CONFORMATIONAL RESTRICTIONS OF PEPTIDES *via* BACKBONE MODIFICATION: SOLUTION AND CRYSTAL-STATE ANALYSIS OF Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 (*) (**)

ANNA MARIA PIAZZESI, RENATO BARDI, MARCO CRISMA, GIAN MARIA BONORA and CLAUDIO TONIOLO (°)
CNR, Centro di Studio sui Biopolimeri, Dipartimento di Chimica Organica, Università di Padova, I-35131 Padova, Italy

VIRANDER SINGH CHAUHAN and PARAMJEET KAUR
International Center for Genetic Engineering and Biotechnology, New Delhi, 110067 India

KUCHIBHOTLA UMA and PADMANABHAN BALARAM
Molecular Biophysics Unit, Indian Institute of Science, Bangalore, 560012 India

Summary — An N^{α}-protected model tripeptide amide containing, in the central position, an α,β-dehydrophenylalanine (Z-configurational isomer), Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 (Boc, *tert*-butyloxycarbonyl), has been synthesized by solution methods and fully characterized. IR absorption and ^1H NMR studies provided evidence for the occurrence of a significant population of a conformer containing two consecutive, intramolecularly H-bonded (type II-III') β-bends in solution. However, an X-ray diffraction analysis clearly indicates that only the type-II β-bend structure survives in the crystal state.

In the last few years α,β-dehydro α-amino acids (ΔAA) have become one of the most promising conformational restrictions in the study of structure-activity relationships in peptides. A number of peptide analogues substituted with ΔAA residues have been prepared, showing significantly enhanced biological potency and/or metabolic stability with respect to their saturated counterparts. Several review-articles have appeared, covering the occurrence, synthesis, and biological properties of ΔAA-containing peptides[2-5]. In addition, a large body of conformational studies have been devoted to this type of peptide modification in view of its numerous potential consequences: (*i*) rigid side-chain orientation, (*ii*) change in the α-carbon geometry, and (*iii*) perturbation in the electronic distribution of the adjacent-CONH- bonds reflecting their conjugation with the π-electrons of the $C^{\alpha}=C^{\beta}$ double bond (for recent reviews and articles see refs. 6-11).

In this report we describe the solution and crystal-state structural analysis (by IR absorption, ^1H NMR and X-ray diffraction) of a well-characterized, N^{α}-protected, synthetic model tripeptide amide containing a central ΔPhe (Z-configurational isomer) residue, namely Boc-L-Pro-Δ^Z-Phe-Gly-NH_2. The results obtained allowed us to establish that the L-Pro-Δ^Z-Phe sequence has a strong tendency to form a type-II β-bend conformation. The folding of the Δ^Z-Phe-Gly sequence into a type-III' β-bend conformer, however, is less stable, being significantly populated only in solution.

(*) This is part CCXXXIII in the series «Linear Oligopeptides». Paper CCXXXII in this series is ref. 1.

(**) Atomic coordinates for this structure have been deposited with the Cambridge Crystallographic Data Centre and can be obtained on request from: The Director, Cambridge Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

(°) To whom correspondence should be addressed.

EXPERIMENTAL

PREPARATION

Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 - To a solution of Boc-L-Pro-Δ^Z-Phe-OH azlactone[8] (0.9 g, 2.5 mmol) in tetrahydrofuran (10 ml) a solution of HBr·H-Gly-NH_2 [prepared from Z-Gly-NH_2[12] (0.68 g, 3 mmol) with a 30% solution of HBr in acetic acid] and *N*-methylmorpholine (0.35 ml, 3 mmol) in tetrahydrofuran (10 ml) was added, and the mixture stirred at room temperature for 24 h. The solvent was removed *in vacuo*, and the remaining oily residue, dissolved in ethyl acetate (30 ml), was washed successively with saturated $NaHCO_3$, 10% citric acid, water, and dried (Na_2SO_4). The solvent was removed under reduced pressure to give the tripeptide which was purified on a silica gel column (30 × 2.5 cm) using a 1:1 chloroform-petroleum ether solution as eluant, followed by recrystallization from ethyl acetate-petroleum ether to give 0.5 g (46%) of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2. M.p. 223-4 °C; $[\alpha]_D^{22} = -131.5$ (c=1.15 mg cm^{-3}, MeOH); R_{f1}=0.50; R_{f2}=0.80. *Elem. anal.*, found % (calcd. for $C_{21}H_{28}N_4O_5$): C, 60.7 (60.6); H, 6.8 (6.8); N, 13.5 (13.5). ^1H NMR (400 MHz, $CDCl_3$), δ: 7.81 (t, 1 H, Gly, NH); 7.62 (s, 1 H, ΔPhe NH); 7.50-7.35 (m, 6 H, ΔPhe aromatic protons and ΔPhe $C^{\beta}H$); 7.04 and 5.33 (2 s, 2 H, primary amide NH_2); 4.20 (t, 1 H, Pro $C^{\alpha}H$); 4.11-3.89 (m, 2 H, Gly CH_2); 3.50 (m, 2 H, Pro $C^{\delta}H_2$); 2.27-1.94 (m, 4 H, Pro $C^{\beta}H_2$ and $C^{\gamma}H_2$); 1.43 (s, 9 H, Boc CH_3). ^{13}C NMR (50 MHz, $CDCl_3$ + CD_3OD), δ: 175.37 (Pro C=O); 173.98 (ΔPhe C=O); 166.32 (Gly C=O); 156.28 (Boc C=O); 134.15 (ΔPhe α-C), 132.2 (ΔPhe β-C); 129.16 (ΔPhe aromatic carbons); 81.54 (Boc quaternary carbon); 60.99 (Pro α-C); 49.01 (Pro δ-C); 43.48 (Gly α-C); 30.22 and 25.99 (Pro β-C and γ-C); 28.04 (Boc methyls). IR absorptions ($CDCl_3$): 3488 (primary amide NH), 3452 (sh), 3405 (sh), 3350 (peptide N-H); 1673, 1638 cm^{-1} (urethane, peptide and amide C=O, and C=C). UV absorption (MeOH), λ_{max}: 278 nm (ε = 17,170).

METHODS

IR absorption spectra were recorded using a Perkin-Elmer model 580 B spectrophotometer equipped with a Perkin-Elmer

model 3600 IR data station. The band positions are accurate to ± 1 cm^{-1}. Cells with path lengths 0.1, 1.0 and 10 mm (with CaF_2 windows) were used. Spectrograde deuteriochloroform (99.8% deuteriated) was purchased from Merck.

The ^1H NMR spectra were recorded by using a Bruker AM-400 spectrometer. Measurements were carried out in deuteriochloroform (99.96% deuteriated; Merck) and bis(trideuteriomethyl) sulphoxide (99.96% deuteriated; Stohler) with tetramethylsilane as the internal standard.

X-Ray diffraction

Colourless crystals of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 were grown by slow evaporation from a methanolic solution. X-Ray diffraction data were collected on a Philips PW 1100 four-circle diffractometer from a crystal of approximate dimensions 0.1 × 0.1 × 0.3 mm. Accurate unit-cell parameters and crystal orientation matrices (together with their e.s.d.'s) were obtained from least-squares refinement of the 2θ, ω, χ, and φ values of 25 carefully centred reflections with $9 < \theta < 14°$. The h, k, l ranges measured were –24 to 24, 0 to 6, and 0 to 9, respectively. The θ/2θ scan mode (scan width 1.20°, scan speed 0.02° s^{-1}) and a Mo-K_α radiation monochromatized by a graphite crystal (λ = 0. 71069 Å) were used. During data collection three standard reflections ($\bar{2}22$, 312, 511) were measured every 180 min to check the stability and orientation of the crystal and the electronics.

2081 unique reflections were obtained up to 2θ=50°, of which 1276 have intensities equal to or greater than 3σ(I). Intensities were corrected for Lorentz and polarization effects and put on an absolute scale by Wilson's method. No absorption correction was applied. The structure was solved by direct methods using MULTAN 80[13]. The crystallographic data are summarized in table 1.

TABLE 1 - CRYSTAL DATA FOR Boc-L-Pro-Δ^Z-Phe-Gly-NH_2

Molecular formula	$C_{21}H_{28}N_4O_5$
Molecular weight	416.5
d (calcd.), $g.cm^{-3}$	1.268
d (exptl.), $g.cm^{-3}$	1.25
Crystal system	monoclinic
Space group	$P2_1$
Z	2
a, Å	21.354(5)
b, Å	5.709(3)
c, Å	9.008(3)
ß, deg	96.7(1)
Reflections (measured)	2081
Reflections [$I \geq 3\sigma(I)$]	1276
R value	0.053
R_w value	0.044

The E maps of the set of phases with the best combined figure of merit showed several fragments of the molecule; the remaining parts were subsequently revealed by Fourier difference maps. Refinement was carried out by block-matrix least-squares allowing the non-H atoms to vibrate anisotropically. The scattering factors were taken from the International Tables for X-ray Crystallography[14]. The H atoms were localized in the difference Fourier maps, but they were introduced in calculated positions and refined in the last cycle. The quantity minimized was $\Sigma w(|F_o|-|F_c|)^2$ with $w = [\sigma^2(F_c + 0.00015\,F_o^2]^{-1}$. For all calculations the SHELX-76 program[15] was used. The final conventional unweighted R factor was 0.053 and the R factor weighted by w was 0.044; $(\Delta/\sigma)_{max}$ in final refinement cycle for positional parameter of the non–H atoms 0.20; maximum and minimum heights in final difference Fourier syntheses ± 0.2 e $Å^{-3}$. Table 2 gives the final positional parameters and equivalent temperature factors for non-H atoms. Additional tables (structure factors, final positional parameters for H atoms) have been deposited as *Supplementary Material*.

TABLE 2 - FINAL POSITIONAL PARAMETERS (×10^4) AND EQUIVALENT TEMPERATURE FACTORS (×10^3) OF NON-HYDROGEN ATOMS FOR Boc-L-Pro-Δ^Z-Phe-Gly-NH_2

Atom	x/a	y/b	z/c	U_{eq} [a]
O(1)	3331(2)	7105(8)	5510(4)	57(2)
O(2)	3281(2)	6833(8)	8015(4)	51(1)
O_1	1879(2)	4775(8)	7916(4)	54(1)
O_2	2297(2)	1175(7)	11636(4)	43(1)
O_3	4270(2)	4012(9)	10418(5)	65(2)
N_1	2498(2)	8464(9)	6448(4)	39(2)
N_2	2086(2)	6951(8)	10027(5)	34(1)
N_3	2972(2)	3338(8)	10480(5)	41(2)
N_4	4405(2)	156(11)	9943(7)	86(3)
C(1)	3704(4)	3150(14)	5794(11)	146(5)
C(2)	4013(3)	5947(17)	3850(8)	112(4)
C(3)	4434(3)	6397(14)	6519(8)	90(3)
C(4)	3879(3)	5581(10)	5439(7)	61(3)
C(5)	3053(3)	7405(10)	6769(7)	43(2)
C_α	2091(2)	8881(9)	7600(6)	37(2)
C_β	1462(3)	9572(11)	6728(6)	48(2)
C_γ	1651(3)	10630(11)	5299(6)	51(2)
$C_{\delta 1}$	2196(3)	9081(11)	4945(6)	54(2)
$C_{\delta 2}$	2012(2)	6648(11)	8505(6)	38(2)
C_α	1972(2)	5039(10)	10980(5)	31(2)
C_β	1511(2)	4949(10)	11826(5)	38(2)
C_δ	992(2)	6589(10)	12013(6)	36(2)
$C_{\delta 1}$	609(3)	6074(11)	13120(7)	54(2)
$C_{\delta 2}$	102(3)	7461(15)	13351(8)	72(2)
C_ζ	-52(3)	9359(14)	12466(9)	70(3)
$C_{\epsilon 1}$	322(3)	9936(12)	11370(7)	62(3)
$C_{\epsilon 2}$	842(3)	8564(11)	11141(7)	50(2)
C_α	2433(2)	3042(11)	11045(6)	35(2)
C_α	3408(2)	1364(11)	10519(6)	55(3)
C_3	4068(3)	2014(15)	10295(7)	55(2)

[a] $U_{eq} = \frac{1}{3}\Sigma_i \Sigma_j U_{ij}\ a_i^*\ a_j^*\ a_i\ a_j$.

RESULTS AND DISCUSSION

Solution conformation

The preferred conformation in solution of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 has been investigated in solvents of different polarity ($CDCl_3$ and DMSO) at various concentrations by using IR absorption and ^1H NMR techniques.

The IR absorption spectra (not shown) were taken over the concentration range 0.2 - 20 mM in $CDCl_3$ solution. In the amide A region a strong absorption is seen near 3490 cm^{-1} (free primary amide, first N-H stretching mode)[16] followed by shoulders at approximately 3450 cm^{-1} (free peptide N-Hs)[12-14] and 3400 cm^{-1} (free primary amide, second N-H stretching mode)[12], and by a much more intense absorption near 3350 cm^{-1} (H-bonded N-Hs)[16-18]. The latter band is largely concentration independent and, therefore,

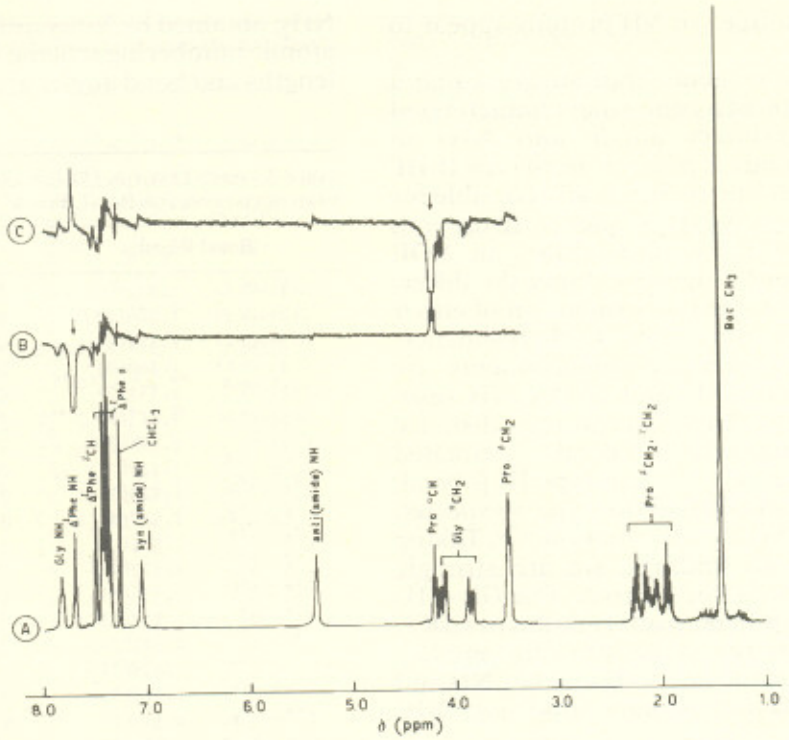

Fig. 1 - (A) 400 MHz ^{1}H NMR spectrum of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 in $CDCl_3$ (conc. 10 m*M*); (B) difference NOE spectrum obtained by saturation of the Δ^Z-Phe NH resonance; (C) difference NOE spectrum obtained by saturation of the L-Pro C^α-H resonance.

may be assigned to intramolecularly H-bonded conformers. In the C=O stretching region the absence of a band or shoulder at 1710-1690 cm^{-1} (at 2 m*M* concentration) indicates that the tertiary urethane carbonyl is intramolecularly H-bonded[19]. In this region, in addition to an absorption near 1670 cm^{-1} (H-bonded tertiary urethane, and free peptide and amide carbonyls), a band is visible at about 1640 cm^{-1} (H-bonded peptide carbonyls)[16,19,20]. In summary, we suggest that an intramolecularly H-bonded structure may in fact predominate in $CDCl_3$ solution for this N^α-Boc tripeptide amide.

In order to get more detailed information on the preferred conformation of the N^α-protected tripeptide amide in solution a ^{1}H NMR study was performed in $CDCl_3$ and DMSO. The resonance assignments in the ^{1}H NMR spectra of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 (for $CDCl_3$ see figure 1) are straightforward and made on the basis of chemical shift values, peak multiplicities, homonuclear spin decouplings, and a complete solvent titration. In DMSO (not shown) there is an additional isomer, corresponding to the *cis* orientation about the tertiary urethane linkage[19]. The population of *cis-trans* is about 30:70.

The delineation of inaccessible (or intramolecularly H-bonded) NH groups in $CDCl_3$ solution was carried out by using solvent (DMSO) dependence of NH chemical shifts at low concentration (3 m*M*). From figure 2A it is clear that the chemical shifts of the Gly and primary amide *anti* NH protons are only slightly sensitive to the addition of the strong H-bonding acceptor solvent DMSO[21], a behaviour characteristic of shielded protons. Conversely, the

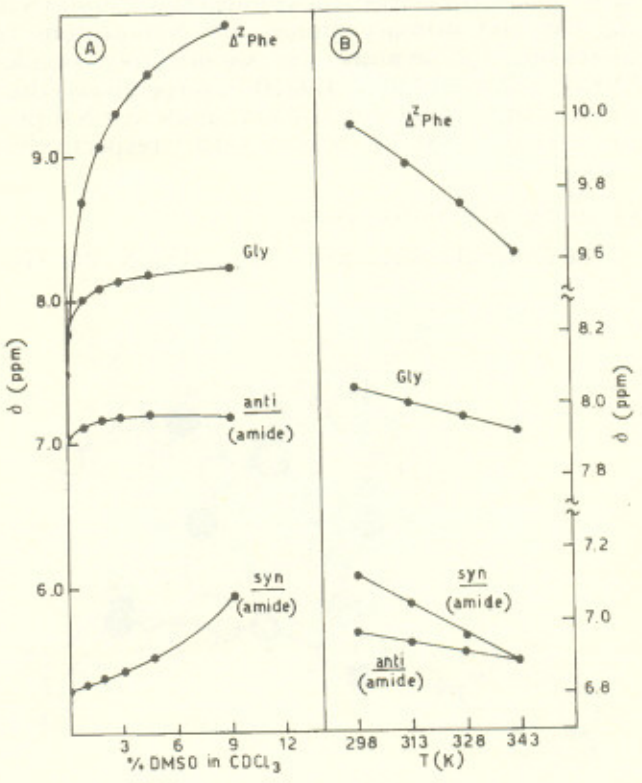

Fig. 2 - Plot of NH chemical shifts in the ^{1}H NMR spectrum of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 (A) in $CDCl_3$ solution (conc. 3 m*M*) as a function of increasing percentages of DMSO (*v/v*), and (B) in DMSO (conc. 6 m*M*) as a function of heating.

ΔPhe and primary amide *syn* NH protons appear to be solvent exposed.

The spectroscopic evidence thus suggests that a significant population of a conformer characterized by the Gly and primary amide *anti* NHs in intramolecular H-bonds is present. Both type II-III' and III-III β-bents are stereochemically possible for the Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 sequence. In order to distinguish between these possibilities an NOE study was undertaken[22]. Figure 1 shows the difference NOE spectra obtained by irradiation of either the Δ^Z-Phe NH or the L-Pro $C^\alpha H$ resonance. Significant (≃ 10%) intensity enhancements are observed on the L-Pro $C^\alpha H$ and Δ^Z-Phe NH resonances, respectively. The interproton distance $C_{i+1}H$-$N_{i+2}H$ in an ideal type-II β-bend is estimated to be = 2.2 Å, while it is = 3.5 Å in type-III β-bend. Therefore, a significant NOE of this type is expected only in the type-II β-bend conformation[23,24]. Taking together all the above findings, we are strongly inclined to conclude that Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 favours a type-II-III' β-bend conformation in $CDCl_3$ solution, stabilized by two consecutive intramolecular H-bonds between the Boc C=O and Gly NH and between the Pro C=O and primary amide *anti* NH, respectively.

A study of the temperature dependences of NH chemical shifts in DMSO (at 6 m*M* concentration) (figure 2B) clearly suggests that this highly folded conformation is still the prevailing one for the *trans*-urethane form of the tripeptide even in this polar solvent. In fact, the sensitivities of the chemical shifts of the Gly and primary amide *anti* NH protons to increasing temperature are significantly smaller ($\Delta\delta/\Delta T = -2.7\times10^{-3}$ and -1.9×10^{-3}, respectively) than those of the Δ^Z-Phe and primary amide *syn* NH protons ($\Delta\delta/\Delta T = -8.0\times10^{-3}$ and -5.3×10^{-3}, respectively).

Crystal-state conformation

The molecular structure of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2, obtained by X-ray diffraction analysis, with the atomic numbering scheme is shown in figure 3. Bond lengths and bond angles are given in table 3. The torsion angles[25] are listed in table 4. The geometry of the intra- and intermolecular H-bonds is reported in table 5. Bond lengths and bond angles are in agreement with previously described results for the geometry of the Boc urethane[26], and amide groups[27], Pro[19,28-30], Δ^Z-Phe[31-34] and Gly residues, and the peptide unit[35]. However, in the Boc-protecting group the *tert*-butyl C(1)-C(4), C(2)-C(4), and C(3)-C(4) bonds show some shortening, being associated with atoms having high thermal parameters [the U_{eq} values for the C(1), C(2), and C(3) atoms are in the range of 90 to 146 Å^2].

The existence of the C^α_2=C^β_2 bond in the Δ^Z-Phe

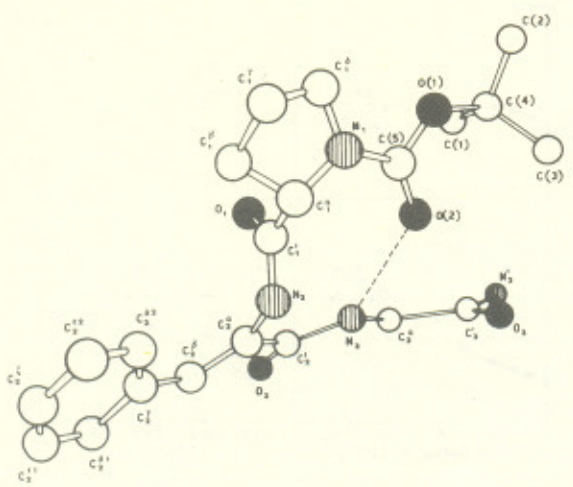

Fig. 3 - Molecular structure of Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 with numbering of the atoms. The intramolecular H-bond is shown as a dashed line.

TABLE 3 - BOND LENGTHS (Å) AND ANGLES (deg) AND ESTIMATED STANDARD DEVIATIONS FOR Boc-L-Pro-Δ^Z-Phe-Gly-NH_2

Bond lengths		Bond angles	
C(1)-C(4)	1.482(9)	C(1)-C(4)-C(2)	114.5(6)
C(2)-C(4)	1.506(10)	C(2)-C(4)-C(3)	110.3(5)
C(3)-C(4)	1.515(9)	C(1)-C(4)-C(3)	110.3(6)
C(4)-O(1)	1.466(6)	C(1)-C(4)-O(1)	108.9(6)
O(1)-C(5)	1.351(7)	C(2)-C(4)-O(1)	101.6(4)
C(5)-O(2)	1.215(7)	C(3)-C(4)-O(1)	110.9(4)
C(5)-N	1.332(7)	C(4)-O(1)-C(5)	123.0(4)
N - C^1_α	1.449(7)	O(1)-C(5)-O(2)	125.0(5)
C^1_α - C^1_β	1.525(7)	O(1)-C(5)-N	110.1(5)
C^1_β - C^1_γ	1.519(8)	O(2)-C(5)-N^1	124.9(6)
C^1_γ - C^1_δ	1.524(9)	C(5)-N^1 - C^1_α	120.8(4)
C^1_δ - N^1	1.472(7)	C(5)-N^1 - C^1_δ	126.2(5)
C^1_α - $C^1_\prime$	1.533(8)	N - C^1_α - C^1_β	103.8(4)
$C^1_\prime$ - O^1	1.213(7)	C^1_α - C^1_β - C^1_γ	103.5(4)
$C^1_\prime$ - N^1	1.372(7)	C^1_β - C^1_γ - C^1_δ	103.4(5)
N^1 - C^2_α	1.428(7)	C^1_γ - C^1_δ - N^1	102.0(4)
C^2_α - C^2_β	1.315(7)	C^1_δ - N^1 - C^1_α	112.6(4)
C^2_β - C^2_1	1.475(7)	N^1 - C^1_α - $C^1_\prime$	110.9(4)
C^2_1 - $C^2_{\zeta_1}$	1.393(8)	C^1_β - C^1_α - $C^1_\prime$	110.1(4)
$C^2_{\zeta_1}$ - $C^2_{\eta_1}$	1.377(10)	C^1_α - $C^1_\prime$ - O^1	122.4(5)
$C^2_{\eta_1}$ - C^2_1	1.362(11)	C^1_α - $C^1_\prime$ - N^1	114.5(5)
C^2_1 - $C^2_{\zeta_2}$	1.380(10)	O^1 - $C^1_\prime$ - N^2	122.8(5)
$C^2_{\zeta_2}$ - $C^2_{\delta_2}$	1.394(9)	$C^1_\prime$ - N^1 - C^2_α	119.8(4)
$C^2_{\delta_2}$ - C^2_γ	1.390(8)	N^1 - C^2_α - C^2_β	124.8(5)
C^2_α - $C^2_\prime$	1.504(8)	C^2_α - C^2_β - C^2_γ	131.7(5)
$C^2_\prime$ - O^2	1.241(7)	C^2_β - C^2_γ - $C^2_{\delta_2}$	117.4(5)
$C^2_\prime$ - N^2	1.324(7)	$C^2_{\delta_2}$ - C^2_γ - $C^2_{\delta_1}$	117.4(5)
N^2 - C^3_α	1.459(7)	C^2_γ - $C^2_{\delta_1}$ - $C^2_{\eta_1}$	121.7(6)
C^3_α - $C^3_\prime$	1.494(8)	$C^2_{\delta_1}$ - $C^2_{\eta_1}$ - C^2_γ	120.6(7)
$C^3_\prime$ - O^3	1.220(10)	$C^2_{\eta_1}$ - C^2_1 - $C^2_{\zeta_1}$	119.2(6)
$C^3_\prime$ - $N^3_\prime$	1.342(10)	C^2_1 - $C^2_{\zeta_1}$ - $C^2_{\delta_1}$	120.6(6)
		$C^2_{\eta_1}$ - $C^2_{\zeta_1}$ - C^2_γ	120.5(6)
		$C^2_{\delta_2}$ - C^2_γ - C^2_β	125.2(5)
		N^2 - C^2_α - $C^2_\prime$	116.5(4)
		C^2_β - C^2_α - $C^2_\prime$	118.6(5)
		C^2_α - $C^2_\prime$ - O^2	118.9(5)
		C^2_α - $C^2_\prime$ - N^2	118.9(5)
		O^2 - $C^2_\prime$ - N^2	122.2(5)
		$C^2_\prime$ - N^2 - C^3_α	118.2(5)
		N^2 - C^3_α - $C^3_\prime$	114.5(5)
		C^3_α - $C^3_\prime$ - O^3	123.3(6)
		O^3 - $C^3_\prime$ - $N^3_\prime$	124.5(6)
		C^3_α - $C^3_\prime$ - $N^3_\prime$	112.2(6)

TABLE 4 - TORSION ANGLES (deg) AND ESTIMATED STANDARD DEVIATIONS FOR Boc-L-Pro-Δ^Z-Phe-Gly-NH_2

C(1)-C(4)-O(1)-C(5)	61.0(7)	N_2 $-C_2^{\alpha}$ $-C_2^{\beta}$ $-C_2^{\gamma}$	2.6(10)
C(2)-C(4)-O(1)-C(5)	-177.8(6)	C_2^{α} $-C_2^{\beta}$ $-C_2^{\gamma}$ $-C_2^{\delta 1}$	-174.1(6)
C(3)-C(4)-O(1)-C(5)	-60.6(7)	$C_2^{\delta 2}$ $-C_2^{\gamma}$ $-C_2^{\delta 1}$ $-C_2^{\varepsilon 1}$	-0.2(9)
C(4)-O(1)-C(5)-O(2)	14.2(9)	$C_2^{\delta 2}$ $-C_2^{\gamma}$ $-C_2^{\delta 1}$ $-C_2^{\varepsilon 2}$	-0.8(9)
C(4)-O(1)-C(5)-N_1	-167.1(5)	C_2^{α} $-C_2^{\beta}$ $-C_2^{\gamma}$ $-C_2^{\delta 2}$	8.1(10)
O(1)-C(5)-N_1 $-C_1^{\alpha}$	176.8(5)	C_2^{β} $-C_2^{\gamma}$ $-C_2^{\delta 1}$ $-C_2^{\varepsilon 1}$	-178.1(6)
O(1)-C(5)-N_1 $-C_1^{\delta}$	4.3(8)	C_2^{γ} $-C_2^{\delta 2}$ $-C_2^{\varepsilon 2}$ $-C_2^{\gamma}$	1.9(11)
O(2)-C(5)-N_1 $-C_1^{\alpha}$	-4.4(9)	$C_2^{\delta 1}$ $-C_2^{\varepsilon 1}$ $-C_2^{\varepsilon 2}$ $-C_2^{\delta 2}$	-2.6(12)
O(2)-C(5)-N_1 $-C_1^{\delta}$	-177.0(6)	$C_2^{\delta 1}$ $-C_2^{\gamma}$ $-C_2^{\varepsilon 1}$ $-C_2^{\delta 1}$	1.7(11)
C(5)-N_1 $-C_1^{\alpha}$ $-C_1^{\beta}$	-167.3(5)	C_2^{γ} $-C_2^{\varepsilon 1}$ $-C_2^{\delta 1}$ $-C_2^{\gamma}$	0.1(10)
C(5)-N_1 $-C_1^{\alpha}$ $-C_1'$	-49.1(7)	$C_2^{\varepsilon 2}$ $-C_2^{\delta 2}$ $-C_2^{\gamma}$ $-C_2^{\beta}$	177.0(6)
N_1 $-C_1^{\alpha}$ $-C_1^{\beta}$ $-C_1^{\gamma}$	-27.3(6)	C_1' $-N_2$ $-C_2^{\alpha}$ $-C_2'$	68.8(7)
C_1^{β} $-C_1^{\alpha}$ $-C_1'$ $-O_1$	65.6(7)	N_2 $-C_2^{\alpha}$ $-C_2'$ $-O_2$	-167.3(5)
C_1^{β} $-C_1^{\alpha}$ $-C_1'$ $-N_2$	-112.7(5)	N_2 $-C_2^{\alpha}$ $-C_2'$ $-N_3$	13.2(7)
C_1' $-C_1^{\alpha}$ $-C_1^{\beta}$ $-C_1^{\gamma}$	-146.1(5)	C_2^{γ} $-C_2^{\beta}$ $-C_2^{\alpha}$ $-C_2'$	-179.8(6)
C_1^{α} $-C_1^{\beta}$ $-C_1^{\gamma}$ $-C_1^{\delta}$	38.2(6)	C_2^{β} $-C_2^{\alpha}$ $-C_2'$ $-O_2$	14.9(8)
C_1^{β} $-C_1^{\gamma}$ $-C_1^{\delta}$ $-N_1$	-33.7(6)	C_2^{β} $-C_2^{\alpha}$ $-C_2'$ $-N_3$	-164.6(5)
C_1^{γ} $-C_1^{\delta}$ $-N_1$ $-C(5)$	-169.6(6)	C_2^{α} $-C_2'$ $-N_3$ $-C_3^{\alpha}$	-178.5(5)
C_1^{γ} $-C_1^{\delta}$ $-N_1$ $-C_1^{\alpha}$	17.3(6)	O_2 $-C_2'$ $-N_3$ $-C_3^{\alpha}$	1.9(8)
C_1^{δ} $-N_1$ $-C_1^{\alpha}$ $-C_1^{\beta}$	6.2(6)	C_2' $-N_3$ $-C_3^{\alpha}$ $-C_3'$	-162.6(5)
C_1^{δ} $-N_1$ $-C_1^{\alpha}$ $-C_1'$	124.4(5)	N_3 $-C_3^{\alpha}$ $-C_3'$ $-O_3$	16.5(9)
N_1 $-C_1^{\alpha}$ $-C_1'$ $-O_1$	-48.8(7)	N_3 $-C_3^{\alpha}$ $-C_3'$ $-N_3'$	-163.4(6)
N_1 $-C_1^{\alpha}$ $-C_1'$ $-N_2$	133.0(5)		
C_1^{α} $-C_1'$ $-N_2$ $-C_2^{\alpha}$	174.8(5)		
O_1 $-C_1'$ $-N_2$ $-C_2^{\alpha}$	-3.4(8)		
C_1' $-N_2$ $-C_2^{\alpha}$ $-C_2^{\beta}$	-113.5(6)		

TABLE 5 - GEOMETRY OF THE HYDROGEN BONDS IN THE CRYSTAL OF THE Boc-L-Pro-Δ^Z-Phe-Gly-NH_2

Donor	Acceptor	Symm. equivalence	Distances (Å)		Angle (°)
D-H	A	of A	D...A	H...A	D-H...A
N_3-H	O(2)	x, y, z	3.11(1)	2.26(4)	154(4)
N_3'-H syn	O_3	1-x, y-½, 2-z	2.96(1)	2.11(4)	146(4)
N_2-H	O_2	x, y+1, z	2.82(1)	1.92(4)	169(4)

The Boc-group[26] and all peptide bonds[35] are *trans* ($\omega \approx 180°$) with [$\Delta\omega$] never exceeding 6°. In addition to the ω_0 torsion angle, the conformation of the *tert*-butyloxycarbonyl N^{α}-protecting group is described by the θ^1 angle. In this peptide θ^1 has a value of –167.1(5)°, and it allows us to classify the urethane moiety as the common type *b*[26].

The pyrrolidine ring of the Pro residue exhibits an approximate C_2 (half-chair) symmetry[28] with the twofold axis passing through the N_1 atom and the midpoint of the C_1^{β}-C_1^{γ} bond. This structure may be designated as C_2-C^{γ}-*exo* (the C_1^{γ} and C_1 atoms are located on the opposite side of the N_1-C_1^{α}-C_1^{δ} plane) or conformation *A* according to Balasubramanian *et al.*[30].

den for Phe [χ^1=2.6(10)°][29,36,37] and in shorter N_2-C^{α}, C_2^{α}-C_2^{β}, and C_2^{β}-C_2^{γ}, distances [1.428(7), 1.315(7), and 1.475(7) Å]. The bond angles are also significantly affected, especially those around the C_2^{α} and C_2^{β} atoms; in particular, the N_2-C_2^{α}-C_2^{β} and C_2^{α}-C_2^{β}-C_2^{γ} angles noticeably exceed the standard trigonal value [124.8(5) and 131.7(5)°, respectively]. These values are comparable to those found in other Δ^Z-Phe-containing peptides[31-34]. Also the $\chi^{2,1}$ and $\chi^{2,2}$ torsion angles of the Δ^Z-Phe side chain [–174.1(6)° and 8.1(10)°] are unusual for a Phe residue[29,36,37].

The backbone conformation is folded at the -L-Pro-Δ^Z-Phe-sequence. The ϕ, ψ values for the L-Pro [–49.1(7), 133.0(5)°] and Δ^Z-Phe [68.8(7), 13.2(7)°] residues lie close to the values expected for an ideal type-II β-bend[38-40]. The conformation of the Δ^Z-Phe residue is found in the minimum energy region II of the map of Ac-Δ^Z-Phe-NHMe[41]. An intramolecular N-H···O=C bond is observed between the Gly-NH and the Boc C=O groups. The N_3···O(2) distance, 3.11(1) Å, is at the limit of the range of values reported for a large number of peptide structures[42,43]. The Gly^3 residue is almost completely extended [ϕ_3, ψ_3 = –162.6(5), –163.4(6)°].

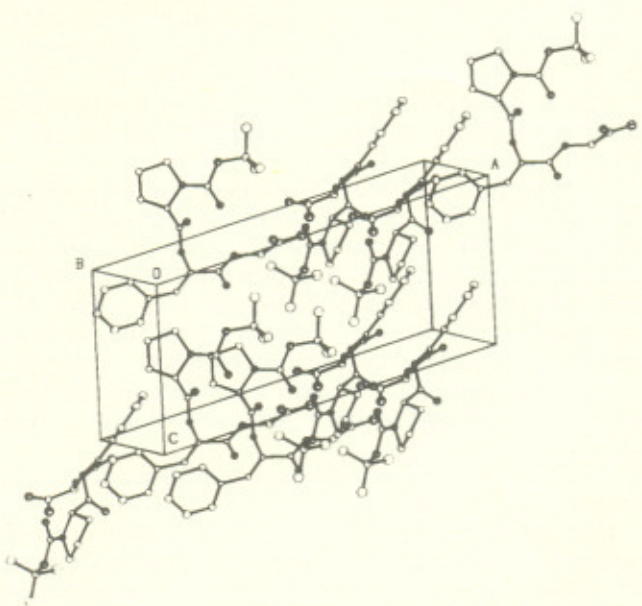

Fig. 4 - Mode of packing of the Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 molecules in the crystal.

The puckering coordinates for the pyrrolidine ring are: q_2=0.373(5)Å and φ_2=–81.7(8)°[44].

The crystal structure is characterized by two intermolecular H-bonds, involving the Δ^Z-Phe NH and C=O groups of symmetry related (x, y+1, z) molecules, and the *syn* carboxamido NH and Gly C=O groups of symmetry related (1–x, y–1/2, 2–z) molecules. The N···O distances of 2.82(1)Å, and 2.96 (1) Å respectively, are within the average range of observed intermolecular H-bonds in peptide structures[42,43].

CONCLUSIONS

The results described here favour the conclusion that in solution the N^{α}-protected tripeptide amide Boc-L-Pro-Δ^Z-Phe-Gly-NH_2 preferentially folds into two consecutive (type II-III') β-bend conformations[45,46]. However, the left-handed helical type-III' β-bend at the Δ^Z-Phe-Gly sequence is disrupted in the crystal state, where the Gly residue adopts an almost fully extended conformation.

Ten X-ray diffraction structures solved to date possess the Δ^Z-Phe residue in an N^{α}-acylated C^{α}-amidated dipeptide (or longer peptide) unit[8,10,11,31-34,47,48],a situation preferable for evaluating β-bend propensities of a given residue in a polypeptide chain. Specific points which emerge from an analysis of those structures are listed below:

(*i*) When a Δ^Z-Phe residue is in the i+2 position of a potential β-bend, this conformation is usually adopted by the peptide. In general, in these folded peptides the β-bend which is formed is of type II, *i.e.* the achiral Δ^Z-Phe residue occupies a position ideal for a D-residue.

(*ii*) Also when a Δ^Z-Phe residue is in the i+1 position of a potential β-bend, this conformation (types I/III; not, however, type II) is usually found in the peptide. These findings, (*i*) and (*ii*), indicate that the Δ^Z-Phe residue may serve as a better nucleus for peptide folding as compared to its saturated counterpart (L-Phe)[49].

Interestingly however, S-shaped conformations characterize the Δ^Z-Phe-Δ^Z-Phe sequence. This result might suggest the absence of an overwhelming stabilization of a series of β-bend structures (3_{10}-helix) produced by the presence of *consecutive* Δ^Z-Phe residues. In this context it is worth noting that energy calculations have indicated that a *single* Δ^Z-Phe residue can be easily accommodated in helical conformations and at both positions of type-I and -II β-bends[41].

(*iii*) Fully-extended and β-sheet conformations are unobserved, as yet.

The present findings are in reasonably good agreement with the aforementioned conclusions in the sense that: (*i*) a type-II β-bend conformation is actually formed by the L-Pro-Δ^Z-Phe sequence; and (*ii*) the Δ^Z-Phe-Gly sequence, a potentially good candidate for type-III' β-bend formation, although preferring an open, partially extended structure in the crystal state, is folded in solution.

In any case, there is little doubt that further exploration of the stereochemistry of the Δ^Z-Phe residue will provide greater insight into its conformational behaviour.

Received April 13th 1990

REFERENCES

(1) C. Toniolo, M. Crisma, E.L. Becker, *Farmaco Ed. Sci., 45*, 921 (1990).

(2) C.H. Stammer, «Chemistry and biochemistry of amino acids, peptides, and proteins», Vol. 6, B. Weinstein, Ed., Dekker, New York, 1982, p. 33.

(3) A.F. Spatola, «Chemistry and biochemistry of amino acids, peptides, and proteins», Vol. 7, B. Weinstein, Ed., Dekker, New York, 1983, p. 267.

(4) K. Noda, Y. Shimonigashi, N. Izumiya, «The peptides: analysis, synthesis, biology», Vol. 5, E. Gross, J. Meienhofer, Eds., Academic Press, New York, 1983, p. 285.

(5) U. Schmidt, A. Lieberknecht, J. Wild, *Synthesis*, 159 (1988).

(6) A. Aubry, G. Boussard, M.T. Cung, M. Marraud, B. Vitoux, *J. Chim. Phys. Physicochim. Biol., 85*, 345 (1988).

(7) A. Aubry, M. Marraud, *Biopolymers, 28*, 109 (1989).

(8) H.C. Patel, T.P. Singh, V.S. Chauhan, P. Kaur, *Biopolymers, 29*, 509 (1990).

(9) P. Narula, H.C. Patel, T.P.Singh, V.S. Chauhan, *Biopolymers, 29*, 935 (1990)

(10) M.R. Ciajolo, A. Tuzi, P.A. Temussi, C.R. Pratesi, A. Fissi, O. Pieroni, Proc. 2nd Naples Workshop on Bioactive Peptides, Capri, Italy, 1990, p. 78.

(11) V. Busetti, M. Crisma, C. Toniolo, S. Salvadori, in preparation.

(12) R. Camble, R. Garner, G.T. Young, *J. Chem. Soc. (C)*, 1911 (1969).

(13) P. Main, S.J. Fiske, S.E. Hull, L. Lessinger, G. Germain, J.P. Declercq, M.M. Woolfson, «MULTAN 80. A system of computer programs for the automatic solution of crystal structures from X-ray diffraction data», University of York, England, and University of Louvain, Belgium, 1980.

(14) «International tables for X-ray crystallography», Vol. IV, 2nd ed., The Kynoch Press, Birmingham, 1974.

(15) G.M. Sheldrick, «SHELX 76. Program for crystal structure determination», University of Cambridge, England, 1976.

(16) L.J. Bellamy, «The infrared spectra of complex molecules», Methuen, London, 1966, p. 203.

(17) C.P. Rao, R. Nagaraj, C.N.R. Rao, P. Balaram, *Biochemistry, 19*, 425 (1980).

(18) G.M. Bonora, C. Mapelli, C. Toniolo, R.R. Wilkening, E.S. Stevens, *Int. J. Biol. Macromol., 6*, 179 (1984).

(19) E. Benedetti, A. Bavoso, B. Di Blasio, V. Pavone, C. Pedone, C. Toniolo, G.M. Bonora, *Biopolymers, 22*, 305 (1983).

(20) M.H. Baron, C. De Loze, C. Toniolo, G.D. Fasman, *Biopolymers, 17*, 2225 (1978).

(21) D. Martin, G. Hauthal, «Dimethyl sulphoxide», Van Nostrand-Reinhold, Wokingham, England, 1975.

(22) A.A. Bothner-By «Magnetic resonance in biology», R.G. Shulman, Ed., Academic Press, New York, 1979, p. 177.

(23) M.A. Khaled, D.W. Urry, *Biochem. Biophys. Res. Commun., 70*, 485 (1976).

(24) C. Toniolo, M. Crisma, G. Valle, G.M. Bonora, S. Polinelli, E.L. Becker, R.J. Freer, Sudhanand, R.B. Rao, P. Balaram, M. Sukumar, *Peptide Res., 2*, 275 (1989).

(25) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry, 9*, 3471 (1970).

(26) E. Benedetti, C. Pedone, C. Toniolo, G. Nemethy, M.J.

POTTLE, H.A. SCHERAGA, *Int. J. Pept. Protein Res., 16,* 156 (1980).

(27) P. CHAKRABARTI, J.D. DUNITZ, *Helv. Chim. Acta, 65,* 1555 (1982).

(28) T. ASHIDA, M. KAKUDO, *Bull. Chem. Soc. Jpn., 47,* 1129 (1974).

(29) T. ASHIDA, Y. TSUNOGAE, I. TANAKA, T. YAMANE, *Acta Crystallogr., Sect. B, 43,* 212 (1987).

(30) R. BALASUBRAMANIAN, A.V. LAKSHMINARAYANAN, M.N. SABESAN, G. TEGONI, K. VENKATESAN, G.N. RAMACHANDRAN, *Int. J. Protein Res., 3,* 25 (1971).

(31) A. AUBRY, G. BOUSSARD, M. MARRAUD, *C.R. Acad. Sci. Paris, Série II, 229,* 1031 (1984).

(32) M.L. GLOWKA, *Acta Crystallogr., Sect. C, 44,* 1639 (1988).

(33) T.P. SINGH, M. HARIDAS, V.S. CHAUHAN, A. KUMAR, D. VITERBO, *Biopolymers, 27,* 1333 (1988).

(34) T.P. SINGH, P. NARULA, V.S. CHAUHAN, P. KAUR, *Biopolymers, 28,* 1287 (1989).

(35) E. BENEDETTI, «Chemistry and biochemistry of amino acids, peptides, and proteins», Vol. 6, B. WEINSTEIN, Ed., Dekker, New York, 1982, p. 105.

(36) E. BENEDETTI, G. MORELLI, G. NEMETHY, H.A. SCHERAGA, *Int. J. Pept. Protein Res., 22,* 1 (1983).

(37) R.O. GOULD, A.M. GRAY, P. TAYLOR, M.D. WALKINSHAW, *J. Am. Chem. Soc., 107,* 5921 (1985).

(38) C.M. VENKATACHALAM, *Biopolymers, 6,* 1425 (1968).

(39) C. TONIOLO, *C.R.C., Crit. Rev. Biochem., 9,* 1 (1980).

(40) G.D. ROSE, L.M. GIERASCH, J.A. SMITH, *Advan. Protein Chem., 37,* 1 (1985).

(41) D. AJÒ, V. BUSETTI, G. GRANOZZI, *Tetrahedron, 38,* 3329 (1982).

(42) C. RAMAKRISHNAN, N. PRASAD, *Int. J. Protein Res., 3,* 209 (1971).

(43) R. TAYLOR, O. KENNARD, W. VERSICHEL, *Acta Crystallogr., Sect. B, 40,* 280 (1984).

(44) D. CREMER, J.A. POPLE, *J. Am. Chem. Soc., 97,* 1354 (1975).

(45) K. UMA, P. BALARAM, P. KAUR, A.K. SHARMA, V.S. CHAUHAN, *Int. J. Biol. Macromol., 11,* 169 (1989).

(46) P. KAUR, K. UMA, P. BALARAM, V.S. CHAUHAN, *Int. J. Peptide Protein Res., 33,* 103 (1989).

(47) O. PIERONI, G. MONTAGNOLI, A. FISSI, S. MERLINO, F. CIARDELLI, *J. Am. Chem. Soc., 97,* 6820 (1975).

(48) O. PIERONI, A. FISSI, S. MERLINO, F. CIARDELLI, *Israel J. Chem., 15,* 22 (1977).

(49) P.Y. CHOU, G.D. FASMAN, *Biophys. J., 26,* 367 (1979).