The Crystal Structure of [*N*-(*t*-Pentyloxycarbonyl)-L-prolyl]-L-alanylglycine

Yasuyuki YAMADA, Isao TANAKA,* and Tamaichi ASHIDA

Department of Applied Chemistry, Faculty of Engineering, Nagoya University, Chikusa, Nagoya 464

(Received May 29, 1980)

The crystal structure of the title compound was determined by the X-ray method. The space group is $P2_1$, with $a=18.741(3)$, $b=12.602(1)$, $c=9.553(1)$ Å, $\beta=118.46(2)°$, and $Z=4$. The structure was determined successfully by the vector-space-search method, using a dimer structure of Boc–Pro–Val–Gly–OH as a rigid group. The final R value was 0.075 for 2025 non-zero reflections. Two independent molecules related by a pseudo two-fold axis are linked by β-sheet-type hydrogen bonds. This kind of dimerization commonly occurs in all the Boc–Pro–X–Gly–OH peptides except Boc–Pro–Leu–Gly–OH. The packing similarity of these peptides is discussed.

A series of structure analyses of the [*N*-(*t*-butoxycarbonyl)]-L-prolyl-X-glycine (Boc–Pro–X–Gly–OH, X; any amino acid residue) type oligopeptides have shown that these peptides take either one of two conformations, a folded or an extended one. The folded conformation of Boc–Pro–Leu–Gly–OH[1] is designated as the β-turn structure[2] because of the presence of a 4-1-type intramolecular hydrogen bond. On the other hand, the extended conformation of Boc–Pro–Ile–Gly–OH[3] and Boc–Pro–Val–Gly–OH,[4] both with a branched side chain at the C^β carbon, is characterzied by the formation of β-sheet-type intermolecular hydrogen bonds. These analyses, therefore, seem to illustrate the fact that a slight difference in the side chain results in the complete change of the main-chain folding.

The present study is concerned with the crystal analysis of peptides containing an alanyl residue at the X site. The alanyl residue having only a methyl group as a side chain seems to exert the least steric effect on the main-chain folding. We could expect, therefore, to get some clue as to whether or not the branched structure at the C^β atom has some definite effect on the main-chain folding. Crystals of [*N*-(*t*-butoxycarbonyl)]alanyl peptide, Boc–Pro–Ala–Gly–OH, however, contained four molecules in an asymmetric unit, furthermore, they were twinned. Therefore, we gave up trying to analyze it; as an alternative, [*N*-(*t*-pentyloxycarbonyl)-L-prolyl]-L-alanylglycine (Poc–Pro–Ala–Gly–OH) was studied.

Experimental

The synthesis was performed by first preparing the [*N*-(*t*-butoxycarbonyl)alanyl]glycine benzyl ester (Boc–Ala–Gly–OBzl), followed by the conventional removal of the Boc group and the addition of N-protected proline Poc–Pro–OH, leading to Poc–Pro–Ala–Gly–OBzl. After the removal of the benzyl group by catalytic reduction, the aimed peptide was crystallized from an ethylacetate solution. The crystal data are given in Table 1.

Table 1. Crystal data

Molecular formula	$C_{16}H_{27}O_6N_3$
Molecular weight	357.40
Space group	$P2_1$
Cell constants	$a=18.741(3)$ Å
	$b=12.602(1)$ Å
	$c=9.553(2)$ Å
	$\beta=118.46(2)°$
Density obsd	1.21 g/cm³
calcd	1.20 g/cm³ (for $Z=4$)

A crystal with dimensions of 0.14 mm × 0.07 mm × 0.06 mm was used for the X-ray experiment. A Hilger & Watts four-circle diffractometer equipped with Ni-filtered Cu K_α radiation was used for the data collection. The ω-2θ step scanning mode was adopted. Data were collected up to $2\theta=100°$ for total 2169 reflections, of which 2025 were non-zero. Lorentz and polarization corrections were applied, but no absorption correction was made.

Structure Determination

The direct method was first applied, but did not give reasonable results. The vector-space-search program RICS was next applied on the assumption that the molecule has either a β-turn or an extended conformation.[4] The extended model, 20 atoms of Boc–Pro–Val–Gly–OH excluding the N-terminal and side-chain carbons, seemed to be probable, but the partial structure tangent refinement based on these atomic coordinates did not converge properly. It was further assumed that two independent molecules exist as a dimer, as is observed in Boc–Pro–Ile–Gly–OH[3] and Boc–Pro–Val–Gly–OH.[4] This assumption was quite reasonable in view of the similar packing pattern of the two previously analyzed peptides. The phases calculated from the 40 atoms thus obtained easily led to the whole structure. All the H atoms were located on the difference Fourier map.

The structure was refined by the block-diagonal least-squares program HBLS V.[5] The function minimized was $\sum\omega(|F_o|-|F_c|)^2$, where the weights, ω, were unity for all the non-zero reflections and zero for the zero reflections. The final R value was 0.097 for all the reflections and 0.075 for the non-zero reflections. The atomic scattering factors were taken from the International Tables for X-Ray Crystallography.[6] The final positional parameters are given in Tables 2 and 3.† The isotropic temperature factor, $B=4.4$ Å², was assigned to all the hydrogen atoms.

Discussion

The structure of Poc–Pro–Ala–Gly–OH is charac-

† The F_o and F_c tables and anisotropic temperature factor table are kept as Document No. 8109 at the Chemical Society of Japan.

TABLE 2. THE ATOMIC POSITIONAL PARAMETERS, WITH THEIR e.s.d.'s IN PARENTHESES ($\times 10^4$)

	X	Y	Z	B_{eq}/Å^2 a)
O(1)A	4436(3)	6399(6)	10922(7)	5.2
O(2)A	5758(3)	6571(6)	12799(7)	5.6
O(3)A	4713(3)	4313(6)	9165(7)	5.7
O(4)A	2378(4)	4363(5)	3984(8)	6.3
O(5)A	1179(4)	2280(6)	937(8)	6.3
O(6)A	1294(4)	2200(8)	3367(9)	8.1
N(1)A	5387(4)	6328(6)	10199(8)	4.2
N(2)A	3618(4)	4989(6)	7085(8)	4.6
N(3)A	2927(4)	2756(6)	4661(9)	4.9
C(1)A	3912(10)	4800(20)	12500(25)	13.9
C(2)A	4253(9)	5850(20)	13142(17)	12.1
C(3)A	4355(14)	7704(20)	12720(32)	16.6
C(4)A	3185(8)	6767(14)	10709(19)	9.8
C(5)A	4072(7)	6660(12)	11932(14)	7.2
C(6)A	5230(5)	6428(8)	11413(11)	4.6
C(7)A	4760(5)	6159(8)	8534(10)	4.1
C(8)A	5223(7)	6277(11)	7635(12)	6.8
C(9)A	6097(7)	6150(22)	8851(15)	13.5
C(10)A	6213(5)	6381(9)	10395(12)	5.5
C(11)A	4382(5)	5080(9)	8322(10)	4.6
C(12)A	3168(5)	3990(8)	6745(11)	5.2
C(13)A	2524(7)	4039(12)	7237(15)	7.5
C(14)A	2804(5)	3750(7)	5016(10)	4.3
C(15)A	2492(6)	2370(9)	3041(12)	5.5
C(16)A	1585(5)	2298(7)	2494(12)	5.3
O(1)B	10520(3)	−138(6)	8990(7)	5.0
O(2)B	11433(3)	−333(6)	11617(7)	6.0
O(3)B	9628(3)	1919(6)	9377(8)	6.4
O(4)B	6979(4)	1899(6)	4716(8)	6.5
O(5)B	5514(4)	3982(6)	2162(8)	6.3
O(6)B	6779(4)	3979(7)	2461(8)	7.3
N(1)B	10103(4)	−61(6)	10788(7)	4.1
N(2)B	8538(4)	1258(6)	7260(8)	4.7
N(3)B	7371(4)	3540(7)	5730(8)	5.5
C(1)B	11142(32)	1392(34)	7461(70)	35.6
C(2)B	11696(12)	380(22)	8840(28)	16.1
C(3)B	11381(12)	−1516(15)	8874(21)	12.1
C(4)B	10459(13)	−496(18)	6532(16)	15.2
C(5)B	11052(7)	−416(10)	8314(14)	6.4
C(6)B	10727(5)	−204(8)	10549(10)	4.6
C(7)B	9270(5)	101(8)	9515(11)	4.6
C(8)B	8776(6)	64(10)	10398(14)	6.4
C(9)B	9345(8)	154(23)	12046(16)	13.8
C(10)B	10166(6)	−91(10)	12332(12)	5.5
C(11)B	9163(5)	1142(8)	8730(11)	4.9
C(12)B	8406(5)	2259(8)	6376(12)	5.1
C(13)B	8683(7)	2128(13)	5161(15)	8.2
C(14)B	7510(6)	2518(8)	5560(11)	4.9
C(15)B	6554(6)	3955(9)	4755(11)	5.6
C(16)B	6322(6)	3968(9)	2994(11)	5.6

a) The equivalent isotropic temperature factor as defined by W.C. Hamilton (*Acta Crystallogr.*, **12**, 609 (1959)).

TABLE 3. HYDROGEN POSITIONAL PARAMETERS, WITH THEIR e.s.d.'s IN PARENTHESES ($\times 10^3$)

	x	y	z	Bonded to
H(1)A	403(5)	423(9)	1341(11)	C(1)A
H(2)A	411(5)	448(8)	1159(11)	C(1)A
H(3)A	327(5)	487(8)	1182(11)	C(1)A
H(1)B	1149(5)	196(9)	739(11)	C(1)B
H(2)B	1072(5)	146(8)	782(11)	C(1)B
H(3)B	1104(5)	89(8)	635(11)	C(1)B
H(4)A	483(5)	580(8)	1372(11)	C(2)A
H(5)A	395(5)	613(9)	1387(11)	C(2)A
H(4)B	1193(5)	61(9)	1006(11)	C(2)B
H(5)B	1224(5)	7(8)	884(11)	C(2)B
H(6)A	407(5)	790(9)	1330(11)	C(3)A
H(7)A	429(5)	824(8)	1191(11)	C(3)A
H(8)A	491(5)	772(8)	1346(11)	C(3)A
H(6)B	1161(6)	−166(8)	823(11)	C(3)B
H(7)B	1091(5)	−194(8)	850(11)	C(3)B
H(8)B	1169(5)	−152(8)	1023(11)	C(3)B
H(9)A	282(5)	677(8)	1142(11)	C(4)A
H(10)A	298(5)	604(8)	1002(11)	C(4)A
H(11)A	312(5)	736(8)	981(11)	C(4)A
H(9)B	1076(5)	−57(8)	590(11)	C(4)B
H(10)B	1019(5)	5(9)	611(11)	C(4)B
H(11)B	994(5)	−118(8)	622(11)	C(4)B
H(12)A	436(5)	671(8)	822(11)	C(7)A
H(12)B	907(5)	−51(8)	858(11)	C(7)B
H(13)A	522(5)	695(9)	723(11)	C(8)A
H(14)A	512(5)	584(8)	681(11)	C(8)A
H(13)B	853(5)	−56(8)	1035(11)	C(8)B
H(14)B	832(5)	55(8)	1006(11)	C(8)B
H(15)A	649(5)	657(8)	849(11)	C(9)A
H(16)A	632(5)	541(8)	892(11)	C(9)A
H(15)B	915(5)	−17(8)	1288(11)	C(9)B
H(16)B	934(6)	90(9)	1247(11)	C(9)B
H(17)A	652(5)	707(9)	1074(11)	C(10)A
H(18)A	666(5)	586(8)	1141(11)	C(10)A
H(17)B	1031(5)	−75(9)	1288(11)	C(10)B
H(18)B	1060(5)	44(8)	1321(11)	C(10)B
H(19)A	340(5)	562(8)	643(11)	N(2)A
H(19)B	824(6)	61(8)	679(11)	N(2)B
H(20)A	354(6)	343(8)	731(11)	C(12)A
H(20)B	874(5)	286(8)	710(11)	C(12)B
H(21)A	274(5)	408(8)	851(11)	C(13)A
H(22)A	221(5)	329(8)	702(11)	C(13)A
H(23)A	208(5)	471(8)	661(11)	C(13)A
H(21)B	922(5)	182(8)	547(11)	C(13)B
H(22)B	862(5)	269(8)	454(11)	C(13)B
H(23)B	837(5)	146(8)	433(11)	C(13)B
H(24)A	329(5)	228(8)	539(11)	N(3)A
H(24)B	770(6)	383(8)	661(11)	N(3)B
H(25)A	271(5)	162(8)	304(11)	C(15)A
H(26)A	257(5)	290(8)	225(11)	C(15)A
H(25)B	646(5)	471(8)	516(11)	C(15)B
H(26)B	616(6)	348(8)	510(11)	C(15)B
H(27)A	58(5)	214(9)	30(11)	O(5)A
H(27)B	526(5)	413(9)	88(10)	O(5)B

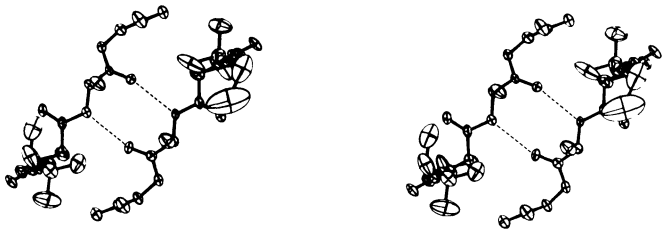


Fig. 1. ORTEP drawing of the dimer structure of Poc-Pro-Ala-Gly-OH *via* β-sheet type hydrogen bonds.

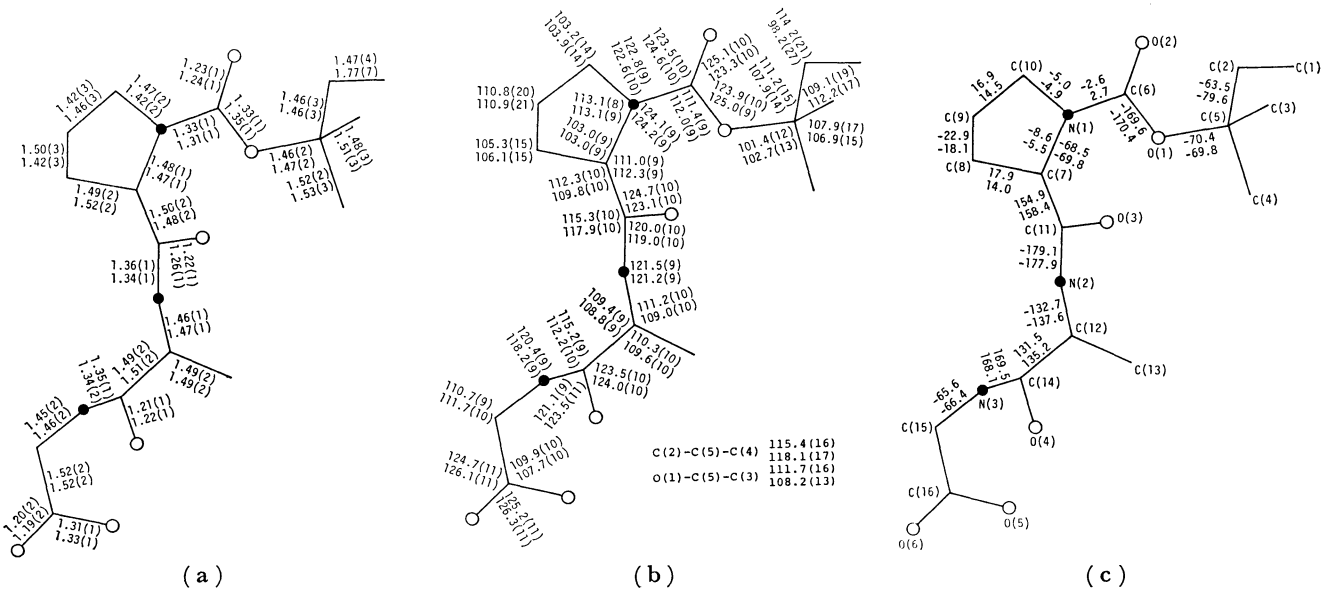
terized by dimer formation with a pair of β-sheet-type hydrogen bonds. The ORTEP drawing[7] of the dimer, viewed along a pseudo-twofold rotation axis, is given in Fig. 1. The bond lengths, bond angles, and torsion angles are shown in Fig. 2. Some abnormal bond lengths in the pentyloxy group may be ascribed to its extremely large thermal motion.

Among the Boc(Poc)-Pro-X-Gly-OH type peptides, those which have Ala, Val, and Ile at the X position crystallize with a similar packing pattern. The crystal-lattice transformation of the alanyl peptide by $a'=a+c$, $b'=b$, and $c'=c$ clearly shows the close relationship between the crystal packings of the alanyl and valyl peptides. Although the space group of the isoleucyl peptide, $P2_12_12$, differs from those of the valyl and alanyl peptides, $P2_1$, the former crystal packing is still close to the latter. That is to say, although the alanyl and valyl peptides are devoid of the twofold axis, there still remains a pseudo-twofold rotation symmetry. The packing similarity is illustrated in Fig. 3, in which dotted lines are drawn in the alanyl and valyl peptide crystals in order to coincide with the unit cell of the isoleucyl peptide.

The similarity of the molecular packing is again obvious in Table 4, which shows the hydrogen-bonding system. No significant difference in the hydrogen-bonding pattern is observed for the two independent molecules (A) and (B); the patterns differ only in the point in which the C-terminal of the molecule interacts with water for the two independent molecules (A) and (B) of Boc-Pro-Val-Gly-OH. Thus, the hydrogen-bonding patterns of three peptides, Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH, and Poc-Pro-Ala-Gly-OH, differ only in the interactions with water molecules, the β-sheet-type hydrogen bonds being completely reserved. The lack of a strict twofold axis in the alanyl and valyl dimer should, therefore,

Table 4. The hydrogen-bond networks in the Boc(Poc)-Pro-X-Gly-OH system

Donor	Acceptor	Distance/Å
	(Boc-Pro-Ile-Gly-OH)	
Gly COOH	Pro C=O	2.60(1)
Ile NH	Ile C=O	2.83(1)
Gly NH	Boc C=O	2.87(1)
Water	Boc C=O	2.80(3)
	(Boc-Pro-Val-Gly-OH)	
Gly(A) COOH	Pro(B) C=O	2.619(8)
Val(A) NH	Val(B) C=O	2.949(7)
Gly(A) NH	Water	2.911(12)
Water	Boc(B) C=O	2.920(10)
Gly(B) COOH	Pro(A) C=O	2.582(9)
Val(B) NH	Val(A) C=O	2.922(7)
Gly(B) NH	Boc(B) C=O	2.942(8)
Water	Boc(A) C=O	2.760(11)
	(Poc-Pro-Ala-Gly-OH)	
Gly(A) COOH	Pro(B) C=O	2.60(1)
Ala(A) NH	Ala(B) C=O	2.86(1)
Gly(A) NH	Poc(A) C=O	2.91(1)
Gly(B) COOH	Pro(A) C=O	2.56(1)
Ala(B) NH	Ala(A) C=O	2.85(1)
Gly(B) NH	Poc(B) C=O	2.84(1)



(a) (b) (c)

Fig. 2. (a) Bond lengths (Å), (b) bond angles (°), and (c) torsion angles (°) with atom numbering system. Upper figures are for (A) and lower for (B) molecule.

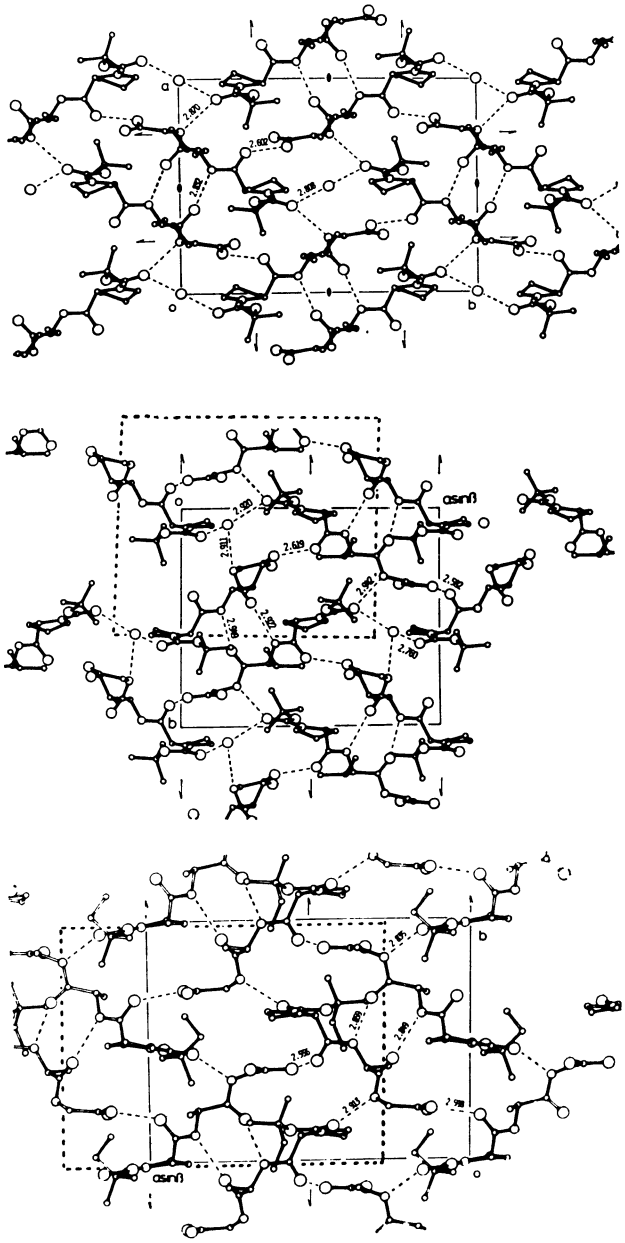


Fig. 3. The comparison of the crystal structures of Boc–Pro–Ile–Gly–OH (top), Boc–Pro–Val–Gly–OH (center), and Poc–Pro–Ala–Gly–OH (bottom). Dotted lines in valyl and alanyl peptide crystals are drawn in order to coincide with the unit cell of isoleucyl peptide.

be attributed to inter-dimer interactions at the C-terminal.

Boc–Pro–Leu–Gly–OH is special in terms of its molecular interaction as well as its molecular conformation. In other words, all three substances besides the leucyl peptide are stabilized by the intermolecular hydrogen bonds, whereas the leucyl peptide is stabilized by the intramolecular one. The dimer structure of the alanyl peptide implies that the C^{β} branched side-chain structure of the valyl and isoleucyl residues is not the only reason for them not to take a β-turn structure, and that the intermolecular interaction emerges as the more important stabilization factor.

No appropriate reason has been given for the uniqueness of the leucyl peptide, but on the basis of these crystal structure analyses, together with other examples, Z–Gly–Pro–Leu–Gly–Pro–OH[8] and Z(*p*-Br)–Gly–Pro–Leu–Gly–OH,[9] we may conclude that there is a stronger tendency for the leucyl residue to be involved in the third site of the β-turn.

References

1) T. Ashida, I. Tanaka, Y. Shimonishi, and M. Kakudo, *Acta Crystallogr., Sect. B*, **33**, 3054 (1977).
2) C. M. Venkatachalam, *Biopolymers*, **6**, 1425 (1968).
3) Y. Yamada, I. Tanaka, and T. Ashida, *Acta Crystallogr., Sect. B*, **36**, 331 (1980).
4) I. Tanaka and T. Ashida, *Acta Crystallogr., Sect. B*, **36**, 2164 (1980).
5) T. Ashida, UNICS-Osaka, The Computation Center, Osaka Univ. 55—61 (1973).
6) "International Tables for X-Ray Crystallography," Birmingham, Kynoch Press (1974), Vol. IV.
7) C. K. Johnson, ORTEP. Oak Ridge National Laboratory Report ORNL-3794 (1965).
8) S. Bando, N. Tanaka, T. Ashida, and M. Kakudo, *Acta Crystallogr., Sect. B*, **34**, 3447 (1978).
9) T. Ueki, T. Ashida, M. Kakudo, Y. Sasada, and Y. Katsube, *Acta Crystallogr., Sect. B*, **25**, 1840 (1969).