The X-ray Studies of the Secondary Structures in Oligopeptides

1979

Isao Tanaka

The X-ray Studies of the Secondary Structures in Oligopeptides

A Doctoral Thesis

Submitted by

Isao Tanaka

Faculty of Science

Osaka University

1979

Acknowledgements

The work of this thesis was carried out under the guidance of Professor Masao Kakudo at Institute for Protein Research, Osaka University. The author is deeply indebted to Professor Kakudo for his cordial guidances and encouragements. His cordial thanks are also due to Professor Tamaichi Ashida at Nagoya University and Associated Professor Nobuo Tanaka at Osaka University for their incessant encouragements and valuable suggestions.

In synthesizing peptide samples, the author is greatly indebted to Associated Professor Yasutsugu Shimonishi and his colleagues for their kind advices.

This work was started in author's school days in Osaka University and continued after his moving to Nagoya University. It was his best fortune to have met many good friends during this period. Without their friendship, this work would not be completed.

Contents

I. INTRODUCTION

The protein crystallography, owing to its rapid progress in this score, has elucidated no less than 100 protein structures and has not only made the visual explanation of the enzymatic activity possible but also contributed to the correlative study between the amino acid sequences and the three dimensional structures of proteins.

Among these proteins, the most well studied group, the serine protease, gives us a good example of the structural homogeneity; in spite of the wide variety of its amino acid sequences the whole three dimensional structures are quite similar with each other and the active site structures are intrinsically identical.

On the other hand the comparative study of the amino acid sequences of the proteins such as ferredoxin and cytochrome c in the different species tells us the protein has both exchangeable and unexchangeable residues.

These facts suggest that the protein has developed through the process of evolution by the exchange of its constituent amino acids to adapt itself to various surroundings with its three dimensional structure of the active site kept unchanged and with the function

conserved.

These unexchangeable or requisite amino acid residues may be divided into two groups, those which directly participate in the enzymatic activity and those which keep whole structure unchanged. Therefore, if in future the artificial molecular design for a more effective enzyme could be possible, it must be based on not only the chemical study of amino acids which have appropriate functional groups for activity but also the structural study of the other amino acids which provide these functional groups with spatially ideal position.

In this thesis, for the first step of the latter study, several oligopeptides whose amino sequences are mutually related were analyzed by the X-ray method and the variability and/or invariability of the three dimensional structure of these peptides were studied.

In comparison with the direct analysis of the proteins, the analysis of oligopeptides as a means of structural study of proteins is more favorable with two respects. First, since the sample is obtained by the synthesis any sequence could be realized, which make systematic study possible, contrasting to the study of protein crystallography which can only elucidate the masterpiece through hundreds millions

years of evolution. Second, because of the low molecular weight high resolutional analysis is possible, which leads to the more precise study of the conformation.

However, a distinctive difference between oligopeptides and proteins is in the molecular surrounding. The soluble protein is enveloped by water, that is difficult to realize in peptide crystals. The results of this research, therefore, would not be able to directly apply to the protein structure; details will be discussed later.

Boc(Aoc)-Pro-X-Gly-OH* was selected as the objective peptides for the investigation, in which X is any amino acid residue such as Ala, Val, Leu, Ile, and Pro. Instead of Boc-Pro-Pro-Gly-OH, Boc-Pro-Pro-Gly-NH_2 was studied in this research, since the former had been already analyzed.(1)

Although these peptides are tripeptides, they could also be designated as tetrapeptides since there is an additional peptide linkage between Boc(Aoc) and Pro groups. For convenience, the position of Boc is called first, Pro is second, X is third and Gly is fourth, throughout the thesis.

Among the analogous peptides, a series of oligo-

peptides, Z-Gly-Pro-Leu-Gly-Pro-OH,[2] Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH,[3] and Z(p-Br)-Gly-Pro-Leu-Gly-OH,[4] had been studied by the X-ray method and shown to have the β-turn structure[5] which is stabilized by the 4-1 type hydrogen bond. The structure of a peptide with a similar sequence S-benzyl-Cys-Pro-Leu-Gly-NH_2[6] had also been reported to take the β-turn structure.

These linear peptides have a proline residue at the second site of the turn. The protein structure analyses at atomic resolution have shown a tendency for proline to play an important role at the second site in the sequential four peptides when it folds into a β-turn.[7] These facts stimulate us to investigate more precisely the structure containing proline at the second site of a four consecutive amino residues.

Although these analyses were carried out for the purpose of the structural investigation of peptides, the accompanying progress of the method of the X-ray analysis was also important, since no definite method was available to oligopeptide. Some were analyzed by the well developed direct method, but others met difficulty with this method. For those peptides, the vector space search method, whose application had

been limited to more rigid molecule, turned out to be very powerful.

---

*The amino acid residues are of L-configuration. For a simpler description, the customary L-designation for individual amino acid residues is omitted. Standard abbreviations for amino acid derivatives and peptides are used according to the IUPAC-IUB Commission on Biochemical Nomenclature. Other abbreviation used are: DCCD, N,N'-dicyclohexylcarbodiimide; TEA, triethylamine; TFA, trifluoroacetic acid; Z, carbobenzoxy; Boc, t-butoxycarbonyl; Aoc, t-amyloxycarbonyl; OBzl, benzyl ester.

II. CRYSTAL STRUCTURE ANALYSIS

1. Boc-Pro-Leu-Gly-OH

1-1. Introduction

Most of the reported oligopeptides which fold into the β-turn in the solid state take leucyl residue at the third site as well as prolyl residue at the second site, such as Z-Gly-Pro-Leu-Gly-Pro-OH,[2] Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH,[3] Z(p-Br)-Gly-Pro-Leu-Gly-OH,[4] and S-benzyl-Cys-Pro-Leu-Gly-NH_2.[6] In the light of this fact, Boc-Pro-Leu-Gly-OH was first analyzed as one of the constituents of the Boc-Pro-X-Gly-OH group. Though the β-turn is possible for this peptide conformation, the C=O bond of Boc and the N-C^{α} bond of Pro should be cis in that case, contrary to the trans arrangement found in all the Boc-Pro structures reported so far.[9-12]

1-2. Experimental

The process commonly applied to the synthesis of Boc-Pro-X-Gly-OH is shown in Table 1.

The first condensation product Boc-Leu-Gly-OBzl was obtained only in oily form, and was provided to the next step immediatly after synthesis. Boc-Pro-Leu-Gly-OBzl was obtained as fine needle crystal,

(mp=114.5-115.5°C). Boc-Pro-Leu-Gly-OH was first obtained as fine crystal but has grown to large crystal within several days by slow evaporation of an ethylacetate solution. Crystal data are given in Table 2.

Intensity measurements were made on a Rigaku four-circle diffractometer with Ni-filtered Cu K_α radiation, ω-2θ scan with range $\Delta(2\theta)=(1.4+0.15\tan\theta)°$ and speed 2° min^{-1}, and background counts for 10s each at both ends of the scan. 1563 reflections with $2\theta<115°$ were collected, of which 1498 were nonzero. Lorents and polarization factors were applied, but no absorption corrections were made. The crystal size was 0.3 × 0.2 × 0.16mm.

1-3. Structure determination

The crystal structure was solved by the direct method with the program MULTAN.[13] The parameters were refined by the block-diagonal least-squares program HBLS V,[14] the H atoms being included in the refinement. The final R value was 0.042 for all the reflections, and 0.037 for nonzero reflections. A few strong reflections suffered seriously from extinction, but no correction was tried. The atomic scattering factors were taken from International

Tables for X-ray Crystallography.[15] The function minimized was $\Sigma\omega(\Delta F)^2$, with $\omega=0.6$ for $|F_o|=0$, and $\omega=\{\sigma^2(F)-0.056|F_o|+0.004|F_o|^2\}^{-1}$ for $|F_o|>0$, where $\sigma(F)$ is the standard deviation based on counting statistics. The final positional and thermal parameters are listed in Tables 3, 4, 5, and 6.

1-4. Discussion

The crystal structure is shown in Fig. 1, and the hydrogen bonds are listed in Table 7. The water molecule plays a role in forming the hydrogen bond network. The O(6)-H(31)··W(1) hydrogen bond of 2.568Å is distinctively short. The intramolecular bond lengths and angles are shown in Fig. 2 along with the torsion angles, defined by the IUPAC-IUB Commission on Biochemical Nomenclature.[16]

An ORTEP drawing[17] of a molecule (Fig. 3) shows that the main-chain folding is classified into β-I turn,[5] which is essentially the same as that found in Z-Gly-Pro-Leu-Gly-Pro-OH,[2] Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH,[3] Z(p-Br)-Gly-Pro-Leu-Gly-OH,[4] and S-benzyl-Cys-Pro-Leu-Gly-Pro-NH_2.[5] Thus it has been shown that all the linear oligopeptides having the X-Pro-Leu-Gly sequence are folded at the Pro-Leu part in the β-I turn with the intramolecular hydrogen

bond between N-H of Gly and C=O of X. The mean torsion angles (ϕ, ψ) of these peptides are (-64.1, -24.7°) for Pro and (-106.6, 12.7°) for Leu.

The $NC^{\alpha}C'$ angle of Pro, 115.8° in this peptide is extraordinarily larger than usual. This is a common characteristic of the $NC^{\alpha}C'$ 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 $NC^{\alpha}C'$ angle in peptides given by Marsh & Donohue[18] is 111°, and the mean for about 20 Pro residues having other conformations is 110.5°. This widening of the angles 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°.

The pyrrolidine-ring of the present peptide takes C_2-C^{γ}-exo conformation following the notation for that ring in peptides.[19]

The leucyl side chains so far analyzed can be classified into three trans-zigzag conformations,

i.e., C' to $C^{\delta 1}$ (type I), N to $C^{\delta 2}$ (type II), and C' to $C^{\delta 2}$ (type III). The conformation of the leucyl side chain in the present crystal is type I, which is found overwhelmingly without any obvious relationships to the main chain torsion angles.[20] Actually, the β-turn conformation is also accompanied with type II or type III.[5] It seems the choice depends on firstly energy preference of the side chain conformation and secondly the crystal packing requirments.

---

*The mean values were obtained from the following structures: for Pro in the β-I turn, Z-Gly-Pro-Leu-Gly-Pro-OH,[2] Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH,[3] Z(p-Br)-Gly-Pro-Leu-Gly-OH,[4] S-Benzyl-Cys-Pro-Leu-Gly-NH_2,[6] Boc-Pro-Leu-Gly-OH, Boc-Pro-Pro-Gly-NH_2, and cyclochlorotine[21]; for non-Pro in the β-I turn, Li antamanide,[22] Na antamanide,[23] tuberactinomycin O,[24] tuberactinomycin N,[25] and cyclo(Gly_4-Ala_2);[26] for β-II turn, Pro-Leu-Gly-NH_2,[27] ferrichrome A,[28] and isobutyl-Pro-Ala-isopropylamide.[29]

Table 1. The process of Boc-Pro-X-Gly-OH synthesis

$$\text{Boc-X-OH} + \text{H-Gly-OBzl}\cdot\text{Tos} \xrightarrow[\text{TEA}]{\text{DCCD}} \text{Boc-X-Gly-OBzl}$$

$$\text{Boc-X-Gly-OBzl} \xrightarrow[\text{TFA}]{} \text{H-X-Gly-OBzl}$$

$$\text{Boc-Pro-OH} + \text{H-X-Gly-OBzl} \xrightarrow[\text{TEA}]{\text{DCCD}} \text{Boc-Pro-X-Gly-OBzl}$$

$$\text{Boc-Pro-X-Gly-OBzl} \xrightarrow[\text{Pd/H}_2]{} \text{Boc-Pro-X-Gly-OH}$$

Table 2. Crystal data

(Boc-Pro-Leu-Gly-OH·H_2O)

Molecular formula	$C_{18}H_{31}N_3O_6\cdot H_2O$
Molecular weight	403.46
Space group	$P2_1$
Cell constant	
	a=10.036(1)Å
	b=18.142(2)
	c= 6.191(1)
	β=101.77(2)°
Density	
obs.	1.208g/cm^3
calc.	1.204 (for Z=2)
mp	105-107°C

Table 3. The atomic positional parameters and their e.s.d.'s in parentheses ($\times 10^4$)

O(1)	890(2)	2604(1)	-473(3)
O(2)	1607(2)	2189(1)	3044(3)
O(3)	3849(2)	503(1)	4705(3)
O(4)	6522(2)	2765(1)	6437(3)
O(5)	4740(2)	2776(1)	10413(3)
O(6)	5109(2)	3957(1)	9875(3)
N(1)	1836(2)	1500(1)	98(3)
N(2)	4425(2)	1487(1)	2872(3)
N(3)	4268(2)	2598(1)	5752(3)
C(1)	-858(3)	3153(2)	1268(5)
C(2)	1452(3)	3761(2)	1533(6)
C(3)	-177(3)	3658(2)	-2086(5)
C(4)	327(3)	3309(2)	121(5)
C(5)	1440(2)	2116(1)	1049(4)
C(6)	2294(2)	846(1)	1428(4)
C(7)	2462(3)	279(1)	-324(5)
C(8)	1375(3)	489(2)	-2284(5)
C(9)	1440(3)	1322(2)	-2273(5)
C(10)	3591(2)	936(1)	3139(4)
C(11)	5670(2)	1662(1)	4453(4)
C(12)	6914(3)	1657(1)	3390(5)
C(13)	7150(3)	929(2)	2310(5)
C(14)	7447(4)	305(2)	3960(8)
C(15)	8305(3)	1026(2)	1032(8)
C(16)	5523(2)	2393(1)	5597(4)
C(17)	4030(3)	3285(2)	6791(4)
C(18)	4671(2)	3301(1)	9226(4)
W(1)	6474(2)	4082(1)	3902(3)

Table 4. Hydrogen positional parameters and their e.s.d.'s in parentheses ($\times 10^3$)

	x	y	z	bonded to
H(1)	-158(3)	291(2)	7(6)	C(1)
H(2)	-121(3)	361(2)	160(5)	C(1)
H(3)	-49(3)	299(2)	279(5)	C(1)
H(4)	222(3)	380(2)	63(5)	C(2)
H(5)	186(3)	363(2)	328(5)	C(2)
H(6)	115(3)	424(2)	165(6)	C(2)
H(7)	63(2)	374(1)	-286(4)	C(3)
H(8)	-71(3)	336(2)	-302(5)	C(3)
H(9)	-65(3)	409(2)	-198(5)	C(3)
H(10)	158(2)	71(1)	227(3)	C(6)
H(11)	340(3)	31(2)	-75(5)	C(7)
H(12)	242(3)	-19(2)	30(5)	C(7)
H(13)	148(2)	28(2)	-363(4)	C(8)
H(14)	40(3)	32(2)	-195(5)	C(8)
H(15)	51(3)	160(2)	-284(6)	C(9)
H(16)	215(3)	157(2)	-305(5)	C(9)
H(17)	428(3)	179(2)	169(5)	N(2)
H(18)	579(2)	129(1)	567(4)	C(11)
H(19)	770(2)	178(1)	468(4)	C(12)
H(20)	683(3)	205(2)	241(5)	C(12)
H(21)	622(3)	73(2)	115(5)	C(13)
H(22)	678(4)	25(3)	502(7)	C(14)
H(23)	829(4)	46(3)	481(7)	C(14)
H(24)	751(4)	-20(3)	315(8)	C(14)
H(25)	846(3)	50(2)	40(6)	C(15)
H(26)	916(4)	119(2)	220(8)	C(15)
H(27)	814(4)	146(3)	-26(7)	C(15)
H(28)	363(2)	238(1)	503(4)	N(3)
H(29)	303(2)	333(2)	689(4)	C(17)

Table 4. (continued)

H(30)	436(2)	370(1)	614(4)	C(17)
H(31)	560(3)	397(1)	1143(4)	O(6)
H(32)	630(3)	447(2)	473(6)	W(1)
H(33)	656(3)	379(2)	484(5)	W(1)

Table 5. The thermal parameters ($\times 10^4$) and their e.s.d.'s in parentheses. Thermal parameters are in the form: $\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)\}$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)	113(2)	22(1)	260(5)	26(2)	45(5)	11(3)
O(2)	96(2)	26(1)	211(5)	15(2)	42(5)	-3(3)
O(3)	118(2)	22(1)	337(6)	-12(2)	21(6)	66(3)
O(4)	95(2)	25(1)	403(7)	-16(2)	-11(6)	-45(4)
O(5)	172(3)	23(1)	305(6)	-25(2)	61(6)	29(3)
O(6)	199(3)	21(1)	313(6)	-35(2)	8(7)	-15(4)
N(1)	89(2)	17(1)	219(6)	8(2)	53(6)	5(3)
N(2)	74(2)	15(1)	253(6)	-8(2)	53(6)	4(3)
N(3)	83(2)	22(1)	279(7)	-7(2)	33(6)	-38(4)
C(1)	128(4)	34(1)	412(11)	40(3)	183(11)	30(6)
C(2)	147(4)	23(1)	479(12)	2(3)	-29(12)	-13(6)
C(3)	139(4)	33(1)	388(11)	45(3)	69(11)	47(6)
C(4)	98(3)	20(1)	307(8)	25(3)	77(9)	5(5)
C(5)	70(3)	21(1)	242(7)	2(2)	38(7)	15(4)
C(6)	81(3)	17(1)	257(7)	-4(2)	83(7)	-3(4)
C(7)	117(3)	19(1)	327(9)	-2(3)	68(9)	-29(5)
C(8)	129(4)	32(1)	278(9)	-6(3)	87(9)	-56(5)
C(9)	149(4)	31(1)	243(8)	11(3)	89(9)	-12(5)
C(10)	77(3)	16(1)	268(8)	4(2)	93(7)	-4(4)
C(11)	79(3)	21(1)	260(8)	4(2)	31(8)	6(4)
C(12)	81(3)	23(1)	406(10)	3(3)	77(9)	3(5)
C(13)	99(3)	28(1)	500(12)	-1(3)	143(10)	-62(7)
C(14)	207(6)	28(1)	874(22)	52(4)	272(18)	8(9)
C(15)	184(6)	49(2)	907(23)	2(5)	486(19)	-112(10)
C(16)	82(3)	21(1)	257(8)	-5(2)	32(7)	-5(4)
C(17)	114(3)	20(1)	298(9)	14(3)	47(9)	-8(5)
C(18)	92(3)	19(1)	284(9)	1(3)	69(8)	-5(5)
W(1)	222(4)	25(1)	321(7)	-14(2)	73(8)	-29(3)

Table 6. The isotropic thermal parameters of hydrogen atoms and their e.s.d.'s in parentheses (×10)

	B		B		B
H(1)	53(8)	H(2)	52(9)	H(3)	35(7)
H(4)	51(8)	H(5)	48(8)	H(6)	73(11)
H(7)	29(6)	H(8)	45(7)	H(9)	40(7)
H(10)	6(5)	H(11)	38(7)	H(12)	31(6)
H(13)	19(6)	H(14)	37(7)	H(15)	50(8)
H(16)	48(8)	H(17)	36(7)	H(18)	21(6)
H(19)	17(5)	H(20)	37(7)	H(21)	38(7)
H(22)	94(13)	H(23)	94(13)	H(24)	96(13)
H(25)	70(11)	H(26)	90(12)	H(27)	87(12)
H(28)	12(5)	H(29)	25(6)	H(30)	11(5)
H(31)	28(6)	H(32)	56(9)	H(33)	49(8)

Table 7. Hydrogen bonds

Donor	Acceptor	Distance(Å)
N(3)	O(2)	2.945
N(2)	$O(5)^{i}$	2.843
O(6)	$W(1)^{ii}$	2.598
W(1)	O(4)	2.855
W(1)	$O(3)^{iii}$	2.758

symmetry code

i x,y,-1.0+z

ii x,y,1.0+z

iii 1.0-x,0.5+y,1.0-z

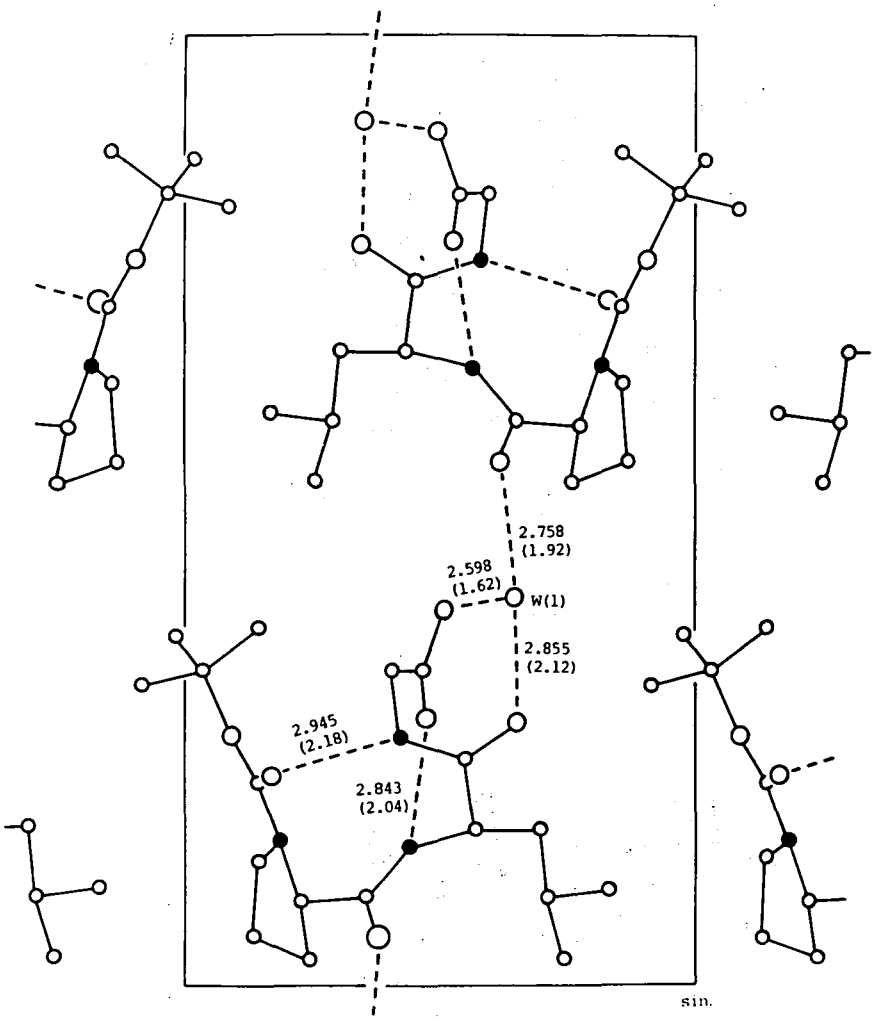
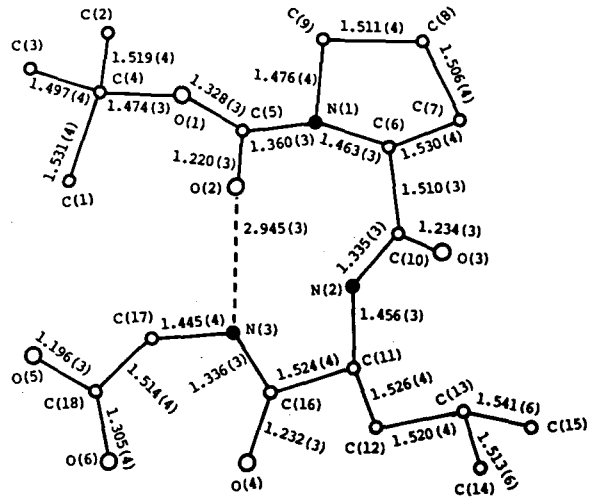
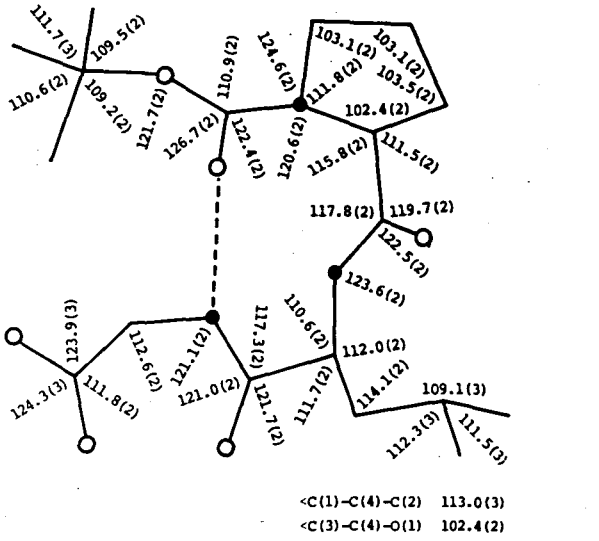


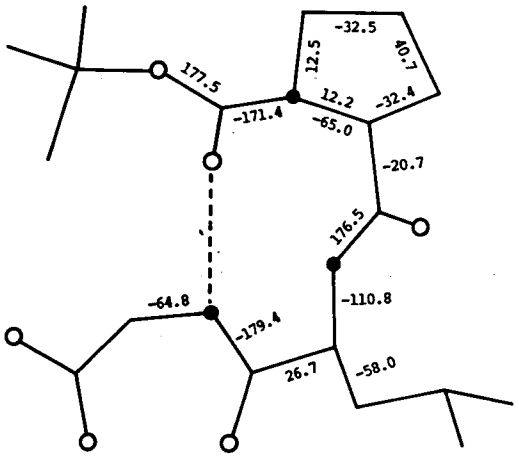
Fig. 1. Crystal structure of Boc-Pro-Leu-Gly-OH. Hydrogen-bond lengths and H···acceptor distances (in parentheses) are shown.


C(2)
C(3)
1.519(4)
C(4)
1.497(4)
1.474(3)
1.328(3)
O(1)
C(5)
1.531(4)
C(1)
1.220(3)
O(2)
1.360(3)
1.463(3)
C(9)
1.511(4)
C(8)
1.476(4)
1.506(4)
N(1)
C(7)
C(6)
1.530(4)
1.510(3)
2.945(3)
1.335(3)
1.234(3)
C(10)
O(3)
N(2)
1.456(3)
C(17)
1.445(4)
N(3)
1.196(3)
O(5)
1.514(4)
1.336(3)
1.524(4)
C(11)
C(18)
1.305(4)
O(6)
C(16)
1.232(3)
O(4)
1.526(4)
C(13)
1.541(6)
C(15)
C(12)
1.520(4)
1.513(6)
C(14)

(a)


111.7(3)
109.5(2)
110.6(2)
109.2(2)
121.7(2)
126.7(2)
110.9(2)
124.6(2)
122.4(2)
120.6(2)
103.1(2)
103.1(2)
111.8(2)
103.5(2)
102.4(2)
115.8(2)
111.5(2)
117.8(2)
119.7(2)
122.5(2)
123.6(2)
110.6(2)
112.0(2)
114.1(2)
109.1(3)
112.3(3)
111.5(3)
111.7(2)
117.3(2)
121.7(2)
121.0(2)
112.6(2)
121.1(2)
123.9(3)
124.3(3)
111.8(2)
<C(1)-C(4)-C(2) 113.0(3)
<C(3)-C(4)-O(1) 102.4(2)

(b)



(c)

Fig. 2. (a) Bond lengths(Å), (b) angles(°) and (c) torsion angles(°) of Boc-Pro-Leu-Gly-OH

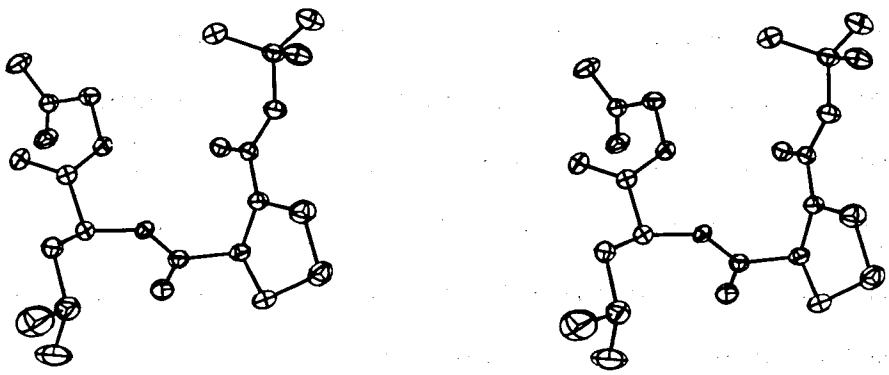


Fig. 3. A stereodrawing of the Boc-Pro-Leu-Gly-OH molecule. The thermal ellipsoids are scaled to include 30% probability.

2. Boc-Pro-Pro-Gly-NH_2

2-1. Introduction

Boc-Pro-Pro-Gly-OH as one of the Boc-Pro-X-Gly-OH type compounds has already been studied by the X-ray method and no intramolecular hydrogen bond, as is found in Boc-Pro-Leu-Gly-OH, has been reported.[1] Here in this section the crystal and molecular structure of Boc-Pro-Pro-Gly-NH_2 was studied. The amide group at the C-terminal provides the present compound with an additional hydrogen bonding capacity, which turned out to be very important.

2-2. Experimental

Boc-Pro-Pro-Gly-NH_2 was prepared by the amination of p-nitrophenylester previously obtained by the condensation reaction of Boc-Pro-Pro-Gly-OH with p-nitrophenol, and crystallized from an ethanol solution. The crystal data are given in Table 8. The crystal of dimensions 0.4 × 0.3 × 0.2 mm was used in the experiment. The lattice parameters were determined by a least-squares refinement from 2θ's of high-angle reflections. The intensity data were collected on a computer-controlled Hilger & Watts four circle diffractometer using Ni-filtered Cu K_α radiation. The ω-2θ step scanning method was adopted. The

symmetrical A setting for the reflections with $0<2\theta<100°$ and fixed-χ setting[30] for those with $100°<2\theta<144°$ were adopted. Of all 1973 reflections up to 144°(2θ) excluding some reflections in the blind regions of the device, 1960 were significantly above background.

Table 8. Crystal data
(Boc-Pro-Pro-Gly-NH_2)

Molecular formula	$C_{17}H_{28}N_4O_5$
Molecular weight	404.48
Space group	$P2_1$
Cell constant	
	a=10.243(1)Å
	b= 9.941(4)
	c=10.550(1)
	β=114.92(1)°
Density	
obs.	1.27g/cm^3
calc.	1.29 (for Z=2)
mp	216-217°C
$\{\alpha\}_D^{24}$	-122.2°(c 1.101,AcOH)

2-3. Structure determination

The structure was solved by the direct method using the program MULTAN.[13] All the non-hydrogen atoms were located on the E-map. The structure was refined by the block-diagonal least-squares method.[14] The function minimized was $\Sigma\omega(|F_o|-|F_c|)^2$, where weights ω were unity for all the non-zero reflections and zero for zero reflections. All the H atoms were found on the ΔF synthesis map, and were included in further refinement. At the final stage, an extinction correction[8] was applied to a few seriously affected reflections. The final R value was 0.041. The final positional and thermal parameters are listed at Tables 9, 10, 11, and 12. All the atomic scattering factors were taken from International Tables for X-ray Crystallography.[15]

2-4. Discussion

The crystal structure projected along the b axis is given in Fig. 4. The bond lengths and angles, and torsion angles are shown in Figs. 5 and 6. Figure 7 illustrates an ORTEP[17] drawing of the molecule with thermal ellipsoids enclosing 50% probability.

The most striking feature concerning the main chain conformation is that the two consecutive prolyl

residues take different conformations, namely an extended (polyproline II type) conformation for the first prolyl residue Pro(1), (ϕ=-60.9°, ψ=156.3°), and a folded (β-turn type) conformation for the second prolyl residue Pro(2), (ϕ=-64.9°, ψ=-23.0°). In the molecules having consecutive prolyl residues, it is impossible for the first prolyl residue to take the folded conformation because of the steric repulsion to the second pyrrolidine ring. This is consistent with the extended conformation found in Aoc-Pro_3-OH[11] Boc-Pro_4-OBzl,[9] and polyproline II.[31] On the other hand the folded conformation for the second prolyl residue in the present peptide is not only free from such steric hindrance but also largely stabilized by the 4-1 hydrogen bond between the first prolyl and the C-terminal amide group.

There are two remarks on the β-turn of this molecule. The first one is that the Pro-Gly sequence at the second and third sites of the turn takes the type I folding, but not type II folding. As suggested by Venkatachalam[5] both I and II types are permissible for the peptides having a glycyl residue at the third site and actually the type II folding had been found in the tripeptide H-Pro-Leu-Gly-NH_2,[27] as well as in many proteins. Although

Venkatachalam expected Pro-Gly sequence could be accomodated in the type II folding,[5] which may be less favorable than the type I folding because of the steric repulsion between the pyrrolidine ring and the main chain. The second one is the ability of the C-terminal amide group to make the β-turn. This kind of β-turn has commonly been observed in the crystal structures such as S-benzyl-Cys-Pro-Leu-Gly-NH_2,[6] and H-Pro-Leu-Gly-NH_2,[27] suggesting it is one of the favorable conformation.

The $NC^{\alpha}C'$ angles, 111.5° for Pro(1) and 113.6° for Pro(2), are slightly but apparently different. A widening of the latter angle is caused by the 4-1 hydrogen bond formation as pointed out in section 1. Following the notation of the pyrrolidine ring conformation,[19] the rings are C_s-C^{γ}-exo for Pro(1) and C_s-C^{γ}-endo for Pro(2). Since both A (C^{γ}-exo) and B (C^{γ}-endo) conformations[32] are found in either main chain conformations, polyproline II type and β-turn type, it seems unlikely to exist any close relationships between the main chain conformations and the pyrrolidine ring conformations.

Table 9. The atomic positional parameters and their e.s.d.'s in parentheses ($\times 10^4$)

	x	y	z
O(1)	6850(2)	2020(2)	9732(2)
O(2)	7104(2)	3221(2)	11662(2)
O(3)	4462(2)	3772(2)	7502(2)
O(4)	3197(2)	4281(3)	3689(3)
O(5)	70(2)	6812(3)	4579(2)
N(1)	4922(2)	2565(3)	10052(2)
N(2)	3816(2)	1797(3)	6349(2)
N(3)	1716(2)	3575(2)	4636(2)
N(4)	2183(3)	5962(3)	6123(3)
C(1)	8344(4)	1499(5)	8642(4)
C(2)	8576(4)	3732(4)	9769(5)
C(3)	9402(3)	1641(5)	11255(4)
C(4)	8325(3)	2242(3)	9885(3)
C(5)	6360(3)	2655(3)	10566(3)
C(6)	4011(3)	1904(3)	8728(2)
C(7)	2482(3)	2146(5)	8663(3)
C(8)	2711(4)	2391(6)	10137(5)
C(9)	4130(3)	3086(4)	10822(3)
C(10)	4150(3)	2573(3)	7497(3)
C(11)	3943(3)	2343(3)	5111(3)
C(12)	3669(4)	1113(4)	4155(3)
C(13)	2772(4)	166(4)	4614(3)
C(14)	3442(3)	355(3)	6196(3)
C(15)	2908(3)	3493(3)	4415(3)
C(16)	690(3)	4653(3)	4064(3)
C(17)	965(3)	5895(4)	4955(3)

Table 10. Hydrogen positional parameters and their e.s.d.'s in parentheses ($\times 10^3$)

	x	y	z	bonded to
H(1)	820(4)	53(4)	872(4)	C(1)
H(2)	760(3)	193(5)	775(3)	C(1)
H(3)	935(4)	154(4)	867(4)	C(1)
H(4)	952(4)	385(5)	970(4)	C(2)
H(5)	770(4)	418(4)	887(4)	C(2)
H(6)	865(4)	424(4)	1047(4)	C(2)
H(7)	900(4)	73(4)	1132(4)	C(3)
H(8)	944(4)	222(5)	1221(4)	C(3)
H(9)	1033(4)	161(5)	1107(4)	C(3)
H(10)	460(4)	277(5)	1187(4)	C(6)
H(11)	395(4)	416(4)	1079(4)	C(7)
H(12)	170(4)	292(4)	1013(4)	C(7)
H(13)	272(5)	131(5)	1061(5)	C(8)
H(14)	199(4)	128(5)	821(4)	C(8)
H(15)	207(5)	297(6)	809(5)	C(9)
H(16)	419(3)	89(4)	873(3)	C(9)
H(17)	271(4)	15(4)	661(3)	C(11)
H(18)	425(3)	-26(4)	666(3)	C(12)
H(19)	160(4)	42(4)	418(4)	C(12)
H(20)	294(4)	-94(4)	447(4)	C(13)
H(21)	327(4)	138(4)	316(4)	C(13)
H(22)	472(4)	71(4)	442(3)	C(14)
H(23)	493(3)	264(3)	538(3)	C(14)
H(24)	152(3)	309(4)	513(3)	N(3)
H(25)	74(4)	492(4)	316(3)	C(16)
H(26)	-31(4)	433(4)	389(3)	C(16)
H(27)	263(4)	521(4)	645(3)	N(4)
H(28)	242(4)	677(5)	662(4)	N(4)

Table 11. The thermal parameters ($\times 10^4$) and their e.s.d.'s in parentheses. Thermal parameters are in the form: $\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)\}$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)	72(2)	101(2)	85(2)	0(4)	51(3)	-61(4)
O(2)	94(2)	117(3)	79(2)	-12(5)	37(4)	-70(4)
O(3)	115(3)	82(2)	96(2)	-39(5)	35(4)	21(4)
O(4)	150(3)	172(4)	181(4)	88(7)	193(6)	215(7)
O(5)	142(3)	118(3)	147(3)	118(6)	65(6)	-7(6)
N(1)	73(2)	107(3)	67(2)	1(5)	46(4)	-26(5)
N(2)	88(2)	78(2)	69(2)	10(6)	59(4)	23(5)
N(3)	78(2)	76(3)	81(2)	17(5)	50(4)	27(5)
N(4)	139(4)	104(4)	103(3)	42(7)	25(6)	-34(6)
C(1)	141(5)	161(6)	144(5)	-14(10)	158(9)	-108(9)
C(2)	188(6)	106(5)	209(7)	-54(10)	232(11)	-52(10)
C(3)	98(4)	177(7)	134(5)	63(9)	55(7)	-10(10)
C(4)	77(3)	95(4)	107(3)	-4(6)	74(6)	-50(6)
C(5)	85(3)	79(3)	73(3)	14(6)	50(5)	1(6)
C(6)	87(3)	98(4)	65(3)	-31(6)	40(5)	-5(6)
C(7)	81(3)	227(8)	113(4)	-81(9)	73(6)	-9(10)
C(8)	136(5)	252(10)	190(6)	-145(12)	188(10)	-193(14)
C(9)	102(4)	144(5)	134(4)	-62(8)	140(7)	-98(9)
C(10)	68(3)	89(3)	71(3)	8(6)	28(5)	18(6)
C(11)	93(3)	108(4)	78(3)	46(6)	84(5)	61(6)
C(12)	205(6)	135(5)	100(4)	124(10)	159(9)	26(8)
C(13)	192(6)	98(4)	91(4)	42(9)	82(8)	-23(7)
C(14)	127(4)	79(4)	86(3)	8(7)	62(6)	13(6)
C(15)	95(3)	104(4)	78(3)	26(6)	71(6)	46(6)
C(16)	83(3)	88(3)	92(3)	40(6)	32(6)	17(6)
C(17)	101(4)	99(4)	90(3)	30(7)	76(6)	11(6)

Table 12. The isotropic thermal parameters of hydrogen atoms and their e.s.d.'s in parentheses (×10)

	B		B		B
H(1)	40(10)	H(2)	35(9)	H(3)	36(10)
H(4)	44(11)	H(5)	44(11)	H(6)	43(10)
H(7)	46(11)	H(8)	47(11)	H(9)	45(11)
H(10)	49(11)	H(11)	44(11)	H(12)	44(11)
H(13)	68(14)	H(14)	51(12)	H(15)	77(16)
H(16)	28(8)	H(17)	34(9)	H(18)	29(9)
H(19)	39(10)	H(20)	46(11)	H(21)	41(10)
H(22)	40(10)	H(23)	19(7)	H(24)	24(8)
H(25)	32(9)	H(26)	34(9)	H(27)	32(9)
H(28)	44(10)				

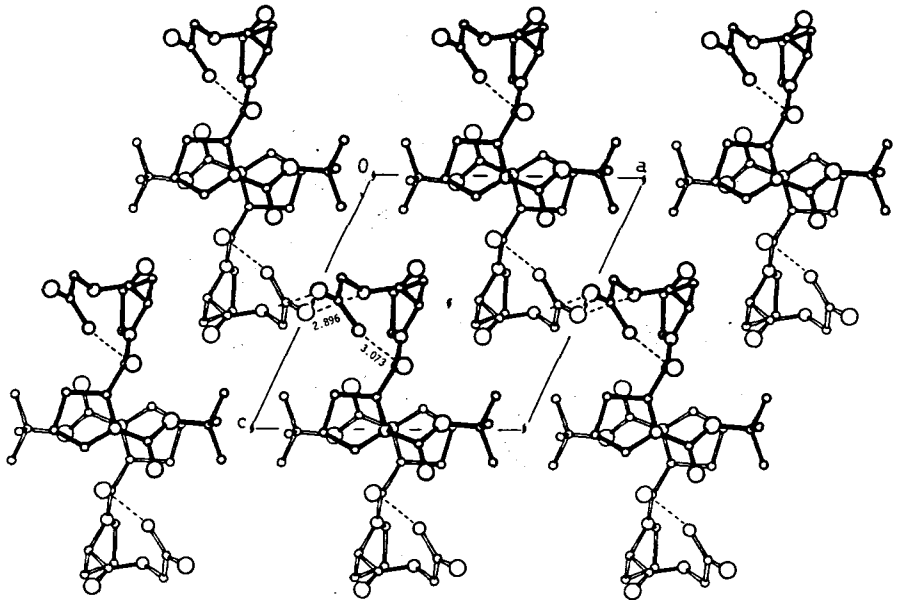


Fig. 4. The crystal structure of Boc-Pro-Pro-Gly-NH_2 viewed along the b axis. Both intra (3.073Å) and intermolecular (2.894Å) hydrogen bonds are shown with dashed lines.

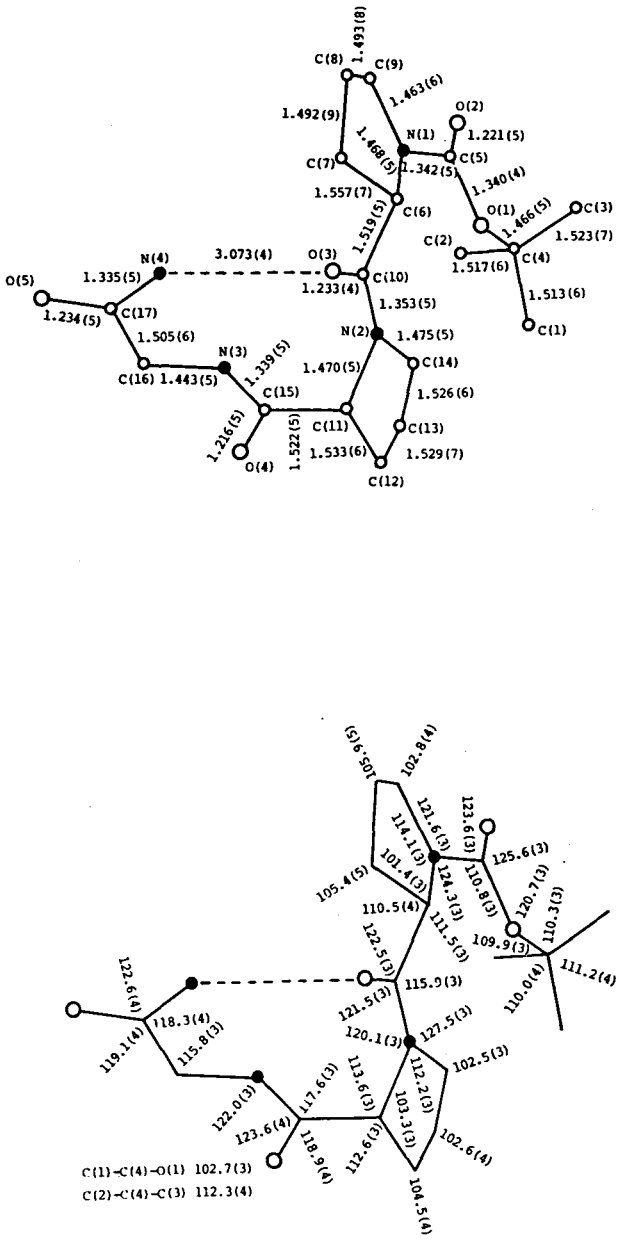


Fig. 5. Bond lengths and angles of Boc-Pro-Pro-Gly-NH_2.

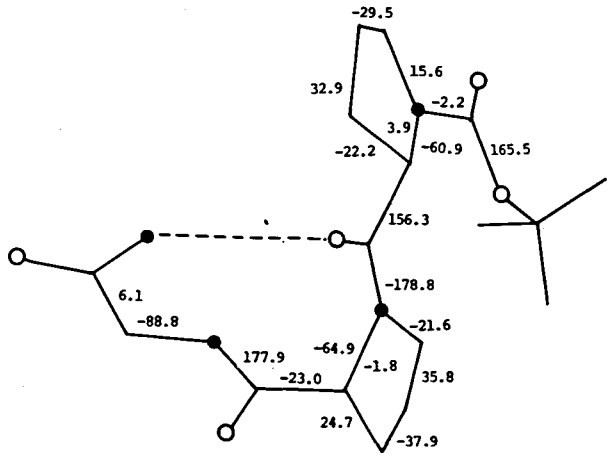


Fig. 6. Conformational angles of the Boc-Pro-Pro-Gly-NH_2 molecule.

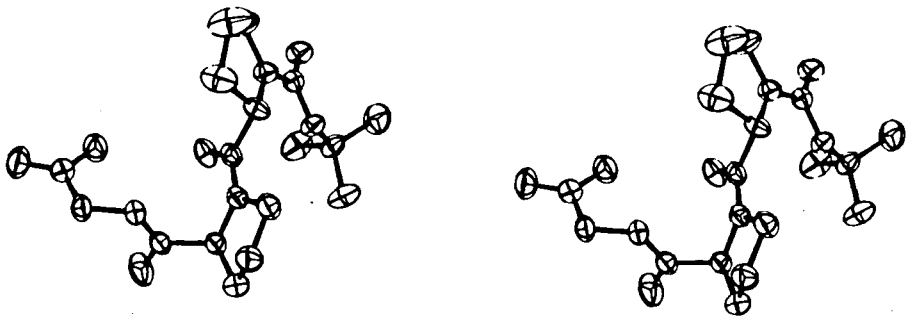


Fig. 7. Stereo drawing of the Boc-Pro-Pro-Gly-NH_2 molecule with thermal ellipsoids drawn to enclose 50% probability.

3. Boc-Pro-Ile-Gly-OH

3-1. Introduction

In the previous two sections, some of the features of the β-turn conformation in the two different peptides were described. These analyses, together with some other examples, seemed to prove the preference of Pro-X peptides for the β-turn. Our attention is then called off to the analysis of Boc-Pro-Ile-Gly-OH whose chemical structure differs from the leucyl peptide only on the branched position of the side chain. The forked structure at the C^{β} atom of the isoleucyl residue may put more ristriction on the main chain folding rather than that at the C^{γ} atom of the leucyl residue.

3-2. Experimental

The synthesis of Boc-Pro-Ile-Gly-OH was done according to the scheme in Table 1. All of the intermediate peptides were obtained as crystals. They are Boc-Ile-Gly-OBzl (mp=97.5-99.0°C) and Boc-Pro-Ile-Gly-OBzl (mp=150.0-152.0°C). The final product Boc-Pro-Ile-Gly-OH was crystallized from an ethylacetate solution. The crystal data are given in Table 13.

The crystal of dimensions 0.20 × 0.25 × 0.30 mm

was used for the X-ray experiment. Hilger & Watts automatic diffractometer with Ni-filtered Cu K_α radiation was used for data collection. The ω-2θ step scanning mode was adopted. The reflections up to 114° (2θ) amount to 2421 of which 2201 were non-zero. Lp corrections were made as usual.

3-3. Structure determination

Structure determination was done by the program MULTAN.[13] Although the calculation using all of the reflections, whose normalized structure factors were above 1.4, did not give a meaningful solution, the calculation elimination one strong low angle reflection (0 0 1), E=2.80, converged to a meaningful result. From the first E-map all of the 24 non-hydrogen atoms excluding three terminal carbon atoms were properly located. These three missing atoms were located at the successively calculated difference Fourier map. The refinement was carried out by the program HBLS V and the consequently obtained difference map showed one large residual peak on the twofold axis. This peak was designated as a water molecule since its position is appropriate for hydrogen bond formation with the carbonyl oxygen atom, the distance being 2.8 Å. However, the peak was

distinctively lower than the expected height showing some disordered structure. Its occupancy was estimated as 0.5 and has been fixed throughout the refinements. The function minimized was $\Sigma\omega(|F_o|-|F_c|)^2$, where weights ω were 0.35567 for $F_o=0$ and $\{\sigma^2(F)+0.15433|F_o|-0.00019|F_o|^2\}^{-1}$ for $|F_o|>0$. Three reflections (0 0 1), (4 1 0) and (2 2 0) were eliminated because of being seriously influenced by the extinction effect. The final R value including all the hydrogen atoms was 0.109 for all the reflections and 0.093 for non-zero reflections. The atomic scattering factors were taken from International Tables for X-ray Crystallography.[15] The final positional and thermal parameters are listed in Tables 14, 15 and 16. The isotropic temperature factor B=4.886 was assigned to all the hydrogen atoms.

3-4. Discussion

As is illustrated by the ORTEP[17] drawing in Fig. 8, the molecule takes an extended conformation rather than the folded β-turn conformation. Such molecular structure is strongly related to its packing scheme in crystalline state especially in the vicinity of the crystalline twofold axis. Two

molecules related by a twofold axis link together by the NH···O hydrogen bonds of isoleucyl residues as is shown in Fig. 9. This hydrogen bonding pattern is that of the well-known antiparallel β-sheet structure. In the present case, however, two molecules are not in a plane in order to avoid the interaction between the Boc group and the COOH group. Such kind of dimerization between two molecules related by twofold or quasi-twofold axis sometimes occurs in peptide crystals e.g., Z-Gly-Pro-OH[33] and Z-Gly-D,L-Pro-OH.[34] Some structural comparison within these dimers will be made in chapter III.

It might be interesting that the crystal and molecular structures of Boc-Pro-Leu-Gly-OH and Boc-Pro-Ile-Gly-OH are totally different, while crystal packing of Leucine and Isoleucine[36] are closely related. This fact suggests that in the amino acid crystal the forked position of the side chain does not affect so much the molecular structure, but in the peptide crystal it plays important role on the main chain folding and therefore on the crystal packing as well.

The water molecule on the twofold axis acts as a hydrogen donor but not as a hydrogen acceptor. This is one of the main difference from Boc-Pro-Val-Gly-OH

crystal in which an additional hydrogen bond exists between water and glycyl residue. The observed disorder may be attributed to this lack of additional hydrogen bond. The hydrogen bond lengths are listed at Table 17.

Figure 10 shows the bond lengths, angles and torsion angles. There is no significant deviation from the standard peptide dimension. The side chain torsion angles of the isoleucyl residue indicate that its conformation is designated as trans form,[37] or C^{δ}-$C^{\gamma 1}$-C^{β}-C^{α}-C' trans-zigzag form. The pyrrolidine ring takes a C^{γ}-endo form.

Table 13. Crystal data
(Boc-Pro-Ile-Gly-OH·1/4H_2O)

Molecular formula	$C_{18}H_{31}N_3O_6 \cdot 1/4H_2O$
Molecular weight	389.46
Space group	$P2_12_12$
Cell constant	
	a=12.909(1)Å
	b=17.567(2)
	c=10.055(3)
Density	
obs.	1.14g/cm^3
calc.	1.14 (for Z=4)
mp	157-158°C

Table 14. The atomic positional parameters and their e.s.d.'s in parentheses ($\times 10^4$)

	x	y	z
O(1)	4267(4)	2500(2)	2116(4)
O(2)	4097(4)	3764(2)	1721(4)
O(3)	6591(3)	2195(2)	352(5)
O(4)	6089(3)	-465(2)	-937(5)
O(5)	8159(3)	-1402(2)	-1134(4)
O(6)	7696(5)	-1817(2)	-3138(5)
N(1)	4660(3)	2954(2)	153(5)
N(2)	5804(3)	1058(2)	24(5)
N(3)	7582(4)	42(2)	-1672(5)
C(1)	4544(12)	3074(7)	4297(10)
C(2)	3978(10)	1717(5)	3928(11)
C(3)	2773(8)	2768(7)	3506(13)
C(4)	3871(6)	2531(4)	3503(7)
C(5)	4319(4)	3115(3)	1345(7)
C(6)	4796(4)	2185(2)	-420(6)
C(7)	4841(7)	2332(4)	-1896(6)
C(8)	5297(7)	3149(4)	-2004(8)
C(9)	4774(5)	3547(3)	-889(7)
C(10)	5817(4)	1819(2)	47(5)
C(11)	6747(4)	614(2)	265(6)
C(12)	6769(5)	238(3)	1620(7)
C(13)	6766(7)	864(4)	2743(8)
C(14)	7736(6)	-261(4)	1772(8)
C(15)	6650(10)	517(5)	4098(10)
C(16)	6775(4)	15(3)	-836(6)
C(17)	7642(5)	-490(3)	-2763(7)
C(18)	7835(5)	-1315(3)	-2353(7)
O(W)	5000(0)	5000(0)	3048(2)

Table 15. Hydrogen positional parameters and their e.s.d.'s in parentheses ($\times 10^3$)

	x	y	z	bonded to
H(1)	448(5)	360(4)	398(7)	C(1)
H(2)	515(5)	284(4)	411(7)	C(1)
H(3)	437(5)	293(3)	498(7)	C(1)
H(4)	467(5)	139(4)	384(6)	C(2)
H(5)	369(5)	161(4)	505(6)	C(2)
H(6)	327(5)	155(4)	332(6)	C(2)
H(7)	245(5)	293(4)	448(7)	C(3)
H(8)	254(5)	312(4)	310(7)	C(3)
H(9)	213(5)	249(4)	276(7)	C(3)
H(10)	418(5)	185(4)	-27(7)	C(6)
H(11)	521(5)	197(4)	-238(7)	C(7)
H(12)	404(5)	233(4)	-235(7)	C(7)
H(13)	614(5)	314(4)	-185(7)	C(8)
H(14)	482(5)	339(4)	-281(6)	C(8)
H(15)	398(5)	374(4)	-122(7)	C(9)
H(16)	526(5)	402(4)	-48(7)	C(9)
H(17)	515(6)	81(4)	-26(6)	N(2)
H(18)	735(6)	90(4)	17(7)	C(11)
H(19)	611(6)	-10(4)	171(6)	C(12)
H(20)	745(5)	121(4)	270(7)	C(13)
H(21)	619(5)	115(4)	261(7)	C(13)
H(22)	764(5)	-63(4)	122(7)	C(14)
H(23)	832(5)	-12(4)	136(6)	C(14)
H(24)	772(5)	-50(4)	278(7)	C(14)
H(25)	587(6)	29(4)	390(7)	C(15)
H(26)	724(5)	14(4)	431(6)	C(15)
H(27)	653(6)	96(4)	482(7)	C(15)
H(28)	819(5)	39(4)	-198(6)	N(3)
H(29)	823(5)	-39(4)	-357(7)	C(17)

Table 15. (continued)

H(30)	687(5)	-47(4)	-343(7)	C(17)
H(31)	814(5)	-177(4)	-88(7)	O(5)
H(32)	495(12)	447(8)	241(14)	O(w)

Table 16. The thermal parameters ($\times 10^4$) and their e.s.d.'s in parentheses. Thermal parameters are in the form: $\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)\}$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)	122(4)	29(1)	149(5)	17(4)	19(8)	-5(4)
O(2)	115(3)	26(1)	179(6)	39(3)	-1(9)	-25(4)
O(3)	72(2)	22(1)	214(6)	-1(2)	-40(7)	-13(4)
O(4)	69(2)	30(1)	237(8)	-15(3)	39(8)	-38(5)
O(5)	107(3)	21(1)	174(5)	16(3)	-67(8)	-2(4)
O(6)	186(6)	40(1)	196(8)	22(6)	-112(12)	-57(6)
N(1)	83(3)	20(1)	148(6)	13(3)	-10(8)	5(4)
N(2)	59(2)	18(1)	162(6)	5(2)	8(7)	-13(4)
N(3)	91(3)	24(1)	152(5)	-7(4)	45(8)	14(5)
C(1)	273(18)	109(7)	148(12)	-93(20)	-64(24)	12(16)
C(2)	213(12)	58(3)	262(16)	49(12)	179(25)	93(14)
C(3)	121(7)	105(6)	303(19)	48(13)	142(23)	83(20)
C(4)	110(5)	45(2)	129(8)	5(6)	23(12)	-2(8)
C(5)	66(3)	27(1)	164(8)	14(4)	2(9)	-7(6)
C(6)	82(4)	20(1)	149(7)	6(4)	-11(10)	-19(5)
C(7)	143(7)	41(2)	116(7)	4(7)	-35(13)	-14(7)
C(8)	148(8)	48(2)	166(10)	-10(8)	-23(15)	32(9)
C(9)	105(5)	32(2)	182(9)	6(5)	-26(12)	45(8)
C(10)	68(3)	23(1)	123(6)	3(3)	-13(8)	-8(5)
C(11)	68(3)	22(1)	144(7)	-4(3)	-7(9)	-8(5)
C(12)	87(4)	31(1)	174(8)	9(5)	26(12)	-2(7)
C(13)	144(7)	45(2)	167(9)	42(8)	-49(16)	-5(9)
C(14)	92(5)	47(2)	210(11)	29(6)	-33(13)	45(9)
C(15)	204(12)	63(3)	188(12)	34(12)	-83(22)	0(12)
C(16)	61(3)	21(1)	182(8)	4(3)	-19(9)	4(6)
C(17)	108(5)	24(1)	172(9)	10(5)	30(12)	10(6)
C(18)	84(4)	27(1)	161(8)	4(4)	21(10)	10(6)
O(w)	136(17)	73(9)	392(47)	-26(21)	0(0)	0(0)

Table 17. Hydrogen bonds

Donor	Acceptor	Distance(Å)
O(5)	O(3)^{i}	2.602
N(2)	O(4)^{ii}	2.832
N(3)	O(2)^{iii}	2.870
O(w)	O(2)	2.808

symmetry code

i 1.5-x,-0.5+y,-z

ii 1.0-x,-y,z

iii 0.5+x,0.5-y,-z

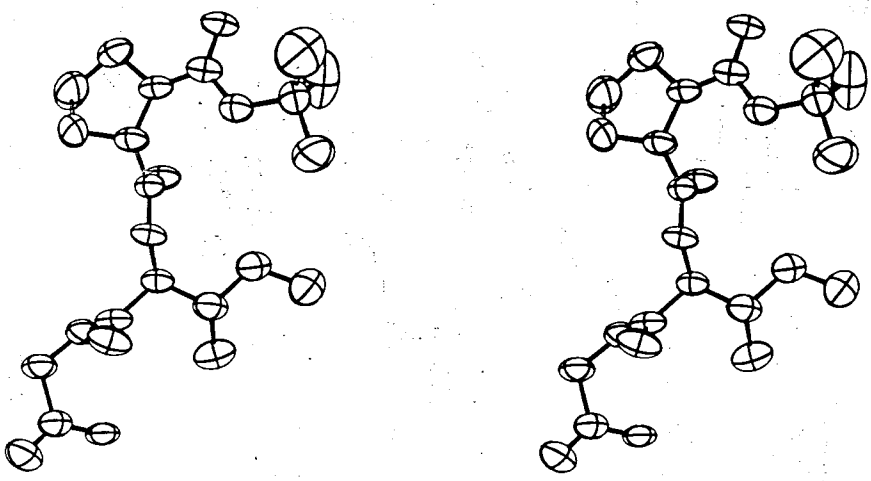


Fig. 8. Stereo drawing of the Boc-Pro-Ile-Gly-OH molecule with thermal ellipsoids drawn to enclose 50% probability.

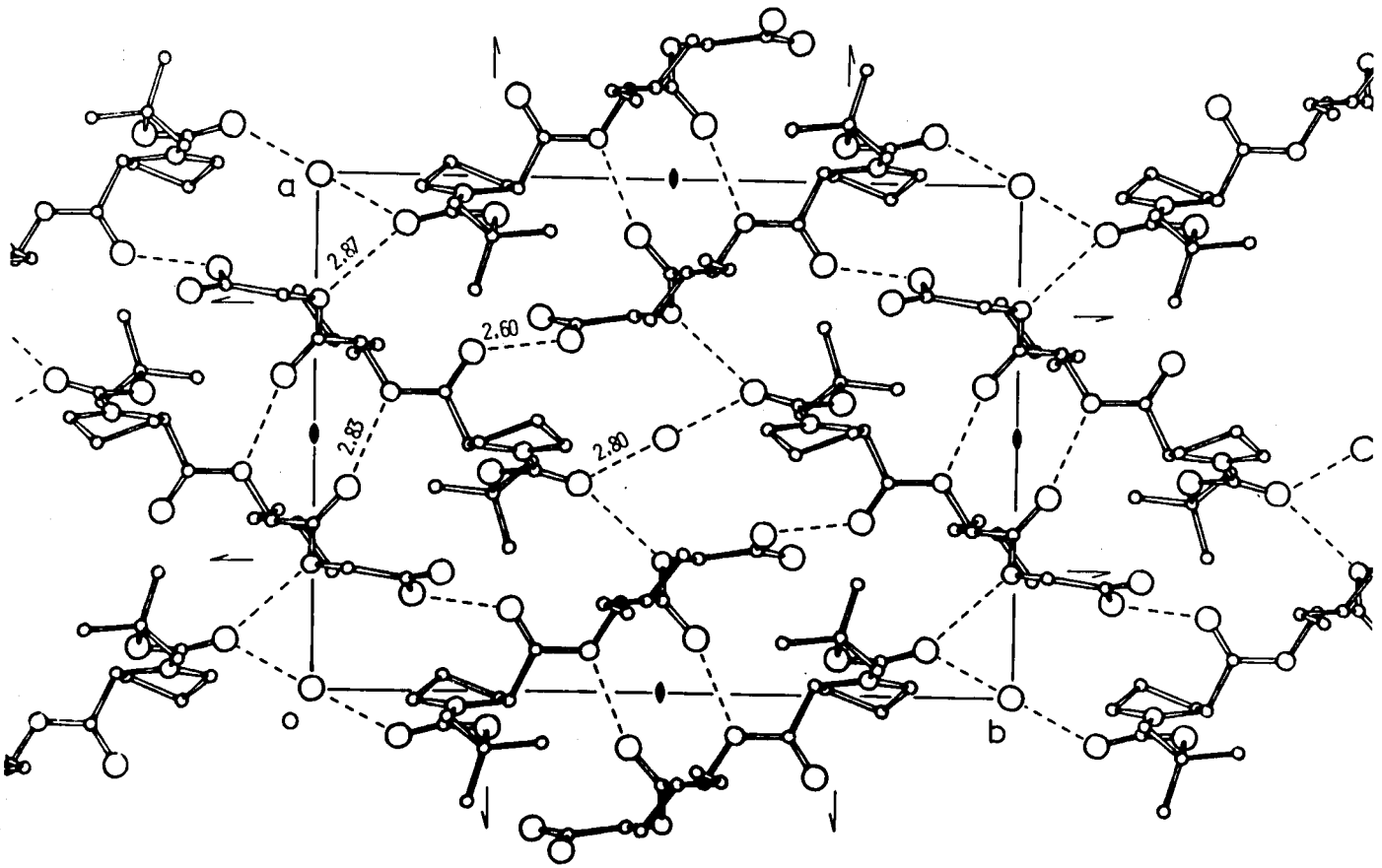
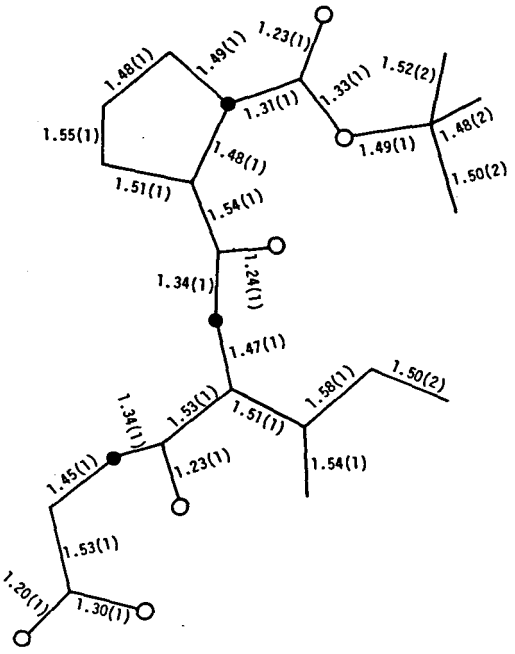
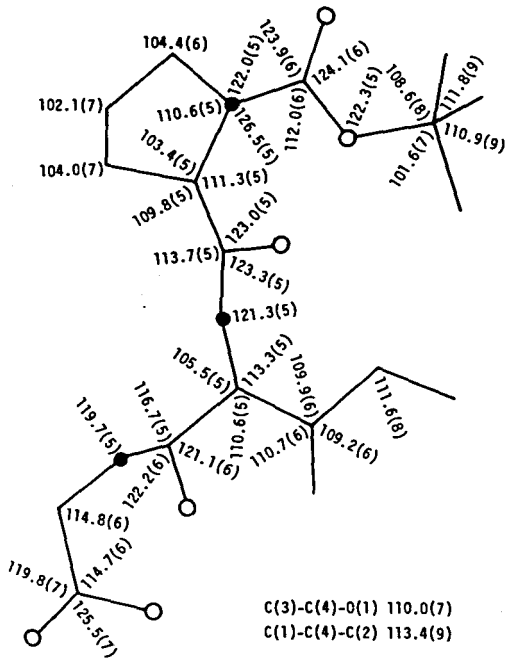


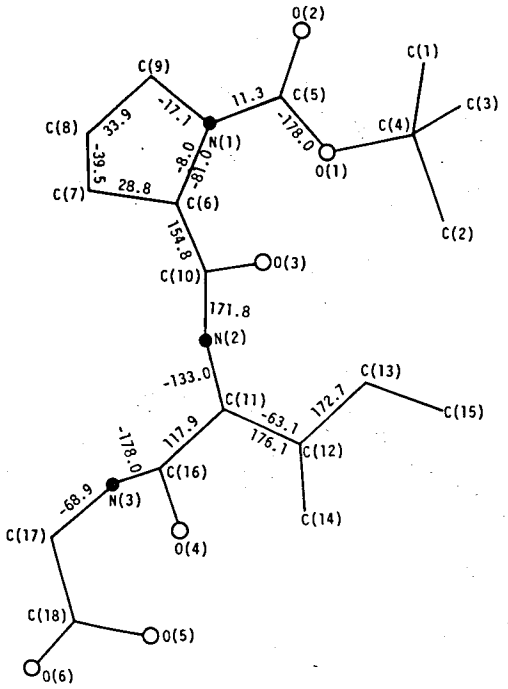
Fig. 9. The crystal structure of Boc-Pro-Ile-Gly-OH viewed along the c axis. The hydrogen bonds are shown by the broken lines.


1.23(1)
1.49(1)
1.48(1)
1.55(1)
1.31(1)
1.33(1)
1.52(2)
1.48(2)
1.49(1)
1.50(2)
1.48(1)
1.51(1)
1.54(1)
1.24(1)
1.34(1)
1.47(1)
1.50(2)
1.58(1)
1.34(1)
1.53(1)
1.51(1)
1.45(1)
1.23(1)
1.54(1)
1.53(1)
1.20(1)
1.30(1)

(a)


104.4(6)
123.9(6)
122.0(5)
124.1(6)
102.1(7)
110.6(5)
126.5(5)
112.0(6)
122.3(5)
108.6(8)
111.8(9)
110.9(9)
101.6(7)
103.4(5)
104.0(7)
111.3(5)
109.8(5)
123.0(5)
113.7(5)
123.3(5)
121.3(5)
105.5(5)
113.3(5)
109.9(6)
111.6(8)
116.7(5)
119.7(5)
110.6(5)
121.1(6)
110.7(6)
109.2(6)
122.2(6)
114.8(6)
114.7(6)
119.8(7)
125.5(7)
C(3)-C(4)-O(1) 110.0(7)
C(1)-C(4)-C(2) 113.4(9)

(b)



(c)

Fig. 10. (a) Bond lengths(Å), (b) angles(°) and (c) torsion angles(°) of Boc-Pro-Ile-Gly-OH.

4. Boc-Pro-Val-Gly-OH and Aoc-Pro-Val-Gly-OH

4-1. Introduction

The throughly different conformations found in the Boc-Pro-Leu-Gly-OH and Boc-Pro-Ile-Gly-OH crystals prompted us to study further the structure of Boc-Pro-X-Gly-OH peptides. The fourth peptide analyzed was Boc-Pro-Val-Gly-OH having a forked isopropyl group.

Boc-Pro-Val-Gly-OH crystallizes with two independent molecules in an asymmetric unit, making the analysis extremely difficult. This difficulty could sometimes be avoided by a slight chemical modification of the molecule. From this viewpoint Boc group was substituted by Aoc group, expecting some effect on the crystal packing. Unfortunately, this trial was not fruitful because Aoc-Pro-Val-Gly-OH also crystallizes into similar packing mode. The same but successful trial will be seen in the following section.

4-2. Experimental

The synthesis was set about with preparing Boc-Val-Gly-OBzl (mp=73.0-75.5°C) followed by the addition of two differently N-protected prolines Boc-Pro-OH and Aoc-Pro-OH, leading to Boc-Pro-Val-

Gly-OBzl (mp=80.0-81.0°C) and Aoc-Pro-Val-Gly-OBzl (mp=96.0-97.5°C) respectively. Each of the final products Boc-Pro-Val-Gly-OH and Aoc-Pro-Val-Gly-OH was crystallized from an ethylacetate solution. Crystal data are given in Tables 18 and 19.

The X-ray intensity data were collected only for Boc-Pro-Val-Gly-OH. The crystal of dimensions 0.15 × 0.15 × 0.3 mm was used for this purpose. Intensity measurements were made on a Rigaku four-circle diffractometer with Ni-filtered Cu K_α radiation up to $2\theta=120°$. In total, 3233 independent reflections were collected, 375 of which were considered to be non-observed. Lp corrections were made but no absorption correction was made.

4-3. Structure determination

In order to solve the structure, a direct method was first applied using the program MULTAN[13] and later MULTAN74 which was a revised version of the former. This attempt, however, had to be given up because no reasonable molecular fragment was obtained in the E-map.

The vector space search method was next applied assuming that this peptide molecule takes a similar conformation to either Boc-Pro-Leu-Gly-OH or Boc-Pro-Ile-Gly-OH. The trial using Boc-Pro-

Ile-Gly-OH as rigid group was straightforwardly successful and the phases based on 20 atomic coordinates, thus obtained, were used as starting phases for tangent refinement. The detailed description of the vector space search program RICS used for this experiment will be given later in chapter IV.

The structure was refined by the program HBLS V. The function minimized was $\Sigma\omega(|F_o|-|F_c|)^2$, where weights ω were unity for all the non-zero reflections and zero for zero reflections. The final R value including all the hydrogen atoms was 0.084 for all the reflections and 0.064 for non-zero reflections. The final positional and thermal parameters are listed in Tables 20, 21 and 22. All the atomic scattering factors were taken from International Tables for X-ray Crystallography.[15] The isotropic temperature factor B=5.141 was assigned to all the hydrogen atoms.

4-4. Discussion

The crystal structure projected along the b axis is given in Fig. 11. The bond lengths, angles and torsion angles are given in Fig. 12. As the structure analysis suggests itself, this peptide possesses a very similar conformation to that of Boc-Pro-Ile-Gly-OH. The structure of one molecule

(molecule(A)) by the ORTEP drawing[17] is shown in Fig. 13. The overall shape of the other molecule (molecule(B)) does not differ so much from that of (A). The more precise comparison in terms of the torsion angles and also by the least squares fitting programs (described in the next chapter) revealed that the conformational similarity between Boc-Pro-Val-Gly-OH (A) and Boc-Pro-Ile-Gly-OH is larger than that between Boc-Pro-Val-Gly-OH (A) and (B). This observation seems to suggest that conformational fluctuation to this extent (18.4° at most) is not directly caused by the difference in the residue but by the crystal packing requirement.

The observed similarity between the molecular structures of Boc-Pro-Vla-Gly-OH and Boc-Pro-Ile-Gly-OH extends further to molecular interactions in the crystal. That is to say, the valyl peptides dimerize by the NH···O hydrogen bonds in a similar way as in the isoleucyl peptides, although the former does not have an exact twofold rotation axis within the dimer. The abundant van der Waals contacts within the dimer suggest a plausible process for crystallization which involves dimerization at an early stage. The water molecule, which exists on the twofold axis with disordered structure in the isoleucyl crystal, occurs

in the present crystal with full occupancy at general position. Two molecules (A) and (B) differs mainly in the scheme of the hydrogen bondings to the water molecules (Table 23).

It may be worthwhile to note a commonly observed phenomenon of the side chain conformations in valine,[37] isoleucine[36] and Boc-Pro-Val-Gly-OH. All of these crystals contain two different molecules with side chain conformations being different. The side chains of the present molecules are trans for (A) molecule and gauche II for (B), while trans and gauche I for both valine and isoleucine.[37] This phenomenon may reflect the fact that these different side chain conformations are almost equally stable.

There is no doubt that Aoc-Pro-Val-Gly-OH has a similar packing pattern in view of the consistency of cell dimensions and space groups.

Table 18. Crystal data

(Boc-Pro-Val-Gly-OH·1/2H_2O)

Molecular formula	$C_{17}H_{29}O_6N_3 \cdot 1/2H_2O$
Molecular weight	371.43
Space group	$P2_1$
Cell constants	
	a=15.783(2)Å
	b=13.428(2)
	c= 9.815(2)
	β=90.94(1)°
Density	
obs.	1.19g/cm^3
calc.	1.21 (for Z=4)
mp	171-172°C

Table 19. Crystal data

(Aoc-Pro-Val-Gly-OH·1/2H_2O)

Molecular formula	$C_{19}H_{31}O_6N_3 \cdot 1/2H_2O$
Molecular weight	385.50
Space group	$P2_1$
Cell constants	
	a=16.254(1)Å
	b=13.575(1)
	c= 9.829(1)
	β=92.42(1)°
Density	
obs.	1.21g/cm^3
calc.	1.21 (for Z=4)
mp	154-155°C

Table 20. The atomic positional parameters and their e.s.d.'s in parentheses ($\times 10^4$)

	x	y	z
O(1)A	588(2)	5999(4)	-1394(4)
O(2)A	-818(2)	6162(5)	-1913(4)
O(3)A	416(2)	3905(3)	578(5)
O(4)A	2693(3)	4223(3)	3895(5)
O(5)A	3588(3)	2112(4)	4666(5)
O(6)A	3521(5)	2562(8)	6781(7)
N(1)A	-353(2)	5739(4)	195(5)
N(2)A	1504(2)	4562(4)	1762(5)
N(3)A	2011(3)	2776(5)	4098(7)
C(1)A	624(7)	7381(10)	-2976(12)
C(2)A	1840(5)	6282(10)	-2544(9)
C(3)A	613(6)	5669(12)	-3816(10)
C(4)A	894(4)	6346(7)	-2717(7)
C(5)A	-226(3)	5986(6)	-1111(6)
C(6)A	322(3)	5634(5)	1225(6)
C(7)A	-171(4)	5723(7)	2551(7)
C(8)A	-1034(5)	5438(11)	2205(8)
C(9)A	-1216(3)	5680(6)	768(6)
C(10)A	753(3)	4622(5)	1125(5)
C(11)A	1947(3)	3615(4)	1903(7)
C(12)A	2724(4)	3515(6)	962(8)
C(13)A	3130(5)	2521(6)	1116(9)
C(14)A	2406(6)	3687(8)	-523(9)
C(15)A	2262(3)	3565(5)	3382(7)
C(16)A	2269(5)	2671(8)	5512(10)
C(17)A	3204(5)	2436(7)	5717(9)
O(1)B	4411(2)	-437(3)	6363(4)
O(2)B	3324(2)	-485(3)	4826(5)
O(3)B	5208(2)	1785(3)	5057(4)

Table 20. (continued)

O(4)B	8062(2)	1439(3)	6832(4)
O(5)B	9794(3)	3539(5)	8179(5)
O(6)B	8476(3)	3422(5)	8939(5)
N(1)B	4647(3)	-86(4)	4217(4)
N(2)B	6370(2)	1103(3)	6072(5)
N(3)B	7860(3)	3048(4)	6356(5)
C(1)B	4636(6)	-879(9)	8625(9)
C(2)B	3353(7)	40(10)	7915(12)
C(3)B	3523(7)	-1754(8)	7261(10)
C(4)B	3945(4)	-783(5)	7541(7)
C(5)B	4074(3)	-352(4)	5117(6)
C(6)B	5553(3)	63(5)	4524(6)
C(7)B	5930(5)	147(8)	3093(8)
C(8)B	5185(8)	453(11)	2201(9)
C(9)B	4404(5)	78(6)	2769(7)
C(10)B	5669(3)	1057(5)	5253(6)
C(11)B	6637(3)	2043(4)	6698(6)
C(12)B	6352(4)	2159(5)	8202(6)
C(13)B	5385(4)	2176(7)	8290(8)
C(14)B	6743(5)	1393(7)	9144(8)
C(15)B	7585(3)	2135(4)	6616(5)
C(16)B	8748(4)	3299(5)	6552(7)
C(17)B	8981(4)	3432(6)	8028(7)
O(W)	1783(4)	621(6)	4156(8)

Table 21. The thermal parameters ($\times 10^4$) and their e.s.d.'s in parentheses. Thermal parameters are in the form: $\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)\}$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)A	40(1)	121(4)	123(5)	-12(4)	-13(4)	31(8)
O(2)A	43(2)	168(6)	151(6)	25(5)	-60(5)	26(10)
O(3)A	55(2)	79(3)	204(7)	-28(4)	-73(6)	-38(8)
O(4)A	61(2)	83(3)	193(7)	-4(4)	-89(6)	25(8)
O(5)A	54(2)	104(4)	209(7)	24(5)	-37(7)	-5(10)
O(6)A	146(5)	289(13)	203(10)	239(15)	-40(12)	-46(18)
N(1)A	29(1)	98(4)	138(6)	2(5)	-33(5)	27(9)
N(2)A	38(2)	68(3)	128(6)	-11(4)	-42(5)	12(8)
N(3)A	45(2)	91(4)	273(11)	10(6)	17(8)	134(12)
C(1)A	119(7)	178(12)	265(18)	36(16)	101(18)	200(25)
C(2)A	50(3)	263(16)	190(12)	-68(12)	-5(10)	182(23)
C(3)A	102(6)	291(18)	183(13)	-199(19)	57(14)	-152(26)
C(4)A	58(3)	132(7)	122(8)	-46(8)	-1(8)	30(13)
C(5)A	38(2)	110(6)	116(7)	5(6)	-19(6)	0(11)
C(6)A	40(2)	79(4)	112(7)	-11(6)	-27(6)	-10(9)
C(7)A	50(3)	138(7)	127(8)	19(8)	-18(8)	-20(13)
C(8)A	53(3)	292(18)	163(11)	-10(14)	37(10)	-88(24)
C(9)A	34(2)	131(7)	142(8)	35(7)	-8(7)	-17(13)
C(10)A	41(2)	82(4)	102(6)	-23(6)	-24(6)	-16(9)
C(11)A	36(2)	57(4)	178(9)	-6(5)	-33(7)	-5(10)
C(12)A	54(3)	85(5)	189(10)	5(7)	18(9)	-25(13)
C(13)A	69(4)	90(6)	251(14)	39(8)	11(12)	-64(16)
C(14)A	101(5)	127(8)	183(13)	52(12)	34(13)	6(17)
C(15)A	37(2)	68(4)	183(9)	9(6)	0(8)	50(11)
C(16)A	61(4)	177(11)	256(15)	82(11)	54(12)	243(22)
C(17)A	80(4)	127(8)	215(13)	96(10)	18(12)	102(17)
O(1)B	40(1)	90(3)	123(5)	-17(4)	-25(4)	21(7)
O(2)B	34(1)	83(3)	227(7)	-35(4)	-56(5)	15(8)

Table 21. (continued)

O(3)B	40(2)	75(3)	184(6)	9(4)	-58(5)	-32(7)
O(4)B	42(1)	69(3)	182(6)	-5(4)	-29(5)	17(7)
O(5)B	51(2)	145(5)	192(7)	-37(6)	-76(6)	-85(11)
O(6)B	64(2)	180(6)	174(7)	-55(7)	-30(7)	-111(11)
N(1)B	44(2)	79(3)	110(5)	-24(4)	-31(5)	-13(7)
N(2)B	30(1)	55(3)	143(6)	-11(4)	-37(5)	-25(7)
N(3)B	48(2)	65(3)	151(7)	-19(5)	-63(6)	10(8)
C(1)B	95(5)	154(9)	157(11)	-29(12)	-26(12)	89(17)
C(2)B	92(6)	177(12)	300(19)	50(14)	118(18)	-19(26)
C(3)B	151(8)	111(8)	213(14)	-115(14)	-16(17)	57(18)
C(4)B	46(3)	84(5)	163(9)	-20(6)	15(8)	22(12)
C(5)B	36(2)	59(4)	146(8)	-15(5)	-37(7)	0(9)
C(6)B	39(2)	91(5)	137(8)	-25(6)	5(7)	-69(11)
C(7)B	75(4)	164(9)	183(11)	-74(11)	83(12)	-147(18)
C(8)B	143(8)	255(17)	142(11)	-231(31)	-14(15)	-10(22)
C(9)B	85(4)	96(6)	138(8)	-50(9)	-51(10)	-11(13)
C(10)B	39(2)	81(4)	113(7)	-28(6)	4(6)	-38(10)
C(11)B	40(2)	57(4)	132(7)	-2(5)	-46(7)	-34(9)
C(12)B	62(3)	82(5)	125(8)	-10(7)	-10(8)	-40(11)
C(13)B	58(3)	118(7)	184(11)	-29(9)	19(10)	-83(15)
C(14)B	70(4)	120(7)	151(10)	17(9)	5(10)	11(14)
C(15)B	47(2)	55(3)	113(7)	-14(5)	-50(7)	-15(9)
C(16)B	49(3)	92(5)	152(8)	-45(7)	-54(8)	-1(12)
C(17)B	48(3)	84(5)	179(10)	-28(7)	-33(8)	-32(12)
O(W)	97(4)	158(7)	343(13)	-38(9)	-117(11)	16(17)

Table 22. Hydrogen positional parameters and their e.s.d.'s in parentheses ($\times 10^3$)

	x	y	z	bonded to
H(1)	168(4)	520(5)	210(7)	N(2)A
H(2)	673(4)	45(5)	615(7)	N(2)B
H(3)	161(4)	227(6)	357(7)	N(3)A
H(4)	749(4)	356(6)	613(7)	N(3)B
H(5)	77(4)	755(5)	-398(7)	C(1)A
H(6)	84(4)	779(5)	-216(7)	C(1)A
H(7)	-3(4)	746(5)	-303(7)	C(1)A
H(8)	439(4)	-105(5)	947(7)	C(1)B
H(9)	489(4)	-21(5)	874(7)	C(1)B
H(10)	505(4)	-152(6)	830(7)	C(1)B
H(11)	217(4)	643(5)	-331(7)	C(2)A
H(12)	197(4)	553(5)	-229(7)	C(2)A
H(13)	202(4)	654(5)	-165(7)	C(2)A
H(14)	298(4)	-22(6)	873(7)	C(2)B
H(15)	278(4)	9(6)	712(7)	C(2)B
H(16)	359(4)	63(6)	800(7)	C(2)B
H(17)	88(4)	588(6)	-464(7)	C(3)A
H(18)	7(4)	545(5)	-381(7)	C(3)A
H(19)	79(4)	501(6)	-353(7)	C(3)A
H(20)	330(4)	-209(6)	810(7)	C(3)B
H(21)	396(4)	-223(5)	691(7)	C(3)B
H(22)	305(4)	-173(6)	652(7)	C(3)B
H(23)	78(4)	616(5)	111(7)	C(6)A
H(24)	585(4)	-49(6)	509(7)	C(6)B
H(25)	-23(4)	652(6)	287(7)	C(7)A
H(26)	12(4)	532(5)	336(7)	C(7)A
H(27)	619(4)	-53(6)	281(7)	C(7)B
H(28)	644(4)	63(6)	302(7)	C(7)B
H(29)	-148(4)	556(6)	296(7)	C(8)A

Table 22. (continued)

H(30)	-103(4)	462(6)	236(7)	C(8)A
H(31)	526(4)	33(5)	116(7)	C(8)B
H(32)	522(4)	126(5)	215(7)	C(8)B
H(33)	-150(4)	635(5)	64(7)	C(9)A
H(34)	-155(4)	512(5)	22(7)	C(9)A
H(35)	432(4)	-60(6)	233(7)	C(9)B
H(36)	389(4)	55(5)	275(7)	C(9)B
H(37)	155(4)	298(6)	167(7)	C(11)A
H(38)	630(4)	256(5)	620(7)	C(11)B
H(39)	318(4)	408(6)	129(7)	C(12)A
H(40)	663(4)	288(6)	856(7)	C(12)B
H(41)	358(4)	244(5)	34(7)	C(13)A
H(42)	268(4)	200(6)	97(7)	C(13)A
H(43)	336(4)	229(5)	209(7)	C(13)A
H(44)	519(4)	227(5)	933(6)	C(13)B
H(45)	512(4)	153(6)	797(7)	C(13)B
H(46)	518(4)	272(5)	767(7)	C(13)B
H(47)	297(4)	357(5)	-114(7)	C(14)A
H(48)	221(4)	441(6)	-62(7)	C(14)A
H(49)	198(4)	315(6)	-81(7)	C(14)A
H(50)	651(4)	163(6)	1017(7)	C(14)B
H(51)	738(4)	149(5)	911(7)	C(14)B
H(52)	648(4)	67(6)	880(7)	C(14)B
H(53)	194(4)	217(6)	604(7)	C(16)A
H(54)	221(4)	340(6)	601(7)	C(16)A
H(55)	887(4)	395(5)	599(7)	C(16)B
H(56)	912(4)	260(5)	611(7)	C(16)B
H(57)	414(4)	202(6)	481(7)	O(5)A
H(58)	1001(4)	359(6)	906(7)	O(5)B
H(59)	235(4)	20(5)	440(7)	O(w)
H(60)	140(4)	83(6)	328(7)	O(w)

Table 23. Hydrogen bonds

Donor	Acceptor	Distance(Å)
O(5)A	O(3)B	2.619
O(5)B	O(3)A^{i}	2.582
N(2)A	O(4)B^{ii}	2.949
N(2)B	O(4)A^{iii}	2.922
N(3)A	O(w)	2.911
N(3)B	O(2)B^{ii}	2.942
O(w)	O(2)B	2.920
O(w)	O(2)A^{iv}	2.760

symmetry code

i 1.0+x,y,1.0+z

ii 1.0-x,0.5+y,1.0-z

iii 1.0-x,-0.5+y,1.0-z

iv -x,-0.5+y,-z

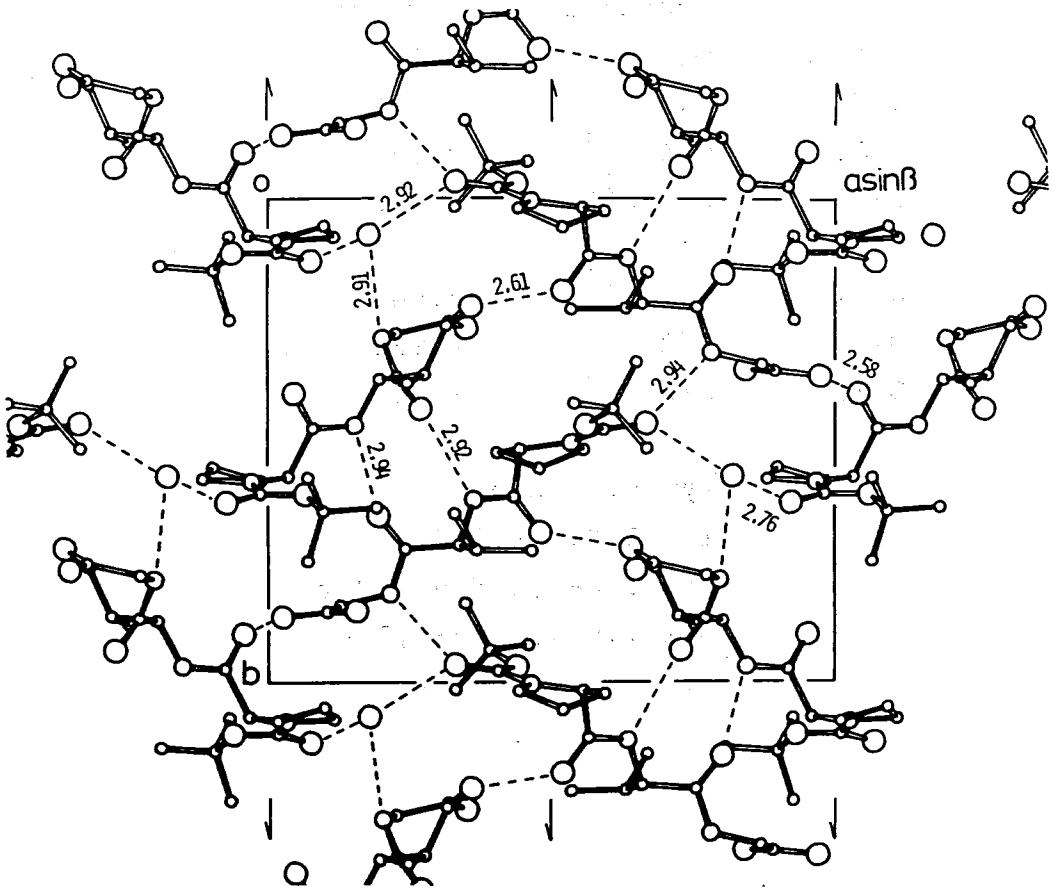
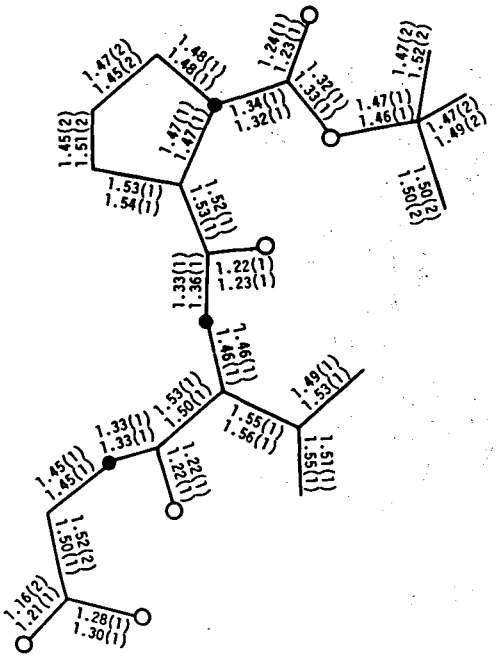
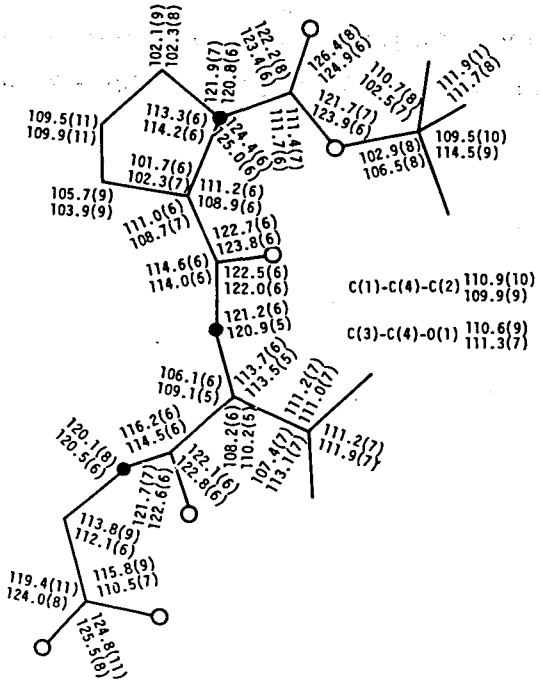


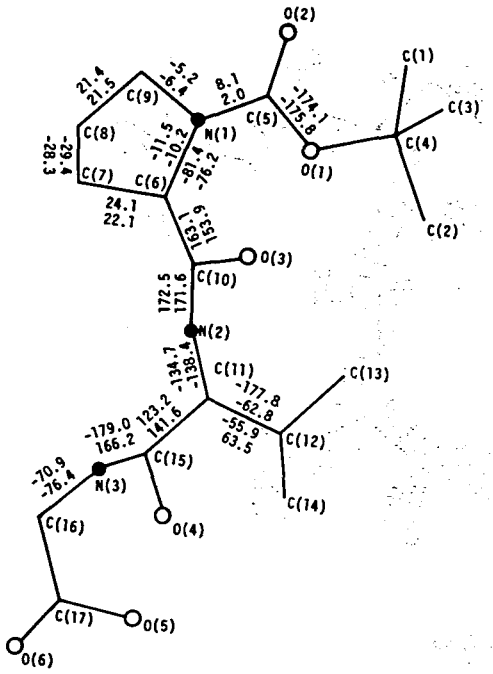
Fig. 11. The crystal structure of Boc-Pro-Val-Gly-OH viewed along the c axis. The hydrogen bonds are shown by the broken lines.


1.47(2) 1.45(2)
1.49(1) 1.48(1)
1.24(1) 1.23(1)
1.32(1) 1.33(1)
1.47(2) 1.52(2)
1.45(2) 1.51(2)
1.47(1) 1.47(1)
1.34(1) 1.32(1)
1.47(1) 1.46(1)
1.47(2) 1.49(2)
1.53(1) 1.54(1)
1.52(1) 1.51(1)
1.50(2)
1.33(1) 1.36(1)
1.22(1) 1.23(1)
1.46(1)
1.49(1) 1.53(1)
1.53(1) 1.50(1)
1.55(1) 1.56(1)
1.33(1) 1.33(1)
1.45(1) 1.45(1)
1.22(1)
1.55(1)
1.51(1)
1.52(2) 1.50(1)
1.16(2) 1.21(1)
1.28(1) 1.30(1)

(a)


102.1(9) 102.3(8)
122.2(8) 123.4(6)
126.4(8) 124.9(6)
121.9(7) 120.8(6)
110.7(8) 102.5(7)
111.9(1) 111.7(8)
109.5(11) 109.9(11)
113.3(6) 114.2(6)
121.7(7) 123.9(6)
111.4(7) 111.7(6)
124.4(6) 125.0(6)
109.5(10) 114.5(9)
102.9(8) 106.5(8)
101.7(6) 102.3(7)
105.7(9) 103.9(9)
111.2(6) 108.9(6)
111.0(6) 108.7(7)
122.7(6) 123.8(6)
114.6(6) 114.0(5)
122.5(6) 122.0(6)
C(1)-C(4)-C(2) 110.9(10) 109.9(9)
121.2(6) 120.9(5)
C(3)-C(4)-O(1) 110.6(9) 111.3(7)
106.1(6) 109.1(5)
113.7(6) 113.5(5)
111.2(7) 111.0(7)
116.2(6) 114.5(6)
120.1(8) 120.5(6)
108.2(6) 110.2(5)
107.4(7) 113.1(7)
111.2(7) 111.9(7)
121.7(7) 122.6(6)
122.1(6) 122.8(6)
113.8(9) 112.1(6)
119.4(11) 124.0(8)
115.8(9) 110.5(7)
124.8(11) 125.5(8)

(b)



(c)

Fig. 12. (a) Bond lengths (Å), (b) angles (°) and (c) torsion angles (°) of Boc-Pro-Val-Gly-OH.

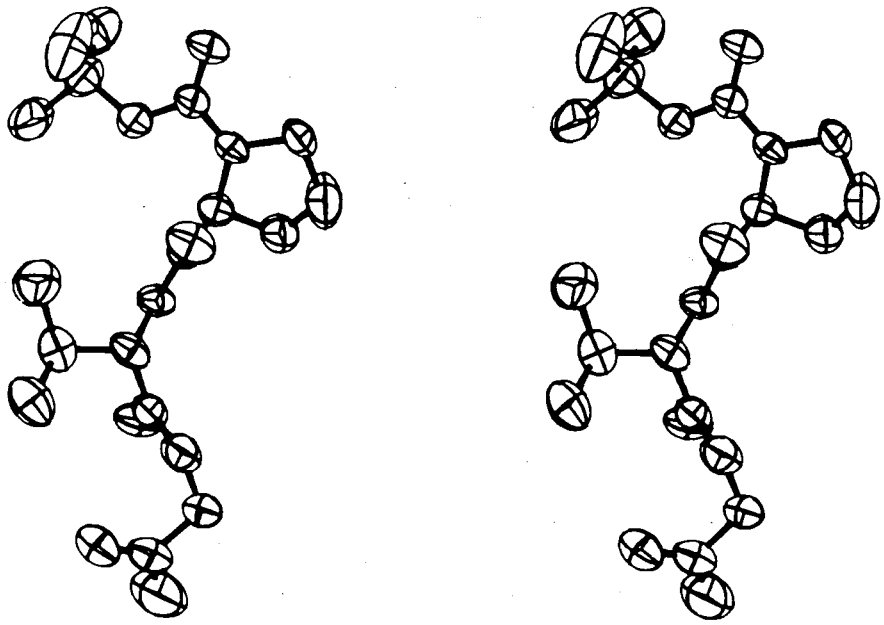


Fig. 13. Stereo drawing of the Boc-Pro-Val-Gly-OH (A) molecule with thermal ellipsoids drawn to enclose 50% probability.

5. Boc-Pro-Ala-Gly-OH and Aoc-Pro-Ala-Gly-OH

5-1. Introduction

The last peptide to be analyzed was the one which contain alanyl residue at the third site. The alanyl residue having only a methyl group as a side chain seems to cause the least steric effect on the main chain folding. It could be expected, therefore, to get some clue whether the branched structure at the C^{β} atom has some definitive effect on the main chain folding or not.

But the analysis was not made straightforwardly, since Boc-Pro-Ala-Gly-OH crystal seemed to contain four molecules in an asymmetric unit, and what was worse, the X-ray photograph showed a twin structure. So the analysis was given up as impossible, and in place of this peptide Aoc-Pro-Ala-Gly-OH was synthesized. This crystal contained two molecules in an asymmetric unit, but showed no twin structure.

In this section, the crystal structure analysis of Aoc-Pro-Ala-Gly-OH will be described.

5-2. Experimental

The processes taken for the synthesis of the two valyl peptides in section 4, were again adopted for the present peptides. The peptides intermediately

obtained as crystals were Boc-Ala-Gly-OBzl (mp=86.0-87.0°C) and Aoc-Pro-Ala-Gly-OBzl (mp=93.0-94.0°C). Boc-Pro-Ala-Gly-OBzl was obtained only as oily form. The aimed peptides were crystallized from an ethylacetate solution. The crystal data are given in Tables 24 and 25.

Intensity measurement was done only for Aoc-Pro-Ala-Gly-OH using a Hilger & Watts four-circle diffractometer equipped with Ni-filtered Cu K_α radiation. 2169 reflections with $2\theta<100°$ were collected, of which 2025 were nonzero. The ω-2θ step scanning mode was adopted. The crystal size was 0.14 × 0.07 × 0.06 mm.

5-3. Structure determination

The program MULTAN[13] was first applied but did not give a reasonable result. The program RICS was next applied using the same rigid groups as those used for the analysis of Boc-Pro-Val-Gly-OH. The extended conformation seemed to be probable but the partial structure tangent refinement based on the 20 atoms obtained could not converge properly. It was further assumed that two independent molecules exist as a dimer in such a way as observed in Boc-Pro-Ile-Gly-OH and Boc-Pro-Val-Gly-OH. This

assumption was not daring in view of the similar packing pattern of the two previously analyzed peptides. The phases calculated from 40 atoms thus obtained easily lead to the whole structure.

The structure was refined by the block-diagonal least squares program HBLS V. The function minimized was $\Sigma\omega(|F_o|-|F_c|)^2$ where weights ω were unity for all the non-zero reflections and zero for zero reflections. All the atomic scattering factors were taken from International Tables for X-ray Crystallography.[15] The final positional and thermal parameters are given in Tables 26, 27 and 28. The isotropic temperature factor B=4.398 was assigned to all the hydrogen atoms.

5-4. Discussion

The crystal structure of Aoc-Pro-Ala-Gly-OH is also characterized by the dimer formation with β-sheet type hydrogen bonds as in Fig. 14. The extended ocnformation of one molecule (molecule (A)) is given in Fig. 15. Bond lengths, bond angles and torsion angles are shown in Fig. 16. Some abnormal bond lengths in the amyloxy group are probably due to its extremely large thermal motion.

After all, in the Boc-Pro-X-Gly-OH peptides, those

which have Ala, Val and Ile at X positions were found to crystallize with similar packing pattern. The crystal lattice transformation of the alanyl peptide by a'=a+c, b'=b and c'=c clearly shows close relationship between the crystal packings of alanyl peptide and valyl peptide. Although the space group of isoleucyl peptide, $P2_12_12$, differs from those of valyl and alanyl peptides, $P2_1$, the former crystal packing is still close to the latters. 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 are obvious in Fig. 17, in which dotted lines in the alanyl and valyl peptide crystals are drawn in order to coincide with the unit cell of the isoleucyl peptide.

The similarity of the molecular packing is again obvious in Table 30 which shows hydrogen bonding system. No difference in hydrogen bonding pattern is seen for the alanyl peptides (A) and (B), and patterns differ only in the point in which the C-terminal of the molecule interacts on water in the case of valyl peptides (A) and (B). Thus, hydrogen bonding pattern of three different peptides differs only in the interaction to water molecules, with β-sheet type hydrogen bonds being completely reserved.

The lack of the strict twofold axis in the alanyl and valyl dimers, should, therefore, attribute to inter-dimer interactions at the C-terminal.

The hydrogen bonding patterns of the Boc-Pro-Leu-Gly-OH and Boc-Pro-Pro-Gly-NH_2 peptides are also listed in Table 31, which shows different nature of these crystals from others. In the case of the prolyl peptide, the difference is quite natural in view of the different molecular nature, i.e., proline as imino acid and C-terminal amide group. But in the leucyl peptide no appropriate reason could be found for such a different packing. It might be possible that β-turn of the leucyl peptide is one of the two stable crystal state. But at least no second state cannot be observed either for the leucyl or other peptides for the time being. With these evidences, together with other leucyl peptides which fold into β-turn, we should conclude that there are stronger tendency for the leucyl peptide to make β-turn with being itself at the third position.

Table 24. Crystal data
(Boc-Pro-Ala-Gly-OH·H_2O)

Molecular formula	$C_{15}H_{25}O_6N_3 \cdot H_2O$
Molecular weight	343.38
Space group	$P2_1$
Cell constants	
	a=16.39Å
	b=19.66
	c=12.79
	β=96°
Density	
obs.	1.18g/cm^3
calc.	1.17 (for Z=8)
mp	144.5-146.0°C

Table 25. Crystal data
(Aoc-Pro-Ala-Gly-OH)

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)
	c= 9.553(2)
	β=118.46(2)°
Density	
obs.	1.21g/cm^3
calc.	1.20 (for Z=4)
mp	145.0-146.0°C

Table 26. The atomic positional parameters and their e.s.d.'s in parentheses ($\times 10^4$)

	x	y	z
O(1)A	4436(3)	6399(6)	10922(7)
O(2)A	5758(3)	6571(6)	12799(7)
O(3)A	4713(3)	4313(6)	9165(7)
O(4)A	2378(4)	4363(5)	3984(8)
O(5)A	1179(4)	2280(6)	937(8)
O(6)A	1294(4)	2200(8)	3367(9)
N(1)A	5387(4)	6328(6)	10199(8)
N(2)A	3618(4)	4989(6)	7085(8)
N(3)A	2927(4)	2756(6)	4661(9)
C(1)A	3912(10)	4800(20)	12500(25)
C(2)A	4253(9)	5850(20)	13142(17)
C(3)A	4355(14)	7704(20)	12720(32)
C(4)A	3185(8)	6767(14)	10709(19)
C(5)A	4072(7)	6660(12)	11932(14)
C(6)A	5230(5)	6428(8)	11413(11)
C(7)A	4760(5)	6159(8)	8534(10)
C(8)A	5223(7)	6277(11)	7635(12)
C(9)A	6097(7)	6150(22)	8851(15)
C(10)A	6213(5)	6381(9)	10395(12)
C(11)A	4382(5)	5080(9)	8322(10)
C(12)A	3168(5)	3990(8)	6745(11)
C(13)A	2524(7)	4039(12)	7237(15)
C(14)A	2804(5)	3750(7)	5016(10)
C(15)A	2492(6)	2370(9)	3041(12)
C(16)A	1585(5)	2298(7)	2494(12)
O(1)B	10520(3)	-138(6)	8990(7)
O(2)B	11433(3)	-333(6)	11617(7)
O(3)B	9628(3)	1919(6)	9377(8)
O(4)B	6979(4)	1899(6)	4716(8)

Table 26. (continued)

O(5)B	5514(4)	3982(6)	2162(8)
O(6)B	6779(4)	3979(7)	2461(8)
N(1)B	10103(4)	-61(6)	10788(7)
N(2)B	8538(4)	1258(6)	7260(8)
N(3)B	7371(4)	3540(7)	5730(8)
C(1)B	11142(32)	1392(34)	7461(70)
C(2)B	11696(12)	380(22)	8840(28)
C(3)B	11381(12)	-1516(15)	8874(21)
C(4)B	10459(13)	-496(18)	6532(16)
C(5)B	11052(7)	-416(10)	8314(14)
C(6)B	10727(5)	-204(8)	10549(10)
C(7)B	9270(5)	101(8)	9515(11)
C(8)B	8776(6)	64(10)	10398(14)
C(9)B	9345(8)	154(23)	12046(16)
C(10)B	10166(6)	-91(10)	12332(12)
C(11)B	9163(5)	1142(8)	8730(11)
C(12)B	8406(5)	2259(8)	6376(12)
C(13)B	8683(7)	2128(13)	5161(15)
C(14)B	7510(6)	2518(8)	5560(11)
C(15)B	6554(6)	3955(9)	4755(11)
C(16)B	6322(6)	3968(9)	2994(11)

Table 27. Hydrogen positional parameters and 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

Table 27. (continued)

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

Table 28. The thermal parameters ($\times 10^4$) and their e.s.d.'s in parentheses. Thermal parameters are in the form: $\exp\{-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)\}$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
O(1)A	31(2)	101(6)	170(11)	-3(7)	41(9)	-42(15)
O(2)A	41(2)	112(7)	122(10)	-51(7)	19(9)	-64(14)
O(3)A	43(3)	84(6)	139(10)	0(7)	-22(9)	54(14)
O(4)A	57(3)	50(5)	176(12)	19(7)	-34(10)	-20(13)
O(5)A	44(3)	92(6)	172(12)	-7(7)	-17(10)	-1(15)
O(6)A	48(3)	174(10)	230(15)	-41(10)	55(12)	-77(22)
N(1)A	28(2)	83(7)	109(11)	-19(7)	17(10)	-38(15)
N(2)A	33(3)	62(6)	141(13)	-2(7)	0(10)	-8(15)
N(3)A	34(3)	64(6)	150(13)	14(8)	-8(11)	-28(16)
C(1)A	102(10)	262(29)	595(56)	82(30)	326(43)	333(69)
C(2)A	67(7)	329(32)	265(30)	-82(26)	110(25)	-11(52)
C(3)A	164(16)	219(26)	865(80)	-123(35)	552(65)	-488(79)
C(4)A	59(6)	186(19)	450(38)	47(19)	211(27)	107(46)
C(5)A	58(5)	149(14)	247(23)	-12(15)	146(20)	-90(30)
C(6)A	40(4)	69(8)	143(15)	-22(10)	43(13)	-72(20)
C(7)A	36(4)	66(7)	104(14)	-3(9)	22(12)	15(18)
C(8)A	76(6)	121(12)	141(17)	-54(15)	95(18)	1(26)
C(9)A	44(5)	474(41)	198(24)	-30(27)	91(20)	-104(57)
C(10)A	32(4)	99(10)	214(20)	-14(11)	61(15)	-44(25)
C(11)A	38(4)	90(9)	87(13)	19(10)	9(12)	25(19)
C(12)A	39(4)	54(7)	163(17)	-20(10)	-19(14)	-24(14)
C(13)A	58(5)	137(14)	257(24)	-70(16)	108(10)	-69(31)
C(14)A	28(3)	51(7)	125(15)	-4(8)	-33(12)	-26(17)
C(15)A	43(4)	79(9)	186(19)	-2(11)	36(15)	-37(22)
C(16)A	39(4)	45(7)	194(18)	-6(10)	-10(14)	-13(20)
O(1)B	47(3)	99(6)	125(10)	45(7)	71(9)	22(13)
O(2)B	29(2)	125(7)	174(12)	49(7)	14(9)	88(16)
O(3)B	40(2)	70(5)	216(13)	2(7)	-32(10)	12(15)

Table 28. (continued)

O(4)B	46(3)	63(5)	208(13)	-31(7)	-42(10)	-33(15)
O(5)B	46(3)	104(7)	173(12)	6(8)	14(10)	45(15)
O(6)B	45(3)	153(9)	206(13)	29(9)	53(11)	79(19)
N(1)B	34(3)	86(7)	95(10)	26(8)	38(10)	17(15)
N(2)B	36(3)	54(6)	120(12)	-32(7)	-39(10)	23(14)
N(3)B	40(3)	79(7)	133(13)	2(8)	-29(11)	-17(17)
C(1)B	470(58)	329(56)	2177(260)	-16(93)	1750(231)	308(198)
C(2)B	118(12)	340(37)	729(69)	-155(36)	482(53)	-393(88)
C(3)B	156(14)	161(19)	410(40)	152(28)	314(41)	85(48)
C(4)B	213(16)	304(29)	218(26)	381(39)	337(37)	276(41)
C(5)B	60(5)	101(11)	238(22)	2(13)	124(19)	-59(27)
C(6)B	48(4)	77(8)	110(14)	31(11)	53(13)	32(18)
C(7)B	32(3)	60(7)	162(17)	16(9)	15(13)	39(20)
C(8)B	37(4)	108(11)	296(25)	22(12)	110(18)	87(29)
C(9)B	70(7)	449(40)	228(27)	125(31)	158(25)	133(58)
C(10)B	46(4)	111(11)	146(17)	24(12)	70(15)	50(23)
C(11)B	28(3)	78(8)	156(17)	5(10)	0(13)	-54(21)
C(12)B	34(4)	63(8)	198(19)	-27(9)	21(14)	56(21)
C(13)B	61(6)	151(15)	297(27)	1(17)	126(22)	140(37)
C(14)B	50(4)	60(7)	129(16)	2(10)	24(14)	9(19)
C(15)B	49(4)	84(9)	114(15)	10(11)	-16(14)	28(20)
C(16)B	44(4)	81(9)	145(17)	17(11)	-6(14)	46(21)

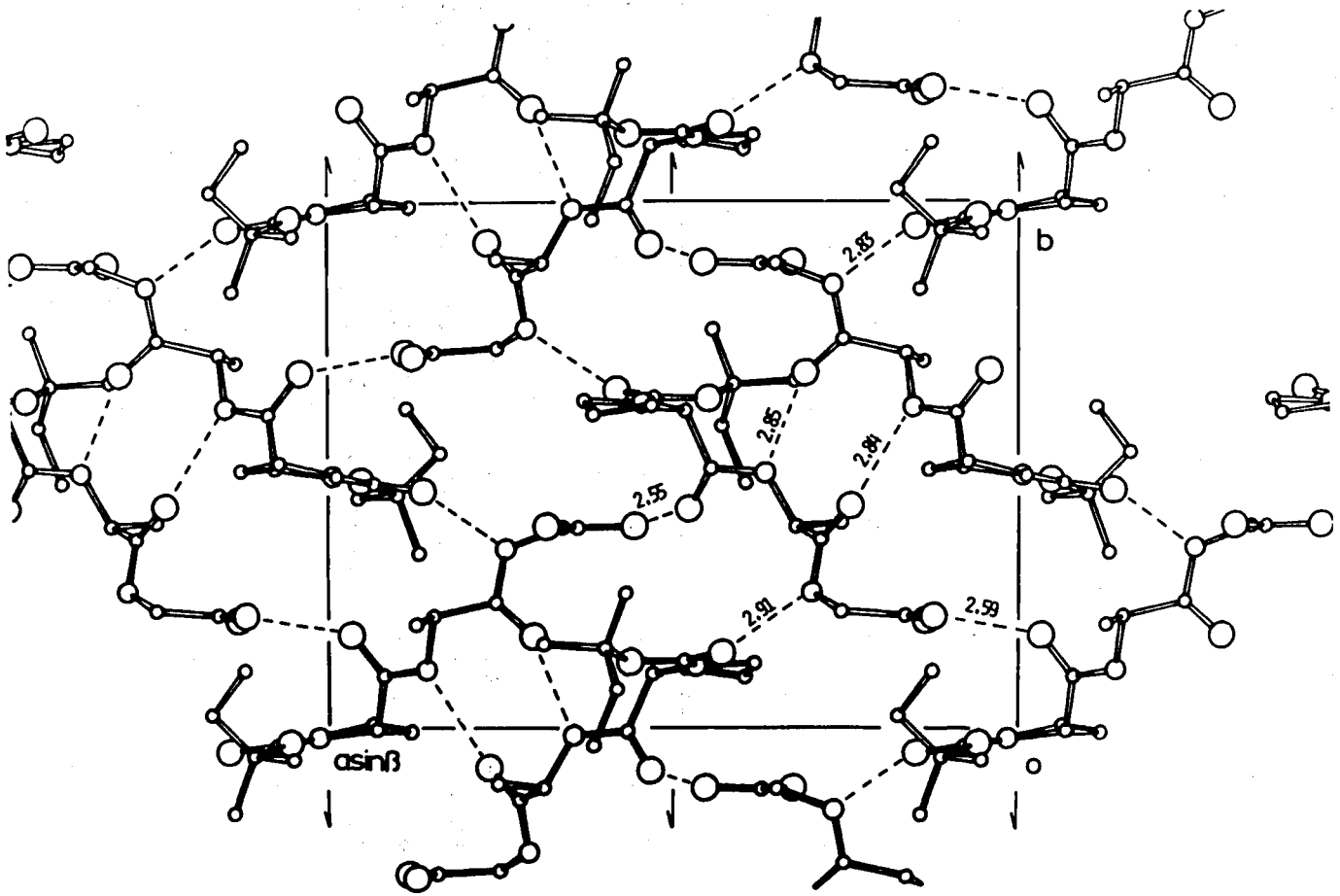


Fig. 14. The crystal structure of Aoc-Pro-Ala-Gly-OH viewed along the c axis. The hydrogen bonds are shown by the broken lines.

Table 29. Hydrogen bonds

Donor	Acceptor	Distance(Å)
O(5)A	$O(3)B^{i}$	2.598
O(5)B	$O(3)A^{ii}$	2.556
N(2)A	$O(4)B^{iii}$	2.859
N(2)B	$O(4)A^{iv}$	2.849
N(3)A	$O(2)A^{v}$	2.913
N(3)B	$O(2)B^{vi}$	2.853

symmetry code

i	-1+x,y,-1+z
ii	x,y,-1+z
iii	1-x,0.5+y,1-z
iv	1-x,-0.5+y,1-z
v	1-x,-0.5+y,2-z
vi	2-x,0.5+y,2-z

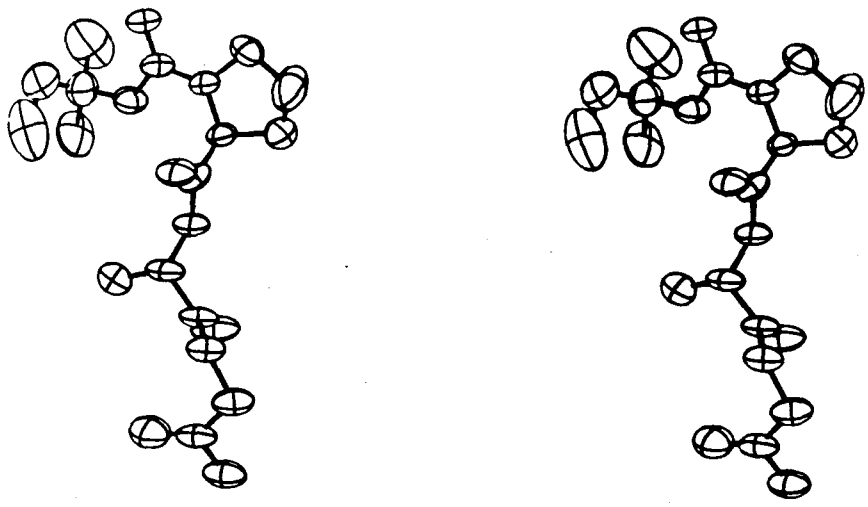


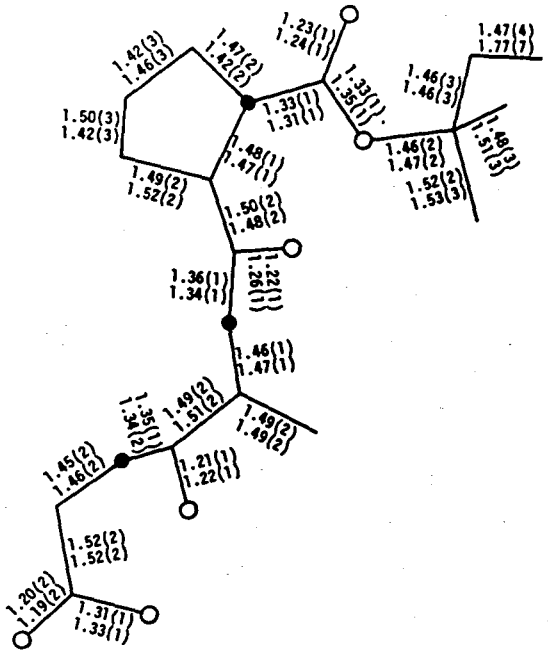
Fig. 15. Stereo drawing of the Aoc-Pro-Ala-Gly-OH with thermal ellipsoids drawn to enclose 50% probability.

Table 30. The hydrogen bond networks in Boc-Pro-X-Gly-OH systems (1)

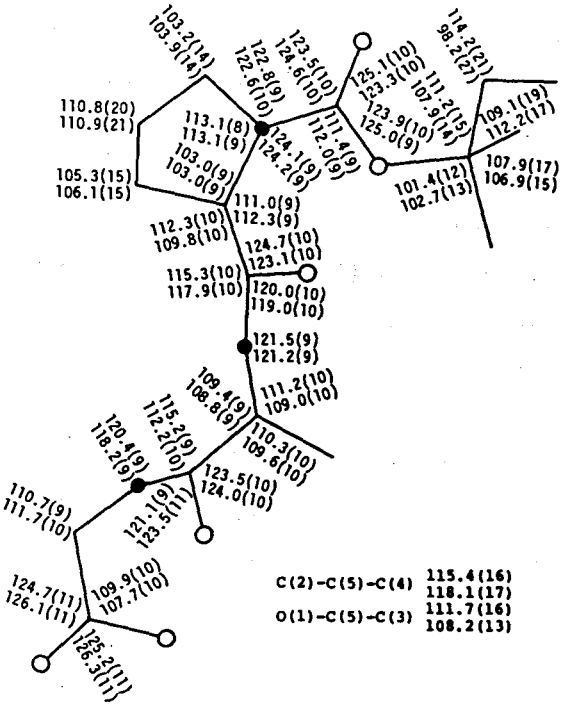
Donor	Acceptor	Distance(Å)
(Boc-Pro-Leu-Gly-OH)		
Gly COOH	W	2.598
Leu NH	Gly C=O	2.843
Gly NH	Boc C=O	2.945
W	Leu C=O	2.855
W	Pro C=O	2.758
(Boc-Pro-Pro-Gly-NH_2)		
NH_2	Pro(I) C=O	3.073
Gly NH	Gly C=O	2.896

Table 31. The hydrogen bond networks in Boc-Pro-X-Gly-OH systems (2)

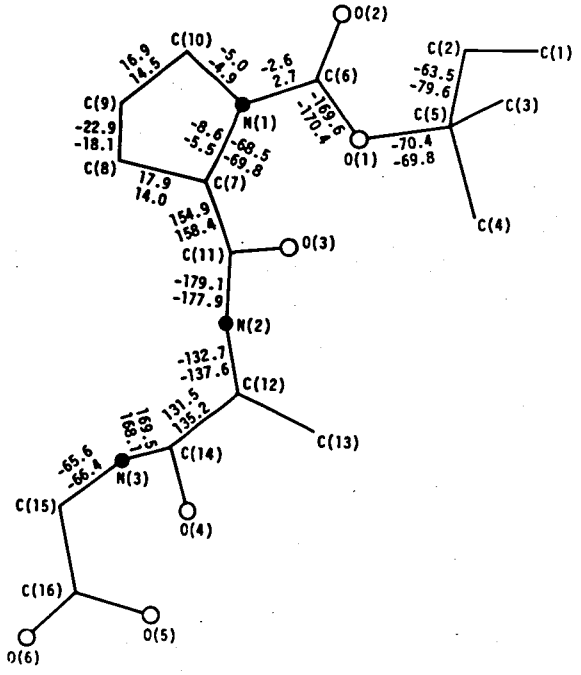
Donor	Acceptor	Distance(Å)
(Boc-Pro-Ile-Gly-OH)		
Gly COOH	Pro C=O	2.602
Ile NH	Ile C=O	2.832
Gly NH	Boc C=O	2.870
W	Boc C=O	2.808
(Boc-Pro-Val-Gly-OH)		
Gly(A) COOH	Pro(B) C=O	2.619
Val(A) NH	Val(B) C=O	2.949
Gly(A) NH	W	2.911
W	Boc(B) C=o	2.920
Gly(B) COOH	Pro(A) C=O	2.582
Val(B) NH	Val(A) C=O	2.922
Gly(B) NH	Boc(B) C=O	2.942
W	Boc(A) C=O	2.760
(Boc-Pro-Ala-Gly-OH)		
Gly(A) COOH	Pro(B) C=O	2.598
Ala(A) NH	Ala(B) C=O	2.859
Gly(A) NH	Aoc(A) C=O	2.913
Gly(B) COOH	Pro(A) C=O	2.556
Ala(B) NH	Ala(A) C=O	2.849
Gly(B) NH	Aoc(B) C=O	2.835


1.23(1)
1.24(1)
1.47(4)
1.77(7)
1.42(3)
1.46(3)
1.47(2)
1.42(2)
1.33(1)
1.35(1)
1.46(3)
1.46(3)
1.50(3)
1.42(3)
1.33(1)
1.31(1)
1.48(1)
1.47(1)
1.46(2)
1.47(2)
1.48(3)
1.51(3)
1.49(2)
1.52(2)
1.52(2)
1.53(3)
1.50(2)
1.48(2)
1.22(1)
1.21(1)
1.36(1)
1.34(1)
1.46(1)
1.47(1)
1.49(2)
1.51(2)
1.49(2)
1.49(2)
1.35(1)
1.34(1)
1.45(2)
1.46(2)
1.21(1)
1.22(1)
1.52(2)
1.52(2)
1.20(2)
1.19(2)
1.31(1)
1.33(1)

(a)


103.2(14)
103.9(14)
122.8(9)
122.6(10)
123.5(10)
124.6(10)
125.1(10)
123.3(10)
114.2(21)
98.2(21)
110.8(20)
110.9(21)
113.1(8)
113.1(9)
124.1(9)
124.2(9)
111.4(9)
112.0(9)
123.9(10)
125.0(9)
111.2(15)
107.9(14)
109.1(19)
112.2(17)
105.3(15)
106.1(15)
103.0(9)
103.0(9)
111.0(9)
112.3(9)
101.4(12)
102.7(13)
107.9(17)
106.9(15)
112.3(10)
109.8(10)
124.7(10)
123.1(10)
115.3(10)
117.9(10)
120.0(10)
119.0(10)
121.5(9)
121.2(9)
109.4(9)
108.8(9)
111.2(10)
109.0(10)
115.2(9)
112.2(10)
120.4(9)
118.2(9)
110.3(10)
109.6(10)
123.5(10)
124.0(10)
110.7(9)
111.7(10)
121.1(9)
123.5(11)
C(2)-C(5)-C(4) 115.4(16)
118.1(17)
O(1)-C(5)-C(3) 111.7(16)
108.2(13)
124.7(11)
126.1(11)
109.9(10)
107.7(10)
125.2(11)
126.3(11)

(b)



(c)

Fig. 16. (a) Bond lengths (Å), (b) angles (°) and (c) torsion angles (°) of Aoc-Pro-Ala-Gly-OH.

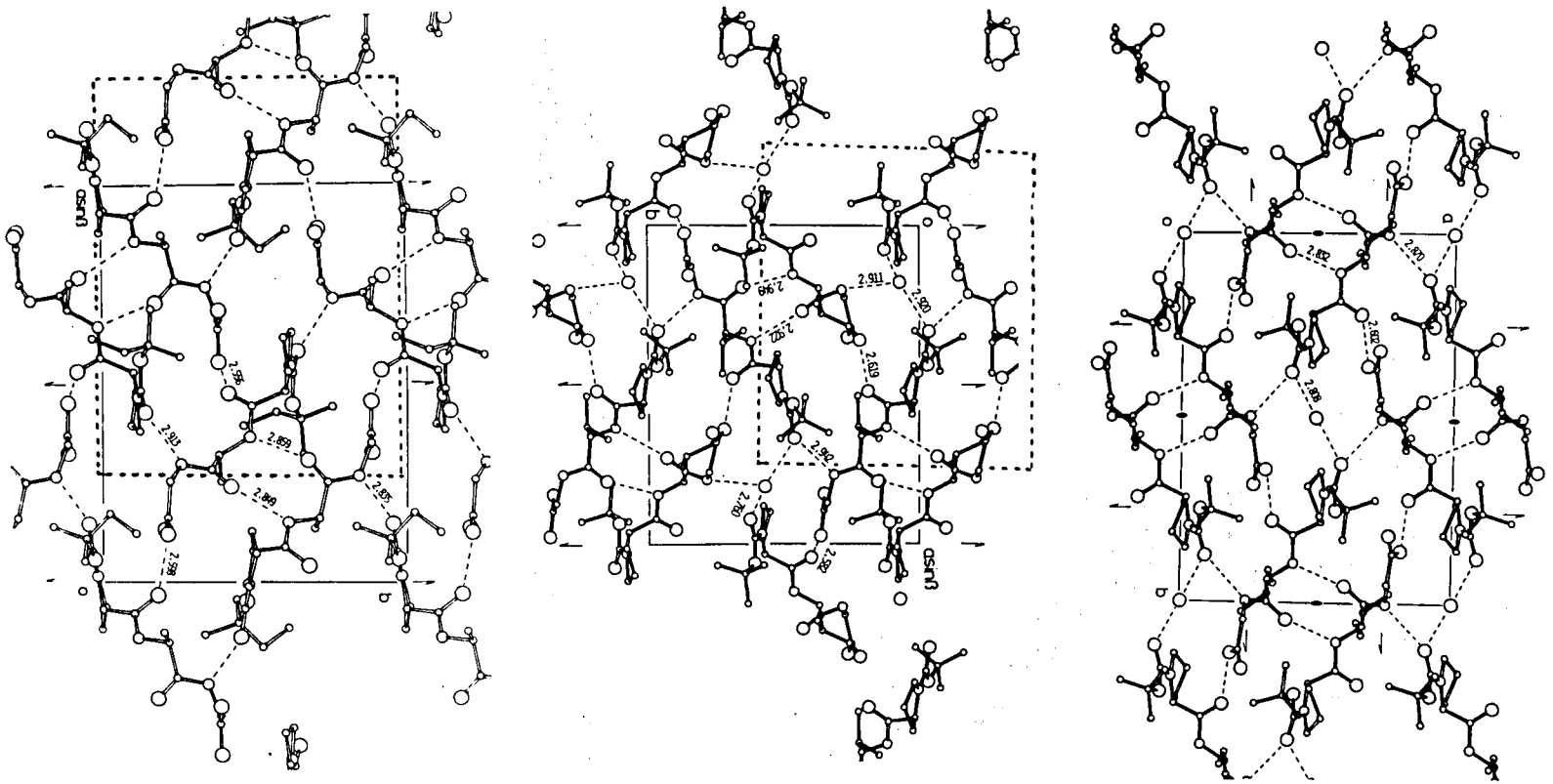


Fig. 17. The comparison of the crystal structures of Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH and Aoc-Pro-Ala-Gly-OH. Dotted lines in valyl and alanyl peptides are drawn in order to coincide with the unit cell of isoleucyl peptide.

III. CONFORMATIONAL COMPARISON OF THE PEPTIDE SECONDARY STRUCTURE

1. Introduction

It is well-known that the tertiary structures of the proteins, especially of the globular proteins, are stabilized by the well-developed intramolecular interactions, and preservation of the proper structures is considered to be requisite for their efficient biological activities. On the other hand, the peptide structures, whose intramolecular interactions are fewer in comparison with proteins, have been naturally considered to be more flexible than the proteins. In this chapter such a peptide conformation is examined on the basis of the present crystal structure analyses.

One of the interesting results obtained in these studies was that all of the present oligopeptides take either a β-turn structure or a β-sheet structure, which have been considered, together with the α-helix, to be three of the most important secondary structures in the proteins. Our attention will be naturally focused on these two structural units, the β-turn and the β-sheet.

2. β-turn

The α-helix and the β-sheet, two most familiar secondary structures of the proteins, were first predicted by Pauling and Corey.[38][39] Both structures were then found to occur in such repeating polypeptides as α-keratins[40] and silks.[41] They were also proved to be important constituents in the globular proteins, as was found in the structures of myoglobin[42] and lysozyme.[43]

On the other hand the β-turn, the third secondary structure, has different history. The β-turn structure was first found in the cyclic peptides, cyclohexaglycyl[35] and ferrichrome A,[28] and in the linear oligopeptides, Z(p-Br)-Gly-Pro-Leu-Gly-OH[4] and Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH.[3] The theoretical study of these conformations was made later by Venkatachalam,[5] who classified the β-turn into three types I, II, and III. As can be seen in Fig. 18, type I and type II, which are related by a 180° twist in the second peptide unit, involve a peptide chain turn; while type III, which could be regarded as a subset of type I, is a part of a 3_{10} helix.

The importance of the concept turn in the globular proteins was suggested by Lewis et al.,[44] Kuntz[45] and Crawford[7] by the studies concerned

with turning position and amino acid residues. Their works commonly showed that turns occur at molecular surfaces and are abundant of hydrophilic residues. More recently, Rose et al.,[46] and Chau et al.,[47] statistically studied the correlations between turn and amino residues using a larger number of globular proteins of known three dimensional structure. Furthermore, Rose[48] presented a simple, non-empirical hypothesis to account for turn position using Nozaki and Tanford hydrophobicity scale.[49]

In contrast to the well-defined α-helix or β-sheet the third secondary structure β-turn is difficult to define clearly, probably due to the fact that the turn involves only a few residues rather than the repeating structures. Nevertheless the turns are important, since the turns constitute recognizable structural units in proteins and they are situated at the solvent-accessible surface of the molecules.[45] When the turn refers to such local structure where protein main chain changes its direction almost 180°, the β-turn is a subset which completes the turn with four consecutive residues. Instead of the term β-turn, hairpin turn, β-bend, reverse turn, U-turn or 1-4 turn are sometimes used to indicate the same structure. The 4-1 hydrogen bond is not necessarily

requisite for the β-turn.

Fig. 19 shows ORTEP drawing[17] of the Boc-Pro-Leu-Gly-OH and Boc-Pro-Pro-Gly-NH_2 peptides which take the β-turn structure. Although two molecular structures are different due to the different amino acid compositions, the structures of the central parts of the molecules which are reinforced by the 4-1 hydrogen bonds are surprisingly similar with each other. Both structures could be referred to the β-turn type I.

In contrast to the proteins only a few β-turn structures have been found so far in peptides mainly due to the limited number of analysis. The torsion angles (ϕ, ψ) of the second and third residues in the β-turns are listed in Table 29 and are plotted in the (ϕ, ψ) map in Fig. 20. It is natural that the structural flexibility of the second residue is smaller than that of the third residue because the former sites are mostly occupied by proline. Therefore, the following discussion about the structural variety of the β-turn is mainly concerned with the third residue.

The variety of the torsion angles may result from the crystal packing requirement as well as from the intrinsic nature of molecules. The three peptides Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH, Z-Gly-Pro-Leu-Gly-Pro-OH

and Z(p-Br)-Gly-Pro-Leu-Gly-OH whose (ϕ_3, ψ_3) values are compatible with each other, show probably the typical β-turn sturcture for the Pro-X system. Boc-Pro-Leu-Gly-OH and Boc-Pro-Pro-Gly-NH_2 are special cases, because the former accomodates Boc group at the first site and the latter does an NH_2 group at the fourth site. The rather different (ϕ_3, ψ_3) values for S-benzyl-Cys-Pro-Leu-Gly-NH_2 (A) and (B) are attributed to their capacity for making another 4-1 hydrogen bond between carbonyl oxygen on the prolyl residue and the C-terminal amide group. Actually, in S-benzyl-Cys-Pro-Leu-Gly-NH_2 (B), two subsequent 4-1 hydrogen bonds are formed, so the conformation may be referred to the 3_{10} helix rather than the β-turn.

Figure 20 illustrates also a correlation between (ϕ_2, ψ_2) and (ϕ_3, ψ_3) angles, which can be explained qualitatively by the Newman projections of three typical molecules in Fig. 21. Torsion angles (ϕ_2, ψ_2) and (ϕ_3, ψ_3) seem to change coordinately not to disturb the 4-1 hydrogen bonding.

A quantitative comparison of these structures was made in terms of the average distance between the atoms in one molecule and the corresponding atoms in the other after refining the rotational and translational parameters in order to minimize the

summation of the square distances between the corresponding atoms. For this purpose, program CMROLS[50] was used and the results are listed in Table 30. The structural deviation of S-benzyl-Cys-Pro-Leu-Gly-NH_2 (A) and (B) from the others is again obvious.

The distinctive feature of the β-turn in oligopeptides is that it is simpler in general, in comparison with that in proteins whose structure strongly depends upon the interaction of the other parts of the molecule. The 4-1 type hydrogen bonds exist without fail in oligopeptides. An ambiguous structure, which is diffecult to identify, has not been found. Although Chau et al. locate the isolated tetrapeptides having C_1^α-C_4^α distances below 7Å as β-turn candidate, these values are within rather narrow range, from 4.977Å (Z-Gly-Pro-Leu-Gly-Pro-OH) to 5.705Å (S-benzyl-Cys-Pro-Leu-Gly-NH_2) in the above peptides.

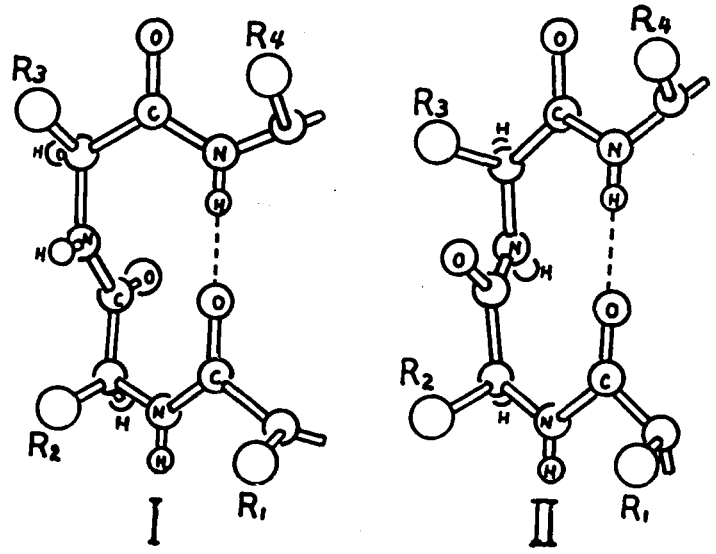


Fig. 18. The β-turn type I and II.

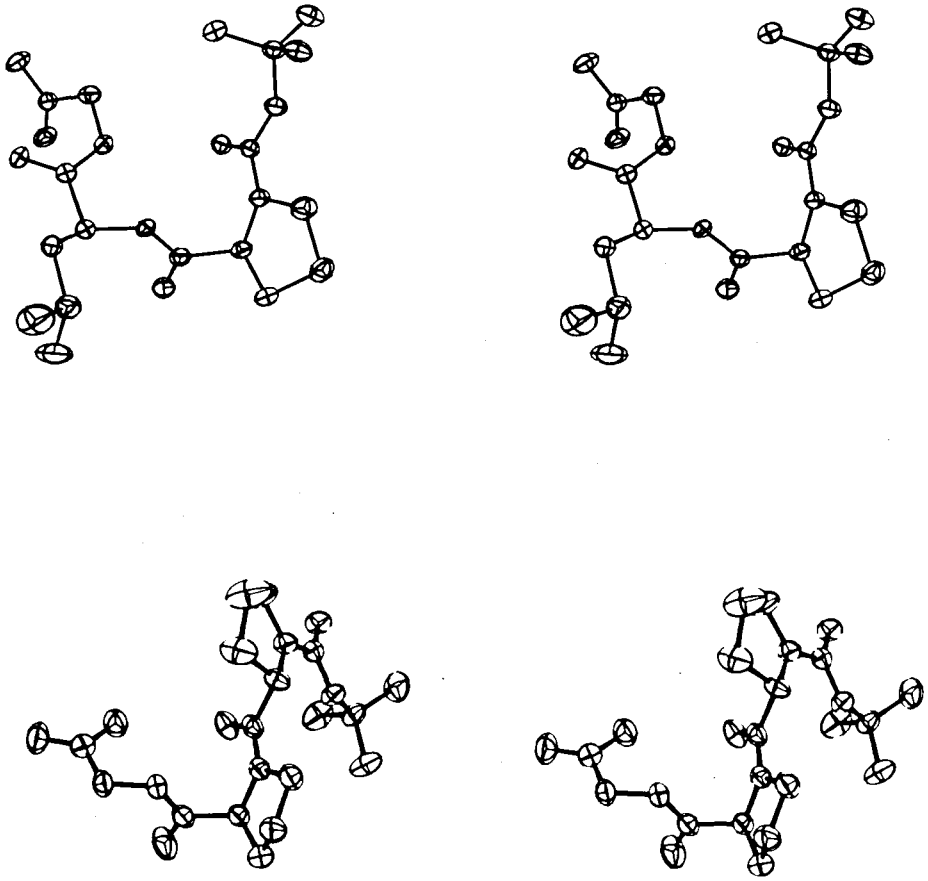


Fig. 19. Stereo drawing of the two β-turn structures. Boc-Pro-Leu-Gly-OH (upper) and Boc-Pro-Pro-Gly-NH_2 (lower).

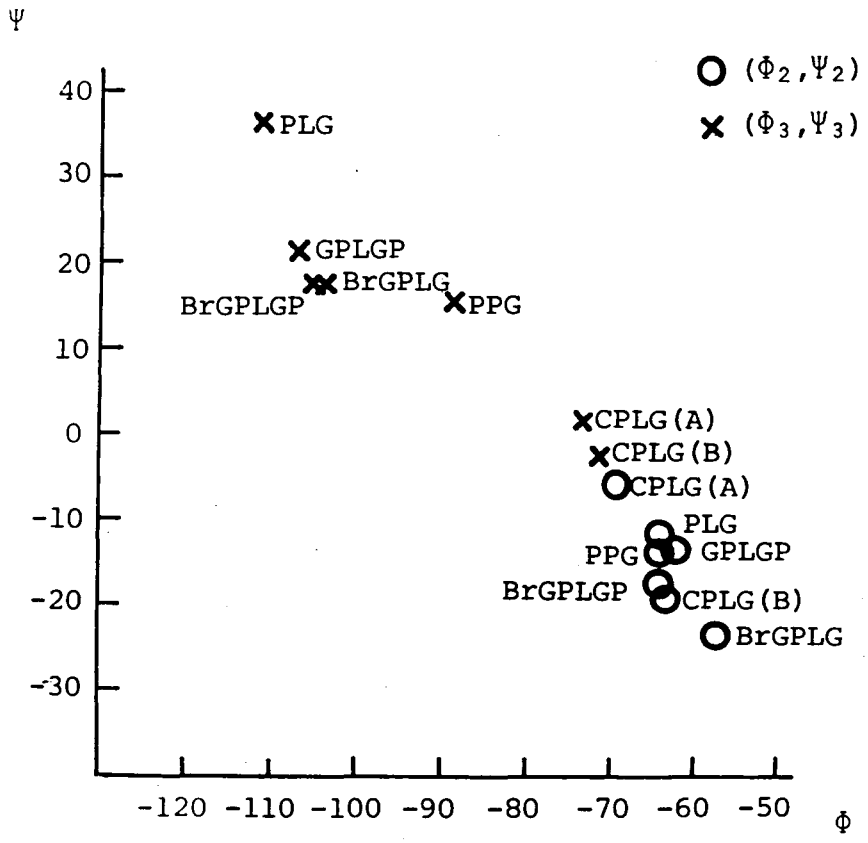


Fig. 20. Ramachandran plot of the main-chain torsion angles of the β-turn. Open circles refer to the angles (ϕ_2, ψ_2), the torsion angles for the second residue, and crosses refer to the angles (ϕ_3, ψ_3), the torsion angles for the third residue. PPG(Boc-Pro-Pro-Gly-NH_2), PLG(Boc-Pro-Leu-Gly-OH), GPLGP(Z-Gly-Pro-Leu-Gly-Pro-OH), BrGPLGP(Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH), BrGPLGP(Z(p-Br)-Gly-Pro-Leu-Gly-OH), CPLG(A) and CPLG(B) (S-Bzl-Cys-Pro-Leu-Gly-NH_2 A and B molecules).

Table 32. The main chain torsion angles in the β-turn

R_1 R_2 R_3 R_4	ϕ_2	ψ_2	ϕ_3	ψ_3
(p-Br)Z-Gly-Pro-Leu-Gly-OH	-57.7	-33.4	-104.0	8.0
(o-Br)Z-Gly-Pro-Leu-Gly-Pro-OH	-65.0	-26.7	-104.7	8.4
Z-Gly-Pro-Leu-Gly-Pro-OH	-63.2	-23.0	-107.3	12.0
S-Bzl-Cys-Pro-Leu-Gly-NH_2 (A)*	-70.2	-16.0	-74.1	-8.5
S-Bzl-Cys-Pro-Leu-Gly-NH_2 (B)*	-63.9	-28.9	-71.6	-11.9
Boc-Pro-Leu-Gly-OH	-65.0	-20.7	-110.8	26.7
Boc-Pro-Pro-Gly-NH_2	-64.9	-23.0	-88.8	6.1
H-Pro-Leu-Gly-NH_2	-61.2	127.8	71.8	-

*Some of the torsion angles for these molecules were incorrectly given in the original paper.[6]
The angles recalculated from the published atomic coordinates are given here.

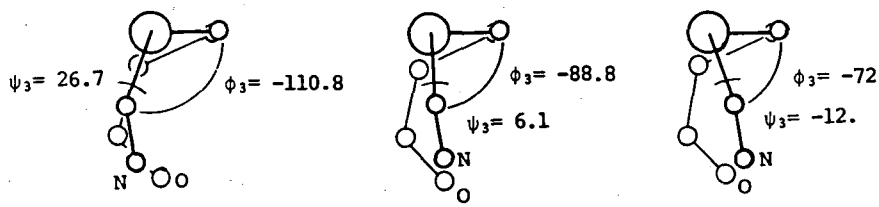


Fig. 21. Newman projection of the β-turn nine-membered ring viewed along the bond C^{α}-N of the third residue. The molecules are (left to right) Boc-Pro-Leu-Gly-OH, Boc-Pro-Pro-Gly-NH_2 and S-Bzl-Cys-Pro-Leu-Gly-NH_2(B).

Table 33. The comparison of the β-turn in various peptides. The mean distance (Å) between the atoms in one molecule and the corresponding atoms in the other. The comparison is made after refinement of the rotational and translational parameters in order to minimize the summation of the square distances between the corresponding atoms.
The upper figure; 9 atoms in the 10 membered β-turn ring excluding a hydrogen atom.
The lower figure; 16 atoms (14 atoms in parentheses), 9 atoms plus directory bonded atoms to them

	PPG	PLG	GPLGP	BrGPL GP	BrGPL G	CPLG (A)	CPLG (B)
PPG	-						
	-						
PLG	0.12	-					
	(0.17)	-					
GPLGP	0.12	0.05	-				
	(0.13)	0.12	-				
BrGPLGP	0.14	0.08	0.05	-			
	(0.15)	0.16	0.10	-			
BrGPLG	0.16	0.08	0.07	0.04	-		
	(0.17)	0.14	0.12	0.07	-		
CPLG(A)	0.12	0.23	0.23	0.25	0.27	-	
	(0.16)	0.42	0.38	0.41	0.45	-	
CPLG(B)	0.08	0.19	0.19	0.21	0.23	0.06	-
	(0.13)	0.41	0.36	0.38	0.42	0.09	-

3. β-sheet

The β-sheet in proteins is rich in variety depending on the directions of main chains, i.e., parallel or antiparallel, on the number of hydrogen bonds and on the wide variety of twist angles. The β-sheet in oligopeptides described in this section, however, has a rather fixed structures with only the minimum numbers of hydrogen bonds (one pair). Unlike the antiparallel β-sheets in proteins, β-sheet structures in the present oligopeptides involve two molecules and they may be referred to a dimer by the β-sheet type hydrogen bonding. In this section the three dimer structures in Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH and Aoc-Pro-Ala-Gly-OH will be studied together with similar structures found in Z-Gly-Pro-OH[33] and Z-Gly-D,L-Pro-OH.[34]

The ORTEP stereo drawings[17] of the five dimer peptides are shown in Fig. 22. Since all of these dimers, except Z-Gly-D,L-Pro-OH crystals, consist of the identical molecules, it is possible for the dimers to possess twofold symmetry. But an exact twofold symmetry exists only in Boc-Pro-Ile-Gly-OH and Z-Gly-Pro-OH dimers. The ORTEP drawings in Fig. 22 are viewed parallel to

the twofold axis for these two cases and perpendicularly to the vectors O-O and N-N involved in hydrogen bondings for others. There is no essential difference in structures of three dimers Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH and Aoc-Pro-Ala-Gly-OH in spite of the difference in symmetry. Since Boc-Pro-Val-Gly-OH and Aoc-Pro-Ala-Gly-OH deviate from twofold symmetry only in the atoms which are far from the center of molecules, they are still close to Boc-Pro-Ile-Gly-OH in the neighborhood around the center of dimer. On the other hand, Z-Gly-Pro-OH and Boc-Pro-Ile-Gly-OH, which commonly possess twofold symmetry, are more different especially in the directions of the C=O and N-H bonds from the main chain, as is shown in Fig. 22. These bonds protrude from the sheet of the main chains most in Z-Gly-Pro-OH and least in Z-Gly-D,L-Pro-OH.

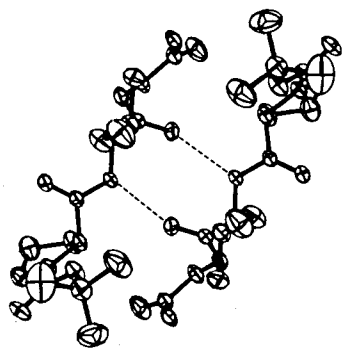
The main chain torsion angles (ϕ, ψ) for these peptides are listed in Table 34. The angles for three peptides Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH and Aoc-Pro-Ala-Gly-OH are very close to the typical antiparallel β-sheet angles (-139°, 135°), but those for Z-Gly-Pro-OH and Z-Gly-D,L-Pro-OH indicate that the former is closer to the collagen type and the latter to the fully extended type. Thus, dimer

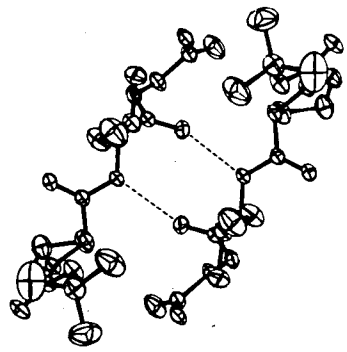
structures in peptides occur with such variety of main chain torsion angles.

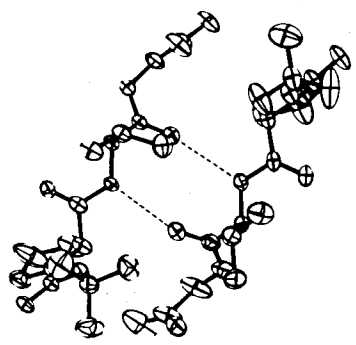
The comparison of these dimers by program CMROLS[50] (described in the preceding chapter) were summarized in Table 35. It is again obvious that these dimers can be divided into the three different groups. The values in Table 35 can be compared with those in Table 33 for the β-turn. The dimers linked by two hydrogen bonds have much more varieties in structures than the β-turn structure with one hydrogen bond.

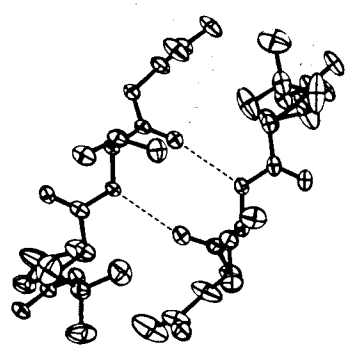
Levitt et al. pointed out in their study[51] to analyze automatically and objectively the secondary structures of the proteins that while the criterion based on the C^{α}-C^{α} distance is particularly good at delineating the clear β-sheets, the H-bond method is better able to detect shorter and more irregular β-sheet regions. The dimer structures in peptides validates his description. Z-Gly-Pro-OH and Z-Gly-D,L-Pro-OH can or cannot be classified as β-sheet depending on the criterions used to identify the structure. If more emphasis is put on the hydrogen bonding pattern they could be called β-sheet. On the other hand they would not be classified as β-sheet when the C^{α}-C^{α} distances or main chain

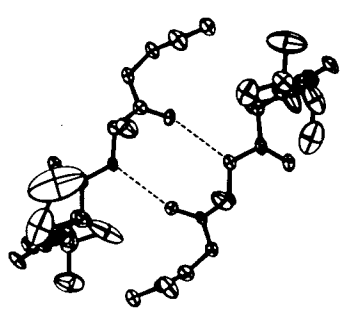
torsion angles are used as the criterions.

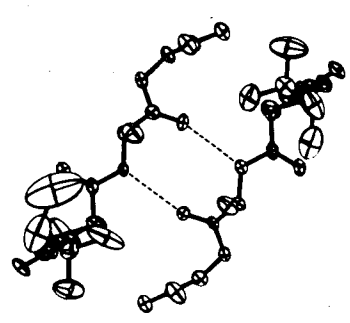












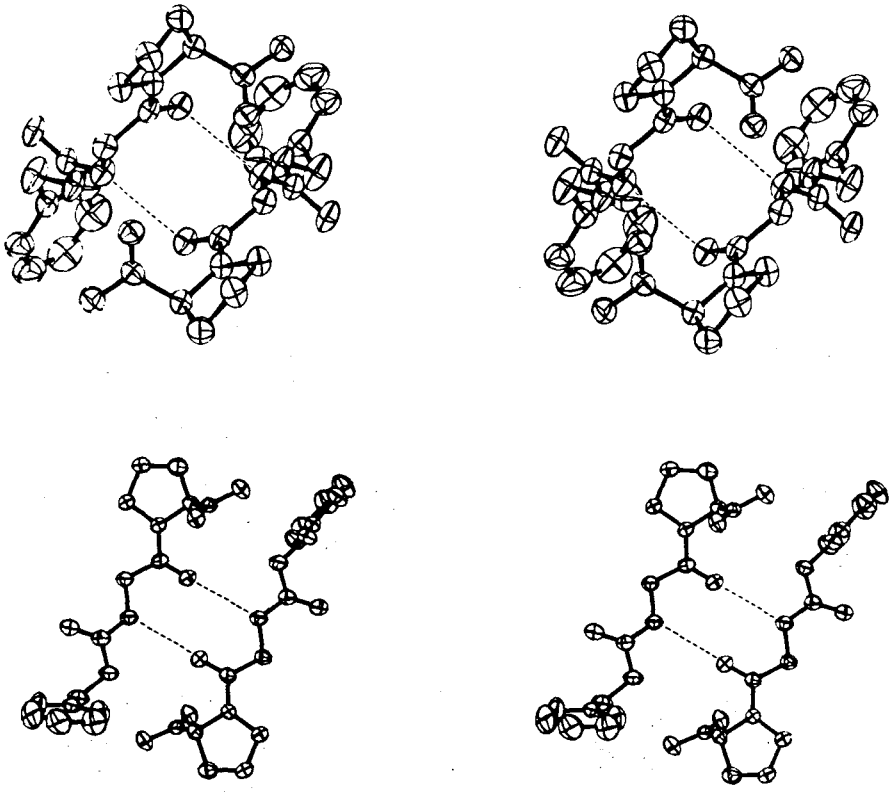


Fig. 22. Stereo drawing of the dimer structures of five peptides. From the top of the previous page, Boc-Pro-Ile-Gly-OH, Boc-Pro-Val-Gly-OH, Aoc-Pro-Ala-Gly-OH, Z-Gly-Pro-OH and Z-Gly-D,L-Pro-OH.

Table 34. Main chain torsion angles for the β-sheet type dimer

		Φ	Ψ
Boc-Pro-Ile-Gly-OH		-133.0	117.9
Boc-Pro-Val-Gly-OH	(A)	-134.7	123.2
	(B)	-138.4	141.6
Aoc-Pro-Ala-Gly-OH	(A)	-132.7	131.5
	(B)	-137.6	135.2
(L) Z-Gly-Pro-OH		-107.7	176.4
(DL) Z-Gly-Pro-OH	(A)	-173.6	171.8
	(B)	-175.7	170.8

Table 35. The comparison of the β-sheet type dimer in various peptides. The mean distance (Å) between the atoms in one dimer and the corresponding atoms in the other. The comparison is made after refinement of the rotational and translational parameters in order to minimize the summation of the square distances between the corresponding atoms.

The upper figure; 8 atoms in the 10 membered β-sheet dimer ring excluding hydrogen atoms.

The lower figure; 14 atoms, 8 atoms plus directory bonded atoms to them.

	PVG	PIG	PAG	L-GP	DL-GP
PVG	-				
	-				
PIG	0.14	-			
	0.24	-			
PAG	0.14	0.23	-		
	0.22	0.42	-		
L-GP	0.76	0.81	0.63	-	
	1.52	1.67	1.33	-	
DL-GP	0.47	0.60	0.41	0.77	-
	0.52	0.73	0.39	1.32	-

IV. THE APPLICATION OF THE PEPTIDE SECONDARY STRUCTURE TO THE CRYSTAL STRUCTURE ANALYSIS

1. Introduction

A systematic Patterson function interpretation by computer based on a known rigid group was first carried out by Nordman & Nakatsu.[52] Since then this method has played a role mainly in solving crystal structures of fused ring systems such as aromatic derivatives. In this chapter the application of the vector space search method to more complicated system, both in the number of atoms and in rigidity of the rigid group, will be described.

Since the conformation of the oligopeptide in the crystalline state has usually been difficult to presume, the vector space search method seemed to be insensate. But in Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH[3] and Z(p-Br)-Gly-Pro-Leu-Gly-OH[4] the β-turn part was found so similar that the part could be regarded as a rigid group. So in solving the crystal structure of Z-Gly-Pro-Leu-Gly-Pro-OH,[2] the vector space search method using the β-turn as a rigid group was first successfully applied.

In this chapter, the computer program RICS for the vector space search method and its application to

some peptide crystals will be described.

2. The vector space search program RICS

Program RICS (Rigid Group Convolution and Search) was first written in an attempt to solve the crystal structure of phenyl α-maltoside.[53] Since then the program was revised in two main points.

The vector space search method depends strongly on the molecules constituting the crystal, especially on the number of the atoms and the shape of the rigid group. The first improvement was made by rewriting the program in such a way to require commands which make it possible to meet various needs. The commands are to be arranged in the most appropriate way for each case. A command consists of four integers; two as main command and two as subcommand.

Main commands are:

00 Initial Control Constant Change

01 Crystal Information Input

02 Control Data for Patterson Function Input

03 Rigid Group Input

04 F Modification

05 Patterson Function Calculation

06 Rotation Search Execution

07 Translation Search Execution

08 Cross Vector Function Execution

Subcommand is used as a supplement.

The second improvement was made on the translation search. Of the two most important calculations in the vector space search method, the rotation search is executed in the real space, while the translation search is in the reciprocal space. Originally RICS worked both in the real space, but the translation search in the real space was so time consuming that it was altered to calculate in the reciprocal space. The calculation time is one of the most important points because a number of probable orientations found in the rotation search must be checked in the following translation search. The phase modulated translation function[54] was applied for the translation search rather than the conventional translation function.[55][57] Once the orientation of a rigid group is given, one may define its molecular transform $F_1(h)$ as

$$F_1(h)=|F_1(h)|\exp(i\phi_1)=\Sigma_j f_{1j}\exp(2\pi i h\cdot r_{1j}) \qquad (1)$$

Here ϕ_1 is phase angle of the reflection h calculated from the first molecular fragment, whose j-th atomic coordinate is r_{1j} and the atomic scattering factor is f_{1j}. Subscript 2 indicates the second molecular fragment, usually generated by the symmetry operation of the first fragment. The phase modulated translation function $\Phi(r)$ is written in the form

$$\Phi(r)=G(h)\exp(-i(\phi_1-\phi_2))\exp(-2\pi ih\cdot r) \quad (2)$$

Here G(h) is calculated from the observed structure factor $|F(h)|$ as

$$G(h)=|F(h)|^2-|F_1(h)|^2-|F_2(h)|^2 \quad (3)$$

on the other hand, the conventional translation function T(r), which is modulated both in phase and amplitude, is written as

$$T(r)=G(h)|F_1(h)||F_2(h)|\exp(-i(\phi_1-\phi_2))\exp(-2\pi ih\cdot r) \quad (4)$$

As Langs proved,[54] the magnitude of spurious peaks produced by the phase-modulated translation function is expected to be a half that produced by the conventional translation function.

3. Application

3-1. Boc-Pro-Leu-Gly-OH

Although Boc-Pro-Leu-Gly-OH structure was solved by a direct method, the structure could also have been solved by the vector space search method if the β-turn was properly anticipated. As a test calculation the vector space search method was applied on this peptide crystal, using three different β-turn fragments of Boc-Pro-Leu-Gly-OH itself, Z-Gly-Pro-Leu-Gly-Pro-OH and Boc-Pro-Pro-Gly-NH_2. The main purpose of this attempt is in the establishment of the most effective route of the vector space search method when a rigid group is accompanied with some structural ambiguity.

The results of the rotation search using 16 atoms (14 atoms for the model Boc-Pro-Pro-Gly-NH_2), which constitute β-turn, are shown in Table 36. The rigid groups used for this calculation were the same as those used for the calculation in Table 33. The β-turn of the model structures are shown to be 0.12Å (Z-Gly-Pro-Leu-Gly-Pro-OH) and 0.17Å (Boc-Pro-Pro-Gly-NH_2) away, in average, from that of Boc-Pro-Leu-Gly-OH. A weighted sum function was used as a measure of fit.[53] The highest peaks in each rotation maps correspond to the correct orientation though they are not so prominent. The phase modulated

translation function was then calculated for each of the orientations which correspond to the five highest peaks in each map. Among them, only the translation maps for the correct orientation are given in Fig. 23. These maps show clearly the correct translation vectors. Other maps for the wrong orientations were commonly featureless. The partial structure-tangent refinement method[56] was next carried out and the E maps gave all of the non-hydrogen atoms.

Since the β-turn has some flexibility in its structure, the model may be divided into smaller but more rigid fragments. From this viewpoint the β-turn model was divided into three smaller groups, each of which consists of five or seven atoms and the rotation search was executed. This result, however, was worse in two points. First, the orientation parameters obtained were less accurate than those obtained by the larger models. This is because, when a smaller rigid group is used, only the vicinity of the origin of the Patterson map is more intensively used, which has severe overlaps of vectors.

Thus, the rigid group should be carefully chosen taking account of the two sometimes conflicting requirements, rigidity and size. In the following sections, the actual examples by the use of the less

rigid but larger groups will be shown

3-2. Boc-Pro-Val-Gly-OH

The crystal has two peptide molecules and one water in an asymmetric unit, so the non-hydrogen atoms amount to 53. After several unsuccessful attempts by the direct method, the vector space search method was applied on the assumption that the molecule has either a β-turn conformation or an extended conformation. For the β-turn model, 16 atoms of Boc-Pro-Leu-Gly-OH, which are the same as used in the calculation in Table 33, were used and for the extended model 20 atoms of Boc-Pro-Ile-Gly-OH were used excluding four carbon atoms in the Boc group, C^{δ} in the isoleucyl residue and two carboxyl oxgen atoms.

Five highest peaks were selected from each of the rotation maps and directly checked by the translation search. While all of the translation maps for the β-turn model were featureless, two maps for the extended model showed prominent peaks. The results of the rotation search by the extended model are listed in Table 37, and two translation maps for the correctly oriented models are shown in Fig. 24. The whole structure was solved by the partial

structure-tangent refinement method based on one of the two molecular fragments thus obtained.

Although the structure could thus be solved, an alternative method based on the determination of the relative disposition of the two fragments was examined. In formula (2), when ϕ_1 and ϕ_2 are calculated from each of the fragments including the ones transformed by the crystal symmetry (in this case 2_1 operation), $\Phi(r)$ gives the shift parameter needed for the first fragment. The inter-fragments vectors will be found in one of the four lines, $\Phi(0,y,0)$, $\Phi(0,y,1/2)$, $\Phi(1/2,y,0)$, and $\Phi(1/2,y,1/2)$ in the case of $P2_1$ space symmetry. Figure 25 shows $\Phi(0,y,1/2)$ in which the highest peak corresponds to the correct inter-fragment vector.

The program CMROLS[50] showed later that the average deviation of the rigid model from Boc-Pro-Val-Gly-OH (A) and (B) was 0.125Å and 0.554Å, respectively. It may be worthwhile to note that even the (B) molecule could be correctly located by the vector space search method. Some of the torsion angles of the (B) molecule are nearly 20 degrees away from those in the used model.

3-3. Aoc-Pro-Ala-Gly-OH

In solving this crystal by the vector space search method, the same models which was used for the analysis of Boc-Pro-Val-Gly-OH was first tried. These rigid groups, however, were not sufficient to solve the structure.

A clue for getting to its structure was found in the detailed investigation of the crystal structures of Boc-Pro-Ile-Gly-OH and Boc-Pro-Val-Gly-OH. As was described in chapter II, the structures of Boc-Pro-Ile-Gly-OH and Boc-Pro-Val-Gly-OH were similar not merely as a single molecule but as a dimer. Then a rigid group model of the dimer structure seemed to be promising. The result of the rotation search is summarized in Table 35. The translation map made for the highest peak in Table 35 is shown in Fig. 26. The partial structure-tangent refinement method using 40 atoms thus obtained, gave easily the whole structure.

Later, the vector space search method of using extended conformation model was found to have been correct. It was at the extension process from the fragments to the whole structure that had failed to work. To establish a proper criterion for selecting a correct fragment is the urgent need for the vector

space search method. Then putting the vector space search and the direct method together seems to widen the crystallographic horizon.

Table 36. Boc-Pro-Leu-Gly-OH Rotation Search

a. rigid model; Boc-Pro-Leu-Gly-OH β-turn 16 atoms

Peak No.	θ	χ	ϕ	Peak Height
1*	-8.3	230.3	57.0	62
2	54.6	289.5	96.8	40
3	-19.5	174.2	71.1	40
4	10.5	348.2	59.1	39
5	-17.1	187.1	79.1	39

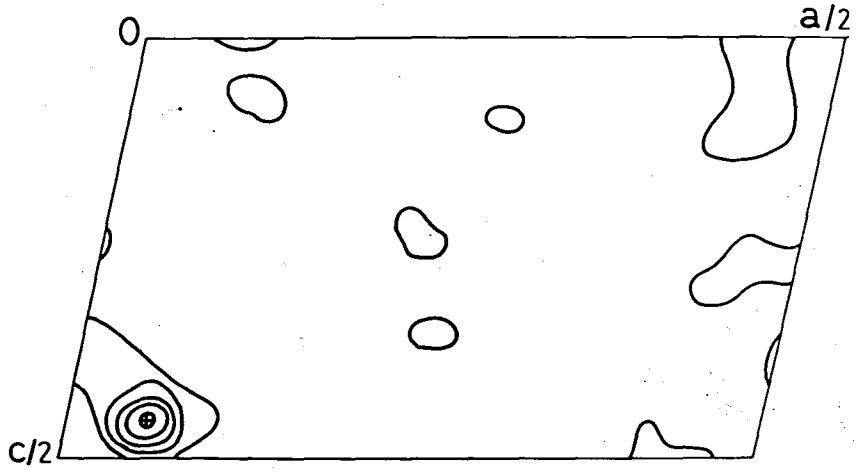
b. rigid model; Z-Gly-Pro-Leu-Gly-Pro-OH β-turn 14 atoms

Peak No.	θ	χ	ϕ	Peak Height
1*	-11.4	231.2	62.0	53
2	-59.6	78.8	133.2	44
3	-2.1	30.8	129.3	41
4	-9.2	14.6	179.0	41
5	9.0	226.1	59.9	40

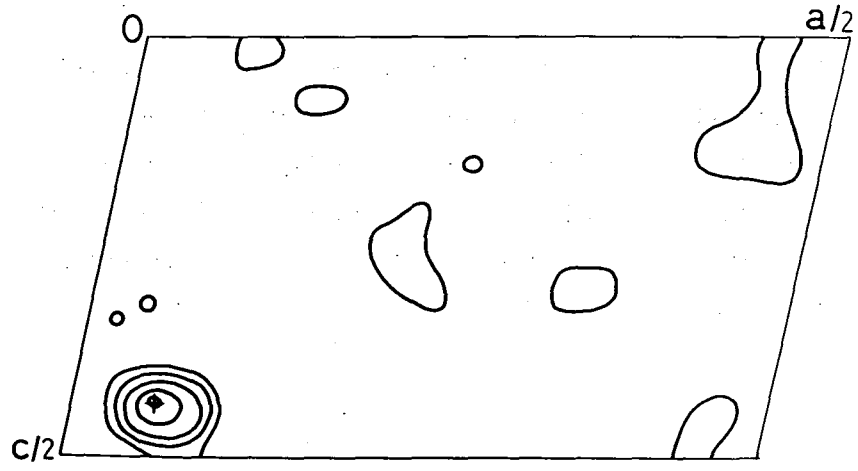
c. rigid model; Boc-Pro-Pro-Gly-NH_2 β-turn 16 atoms

Peak No.	θ	χ	ϕ	Peak Height
1*	-14.4	224.7	64.3	43
2	-2.7	232.0	63.0	41
3	53.1	131.5	179.1	40
4	55.3	345.0	161.0	39
5	53.9	316.1	146.3	39

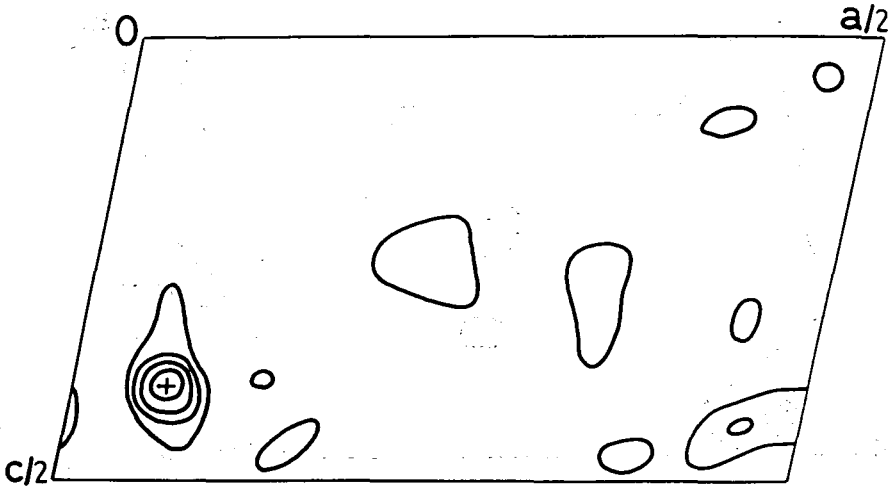
* correct peak



(a)



(b)



(c)

Fig. 23. Translation search of Boc-Pro-Leu-Gly-OH crystal. Rigid model is β-turn from Boc-Pro-Leu-Gly-OH (a), Z-Gly-Pro-Leu-Gly-Pro-OH (b) and Boc-Pro-Pro-Gly-NH_2 (c). The correct translation vector is marked by + in each map.

Table 37. Boc-Pro-Val-Gly-OH Rotation Search

rigid model; Boc-Pro-Ile-Gly-OH 20 atoms

Peak No.	θ	χ	ϕ	Peak Height
1*	8.3	273.8	159.7	80
2*	-5.4	93.3	143.8	79
3	-1.2	263.1	9.2	77
4	-5.6	91.8	117.8	72
5	-10.9	83.3	27.9	68

*correct peak

Table 38. Aoc-Pro-Ala-Gly-OH Rotation Search

rigid model; Boc-Pro-Val-Gly-OH dimer 38 atoms

Peak No.	θ	χ	ϕ	Peak Height
1*	7.5	0.7	178.6	135
2*	-8.0	181.1	120.6	133
3	26.7	242.8	149.6	102
4	-21.4	55.5	155.1	99
5	63.7	72.4	16.9	93

*correct peak

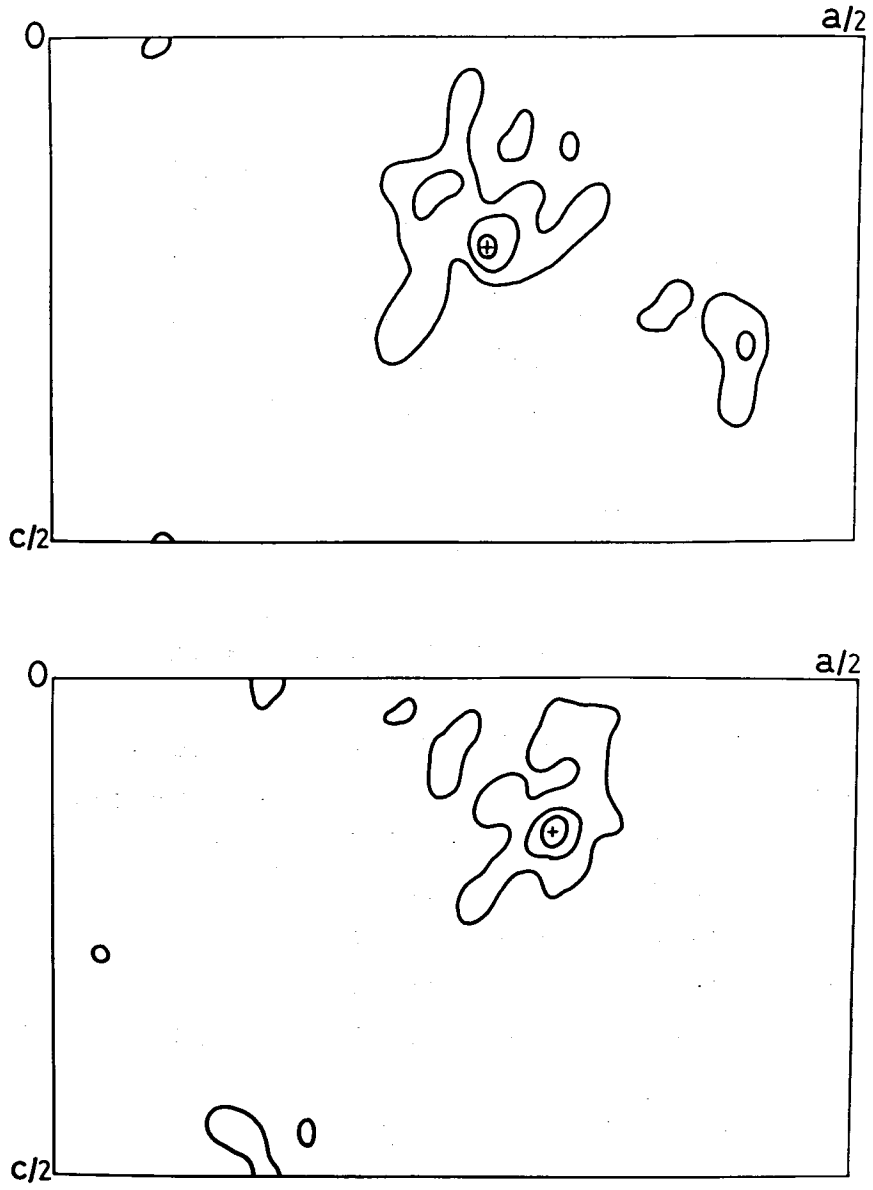


Fig. 24. The translation search of Boc-Pro-Val-Gly using 20 atoms of Boc-Pro-Ile-Gly-OH as rigid group. The orientation parameters are from peak No 1 and 2 in Table 37. The + mark in each map corresponds to the correct translation vector for the independent molecule (A) and (B) respectively.

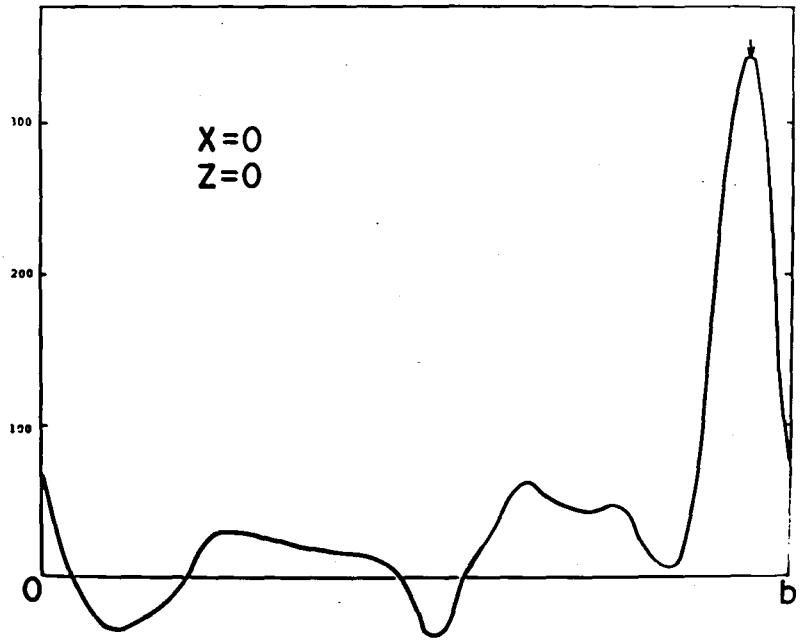


Fig. 25. The translation search of the inter-fragment vector of Boc-Pro-Val-Gly-OH (A) and (B). The arrow indicates the correct vector.

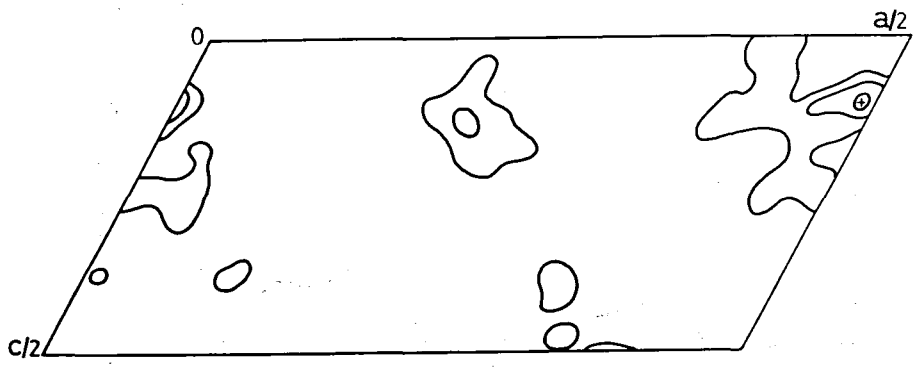


Fig. 26. The translation srarch of Aoc-Pro-Ala-Gly-OH using the dimer as rigid group.
The orientation parameter is that of the peak number 1 in Table 38.
The correct translation vector is marked by the cross.

V. CONCLUSION

By a series of analyses of the mutually related peptides, the concept of the secondary structure, which was developed to interpret the protein structure, has been proved to be important also in these small peptide molecules. They are the β-turn and the β-sheet structures. The lack of the α-helix, one of the most important secondary structures, is attributable to the fact that the α-helix is based on the hydrogen bonds between more distant amino residues (5-1 interaction). It is no doubt that the α-helix is also to be seen in some longer peptides.

The nature of the secondary structures observed in the peptides were studied in chapter III and summarized in Tables 32 to 35 and in Figure 20. Although this is a comparison within a small number of biased group of peptides, the β-turn having proline at the second site and the β-sheet having a pair of hydrogen bonds, it is important that the dispersion of the torsion angles given in these Tables and Figure does not have any ambiguity which is inevitable in protein structure analyses. A distinctive feature of the torsion angle dispersion is in the coordinated changes of the main chain angles so as not to disturb

hydrogen bonding. It was also shown that a β-sheet structure, which is stabilized by the two hydrogen bonds, fluctuates more than a β-turn structure with one intramolecular hydrogen bond.

The knowledge about the effect of the X residue on the conformation of the Boc-Pro-X-Gly-OH molecule was also obtained to some extent. Among the examined peptides, those having leucine at X site takes the β-turn conformation, while those having isoleucine, valine or alanine commonly take β-sheet structure. Boc-Pro-Pro-Gly-NH_2 is another example of the β-turn, which is quite different in nature, owing to the C-terminal amide group and consecutive prolyl residue. The fact that among those peptides which commonly have hydrophobic residue at X site, only leucyl peptide takes the β-turn conformaion cannot be dismissed as a trivial matter. Because some other peptides such as Z-Gly-Pro-Leu-Gly-Pro-OH,[2] Z(o-Br)-Gly-Pro-Leu-Gly-Pro-OH,[3] Z(p-Br)-Gly-Pro-Leu-Gly-OH,[4] and S-benzyl-Cys-Pro-Leu-Gly-NH_2 commonly have the same β-turn structure with leucine at the third site. This shows that leucine is more favorable than other hydrophobic residues for the β-turn third site, perhaps because of the stereochemical requirement. This conclusion, however, seemingly contradicts

against the conclusion[47] obtained from protein crystal structures. That is, in proteins all the hydrophobic residues, including leucine, are hard to be involved in the β-turn. In this respect the difference of the driving force to build a β-turn is suggested. In peptides stereochemical requirements and hydrogen bond formation are most important in determining molecular structures, but in proteins, in addition to these requirements, the hydrophobic interaction caused from surrounding aqueous solution is also important. Leucine may be such an residue that is stereochemically preferable at the third site but does not occur at the third site in proteins only because of its hydrophobicity.

In addition to these rather general conclusion, some detailed structural rule was also established. They are firstly a widening of the $NC^{\alpha}C'$ angle in the β-turn type I at the second site (chapter II 1-4) and the irrelevance of the side chain conformation to the main chain conformation (chapter II 1-4, 3-4). The former point must be taken into consideration in the high resolutional protein structure analysis. The latter indicates that several possible conformations must be equally checked in the analysis of the fibrous protein.

From the X-ray analytical point of view, the vector space search method was proved to be a useful method to apply with the peptide crystals owing to the unexpected fact that the peptides take either one of the rather small number of the secondary structures. Especially, for the study of a series of peptides the vector space search method would be a powerful method. The wider application would be promised in future as a more reliable anticipation of the tertiary structure becomes available.

References

1) J.H. Hodsdon, B. Shaw, J.M. Schurr and L.H. Jensen
Amer. Cryst. Assoc., Abstr. Papers (Winter Meeting) 66 (1972)

2) S. Bando, N. Tanaka, T. Ashida and M. Kakudo
Acta Cryst., B34, 3447-3449 (1978)

3) T. Ueki, S. Bando, T. Ashida and M. Kakudo
Acta Cryst., B27, 2219-2231 (1971)

4) T. Ueki, T. Ashida, M. Kakudo, Y. Sasada and Y. Katsube
Acta Cryst., B25, 1840-1849 (1969)

5) C.M. Venkatachalam
Biopolymers, 6, 1425-1436 (1968)

6) A.D. Rudko and B.W. Low
Acta Cryst., B31, 713-725 (1975)

7) J.L. Crawford, W.N. Lipscomb and C.G. Schellman
Proc. Nat. Acad. Sci. USA 70, 538-542 (1973)

8) W.H. Zachariasen
Acta Cryst., 23, 558-564 (1967)

9) T. Matsuzaki
Acta Cryst., B30, 1029-1036 (1974)

10) E. Benedetti, M.R. Ciajolo and A. Maisto
Acta Cryst., B30, 1783-1788 (1974)

11) G. Kartha, T. Ashida and M. Kakudo
Acta Cryst., B30, 1861-1866 (1974)

12) F. Robert

Acta Cryst., B32, 2361-2369 (1976)

13) G. Germain, P. Main and M.M. Woolfson

Acta Cryst., A27, 368-376 (1971)

14) T. Ashida

UNICS-Osaka, The Computation Center, Osaka Univ.

55-61 (1973)

15) International Tables for X-Ray Crystallography

Vol. IV, Birmingham: Kynoch Press. (1974)

16) IUPAC-IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE

J. Mol. Biol., 52, 1-17 (1970)

17) C.K. Johnson

ORTEP. Oak Ridge National Laboratory Report

ORNL-3794 (1965)

18) R.E. Marsh and J. Donohue

Advance Protein Chemistry, 22, 235-256 (1967)

19) T. Ashida and M. Kakudo

Bull. Chem. Soc. Japan, 47, 1129-1133 (1974)

20) H. Sugino, I. Tanaka and T. Ashida

Bull. Chem. Soc. Japan, 51, 2855-2861 (1978)

21) H. Yoshioka, K. Nakatsu, M. Sato and T. Tatsuno

Chem. Lett., 1319-1322 (1973)

22) I.L. Karle

J. Amer. Chem. Soc., 96, 4000-4006 (1974)

23) I.L. Karle

Biochemistry, 13, 2155-2162 (1974)

24) H. Yoshioka, T. Aoki, H. Goko, K. Nakatsu, T. Noda, H. Sakakibara, T. Take, A. Nagata, J. Abe, T. Wakamiya, T. Shiba and T. Kaneko
Tetrahedron Lett, 2043-2046 (1971)

25) H. Yoshida and K. Nakatsu
private communication (1976)

26) I.L. Karle, J.W. Gibson and J. Karle
J. Amer. Chem. Soc., 92, 3755-3760 (1970)

27) L.L. Reed and P.L. Johnson
J. Amer. Chem. Soc., 95, 7523-7524 (1973)

28) A. Zalkin, J.D. Forrester and D.H. Templeton
J. Amer. Chem. Soc., 88, 1810-1814 (1966)

29) A. Aubry
PhD Thesis Univ. of Nancy (1976)

30) U.W. Arndt and B.T.M. Willis
Single Crystal Diffractometry, Cambridge Univ. Press (1966)

31) V. Sasisekharan
Acta Cryst., 12, 897-903 (1959)

32) R. Balasubramanian, A.V. Lakshminarayanan, M.N. Sabesan, G. Tegoni, K. Venkatesan, G.N. Ramachandran
Int. J. Protein Res., 3, 25 (1971)

33) I. Tanaka, T. Kojima, T. Ashida, N. Tanaka and M. Kakudo
Acta Cryst., B33, 116-119 (1977)

34) T. Kojima, T. Yamane and T. Ashida
Acta Cryst., B34, 2896-2898 (1978)

35) I.L. Karle and J. Karle
Acta Cryst., 16, 969-975 (1963)

36) K. Torii and Y. Iitaka
Acta Cryst., B27, 2237-2246 (1971)

37) K. Torii and Y. Iitaka
Acta Cryst., B26, 1317-1326 (1970)

38) L. Pauling, R.B. Corey and H.R. Branson
Proc. Nat. Acad. Sci. USA, 37, 205-211 (1951)

39) L. Pauling and R.B. Corey
Proc. Nat. Acad. Sci. USA, 37, 729-738 (1951)

40) M.F. Perutz
Nature (London) 167, 1053-1054 (1951)

41) R.E. Marsh, R.B. Corey and L. Pauling
Biochem. Biophys. Acta, 16, 1-34 (1955)

42) F.C. Kendrew, R.E. Dickerson, B.E. Strandberg,
R.G. Hart, D.R. Davies, D.C. Phillips and
V.C. Shore
Nature (London) 185, 422-427 (1960)

43) C.C.F. Blake, D.F. Koenig, G.A. Mair, A.C.T. North,
D.C. Phillips and V.R. Sarma
Nature (London) 206, 757-761 (1965)

44) P.N. Lewis, F.A. Monany and H.A. Scheraga
Proc. Nat. Acad. Sci. USA, 68. 2293-2297 (1971)

45) I.D. Kuntz

J. Amer. Chem. Soc., 94, 4009-4012 (1972)

46) G.D. Rose and J.P. Seltzer

J. Mol. Biol., 113, 153-164 (1977)

47) P.Y. Chou and G.D. Fasman

J. Mol. Biol., 115, 135-175 (1977)

48) G.D. Rose

Nature, 272, 586-590 (1978)

49) Y. Nozaki and C. Tanford

J. Biol. Chem., 249, 2211-2217 (1971)

50) H. Itoh

Master's thesis, Nagoya Univ. (1977)

51) M. Levitt and J. Greer

J. Mol. Biol., 114, 181-293 (1977)

52) C.E. Nortman and M. Nakatsu

J. Amer. Chem. Soc., 85, 353 (1963)

53) I. Tanaka

Master's Thesis, Osaka Univ. (1973)

54) D.A. Langs

Acta Cryst., A31, 543-550 (1975)

55) V. Vand and R. Pepinsky

Z. Kristallogr., 108, 1-14 (1956)

56) J. Karle

Acta Cryst., B24, 182-186 (1968)

57) P. Tollin

Acta Cryst., 21, 613-614 (1966)

List of publications

I. Related papers

1. The Crystal Structure of tert-Butoxycarbonyl-L-prolyl-L-leucylglycine Hydrate
   T. Ashida, I. Tanaka, Y. Shimonishi and M. Kakudo
   Acta Cryst., B33, 3054-3059 (1977)
2. The Crystal and Molecular Structure of tert-Butyloxycarbonyl-L-prolyl-L-prolyl-glycinamide
   I. Tanaka, T. Ashida, Y. Shimonishi and M. Kakudo
   Acta Cryst., B35, 110-114 (1979)
3. The Crystal and Molecular Structure of tert-Butyloxycarbonyl-L-prolyl-L-isoleucylglycine
   Y. Yamada, I. Tanaka and T. Ashida
   Acta Cryst., to be published
4. The Crystal and Molecular Structure of tert-Butyloxycarbonyl-L-prolyl-L-valylglycine hemihydrate
   I. Tanaka and T. Ashida
   Acta Cryst., to be published
5. The Crystal Structure of tert-Amyloxycarbonyl-L-prolyl-alanylglycine
   Y. Yamada, I. Tanaka and T. Ashida
   to be published

II. Other papers

1. The Crystal and Molecular Structure of phenyl α-Maltoside

   I. Tanaka, N. Tanaka, T. Ashida and M. Kakudo

   Acta Cryst., B32, 155-160 (1976)

2. Benzyloxycarbonylglycyl-L-proline

   I. Tanaka, T. Kojima, T. Ashida, N. Tanaka and M. Kakudo

   Acta Cryst., B33, 116-119 (1977)

3. The Crystal Structures of Complexes of Cyclodisarcosyl with Metal Salts.

   I. The 2:1 Complex of Cyclodisarcosyl with Lithium Perchlorate

   N. Takahashi, I. Tanaka, T. Yamane, T. Ashida, T. Sugihara, Y. Imanishi and T. Higashimura

   Acta Cryst., B33, 2132-2136 (1977)

4. Cyclo-L-leucyl-L-histidyl Monohydrate

   I. Tanaka, T. Iwata, N. Takahashi and T. Ashida

   Acta Cryst., B33, 3902-3904 (1977)

5. Crystal Structures of N-(benzyloxycarbonyl) prolylleucine Ethyl Ester and N-(t-Butoxycarbonyl) prolylleucine Benzyl Ester

   H. Sugino, I. Tanaka and T. Ashida

   Bull. Chem. Soc. Japan, 51, 2855-2861 (1978)

The X-ray Studies of the Secondary Structures in Oligopeptides

1979

Isao Tanaka