A Comparison of the pH, Urea, and Temperature-denatured States of Barnase by Heteronuclear NMR: Implications for the Initiation of Protein Folding

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · NOVEMBER 1995

Impact Factor: 4.33 · DOI: 10.1006/jmbi.1995.0618 · Source: PubMed

5 AUTHORS, INCLUDING:

Vickery L Arcus
The University of Waikato
72 PUBLICATIONS **1,972** CITATIONS
SEE PROFILE

Stéphane Vuilleumier
University of Strasbourg
116 PUBLICATIONS **2,824** CITATIONS
SEE PROFILE

Mark Bycroft
Medical Research Council (UK)
102 PUBLICATIONS **6,715** CITATIONS
SEE PROFILE

Alan Fersht
University of Cambridge
629 PUBLICATIONS **52,821** CITATIONS
SEE PROFILE

J. Mol. Biol. (1995) **254**, 305–321

JMB

A Comparison of the pH, Urea, and Temperature-denatured States of Barnase by Heteronuclear NMR: Implications for the Initiation of Protein Folding

Vickery L. Arcus, Stephane Vuilleumier, Stefan M. V. Freund Mark Bycroft and Alan R. Fersht*

MRC Unit for Protein Function and Design Cambridge Centre for Protein Engineering, University Chemical Laboratory Lensfield Road, Cambridge CB2 1EW, UK

**Corresponding author*

The denatured states of barnase that are induced by urea, acid, and high temperature and acid have been assigned and characterised by high resolution heteronuclear NMR. The assignment was completed using a combination of triple-resonance and magnetisation-transfer methods. The latter was facilitated by selecting a suitable mutant of barnase (Ile→Val51) which has an appropriate rate of interconversion between native and denatured states in urea. ^{3}J NH-C$^{\alpha}$H coupling constants were determined for pH and urea-denatured barnase and intrinsic ''random coil'' coupling constants are shown to be different for different residue types. All the denatured states are highly unfolded. But, a consistent series of weak correlations in chemical shift, NOESY and coupling constant data provides evidence that the acid-denatured state has some residual structure in regions that form the first and second helices and the central strands of β-sheet in the native protein. The acid/temperature-denatured states has less structure in these regions, and the urea-denatured state, less still. These observations may be combined with detailed analyses of the folding pathway of barnase from kinetic studies to illuminate the relevance of residual structure in the denatured states of proteins to the mechanism of protein folding. First, the folding of barnase is known to proceed in its later stages through structures in which the first helix and centre of the β-sheet are extensively formed. Thus, embryonic initiation sites for these do exist in the denatured states and so could well develop into true nuclei. Second, it has been clearly established that the second helix is unfolded in these later states, and so residual structure in this region of the protein is non-productive. These data fit a model of protein folding in which local nucleation sites are latent in the denatured state and develop only when they make interactions elsewhere in the protein that stabilise them during the folding process. Thus, studies of the structure of denatured states pinpoint where nucleation sites may be, and the kinetic and protein engineering studies show which ones are productive.

Keywords: unfolded proteins; NMR; protein folding; nucleation; initiation

Introduction

The extent of unfolding of denatured states of proteins under different conditions has long been of interest because of the possible relevance of their conformations to the protein folding pathway. Critical elements along the reaction co-ordinate for protein folding are; the starting point or the denatured state; the final folded state; partially

Present address: S. Vuilleumier, Institut für Mikrobiologie, ETH, CH-8092 Zürich, Switzerland.

Abbreviations used: NMR, nuclear magnetic resonance; 2D, 2-dimensional; 3D, 3-dimensional; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; HSQC, heteronuclear single quantum coherence; HMQC, heteronuclear multiple quantum coherence.

0022–2836/95/470305–17 $12.00/0

stable intermediate states; and the intervening high energy transition states. Conformations appearing late along the reaction co-ordinate are well characterised: the structures of a large number of folded proteins have been determined by both high resolution NMR and X-ray crystallography at the atomic level. In addition, a number of late intermediates on the reaction co-ordinate for several proteins have been characterised by protein engineering (Matouschek *et al*., 1990) and quench-flow NMR techniques (Roder *et al*., 1988; Udgaonkar & Baldwin, 1988; Bycroft *et al*., 1990a; Kuszewski *et al*., 1994). The major transition state of folding for the proteins chymotrypsin inhibitor 2 and barnase have also been mapped at the level of individual residues using protein engineering (Serrano *et al*., 1992b; Otzen *et al*., 1994; Prat Gay *et al*., 1994) and recent theoretical approaches have lent weight to these empirical results (Caflisch & Karplus, 1994; Li & Daggett, 1994). In contrast, little is known about the factors that govern early folding events or the conformations that occupy this extreme on the reaction co-ordinate. Characterisation of the ensemble of conformations that populate the energy surface at this extreme of the reaction co-ordinate is of central importance in understanding how protein folding is initiated and how the initiation precipitates a funnelling of the energy surface such that the protein rapidly folds to a unique late intermediate or to the major transition state saddle point. Direct access to these states is precluded by their negligible population under conditions favourable for folding. However, under conditions of low pH, high temperatures or high concentrations of denaturant, denatured states are fully populated, and, if it is assumed that these closely resemble early states on the protein folding pathway, then their structural characterisation will provide insights into factors governing early folding events.

Characterisation by NMR requires, initially, the assignment of all the resonances in a protein. This process is made particularly difficult for denatured proteins as corresponding nuclei for residues of the same type experience very similar magnetic environments, and the wide dispersion of chemical shifts observed for native proteins is lost on denaturation. We have utilised a combined approach in circumventing the problem of chemical shift degeneracy in the assignment of a denatured protein (Arcus *et al*., 1994). This approach utilises both ''magnetisation transfer'' (Dobson & Evans, 1984; Wider *et al*., 1991) and ''triple resonance'' (Grzesiek & Bax, 1992b; Logan *et al*., 1993; Bax, 1994) NMR methods to give independent sequence-specific assignments. Here, we report the complete backbone ^{15}N, ^{13}C and ^{1}H assignments and near complete side-chain ^{1}H assignments for urea-denatured barnase (Figure 1, Table 1) and from these assignments evidence for ''non-random'' conformations in solution is assessed with reference to deviations from random coil chemical shift values, NOESY and coupling constant data. In addition, the complete amide assignments for a temperature-denatured state at low pH have been determined. The complete assignment for pH-denatured barnase is reported elsewhere (Arcus *et al*., 1994).

Comparisons are then made across the NMR parameters (chemical shift, NOESY contacts, coupling constants and amide exchange) between the pH, urea and pH/temperature-denatured states of barnase and we draw hypotheses from these data in terms of early events on the folding pathway of this protein.

Results

The assignment process for urea-denatured barnase

The transfer of heteronuclear magnetisation could not be observed at mixing times up to 500 ms at the transition midpoint for urea-denaturation of wild-type barnase ([urea] = 3.5 M). The unfolding rate constant under these conditions ($0.12\ s^{-1}$) is very similar to that at the transition midpoint for pH-denaturation ($0.10\ s^{-1}$) where magnetisation transfer peaks, albeit quite weak, are observed (Arcus *et al*., 1994). The ratio of the intensities for the magnetisation transfer peaks *versus* the HSQC correlation peaks is estimated to be 5% at a mixing time of 500 ms (Ernst *et al*., 1987). Thus, even small perturbations from the measured rate constants for unfolding at urea concentrations for 50% unfolding [k_u([urea]50%)] could cause the disappearance of the magnetisation transfer cross peaks under urea-denaturing conditions. To overcome this problem a suitable, conservative barnase mutant, I51V, was chosen with a significantly higher k_u([urea]50%) of $0.49\ s^{-1}$. Correspondingly, at the midpoint for unfolding of this mutant ([urea] = 2.5 M) the symmetry related magnetisation transfer cross peaks were readily observed in the ^{15}N-^{1}H exchange-relayed-HSQC spectrum.

The ^{15}N and ^{1}H amide resonances for native I51V were assigned in a standard manner from 2D and 3D ^{15}N-^{1}H TOCSY and NOESY spectra and with reference to the full wild-type barnase assignments at pH 4.5 (Bycroft *et al*., 1990b). These amide resonances were then transferred to the urea-denatured state *via* heteronuclear magnetisation transfer at the transition midpoint ([urea] = 2.5 M). This experiment gave 36 (of a possible 107) sequence specific ^{15}N, ^{1}H amide assignments for urea-denatured I51V barnase. HSQC spectra for urea-denatured wild-type barnase and urea-denatured I51V were superimposable with the exception of five correlation-peaks. Three of these peaks could be assigned: Gly48, Ser50 and Gly52, residues close to the site of mutation. Given the similarity of the spectra, the 36 sequence-specific assignments for urea-denatured I51V were transferred directly to urea-denatured wild-type barnase.

Sequence-specific assignment was limited to 36

amides from the magnetisation transfer method due to extreme resonance overlap in the centre of the ^{1}H-^{15}N spectral plane. This arises from the presence of two complete sets of HSQC correlation peaks for the equally populated native and denatured states and a third set of weak symmetry-related cross peaks reflecting the transfer heteronuclear magnetisation.

^{15}N-^{1}H assignments for these 36 residues could be extended to their $C^{\alpha}H$ and aliphatic side-chain protons *via* 2D and 3D ^{15}N-^{1}H TOCSY-HMQC spectra (Marion *et al*., 1989a) acquired at 5.5 M urea where the denatured state is fully populated. In addition, these TOCSY spectra provided residue-type assignments for a further 32 spin systems (from unambiguous coincidence of observed chemical shifts with random coil values (Wüthrich, 1986)). The remaining spin systems belong to aromatic, aspartate and asparagine residues whose random coil proton chemical shifts are closely overlapping.

These sequence-specific and residue type assignments from magnetisation transfer and ^{15}N-^{1}H TOCSY-HMQC spectra were confirmed, augmented and the assignment was completed using six (three pairs) of triple resonance experiments.

In principle, the two pairs of NMR experiments HNCO, HN(CA)CO and CBCA(CO)NH, CBCANH give independent stepwise correlation of backbone ^{15}N, ^{1}H and ^{13}C chemical shifts (Figure 2). Pragmatically, for a denatured protein, the triple resonance assignment process begins with the identification of residue type pairs, the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ random coil chemical shifts from CBCA(CO)NH and CBCANH spectra unambiguously define a residue type (Grzesiek & Bax, 1993). Where these residue-type pairs are unique in the sequence, the sequence-specific assignment follows directly. This breaks down if a residue-type pair occurs more than once in the protein sequence and equivalent sequence pairs must be assigned by sequential correlation of the ^{13}CO chemical shifts in the case of the HNCO, HN(CA)CO pair and, independently, correlation of the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts from CBCA(CO)NH, CBCANH spectra (Figure 2). The HN(CO)CA and HNCA spectra provide increased resolution in the carbon dimension in comparison to their CBCA counterparts and help in the demarcation of near degenerate $^{13}C^{\alpha}$ chemical shifts.

The chemical shift degeneracy seen in the spectra for denatured proteins precludes sequence-specific assignment by traditional methods. The combined assignment process we have used here is a juxtaposition of: (1) utilising the residue-type information harboured in the ^{1}H and ^{13}C chemical shift degeneracy, (2) circumventing the degeneracy by using independent pairs of triple resonance experiments for sequential correlation of backbone chemical shifts, and (3) confirmation of isolated sequence-specific assignments from heteronuclear magnetisation transfer experiments.

NMR derived parameters for urea-denatured barnase

The majority of ^{1}H resonances for urea-denatured barnase are in close agreement with random coil values (Figure 1 and Table 1 for the complete assignment). Those resonances which deviate to a significant extent ($\Delta\delta > 0.3$ ppm for labile protons and $\Delta\delta > 0.1$ ppm for non-labile protons) are clustered in three regions. These are:

Gly9 to Tyr17	(G9-NH) (V10-NH) (A11-Me) (Y13-βH) (L14-NH, βH, δMe) (T16-αH, βH, γMe) (Y17-βH)
Arg69 to Ile76	(R69-βH, γH) (R72-αH, βH, γH) (E73-αH) (I76-αH, βH, γMe)
Arg87 to Ile96	(R87-αH, γH) (I88-αH, βH, γMe) (L89-βH) (S91-βH) (L95-NH, αH, βH, δMe) (I96-αH, βH, γMe)

Figure 3 shows deviations from random coil chemical shift values for all backbone nuclei except ^{13}CO. The above clustering of deviations for ^{1}H chemical shifts is not reflected in the backbone nuclei and no discernible trends are evident for these deviations. Thus the above clustering presents only weak evidence for the presence of local structured conformations. Conversely, the broad agreement with random coil values for resonances of urea-denatured barnase is strong evidence for the dominance of random coil conformations in solution.

The ^{1}H-^{15}N HMQC-NOESY spectra are also dominated by intra-residue and H^{α}-NH ($i, i + 1$) NOE contacts typical of random coil conformations. The consistent observation of H^{α}-NH ($i, i + 1$) NOE contacts along the length of the sequence provides further confirmation of sequence-specific assignments. A number of NOESY peaks that are inconsistent with random coil configurations are observed in these spectra. These include strong sequential NH-NH ($i, i + 1$) NOESY peaks for (Asp8, Gly9), (Val10, Ala11), (Gln15, Thr16), (Val45, Ala46) and (Ile96, Tyr97). The aromatic ring protons of Trp35 and Trp94 show contacts to methylene and methyl protons in homonuclear NOESY spectra, but the aliphatic protons cannot be unambiguously assigned here. Tyr17 δH shows a contact with Thr16 Me protons and this is accompanied by an upfield shift for Thr16 Me to 1.05 ppm.

The observation of strong sequential NH-NH ($i, i + 1$) NOE contacts and aromatic to side-chain NOE contacts in regions of the protein which coincide with side-chain ^{1}H chemical shift deviations from random coil values strongly suggests the presence of local non-random conformations in these regions (Asp8 to Tyr17 and Arg87 to Ile96). These two regions span the first helix and central strands of β-sheet in the native protein and these elements of secondary structure are shown to be well formed in the folding intermediate from protein engineering experiments. In addition, these two regions have been predicted theoretically to

constitute initiation sites for folding based on the burial of hydrophobic area in the native state (Moult & Unger, 1991; Serrano *et al*., 1992a).

^3J NH-C$^\alpha$H coupling constants were measured for 73 residues in urea-denatured barnase (Figure 9A). Significantly, ^3J NH-C$^\alpha$H coupling constants for residues of the same type are closely coincident (the widest spread of values is 1.5 Hz for isoleucine) and between residue types are markedly different. For example, alanine appears to have a ''random coil'' coupling constant of 5.8 Hz whereas valine has a ''random coil'' value of 8.2 Hz. It is deviations from these random coil values which are significant for small populations of structured states in solution.

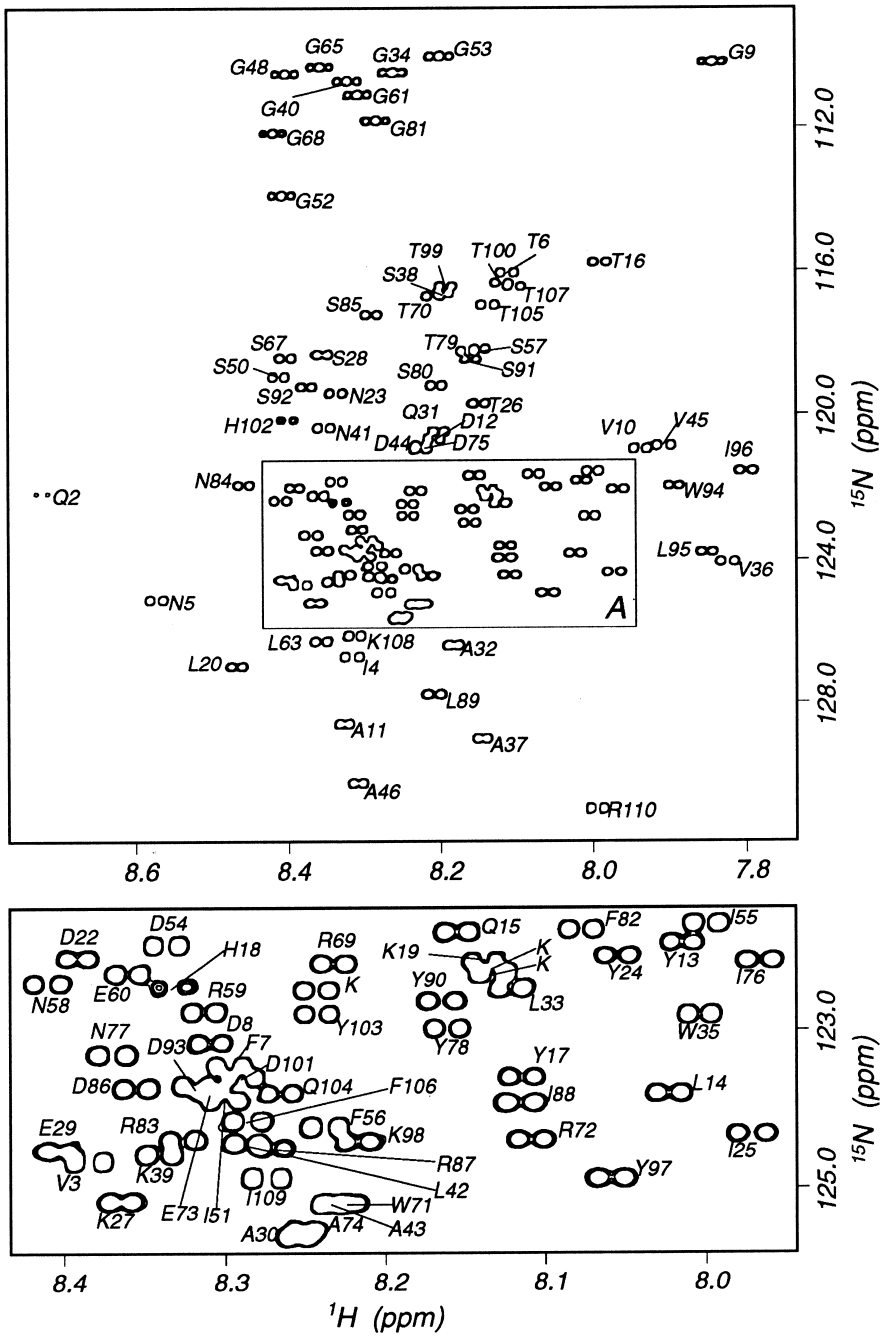

Figure 1. The 2D ^{15}N-^{1}H HSQC spectrum of uniformly ^{15}N-labelled barnase in the urea-denatured state (2.9 mM protein, 5.5 M urea / 10% ^2H_2O as solvent, pH 4.5 and a temperature of 30°C). All backbone amides are assigned with the exception of Ala1, which is not seen. In addition, Lys49, Lys62 and Lys66 cannot be differentiated and these are labelled with just the one-letter amino acid code, K. Boxed area A above is shown on an expanded scale below. The amide peaks in this spectrum are unusually sharp and this results in visible ^3J NH-C$^\alpha$H coupling patterns for each amide peak.

Table 1. ^{1}H, ^{15}N and ^{13}C chemical shifts for urea-denatured barnase ([urea] = 5.5 M, pH 4.5, 30°C)

	Chemical shift (ppm)[a]							
Residue	^{13}CO	$^{13}C^{\alpha}$[b]	$^{13}C^{\beta}$[b]	^{15}N	NH	$C^{\alpha}H$	$C^{\beta}H$	Other protons[c]
A1	171.6	49.7	16.9					
Q2	172.9	53.9	27.4	122.19	8.72	4.43	2.35, 1.98	γH, 2.04
V3	173.8	60.5	30.5	124.73	8.38	4.14	2.01, 0.91	
I4	173.7	58.9	36.7	126.71	8.32	4.22	1.82	γH, 1.15; γCH_3, 0.86
N5	173.2	51.0	37.0	125.10	8.57	4.81	2.74	
T6	172.3	59.8	68.0	116.01	8.11	4.30	4.20	γCH_3 1.13
F7	172.4			123.47	8.30	4.75	3.16, 3.02	
D8	174.1	52.0	38.4	123.18	8.31	4.58	2.65	
G9	172.0	43.4		110.16	7.83	3.92		
V10	174.0	60.6	30.4	120.92	7.92	4.09	2.08	γCH_3, 0.92
A11	175.5	50.6	16.7	128.59	8.33	4.29	1.29	
D12	174.2	52.4	38.9	120.70	8.21	4.63	2.76, 2.65	
Y13	174.0	56.5	36.2	121.82	8.01	4.49	3.05	
L14	175.5	53.5	40.0	123.84	8.02	4.27	1.55	γH, 1.62; δCH_3, 0.85, 0.77
Q15	174.2	54.1	26.8	121.73	8.16	4.30	1.99, 2.05	γH, 2.34
T16	172.1	60.3	68.0	115.74	8.00	4.23	4.09	γCH_3, 1.05
Y17	173.4	56.0	36.5	123.68	8.11	4.53	2.96, 2.87	
H18	173.9	52.6	38.8	122.42	8.33	4.64	3.22, 3.08	
K19	174.4	54.0	28.5	122.16	8.13	4.39	1.80, 1.73	γH, 1.41
L20	173.4			126.30	8.36	4.62	1.59	δCH_3, 0.96, 0.93
P21	174.6	61.1	29.8					
D22	174.1	52.7	38.8	122.03	8.39	4.53	2.64	
N23	172.7	51.1	36.3	119.41	8.34	4.67	2.75	
Y24	173.5	56.2	36.6	122.00	8.05	4.54	3.00	
I25	174.0	59.1	36.6	124.32	7.97	4.19	1.81	γH, 1.14; γCH_3, 0.86
T26	172.5	59.9	68.1	119.79	8.15	4.32	4.25	γCH_3, 1.24
K27	174.6	54.7	30.9	125.25	8.36	4.33	1.80	γH, 1.43
S28	173.5	56.7	61.8	118.37	8.35	4.43	3.88	
E29	174.3	54.8	27.7	124.57	8.40	4.31	2.07, 1.98	γH, 2.34
A30	176.0	50.9	16.6	125.71	8.25	4.21	1.37	
Q31	174.1	54.2	27.1	120.50	8.20	4.25	2.00	γH, 2.34
A32	175.8			126.37	8.18	4.30	1.38	
L33	176.0	53.4	40.2	122.41	8.12	4.31	1.61	γH, 1.48; δCH_3, 0.88
G34	171.7	43.2		110.49	8.27	3.94		
W35	174.0	55.7	27.4	122.79	8.00	4.65	3.24	
V36	173.2	59.9	31.0	124.03	7.82	3.98	1.91	γCH_3, 0.83
A37	175.7	50.5	16.7	129.01	8.14	4.15	1.36	
S38	172.7	56.2	62.0	116.63	8.20	4.40	3.92, 3.85	
K39	174.9	54.6	30.8	124.62	8.34	4.34	1.86, 1.74	γH, 1.42
G40	171.6	43.2		110.70	8.32	3.95		
N41	173.4	51.2	36.8	120.38	8.35	4.73	2.82, 2.76	
L42	175.1	53.5	40.1	124.49	8.29	4.27	1.62	δCH_3, 0.92
A43	175.4	50.6	16.9	125.23	8.23	4.30	1.38	
D44	174.0	52.3	38.8	120.89	8.22	4.64		
V45	173.4	59.8	30.6	120.90	7.90	4.15	2.08	γCH_3, 0.92
A46	173.4	48.4	15.7	130.26	8.31	4.58	1.35	
P47	175.6							
G48	172.0	43.1		110.48	8.40	4.01, 3.95		
K49	174.5	54.2	31.0					
S50	172.6	56.1	62.0	119.00	8.42	4.54	3.88	
I51	174.6	59.4	36.7	123.93	8.30		1.90	γH, 1.19; γCH_3, 0.86
G52	172.3	43.0		113.43	8.41	4.02, 3.99		
G53	171.7	43.2		109.98	8.20	3.97		
D54	174.1	52.1	39.1	121.85	8.34		2.67	
I55	174.1	59.7	36.5	122.08	7.96	4.10	1.79	γH, 1.11; γCH_3, 0.92 δCH_3, 0.74
F56	173.9	55.7	37.2	124.28	8.24	4.70	3.22, 3.01	
S57	172.3	56.4	62.0	118.19	8.15	4.44	3.88	
N58	173.4	51.3	36.6	122.38	8.41	4.74	2.85	
R59	174.4			122.76	8.31	4.32	1.89, 1.77	γH, 1.64
E60	174.7			122.26	8.35	4.32	2.11, 1.99	γH, 2.35
G61		43.2		111.08	8.31	4.00, 3.93		
K62	174.1							
L63	173.2			126.99	8.47	4.56	1.61	γH, 1.48
P64	175.4	61.3	29.9					
G65	172.0	43.0		110.28	8.36	4.04	3.92	
K66	174.7	54.4	31.0					
S67	172.9	56.4	62.1	118.43	8.40	4.49	3.90	
G68	171.8	43.1		112.13	8.42	4.05, 3.99		

continued overleaf

Table 1. *continued*

Residue	Chemical shift (ppm)[a]							
	^{13}CO	$^{13}C^{\alpha}$[b]	$^{13}C^{\beta}$[b]	^{15}N	NH	$C^{\alpha}H$	$C^{\beta}H$	Other protons[c]
R69	174.4	54.2	28.8	122.13	8.23	4.39	1.71, 1.54	γH, 1.38
T70	172.1	59.7	68.2	116.71	8.21	4.38	4.20	γCH_3, 1.17
W71	173.7		27.4	125.23	8.23	4.68	3.26	
R72	173.8	53.9	28.8	124.43	8.11	4.23	1.72, 1.58	γH, 1.40
E73	174.0	54.8	27.6	123.83	8.31	4.15	1.91	γH, 2.23
A74	176.0	50.6	16.5	125.62	8.25	4.27	1.38	
D75	174.3			120.93	8.22	4.56	2.67	
I76	173.6	59.4	36.6	121.63	8.00	4.10	1.75	γH, 1.01; γCH_3, 0.73
N77	173.0	51.1	36.8	123.35	8.37	4.72	2.75, 2.68	
Y78	174.1	56.4	36.7	123.00	8.16	4.59	3.08, 2.94	
T79	172.5	59.8	68.2	118.26	8.16	4.32	4.13	γCH_3, 1.18
S80	172.9	56.5	61.9	119.23	8.21	4.44	3.88	
G81	171.7	43.1		111.80	8.28	3.95, 3.91		
F82	173.7	55.9	37.5	121.64	8.07	4.60	3.10, 3.01	
R83	173.8	53.9	28.8	124.44	8.32	4.32	1.81, 1.69	γH, 1.53
N84	173.4	51.6	36.7	121.94	8.45	4.69	2.81	
S85	172.2	56.5	61.9	117.24	8.29	4.44	3.95, 3.84	
D86				123.72	8.32	4.63	2.72	
R87	173.9	54.1	28.5	124.53	8.27	4.27	1.77	γH, 1.42
I88	173.7	59.2	36.5	123.90	8.11	4.06	1.74	γH, 1.11; γCH_3, 0.70
L89	174.7	52.8	40.3	127.77	8.21	4.33	1.50	γH, 1.39; δCH_3, 0.80, 0.87
Y90	173.9	56.0	36.6	122.63	8.16	4.58	3.05, 2.87	
S91	172.7	56.2	62.1	118.51	8.16	4.43	3.75	
S92	172.3			119.25	8.37	4.40	3.89, 3.78	
D93	174.0	52.9	38.5	123.77	8.36	4.72	2.72	
W94	174.1	55.6	27.0	122.02	7.89	4.62	3.29	
L95	175.0	53.5	40.0	123.80	7.85	4.19	1.46	γH, 1.38 δCH_3, 0.94, 0.75
I96	174.0	59.4	36.3	121.57	7.80	4.04	1.72	γH, 1.07; γCH_3, 0.73
Y97	173.6	55.8	36.6	124.91	8.06	4.60	3.05, 2.88	
K98	174.4	54.0	31.0	124.47	8.22	4.38	1.77, 1.70	γH, 1.36
T99	172.7	59.8	68.2	116.44	8.19	4.43	4.29	γCH_3, 1.23
T100	172.2	59.7	68.0	116.33	8.11	4.37	4.29	γCH_3, 1.18
D101	173.8	52.3	39.0	123.62	8.29	4.61	2.69	
H102	172.0	53.4	26.6	120.17	8.40	4.64	3.16, 3.07	
Y103	172.1	56.0	37.0	122.77	8.24	4.56	3.03, 2.93	
Q104	173.6	53.7	27.5	123.82	8.27	4.60	2.36	
T105	172.1	59.8	68.2	116.94	8.14	4.30	4.14	γCH_3, 1.15
F106	174.1	56.5	36.7	124.19	8.29	4.65	3.15, 3.01	
T107	171.8	59.7	68.1	116.38	8.09	4.36	4.25	γCH_3, 1.18
K108	174.1	54.3	30.9	126.12	8.31	4.37	1.82, 1.72	γH, 1.40
I109	173.3	59.5	36.5	124.95	8.27	4.16	1.85	γCH_3, 0.91
R110	178.8	55.6	29.3	131.90	8.00	4.20	1.83, 1.73	γH, 1.59; δH, 3.19

[a] Proton chemical shift values are relative to internal 3-trimethylsilyl-d^4-propanoic acid. ^{15}N chemical shift values are relative to external [^{15}N] ammonium chloride and ^{13}C chemical shifts are indirectly referenced to the solvent ^1H frequency.

[b] ^{13}C chemical shifts are not included where the assignment is ambiguous. This is most commonly as a result of sequence degeneracy, which in turn causes chemical shift degeneracy. For example, Leu20 and Leu63 are both preceded by lysine and followed by proline.

[c] It is not possible to distinguish β and γ protons for leucine side-chains. The ring protons of proline, phenylalanine, tryptophan, histidine and tyrosine could not be assigned due to spectral overlap.

Deviations from random coil chemical shift values between denatured states

The factors affecting chemical shift are numerous and difficult to quantify, although recent statistical, semi-empirical and theoretical approaches have endeavoured to tap the structural information inherent in chemical shift values (Wishart *et al*., 1991; De Dios *et al*., 1993; Osapay *et al*., 1994; Williamson *et al*., 1995). Significant deviations in chemical shifts from values for model compounds give preliminary evidence for ''non-random'' conformations (Wüthrich, 1986). Figure 4 shows the deviations from random coil values for the backbone nuclei (excluding ^{13}CO) of pH-denatured barnase (cf. Figure 3 for urea-denatured barnase). The clustering of negative deviations for ^{15}N chemical shift values in the regions Leu14 to Lys19, Ser28 to Gly34 and Leu95 to Thr100 provides initial weak evidence that indicates the presence of stable structured conformations in regions of the native protein which contain the first helix (Thr6 to His18), the second helix (Thr26 to Gly34) and the central strands of the β-sheet (Asp86 to Thr99). Residual structure in the region of the first helix of barnase is reinforced by large deviations for the amide protons of Gly9 and Val10 (Figure 4B) and a marked clustering of positive deviations for the $^{13}C^{\alpha}$ chemical shifts of Tyr13 to Thr16 (Figure 4C). Significant deviations in the region of the second helix are also seen for the $^{13}C^{\alpha}$ chemical shifts of Lys27 to Ala32 (Figure 4C). The presence of residual structure in the region of the central strands of the β-sheet is suggested by negative deviations for the

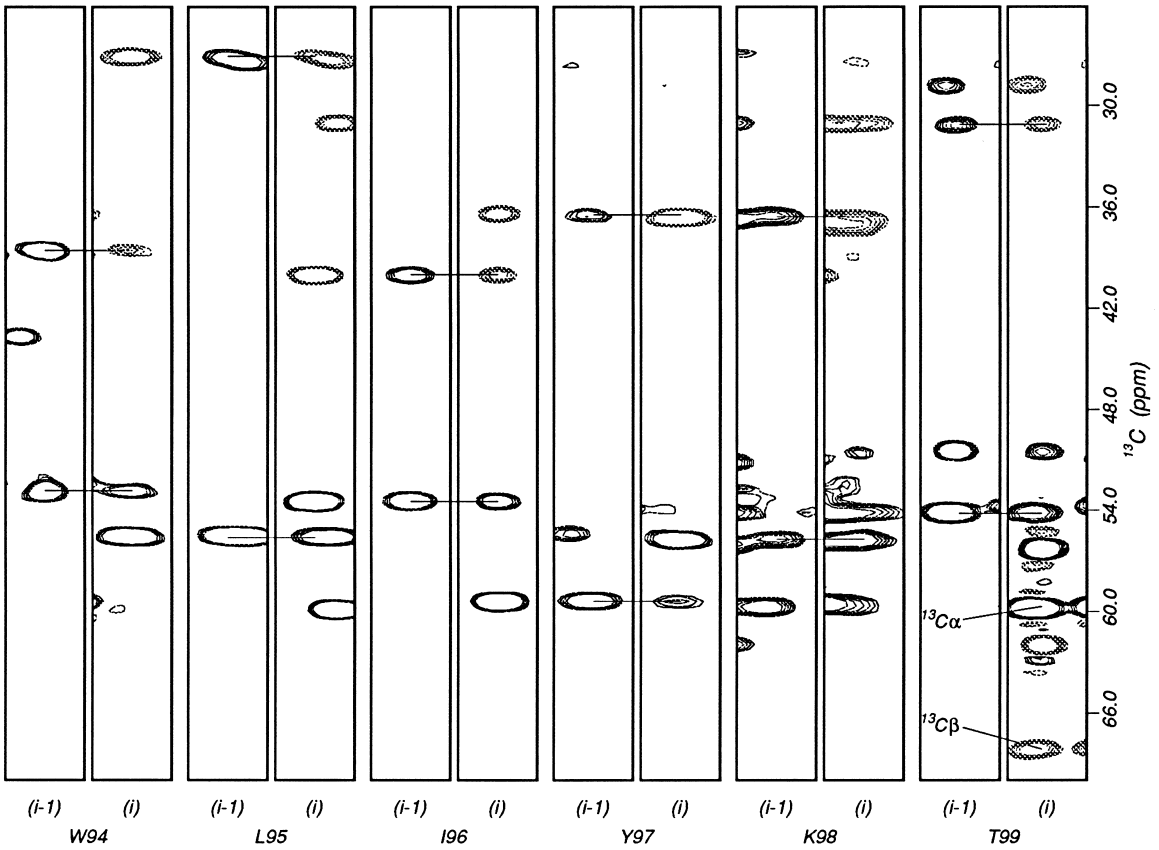

Figure 2. Selected paired regions of the 3D CBCA(CO)NH and CBCANH spectra acquired with uniformly ^{15}N, ^{13}C-labelled barnase in the urea-denatured state (2.5 mM protein, 5.5 M urea / 10% $^{2}H_2O$ as a solvent, pH 4.5 and 30°C) showing sequential connection of $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts. The left strip of each pair is taken from the CBCA(CO)NH spectrum which correlates the amide ^{1}H and ^{15}N resonances of residue *i* with the $^{13}C^{\alpha}$ and $^{\beta}C^{\beta}$ resonances of residue $i-1$. The right strip of each pair is taken from the CBCANH spectrum which correlates the amide ^{1}H and ^{15}N resonances of residue *i* with the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances of both residue $i-1$ and residue *i*. Resonances for $^{13}C^{\alpha}$ appear in this spectrum as positive peaks, and the contours are shown as continuous lines. $^{13}C^{\beta}$ resonances appear as negative peaks, and their contours are shown as broken lines. $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances for residue $i-1$ are connected between spectra with horizontal lines. Sequential residue assignments are given below each of the paired strips.

amide protons of Leu 95 and Ile96 (Figure 4B), and negative deviations for the αH protons of Asp93, Leu95 and Ile96 (Figure 4D).

These backbone chemical shift deviations can be directly compared with those for urea-denatured barnase (Figure 3). There are only marginal differences for the backbone proton nuclei between the two denatured states, but for the heteronuclei, $^{13}C^{\alpha}$ and ^{15}N, clear trends emerge (Figure 5). The positive clustering seen in the regions of the first and second helices and to a lesser extent the fourth β-strand for the $^{13}C^{\alpha}$ chemical shifts of pH denatured barnase are lost under urea-denaturation (Figure 5). This comparison is mirrored to some extent for the negative clustering of ^{15}N random coil deviations in the three regions of pH-denatured barnase (Figure 5).

Positive deviations from random coil values for $^{13}C^{\alpha}$ chemical shifts have been shown to be indicative of helices (Spera & Bax, 1991). Figure 4C would suggest that for pH-denatured barnase, helical conformations are populated in those regions where helices are present in the native state and that evidence for residual structure (positive deviations) between residues Ser92 and Ile96 is inconsistent with native-like structure. Figure 5 directly implies that these structured regions in pH-denatured barnase are either less structured or populated to a lesser extent in urea-denatured barnase.

Quantifying greatest chemical shift changes for temperature denaturation: the normalised vector in the ^{1}H-^{15}N plane

Figure 6 shows the incremental changes in chemical shifts in the ^{1}H-^{15}N plane with rising temperature for two of the three tryptophan residues and two lysine residues in pH-denatured barnase (pH 2.0, 30 to 58°C). The adjacent vectors describe the corresponding translation, which, in most cases, is closely linear. To determine the magnitude of the vector such that the value would not be biased towards one dimension, horizontal and vertical components were normalised (this involves dividing each of the horizontal and vertical

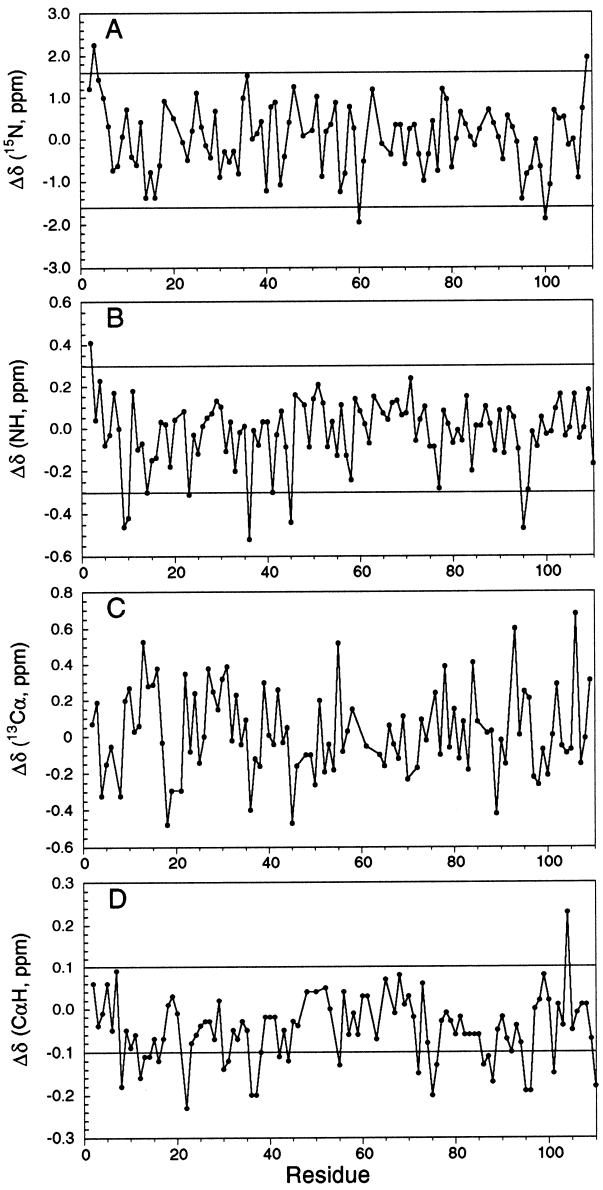

Figure 3. The deviations from random coil values for the chemical shifts for backbone nuclei of urea-denatured barnase. A, Deviations for the amide ^{15}N resonances. Random coil values were calculated by ''sequence correction'' from Braun *et al*. (1994). B, and D, Deviations for amide NH and C$^{\alpha}$H protons, respectively. Random coil values are from Wüthrich (1986). C, Deviations for ^{13}C$^{\alpha}$ chemical shifts. Random coil values are from Wishart *et al*. (1995). In all cases, random coil values were corrected for the three residues followed by proline as described by Wishart *et al*. (1995). The horizontal lines show the accepted ranges outside which deviations are considered to be significant (all ^{13}C$^{\alpha}$ chemical shifts lie within this range). Deviations for both ^{15}N and ^{13}C$^{\alpha}$ were centred about zero to normalise referencing differences.

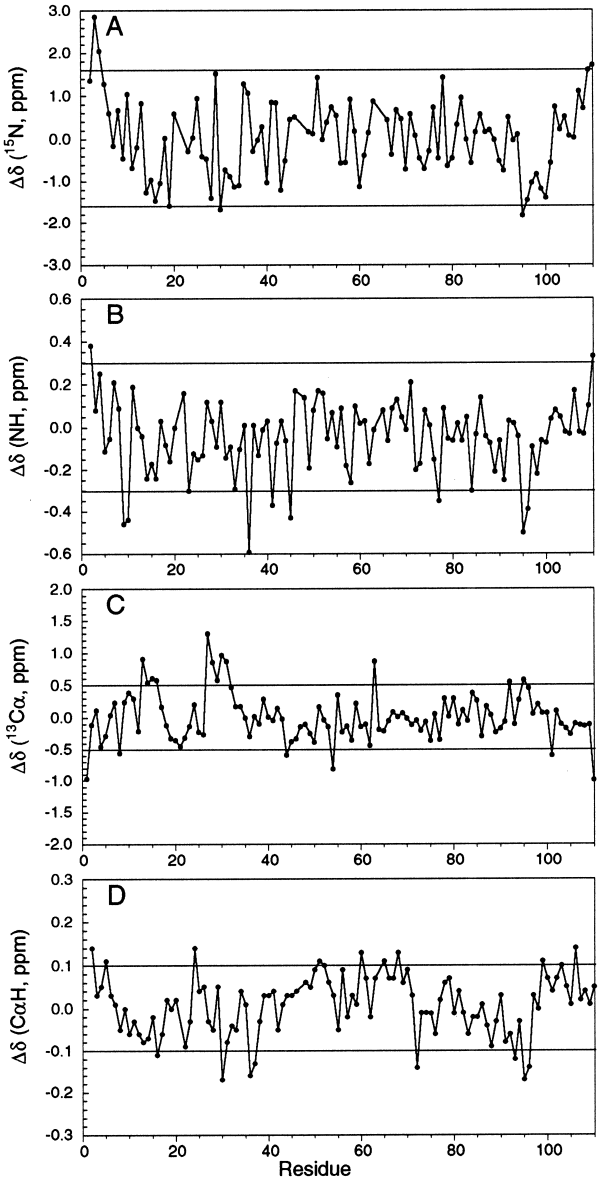

Figure 4. The deviations from random coil values for the chemical shifts for backbone nuclei of pH-denatured barnase. A, Deviations for the amide ^{15}N resonances. Random coil values were calculated by ''sequence correction'' from Braun *et al*. (1994). B and D, Deviations for amide NH and C$^{\alpha}$H protons, respectively. Random coil values are from Wüthrich (1986). C, Deviation for 13C$^{\alpha}$ chemical shifts. Random coil values are from Wishart *et al*. (1995). In all cases random coil values were corrected for the three residues followed by proline (Wishart *et al*., 1995). The horizontal lines show the accepted ranges outside which deviations are considered to be significant. Deviations for both ^{15}N and 13C$^{\alpha}$ were centred about zero to normalise referencing differences.

components by the greatest values found for each dimension). Figure 7 shows the normalised vector magnitudes across the protein describing the changes in the ^{1}H-^{15}N plane with increasing temperature. Regions of comparatively larger changes are seen in the first third of the protein, Thr6 to Lys39, and in the last fifth, Ile88 to Phe107 as well as isolated spikes around Ala46, Ser57 and Leu63. Assuming that the vector magnitudes are in some sense a measure of unfolding of the protein on increasing the temperature, then these data are in broad agreement with results from the above

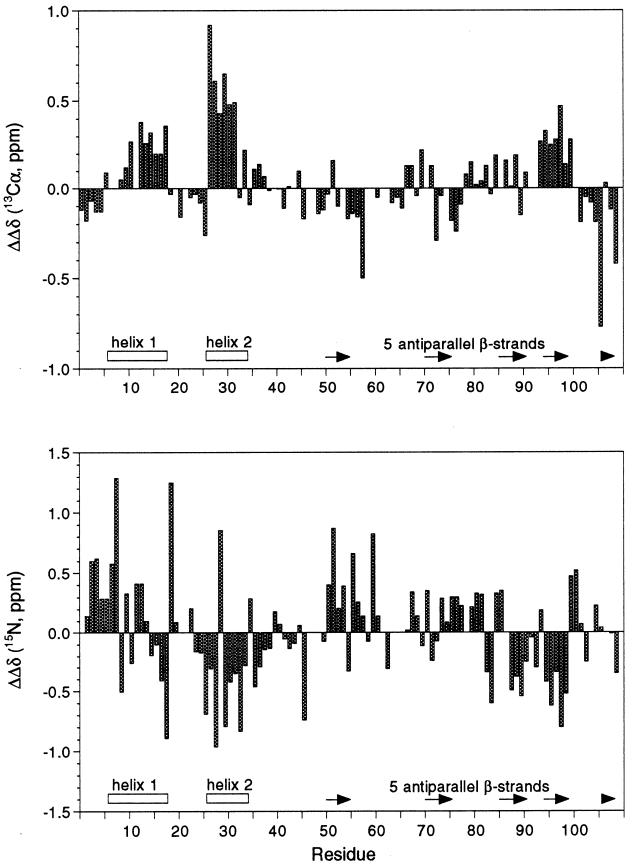

Figure 5. A comparison of the deviations from random coil chemical shifts between pH and urea-denatured barnase for the backbone heteronuclei $^{13}C^{\alpha}$ and amide ^{15}N. Elements of secondary structure for the native protein are given along the residue axis.

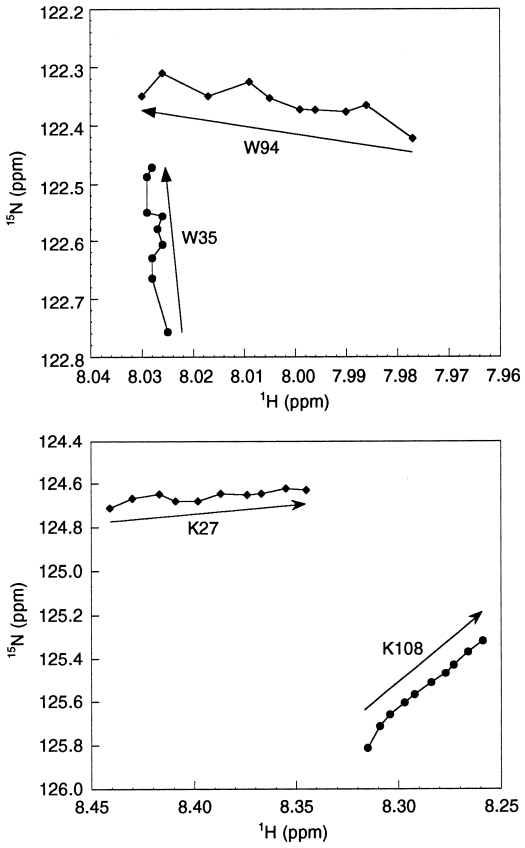

Figure 6. A graphical representation of the changes in the ^{1}H-^{15}H plane for pH-denatured barnase with increasing temperature. The linked points show the movement of the correlation peaks for ten temperature values between 30 and 58°C. The vectors describe the overall movement and the normalised magnitude of these vectors are shown in Figure 7 for all backbone amides. These changes are relative to Gln2, which was used as an internal reference.

analysis in terms of deviations from random coil values.

Degeneracy analysis

The above comparisons of chemical shift values for the different denatured states of barnase may be criticised insofar as the conditions of solvent and temperature may make a significant contribution which cannot be assessed. The effects of solvent and temperature are normalised if it is assumed that the key determinants of amide ^{1}H and ^{15}N chemical shifts are the side-chain types of its own and the preceding residue (Braun *et al*., 1994; Wishart *et al*., 1995). It follows that for a typical ''random coil'', pairs of residues that are equivalent in sequence should also show degeneracy in the ^{1}H-^{15}N plane. This can be seen in Figure 6 where Lys27 and Lys108 are convergent (each residue is preceded by threonine) with increasing temperature. Although Trp35 and Trp94 are not part of equivalent sequence-pairs they also typify a trend for residues of the same type to move to near degenerate positions with increasing temperature or urea. The magnitude of the normalised vector (see above discussion) between equivalent sequence-pair residues provides a quantity which is independent of sample conditions and thus may be compared between the different denatured states (Figure 8). Of the 24 equivalent sequence-pairs in barnase, 18 show an increase in degeneracy (reflecting a greater population of ''random coil'' conformations) on going from pH-denatured barnase to the urea-denatured state and similarly, 21 residue pairs move to more degenerate positions on going from pH 2.0 and 30°C to 58°C. It is not clear which residue of the pair is ''unfolding'' under increasing temperature or under urea but a general trend is clear that the pH-denatured state must contain greater populations of structured states when compared with urea or pH/temperature-denatured barnase.

Random coil coupling constants for residues of the same type

^{3}J NH-$C^{\alpha}H$ coupling constants have been determined for 85 residues in pH-denatured barnase and for 73 residues in urea-denatured barnase and their comparison is summarised in Figure 9 according to

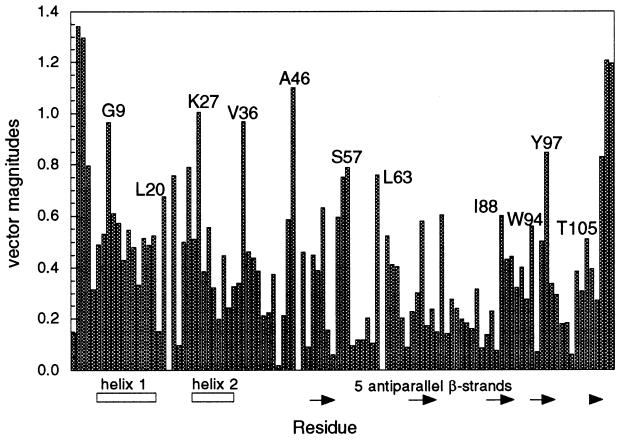

Figure 7. Vector magnitudes for the normalised vector ([^{1}H, ^{15}N], 30°C; [^{1}H, ^{15}N], 58°C) for pH-denatured barnase. Elements of secondary structure for the native protein are given along the (Residue) axis. Selected residues are labelled with the one-letter code and the vector magnitudes are in arbitrary units.

both sequence (Figure 9B) and residue type (Figure 9A). It is initially clear that residues of the same type have closely coincident coupling constant values and that these values vary significantly between differing residue types. For example, it is apparent that alanine has a random coil ^3J value of 5.7 Hz in comparison with valine, 8.1 Hz, and threonine, 7.6 Hz. Residues thought to have ^3J NH-C$^\alpha$H values significantly lower than these random coil values for pH-denatured barnase are confirmed as deviations for Tyr13, Leu14, Ser28, Ala30 and Val36 are absent in urea. In terms of the sequence, differences greater than 0.5 Hz are limited to the first and last thirds of the protein (the

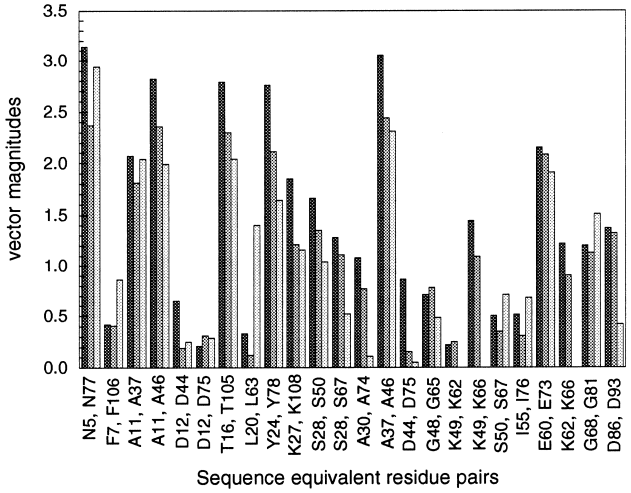

Figure 8. A comparison of the normalised vector magnitudes for the vector ([^{1}H, ^{15}N], X_1; [^{1}H, ^{15}N], X_2) for pH (left), pH/temperature (centre) and urea-denatured (right) barnase, where X_1 and X_2 are the second members of equivalent sequence-pairs. For example X_1 = A11 and X_2 = A46 as these two residues are both preceded by valine residues. The second members of all equivalent sequence-pairs are given along the *x*-axis.

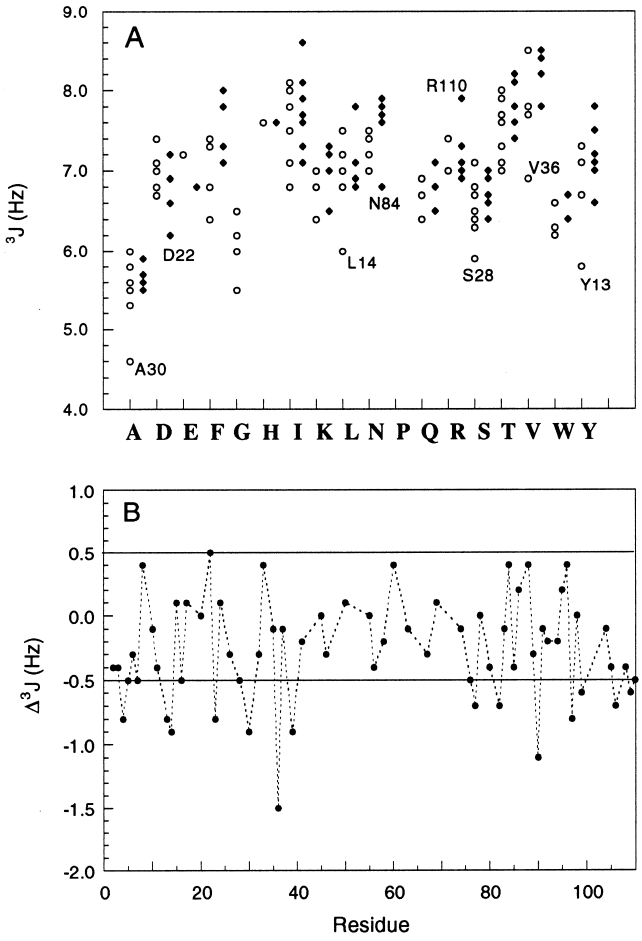

Figure 9. ^3J NH-C$^\alpha$H coupling constants for pH and urea-denatured barnase. A, The ^3J coupling constants arranged according to residue type. The open circles are those for pH-denatured barnase, the filled diamonds are those for the urea-denatured protein. A number of coupling constants are degenerate and, for example, the six alanine values appear as four values for urea-denatured barnase as there are three residues with a ^3J coupling constant of 5.9 Hz. Selected residues are shown as their one-letter code. B, The difference along the sequence between coupling constants for pH and urea-denatured barnase.

regions of the first and second helices and central strands of the β-sheet). Note that the correlation peaks in the HSQC spectrum (Figure 1) are sufficiently sharp that the ^3J NH-C$^\alpha$H coupling can be seen in the proton dimension.

NOESY data

NOESY spectra provide definitive evidence for the close proximity of protons and the presence of non-random conformations in denatured states (Neri *et al*., 1992; Logan *et al*., 1994). For both pH and urea-denatured barnase the NOESY spectra are dominated by random coil intra-residue contacts as well as sequential H$^\alpha$, NH (i, i + 1) contacts and no medium range (i, i + 2) or (i, i + 3) contacts are seen that can be unambiguously assigned. How-

ever, the presence of structured conformations in pH-denatured barnase is supported by sequential NH-NH ($i, i + 1$) contacts for Asp8 to Tyr17, Thr26 to Ser38, Tyr78 to Phe82 and Tyr90 to Tyr97 as well as aromatic to side-chain contacts for Tyr17, Trp35 and Trp94. These aromatic to side-chain contacts are also seen for urea-denatured barnase but the number of sequential NH-NH ($i, i + 1$) contacts is reduced to just five, namely Asp8–Gly9, Val10–Ala11, Gln15–Thr16, Val45–Ala146 and Ile96–Tyr97, suggesting that residual structure may persist under these conditions in the region of the first helix but that structure in the regions of the second helix is absent in urea and in the region of the β-sheet is greatly reduced.

Amide exchange

The protection of amide protons against exchange with the bulk solvent with reference to calculated intrinsic exchange rates (Bai *et al*., 1993) from model compounds provides clear evidence that these amides are involved in hydrogen-bonded or buried conformations. The measurement of amide exchange rates for denatured proteins and subsequent calculation of protection factors is routine. The observation of very sharp lines seen for the denatured NMR spectra of barnase (Figure 1) suggests that any structured conformations are in fast exchange with random coil conformations. If a particular amide proton is involved in a structured conformation that affords significant protection but this conformation is in rapid exchange with unstructured conformations, then the amide exchange rate will be dominated by exchange from the unstructured state. For example, if the structured conformation has a population of 75% in solution, then the upper limit for its protection factor is just 4, and small protection factors would thus represent significant populations of structured states. In an effort to determine the reliability of the small protection factors seen for pH-denatured barnase, we acquired a set of amide exchange rates at 25°C and pH 2.0, and a complementary set at pH 2.0, 25°C and 1 M urea. If the presence of 1 M urea was to cause significant redistribution of the populations of structured and unstructured states in solution, then a change in protection factors reflecting the change in populations would be expected.

pH-denatured barnase shows no protection factors greater than 8 (Figure 10A) and the corresponding protection factors in 1 M urea are all less than 4 (Figure 10B). However, the protection factors are closely correlated (Figure 10C) and this clearly reveals the limits in the accuracy of the determined protection factors. The reduced values in 1 M urea are an artifact of the calculation insofar as a constant ratio of the two protection factors across the sequence implies that there are no relative differences between the two states. It is possible that 1 M urea does not have a sufficiently large effect on the population distribution in solution to alter

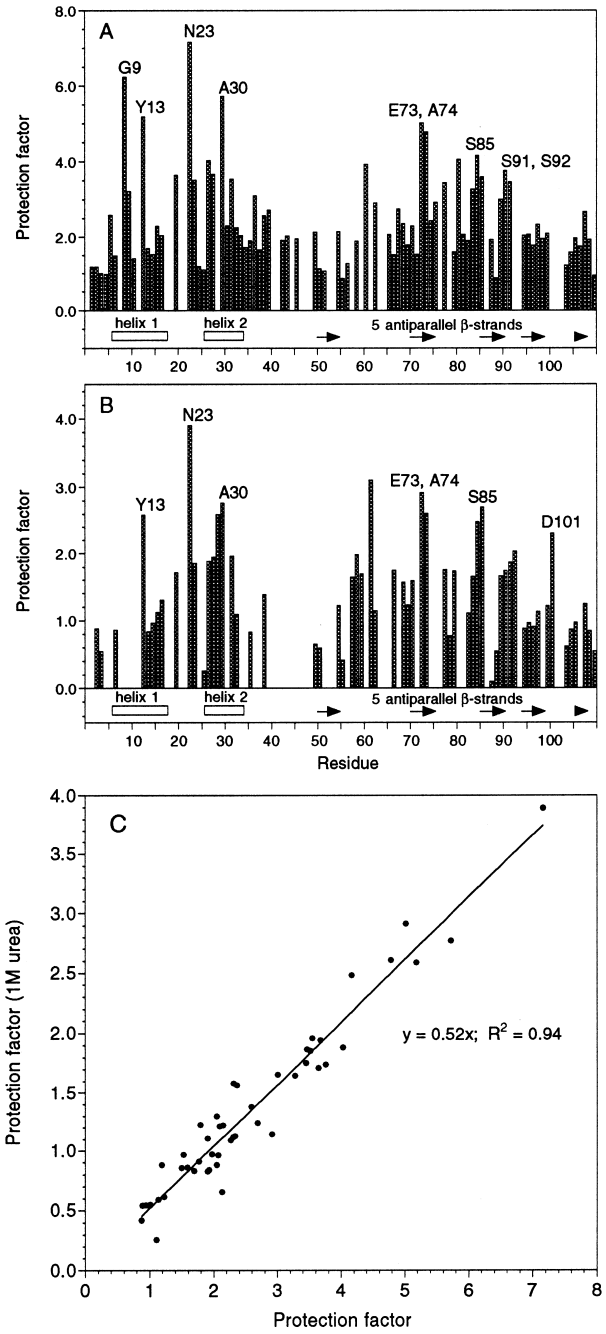

Figure 10. Protection factors along the sequence for pH-denatured barnase. A, Protection factors for barnase unfolded at pH 2.0, 25°C. B, Protection factors for barnase unfolded at pH 2.0, 25°C and 1 M urea concentration. C, Correlation between protection factors from A and B above.

the observed exchange rates although it is more likely that the range of protection factors for pH-denatured barnase are encompassed by the intrinsic error associated with the approach.

A similar range of protection factors have been observed for denatured states of hen egg-white lysozyme where residual structure is postulated from deviations in random coil chemical shifts and far UV CD spectroscopy (Buck *et al*., 1994). These

data are consistent with the presence of residual structure in denatured proteins as protection factors between 3 and 10 theoretically represent proportionately large populations of protected groups. Protection factors greater than 10 would not be expected under conditions where structured states are in fast exchange with fully unfolded conformations.

Discussion

The NMR spectra of the acid, urea, and acid/temperature-denatured states of barnase closely approximate those expected for random coil conformations. The minimal interpretation of our data is thus that the denatured states of barnase are highly unfolded. However, we have made a number of weak correlations between deviations from random coil behaviour and possible residual structure. These correlations are consistent with the presence of a small amount of residual structure against a large background of unfolded conformations. We have also attempted to derive significant comparisons between the different denatured states which are independent of the widely different solvent conditions.

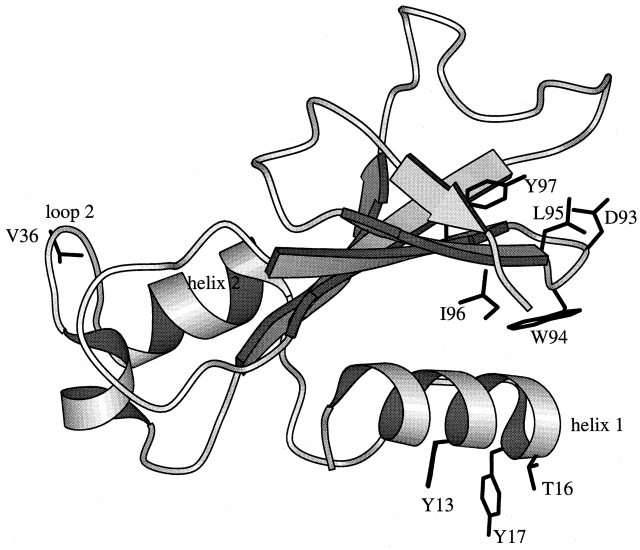

Figure 11. A view of native barnase showing loop regions and regions of secondary structure. The three regions implicated in residual structure formation in the pH-denatured state are highlighted. The identification of these regions in the native protein does not imply that native structure is present in the denatured states (see Discussion).

Assignment

We used a combination of magnetisation transfer and triple resonance NMR methods for the assignment of both urea and pH-denatured barnase. Suitable conditions for the magnetisation transfer experiment could not be found for urea-denatured wild-type barnase because the rate constants for folding and unfolding are in a range that is incompatible with the experiment. However, we showed that the use of a conservative mutation to bring the exchange rates into a suitable range provides an easily accessible alternative. The acid/temperature-denatured state was assigned by following the incremental changes in chemical shift with increasing temperature (from 30 to 58°C, in 3 deg. steps).

Characterisation

The poor dispersion of chemical shifts that complicates the NMR assignment of denatured states also complicates their characterisation. Whereas distance constraints derived from NOESY spectra are central in determining the structure of native proteins, NOESY spectra for denatured proteins are difficult to assign because of spectral overlap, and subsequently difficult to interpret as they represent time averaged data from an ensemble of rapidly interconverting conformations (Wüthrich, 1994). The strategy used here in analysing the NOESY, chemical shift, coupling constant and amide exchange data has been to assess deviations from ideal random coil behaviour. These deviations are expected to be small given the assumed dominance of random coil conformations in solution and thus, the significance of deviations must also be assessed. For example, the observed deviations of amide exchange rates for pH-denatured barnase (i.e. marginally high protection factors) all lie within the error associated with the experiment (Figure 10). Conversely, small changes of chemical shifts in the ^{1}H-^{15}N plane (Figure 8) can be reliably assessed between denatured states using equivalent sequence-pairs.

From this approach, it is clear that both pH/temperature and urea-denatured barnase are more ''unfolded'' than the pH-denatured protein. A comparison of the deviations from random coil values between urea and pH-denatured barnase suggests that the urea-denatured state is unfolded to a greater extent in the regions of the first and second helices and the central strands of β-sheet (Figures 5 and 11). This trend is mirrored by the pH/temperature-denatured state where the greatest changes in the ^{1}H-^{15}N plane on increasing temperature are also broadly in these regions (Figure 7). An analysis of the degree of degeneracy in the ^{1}H-^{15}N plane for the three denatured states between equivalent sequence-pairs also implies an unfolding on going from pH to urea or pH/temperature-denatured barnase.

The minor dispersion of ^{3}J NH-C$^{\alpha}$H coupling constants seen for pH-denatured barnase is lost on urea denaturation (Figure 9). The residues implicated are Tyr13, Leu14, Ser28, Ala30 and Val36, all contained in the first and second helices in the native protein with the exception of Val36, which is two residues away from the C terminus of the second helix. The small coupling constants seen for

these residues are consistent with populations of helical conformations. Non-random coil NOESY contacts are also more prevalent in these regions for pH-denatured barnase compared with the urea-denatured state.

Coupling constants in random coils

An observation of some importance is that ^{3}J NH-C$^{\alpha}$H coupling constants for different residue types cluster around different values in fully denatured states (Figure 9A). If the key determinant for the small differences between these coupling constants is differences in the dihedral angle, ϕ, and not, say, the differences in electronegativity of the side-chain groups, then this would strongly support differing ϕ propensities for each amino acid type in the denatured state. This has recently been suggested by Swindells *et al*. (1995).

Implications for folding pathways

The minimal interpretation of our data is that if there are initiation sites for protein folding in the denatured states of barnase, then they are only poorly occupied. We can construct a more detailed analysis by combining the NMR data with the extensive kinetic studies that have been performed previously. From the NMR analyses, there is weak evidence for the persistence of some residual structure in the urea-denatured state from sequential NH-NH (i, i + 1) NOE contacts and aromatic to side-chain contacts. These are in the centre of the first helix (Asp8 to Thr16) and the tight turn between the central strands of β-sheet (Trp94 to Tyr97) (Figure 11). Barnase folds *via* the rapid formation of an intermediate followed by the rate determining (major) transition state. An analysis by protein engineering and kinetics shows that the C-terminal region of the first helix and the centre of the β-sheet are well formed in the major transition state for folding and are only slightly weaker in the preceding folding intermediate (Fersht, 1993). The second helix forms very late (Matthews & Fersht, 1995). Thus, the residual structure in the first helix and the β-sheet in the denatured state may represent initiation points for the folding of barnase. On proceeding from a purely solvent-denatured state (5.5 M urea) to one which more closely resembles conditions favourable for folding (low pH, i.e. ~50 mM HCl) there occurs a local collapse of structure in these regions as well as the development of structure in a third region (the second helix) and the structured conformations are populated to a greater extent. It is not certain whether these structured regions are native-like, although the $^{13}C^{\alpha}$ deviations from random coil chemical shifts, NOESY contacts and coupling constant data for pH-denatured barnase are consistent with native-like structure for the first and second helices and non-native like hydrophobic clustering around Trp94, Leu95 and Ile96 in the central region of the β-sheet.

These observations may be analysed according to the nucleation-condensation mechanism that is found to describe the folding of the chymotrypsin inhibitor 2 (CI2) (Itzhaki *et al*., 1995) or the nucleation growth-mechanism observed in lattice simulations (Abkevich *et al*., 1994). In both cases, nucleation sites are only weakly occupied in denatured states but develop in the transition state for folding. This has been shown to be optimal for fast folding (Fersht, 1995). The criterion for an embryonic initiation site composed of adjacent residues in a sequence to develop into a nucleus in the transition state is that the nucleus makes more extensive interactions with residues further removed in sequence. Thus, the first helix and the centre of the β-sheet develop into nuclei because they do make those long-range interactions, which have been identified in the protein engineering analysis of the folding intermediate (Matouschek *et al*., 1992; Fersht, 1993). But the second helix does not (Matthews & Fersht, 1995). The NMR analysis shows what regions in the denatured protein have a tendency to form structure. The protein engineering method then directly pinpoints which of these regions of structure are central in searching for crucial tertiary interactions to reach the major transition state and facilitate folding to the native state.

An important feature of the NMR studies is that they identify which regions of structure do have a tendency to form regular structure in the absence of tertiary interactions. This is relevant to the Levinthal (1968) paradox, which emphasises the large number of conformations that would have to be explored by a denatured polypeptide chain during folding in an unbiased search. A tendency to form native-like structure is important in reducing the conformational space. But, too large a tendency is unfavourable, since the denatured state and folding intermediates can become too stable and so slow down rates of folding and lower stability (Fersht, 1995). It is likely, therefore, that NMR analyses of denatured states will frequently fail to detect local conformational preferences because they have evolved to be weak. Even where there is clear evidence for a hydrophobic cluster in the 434-repressor in the presence of denaturant, it is in equilibrium with the random coil conformation (Neri *et al*., 1992).

Materials and Methods

Sample preparation

Wild-type barnase was purified from cultures of *Escherichia coli* (BL21) containing the plasmid pMT410 (Paddon & Hartley, 1987) as described previously (Mossakowska *et al*., 1989; Jones *et al*., 1993). Separate protein samples were produced of uniformly ^{15}N-labelled and uniformly ^{15}N, ^{13}C-labelled. Uniformly ^{15}N-labelled mutant protein, Ile51 → Val51 (I51V) was prepared in

exactly the same manner as above except that elution from the trisacryl resin required 7 M urea. The eluant was exhaustively dialysed against water as above. Yields of I51V were approximately 6.5 mg l^{-1}.

NMR samples for wild-type barnase and the mutant protein, I51V, were prepared by dissolving lyophilised protein in 0.5 ml of, respectively, a 5.5 M urea solution containing 10% ^2H_2O and a 2.5 M urea solution containing 10% ^2H_2O. Small aliquots of a 0.25 M solution of HCl were used to adjust the measured pH to 4.5($\pm$0.1). The concentrations for samples of uniformly ^{15}N-labelled barnase were 2.9 mM, for ^{15}N/^{13}C-labelled barnase, 2.5 mM, and for the mutant protein I51V, 2.2 mM.

NMR spectroscopy

All NMR spectra were recorded on a Bruker AMX500 spectrometer fitted with a fourth channel and a triple resonance probe. Data were processed on a Silicon Graphics Indigo^2 workstation using the Felix software package (Biosym Technologies) or on Bruker X32 workstations using UXNMR. ^{1}H chemical shifts are referenced to internal sodium 3-(trimethylsilyl)-d^4-propionate, ^{15}N chemical shifts are referenced to external [^{15}N]ammonium chloride (24.93 ppm) and ^{13}C chemical shifts are indirectly referenced to the solvent ^{1}H frequency (Bax & Subramanian, 1986).

Heteronuclear ^{15}N-^{1}H NMR spectroscopy

2D ^{15}N-^{1}H TOCSY-HMQC spectra (Marion *et al*., 1989a) were recorded at 0 M urea for native I51V and at 5.5 M urea for denatured wild-type barnase with 4096 complex data points in t_2 and 256 t_1 increments. Spectral widths were 8064 Hz in t_2 and 2000.0 Hz in t_1 for native I51V or 1300 Hz in t_1 for urea-denatured wild-type barnase. 128 scans were recorded for each value of t_1 and mixing times of 55 or 30 ms were used. The experimental data were zero filled to give 4096×1024 matrices and processed using gaussian multiplication in t_2 and a shifted ($\pi/5$) sine bell function in t_1 prior to Fourier transformation.

3D ^{15}N-^{1}H TOCSY-HMQC spectra (Marion *et al*., 1989a) were recorded at 0 M urea for native I51V and at 5.5 M urea for denatured wild-type barnase with 512 complex data points in t_3, 64 and 256 increments to t_{max} values of 20 ms and 19.7 ms in t_2 and t_1, respectively. Spectral widths were 4000 Hz in t_3 (the offset was switched to the centre of the amide region immediately prior to acquisition and following on-resonance presaturation of the water signal), 1600 Hz in t_2 and 6500 Hz in t_1. The TOCSY mixing times were 55 ms and 8 scans per t_1/t_2 value were collected. Linear prediction was used to extend the time domain data to 82 points in the t_2 dimension. Sine bell apodisation was used in all three dimensions prior to Fourier transformation. The final 3D matrices contained $512 \times 128 \times 512$ real data points.

3D ^{15}N-^{1}H NOESY-HMQC spectra (Marion *et al*., 1989a) were acquired for native I51V and for urea denatured-barnase in the same manner as above. NOESY mixing times of 150 ms for folded I51V and 300 ms for urea-denatured barnase were used.

A 3D ^{15}N-^{1}H HMQC-NOESY-HMQC (Frenkiel *et al*., 1990) spectrum was acquired for urea-denatured barnase with 1024 complex data points in t_3 and 64 increments to a t_{max} value of 15.8 ms in both t_2 and t_1 dimensions. Spectral widths were 2200 Hz in t_3 (offset switching as above), 2028 Hz in both t_2 and t_1. The NOESY mixing time was 300 ms and 32 scans per t_1/t_2 value were collected.

2D ^{15}N-^{1}H exchange-relayed-HSQC spectra (Wider *et al*., 1991) were acquired at 2.4 M urea for I51V and at 3.6 M urea for wild-type barnase with 4096 complex data points in t_2 and 256 t_1 increments. 160 scans per t_1 increment were collected and the spectral widths were 8064 Hz in t_2 and 1520 Hz in t_1. The two interleaved data sets were processed in an identical manner. The scaled spectra were then combined to give a difference spectrum (Wider *et al*., 1991).

Triple resonance spectroscopy

Triple resonance experiments were acquired with uniformly ^{15}N, ^{13}C-labelled barnase in the urea-denatured state (2.5 mM protein concentration, 5.5 M urea/10% ^2H_2O as a solvent, pH 4.5 and a temperature of 30°C). In each case, the spectral width in the ^{1}H dimension was 6024 Hz and suppression of the water signal was achieved by spin lock pulses (Messerle *et al*., 1989). The water signal was additionally suppressed during data processing by convolution of the t_3 (^{1}H) time domain data (Marion *et al*., 1989b). The spectral width in the t_1 (^{15}N) dimension was 1572 Hz and quadrature detection was achieved in t_1 (^{15}N) and t_2 (^{13}C) by the States-TPPI method. In all cases constant time evolution in t_1 (^{15}N) (Grzesiek & Bax, 1992b) was employed. Mirror image linear prediction (Zhu & Bax, 1992) was used to double the time domain data during processing. 16 scans per t_1/t_2 increment were collected in each experiment.

Data for the HNCO and HN(CA)CO experiments (Bax & Pochapsky, 1992; Grzesiek & Bax, 1992b) were acquired with time-domain matrices of $32^*(t_1, {}^{15}N) \times 32^*(t_2, {}^{13}C) \times 512^*(t_3, {}^1H)$ data points, where n^* refers to n complex data points. The spectral width in the t_2 (^{13}C) dimension was 1282 Hz.

Data for the HNCA and HN(CO)CA experiments (FarmerII *et al*., 1992; Grzesiek & Bax, 1992b) were acquired with time-domain matrices of $23^*(t_1, {}^{15}N) \times 32^*(t_2, {}^{13}C) \times 512^*(t_3, {}^1H)$ data points and the spectral width in the t_2 (^{13}C) dimension was 3100 Hz. Data for the CBCA(CO)NH and CBCANH experiments (Grzesiek & Bax, 1992a, 1993) were acquired with time-domain matrices of $16^*(t_1, {}^{15}N) \times 58^*(t_2, {}^{13}C) \times 512^*(t_3, {}^1H)$ data points and here the t_2 (^{13}C) spectral width was 8928 Hz.

For all spectra, linear prediction was used to extend the time domain data in the t_2 (^{13}C) dimension increasing the number of complex points by one third. Mirror image linear prediction in t_1 (^{15}N) was used to double the number of complex points in this dimension. This was followed by sine bell ($\pi/3$) apodisation in all three dimensions and Fourier transformation. The 3D matrices contained $2048 \times 128 \times 256$ real data points for HNCO and HN(CA)CO spectra, $2048 \times 256 \times 128$ data points for CBCA(CO)NH and CBCANH spectra, and $2048 \times 128 \times 128$ data points for HNCA and HN(CO)CA spectra. In all cases the limited chemical shift dispersion in the ^{1}H dimension enabled seven eighths of the 2048 points to be discarded and the final 3D matrices contained 256 data points in this dimension.

^3J coupling constants

^3J HN-C$^\alpha$H coupling constants were measured for pH and urea denatured barnase from a 2D ^{15}N-^{1}H HSQC spectrum as described by Stonehouse & Keeler (1995).

Temperature denaturation

A series of twelve 2D ^{15}N-^{1}H HSQC spectra were acquired for pH-denatured barnase at temperatures from 28°C up to 58°C. The assignments for pH-denatured barnase at 30°C were followed incrementally to 58°C. This yielded a complete ^{15}N-^{1}H amide assignment for a temperature-denatured state at low pH. The temperature coefficient for Gln2 was determined in both dimensions (–6.28 ppb K^{-1} in ^{1}H and –10.26 ppb K^{-1} in ^{15}N) and then all spectra were re-referenced with Gln2 at δ = 8.720 and δ = 121.77 ppm in ^{1}H and ^{15}N dimensions respectively. Thus, changes in chemical shift reported here are changes relative to Gln2.

Amide exchange

Amide proton exchange rates for the pH-denatured form of barnase were measured at pH 2.0 and 25°C by dissolving lyophilised protein in $^{2}H_2O$ containing 0.03 M ^{2}HCl and following ^{1}H/^{2}H exchange with a series of ^{15}N-^{1}H HSQC spectra. ^{15}N-^{1}H HSQC spectra were acquired with 1024 complex data points in t_2 and 128 t_1 increments. The pulse sequence was adapted to allow the collection of just two scans per increment (Marion *et al*., 1989c) and this gave a total experiment time of five minutes 42 seconds. A total of 28 spectra were acquired sequentially with the final time point collected at 168.5 minutes. A second set of amide proton exchange rates were measured at pH 2.0, 25°C and 1 M urea. Data for each amide were fitted to a single exponential to determine the rate constant for exchange, k_{ex}, and protection factors were calculated from the ratio k_{rc}/k_{ex}, where k_{rc} is the calculated intrinsic rate constant (Bai *et al*., 1993).

Acknowledgements

V.L.A. is supported by a Prince of Wales Scholarship from the Cambridge Commonwealth Trust. S.V. was supported by a grant from the Swiss National Research Foundation (grant 0823A-031993).

References

Abkevich, V, I., Gutin, A. M. & Shakhnovich, E. I. (1994), Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry*, **33**, 10026–10036.

Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M. & Fersht, A. R. (1994). Toward solving the folding pathway of barnase: the complete backbone ^{13}C, ^{15}N and ^{1}H NMR assignments of its pH-denatured state. *Proc. Natl Acad. Sci. USA*, **91**, 9412–9416.

Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 75–86.

Bax, A. (1994). Multidimensional nuclear magnetic resonance methods for protein studies. *Curr. Opin. Struct. Biol.* **4**, 738–744.

Bax, A. & Pochapsky, S. (1992). Optimized recording of heteronuclear multidimensional NMR spectra using pulsed field gradients. *J. Magn. Reson.* **99**, 638–643.

Bax, A. & Subramanian, S. (1986). Sensitivity-enhanced two-dimensional heteronuclear shift correlation NMR spectroscopy. *J. Magn. Reson.* **67**, 565–569.

Braun, D., Wider, G. & Wüthrich, K. (1994). Sequence corrected ^{15}N ''random coil'' chemical shifts. *J. Am. Chem. Soc.* **116**, 8466–8469.

Buck, M., Radford, S. E, & Dobson, C. M. (1994). Amide hydrogen exchange in a highly denatured state: hen egg-white lysozyme in urea. *J. Mol. Biol.* **237**, 247–254.

Bycroft, M., Matouschek, A., Kellis, J. T. J., Serrano, L. & Fersht, A. R. (1990a). Detection and characterisation of a folding intermediate by NMR. *Nature*, **346**, 488–490.

Bycroft, M., Sheppard, R. N., Lau, F. T.-K. & Fersht, A. R. (1990b). Sequential assignment of the ^{1}H nuclear magnetic resonance spectrum of barnase. *Biochemistry*, **29**, 7425–7432.

Caflisch, A. & Karplus, M. (1994). Molecular dynamics simulation of protein denaturation: solvation of the hydrophobic cores and secondary structure of barnase. *Proc. Natl Acad. Sci. USA*, **91**, 1746–1750.

De Dios, A. C., Pearson, J. G. & Oldfield, E. (1993). Secondary and tertiary structural effects on protein NMR chemical-shifts: an *ab initio* approach. *Science*, **260**, 1491–1496.

Dobson, C. M. & Evans, P. A. (1984). Protein folding kinetics from magnetisation transfer nuclear magnetic resonance. *Biochemistry*, **23**, 4267–4270.

Ernst, R. R., Bodenhausen, G. & Wokaun, A. (1987). *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

FarmerII, B. T., Venters, R. A., Spicer, L. D., Wittekind, M. G. & Müller, L. (1992). A refocused and optimized HNCA: increased sensitivity and resolution in large macromolecules. *J. Biomol. NMR*, **2**, 195–202.

Fersht, A, R. (1993). Protein folding and stability: the pathway of folding of barnase. *FEBS Letters*, **325**, 5–16.

Fersht, A. R. (1995). Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl Acad. Sci. USA*, in the press.

Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B. & Feeney, J. (1990). HMQC-NOESY-HMQC, a three-dimensional NMR experiment which allows detection of nuclear Overhauser effects between protons with overlapping signals. *J. Magn. Reson.* **90**, 420–425.

Grzesiek, S. & Bax, A. (1992a). An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Reson.* **99**, 201–207.

Grzesiek, S. & Bax, A. (1992b). Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432–440.

Grzesiek, S. & Bax, A. (1993). Amino-acid type determination in the sequential assignment procedure of uniformly C-13/N-15-enriched proteins. *J. Biomol. NMR*, **3**, 185–204.

Itzhaki, L., Otzen, D. & Fersht, A. R. (1995). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**, 260–288.

Jones, D. N. M., Bycroft, M., Lubienski, M. J. & Fersht, A. R. (1993). Identification of the barstar binding site of barnase by NMR spectroscopy and hydrogen-deuterium exchange. *FEBS Letters*, **331**, 165–172.

Kuszewski, J., Clore, G. M. & Gronenborn, A. M. (1994). Fast folding of a prototypic polypeptide. The

immunoglobulin binding domain of streptococcal protein-G. *Protein Sci.* **3**, 1945–1952.

Levinthal, C. (1968). Are there pathways for protein folding? *J. Chem. Phys.* **85**, 44–45.

Li, A. J. & Daggett, V. (1994). Characterization of the transition-state of protein unfolding by use of molecular dynamics – Chymotrypsin inhibitor-2. *Proc. Natl Acad. Sci. USA*, **91**, 10430–10434.

Logan, T. M., Olejniczak, E. T., Xu, R. X. & Fesik, S. W. (1993). A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. *J. Biomol. NMR*, **3**, 225–231.

Logan, T. M., Thériault, Y. & Fesik, S. W. (1994). Structural characterisation of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* **236**, 637–648.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M. & Clore, G. M. (1989a). Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β. *Biochemistry*, **28**, 6150–6156.

Marion, D., Ikura, M. & Bax, A. (1989b). Improved solvent supression in one and two-dimensional NMR spectra by convolution of time domain data. *J. Magn. Reson.* **84**, 425–430.

Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989c). Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393–399.

Matouschek, A., Kellis, J. T. J., Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterized by protein folding. *Nature*, **346**, 440–445.

Matouschek, A., Serrano, L. & Fersht, A. R. (1992). The folding of an enzyme IV: Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure. *J. Mol. Biol.* **224**, 819–835.

Matthews, J. M. & Fersht, A. R. (1995). Exploring the energy surface of protein folding by structure-reactivity relationships and engineered proteins: observation of Hammond behaviour for the gross structure of the transition state and anti-Hammond behaviour for structural elements for unfolding / folding of barnase. *Biochemistry*, **34**, 6805–6814.

Messerle, B. A., 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.

Mossakowska, D. E., Nyberg, K. & Fersht, A. R. (1989). Kinetic characterization of the recombinant ribonuclease from *Bacillus amyloliquefaciens* (barnase) and investigation of key residues in catalysis by site-directed mutagenesis. *Biochemistry*, **28**, 3843–3850.

Moult, J. & Unger, R. (1991). An analysis of protein folding pathways. *Biochemistry*, **30**, 3816–3824.

Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*, **257**, 1559–1563.

Osapay, K., Theriault, Y., Wright, P. E. & Case, D. A. (1994). Solution structure of carbonmonoxy-myoglobin determined from NMR distance and chemical-shift constraints. *J. Mol. Biol.* **244**, 183–197.

Otzen, D. E., Itzhaki, L. S., elMasry, N. F., Jackson, S. E. & Fersht, A. R. (1994). Structure of the transition state for the folding / unfolding of barley chymotrypsin inhibitor 2 and its implications for mechanisms of folding. *Proc. Natl Acad. Sci. USA*, **91**, 10422–10425.

Paddon, C. J. & Hartley, R. W. (1987). Expression of *Bacillus amyloliquefaciens* extracellular ribonuclease (barnase) in *Escherichia coli* following an inactivating mutation. *Gene*, **53**, 11–19.

Prat Gay, G. D., Ruiz-Sanz, J., Davis, B. & Fersht, A. R. (1994). The structure of the transition state for the association of 2 fragments of the barley chymotrypsin inhibitor 2 to generate native like protein: implications for mechanisms of protein folding. *Proc. Natl Acad. Sci. USA*, **91**, 10943–10946.

Roder, H., Elove, G. A. & Englander, S. W. (1988). Structural characterisation of folding intermediates in cytochrome *c* by H-exchange labelling and proton NMR. *Nature*, **335**, 700–704.

Serrano, L., Kellis, J. T., Cann, P., Matouschek, A. & Fersht, A. R. (1992a). The folding of an enzyme II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**, 783–804.

Serrano, L., Matouschek, A. & Fersht, A. R. (1992b). The folding of an enzyme III. Structure of the transition state for unfolding of barnase analysed by protein engineering procedure. *J. Mol. Biol.* **224**, 805–818.

Spera, S. & Bax, A. (1991). Empirical correlation between protein backbone conformation and C^{α} and $C^{\beta 13}C$ nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492.

Stonehouse, J. & Keeler, J. (1995). A convienient and accurate method for the measurement of the values of spin-spin coupling constants. *J. Magn. Reson. ser. A*, **112**, 43–57.

Swindells, M. B., MacArthur, M. W. & Thornton, J. M. (1995). Intrinsic ϕ, ψ propensities of amino acids, derived from the coil regions of known structures. *Nature Struct. Biol.* **2**, 596–603.

Udgaonkar, J. B. & Baldwin, R. L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*, **335**, 694–699.

Wider, G., Neri, D. & Wüthrich, K. (1991). Studies of slow conformational equilibria in macromolecules by exchange of heteronuclear longitudinal 2-spin-order in a 2D difference correlation experiment. *J. Biomol. NMR*, **1**, 93–98.

Williamson, M. P., Kikuchi, J. & Asakura, T. (1995). Application of ^{1}H-NMR chemical shifts to measure the quality of protein structures. *J. Mol. Biol.* **247**, 541–546.

Wishart, D. S., Sykes, B, D. & Richards, F. M. (1991). Relationship between nuclear-magnetic-resonance chemical-shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–333.

Wishart, D. S., Bigham, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. (1995). ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbour effects. *J. Biomol. NMR*, **5**, 67–81.

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

Wüthrich, K. (1994). NMR assignments as a basis for structural characterization of denatured states of globular proteins. *Curr. Opin. Struct. Biol.* **4**, 93–99.

Zhu, G. & Bax, A. (1992). Improved linear prediction of damped NMR signals using modified forward backward linear prediction. *J. Magn. Reson.* **100**, 202–207.

Edited by J. Karn

(Received 16 August 1995; accepted 15 September 1995)