Protein Science (1997), 6:162–174. Cambridge University Press. Printed in the USA.

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

MARINA RAMÍREZ-ALVARADO, LUIS SERRANO, AND FRANCISCO J. BLANCO
European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

(RECEIVED July 19, 1996; ACCEPTED September 30, 1996)

Abstract

The solution conformation of three peptides corresponding to the two β-hairpins and the α-helix of the protein L B1 domain have been analyzed by circular dichroism (CD) and nuclear magnetic resonance spectroscopy (NMR). In aqueous solution, the three peptides show low populations of native and non-native locally folded structures, but no well-defined hairpin or helix structures are formed. In 30% aqueous trifluoroethanol (TFE), the peptide corresponding to the α-helix adopts a high populated helical conformation three residues longer than in the protein. The hairpin peptides aggregate in TFE, and no significant conformational change occurs in the NMR observable fraction of molecules. These results indicate that the helical peptide has a significant intrinsic tendency to adopt its native structure and that the hairpin sequences seem to be selected as non-helical. This suggests that these sequences favor the structure finally attained in the protein, but the contribution of the local interactions alone is not enough to drive the formation of a detectable population of native secondary structures. This pattern of secondary structure tendencies is different to those observed in two structurally related proteins: ubiquitin and the protein G B1 domain. The only common feature is a certain propensity of the helical segments to form the native structure. These results indicate that for a protein to fold, there is no need for large native-like secondary structure propensities, although a minimum tendency to avoid non-native structures and to favor native ones could be required.

Keywords: peptide structure; protein folding; protein G; protein L; secondary structure; ubiquitin

Experimental and theoretical analysis of protein folding suggest that there is an energy gap between the native and the other possible conformations of a protein (Anfinsen, 1973; Bryngelson et al., 1995), and that this is a necessary and sufficient condition for folding (Sali et al., 1994). The occurrence of this gap requires the set of interactions to be stronger in the native structure than in any of the possible non-native conformations. There are two types of interactions: local, which occur between amino acids close in the primary sequence, and non-local (Dill, 1990). Local interactions participate to define secondary structure, while non-local interactions are involved in defining the tertiary structure. A very important question refers to the roles of local and non-local interactions in protein folding and stability. An indirect way to attain some information relevant to this point comes from the conformational analysis of protein fragments. Short fragments of proteins allow the study of local interactions isolated from protein context (Dyson & Wright, 1993) and therefore indicate the importance of these interactions in determining protein secondary structure elements. If local interactions are important in protein folding and stability, it should be expected that isolated secondary structure elements will tend to populate native-like structures in the absence of tertiary contacts.

A large number of examples exist on helical protein fragments that are able to adopt a significant population of the native α-helix structure in water. The best example is the 13-residue C-peptide of ribonuclease A that forms up to 25% of the native helical structure in water (Brown & Klee, 1971). This case is, however, far from being general, since normally, native helices have marginal stability in short protein fragments (Jiménez et al., 1987, 1993; Dyson et al., 1988, 1992a; Segawa et al., 1991; Muñoz et al., 1995a,

Reprint requests to: Francisco J. Blanco, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany; e-mail: blanco@embl-heidelberg.de.

Abbreviations: NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; COSY, 2D scalar correlated spectroscopy; DQF-COSY, double quantum filtered 2D scalar correlated spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; ROESY, 2D rotating frame NOE spectroscopy; CD, circular dichroism; TFE, 2,2,2-trifluoroethanol; TSP, sodium 3-trimethylsilyl (2,2,3,3-$^2H_4$) propionate; ppm, parts per million; ppb, parts per billion.

1995b), although higher populations are found in peptides with protected chain termini (Kuroda, 1993; Waltho et al., 1993). Similar results are reported for β-turns (Blanco et al. 1991; Sönnichsen et al., 1992; Shin et al., 1993), although a pentapeptide from a long α-helix of influenza virus hemagglutinin shows a 40% population of a type II β-turn structure (Dyson et al., 1988). Hairpin structures have been characterized in modified protein fragments (Blanco et al., 1993; de Alba et al., 1995; Searle et al, 1995) or de novo designed peptides (de Alba et al., 1996; Ramírez-Alvarado et al., 1996), but protein fragments corresponding to β-hairpin structures have usually been found to be unstructured or aggregating (Dyson et al., 1992b; Viguera et al., 1996). A single and remarkable exception is the 16-residue fragment at the C-terminus of protein G immunoglobulin G (IgG) binding B1 domain. This fragment adopts the native hairpin structure with a population of 40% in water solution (Blanco et al., 1994a).

Comparison of the pattern of conformational tendencies along the sequence is another way to investigate the role of local interactions in protein folding and stability. In principle, if proteins with the same structure fold in a similar way and a particular secondary structure element is stable, then part of this folding nucleus on which the rest of the protein condenses would be conserved (Fersht, 1995) and this element should be expected to be stable in all these proteins. Alternatively, if structurally related proteins fold in different ways, they should not have a common pattern. A comparison of the helical propensities has been done with the isolated peptides spanning the helices of three structurally similar α/β parallel proteins with no evolutionary relationship (Muñoz et al., 1995a), as well as with two members of two evolutionary related families (Muñoz et al., 1995b). These studies found that specific α-helical propensity profiles are not conserved, indicating that a conserved secondary structure tendency is not required to determine their final three-dimensional structure. A similar comparison for a set of proteins with β-sheet structure that can be split into isolated β-hairpins is of interest.

The Protein G B1 domain contains two β-hairpins connected by an α-helix (Gronenborn et al., 1991). As shown in Figure 1, two other small proteins/domains that possess essentially the same fold have been found: ubiquitin (Vijay-Kumar et al., 1987), with a long insertion between the third and the fourth β-strands and 12% sequence identity with protein G (Kraulis, 1991a), and the immunoglobulin light chain-binding B1 domain of protein L from *Peptostreptococcus magnus* (Wikström et al., 1994). Other protein domains show structural similarities, but they have different connectivities between their secondary structure elements (Sauer-Eriksson et al., 1995), so they represent other fold types. Protein G and Protein L B1 domains have essentially the same structure in spite of a very low sequence identity (16%) and their different binding properties (they bind to different regions of IgG, although their binding surfaces are similar; Wikström et al., 1995).

Experimental analysis of the protein G B1 domain and ubiquitin fragments spanning their secondary structure elements indicate that the patterns of secondary structure propensities in these sequences have several features in common. The peptides corresponding to their α-helices adopt a native helical structure in alcohol/water mixtures, but in protein G the first hairpin adopts a high population of native structure in water, while in ubiquitin is the second hairpin, which shows an intrinsic tendency to form a native hairpin structure, although only in 60% methanol (Cox et al., 1993; Blanco & Serrano, 1995). The B1 domain of protein L constitutes an interesting target to characterize its secondary structure propensities when compared with the reported results of ubiquitin and protein G B1 domain. In this work we have addressed this point by analyzing the conformational properties of the peptides corresponding to all the secondary structure elements of the protein L B1 domain in different solvent conditions by CD and NMR.

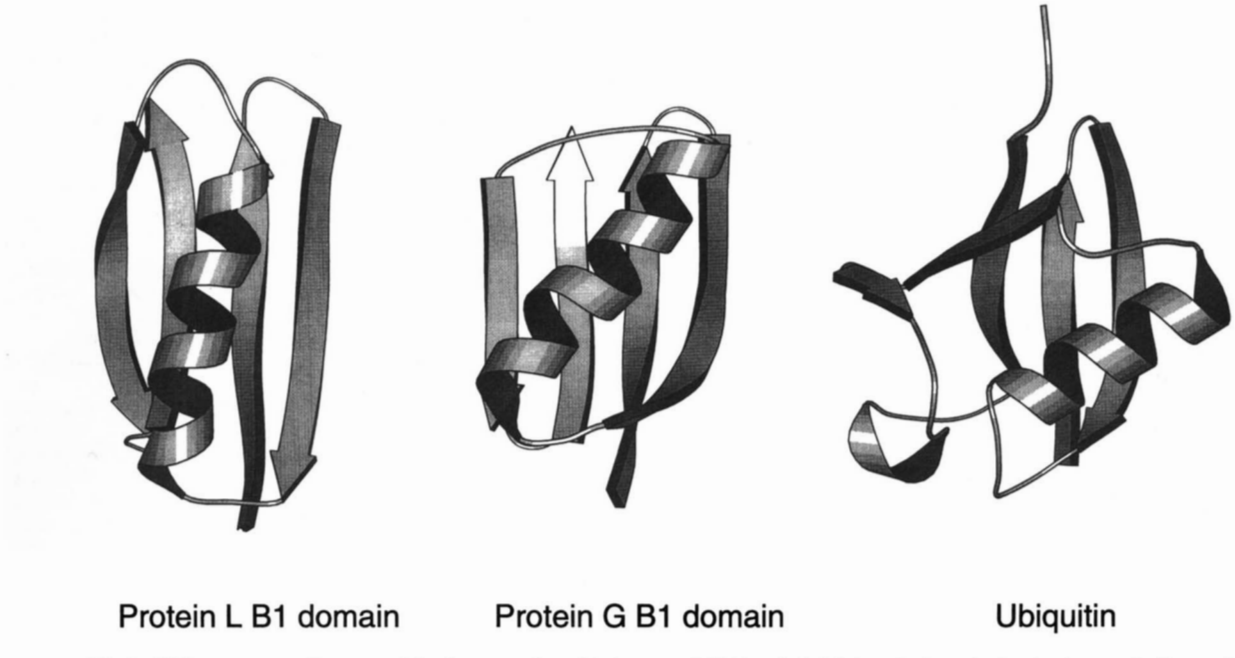

Fig. 1. Ribbon structure diagrams of the three proteins with the same fold. Protein L B1 domain (protein data bank entry 2ptl), protein G B1 domain (2gb1) and ubiquitin (1ubq) are at the left, middle and right position, respectively. The plots were generated with the program MOLSCRIPT (Kraulis, 1991b).

Results

Protein L B1 (95–115) fragment

This peptide was not soluble at pH 6.0 and it was finally studied at pH = 4.1 where it appeared to be monomeric, since the CD spectra at 500 and 10 μM were almost identical and neither chemical shifts nor line width changes were observed in 1D NMR spectra between 100 and 650 μM (data not shown). The far ultraviolet CD spectrum (Fig. 2A) was close to the one typical of a random coil conformation but with a minimum slightly red shifted (198 nm). The negative ellipticity observed at around 218 nm in aqueous solution was close to zero in 6 M urea, indicating the presence of a small population of non-random structures in pure water. A conformational change to a more ordered structure was induced in the presence of 30% trifluoroethanol. The maximum at 190 nm and a minimum at 208 nm together with a broad minimum at 219 nm suggested the presence of a mixture of helix and/or β-sheet in equilibrium with random coil conformations. The CD spectrum in TFE did not significantly changed from 350 to 12 μM peptide concentration.

A summary of the NMR analysis of this peptide is in Figure 2B and C, and the assignment of the ^{1}H resonances is in Table 1. Strong sequential $d_{\alpha N}(i, i + 1)$ and weak $d_{NN}(i, i + 1)$ NOEs were observed at most positions along the peptide sequence, and the conformational shifts of the C_α protons are small, both patterns being typical of mainly random coil peptides. However, three non-sequential NOEs showed the presence of locally folded conformations. A $d_{\alpha N}(i, i + 2)$ observed between N105 and S107 indicated the formation of a turn-like structure in the central region of the peptide. Two other $(i, i + 2)$ NOEs were observed between the side chains of A111 and F113 and between L101 and F103 (Fig. 3A). The intensity of these NOES was very weak so the population of the local structures around these residues is very low. In the presence of 6 M urea the three non-sequential NOEs were not observed. Another non-random local structure was detected at the C-terminal end of the peptide by the upfield shift of the amide protein chemical shift of G115 (7.84 ppm) with respect to the random coil value (8.66 ppm), and its low temperature coefficient (−3.4 ppb as compared with −7.0 in the pentapeptide GGGGG, Merutka et al., 1995). These values are nearly the same in 6 M urea (7.86 ppm, −2.6 ppb/K). These non-random values could be due to an interaction between the aromatic side chain of Phe 113 and the amide proton of G115, as previously described in other peptides (Kemmink et al., 1993; Kemmink & Creighton, 1995a).

In the presence of 30% TFE, the signals were much broader than in water. This could be due to a major conformational change to a more compact and rigid structure or chemical exchange between random and folded structures, as suggested by the CD spectrum. However, the chemical shifts of most of the signals were very similar (Fig. 2B). The broad lines were likely due to an oligomerization equilibrium that rendered the high molecular weight aggregates not observable and the low molecular weight species (including the monomeric ones) with broad signals, which further sharpened by increasing the temperature. In these conditions, the interpretation of NOE data cannot be conclusive, as these could be transferred NOEs coming from the aggregated molecules (Dyson & Wright, 1991). From the chemical shift data, we can conclude that TFE is not producing any major conformational change in the monomeric peptide. In the aggregated species some secondary structure is induced by TFE, as measured in the CD spectrum. The chemical shift and the temperature coefficient of the G115 amide proton is almost the same as in water; hence, the interaction between F113–G115 seems to be present to the same extent in any of the three conditions.

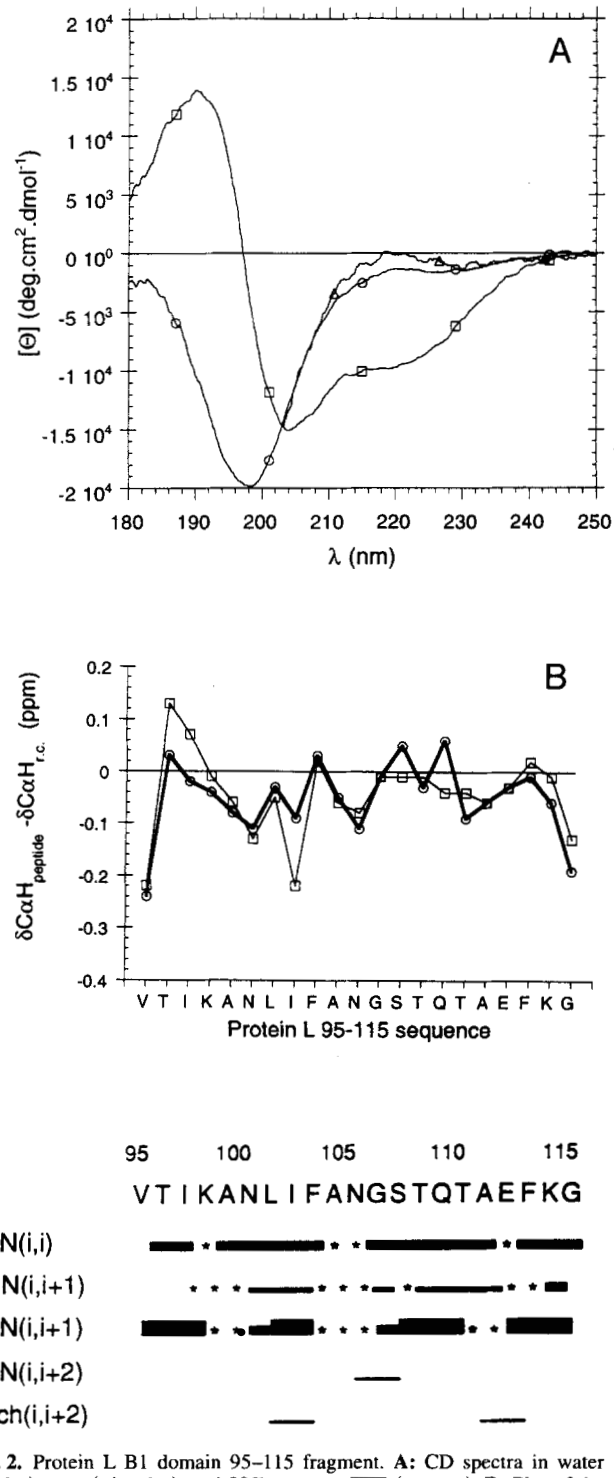

Fig. 2. Protein L B1 domain 95–115 fragment. **A:** CD spectra in water (circles), urea (triangles), and 30% aqueous TFE (squares). **B:** Plot of the conformational shifts of C_α protons in water (circles) and 30% TFE (squares). **C:** Pattern of the NOEs observed in NOESY and ROESY spectra in water. NOEs involving side-chain protons are grouped as dsch. The asterisks indicate NOEs that could not be observed, if present, due to resonance overlapping. All data were obtained at pH = 4.1, 278 K.

Table 1. *Proton chemical shifts (in ppm) and amide signal temperature coefficients ($\Delta\delta/\Delta T$, in ppb/K) of protein L B1 domain (95–115) fragment*[a]

Residue	NH	CαH	CβH	CγH	CδH	Others	$\Delta\delta/\Delta T$
Val 95	—	3.89	2.22	1.02			—
	—	(3.91)	(2.28)	(1.08)			—
Thr 96	8.83	4.42	4.08	1.21			−7.7
	(8.61)	(4.52)	(4.22)	(1.28)			(−5.8)
Ile 97	8.68	4.16	1.85	1.20, 1.49	0.86	CγH$_3$ 0.86	−9.9
	(8.48)	(4.25)	(1.92)	(1.52)	(0.96)	(0.96)	(−9.5)
Lys 98	8.61	4.28	1.82	1.43, 1.45	1.69	CεH 3.00	−8.9
	(8.24)	(4.31)	(1.91)	(1.52)	(1.80)	(3.04)	(−7.1)
						NεH 7.62	
						(7.72)	
Ala 99	8.52	4.26	1.38				−7.4
	(8.24)	(4.28)	(1.50)				(−6.8)
Asn 100	8.57	4.65	2.74, 2.81			NγH 7.70, 7.00	−7.7
	(8.26)	(4.63)	(2.88)			(7.58, 6.83)	(−3.7)
Leu 101	8.29	4.32	1.61	1.46	0.94, 0.85		−8.6
	(8.01)	(4.30)	(1.7, 1.8)	(1.58)	(0.96)		(−5.9)
Ile 102	8.09	4.09	1.82	1.38, 1.12	0.78	CγH$_3$ 0.82	−7.7
	(7.86)	(3.96)	(1.80)	(1.50, 1.12)	(0.76)	(0.80)	(−7.7)
Phe 103	8.48	4.65	2.96, 3.14		7.24	CεH 7.33	−9.9
	(8.11)	(4.60)	(3.10, 3.26)		(7.30)	(n.a.)	(−10.2)
						CζH 7.23	
						(n.a.)	
Ala 104	8.47	4.29	1.38				−7.7
	(8.24)	(4.28)	(1.50)				(−9.3)
Asn 105	8.52	4.65	2.84			NγH 7.75, 7.04	−6.1
	(8.34)	(4.68)	(2.90)			(7.72, 6.93)	(−7.1)
Gly 106	8.52	4.00					−6.8
	(8.49)	(4.00)					(−7.1)
Ser 107	8.32	4.54	3.91				−5.5
	(8.30)	(4.48)	(3.91, 4.00)				(5.5)
Thr 108	8.38	4.36	4.26	1.21			−7.4
	(8.20)	(4.38)	(4.32)	(1.26)			(−4.9)
Gln 109	8.51	4.42	1.98, 1.21	2.34		NεH 7.60, 6.93	−5.8
	(8.38)	(4.32)	(2.10, 2.20)	(2.40)		(7.52, 6.83)	(−4.9)
Thr 110	8.32	4.30	4.25	1.24			−7.7
	(8.12)	(4.35)	(4.31)	(1.30)			(−5.9)
Ala 111	8.50	4.28	1.38				−8.6
	(8.26)	(4.28)	(1.46)				(−10.2)
Glu 112	8.38	4.26	1.94	2.26, 2.36			−7.4
	(8.22)	(4.26)	(2.02)	(2.32, 2.41)			(−6.5)
Phe 113	8.40	4.61	3.10		7.27	CεH 7.34	(−8.3)
	(8.14)	(4.64)	(3.15, 3.22)		(7.32)	(n.a.)	(−7.1)
						CζH 7.32	
						(n.a.)	
Lys 114	8.47	4.26	1.82	1.38	1.66	CεH 2.98	−8.3
	(8.26)	(4.36)	(1.80, 1.90)	(1.48, 1.52)	(1.74)	(3.04)	(−6.8)
						NεH 7.62	
						(7.72)	
Gly 115	7.84	3.82					−3.4
	(7.73)	(3.88)					(−3.4)
NH_2	7.22, 7.54						
	(7.15, 7.47)						

[a]Chemical shifts are referenced to TSP. Conditions: 278 K, pH = 4.12, H_2O 10% ^2H_2O (by vol.) or H_2O 30% perdeuterated TFE (in parentheses). An em dash indicates protons that exchange with solvent too fast to be observed. n.a. means not assigned.

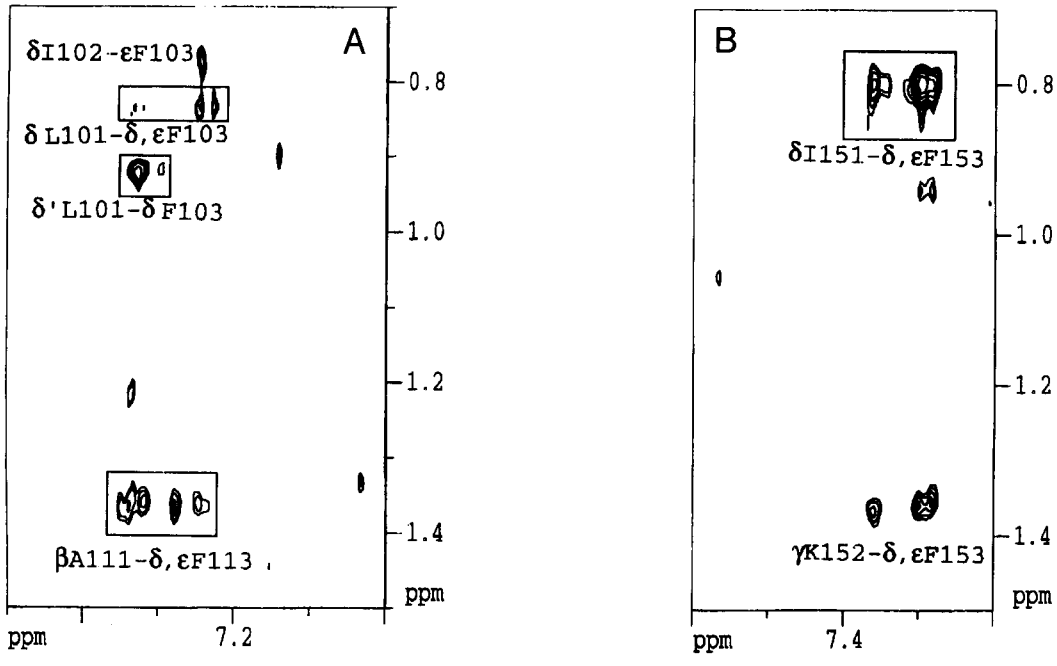

Fig. 3. Regions of 500 MHz ROSESY spectra of protein L B1 domain (95-115) and (136-155D) fragments. **A:** ROESY region corresponding to 95-115 fragment; **B:** ROESY region corresponding to 136-155D fragment. Nonsequential NOES are boxed.

Protein L B1 (114–138) fragment

This peptide was insoluble at pH above 3.0. At this pH, high concentration peptide samples could be prepared (5 mM) but the broad line widths of the NMR signals clearly indicated that the peptide was aggregating. At pH = 2.4 and 500 μM peptide concentration, the line widths were sharper, although still a bit broader than in a fivefold diluted sample at the same pH. This meant that aggregation was still taking place, although it was less severe. These conditions were finally selected for the NMR experiments. The CD spectra in water and TFE at pH = 2.4 were concentration independent over the range 600–12 μM, although the peptide slightly aggregated in water at 500 μM, as seen by the NMR signal line widths, which are a more sensible indicator of self-association than changes in the CD bands. The far-ultraviolet CD spectra of this peptide are shown in Figure 4A. In aqueous solution, the spectrum had a minimum of 196, typical from random coil conformations. There was, however, negative ellipticity at 222 nm that became positive in the presence of 6 M urea. The possible helical population calculated from this value with the method of Chen et al. (1974) is ~4%. The CD spectrum in 30% trifluoroethanol was a typical mixture of α-helix and random coil. The helical population calculated by the method of Chen et al. (1974) from the ellipticity at 222 nm is 42%.

The NMR assignment for this peptide is given in Table 2. The pattern of sequential and intraresidue NOEs between backbone protons indicated that the peptide was mainly a random coil in water, although a few weak non-sequential NOES were also found in the NOESY spectrum (data not shown). However, the peptide aggregation and the absence of these NOEs in the ROESY spectrum preclude any definitive conclusion about the conformational state of the peptide in the monomeric state. In addition, the conformational shifts profile could be more distorted than usual by the presence of the aromatic side chains of the two tyrosines at positions 125 and 127. In fact in the presence of 6 M urea, the chemical shifts of this region barely changed, suggesting that the negative conformational shifts were mainly due to the aromatic rings.

In 30% TFE, several $d_{\alpha N}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ appeared in the region T116 to D134 (Fig. 4C), and together with the conformational shifts of C_α protons (negative from T116 to K132; Fig. 4B), indicated the presence of an α-helical conformation spanning these residues. The data clearly indicated that the helix is more populated from A124 to D134, as the values of the conformational shifts were more negative (Williamson, 1990).

Protein L B1(136–155D) fragment

The peptide corresponding to the second β-hairpin of protein L was rather insoluble at acidic pH values. Solubility was high at pH = 6.1 and both CD and 1D NMR spectra were concentration independent at this pH. This peptide was conformationally very similar to the one corresponding to the first β-hairpin. The far-ultraviolet CD spectrum in water is close to the one typical of a random coil conformation with a minimum at 196 nm (Fig. 5A). The addition of urea resulted in a small change of the spectrum shifting the negative ellipticity around 220 nm toward zero or positive values. In the presence of 30% trifluoroethanol, the spectrum has a small maximum at 190 nm and a minimum around 204 nm. The shape of this spectrum changed with the sample concentration, being the ellipticity around 214 nm less negative in the diluted sample (Fig. 5A). This result indicates that the peptide adopts a folded conformation in TFE (possibly in a β-sheet structure), but in an oligomeric state. At low concentrations the population of random coil conformations increased, probably due to the increase of the proportion of monomers that remained basically unfolded.

The NMR analysis in water showed a pattern of backbone intraresidue, sequential NOEs, and a conformational shifts profile

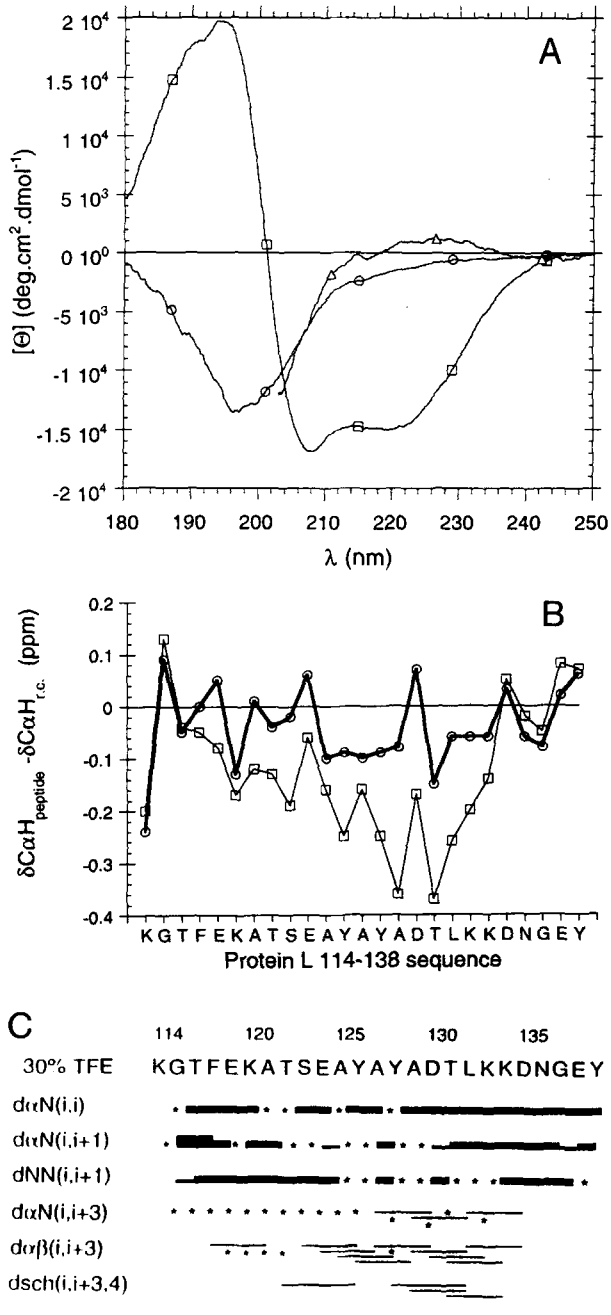

Fig. 4. Protein L B1 domain 114–13B fragment. **A:** CD spectra in water (circles), urea (triangles), and 30% aqueous TFE (squares). **B:** Plot of the conformational shifts of C_α protons in water (circles), 30% TFE (squares). **C:** Pattern of the NOEs connectivities measured in NOESY spectrum in 30% TFE. NOEs involving side-chain protons other than $C_\beta H$ are grouped as dsch. The asterisks indicate NOEs that could not be observed, if present, due to resonance overlapping. All data were obtained at pH = 2.4, 278 K.

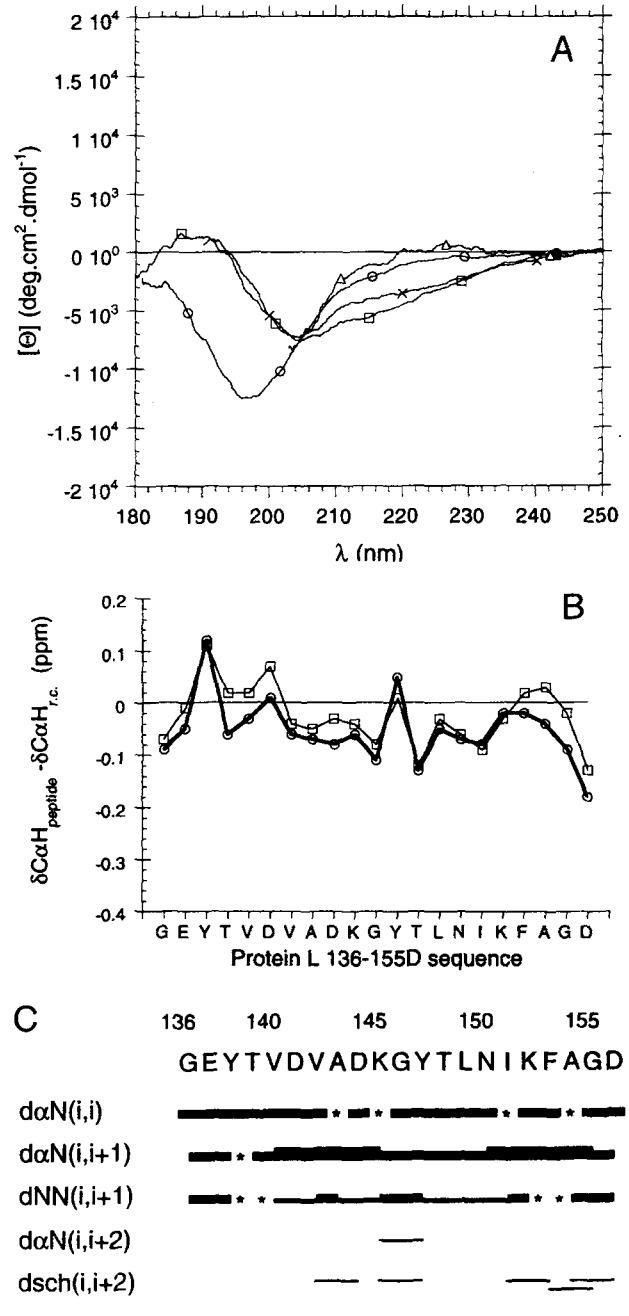

Fig. 5. Protein L B1 domain 136–155D fragment. **A:** CD spectra in water (circles), urea (triangles), and 30% aqueous TFE (squares). **B:** Plot of the conformational shifts of C_α protons in water (circles), 30% TFE (squares). **C:** Pattern of the NOEs observed in NOESY and ROESY spectra in water. NOEs involving side-chain protons are grouped as dsch. The asterisks indicates NOEs that could not be observed if present due to resonance overlapping. All data obtained at pH = 6.1, 278 K except the chemical shifts in 30% TFE, measured at 295 K.

typical of mainly random coil peptides (Fig. 5B, C). Several weak non-sequential NOEs were indicative of the presence of a low population of locally folded structures. Two NOEs between K145 and Y147, including a $d_{\alpha N}(i, i + 2)$ NOE, indicates the existence of a turn-like conformation in this region, the same where the native turn exists. Other observed NOEs involving the side chains of residues V142–D144, I151–F153, and A154–D156 also suggested that some preferred conformations that were more extended than the average in the random coil could be present in these regions. The chemical shift of the amide proton of G155 (7.86 ppm) is upfield shifted with respect to the random coil value and its temperature coefficient is also lower (−1.7 ppb/K). This suggests the presence of an interaction between F153 and G155 similar to the one found in the C-terminal end

of the 95–115 fragment. The NOE between the side chain of F153 and the $C_\alpha H$ of G155 supports this conclusion. This interaction is destabilized but still present in 6 M urea since the chemical shift and the temperature coefficient of G155 (7.90 ppm and −3.1 ppb/K, respectively) remained far away from random coil values. The non-sequential NOEs found in water were not present in 6 M urea.

Addition of 30% TFE at 278 K produced a severe line broadening, but no apparent peptide precipitation occurred. Two-dimensional spectra were devoid of crosspeaks and only a weak diagonal appeared. We performed several 1D spectra at different TFE concentrations (0, 5, 10, 12.5, 20, and 25% by vol.). Only the chemical shifts of exchangeable protons moved during the titration, others remaining almost invariant. The signals got slightly broadened up to a 20% TFE concentration and then at 25% a large increase in the line widths occurred, similar to that at 30% TFE. These results indicated that the structure of the peptide was not significantly changing, but an aggregation process appeared at high TFE concentrations. TOCSY spectra were successfully recorded in 30% TFE at 295 K. A NOESY spectrum was still not useful, but as the chemical shifts of C_α protons were in general very similar to those in water (Table 3), the assignment could be done with the help of a TOCSY spectrum recorded in water at the same temperature and another TOCSY spectrum in 30% TFE at 310 K.

Intrinsic secondary structure propensity prediction

Figure 6 shows the predicted helical and extended behavior of the three fragments of protein L B1 domain. For the first hairpin, a mixed tendency to be helical and extended was found all along the sequence except for the turn region. Both tendencies were not very marked, with the exception of the first four residues in which the propensity of being extended is high. However, the tendency for helix formation was always lower than the tendency for β-strand. For the protein L 114-138 fragment a mixed secondary structure propensity was also predicted, with an irregular distribution in the β-strand tendency. In this peptide, however, the tendency for the adoption of a helical structure was larger than the β-strand tendency, in the region 122–134. For the first eight residues the β-strand tendency predominated and the helical tendency was low. The third peptide showed a negligible propensity for the helical structure and a large tendency to be extended for the residues that form the two strands in the hairpin of protein L.

Discussion

Conformation of the isolated protein L fragments

The protein L fragments corresponding to the two β-hairpins have similar conformational properties. Locally folded preferred structures are detected in water solution both by CD and NMR, although their population is very low. There are both native and non-native local interactions. Both peptides show a preference for turn conformations at their central regions, the same where the native turns connecting the antiparallel β-sheet strands exist in the protein structure. NOEs of the type $d_{\alpha N}(i, i + 2)$ are found at residues N105–S107 and K145–Y147. In protein L residues F103–S107 form a five residue turn and residues D144–Y147 form a type I′ β-turn. In the second β-hairpin peptide an additional NOE is observed between K145 and Y147 side chains that are also close in the protein L B1 domain structure. It seems that the tendency of the chain to bend at these points is determined by the local interactions as it is maintained in the isolated peptides. Other preferred local conformations involve i, $i + 2$ interactions between apolar and aromatic residues. This kind of interaction is relatively frequent in peptides corresponding to protein β-strands (Dyson et al., 1992b; Kemmink & Creighton, 1993) and could indicate the preference for more extended conformations than the average in the random coil. In the protein L structure, the residue pairs L101–F103, A111–F113, and I151–F153 are in close contact, pointing to the same side of the β-sheet. These interactions are detected in the peptides in water, and then could contribute to the final formation of the β-sheet in the whole protein. The interaction between V142 and D147 side chains is not present in the protein L. At the C-terminal end of both peptides another local folded conformation is manifested by the anomalous values of the amide proton chemical shifts and temperature coefficients of the amide protons of G115 and G155, respectively. This is due to an interaction of the amide protons of the glycines with the aromatic side chains of F113 and F154. This type of interaction occurs in peptides with the sequence X-Z-G, where X is an aromatic residue, Z is any amino acid except proline and G is glycine, and has been analyzed in detail in peptides derived from a fragment of bovine pancreatic

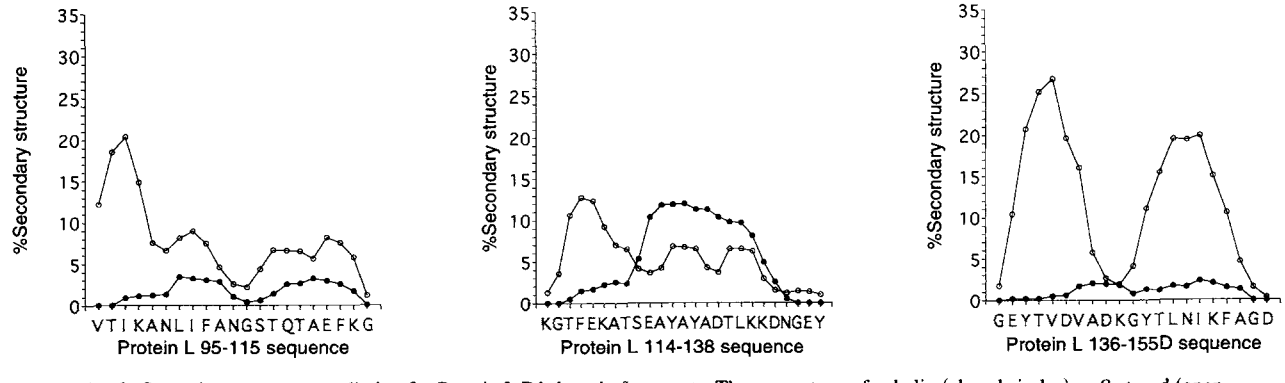

Fig. 6. Secondary structure prediction for Protein L B1 domain fragments. The percentage of α-helix (closed circles) or β-strand (open circles) per residue, calculated with the programs AGADIR and AMPURIA are plotted versus the peptide's sequence for protein L 95–115 (left), protein L 114–137 (middle), and protein L 136–155D (right) fragments.

Table 2. *Proton chemical shifts (in ppm) and amide signal temperature coefficients ($\Delta\delta/\Delta T$, in ppb/K) of protein L B1 domain (114–137) fragment*[a]

Residue	NH	CαH	CβH	CγH	CδH	Others	$\Delta\delta/\Delta T$
Lys 114	— (—)	4.08 (4.12)	1.74, 1.78 (2.00)	1.38 (1.52)	1.74 (1.76)	CεH 3.03 (3.05) NεH 7.66 (7.70)	— (—)
Gly 115	8.90 (8.86)	4.10 (4.14)					(−7.1) (−5.6)
Thr 116	8.36 (8.18)	4.34 (4.35)	4.14 (4.25)	1.18 (1.24)			−8.04 (−5.0)
Phe 117	8.55 (8.54)	4.62 (4.57)	2.86 (3.14)		7.23 (7.26)	CεH 7.34 (7.36) CζH 7.33 (7.34)	−8.4 (−10.2)
Glu 118	8.32 (8.38)	4.34 (4.21)	1.87, 2.00 (2.05, 2.11)	2.40 (2.49)			−5.0 (−6.2)
Lys 119	8.48 (8.30)	4.19 (4.15)	1.74, 1.82 (1.84, 1.92)	1.48 (1.50)	1.72 (1.82)	CεH 3.02 (2.98) NεH 7.64 (7.64)	−8.7 (−6.2)
Ala 120	8.59 (8.42)	4.35 (4.22)	1.41 (1.48)				−9.0 (−9.9)
Thr 121	8.31 (8.00)	4.35 (4.26)	4.28 (4.36)	1.22 (1.20)			−8.4 (−4.0)
Ser 122	8.43 (8.16)	4.47 (4.30)	3.88, 3.96 (4.02, 4.07)				−7.4 (−3.1)
Glu 123	8.52 (8.31)	4.35 (4.23)	1.94, 204 (2.16)	2.44 (2.50, 2.62)			−8.04 (−5.3)
Ala 124	8.35 (8.11)	4.24 (4.18)	1.32 (1.49)				−6.5 (−4.0)
Tyr 125	8.17 (8.12)	4.47 (4.31)	2.95 (3.13)		7.07 (7.17)	CεH 6.80 (6.83)	−11.8 (−9.9)
Ala 126	8.13 (8.07)	4.24 (4.18)	1.28 (1.54)				−5.9 (−7.1)
Tyr 127	8.17 (8.39)	4.47 (4.31)	3.02 (3.16)		7.12 (7.14)	CεH 6.82 (6.82)	−9.0 (−12.1)
Ala 128	8.24 (8.38)	4.26 (3.98)	1.38 (1.52)				−6.8 (−5.3)
Asp 129	8.42 (8.50)	4.70 (4.46)	2.86, 2.94 (2.85, 2.92)				−6.2 (−7.7)
Thr 130	8.11 (7.99)	4.24 (4.02)	4.24 (4.34)	1.22 (1.28)			−5.6 (−3.7)
Leu 131	8.11 (7.94)	4.29 (4.09)	1.58, 1.66 (1.72)	1.50 (1.50)	0.84, 0.90 (0.80)		−5.6 (−4.3)
Lys 132	8.24 (7.98)	4.26 (4.12)	1.74, 1.82 (1.90)	1.39, 1.45 (1.44, 1.56)	1.66 (1.70)	CεH 2.99 (2.99) NεH 7.60 (7.64)	−6.2 (−4.3)
Lys 133	8.41 (8.06)	4.26 (4.18)	1.78 (1.93)	1.39, 1.45 (1.46, 1.54)	1.67 (1.74)	CεH 2.99 (3.02) NεH 7.60 (7.64)	−7.7 (−3.7)
Asp 134	8.62 (8.46)	4.66 (4.68)	2.86, 2.96 (2.98)				−6.8 (−5.6)
Asn 135	8.51 (8.16)	4.70 (4.74)	2.86 (2.86, 2.96)			NγH 7.70, 6.96 (7.64, 6.84)	−5.9 (−0.3)
Gly 136	8.42 (8.26)	4.00, 3.90 (3.92, 4.01)					−5.9 (−2.5)
Glu 137	8.08 (8.03)	4.31 (4.37)	1.89, 1.96 (1.94, 2.04)	2.26, 2.32 (2.38)			−4.6 (−3.2)
Tyr 138	8.35 (8.08)	4.62 (4.63)	2.93, 3.18 (2.99, 3.14)		7.12 (7.14)	CεH 6.82 (6.86)	−8.3 (−5.6)

[a]Chemical shifts are referenced to TSP. Conditions: 278 K, pH = 2.4, H_2O 10% ^2H_2O (by vol.) or H_2O 30% perdeuterated TFE (in parentheses). Am em dash indicates protons that exchange with solvent too fast to be observed. n.a. means not assigned.

Table 3. *Proton chemical shifts (in ppm) and amide signal temperature coefficients ($\Delta\delta/\Delta T$, in ppb/K) of protein L B1 domain (136–155D) fragment*[a]

Residue	NH	CαH	CβH	CγH	CδH	Others	$\Delta\delta/\Delta T$
Acetyl		2.09 (2.10)					
Gly 136	8.40 (8.16)	3.92 (3.94)					−8 (−8.0)
Glu 137	8.56 (8.51)	4.24 (4.28)	1.93, 1.86 (1.93, 1.88)	2.16 (2.18)			−6.3 (−7.5)
Tyr 138	8.42 (8.12)	4.68 (4.67)	3.12, 2.95 (3.14, 3.00)		7.13 (7.14)	CεH 6.84 (6.86)	−7.7 (−6.2)
Thr 139	8.17 (7.90)	4.33 (4.41)	4.33 (7.88)	1.17 (1.20)			−5.6 (−5.0)
Val 140	8.27 (7.97)	4.10 (4.15)	2.13 (2.11)	0.96 (0.96)			−8.2 (−8.1)
Asp 141	8.58 (8.35)	4.64 (4.70)	2.61, 2.76 (2.64, 2.78)				−7.6 (−6.2)
Val 142	8.27 (7.96)	4.07 (4.09)	2.09 (2.18)	0.94 (0.96)			−8.8 (−6.2)
Ala 143	8.47 (8.20)	4.27 (4.29)	1.41 (1.44)				−6.8 (−5.0)
Asp 144	8.34 (8.12)	4.55 (4.60)	2.70, 2.77 (2.78)				−6.8 (−4.3)
Lys 145	8.48 (8.03)	4.26 (4.31)	1.74, 1.85 (1.72)	1.44, 1.48 (1.47, 1.52)	1.74 (1.72)	CεH 3.03 (3.04) NεH 7.64 (—)	−9.3 (−6.8)
Gly 146	8.60 (8.43)	3.90 (3.93)					−6.7 (−5.6)
Tyr 147	8.10 (7.88)	4.61 (4.57)	3.07 (3.10)		7.12 (7.14)	CεH 6.83 (6.86)	−5.9 (−5.0)
Thr 148	8.20 (7.92)	4.26 (4.27)	4.26 (4.32)	1.17 (1.20)			−5.9 (−5.6)
Leu 149	8.24 (7.91)	4.30 (4.32)	1.63 (1.67)	1.63 (1.67)	0.92, 0.98 (0.96)		−7.7 (−5.6)
Asn 150	8.53 (8.16)	4.69 (4.70)	2.75, 2.85 (2.78, 2.84)			NγH 7.72, 7.01 (6.79, 7.50)	−6.5 (−5.0)
Ile 151	8.18 (7.78)	4.10 (4.09)	1.83 (1.84)	1.14, 1.44 (1.14, 1.45)	0.88 (0.86)	$C\gamma H_3$ 0.82 (0.82)	−9.3 (−5.0)
Lys 152	8.43 (8.02)	4.30 (4.29)	1.71 (1.72)	1.34, 1.40 (1.30, 1.36)	1.68 (1.64)	CεH 3.00 (2.98) NεH 7.64 (—)	−9.1 (−7.5)
Phe 153	8.56 (8.16)	4.60 (4.64)	3.04, 3.12 (3.04, 3.22)		7.32 (n.a.)	CεH 7.36 (n.a.) CζH 7.34 (n.a.)	−9.9 (−8.7)
Ala 154	8.44 (8.02)	4.30 (4.37)	1.39 (1.42)				−7.4 (−6.2)
Gly 155	7.88 (7.89)	3.92 (3.99)					−1.7 (−5.6)
Asp 156	8.04 (7.8)	4.45 (4.50)	2.64, 2.73 (2.66, 2.74)				−5.9 (−5.6)

[a]Chemical shifts are referenced to TSP. Conditions: 278 K, pH = 6.1, H_2O 10% ^2H_2O (by vol.) or H_2O 30% TFE at 295 K (in parentheses). An em dash indicates protons that exchange with solvent too fast to be observed. n.a. means not assigned.

trypsin inhibitor (BPTI) (Kemmink et al., 1993; Kemmink & Creighton, 1995a). This interaction is characterized by a perpendicular orientation of the N–H bond with respect to the aromatic side chain plane. In protein L B1 domain, the local structure in these regions is different, with the N–H and the aromatic side chain pointing to opposite directions of the β-sheet, which result in much more normal values of the glycines amide proton chemical shifts (Wikström et al., 1993).

With respect to the fragment corresponding to the α-helix, it also has a small population of folded conformers in water as seen by CD and NMR. However, the peptide show signs of aggregation even in the most favorable conditions, a rather extreme pH value, and moderate peptide concentrations. These structures could be induced by the self-association of the peptide molecules, so that the conclusions that can be drawn about the intrinsic conformational tendency of the sequence are not so clear.

A major problem with the detection of non-random structures in short linear peptides is that their population is often very low. A useful technique to amplify the intrinsic conformational tendencies of the sequence, increasing the population of the folded structures, is the use of trifluoroethanol as a cosolvent at concentrations around 30% by volume. Aqueous TFE stabilizes native α-helices (Tamburro et al., 1968) and β-hairpins (Blanco et al., 1994b), although non-native structures can also be stabilized (Sönnichsen et al., 1992). The effect of TFE is not clearly understood. Besides its different dielectric constant and hydrogen bonding properties as compared with pure water (Llinás & Klein, 1975; Nelson & Kallenbach, 1986), the heterogeneity of TFE–water mixtures resulting from the immiscibility of the two components could be very important, as recently detected by small-angle X-ray scattering measurements (Kuprin et al., 1995). This study suggest that alcohols (specially fluorinated ones) behave as other amphiphilic molecules, and at relatively high concentrations can form some kind of pool of extended structures. It could be then that peptide molecules associate with these pools of TFE molecules in a way that the new environment stabilizes secondary structure formation. In fact, several helical peptides behave similarly in aqueous TFE and in SDS micelles (Rizo et al., 1993). The two hairpin fragments of protein L show a strange behavior in 30% TFE. Their CD spectra are similar to the ones observed in peptides with β-hairpin or mixed hairpin-helical structures plus random conformations (Blanco et al., 1994b; Blanco & Serrano 1995). However, the NMR line widths indicate that the peptides aggregate. What is probably happening is that there are peptide molecules that form oligomers adopting a folded structure and molecules that remain in a monomeric state or form low molecular weight aggregates. Both species would give rise to CD signal, but only the last ones could be observed by NMR due to the short relaxation time of the high molecular weight oligomers. On the other hand, the fragment corresponding to the α-helix shows a much more normal behavior in TFE. The line widths increase with respect to the aqueous solution experiment but to a lower extent than in the case of the hairpin peptides. In this peptide, the change is associated with the formation of a high population of α-helical structure in equilibrium with random coil conformations as shown by the CD spectrum and the NMR data (chemical shifts and NOEs). AGADIR predicts a tendency of being helical in water for residues T116 to D134, exactly the same region found to be helical in 30% TFE. For residues 116 to 121, AGADIR predicts less α-helix population than in the rest of the helical region, and that it competes with a relatively high tendency to be extended, and in fact, the helix is more frayed at its N-terminus end as manifested in the chemical shifts and NOEs of these residues. The helix in the protein L B1 domain structure starts at T116 (the N-cap residue) and the last residue with helical dihedral ϕ,ψ angles is L131. K132 has positive angles, directing the chain towards the second hairpin, and K133 has, again, helical angles. The helix in the isolated peptide starts at the same N-capping residue, but extends until D134. In the absence of tertiary interactions, it is not surprising that the helix also spans the two consecutive lysines (with high intrinsic tendency to form α-helices) and stops at D134, a poor helix former, instead of at L131.

Implications for protein L B1 domain folding

Protein L fragments show low secondary structure tendencies in aqueous solution in general. Although there are no signs of formation of the native β-hairpin structure in the isolated peptides corresponding to the protein hairpins, certain related conformational features are retained, namely the turns and side chain–side chain interactions in three of the four β-strands. The importance of these local structures for the folding of the protein is supported by its destabilization in the presence of 6 M urea. They also have a non-native interaction at the C-terminus end that has to be disrupted for the complete protein to be folded. More interestingly, TFE does not promote helix formation and AGADIR predicts a very low helical potential for these sequences, much lower that the tendency to be extended, thus suggesting a selection against non-native secondary structure propensities. These sequences then, have an intrinsic potential to form the native hairpin structure, but tertiary interactions with the rest of the protein are necessary for the sequence to form the overall hairpin structure. The sequence corresponding to the α-helix could form some amount of helical structure in water, but it is difficult to interpret its importance for the folding of the protein since the peptide has some tendency to self-associate. The intrinsic propensity for forming an α-helical structure is revealed in the presence of 30% TFE, although the helix is three residues longer at the C-terminus end. In this case, only a change in the environment would be needed to stabilize the helical conformation, but tertiary interactions again play an essential role in defining the final length of the helix in the whole protein.

It has been proposed that optimal folding occurs when there are only a few local interactions and, therefore, isolated fragments of secondary structure should be inherently weakly structured (Fersht, 1995; Govindarajan & Goldstein, 1995). Experimental analysis also show that there is a certain range of stability for a local structural element that permits the protein to attain its thermodynamics and folding characteristics: if local interactions are too weak, the folded state is destabilized, and if too strong, the cooperativity of folding is reduced due to a decrease in the energy gap between the folded state and alternative partially folded structures (Muñoz et al., 1996; Muñoz & Serrano, 1996). Some evolutionary pressure could therefore oppose the accumulation of nucleation sites in the denatured state because these sites should become stable only after interacting with the other parts of the protein (Fersht, 1995). Our experimental results with the protein L B1 domain fragments support this view. There are indications that the local sequence favors the structure finally attained in the protein, but the contribution of the local interactions alone is not high enough to drive the formation of experimentally detectable population of the native secondary structures.

Comparison with the B1 domain of protein G and ubiquitin

These three proteins have a similar fold and very low sequence identity. The main difference exists in a long loop in ubiquitin instead of the short turns connecting the two β-strands in the second β-hairpins of protein G and protein L. The conformational propensities of their fragments permit a comparison of the intrinsic secondary structure propensities of these sequences. The second β-hairpin of protein G forms around 40% population of the native structure in water, the α-helix forms a high population of helical structure in 30% TFE, and the first β-hairpin attain its native structure also in 30% TFE (Blanco et al., 1994b; Blanco & Serrano, 1995). None of the hairpins of protein L or ubiquitin are stable in water. Peptides corresponding to the first β-hairpin and the helix of ubiquitin adopt the native structures only in the presence of 60% aqueous methanol, while the C-terminus of the protein adopts non-native helical structures in the A state of the protein (a partially denatured state in 60% aqueous methanol, pH = 2.0; Cox et al., 1993) or in a peptide (Muñoz & Serrano, 1994). The hairpins of protein L do not form β-hairpin structures in water and their conformational state in TFE is uncertain as both aggregate. The tendency of the α-helical sequence to form the native α-helix structure, observed in mixed alcohol–water mixtures in the three cases, is the only common feature. This shows that an overall similar pattern of secondary structure tendency is not conserved within this structurally related set of proteins, and that this is not necessary to reach their final fold, although a minimum tendency to avoid non-native structures and to favor native ones could be required.

The protein G B1 domain seems to be a special case, as the local interactions alone appear to determine their native secondary structures. The balance of local versus non-local interactions in this protein is different than the one corresponding to ubiquitin and protein L B1 domain. The comparison of IgG binding surfaces of the protein G and protein L B1 domains suggests a possible common ancestral domain from which they have diverged (Wikström et al., 1995). Their sequences are highly dissimilar, but still their structures are essentially the same with only small differences in the length of the secondary structure elements and the orientation of the helix with respect to the β-strands. Their folded states show very similar dynamic properties (Barchi et al., 1994; Wikström et al., 1996) and their unfolding processes can be described as two-state transitions (Alexander et al., 1992; Yi & Baker, 1996). However, their pattern of secondary structure propensities is different: the salient feature in protein G B1 domain, the stable hairpin structure in the isolated second hairpin peptide, is not observed in protein L B1 domain.

Random mutagenesis and phage display technology has been applied on both domains to obtain mutant proteins that retain the correct three-dimensional structure, selected through IgG binding ability. For protein G B1 domain, all the mutants were less stable than the wild type and could be grossly classified in mutations altering tertiary interactions and mutations destabilizing secondary structures (O'Neil et al., 1995). In the case of protein L B1 domain, only the first hairpin has been mutagenized and all the mutants displayed reduced stability compared with the wild type, but up to nine simultaneous mutations were allowed (Gu et al., 1995). However, natural selection operates in a different way, and both folding characteristics and stability should be maintained. The secondary structure propensity has not been conserved in these two proteins through evolution, and mutations that locally destabilize the secondary structures could be compensated by mutations in other regions of the protein stabilizing the secondary structure by tertiary interactions, as was concluded for a similar analysis of α-helix propensity in evolutionarily related α/β parallel proteins (Muñoz et al., 1995b) In this way, the overall balance of local and non-local interactions can change within the range that ensures the folding and stability characteristics necessary for proper functionality in the living organism.

Materials and methods

Peptide design and synthesis

The length of the peptides was chosen to encompass whole secondary structure elements with their complete hydrogen-bonding pattern and van der Waals contacts, as defined in the NMR-determined three-dimensional structure of protein L B1 domain (Wikström et al., 1994). The first β-hairpin spans residues 95–115 and this fragment was selected. The α-helix of protein L starts at residue 117 and ends at residue 131. The following sequence (KKDNG) connects the helix with the second β-hairpin and resembles the sequence at the last turns of the helix in protein G (ANDNG). We decided to analyze a peptide with this sequence included, although it is not in helical conformation in protein L, in order to keep an open possibility for the helix to pass the limits observed in the whole protein. This situation has been found in other helical protein fragments (Jiménez et al., 1994). Three non-helical protein residues were also included at its N-terminus and two more residues at its C-terminus to minimize end chain effects and to favor solubility. The peptide studied then corresponds to residues 114–137 of protein L. The second β-hairpin spans residues 137–154. The corresponding peptide included the two glycines flanking the hairpin ends in the protein plus an additional aspartic acid residue at the C-terminus, not present in the native sequence, that would favor peptide solubility. This residue is a lysine in the complete protein L, and belongs to the connection between domains B1 and B2. This region is homologous to the disordered N-terminal end of the B1 domain (not shown in Fig. 1), and therefore, we do not expect any influence on the structure of the peptide. The peptide corresponding to the second hairpin is then named 136–155D fragment. Attending to solubility criteria, the net charge of the peptides was further adjusted at the pH range where full NMR analysis can be conveniently done (below 7): the first β-hairpin peptide has its C-terminus amidated, the α-helix peptide has both N- and C-termini free and the second β-hairpin peptide has its N-terminus acetylated.

The peptides were synthesized by the EMBL peptide synthesis service using Fmoc chemistry and PyBOP activation at a 0.025 mmol scale. Peptide homogeneity and identity were analyzed by analytical high-performance liquid chromatography, amino acid analysis, and matrix-assisted laser desorption time-of-flight mass spectrometry. The concentration of the peptide samples were determined by ultraviolet absorbance (Gill & Hippel, 1989) and amino acid analysis.

Far-UV circular dichroism spectroscopy

CD spectra were recorded on a Jasco-710 dichrograph calibrated with (1S)-(+)-10-camphorsulphonic acid. CD spectra were obtained in the continuous mode by taking point measurements every 0.2 nm, with 100 nm/min scan rate, a response of one second and

a 1 nm band width. Thirty consecutive scans were averaged. Cells with path lengths of 0.01 cm and 0.5 cm were used for the analysis of sample with peptide concentrations around 500 μM and 10 μM, respectively.

Nuclear magnetic resonance

NMR samples were prepared in H_2O with 10% ^2H_2O (by vol.), using milli Q water from a Millipore water system and ^2H_2O from Cambridge Isotope Chem.), 30% aqueous perdeuterated trifluoroethanol ($CF_3C^2H_2O^2H$, Cambridge Isotope Chem.), and 6 M urea (Merck). Minute amounts of HCl and NaOH were added in order to adjust the pH of the samples; this was measured with an Ingold combination electrode (Wilmad) inside the NMR tube, and isotope effects were not corrected. Sodium 3-trimethylsylyl (2,2,3,3-$^2H_4$)propionate (TSP) was used as an internal reference at 0.00 ppm. NMR experiments were performed on a Bruker AMX-500 spectrometer. The data were processed with the program UX-NMR from Bruker in an Aspect X32 computer. DQFCOSY (Piantini et al., 1982), NOESY (Kumar et al., 1980), and ROESY (Bothner-By et al., 1984) spectra were acquired using standard procedures. NOESY and ROESY spectra were routinely recorded for every peptide and the spectra were jointly analyzed to discard artifactual NOEs as those arising from spin diffusion. The mixing times used were 200 ms and 100 ms for spectra recorded in water and 30% TFE, respectively. TOCSY (Bax & Davis, 1985) spectra were acquired using the standard MLEV17 spin lock sequence and 80 ms mixing time. The spectral width was 5555.55 Hz and the water signal was presaturated during the relaxation delay (one second) and also during the mixing time of NOESY spectra. Temperature dependence of the amide signals was followed through a series of 1D, TOCSY, and ROESY spectra recorded at several temperatures within the 278–308 K range. The peptide ^{1}H-NMR spectra were assigned by the sequential assignment procedure (Wüthrich et al., 1982). Crosspeak intensities were evaluated by visual inspection of the contour levels. The conformational shifts (conformation-dependent chemical shifts dispersion) of the C_α protons were obtained by subtracting the random coil values (Merutka et al., 1995) from the measured ones for each residue.

Secondary structure prediction

The standard one-sequence version (AGADIR1s; Muñoz & Serrano, 1995) of the helix/coil transition algorithm AGADIR (Muñoz & Serrano, 1994) and AMPURIA (Muñoz & Serrano, unpublished) were used to predict α-helical or β-strand tendencies of the peptides of protein L B1 domain. These algorithms have been developed to estimate at a residue level the propensity of a monomeric peptide with no tertiary interaction to populate an α-helical or extended conformation.

The program AGADIR can be run on the world wide web (URL: http://embl-heidelberg.de/ExternalInfo/serrano).

Note added in proof

An interesting experimental study on the first β-hairpin of ferredoxin I appeared after submission of this manuscript. The protein has a structure similar to ubiquitin, and the peptide adopts a native-like hairpin structure in alcohol-water mixtures. (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.)

Acknowledgments

We thank M. Spiegel for his careful reading and corrections. M. Ramírez-Alvarado was supported by a doctoral fellowship from the National and Autonomous University of Mexico (DGAPA-UNAM). F.J. Blanco was supported by a postdoctoral fellowship from the Spanish ministry of Science and Education.

References

Alexander P, Orban J, Bryan P. 1992. Kinetic analysis of folding and unfolding the 56 amino acid IgG-binding domain of streptococcal protein G. *Biochemistry 31*:7243–7248.

Anfinsen CB. 1973. Principles that govern the folding of protein chains. *Science* 181:223–230.

Bax A, Davis DG. 1985. MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson 65*:355–360.

Barchi JJ Jr, Grasberger B, Gronenborn AM, Clore GM. 1994. Investigation of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional ^{1}H-^{15}N nuclear magnetic resonance spectroscopy. *Protein Sci 3*:15–21.

Blanco FJ, Jiménez MA, Rico M, Santoro J, Herranz J, Nieto JL. 1991. Tendamistat (12–26) fragment. NMR characterization of isolated β-turn folding intermediates. *Eur J Biochem 200*:345–351.

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

Blanco FJ, Jiménez MA, Pineda A, Rico M, Santoro J, Nieto JL. 1994b. NMR solution structure of the isolated N-terminal fragment of protein-G B1 domain. Evidence of trifluoroethanol induced native-like β-hairpin formation. *Biochemistry 33*:6004–6014.

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.

Bothner-By AA, Stephens RL, Lee JM, Warren CD, Jeanloz RW. 1984. Structure determination of a tetrassacharide: Transient nuclear Overhauser effect in the rotating frame. *J Am Chem Soc 106*:811–813.

Brown JE, Klee WA. 1971. Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry 10*:470–476.

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. 1995. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct Funct Genet 21*:167–195.

Chen YH, Yang JT, Chau KH. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry 13*:3350–3359.

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, Blanco FJ, Jiménez MA, Rico M, Nieto JL. 1995. Interactions responsible for the pH dependence of the β-hairpin conformational population formed by a designed linear peptide. *Eur J Biochem 233*:283-292.

de Alba E, Jiménez 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.

Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 29*:7133–7155.

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

Dyson HJ, Wright PE. 1991. Defining solution conformations of small linear peptides. *Annu Rev Biophys Biophys Chem 20*:519–538.

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

Dyson HJ, Sayre JR, Merutka G, Shin H-C, Lerner RA, Wright PE. 1992b. Folding of peptide fragments comprising the complete sequence of proteins. Models for Initiation of protein folding. II. Plastocyanin. *J Mol Biol 226*:819–835.

Dyson HJ, Wright PE. 1993. Peptide conformation and protein folding. *Curr Opin Struct Biol 3*:60–65.

Fersht AR. 1995. Optimization of rates of protein folding: The nucleation-condensation mechanism and its implications. *Proc Natl Acad Sci USA 92*:10869–10873.

Gill SC, Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem 182*:319–326.

Govindarajan S, Goldstein RA. 1995. Optimal local propensities for model proteins. *Proteins Struct Funct Genet 22*:413–418.

Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. 1991. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science 253*:657–661.

Gu H, Yi Q, Bray ST, Riddle DS, Shiau AK, Baker D. 1995. A phage display system for studying the sequence determinants of protein folding. *Protein Sci 4*:1108–1117.

Jiménez MA, Herranz J, Nieto JL, Rico, M, Santoro J. 1987. ^{1}H-NMR and CD evidence of folding of the isolated ribonuclease 50–61 fragment. *FEBS Lett 221*:320–324.

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

Jiménez MA, Muñoz V, Rico M, Serrano L. 1994. Helix stop and start signals in peptides and proteins: The capping box does not necessarily prevent helix elongation. *J Mol Biol 242*:487–496.

Kemmink J, Creighton TE. 1993. Local conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor and their roles in folding. *J Mol Biol 234*:861–878.

Kemmink J, van Mierlo CPM, Scheek RM, Creighton TE. 1993. Local structure due to an aromatic-amide interaction observed by ^{1}H-nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor. *J Mol Biol 230*:312–322.

Kemmink J, Creighton TE. 1995a. The physical properties of local interactions of tyrosine residues in peptides and unfolded proteins. *J Mol Biol 245*:251–260.

Kraulis PJ. 1991a. Similarity of protein G and ubiquitin. *Science 254*:581.

Kraulis PJ. 1991b. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.

Kuprin S, Gräslund A, Ehrenberg A, Koch MHJ. 1995. Nonideality of Water–Hexafluoropropanol mixtures as studied by X-ray small angle scattering. *Biochem Biophys Res Commun 217*:1151–1156.

Kuroda Y. 1993. Residual helical structure in the C-terminal fragment of cytochrome c. *Biochemistry 32*:1219–1224.

Llinás M, Klein MP. 1975. Charge relay at the peptide bond: A protein magnetic resonance study of solvation effects on the amide electron desity distribution. *J Am Chem Soc 97*:4731–4737.

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

Muñoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides in solution using empirical parameters. *Nat Struct Biol 1*:399–409.

Muñoz V, Serrano L. 1995. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J Mol Biol 245*:275–296.

Muñoz V, Blanco FJ, Serrano L. 1995a. The distribution of α-helix propensity along the polypeptide chain is not conserved in proteins from the same family. *Protein Sci 4*:1577–1586.

Muñoz V, Serrano L, Jiménez MA, Rico M. 1995b. Structural analysis of peptides encompassing all α-helices of three α/β parallel proteins: Che-Y, flavodoxin and p21-ras: Implications for α-helix stability and the folding of α/β parallel proteins. *J Mol Biol 247*:648–669.

Muñoz V, Cronet P, López-Hernández E, Serrano L. 1995. Analysis of the effect of local interactions on protein stability. *Folding & Design 1*:167–178.

Muñoz V, Serrano L. 1995. Local *vs* non-local interactions in protein folding and stability. An experimentalist's point of view. *Folding & Design 1*: R71–R77.

Nelson JW, Kallenbach NR. 1986. Stabilization of the ribonuclease S-peptide α-helix by trifluoroethanol. *Proteins Struct Funct Genet 1*:211–217.

O'Neil KT, Hoess RH, Raleigh DP, DeGrado WF. 1995. Thermodynamic genetics of the folding of the B1 immunoglobulin-binding domain from streptococcal protein G. *Proteins Struct Funct Genet 21*:11–21.

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

Ramírez-Alvarado M, Blanco FJ, Serrano L. 1996. *De novo* design and structural analysis of a model β-hairpin peptide system. *Nat Struct Biol 3*:604–613.

Rizo J, Blanco FJ, Kobe B, Bruch MD, Gierasch LM. 1993. Conformational behaviour of *Escherichia coli* OmpA signal peptides in membrane mimetic environments. *Biochemistry 32*:4881–4894.

Sali A, Shakhnovich E, Karplus M. 1994. How does a protein fold? *Nature 369*:248–251.

Sauer-Eriksson AE, Kleywegt GJ, Uhlén M, Jones TA. 1995. Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG. *Structure 3*:265–278.

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.

Segawa S-I, Fukuno T, Fujiwara K, Noda Y. 1991. Local structures in unfolded lysozyme and correlation with secondary structures in the native conformation: Helix-forming or breaking propensity of peptide fragments. *Biopolymers 31*:497–509.

Shin HC, Merutka G, Waltho JP, Wright PE, Dyson HJ. 1993. Peptide models of protein folding initiation sites. 2. The G-H turn region of myoglobin acts as a helix stop signal. *Biochemistry 32*:6348–6355.

Sönnichsen FD, Eyk JEV, Hodges RS, Sykes BD. 1992. Effect of trifluoroethanol on protein secondary structure: An NMR and CD study unsing a synthetic actin peptide. *Biochemistry 31*:8790–8798.

Tamburro AM, Scaturin A, Rocchi R, Marchiori F, Borin G, Scoffone E. 1968. Conformational transitions of bovine pancreatic ribonuclease S-peptide. *FEBS Lett 1*:298–300.

Viguera AR, Jiménez MA, Rico M, Serrano L. 1996. Conformational analysis of peptides corresponding to β-hairpins and β-sheet that represent the entire sequence of the α-spectrin SH3 domain. *J Mol Biol 255*:507–521.

Vijay-Kumar S, Bugg CE, Wilkinson KD, Vierstra RD, Hatfield PM, Cook WJ. 1987. Comparison of the three dimensional structures of human, yeast and oat ubiquitin. *J Mol Biol 13*:6396–6399.

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- and H-helices of Myoglobin. *Biochemistry 32*:6337–6347.

Wikström M, Sjöbring U, Kastern W, Björck L, Drakenberg T, Forsén S. 1993. Proton nuclear magnetic resonance sequential assignments and secondary structure of an immunoglobulin light chain-binding domain of protein L. *Biochemistry 32*:3381–3386.

Wikström M, Drakenberg T, Forsén S, Sjöbring U, Björk L. 1994. Three dimensional structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G. *Biochemistry 33*:14011–14017.

Wikström M, Sjöbring U, Drakenberg T, Forsén S, Björk L. 1995. Mapping of the immunoglobulin light chain-binding site of protein L. *J Mol Biol 20*:128–133.

Wikström M, Forsén S, Drakenberg T. 1996. Backbone dynamics of a domain of protein L which binds to immunoglobulin light chains. *Eur J Biochem 235*:543–548.

Williamson MP. 1990. Secondary-structure dependent chemical shifts in proteins. *Biopolymers 29*:1423–1431.

Wright PE, Dyson HJ, Lerner RA. 1988. Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding. *Biochemistry 27*:7167–7175.

Wüthrich K, Wieder G, Wagner G, Braun W. 1982. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution magnetic resonance. *J Mol Biol 155*:311–319.

Yi Q, Baker D. 1996. Direct evidence for a two-state protein unfolding transition from hydrogen-deuterium exchange, mass spectrometry and NMR. *Protein Sci 5*:1060–1066.