The Crystal Structure of *tert*-Butyloxy carbonyl-L-prolyl-L-leucylglycine Hydrate

BY T. ASHIDA AND I. TANAKA

Department of Applied Chemistry, Nagoya University, Chikusa-ku, Nagoya 464, Japan

AND Y. SHIMONISHI AND M. KAKUDO

Institute for Protein Research, Osaka University, Suita 564, Japan

(Received 20 January 1977; accepted 24 March 1977)

Crystals of the title compound are monoclinic, $P2_1$, with $a = 10.036$ (1), $b = 18.142$ (2), $c = 6.191$ (1) Å, $\beta = 101.77$ (2)$^\circ$, $Z = 2$. The final $R$ is 0.037 for 1498 non-zero reflexions. The peptide molecule is folded into a $\beta$-turn of type I, in which Pro and Leu are at the corner and an intramolecular hydrogen bond is formed between NH of Gly and CO of the oxycarbonyl group. The folding is essentially the same as that found in some peptides having an $X$-Pro-Leu-Gly- sequence ($X$: any acyl group). Two common rules in the foldings of peptides are found: (i) the $\text{NC}^\circ\text{C'}$ angle of the second residue, Pro in this peptide, in the four residues involved in the $\beta$-turn of type I is extraordinarily larger than the $\text{NC}^\circ\text{C'}$ angles of the residues having other conformations; (ii) the most probable conformation of the Leu residues in peptides has a *trans* zigzag chain of $\text{C}^\circ\text{t}-\text{C}_v-\text{C}^\circ\text{t}-\text{C'}$. The present peptide, *tert*-butyloxy carbonyl(Boc)-Pro-Leu-Gly, is therefore expected to fold into a $\beta$-I turn conformation. In this conformation, however, the C=O bond of Boc and the N--C$^\circ$ bond of Pro should be *cis*, contrary to the *trans* arrangement found in all the Boc (or *tert*-amyl oxy carbonyl)-Pro structures reported so far (Matsuzaki, 1974; Benedetti, Ciajolo & Maisto, 1974; Kartha, Ashida & Kakudo, 1974; Robert 1976).

Experimental

Crystal data

*tert*-Butyloxy carbonyl-L-prolyl-L-leucylglycine hydrate (needle crystal from an ethyl acetate–hexane solution), $C_{18}H_{31}N_{3}O_{7}$, $M_r = 401.46$. Monoclinic, $P2_1$,
Intensity measurements

Intensity measurements were made on a Rigaku four-circle diffractometer with Ni-filtered Cu Ka radiation, ω-2θ scan with range Δ(2θ) = (1.4 + 0.15 tan θ) ° and speed 2 ° min⁻¹, and background counts for 10 s each at both ends of the scan. 1563 reflexions with 2θ < 115 ° were collected, of which 1498 were non-zero. Lorentz and polarization factors were applied, but no absorption corrections were made. The crystal size was 0.3 x 0.2 x 0.16 mm.

Structure determination

The crystal structure was solved by the direct method with the program MULTAN (Germain, Main & Woolfson, 1971). The parameters were refined by the block-diagonal least-squares method (HBLSV: Ashida, 1973), the H atoms being included in the refinement. The final R value was 0.042 for all the reflexions, and 0.037 if 65 zero reflexions were omitted. A few strong reflexions suffered seriously from extinction, but no correction was tried. The atomic scattering factors were taken from International Tables for X-ray Crystallography (1974). The function minimized was Σ w(AF)², with w = 0.6 for |F₀| = 0, and w = [a²(F) - 0.0561|F₀| + 0.0041|F₀|²]⁻¹ for |F₀| > 0, where σ(F) is the standard deviation based on counting statistics. All the calculations were made on the FACOM 230-60 computer of Nagoya University. The final positional parameters are listed in Tables 1 and 2.*

Discussion

The bond lengths and angles are shown in Fig. 1. Some bond angles are seriously affected by the non-bonded interactions. The torsion angles, defined by the IUPAC-IUB Commission on Biochemical Nomenclature (1970), are also shown in Fig. 1. An

---

* Lists of the temperature factors, anisotropic for non-hydrogen atoms and isotropic for H atoms, and structure factors have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 32628 (11 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.
ORTEP drawing (Johnson, 1965) of a molecule is shown in Fig. 2. The equations of the best planes of several planar groups are listed in Table 3.

![Fig. 1. Bond lengths (Å), angles (°) and torsion angles (°).](image)

### Table 3. Best planes

(a) Equations. \(X = ax + cz \cos \beta, Y = by, Z = cz \sin \beta\).

1. \(-0.9374X - 0.3479Y + 0.0159Z = -2.5430\) Boc-Pro amide
2. \(0.6067X - 0.5925Y - 0.5300Z = -0.0762\) Pro-Leu peptide
3. \(0.0984X + 0.4843Y - 0.8693Z = -0.3855\) Leu-Gly peptide
4. \(-0.9522X + 0.2258Y + 0.2059Z = -0.8440\) Gly carboxyl
5. \(-0.9509X - 0.2291Y + 0.2083Z = -2.3413\) Pro ring

(b) Displacements \((\times 10^3 \text{ Å})\) of atoms from the planes

<table>
<thead>
<tr>
<th>Plane</th>
<th>Atom</th>
<th>Displacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)</td>
<td>O(1)</td>
<td>2</td>
</tr>
<tr>
<td>(I)</td>
<td>C(5)</td>
<td>-13</td>
</tr>
<tr>
<td>(I)</td>
<td>O(2)</td>
<td>40</td>
</tr>
<tr>
<td>(II)</td>
<td>N(1)</td>
<td>-118</td>
</tr>
<tr>
<td>(II)</td>
<td>C(6)</td>
<td>33</td>
</tr>
<tr>
<td>(II)</td>
<td>C(9)</td>
<td>63</td>
</tr>
<tr>
<td>(II)</td>
<td>H(N21)</td>
<td>81</td>
</tr>
<tr>
<td>(III)</td>
<td>C(17)</td>
<td>2</td>
</tr>
<tr>
<td>(III)</td>
<td>C(18)</td>
<td>-7</td>
</tr>
<tr>
<td>(III)</td>
<td>O(5)</td>
<td>3</td>
</tr>
<tr>
<td>(III)</td>
<td>O(6)</td>
<td>2</td>
</tr>
<tr>
<td>(IV)</td>
<td>H(O61)</td>
<td>-85</td>
</tr>
<tr>
<td>(V)</td>
<td>C(17)</td>
<td>2</td>
</tr>
<tr>
<td>(V)</td>
<td>C(18)</td>
<td>-7</td>
</tr>
<tr>
<td>(V)</td>
<td>O(5)</td>
<td>3</td>
</tr>
<tr>
<td>(V)</td>
<td>O(6)</td>
<td>2</td>
</tr>
<tr>
<td>(V)</td>
<td>H(O61)</td>
<td>-85</td>
</tr>
</tbody>
</table>

(c) Dihedral angles (°) between the planes

<table>
<thead>
<tr>
<th>Plane</th>
<th>Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I)---(V)</td>
<td>13.0</td>
</tr>
<tr>
<td>(II)---(III)</td>
<td>76.5</td>
</tr>
<tr>
<td>(III)---(IV)</td>
<td>111.8</td>
</tr>
</tbody>
</table>

* Atoms not included in the calculation of the planes.

**β-Turn**

The main-chain folding, which is characterized by the β-I turn, is essentially the same as that found in Z(p-Br)-Gly-Pro-Leu-Gly (Ueki et al., 1969), Z(o-Br)-Gly-Pro-Leu-Gly-Pro (Ueki et al., 1971), Z-Gly-Pro-Leu-Gly-Pro (Bando et al., 1973) and S-benzyl-Cys-Pro-Leu-Gly-NH₂ (Rudko & Low, 1975). Thus it has been shown that all the linear oligopeptides having the X-Pro-Leu-Gly-sequence \((X: \text{any acyl group})\) are folded at the Pro-Leu part into the β-I turn with the intra-molecular hydrogen bond between NH of Gly and C=O of X. The mean torsion angles \((\phi, \psi)\) of these peptides are \((-64.1, -24.7°)\) for Pro and \((-106.6, 12.7°)\) for Leu.

In order to compare these β-I turns in detail, for each pair of peptides the atomic coordinates of one peptide were fitted to those of the other by adjusting the relative positional and orientational parameters by a least-squares procedure; the quantity minimized was \(\Delta E = \sum_{k} (\Delta_{ik}^{2} + \Delta_{jk}^{2})/2\), where \(\Delta_{ik}\) and \(\Delta_{jk}\) are the positional vectors of the \(k\)th atoms in the \(i\)th and \(j\)th peptides, and the summation was made over nine non-hydrogen atoms in the hydrogen-bonded loop of the β-I turn and five others directly bonded to the atoms in the loop. The root-mean-square deviations, \(\Delta E^2\), are 0.089-0.180 Å for group A \([Z(p-Br)-Gly-Pro-Leu-Gly, Z(o-Br)-Gly-Pro-Leu-Gly-Pro, Z-Gly-Pro-Leu-Gly-Pro and the present peptide]\), 0.117 Å for group B (two molecules
of S-benzyl-Cys-Pro-Leu-Gly-NH₂), and 0.372-0.474 Å between A and B. The similarity of the conformations within each group is striking, while the dissimilarity between the two groups is small but significant, and is a cause or effect of the different conformations of the side chains in the two groups.

The NCαC' angle of Pro, 115.8°, in this peptide is extraordinarily larger than usual. This is a common characteristic of the NCαC' angle of the second residues in the β-I turn, the mean angles being 115-1° for nine Pro residues and 113-6° for eight non-Pro residues,* while the mean for about 20 Pro residues having other conformations is 110-5°. This widening of the angle in the β-I turn, which is due to the steric repulsion between the carbonyl group of the first residue and the peptide group joining the second and third residues, should be taken into account in the model building and refinement of the protein structures. On the other hand, no such widening of the angle is observed in the β-II turn, the mean for three structures being 111-3°.

Pyrrolidine ring

The conformation of the pyrrolidine ring is very similar to those in the β-turns of group A, but is different from those in group B. Following the notation of the pyrrolidine-ring conformation in peptides (Ashida & Kakudo, 1974), the rings in group A are C2-Cv-C~-endo in Z(p-Br)-Gly-Pro-Leu-Gly-Pro and C2-Cv-exo in Z-Gly-Pro-Leu-Gly-Pro; these three conformations are very close to each other and between them no significant potential energy barrier exists. The conformations in the two molecules of S-benzyl-Cys-Pro-Leu-Gly-NH₂ are Cv-C'endo. The potential barrier between Cv-exo (and Cβ-endo) and Cv-endo is significant.

Leu side chain

In all the Leu residues in group A, Cβ1-Cv-Cβ2-C' is trans zigzag, while in group B Cβ1-Cv-Cβ2-C' is trans zigzag in one molecule and Cβ2-Cv-Cβ1-C' is trans zigzag in the other. In almost all the other peptides Cβ1-Cv-Cβ2-C' is trans zigzag: Z-Gly-Pro-Leu (Yamane et al., 1976), Pro-Leu-Gly-NH₂ (Reed & Johnson, 1973), Leu-Pro-Gly (Leung & Marsh, 1958), ilamycin B₁ (Itakura, Nakamura, Takada & Takita, 1974), cyclo-Pro-Leu (Karle, 1972), and acetyl-Leu-ethylamide and acetyl-Leu-isopropylamide (Aubry, 1976). An exception is benzoyl-Leu-Gly ethyl ester (Timmins, 1975) in which Cβ2-Cv-C' is trans zigzag. Thus Leu in the peptides usually prefers the trans zigzag Cβ1-Cv-Cβ2-C' chain; that is, the conformation of Leu as found in the present crystal may be the most common one in proteins.

Boc-Pro bond

With respect to the C(5)-N(1) amide bond the two bonds O(1)-C(5) and N(1)-C(6) are trans. This is unique, since the two bonds are cis in all the other Boc(Aoc)-Pro structures reported so far: Boc-(Pro)₄ benzyl ester (Matsuzaki, 1974), Boc-Pro (Benedetti et al., 1974), Aoc-(Pro)₃ (Kartha et al., 1974), Boc-thiazolidine-4-carboxylic acid (Robert, 1976). The conformational potential energies of the Boc-Pro structures were estimated by a method analogous to that of Levitt & Lifson (1969). The result shows that the bond in question usually prefers the cis form to the trans, the former being more stable than the latter by 1-2 kcal mol⁻¹. Since the energy difference is not very large, the trans form, having a hydrogen bond of medium strength at the C=O group, may be more stable than the cis form, which has no hydrogen bond.
In the present peptide the \textit{trans} form is stabilized by forming an intramolecular N(3)–H(31)···O(2) hydrogen bond which is of essential importance for the \textit{β}-turn.

The bond angles around the N(1) atom of the pyrrolidine ring depend strongly on the configuration, \textit{cis} or \textit{trans}, of the C(5)–N(1) amide bond. The bond angle C(5)–N(1)–C(6) is 3.5° smaller, and the C(5)–N(1)–C(9) angle 3.5° larger than their corresponding mean angles, 124.1 and 121.1° respectively, in the \textit{cis} form of the compounds: Boc-(Pro)$_4$ benzyl ester, Boc-Pro, Aoc-(Pro)$_3$, Boc-Pro-Pro-Gly-NH$_2$ (Tanaka \textit{et al.}, 1977). A similar but more exaggerated deformation has been found in the \textit{cis} peptide bonds in the Gly-Pro structure (Yamane \textit{et al.}, 1976) and in small cyclic peptides (e.g. Kartha & Ambady, 1975). Through this deformation of the angles the non-bonded interactions around the C$\alpha$ atom in the \textit{cis} form are nearly identical with those around the C$\alpha$ atom in the \textit{trans} form.

\textbf{Boc group}

The C(3)–C(4)–O(1) angle of 102.5° is extraordinarily compressed from the regular \textit{sp}$^3$ angle, and, owing to this compression, all the C–C–C angles around the C(4) atom are widened by 2 or 3°, releasing the close contacts among the methyl groups. In the ester group the angles C(4)–O(1)–C(5), 121.7°, and O(1)–C(5)–O(2), 126.7°, are much larger than their corresponding angles in many other ester groups, e.g. 116 and 123° in ethyl carbamate (Bracher & Small, 1967) and 113.5 and 122.6° in Z-Gly-Pro (Tanaka \textit{et al.}, 1977). This is a result of the interaction between the bulky tert-butyl group and the carbonyl group. These significant deformations of the bond angles around C(4), O(1) and C(5) are common in several tert-Boc and tert-Aoc groups reported so far.

\textbf{Crystal structure}

The crystal structure is shown in Fig. 3, and the hydrogen bonds are listed in Table 4. The water molecule plays a role in forming the hydrogen-bond network. The O(6)–H(O61)···W(1) hydrogen bond of 2.598 Å is fairly short. The hydrogen bond between the non-ionized carboxyl group (donor) and the water molecule (acceptor) is usually strong (Ashida, Sasada & Kakudo, 1967).

The authors thank Dr T. Yamane of Nagoya University for his help in the data collection, Mr H. Itoh who wrote the program for the comparison of the conformations and Mr N. Takahashi for the conformational-energy calculation. We also thank Professor K. Nakatsu and Dr H. Yoshioka of Kwansei Gakuin University for sending us their unpublished data on tuberactinomycin N. Part of this research was supported by a Scientific Research Grant from the Ministry of Education.

\textbf{References}


The structures of two isomeric sugars, C_{22}H_{29}NO_{10}, have been studied by X-ray diffraction. Isomer (D) was shown to be ethyl 5-O-acetyl-3-O-benzyl-6-deoxy-6-formylamino-1,2-O-isopropylidene-D-glycero-D-talo-heptofuranuronate and crystallizes in the space group P2_12_12, with a = 18.40 (1), b = 14.82 (1), c = 8.65 (1) Å, Z = 4. Isomer (C) was shown to be ethyl 5-O-acetyl-3-O-benzyl-6-deoxy-6-formylamino-1,2-O-isopropylidene-L-glycero-D-allo-heptofuranuronate and crystallizes in P2_1, with a = 19.50 (1), b = 8.50 (1), c = 14.99 (1) Å, β = 109.0 (1)°, Z = 4. The structures were refined to R(D) = 0.05, R(C) = 0.07. Packing forces cause the distortions which lead to the crystallographic dimerization.

**Introduction**

The discovery of the naturally occurring polyoxins (Suhadolnik, 1970) has led to interest in C–C linked sugar α-amino acids, and several compounds with Cα of an α-amino acid moiety attached to C(1), C(2), C(3) or C(4) of furanosyl sugars have been reported (Bischofberger, Brink, De Villiers, Hall & Jordaan, 1977, and references therein). Recently, similar compounds with an α-amino acid moiety attached to C(5) of a furanosyl sugar have been prepared (Hall, Bischofberger, Brink, De Villiers & Jordaan, to be published) by reacting 3-O-benzyl-1,2-O-isopropylidene-α-D-ribo-pentodialdo-1,4-furanose with ethyl isocyanocacetate in ethanol in the presence of sodium cyanide. Hydrolysis of the two oxazoline derivatives obtained (Hoppe, 1974) gave two ethyl 3-O-benzyl-6-deoxy-6-formylamino-1,2-O-isopropylidene-heptofuranuronates [(A) and (B)] which were separated by column chromatography. Acetylation, with acetic