecular weight determined was found to agree with the expected mass from the amino acid sequence. The protein sample was also highly pure, as judged from the appearance of a single band in SDS and IEF gel electrophoresis (pI 4.6) and a single peak in HPLC.

Uniform ^{15}N-labelling and ^{15}N,^{13}C-labelling (both >98%) were performed in ^{15}N- and ^{15}N,^{13}C-labelled minimal media, using $^{15}NH_4Cl$ and $^{13}C_6$-glucose as N- and C-sources. Fractional ^{13}C-labelling was achieved using a minimal medium containing a mixture of 88% UL-$^{12}C_6$-glucose and 12% UL-$^{13}C_6$-glucose (>98% ^{13}C) and $^{15}NH_4Cl$ (99%) following the protocol given by Senn et al. (1989).

For the acquisition of NMR spectra, 5 to 10 mg of lyophilized, salt-free protein was dissolved in 0.5 ml of 50 mM sodium phosphate buffer (pH 6.5) containing 90% (v/v) H_2O and 10% (v/v) ^2H_2O. NMR samples in pure ^2H_2O were obtained by dissolving 5 mg of protein in ^2H_2O, lyophilizing the protein solution and adding 0.5 ml of pure ^2H_2O (> 99%) containing 50 mM sodium phosphate, pH 6.5. The sample for the measurement of the NH exchange was prepared by dissolving 5 mg of protein in 0.5 ml of ^2H_2O.

NMR measurements

All NMR measurements were performed on Bruker AMX-500 and AMX2-600 spectrometers at 30 °C. The experiments are listed in Table 1, together with the relevant references. For the NH exchange measurement, the sample was transferred to the preshimmed magnet immediately after the protein had been dissolved in ^2H_2O, and 10 ^1H-^{15}N HSQC experiments were recorded with a time interval of 30 min. In addition, spectra were also acquired 6, 10, and 20 h after the start of the exchange.

2D spectra were processed using the program UXNMR (v. 940320.5) on Silicon Graphics Inc. workstations. For the processing of 3D spectra the NMRPipe software package (Delaglio et al., 1995) was applied. Routinely, the time-domain data for ^1H frequencies were zero-filled twice prior to Fourier transformation. For the ^{13}C and ^{15}N time domains the sizes were first doubled by linear prediction and then zero-filled once before Fourier transformation. After processing, the data were transferred into the program XEASY (Bartels et al., 1995) for visual representation and analysis.

Sequential and stereospecific assignments of signals

The resonances of p56 Lck-SH3 were assigned using a combination of four 3D NMR experiments, ^{15}N NOESY-HSQC, ^{15}N TOCSY-HSQC, CBCANH, and CBCA(CO)-NH. Initially, ^1H and ^{15}N resonances of individual residues were identified by the conventional 'sequential assignment method' based on interproton NOEs (Wüthrich, 1986) using ^{15}N NOESY-HSQC and ^{15}N TOCSY-HSQC data. The assignment of connectivities was confirmed by a comparison of CBCANH and CBCA(CO)NH spectra recorded with the uniformly ^{13}C,^{15}N-labelled protein exploiting correlations via chemical bonds. Additional information on side-chain resonances was obtained from 2D ^{13}C CT-HSQC and 3D ^{13}C HCCH-TOCSY experiments.

The stereospecific assignment of diastereomeric β-protons was deduced from 3D HNHB and HN(CO)HB experiments recorded with uniformly ^{15}N- and ^{13}C,^{15}N-labelled SH3-Lck. Stereospecific assignments of prochiral methyl groups of leucine and valine residues were obtained from a comparison of the relative sign of the signals in 2D ^{13}C CT-HSQC experiments recorded using a fractionally ^{13}C-labelled SH3-Lck sample with various constant-time refocusing delays. The side-chain amide protons $H^{\delta1}/H^{\delta2}$ of asparagine and $H^{\varepsilon1}/H^{\varepsilon2}$ of glutamine residues were individually assigned by comparing the relative intensities of the pairs of H^{β}-H^{δ} and H^{γ}-H^{ε} NOEs, respectively (Montelione et al., 1984).

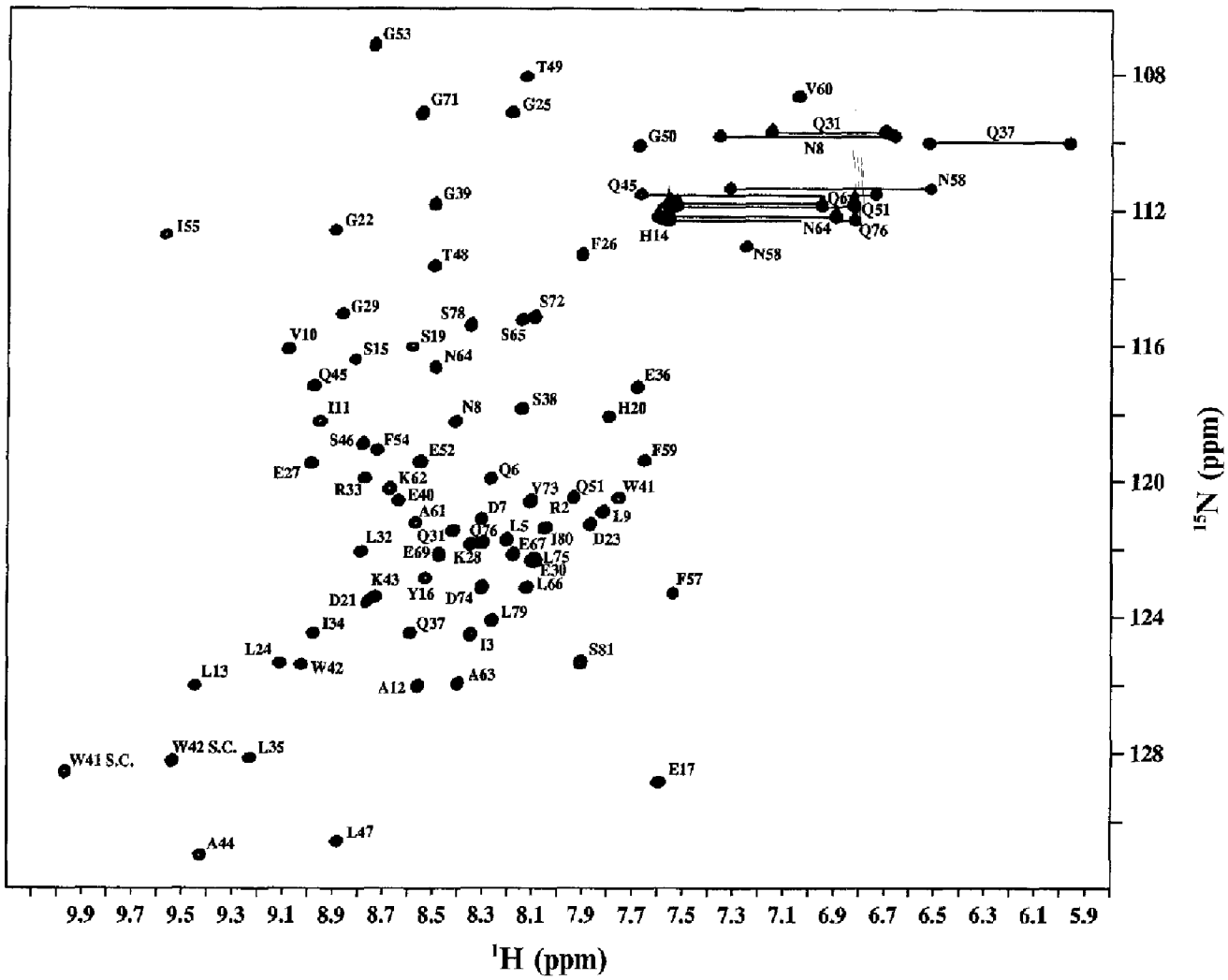

Fig. 1. ^{1}H-^{15}N HSQC spectrum of the 99% ^{15}N-, 12% ^{13}C-labelled SH3 domain of human p56 Lck recorded at 600 MHz. The backbone H-N resonances are labelled by residue types and numbers that correspond to their position in the sequence. Horizontal lines connect the side-chain amide resonances of asparagine and glutamine residues.

Measurements of the proton-detected heteronuclear ^{15}N-{^{1}H}-NOE as a parameter which characterizes the degree of internal protein motion were performed using the pulse sequence of Grzesiek and Bax (1993) and evaluated as stated therein. Large values are considered to be associated with well-ordered parts of the protein backbone, whereas smaller values indicate a disordered and flexible structure.

Collection of the experimental data for the three-dimensional structure calculation

The interproton distance restraints used for the structure calculations were derived from the analysis of three different NOE spectra: (i) NOEs involving NH protons were extracted from a 3D ^{15}N NOESY-HSQC spectrum measured in H_2O; (ii) NOEs between aliphatic protons from a 3D ^{13}C CT-NOESY-HSQC spectrum measured in $^{2}H_2O$; and (iii) distances involving aromatic protons were determined from a 2D NOESY spectrum in $^{2}H_2O$. Volumes of individual cross peaks were integrated using the program XEASY. Upper distance limits were estimated from the cross-peak intensities using the program CALIBA (Güntert et al., 1991a,b). An r^{-6} dependence was assumed for calibrating distances between H^N, H^α and H^β atoms, while all other NOEs involving more peripheral side-chain protons were converted into upper-bound restraints with an r^{-5} dependence (Güntert et al., 1991a, b). Prior to the structure calculations, appropriate pseudoatom corrections were automatically added to the upper-bound distance restraints by DIANA. To avoid an overestimation of NOE intensities due to a contribution of zero-quantum coherence, all intraresidual NOEs between vicinal protons were excluded. $^{3}J_{HNH\alpha}$ coupling constants were determined by a comparison of signal intensities of NH correlations from six 2D HNHA (also called HSQC-J) experiments with various refocusing delays. $^{3}J_{HNH\alpha}$ values were approximated from the delay at which the sign of the individual signals was inverted, following

TABLE 2
^{1}H, ^{15}N AND ^{13}C CHEMICAL SHIFTS OF Lck-SH3 AT 303 K, pH 6.5

Residue	^{15}N	H^{N}	$^{13}C^{\alpha}$	H^{α}	$^{13}C^{\beta}$	H^{β}	$^{13}C^{\gamma}$	H^{γ}	$^{13}C^{\delta}/^{15}N^{\delta}$	H^{δ}	$^{13}C^{\varepsilon}/^{15}N^{\varepsilon}$	H^{ε}	$H^{\varepsilon 3}$	H^{ζ}	$H^{\zeta 2}$	$H^{\zeta 3}$	$H^{\eta 2}$
Met[1]			57.81														
Arg[2]	120.91	7.92	56.01	4.40	30.98	1.71	28.15	1.48	43.06	3.17							
						1.78		1.31									
Ile[3]	124.56	8.35	58.34	4.45	38.52	1.83	26.97	1.47	12.63	0.78							
							17.02	0.91									
Pro[4]			63.13	4.40	32.02	2.23	27.44	1.97	51.04	3.66							
						1.88				3.84							
Leu[5]	121.68	8.21	55.27	4.24	42.37	**1.59**	27.17	1.61	**24.89**	**0.89**							
						1.64			**23.69**	**0.85**							
Gln[6]	120.11	8.27	55.70	4.31	29.43	**2.06**	33.83	2.32			111.95	**7.53**					
						1.91						**6.83**					
Asp[7]	121.18	8.31	54.30	4.56	41.43	2.58											
						2.62											
Asn[8]	118.25	8.40	52.47	4.66	37.85	2.94			109.93	**7.36**							
						2.97				**6.67**							
Leu[9]	120.98	7.82	54.07	5.19	43.68	**1.68**	26.95	1.53	**25.53**	**0.85**							
						1.45			**24.18**	**0.78**							
Val[10]	116.11	9.07	58.61	4.96	36.29	2.03	**21.81**	**0.78**									
							19.62	**0.60**									
Ile[11]	118.22	8.95	58.16	4.91	41.49	1.56	28.24	1.36	13.17	0.79							
								1.07									
							17.29	0.79									
Ala[12]	126.13	8.56	52.33	4.55	21.22	1.67											
Leu[13]	126.05	9.44	55.38	4.09	43.15	**1.27**	26.75	1.40	**25.35**	**0.64**							
						1.15			**22.20**	**0.67**							
His[14]	112.35	7.58	52.94	4.57	33.59	2.53				6.65	8.44						
						2.99											
Ser[15]	116.41	8.81	58.47	4.62	63.81	**4.10**											
						3.92											
Tyr[16]	123.09	8.53	57.44	4.65	42.22	**1.08**				7.08	6.74						
						2.29											
Glu[17]	128.89	7.60	51.73	4.52	30.01	1.72	35.80	2.04									
						1.61											
Pro[18]			63.45	3.99	33.11	2.01	27.42	1.81	50.39	3.55							
								1.66									
Ser[19]	116.21	8.58	58.25	4.40	64.59	3.59											
						3.69											
His[20]	118.21	7.81	53.44	4.73	30.37	3.17				7.19	8.30						
						2.86											
Asp[21]	123.58	8.77	56.20	4.41	40.49	2.57											
						2.61											
Gly[22]	112.75	8.89	45.21	4.39													
				3.93													
Asp[23]	121.47	7.88	54.33	5.33	42.87	2.91											
						3.01											
Leu[24]	125.33	9.11	53.78	4.56	44.71	**1.78**	26.70	1.76	**24.49**	**0.81**							
						0.75			**26.63**	**0.66**							
Gly[25]	109.23	8.18	44.48	4.34													
				3.56													
Phe[26]	113.36	7.91	56.04	5.03	40.08	**2.90**				7.00	7.49			6.84			
						3.30											
Glu[27]	119.61	8.98	53.60	4.77	32.75	**1.95**	36.35	2.38									
						2.14											
Lys[28]	121.87	8.36	58.93	3.37	32.60	1.65	29.71	1.69	24.88	1.22	41.82	2.99					
						1.69				1.07							
Gly[29]	115.11	8.86	44.95	**3.49**													
				4.36													
Glu[30]	122.39	8.11	57.78	4.17	31.98	2.11	37.19	2.50									
						2.39		2.38									
Gln[31]	121.48	8.42	54.95	5.23	30.74	**2.06**	34.98	2.48			109.81	**7.15**					
						2.03		2.11				**6.70**					

TABLE 2
(continued)

Residue	^{15}N	H^N	$^{13}C^\alpha$	H^α	$^{13}C^\beta$	H^β	$^{13}C^\gamma$	H^γ	$^{13}C^\delta/^{15}N^\delta$	H^δ	$^{13}C^\varepsilon/^{15}N^\varepsilon$	H^ε	$H^{\varepsilon 3}$	H^ζ	$H^{\zeta 2}$	$H^{\zeta 3}$	$H^{\eta 2}$
Leu[32]	122.07	8.78	53.26	5.04	45.82	**1.19**	27.36	1.36	**25.99**	**0.08**							
						1.10			**24.99**	**0.62**							
Arg[33]	120.00	8.77	54.03	5.01	32.74	1.61	26.93	1.57	43.35	3.16	109.48	7.51					
						1.70											
Ile[34]	124.52	8.98	58.88	4.25	35.60	1.97	26.75	1.60	10.44	0.74							
								1.30									
							18.72	0.56									
Leu[35]	128.17	9.23	55.40	4.35	42.90	**1.25**	27.53	1.50	**25.40**								
						1.56			**22.50**								
Glu[36]	117.24	7.69	56.35	4.32	33.58	**2.06**	36.46	2.24									
						1.90		2.15									
Gln[37]	124.55	8.59	53.40	3.62	28.84	**1.02**	33.37	0.99			110.13	**6.52**					
						1.17		0.82				**5.96**					
Ser[38]	117.99	8.14	56.90	4.60	63.79	4.36											
						3.85											
Gly[39]	111.84	8.50	45.89	4.08													
				3.75													
Glu[40]	120.55	8.64	56.63	3.81	30.12	**1.71**	36.79	1.97									
						1.85		1.75									
Trp[41]	120.59	7.77	55.31	5.24	30.22	2.95				7.10	128.56	9.97	7.18		7.41	6.77	7.27
Trp[42]	125.45	9.03	53.98	5.44	30.89	**2.98**				6.89	128.28	6.94	7.24		7.44	6.89	7.25
						3.34											
Lys[43]	123.41	8.74	55.95	4.47	34.03	1.61	28.85	1.38	25.53	1.12	41.57	2.64					
						1.67				0.82		2.57					
Ala[44]	131.00	9.42	50.33	5.47	25.73	1.23											
Gln[45]	117.41	8.97	50.40	5.29	33.71	**1.97**	33.69	2.12			111.65	**7.56**					
						1.73						**6.95**					
Ser[46]	119.04	8.79	57.74	4.68	63.30	**4.05**											
						4.21											
Leu[47]	130.53	8.88	57.10	4.29	41.03	1.74	28.82	1.72	**25.57**	**0.86**							
									23.46	**0.81**							
Thr[48]	113.69	8.50	64.61	4.28	69.23	4.09	22.45	1.26									
Thr[49]	108.21	8.13	61.54	4.55	71.32	4.49	21.40	1.27									
Gly[50]	110.25	7.69	45.72	4.23													
				3.83													
Gln[51]	120.57	7.93	56.38	4.23	29.50	**1.94**	34.15	2.43			111.98	**7.57**					
						2.07		2.40				**6.95**					
Glu[52]	119.34	8.54	53.62	5.70	34.19	1.89	35.93	2.11									
						1.82		2.38									
Gly[53]	107.20	8.74	45.46	4.11													
				3.92													
Phe[54]	119.23	8.72	58.68	5.66	41.30	3.16				7.21	7.41			7.35			
						2.98											
Ile[55]	112.77	9.57	57.19	5.21	40.64	1.67	25.62	1.44	15.82	0.44							
								1.04									
							21.40	1.12									
Pro[56]			61.44	3.70	30.28	1.14	27.07	0.63	49.73	2.83							
						1.25											
Phe[57]	123.50	7.55	59.26	2.87	35.51	**1.14**				6.91	7.17			7.21			
						1.49											
Asn[58]	113.17	7.26	52.11	4.00	35.61	**1.45**			111.46	**7.32**							
						2.54				**6.51**							
Phe[59]	119.36	7.66	57.93	4.50	39.39	**3.24**				7.07	7.41			7.30			
						3.63											
Val[60]	108.90	7.05	58.32	5.25	35.69	1.79	**21.95**	**0.40**									
							18.00	**0.69**									
Ala[61]	121.25	8.57	50.55	4.73	23.21	1.37											
Lys[62]	120.17	8.68	57.32	4.28	32.71	1.91	29.20	1.76	25.17	1.61	41.95	3.04					
						1.80				1.54							
Ala[63]	125.93	8.40	52.79	4.21	19.45	1.27											

TABLE 2
(continued)

Residue	^{15}N	H^N	$^{13}C^\alpha$	H^α	$^{13}C^\beta$	H^β	$^{13}C^\gamma$	H^γ	$^{13}C^\delta/^{15}N^\delta$	H^δ	$^{13}C^\varepsilon/^{15}N^\varepsilon$	H^ε	$H^{\varepsilon 3}$	H^ζ	$H^{\zeta 2}$	$H^{\zeta 3}$	$H^{\eta 2}$
Asn[64]	116.72	8.49	53.15	4.69	38.73	**2.82**				112.27	**7.57**						
						2.78					**6.90**						
Ser[65]	115.31	8.15	58.26	4.42	63.99	3.85											
						3.88											
Leu[66]	123.19	8.13	54.84	4.35	42.42	1.55	26.95	1.62	**25.16**	**0.87**							
						1.63			**23.42**	**0.81**							
Glu[67]	122.24	8.18	54.14	4.55	29.85	1.83	35.96	2.26									
						1.98											
Pro[68]			62.79	4.37	30.03	2.25	31.97	2.06	50.51	3.74							
						1.85				3.65							
Glu[69]	122.18	8.47	54.35	4.54	29.56	2.00	36.06	2.31									
						1.86											
Pro[70]			63.48	4.71	31.84	2.28	27.53	1.91	50.76	3.83							
						1.91				3.67							
Gly[71]	109.20	8.54	45.12	3.97													
				3.80													
Ser[72]	115.21	8.10	58.21	4.46	63.90	3.84											
						3.88											
Val[73]	120.55	8.11	62.13	4.14	32.92	2.08	**21.19**	**0.91**									
							20.28	**0.90**									
Asp[74]	123.20	8.31	54.16	4.58	41.17	**2.54**											
						2.68											
Leu[75]	122.37	8.08	54.80	4.34	42.49	1.59	26.91	1.61	**24.93**	**0.90**							
						1.61			**23.85**	**0.84**							
Gln[76]	121.88	8.31	53.56	4.57	28.90	1.90	33.61	2.36			112.40	**7.56**					
						2.05						**6.82**					
Pro[77]			62.84		31.98	2.28											
						1.90											
Ser[78]	115.47	8.35	58.17	4.39	63.74	3.83											
						3.88											
Leu[79]	124.16	8.26	55.10	4.39	42.37	1.58	26.91	1.54	**24.93 23.70**								
						1.63											
Ile[80]	121.54	8.06	60.94	4.22	38.79	1.86	27.14	1.43	12.47								
							17.38	0.90									
Ser[81]	125.34	7.91	59.75	4.24	64.96	3.79											

^1H, ^{13}C and ^{15}N chemical shifts for the SH3 domain of p56 Lck at 303 K and a pH of 6.5. Uncertainties in the shifts are ± 0.02, 0.1 and 0.1 ppm for ^1H, ^{13}C and ^{15}N, respectively. Referencing is made relative to DSS directly for ^1H shifts, and indirectly for ^{13}C and ^{15}N shifts with conversion factors of 0.251449 and 0.101329, respectively (Wishart et al., 1995). Bold values represent stereospecific or regiospecific assignments. In these cases the pro-(*R*) assignment is given in the first place. For the NH_2 group of asparagine and glutamine residues the *E*-resonance is listed first.

the equation:

$$^3J_{HNH\alpha} = 1/4\ T \qquad (1)$$

with an experimental error of 1.0 Hz.

$^3J_{NH\beta}$ and $^3J_{CH\beta}$ coupling constants were estimated from an analysis of signals in 3D HNHB and HN(CO)HB experiments. The theoretical values of −6.0 (± 2.0) Hz and 8.0 (± 2.0) Hz, which are assumed for the staggered conformations, were applied for $^3J_{NH\beta}$ and $^3J_{CH\beta}$ for individual residues, when the corresponding cross peaks were observed in the spectra; 0 (± 2.0) Hz was used when the signals did not appear. $^3J_{HNH\alpha}$, $^3J_{NH\beta}$ and $^3J_{CH\beta}$ coupling constants were subjected to a local conformation analysis by the program HABAS (Güntert et al., 1991a), but no NOE information was applied. The resulting torsion-angle restraints were taken for the structure calculation.

Main-chain hydrogen bonds could be deduced from the NH-exchange experiment and from the secondary structure of SH3-Lck, which was obtained from a preliminary pattern analysis of intensities of sequential and inter-strand NOEs among backbone protons. They were con-

TABLE 3
RELATIVE INTENSITIES OF CROSS PEAKS IN THE CT-HSQC EXPERIMENT OF FRACTIONALLY LABELLED PROTEINS

Labelled protein	Cross-peak intensity		
	$1/4J_{CC}$	$1/2J_{CC}$	$1/J_{CC}$
$Leu^{\delta 1}$ and $Val^{\gamma 1}$	0	−1	1
$Leu^{\delta 2}$ and $Val^{\gamma 2}$	1	1	1
Thr^{γ} and $Ile^{\delta 1}$	0.5	0	1

The sign of the magnetisation is modulated by $\cos^n(2\pi T/J_{CC})$ where n is the number of directly bonded ^{13}C nuclei.

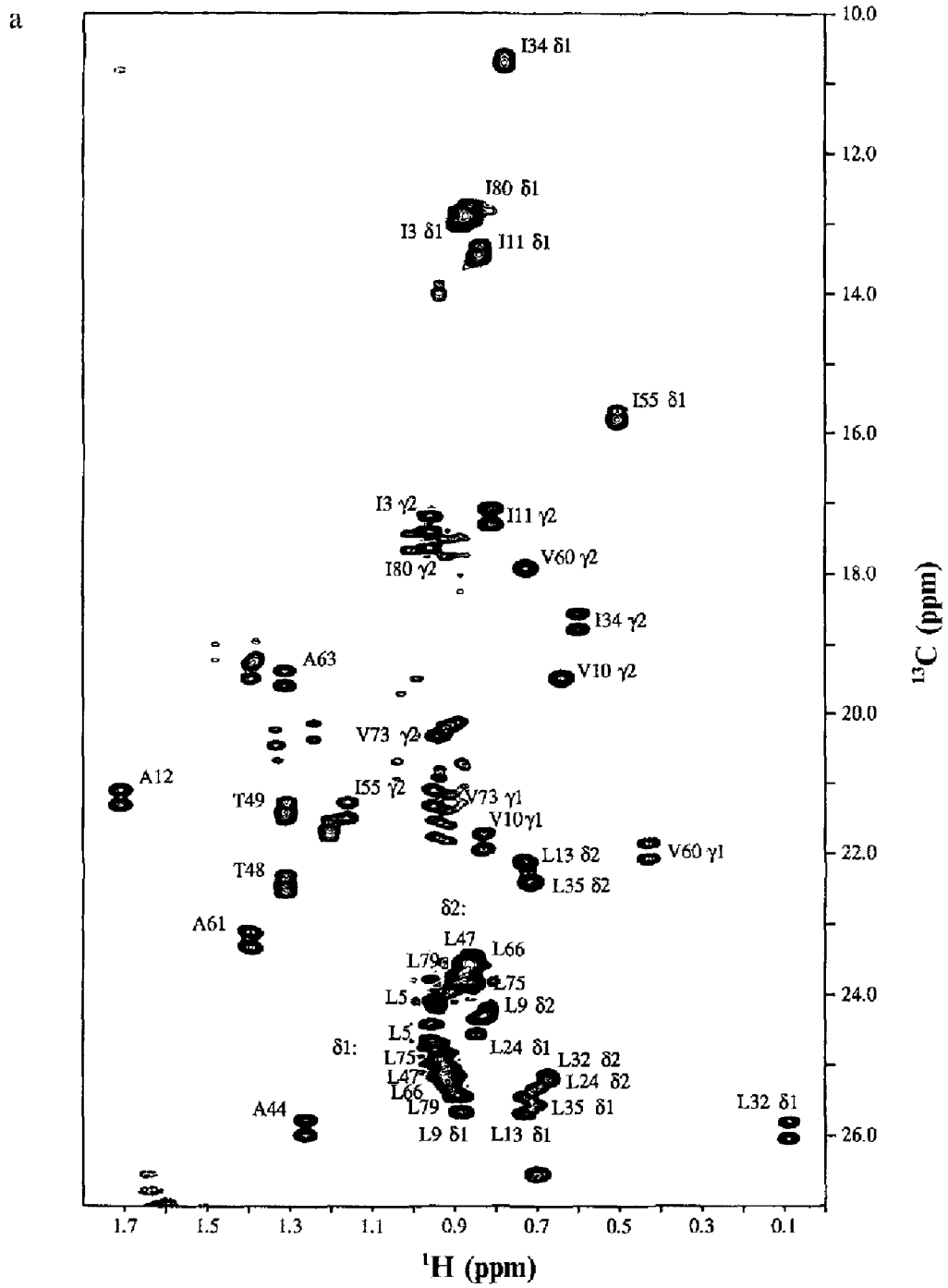

Fig. 2. (a) Methyl region of a 2D ^{1}H-^{13}C HSQC spectrum of the 'fractionally labelled' SH3 domain of human p56 Lck. The pro-*R* prochiral methyl groups of leucine and valine are split up into doublets due to the presence of the ^{13}C-labelled γ-methyl carbon. In addition, Ala-C$^{\beta}$, Thr-C$^{\gamma 2}$, Ile-C$^{\gamma 2}$, and Ile-C$^{\delta 1}$ show the same pattern, whereas Val-C$^{\gamma 2}$ and Leu-C$^{\delta 2}$ appear as singlet peaks. Unlabelled cross peaks are due to impurities of the sample. (b) Constant-time version of the 2D ^{1}H-^{13}C HSQC spectrum of the 'fractionally labelled' SH3 domain of p56 Lck. Here, the prochiral pro-(*S*) methyl groups of leucine and valine are of opposite sign relative to the pro-(*R*) methyl group. Negative contours are indicated by broken lines.

verted into upper-bound distance restraints of 4.0 Å and 2.0 Å and lower-bound distance restraints of 2.7 Å and 1.6 Å from C' to N and O to H^{N}, respectively.

Calculations of three-dimensional structures

Structures were calculated using both the distance-geometry program DIANA (Güntert et al., 1991a,b) and the dynamical simulated-annealing protocol of Nilges (Nilges et al., 1988; Nilges, 1993,1995) on a Silicon Graphics Indy workstation. Initial conformations were obtained by DIANA using only unambiguously defined distance restraints. Additional restraints were incorpor-

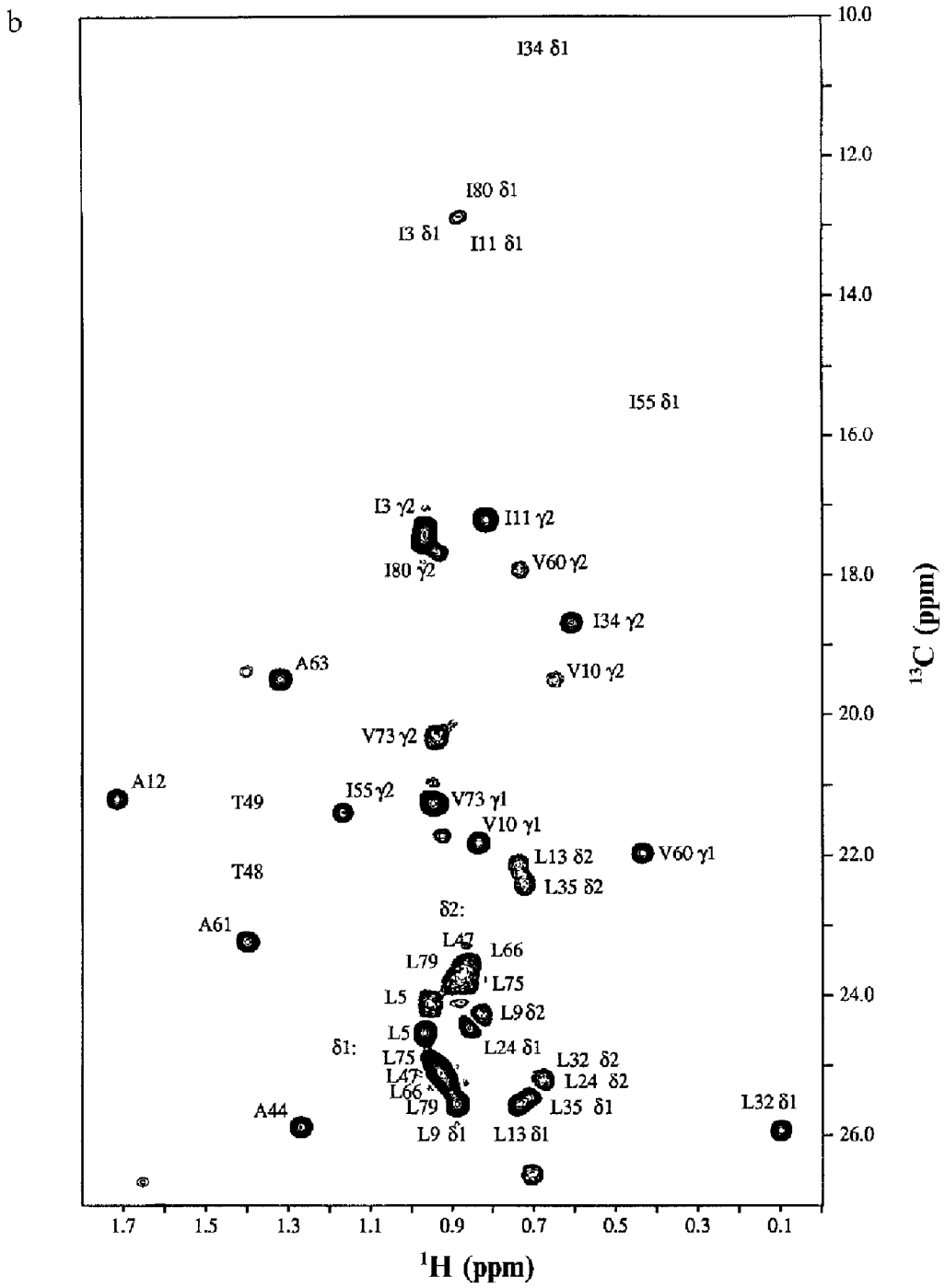

Fig. 2. (continued).

ated in successive stages according to the concept of structure-aided assignment of ambiguous NOE cross peaks (Güntert et al., 1993). The DIANA v. 2.4 program with the subroutine REDAC (Güntert and Wüthrich, 1991) was used to generate intermediate structures. At each stage of the structure-based assignment, ambiguous NOE cross peaks were examined in the five best conformers obtained from 50 initially randomized conformations with one cycle of the REDAC approach with a cutoff value of 50 Å^2 for the final target function. The newly assigned cross peaks were subsequently added to the restraints list for the next stage. The program ASNO (Güntert et al., 1993) and the in-house program SBAeasy3d (H. Hiroaki, unpublished results) were used for this purpose.

The final structure calculation proceeded through the following steps: (i) NOE-derived upper distance bounds, upper and lower distance bounds of hydrogen bonds, and dihedral angular restraints were used as input for the DIANA calculations. Starting from 50 randomized con-

TABLE 4
CATEGORIES AND NUMBERS OF NMR-DERIVED RESTRAINTS USED FOR THE STRUCTURE CALCULATIONS

Category	Number of restraints
Distance restraints	
intraresidual	210
sequential ($\|i-j\|=1$)	231
medium-range ($1<\|i-j\|<5$)	64
long-range ($\|i-j\|>4$)	280
hydrogen-bond	28
Angular restraints	
torsion angle ϕ	67
torsion angle ψ	54
dihedral angle χ_1	25
dihedral angle χ_2	4

Summary of the structural restraints derived from NOEs, coupling constants and the DIANA/REDAC approach. For each hydrogen bond there are two distance restraints: 1.6 Å $< r_{NH-O} <$ 2.0 Å and 2.7 Å $< r_{N-O} <$ 4.0 Å. All hydrogen bonds involving slowly exchanging amide protons are located in the core region of the β-sheets.

formations, a threshold value of 50 Å^2 as the cutoff for the target function was applied to reduce the ranges of the angular restraints in the REDAC routine. This REDAC step was repeated twice; (ii) the program XPLOR (v. 3.2) with the dynamical simulated-annealing protocol (Nilges et al., 1998; Nilges, 1993,1995) was used. The force field included a dihedral energy term, and bond lengths and bond angles were modified to match the corresponding values of the 'parhcsdx.pro' parameter set, which is derived from a statistical survey of X-ray structures of small compounds from the Cambridge Structural Database (Engh and Huber, 1991). The upper limits for the distance bounds were the same as used for DIANA, while the lower limit distances were uniformly set to 2.0 Å. In addition, angular restraints which resulted from the previous two cycles of the REDAC routine were also incorporated into the calculations. Simulated annealing was done in a manner as described by Nilges (1995). For a check of the quality of the structures the program PROCHECK-NMR (MacArthur and Thornton, 1993) was applied. Analysis and visualization of the resulting

TABLE 5
BEST-MATCH FIT FOR VARIOUS SEGMENTS OF THE 25 BEST NMR SOLUTION STRUCTURES (Å)

Residues	Backbone atoms	All heavy atoms
6–64	0.60	1.19
7–63	0.54	1.14
7–16, 24–63	0.42	1.00
over β-sheets	0.26	0.75

Structural statistics for the 25 NMR-derived solution conformations. For the superposition, the weighted sum of mutual squared deviations between corresponding backbone atoms (N, C^α, C', and O) or heavy atoms in the final XPLOR structures is minimized (Gerber and Müller, 1987). None of the structures exhibit distance-restraint violations larger than 0.25 Å or dihedral angle violations > 3°.

TABLE 6
OVERLAY OF AVERAGE NMR STRUCTURE WITH CORRESPONDING SEGMENTS OF THE X-RAY STRUCTURE OF SH2–SH3 (Å)

Residues	Backbone atoms	All heavy atoms
6–64	1.13	1.81
7–63	0.88	1.47
7–16, 24–63	0.80	1.42
over β-sheets	0.49	1.32

The average NMR structure is obtained by averaging the coordinates of the final conformers after rmsd best fit over residues 4 to 71 which correspond to the native protein sequence.

structures was performed on Silicon Graphics Indy workstations using the in-house molecular modelling program MOLOC (Gerber and Müller, 1987,1995; Müller et al., 1988).

Results and Discussion

NMR assignment of p56 Lck-SH3

The ^{1}H, ^{13}C, and ^{15}N NMR chemical shifts of p56 Lck-SH3 at 303 K and pH 6.5 were obtained by a combination of 3D ^{1}H-^{13}C, ^{1}H-^{15}N, and ^{1}H-^{13}C-^{15}N experiments following established methods for small proteins. Figure 1 shows the 2D ^{1}H-^{15}N HSQC spectrum of Lck-SH3 with the assignments indicated. All resonances of the polypeptide backbone and side-chain atoms were assigned with the exception of the ^{13}C chemical shift of carbonyls and carboxyls, the ^{1}H, ^{13}C and ^{15}N shifts of peripheral atoms of arginine and lysine side chains and the ^{13}C resonances of aromatic carbons of phenylalanine, tyrosine, tryptophan and histidine residues. The final result is presented in Table 2, in which the stereospecific assignments of prochiral methylene and methyl groups are also indicated. In total, 27 methylene groups, all the methyl groups of valine and leucine, and all nine pairs of side-chain amide protons of asparagine and glutamine were assigned stereo-

TABLE 7
AVERAGE AND STANDARD DEVIATION OF THE XPLOR ENERGY TERMS FOR THE 25 BEST STRUCTURES

Energy term	Average ± standard deviation (kcal mol^{-1})
E_{tot}	130.53 ± 3.06
E_{bond}	4.16 ± 0.27
E_{angl}	40.45 ± 0.87
E_{dihed}	43.50 ± 2.87
E_{impr}	4.94 ± 0.20
E_{repel}	17.66 ± 1.74
E_{NOE}	19.23 ± 1.98
E_{CDIH}	0.58 ± 0.17

The values for the XPLOR energy terms are obtained with force constants of 4 kcal mol^{-1} Å^{-4} (E_{repel}), 50 kcal mol^{-1} Å^{-2} (E_{NOE}) and 200 kcal mol^{-1} (E_{CDIH}). The numbers are the arithmetic mean and standard deviation of the energy terms for the 25 final conformations.

specifically by the various NMR experiments. Using the program GLOMSA (Güntert et al., 1991a,b), one additional stereospecific assignment for Gly[28] was obtained.

A novel combination of the constant-time 2D ^{1}H-^{13}C correlation experiment and fractionally ^{13}C-labelled sample, which was initially described by Senn et al. (1989), was applied to the stereospecific assignment of prochiral methyl groups. For a fractionally ^{13}C-labelled protein the resonance of the pro-*R* methyl carbon (γ^1 and δ^1 of valine and leucine, Fig. 2) is a doublet in the NMR ^{13}C dimension, due to the presence of a directly bonded ^{13}C nucleus which causes a splitting with a ^{1}J ^{13}C-^{13}C coupling of ~ 33 Hz. The ^{13}C NMR signal of the pro-*S* methyl is a singlet because it has a ^{12}C nucleus as its directly bonded neighbour. In the CT-HSQC experiment the number of directly bonded ^{13}C nuclei is reflected in the sign of the cross peaks. If the CT delay (2T) is adjusted to $1/J_{CC}$, the sign of the ^{13}C magnetisation is opposite for carbons coupled to an odd versus an even number of ^{13}C-coupling partners (Table 3). The fractional labeling pattern of the γ-CH_3 group of threonine and the δ^1-CH_3 group of isoleucine is such that the directly bonded carbon has an equal prob-

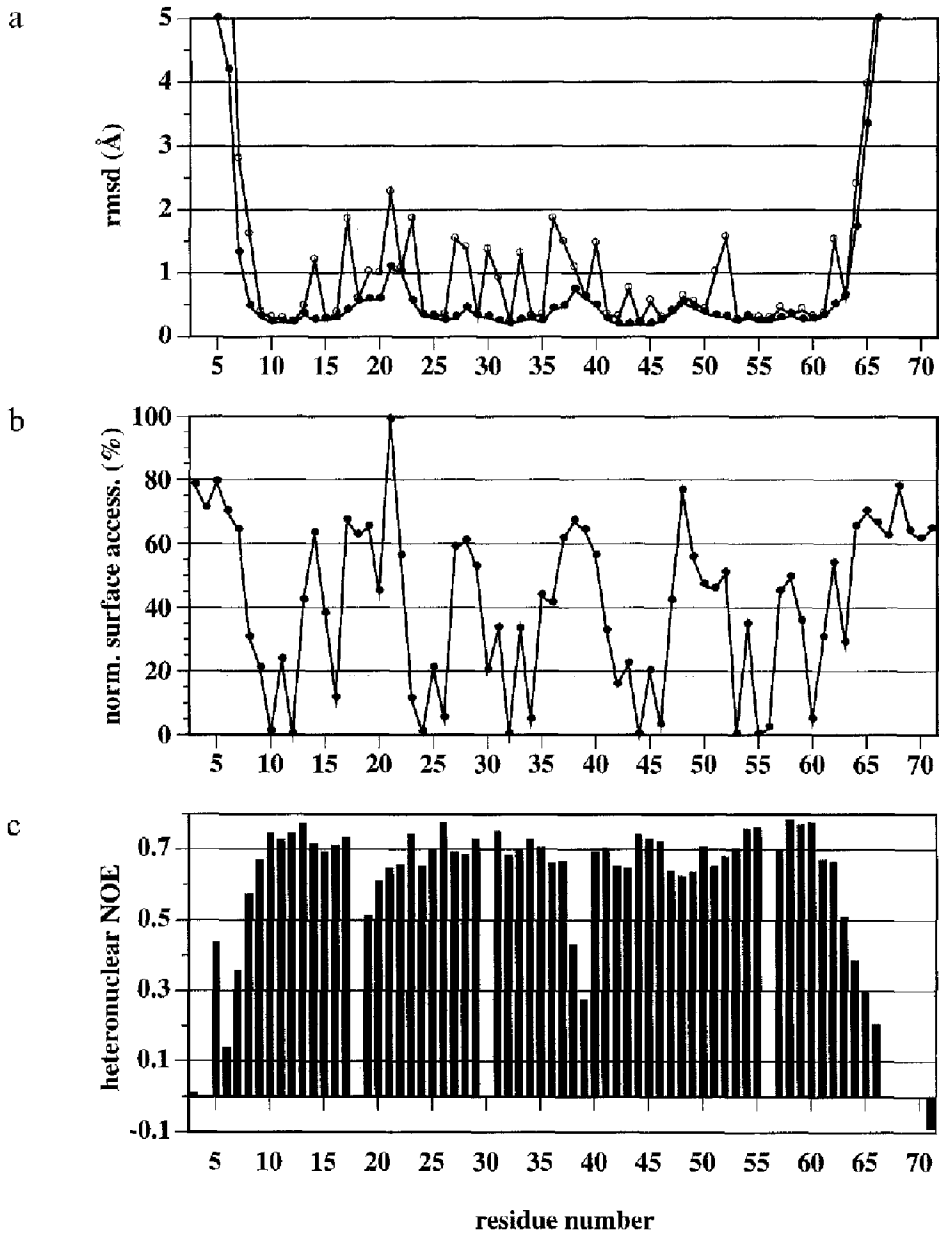

Fig. 3. (a) Plot of the rmsd distribution for residues 4 to 71 of the final 25 NMR conformations obtained after the best fit (Gerber and Müller, 1987) over this segment for backbone atoms N, C^{α}, C', and O (filled circles) or all heavy atoms (open circles); (b) plot of the relative mean surface accessibility for the individual amino acid residues in the 25 NMR conformations. The water-accessible surface area of each residue Xxx was calculated using a probe radius of 1.4 Å and normalized by its maximum accessible surface area of a fully extended tripeptide Gly-Xxx-Gly (Kabsch and Sander, 1983); (c) distribution of ^{15}N-{^{1}H}-detected heteronuclear NOE values over the sequence of p56 Lck SH3. Missing bars are due to proline residues (at positions 4, 18, 56, 68, 70) or unresolved cross peaks (at position 30).

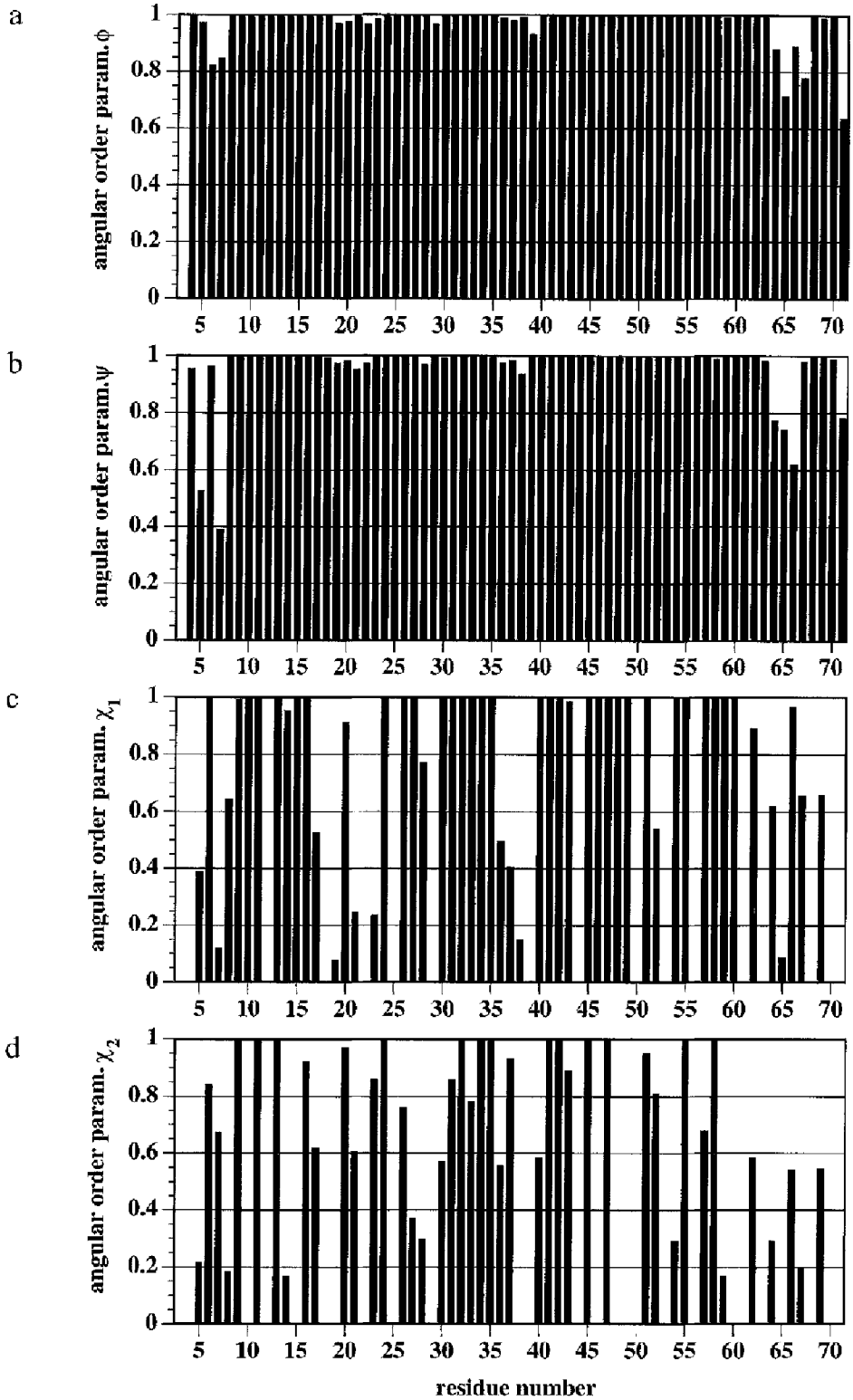

Fig. 4. Angular order parameters for φ (a), ψ (b), χ_1 (c) and χ_2 (d) calculated from the 25 best NMR conformers.

ability to be ^{12}C or ^{13}C (Szyperski, 1995). This leads to a cancellation of their signals (Fig. 2).

Assignment of NOESY cross peaks and structure determination

For the first stage of the structure calculations, which was performed using the distance-geometry program DIANA (Güntert et al., 1991a,b), an initial set of 402 upper-limit distance restraints obtained from unambiguous NOEs, together with 150 dihedral angle restraints derived by the program HABAS (Güntert et al., 1991a), were applied. Assignments of ambiguous NOESY cross

peaks were achieved according to the concept of structure-aided assignment (Güntert et al., 1993) where unassigned NOE peaks were examined against the intermediate DIANA structures. In the initial phase of the structure-based assignment, peaks which had only one tentative assignment to a proton pair with a maximum distance below 5.0 Å in each of the five best structures were assigned. Usually 10 to 30 newly assigned peaks were found in consecutive rounds of structure calculations. At the same time, all the assigned peaks were examined against the best 20 intermediate structures and revised when the corresponding distance restraints were violated in at least 14 of the structures. This procedure was repeated more than 20 times. In the last step, hydrogen bonds were incorporated as additional distance restraints. Finally, a total of 1017 independent cross peaks from ^{1}H-^{15}N 3D NOESY-HSQC, ^{1}H-^{13}C 3D CT-HSQC-NOESY and from 2D NOESY spectra were assigned. As the information content of intraresidue NOEs between neighbouring protons tends to be biased by contributions from zero-quantum coherences and spin diffusion, this results in unreliable distance restraints. In order to avoid an incorrect interpretation of NOE intensities due to spurious contributions of coupling constants, all interproton distance restraints among vicinal protons were excluded, thereby reducing the number of NOE-derived distance restraints to 785.

The various types of restraints for the structure calculations are summarized in Table 4. In the final stage 100 random conformations were subjected to simulated annealing using XPLOR (Nilges et al., 1988; Nilges, 1993, 1995), and the 25 structures with the lowest value of the total energy term were selected to represent the solution conformation of p56 Lck-SH3. Tables 5–7 give an overview of the structural statistics. The atomic rms deviations for backbone and side-chain atoms, together with the normalized solvent-accessible surface area and the ^{15}N-{^{1}H}-NOE values are plotted on a per-residue basis

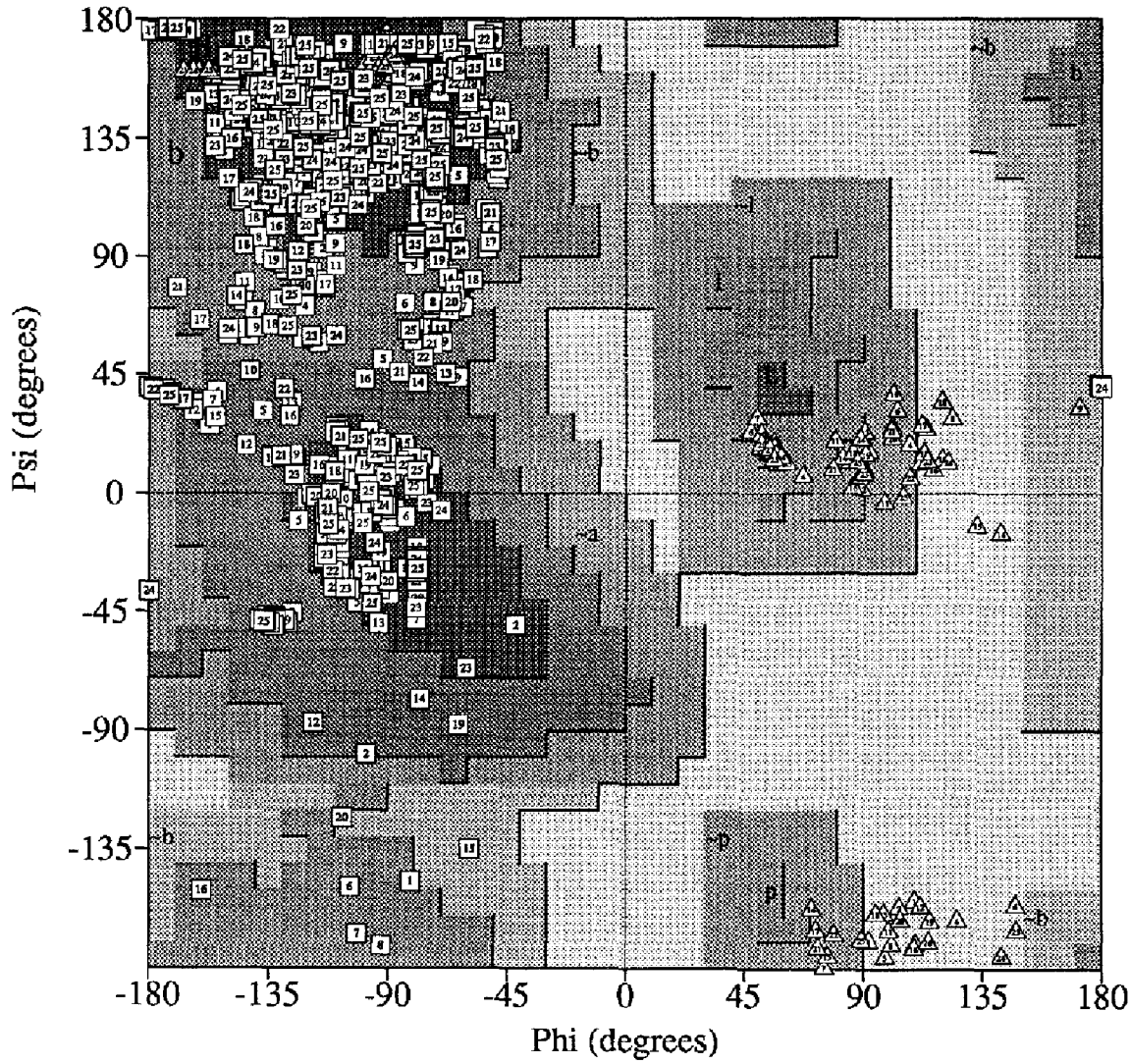

Fig. 5. Ramachandran plot for the 25 NMR conformations (residues 6 to 65). Glycine residues are shown as triangles. The plot was generated using the program PROCHECK-NMR (MacArthur and Thornton, 1993).

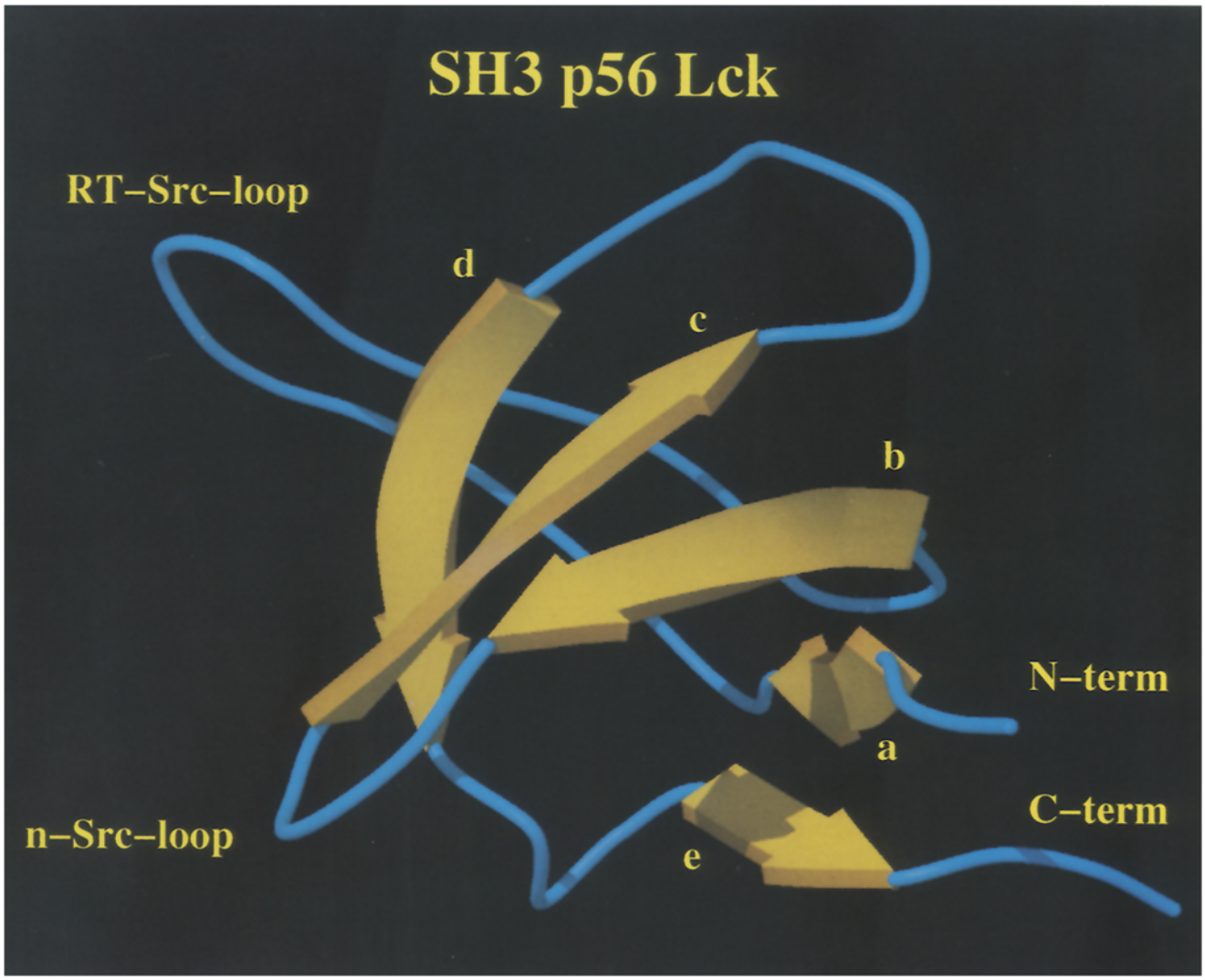

Fig. 6. Schematic representation of the overall fold of the SH3 domain of p56 Lck. The β-strands, which are shown as ribbons, the loops, and the termini of the molecule are labelled. The figure was generated using the program RASTER3D (Merritt and Murphy, 1994).

in Fig. 3. Regions of low internal flexibility are characterized by large values of the heteronuclear NOEs. Theoretically, this NOE can vary from –3.6, for small molecules in the fast-tumbling regime of overall motion, to +0.82 for large molecules in the slow-tumbling regime (Kay et al., 1989). Therefore, residues 9 to 62 with values of the heteronuclear NOE larger than 0.6 are considered to delineate the structurally well-defined region of the protein. Within this part of the molecule, only Ser^{38}/Gly^{39} and, to a lesser extent, Ser^{19} show an increased internal flexibility. Concomitantly, the rms deviation of 0.54 Å for the backbone atoms N, C^{α}, C', and O is low for the segment 7–63. Figure 4 gives the angular order parameters (Hyberts et al., 1992) for the torsion angles ϕ, ψ, χ_1, and χ_2. A Ramachandran plot of residues 6–65 of the final 25 structures is shown in Fig. 5, with the backbone torsion angles for all non-glycine residues lying within the allowed regions.

The solution conformation of Lck-SH3

The schematic representation of the solution structure of the SH3 domain of p56 Lck is shown in Fig. 6. It comprises five β-strands (a to e) which are arranged into a five-membered antiparallel β-sheet of concave shape. The strands are formed by residues Leu^{9}–Ala^{12}, Gln^{31}–Glu^{36}, Trp^{42}–Ser^{46}, Glu^{52}–Ile^{55}, and Val^{60}–Lys^{62}. Strand b is a twisted β-pleated sheet, with hydrogen bonds shared between strands a and c. Furthermore, strands a and c are connected by H bonds to e and d, respectively. Three tight turns are found in the structure. The segment His^{20}–Asp^{23} forms a type-II turn, as do residues Glu^{27}–Glu^{30}, whereas the segment Leu^{47}–Gly^{50} forms a turn of type I. The presence of a helical half turn in the region Phe^{57}–Phe^{59}, as indicated by the pattern of sequential NOEs, is confirmed by the structure calculations. The residues Pro^{4}–Gln^{6} and Asn^{64}–Gly^{71} which are part of the native sequence of the protein are flexible, as are their N- and C-terminal exten-

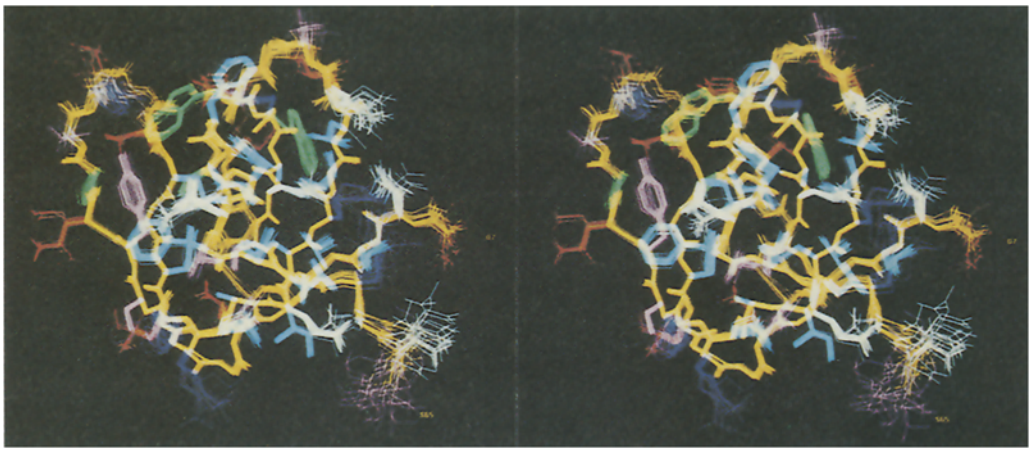

Fig. 7. Stereoview of the polypeptide (N, C^{α}, C', O; in orange) of residues 7 to 65 of p56 Lck SH3. The side chains are coloured according to residue type: red: Asp, Glu; blue: Arg, Lys, His; cyan: Ala, Ile, Leu, Phe, Val; green: Trp, Pro; pink: Tyr, Thr, Ser; white: Asn, Gln.

sions Met^{1}–Ile^{3} and Ser^{72}–Ser^{81} that originate from the expression system and do not belong to the native sequence of SH3.

The atomic rmsd values for different sets of superpositions are given in Tables 5 and 6. The best-defined regions of the structure are delineated by residues 7 to 16 and 24 to 63 with an rmsd value of 0.42 Å, whereas residues 17–23 are less well defined. The individual amino acid side-chain conformations are determined with variable precision in the solution structure (Fig. 7). The group of residues with the 'best-defined' side chains (neglecting glycine) is characterized by an rmsd value below 0.8 Å when all heavy atoms in the range Asp^{7}–Ala^{63} are superimposed for minimum mutual deviation. It comprises Leu^{9}–Leu^{13}, Ser^{15}, Tyr^{16}, Leu^{24}, Phe^{26}, Leu^{32}, Ile^{34}, Leu^{35}, Trp^{41}–Thr^{49}, Phe^{54} Ala^{61}, and Ala^{63}. Most of them belong to the category of 'large, non-polar residues' reflecting their hydrophobic packing into the core of the protein. Only Ser^{15}, Gln^{45}, Ser^{46}, Thr^{48}, and Thr^{49} represent polar side chains which are exposed to various degrees to the surface of the protein.

Comparison of NMR and X-ray structures of Lck-SH3

Figure 8 shows the 25 selected NMR structures calculated by XPLOR superimposed on the X-ray coordinates (Eck et al., 1994) at the positions of the backbone atoms (N, C^{α}, C', O) for the residues that are involved in the formation of the β-sheet. Essentially, the largest part of the X-ray structure lies inside the envelope defined by the ensemble of the solution structures. The main-chain trace and also the orientation of the interior side chains of the protein are almost identical. Differences between the solution and crystal structure are observed at the positions of Ser^{38}/Gly^{39}, Thr^{48}/Thr^{49} and for the 'ab-loop'-connecting β-strands a and b. These segments also correspond to the least well-defined portions in the backbone

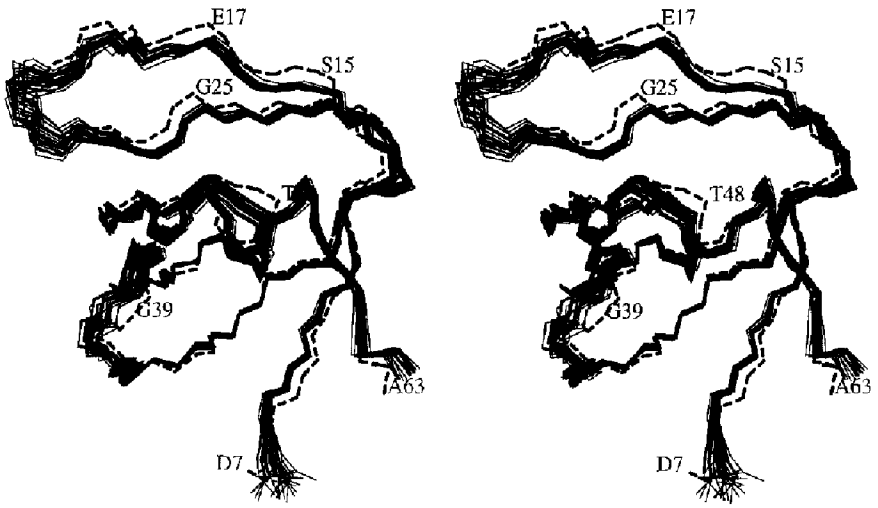

Fig. 8. Stereoview showing residues 7 to 63 of the backbone atoms (N, C^{α}, C') of the 25 NMR conformations after a best-fit superposition on the X-ray coordinates (Eck et al., 1994) for the residues that are involved in the formation of the β-sheet (9–12, 31–36, 42–46, 52–55, 60–62). The backbone conformation of the crystal structure is indicated by broken lines.

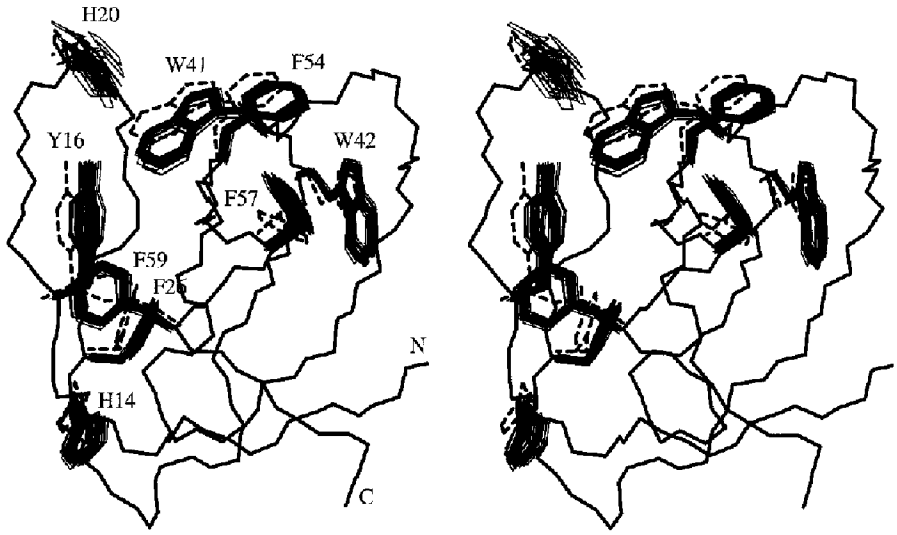

Fig. 9. Stereoview illustrating the clustering of the aromatic residues. The location of their side chains and of one arbitrarily selected backbone of the 25 NMR structures is shown with solid lines. The orientation of the aromatic side chains in the crystal structure is indicated by broken lines.

superposition of the NMR structures alone. While the differences at residues 38/39 correlate with the locally increased flexibility of the NMR conformers (see Fig. 3), the threonines at positions 48 and 49 form a well-defined turn in both the aqueous and crystalline environment. In the X-ray structure, however, the torsion angles of Thr^{48} lie well within the preferred region of the Ramachandran plot, whereas the ϕ angle of Thr^{48} is slightly shifted towards more negative values in the solution structures. The differences involving the 'ab loop' can be rationalized as a small displacement of this structural element in the X-ray structure of the SH3–SH2-regulatory domains liganded with the phosphopeptide. There, this loop of the SH3 module lies at the dimer interface to the SH2 domain of the other molecule in the asymmetric unit and is in contact with the C-terminal tail of the phosphopeptide.

Comparison to other SH3 domains

In published structures of SH3 domains, the general topology is described as a perpendicular arrangement of two three-membered β-sheets plus two additional short β-strands within the 'ab loop', and a 3_{10} helix. In most of the structures, the β-strand b is longer than that found in Lck-SH3. Furthermore, it is kinked and arranged in such a way that it is part of both sheets (Kohda et al., 1994). However, in both the solution and X-ray structure of Lck-SH3 the topology resembles more a continuous twisted five-membered β-sheet. The additional two short β-strands in residue ranges Ser^{15}–Glu^{17} and Gly^{25}–Glu^{27}, which have been observed in the other SH3 domains, are found neither in the calculated NMR structures of Lck nor in the crystal structure of the SH3–SH2 fragment. The presence of Pro^{18} in this part of the sequence interferes with the formation of the short β-sheet and leads to an increased flexibility of this segment (see Figs. 3a and 3c). This residue is a specific feature of Lck-SH3, whereas alanine or serine are predominantly found at that position in other SH3 domains. The residues 57–59 would correspond to the aforementioned 3_{10} helix, which is often observed in other SH3 domains. Although its existence could be predicted from a preliminary NOE analysis, no consistent hydrogen bonding indicative of such a helix was found during the intermediate stages of the structure calculation. Furthermore, no slowly exchanging hydrogen was observed by the H–D exchange experiment for this segment, and consequently no restraints for hydrogen bonding were applied for these residues. Therefore, the putative 3_{10} helix did not converge into its ideal hydrogen-bonded geometry, but the main chain shows a half turn of helical conformation.

Structure of the putative binding site

Several studies on the structures of SH3–ligand complexes have been published (for a review see Chen and Schreiber, 1995) and confirmed the previous suggestions that the peptide ligands can bind in two different orientations to SH3 domains (Chen et al., 1993; Yu et al., 1994). According to these reports, the ligand interaction site of SH3 modules is formed by a shallow hydrophobic patch located above the short helical half turn and between two flanking loops. In the Lck-SH3 structure, these loops correspond to residues 17–23 (the 'ab loop') and residues 37–40 (the 'bc loop'). They are also called 'RT-Src loop' (because this loop in the tyrosine kinase Src contains arginine and threonine residues whose mutation leads to cell transformation) and 'n-Src loop' (the site of an insertion in the sequence of neuronal Src). Both loops are flexible, with Gly^{39} being the residue with the lowest heteronuclear ^{15}N-$\{^1H\}$-NOE value, and hence the highest internal mobility found within the structured part of the

protein. Thus, the Lck-SH3 ligand binding site seems to be characterized by a local flexibility of these loops. Besides that, in the various SH3 sequences the length of the n-Src loop varies from five to eight residues, with the SH3 domain of PI3K containing an insertion of 15 amino acids. This might be relevant for the discrimination of peptide ligand binding and hence for specificity.

The hydrophobic patch which is framed by the loops comprises conserved aromatic amino acids that correspond to His^{14}, Tyr^{16}, Trp^{41}, Phe^{54}, and Phe^{59} (Fig. 9). His^{14} is a residue which is unique to Lck-SH3, as its position is normally occupied by well-conserved tyrosine or phenylalanine residues in other SH3 domains. The side chain of His^{14} is rather flexible (local rmsd of all its heavy atoms: 1.20 Å, see Fig. 3a) and exposed to the solvent. It forms a salt bridge with an aspartic acid of the SH2 domain at the interdomain surface in the crystal structure (Eck et al., 1994). The other residues mentioned are found to be involved in a network of aromatic interactions such that they are arranged to form three binding pockets for the poly-proline ligand. The well-conserved residues Asp^{23} and Trp^{41}, which establish the first pocket, are positioned to accommodate the positively charged amino acid found at either the N- or C-terminus of the proline-rich peptide. At the same time, Trp^{41} – together with Tyr^{16} and Phe^{54} – is thought to anchor the first proline of the P-X-X-P consensus sequence in the second binding pocket (Chen and Schreiber, 1995). His^{20} near the apex of the 'ab loop' which in the X-ray structure forms a salt bridge with the C-terminal carboxyl group of the phosphotyrosine peptide is in spatial proximity, and its side chain might also be involved in the fixation of the proline-rich peptide. Finally, His^{14} and Phe^{59} – the latter residue is located close to Tyr^{16} and the less well-conserved Phe^{26} – delineate the third binding pocket.

The other aromatic amino acids, Trp^{42} and Phe^{57}, are structurally well-defined via aromatic interactions. Since only the sequence of Lck shows a phenylalanine residue at position 57 whereas serine or alanine are predominantly found in other SH3 domains, it seems that they are not directly involved in the interaction with the poly-proline peptide ligand.

Conclusions

The overall conformational similarity of the SH3 module of p56 Lck to all other published structures of SH3 domains is remarkable. Only small differences have been observed in the RT-Src loop connecting the β-strands a and b. Here, the mini-β-sheet present in the SH3 domains of Grb2C, Fyn, PLC-γ, spectrin, Src, and PI3K – as described in Kohda et al. (1994) – has not been found in the NMR solution structure of Lck. This corresponds to the X-ray analysis of the same protein (Eck et al., 1994) and of Grb2N (Guruprasad et al., 1995).

Also the solution and crystal structures of Lck are almost identical. Minor differences have been observed in the region of the RT-Src and n-Src loops. Correspondingly, these loops are the regions with the highest flexibility in the protein, except for the termini. As they have been implicated in the ligand binding of proline-rich peptides, their inherent increased flexibility might be needed for a fast and selective adaptation of the binding surface in protein–protein interaction.

Acknowledgements

We are grateful to Dr. P. Burn for providing the clone of the SH3 domain of p56 Lck, to Drs. S. Harrison and M. Eck for the refined coordinates of the SH2 SH3 complex and to F. Delaglio for the NMRPipe processing software. We also thank B. Gsell for the excellent biochemistry support.

References

Alexandropoulos, K., Cheng, G. and Baltimore, D. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 3110–3114.

Anil Kumar, Wagner, G., Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.*, **95**, 1 6.

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

Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) *J. Magn. Reson.*, **88**, 425–431.

Bax, A., Vuister, G.W. and Grzesiek, S. (1994) *Methods Enzymol.*, **239**, 79–105.

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

Borchert, T.V., Mathieu, M., Zeelen, J.P., Courtneidge, S.A. and Wierenga, R.K. (1994) *FEBS Lett.*, **341**, 79 85.

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

Chen, J.K. and Schreiber, S.L. (1995) *Angew. Chem. Int. Ed. Engl.*, **34**, 953 969.

Cohen, G.B., Ren, R. and Baltimore, D. (1995) *Cell*, **80**, 237–248.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Driscoll, P.C., Clore, G.M., Marion, D., Wingfield, P.T. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 3542–3556.

Eck, M.J., Atwell, S.K., Shoelson, S.E. and Harrison, S.C. (1994) *Nature*, **368**, 764–769.

Engh, R.A. and Huber, R. (1991) *Acta Crystallogr.*, **A47**, 392–400.

Feng, S., Chen, J.K., Yu, H., Simon, J.A. and Schreiber, S.L. (1994) *Science*, **266**, 1241 1247.

Gerber, P.R. and Müller, K. (1987) *Acta Crystallogr.*, **A43**, 426–428.

Gerber, P.R. and Müller, K. (1995) *J. Comput.-Aided Mol. Design*, **9**, 251–268.

Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870 7872.

Grzesiek, S. and Bax, A. (1992a) *J. Magn. Reson.*, **99**, 201–207.

Grzesiek, S. and Bax, A. (1992b) *J. Am. Chem. Soc.*, **114**, 6291–6293.

Grzesiek, S., Ikura, M., Clore, G.M., Gronenborn, A.M. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 215–221.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Güntert, P., Braun, W. and Wüthrich, K. (1991a) *J. Mol. Biol.*, **217**, 517–530.

Güntert, P., Qian, Y.Q., Otting, G., Müller, M., Gehring, W. and Wüthrich, K. (1991b) *J. Mol. Biol.*, **217**, 531–540.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

Güntert, P., Berndt, K. and Wüthrich, K. (1993) *J. Biomol. NMR*, **3**, 601–606.

Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T.L., Gout, I. and Waterfield, M.D. (1995) *J. Mol. Biol.*, **248**, 856–866.

Hyberts, S.G., Goldberg, M.S., Havel, T.F. and Wagner, G. (1992) *Protein Sci.*, **1**, 736–751.

Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 204–209.

Kabsch, W. and Sander, Ch. (1983) *Biopolymers*, **22**, 2577–2637.

Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Klausner, R.D. and Samelson, L.E. (1991) *Cell*, **64**, 875–878.

Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J. and Inagaki, F. (1993) *Cell*, **72**, 953–960.

Kohda, D., Terasawa, H., Ichikawa, S., Ogura, K., Hatanaka, H., Mandiyan, V., Ullrich, A., Schlessinger, J. and Inagaki, F. (1994) *Structure*, **2**, 1029–1040.

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

Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax, A. (1994) *J. Biomol. NMR*, **4**, 871–878.

Lim, W.A., Richards, F.M. and Fox, R.O. (1994) *Nature*, **372**, 375–379.

MacArthur, M.W. and Thornton, J.M. (1993) *Proteins*, **17**, 232–251.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989b) *J. Am. Chem. Soc.*, **111**, 1515–1517.

Merritt, E.A. and Murphy, M.E.P. (1994) *Acta Crystallogr.*, **D50**, 869–873.

Montelione, G.T., Arnold, E., Meinwald, Y.C., Stimson, E.R., Denton, J.B., Huang, S.G., Clardy, J. and Scheraga, H.A. (1984) *J. Am. Chem. Soc.*, **106**, 7946–7958.

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

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

Müller, K., Ammann, H.J., Doran, D.M., Gerber, P.R., Gubernator, K. and Schrepfer, G. (1988) *Commun. Soc. Chim. Belg.*, **97**, 655–667.

Neri, D., Szyperski, T., Otting, G., Senn, H. and Wüthrich, K. (1989) *Biochemistry*, **28**, 7510–7516.

Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.

Nilges, M. (1993) *Proteins*, **17**, 297–309.

Nilges, M. (1995) *J. Mol. Biol.*, **245**, 645–660.

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

Ostergaard, H.L., Shackelford, D.A., Hurley, T.R., Johnson, P., Hyman, R., Sefton, B.M. and Trowbridge, I.S. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8959–8963.

Pawson, T. (1995) *Science*, **373**, 573–580.

Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–485.

Rudd, C.E., Anderson, P., Morimoto, C., Streuli, M. and Schlossman, S.F. (1989) *Immunol. Rev.*, **111**, 225–266.

Sefton, B.M. (1991) *Oncogene*, **6**, 683–686.

Senn, H., Werner, B., Messerle, B.A., Weber, C., Traber, R. and Wüthrich, K. (1989) *FEBS Lett.*, **249**, 113–118.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., Birge, R.B., Fajardo, J.E., Chou, M.M., Hanahusa, H., Schaffhausen, B. and Cantley, L.C. (1993) *Cell*, **72**, 767–778.

Szyperski, T. (1995) *Eur. J. Biochem.*, **232**, 433–448.

Veillette, A., Foss, F.M., Sausville, E.A., Bolen, J.B. and Rosen, N. (1987) *Oncogene Res.*, **1**, 357–374.

Veillette, A., Bookman, M.A., Horak, E.M. and Bolen, J.B. (1988) *Cell*, **55**, 301–308.

Veillette, A., Bookman, M.A., Horak, E.M., Samelson, L.E. and Bolen, J.B. (1989) *Nature*, **338**, 257–259.

Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.

Wittekind, M., Mapelli, C., Farmer II, B.T., Suen, K.-L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C.A. and Müller, L. (1994) *Biochemistry*, **33**, 13531–13539.

Wu, X., Knudsen, B., Feller, S.M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H. and Kuriyan, J. (1995) *Structure*, **3**, 215–226.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.

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

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

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