(35) Shiojl, Y.; Ohnishi, S. I.; Nitta, I. *J. Polym. Sci., Part A* **1963**, *1*, 3373.
(36) Boyarchuk, Y. M. *Vysokomol. Soedin., Ser. A* **1969**, *11*, 2161.
(37) Butler, G. B.; Crawshaw, A.; Miller, W. L. *J. Am. Chem. Soc.* **1958**, *80*, 3615.
(38) Solomon, D. H.; Hawthorne, D. G. *J. Macromol. Sci., Rev. Macromol. Chem.* **1976**, *C15* (1), 143.
(39) Lancaster, J. E.; Baccei, L.; Panzer, H. P. *J. Polym. Sci., Part B* **1976**, *14*, 549.
(40) Lamb, R. C.; Ayers, P. W.; Toney, M. K. *J. Am. Chem. Soc.* **1963**, *85*, 3483.
(41) Walling, C.; Pearson, M. S. *J. Am. Chem. Soc.* **1964**, *86*, 2262.
(42) Butler, G. B.; Raymond, M. A. *J. Polym. Sci., Part A* **1965**, *3*, 3413.
(43) Capon, B.; Rees, C. W. *Annu. Rep. Prog. Chem.* **1964**, *61*, 261.
(44) Struble, D. L.; Beckwith, A. L. J.; Gream, G. E. *Tetrahedron Lett.* **1968**, 3701.

Preferred Conformations of Linear Homooligoprolines. *N-tert*-Butyloxycarbonyl-D-prolyl-D-prolyl-L-proline[1]

Alfonso Bavoso,*[2a] Ettore Benedetti,[2a] Benedetto Di Blasio,[2a] Vincenzo Pavone,[2a] Carlo Pedone,[2a] Claudio Toniolo,[2b] and Gian Maria Bonora[2b]

Istituto Chimico, Università di Napoli, 80134 Napoli, Italy, and Centro di Studi sui Biopolimeri, CNR, Istituto di Chimica Organica, Università di Padova, 35100 Padova, Italy.
Received May 12, 1981

ABSTRACT: An analysis of the preferred conformations of *N-tert*-butyloxycarbonyl-D-prolyl-D-prolyl-L-proline was performed in the solid state and in solution using X-ray diffraction and infrared absorption. In the solid state the stereosequence of the ω angles is cis,trans,cis and that of the ϕ, ψ angles F*,F*,F. The molecules do not form an intramolecularly hydrogen-bonded oxy-C_{10} peptide conformation, but rather they are held together through intermolecular O–H···O=C (Pro_1–Pro_2 peptide carbonyl) bonds. In solvents of low polarity the following are observed: (i) in concentrated solution the type of association which occurs is different from that found in the solid state and (ii) in dilute solution there is no evidence for the onset of intramolecularly hydrogen-bonded forms. In solvents with strong hydrogen-bonding acceptor and donor properties only solvated species are seen.

In 1974, on the basis of infrared (IR) absorption evidence, Deber[3] proposed the presence of oxy analogues of the 4←1 intramolecularly hydrogen-bonded peptide conformations (β turns, C_{10} ring structures)[4,5] in the solid state for *t*-Boc-Gly-L-Pro-OH (*t*-Boc = *tert*-butyloxycarbonyl), *t*-Boc-L-Pro-Gly-OH, *t*-Boc-L-Pro-D-Pro-OH, and *t*-Boc-L-Pro-L-Pro-L-Pro-D-Pro-OH.

Since purely on the basis of the IR absorption data it is impossible to rule out unequivocally intermolecular hydrogen-bonding effects in the crystals of these acids, following Deber's suggestion[3] we undertook an X-ray diffraction analysis of the four aforementioned oligopeptides. The results already published indicate that the molecules of *t*-Boc-L-Pro-Gly-OH,[6] *t*-Boc-Gly-L-Pro-OH,[7] and *t*-Boc-L-Pro-D-Pro-OH[8] do not adopt an oxy-C_{10} conformation but rather that they are held together through intermolecular O–H···O=C hydrogen bonds, with either the urethane or the peptide carbonyl group acting as the acceptor.

In this paper we describe first the solid-state preferred conformation of the enantiomer of the last peptide examined by Deber,[3] namely, *t*-Boc-D-Pro-D-Pro-L-Pro-OH, using IR absorption and X-ray diffraction. The ϕ, ψ conformational region (A or F),[9] urethane and peptide bond conformations (cis or trans),[10-12] pyrrolidine puckering (A or B; C_2 or C_s symmetry),[13,14] and hydrogen bond (inter- or intramolecular, with urethane, peptide, or acid carbonyl as the acceptor)[4,5] will be discussed. The IR absorption properties of *t*-Boc-D-Pro-D-Pro-L-Pro-OH in solvents of widely divergent polarity and hydrogen-bonding capability are also reported. The conclusions are facilitated by comparison with the results obtained for the tripeptide benzyl ester. The trans-II′ oxy-C_{10} structure for a D-Pro-L-Pro sequence occurring at the C terminus of a polypeptide chain is illustrated in Figure 1.

This study is also part of our continuing, systematic analysis of the conformation of linear homooligoprolines (to the tetrapeptide) having all possible chiral sequences and different N-protecting and C-terminal groups.[8,15,16]

Experimental Section

Synthesis of Peptides. *t*-Boc-D-Pro-D-Pro-L-Pro-OH was prepared by catalytic (Pd/C) hydrogenation of *t*-Boc-D-Pro-D-Pro-L-Pro-OBzl (OBzl = benzyloxy) in *tert*-butyl alcohol: mp 202–203 °C, after recrystallization from methanol–diethyl ether; $[\alpha]^{20}_D$ +40.5° (*c* 1.45; methanol); TLC, R_{f_1} 0.50 (SiO_2, Merck; 3:1:1 *n*-butyl alcohol–water–acetic acid), R_{f_2} 0.10 (SiO_2, Merck; 9:1 chloroform–ethanol). Anal. Calcd for $C_{20}H_{31}N_3O_6$: C, 58.7; H, 7.6; N, 10.3. Found: C, 58.1; H, 7.6; N, 10.2.

In turn, *t*-Boc-D-Pro-D-Pro-L-Pro-OBzl was prepared from *t*-Boc-D-Pro-D-Pro-OH[8] and HCl·H-Pro-OBzl in anhydrous chloroform in the presence of *N*-methylmorpholine and isobutyl chloroformate:[17] oil, after precipitation from ethyl acetate–petroleum ether; $[\alpha]^{20}_D$ +26.9° (*c* 1.2; methanol); R_{f_1} 0.75; R_{f_2} 0.90. Anal. Calcd for $C_{27}H_{37}N_3O_6$: C, 64.9; H, 7.5; N, 8.4. Found: C, 64.2; H, 7.4; N, 8.3.

Infrared Absorption. Infrared absorption spectra were recorded with a Perkin-Elmer Model 580 spectrophotometer. For the solid-state measurements the KBr disk technique was used. For the solution measurements 0.2- and 0.1-mm cells with CaF_2 windows were employed. Trimethyl phosphate (TMP), methylene-d_2 chloride (99.8% D), deuteriochloroform (99.8% D), deuterium oxide (99.9% D), *p*-dioxane, and acetonitrile were purchased from Merck. The band positions are accurate to ±1 cm^{-1}.

X-ray Diffraction. Crystals of *t*-Boc-D-Pro-D-Pro-L-Pro-OH in the form of colorless plates were grown from 1:1 methanol–acetone solution.

A CAD4 Enraf-Nonius diffractometer of the Centro di Metodologie Chimico Fisiche of the University of Naples, equipped with PDP 8/E and PDP 11/34 digital computers, was used for the data collection, structure determination, and refinement. The SDP (structure determination program) package of crystallographic programs was used in all stages of calculations. The

0024-9297/82/2215-0054$01.25/0 © 1982 American Chemical Society

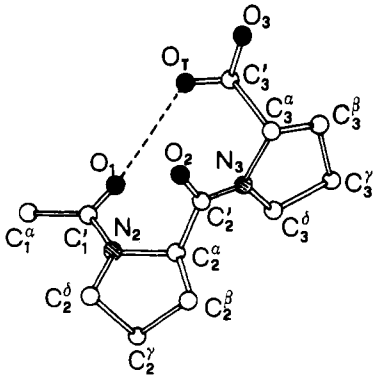

Figure 1. Trans-II′ oxy-C_{10} structure of a D-Pro-L-Pro sequence occurring at the C terminus of a polypeptide chain.

Table I
Crystal Data for *t*-Boc-D-Pro-D-Pro-L-Pro-OH

molecular formula	$C_{20}H_{31}N_3O_6$
molecular weight	409.49
crystal system	tetragonal
space group	$P4_1$
Z, molecules/unit cell	4
a, Å	12.180 (1)
c, Å	14.411 (2)
V, Å^3	2137.9
density, by flotation, g·cm^{-3}	1.27
density, calcd, g·cm^{-3}	1.272
radiation	Cu Kα, λ = 1.5418 Å
no. of independent reflections	1990
temp, °C	23, ambient
final R value	0.049

procedures employed for the determination of the unit cell parameters and data collection are similar to those already described in previous papers.[18,19] Crystallographic data are reported in Table I.

Data collection was carried out with the θ–2θ scan mode and Cu Kα radiation. A total of 2215 reflections in the range 1–140° of 2θ were measured, 1990 of which had a net intensity greater than $3.0\sigma(I)$ and were considered as "observed". All reflections were corrected for Lorentz and polarization effects.

The structure was determined by means of the direct multisolution method in the form programmed by Germain, Main, and Woolfson in MULTAN,[20] included in the SDP package. A total of 256 reflections with $E \geq 1.5$ were employed with the 2000 largest Sayre relationships. The set of phases having the highest combined figure of merit led to an E map from which 20 of the 29 nonhydrogen atoms were recovered. The remaining atoms in the independent unit were subsequently located from a Fourier synthesis phased with the positions of the atoms previously determined.

The structure was then refined by a least-squares procedure to a final R value of 0.045 for the 1990 observed reflections, with anisotropic temperature factors for C, N, and O atoms and isotropic temperature factors for hydrogen atoms in their calculated stereochemically expected positions. The parameters of the hydrogen atoms were kept fixed with an individual isotropic temperature factor for each hydrogen atom equal to that of the corresponding carrier atom. Atomic scattering factor for all atoms were calculated from Cromer and Waber.[21] In the least-squares refinement a weight $w = 1/\sigma^2(F_o)$ was employed and only those reflections with $F_o^2 \geq 3\sigma(F_o^2)$ were considered. Refinement was ended when the maximum shift in the atomic coordinates and anisotropic temperature parameters were less than $^1/_5$ and $^1/_3$ of the corresponding standard deviations, respectively.

The final atomic parameters and their standard deviations are presented in Table II.

Results and Discussion

Solid-State Conformational Analysis. The IR absorption spectra in the 1800–1580-cm^{-1} region of *t*-Boc-D-Pro-D-Pro-L-Pro-OH and its benzyl ester in the solid state

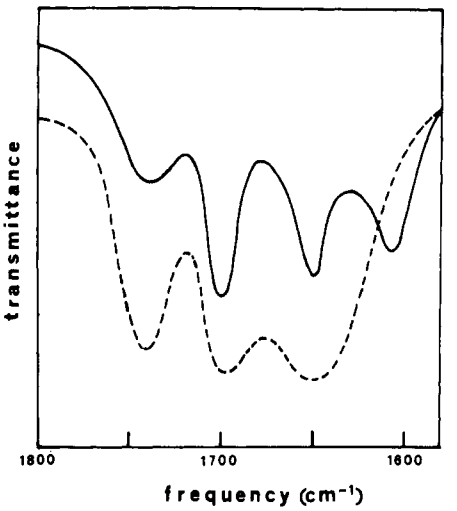

Figure 2. Infrared absorption spectra in the 1800–1580-cm^{-1} region of *t*-Boc-D-Pro-D-Pro-L-Pro-OH (solid line) and its benzyl ester (dashed line) in the solid state.

are shown in Figure 2. The curve of the free acid, which matches closely that previously reported by Deber,[3] does not change on changing the solvent from which the solid material is obtained (including the solvent mixture employed for growing the crystals for the X-ray diffraction analysis). The main difference between the two spectra in Figure 2 rests in the position of the tertiary peptide carbonyl bands which are found at 1649 and 1606 cm^{-1} in the free acid and at 1650 cm^{-1} in the ester. The large shift to lower frequency of a band associated with one of the two peptide groups could be attributed either to intramolecular[3] or to intermolecular OH···O=C hydrogen-bond formation. Absorptions in the 1612–1603-cm^{-1} frequency range have also been observed for the tertiary amide or peptide C=O bands of Ac-L-Pro-OH[22] (Ac = acetyl), Ac-L-Sar-OH,[18] *t*-Boc-L-Pro-L-Pro-OH,[3,8] *t*-Boc-L-Pro-Sar-OH,[3,18] *t*-Boc-L-Pro-D-Pro-L-Pro-OH,[22] and *t*-Boc-L-Pro-L-Pro-D-Pro-OH.[3] Interestingly, both the carbonyl peptide group of *t*-Boc-L-Pro-Sar-OH[18] and the carbonyl amide group of the *N*-acetyl derivative of the higher homologue of proline, namely, *N*-acetyl-DL-pipecolic acid,[23] are intermolecularly hydrogen bonded (obviously with the carboxylic acid OH group) in the corresponding crystal structures, as shown by X-ray diffraction.

Conversely, the positions of the carbonyl bands of the un-ionized carboxylic acid (1740 cm^{-1})[3,6,8,16,18] and the tertiary urethane (1701 cm^{-1})[3,6–8,16,18] of *t*-Boc-Ld-Pro-D-Pro-L-Pro-OH indicate that these two C=O groups are not involved in hydrogen-bond formation.

Consistent with this analysis, the absorption band due to the stretching of the free OH group of the carboxylic acid moiety (3600–3500 cm^{-1}) is absent for *t*-Boc-D-Pro-D-Pro-L-Pro-OH. This finding is in line with Deber's results.[3]

The X-ray diffraction analysis of *t*-Boc-D-Pro-D-Pro-L-Pro-OH allowed us to solve the ambiguity of the conformational assignment made on the basis of its IR absorption properties. The molecular structure is illustrated in Figure 3, where bond lengths and bond angles are also given together with the internal rotation angles defining the conformation of the peptide main chain.[24] The complete list of internal rotation angles is reported in Table III.

The values of bond lengths and bond angles, with estimated standard deviations of 0.005 Å and 0.4°, respectively, are in agreement with literature data on the geometry of proline residues,[11–14,25] the peptide unit,[11,12] and *t*-Boc urethanyl derivatives.[10]

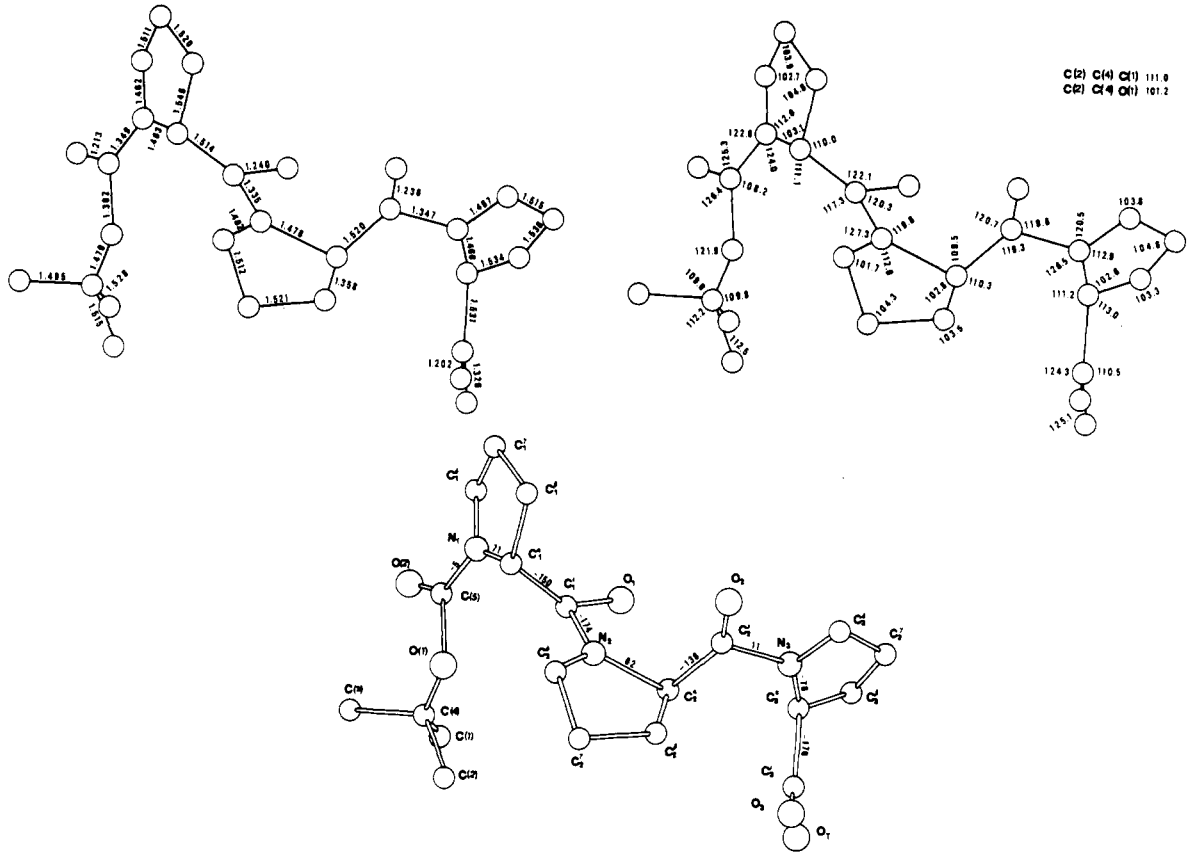

Figure 3. Molecular structure of *t*-Boc-D-Pro-D-Pro-L-Pro-OH showing bond lengths, bond angles, and the internal rotation angles of the main chain.

The C–O distances in the carbonyls (Figure 3) correlate well with the IR absorption data, the highest frequency in the spectrum (1740 cm^{-1}) being identified with the carboxylic acid grouping C_3'–O_3, which shows the shortest C–O distance (1.202 Å), and the lowest frequencies (1649 and 1606 cm^{-1}) with the amide C_1'–O_1 and C_2'–O_2 groupings, which show the longest distances (1.240 and 1.236 Å, respectively).

The stereosequence of the urethane and two peptide bonds (ω angles) is cis,trans,cis. The occurrence of a tertiary *t*-Boc urethane group in the cis conformation has recently been shown to be frequent[10] and minimum energy calculations have demonstrated nearly equal energies for the cis and trans arrangements.[10]

Also the occurrence of a tertiary amide or peptide X–Y bond (Y = N-alkylated α-amino acid residue) in the cis conformation is not a rare event in linear peptides.[26-32] In fact, the difference in energy between the cis and trans conformations of this type of bond is estimated to be low, although higher than that of the tertiary *t*-Boc urethane bond.[10]

From the linear peptide structures reported in the literature in which a cis amide or peptide bond preceding an *N*-alkylamino acid residue has been observed[26-32] we have calculated the average values for bond lengths and bond angles for a cis unit preceding an N-alkylated α-amino acid residue. Clearly, the geometry of a noncyclic residue (Sar, MeTyr (Me)) has to be considered more flexible than that of a Pro (or 4-Hyp) residue, where the ring closure exerts some restrictions. The values given the Figure 4 show that in the cis arrangement the main changes with respect to the trans arrangement occur on the nitrogen atom. The C′–N–C^α angle is larger than the C′–N–C angle by about 5° in the cis unit, but in the trans unit it is the C′–N–C angle which is larger than in the C′–N–C^α angle by 5°. A

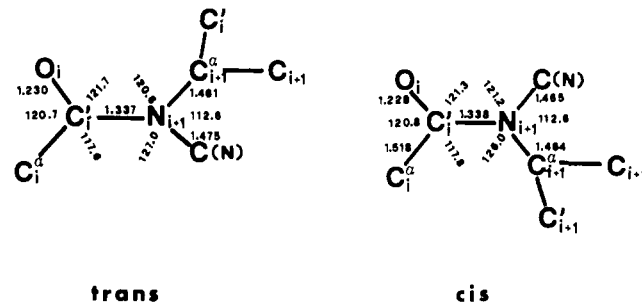

Figure 4. Right: bond lengths (in Å) and bond angles (in degrees) of a cis peptide unit preceding an N-alkylated α-amino acid residue. The average angles at the nitrogen pertain to propyl residues; for *N*-methyl residues the C_i'–N_{i+1}–C(N) and C_{i+1}^α–N_{i+1}–C(N) angles, where C(N) is the methyl linked to nitrogen, became 118.4 and 116.4°, respectively. Left: bond lengths (in Å) and bond angles (in degrees) of a trans peptide unit preceding a prolyl residue (from ref 11).

tentative explanation can be that the steric repulsion between the C^α atom (of the preceding residue) and the C^α atom of the N-alkylated amino acid residue in the cis arrangement is greater than that between the oxygen atom of the carbonyl group and the C^α atom of the N-alkylated amino acid residue in the trans arrangement.

The internal rotation angles ϕ_1,ψ_1 (71°,−150°), ϕ_2,ψ_2 (62°,−138°), and ϕ_3,ψ_3* (−78°,176°) (the ψ_3* notation corresponds to the N_3–C_3–C_3'–O_T internal rotation angle) or the proline residues fall in one of the calculated minimum-energy regions (the F* region for the D residues and the enantiomeric F region for the L residue, according to the one-letter code introduced by Zimmerman et al.[9]).

In the N-protecting *t*-Boc group the C(4)–O(1) bond is in the usual trans arrangement relative to the C(5)–N_1 bond; this feature, accompanied by the cis arrangement

Table II
Final Positional[a] and Thermal Parameters[a,b] and Their Estimated Standard Deviations for t-Boc-D-Pro-D-Pro-L-Pro-OH

atom	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
C(1)	5875 (5)	5905 (5)	2788 (5)	92 (4)	75 (4)	83 (4)	−37 (7)	−17 (8)	13 (7)
C(2)	5203 (4)	7865 (5)	2767 (4)	63 (4)	86 (4)	67 (3)	0 (7)	−13 (6)	4 (7)
C(3)	6273 (5)	7145 (4)	1435 (4)	88 (4)	73 (4)	52 (3)	−8 (7)	−13 (6)	−9 (6)
C(4)	6093 (4)	7077 (4)	2460 (4)	60 (3)	69 (3)	48 (3)	−16 (6)	−20 (5)	−1 (6)
O(1)	7060 (2)	7529 (3)	2951 (2)	54 (2)	67 (2)	56 (2)	17 (4)	−13 (4)	−37 (4)
C(5)	8069 (4)	7052 (4)	2891 (4)	64 (3)	73 (4)	45 (3)	15 (6)	−6 (5)	−14 (5)
O(2)	8288 (3)	6197 (3)	2500 (0)	85 (3)	75 (2)	73 (2)	39 (5)	−16 (5)	−67 (4)
N_1	8789 (3)	7681 (3)	3363 (3)	53 (3)	64 (3)	45 (2)	21 (5)	−3 (4)	−27 (4)
C_1^{α}	8513 (4)	8739 (3)	3784 (3)	59 (3)	51 (3)	32 (2)	21 (5)	−5 (4)	−2 (4)
C_1'	7776 (4)	8590 (4)	4620 (3)	57 (3)	48 (3)	35 (2)	6 (5)	−8 (5)	−7 (4)
O_1	7878 (3)	7803 (2)	5159 (2)	101 (3)	42 (2)	41 (2)	33 (4)	8 (4)	17 (3)
C_1^{β}	9642 (4)	9190 (4)	4090 (4)	65 (3)	66 (3)	48 (3)	−7 (6)	−16 (5)	1 (6)
C_1^{γ}	10473 (4)	8558 (4)	3512 (4)	54 (3)	80 (4)	71 (3)	2 (7)	−9 (6)	25 (6)
C_1^{δ}	9963 (4)	7434 (4)	3409 (4)	52 (3)	71 (3)	67 (3)	34 (6)	2 (6)	−8 (6)
N_2	7068 (3)	9397 (3)	4810 (3)	54 (2)	47 (2)	32 (2)	14 (4)	6 (4)	13 (4)
C_2^{α}	6407 (3)	9354 (3)	5669 (3)	51 (3)	54 (3)	29 (2)	−4 (5)	3 (4)	2 (4)
C_2'	7164 (3)	9363 (4)	6507 (3)	49 (3)	56 (3)	35 (2)	8 (5)	−3 (4)	3 (5)
O_2	7924 (3)	10033 (3)	6563 (3)	67 (2)	77 (2)	55 (2)	57 (4)	18 (4)	22 (4)
C_2^{β}	5742 (4)	10426 (4)	5623 (3)	60 (3)	73 (3)	43 (2)	32 (6)	5 (5)	5 (5)
C_2^{γ}	5662 (4)	10658 (4)	4588 (4)	63 (3)	75 (4)	50 (3)	44 (6)	−9 (6)	13 (6)
C_2^{δ}	6777 (4)	10342 (4)	4211 (3)	76 (4)	50 (3)	42 (2)	31 (6)	3 (5)	24 (5)
N_2	6950 (3)	8691 (3)	7226 (2)	46 (2)	66 (3)	32 (2)	−6 (5)	−7 (4)	5 (4)
C_3^{α}	6152 (4)	7790 (4)	7241 (3)	52 (3)	53 (3)	41 (2)	−5 (5)	8 (5)	10 (5)
C_3'	4981 (4)	8223 (4)	7368 (3)	55 (3)	69 (3)	29 (2)	−11 (6)	−8 (5)	4 (5)
O_3	4741 (3)	9180 (3)	7392 (3)	61 (2)	64 (2)	57 (2)	11 (4)	18 (4)	4 (4)
O_T	4277 (3)	7398 (3)	7446 (3)	52 (2)	68 (2)	87 (2)	−20 (4)	−5 (4)	11 (5)
C_3^{β}	6546 (4)	7093 (4)	8062 (4)	63 (3)	81 (4)	62 (3)	19 (7)	4 (6)	59 (6)
C_3^{γ}	7044 (4)	7944 (5)	8726 (4)	60 (3)	115 (5)	43 (2)	35 (7)	−2 (5)	40 (6)
C_3^{δ}	7571 (4)	8788 (4)	8095 (3)	65 (3)	96 (4)	33 (2)	2 (7)	−30 (5)	2 (5)

atom	x	y	z	β, Å^2	atom	x	y	z	β, Å^2
H1C(1)	5172	5573	2442	5.01	HC_2^{α}	5892	8637	5707	3.00
H2C(1)	5709	5919	3541	5.01	$H1C_2^{\beta}$	4919	10297	5948	3.78
H3C(1)	6584	5398	2671	5.01	$H2C_2^{\beta}$	6132	11086	6000	3.78
H1C(2)	4427	7645	2447	4.38	$H1C_2^{\gamma}$	5465	11511	4454	4.15
H2C(2)	5410	8700	2559	4.38	$H2C_2^{\gamma}$	5001	10156	4269	4.15
H3C(2)	5107	7845	3514	4.38	$H1C_2^{\delta}$	7374	10993	4283	3.55
H1C(3)	5573	6818	1073	4.25	$H2C_2^{\delta}$	6756	10086	3485	3.55
H2C(3)	6999	6671	1247	4.25	HC_3^{α}	6209	7323	6600	3.09
H3C(3)	6402	7991	1231	4.25	$H1C_3^{\beta}$	5855	6651	8371	4.56
HC_1^{α}	8122	9276	3274	2.86	$H2C_3^{\beta}$	7151	6493	7841	4.56
$H1C_1^{\beta}$	9728	10072	3958	3.85	$H1C_3^{\gamma}$	7648	7554	9189	4.31
$H2C_1^{\beta}$	9795	9057	4829	3.85	$H2C_3^{\gamma}$	6405	8297	9175	4.31
$H1C_1^{\gamma}$	11277	8517	3888	4.35	$H1C_3^{\delta}$	8427	8631	7998	4.00
$H2C_1^{\gamma}$	10626	8959	2860	4.35	$H2C_3^{\delta}$	7475	9617	8391	4.00
$H1C_1^{\delta}$	10153	6892	3998	4.10	HO_T	3477	7550	7533	4.41
$H2C_1^{\delta}$	10249	7018	2782	4.10					

[a] The parameters are multiplied by 10^4. [b] The form of the anisotropic thermal parameter is $\exp[-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + \beta_{12}hk + \beta_{13}hl + \beta_{23}kl)]$.

of the O(1)–C(5) bond relative to the N_1–C_1^{α} bond, allows one to classify the urethane moiety of *t*-Boc-D-Pro-D-Pro-L-Pro-OH as type a.[10,16]

The pyrrolidine rings of both D-proline residues exhibit C_s approximate symmetry with C^{γ}-exo[14] (conformation A according the Balasubramanian et al.,[13] provided that reflection is made in the value of the internal rotation angles of the D residues), while the ring of the L-proline residue presents C_2 symmetry with C^{β}-exo and C^{γ}-endo[14] (conformation B[13]). The observed difference in the conformation of the three pyrrolidine rings again confirms the great flexibility of this five-membered annular system.[8,11,12,33]

The carboxylic acid group adopts the preferred synplanar conformation[34,35] with respect to the C_3^{α}–N_3 *bond*, since the O_3–C_3'–C_3^{α}–N_3 internal rotation angle has a value of 4°.

In the structure of *t*-Boc-D-Pro-D-Pro-L-Pro-OH the OH group of the carboxylic acid moiety of the molecule in position (x, y, z) is intermolecularly hydrogen bonded to the oxygen atom of the carbonyl group of the first peptide unit (O_1) of the symmetry-related molecule in position $(y, -x, {}^3/_4+z)$. The O···O distance is 2.61 Å, i.e., within the most probable range for O–H···O hydrogen bonds in the crystal state with a carbonyl group as the hydrogen acceptor.[36] The arrangement of the molecules and the symmetry of the crystal along the fourfold axis (*c* axis) induce the formation of a helix of hydrogen-bonded molecules,[35] additionally stabilized by nonbonding interactions, as shown in Figure 5. The helices, then, pack with each other through van der Waals interactions.

In summary, in *t*-Boc-D-Pro-D-Pro-L-Pro-OH, as in *t*-Boc-Gly-L-Pro-OH,[7] *t*-Boc-L-Pro-Gly-OH,[6] and *t*-Boc-L-Pro-D-Pro-OH,[8] the oxy analogues of the 4←1 intramolecularly hydrogen-bonded folded peptide conformations[4,5] proposed by Deber[3] on the basis of an IR absorption investigation, appear to be absent in the solid state. Thus, great caution should be used in any attempt to correlate the IR absorption spectral shifts to conformational preferences of oligopeptides.

Table III
Internal Rotation Angles (Deg) for
t-Boc-D-Pro-D-Pro-L-Pro-OH

C(3)-C(4)-O(1)-C(5)	−64	N_2-C_2^α-C_2'-O_2	47
C(2)-C(4)-O(1)-C(5)	179	N_2-C_2^α-C_2'-N_3	−138
C(1)-C(4)-O(1)-C(5)	60	N_2-C_2^α-C_2^β-C_2^γ	27
C(4)-O(1)-C(5)-O(2)	−5	C_2^α-C_2'-N_3-C_3^α	11
C(4)-O(1)-C(5)-N_1	176	C_2^α-C_2'-N_3-C_3^δ	−171
O(1)-C(5)-N_1-C_1^α	−5	C_2^α-C_2^β-C_2^γ-C_2^δ	−39
O(1)-C(5)-N_1-C_1^δ	−177	C_2^α-N_2-C_2^δ-C_2^γ	−18
O(2)-C(5)-N_1-C_1^α	176	C_2^β-C_2^γ-C_2^δ-N_2	35
O(2)-C(5)-N_1-C_1^δ	4	C_2^β-C_2^α-N_2-C_2^δ	−6
C(5)-N_1-C_1^α-C_1'	71	C_2^β-C_2^α-C_2'-O_2	−65
C(5)-N_1-C_1^α-C_1^β	−171	C_2^β-C_2^α-C_2'-N_3	110
C(5)-N_1-C_1^δ-C_1^γ	150	C_2^γ-C_2^β-C_2^α-C_2'	144
N_1-C_1^α-C_1^β-C_1^γ	20	C_2^δ-N_2-C_2^α-C_2'	−123
N_1-C_1^α-C_1'-O_1	36	C_2'-N_3-C_3^α-C_3'	−78
N_1-C_1^α-C_1'-N_2	−150	C_2'-N_3-C_3^α-C_3^β	161
C_1^α-C_1^β-C_1^γ-C_1^δ	−34	C_2'-N_3-C_3^δ-C_3^γ	177
C_1^α-C_1'-N_2-C_2^δ	11	O_2-C_2'-N_3-C_3^δ	3
C_1^α-C_1'-N_2-C_2^α	−174	O_2-C_2'-N_3-C_3^α	−174
C_1^β-C_1^γ-C_1^δ-N_1	34	N_3-C_3^α-C_3'-O_3	4
C_1^β-C_1^α-C_1'-O_1	−77	N_3-C_3^α-C_3'-O_T	−176
C_1^β-C_1^α-C_1'-N_2	97	N_3-C_3^α-C_3^β-C_3^γ	32
C_1^γ-C_1^δ-N_1-C_1^α	−23	N_3-C_3^δ-C_3^γ-C_3^β	25
C_1^γ-C_1^β-C_1^α-C_1'	139	C_3^α-C_3^β-C_3^γ-C_3^δ	−36
C_1^δ-N_1-C_1^α-C_1'	−116	C_3^α-N_3-C_3^δ-C_3^γ	−5
C_1^δ-N_1-C_1^α-C_1^β	2	C_3^β-C_3^α-N_3-C_3^δ	−17
C_1'-N_2-C_2^α-C_2'	62	C_3^β-C_3^α-C_3'-O_3	119
C_1'-N_2-C_2^α-C_2^β	179	C_3^β-C_3^α-C_3'-O_T	−61
C_1'-N_2-C_2^δ-C_2^γ	157	C_3^γ-C_3^β-C_3^α-C_3'	−88
O_1-C_1'-N_2-C_2^α	0	C_3^δ-N_3-C_3^α-C_3'	104
O_1-C_1'-N_2-C_2^δ	−174		

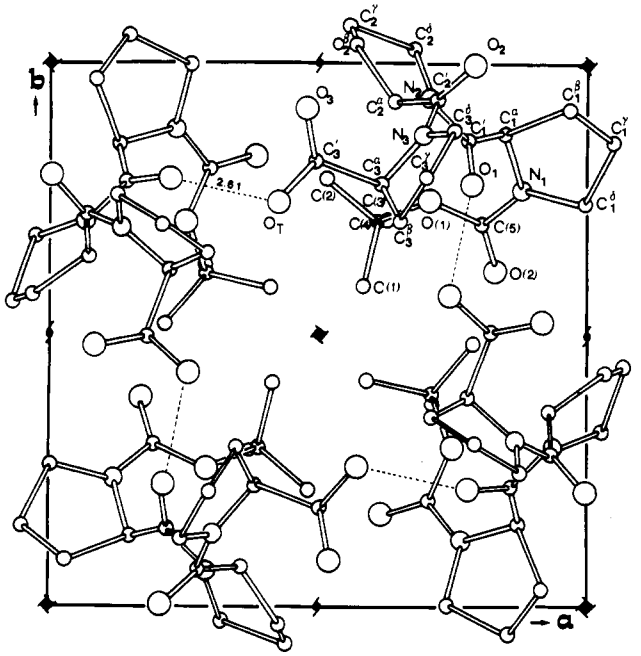

Figure 5. Mode of packing of *t*-Boc-D-Pro-D-Pro-L-Pro-OH as projected down the *c* axis. The hydrogen bonds are indicated as dashed lines.

A comparison of the structure of *t*-Boc-D-Pro-D-Pro-L-Pro-OH with that of *t*-Aoc-L-Pro-L-Pro-L-Pro-OH[37] (the replacement of the *t*-Boc group by the *t*-Aoc group, a closely related urethanyl N-protecting group, does not seem to have any influence on the conformationl preferences of amino acid derivatives and peptides[15,16,28]) indicates that (i) the conformational sequence of the urethane and peptide bonds (ω angles) is cis,trans,cis in the former, while cis,trans,trans in the latter, (ii) in both N-protected tripeptides the disposition of the uretane moiety is of type a,[10,16] (iii) the sets of ϕ,ψ angles of all proline residues fall in the F (or F*) conformational region,[9] (iv) the type of

Table IV
Infrared Absorption Data (cm^{-1}) for
t-Boc-D-Pro-D-Pro-L-Pro-OX (X = H, Bzl) in
the 1800–1550-cm^{-1} Region

solvent	*t*-Boc-D-Pro-D-Pro-L-Pro-OH	*t*-Boc-D-Pro-D-Pro-L-Pro-OBzl
CD_2Cl_2 [a]	1747, 1692, 1653, 1628 (s) [c]	1745, 1690, 1650
CD_2Cl_2 [b]	1749, 1689, 1656, 1627	1743, 1691, 1651
$CDCl_3$ [b]	1746, 1681, 1653, 1632 (s) [c]	1741, 1683, 1649
CH_3CN [b]	1749, 1694, 1654, 1631 [c,d]	1745, 1694, 1653
p-dioxane [b]	1746, 1696, 1653, 1630 (s) [c]	1745, 1698, 1655
TMP [b]	1733, 1695, 1653	1745, 1696, 1654
D_2O [b]	1711, 1643, 1617	1727, 1641, 1620

[a] Concentration ~3 × 10^{-3} M. [b] Concentration ~3 × 10^{-2} M. [c] s, shoulder. [d] Weak band.

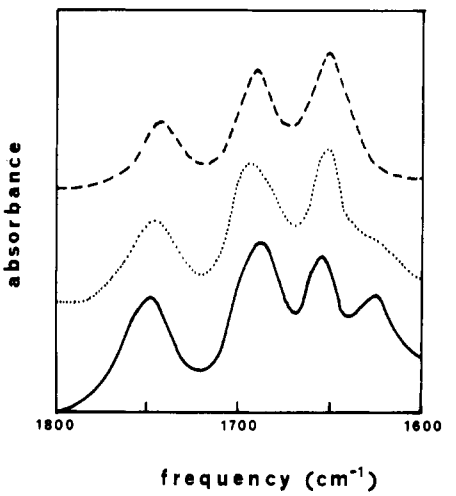

Figure 6. Infrared absorption spectra in the 1800–1600-cm^{-1} region of *t*-Boc-D-Pro-D-Pro-L-Pro-OH (concentration 3.1 × 10^{-2} M, full line; concentration 3.5 × 10^{-3} M, dashed line) in CD_2Cl_2.

puckering of the three pyrrolidine rings partially diverges in the sense that in *t*-Boc-D-Pro-D-Pro-L-Pro-OH the sequence is A,A,B,[13] while in *t*-Aoc-L-Pro-L-Pro-L-Pro-OH it is B,B,B,[13] and (v) the intermolecular O–H···O hydrogen bond has the carbonyl oxygen of the first peptide group as the acceptor in both compounds despite their different chiral sequences (D,D,L vs. L,L,L). Clearly, the structure of other linear homotriprolines should be solved by X-ray diffraction before any generalization on their preferred conformations is made. In this context it is of interest to recall that, as mentioned above, the IR absorption spectrum of *t*-Boc-L-Pro-D-pro-L-Pro-OH[22] matches closely that of its diastereoisomer *t*-Boc-D-Pro-D-Pro-L-Pro-OH discussed in this paper, including the presence of the low-frequency carbonyl peptide band near 1605 cm^{-1}.

Solution Conformational Analysis. In order to investigate the conformational preferences of *t*-Boc-D-Pro-D-Pro-L-Pro-OH in solution we performed an IR absorption analysis on the free acid tripeptide and its benzyl ester in a variety of solvents of divergent polarity (CD_2Cl_2, $CDCl_3$, CH_3CN, *p*-dioxane, TMP, and D_2O) and at different concentrations (Table IV). The spectra in CD_2Cl_2 are shown in Figure 6. The results in $CDCl_3$ and *p*-dioxane are in agreement with those reported by Deber.[3]

In solvents of relatively low polarity (CD_2Cl_2, $CDCl_3$, CH_3CN, and *p*-dioxane), in addition to the bands of the free C=O groups of the acid moiety at 1749–1746

cm^{-1} [8,18,38,39] (the free ester moiety of t-Boc-D-Pro-D-Pro-L-Pro-OBzl has its C=O band at 1745–1741 cm^{-1} [8]), of the tertiary urethane moiety at 1698–1681 cm^{-1},[8,18] and of the tertiary peptide moieties at 1656–1649 cm^{-1},[8,40] a band (or a pronounced shoulder) near 1630 cm^{-1} is seen in the spectra of t-Boc-D-Pro-D-Pro-L-Pro-OH (Table IV andd Figure 6). We assign this latter absorption, which is absent in the spectra of the benzyl ester and tends to disappear with decreasing petide concentration (Table IV and Figure 6), to C=O groups of tertiary peptide moieties of molecules linked intermolecularly through O–H···O=C (peptide) hydrogen bonds. Interestingly, however, this type of association is weaker than that found in the solid state, where the corresponding absorption is visible at 1606 cm^{-1} (Figure 2). Also, it is different from that observed for N-protected dipeptides t-Boc-D-Pro-L-Pro-OH, t-Boc-L-Pro-L-Pro-OH, and Z-L-Pro-L-Pro-OH in concentrated $CDCl_3$ solution, where the acceptor of the intermolecular hydrogen bond was not identified with the peptide carbonyl.[8] Concerning the possible occurrence of intramolecularly hydrogen-bonded structures[4,5] (e.g., the oxy-C_{10} form shown in Figure 1) for t-Boc-D-Pro-D-Pro-L-Pro-OH in dilute solutions, where association is almost completely prevented, we agree with Deber[3] that there are no shifted frequencies indicating the onset of specific structural features.

In TMP, a strong hydrogen-bonding acceptor, all carbonyl groups of t-Boc-D-Pro-D-Pro-L-Pro-OH and its benzyl ester are free. The shift to lower frequency (1733 cm^{-1}) of the C=O absorption of the COOH moiety of the former compound should be attributed to the solvation of the hydroxyl group by TMP through O–H···O=P hydrogen bonds (Table IV).[8,16,18] In D_2O, which acts mainly as a hydrogen-bonding donor, the molecules of t-Boc-D-Pro-D-Pro-L-Pro-OH and its benzyl ester are strongly solvated, as indicated by the location of the bands of the acid C=O group at 1711 cm^{-1} [8,16,18,41] (the band of the ester C=O group of t-Boc-D-Pro-D-Pro-L-Pro-OBzl is seen at 1727 cm^{-1} [8]), of the tertiary urethane C=O groups at 1643–1641 cm^{-1},[6,8,16,18] and of the tertiary peptide C=O groups at 1620–1617 cm^{-1} [8,18] (Table IV). It may be concluded that t-Boc-D-Pro-D-Pro-L-Pro-OH and its benzyl ester are fully solvated in polar solvents.

References and Notes

(1) This work is part 85 of the series "Linear Oligopeptides". Part 84: Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *J. Biol. Chem.*, in press.
(2) (a) Istituto Chimico, Università di Napoli. (b) Istituto di Chimica Organica, Università di Padova.
(3) Deber, C. M. *Macromolecules* **1974**, *7*, 47.
(4) Toniolo, C. In "Bioorganic Chemistry", van Tamelen, E. E., Ed., Academic Press: New York, 1977; Vol. III, pp 265–91.
(5) Toniolo, C. *CRC Crit. Rev. Biochem.* **1980**, *9*, 1.
(6) Benedetti, E.; Pavone, V.; Toniolo, C.; Bonora, G. M.; Palumbo, M. *Macromolecules* **1977**, *10*, 1350.
(7) Benedetti, E.; Palumbo, M.; Bonora, G. M.; Toniolo, C. *Macromolecules* **1976**, *9*, 417.
(8) Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Macromolecules* **1980**, *13*, 1454.
(9) Zimmerman, S. S.; Pottle, M. S.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1.
(10) Benedetti, E.; Pedone, C.; Toniolo, C.; Némethy, G.; Pottle, M. S.; Scheraga, H. A. *Int. J. Pept. Protein Res.* **1980**, *16*, 156.
(11) Benedetti, E. In "Peptides"; Goodman, M., Meienhofer, J., Eds.; Wiley: New York, 1977; pp 257–73.
(12) Benedetti, E. In "Chemistry and Biochemistry of Amino Acids, Peptides and Proteins"; Weinstein, B., Ed.; Marcel Dekker: New York, in press.
(13) Balasubramanian, R.; Lakhsminarayanan, A. V.; Sabesan, M. N.; Tegoni, G.; Venkatesan, K.; Ramachandran, G. N. *Int. J. Protein Res.* **1971**, *3*, 25.
(14) Ashida, T.; Kakudo, M. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1129.
(15) Benedetti, E.; Ciajolo, M. R.; Maisto, A. *Acta Crystallogr., Sect. B* **1974**, *30*, 1783.
(16) Benedetti, E.; Ciajolo, A.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Int. J. Pept. Protein Res.* **1979**, *14*, 130.
(17) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. *J. Am. Chem. Soc.* **1967**, *89*, 5012.
(18) Benedetti, E.; Ciajolo, A.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Macromolecules* **1979**, *12*, 438.
(19) Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M. *Int. J. Biol. Macromol.* **1980**, *2*, 217.
(20) Germain, G.; Main, P.; Woolfson, M. M. *Acta Crystallogr., Sect. A* **1971**, *27*, 368.
(21) Cromer, D. T.; Waber, J. T. In "International Tables for X-Ray Crystallography"; Kynoch Press: Birmingham, England, 1974; Vol. IV, Table 2.2B.
(22) Toniolo, C.; Bonora, G. M. unpublished results.
(23) Rae, I. D.; Raston, C. L.; White, A. H. *Aust. J. Chem.* **1980**, *33*, 215.
(24) IUPAC–IUB Commission on Biochemical Nomenclature *Biochemistry* **1970**, *9*, 3471.
(25) De Tar, D. L. F.; Luthra, N. P. *J. Am. Chem. Soc.* **1977**, *99*, 1232.
(26) Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, G.; Felix, A.; Goodman, M. *Biopolymers* **1981**, *20*, 283.
(27) Carbay-Jaureguiberry, C.; Arnoux, B.; Prangé, T.; Wehri-Altenburger, S.; Pascard, C.; Roques, B. P. *J. Am. Chem. Soc.* **1980**, *102*, 1827.
(28) Kojima, T.; Kido, T.; Itoh, H.; Yamane, T.; Ashida, T. *Acta Crystallogr., Sect. B* **1980**, *36*, 326.
(29) Benedetti, E.; Di Blasio, B; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M., unpublished results.
(30) Bates, R. B.; Hruby, V. J.; Kriek, G. R. *Acta Crystallogr., Sect. B* **1979**, *35*, 188.
(31) Yamane, T.; Ashida, T.; Shimonishi, K.; Kakudo, M.; Sasada, Y. *Acta Crystallogr., Sect B* **1976**, *32*, 2071.
(32) Giordano, F., personal communication.
(33) Madison, V. *Biopolymers* **1977**, *16*, 2671.
(34) Dunitz, J. D.; Strickler, P. In "Structural Chemistry and Molecular Biology"; Rich, A.; Davidson, N., Eds.; W. H. Freeman: San Francisco, 1968; pp 595–602.
(35) Leiserowitz, L. *Acta Crystallogr., Sect. B* **1976**, *32*, 775.
(36) Mitra, J.; Ramackhrishnan, C. *Int. J. Pept. Protein Res.* **1977**, *9*, 27.
(37) Kartha, G.; Ashida, T.; Kakudo, M. *Acta Crystallogr., Sect B* **1974**, *30*, 1861.
(38) Bulmer, J. T.; Shurvell, H. F. *J. Phys. Chem.* **1973**, *77*, 256.
(39) MacPhee, J. A.; Dubois, J. E. *Tetrahedron Lett.* **1978**, 2225.
(40) Boussard, G.; Marraud, M.; Néel, J.; Maigret, B.; Aubry, A. *Biopolymers* **1977**, *16*, 1033.
(41) Wolfenden, R. *Biochemistry* **1978**, *17*, 201.