

Assignment of the backbone ^1H and ^{15}N NMR resonances and secondary structure characterization of barstar

Michael J. Lubienski, Mark Bycroft*, David N.M. Jones, 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, UK

Received 22 June 1993; revised version received 20 August 1993

Barstar, a polypeptide inhibitor of ribonucleases, has been studied by 2D and 3D NMR techniques using uniformly ^{15}N -labeled protein. Backbone ($^{15}\text{NH}-\text{C}_\alpha\text{H}-\text{C}_\beta\text{H}$) resonances were assigned for all but 5 of the 89 residues. Dihedral angle and deuterium exchange studies were used in conjunction with medium range inter-proton NOEs to characterize the secondary structure of barstar. The protein is composed of four α -helices and three short stretches of extended strand. By further analysis of the NOE data three of the helices were found to be parallel to each other with the single disulphide bond linking the second and fourth helices at their C-terminal ends.

Sequence-specific NMR assignment; Barnase; Barstar

1. INTRODUCTION

Barstar is an intracellular, 89 residue polypeptide inhibitor of barnase, an extracellular endoribonuclease from *Bacillus amyloquefaciens* [1]. The one-to-one interaction between the proteins has evolved to be extremely tight (K_d of 2×10^{-14}) [2]. This interaction is presumably essential to avoid degradation of intracellular RNA by any internally-expressed barnase.

The interaction is ideal for a model system to study molecular recognition by NMR, the complex being only 199 amino acids. Crystal [3] and solution [4] structures for barnase are available. To date there is no structural information on barstar. Far-UV circular-dichroism measurements on barstar indicate that the protein is predominately helical (unpublished data). Consistent with this, we have found that 2D homonuclear NMR spectra are poorly dispersed in the C_αH region. Assignment of mainly helical proteins is difficult by conventional 2D homonuclear techniques due to spectral overlap and chemical-shift degeneracies. To simplify the task, the protein can be uniformly ^{15}N labeled so that a variety of three-dimensional NMR techniques can be used.

In this paper we describe the assignment of the back-

bone ^1H and ^{15}N resonances, and many of the side-chain ^1H resonances by two- and three-dimensional homonuclear and heteronuclear NMR. In addition, using NOE, $^3J_{\text{HN}\alpha}$ coupling constant, and hydrogen-exchange data, we also outline the regions of secondary structure within the protein, and their possible topology.

2. MATERIALS AND METHODS

2.1. Barstar preparation

Barstar was purified from the *E. coli* strain BL21(DE3)[pLysE] harbouring the barstar expression plasmid pML2bs. The method of preparation will be described elsewhere. For unlabeled barstar, the bacteria were grown in $2 \times \text{TY}$ medium. For uniformly ^{15}N -labeled protein bacteria were grown in Celstone-N medium (Martek). A 3–4 mM sample of pure barstar could be produced from half a litre of culture. Unlabeled or labeled samples were produced by extensively dialysing the protein against 10 mM potassium phosphate pH 6.6 and concentrating it in an Amicon 8MC concentrator with a YM2 membrane to 2.5–3.5 mM. H_2O samples contained 10% D_2O . For D_2O samples the solvent was extensively exchanged for 10 mM deuterated potassium phosphate pH 6.6 (uncorrected) in the concentrator. All samples contained 0.05% sodium azide.

2.2. NMR measurements

All NMR experiments were carried out on a Bruker AMX 500 spectrometer equipped with a triple-resonance $^1\text{H}/^{15}\text{N}/^{13}\text{C}$ probe optimised for proton detection. All experiments were performed at 305 and 283 K. Homonuclear COSY [5], TOCSY [6,7] and NOESY [8,9] experiments were recorded with 2048 complex data points in t_2 and 512 increments in t_1 with a spectral width of 8000 Hz in both dimensions. 64 transients for each increment were collected. Phase-sensitive two-dimensional spectra were obtained by the time-proportional phase incrementation method (TPPI) [10,11]. Solvent suppression was achieved by on-resonance presaturation of the solvent signal. NOESY and TOCSY mixing-times were 150 and 50 ms, respectively. Spectra were zero-filled to 2048×1024 real data points along f_2 and f_1 respectively. Mild ($\pi/3$) sine-bell window functions were used in both dimensions.

*Corresponding author.

Abbreviations: NMR, nuclear magnetic resonance; 2D, two-dimensional; 3D, three-dimensional; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional NOE-correlated spectroscopy; TOCSY, total correlated spectroscopy; HMQC, ^1H -detected heteronuclear multiple quantum correlation; HSQC, 2D ^1H -detected heteronuclear single quantum correlation.

3. RESULTS AND DISCUSSION

The 3D NOESY-HMOC was used to connect the

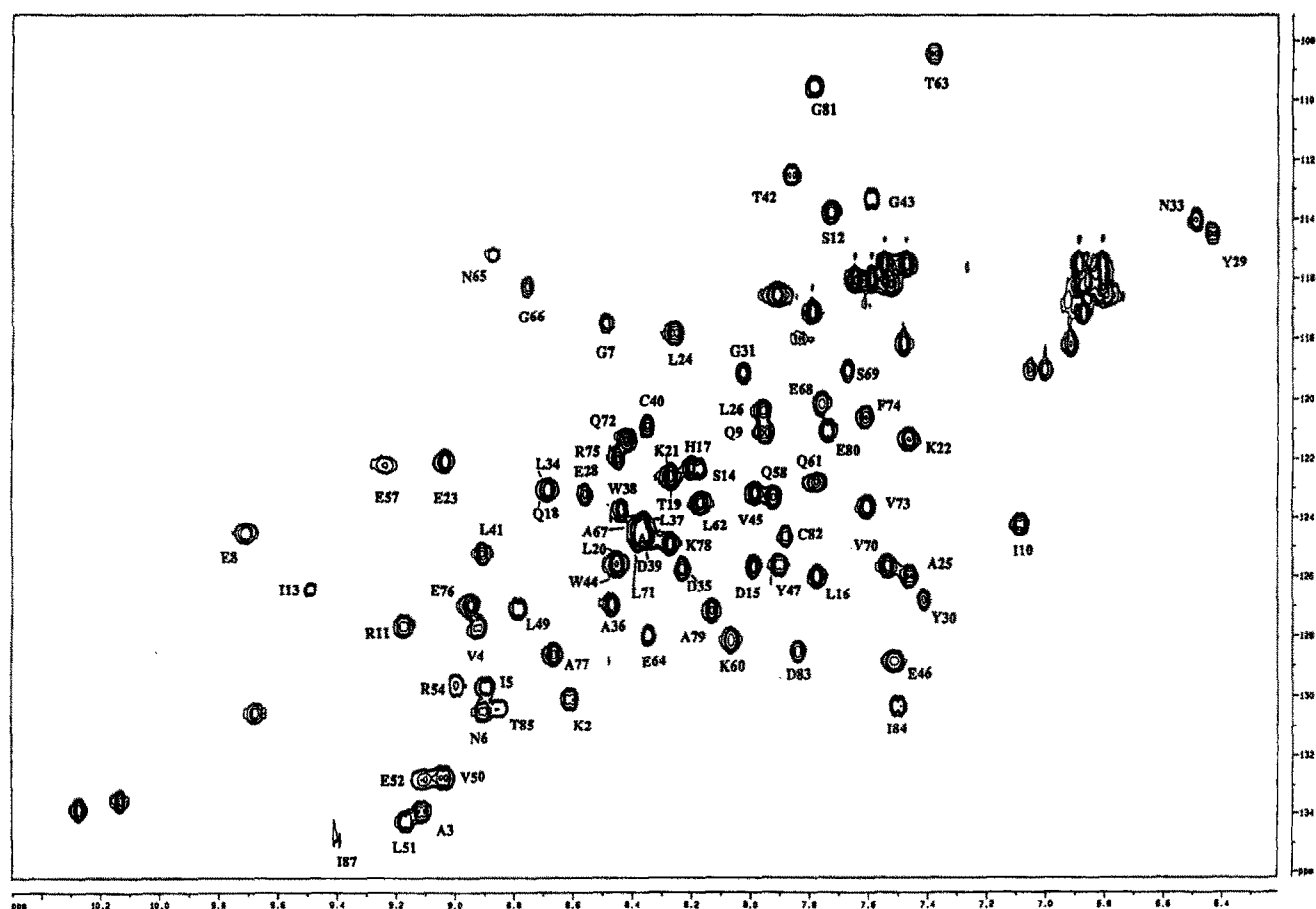


Fig. 1. The ^1H - ^{15}N HSQC spectrum of ^{15}N -labeled barstar recorded at pH 6.6 and 32°C. The ^1H - ^{15}N backbone assignments listed in this paper are indicated.

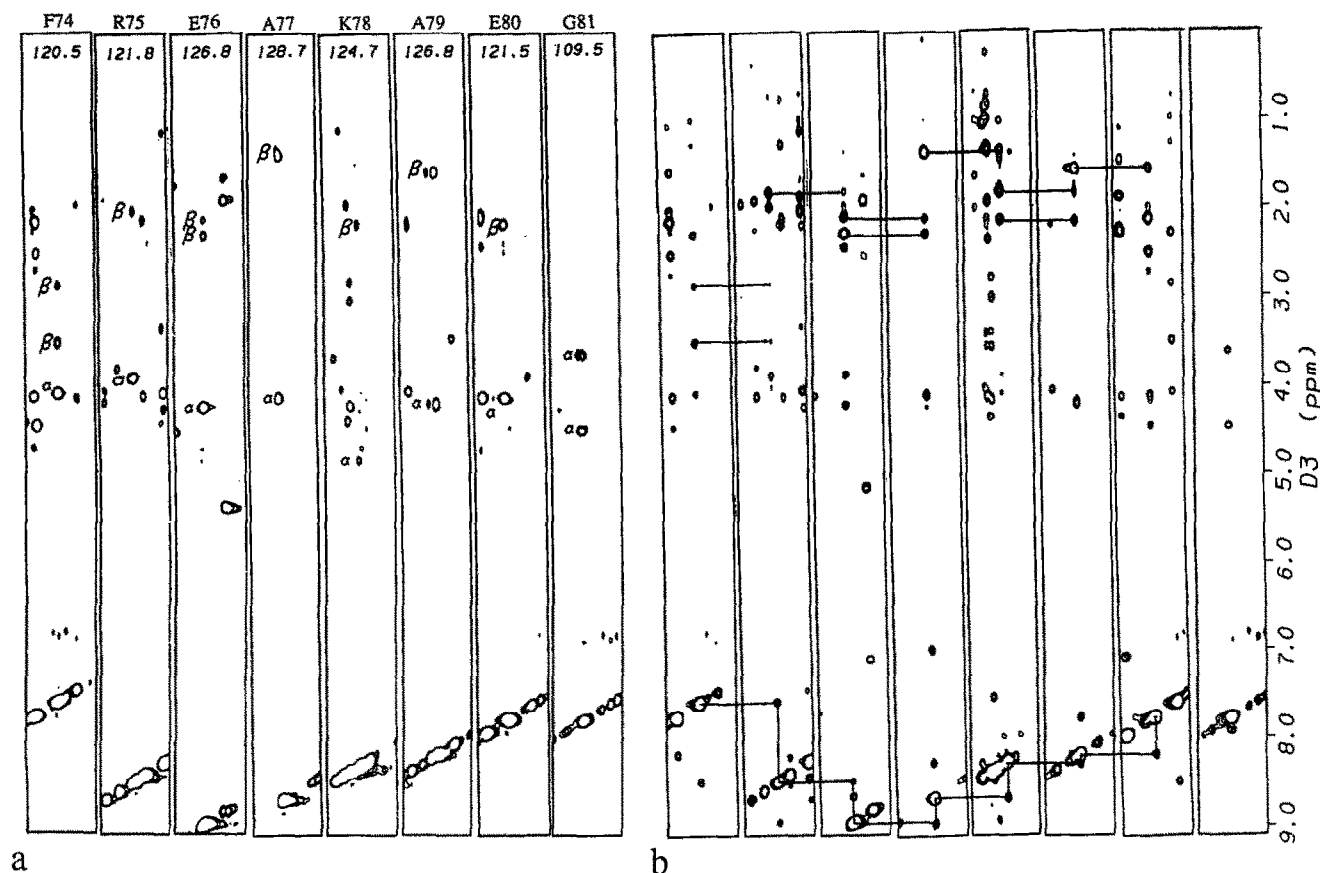


Fig. 2. Series of amide strips taken from identical regions of the 3D TOCSY-HMQC spectrum (a) and 3D NOESY-HMQC spectrum (b) both recorded at 32°C. The residue assignments and ^{15}N chemical shifts are given above the TOCSY strips. Also in the TOCSY, α and β indicate the intra-residue $\text{NH-C}_\alpha\text{H}$ and $\text{NH-C}_\beta\text{H}$ cross-peaks respectively. In the NOESY spectrum, NH-NH and $\text{C}_\beta\text{H-NH}$ sequential connectivities are indicated by solid lines. The NH-NH NOE between E80 and G81 was not observed due to the chemical shift degeneracy of their amide proton shifts. The NOE was observed, however, in the 3D HMQC-NOESY-HMQC spectrum.

spin-systems. The observed NOEs were mainly NH-NH rather than $\text{C}_\alpha\text{H-NH}$ as was expected for a highly α -helical protein. In places where the NH-NH sequential NOE was not observed due to the amide proton shifts of both NHs being equal, connectivities were made using the HMQC-NOESY-HMQC experiment. $\text{C}_\beta\text{H-NH}$ NOEs were also useful in sequentially linking the spin-systems. Once stretches of connectivity between spin-systems were found, the chains of spin-system amino acid identities were fitted into the primary sequence of barstar. An example of this procedure is shown in Fig. 2. Stretches of connectivity were broken either by a proline residue or by the next residue being one of those for which no $^1\text{H-}^{15}\text{N}$ cross-peak was observed in the 2D HSQC spectrum. In addition, stretches of connectivity in loop regions were sometimes broken due to weak or ambiguous NOEs. At the end of the assignment, the remaining spin-systems could be inserted into the primary sequence by a process of elimination. In this way, all the residues corresponding to an HSQC peak were fitted into the primary structure of the protein (Fig. 1). The full assignments are listed in Table

I. Interestingly, a cross-peak was observed for the N-terminal residue in the 2D HSQC spectrum (at 283K) which is an indication that one of its amine protons is hydrogen-bonded.

The configuration of the two proline residues was determined by examining the 3D spectra and also the 2D homonuclear spectra. A medium $d_{\text{HN}\alpha}(i, i+1)$ NOE between leucine 26 and proline 27, and a strong $d_{\alpha\alpha}(i, i+1)$ NOE between tyrosine 47 and proline 48 indicates that the prolines are in the *trans* and *cis* configurations, respectively.

Short and medium range NOEs were analysed in order to give insight into the secondary structure of the protein. Stretches of $d_{\text{HN-HN}}$, $d_{\alpha\text{N}}(i, i+3)$, $d_{\alpha\beta}(i, i+3)$ and $d_{\alpha\text{N}}(i, i+4)$ connections are characteristic of α -helical regions. Stretches of strong $d_{\alpha\text{N}}$ connectivity are characteristic of extended strand [20]. Further information on secondary structure was obtained from coupling constant and hydrogen-deuterium exchange data. $^3J_{\text{HN}\alpha}$ coupling constants under 6 Hz are characteristic of a helical backbone configuration, while those over 7.5 Hz are indicative of an extended backbone configuration.

Table I
 ^1H and ^{15}N Chemical shifts of barstar in water at 32°C (pH 6.6)

Amino acid ^a	Chemical shift (ppm) ^{b,c,d}						
	^{15}N	NH	αH	βH	γH	δH	Other
Lys 1	130.9	8.43	4.44	1.74, 1.65	1.20		
Lys 2	130.2	8.63	5.31	1.81, 1.60	1.39, 1.24		
Ala 3	133.9	9.12	4.96	1.16 (Me)			
Val 4	127.7	8.93	4.90	1.85	0.88(Me) ¹ , 0.71(Me) ²		
Ile 5	129.7	8.90	3.89	1.16	0.71, 0.48, -0.02(Me)	-0.16(Me)	
Asn 6	130.5	8.92	4.90	2.70 ² , 2.87 ³	7.49, 6.92		$\gamma^{15}\text{N}$, 115.5
Gly 7	117.5	8.49	4.55, 3.72				
Glu 8	124.6	9.73	4.25	2.13, 1.98	2.33		
Gln 9	121.2	7.96	4.40	2.36, 2.00			
Ile 10	124.2	7.09	4.18	2.28	0.94(Me)		
Arg 11	127.6	9.18	4.56	2.11, 1.81	1.68		
Ser 12	113.8	7.72	4.79	4.65, 4.18			
Ile 13	126.4	9.50	4.18	2.13	1.69, 1.02(Me)		
Ser 14	122.3	8.18	4.55	4.03, 4.00			
Asp 15	125.7	8.00	4.49	3.07 ² , 2.95 ³			
Leu 16	126.0	7.78	2.29, 1.57	1.69	0.79(Me), 0.69(Me)		
His 17	122.3	8.21	4.41	3.33 ² , 3.08 ³			
Gln 18	123.0	8.69	4.11	2.16, 2.31			
Thr 19	122.7	8.27	4.34	4.12	1.23(Me)		
Leu 20	125.6	8.46	3.72	2.12 ² , 1.12 ³	1.78	0.83(Me), 0.69(Me)	
Lys 21	122.6	8.27	4.12	2.14, 1.96			
Lys 22	121.4	7.46	4.21	2.05, 1.60			
Glu 23	122.1	9.04	4.13	1.76 ² , 2.17 ³			
Leu 24	117.8	8.26	4.35	1.75, 1.63		0.87(Me)	
Ala 25	126.0	7.46	4.18	1.46(Me)			
Leu 26	120.4	7.97	4.02	1.24 ² , 1.04 ³	1.58	0.59(Me), -0.14(Me)	
Pro 27			3.97	1.97, 1.85	2.15	3.40	
Glu 28	123.2	8.57	3.84	2.00, 1.98	2.28		
Tyr 29	114.4	6.44	4.58	3.32 ² , 2.83 ³		7.02	HE, 6.82
Tyr 30	126.8	7.41	4.29	3.09 ² , 2.57 ³		7.00	HE, 6.81
Gly 31	119.1	8.05	3.80, 3.34				
Glu 32							
Asn 33	114.0	6.49	4.46	3.06 ² , 2.78 ³	7.01, 7.05		$\gamma^{15}\text{N}$, 119.1
Leu 34	123.0	8.69	4.23	2.07, 2.00	1.64	1.05(Me), 0.95(Me)	
Asp 35	125.8	8.23	4.60	2.82 ² , 2.65 ³			
Ala 36	126.9	8.48	4.38	1.39(Me)			
Leu	124.1	8.37	4.27	2.25		0.90(Me)	
Trp 38	123.8	8.45	4.25	3.03 ² , 2.88 ³		7.40	HE1, 10.14; HE3, 7.60; $\epsilon^{15}\text{N}$, 133.6; HH2, 6.98; HZ2, 7.57; HZ3, 7.04
Asp 39	124.7	8.35	4.47	3.17, 2.88			
Cys 40	120.9	8.35	4.74	3.17, 2.88			
Leu 41	125.2	8.91	4.21	1.66, 1.23	1.75	0.71(Me), 0.24(Me)	
Thr 42	112.5	7.86	4.27	4.15	0.31(Me)		
Gly 43	113.4	7.59	4.64, 3.85				
Trp 44	125.7	8.46	4.66	3.60, 3.46		6.84	HE1, 10.28; HE3, 7.78; $\epsilon^{15}\text{N}$, 133.9; HH2, 7.36; HZ2, 7.52; HZ3, 7.28
Val 45	123.2	7.99	3.71	1.92	0.97(Me) ¹ , 0.62(Me) ²		
Glu 46	128.8	7.52	4.15	2.02, 1.72			
Tyr 47	125.5	7.91	4.48	3.29, 2.65		7.12	HE, 6.69
Pro 48			5.21	3.66, 2.66	2.58, 2.10	3.82, 3.60	
Leu 49	127.1	8.80	5.41	1.99 ² , 1.74 ³	1.48	1.02(Me), 0.87(Me)	
Val 50	132.9	9.04	4.71	1.98	0.99 ¹ , 0.88 ²		
Leu 51	134.3	9.17	4.97	1.79, 0.78	1.15	0.39(Me), 0.10(Me)	
Glu 52	132.8	9.11	5.18	2.16 ² , 2.03 ³			
Trp 53			5.21	3.61, 2.71		7.28	HE1, 9.68; HE3, 7.19; $\epsilon^{15}\text{N}$, 130.6; HH2, 7.30; HZ2, 71.8; HZ3, 6.84

Table I (continued)
¹H and ¹⁵N Chemical shifts of barstar in water at 32°C (pH 6.6)

Amino acid ^a	Chemical shift (ppm) ^{b,c,d}						
	¹⁵ N	NH	αH	βH	γH	δH	Other
Arg 54	129.3	8.99	4.87	2.08			
Gln 55							
<u>Phe 56</u>			3.68	2.71		5.74	HE, 6.28; HZ, 6.61
Glu 57	122.2	9.23	3.65	2.06			
Gln 58	123.3	7.93	4.17	2.50, 2.18			
Ser 59							
Lys 60	127.9	8.06	3.53	1.90, 1.43			
Gln 61	122.8	7.78	4.18	2.58, 2.26			
Leu 62	123.6	8.17	4.35	1.97, 1.75	1.94	<u>1.24</u> (Me), 1.00(Me)	
Thr 63	108.4	7.39	4.61	4.16	1.12(Me)		
Glu 64	128	8.35	4.10	2.28, 2.24			
Asn 65	115.3	8.89	4.28	3.17	7.54, 6.88		γ ¹⁵ N, 116.2
Gly 66	116.2	8.76	4.14, 3.91				
Ala 67	124.4	8.39	4.19	1.18(Me)			
Glu 68	119.9	7.77	4.17	2.34, 2.12			
Ser 69	119.2	7.68	4.38	4.15			
Val 70	125.6	7.54	3.76	2.47	1.13(Me) ¹ , 0.97(Me) ²		
Leu 71	124.8	8.36	4.40	2.04, 1.48			
Gln 72	121.4	8.42	3.99	2.36, 2.28			
Val 73	123.6	7.61	3.92	2.38	1.08(Me) ¹ , 1.38(Me) ²		
<u>Phe 74</u>	120.6	7.61	4.14	3.57 ² , 2.96 ³		7.55	HE, 7.15; HZ, 7.10
Arg 75	122.1	8.46	3.97	2.11			
Glu 76	126.8	8.95	4.28	2.39, 2.20	2.55		
Ala 77	128.6	8.67	4.17	1.48(Me)			
Lys 78	124.9	8.28	4.87	1.96 ² , 2.27 ³			
Ala 79	127.1	8.15	4.24	1.68(Me)			
Glu 80	121.1	7.75	4.49	2.58 ² , 2.24 ³			
Gly 81	109.6	7.79	4.52, 3.69				
Cys 82	124.6	7.89	3.83	2.10 ² , 1.64 ³			
Asp 83	128.5	7.85	4.55	2.58 ² , 2.87 ³			
Ile 84	130.4	7.50	5.14	1.55	1.16, 0.95(Me)		
Thr 85	130.2	8.84	4.11	4.62	1.20(Me)		
<i>Ile 86</i>	132.3	9.16	4.57	1.63	0.63(Me)		
<i>Ile 87</i>	135.6	9.36	4.19	1.62	0.75(Me)		
<i>Leu 88</i>	133.1	8.66	5.06	1.51, 1.23			
<i>Ser 89</i>	134.9	9.10	5.09	3.45			

^a Residues in italics were only assignable using spectra recorded at 283 K. Underlined residues did not exhibit either an ¹H-¹⁵N crosspeak in the HSQC experiment or a fingerprint peak in the homonuclear 2D COSY experiment at 305 or 283 K and their backbone assignments could therefore not be obtained.

^b Protons that were stereospecifically assigned have superscript numbers after their assignments. HB2 and HB3 in the βH column are indicated by ² and ³, respectively. HG1 and HG2 in the γH column are indicated by ¹ and ², respectively.

^c Methyl proton assignments have (Me) after them.

^d Underlined assignments are made tentatively.

Amide protons that do not exchange when placed in a D₂O buffer are thought to be either involved in hydrogen bonds or significantly shielded from solvent molecules. Fig. 3 shows a summary of all this information.

Although absolutely defining the start and end of elements of secondary structure awaits further structural analysis, the data suggest the following features. The protein consists of three well-defined α-helices (residues 15–26, 33–42, 69–80), one less-defined short helix (around residues 60–65), two well-defined stretches of extended strand (residues 1–5, 49–53) and one less-defined stretch (residues 85–89). The helices were termed

1, 2, 3 and 4 in order of position in the primary structure.

Analysis of the 2D proton and 3D heteronuclear NOESY experiments yielded 813 NOEs: 343 intra-residue NOEs, 380 [*i*–*j*] < 5 and 90 [*i*–*j*] ≥ 5. The distribution of the NOEs is indicated by the diagonal plot in Fig. 4. The pattern of long range NOEs was used to define a tentative topology for the secondary structure elements of barstar (Fig. 5). Helices 1, 2 and 4 form a parallel helix bundle and are connected by long overhand connections. The dipoles of the three helices are orientated in the same direction which may have a func-

tional significance. The third, short, helix is situated in the middle of a long loop. The disulphide bond in barstar is consistent with the observed NOEs. It connects the second and fourth helices at their C-terminal ends.

Too few NOEs were observed between the extended strands to be able to orientate them unequivocally. Those NOEs that were assigned are consistent with the first and second stretches of extended strand forming a parallel sheet. Even at 283 K, the HSQC cross-peaks of the third strand are rather broad, and long-range NOEs from their amide protons are not observed. However,

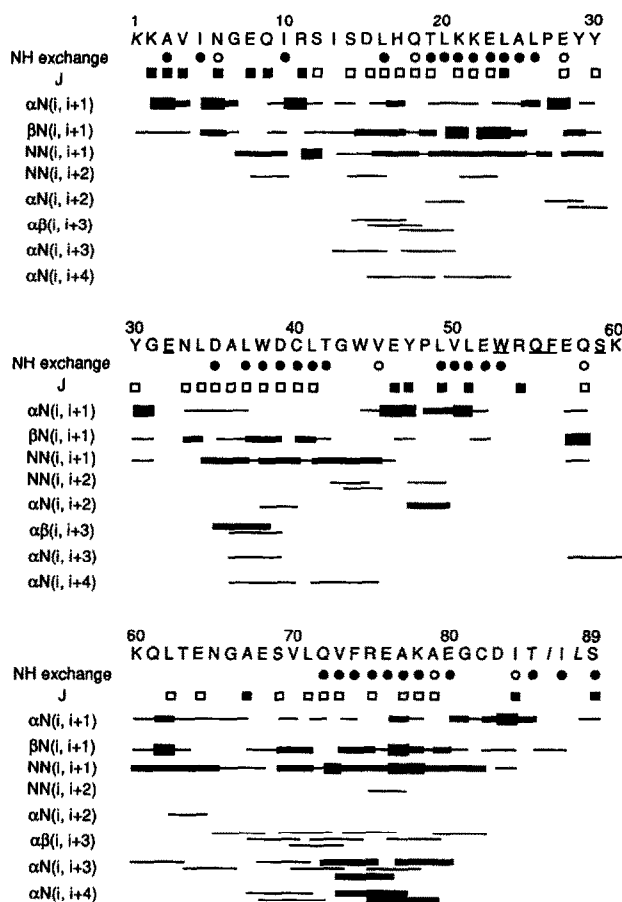


Fig. 3. The sequence of barstar together with a summary of the short- and medium-range NOEs involving NH, $C_\alpha H$ and $C_\beta H$ protons. The NOE intensities are represented by the thickness of the solid bars (the thickest bars representing strong NOEs). The $NN(i, i+1)$ NOE from residue 26 to 27 represents an NOE from the amide proton of L26 to the δ proton of P27. Solid circles represent amide protons that are not fully exchanged 1 h after the protein had been placed into the D_2O buffer (32°C) described in the methods and materials section. Open circles represent amide protons almost fully exchanged under the above conditions. The circle under W53 refers to the indole proton of the side chain of W53. Residues with accurately measured $^3J_{NH\alpha} > 7.5$ Hz are indicated by solid squares, $^3J_{NH\alpha} < 6.0$ Hz by open squares. Residues in italics had an observable 1H - ^{15}N HSQC cross-peak only at 283 K and so NH exchange and J measurements were not obtained for these residues. Underlined residues did not have observable 1H - ^{15}N HSQC cross-peaks at 305 or 283 K.

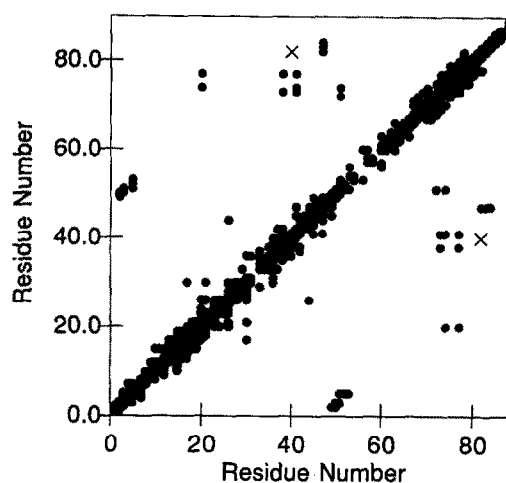


Fig. 4. A diagonal plot showing pairs of amino acids for which an NOE effect has been observed. The disulphide connectivity is indicated by a cross.

NOEs between the N-terminal ends of the latter two strands may indicate that the three strands form a parallel β -sheet with the second strand in the centre. The hydrogen exchange data supports this: residues 49–52 have protected amide protons along its length and could, therefore, be the central strand of a β -sheet, and the first and third extended strands have every second amide proton protected and could, therefore, be the peripheral strands of a β -sheet.

At this stage we have not assigned enough long-range NOEs to define the tertiary structure of barstar mainly

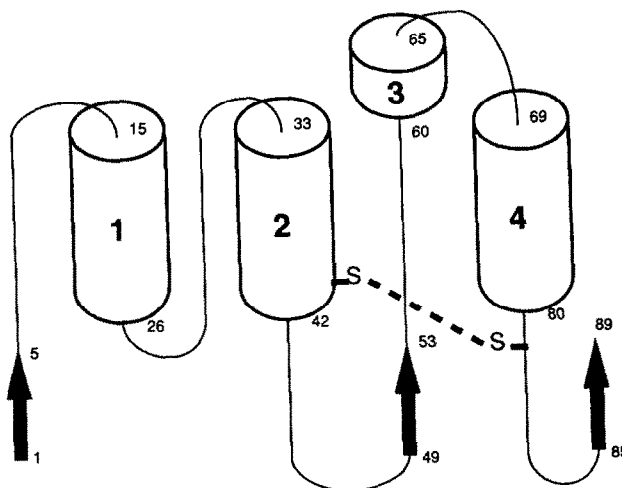


Fig. 5. A schematic diagram showing the secondary structure topology of barstar. The start and end of each piece of secondary structure is labeled with an amino acid number. Helices are represented by cylinders, and extended strands by arrows. The diagram shows the orientation of the elements of secondary structure but does not represent how they pack.

due to the overlap in the methyl region caused by the large number of leucines in the protein. ^1H - ^{13}C heteronuclear NMR will be a necessary step in alleviating this problem (as has been found in other helix-bundle proteins [21,22]). Using the above assignments, we are in the position to use NMR to determine the structure and dynamics of barnase in solution as well as to study its extremely tight interaction with barnase.

REFERENCES

- [1] Hartley, R.W. (1989) *TIBS* 14, 450-454.
- [2] Schreiber, G.E. and Fersht, A.R. (1993) *Biochemistry* 32, 5145-5150.
- [3] Mauguen, Y., Hartley, R.W., Dodson, E.J., Dodson, G.G., Bricogne, G., Chothia, C. and Jack, A. (1982) *Nature* 297, 162-164.
- [4] Bycroft, M., Ludvigsen, S., Fersht, A.R. and Poulsen, F.M. (1991) *Biochemistry* 30, 8697-8701.
- [5] Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976) *J. Chem. Phys.* 64, 2229-2246.
- [6] Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.* 53, 521-528.
- [7] Bax, A. and Davies, D.G. (1985) *J. Magn. Reson.* 65, 355-360.
- [8] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.* 71, 4546-4553.
- [9] Kumar, A., Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.* 95, 1-6.
- [10] Redfield, A.G. and Kuntz, S.D. (1975) *J. Magn. Reson.* 19, 250-254.
- [11] Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967-974.
- [12] Bodenhausen, G. and Ruben, D.L. (1980) *Chem. Phys. Lett.* 69, 185-188.
- [13] Bax, A., Mitsuhashi, I., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn. Reson.* 86, 304-318.
- [14] Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.* 90, 420-425.
- [15] Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) *Biochemistry* 28, 6150-6156.
- [16] Kadhodaie, M., Rivas, O., Tan, M., Mohebbi, A. and Shaka, A.J. (1991) *J. Magn. Reson.* 91, 437-443.
- [17] Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.* 64, 547-553.
- [18] Messerle, B.A., Wider, G., Otting, G., Weber, C. and Wüthrich, K. (1989) *J. Magn. Reson.* 85, 608-613.
- [19] Billeter, M., Neri, D., Otting, G., Qian, Y.Q. and Wüthrich, K. (1992) *J. Biomol. NMR* 2, 257-274.
- [20] Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
- [21] Smith, L.J., Redfield, C., Boyd, J., Lawrence, G.M.P., Edwards, R.G., Smith, R.A.G. and Dobson, C.M. (1992) *J. Mol. Biol.* 224, 899-904.
- [22] Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1992) *Science* 256, 1673-1677.