A unique example of a core-modified bis-proline peptide self-assembling into an infinite hydrogen-bonded β-sheet ribbon: crystal structure of Z-ProNH$(CH_2)_2$NHPro-Z

Darshan Ranganathan,*[a] M. Gopi Kumar,[a] R. S. K. Kishore[a] and Isabella L. Karle*[b]

[a] *Discovery Laboratory, Indian Institute of Chemical Technology, Hyderabad 500 007, India. E-mail: ranganathan@iict.ap.nic.in*

[b] *Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5341, USA*

Received (in Columbia, MO, USA) 6th November 2000, Accepted 5th December 2000
First published as an Advance Article on the web 23rd January 2001

The crystal structure of Z-ProNH$(CH_2)_2$NHPro-Z exhibits an extended backbone, with occurrence of both a *cis* and a *trans* conformation preceding the two prolyl residues, that self-assembles into an infinite hydrogen-bonded β-sheet ribbon.

Proline, the only amino acid found in proteins with a secondary amino group in a five-membered ring, is known to play an important role in the folding of natural proteins.[1] *N*-Acyl prolines lacking amide NHs display energetically similar *cis*- and *trans*-isomeric forms[2] and both forms are known to occur in proteins, for example, ribonuclease[3] or in bioactive peptides such as bradykinin.[4] Being a cyclic amino acid, proline imposes a certain degree of conformational constraint in the protein or peptide backbone causing bends or breaking helices,[5] and consequently proline-containing peptides are generally not known to display extended structures. We describe herein a unique example of a core-modified bis-proline peptide that is preceded by both a *cis* as well as a *trans* amide bond in its framework and self-assembles to form a hydrogen-bonded infinite β-sheet ribbon in the solid state.

The 1,2-ethylenediamine spacer-linked bis-proline **1** [Fig. 1(a)] was prepared[6] in a single step by coupling Z-proline with 1,2-diaminoethane using a direct azide coupling (diphenyl phosphoryl azide) procedure. The bis-Pro peptide **1** was fully characterized[7,8] by spectral and analytical data.

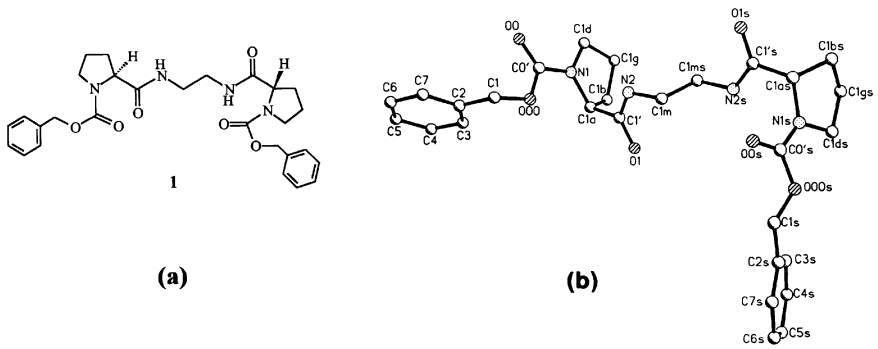

Fig. 1 (a) Molecular formula of **1**. (b) X-Ray structure of **1**. The numbering is the same for each half except that the letter 's' is added for atoms on one side.

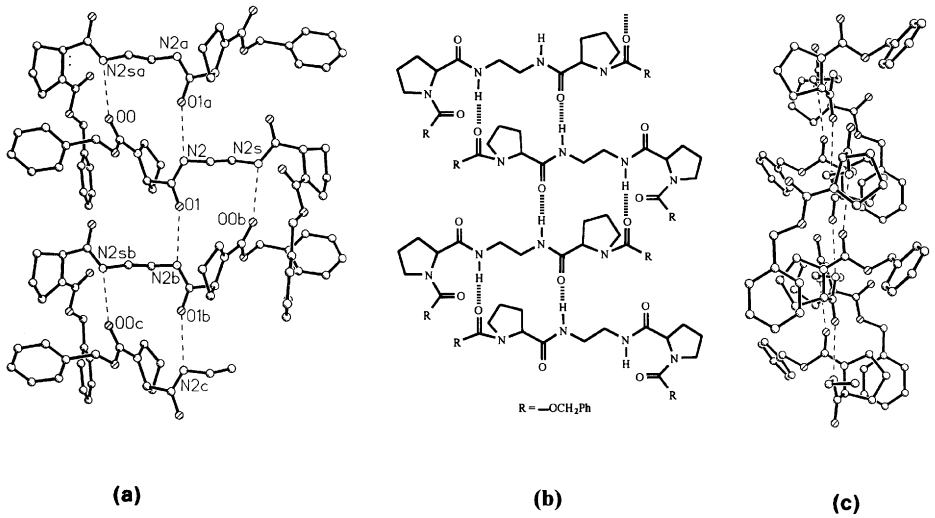

Fig. 2 (a) Flat ribbon formation by connecting the molecules by two types of hydrogen bonds, N(2)H⋯O(1) and N(2s)H⋯O(0). The molecules are repeated by a vertical two-fold screw axis. (b) Schematic representation of the ribbon assembly of **1**. (c) A side view of the ribbon. The dashed lines represent hydrogen bonds. Both surfaces of the ribbon are largely hydrophobic by virtue of the extended placement of the benzyl moieties and prolyl rings.

Suitable crystals for X-ray diffraction were obtained from a mixture of ethyl acetate and hexane. The crystal structure of **1** showed a backbone extended in two sections with a sharp bend at Pro(1s) [Fig. 1(b)]. Interestingly, the two halves of the molecule are not symmetric. On one side of the molecule the Z-L-Pro (1) amide bond C0′–N1 has the *cis* conformation, whereas on the other side the Z-L-Pro (1s) amide bond C0′s–N1s has the *trans* conformation. A similar occurrence of a *cis* amide bond preceding a Pro residue had been noted in 1974 for the Boc-Pro (1) amide bond in Boc-(L-Pro)$_4$-OBzl.[9]

The hydrogen-bonding pattern [Fig.2(a)] in the crystal lattice showed the formation of an infinite ribbon assembly around a two-fold screw axis (space group $P2_1$). There are only two independent hydrogen bonds, N(2)H··O(1) and N(2s)H··O(0) that connect the molecules into ribbons. Fig. 2(b) shows the schematic picture of the ribbon assembly. A side view of the ribbon in Fig. 2(c) shows a relatively flat structure. The protruding phenyl and pyrrolidine rings on either side impart a hydrophobic surface to the ribbon. Table 1 presents the hydrogen bond parameters and torsional angles in **1**. In solution-state conformational studies of **1**, while ^{1}H NMR (variable temperature measurements) has shown agreement with the solid-state structure in exhibiting high dδ/dT values (> 4 ppb K^{-1}) indicating the absence of any intramolecular hydrogen bonding, the FTIR spectrum in chloroform solution at 297 K showed two bands in the NH stretch region. The band at *ca*. 3430 cm^{-1} is assigned to the non hydrogen-bonded NH and the concentration independent, major band at *ca*. 3340 cm^{-1} is attributed to an internally hydrogen-bonded NH, possibly, in a seven-membered ring.

Table 1 Structural characteristics of **1**

(a) Intermolecular hydrogen bonds (Å) and angles (°)					
N2···O1[a]	2.899		N2s···O0[b]	2.877	
H2···O1[a]	2.10		H2s···O0[b]	2.09	
N2···O1=C	153		N2s···O0=C	152	
(b) Selected torsional angles (°)[c]					
C2C1O00C0′		+83	C2sC1sO00sC0′s		+155
C1O00C0′N1	ψ_o	179	C1sO00sC0′sN1s	ψ_{os}	−176
O00C0′N1C1a	ω_o	+12	O00sC0′sN1sC1as	ω_{os}	−171
C0′N1C1aC1′	ϕ_1	−90	C0′sN1sC1asC1′s	ϕ_{1s}	−82
N1C1aC1′N2	ψ_1	−7	N1sC1asC1′sN2s	ψ_{1s}	−12
C1aC1′N2C1m	ω_1	−173	C1asC1′sN2sC1ms	ω_{1s}	177
C1′N2C1mC1ms		+92	C1′sN2sC1msC1m		−99
N2C1mC1msN2s		−175			

[a] Symmetry equivalent $-x$, $-0.5 + y$, $1 - z$. [b] Symmetry equivalent $-x$, $+0.5 + y$, $1 - z$. [c] Conventions for labels in ref. 10.

The present results demonstrate that appropriate core inserts can modify the conformational behavior of proline residues in a peptide.

This work was financially supported by the Department of Science and Technology, New Delhi, National Institute of Health (GM-30902), USA, and the Office of Naval Research. Helpful advice from Professor S. Ranganathan is acknowledged. D. R. is an honorary faculty member of Jawaharlal Nehru Center for Advanced Scientific Research.

Notes and References

1 For reviews on protein folding, see: T. E. Creighton, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 5082; G. T. Montelione and H. A. Scheraga, *Acc. Chem. Res.*, 1989, **22**, 70; J. King, *Chem. Eng. News*, 1989, **67**, 32; R. L. Baldwin, *Trends Biochem. Sci.*, 1989, **14**, 291; L. M. Gierasch and J. King, *Protein Folding; Deciphering the Second Half of the Genetic Code*, AAAS Pub., Washington D.C., 1990; T. E. Creighton, *Biochem. J.*, 1990, **270**, 1; R. Jaenicke, *Biochemistry*, 1991, **30**, 3147; C. M. Dobson, A. Sali and M. Karplus, *Angew. Chem., Int. Ed.*, 1998, **37**, 868.

2 W. S. Blair and B. L. Semler, *Curr. Opin. Cell. Biol.*, 1991, **3**, 1039 and references therein.

3 P. N. Lewis, F. A. Momany and H. A. Scheraga, *Biochim. Biophys. Acta*, 1973, **303**, 211.

4 R. E. London, J. M. Stewart, R. Williams, J. R. Cann and N. A. Matwiyoff, *J. Am. Chem. Soc.*, 1979, **101**, 2455.

5 J. L. Crawford, W. N. Lipscomb and C. G. Schollman, *Proc. Natl. Acad. Sci., USA*, 1973, **70**, 538; P. Y. Chou and G. D. Fasman, *Biochemistry*, 1974, **13**, 222; C. M. Venkatachalam, *Biopolymers*, 1969, **6**, 1425.

6 The N$^{\alpha}$Z protected proline azide, generated *in situ*, directly from Z-Pro (5 mmol) and diphenylphosphoryl azide (DPPA, 5 mmol) in dry CH_2Cl_2–DMF (5 mL, 3:2) was admixed with 1,2-diaminoethane (2.5 mmol) and the reaction mixture was left stirred for 8 h at 0 °C and 12 h at room temp., the solvents evaporated *in vacuo*, residue mixed with water (*ca*. 50mL), extracted with ethyl acetate (2 × 50 mL), organic extract washed with aq. bicarbonate solution and dried (anhyd. $MgSO_4$). The residue, after solvent removal *in vacuo*, was purified on a short column of silica gel using ethyl acetate–hexane as eluents.

7 *Selected data* for **1**. Yield 80%; mp 158–160 °C; δ_H (200 MHz $CDCl_3$) 1.80–2.31 (m, 8H), 3.05–3.71 (m, 8H), 4.20 (m, 2H), 5.15 (m, 4H), 6.6–7.0 (br, 2H), 7.35 (br s, 10H); FAB-MS (*m*/*z*) 523 (MH)$^+$, 545 (M+Na$^+$).

8 Crystal data for **1**: $C_{28}H_{34}N_4O_6$, space group $P2_1$, a = 9.992(2), b = 10.251(2), c = 13.325(3), β = 90.25(2)°, V = 1364.9(5) Å^3, D_c = 1.272 g cm^{-3}, Cu-Kα radation, λ = 1.54178 Å. Least-squares refinement on F, R = 0.0578 for 1652 data [$F > 3.0\sigma(F)$]. Data collection at 293 °C. CCDC 182/1884.

9 T. Matsuzaki, *Acta Crystallogr., Sect. B*, 1974, **30**, 1029.

10 IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, 1970, **9**, 3471.