Structural characterization of a mutant peptide derived from ubiquitin: Implications for protein folding

ROSA ZERELLA,¹ PEI-YEH CHEN,^{2,3} PHILIP A. EVANS,² ANDREW RAINE,²
AND DUDLEY H. WILLIAMS¹

¹Cambridge Centre for Molecular Recognition and University Chemical Laboratory, Lensfield Road,
Cambridge CB2 1EW, United Kingdom

²Cambridge Centre for Molecular Recognition and Department of Biochemistry, Tennis Court Road,
Cambridge CB2 1QW, United Kingdom

(RECEIVED May 31, 2000; FINAL REVISION August 15, 2000; ACCEPTED August 24, 2000)

Abstract

The formation of the N-terminal β -hairpin of ubiquitin is thought to be an early event in the folding of this small protein. Previously, we have shown that a peptide corresponding to residues 1–17 of ubiquitin folds autonomously and is likely to have a native-like hairpin register. To investigate the causes of the stability of this fold, we have made mutations in the amino acids at the apex of the turn. We find that in a peptide where Thr9 is replaced by Asp, U(1–17)T9D, the native conformation is stabilized with respect to the wild-type sequence, so much so that we are able to characterize the structure of the mutant peptide fully by NMR spectroscopy. The data indicate that U(1–17)T9D peptide does indeed form a hairpin with a native-like register and a type I turn with a G1 β -bulge, as in the full-length protein. The reason for the greater stability of the U(1–17)T9D mutant remains uncertain, but there are nuclear Overhauser effects between the side chains of Asp9 and Lys11, which may indicate that a charge–charge interaction between these residues is responsible.

Keywords: β -hairpin; β -sheet; NMR; peptide; peptide conformations; structure; ubiquitin; X-PLOR

The β -hairpin is one of the most commonly occurring structural motifs in globular protein folds: most antiparallel β -sheets contain at least some strand pairs that are contiguous in sequence and connected by a relatively tight turn. There has also been much speculation and some experimental evidence to support the idea that hairpins may play a key role in initiating the assembly of β -structures in folding (Ptitsyn, 1991; Muñoz et al., 1997). In recent years, there has been considerable interest in using β -hairpins as model systems for developing our understanding of β -structure and several examples both of protein fragments and of de novo designed peptides capable of folding autonomously as hairpins have now been identified (Blanco et al., 1994; Searle et al., 1995; Ramirez-Alvarado et al., 1996). The lessons learned from exploring the effects of sequence variations in these systems have been reviewed recently (Griffiths-Jones et al., 1998; Ramirez-Alvarado et al., 1999).

Sequence variations cannot only alter the stability of hairpin structures, but also the register of the β -strands. In the arms of the hairpin, a high β -propensity of individual residues is important for stability. However, specific pairwise contacts, especially between

side chains interacting across the hairpin, are also important and optimization of these is hypothesized to be an important contribution in fixing the strand register (Smith & Regan, 1995; Wouters & Curmi, 1995; Hutchinson et al., 1998). The turn sequence is also important: it is clear from analysis of hairpins within intact protein structures that some turn types are prevalent in β -hairpins. Sequence compatibility with these preferred structures has been shown in peptide studies to be an important determinant of structural specificity and stability (Ramirez-Alvarado et al., 1997; De Alba et al., 1999; Griffiths-Jones et al., 1999). The structural mechanisms by which β -structures are stabilized, and specific strand registers are selected, continue to be the subject of considerable interest.

For the small globular protein ubiquitin, it has been observed that the full-length protein adopts an unusual structure in aqueous methanol: the N-terminal 35 residues form a native-like structure while the remaining residues form a nonnative structure (Harding et al., 1991; Stockman et al., 1993). This suggests particular stability of the N-terminal region of the protein. This finding prompted the investigation of the potential for independent folding of peptide fragments derived from the N-terminus of ubiquitin. Later work showed that a peptide comprising the N-terminal 17 residues of the protein forms a native-like β -hairpin in both aqueous methanol and water, albeit at low apparent population (Cox et al., 1993; Zerella et al., 1999). Furthermore, protein engineering studies have re-

Reprint requests to: Dudley H. Williams, Cambridge Centre for Molecular Recognition, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom; e-mail: dhw1@cam.ac.uk.

³Present address: Institute of Chemistry, Academia Sinica, Taipei, Taiwan.

vealed that residues 1–17 are involved in a productive intermediate populated at an early stage of the folding process (Khorasanizadeh et al., 1996), suggesting that autonomous folding of this structural element might well play an important role in the folding mechanism of this protein.

Hairpin-forming peptides are proving valuable as model systems, but it is important to recognize that they may not be fully representative of β -structures in globular proteins: an isolated hairpin lacks the tertiary interactions that would help to define its geometry in a folded protein. An isolated hairpin may form non-native intramolecular interactions that will compensate in part for the absence of tertiary interactions. The nature of these interactions could have a profound effect on folding trajectories, and knowledge of them can be of assistance in understanding both folding and protein design.

Determining three-dimensional (3D) structures of isolated structural elements is complicated by the fact that they are typically only marginally stable in aqueous solution. Even so, in cases where the apparent population of a folded hairpin is reasonably high, it has been possible to determine structures from NMR data. In these instances, regular hydrogen-bonded structures have generally been observed, resembling hairpins within globular domains. It should be noted, however, that the peptides for which such structures have been reported have all had novel, designed sequences and, therefore, no direct comparison with a structure adopted in a protein context has been possible.

The apex of the N-terminal hairpin of ubiquitin exemplifies one of the turn types that occur frequently in β -hairpins, a five residue “turn” comprised of a type I turn with a G1 β -bulge. The importance of this structural element in fixing the strand register has been illustrated by experiments which showed that its replacement by an alternative sequence could lead to a quite different, well-populated hairpin with a shifted strand register (Searle et al., 1995). The strand sequences do not obviously favor any particular register when pairwise interactions are considered, so this result led us to conclude previously that the turn sequence probably is important to the structural specificity of the hairpin in this instance.

The present study was to investigate the reasons for the stability of the observed native-like turn in the isolated peptide U(1–17): the strongest residue preference associated with the type I turn plus G1 β -bulge is for a Gly at the fourth position (residue 10 in the sequence), where the backbone torsion angle ϕ is positive, making accommodation of a side-chain difficult. However, in the context of the surrounding residues, Gly10 could also be compatible with alternative structures, in particular a type I' turn—also known to be a favored geometry in β -hairpins. The strong preference of type I' turns for a Gly at position *three* would suggest that formation of such a turn would result in a quite different hairpin with its strands displaced by one residue. This displacement produces a register analogous to that observed in the stable structure of the chimeric hairpin studied by Searle et al. (1995) (Fig. 1). In our earlier study of the U(1–17) peptide, we found no evidence for such a register shift.

Other sequence features were identified as potential drivers of the formation of the native turn. One clear candidate is the residue preceding the bulge residue, Thr9. Such a β -branched side chain would not be stereochemically favorable at this position in a type I' turn, where the ϕ torsion angle is positive, but is clearly acceptable in the type I plus G1 β -bulge turn of ubiquitin. Furthermore, Thr9 is the only residue in the turn sequence that deviates significantly from the consensus for a type I' turn: residues 7 to 11 in the

native sequence of ubiquitin are Thr-Leu-Thr-Gly-Lys, with Gly10 at position 4 in the turn, while the consensus for a type I' turn is Leu-{Asn or Asp, or Gly}-Gly-Lys with Gly10 at position 3 of the turn (Hutchinson & Thornton, 1994). Consequently, we have mutated Thr9 in the U(1–17) peptide to probe the contribution of this residue to the stability of the hairpin structure in isolation and implicitly in the folding pathway of full-length ubiquitin (Chen, 1998).

Previously (Zerella et al., 1999), we found that wild-type peptide U(1–17) has a relatively low apparent folded population in aqueous solution and gives NMR spectra with limited dispersion and weak nuclear Overhauser effect (NOE) intensities, making the peptide refractory to structural analysis. The current study finds that a peptide with Thr9 replaced by Asp, U(1–17)T9D, forms a hairpin in solution with a much increased apparent folded population. This has enabled us to make a full NMR structural analysis of the modified peptide and to examine the consequences of the change.

Results

¹H NMR spectra of U(1–17)T9D

U(1–17)T9D is readily soluble in water at pH values below about 4.0 and gave ¹H NMR spectra with greater dispersion than wild-type U(1–17). The spectra were assigned by the sequential method (Wüthrich, 1986).

Not only are the signals considerably more dispersed than those of unfolded peptides, and those of U(1–17), but there is clear evidence for a significant population of β -hairpin structure in the spectra of U(1–17)T9D. Figure 2 shows that the residues in the arms of the hairpin exhibit marked shifts of H_{α} resonances to low field of the values measured in peptides corresponding to the arms individually. The ³J_{HN α couplings for these residues, listed in Table 1, also have large values characteristic of β -structure. Furthermore, it appears that the register of the hairpin is native-like: the cross-strand H_{α} - H_{α} NOEs shown in Figure 3, for example, are strongly indicative of this. No cross-strand connectivities indicative of structures with alternative registers are observed. In addition, it was noted that the side-chain methyl resonances of Thr12 and Thr14 are shifted significantly to high field of the random coil position, suggestive of a close cross-strand interaction with Phe4, as is observed in the parent peptide U(1–17) (Zerella et al., 1999); no other side-chain resonances appeared to be significantly shifted through ring current effects that further supports the conclusion that only the native-like hairpin is significantly populated.}

The spectra of U(1–17)T9D were found to show a marked pH dependence in that the differences in chemical shifts and coupling constants from those anticipated for an unstructured state were consistently larger at pH 3.8 than at pH 2.0 (Fig. 2). This behavior is in contrast to that of the parent peptide, U(1–17), for which no significant pH dependence was detected over this pH range. It therefore seems likely that deprotonation of the Asp9 side chain, not present in the parent peptide, results in more favorable charge-charge interactions in the hairpin at the higher pH value.

Population of the β -hairpin

Comparison of the H_{α} chemical shifts and ³J_{HN α coupling constants for U(1–17)T9D and for the U(1–17) parent peptide shows that the shift deviations from random coil values are increased}

Fig. 1. **A:** β -hairpin conformation observed in solution for the peptide U(1-17)T9D showing interstrand hydrogen bond and relative orientation of the H_α atoms. **B:** Schematic drawing of possible alternative β -hairpin structure for the peptide U(1-17)T9D. No side chains are shown. The C-terminus of the peptide is protected by amidation.

throughout the peptide as a result of the mutation in the turn region, indicating that the hairpin is folded (Fig. 2; Table 1). As in our previous study (Zerella et al., 1999), the baseline values used for the unstructured state are those from the short peptides with the correct sequence for the individual arms of the peptide to account for the local effects of sequence in the unstructured state.

An accurate determination of the population of the folded hairpin is difficult. However, if we assume a two-state model (Zerella et al., 1999), we obtain estimates of $\sim 64\%$. This is significantly greater than the 18% estimated for the wild-type U(1-17). The greater population of the native-like hairpin in the U(1-17)T9D mutant, relative to native U(1-17), is striking.

Determination of the hairpin structure

The greater dispersion in the resonances of U(1-17)T9D compared to the parent peptide, as well as the increased intensity of medium- and long-range NOEs, made it possible to identify unambiguously a much more substantial list of distance restraints for

U(1-17)T9D than for the wild-type U(1-17). These comprised in total 142 intrasidue, 83 sequential, and 102 long-range connectivities. These data were used to calculate structures of U(1-17)T9D using dynamic simulated annealing (Brünger, 1992). No dihedral angle restraints were included, although the backbone $^3J_{HN\alpha}$ couplings had been accurately measured. The final structures, however, are consistent with these values.

Initially, all NOEs identified in the NOE spectroscopy (NOESY) spectra at pH 3.8 were interpreted as distance restraints. This approach effectively ignores the fact that the hairpin is not 100% populated. We expected that at least some short interproton distances, which are necessary for significant NOE intensity, would arise from the unfolded state. Structure calculations yielded an ensemble of 30 structures, all resembling hairpins with a native-like register, but the structures with no NOE violations typically contained several residues with very unusual backbone torsion angles. This lack of convergence to an acceptable solution suggested that the restraints derived from the NOE data were indeed not compatible with a single structure. Close study of the NMR

Fig. 2. Plot of the deviation of H_{α} chemical shifts ($\Delta\delta H_{\alpha}$) observed for U(1–17) and U(1–17)T9D from the values observed for peptides U(1–7) and U(11–17) corresponding to the two isolated arms of the hairpin (Zerella et al., 1999): U(1–17) in $H_2O:D_2O$ (9:1), pH 3.8 (solid bars); U(1–17)T9D in $H_2O:D_2O$ (9:1), pH 3.8 (clear bars); and U(1–17)T9D in $H_2O:D_2O$ (9:1), pH 2.0 (hatched bars).

data showed that the $^3J_{HN\alpha}$ data supported the presence of native-like β -strand ϕ/ψ angles for residues expected to be in the arms of the hairpin. Moreover, NOE violations consistently arose from a set of interactions between Ile3 and Phe4, and between Phe4 and Val5. In a normal β -strand, residues 3 and 5 would lie on one face of the structure, and residue 4 would lie on the other. Thus, these side chains would be too far apart to yield NOEs. On the other hand, the unfolded state might permit sequential interactions among these residues, particularly as they are all hydrophobic. Thus, it was concluded that the sequential side-chain NOEs observed in the NOESY spectra almost certainly reflect the presence of some unfolded peptide. Therefore, to explore only the folded state, the 12 distance restraints between side-chain protons of residues 3, 4, and 5 were eliminated. Subsequent structure calculations yielded an ensemble of 60 structures, of which 28 violated no NOE distance restraints greater than 0.5 Å. The program PROCHECK (Laskowski et al., 1993) was used to demonstrate that in 22 of the 28 structures, all ϕ/ψ angles lie in the “most favored” region of the Ramachandran map; in the remaining structures, there are few exceptions of ϕ/ψ angles in the “additionally allowed” region. For one structure in this set, the backbone angles of residue 11 lay in a “disallowed”

region; this structure also had very unusual backbone torsion angles for residues 8 and 11, which were, moreover, inconsistent with measured $^3J_{HN\alpha}$ couplings. Thus, the ensemble of converged structures was taken to be the 27 structures with no NOE violations and in agreement with the observed dihedral angle couplings. These 27 structures have been deposited in the Protein Data Bank, under code 1E0Q.

The hairpin structure

The U(1–17)T9D peptide forms a β -hairpin with native-like register in solution. The structure is well defined, having a backbone (C_{α}, N, C, O) root-mean-square deviation (RMSD) of 0.59 Å (Fig. 4). Considering only the β -strands (residues 1–7 and 11–17), the backbone RMSD is 0.4 Å; for the residues of the turn (residues 7–11), the backbone RMSD is 0.28 Å. Although the structure is relatively well defined, the NOEs upon which it is defined are themselves time-averaged parameters. The relatively small RMSDs are, therefore, not inconsistent with a structure that is more dynamic than is the case for a β -hairpin in a protein structure. Indeed, the longer hydrogen bonds observed in U(1–17)T9D compared to those found

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

Residue	$^3J_{\text{HN}\alpha}$ (Hz)			
	Random coil values ^a	U(1–7) and U(11–17) ^b	U(1–17)T9D in H ₂ O/D ₂ O ^c	U(1–17) in H ₂ O/D ₂ O ^c
Gln2	7.1	6.7	7.8	7.0 (6.9)
Ile3	7.6	7.8	9.6	8.6
Phe4	7.5	7.5	8.9	8.1 (8.1)
Val5	7.7	8.3	9.4	8.5
Lys6	7.1	6.7	8.5	7.0
Thr7	7.6	—	8.6	7.9
Lue8	7.1	—	5.9	6.4
Thr9	7.6	—	—	7.8
Gly10	—	—	7.2	—
Lys11	7.1	6.5	8.5	7.3
Thr12	7.6	7.5	8.7	7.7
Ile13	7.6	7.9	9.1	8.2
Thr14	7.6	7.9	9.2	8.2
Leu15	7.1	7.1	8.5	7.4 (7.5)
Glu16	6.8	6.9	8.2	7.4
Val17	7.7	7.7	8.8	8.0

^aValues derived from the COIL data set as listed in Smith et al. (1996).

^bValues were obtained from measurements on 1D spectra. The errors of all measurements are ± 0.1 Hz.

^cValues were obtained for the peptides at pH 3.8 using the method of Titman and Keeler (1990). Numbers in parenthesis were obtained from measurements on 1D spectra. The errors of all the measurements are ± 0.2 Hz.

in the corresponding element of the crystal of ubiquitin (Table 2) are consistent with more dynamic behavior of U(1–17)T9D.

The backbone conformation is well defined with the exception of residue 7. Five of the 27 converged structures diverge at this point from the consensus structure. However, the ϕ/ψ angles of these residues do not lie in disallowed regions of Ramachandran space, nor can they be excluded on the basis of measured $^3J_{\text{HN}\alpha}$ couplings.

Distances between heavy atoms and geometries of backbone atoms in the final ensemble are consistent with five hydrogen bonds in the peptide (Fig. 4). However, as there is unlikely to be any H_N exchange protection, hydrogen bonding constraints were

not included in the calculations. As a test, a calculation was performed with the addition of H-bonding constraints. The resulting structures were very similar to the final set and did not show any additional violations of NOE constraints. Thus, it is concluded that the pattern of hydrogen bonding in U(1–17)T9D is likely to be the same as that seen in the structure of full length ubiquitin.

In 22 of the 27 structures in the ensemble the ϕ/ψ angles of the residues at positions $i + 1$ and $i + 2$ in the turn (Leu8 and Asp9) most closely match those expected for the native-like type I turn (Hutchinson & Thornton, 1994), rather than matching those characteristic of any other type of β -turn. However, ϕ_{i+2} , which is that of the mutated residue Asp9, is consistently around -38° compared with the ideal value of 0° . The remaining five structures are those whose ϕ/ψ angles are not 100% in the most favored region of the Ramachandran map.

Side-chain conformations are well defined for several residues, notably Ile3, Phe4, Val5, Lys6, Thr12, and Thr14. Other residues are less well defined, particularly Thr7, Leu8, Ile13, and Leu15. The observed side-chain definition corresponds, as expected, to the number of restraints per residue. Interestingly, Asp9, which is the mutant residue in this peptide and seems to confer unusual stability on the structure, shows NOE interactions with the side chain of Lys11. Although there is no direct evidence for a charge–charge interaction between these two side chains, their proximity indicates that this is a possibility. The well-defined side chains are all involved in local packing interactions. Ile3 and Val5 are packed together, but while residues Ile13 and Leu15 are also in van der Waals contact with one another (and with Ile3 and Val5), there are insufficient NOE data resolved to define the packing interaction unambiguously. On the other face of the hairpin, the methyl group of Thr14 is packed against the ring of Phe4, and the methyl group of Thr12 makes contact with the side-chain methylene groups of Lys6 (Fig. 5). Interestingly, while the consensus structure is for the γ -methyl of Thr14 to pack against the ring face of Phe4, a few alternative structures show the Thr14 β -CH group facing the Phe4 ring. Although the NOE data do not discriminate between these possibilities, the packing of the methyl group against the face of the aromatic ring is also reflected by its high field chemical shift (1.06 ppm in U(1–17)T9D vs. 1.17 ppm in the peptide corresponding to the second arm of the hairpin, U(11–17)).

Discussion

Motivation for selection of the U(1–17)T9D mutation

Studies of protein folding have suggested that early formation of local structures may drive later folding processes. In this work, we have examined the structure of a peptide closely analogous to one thought to play such a role in the folding of ubiquitin. Earlier work on a peptide with the native sequence demonstrated that a significant degree of structure was present and that the structure was native-like in its register (Zerella et al., 1999). However, the wild-type peptide did not yield adequate NMR data to perform a full structural analysis. Other work on mutant peptides that demonstrated surprising conformational switching behavior (Searle et al., 1995) suggested that mutations in the loop sequence could modulate the stability of the peptide's folded structure. In an effort to explore these effects further, the U(1–17)T9D mutant was prepared and demonstrated to possess an exceptionally stable structure for a β -hairpin. An NMR structural analysis of the peptide is reported here. More recent work (P.-Y. Chen & P.A. Evans, un-

Table 2. $\text{N}\cdots\text{O}$ distances across the hairpin for groups putatively involved in hydrogen bonds

Hydrogen bonding	Hydrogen bonding donor	Average $\text{N}\cdots\text{O}$ distances in NMR converged structures (Å)	Distances in crystal structure of ubiquitin (Å)
Met1 O	Val17 N	4.0 ± 0.4	2.71
Leu15 O	Ile3 N	3.4 ± 0.2	2.95
Ile3 O	Leu15 N	3.13 ± 0.08	2.97
Ile13 O	Val5 N	3.7 ± 0.4	2.72
Val5 O	Ile13 N	3.8 ± 0.3	2.79
Lys11 O	Thr7 N	3.6 ± 0.2	2.95

Fig. 3. Region of a NOESY spectrum of a 1 mM solution of U(1-17)T9D in D₂O at pH 3.8 and 275 K. Some important NOEs are highlighted. Asterisk-labeled peaks are interstrand main chain-main chain NOEs.

publ. data) has shown that the T9D mutation in full length ubiquitin may actually enhance the rate of folding, showing that the mutation does not interfere with formation of the native structure, and suggesting that the enhanced stability of the mutant β -hairpin may help to drive folding.

The structure of U(1-17)T9D

The high quality of the spectra of the U(1-17)T9D peptide permitted a thorough structural analysis. The density of NMR constraints is similar to that observed in well-defined protein structures,

Fig. 4. Superposition of the backbone atoms of the converged 27 structures of U(1-17)T9D. The average backbone RMSD is 0.59 Å. The structures are colored by atom (oxygen in red, carbon in black, nitrogen in blue, and hydrogen in grey). Key residues mentioned in the text are numbered.

Fig. 5. A close up view of the side-chain interactions on the hydrophilic face of the β -hairpin for (A) the NMR structure of U(1–17)T9D and (B) the crystal structure of ubiquitin. **A:** In the NMR structure, note the interactions between the Phe4 ring and the γ -methyl of Thr14, and between Lys6 and the γ -methyl of Thr12. Note also the absence of an interaction between Gln2 and Phe4. **B:** In the crystal structure, note the different interactions involving the above side chains: Phe4 with Thr12, Thr14, and Gln2.

and the precision of the ensemble of calculated structures reflects this (Fig. 4).

The structure itself reveals several interesting features. First, the register of the hairpin is native-like. Furthermore, the geometry of the turn itself is also native-like, in that the ϕ/ψ angles most closely resemble the native type I turn rather than any other type. The major deviation from the native backbone conformation in the turn is the ψ angle of the mutated residue Asp9, which adopts a

value of approximately -40° rather than approximately 0° . This is perhaps not surprising, as the side chain of Asp9 will make different interactions with its neighbors than that of the β -branched Thr in the native sequence.

Since cooperativity is important in determining the overall effect of individual interactions on structural specificity and stability, it is impossible to quantify the individual contributions. However, hydrophobic packing seems to be important in determining the structural specificity displayed by the structured peptide and also in determining its thermodynamic stability. On both faces of the peptide, packing interactions were observed between residues lying on both strands of the hairpin. One set of these interactions was sufficiently well defined to be unambiguously non-native. The other set (involving residues Ile3, Val5, Ile13, and Leu15) was not well defined due to poor dispersion in the NMR spectra. This lack of dispersion may reflect a dynamic structure, but one would not necessarily expect a great variation in chemical shift of these amino acid types. Given that these residues form part of the hydrophobic core of the full-length protein, it is possible that they are indeed not well structured in the peptide. Another determinant of the hairpin structure appears to be hydrogen bonding. Although H-bonding restraints were not included in the structure calculations, the $N \cdots O$ distances across the backbone are clearly indicative of five native-like hydrogen bonds (Table 2). The pH dependence of the hairpin stability suggests that Asp9 must make some kind of stabilizing interaction, at least in its deprotonated state (see Fig. 2). In the full length protein, however, the side chain NH_3^+ of Lys11 makes a salt bridge with the side chain COO^- of Glu34. The NOEs between the side chain of Lys11–Asp9 in U(1–17)T9D and the new pH dependence would strongly suggest that the stabilizing effect of Asp9 is due to a charge–charge interaction between the NH_3^+ of Lys11 and the COO^- of Asp9.

Comparison of NMR and X-ray crystal structures

The converged NMR structures obtained for the U(1–17)T9D peptide are similar to that of the N-terminal hairpin in the crystal structure of the full length protein (Fig. 6). The average pairwise RMSD between the converged structures and the crystal structure is $1.59 \pm 0.2 \text{ \AA}$ for backbone atoms. As a further check on the similarity between the structures, we compared interproton distances in the crystal structure with the NMR distance restraints used in the structure calculations. A total of 16 pairs with distances differing by more than $0.5 \text{ \AA}$ were observed. Of the 13 such differences in the β -strands, 12 were side chain–side chain contacts

Fig. 6. Superposition of the backbone atoms of the converged 27 structures of U(1–17)T9D (grey) on the first 17 residues of the crystal structure of ubiquitin (black). The average pairwise RMSD between the converged structures and the crystal structure is $1.59 \text{ \AA}$ for backbone atoms.

involving residues 4, 6, 12, and 14, which we have already noted as appearing to adopt distinctly nonnative conformations in the peptide.

Within this small group of residues, there are some striking differences between the peptide in solution and the peptide in the full length protein. The Phe4–Thr14 and Lys6–Thr12 interactions, which are characteristic of the peptide structure, are quite different in the crystal structure (Fig. 5). In the peptide, the γ -methyl of Thr14 packs against the ring face of Phe4. Instead, in the crystal structure, the methyl packs against the ring edge, while the ring face of Phe4 is packed against the carboxamide group of Gln2 in a π - π interaction. In the peptide, the γ -methyl of Thr12 packs against the side chain of Lys6, but in the crystal structure, the methyl group is packed against the β -methylene group of Phe4, thus pointing the threonine hydroxyl group at the side chain of Lys6. As might be expected, the positions of the side chains in the NMR structures seem to be more favorable to solvation of peptide: the outward-pointing hydroxyl of Thr12 and carboxamide of Gln2 are consistent with greater solvation.

Conclusions

Structural analysis of the U(1–17)T9D peptide reveals a structure that is native-like overall, yet distinctly nonnative in some of its details. These conclusions, together with the observation that the U(1–17)T9D mutation, does not impede folding in the full length protein (P.-Y. Chen & P.A. Evans, unpubl. data), supporting the hypothesis that folding of the N-terminal 17 residues could play a role in the early folding of ubiquitin.

The origin of the stabilizing effect of the U(1–17)T9D mutation is uncertain. The pH dependence indicates that electrostatic interactions contribute to this effect when Asp9 is deprotonated. Although the exact interactions cannot be identified, it seems likely that charge–charge interaction with the side chain of Lys11 occurs, and this may be the prime cause of the peptide's greater stability in the structured form.

It is well known that only certain types of β -turn are favored in β -hairpins, most notably types I' and II', as well as some more extended turns, most notably a type I turn with a G1 β -bulge. These turns have distinct sequence preferences dictated by local stereochemical restrictions, and it might seem likely that ensuring compatibility with these could be an effective way of defining structural features of the hairpin formed. Indeed, the structure of the 1–17 peptide can be significantly altered by other changes to the loop sequence (Searle et al., 1995). In this case, however, the limited structural data available on the wild-type peptide are in close correspondence with the data on U(1–17)T9D; it appears that the two peptides are very similar.

Materials and methods

Peptide synthesis and purification

The peptide was synthesized (0.1 mmol scale) as described previously (Zerella et al., 1999), by the continuous flow solid phase Fmoc-polyamide method. NovaSyn KR resin was used on a NovaSyn Crystal synthesizer. The crude peptide was dissolved in 5% acetic acid and purified on a Waters Deltapak reverse-phase C₁₈ column using 0.1% TFA/H₂O and 0.1% TFA/acetonitrile as eluents. The final purified lyophilized products gave satisfactory

analysis by amino acid composition and positive ion electrospray mass spectrometry (VG BioQ).

Aggregation state of peptide

The effects of concentration on conformational properties of U(1–17)T9D were investigated. The peptide was soluble in D₂O up to a concentration of at least 3 mM in the pH range 2–4. The effects of varying the concentrations of the ubiquitin-derived peptide was then investigated using one-dimensional (1D) ¹H NMR spectra recorded from solutions at pH values of 2.0 and 3.8 and peptide concentrations in the range 0.1–3 mM. The chemical shifts and linewidths of the resonances, especially for those diagnostic of β -hairpin formation, were independent of concentrations over a wide range of temperature (273 to 294 K), strongly indicating that this peptide is monomeric under these conditions. No NMR spectra showed any time-dependent changes that might also have indicated intermolecular associations.

NMR spectroscopy

Samples were prepared by dissolving the peptide in 600 μ L of unbuffered 9:1 H₂O:D₂O or D₂O. The concentrations of all peptide samples were in the range of 1.0–2.0 mM. pH values were measured using a Corning 240 pH electrode and were adjusted by adding aliquots of D₂O solution of NaOD or DCl to a final pH of 3.8 or 2.0 and were not corrected for isotope and solvent effects. All spectra were referenced to the singlet resonance of internal dioxane (δ 3.74 ppm).

All spectra were recorded on Bruker DRX 500 and 600 and Varian Unity Plus 600 MHz spectrometers. For 1D experiments, 32k data points were collected and the spectra were processed using an exponential window function (typically 1–2 Hz) to improve the signal-to-noise ratio. Two-dimensional spectra were recorded with 2 or 4k complex data points in *f*₂ and 512 or 1k real points in *f*₁. The spectra were acquired in phase-sensitive mode using time proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983) to achieve quadrature detection in *f*₁. Spectral widths were 5,000 Hz. NOESY (Jeener et al., 1979), total correlation spectroscopy (TOCSY) (Braunschweiler & Ernst, 1983), and double-quantum filtered correlation spectroscopy (DQF-COSY) (Piantini et al., 1982) spectra were collected. For NOESY spectra, relatively long mixing times were typically used (τ_m = 300 ms) to allow the build-up of medium- and long-range NOEs. Spectra with shorter mixing times (τ_m = 100, 200 ms) were also recorded to avoid misassignment of resonances as a result of spin diffusion. TOCSY experiments were performed employing a MLEV17 sequence and a spin locking field of 8.0 kHz and a mixing time ranging between 60 and 100 ms. For samples in D₂O, suppression of the residual water signal was achieved by presaturation during the relaxation delay of 2.0 s and during the mixing time. For samples in 9:1 H₂O:D₂O, suppression of the intense solvent resonance was achieved by using the WATERGATE sequence (Piotto et al., 1992).

The method of Titman and Keeler (1990) was used to measure the homonuclear ³J_{HN α couplings. NOESY and DQ-correlation (Braunschweiler & Ernst, 1983) spectra were acquired with 4k $\times$ 1k data points and 80 scans per increment to ensure a good signal-to-noise ratio. Spectral widths of 5,000 and 9,000 Hz were used in the *t*₁ and *t*₂ dimensions of the DQ correlation. A final matrix of 4k $\times$ 2k data points was obtained. FELIX 2.3 (Biosym, San Diego,}

California) was used to process and analyze the data. The accuracy of the couplings obtained by this method was estimated to be within 0.2 Hz of the true value.

Population estimates

Population estimates were made on the basis of deviation of H_{α} chemical shift and $^3J_{HN\alpha}$ coupling constants from values for random coil. Values for random coil chemical shifts and $^3J_{HN\alpha}$ coupling constants were taken from short peptides of sequences corresponding to the arms of the β -hairpin, U(1–7) and U(11–17) (Zerella et al., 1999). The values for chemical shifts and coupling constants for the structured state are based on an average over a large number of proteins. Populations were calculated as in previous studies (Zerella et al., 1999).

Structure calculation

Structures were calculated from the experimental constraints using the program X-PLOR 3.8 (Brünger, 1992). Simulated annealing from random x, y, z coordinates was performed using derivatives of scripts random.inp, sa.inp, and refine.inp from the standard library. The parameters used were those from version 5.1 of Nilges' revised topallhdg.pro and parallhdg.pro with the parameter par-non bonded set to "PROLSQ" (Linge & Nilges, 1999).

Coordinates of the ensemble of 27 structures of U(1–17)T9D have been deposited in the Protein Data Bank (reference code 1E0Q).

Acknowledgments

We are grateful to Dr. Tom Frenkiel at the MRC, London, for helpful advice and access to NMR instrumentation; and to Dr. Len Packman at the Biochemistry Department, Cambridge University, for peptide synthesis. This work was supported by the BBSRC through the Cambridge Centre for Molecular Recognition (CCMR). R.Z. gratefully acknowledges the European Community for a Training and Mobility for Researcher Fellowship. P.-Y.C. was supported by a scholarship from the Ministry of Education, Taiwan, Republic of China. P.A.E. acknowledges the support of the Lister Institute of preventive medicine.

References

Blanco FJ, Rivas G, Serrano L. 1994. A short linear peptide that folds into a native stable β -hairpin in aqueous-solution. *Nature Struct Biol* 1:584–590.
 Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing—Application to proton correlation spectroscopy. *J Magn Reson* 53:521–528.
 Brünger AT. 1992. *X-PLOR. A system for X-ray crystallography and NMR*. New Haven, CT: Yale University Press.
 Chen P-YR. 1998. The role of the N-terminal β -hairpin in the folding of ubiquitin [PhD Thesis]. Cambridge, UK: University of Cambridge.
 Cox JPL, Evans PA, Packman LC, Williams DH, Woolfson DN. 1993. Dissecting the structure of a partially folded protein—Circular-dichroism and NMR studies of peptides from ubiquitin. *J Mol Biol* 234:483–492.
 De Alba E, Rico M, Jimenez MA. 1999. The turn sequence directs β -strand alignment in designed β -hairpins. *Protein Sci* 8:2234–2244.

Griffiths-Jones SR, Maynard AJ, Searle MS. 1999. Dissecting the stability of a β -hairpin peptide that folds in water: NMR and molecular dynamics analysis of the β -turn and β -strand contributions to folding. *J Mol Biol* 292:1051–1069.
 Griffiths-Jones SR, Sharman GJ, Maynard AJ, Searle MS. 1998. Modulation of intrinsic ϕ, ψ propensities of amino acids by neighbouring residues in the coil regions of protein structures: NMR analysis and dissection of a β -hairpin peptide. *J Mol Biol* 284:1597–1609.
 Harding MM, Williams DH, Woolfson DN. 1991. Characterization of a partially denatured state of a protein by 2-dimensional NMR—Reduction of the hydrophobic interactions 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.
 Hutchinson EG, Thornton JM. 1994. A revised set of potentials for β -turn formation in proteins. *Protein Sci* 3:2207–2216.
 Jeener J, Meier BH, Bachmann P, Ernst RR. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Phys Chem* 71:4546–4553.
 Khorasanizadeh S, Peters ID, Roder H. 1996. Evidence for a 3-state model of protein folding from kinetic-analysis of ubiquitin variants with altered core residues. *Nature Struct Biol* 3:193–205.
 Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
 Linge JP, Nilges M. 1999. Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation. *J Biomol NMR* 13:51–59.
 Marion D, Wüthrich K. 1983. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of H1-H1 spin-spin coupling-constants in proteins. *Biochem Biophys Res Com* 113:967–974.
 Muñoz V, Thompson PA, Hofrichter J, Eaton WA. 1997. Folding dynamics and mechanism of β -hairpin formation. *Nature* 390:196–199.
 Piantini U, Sorensen 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.
 Pitsyn OB. 1991. How does protein synthesis give rise to the 3D-structure. *FEBS Lett* 285:176–181.
 Ramirez-Alvarado M, Blanco FJ, Niemann H, Serrano L. 1997. Role of β -turn residues in β -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. *Nat Struct Biol* 3:604–612.
 Ramirez-Alvarado M, Kortemme T, Blanco FJ, Serrano L. 1999. β -hairpin and β -sheet formation in designed linear peptides. *Bioorg Med Chem* 7:93–103.
 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. *Nat Struct Biol* 2:999–1006.
 Smith CK, Regan L. 1995. Guidelines for protein design—The energetics of β -sheet side chain interactions. *Science* 270:980–982.
 Stockman BJ, Euvrard A, Shaill 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.
 Wouters MA, Curmi PMG. 1995. An analysis of side chain interactions and pair correlations within antiparallel β -sheets—The difference between backbone hydrogen-bonded and non-hydrogen-bonded residue pairs. *Proteins* 22:119–131.
 Wüthrich K. 1986. *NMR of protein and nucleic acids*. New York: John Wiley & Sons.
 Zerella R, Evans PA, Ionides JMC, Packman LC, Trotter BW, Mackay JP, Williams DH. 1999. Autonomous folding of a peptide corresponding to the N-terminal β -hairpin from ubiquitin. *Protein Sci* 8:1320–1331.