A Designed Three Stranded β-Sheet Peptide as a Multiple β-Hairpin Model

Chittaranjan Das,[†] S. Raghothama,[‡] and P. Balaram*[,†]

Molecular Biophysics Unit and
Sophisticated Instrumentation Facility
Indian Institute of Science, Bangalore-560012, India

Received October 29, 1997

β-Hairpins are important elements in protein structure,[1] implicated in the nucleation of β-sheets in protein folding.[2] Multiple hairpin structures, in which successive strands are connected by very short loops (2−5 residues), are found in the trans-membrane β-barrel structures formed by the bacterial porins.[3] Considerable recent attention has been directed toward the synthesis of isolated β-hairpins, stable in both aqueous and organic solvents.[4] In most cases the central nucleating segment has a high propensity to form either type I′ or type II′ β-turn structures, a feature that occurs widely in protein structures.[1] The design of peptides that mimic folding patterns observed in proteins, from first principles, provides a rigorous test of our understanding of the factors that determine polypeptide stereochemistry. While helical structures have been the subject of a very large number of investigations,[5] β-sheets have been much less studied by *de novo* design approaches,[6] although the very first examples of synthetic protein design targeted β-sheets.[7] The major problem in β-sheet design is the limited solubility and pronounced tendency of extended peptide strands to aggregate in solution. Further, strong residue pair correlations are not observed in β-hairpins in proteins[1c,8] precluding specific sequence choices in "first principle design" approaches. Since in protein structures there is a preponderance of β-branched residues (Val, Ile, Thr) in β-sheets,[9] designed sequences have tended to incorporate these residues. Recent studies establish that the type II′ β-turn conformation adopted by a central ^{D}Pro-Xxx segment can stabilize β-hairpin formation,[4g−k,10] a feature that results from the constraint of the pyrrolidine ring restricting the ϕ value in ^{D}Pro to ca. +60°. Since interstrand hydrogen bonding and β-turn nucleation appear to be the key determinants of β-hairpin structures, we decided to explore the possibility of constructing a well-defined synthetic three-stranded β-sheet peptide. The results presented in this communication provide definitive characterization of stable multiple hairpin structures in organic solvents, in which the three strands of a β-sheet are connected by two tight, ^{D}Pro-Gly β-turns.

Table 1. NMR Parameters for Peptide **1** in $CDCl_3$

residue	NH δ (ppm)	$C^{\alpha}H$ δ (ppm)	$^{3}J_{NH-C^{\alpha}H}$ (Hz)	Δδ (ppm)[a]
Leu (1)	6.70	4.15	8.7	0.03
Phe (2)	6.66	5.30	8.5	1.24
Val (3)	8.70	4.40	9.2	−0.29
D-Pro (4)		4.25		
Gly (5)	5.51	3.67		2.08
		4.31		
Leu (6)	7.20	4.78	*b*	0.36
Val (7)	7.70	4.87	9.7	−0.20
Leu (8)	8.70	4.91	9.2	−0.10
Ala (9)	8.62	4.97	8.7	−0.24
D-Pro (10)		4.35		
Gly (11)	6.40	3.62		1.59
		4.12		
Phe (12)	7.62	4.70	8.9	0.06
Val (13)	6.56	4.85	9.6	1.37
Leu (14)	8.79	4.84	9.0	−0.26

[a] $\Delta\delta = \delta(NH),CDCl_3 - \delta(NH),CDCl_3/DMSO$ (16%). [b] Coupling constant not determined because of resonance overlap.

The fourteen residue peptide Boc Leu-Phe-Val-^{D}Pro-Gly-Leu-Val-Leu-Ala-^{D}Pro-Gly-Phe-Val-Leu-OMe (**1**)[11] was designed such that the N-terminal strand would be composed of three residues (1−3), the central strand of four residues (6−9), and the C-terminal strand of three residues (12−14), with two β-turn nucleating ^{D}Pro-Gly segments. The peptide showed high solubility in a wide range of organic solvents, including the apolar solvent chloroform. The 500 MHz ^{1}H NMR spectrum in $CDCl_3$ was extremely well resolved with sharp resonances. The lack of concentration dependence of NH chemical shifts over the range 7.1 to 0.2 mM suggests the absence of aggregation effects. The wide dispersion of NH and $C^{\alpha}H$ chemical shifts facilitates complete assignment and is indicative of a well-defined conformational species in solution. Sequence specific assignment of all backbone proton resonance was readily achieved by using a combination of TOCSY and ROESY experiments.[12] The backbone proton chemical shifts and coupling constants are summarized in Table 1. All the vicinal $J_{HNC^{\alpha}H}$ values for the potential strand residues are greater than 8.5 Hz, with several values greater than 9 Hz providing clear evidence for $\phi \sim -120°$ throughout the strand segments.[13] Upon addition of the free radical 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) only two resonances, Gly5

* Address correspondence to this author.
[†] Molecular Biophysics Unit.
[‡] Sophisticated Instrumentation facility.

(1) (a) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170−174. (b) Sibanda, B. L.; Blundell, T. L.; Thornton, J. M. *J. Mol. Biol.* **1989**, *206*, 759−777. (c) Gunasekaran, K.; Ramakrishnan, C.; Balaram, P. *Protein Eng.* In press.
(2) Ptitsyn, O. B. *FEBS Lett.* **1981**, *131*, 197−201.
(3) Schulz, G. E. *Curr. Opin. Struct. Biol.* **1996**, *6*, 485−490.
(4) (a) Alvarado, M. R.; Blanco, F. J.; Serrano, L. *Nature Struct. Biol.* **1996**, *3*, 604−612. (b) Blanco, F. J.; Rivas, G.; Serrano, L. *Nature Struct. Biol.* **1994**, *1*, 584−590. (c) Maynard, A. J.; Searle, M. S. *J. Chem. Soc., Chem. Commun.* **1997**, 1297−1298. (d) Searle, M. S.; Williams, D. H.; Packman, L. C. *Nature Struct. Biol.* **1995**, *2*, 999−1006. (e) Alba, E.; Jimenez, M. A.; Rico, M.; Neito, J. L. *Folding Design* **1996**, *1*, 133−144. (f) Alba, E.; Jimenez, M. A.; Rico, M. *J. Am. Chem. Soc.* **1997**, *119*, 175−183. (g) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 2303−2304. (h) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 6975−6985. (i) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4105−4106. (j) Awasthi, S. K.; Raghothama, S.; Balaram, P. *Biochem. Biophys. Res. Commun.* **1995**, *216*, 375−381. (k) Karle, I. L.; Awasthi, S. K.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8189−8193.
(5) (a) DeGrado, W. F. *Adv. Protein Chem.* **1988**, *39*, 51−124. (b) Betz, S. F.; Raleigh, D. P.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 601−610. (c) Chakrabarthy, A.; Baldwin, R. L. *Adv. Protein Chem.* **1995**, *46*, 141−176.
(6) (a) Richardson, J. S.; Richardson, D. C.; Tweedy, N. B.; Gernert, K. M.; Quinn, T. P.; Hecht, M. H.; Erickson, B. W.; Yan, Y.; McClain, R. D.; Donlan, M. E.; Surles, M. C. *Biophys. J.* **1992**, *63*, 1186−1209. (b) Quinn, T. P.; Tweedy, N. B.; Williams, R. W.; Richardson, J. S.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8747−8751. (c) Yan, Y.; Erickson, B. W. *Protein Sci.* **1994**, *3*, 1069−1073.
(7) (a) Moser, R.; Thomas, R. M.; Gutte, B. *FEBS Letts.* **1983**, *157*, 247−251. (b) Gutte, B.; Klauser, S. *Peptides: Synthesis, Structure, and Application*; Gutte, B., Ed.; Academic Press: New York, 1995; pp 363−369.
(8) Wouters, M. A.; Curmi, P. M. G. *Proteins: Struct. Funct. Genet.* **1995**, *22*, 119−131.
(9) (a) Minor D. E.; Kim, P. S. *Nature* **1994**, *367*, 660−663. (b) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510−5517.
(10) (a) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342−342. (b) Raghothama, S. R.; Awasthi, S. K.; Balaram, P. *J. Chem. Soc., Perkin Trans. 2*. In press.
(11) The 14-residue peptide (**1**) was synthesized by conventional solution phase methods and purified by reverse phase medium-pressure liquid chromatography (C_{18}, 40−60 μm) followed by HPLC (C_{18}, 10 μm) with methanol−water gradients. The peptide was characterized by its FAB mass spectrum (MH^{+} 1556, $M_{calcd.}$ 1555) and by complete assignment of the 500 MHz ^{1}H NMR spectrum. All NMR studies were carried out as previously described.[14]
(12) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(13) Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741−751.
(14) Raghothama, S.; Chaddha, M.; Balaram, P. *J. Phys. Chem.* **1996**, *100*, 19666−19671.

S0002-7863(97)03739-6 CCC: $15.00
Published on Web 05/27/1998

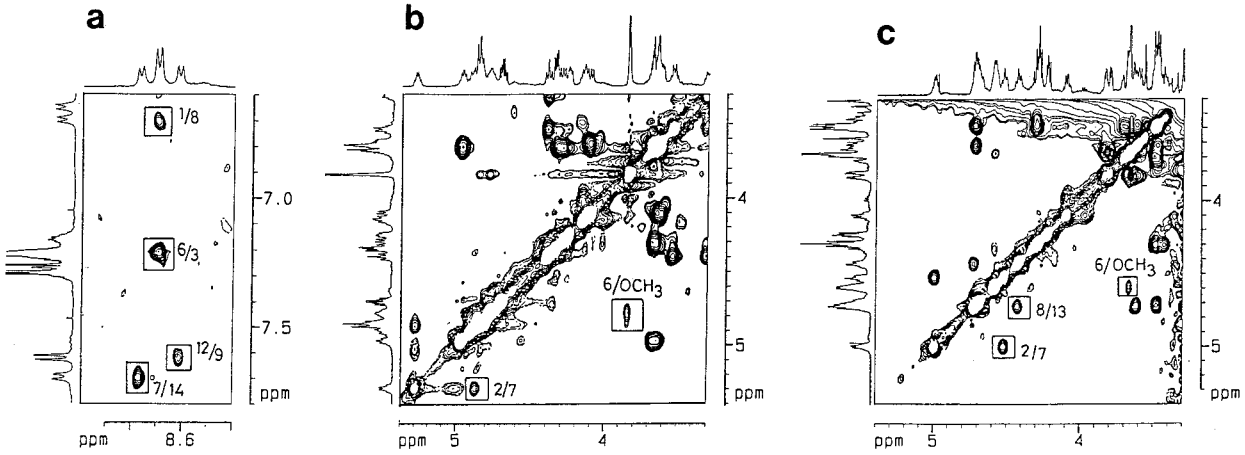

Figure 1. Partial 500 MHz ROESY spectra of peptide **1** (2 mM, 300 K) in $CDCl_3$: (a) interstrand NH↔NH NOEs; (b) $C^{\alpha}H \leftrightarrow C^{\alpha}H$ NOEs; (c) $C^{\alpha}H \leftrightarrow C^{\alpha}H$ NOEs observed on addition of DMSO-d_6 (16%) to a $CDCl_3$ solution of peptide **1**.

and Gly11, broaden considerably. Addition of the strongly hydrogen bonding solvent, dimethyl sulfoxide ($(CD_3)_2SO$) to $CDCl_3$ solutions of peptides should perturb only exposed NH groups capable of solute−solvent hydrogen bonding.[14] The solvent shifts ($\Delta\delta$ values) listed in Table 1 clearly demonstrate that only 4 NH groups, Phe2, Gly5, Gly11, and Val13, experience large downfield shifts indicating their exposure. The remaining 8 NH groups appear to be solvent inaccessible, strongly supporting their involvement in intramolecular hydrogen bond formation. Two critical $N_iH \leftrightarrow N_{i+1}H$ NOEs (Gly5-Leu6 and Gly11-Phe12) establish backbone dihedral angle values in the helical region for residues 11 and 12 ($\phi \sim -60° \pm 30°$; $\psi \sim -40° \pm 30°$), consistent with these residues occupying the $i + 2$ position of type II′ β-turns. A striking feature of the NOE spectra is the absence of any interresidue $N_iH \leftrightarrow N_{i+1}H$ NOEs in the strand segments. This observation, together with the fact that intense interresidue $C_i^{\alpha}H \leftrightarrow N_{i+1}H$ NOEs are observed for all potential strand residues, provides support for a structure consisting exclusively of strand segments linked by internal β-turns. Evidence for the multiple hairpin conformation is obtained by the observation of the cross strand NOEs between the amide protons Leu1/Leu8, Val3/Leu6, Val7/Leu14, and Ala9/Phe12 (Figure 1a). The interstrand $C^{\alpha}H \leftrightarrow C^{\alpha}H$ NOE between Phe2/Val7 (Figure 1b) firmly establishes the N-terminal hairpin formed between residues 1 and 8. Also shown in Figure 1 is the ROESY spectrum of peptide 1 obtained after addition of 16% $(CD_3)_2SO$ to a $CDCl_3$ solution. Under these conditions, both the Phe2/Val7 and Leu8/Val13 $C^{\alpha}H \leftrightarrow C^{\alpha}H$ NOEs are observed, demonstrating the presence of *both* the hairpin segments in the molecule. An important long-range NOE is also observed between the methyl group of the C-terminal ester function and Leu6 $C^{\alpha}H$, confirming that the antiparallel strands do not fray, even at the termini.

The wide chemical shift dispersion, identification of eight solvent shielded NH groups and four solvent exposed NH groups, the high $J_{HNC^{\alpha}H}$ values for all strand residues, and the observation of as many as seven critical interstrand long-range NOEs provide conclusive evidence for the three-stranded β-sheet structure shown in Figure 2a. By using 35 NOE constraints and 8 interstrand hydrogen-bonding constraints a restrained molecular dynamic simulation converged to a family of three-dimensional structures (mean RMSD for backbone atoms 0.66 Å), illustrated in Figure 2b.

The absence of aggregation even in a poorly solvating medium like chloroform merits special comment. Earlier studies in this laboratory have demonstrated isolated β-hairpin formation in the octapeptide (Boc-Leu-Val-Val-^{D}Pro-Gly-Leu-Val-Val-OMe) in organic solvents[4j,10b] and in crystals.[4k] If only residues 1 to 8 are considered in an isolated hairpin two NH groups on the strands will be exposed, while four will be internally hydrogen bonded. Expansion of this motif to the three-stranded structure shown in

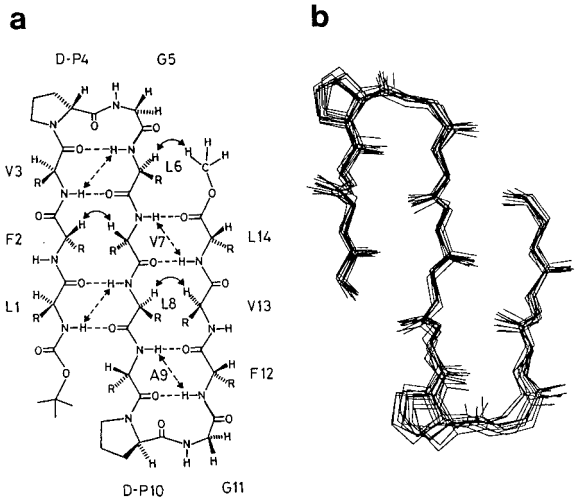

Figure 2. (a) Summary of long-range NOEs (between backbone protons) observed for peptide **1**. Solid arrows indicate strong NOEs and broken arrows indicate weak NOEs. Observed hydrogen bonding interactions are shown as broken lines. (b) Superposition of 10 structures of peptide **1** obtained from molecular dynamics calculations with NOE-derived restraints. The backbone traces (including D-Pro side chain atoms) are shown.

Figure 2a results in as many as eight internally hydrogen bonded NH groups, while only two NH groups are exposed (residue 2 and 13) in strand segments (the Gly NH groups in the turns do not contribute to extended two-dimensional sheet formation). The presence of only one free NH group on the two outer strands limits aggregation, a process facilitated by cooperative formation of several hydrogen bonds in sheets. In the present case, any enthalpic gain by formation of limited intermolecular hydrogen bonds may be offset by entropic losses due to desolvation and restriction of side chain flexibility and molecular motion. Extension of this design strategy to larger, intramolecular β-sheets appears possible. The demonstration of stable multiple hairpin formation in organic solvents is of particular relevance in mimicking porin like structures. Significantly, β-sheets in proteins are also largely buried in hydrophobic environments, consistent with the importance of cross-strand hydrogen bonds in maintaining structural integrity.

Acknowledgment. We acknowledge S. Kumar Singh and K. Gunasekaran for their help in molecular dynamics based structure calculations

Supporting Information Available: Spectra of 500 MHz ^{1}H NMR showing nitroxide induced line broadening and 2D ROESY (2 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA973739Q