Eur Biophys J (1996) 24: 371–380

ARTICLE

Alessandro Pintar · Meike Hensmann
Kornelia Jumel · Maureen Pitkeathly
Stephen E. Harding · Iain D. Campbell

Solution studies of the SH2 domain from the fyn tyrosine kinase: secondary structure, backbone dynamics and protein association

Abstract The SH2 domain from Fyn tyrosine kinase, corresponding to residues 155–270 of the human enzyme, was expressed as a GST-fusion protein in a pGEX-*E. coli* system. After thrombin cleavage and removal of GST, the protein was studied by heteronuclear NMR. Two different phosphotyrosyl-peptides were synthesized and added to the SH2 domain. One peptide corresponded to the regulatory C-terminal tail region of Fyn. Sequence-specific assignment of NMR spectra was achieved using a combination of ^{1}H-^{15}N-correlated 2D HSQC, ^{15}N-edited 3D TOCSY-HMQC, and ^{15}N-edited 3D NOESY-HMQC spectra. By analysis of the α-proton chemical shifts and NOE intensities, the positions of secondary structural elements were determined and found to correspond closely to that seen in the crystal structure of the, homologous, Src-SH2 domain.

To investigate the internal dynamics of the protein backbone, T_1 and T_2 relaxation parameters were measured on the free protein, as well as on both peptide complexes. Analytical ultracentrifugation and dynamic light scattering were employed to measure the effect of concentration and peptide-binding on self-association. The results suggest that, at NMR-sample concentrations, the free protein is present in at least dimeric form. Phosphopeptide binding and lower concentration significantly, but not completely, shift the equilibrium towards monomers. The possible role of this protein association in the regulation of the Src-family tyrosine kinases is discussed.

Key words NMR · Dynamics · Src-homology domain · Secondary structure

Abbreviations

SH	Src homology
GST	glutathione-S-transferase
IPTG	isopropyl-β-D-galactopyranoside
DTT	dithiothreitol
PMSF	phenyl-methyl-sulphonyl-fluoride
TBS	50 mM Tris, 150 mM NaCl, 5 mM DTT, pH 8.0
MWCO	molecular weight cut off
NMR	nuclear magnetic resonance
HSQC	heteronuclear single-quantum correlation
NOESY	nuclear Overhauser effect spectroscopy

Introduction

SH2 domains are protein modules of around 100 amino acids length. They are found in many of the components of intracellular signalling pathways, such as PI 3′kinase, GAP and PLC-γ reviewed in (Pawson and Schlessinger 1993; Pawson 1995; Cohen et al. 1995; Schlessinger 1994). Their hallmark is the ability to bind phosphotyrosine sites, created by the initial events of signal transduction (Songyang and Cantley 1995). The structures of several different SH2 domains have been determined, both free and in complex with phosphopeptides (Waksman et al. 1992; Booker et al. 1992; Waksman et al. 1993; Eck et al. 1993; Xu et al. 1995; Pascal et al. 1994). They exhibit a consensus structure, consisting of a triple-stranded β-sheet followed by a small, double-stranded β-sheet flanked by two α-helices, and with two short N- and C-terminal β-strands. The peptide binding surface, located on one side of the molecule, involves residues from the large β-sheet, the N-terminal α-helix and several loops. The interaction

A. Pintar · M. Hensmann · I. D. Campbell (✉)
Department of Biochemistry, Oxford University,
South Parks Road, Oxford OX1 3QU, UK
(Tel: 01865-275345, Fax: 01865-275253)

K. Jumel · S. E. Harding
National Centre for Macromolecular Hydrodynamics,
Department of Applied Biochemistry & Food Science,
University of Nottingham, Sutton Bonington, LE12 5RD, UK

M. Pitkeathly · I. D. Campbell
Oxford Centre for Molecular Sciences,
New Chemistry Laboratory, University of Oxford,
South Parks Road, Oxford OX1 3QT, UK

resembles the insertion of a two-pronged plug into a socket.

$p59^{fyn}$ is a member of the Src-family of protein tyrosine kinases (Kawakami et al. 1986). These have a consensus modular organisation consisting of a unique N-terminal domain, followed by an SH3, and SH2, and a tyrosine kinase domain. A myristylation site at the N-terminus locates the molecule at the cytoplasmic side of the membrane. The different members of the family provide primary downstream protein kinase activities for a variety of membrane receptors (Veillette and Davidson 1992). Fyn associates, via its unique N-terminal domain, with the ζ-chain of the T-cell receptor (TcR) and appears to be one of the main enzyme activities associated with TcR-signalling (Cooke et al. 1991). Gene-knockout mice, lacking the kinase, are partially defective in TcR/CD3-mediated signalling, particularly in the thymus (Appleby et al. 1992; Stein et al. 1992).

Here, we report sequential resonance assignment, secondary structure determination and ^{15}N NMR relaxation studies of the Fyn-SH2 domain, both alone and complexed with two different peptides. One peptide, called here "tail", corresponds to the inhibitory C-terminal phosphotyrosine site of Fyn, while the "specific" peptide represents a phosphotyrosine sequence from hamster middle-T antigen that has been shown to bind to Fyn with high affinity (Maniatis et al. 1982; Cantley et al. 1991). The ^{15}N relaxation data allow examination of the differential dynamics of the free and peptide-bound forms; these studies are complemented by investigations of the self-association state of the protein by sedimentation equilibrium and dynamic light scattering.

Materials and methods

Protein expression and purification

The Fyn-SH2 domain was expressed in *E. coli* strain BL21 as a fusion protein with glutathione transferase, using pGEX2T plasmids (Pharmacia). The construct was a kind gift of Dr. S. Courtneidge (EMBL, Heidelberg). For the preparation of uniformly ^{15}N-labelled samples, bacteria were grown at 37 °C in a M9 minimal medium (Courtneidge et al. 1991) containing $^{15}NH_4Cl$ as the only nitrogen source. Expression was induced after 5 h. ^{15}N-leucine and ^{15}N-glycine specifically labelled samples were prepared from a M9 medium containing all the amino acids except leucine or glycine, respectively, which were supplemented as their ^{15}N-labelled isotopes (Sigma). Expression was induced with IPTG (isopropyl-β-D-galactopyranoside, Sigma). Four hours after induction, cells were harvested by centrifugation at 5000 rpm for 15 min. Induction times longer than 3–5 h did not improve the amount of protein produced. The supernatant was discarded and the pellet resuspended in Triton-Tris-buffered saline (TTBS: 50 mM Tris, 150 mM NaCl, 0.1% (v/v) Triton X-100; pH 8.0) containing 5 mM DTT (dithiothreitol, Sigma). PMSF (phenyl-methyl-sulphonyl-fluoride, Sigma) was added as a protease inhibitor, and cells were lysed by sonication. The lysed material was centrifuged at 10,000 rpm for 30 min at 4 °C. The supernatant was loaded on a glutathione-Sepharoe 4B (Pharmacia) column, the column washed with TTBS, then with TBS (50 mM Tris, 150 mM NaCl, 5 mM DTT, pH 8.0), and the fusion protein eluted with a 10 mM solution of reduced glutathione in TBS. DTT was routinely used through the purification procedure and in the NMR sample preparation, because of three free cysteines (C106, C107 and C113). The final protein yield from a 5 litre culture of was typically 15 mgs.

Pooled fractions containing GST-SH2 fusion protein were exposed to 2.5 mM $CaCl_2$, and 100 units of human thrombin (Sigma) for cleavage overnight. Cleavage was confirmed by SDS-PAGE (PhastSystem, Pharmacia), and the reaction stopped by addition of PMSF. The solution was concentrated in an Amicon stirred-cell using a 3000 MWCO membrane, loaded onto a Sepharose 100 HR (Pharmacia) size exclusion column, and eluted with TBS buffer. Fractions containing protein were pooled and concentrated with a Centriprep device (Amicon). In our construct, shown in Fig. 1, there is a potential thrombin cleavage PRGT site at residues 36–39 but no evidence of corresponding fragments was seen in SDS-PAGE. A C-terminal PRLTD sequence was, however, effectively cleaved, and the molecular mass of the purified protein is consistent with that of a polypeptide lacking the last three C-terminal residues. The identity of the Fyn-SH2 product was confirmed by N-terminal sequencing and electro-spray mass spectrometry; purity was assessed to be >95%.

For the preparation of NMR samples, the concentrated solution was desalted using a prepacked DG10 column (Biorad) previously equilibrated with 10 mM phosphate buffer containing NaCl (30 mM) and DTT (5 mM), pH 6.0, and eluted with the same buffer. Fractions containing the protein were pooled and concentrated to 0.5 ml using a Centricon device (Amicon). The final solution (approx. 1.5 mM) was made up to 10% D_2O and transferred to an NMR tube.

Preparation of "tail" and "specific" peptides

Both peptides were synthesized on an Applied Biosystems-430A peptide synthesizer using Fmoc chemistry. Thr/Glu residues were protected by ^{t}Bu and Asn/Gln residues by Trityl protecting groups. Couplings were HBTU mediated using a 10 fold excess of the amino acid, and capping was carried out subsequent to each coupling reaction. Phosphotyrosine was incorporated directly into the peptide by the use of N-α–Fmoc-O-phospho-L-tyrosine (purchased from Novabiochem) thus eliminating the need for post-synthetic phosphorylation. The "tail" peptide sequence corresponds to the inhibitory C-terminal region of Fyn, ATEPQ(pY)-QPGEN (Pawson and Schlessinger 1993), while "specific" represents the phosphotyrosine sequence, EPQ(pY)EEI-PIYL, that has been shown to bind to Fyn with high affinity (Maniatis et al. 1982; Cantley et al. 1991). Peptides

were deprotected in 85% TFA (5% H_2O, 5% ethane-dithiol, 5% thioanisol), separated from the resin, and precipitated in iced ether. Each was then purified in one step by reverse phase HPLC.

NMR spectroscopy

All spectra were acquired at 30 °C on home-built 500 and 600 MHz spectrometers, operating at 500.1 or 600.2 MHz for ^1H, and 50.7 or 60.8 MHz for ^{15}N, respectively. ^{15}N-decoupling during acquisition was achieved by GARP-1 phase modulation (Shaka et al. 1985). Sign discrimination in t_1 was achieved using the States time-proportional phase incrementation method (Marion et al. 1989) and presaturation during the recycling delay and spin lock pulses were used to suppress the solvent signal (Messerle et al. 1989). ^{15}N-^1H 2D heteronuclear single quantum correlation (HSQC) experiments with sweep widths of 1300 Hz in the ^{15}N dimension and 8000 Hz in the ^1H dimension were acquired with 32 to 128 scans per increment and $512(t_2)\times128(t_1)$ complex points. ^{15}N-edited, 3D NOESY-HMQC and 3D TOCSY-HMQC experiments were acquired with sweep widths of 1300 Hz in the ^{15}N dimension, and 8000 Hz in both the directly and indirectly detected ^1H dimensions, using 512 complex points in the ^1H acquisition dimension, 128 complex points in the indirected detected ^1H dimension, and 32 complex points in the ^{15}N dimension. All 3D spectra were recorded with 8 scans per increment. A mixing time of 150 ms was used for the 3D NOESY-HMQC experiment, and a spin-lock time of 40 ms (DIPSI-3) for the 3D TOCSY-HMQC. ^1H-^1H 2D NOESY (at different mixing times) and TOCSY spectra on the ^{15}N labelled sample were recorded using a SCUBA sequence (Brown et al. 1988). Broadband ^{15}N decoupling was used during t_1 and the acquisition time. These experiments were recorded using 32 scans per increment, $400(t_1)\times512(t_2)$ complex points, and a sweep width of 8000 Hz.

T_1 and T_2 measurements on the Fyn-SH2 domain were acquired with spectral widths of 12500 Hz (^1H) and 1300 Hz (^{15}N) at 500 MHz. 2048 t_2 data points using pulse schemes previously described (Kay et al. 1989; Buck et al. 1995) and 200–230 t_1 increments were recorded for each time point. T_1 measurements were carried out with relaxation delays of 20, 80, 160, 220, 300, 400, 562, 762, and 1004 ms. Incrementation of the pulse sequence in the T2 measurements resulted in the following relaxation delays: 6.2, 18.8, 31.4, 44.0, 62.8, 81.8, 100.6, and 125.8 ms.

Data were processed using the FELIX 2.3 software package (Biosym) on Sun workstations. Typically, the solvent resonance was eliminated by subtraction of the low frequency component from the FID, a shifted sinebell squared window function was applied, data were zero-filled once and Fourier transformed. A first order baseline correction was usually applied. For 3D spectra, data in the ^{15}N dimension were extended from 32 to 64 complex points using linear prediction. 3D spectra analysis (extraction of 2D strips from 3D spectra and automatic strip comparison) was performed using the XEASY software package (Bartels et al. 1995).

Analytical ultracentrifugation

Sedimentation equilibrium experiments on the Fyn-SH2 domain were performed using a Beckman (Palo Alto, USA) XL-A analytical ultracentrifuge equipped with a monochromator and scanning absorption optics. A rotor speed of 25000 rev/min, a temperature of 25.0 °C and a scanning wavelength of 278 nm were employed. Apparent (i.e. corresponding to a particular loading concentration) weight average molar masses $M_{w,app}$ were evaluated using the M* procedure (Creeth and Harding 1982) and the routine MSTARA (Harding et al. 1992) and a partial specific volume of 0.732 ml/g (calculated from the amino acid sequence). A phosphate chloride buffer (pH = 6.0, I = 0.10) was used as solvent. Baselines were obtained by overspeeding at 50000 rev/min at the end of an ultracentrifuge.

Dynamic Light Scattering

Dynamic light scattering experiments were performed using a fixed-angle (90°) DynaPro 801-photometer (Claes et al. 1992). Highly focussed light from a near infra-red semiconductor laser (780 nm) is scattered by the solution and the intensity fluctuations through Brownian diffusivity collected by an APD (avalanche photodiode) via an optical fibre and then the photon output is analysed via an autocorrelator assuming spherical particle characteristics to give an estimate of the apparent z-average diffusion coefficient; from this, after a Mark-Houwink type of calibration using standard globular proteins (Claes et al. 1992; Harding 1995), an estimate for the apparent molar mass is obtained. Several different values of solvent pH were used (6.0, 6.5, 7.0 and 7.4) and different amounts of specific peptide. A loading concentration of 8 mg/ml was employed.

Results

Sequential assignment

The SH2 domain consensus sequence spans residues from W16 to C113 in our construct (see Fig. 1), leaving an additional 9 residues at the N-terminus and 6 residues at the C-terminus. G1 and S2 come from the thrombin cleavage site; residues P3 to M5 are non-native residues introduced into the construct; A6 to E15 as well as H114 to R119 correspond to the actual linker regions found in full-length Fyn.

^1H-^{15}N-correlated 2D HSQC, ^{15}N-edited 3D TOCSY-HMQC, ^{15}N-edited 3D NOESY-HMQC spectra, together with ^1H-^1H 2D spectra were used to assign the ^1H and ^{15}N

backbone resonances (see Table 1 for the assignment both with and without bound peptides). A section of the NOESY strip-plots and an HSQC spectrum of the Fyn-SH2 domain are shown in Figs. 2 and 3. In the HSQC spectrum, the overall dispersion of the signals in both the ^{15}N and ^{1}H dimensions is good, although at least 14 peaks appear as pairs of very similar chemical shifts in both the ^{15}N and the ^{1}H dimensions. The sharp and very intense peaks in the central portion of the spectrum correspond to unstructured tail residues. Only limited information about the spin system types could be extracted from the 3D TOCSY-HMQC spectra, due to the relatively poor magnetization transfer from the NH to the α and side chain protons, caused by self-association of the SH2 domain (see below). Only the 7 alanines could be identified and, based on the ^{15}N chemical shift, some of the glycines. Hence, the sequential assignment procedure had to rely almost entirely on the identification of the sequential NOEs in the 3D NOESY-HMQC. The identification of unique, easily recognizable residues pairs, such as G50-A51, G78-G79, A102-A103, A103-G104, provided useful anchor points. In this way, nearly 50% of the backbone could be assigned.

To confirm these assignments, and to provide further anchor points, ^{15}N-leucine and ^{15}N-glycine samples were prepared, and the corresponding HSQC spectra acquired. The degree of incorporation of ^{15}N-leucine was fairly low, and a certain degree of random labelling was observed. It was, however, possible to identify the 10 leucine residues (L21, L30, L31, L41, L54, L75, L91, L94, L105, L109) from a HSQC spectrum acquired with 256 scans. For the ^{15}N-glycine specifically labelled sample, the degree of incorporation was high, and the 10 glycines (G19, G22, G34, G38, G50, G64, G78, G79, G104, G116) could easily be identified, as well as 7 serines (S2, S10, S32, S45, S53, S55, S99). With the identification of serine and leucine residues, and the confirmation of the glycines, it was possible to assign 75% of the backbone resonances of the structured region of the polypeptide chain.

Further progress in the sequential assignments was achieved by adding phosphopeptide to the ^{15}N-uniformly labelled sample, and recording a new set of NMR spectra. The induced shifts upon peptide binding resolved some of the peak overlap and the chemical shift changes of the HN resonances in the ^{1}H and ^{15}N frequencies can be correlated with the residues that, from SH2 models, are known to be involved in peptide binding. The shifts differences observed on binding the two peptides are shown in Fig. 4.

```
                    10                  20
G S P H M A P V D S I Q A E E W Y F G K
                    30                  40
L G R K D A E R Q L L S F G N P R G T F
                    50                  60
L I R E S E T T K G A Y S L S I R D W D
                    70                  80
D M K G D H V K H Y K I R K L D N G G Y
                    90                  100
Y I T T R A Q F E T L Q Q L V Q H Y S E
                    110                 120
R A A G L C C R L V V P C H K G M P R L T D
```

Fig. 1 The amino acid sequence of the Fyn-SH2 domain used in this study. The residue numbering used throughout the text is indicated above the sequence

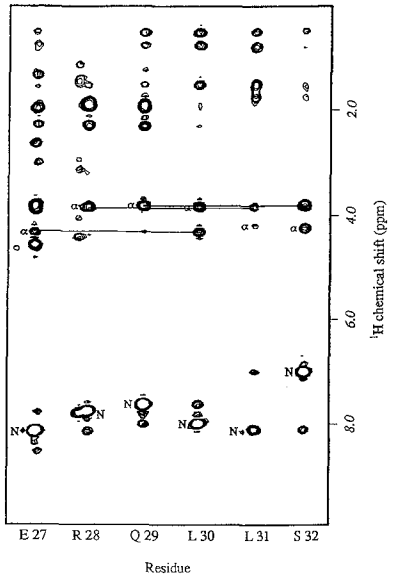

Fig. 2 A part of the strip-plots extracted from a 3D-NOESY-HMQC experiment. The strips shown correspond to residues in the first α-helix. Strong Cα-NH (i, i+3) and NH-NH(i, i+1) connectivities, typical for α-helices, are evident

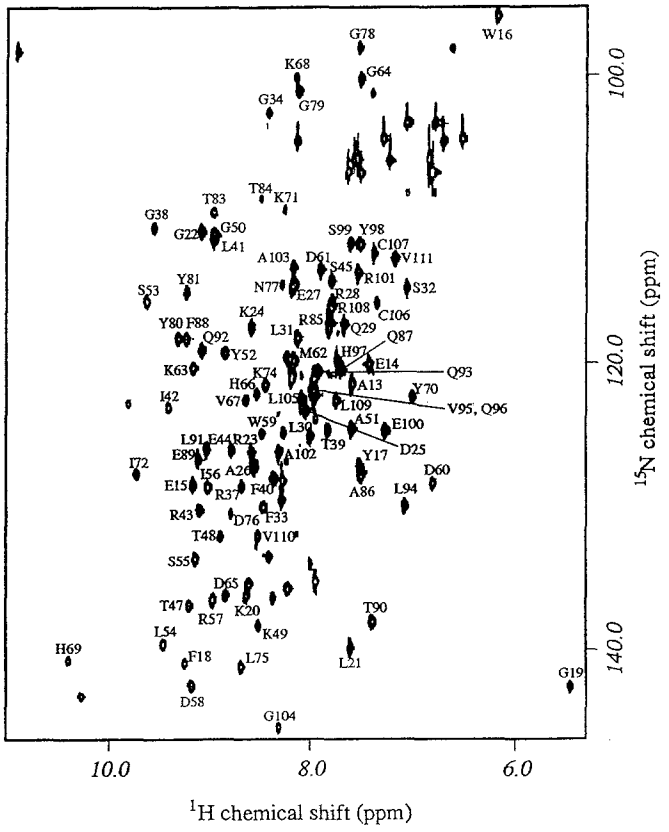

Fig. 3 The assigned ^{15}N-^{1}H HSQC spectrum of the free Fyn-SH2 domain. Peaks corresponding to side-chain amides are not labelled

Table 1 The {^{1}H-^{15}N} HSQC resonance assignments for the Fyn SH2 alone and complexed with either peptide (unassigned resonances are indicated by 'x')

Residue		^{1}H free	^{15}N free	^{1}H Pep1	^{15}N Pep1	^{1}H Pep2	^{15}N Pep2
A	13	7.59	121.44	7.59	121.69	7.57	121.53
E	14	7.41	120.08	7.44	120.14	7.42	120.09
E	15	9.16	128.58	9.15	128.41	9.15	128.37
W	16	6.17	95.82	6.15	95.84	6.15	95.91
Y	17	7.51	127.24	7.49	127.35	7.48	127.16
F	18	9.23	140.96	9.22	140.43	9.25	140.44
G	19	5.46	142.66	5.49	142.82	5.47	142.72
K	20	8.63	136.30	8.68	137.10	8.68	137.15
L	21	7.59	139.90	7.56	139.31	7.54	139.75
G	22	9.06	110.89	9.08	111.05	9.07	110.75
R	23	8.58	126.25	9.57	125.91	8.50	125.51
K	24	8.57	117.53	8.70	118.67	8.73	118.78
D	25	8.04	123.34	8.10	123.28	8.12	123.08
A	26	8.55	127.34	8.56	127.30	8.54	127.49
E	27	8.15	114.63	8.17	114.90	8.17	114.90
R	28	7.77	116.05	7.81	116.21	7.79	116.17
Q	29	7.65	117.36	7.67	117.88	7.66	117.48
L	30	7.99	125.14	8.00	125.50	7.99	125.31
L	31	8.12	118.31	8.09	118.08	8.11	118.16
S	32	7.04	114.82	7.04	115.01	7.03	114.97
F	33	8.51	132.14	8.50	131.94	8.49	132.02
G	34	8.41	102.60	8.43	102.75	8.42	102.76
N	35	x	x	x	x	x	x
P	36	x	x	x	x	x	x
R	37	8.67	128.66	8.67	128.74	8.66	128.70
G	38	9.55	110.60	9.55	110.89	9.53	110.85
T	39	7.82	124.67	7.81	124.36	7.80	124.10
F	40	8.35	128.14	8.35	128.20	8.36	128.25
L	41	8.94	111.29	8.81	110.82	8.79	110.65
I	42	9.40	123.12	9.42	122.95	9.40	122.96
R	43	9.08	130.36	9.09	130.21	9.13	130.50
E	44	8.77	126.10	8.82	126.36	8.81	126.35
S	45	7.78	114.32	7.86	114.67	7.90	114.81
E	46	x	x	x	x	x	x
T	47	9.19	136.98	9.05	136.99	9.90	135.38
T	48	8.88	132.17	9.51	132.25	9.70	132.23
K	49	8.51	138.39	8.51	138.51	8.51	138.65
G	50	8.94	111.29	8.92	111.14	8.93	111.20
A	51	7.57	124.59	7.56	124.63	7.55	124.74
Y	52	8.83	119.29	8.81	117.91	8.89	118.28
S	53	9.63	115.80	9.75	115.85	9.79	117.27
L	54	9.45	139.69	9.60	140.91	9.68	140.38
S	55	9.13	133.77	9.15	134.02	9.13	133.73
I	56	9.00	128.74	9.02	129.23	9.04	129.42
R	57	8.95	136.60	8.97	136.94	8.97	136.83
D	58	9.16	142.63	9.14	142.53	9.13	142.48
W	59	8.47	124.94	8.39	124.47	8.43	124.74
D	60	6.78	128.46	6.72	128.07	6.72	128.20
D	61	7.89	113.61	7.87	112.95	7.87	113.24
M	62	8.15	119.88	8.16	120.17	8.16	120.04
K	63	9.15	120.35	9.18	120.38	9.18	120.48
G	64	7.49	100.26	7.46	100.15	7.46	100.09
D	65	8.83	136.28	8.85	136.55	8.83	136.34
H	66	8.53	122.13	8.51	121.95	8.51	122.26
V	67	8.63	122.59	8.68	122.19	8.67	122.30
K	68	8.13	100.17	8.25	102.24	8.23	104.68
H	69	10.45	140.97	x	x	x	x
Y	70	6.98	122.38	6.97	122.41	6.96	122.29
K	71	8.20	109.50	x	x	x	x
I	72	9.72	127.83	9.91	126.36	9.92	125.98
R	73	x	x	x	x	x	x
K	74	8.44	121.53	8.41	121.44	8.38	121.13
L	75	8.68	141.25	8.64	141.45	8.60	141.41
D	76	8.78	130.58	8.79	131.06	8.79	131.53
N	77	8.27	114.64	8.28	114.58	8.31	114.76
G	78	7.51	98.01	7.49	98.02	7.49	98.02
G	79	8.11	101.10	8.13	101.35	8.15	101.55
Y	80	9.30	118.33	9.30	118.17	9.30	117.87
Y	81	9.21	115.15	9.17	114.84	9.13	114.76

Table 1 (continued)

Residue	^{1}H free	^{15}N free	^{1}H Pep1	^{15}N Pep1	^{1}H Pep2	^{15}N Pep2
I 82	8.22	135.80	8.16	135.86	8.19	135.55
T 83	8.95	109.51	x	x	8.92	113.32
T 84	8.44	108.98	x	x	x	x
R 85	7.80	117.22	7.85	118.62	7.74	118.40
A 86	7.49	127.97	7.68	129.85	7.93	132.47
Q 87	7.69	120.42	7.69	119.68	7.69	118.36
F 88	9.22	118.37	9.21	118.77	9.23	118.97
E 89	9.08	126.75	9.09	126.75	9.09	126.55
T 90	7.38	138.11	7.37	138.19	7.34	137.77
L 91	9.01	126.03	8.96	126.04	8.91	126.75
Q 92	9.06	119.16	9.08	119.13	9.10	118.93
Q 93	7.91	120.62	7.90	120.54	7.93	120.52
L 94	7.06	130.05	7.04	130.02	7.06	130.11
V 95	7.97	121.84	7.96	121.94	7.97	121.71
Q 96	7.94	122.23	7.96	122.18	7.99	122.14
H 97	7.72	119.85	7.73	121.29	7.68	119.73
Y 98	7.50	111.79	7.46	111.84	7.44	112.30
S 99	7.59	111.71	7.62	111.51	7.56	110.33
E 100	7.25	124.75	7.25	124.99	7.27	126.19
R 101	7.52	113.80	7.52	113.91	7.47	113.15
A 102	8.31	126.25	8.30	126.25	8.33	125.64
A 103	8.15	113.39	8.15	113.69	8.22	110.55
G 104	8.31	145.64	8.29	146.37	8.12	141.76
L 105	8.08	122.58	8.06	121.40	8.01	123.05
C 106	7.34	115.90	7.39	116.22	7.33	117.32
C 107	7.36	112.39	7.36	112.50	7.31	112.86
R 108	7.80	116.40	7.78	117.17	7.79	117.07
L 109	7.74	122.63	7.67	122.61	7.82	122.90
V 110	8.46	130.12	8.38	130.86	8.37	130.91
V 111	7.15	112.78	7.71	113.09	7.16	112.94

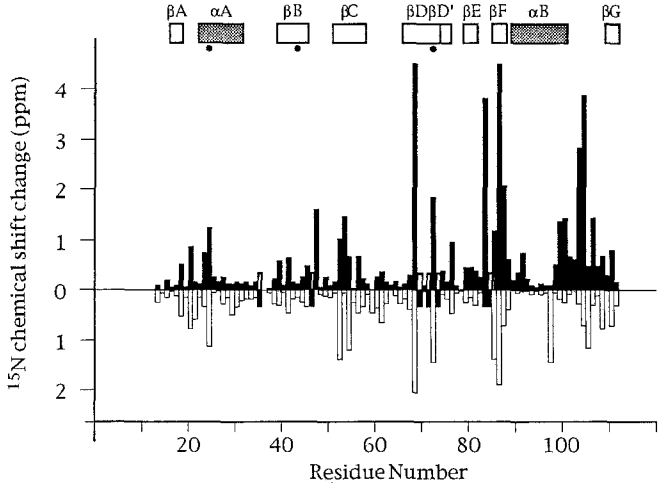

Fig. 4 The chemical shift changes observed in the ^{15}N dimension upon binding of the specific (*black, pointing up*) and regulatory (*open, pointing down*) peptides. Unassigned residues are represented by bars of arbitrary height with opposite shading. The location of secondary structure elements is undicated above the plot. *Black circles* indicate residues expected to be involved in anchoring the phosphotyrosine

Secondary structure

The position of secondary structural elements was determined from the deviations of the α proton chemical shifts from their random coil values, and from the relative intensities of the sequential αN(i, i+1) and NN(i, i+1) NOE connectivities. A Wishart (Wishart et al. 1992) plot of the chemical shifts values is reported in Fig. 5A and a schematic representation of the sequential NOEs in Fig. 5B. In this last plot, a value of +1 is assigned when the αN(i, i+1) NOE is stronger than the corresponding NN(i, i+1), a value of –1 when the NN(i, i+1) NOE is stronger than the corresponding αN(i, i+1) NOE, and a value of 0 when the two peaks are more or less of the same intensity. Stretches of +1 are indicated of extended structure (β-strand), whereas stretches with values of –1 are indicative of helical regions or tight turns. This simplified representation of sequential NOEs is effective in rendering the overall secondary structure features and matches well with the Wishart plot.

No major differences were observed in the NOE patterns of strips extracted from the 3D NOESY-HMQC spectra of the protein in its free- and peptide-bound form, suggesting that no changes in secondary structure occur upon peptide binding.

Dynamic behaviour with and without peptide

Measurements of the T_1 and T_2 relaxation parameters were performed on the free protein as well as on both peptide complexes. Figure 6 shows a comparison of the T_1/T_2 profiles between the free protein and the two complexes. It is immediately apparent that there is a global change in the T_1/T_2 ratio, arising largely from a significant drop in T_2

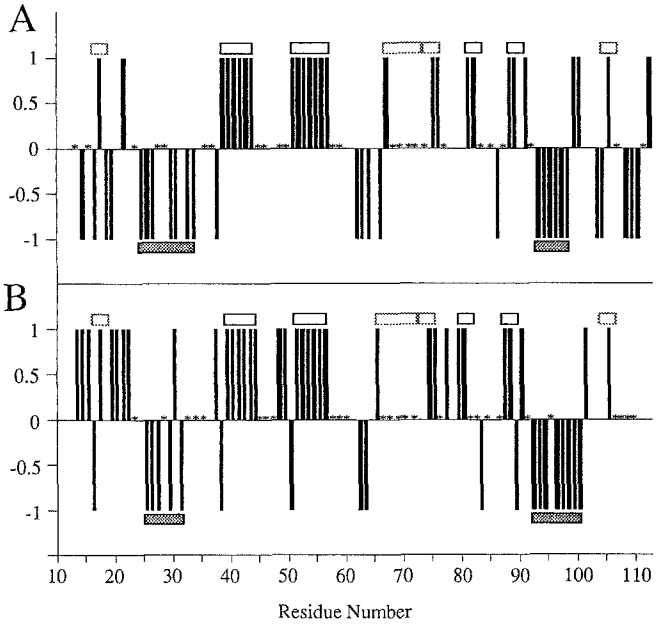

Fig. 5 A Wishart plot (**A**), based on the α-proton chemical shifts, and the distribution of αN(i,i+1) and NN(i,i+1) NOE intensities (**B**) along the backbone of the Fyn-SH2 domain. In (A), stretches of the consecutive residues assigned "+1" values corresponded to extended regions of the polypeptide chain (β-strands), those assigned "–1" values to α-helical regions. In (B), "+1" is assigned when the αN(i,i+1) NOE is stronger than the NN(i,i+1), "–1" for the opposite case. Where both intensities are similar, a value of "0" is given. In this plot, stretches of residues with a "2+1" values correspond to extended regions of the polypeptide chain (β-strands), while those with "–1" values correspond to α-helical regions. In both plots, asterisks indicate residues with unassigned or not unambiguously assigned backbone resonances. Inferred secondary structural elements are indicated by bars (β-strands: *white*, α-helices: *shaded*)

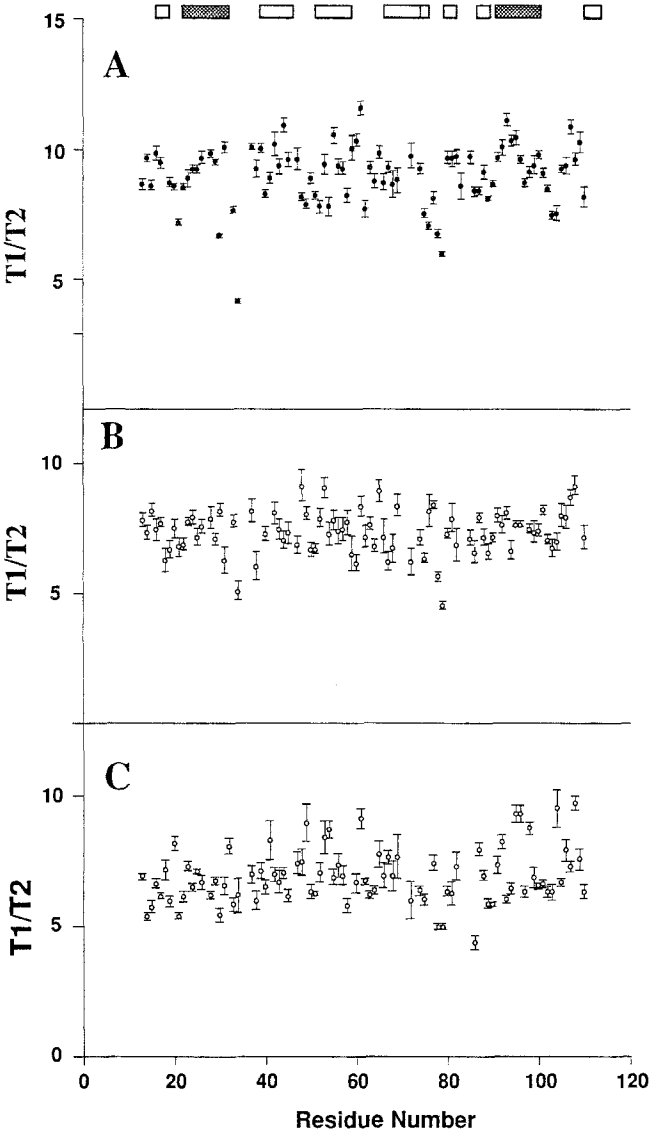

Fig. 6 Plots of the T_1/T_2 ratio against amino acid sequence for **A** the free protein (filled circles), **B** the tail-peptide-bound protein (open circles), and **C** the specific peptide-bound form (*open circles*) bound forms. The position of secondary structural elements is indicated above the plots

values (results not shown) from an average of about 90 ms to about 70 ms upon binding to either peptide. The change in T_1/T_2 ratio can be directly related to a change in apparent overall correlation time. The average T_1/T_2 ratio for the free protein is 9.1 ± 0.7, for the specific peptide-bound form it is 6.8 ± 0.6, and for the tail peptide-bound from 7.4 ± 0.5 (these averages were calculated not including values that were more than two standard deviations different from the average). Using standard Lipari & Szabo treatment (Lipari and Szabo 1982), assuming isotropic motion, this corresponds to apparent τ_c values of 11.1 ± 0.4 ns, 9.6 ± 0.4 ns and 10.0 ± 0.3 ns respectively. Comparison with other NMR studies (e.g. Farrow et al. 1994) suggests that these correlation times correspond to effective molecular weights in solution of approximately 25 kDa for the free protein and around 22 kDa for the peptide-bound complexes.

In the free protein, slightly higher than average mobility (lower T_1/T_2 ratio) is detected in most of the loops, but particularly between strands βD and βE. This region retains its increased flexibility in both peptide-bound forms. In contrast, the T_2 values for residues in the BC loop seem to fall within the average in the peptide bound protein.

Assay of sample associative phenomena

To investigate the possibility of concentration, or pH, dependent dimerization of the SH2 domain, sedimentation equilibrium and dynamic light scattering experiments were performed. Figure 7 shows a plot of the apparent molecular weight, $M_{w,app}$, as a function of concentration, c, obtained from the sedimentation equilibrium experiments on the SH2 domain. Since the molecular weight of the SH2 monomer, as found from mass spectrometry, is 13,983 Da the protein is clearly showing dimerization behaviour, even at the lowest concentration examined (0.23 mg/ml). The strength of the dimerisation can be estimated by fitting the $M_{w,app}$ (Da = g/mol) as a function of concentration, c, in g/l

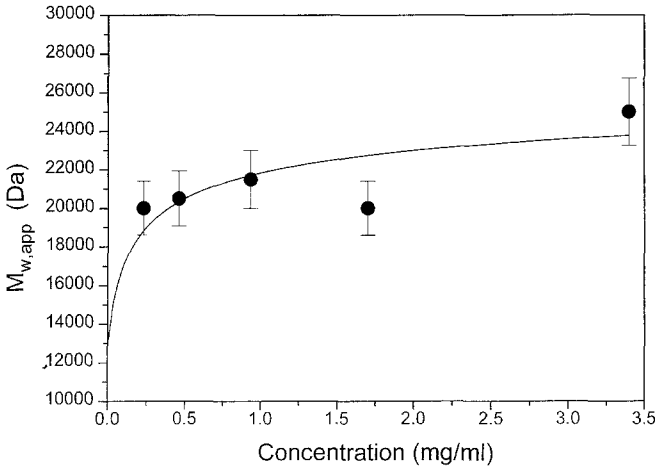

Fig. 7 Plot of weight average apparent molecular weight, $M_{r,app}$, from sedimentation equilibrium analysis versus ultracentrifuge cell loading concentration, c (mg/ml). The line fitted is for a non-ideal dimerisation with the monomer molecular weight, $M_1 = 13$ kDa

(Deonier and Williams 1970) in terms of the dimerisation constant, k_2 (l/g):

$$\{R^2/[(2(1\text{-}BM_1Rc)\text{-}R]^2\} - 1 = 4k_2c$$

where $R = M_{w,app}/M_1$ and B is the 2nd thermodynamic (non-ideality) viral coefficient. From Fig. 7, an estimate for k_2 of (6.4 ± 3.6)l/g is obtained, corresponding to a molar dimerisation constant $K_2 = (M_1/2) \cdot k_2$ of (0.041 ± 0.023)μM^{-1} (Kim et al. 1977) and a molar dissociation constant, K_D of (25 ± 14)μM.

For measurements at concentrations comparable to that used for NMR, dynamic light scattering measurements were employed at 8 mg/ml. At these concentrations, application of the absorption optical system of the ultracentrifuge for sedimentation equilibrium analysis, becomes unreliable, even with the use of short path length cells and "off absorption peak" wavelengths. Provided the protein is not too asymmetric, dynamic light scattering at a fixed angle (90°) can provide a reasonable estimate of the apparent molecular weight, $M_{r,app}$, (Claes et al. 1992), relative to globular protein standards.

Comparative dynamic light scattering measurements were performed in phosphate buffer (I = 0.10), as for sedimentation equilibrium analysis, but with a range of pH's and with the protein in the presence of and absence of the specific peptide. In the absence of peptide the protein has an apparent relative molecular weight $M_{r,app}$, of ≈35 kDa at all the pH's examined (6.0, 6.5, 7.0 and 7.4). Upon addition of peptide, the $M_{r,app}$ decreases in each case to ≈30 kDa, but most dramatically at pH 7.4 (22 kDa). A third measurement was conducted at pH 6.0, after addition of an excess of peptide (ratio of specific peptide to protein 3:2). Under these conditions the apparent relative molecular weight (24 kDa) is similar to that for the lower peptide concentration at pH 7.4. Control experiments of the free peptide alone showed that it made no significant contribution to the autocorrelation data.

Discussion

Comparison of the secondary structural assignments obtained here (see Fig. 5) with those expected from homology with the Src SH2 domain (Fig. 8), shows not major differences. There are, however, a few discrepancies. The first short β-strand (W16-F18) is not well defined, despite the fact that the NH resonances of residues W16 and G19 are strongly shifted. Strands β2, β3 and β4 seem to be shifted by one or two residues towards the amino terminus. Finally, helix B might be shorter than expected, extending from residue 92 to 97.

The results from NMR dynamics measurements, analytical ultracentrifugation and dynamic light scattering suggest that the free protein forms at least dimers in solution not only at the pH and concentrations encountered in the NMR samples but also as low as 50 μM. It behaves like a protein of larger size in all the studies presented here. Binding of either peptide to the SH2 domain appears to sigificantly, but not completely, shift the equilibrium towards the monomeric form. The apparent molecular weight under those conditions is decreased, but is still larger than the expected value (14 kD). These results are consistent with the observation that the spectra of the protein complexed with the peptide appear less broadened than those of the free protein. A similar phenomenon was observed in an NMR study of the SH2 domain from PLC-γ(Farrow et al. 1994). In that study, however, the peptide seemed to shift the monomer-dimer equilibrium completely to monomers, since the apparent molecular weight upon peptide addition decreased to the expected value of around 13 kDa. In the case of the Fyn SH2, the extent of dimerisation in the absence of phosphopeptide can be explained in terms of a molar dissociation constant of around 25 μM (see Fig. 7). This is low enough that the dimerisation could have some in vivo significance.

Judging from the apparent disrupting effect of the peptide, dimerisation is likely to engage part of the same surface of the protein as peptide binding. In particular a hydrophobic patch defining the (+3) binding pocket, accommodating an isoleucine in the structures of Src and Lck complexed with high-affinity peptides, may be involved. This would also provide an explanation for the even stronger than average broadening of the NMR resonances that was observed in this region of the protein and the great difficulties in assigning residues along the βD sheet. The peptide appears to have a stronger disrupting effect on self-association at higher pH. This may be due to the protonation state of the tyrosyl phosphate group, which will be more negatively charged at higher pH and allow stronger binding of the peptides.

The tendency of the SH2 domain to form dimers in solution may have implications for the regulation of the Fyn tyrosine kinase in vivo (Panchamoorthy et al. 1994). It corresponds well to observations made in the crystal structure of the regulatory (SH3 and SH2) domains of Lck, another member of the Src-family of tyrosine kinases with strong sequence homology to Fyn (Eck et al. 1993). In the crys-

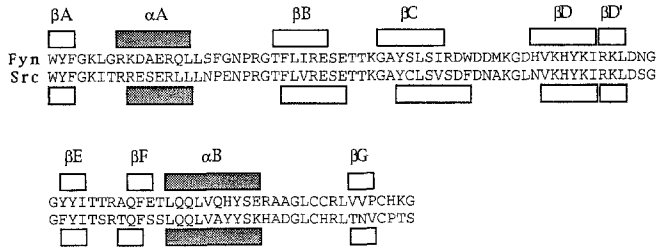

Fig. 8 A comparison of the Src and Fyn-SH2 sequences with known (Src, from Waksman et al. 1992) and proposed (Fyn) secondary structural elements. Light grey shading in the proposed (Fyn) elements indicates uncertainties arising from discrepancies between the Wishart and NOE plots (Fig. 5)

tal, the domains are present as head-to-tail dimers. The peptide binding surface of the SH2 domain is seen to form part of the dimer interaction surface. A second structure of the domains in complex with the tail peptide shows that this peptide does not appear to dissociate the dimer. In our case, the SH3 domain is not present and the specific peptide could engage the entire potential dimer interaction surface. It is thus conceivable that, while the tail peptide might promote dimer formation of the whole tyrosine kinase in vivo, the specific peptide could disrupt the dimer and allow a reorganisation of the modules with respect to each other that could lead to enzyme activation.

The dynamics measurements indicate that there is only a small extent of loop mobility; the BC loop in the peptide-bound protein appears to lose the small amount of flexibility it had in the uncomplexed form. This is consistent with the mode of peptide binding described elsewhere (Waksman et al. 1993; Eck et al. 1993), involving closure of this loop over the phosphotyrosine site. It is also consistent with the kinds of shifts observed when the peptides bind (Fig. 4). The low degree of internal flexibility correlates well with the results of the heteronuclear nOe measurements on the p85 SH2 domain (Hensmann et al. 1994), as well as with the observed dynamic behavior of the PLC-γSH2 domain. This is not surprising, since the loops in these compact domains are relatively short, in contrast to, for example, the highly mobile RGD loop in the fibronectin type III module (Main et al. 1992).

Acknowledgements This is a contribution from the Oxford Centre for Molecular Sciences that is supported by MRC, BBSRC and EPSRC. The work was also supported by the Wellcome Trust and EMBO and EU made contributions to Alessandro Pintar's financial support. Many colleagues have contributed to this work but we especially thank Craig Morton for help with the sample preparation and Jonathan Boyd for advice on NMR.

References

Appleby MW, Gross JA, Cooke MP, Levin SD, Qian X, Perlmutter RM (1992) Defective T cell receptor signalling in mice lacking the thymic isoform of p59^{fyn}. Cell 70:751–763

Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K (1995) The program XEASY for computer supported NMR spectral analysis of biological molecules. J Biomolec NMR 5:1–10

Booker GW, Breeze AL, Downing AK, Panayotou G, Gout I, Waterfield MD, Campbell ID (1992) Structure of on SH2 domain of the p85α subunit of phosphatidylinositol-3-OH kinase. Nature 358:684–687

Brown SC, Weber PL, Mueller L (1988) Toward complete ^{1}H NMR spectra in proteins. J Mag Res 77:166–169

Buck M, Boyd J, Redfield C, MacKenzie DA, Jeenes DJ, Archer DB, Dobson CM (1995) Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. Biochem 34:404–4055

Cantley LC, Auger KR, Carpenter C, Duckworth B, Graziani A, Kapellar R, Soltoff S (1991) Oncogenes and signal transduction. Cell 64:281–302

Claes P, Dunfor M, Kenney A, Vardy P (1992) An on-line dynamic light scattering instrument for macromolecular characterisation. In: Harding SE, Sattelle DB, Bloomfield VA (eds) Laser light scattering in biochemistry. Royal Soc Chem, Cambs

Cohen GB, Den R, Baltimore D (1995) Modular binding domains in signal transduction proteins. Cell 80:237–248

Cooke MP, Abraham KM, Forbush KA, Perlmutter RM (1991) Regulation of T cell receptor signalling by a src family protein-tyrosine kinase (p59^{fyn}). Cell 65:281–291

Courtneidge SA, Goutebroze L, Cartwright A, Heber A, Scherneck S, Feunteun J (1991) Identification and characterization of the hamster polyomavirus middle T antigen. J Virol 65:3301–3308

Creeth JM, Harding SE (1982) Some observations on a new type of point average molecular weight. J Biochem Biophys Meth 7:25–34

Deonier RC, Williams JW (1970) Self-association of muramidase (lysozyme) in solution at 25°, pH 7.0, and I=0.20. Biochem 9:4260–4267

Eck MJ, Shoelson SE, Harrison SC (1993) Recognition of a high-affinity phosphotyrosyl peptide by the src homology-2 domain pf p56^{lck}. Nature 362:87–91

Eck MM, Atwell SK, Shoelson SE, Harrison SC (1994) Structure of the regulatory domains of the Src-family tyrosine kinase Lck. Nature 368:764–769

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. Biochem 33:5984–6003

Harding SE (1995) On the hydrodynamic analysis of macromolecular conformation. Biophys Chem 55:69–93

Harding SE, Horton JC, Morgan PJ (1992) MSTRA: A FORTRAN programme for the model independent molecular weight analysis of macromolecules using low speed or high speed sedimentation equilibrium. In: Harding SE, Rowe AJ, Horton JC (eds) Analytical ultracentrifugation in biochemistry and polymer science. Royal Soc Chem, Cambs

Hensmann M, Booker GW, Panayotou G, Boyd J, Linacre J, Waterfield M, Campbell ID (1994) Phosphopeptide binding to the N-terminal SH2 domain of the p85α subunit of PI 3′-kinase: a heteronuclear NMR study. Prot Sci 3:1020–1030

Kawakami T, Pennington CY, Robbins KC (1986) Isolation and oncogenic potential of a novel human src-like gene. Mol Cell Biol 6:4105–4201

Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studies by ^{15}N inverse detected heteronuclear NMR: application to staphylococcal nuclease. Biochem 28:8972–8979

Kim H, Deonier RC, Williams JW (1977) The investigation of self-association reactions by equilibrium ultracentrifugation. Chem Rev 77:659–690

Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. J Am Chem Soc 104:4546–4570

Main AL, Baron M, Harvey TS, Boyd J, Campbell ID (1992) The three dimensional structure of the tenth type III module of fibro-

nectin: an insight into RGD mediated interactions. Cell 71:671–678
Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York, USA
Marion D, Ikura M, Tschudin R, Bax A (1989) Rapid recording of 2D NMR spectra without phase-cycling. Application to the study of hydrogen exchange in proteins. J Mag Res 85:393–399
Messerle BA, Wider G, Otting G, Weber C, Wüthrich K (1989) Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with H_2O solutions. J Magn Reson 85:608–613
Panchamoorthy GM, Fukazawa T, Stolz L, Payne G, Reedquist K, Shoelson S, Songyang Z, Cantley L, Walsh C, Band H (1994) Physical and functional interactions between SH2 and SH3 domains of the Src family protein tyrosine kinase $p59^{fyn}$. Mol Cell Biol 14:6372–6385
Pascal SM, Singer AU, Gish G, Yamazaki T, Shoelson SE, Pawson T, Kay LE, Forman-Kay JD (1994) Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-γ1 complexed with a high affinity binding peptide. Cell 77:461–472
Pawson T (1995) Protein modules and signalling networks. Nature 373:573–580
Pawson T, Schlessinger J (1993) SH2 and SH3 domains. Curr Biol 3:434–442
Schlessinger J (1994) SH2/SH3 signalling proteins. Curr Op Gen Dev 4:25–30
Shaka AJ, Barker PB, Freeman R (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. J Mag Res 64:547–552
Songyang Z, Cantley LC (1995) Recognition and specificity in protein tyrosine kinase-mediated signalling. TIBS 20:470–475
Stein PL, Lee HM, Rich S, Soriano P (1992) $p59^{fyn}$ mutant mice display differential signalling in thymocytes and peripheral T cells. Cell 70:741–750
Veillette A, Davidson D (1992) Src-related protein tyrosine kinases and T-cell receptor signalling. TIG 8:61–66
Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M, Resh MD, Rios CB, Silverman L, Kuriyan J (1992) Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. Nature 358:646–653
Waksman G, Shoelson SE, Pant N, Cowburn D, Kuryian J (1993) Binding of a high affinity phosphotyrosyl peptide to the src SH2 domain of $p56^{lck}$. Cell 72:779–790
Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochem 31:1647–1651
Xu RX, Word M, Davis DG, Ring MJ, Willard Jr. DH, Gampe Jr. RT (1995) Solution structure of the human pp60c-src SH2 domain complexed with a phosphorylated tyrosine pentapeptide. Biochem 34:2107–2121