Protein Science (1999), *8*:1320–1331. Cambridge University Press. Printed in the USA.

Autonomous folding of a peptide corresponding to the N-terminal β-hairpin from ubiquitin

ROSA ZERELLA,[1] PHILIP A. EVANS,[2] JOHN M.C. IONIDES,[3] LEN C. PACKMAN,[2]
B. WESLEY TROTTER,[1,4] JOEL P. MACKAY,[1,5] AND DUDLEY H. WILLIAMS[1]

[1]Cambridge Centre for Molecular Recognition and University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom
[2]Cambridge Centre for Molecular Recognition and Department of Biochemistry, Tennis Court Road, Cambridge CB2 1QW, United Kingdom
[3]Cambridge Centre for Molecular Recognition and MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

(RECEIVED November 16, 1998; ACCEPTED February 14, 1999)

Abstract

The N-terminal 17 residues of ubiquitin have been shown by ^{1}H NMR to fold autonomously into a β-hairpin structure in aqueous solution. This structure has a specific, native-like register, though side-chain contacts differ in detail from those observed in the intact protein. An autonomously folding hairpin has previously been identified in the case of streptococcal protein G, which is structurally homologous with ubiquitin, but remarkably, the two are not in topologically equivalent positions in the fold. This suggests that the organization of folding may be quite different for proteins sharing similar tertiary structures. Two smaller peptides have also been studied, corresponding to the isolated arms of the N-terminal hairpin of ubiquitin, and significant differences from simple random coil predictions observed in the spectra of these subfragments, suggestive of significant limitation of the backbone conformational space sampled, presumably as a consequence of the strongly β-structure favoring composition of the sequences. This illustrates the ability of local sequence elements to express a propensity for β-structure even in the absence of actual sheet formation. Attempts were made to estimate the population of the folded state of the hairpin, in terms of a simple two-state folding model. Using published "random coil" values to model the unfolded state, and values derived from native ubiquitin for the putative unique, folded state, it was found that the apparent population varied widely for different residues and with different NMR parameters. Use of the spectra of the subfragment peptides to provide a more realistic model of the unfolded state led to better agreement in the estimates that could be obtained from chemical shift and coupling constant measurements, while making it clear that some other approaches to population estimation could not give meaningful results, because of the tendency to populate the β-region of conformational space even in the absence of the hairpin structure.

Keywords: β-hairpin; β-sheet; chemical shifts; coupling constants; peptide conformations; ubiquitin

The cooperativity of interactions defining a protein structure is one of the most difficult aspects of the folding problem. Since the stability of every interaction depends, in principle, on cooperation with every other interaction, it is not straightforward to extract information about, for example, the extent to which the conformation of a particular residue is specified by the local sequence or, conversely, by the requirements of the wider tertiary structure. It is clear, however, that global cooperativity is not an absolute requirement for structural ordering, so that independent folding of substructures, such as isolated secondary structural elements, is often detectable.

There have been many studies of peptide fragments that can form α-helices and β-turns in solution (Dyson et al., 1988, 1992; Scholtz & Baldwin, 1992). These fragments have proved to be invaluable for systematic studies of the role that individual residues play in ensuring their stability. Their ability to fold independently means that they might play an important role in dictating the structures of other parts of the molecule through the tertiary structural constraints they place and they might also help to direct early events in folding pathways.

Reprint requests to: Philip A. Evans, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Old Addenbrooke's Site, Cambridge, CB2 1GA, England; e-mail:pe@mole.bio.cam.ac.uk.
[4]Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.
[5]Present address: University of Sydney, Department of Biochemistry, Sydney, New South Wales 2006, Australia.

Abbreviations: CD, circular dichroism; DQF-COSY, double-quantum filtered correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; RP-HPLC, reversed-phase high-performance liquid chromatography; TOCSY, total correlation spectroscopy.

Although rather less work has been done in the case of β-sheet structures, it is now apparent that some β-hairpins can fold autonomously into native-like structures, either in aqueous solution or in the presence of an organic cosolvent (Cox et al., 1993; Blanco et al., 1994a, 1994b; Searle et al., 1995, 1996). *De novo* designed hairpins have also been shown to fold in some instances (Blanco et al., 1993; de Alba et al., 1996; Ramirez-Alvarado et al., 1996; Maynard et al., 1998). The work reported in this paper concerns an N-terminal 17-residue segment of ubiquitin, which, in the native structure, forms a β-hairpin (Fig. 1).

A solution of ubiquitin at pH 2 in 3:2 methanol:water was found to exist in a partially structured "A-state" in which most of the residues 1–35 (spanning the N-terminal hairpin and the native α-helix) exhibit $C_\alpha H$ shifts whose pattern and relative magnitude approximately mirror (40–100%) those found in the native structure (Harding et al., 1991; Stockman et al., 1993). On the other hand, the C-terminal segment in the A-state is not at all native like, but is composed almost entirely of non-native helical structure (Stockman et al., 1993).

To assess the extent to which the native-like structures in the N-terminal part of the molecule might be interdependent for their stability, peptide fragments corresponding to residues 1–21 and 1–35 were examined under the same mixed solvent conditions. It was found that the chemical shifts of residues within the hairpin were essentially identical in both of these peptides to those in the intact A-state, suggesting that this is, in fact, an autonomously folding element under these conditions (Cox et al., 1993). It seemed from CD and one-dimensional NMR studies of the same peptides, however, that under purely aqueous conditions a much lower level of order existed, presumably because the alcohol component alters the solvent properties so that the driving force for the formation of intramolecular hydrogen bonds is increased. Nonetheless, the persistence of even a modest level of native-like structure in water would be of considerable interest in relation to an understanding of the folding of this protein. We have therefore investigated the behavior of the hairpin sequence U(1–17) in more detail, as reported in this paper.

Results

NMR evidence for folding of U(1–17) in aqueous solution

^{1}H NMR spectra of the peptide U(1–17) in aqueous solution at 2 °C were characterized by sharp resonances with sufficient dispersion in two-dimensional spectra to allow complete assignment using the sequential method of Wüthrich (1986). Full details of the assignments are available as supplementary material.

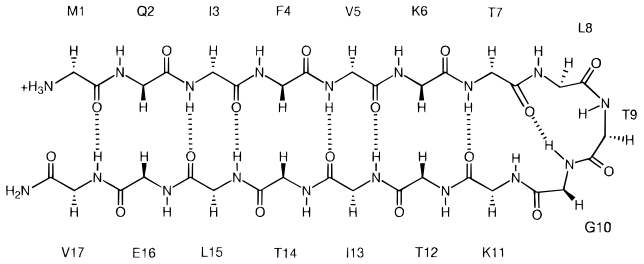

Fig. 1. Schematic drawing of the β-hairpin structure formed by the N-terminal segment of the ubiquitin in its native state. Interstrand hydrogen bonding and relative orientation of the $C_\alpha H$ protons are indicated.

The fingerprint regions of the NOESY spectra of U(1–17) in water and in 3:7 methanol:water are compared in Figure 2. It is clear that the chemical shift dispersion, both of $C_\alpha H$ and NH protons, is markedly diminished in the purely aqueous system, suggesting a lower level of secondary structure formation by the peptide under these conditions. However, closer examination of the NMR data provided a series of indications that, even in water, there is some population of a β-hairpin broadly similar to the structure formed in aqueous methanol and in native ubiquitin (Vijaykumar et al., 1987; Cox et al., 1993). The $C_\alpha H$ shifts are predominantly downfield of the values expected for an unstructured peptide, for residues 1–6 and 12–17, corresponding to the residues of the native β-strands (Fig. 3): this is typical of residues in a β-sheet (Wishart et al., 1991). Two clear exceptions are the $C_\alpha H$ protons of Ile3 and Val5, both of which are actually upfield shifted.

Table 1 shows that for residues in the sequences 2–7 and 12–17, the coupling constants $^3J_{HN\alpha}$ are mostly substantially greater than would be expected for these residues in a substantially unfolded peptide (Smith et al., 1996), which is again indicative of β-structure (Wüthrich, 1986). There is also a network of NOEs supporting a roughly native-like folded structure (Fig. 4). These include a substantial number of cross-strand NOEs, mainly involving side chains, but also including main-chain NOEs between the $C_\alpha H$ protons of Phe4 and Thr14, and between the NH protons of Thr7 and Lys11. These demonstrate that the register of the β-sheet interactions is indeed native like. Other cross-strand backbone NOEs that might have been expected on the basis of the native hairpin structure were not observable owing to severe spectral overlap. All but a few of the side-chain NOEs identified were native like to the extent that they involve residues that are notionally adjacent in an idealized β-hairpin conformation. There were no long-range NOEs suggestive of non-native interstrand interactions, indicating that there is no significant population of hairpin structures with alternative strand registers. Nonetheless, there are significant differences of detail from the NOE pattern actually observed in spectra of native ubiquitin. This is illustrated for the case of the aromatic ring of Phe4, in Figure 5. The NOEs from Phe4 to the methyl groups of Ile3 are not consistent with the native structure since these side chains are on opposite faces of the hairpin, and therefore would be expected to be too distant from one another. These effects are thus most likely to arise from unfolded or, at least, nonhairpin conformations in equilibrium with the native-like structure. More subtly, there are strong NOEs to the two Thr side chains on the opposite strand, residues 12 and 14. These are compatible with a native-like backbone folding (Fig. 1) but in the native structure the orientation of the ring is such that Thr12 is not close and the corresponding NOEs are not observed. This suggests, therefore, that although the backbone framework of the hairpin is native like in the peptide, there are significant differences of detail in the side-chain interactions supporting it.

The NMR results are qualitatively similar to those obtained previously for this peptide in aqueous methanol solution, where the formation of such a hairpin structure was also inferred (Cox et al., 1993). However, the disparity in the chemical shift perturbations observed, evident in Figure 2, suggested that the population of the structured state might be much more limited in purely aqueous solution. The feasibility of using various conformation dependent NMR parameters to quantify the population of the hairpin, assum-

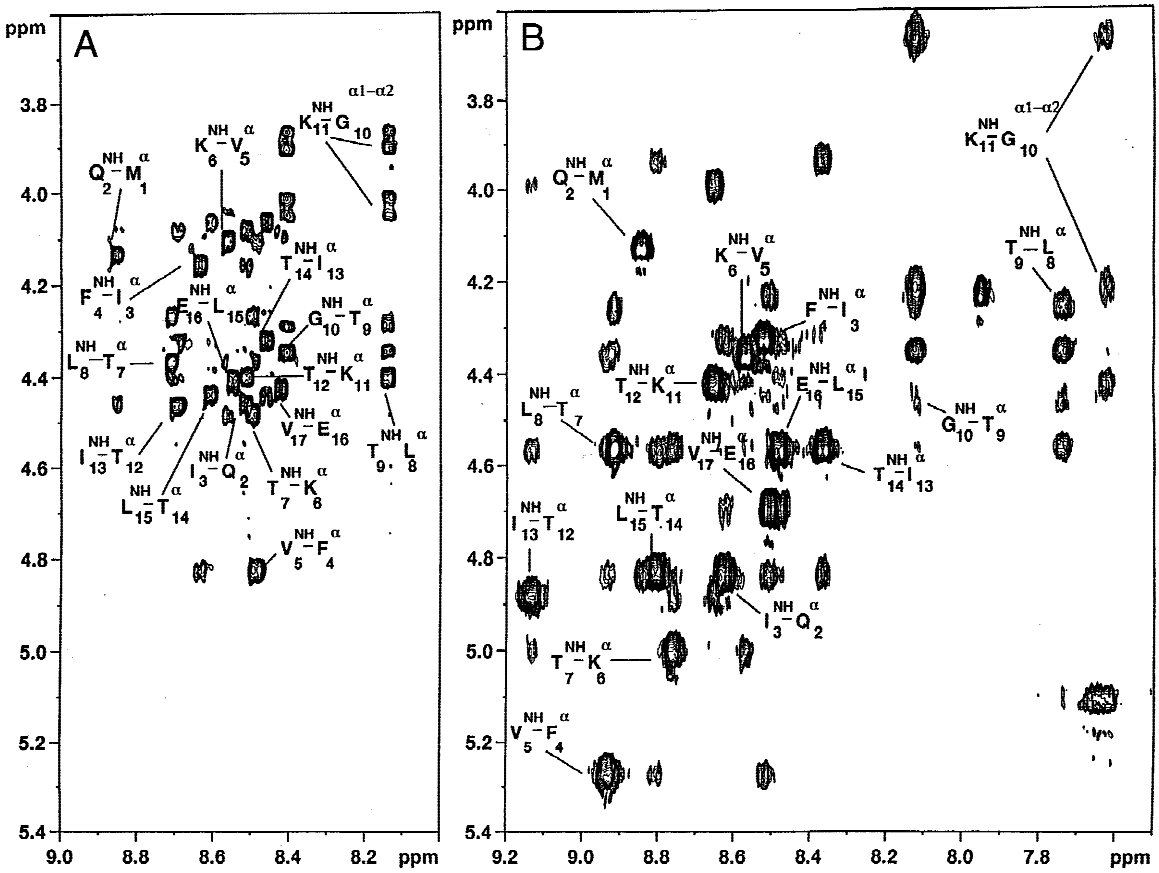

Fig. 2. A: Fingerprint region of the 200 ms NOESY spectrum of U(1–17) in H_2O:D_2O (9:1) (2 °C, 1 mM, pH 3.8). **B:** Fingerprint region of the 200 ms NOESY spectrum of U(1–17) in CD_3OD:H_2O (3:7) (2 °C, 1 mM, pH 3.8). Assignments of intraresidue cross peaks are indicated. Note that the chemical shift scale is the same in both spectra.

ing a simple two-state (folded and unfolded) model was therefore investigated.

Population of the hairpin structure

There are a number of problems to be addressed in trying to quantify the population of a β-hairpin structure. First, there is the issue of whether a simple two-state model is really applicable, or whether the system needs to be treated as a more complex dynamic ensemble. The most obvious test of this would seem to be to check the consistency of calculated populations derived from alternative experimental parameters, assuming the simple two-state model. The other problem is to find suitable limiting values for these parameters, characteristic of fully populated structured and unfolded states. For the structured state, the most straightforward approach was to assume that it was native like in detail and therefore take values from the spectrum of intact ubiquitin in its native state. For the unfolded state, there is now a substantial body of data, derived from unstructured peptides and from irregular regions of proteins, which can be used to provide estimated NMR parameters.

Figure 6 presents estimates of the apparent percentage population of β-hairpin structure for individual residues in U(1–17), cal-

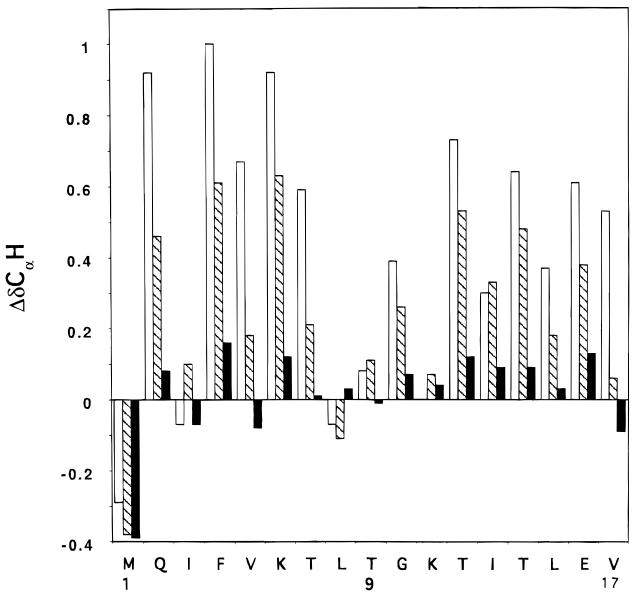

Fig. 3. Deviations of observed $C_\alpha H$ chemical shifts from random coil values ($\Delta\delta C_\alpha H$). Clear bars: Native ubiquitin; hatched bars: U(1–17) in CD_3OD:H_2O (3:7); black bars: U(1–17) in H_2O:D_2O (9:1).

Table 1. *$^3J_{HN\alpha}$ coupling constants*

	$^3J_{HN\alpha}$ (Hz)					
Residue	Random coil values[a]	U(1–7) and U(11–17)[b]	U(1–17) in H_2O/D_2O[c]	U(1–17) in CD_3OD/H_2O[d]	Native ubiquitin[c]	Database[e]
Met1	—	—	—	—	—	—
Gln2	7.1	6.7	7.0 (6.9)	7.9	8.0	9.4
Ile3	7.6	7.8	8.6	9.2	8.7	9.7
Phe4	7.5	7.5	8.1 (8.1)	8.8	9.9	8.5
Val5	7.7	8.3	8.5	9.0	9.6	9.7
Lys6	7.1	6.7	7.0	8.0	8.9	8.8
Thr7	7.6	—	7.9	8.0	8.2	7.0
Leu8	7.1	—	6.4	5.3	5.3	—
Thr9	7.6	—	7.8	—	—	—
Gly10	—	—	—	—	—	—
Lys11	7.1	6.5	7.3	—	—	—
Thr12	7.6	7.5	7.7	8.1	9.4	8.8
Ile13	7.6	7.9	8.2	9.1	9.6	9.6
Thr14	7.6	7.9	8.2	8.8	9.5	9.6
Leu15	7.1	7.1	7.4 (7.5)	8.7	9.5	9.7
Glu16	6.8	6.9	7.4	7.9	9.6	9.6
Val17	7.7	7.7	8.0	9.1	9.5	9.6

[a]Values derived from the COIL data set as listed in Smith et al. (1996).
[b]Values obtained from one-dimensional spectra. Estimated error margins are ±0.1 Hz.
[c]Values obtained using the method of Titman and Keeler (1990). Estimated error margins are ±0.1 Hz. Values in parentheses were determined from one-dimensional spectra.
[d]Estimated error margins ±0.2 Hz.
[e]Values calculated from average ϕ-values, determined from crystallographic coordinates, for 14 β-hairpins containing G1 β-bulges (Sibanda & Thornton, 1991).

culated on this basis from various NMR parameters. The $C_\alpha H$ chemical shift deviations from estimated unfolded state values (Bundi & Wüthrich, 1979) were assumed to be proportional to the population of a structured state in the spectrum of which these shift deviations would be the same as they are in the case of intact, native ubiquitin. The values obtained are generally substantially lower in water than in aqueous methanol, as expected, so that the estimated β-sheet content, averaged over the residues 2–6 and 12–16 (but omitting Ile3, which is anomalous, as discussed above), rises from 14 to 63% in the presence of the cosolvent. It is clear that in both cases, however, there is wide variation from residue to residue. The conformation-dependent $^3J_{HN\alpha}$ coupling constants were used in a similar way to provide alternative population estimates. The coupling constants observed in native ubiquitin were used as estimated values for the folded state of the hairpin and the values given by Smith et al. (1996) were used for the unfolded state. Essentially the same results could be obtained using the random coil values given by Serrano (1995). The apparent average β-sheet content for U(1–17), averaged over the β-strand residues, was found to be 31% in water and 67% in aqueous methanol. There is a clear disparity here in that the large increase in population suggested by the chemical shift changes when the cosolvent is introduced is not mirrored in the coupling constants, whose averaged values suggest little change in the population. In both cases, how-

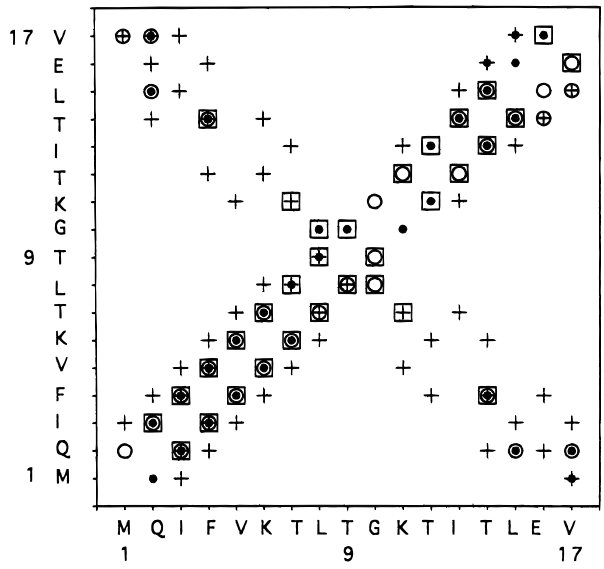

Fig. 4. Summary of NOE data for U(1–17) in H_2O:D_2O (9:1). Cross peaks of significant intensity in NOESY spectra are classified thus: □, main chain–main chain; ○, main chain–side chain; ●, side chain–main chain; +, side chain–side chain. The string of NOEs running perpendicular to the main diagonal is characteristic of antiparallel β-hairpin structure.

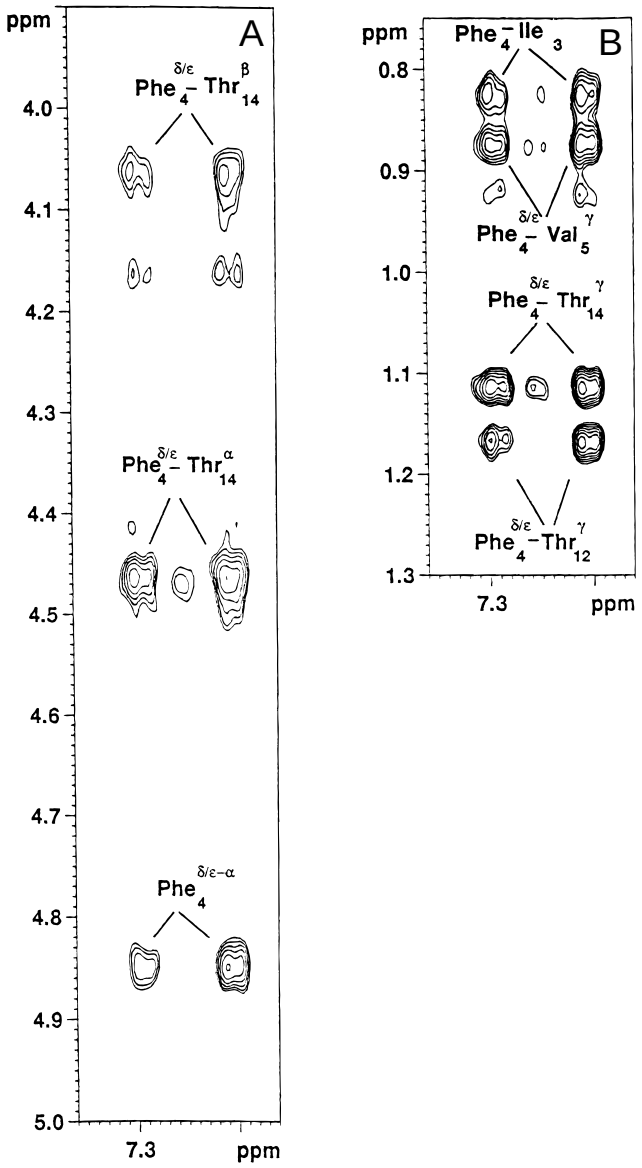

Fig. 5. Portions of the 200 ms NOESY spectrum of U(1–17) in D_2O (2 °C, 1 mM, pH 3.8), showing NOEs between the side-chain aromatic protons of Phe4 and (**A**) $C_\alpha H$ and $C_\beta H$ protons and (**B**) methyl group protons.

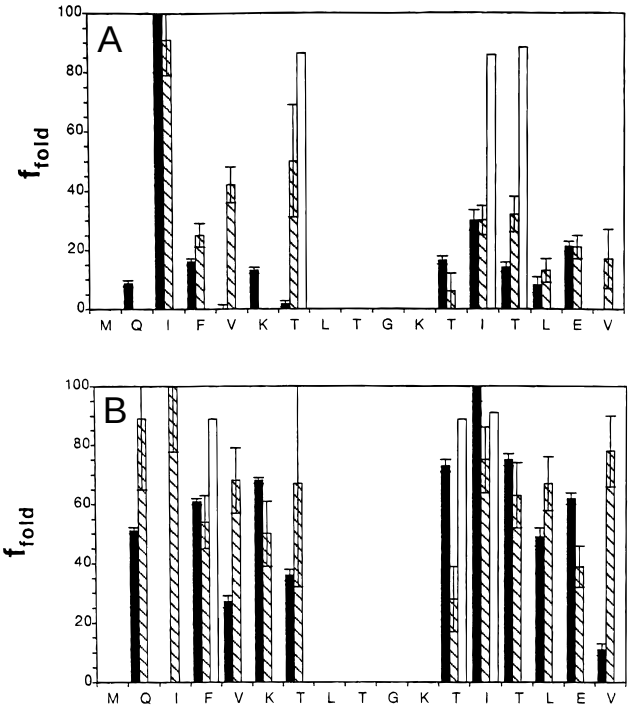

Fig. 6. Apparent population of a native like β-hairpin state of U(1–17), based on various NMR parameters. Limiting values for the structured and unfolded states were estimated based on native ubiquitin and on literature random coil values, respectively (see text). Black bars: $C_\alpha H$ chemical shift perturbations; hatched bars: $^3J_{HN\alpha}$ values; clear bars: $d_{NN}/d_{\alpha N}$ NOE intensity ratios. **A:** U(1–17) in H_2O:D_2O (9:1). **B:** U(1–17) in CD_3OD:H_2O (3:7).

ever, there is considerable variation in the population estimate from residue to residue.

The intensities of NOEs between various pairs of backbone protons also depend strongly on the local conformation and methods for estimation of the population of a particular secondary structure from the relative intensities of such effects have been devised accordingly. Values of the ratio $I_{\alpha N}(i,i)/I_{\alpha N}(i,i+1)$ were measured in the NOESY spectrum of U(1–17), for those cases where the cross peaks were sufficiently well resolved for quantitative analysis, and the results are presented in Table 2. The average value of this parameter, for the peptide residues in the β-sheet region, was approximately 0.2 in both aqueous and aqueous methanol solutions, which closely matches the value of 0.24 predicted for typical β-sheet structures (Saulitis & Liepins, 1990). This is inconsistent with a value intermediate between the α-helix and β-sheet extremes, as might reasonably have been expected for a peptide with a significantly populated unfolded state (Fiebig et al., 1996).

The sequential d_{NN} NOEs were, as expected, much weaker than the $d_{\alpha N}$ effects, for protons in the arms of the hairpin structure but a number of them could nonetheless be identified. Values of the ratio $I_{NN}(i,i+1)/I_{\alpha N}(i,i+1)$ for several protons in U(1–17) are included in Table 2. Comparison of these with the corresponding NOE intensity ratios predicted for the native state from the crystal structure, and with an estimated value for the unfolded state of 0.4 (Fiebig et al., 1996), allowed the apparent population of the β-structured state to be estimated for each residue pair. The average value obtained was 87% in aqueous solution and 89% in aqueous methanol.

The intensities of interstrand NOEs can also be used to estimate populations of a β-structure. In some ways this is a more straightforward approach since it is reasonable to assume that such a long-range interaction would show no NOE intensity in the absence of the hairpin and therefore the problem of choosing an appropriate extreme value characteristic of the unfolded state is trivial. However, NOEs involving side chains could not be used for this purpose since, as discussed above, their interactions are clearly not native like in detail. The only backbone NOE of this type, which was well enough resolved in the spectrum of U(1–17) to permit quantitative analysis, was that between the $C_\alpha H$ protons of Phe4 and Thr14 (see Fig. 7). The intensity of this NOE cross peak, normalized relative to the intraresidue $C_\alpha H$-$C_\alpha H$ NOE of Gly10,

Table 2. *NOE intensity ratios for protons in U(1–17) and U(1–7)*

	$I_{\alpha N(i,i)}/I_{\alpha N(i,i+1)}$			$I_{NN(i,i+1)}/I_{\alpha N(i,i+1)}$		
Residue	Theoretical[a]	U(1–17)[b]	U(1–7)[c]	Theoretical[a]	U(1–17)[b]	U(1–7)[c]
Met1	0.19	0.19	—	—	—	—
Gln2	—	—	0.23	—	—	—
Ile3	0.26	0.18	0.15	—	—	0.06
Phe4	0.24	(0.25)	0.15	0.04	(0.08)	0.08
Val5	0.24	0.20	0.15	—	—	0.07
Lys6	—	—	0.13	—	—	0.07
Thr7	0.20	(0.21)	—	0.03	0.09	0.06
Leu8	—	—	—	3.85	0.32 (0.52)	0.06
Thr9	—	—	—	—	—	—
Gly10	—	—	—	—	—	—
Lys11	—	—	—	—	—	—
Thr12	0.18	0.21(0.18)	—	0.02	(0.06)	—
Ile13	0.20	0.22(0.20)	—	0.02	0.07 (0.05)	—
Thr14	—	—	—	0.02	0.06	—
Leu15	—	—	—	—	—	—
Glu16	0.11	0.21	—	—	—	—
Val17	—	—	—	—	—	—

[a]Values based on interproton distances derived from crystallographic coordinates of native ubiquitin, with proton positions generated using the program QUANTA (Molecular Simulations).
[b]Data obtained from NOESY spectra (mixing time 100 ms) of U(1–17) in H_2O:D_2O (9:1) and in MeOH:H_2O (3:7) (values in parentheses).
[c]Data obtained from a NOESY spectrum (mixing time 100 ms) of U(1–7) in H_2O:D_2O (9:1).

was determined by volume integration in the peptide spectra. Several different mixing times were used to ensure that the results were not distorted by spin diffusion effects. Using 2.3 Å as an estimate of the $C_\alpha H$-$C_\alpha H$ distance in a 100% structured hairpin, the apparent fraction of folded peptide, assuming a two-state equilibrium, was calculated to be around 60%.

It is clear that there is wide variation in the estimates of the hairpin population obtained using the various methods described and, indeed, for different protons using the same method. This suggests either that the two-state model, or at least our choice of reference states, can only be at best a rather crude approximation to reality. We therefore sought alternative ways to derive estimates of limiting values of NMR parameters characteristic of the unfolded and folded states.

NMR spectra of U(1–7) and U(11–17)

The use of published "random coil" NMR parameters, derived from spectra of short, structureless peptides, as estimates of the values characteristic of the unfolded state of U(1–17) might not be appropriate if, for example, there were significant local conformational preferences even in the absence of hairpin formation. In this case, it should be possible to obtain better values by setting up a more directly relevant experimental model: since both arms of a β-hairpin must, by definition, be present for the β-structure to form, then a peptide subfragment corresponding to just a single arm should provide a good model for the corresponding segment

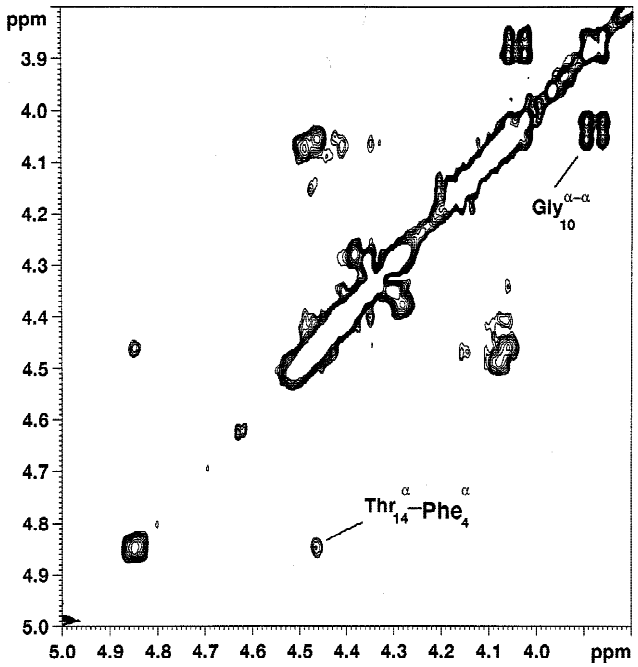

Fig. 7. Part of the 200 ms NOESY spectrum of U(1–17) in D_2O (2 °C, 1 mM, pH 3.8), showing the interstrand $C_\alpha H$-$C_\alpha H$ NOE consistent with a native like β-hairpin register (see Fig. 1). The intraresidue NOE between the $C_\alpha H$ protons of Gly10, which was used a reference peak for intensity calibration, is also shown.

of the unfolded state of the full hairpin peptide (Maynard et al., 1998).

The spectra of the two short peptides, each corresponding to one arm of the U(1–17) β-hairpin, were assigned. Examination of the NOESY data revealed only sequential and intraresidue NOEs. $C_\alpha H$ chemical shifts (Fig. 8) and coupling constants (Table 1) were determined and, as expected, were found to be markedly closer to the model "random coil" values than those measured for U(1–17), confirming that the deviations from these values observed for the larger peptide are indeed, at least in good part, a reflection of hairpin formation.

Comparison with the subfragment provides insight into some of the apparently anomalous features noted for U(1–17). For example, the upfield $C_\alpha H$ shift perturbations observed for Ile3 and Val5 appear to violate the usual rule that these protons experience downfield shift perturbations in a β-sheet structure. It turns out, however, that both of these protons are even more upfield shifted in U(1–7), so that we can now attribute this effect to a local interaction, within residues 1–7, in the unfolded state and see that the effect of hairpin formation is, as would have been expected, actually to superimpose on this a modest downfield shift. In the case of Ile3, at least, a similarly unexpected shift is also observed in the spectrum of native ubiquitin, and it seems likely that in all of these systems these perturbations may be due to the ring current from the adjacent Phe4. Qualitatively similar effects of aromatic residues on the backbone proton chemical shifts of near neighbors in the sequence have, indeed, been noted in the spectra of small, unfolded peptides (Merutka et al., 1995).

Figure 8 shows that, these few anomalous residues excepted, the $C_\alpha H$ chemical shifts in the short peptides are mostly rather close to published "random coil" values and that most of the residual discrepancies are actually upfield, in the opposite sense to those characteristic of β-structure. The backbone *J*-couplings show rather more significant deviations from model random coil values and, in general, these deviations are in the same direction as those observed in the spectrum of U(1–17), probably reflecting some preference for the population of β-space even for residues in the small peptides. The relative intensities of sequential NOEs in the spectra of the peptide subfragments differ dramatically from the values expected from the random coil model: in both U(1–7) and U(11–17) the ratios $I_{\alpha N}(i,i)/I_{\alpha N}(i,i+1)$ were found generally to be close to 0.2 for all residue pairs for which they could be measured. This value is essentially the same as that determined for the corresponding residues in U(1–17) and is close to the expected value for an idealized β-sheet. Similarly, the values of the ratio $I_{NN}(i,i+1)/I_{\alpha N}(i,i+1)$ were found to be very similar to those obtained in the case of U(1–17) (Table 2).

It is reasonable to assume that the deviations from random coil values observed for NMR parameters of U(1–7) and U(11–17) would be reflected in the spectra of the unfolded state of U(1–17), in the hypothetical absence of hairpin formation. We should therefore be able to improve our estimates of the β-hairpin population for U(1–17) by adopting these, rather than literature-derived random coil values, as estimates for the unfolded state values. It is then immediately clear that population estimates based on sequential NOE intensities are not meaningful, since there is virtually no detectable difference in the intensity ratios between U(1–17) and the smaller peptides. Thus, the very large values obtained for the hairpin population using the original analysis of these data can be dismissed as artefactual.

The revised apparent hairpin populations derived from the chemical shift and *J*-coupling data using this new baseline are summarized in Figure 9A. There is still some variation in the apparent fractional populations calculated for different residues, but only

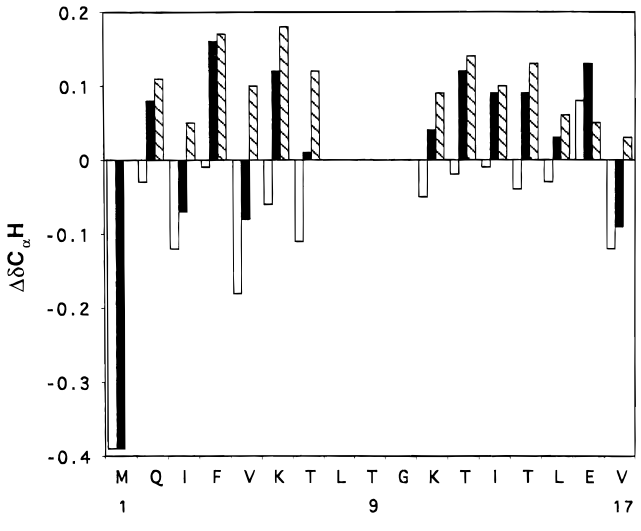

Fig. 8. $C_\alpha H$ chemical shift differences: U(1–7) and U(11–17) from literature random coil values (clear bars); U(1–17) from literature random coil values (black bars); U(1–17) from U(1–7) and U(11–17) (hatched bars).

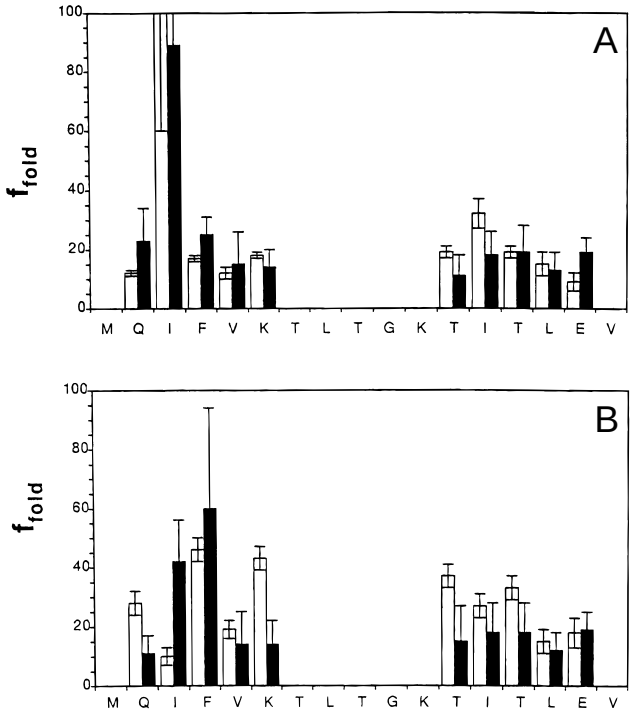

Fig. 9. Hairpin population estimates for U(1–17) in H_2O:D_2O (9:1), obtained using data from the spectra of U(1–7) and U(11–17) to model the unfolded state. In each case the black bars represent values derived from $^3J_{HN\alpha}$ data; clear bars are values derived from $C_\alpha H$ shifts perturbations. **A:** Values obtained using native ubiquitin to model the structured state. **B:** Values obtained using an average over a number of proteins to model the structured state (see text).

Ile3 now stands out as being grossly out of line with the others. The very large population estimates obtained for this residue seem to be related primarily to anomalous native state chemical shifts and coupling constants. Otherwise, there is now reasonable agreement between the population estimates obtained for different residues and between those based on coupling constants and those based on $C_\alpha H$ chemical shifts: the average population estimated for residues within the β-strands (omitting Ile3) are 17 and 18%, respectively, when native ubiquitin is used to provide the folded state baseline.

Alternative ways to obtain estimates of NMR parameters characteristic of the structured state of the hairpin peptide include using theoretical values for an idealized structure or average database values from a number of β-structures. The rationale for the latter approach would be that the superimposed effects of specific tertiary interactions would be diluted or, indeed, averaged away. To see whether further improvement in the consistency of population estimates might thus be obtained in the U(1–17) system, we adopted an approach similar to that used by Ramirez-Alvarado et al. (1996), in their study of a hairpin formed by a designed peptide, with the results shown in Figure 9B. This approach had the desired effect of removing the extreme outlier, Ile3, which had resulted when native ubiquitin was used to provide the reference values. Otherwise, however, the general level of consistency from residue to residue and, more particularly, between the values derived from *J*-couplings and from chemical shifts is actually less good using this approach. This mainly reflects a general increase in the population estimates based on chemical shifts, which is not matched in those based on coupling constants.

Discussion

An autonomously folding β-hairpin

The NMR data obtained for the U(1–17) peptide in aqueous solution demonstrate unequivocally that a native-like β-hairpin is present. This is of great interest in relation to protein folding, because it reinforces the conclusion that the conformation of this part of the ubiquitin molecule is encoded in its own sequence, rather than being imposed by the requirements of the overall native structure. Knowledge of this type is invaluable in trying to construct a hierarchical understanding of folded protein structures.

The sequences both of the arms of the hairpin and of the turn are likely to be important factors favoring autonomous folding of the hairpin structure. The residues of the two arms include a high proportion of those which show a pronounced preference for β-structure: 7 out of 14 are β-branched residues and one of the remainder is aromatic, while only Glu16 has a significantly low β-propensity. On the other hand, the pairing of residues across the strands does not feature many of the combinations that have been shown to be especially favorable on the basis of statistical analysis of the structural database (Wouters & Curmi, 1995; Hutchinson et al., 1998). Thus, despite the prevalence of β-branched residues, there is only one cross-strand pair of them (Val5–Ile13), and these are a hydrogen bonded pair, which means that the orientation of the side chains is not optimal for close interaction, and there is no statistical evidence for a significantly favored interaction between them in such a position (Hutchinson et al., 1998). There are no other obvious features, such as disulfide bridges or charge pairs, that might help to stabilize a β-structure. These considerations suggest that the β-hairpin is favored predominantly because of the overall composition of the strands, rather than a network of specifically favored interactions between them.

The NMR data suggest that there is no significant population of alternative hairpin structures and the question therefore arises of how the register of the strands is specified. This contrasts with some designed peptides, for which coexistence of populations of hairpins with alternative registers has been observed (de Alba et al., 1996), and it is especially interesting in relation to a modified version of U(1–17), in which the five residue turn was replaced with a quite different four residue sequence, which was found to form a well-populated hairpin with a five residue turn and a register shifted by one residue (Searle et al., 1995). The latter result is consistent with our conclusion that the strand sequences in U(1–17) favor β-structure in general but do not afford strong pairing preferences. In this case, therefore, it seems likely that it is the turn sequence that has the dominant role in determining the register. A similar conclusion has been drawn for a number of hairpin peptide systems (Ramirez-Alvarado et al., 1997a; Gellman, 1998), although in other cases, residue pairing preferences would certainly have a strong influence as well.

The G1 β-bulge that occurs in native ubiquitin has been observed to be especially favored in β-hairpin structures, on the basis of statistical evidence (Sibanda & Thornton, 1991) and of studies of designed peptides (de Alba et al., 1996). It seems likely, therefore, that it is the formation of this structure that favors the native-like strand register observed in U(1–17). This particular turn type has an overwhelming preference for Gly at the fourth position, which is compatible with the sequence Thr-Leu-Thr-Gly-Lys that occurs in U(1–17) in its native-like hairpin conformation, and it can readily be seen how this preference could force the alternative register observed when the native turn was replaced with a four residue sequence that happened to have the requisite Gly in the fourth position.

The autonomous folding of U(1–17) in aqueous solution confirms the conclusion from our earlier study in aqueous methanol, that this part of the ubiquitin structure is strongly encoded in its own sequence (Cox et al., 1993). Preliminary studies of a peptide spanning the C-terminal half of the molecule (residues 36–76) suggest, by contrast, that native like β-sheet is not significantly populated in aqueous solution (J.P. Mackay, unpubl. results), and this part of the molecule was observed to form non-native secondary structure in the aqueous methanol denatured state (Stockman et al., 1993). Thus it is plausible to suggest that the N-terminal part of the molecule, possibly also including the α-helix, forms an autonomously folding platform, which then may play an important role in dictating the remainder of the fold through tertiary structural constraints.

A number of protein domains have a similar tertiary structure to ubiquitin, which has been described as the ubiquitin α/β roll "superfold" (Orengo & Thornton, 1993). A particularly interesting example is the B1 domain of streptococcal protein G, which has been subject to a similar "dissection" into peptide fragments (Blanco et al., 1994a, 1994b; Blanco & Serrano, 1995). Remarkably, in this case, the pattern of stability of autonomously folding structural elements is quite different from that observed in ubiquitin. The N-terminal hairpin, analogous to U(1–17), appears to be essentially unfolded in aqueous solution, whereas a C-terminal segment forms a well-populated β-hairpin. In ubiquitin an elaboration of the fold means that there is no direct analogue of this second hairpin but, in any case, the C-terminal portion of the protein shows no evidence of a tendency to fold in isolation. In the equiv-

alent domain of a related molecule, protein L, neither hairpin appears to have significant ability to fold independently under aqueous conditions (Ramirez-Alvarado et al., 1997b). Thus it seems that the conserved overall folding topology is not reflected in a common pattern of local encoding of structure, and it is thus clear that it may be dangerous to try to generalize conclusions about the organization of folding from a single protein, even to its immediate structural homologues.

Conformational properties of isolated strands

The peptides U(1–7) and U(11–17), each corresponding to an isolated arm of the hairpin, were investigated with the aim of providing a model for the behavior of the parent peptide, U(1–17), in the absence of hairpin formation. Unfolded states are generally supposed to approximate reasonably well to the so called random coil model and estimates of NMR parameters characteristic of such a state have generally been made from spectra of small, apparently unstructured peptides. It has been recognized that different residue types may have different average conformational properties even in such a highly disordered state and NMR observables have therefore now been measured for each of the common amino acid residues in the context of suitable small peptides (Wüthrich, 1986; Merutka et al., 1995; Fiebig et al., 1996; Smith et al., 1996). This now provides a good baseline for assessing deviations from random behavior in experimental systems.

The NMR data obtained for U(1–7) and U(11–17) do, in fact, show considerable deviations from the random coil baseline. This is most marked for the backbone NOE intensity ratios, which suggest a very high population of β-sheet structure, according to conventional analysis, even though these short peptides are clearly unable to form such a secondary structure. Several of the backbone *J*-couplings are affected in a similar way, though less markedly. On the other hand, the $C_\alpha H$ chemical shifts are actually rather close to the predicted random coil values and do not hint at any significant β-sheet population. Rather similar behavior has been noted in an analogous study of a designed, autonomously folding β-hairpin, suggesting that this is likely to be a general feature of sequences with this type of structural propensity (Maynard et al., 1998).

There is a rather marked difference between the NOE intensity ratios we observe for U(1–7) and U(11–17) and those that have been reported in a number of other small, unstructured peptides, suggesting that there is a real qualitative difference in the sampling of conformational space by these strongly β-favoring peptide segments of ubiquitin. It seems likely that in both peptide subfragments there are significant constraints on the sampling of backbone conformations, so that residues may be confined to the β-region of conformational space for a greater proportion of the time than would be expected in such a small, unstructured peptide. In that case, the relative insensitivity of other NMR observables to the nonrandom conformational distribution would have to be interpreted in terms of their sensitivity to different aspects of β-structure.

The lack of any significant effect on the $C_\alpha H$ shifts in the spectra of U(1–7) and U(1–17) suggests, as Maynard et al. (1998) have pointed out, that the downfield shifts characteristically observed for protons in β-structure result largely from proximity to the carbonyl groups of the opposite, hydrogen bonded strand, rather than those adjacent in the sequence. However, since downfield shifts, albeit somewhat smaller, are observed also for $C_\alpha H$ protons directed outward from the hairpin structure, there may well be an additional effect operative in a bona fide β-structure, which is diminished in an isolated strand, perhaps related the fact that residues are likely to be more geometrically constrained by β-sheet hydrogen bonding than by the more general steric effects that are operative in isolated segments of the structure. For the backbone *J*-couplings measured for the short peptides, there is an intermediate situation: these are certainly indicative of substantial conformational freedom but, in contrast to the $C_\alpha H$ shifts, there are nonetheless clear differences from random coil values. This would suggest that an increased preference for the broad region of β-space has a measurable effect on this parameter, as observed in the short peptides, but that the additional conformational constraints imposed by hydrogen bonding are needed to produce the larger effects characteristic of β-sheet structure.

In these short peptides, the restricted conformational freedom we have thus inferred can only reflect the optimization of interactions between residues close by in the sequence. The suggestion is, therefore, that even in the absence of any interstrand interactions, there can be a substantial effect on the backbone conformational distribution as a result of the cooperative effect between residues of high β-propensity clustered together in the sequence. This is probably analogous to the "nascent helix," which may result when residues of high α-propensity are predominant in a local sequence (Dyson et al., 1992). These effects may be quite important in relation to folding equilibria, since in parts of the sequence such as these the configurational entropy loss on forming the native structure of a protein will be significantly smaller, and therefore less destabilizing, than would be the case if the full backbone conformational freedom predicted by the classical random coil model prevailed.

Population analysis

Quantitative analysis of the degree of folding present in β-hairpin peptides is much less straightforward than in the case of α-helical structures, as has been discussed recently by Ramirez-Alvarado et al. (1997a). In the present study, we have adopted a simplified model, which assumes that the equilibrium can be described in terms of just two well-defined states. It was not possible to obtain spectra of either of these states in isolation, since the equilibrium cannot easily be shifted fully in either direction, thus making it necessary to estimate baseline values for NMR parameters in the putative unfolded and folded states of the peptide.

Using our initial choice of models for these limiting states—native ubiquitin to model the folded state and literature "random coil" NMR parameters to describe the unfolded state—we were unable to arrive at any consensus as to the actual population of hairpin structure by U(1–17): values obtained using different residues as probes varied widely along the sequence and estimates based on different NMR parameters were strikingly different. This variation could not be explained on the basis of any simple structural model such as strand fraying, suggesting either that our choice of reference states or the two-state model itself might be inadequate to describe the conformational equilibrium.

Use of the short peptides U(1–7) and U(11–17) as a model revealed, as we have seen, that at least some of the NMR parameters characteristic of the unfolded state are not well predicted by literature "random coil" values. Most dramatically, the backbone NOE intensity ratios, which had given the largest population estimates in the original analysis of the U(1–17) data, were found to be virtually identical for the short peptides and for U(1–17) and could therefore be dismissed as useless for the estimating hairpin

population. A new unfolded state baseline based on these peptide subfragments did lead to a significant improvement in the level of agreement between population estimates based on chemical shifts and coupling constants: neglecting a solitary outlier, the parameters measured suggest a population of the hairpin in the region of 20% (Fig. 9A). There remains at least one obvious anomaly and a degree of variation from residue to residue, however.

The residual scatter in the population estimates presumably reflects, at least in part, the problems in finding a good model for the folded state of the hairpin. Our initial model implicitly assumed that the fully folded U(1–17) peptide has the same conformation as the corresponding residues in the full length ubiquitin protein. In reality, of course, there is lot of scope for conformational adjustments in β-structure and the average hairpin conformation and the effects of dynamic averaging could be quite different when not constrained by tertiary interactions with the remainder of the protein, so that this is almost certainly not a safe assumption. An alternative approach, using average native state parameters derived from a number of β-hairpins in various proteins led to many differences of detail in the estimated populations, including suppression of the most prominent outlier in the distribution, but the overall level of agreement between estimates based on *J*-couplings and $C_\alpha H$ chemical shifts was not improved by this approach (Fig. 9B).

A third option would be to attempt to find conditions under which the folded state of the U(1–17) peptide is fully populated and use NMR variables determined under these conditions to provide a baseline for the fully structured peptide. U(1–17) in methanol:water mixtures appears to be more fully structured than it is in water alone, and one possibility might therefore have been to view the effect of the cosolvent as simply increasing the population of the hairpin structure, rather than altering the structure itself, and thus to use NMR parameters measured for the peptide in the mixed solvent as estimates for the limiting values. In practice, however, this approach was not feasible because experiments revealed that at the volume composition used in this work (3:7 CD_3OD:H_2O) neither the backbone chemical shifts nor coupling constants had yet reached plateau values and, in the presence of the greater proportion of alcohol that would presumably be needed to force the equilibrium fully over in favor of the structured state, the peptide was found to aggregate strongly. Moreover, the balance of forces driving folding will be quite different in aqueous solution and in a mixed aqueous:organic phase, with the consequence that the details of the average structure of the folded state of the peptide may also be significantly altered. There is some evidence to support this in comparing the spectra of U(1–17) obtained under the different solvent conditions. This approach was therefore not pursued further.

The other type of structurally sensitive parameter that we investigated was an interstrand NOE between two $C_\alpha H$ protons. Comparing the intensity of this strong and clearly resolved NOE with what would be expected from either an idealized β-hairpin or from the hairpin as it is in the native ubiquitin structure suggested folded state populations in the region of 60%. These values are clearly much greater than the other estimates and might, again, suggest a problem with the native state model. However, the major problem with NOE intensities is the r^{-6} distance dependence of the interaction, which leads to considerable uncertainty in the population (Ramirez-Alvarado et al., 1997b). Therefore we cannot draw any strong conclusions from this apparent anomaly.

A significant underlying limitation of the two-state analysis is likely to be the assumption that the structured state can adequately be described in terms of a single conformation. In a native protein structure, the conformation of a particular peptide segment is frequently tightly defined because the close packing against other parts of the structure limit the scope for large amplitude fluctuations. There are obviously exceptions—poorly defined loops and terminal segments are common—but in most instances it does mean that the concept of a single "average structure" is adequate for interpretation of most experimental data. However, it is less clear that this will be the case for a small peptide fragment, where the number of constraints cooperating to define its conformation will necessarily be much smaller. For example, in a somewhat analogous antibiotic/peptide binding system, a clear dependence of the limiting values of NMR parameters on the number of interactions constraining the bound state was demonstrable (Williams et al., 1997). Thus, the two-state model may not be adequate to account for the data, whatever structure is assumed for the folded state. Instead the NMR parameters measured may need to be understood as averages over a broad distribution of conformations in rapid equilibrium with one another. Since all measurable parameters will vary in different ways as a result of conformational fluctuations, it would not be surprising that any attempt to force the data to fit to a simple two-state model would produce inconsistent results. The evidence of the analysis we have carried out here, once the improved model for the unfolded state is incorporated, is that this may not be an unmanageably great problem, since there was a reasonable degree of agreement between population estimates based on different parameters. However, it remains to be seen whether a more detailed understanding of the conformational properties of hairpins such as this is attainable within the constraints of a model that assumes a single, static structure for the hairpin.

Materials and methods

Peptide synthesis and purification

Peptides were synthesized using the continuous flow Fmoc-polyamide method and purified by RP-HPLC, as described for related peptides (Cox et al., 1993). The peptide denoted U(1–17), corresponding to the first 17 residues of ubiquitin, was synthesized as the C-terminal amide; the peptide denoted U(11–17) was protected as the amide at both the N- and C-termini while the peptide U(1–7) was not protected. The final products were characterized by amino acid analysis and positive ion electrospray mass spectrometry. Intact ubiquitin was purchased from Fluka (Buchs, Switzerland) and was used without further purification.

Peptide aggregation

^{1}H NMR spectra of U(1–17) in aqueous solution at pH 3.8 and low temperatures were characterized by sharp resonances and were not observed to change significantly between 0 and 7 °C. As the temperature was raised further, however, a degree of broadening became apparent. This effect was entirely reversible and was found to be concentration dependent. It may be attributed to limited self-association of the peptide, which appears to be minimized at low temperature. Over the concentration range 0.1 to 3.0 mM, there was no significant variation in linewidths or, more importantly, in chemical shifts, even of protons whose resonances were later shown to be highly sensitive to secondary structure formation, in spectra recorded at 2 °C. All further studies were, therefore,

pursued at low temperature (2 °C), where we could confidently interpret the spectra in terms of monomeric peptide structures.

NMR spectroscopy

Samples for NMR spectroscopy were prepared by dissolving the peptide or full length protein in unbuffered 9:1 H_2O:D_2O or D_2O. For aqueous methanolic solutions, the samples were first dissolved in the appropriate volume of H_2O or D_2O and then CD_3OD was added in the appropriate proportion. Solution pH values were measured using a Corning 240 pH electrode and were adjusted by adding aliquots of D_2O solution of NaOD or DCl to a final pH of 3.8 (uncorrected for isotope and solvent effects). All spectra were referenced to the singlet resonance of internal dioxan (3.74 ppm relative to TSP). Spectra were recorded on Bruker DRX 500 and Varian Unity 600 spectrometers.

For one-dimensional experiments, 32k data points were collected, while two-dimensional spectra were defined by 2k or 4k complex data points in f2 and 512 or 1k real points in f1. The spectra were acquired in phase-sensitive mode using time proportional phase incrementation (Marion & Wüthrich, 1983) to achieve quadrature detection in f1. Spectral widths were 10 ppm for U(1–17) and 12 ppm for the native protein. NOESY (Jeener et al., 1979), TOCSY (Braunschweiler & Ernst, 1983), and DQF-COSY (Piantini et al., 1982) spectra were collected. For the peptide samples, relatively long mixing NOESY times were generally used (t_m = 300 ms) to allow build up of medium- and long-range NOEs, although spectra with shorter mixing times (t_m = 100, 200 ms) were also recorded to avoid misinterpretation as a result of spin diffusion. NOESY mixing times for the native protein spectra were 50, 80, and 120 ms. Suppression of intense solvent resonances was achieved by presaturation or use of the WATERGATE sequence (Piotto et al., 1992). Data were processed on a Silicon Graphics Indy workstation using XWINNMR software (Bruker, Karlsruhe, Germany); shifted squared sinebell window weighting functions (phase shift of $\pi/2$ to $\pi/6$) were generally used in both dimensions. Baseline corrections were applied to the processed spectra in both dimensions.

$^3J_{HN\alpha}$ spin–spin couplings were measured using the method of Titman and Keeler (1990). NOESY and DQ-correlation (Braunschweiler & Ernst, 1983) experiments were acquired with 4k × 1k datapoints. For comparison, a one-dimensional spectrum with 32k acquired points was used to measure the $^3J_{HN\alpha}$ values for resolved resonances. Cross-peak intensities in the NOESY spectra were obtained by calculating peak volume integrals using the program ANSIG (Kraulis, 1989). Interproton distances in the native state were calculated from the crystal structure using the program QUANTA (Molecular Simulations, Burlington, Massachusetts).

Calculations of population of folded structure

Apparent fractions of folded peptide were calculated for each residue, using three types of NMR parameter: $C_\alpha H$ chemical shifts, $^3J_{HN\alpha}$ couplings and backbone NOE intensities. A simple two-state model was assumed in which the folded peptide and a random coil state are in rapid equilibrium in the aqueous solution. Thus, the apparent population of folded peptide based on the $\Delta\delta C_\alpha H$ values was estimated according to the equation: $f_{fold} = \Delta\delta C_\alpha H_p / \Delta\delta C_\alpha H_{ref}$; where $\Delta\delta C_\alpha H_p$ is the value for the peptide and $\Delta\delta C_\alpha H_{ref}$ is the value for the chosen reference state quantified (Jimenez et al., 1993; Blanco et al., 1994b; Blanco & Serrano, 1995). The random coil chemical shift values used were those given by Bundi and Wüthrich (1979).

The apparent fractional populations of folded peptide based on the coupling constants values for each residue was obtained using the equation: $f_{fold} = \Delta^3 J_{HN\alpha\ p} / \Delta^3 J_{HN\alpha\ ref}$; where $\Delta^3 J_{HN\alpha\ p}$ represents the deviation of the experimental coupling constant for the peptide from the value predicted for a random coil and $\Delta^3 J_{HN\alpha\ ref}$ represents the corresponding deviation from random coil for a fully-folded reference state (Waltho et al., 1993; Ramirez-Alvarado et al., 1996).

Values of the ratio of backbone NOE intensities, $I_{NN}(i,i+1)/I_{\alpha N}(i,i+1)$ were also used to calculate f_{fold}, according to the equation (Bradley et al., 1990):

$$f_{fold} = \frac{\Delta[I_{NN}(i,i+1)/I_{\alpha N}(i,i+1)]_p}{\Delta[I_{NN}(i,i+1)/I_{\alpha N}(i,i+1)]_{ref}}. \quad (1)$$

The random coil value of the ratio was taken to be 0.4 (Fiebig et al., 1996) while values for the folded state were estimated from the native ubiquitin crystal structure.

As an alternative approach to estimating the values of NMR parameters characteristic of the hairpin structure, average values of backbone coupling constants and chemical shift perturbations derived from a number of folded proteins were computed. For the coupling constants $^3J_{HN\alpha}$, these were calculated from Φ values measured from the crystallographic coordinates of 14 β-hairpins containing G1 β-bulges (Sibanda & Thornton, 1991). The average $C_\alpha H$ shift deviation from the random coil position was taken to be 0.358 ppm for protons in β-structure (Williamson, 1990).

Supplementary material for Electronic Appendix

Tables of the complete ^{1}H NMR assignment of U(1–17), U(1–7), U(11–17) in aqueous and aqueous methanolic solution are available as supplementary material.

Acknowledgments

This work was supported by the BBSRC through the Cambridge Centre for Molecular Recognition (CCMR). R.Z. gratefully acknowledges the British Council and European Community for a Training and Mobility for Researcher Fellowship. J.M.C.I. was supported by a studentship from the MRC. B.W.T. thanks the U.S. Churchill Foundation, for Scholarships. J.P.M. thanks the Association of Commonwealth Universities. P.A.E. is a Lister Institute Research Fellow.

References

Blanco FJ, Jimenez MA, Herranz J, Rico M, Santoro J, Nieto JL. 1993. NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution. *J Am Chem Soc 115*:5887–5888.

Blanco FJ, Jimenez MA, Pineda A, Rico M, Santoro J, Nieto JL. 1994a. NMR solution structure of the isolated N-terminal fragment of protein-G B-1 domain—Evidence of trifluoroethanol induced native-like β-hairpin formation. *Biochemistry 33*:6004–6014.

Blanco FJ, Rivas G, Serrano L. 1994b. A short linear peptide that folds into a native stable β-hairpin in aqueous solution. *Nature Struct Biol 1*:584–590.

Blanco FJ, Serrano L. 1995. Folding of protein-G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur J Biochem 230*:634–649.

Bradley EK, Thomason JF, Cohen FE, Kosen PA, Kuntz ID. 1990. Studies of synthetic helical peptides using circular-dichroism and nuclear-magnetic-resonance. *J Mol Biol 215*:607–622.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing—Application to proton correlation spectroscopy. *J Magn Reson 53*:521–528.

Bundi A, Wüthrich K. 1979. ^{1}H NMR parameters of the common amino acid residues measured in aqueous solution of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers 18*:285–297.

Cox JPL, Evans PA, Packman LC, Williams DH, Woolfson DN. 1993. Dissecting the structure of a partially folded protein—Circular dichroism and nuclear magnetic resonance studies of peptides from ubiquitin. *J Mol Biol 234*:483–492.

de Alba E, Jimenez MA, Rico M, Nieto JL. 1996. Conformational investigation of designed short linear peptides able to fold into β-hairpin structures in aqueous solution. *Folding Design 1*:133–144.

Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. 1992. Folding of peptide-fragments comprising the complete sequence of proteins—Models for initiation of protein folding. 1. Myohemerythrin. *J Mol Biol 226*:795–817.

Dyson HJ, Rance M, Houghten RA, Lerner RA, Wright PE. 1988. Folding of immunogenic peptide fragments of proteins in water solution. 1. Sequence requirements for the formation of a reverse turn. *J Mol Biol 201*:161–200.

Fiebig KM, Schwalbe H, Buck M, Smith LJ, Dobson CM. 1996. Toward a description of the conformations of denatured states of proteins—Comparison of a random coil model with NMR measurements. *J Phys Chem 100*:2661–2666.

Gellman SH. 1998. Minimal model systems for β-sheet secondary structure in proteins. *Curr Opin Chem Biol 2*:717–725.

Harding MM, Williams DH, Woolfson DN. 1991. Characterization of a partially denatured state of a protein by 2-dimensional NMR—Reduction of the hydrophobic interaction in ubiquitin. *Biochemistry 30*:3120–3128.

Hutchinson EG, Sessions RB, Thornton JM, Woolfson DN. 1998. Determinants of strand register in antiparallel β-sheets of proteins. *Protein Sci 7*:2287–2300.

Jeener J, Meier BH, Bachmann P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Phys Chem 71*:4546–4553.

Jimenez MA, Bruix M, Gonzalez C, Blanco FJ, Nieto JL, Herranz J, Rico M. 1993. CD and H-1-NMR studies on the conformational properties of peptide fragments from the C-terminal domain of thermolysin. *Eur J Biochem 211*:569–581.

Kraulis PJ. 1989. ANSIG—A program for the assignment of protein H-1 2D-NMR spectra by interactive computer graphics. *J Magn Reson 84*:627–633.

Marion D, Wüthrich K. 1983. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of H-1-H-1 spin-spin coupling constants in proteins. *Biochem Biophys Res Commun 113*:967–974.

Maynard AJ, Sharman GJ, Searle MS. 1998. Origin of beta-hairpin stability in solution: Structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilization in water. *J Am Chem Soc 120*:1996–2007.

Merutka G, Dyson HJ, Wright PE. 1995. Random coil H-1 chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J Biomol NMR 5*:14–24.

Orengo CA, Thornton JM. 1993. Alpha plus beta folds revisited: Some favored motifs. *Structure 1*:105–120.

Piantini U, Sørensen OW, Ernst RR. 1982. Multiple quantum filters for elucidating NMR coupling networks. *J Am Chem Soc 104*:6800–6801.

Piotto M, Saudek V, Sklenar V. 1992. Gradient-tailored excitation for single-quantum NMR-spectroscopy of aqueous solutions. *J Biomol NMR 2*:661–665.

Ramirez-Alvarado M, Blanco FJ, Niemann H, Serrano L. 1997a. Role of beta-turn residues in beta-hairpin formation and stability in designed peptides. *J Mol Biol 273*:898–912.

Ramirez-Alvarado M, Blanco FJ, Serrano L. 1996. *De novo* design and structural analysis of a model β-hairpin peptide system. *Nature Struct Biol 3*:604–612.

Ramirez-Alvarado M, Serrano L, Blanco FJ. 1997b. Conformational analysis of peptides corresponding to all the secondary structure elements of protein L B1 domain: Secondary structure propensities are not conserved in proteins with the same fold. *Protein Sci 6*:162-174.

Saulitis J, Liepins E. 1990. Quantitative evaluation of interproton distances in peptides by 2-dimensional Overhauser effect spectroscopy. *J Magn Reson 87*:80–91.

Scholtz JM, Baldwin RL. 1992. The mechanism of α-helix formation by peptides. *Ann Rev Biophys Biomol Struct 21*:95–118.

Searle MS, Williams DH, Packman LC. 1995. A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native β-hairpin. *Nature Struct Biol 2*:999–1006.

Searle MS, Zerella R, Williams DH, Packman LC. 1996. Native-like β-hairpin structure in an isolated fragment from ferredoxin—NMR and CD studies of solvent effects on the N-terminal-20 residues. *Protein Eng 9*:559–565.

Serrano L. 1995. Comparison between the phi distribution of the amino acids in the protein database and NMR data indicates that amino acids have various phi propensities in the random coil conformation. *J Mol Biol 254*:322–333.

Sibanda BL, Thornton JM. 1991. Conformation Of β-hairpins in protein structures—Classification and diversity in homologous structures. *Methods Enzymol 202*:59–82.

Smith LJ, Bolin KA, Schwalbe H, Macarthur MW, Thornton JM, Dobson CM. 1996. Analysis of main-chain torsion angles in proteins—Prediction of NMR coupling constants for native and random coil conformations. *J Mol Biol 255*: 494–506.

Stockman BJ, Euvrard A, Scahill TA. 1993. Heteronuclear 3-dimensional NMR spectroscopy of a partially denatured protein—The A-state of human ubiquitin. *J Biomol NMR 3*:285–296.

Titman JJ, Keeler J. 1990. Measurement of homonuclear coupling constants from NMR correlation spectra. *J Magn Reson 89*:640–646.

Vijaykumar S, Bugg CE, Cook WJ. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol 194*:531–544.

Waltho JP, Feher VA, Merutka G, Dyson HJ, Wright PE. 1993. Peptide models of protein folding initiation sites: 1. Secondary structure formation by peptides corresponding to the G-helix and H-helix Of myoglobin. *Biochemistry 32*:6337–6347.

Williams DH, Bardsley B, Tsuzuki W, Maguire AJ. 1998. A limitation of two-state analysis for transitions between disordered and weakly ordered states. *Chem Biol 4*:507–512.

Williamson MP. 1990. Secondary structure dependent chemical shifts in proteins. *Biopolymers 29*:1428–1431.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol 222*:311–333.

Wouters MA, Curmi PM. 1995. An analysis of side chain interactions and pair correlations within antiparallel beta-sheets: The differences between backbone hydrogen-bonded and non-hydrogen-bonded residue pairs. *Proteins 2*:119–131.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: John Wiley & Sons.