Three-Dimensional Solution Structure and ^{13}C Assignments of Barstar Using Nuclear Magnetic Resonance Spectroscopy‡

Michael J. Lubienski, Mark Bycroft,* Stefan M. V. Freund, and Alan R. Fersht

MRC Unit for Protein Function and Design, Cambridge Centre for Protein Engineering, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Received March 1, 1994; Revised Manuscript Received May 9, 1994®

ABSTRACT: We present the high-resolution solution structure and ^{13}C assignments of wild-type barstar, an 89 amino acid residue polypeptide inhibitor of barnase, derived from heteronuclear NMR techniques. These were obtained from measurements on unlabeled, uniformly ^{15}N- and $^{13}C/^{15}N$-labeled, and 10% ^{13}C-labeled barstar samples that have both cysteines (at positions 40 and 82) fully reduced. In total, 30 structures were calculated by hybrid distance geometry–dynamical simulated annealing calculations. The atomic rms distribution about the mean coordinate positions is 0.42 Å for all backbone atoms and 0.90 Å for all atoms. The structure is composed of three parallel α-helices packed against a three-stranded parallel β-sheet. A more poorly defined helix links the second β-strand and the third major α-helix. The loop involved in binding barnase is extremely well defined and held rigidly by interactions from the main body of the protein to both ends and the middle of the loop. This structure will be used to aid protein engineering studies currently taking place on the free and bound states of barstar and barnase.

One of the tightest and fastest protein–protein interactions known is the one-to-one interaction between barnase and barstar (Hartley, 1993; Schreiber & Fersht, 1993). Barnase is an extracellular endoribonuclease from *Bacillus amyloliquefaciens* (Hartley, 1989). Any barnase that is expressed internally is lethal to the cell. *B. amyloliquefaciens*, however, internally expresses an 89-residue polypeptide inhibitor of barnase, barstar, which interacts with barnase with a dissociation constant (K_d) of 2×10^{-14} M (Hartley, 1993; Schreiber & Fersht, 1993). In order to express recombinant barnase in *Escherichia coli*, the barstar gene must be present on the same construct (Hartley, 1989). Barstar has also been used in other systems to neutralize the cytotoxic nature of heterologously expressed barnase, for example, to reverse the effect of male plant sterility caused by the expression of barnase (Mariani *et al.*, 1992).

The interaction between barstar and barnase is being studied in this laboratory as a model system to study molecular recognition by protein engineering (Schreiber & Fersht, 1993), NMR[1] (Jones *et al.*, 1993), and X-ray crystallography (Buckle *et al.*, 1994). Until now, information from protein–protein interactions has come mainly from protease–inhibitor and antibody–antigen complexes (Janin & Chothia, 1990). Recently, the first structure of a ribonuclease-polypeptide inhibitor complex (between barnase and the Cys-40 → Ala and Cys-82 → Ala double mutant of barstar) was solved to 2.6 Å (Guillet *et al.*, 1993) and independently to 2.0 Å in this laboratory (Buckle *et al.*, 1994) by X-ray crystallography. An advantage of the barnase–barstar system is that the complex is comprised of only 199 amino acids and thus can also be readily studied by heteronuclear NMR. The effect of barstar binding to barnase has been analyzed using this approach (Jones *et al.*, 1993).

A requirement of studying the complex is that both partners must be structurally characterized separately. Barnase has already been well characterized with crystal (Buckle & Fersht, 1994; Buckle *et al.*, 1993; Chen *et al.*, 1993; Mauguen *et al.*, 1982) and solution (Bycroft *et al.*, 1991) structures. Until now, there has been no structural information available on the free structure of barstar. The only structural work done on a free ribonuclease inhibitor has been on a porcine protein which is substantially larger than barstar (Kobe & Deisenhofer, 1993).

A previous study by NMR using uniformly ^{15}N-labeled barstar reported the assignment of all but five of the backbone ^{1}H and ^{15}N resonances (Lubienski *et al.*, 1993). The study also showed that the protein is composed of three parallel helices, one poorly-defined helix, and a three-stranded parallel β-sheet using the pattern of the available NOEs and equilibrium hydrogen exchange and dihedral angle information. At that stage, the derivation of a three-dimensional structure was not possible because of the overlap of the resonances in the methyl region of the homonuclear NOESY spectrum.

In this paper, we report the assignment of the ^{13}C resonances of barstar by two- and three-dimensional heteronuclear NMR using 10% ^{13}C-labeled and uniformly $^{13}C/^{15}N$-labeled barstar samples. In addition, we present the high-resolution structure of barstar and comment on its novel structure.

MATERIALS AND METHODS

Barstar Preparation. Construction of the plasmid containing the barstar gene, pML2bs, has been described (Jones *et al.*, 1993). The plasmid was transformed into the *E. coli* strain TG2. To express barstar, a M9 minimal medium (containing 4 g/L glycerol instead of glucose) starter culture containing ampicillin (50 μg/mL) was grown overnight from

‡ The atomic coordinates have been deposited in the Brookhaven Protein Data Bank (1BTA and 1BTB).

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* July 1, 1994.

[1] Abbreviations: pfu, plaque forming units; ppm, parts per million; IPTG, isopropyl β-D-thiogalactoside; NMR, nuclear magnetic resonance; 2D, two dimensional; 3D, three dimensional; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; HMQC, ^{1}H-detected heteronuclear multiple-quantum correlation; FID, free induction decay; rms, root mean square.

0006-2960/94/0433-8866$04.50/0

a frozen stock of TG2[pML2bs]. Five milliliters of the starter culture was added to a 2-L baffled flask containing 500 mL of the M9 media described above. The culture was grown into early log phase at 37 °C with shaking. At A_{600} = 0.3–0.4, 5 pfu of M13 phage containing the gene for T7 polymerase under the control of the *lac* promoter (obtained from the Invitrogen Corp. XPRESS System kit) was added along with 0.2 mM IPTG, and the culture was grown overnight. All subsequent operations were carried out at 4 °C. The cell pellet was resuspended in 50 mM Tris-HCl buffer containing 1 mM phenylmethanesulfonyl fluoride at pH 8. The solution was sonicated and was centrifuged in a Sorvall SS34 rotor (1 h, 18 000 rpm, 4 °C). The proteins in the supernatant were precipitated with 80% ammonium sulfate. This was centrifuged (Sorvall SS34 rotor, 18 000 rpm, 30 min, 4 °C) and the pellet was dissolved in a minimal volume of 50 mM Tris-HCl with 50 mM NaCl, pH 8. The resulting solution was loaded onto a Pharmacia Superdex75 column and eluted at a flow rate of 1 mL/min using the same buffer. Column fractions which fully inhibited barnase at concentrations of 0.1 mg/mL were combined for an ion exchange step using a Pharmacia Mono-Q 10/10 column equilibrated in 50 mM Tris-HCl, pH 8. The column was eluted using the same buffer containing 1 M NaCl. Barstar was eluted between 300 and 350 mM NaCl. The resulting peak was a single band when analyzed by SDS–PAGE. The protein was flash frozen in liquid N_2 and stored at −70 °C. The two cysteines in the protein were shown to be almost 100% reduced using iodoacetamide/iodoacetic acid blocking and the Ellman assay (Hollecker & Creighton, 1989). Sixty to eighty milligrams of pure protein could be obtained per liter of M9. The 10% ^{13}C-labeled protein was prepared by adding 90% of the total glycerol content in the medium as unlabeled material and 10% as 98% atom ^{13}C-labeled glycerol (Martek). The ^{13}C/^{15}N doubly labeled protein was prepared using 98% atom ^{13}C-labeled glycerol as the sole carbon source in the medium and 99.3% atom ^{15}N-labeled ammonium chloride (Isotec Inc.) as the sole nitrogen source. Two 3 mM samples of pure barstar could be produced from half a liter of culture.

NMR Sample Preparation. Samples were produced by extensively dialyzing the products against 10 mM potassium phosphate, pH 6.7, and concentrating the protein in an Amicon Centriprep 3 to 2.5–3.5 mM in 450 μL. To prepare samples in 90% H_2O/10% D_2O, we added to the 450-μL solution 50 μL of 10 mM potassium phosphate, pH 6.7, that had been exchanged into D_2O by repeated lyophilization. To prepare samples in D_2O, the protein was concentrated to approximately 1.5 mM in 1 mL of buffered H_2O and diluted to 5 mL with the deuterated phosphate buffer. This was repeated three times to reduce further the H_2O content. The sample was then concentrated to 500 μL. All samples contained 0.05% sodium azide.

NMR Measurements. All NMR experiments were carried out on a Bruker AMX 500 spectrometer equipped with a triple-resonance ^{1}H/^{15}N/^{13}C probe optimized for proton detection. All spectra were recorded at 305 K. In the 3D HNHB experiment (Archer *et al.*, 1991), the spectral width and number of points acquired were 4000 Hz and 256 complex points in ^{1}H(F_1), 5000 Hz and 190 real points in ^{1}H(F_2), and 2200 Hz and 64 real points in ^{15}N(F_3). Solvent suppression was achieved by on-resonance presaturation of the solvent signal, and the transmitter frequency was switched to the middle of the NH region immediately prior to acquisition. The ^{15}N carrier frequency was placed at 116.3 ppm. Sixteen transients were acquired per increment. Chemical shifts were measured relative to external TSP (0.0 ppm ^{1}H) and ^{15}N ammonium sulfate (24.93 ppm). All other homonuclear and ^{15}N heteronuclear NMR experiments on barstar have been described in Lubienski *et al.* (1993).

The two three-dimensional ^{13}C heteronuclear experiments that were initially performed with the doubly labeled sample in D_2O were the ^{13}C–^{1}H HCCH-COSY (Kay *et al.*, 1990) and the ^{13}C–^{1}H NOESY-HMQC (Ikura *et al.*, 1990) experiments. The COSY experiment was used to assign the ^{13}C shifts of many C^α, C^β, and aromatic carbons, and all methyl carbons. The NOESY experiment was used to assign side chain–side chain and side chain–backbone NOEs. The spectral width and number of points acquired for both experiments were 6410.25 Hz and 256 complex points in ^{1}H(F_1), 4300 Hz and 32 complex points in ^{13}C(F_2), and 6400 Hz and 256 real points in ^{1}H(F_3). Solvent suppression was achieved by on-resonance presaturation of the solvent signal. The ^{13}C carrier frequency was placed at 39.09 ppm. Sixteen transients were acquired per increment. ^{13}C chemical shifts were measured indirectly from the solvent ^{1}H frequency (Bax & Subramanian, 1986). The NOESY experiment had a mixing time of 150 ms. The experiments were processed using FELIX (version 2.10) (Biosym Technologies).

All multidimensional triple-resonance experiments were performed with the same spectrometer with an external Bruker fourth channel. All spectra were acquired at 305 K and were recorded with States-TPPI for the nonacquired dimensions (Marion *et al.*, 1989a). 3D triple-resonance experiments were performed as sets of 2D spectra. The carrier frequencies were set at 4.72 ppm (^{1}H), 121 ppm (^{15}N), 53 ppm (^{13}C aliphatic), and 178 ppm (^{13}C carbonyl). Spectra were processed by FELIX. In some cases, the number of acquired time points in the constant time dimension was doubled using mirror image linear prediction (Zhu & Bax, 1990). Low-frequency deconvolution of the acquired FIDs of the spectra recorded in H_2O was used to reduce the size of the residual H_2O signal (Marion *et al.*, 1989b).

The 3D HNCACB spectrum (Wittekind & Mueller, 1993) was acquired with 32 scans per increment and spectral widths of 8000, 9000, and 1500 Hz as 512 × 64 × 26 complex points in t_1 (^{1}H), t_2 (^{13}C), and t_3 (^{15}N), respectively. After the initial processing of t_1, only the amide region was further processed and transformed into a 3D matrix of 512 × 256 × 128 real points. The constant time HNCA (Farmer *et al.*, 1992) and HNCOCA (Bax & Ikura, 1991) spectra were acquired with spectral widths of 8000, 3500, and 1500 Hz, as 512 × 32 × 32 (26 for the HNCOCA) complex points with 64 scans per FID. Carbonyl decoupling in the HNCA experiment was achieved by low-power GARP decoupling (Shaka *et al.*, 1985). The constant time HNCO (Grzesiek & Bax, 1992) experiment was acquired with spectral widths of 8000, 1650, and 1500 Hz as 512 × 32 × 32 complex points with 32 scans per FID and processed to yield a matrix of 512 × 128 × 128 real points. The aforementioned triple-resonance experiments were performed on a ^{13}C/^{15}N doubly labeled barstar sample in H_2O. The constant time HCACO (Grzesiek & Bax, 1993) spectrum was acquired using a ^{13}C/^{15}N doubly labeled barstar sample in D_2O with spectral widths of 4000, 3150, and 3800 Hz as 512 × 32 × 32 complex points in t_1 (^{1}H), t_2 (^{13}C carbonyl), and t_3 ($^{13}C_\alpha$) and 16 scans per FID. Phase-shifted $(\sin x)/x$ pulses were used for the excitation of the carbonyl region. Processing of the data yielded a final matrix of 512 × 128 × 128 real points.

Details of Structural Restraints. Using previously assigned ^{1}H and ^{15}N resonances, 813 NOE cross-peaks could be

Table 1: ^{1}H, ^{13}C, and ^{15}N Resonance Assignments for Barstar in H_2O at pH 6.7, 305 K[a]

	N′	H^N	C^α	H^α	C′	C^β	H^β
A3	133.90	9.12	47.99	4.96	172.55	20.00	1.16
A25	126.00	7.46	50.70	4.17	175.00	14.78	1.45
A36	126.90	8.48	52.29	4.39	178.84	16.83	1.39
A67	124.40	8.41	53.36	4.07	176.41	17.11	1.17
A77	128.60	8.67	53.69	4.17	178.14	15.33	1.48
A79	127.10	8.15	52.85	4.25	177.39	15.71	1.68

	N′	H^N	C^α	H^α	C′	C^β	H^β
C40	120.90	8.36	61.46	4.47	175.41	25.26	$^{2}3.17/^{3}2.63$
C82	124.60	7.90	57.89	3.82	173.50	25.32	$^{2}2.09/^{3}1.66$

	N′	H^N	C^α	H^α	C′	C^β	H^β
D15	125.70	8.01	55.27	4.49	177.62	39.50	$^{2}3.07/^{3}.294$
D35	125.80	8.24	55.65	4.60	176.17	38.48	$^{3}2.83/^{2}2.70$
D39	124.70	8.39	53.31	4.24	177.79	38.50	3.07/2.88
D83	128.50	7.87	52.10	4.54	171.17	37.17	$^{2}2.58/^{3}2.86$

	N′	H^N	C^α	H^α	C′	C^β	H^β	H^γ
E8	124.60	9.73	55.64	4.25	174.39	26.25	2.13/1.98	2.48
E23	122.10	9.04	56.80	4.13	176.11	28.44	$^{2}1.75/^{3}2.17$	2.43
E28	123.20	8.57	57.14	3.84	174.65	27.63	2.00/1.98	2.27
E32			54.40	2.84	170.44	29.60	1.57/1.98	
E46	128.80	7.51	53.22	4.14	171.12	29.05	2.18/1.80	2.04
E52	132.80	9.08	52.29	5.17	170.06	29.24	$^{3}2.17/^{2}2.02$	
E57	122.20	9.21	57.71	3.67	177.46	26.73	2.04/2.26	
E64	128.00	8.35	56.39	4.10	174.30	25.13	2.26	
E68	119.90	7.78	57.65	4.17	176.47	27.06	$^{3}2.34/^{2}2.21$	2.12
E76	126.80	8.95	57.23	4.28	177.31	26.81	$^{2}2.40/^{3}2.21$	2.55
E80	121.10	7.75	54.47	4.49	176.71	27.06	$^{3}2.58/^{2}2.24$	2.59/2.77

	N′	H^N	C^α	H^α	C′	C^β	H^β	C^δ	H^δ	C^ϵ	H^ϵ	C^ζ	H^ζ
F56			60.00	3.65	176.38	37.30	2.73	129.51	5.74	128.19	6.28	126.87	6.61
F74	120.60	7.60	61.03	4.13	176.59	37.12	$^{2}3.57/^{3}2.95$	129.98	7.55	128.80	7.15	126.90	7.10

	N′	H^N	C^α	C′	H^α
G7	117.50	8.50	46.22		3.72/4.55
G31	119.10	8.05	43.78		3.34/3.81
G43	113.40	7.60	44.36	174.05	3.84/4.65
G66	116.20	8.77	45.69	174.16	3.91/4.13
G81	109.60	7.79	43.05	172.90	3.67/4.50

	N′	H^N	C^α	H^α	C′	C^β	H^β	$C^{\delta 2}$	$H^{\delta 2}$	$C^{\epsilon 1}$	$H^{\epsilon 1}$
H17	122.30	8.22	60.00	4.41	176.54	29.57	$^{2}3.33/^{3}3.08$	117.57	7.10	134.88	7.07

	N′	H^N	C^α	H^α	C′	C^β	H^β	$C^{\gamma 1}$	$H^{\gamma 1}$	C^δ	H^δ	$C^{\gamma 2}$	$H^{\gamma 2}$
I5	129.70	8.89	57.05	3.89	172.54	37.26	1.16	23.40	0.48/0.71	9.85	–0.16	14.40	–0.02
I10	124.20	7.09	57.05	4.19	172.26	34.56	2.28	23.93	1.24/1.63	8.90	0.69	17.17	0.95
I13	126.40	9.48	60.50	4.18	175.11	35.96	2.13	26.32	1.69	11.30	1.00	16.75	1.22
I84	130.40	7.50	57.70	5.13	172.51	40.53	1.56	26.58	1.46	12.90	1.15	15.37	0.95
I86	132.30	9.16		4.64	173.24		1.76	23.78	1.67	12.30	0.85	15.22	0.76
I87	135.60	9.42		4.30	172.08		1.77	24.30	1.41/1.02	11.85	0.79	15.06	0.87

	N′	H^N	C^α	H^α	C′	C^β	H^β	H^γ
K1	130.80	8.50		4.56	171.55		1.84/1.90	1.20
K2	130.20	8.63	52.60	5.31	172.43	32.31	1.81/1.60	1.27
K21	122.60	8.28	57.51	3.40	176.50		$^{2}1.97/^{3}2.13$	
K22	121.40	7.46	63.04	4.21	178.06	30.98	2.05/1.60	
K60	127.90	8.05	59.01	3.52	175.16	29.71	1.89/1.45	
K78	124.90	8.29	56.86	4.88	178.69	31.01	$^{2}1.95/^{3}2.25$	

	N′	H^N	C^α	H^α	C′	C^β	H^β	C^γ	H^γ	$C^{\delta 1}$	$H^{\delta 1}$	$C^{\delta 2}$	$H^{\delta 2}$
L16	126.00	7.78	56.49	4.07	175.97	39.13	$^{3}2.29/^{2}1.57$	24.70	1.42	21.02	0.78	24.27	1.01
L20	125.60	8.46	56.49	3.72	175.94	39.41	$^{2}2.12/^{3}1.12$	23.68	1.77	24.37	0.81	20.50	0.67
L24	117.80	8.27	52.35	4.35	172.75	38.49	1.75/1.63	24.96	1.63	24.17	0.89	22.50	0.83
L26	120.40	7.97	51.40	4.01	170.52	35.85	$^{2}1.22/^{3}1.03$	23.04	1.58	20.62	–0.15	20.83	0.58
L34	123.00	8.68	56.77	4.23	177.20	37.85	$^{3}1.64/^{2}2.00$	24.01	2.07	23.62	1.03	20.75	0.95
L37	124.10	8.37	56.86	4.25	175.88	38.48	$^{2}2.25/^{3}2.30$	24.91	1.64	18.80	0.88	25.08	0.90
L41	125.20	8.91	55.46	4.21	178.06	38.28	$^{2}1.66/^{3}1.23$	23.64	1.77	24.50	0.22	19.45	0.72
L49	127.10	8.80	51.82	5.41	173.41	45.47	$^{3}1.98/^{2}1.73$	25.00	1.51	21.83	0.86	24.54	1.02
L51	134.30	9.48	50.66	4.97	170.90	42.49	$^{3}1.77/^{2}0.78$	24.40	1.15	21.25	0.37	22.40	0.09
L62	123.60	8.19	54.71	4.35	172.45	40.16	$^{2}1.96/^{3}1.74$	24.70	1.93	23.30	1.00	20.90	0.97
L71	124.80	8.35	56.62	4.40	176.93	38.47	$^{3}2.03/^{2}1.45$		1.48	21.05	0.79	24.35	0.37
L88	133.10	8.59	51.54	5.14	172.37		1.67/1.35	23.92	1.41	22.72	0.40	21.35	0.32

Table 1 (Continued)

	N′	H^N	C^α	H^α	C′	C^β	H^β	$H^{\delta 2}$
N6	130.50	8.92	50.61	4.90	173.56	35.58	^{2}2.70/^{3}2.87	7.49/6.92
N33	114.00	6.50	49.71	4.45	176.54	37.54	^{2}3.07/^{3}2.83	7.01/7.05
N65	115.30	8.93	53.73	4.28	175.03	35.85	3.16	7.54/6

	C^α	H^α	C′	C^β	H^β	H^γ	C^δ	H^δ
P27	60.31	3.97	174.68	29.96	^{3}1.20/^{2}0.36	0.80/0.13	47.60	2.77/1.67
P48	60.78	5.21	172.38	33.75	^{2}2.07/^{3}2.65	1.95/2.08	49.45	3.82/3.57

	N′	H^N	C^α	H^α	C′	C^β	H^β	$H\gamma$
Q9	121.20	7.96	52.89	4.40	173.30	26.64	^{2}2.02/^{3}2.33	
Q18	123.00	8.68	57.33	4.10	177.08	25.84	2.31	2.63
Q55				4.25			2.32	2.49
Q58	123.30	7.92	57.42	4.17	177.91	25.88	2.16	
Q61	122.80	7.77	56.77	4.17	177.63	26.25	2.46/2.26	
Q72	121.40	8.42	57.01	4.04	175.62	25.37	^{3}2.32/^{2}2.27	2.56

	N′	H^ν	C^α	H^α	C′	C^β	H^β	H^γ
R11	127.60	9.17	53.78	4.56	173.29	28.59	^{3}2.11/^{2}1.82	1.68
R54	129.30	8.98	53.40	4.89	174.25	29.00	2.09	
R75	122.10	8.46	58.74	3.95	178.23	27.48	^{2}2.12/^{3}1.95	

	N′	H^N	C^α	H^α	C′	C^β	H^β
S12	113.80	7.75	54.90		172.84	64.14	4.53/4.32
S14	122.30	8.19		4.55	172.93	60.28	4.02
S59				4.44	174.87		4.15/3.97
S69	119.20	7.69	59.76	4.38	174.39	61.26	4.15
S89	134.90	9.13		4.54	173.67	63.30	4.10/3.82

	N′	H^N	C^α	H^α	C′	C^β	H^β	$C^{\gamma 2}$	$H^{\gamma 2}$
T19	122.70	8.28		4.12	171.39	66.36	4.31	19.75	1.23
T42	112.50	7.87	61.06	4.28	173.93	68.24	4.10	18.30	0.30
T63	108.40	7.38	59.19	4.61	173.67	68.34	4.61	19.41	1.11
T85	130.20	8.88	60.59	4.10	174.84	68.62	4.11	18.30	1.20

	N′	H^N	C^α	H^α	C′	C^β	H^β	$C^{\gamma 1}$	$H^{\gamma 1}$	$C^{\gamma 2}$	$H^{\gamma 2}$
V4	127.70	8.92	59.10	4.90	172.55	32.13	1.85	18.50	0.71	19.20	0.88
V45	123.20	7.99	61.53	3.71	171.80	31.11	1.92	20.07	0.62	21.63	0.97
V50	132.90	9.06	58.92	4.71	170.69	31.94	1.98	18.82	0.87	19.20	0.99
V70	125.60	7.53	64.89	3.77	175.54	29.61	2.48	20.25	1.12	21.10	0.97
V73	123.60	7.61	65.35	3.93	176.56	29.14	2.38	20.33	1.08	21.97	1.38

	N′	H^N	C^α	H^α	C′	C^β	H^β	$C^{\delta 1}$	$H^{\delta 1}$
W38	123.80	8.47	59.01	4.19	176.08	27.23	^{3}2.63/^{2}3.48	125.38	7.40
W44	125.70	8.49	58.63	4.65	173.04	29.99	^{2}3.59/^{3}3.45	122.92	6.84
W53			52.70	5.24	172.29	26.30	2.53/3.61	121.57	7.28

$C^{\zeta 2}$	$H^{\zeta 2}$	$C^{\eta 2}$	$H^{\eta 2}$	$C^{\epsilon 3}$	$H^{\epsilon 3}$	$C^{\zeta 3}$	$H^{\zeta 3}$
112.98	7.57	121.67	6.98	117.75	7.60	119.02	7.03
112.16	7.51	123.20	7.36	118.69	7.78	120.47	7.28
111.35	7.18	122.80	7.30	118.57	7.19	117.22	6.84

	N′	H^N	C^α	H^α	C′	C^β	H^β	C^δ	H^δ	C^ϵ	H^ϵ
Y29	114.40	6.44	53.20	4.57	172.46	34.26	^{2}3.82/^{3}2.83	131.29	7.03	116.62	6.81
Y30	126.80	7.42	56.02	4.29	175.62	36.80	^{2}3.09/^{3}2.57	131.29	7.01	116.22	6.82
Y47	125.50	7.90	54.94	4.48	176.57	37.48	^{3}3.29/^{2}2.65	130.93	7.12	115.57	6.69

[a] H^β protons that were stereospecifically assigned have a superscript 2 or 3 preceding the shift for $H^{\beta 2}$ and $H^{\beta 3}$, respectively.

identified from the 3D HMQC-NOESY and 2D homonuclear NOESY spectra (Lubienski *et al.*, 1993). This made a total of 1437 assigned NOE cross-peaks when the NOEs from the analysis of the ^{13}C-edited spectra were added. Each NOE restraint was separated into one of three distance ranges depending on whether it was a strong, medium, or weak NOE in the 150-ms NOESY spectra. These ranges were 1.8–2.7, 1.8–3.5, and 1.8–5.0 Å, respectively. Upper limits for NOEs from nonstereospecifically assigned methylene groups and rotationally averaged aromatic protons were corrected for by r^{-6} averaging. In addition, 0.5 Å was added to the upper limit for each methyl group involved in an NOE. All valine γ- and leucine δ-methyl groups were stereospecifically assigned using a 10% ^{13}C-labeled barstar sample in D_2O and the method described by Neri *et al.* (1989).

All ϕ dihedral-angle restraints, derived from accurately measured $^3J_{NH\alpha}$ coupling constants described by Lubienski *et al.* (1993), were set at $-120 \pm 40°$ for $^3J_{NH\alpha}$ coupling constants greater than 7.5 Hz, and $-50 \pm 30°$ for those less than 6.0 Hz. No ψ angles were restrained. All χ_1 angle restraints and β-methylene stereospecific assignments were derived from analysis of $^3J_{\alpha\beta}$ coupling constants in the homonuclear DQF-COSY spectrum, intraresidue HN–HB NOE strengths, and ^{15}N–HB couplings (Archer *et al.*, 1991). The allowed error

of χ_1 angle restraints was ±60°.

Hydrogen-bond restraints were included for backbone amide protons which exchanged for deuterons at a slow rate in equilibrium conditions at pH 6.7 (Lubienski *et al.*, 1993). Restraints were included only for amide protons involved in standard secondary structure hydrogen bonding where the carbonyl acceptor was known. In addition, two *i,i+5* hydrogen bonds (Leu-26 NH → Lys-21 CO; Val-45 NH → Cys-40 CO), common at tight turns at the ends of helices, were included only after initial rounds of structural calculation (using only NOE and dihedral angle constraints) had unambiguously defined the hydrogen-bond acceptors for these two protected amides. Two distance restraints were included for each hydrogen bond: HN → O (1.8–2.0 Å) and N → O (2.7–3.0 Å).

Method of Structure Calculation. Initial structures were calculated so that any misassigned NOEs could be removed and the possibility of the aforementioned *i,i+5* hydrogen bonds at the ends of the first and second helices could be established. All peptide bonds were restrained to be planar and *trans*, with the exception of that between Tyr-47 and Pro-48, which was restrained to be planar and *cis* on the basis of the $C^{\alpha}H(i)$–$C^{\alpha}H(i+1)$ NOE between them (Lubienski *et al.*, 1993). In total, 30 final structures were calculated using a distance geometry–simulated annealing hybrid protocol with the program XPLOR (Brünger, 1992). The first step involved using the experimental constraints and covalent geometry to generate a family of embedded substructures from a linear conformation of barstar using a distance geometry algorithm. The remaining atoms were added, and the system was taken through high-temperature molecular dynamics followed by dynamical simulated annealing (Nilges *et al.*, 1988). A quartic van der Waals' repulsion term and square-well quadratic potential terms, for interproton distance and torsion angle restraints, were used in the process. The van der Waals' hard-sphere radii were set at 0.8 times the CHARMm values (Brooks *et al.*, 1983). No attractive energy terms were used in the protocol. An additional 1000 cycles of restrained Powell energy minimization produced the final structures.

RESULTS

Assignment of Barstar. The majority of the ^{13}C assignments of barstar were obtained from the 3D ^{13}C–^{1}H HCCH-COSY experiment in conjunction with the previously reported proton assignments (Lubienski *et al.*, 1993). In addition, the experiment made possible the extension of the proton assignments, and previously ambiguous assignments were confirmed or corrected. Stereospecific assignments of valine γ-methyls and leucine δ-methyls were easily and unambiguously determined using the 10% ^{13}C fractionally labeled sample and the method of Neri *et al.* (1989).

The triple-resonance experiments were extremely useful for confirming the sequence-specific assignments and assigning the remaining C^{α} and C^{β} ^{13}C resonances as well as those of the backbone carbonyls. Although the proton shifts varied considerably according to chemical environment, the ^{13}C shifts were always within the range distinctive of atom type. This considerably helped the assignment. The full assignments of barstar are shown in Table 1 by residue type. The majority of the gaps in the assignments of barstar are associated with a general broadening of the resonances within a particular area of the molecule, the cause of which is currently being investigated.

The spread of methyl ^{13}C shifts was essential for assigning interresidue side chain–side chain and side chain–backbone

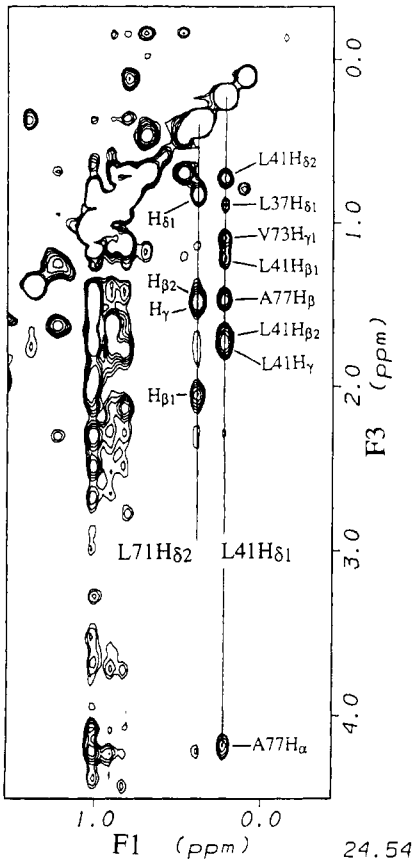

FIGURE 1: Example of a $^{1}H(F_1)$–$^{1}H(F_3)$ plane taken at $^{13}C(F_2)$ = 24.54 ppm of the 150-ms mixing time 3D ^{13}C-separated NOESY spectrum of uniformly ^{15}N–^{13}C labeled barstar. A number of NOEs from Leu-71 $H_{\delta 2}$ (intraresidue) and Leu-41 $H_{\delta 1}$ (intraresidue and long-range) are labeled.

NOEs with the 3D ^{13}C–^{1}H NOESY-HMQC spectrum, which was not possible from the previous homonuclear work because of extreme overlap in the proton resonances. This was especially problematic in the region of the homonuclear NOESY spectrum containing the NOEs from the δ-methyls of the 12 leucines of barstar. An example of the excellent resolution afforded by the 3D ^{13}C–^{1}H NOESY-HMQC spectrum is illustrated in Figure 1. The stereospecific assignment of the methyl groups of the 17 leucine and valine residues was important in order to make full use of the NOE information in the structure calculations, the pattern of NOEs being frequently different for both methyls of a valine or leucine residue.

Analysis of C^{α} and C^{β} Chemical Shifts and Their Significance to Secondary Structure. When the C^{α} and C^{β} shifts were compared to the ^{13}C residue-type specific random coil values derived by Howarth and Lilley (1978), various details were confirmed about barstar. First, prolines 27 and 48 have C^{β} chemical shift values (29.69 and 33.75 ppm) that correspond to the random coil *trans* and *cis* configurations, respectively (30.4 and 32.95 ppm), which agrees with the analysis of the NOE evidence. Second, the reduced state of the two cysteine side-chains (40 and 82) was confirmed by their C^{β} values (25.26 and 25.32 ppm, respectively) compared with the random coil values of 27.6 ppm for cysteines with a free SH group and 38.9–42.8 ppm for those involved in a disulfide bond. In addition, it has been shown that the difference between both the C^{α} and C^{β} shifts of a given residue from their random coil values can indicate the type of secondary structure in which the residue is involved in the folded protein (Spera & Bax, 1991). Subtracting random coil C^{α} shifts from

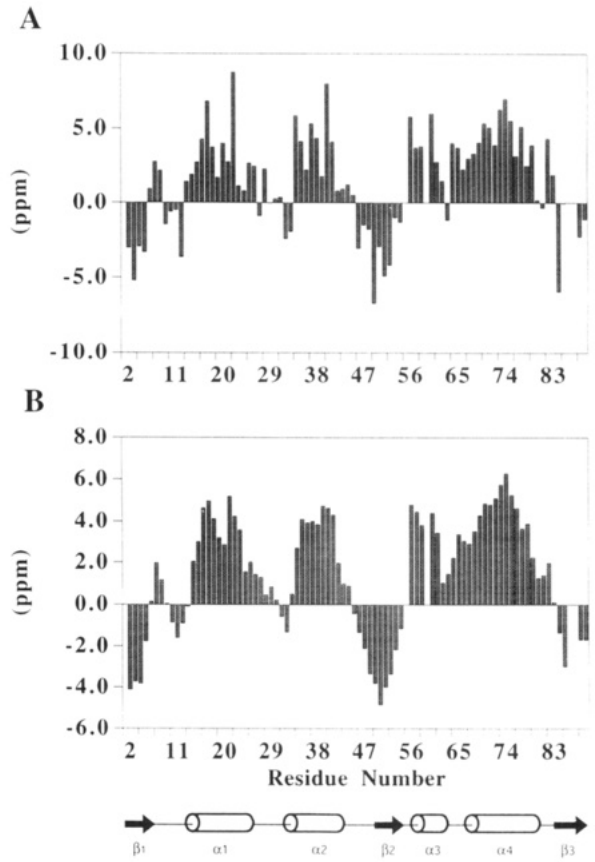

FIGURE 2: (A) Histogram showing the observed secondary shift (the measured ^{13}C chemical shift minus the random coil ^{13}C chemical shift) for each residue's C^{α} minus the secondary shift of its C^{β}. The secondary shifts of residues for which only the C^{α} could be assigned (all glycine residues and residues 21 and 88) and residues for which only the C^{β} could be assigned (14, 19, and 89) are included. Residues 1, 55, 59, 86, and 87 were not included because their C^{α} and C^{β} chemical shifts were not assigned. (B) Histogram enhancing the trends in panel A by applying a three-point smoothing function to the data. A two-point averaging was done for residues with only data from one neighboring residue. Helical and extended regions are observed as positive and negative peaks respectively in both panels A and B. A schematic showing the secondary structure of barstar is shown underneath panel B. Helices are represented by cylinders and extended strands by arrows.

the C^{α} shifts of a protein tends to give a positive secondary shift for C^{α}s in helical backbone conformation (3.09 ± 1.00 ppm) and a negative secondary shift for those in an extended strand conformation (–1.48 ± 1.23 ppm). This trend is opposite in sign for C^{β} secondary shifts (–0.38 ± 0.85 and 2.16 ± 1.91 ppm). By subtracting the C^{β} secondary shift of a residue from its C^{α} secondary shift, this correlation is enhanced. The result of this operation is shown for barstar in Figure 2A. A three-point smoothing function on the data is shown in Figure 2B. The data agree closely with the secondary structure defined by analysis of $^3J_{NH\alpha}$ couplings, hydrogen–deuterium exchange protection data, and sequential and medium-range NOE data (Lubienski *et al.*, 1993).

Structural Restraints and Statistics. A total of 30 structures were calculated using the restraints shown in Table 2. The analysis of the calculated structures is for all residues of barstar as the entire structure is very well defined by the experimental restraints. The average rms deviation of the final 30 structures from the mean structure is 0.42 ± 0.07 Å for the backbone atoms and 0.90 ± 0.07 Å for all heavy atoms. Table 3 gives a summary of the various energy terms and deviations from idealized geometry. None of the structures have NOE violations over 0.3 Å nor dihedral angle violations over 5°. A large, negative Lennard-Jones van der Waals' energy (not included in any of the target functions of the structure calculations) indicates that there are no bad nonbonded contacts.

Table 2: Summary of Structure Calculation Data

restraint type	no. of restraints
all	1613
interproton distances	
intraresidue	455
interresidue ($1 \leq \|i-j\| \leq 4$)	614
interresidue ($\|i-j\| > 4$)	368
hydrogen bond restraints (two per hydrogen bond)	74
ϕ dihedral angle restraints	54
χ_1 torsion angle restraints	48

The structure is very well defined as the statistics in Table 3 demonstrate. Figure 3 shows the atomic rms distributions for the 30 structures, as well as the variation in surface accessibility, as a function of residue number. Using the schematic representation of the secondary structure elements in Figure 3, it can be seen that the backbone of barstar is extremely well defined in most of the regions of secondary structure and most of the connecting loops. Exceptions to this are the N-terminal residue (which is in a β-strand) and the region comprising residues 55–66. This region begins with a sharp bend from the second β-strand to the third helix for which there were far fewer distance restraints than the other helices in the protein. The region ends with a three-residue loop (which exhibits the highest backbone rms deviations in the 30 structures) which connects the third helix to the fourth. The loop connecting the first and second helices (residues 26–32) which, along with the second helix, is responsible for binding barnase, has a well-defined backbone, especially in its middle at Tyr-30 (which will be discussed later).

Figure 3 also shows that the definition of the molecule also extends to the side chains: over 70% of the residues have side chains with rms deviations less than 1 Å. Of those side chains with deviations over 1 Å, all, with the exception of Phe-56, are polar or charged side chains. These include 8 of 11 Glu residues, and all 6 Gln and 3 Arg residues. In addition, the majority of these ill-defined residues have average solvent accessibilities over 40% for the 30 structures. The fractional solvent accessibility of a residue was computed from the total solvent accessibility for the residue in its current conformation and in the tripeptide fragment of which it is the central residue using a solvent probe radius of 1.5 Å (Richmond, 1984).

Ramachandran Plot Analysis. Figure 4A represents the backbone ϕ,ψ angles for the family of 30 structures, ⟨SA⟩. The vast majority of backbone dihedral angles fall within allowed regions of the plot and vary little within the 30 structures. In addition, many of the dihedral angles of residues involved in turns appear in the α_L-helix region of the plot. The residues associated with high backbone rms deviations in Figure 3B are also those for which there is more of a scatter in the Ramachandran plot. For the short loop between the second β-strand and the third helix (residues 55–57), there are two populations of ϕ,ψ values. The minor population exhibits generously allowed or disallowed ϕ,ψ angles, while the majority of the structures have torsion angles in this region that are allowed and that agree with those in the X-ray structure of barstar in the barstar–barnase complex. The second region of scatter is for the loop residues connecting the third and fourth helices (especially residue 64). These residues adopt a variety of allowed ϕ,ψ torsion angles. The plot is simplified by looking at the restrained minimized mean

Table 3: Structural Statistics for 30 Barstar Structures[a]

structural statistics	⟨SA⟩	$\overline{\langle SA \rangle}r$
rms deviations from experimental restraints[b]		
distance (Å)	0.0358 ± 0.0005[c]	0.0389
dihedral angle (deg)	0.54 ± 0.09	0.70
XPLOR energies (kcal mol^{-1})		
F_{vdW}[d]	49.76 ± 4.71	56.64
$F_{L\text{-}J}$[e]	−309.9 ± 23.1	−325.0
deviations from ideal covalent geometry		
bonds (Å)	0.0038 ± 0.0000	0.0041
angles (deg)	0.52 ± 0.01	0.58
impropers (deg)	0.35 ± 0.01	0.42
atomic rms differences (Å)	**backbone**	**all non-H**
⟨SA⟩ vs $\overline{\langle SA \rangle}$	0.42 ± 0.07	0.90 ± 0.07

[a] ⟨SA⟩ is the 30 final structures, $\overline{\langle SA \rangle}$ is the mean atomic structure obtained by averaging the superimposed coordinates for the 30 structures, and $\overline{\langle SA \rangle}r$ is the mean atomic structure after 1000 steps of restrained energy minimization. [b] In none of the structures were distance restraints violated over 0.3 Å nor dihedral angle constraints violated over 5°. [c] All variances are quoted ± one standard deviation. [d] The quadratic van der Waals term was calculated with a force constant of 4 kcal mol^{-1} within the van der Waals radii set to 0.8 times the standard value used in the CHARMm (Brooks *et al.*, 1983). [e] $F_{L\text{-}J}$ was calculated using the CHARMm empirical energy function. Note that this Lennard-Jones term is used as part of the target function in any part of the structure calculations.

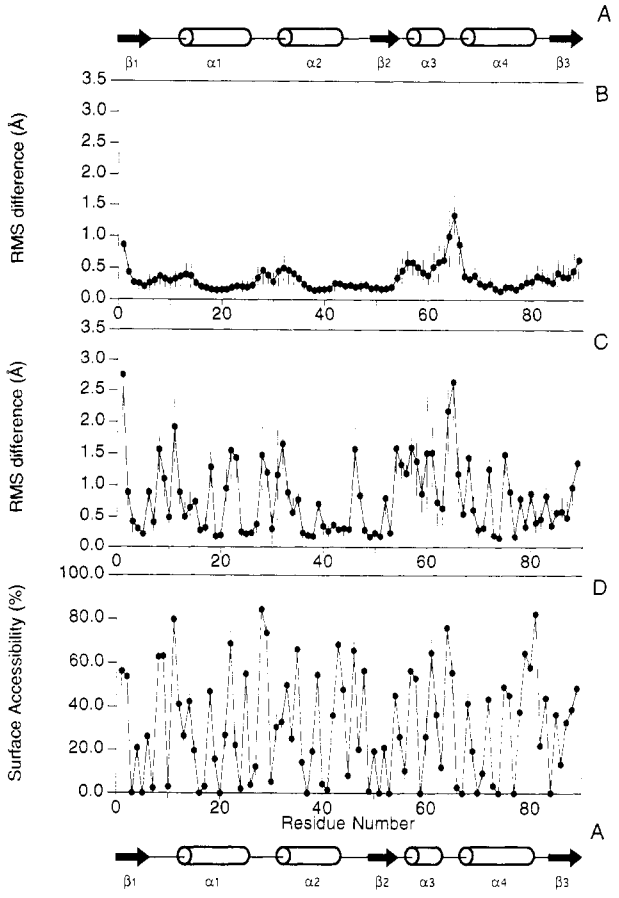

FIGURE 3: (A) Schematic showing the secondary structure of barstar. Helices are represented by cylinders and extended strands by arrows. Residue-based root mean square deviations of the atomic coordinates from the mean atomic structure are shown for (B) the backbone heavy atoms (C^{α}, C′, N, and O) and (C) all non-hydrogen atoms. Error bars represent ± one standard deviation. (D) Residue-based solvent accessibility. This was calculated as fractional accessibilities in percent. The total accessibility is computed for the residue in its current conformation and in the tripeptide fragment of which it is the central residue (Richmond, 1984).

structure of barstar, $\overline{\langle SA \rangle}r$, which adopts the most populated ϕ,ψ angles for the ill-defined residues (Figure 4B). These backbone dihedral angles agree closely with those in the complexed crystal structure of barstar. The residues that have positive ϕ angles in all 30 structures are Ala-25 (ϕ = 55°), Glu-32 (ϕ = 50°), Gln-55 (ϕ = 49.9°), Asn-65 (ϕ = 73.7°), and Gly-81 (ϕ = 113.3°). These residues are involved in tight turns either coming out of helices (Ala-25, Gly-81) or leading into helices (Glu-32, Gln-55, Asn-65). The preponderance of these residues is indicative of how positive ϕ angles are used to connect items of secondary structure in the protein with the shortest loop possible, rather than having longer loops with more energetically favorable ϕ,ψ angles.

DISCUSSION

Description of the Barstar Structure. Barstar is an α/β protein with overall dimensions of approximately 29 × 22 × 21 Å. It is composed of three parallel α-helices (residues 14–25, 33–43, 68–81) packed onto a three-stranded parallel β-sheet (residues 1–6, 49–54, 83–89). It has an additional helix (residues 56–63) linking the second, central, strand and the fourth helix. This third helix is not part of the bundle and is not so well defined by the structural restraints as the other three helices. Figure 5 shows the superposition of the backbone (N, C^{α}, and C) atoms for the 30 final structures of barstar, in two different stereoviews. Figure 5A shows the bundle composed of three of the helices and the β-sheet, with the disordered helix running along the top of the molecule and a disordered loop linking it to the fourth helix. This arrangement is also shown by a schematic drawing of barstar in Figure 6. The first and fourth helices interact with the β-sheet, whereas the second helix packs onto the first and fourth helices. Figure 5B shows the view of the molecule looking down the disordered third helix. From this view, and the view in Figure 6, it can be seen that the first and fourth helices splay away from the β-sheet toward their C-terminal ends. This leaves room for the second helix to pack into the space left between the two helices. The first and second helices are linked by an extremely well defined seven-residue overhand loop. This loop and the second helix protrude from the rest of the molecule and form the region of barstar which fits into the active site of barnase (as is described in the following paper in this issue).

Helices and C-Terminal Turns. The relationship between all four helices can be seen in Table 4. The table illustrates the angles between the helices and the points of closest C^{α} contact. Helix 3 runs along the N-terminal end of the parallel helix bundle that is formed by helices 1, 2, and 4. Equilibrium hydrogen–deuterium exchange data (Lubienski *et al.*, 1993) shows that helices 1, 2, and 4 have protected amide protons

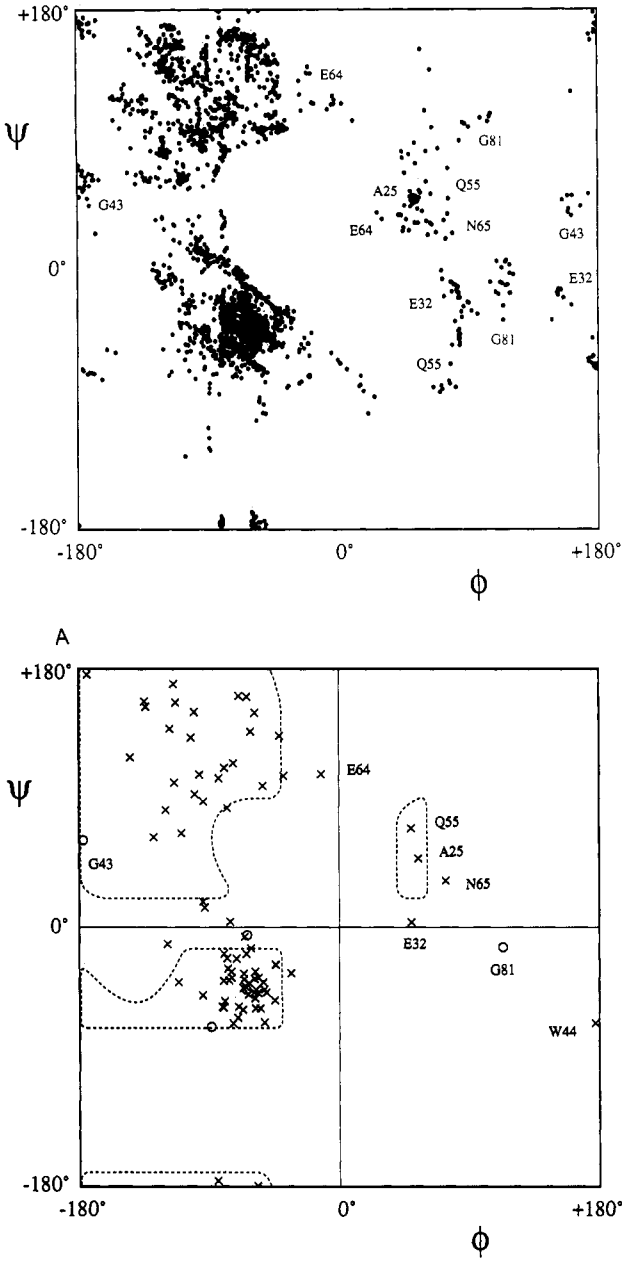

FIGURE 4: (A) Ramachadran plot for the 30 NMR solution structures of barstar. Labels indicate the position of a group of dihedral angles for the given residue. (B) Ramachadran plot for the restrained minimized mean structure of barstar. Glycine residues are represented by circles. Allowed regions are marked.

from residues 18–26, 37–42, and 72–80, respectively, which indicates that these amide protons are involved in hydrogen bonds. No amide protons in the third helix were protected under the experimental conditions reported. However, the conformation of the region is defined by various sequential and medium-range NOEs.

The hydrogen-exchange data would predict the N-cap residues of helices 1, 2, and 4 to be Ser-14, Asn-33, and Glu-68, respectively. Ser, Asn, and Glu have been shown to be good N-cap residues due to their ability to hydrogen-bond to free amide groups (especially i+3) in the first turn of helix, which are exposed and hence stabilize the helix (Presta & Rose, 1988; Serrano *et al.*, 1992). Although this is reflected in the structure for helix 2 (with the side chain of Asn-33 hydrogen-bonded with the amide group of Ala-36), this N-capping is not seen in the structure for helices 1 and 4. Ser-12, however, can hydrogen-bond with the free NH of Asp-15. In addition, the ϕ,ψ angles, and the possibility of backbone carbonyl (i) to amide (i+4) hydrogen bonds, indicate that helix 1 may start at Ile-13 and helix 4 at Ala-67. The N-cap residue for helix 3 is Phe-56. This is supported by its having helical ϕ,ψ angles, its carbonyl being in a reasonable orientation for helical hydrogen bonding, and, indirectly, its positive C^{α}–C^{β} secondary shift (Figure 2).

There are tight turns at the end of helices 1, 2, and 4, each characterized by a stabilizing i,i+5 hydrogen bond and a residue with a positive ϕ angle. The function of the tight turn is both to minimize the length of the chain connecting items of secondary structure and to increase the exposure to solvent of the non-hydrogen-bonded backbone carbonyl groups in the last turn of helix, which stabilizes the helix. Glycine is particularly well suited to being the C-cap residue of a helix because of the large degree of conformational freedom it has to initiate the tight turn (Serrano *et al.*, 1992). In barstar, the C-cap residues for helices 1, 2, 3, and 4 are Ala-25, Gly-43, Thr-63, and Gly-81, respectively.

The amide proton of Ala-25 is hydrogen-bonded to the backbone carbonyls of Lys-21 and Lys-22. In addition, Ala-25 exhibits positive ϕ and ψ angles, and the amide proton of Leu-26 hydrogen bonds with the backbone carbonyl of Lys-21 (and is well protected according to the hydrogen exchange data). This arrangement is a recognized motif at the end of helices termed the "paperclip" (Milner-White, 1988). A similar arrangement is seen at the C-cap of helix 4. Gly-81 has a positive ϕ angle, its amide is hydrogen-bonded to the backbone CO of Lys-78, and the backbone NH of Cys-82 is within hydrogen-bonding distance of the CO of Ala-77. Neither the NH of Gly-81 nor Cys-82 is protected from hydrogen exchange, and so the turn may not be as fixed as that at the end of helix 1.

Helix 2 differs at its C-cap from helices 1 and 4. The NH of Thr-42 donates a hydrogen bond to the CO of Trp-38 in a normal α-helical fashion; however, Gly-43 and Trp-44 have dihedral angles such that they make a wider girth of the helix, and the NH of Val-45 hydrogen bonds with the CO of Cys-40. According to the derived structure, the amide protons of 43 and 44 may still form long hydrogen bonds with the CO groups of Asp-39 and Cys-40, respectively. In terms of hydrogen exchange behavior for residues 42–45, only Thr-42 and Val-45 are protected. The wide helical turn emanating from helix 2 has implications for barstar's interaction with barnase, as is discussed in the following paper (Buckle *et al.*, 1994).

After the C-cap residue of helix 3 (Thr-63) a short, ill-defined loop (comprising residues Glu-64, Asn-65, and Gly-66) links the third helix to helix 4. This is the worst defined region of the structure and is, interestingly, highly disordered in the barstar–barnase complex (Guillet *et al.*, 1993; Buckle *et al.*, 1994). In each of the 30 structures, however, Glu-64 has a positive ψ angle, and Asn-65 a positive ϕ angle. This induces two, tight, right-angled bends in the polypeptide chain. A ^{15}N NMR relaxation study is currently underway to investigate whether the loop is mobile rather than undefined.

β-Sheet. The sheet has the characteristic right-handed twist and is composed of three parallel strands comprising residues from Lys-1–Asn-6, Leu-49–Arg-54, and Asp-83–Ser-89. The strands are arranged in order of sequence, with the second strand central to the other two. Hydrogen exchange studies show the NH groups of strands 1 and 3 to be protected alternately, and those of strand 2 to be protected on both sides. In terms of rms deviation, the beginning of strand 1

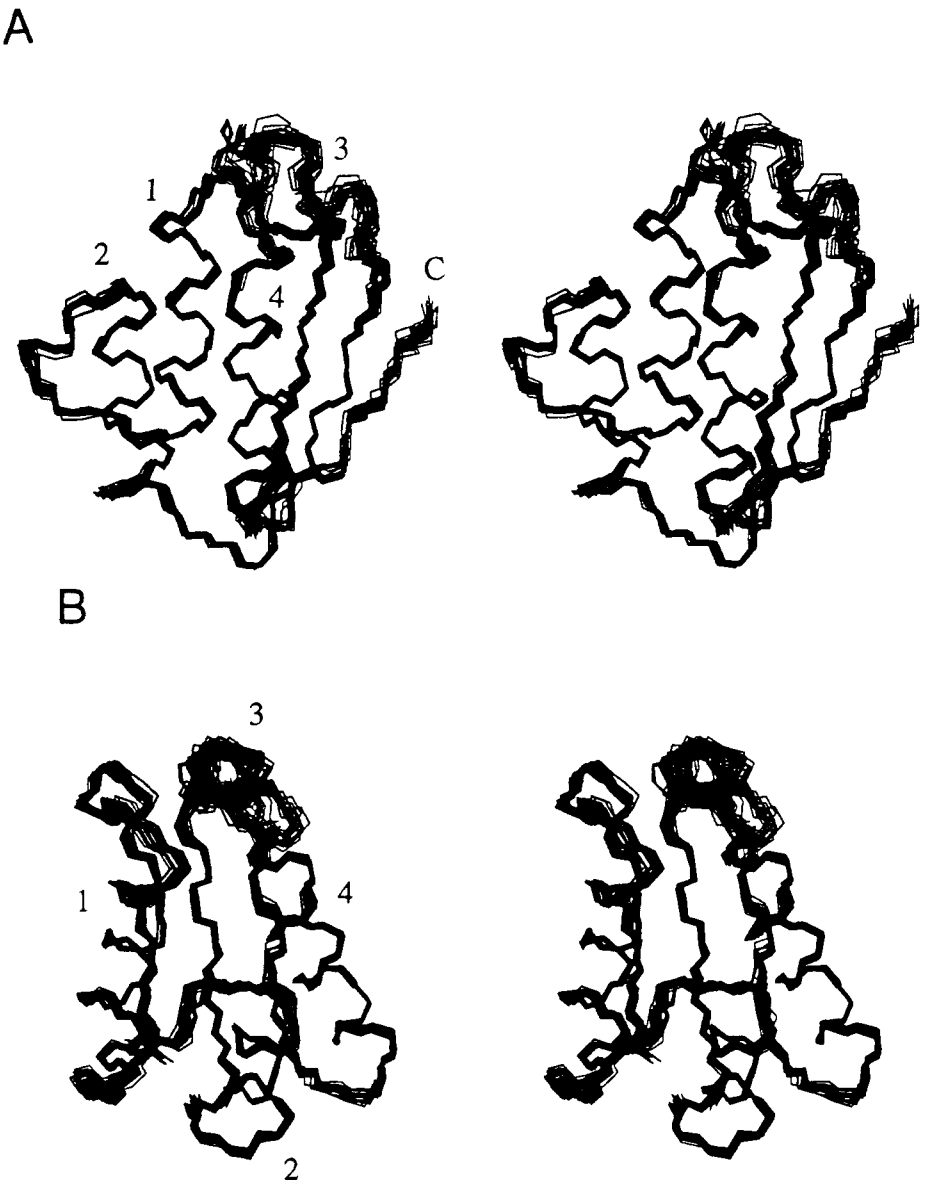

FIGURE 5: Two stereoviews showing the superposition of the C^{α}, C′, and N backbone atoms for the 30 solution structures of barstar. The four helices are numbered 1–4 in order of sequence. In panel A the C-terminus of the protein is also labeled.

and end of strand 3 are the worst defined areas of the sheet.

Strand 1 begins with the terminal NH_2 group hydrogen-bonding with the backbone CO of Pro-48. The strand ends with the CO of Ile-5 hydrogen-bonding with the backbone NH of Arg-54. Although Asn-6 maintains the extended strand conformation, it leads into a helical turn from residues 6–10. This turn is supported with a hydrogen bond from the NH of Ile-10 to the backbone CO groups of Gly-7 and/or Asn-6. The hydrogen bond must be quite stable as the NH of Ile-10 is protected from hydrogen–deuterium exchange. Interestingly, the turn is also indirectly reflected in its positive C^{α}–C^{β} secondary shifts (Figure 2). After Ile-10, the loop becomes more extended in conformation (and exhibits negative secondary shifts) before leading into helix 1.

The loop from helix 2 leads into strand 2 via a *cis*-proline (Pro-48), which creates a right-angled turn in the backbone. The amide protons of the strand are hydrogen-bonded from Leu-49 to Arg-54. Gln-55 has a positive ϕ angle in each of the 30 structures which induces a sharp right-angled turn which immediately leads into helix 3 at residue 56. Finally, strand 3 extends from Asp-83 (the CO of which accepts a hydrogen bond from the NH of Leu-49) to Ser-89 (the NH of which donates a hydrogen bond to the CO of Trp-53).

Side Chains. Barstar contains 24 charged amino acids: 11 Glu, 4 Asp, 6 Lys, and 3 Arg. This distribution gives the protein a net negative charge. These residues are mainly distributed evenly throughout the surface of the protein. The exception to this is the area comprising the second helix and the C-terminal end of the fourth. This region of barstar is highly negatively charged compared with the rest of the protein (containing residues Asp-35, Asp-39, Glu-76, Glu-80, and no positively charged side chains). This has functional implications, and is discussed in the following paper (Buckle *et al.*, 1994). Among several possible salt bridges between the charged side chains, two are certainly defined by the NMR structures: between Glu-57 and Lys-60, and between Glu-52 and Lys-2.

The hydrophobic core in barstar is formed between the side chains of helices 1, 2, and 4 and those of the β-sheet. Figure 7A shows a stereoview of various residues within the core. The buried residues are extremely well defined. The core is composed of 10 of barstar's 12 leucine residues, three of its five valines, four of its six isoleucines, Trp-53, Phe-56, Phe-74, Ala-3, Ala-67, and Ala-77. The large number of leucines forming the core is a common occurrence in helix-bundle proteins. The majority of the leucines in the core are from

FIGURE 6: Schematic picture of the restrained minimized mean structure of barstar drawn with MOLSCRIPT (Kraulis, 1991). The N-terminus is at the bottom of the left-most β-strand, and the C-terminus is at the top of the right-most β-strand.

Table 4: Angles between Helices and Points of Closest C_α Contacts[a]

pairs of helices[b]	closest C_α[c]	distance (Å)[c]	angle (deg)[c]
1 and 2	17/37	7.8	10
1 and 3	13/63	5.1	−125
1 and 4	16/67	6.9	−34
2 and 3	33/63	12.9	131
2 and 4	38/73	5.8	−26
3 and 4	60/67	4.6	122

[a] The analysis is done for the restrained minimized mean structure of barstar using the CCP4 program HELIXANGLE (CCP4, 1979). [b] Helices 1–4 are as described in the text. [c] Pairs of residues are given, one from each respective helix, representing the closest α-carbon distance between the helices. The distance is given in the next column, and the angle between the respective helices is given in the last column.

helix 1, helix 2, and strand 2. The isoleucines are mainly from the peripheral strands. Helix 4 provides two valines and two alanines as well as the central core residue, Phe-74. Figure 3D shows that all the core residues have less than 25% solvent accessibility, the majority having close to 0%. Interestingly, although the benzene ring of Phe-74 is at the heart of the core and has the lowest rms deviations of all barstar atoms, both the δ and ϵ protons have degenerate chemical shifts. It is unlikely that the chemical environment of the ring would be the same on both sides, and so the ring must flip inside the core reasonably fast on the NMR time scale. This would indicate either that the core in barstar is not rigid, but more fluid in character, or that the core is not well packed in the vicinity of Phe-74.

All the helices in barstar are amphipathic with the exception of helix 3. This helix is mainly composed of polar side chains and contributes only Phe-56 to the hydrophobic core (which is 15% exposed to solvent). Nevertheless, the helix is extensively packed: theoretical calculations show that barstar loses 21% of accessible surface area when the third helix associates with the rest of the protein (Guillet *et al.*, 1993). Ser-59 is completely buried, and the hydrogen bond accepted by its hydroxyl oxygen from the indole amide proton of Trp-53 accounts for the protection of this amide from hydrogen exchange. The role of this helix seems to fulfil a requirement of all bundle-type proteins (whether all α-helix or β-sheet in composition): to shield the exposed top and bottom of the bundle from solvent entry. In barstar, this is achieved at the N-terminal end of the helices by the third helix and the loop comprising residues 7–13, and at the C-terminal end by the close association of the tight turning loops from residues 44–48 and 81–83. Especially important at the C-terminal end are the roles of Val-45, which forms part of the hydrophobic core, and Tyr-47, which forms a central "plug" at the bottom of the bundle which efficiently prevents solvent from entering the core. Tyrosine is an efficient amino acid to perform this function because of its ability to interact with both the hydrophobic core and the solvent.

The side chains in the β-sheet are not alternately hydrophobic and hydrophilic as might be expected. Strand 1 positions Val-4 on the outside of the protein, strand 2 does the same with Val-50, and strand 3 exposes Ile-87. It is interesting, however, that Val-4 and Val-50 both have solvent accessibilities less than 20%. This is mainly due to their being packed on the aliphatic part of charged side chains close in proximity. Val-4, for instance, packs onto Glu-52 and Lys-2, which also form a salt bridge. Contacts are also made between the aliphatic protons of both Glu-52 and Lys-2 and Val-50, and in addition between Thr-85 and Val-50. These particularly neat arrangements that are defined by various NOE distance constraints mean that the side chains involved have low rms deviations in each of the 30 structures from the average structure: Val-4 and Val-50 less than 0.4 Å; Lys-2 and Glu-52 less than 1 Å. The function of these hydrophobic clusters on the outside of the protein may be to constrain the strands into forming only one possible arrangement at the end of the folding process.

The two cysteines of barstar are positioned at the C-terminal ends of helices 2 (Cys-40) and 4 (Cys-82). It is has been shown that these residues can form a disulfide bridge which stabilizes the protein and slightly improves the barnase–barstar interaction (Hartley, 1993). The NMR structure of the reduced form of barstar agrees with the observation that the ends of the side chains are too far apart to form a disulfide (Guillet *et al.*, 1993). However, rotations of the χ_1 angles of the two residues, a rotation of helix 2, and a small movement in the loop region around Cys-82 would bring the side chains within bonding distance. Both Cys-40 and Cys-82 are relatively buried (with solvent accessibilities of 5% and 25%, respectively), which explains why the protein must be unfolded before the Ellman assay can detect the free thiols (M. Lubienski, unpublished results). Currently, work is underway to oxidize the protein fully and solve its structure by NMR.

Stabilization of the Binding Loop. The overhand loop between helices 1 and 2 is extremely well defined in the NMR structure, and its conformation is not significantly altered when bound to barnase (discussed further in the following paper). To fix the loop in such a firm conformation, the side chains of the loop form various hydrophobic contacts which "fasten" it at either end and in its middle. Figure 7B shows a stereoview of the side chains of the 30 superimposed structures which form interactions that stabilize the loop. As

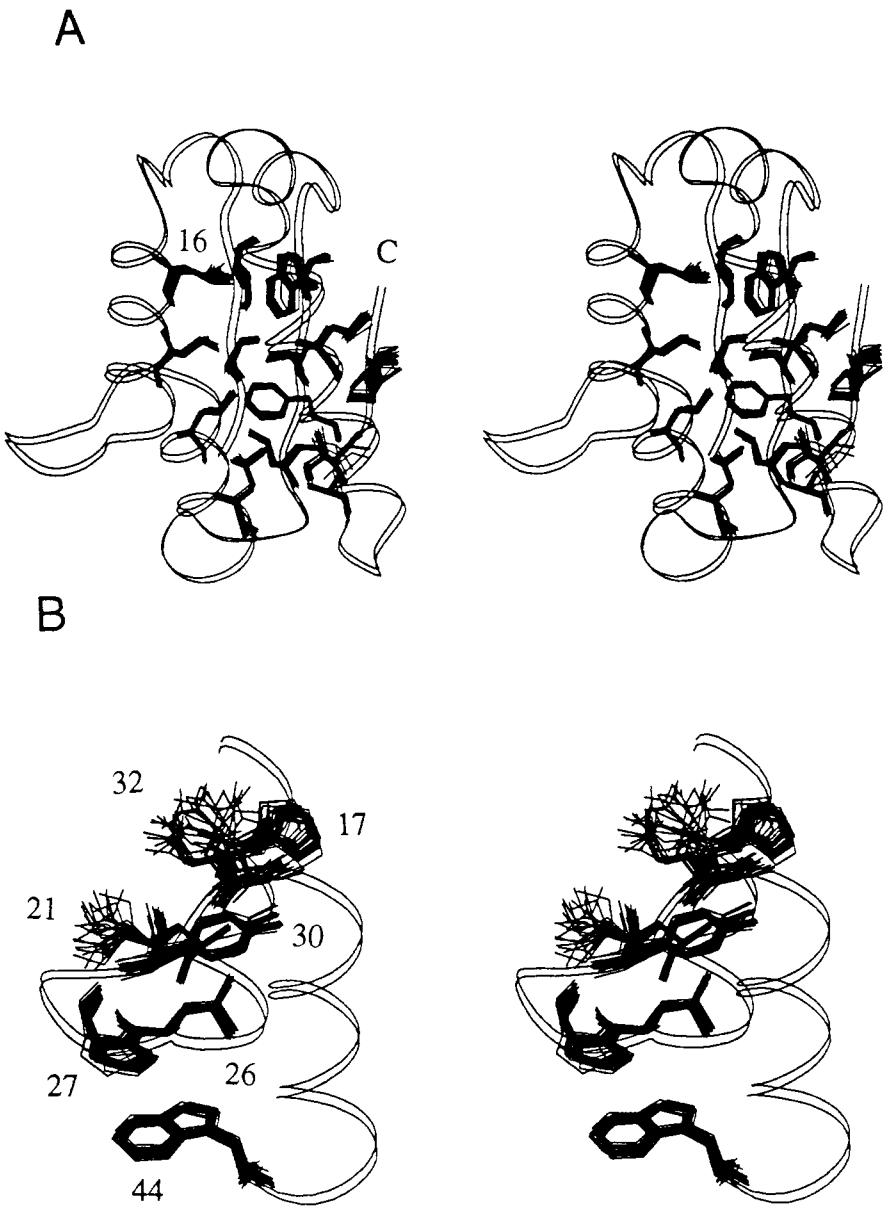

FIGURE 7: Two stereoviews showing the superposition of selected side chains for the 30 solution structures of barstar inserted on a two strand backbone ribbon. (A) Selected hydrophobic core residues: Ala-3, Ile-5, Leu-16, Leu-20, Leu-37, Leu-41, Leu-49, Leu-51, Trp-53, Leu-71, Phe-74, Ala-77, Ile-84, and Ile-86. Leu-16 and the C-terminus of the ribbon are labeled. (B) Interactions stabilizing the inhibitory loop. A ribbon is shown from the N-terminus of helix 1 to the C-terminus of helix 2. The residues shown are His-17, Lys-21, Leu-26, Pro-27, Tyr-30, Glu-32, and Trp-44.

discussed before, the backbone NH of Leu-26 hydrogen-bonds to the CO of Lys-21 to form a tight turn. The side chain of Leu-26 also packs against the aliphatic atoms of Lys-21 as well as forming part of the hydrophobic core. The ring of Pro-27 forms a perfect face to face interaction with the six-membered ring of Trp-44. This interaction is very well defined and is the main reason for the low surface accessibility of the proline residue (10%). The chemical shifts of the ring protons of Pro-27 are immensely shifted upfield because of this interaction (the γ protons resonating at 0.13 and 0.80 ppm for instance). Tyrosines 29 and 30 are interesting in that the former is important for barstar's function of binding barnase, points out into solvent, and has a high rms deviation, whereas the latter is important in maintaining the structure of barstar, points into the core of the protein, stabilizing the middle and C-terminal end of the loop with various interactions, and has an extremely low rms deviation. Tyr-30 packs against the side chains of Leu-26, Leu-20, Ala-36, the aliphatic atoms of Glu-32, and Lys-21, and its hydroxyl accepts a hydrogen bond from the $N^{\delta 1}H$ of the imidazole ring of His-17. These interactions are all well defined by the experimental constraints. The hydrophobic clusters defined by these interactions are not separated from the main hydrophobic core (with the exception of the 27/44 interaction which is solvent exposed) and so may be considered part of it. The fixed nature of this loop is quite reminiscent of the fixed conformation of the inhibitory loop of protease inhibitors.

Relationship of Structure to Function. The molecular architecture of barstar is ideally suited to performing its function of blocking the active site of barnase. The loop and helix comprising residues 26–44 protrude from the rest of the molecule and have a very stable conformation which is perfect for docking onto barnase. The interaction with barnase is described in detail in the following paper (Buckle *et al.*, 1994) as well as the functional implications of the charge distribution of barstar and the wide girth of the last turn of helix 2, which is stabilized by the Trp-44–Pro-27 interaction. The paper also fully discusses the role of barstar's interacting side chains

in forming the tight complex and compares the free NMR barstar structure with the crystal structure of barstar complexed with barnase.

Characterization of the Barstar Fold. The structure adopted by barstar is a novel fold. It is a very economical structure in terms of the preference for tight turns induced by positive ϕ and ψ backbone dihedral angles and a *cis*-proline rather than including longer loops. The fold can be looked on in two different ways: as a four-helix bundle where one helix is replaced with a three-stranded parallel β-sheet or, more obviously, as a classical α/β protein (such as the C-terminal half of the mononucleotide binding domain) with a loop and helix inserted between the first helix and second strand. Another interesting feature is that in barstar the dipoles of helices 1, 2, and 4 are arranged together. No functional significance has been attributed to this.

Concluding Remarks. The availability of $^{13}C/^{15}N$ doubly labeled samples was essential for the structural determination of barstar because the large number of inter-side-chain NOEs that were resolved that were important for defining the relative positions of the helices. The heteronuclear techniques described are currently being used to characterize the nature of the acid-denatured state of barstar, which has been shown to possess properties resembling a molten globule state (M. Lubienski and R. Golbik, unpublished results; Khurana & Udgaonkar, 1994). In addition, ^{15}N-relaxation measurement techniques are being used to define any regions of mobility in the barstar backbone. As the barstar molecule is reasonably free of undefined regions, the protein will be ideal for structural analysis by joint refinement (Shaanan *et al.*, 1992) with any future crystal structure of free barstar. The full assignment and structure of barstar, reported in this paper, will be used in this laboratory for protein engineering and structural studies of the free and complexed state of the molecule by NMR.

ACKNOWLEDGMENT

We thank Laurence Tisi of the University of Cambridge Biochemistry department for useful discussions.

REFERENCES

Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636–641.
Bax, A., & Subramanian, S. (1986) *J. Magn. Reson. 67*, 565–569.
Bax, A., & Ikura, M. (1991) *J. Biomol. NMR 1*, 99–104.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.
Brünger, A. T. (1992) XPLOR Version 3 Manual, Yale University, New Haven, CT.
Buckle, A. M., & Fersht, A. R. (1994) *Biochemistry 33*, 1644–1653.
Bycroft, M., Ludvigsen, S., Fersht, A. R., & Poulsen, F. M. (1991) *Biochemistry 30*, 8697–8701.
Buckle, A. M., Henrick, K., & Fersht, A. R. (1993) *J. Mol. Biol. 234*, 847–860.
Buckle, A. M., Schreiber, G., & Fersht, A. R. (1994) *Biochemistry* (following paper in this issue).
CCP4 (1979) The SERC (U.K.) Collaborative Computing Project Number 4, a Suite of Programs for Protein Crystallography, distributed by Daresbury Laboratories, Warrington WA4 4AD, U.K.
Chen, Y. W., Fersht, A. R., & Henrick, K. (1993) *J. Mol. Biol. 234*, 1158–1170.
Farmer, B. T., Venters, R. A., Spicer, L. D., Wittekind, M. G., & Mueller, L. (1992) *J. Biomol. NMR 2*, 195–202.
Grzesiek, S., & Bax, A. (1992) *J. Magn. Reson. 96*, 432–440.
Grzesiek, S., & Bax, A. (1993) *J. Magn. Reson. Ser. B 102*, 103–106.
Guillet, V., Lapthorn, A., Hartley, R. W., & Mauguen, Y. (1993) *Structure 1*, 165–177.
Hartley, R. W. (1989) *Trends Biochem. Sci. 14*, 450–454.
Hartley, R. W. (1993) *Biochemistry 32*, 5978–5984.
Hollecker, M., & Creighton, T. E. (1989) in *Protein Structure: A Practical Approach* (Creighton, T. E., Ed.) pp 146–148, 150–153, and 156–157, IRL Press, Oxford.
Howarth, O. W., & Lilley, D. M. (1978) *J. Prog. NMR Spectrosc. 12*, 1–40.
Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204–209.
Janin, J., & Chothia, C. (1990) *J. Biol. Chem. 265*, 16027–16030.
Jones, D. N. M., Bycroft, M., Lubienski, M. J., & Fersht, A. R. (1993) *FEBS Lett. 331*, 165–172.
Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888–889.
Khurana, R., & Udgaonkar, J. B. (1994) *Biochemistry 33*, 106–115.
Kobe, B., & Deisenhofer, J. (1993) *Nature 366*, 751–756.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Lubienski, M. J., Bycroft, M., Jones, D. N. M., & Fersht, A. R. (1993) *FEBS Lett. 332*, 81–87.
Mariani, C., Gossele, V., Debeuckeleer, M., Deblock, M., & Goldberg, R. B. (1992) *Nature 357*, 384–387.
Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989a) *Biochemistry 28*, 6150–6156.
Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515–1517.
Mauguen, Y., Hartley, R. W., Dodson, E. J., Dodson, G. G., Bricogne, G., Chothia, C., & Jack, A. (1982) *Nature 297*, 162–164.
Milner-White, E. J. (1988) *J. Mol. Biol. 199*, 503–511.
Neri, D., Szyperski, T., Otting, G., Senn, H., & Wüthrich, K. (1989) *Biochemistry 28*, 7510–7516.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317–324.
Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641.
Richmond, T. J. (1984) *J. Mol. Biol. 178*, 63–89.
Schreiber, G., & Fersht, A. R. (1993) *Biochemistry 32*, 5145–5150.
Serrano, L., Sancho, J., Hirshberg, M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544–559.
Shaanan, B., Gronenborn, A. M., Cohen, G. H., Gilliland, G. L., Veerapandian, B., Davies, D. R., & Clore, G. M. (1992) *Science 257*, 961–964.
Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547–553.
Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490–5492.
Wittekind, M., & Mueller, L. (1993) *J. Magn. Reson. Ser. B 101*, 201–225.
Zhu, G., & Bax, A. (1990) *J. Magn. Reson. 90*, 405–410.