β-Turn Conformation of N-Acetyl-L-prolylglycyl-L-phenylalanine. Crystal Structure and Solution Studies

Samir K. Brahmachari,[1a] T. N. Bhat,[1a] V. Sudhakar,[1a] M. Vijayan,[1a] R. S. Rapaka,[1b] R. S. Bhatnagar,[1b] and V. S. Ananthanarayanan*[1a,c]

Contribution from the Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India, the School of Denistry, University of California, San Francisco, California 94143, and the Department of Biochemistry, Memorial University of Newfoundland, St. John's, Newfoundland A1B 3X9, Canada. *Received November 26, 1979*

Abstract: Pro-Gly segments in peptides and proteins are prone to adopt the β-turn conformation. This paper reports experimental data for the presence of this conformation in a linear tripeptide N-acetyl-L-prolylglycyl-L-phenylalanine both in the solid state and in solution. X-ray diffraction data on the tripeptide crystal show that it exists in the type II β-turn conformation. CD and proton NMR data show that this conformation persists in trifluoroethanol and methanol solutions in equilibrium with the nonhydrogen-bonded structures. Isomerization around the acetyl–prolyl bond is seen to take place in dimethyl sulfoxide solutions of the tripeptide.

The β turn has been recognized as an important structural feature forming, on the average, about 30% of the secondary structure of globular proteins.[2] In addition, functional roles have been assigned to the β turn such as its acting as the site for posttranslational enzymic modifications of proteins.[3,4] An analysis of the X-ray crystallographic data on several globular proteins[2] reveals positional preferences for the amino acid residues to occur within the β turn formed by the tetrapeptide segment R_1-R_2-R_3-R_4. In particular, Pro is found to have the highest preference for the second position of the β turn.[2] With Pro in the second position, Gly happens to be the most preferred residue in the third position of the β turn.[5] It is therefore of interest to examine the details of the conformation of tetrapeptide segments of the type R_1-Pro_2-Gly_3-R_4. One of the important observations in this direction is the fact that the stability of the β turn formed by such segments is governed by the nature of the residues R_1 and R_4 adjoining the Pro-Gly segment (see Figure 1) in several globular proteins.[5] This is supported by our experimental studies (submitted for publication) on a series of tripeptides having the general sequence N-acetyl-Pro-Gly-X-OH, where X = Gly, Ala, Leu, Ile, or Phe. We find that, in solution, the relative thermodynamic stability of the β turn in these compounds is dictated by the nature of the residue, X. These observations are useful not only in understanding the stability of the β turn in globular proteins but also in governing the conformational characteristics of procollagen, which contains a fairly large number of Pro-Gly segments.[6] We have recently shown[3] that the enzymatic hydroxylation of selected proline residues in nascent procollagen is governed by the proportion of the β-turn conformation at the Pro-Gly-X segments which, in turn, depends on the nature of the residue X.

As part of our experimental studies on Pro-Gly-containing tripeptides, we report in this paper the conformation of N-acetyl-Pro-Gly-Phe-OH in the solid state and in solution. Several Pro-Gly-Phe segments are found to occur in nascent procollagen.[6] This segment is also present in the nanopeptide hormone bradykinin. Our results show that the preferred conformation of the tripeptide is the β turn both in the solid state, as determined by X-ray crystallography, and in solution, as judged from CD and NMR data.

Experimental Section

Synthesis. L-Prolylglycyl-L-phenylalanine was synthesized according to the method of Rapaka and Bhatnagar.[7] This peptide (0.64 g, 2 mmol) was dissolved in 3 mL of acetic acid, cooled, treated with acetic anhydride (0.51 g, 5 mmol), and stirred continuously. After about 7 h, the reaction mixture was lyophilized to get a thick oily mass. The oil was crystallized from 95% ethanol–acetone–petroleum ether to get white crystalline needles of N-acetyl-L-prolylglycyl-L-phenylalanine (which will be referred to as the Phe-tripeptide): mp 164–165 °C; $[\alpha]_D$ = −22.1° (c 0.43, MeOH). The compound showed a single spot on TLC in the solvent systems: butanol–acetic acid–water (4:1:1), R_f = 0.49; butanol–acetic acid–pyridine–water (20:1:15), R_f = 0.70. Anal for $C_{18}H_{23}N_3O_5$: C, H, N.

Table I. Crystal Data for N-Acetyl-Pro-Gly-Phe-OH

mol formula	$C_{18}H_{23}O_5N_3$	crys system	monoclinic
mol wt	361.18	space group	$P2_1$
color	white	a, Å	8.915
habit	needle shape	b, Å	8.458
density (measd), g cm^{-3}	1.238 ± 0.005	c, Å	13.216
density (calcd), g cm^{-3}	1.244	β, deg	104.8
radiation, Cu Kα, Å	1.5418	V, Å^3	963.47
μ (Cu Kα), cm^{-1}	6.20		

X-ray Studies. The crystals of the Phe-tripeptide obtained in the synthesis were used for X-ray diffraction measurements. The space group and the unit cell dimensions of the crystal were determined from oscillation and Weissenberg photographs. The density of the crystal, measured by flotation method in a mixture of CCl_4 and benzene, indicated the presence of two tripeptide molecules in the unit cell. The complete crystal data are given in Table I.

The intensity data were collected on a computer-controlled CAD-4 diffractometer from a crystal of dimensions 0.6 mm × 0.2 mm × 0.2 mm to a maximum Bragg angle of 65° with use of Cu Kα radiation. Data were corrected for Lorentz–polarization factor but not for absorption. The intensities were put on the absolute scale by using Wilson's statistics.

^{1}H NMR Data. Proton NMR spectra were recorded on a Brucker WH-270 FT instrument at the Bangalore NMR facility. Chemical shifts were measured in ppm with internal tetramethylsilane (Me_4Si) as reference in all the cases. All the spectra were recorded in the Fourier transform mode with the use of a deuterium lock. For nondeuterated solvents, the homogeneity was adjusted by observing the FID of the solvent signal. In these cases, a gated decoupler was used to suppress the solvent signal and a minimum of 100–400 scans were made to improve the signal-to-noise ratio. Temperature control to ±1 °C was achieved by passing dry N_2 gas through a heating coil for high temperature and through a liquid-nitrogen trap for low temperatures. At each temperature, the sample was equilibrated for 15–20 min before collecting the

* To whom correspondence should be addressed at Memorial University of Newfoundland.

(1) (a) Indian Institute of Science, Bangalore. (b) University of California, San Francisco. (c) Memorial University, St. John's, Newfoundland.

(2) Chou, P. Y.; Fasman, G. D. J. Mol. Biol. 1977, 115, 135–175.
(3) Brahmachari, S. K.; Ananthanarayanan, V. S. Proc. Natl. Acad. Sci. U.S.A. 1979, 76, 5119–5123.
(4) Beely, J. G. Biochem. Biophys. Res. Commun. 1977, 76, 1051–1055.
(5) Brahmachari, S. K.; Ananthanarayanan, V. S. Paper presented at the International Symposium on Biomolecular Structure, Conformation, Function and Evolution, Madras, January, 1978.
(6) Piez, K. A. "Biochemistry of Collagen"; G. N. Ramachandran, G. and N., Reddi, A. H., Eds.; Plenum Press: New York, 1976; pp 45–84.
(7) Rapaka, R. S.; Bhatnagar, R. S. Int. J. Pept. Protein Res. 1975, 7, 119–128.

0002-7863/81/1503-1703$01.25/0 © 1981 American Chemical Society

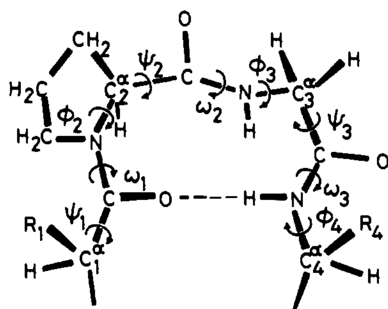

Figure 1. A typical β turn in a tripeptide sequence containing proline in the second and glycine in the third position.

Table II. Final Positional Coordinates ($\times 10^4$) of Nonhydrogen Atoms of *N*-Acetyl-Pro-Gly-Phe-OH[a]

atom	*x*	*y*	*z*
O(1)	5904 (4)	1472 (6)	1944 (3)
C(2)	6950 (6)	2428 (7)	2337 (4)
C(3)	6861 (9)	3371 (9)	3279 (5)
N(4)	8158 (5)	2616 (6)	1924 (3)
C(5)	9404 (8)	3800 (10)	2250 (5)
C(6)	10067 (12)	3902 (13)	1283 (9)
C(7)	9815 (7)	2288 (9)	785 (6)
C(8)	8217 (6)	1777 (6)	965 (4)
C(9)	8141 (5)	−20 (7)	1103 (4)
O(10)	9049 (4)	−679 (6)	1839 (3)
N(11)	7062 (5)	−768 (6)	376 (3)
C(12)	6794 (6)	−2468 (7)	426 (4)
C(13)	5797 (5)	−2967 (6)	1135 (4)
O(14)	5440 (5)	−4385 (5)	1147 (3)
N(15)	5362 (4)	−1866 (6)	1721 (3)
C(16)	4397 (5)	−2228 (6)	2420 (3)
C(17)	5070 (5)	−1467 (7)	3502 (4)
C(18)	6525 (6)	−2250 (7)	4125 (4)
C(19)	6510 (7)	−3203 (10)	4981 (5)
C(20)	7868 (9)	−3904 (10)	5557 (6)
C(21)	9269 (7)	−3679 (9)	5291 (6)
C(22)	9281 (6)	−2776 (10)	4431 (5)
C(23)	7925 (6)	−2066 (9)	3861 (4)
C(24)	2766 (5)	−1582 (7)	1952 (3)
O(25)	2427 (5)	−735 (7)	1194 (3)
O(26)	1762 (4)	−2069 (6)	2467 (3)

[a] Coordinates referred to molecule A in Figure 2.

data. The sample concentration ranged from 3 to 10 mg/mL, depending on the solubility of the material in the solvent used. All the deuterated solvents were obtained from Stohler-Isotope Co.

CD Data. CD spectra of the tripeptide as well as of *N*-acetyl-L-phenylalanine ethyl ester were obtained in trifluoroethanol (TFE) on a Jasco J-20 automatic recording spectropolarimeter. Concentrations used were 0.1–0.5 mg/mL. The spectra were scanned in quartz cells of path length 0.01–0.5 cm from 250 to 200 nm. The molar ellipticity $[\alpha]_{molar}$ is expressed in deg cm^2 mol^{-1}.

Results

X-ray Analysis. The structure was determined by the direct method using the MULTAN program.[8] The positions of nonhydrogen atoms were refined first isotropically and then anisotropically by the full-matrix structure-factor least-squares method to a final R value of 0.070 for 1685 observed reflections ($I > 3\sigma(I)$). In the final cycle, all the least-squares shifts were much less than the corresponding standard deviations. The weighting scheme used was of the form $1/F_o$. The atomic scattering factors were taken from Cramer and Waber.[9] The final positional parameters are given in Table II.

The crystal structure of the molecule projected along the a axis is shown in Figure 2. The crystal structure is stabilized by van der Waals interactions and hydrogen bonds. The hydrogen-bonded contact distances in the structure are listed in Table III. The two NH groups and one OH group in the molecule donate one proton each; likewise, all the carboxyl groups of the peptide units accept one proton each. The crystal structure can be described as consisting of layers of tripeptide molecules stacked perpendicular to the c axis. One face of each layer is wholly hydrophobic, consisting enclusively of phenyl rings, whereas the other face has both hydrophobic and hydrophillic regions.

The bond lengths and valency angles in the molecule are given in Figure 3. These dimensions are normal and do not need special comment. As shown in Figure 3, the tripeptide forms a β turn with a 4 → 1 hydrogen bond between the nitrogen atom of the third peptide unit and the carboxyl oxygen atom of the first peptide unit. The hydrogen-bonded N···O distance (2.867 Å) is normal.[10]

A perspective view of the molecule as seen along the normal to the mean plane of the first and the third peptide units is shown in the Figure 4. The dihedral angles, ϕ, ψ, and ω that define the main-chain conformation of the molecule are given in Table IV. The ω angles for the first, the second, and the third peptide units are −176, −177, and ±180°, respectively. These indicate that the first peptide unit is significantly nonplanar whereas the third unit is planar.

It is clear from the observed dihedral angles (Table IV) that the β turn formed by the molecule is one of the type II predicted by Venkatachalam,[11] with the first and the third peptide units nearly coplanar and the second peptide unit perpendicular to this plane.

The pyrrolidine ring in prolyl residues is known to assume widely different conformations in crystal structures of peptides containing them.[12] In the present structure (Table V), the proline ring adopts

Table III. Intra- and Intermolecular Hydrogen-Bonded Contact Distances (Å) in *N*-Acetyl-Pro-Gly-Phe-OH Crystal Structure

N(15)···O(1) (A)[a]	2.867 ± 0.007
N(11)···O(14) (C)	2.846 ± 0.005
O(26)···O(10) (B)	2.626 ± 0.005

[a] A: x, y, z; B: $x-1, y, z$; C: $1-x, y-1, -z$.

Table IV. Backbone Dihedral Angles (Deg) in *N*-Acetyl-Pro-Gly-Phe-OH and in the Pro-Gly Segment of *cyclo*-(Gly-Pro-Gly-D-Ala-Pro)

dihedral angle[a]	*N*-Ac-Pro-Gly-Phe-OH	theor value[b] type II β turn	Pro-Gly segment of cyclic pentapeptide structure[c]
ω_1	−176	±180	
ϕ_2	−59	−60	−52
ψ_2	+128	+120	+126
ω_2	−177	±180	−179
ϕ_3	+81	+80	+74
ψ_3	−6	0	+12
ω_3	+180	±180	+177
ϕ_4	−107		

[a] Numbering of the dihedral angles is done according to Figure 1. The IUPAC–IUB nomenclature convention is followed: *Biochemistry* 1970, 9, 3471. In the fully extended chain, $\phi_i = \psi_i = \omega_i = 180°$. [b] From ref 11. [c] From ref 29.

Table V. Dihedral Angles (Deg) in the Proline Ring in *N*-Acetyl-Pro-Gly-Phe-OH

bonds	dihedral angle	nature of puckering
N-C$^\delta$-C$^\gamma$-C$^\beta$	−30	
C$^\delta$-C$^\gamma$-C$^\beta$-C$^\alpha$	+36	$^\beta T_\gamma$
C$^\gamma$-C$^\beta$-C$^\alpha$-N	−28	
C$^\beta$-C$^\alpha$-N-C$^\delta$	+10	
C$^\alpha$-N-C$^\delta$-C$^\gamma$	+12	

(8) Germain, G.; Main, P.; Woolfson, M. M. *Acta Crystallogr., Sect. A*, **1971**, *A27*, 368–276.
(9) Cramer, D. T.; Waber, J. T. *Acta Crystallogr.* **1965**, *18*, 104–109.
(10) Ramakrishnan, C.; Prasad, V. *Int. J. Protein Res.* **1971**, *3*, 209.
(11) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425–1426.
(12) Kartha, G.; Ambaby, G. *Acta Crystallogr., Sect. B* **1975**, *B31*, 2035–2039.

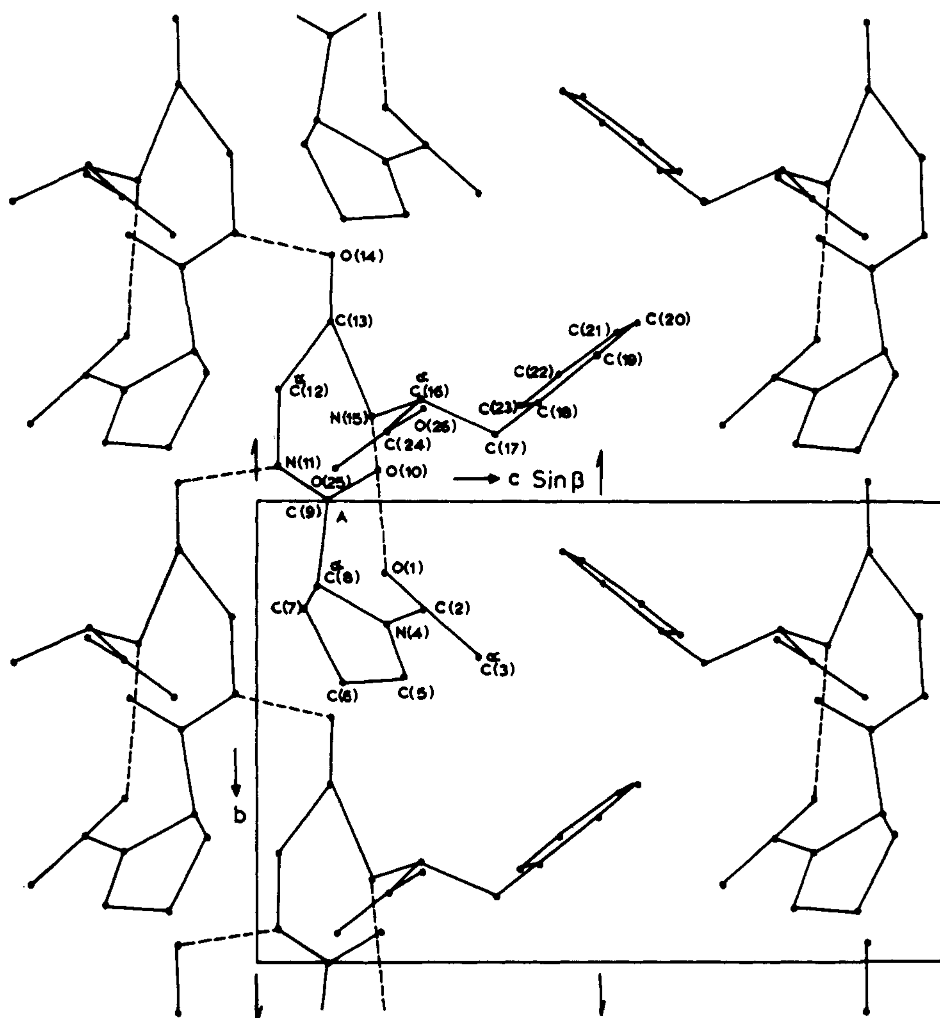

Figure 2. Crystal structure of N-acetyl-Pro-Gly-Phe-OH as viewed along the a axis. The hydrogen bonds are indicated by dashed lines. So that confusion could be avoided, the hydrogen bond that is formed to connect O(26) of the reference molecule to O(10) of the molecule translated by one unit along the a axis is not shown. The positional coordinates are given in Table II for the molecule A.

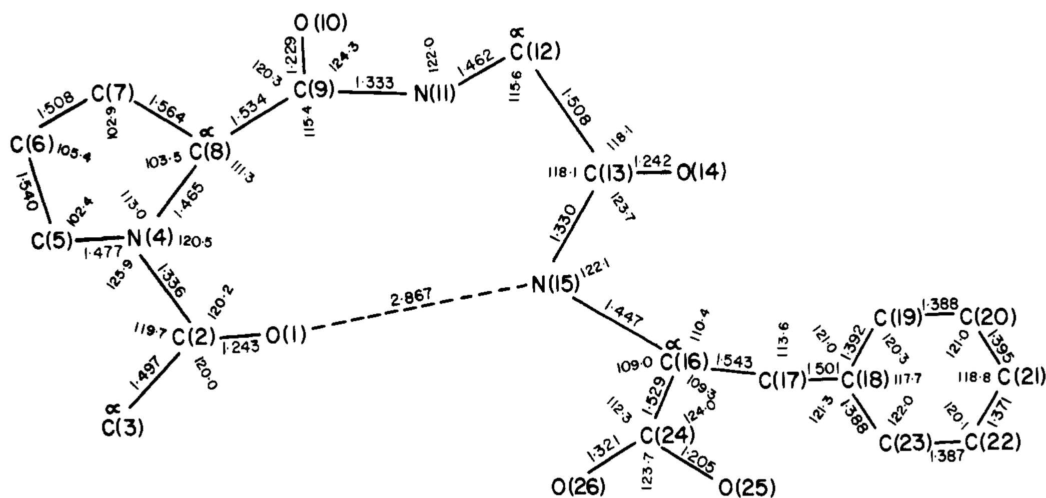

Figure 3. Bond lengths (in Å) and bond angles (in deg) in the Phe-tripeptide molecule. Mean standard deviations for bond lengths and bond angles are 0.008 Å and 0.5°, respectively.

a mixed C^{β}-endo–C^{γ}-exo puckered conformation.[13] The side-chain conformation of phenylalanine is defined[14] by the dihedral angles χ^1 and χ^{21} or χ^{22}. Of these, χ^{21} and χ^{22} are known[15] to have preferred values in the region of ±90°. Thus, the side-chain

(13) (a) Balasubramanian, R.; Lakshminarayanan, A. V.; Sabesan, M. N.; Tegoni, G.; Venkatesan, K.; Ramachandran, G. N. *Int. J. Protein Res.* **1971**, *3*, 25. (b) The relatively large thermal parameters of the C^{γ} atom has been taken to indicate its random thermal agitation as has been demonstrated in proline-containing peptides: Garbay-Jaureguiberry, C.; Arnoux, B.; Prangé, T.; Wehri-Altenburger, S.; Pascard, C.; Roques, B. P. *J. Am. Chem. Soc.* **1980**, *102*, 1827–1837.

(14) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283–437.

(15) Sasisekharan, V.; Ponnuswamy, P. K. *Biopolymers* **1971**, *10*, 583–592.

(16) Brahmachari, S. K.; Ananthanarayanan, V. S.; Brahms, S.; Brahms, J.; Rapaka, R. S.; Bhatnagar, R. S. *Biochem. Biophys. Res. Commun.* **1979**, *86*, 605–612.

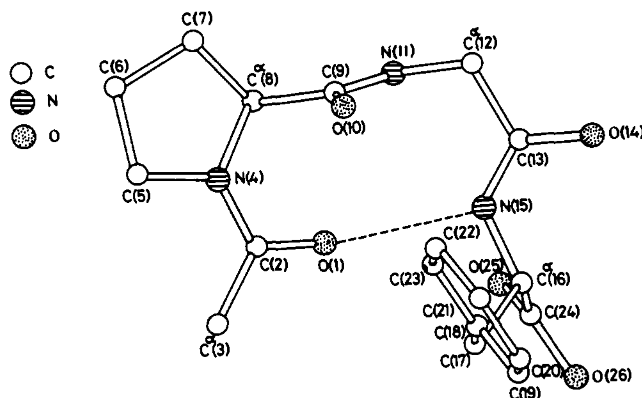

Figure 4. Molecular structure of *N*-acetyl-Pro-Gly-Phe-OH as viewed along the normal to the mean plane of the first and third peptide units. The internal hydrogen bond is indicated by the dashed line.

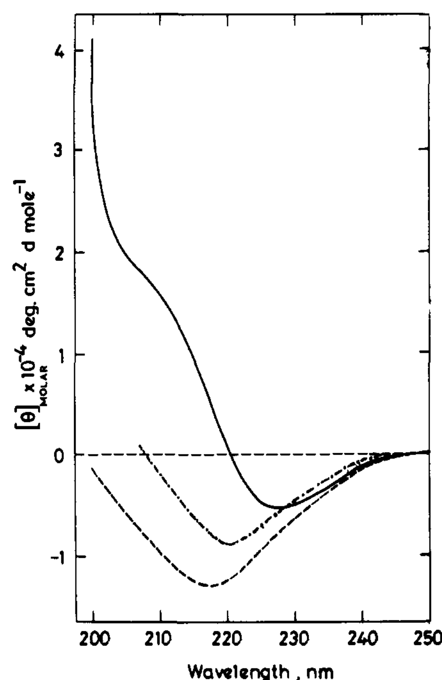

Figure 5. CD spectra of *N*-acetyl-Pro-Gly-Phe-OH in TFE at 25 °C before (—) and after (---) correction for the Phe side-chain contribution (see text) (concentration 0.2 mg/mL). CD spectrum of *N*-acetyl-Pro-Gly-methyl amide (-·-·-) in TFE at 25 °C (redrawn from Stimson et al.[19]) is also given in the figure for comparison.

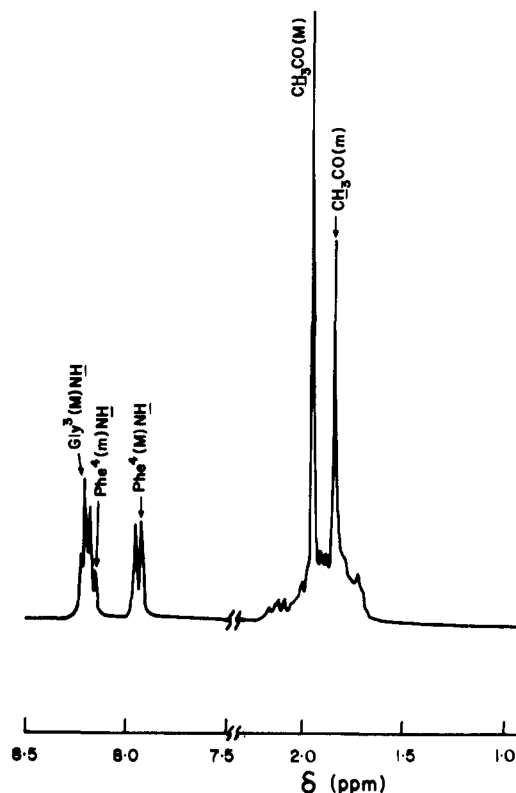

Figure 6. 270-MHz ^{1}H NMR spectrum of the amide region of *N*-acetyl[1]-Pro[2]-Gly[3]-Phe[4]-OH in Me_2SO-d_6. Major (M) and minor (m) resonances of *N*-acetyl protons are also shown.

Table VI. ^{1}H NMR Chemical Shift Data for NH Protons of *N*-Ac[1]-Pro[2]-Gly[3]-Phe[4]-OH[a]

amide proton	chemical shift		
	Me_2SO-d_6	TFE	MeOH
Gly[3] (M) N*H*	8.237, t	7.660, t[b]	8.430, t
Phe[4] (M) N*H*	8.205, d	7.500, d[b]	7.965, d
Phe[4] (m) N*H*	7.980, d		

[a] Recorded at 20 ± 1 °C to an accuracy of ±0.005δ on a 270-MHz spectrometer. Key: d = doublet; t = triplet; m = minor peak; M = major peak. [b] Measured on a 100-MHz spectrometer; ±0.05δ accuracy.

conformation is determined essentially by the value of χ^1 which defines the orientation of the phenyl ring with respect to the two adjacent peptide units. In the present structure, χ^1 has a value of −72° which is close to the −60° corresponding to a conformation with the phenyl ring trans to the terminal carboxyl group and gauche to the nitrogen atom. As is well-known,[14] this conformation leads to the most sterically favorable arrangement.

Solution Studies. While the results presented above demonstrate the presence of a unique type II β-turn conformation for the Phe-tripeptide in the solid state, it was of interest to see whether this conformation persists in solution also. This information was sought to be obtained from CD and ^{1}H NMR spectral studies.

The CD spectrum of the Phe-tripeptide in TFE at 25 °C is shown in Figure 5. TFE was chosen as it is known to stabilize the ordered hydrogen-bonded conformation in peptides and polypeptides.[18,19] The spectrum is seen to be complex due to significant contribution from the phenylalanine side chain in addition to that of the peptide backbone. The observed CD spectrum was then subjected to a correction procedure described by us earlier[20] to eliminate the side-chain contribution. The CD spectrum of the model compound *N*-acetyl-Phe ethyl ester in TFE at 25 °C was subtracted from the tripeptide spectrum on a molar basis with respect to the Phe residue. The corrected spectrum shown in Figure 5 is found to have a broad negative band at 220 nm and resembles that of an analogous peptide *N*-acetyl-Pro-Gly-$NHCH_3$, which has been shown to exist in the β-turn conformation in TFE.[19] However, neither this compound nor the Phe-tripeptide can be taken to be exclusively in the β-turn conformation since their CD spectra only partially resemble that obtained for the nearly 100% β-turn spectrum of *N*-acetyl-Pro-Gly-Leu-OH.[16] It thus appears that the Phe-tripeptide exists as a mixture of β-turn and unordered conformation in solution in the concentration range employed (0.1–0.5 mg/mL). Extension of the CD studies to higher concentrations was hampered by excessive absorption of the phenyl chromophore.

^{1}H NMR Studies. In order to estimate the extent of the hydrogen-bonded β-turn conformation in solution, we resorted to

(17) Boussard, G.; Marraud, M.; Neel, J. *J. Chim. Phys. Phys.-Chim. Biol.* **1971**, *71*, 1081.
(18) Ananthanarayanan, V. S.; Rapaka, R. S.; Brahmachari, S. K.; Bhatanger, R. S. *Biopolymers* **1976**, *15*, 707–716.
(19) Stimson, E. R.; Zimmermann, S. S.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1049–1060.
(20) Brahmachari, S. K.; Ananthanarayanan, V. S.; Rapaka, R. S.; Bhatnagar, R. S. *Biopolymers* **1978**, *17*, 2097–2105.

Table VII. Calculated Mole Fractions of Shielding for *N*-Acetyl-Pro-Gly-Phe-OH from ^{1}H NMR Data

solvent	proton	0° intercept, ppm	$10^3\Delta\delta/\Delta T$, ppm/deg	Y_T[a]
Me_2SO-d_6	Gly^3 (M) N*H*	8.33	−4.7	0.33
	Phe^4 (m) N*H*	8.33	−6.3	≃0.0
	Phe^4 (M) N*H*	8.08	−4.6	0.35
MeOH	Gly^3 N*H*	8.59	−7.4	0.02
	Phe^4 N*H*	8.08	−5.6	0.42

[a] Mole fraction of shielding. Assuming $\Delta\delta/\Delta T$ values for fully shielded and nonshielded N*H* protons, respectively, as follows: -2×10^{-3} ppm/deg and -6×10^{-3} ppm/deg in Me_2SO-d_6; -3.0×10^{-3} ppm/deg and -7.5×10^{-3} ppm/deg in MeOH (see text).

^{1}H NMR spectroscopy of the Phe-tripeptide. Several recent studies[19,21,22] have made use of NMR spectroscopy to demonstrate the presence and the extent of the β-turn conformation in linear peptides containing proline residues.

The ^{1}H NMr spectra of the Phe-tripeptide were obtained in TFE, methanol (MeOH) and Me_2SO-d_6. The spectrum in Me_2SO-d_6 showing the amide region and the acetyl protons is given in Figure 6. The assignments of the resonances were made by comparing the spectra with the published spectra of the peptides containing the same amino acid residues in similar solvents and by use of area ratios, coupling patterns, and decoupling experiments.[23-25] Table VI summarizes the chemical shift (δ) data of the NH protons in Phe-tripeptide obtained at 20 ± 1 °C in Me_2SO-d_6, TFE, and MeOH.

In TFE and MeOH, the NH protons of the Phe-tripeptide show a single peak. On the other hand, in Me_2SO-d_6 additional minor peaks are observed as shown in Figure 6. This clearly indicates that in TFE and MeOH all the signals correspond to a single major (all trans) conformation. In Me_2SO-d_6, additional conformers could arise due to the *cis* ⇌ *trans* isomerization about the acetyl–Pro bond which is quite commonly observed in similar peptides.[19,25,26]

The results in Table VI show that, on going from TFE to the relatively more polar solvent MeOH, there is a relatively larger downfield shift (δ 0.58) of the Gly^{3}N*H* proton of the tripeptide as compared to the Phe^{4}N*H* proton (δ 0.48). This shows that Gly^{3}N*H* is relatively more exposed to the solvent than Phe^{4}N*H*, which appears to be partly solvent shielded. The latter is likely to arise from *intra*molecular hdyrogen bonding similar to that observed in the solid state (Figure 4). We may rule out any solvent-induced conformational changes in going from TFE to MeOH since the CD spectra of the tripeptide in both these solvents at a given temperature have been found to be identical.

In order to obtain the mole fraction of the tripeptide in the hydrogen-bonded conformation, we have studied the temperature dependence of the chemical shifts of the tripeptide NH protons in Me_2SO-d_6 and MeOH. In both the solvents, the temperature coefficients for solvent-shielded and solvent-exposed NH protons have been documented entensively.[19,22,27] The temperature coefficient, $\Delta\delta/\Delta T$, values as well as the 0° intercept values for the NH protons in Me_2SO-d_6 and MeOH (including, both the major (M) and minor (m) resonances associated with the trans and cis conformers, respectively, in Me_2SO-d_6) are listed in Table VII. Using the $\Delta\delta/\Delta T$ values from Urry and Long[22] for the completely solvent-exposed and solvent-shielded NH protons in MeOH, we estimate that, on an average, the Phe^{4}N*H* of the Phe-tripeptide has 42% of intramolecular hydrogen-bonded (i.e., solvent-shielded) character in contrast to Gly^{3}N*H*, which is almost 100% solvent exposed (see Table VII). A similar estimate in Me_2SO-d_6 was also made and is given in Table VII. In this solvent, besides the slight decrease in the hydrogen-bonded character of the Phe^{4}N*H* proton (with *trans* acetyl-Pro bond), one observes a significant extent of solvent shielding of the Gly^{3}N*H* proton also. The latter could arise from hydrogen bonding involving the acetyl carbonyl oxygen and Gly^{3}N*H* proton (i.e., the C_7 conformation,[19] with *trans* acetyl-Pro bond[28]).

Discussion

The results obtained from the X-ray crystallographic analysis of the Phe-tripeptide demonstrate the presence of a type II β-turn conformation in this compound. It may be mentioned that this is the first time such a β turn involving Pro and Gly as the middle residues has been observed at atomic resolution in a linear oligopeptide. The observed values of ϕ and ψ for this peptide is the β-turn region (ϕ_2 = −59°, ψ_2 = 128°, ψ_3 = 81°, ψ_3 = −6°) are indeed so close to those predicted by Venkatachalam[11] for the classical β turn of type II (ϕ_2 = 60°, ψ_2 = 120°, ψ_3 = 80°, ψ_3 = 0°) that the present structure provides a good description of this conformation. The only other instance where a β turn involving a Pro-Gly sequence has been observed in atomic details is in the crystal structure of the cyclic peptide *cyclo*-(Gly_1-Pro_2-Gly_3-D-Ala_4-Pro_5).[29] The main-chain dihedral angles in the Pro_2-Gly_3 turn region of this cyclic peptide are also close to their classical values as in the present structure, in spite of the differences between the two structures, namely, the constraints imposed by the ring in the cyclic peptide and the involvement of the Pro-Gly peptide bond in *inter*molecular hydrogen-bond network in the linear peptide. The crystal structures of three linear peptides containing Pro-Gly sequences, namely, *t*-Boc-Pro-Gly-OH, *N*-Piv-Pro-Gly-NMe_2, and Leu-Pro-Gly-OH, have also been reported recently.[30] None of these molecules assumes a β-turn conformation in spite of the presence of Pro-Gly sequence. This clearly indicates that the presence of two β-turn-favoring residues in the second and third position is a necessary but not sufficient condition for the peptide segment containing them to be stabilized in the β-turn conformation in the respective crystals. This is found to be also true in globular proteins where the stability of the β-turn conformation is significantly dependent on the nature of the residues adjoining the Pro-Gly segment.[5]

An insight into the stability of the β-turn conformation for the Phe-tripeptide in solution can be obtained from the CD and NMR results. However, there are intrinsic difficulties in deducing the preferred conformation(s) of a linear oligopeptide in solution due to the multiplicity of available conformational states. We have, in our treatment of the CD and NMR data, chosen to tackle this problem using reference values for the spectral parameters of particular selected conformations. This approach is necessarily limited in quantitative terms since the chemical shifts and CD parameters can be influenced by the nature of the specific compound or conformation (especially here where an aromatic ring is present). In addition, the populations of the various conformers (the cis and various trans structures) are perturbed by temperature variations. The calculated populations from the spectral data

(21) Kopple, K. D.; Go, A. "Peptides: Chemistry, Structure and Biology"; Walker, R., Meienhofer, J., Eds.; Ann Arbour Science Publisher: Ann Arbour, MI, 1975; pp 139.

(22) Urry, D. W.; Long, M. M. *CRC Crit. Rev. Biochem.* **1976**, *4*, 1–45.

(23) Torchia, D. A. *Macromolecules* **1972**, *5*, 566–569.

(24) (a) Pogliani, L.; Ellenberger, M.; Valat, J.; Bellocq, A. M. *Int. J. Pept. Protein Res.* **1975**, *7*, 345–360. (b) Young, P. E.; Deber, C. M. *Biopolymers* **1975**, *14*, 1547–1549.

(25) Higashijima, T.; Mitsuotasumi; Miyazawa, T. *Biopolymers* **1977**, *16*, 1259–1270.

(26) (a) That the isomerization involves the acetyl–Pro bond is seen from the presence, in Me_2SO-d_6, of the minor resonance of the Gly^{4}N*H* in *N*-acetyl-Pro-Gly-Gly-OH and its absence in Pro-Gly-Gly-OH. In the case of *N*-acetyl-Pro-Gly-Phe-OH, the major and minor resonances of Phe^{4}N*H* tend to coalesce on heating above 80 °C (Brahmachari, S. K. 1979, Ph.D. Thesis). (b) The assignment of the resonances to the cis and trans conformers were made by assuming that the single resonance observed in TFE and MeOH corresponds to the trans conformer (that gives rise to the β-turn CD of Figure 5) and that the additional minor resonance that emerges on the addition of increasing amounts of Me_2SO-d_6 is due to the presence of the cis isomer. These assignments are compatible with the observed smaller temperature coefficient of the chemical shift of the major (trans) resonance when compared with the minor (cis) resonance, the trans conformer being stabilized more by *intra*molecular hydrogen bonds (see below).

(27) Kopple, K. D. *Biopolymers* **1971**, *10*, 1139–1152.

(28) We thank one of the reviewers for suggesting this possibility.

(29) Karle, I. *J. Am. Chem. Soc.* **1978**, *100*, 1286–1289.

(30) Benedetti, E. *Pept., Proc. Am. Pept. Symp., 5th* **1977**, 257–273.

should be viewed in the light of these limitations and taken at best as semiquantitative. The CD data point out that this tripeptide is not completely ordered in TFE solution but exists as a mixture of the β-turn and disordered conformation. From NMR data, the extent of the hydrogen-bonded conformation is found to be about 40% in MeOH or Me_2SO-d_6. It is interesting to compare this value with that obtained for the analogous *N*-acetyl-Pro-Gly-$NHCH_3$, viz., about 68% in Me_2SO-d_6.[19] The lower value for the Phe-tripeptide indicates that the Phe side chain tends to destabilize the β-turn conformation formed by the backbone in solution. In the solid state, however, this effect appears to be offset by intermolecular hydrogen bonding (Figure 2). In globular proteins, the interaction of the Phe side chain with other aromatic side chains may stabilize this conformation. In these proteins, Phe is found to have relatively high preference for the fourth position of the β turn in the R_1-Pro_2-R_3-R_4 sequence comparable to Leu.[2] The results presented here are also of relevance to the conformational requirement of prolyl hydroxylase which hydroxylates selected Pro residues in the Pro-Gly-X segments of nascent procollagen. Two of the authors have recently shown[3] that the β turn is a conformational requirement in these segments for enzymic proline hydroxylation. It is interesting to note that all the seven segments that contain the Pro-Gly-Phe sequence in procollagen are hydroxylated.[6] It is also worth noting that out of the three Pro residues in bradykinin, that in the -Pro-Gly-Phe-segment gets hydroxylated.[31]

Acknowledgment. This work was supported by grants from USPHS-PL480 and the Department of Science and Technology (India) at Bangalore, from USPHA at California, and from the Medical Research Council of Canada. We thank Dr. T. P. Singh for help in the initial stages of the crystallographic studies and the Bangalore FT-NMR Facilities for use of the 270-MHz NMR spectrometer.

Supplementary Material Available: Listings of the observed and calculated structure factors (Table A) and anisotropic thermal parameters (Table B) (8 pages). Ordering information is given on any current masthead page.

(31) Cardinale, G. J.; Udenfriend, S. *Adv. Enzymol. Relat. Areas Mol. Biol.* 1974, *41*, 245-300.

Topography of Cyclodextrin Inclusion Complexes. 15.[1] Crystal and Molecular Structure of the Cyclohexaamylose-7.57 Water Complex, Form III. Four- and Six-Membered Circular Hydrogen Bonds

K. K. Chacko[2] and W. Saenger*

Contribution from the Abteilung Chemie, Max-Planck-Institut für Experimentelle Medizin, D-3400 Göttingen, West Germany. *Received August 11, 1980*

Abstract: Cyclohexaamylose, α-cyclodextrin (α-CD), is a doughnut-shaped molecule composed of six $\alpha(1\rightarrow4)$ linked glucoses. Owing to an annular aperture of 5.0 Å, it is able to form inclusion complexes with a variety of substrate molecules. Two crystal modifications with water as included guest have already been described. A third crystal modification of α-CD with water included was obtained while crystallizing α-CD from 1.2 M $BaCl_2$. The space group in all three hydrate modifications is orthorhombic, $P2_12_12_1$. The cell dimensions of the present crystals are a = 14.356 (5), b = 37.538 (12), and c = 9.400 (4) Å, with four formula units α-CD·7.57H_2O per cell. Of the 7.57 hydration water molecules/asymmetric unit, 5 are located outside the cyclodextrin cavity with an extensive hydrogen bonded network between themselves as well as with the hydroxyl oxygen atoms of α-CD. The remaining 2.57 water molecules are enclosed within the cavity and disordered with four alternative sites of approximately 0.64 occupancy each. There are three "circular" hydrogen bonds observed for this structure involving three of the external water molecules and O-H groups of α-CD. Two of the circles, one six membered and one four membered, are *antidromic*, and one is four-membered *homodromic*. The macrocyclic conformation of α-CD is nearly symmetrical with all the possible O(2)···O(3) intramolecular hydrogen bonds formed. In terms of "induced fit" type complex formation of this model enzyme, the α-CD molecule exists in a relaxed state with "activated" water inside the cavity. In contrast, α-CD in the other two hydrate modifications is observed in a tense state with only four or five of the six O(2)···O(3) hydrogen bonds formed and one glucose rotated to fix the included waters in fully occupied positions by hydrogen bonding to O(6) hydroxyls.

Cyclodextrins are doughnut-shaped, cyclic oligosaccharides obtained by the action of glucanotransferase on the helical starch molecule. They consist of six or more $\alpha(1\rightarrow4)$-linked glucopyranose rings. The hexasaccharide α-cyclodextrin (α-CD) has an annular aperture of 5.0 Å, and hence it is able to form inclusion complexes with a variety of guest molecules small enough to fit inside its cavity. A number of nonisomorphous crystalline modifications of several α-CD inclusion complexes with different guest molecules are known.[1-3] Two hexahydrate modifications of the "empty" α-CD, i.e., the species occurring in aqueous solution without added guest molecule, have already been described.[4,5] In one of these, two water molecules are located inside the cavity.[4] In the second modification,[5] only one water molecule and the primary O(6) hydroxyl of an adjacent α-CD molecule are in the center of the cavity with five of the remaining water molecules located outside. The above hexahydrate modifications have only four[4] and five[5] of the possible intramolecular O(2)···O(3) hydrogen bonds between adjacent glucoses. Here we report of a third crystal modification of α-CD with five water molecules located outside

(1) For part 14, see K. Linder and W. Saenger, *Biochem. Biophys. Res. Commun.*, 92, 933-938 (1980).
(2) W. Saenger in "Environmental Effects on Molecular Structure and Properties", B. Pullman, Ed., D. Reidel Publishing Co., The Netherlands, pp 265-305 (1976); W. Saenger, M. Noltemeyer, P. C. Manor, B. Hingerty, and B. Klar, *Bioorg. Chem.* 5, 187-195 (1976).
(3) R. K. McMullan, W. Saenger, J. Fayos, and D. Mootz, *Carbohyd. Res.* 31, 37-46 (1973).
(4) P. C. Manor and W. Saenger, *J. Am. Chem. Soc.* 96, 3630-3639 (1974).
(5) K. Lindner, Ph.D. Thesis, Universität Göttingen, 1980.

0002-7863/81/1503-1708$01.25/0 © 1981 American Chemical Society