

# Internal Water Bridge and Antiparallel Sheet in the Structure of Benzyloxycarbonyl-L-alanyl-D-phenylalanyl-L-proline Monohydrate

BY C. M. K. NAIR, R. NAGARAJ, S. RAMAPRASAD, P. BALARAM AND M. VIJAYAN

*Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India*

## Abstract

Benzyloxycarbonyl-L-alanyl-D-phenylalanyl-L-proline monohydrate,  $C_{25}H_{29}N_3O_6 \cdot H_2O$ , crystallizes in the orthorhombic space group  $P2_12_12_1$  with four molecules in a unit cell of dimensions  $a = 9.594$  (9),  $b = 9.705$  (4) and  $c = 27.917$  (12) Å. The structure has been refined to an  $R$  value of 0.067 for 2046 observed reflections. All the peptide units in the molecule are *trans* and the prolyl residue is in the  $C_2-C^{\gamma}$ -*exo*- $C^{\beta}$ -*endo* conformation. The lone water molecule in the structure is hydrogen bonded to the carbonyl O atom in the

benzyloxycarbonyl group and to one of the O atoms in the terminal carboxyl group. This internal water bridge, observed for the first time in a linear peptide, provides a model for water-mediated chain-reversal. An interesting feature of the crystal structure is the presence of an antiparallel sheet involving the alanyl and the phenylalanyl residues.

## Introduction

Structural studies on small proline-containing peptides with all L, and mixed L and D sequences are valuable in

elucidating the role of prolyl residues in determining the conformation of proteins, polypeptides and peptide antibiotics. We report here the crystal structure of the monohydrate of benzyloxycarbonyl-L-alanyl-D-phenylalanyl-L-proline, a proline-containing tripeptide with LDL sequence. In view of the importance of bound water molecules in the structure and function of proteins, it was also of interest to investigate the influence of the lone water molecule in the structure on peptide conformation and crystal packing.

### Experimental

The compound was crystallized from an aqueous methanol solution. The space group and the unit-cell dimensions were determined from X-ray diffraction photographs. The cell parameters were subsequently refined on a **CAD-4** four-circle diffractometer. The density of the samples was measured by flotation in aqueous potassium iodide solution. The observed density [ $1.260$  (**5**)  $\text{Mg m}^{-3}$ ] agrees well with that ( $1.262 \text{ Mg m}^{-3}$ ) calculated for four peptide molecules and four water molecules in the unit cell.

Intensity data were collected from a specimen of dimensions  $0.7 \times 0.4 \times 0.3 \text{ mm}$  on a computer-controlled CAD-4 diffractometer employing the  $\omega-2\theta$  scan up to a maximum Bragg angle of  $75^\circ$  using graphite-monochromated  $\text{Cu } K\alpha$  radiation. Of the 2582 reflections collected in this range, 2046 having  $I > 2\sigma(I)$  were used for structure determination and refinement. The intensities were corrected for Lorentz and polarization factors but not for absorption ( $\mu = 0.78 \text{ mm}^{-1}$  for  $\text{Cu } K\alpha$  radiation).

### Structure analysis

Although the structure contained only 35 crystallographically independent non-hydrogen atoms and the space group was favourable, structure solution using direct methods turned out to be somewhat difficult. Attempts at phase determination employing *MULTAN* (Germain, Main & Woolfson, 1971) using normal *E* values as well as those normalized separately in different parity groups and different ranges of Bragg angle were of no avail. *E* values were then calculated using group scattering factors for the two six-membered rings and the pyrrolidine ring in the structure. A chemically sensible fragment containing 12 atoms could be identified with difficulty in the *E* map corresponding to the best set of phases produced by *MULTAN* using these *E* values. This fragment was input with random orientation and random position in the next *MULTAN* (1977 version) run. The subsequent *E* map based on the best set of phases revealed the positions of 29 out of the 35 non-hydrogen atoms. The rest of the atoms were located from a subsequent difference Fourier map.

The structure was refined by the block-diagonal SFLS method using the locally modified version of a

program originally written by Professor R. Shiono. The heavy atoms and the H atoms, located from a difference Fourier map, were assigned anisotropic and isotropic thermal parameters respectively in the final cycles. The refinement converged at  $R = 0.067$ . The weighting function used had the form  $1/(a + bF_o + cF_o^2)$  where  $a = 1.20$ ,  $b = 0.049$  and  $c = -0.0001$ . The scattering factors of the non-hydrogen atoms and the H atoms were taken from Cromer & Waber (1965) and Stewart, Davidson & Simpson (1965) respectively. The computations were carried out first on an IBM 360/44 machine and subsequently on a DEC-1090 computer. The final positional parameters and the equivalent isotropic temperature factors (Hamilton, 1959) of the non-hydrogen atoms are given in Table 1.\*

\* Lists of structure factors, thermal parameters and H atom parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 35734 (18 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 1. *Fractional coordinates ( $\times 10^4$ ) and equivalent isotropic temperature factors (Hamilton, 1959) of the non-hydrogen atoms*

Standard deviations are given in parentheses.

	x	y	z	B ( $\text{\AA}^2$ )
C(1)	4669 (11)	471 (13)	2699 (3)	9.2 (5)
C(2)	4934 (11)	875 (14)	3168 (3)	9.0 (4)
C(3)	5971 (13)	1692 (13)	3274 (3)	9.1 (4)
C(4)	6843 (14)	2205 (13)	2930 (3)	9.8 (5)
C(5)	6588 (12)	1804 (12)	2453 (3)	8.5 (4)
C(6)	5482 (9)	945 (9)	2331 (3)	5.6 (3)
C(7)	5249 (13)	502 (10)	1821 (3)	8.4 (4)
O(8)	5452 (6)	1696 (7)	1514 (2)	6.5 (2)
C(9)	4587 (7)	1794 (8)	1126 (2)	4.7 (2)
O(10)	3680 (5)	973 (6)	1039 (2)	5.8 (2)
N(11)	4878 (5)	2943 (6)	882 (2)	4.3 (2)
C(12)	4131 (6)	3263 (8)	445 (2)	4.0 (2)
C(13)	4139 (9)	4792 (8)	349 (3)	5.7 (3)
C(14)	4821 (6)	2488 (8)	19 (2)	3.8 (2)
O(15)	6088 (4)	2342 (7)	-1 (2)	5.8 (2)
N(16)	3951 (5)	2093 (6)	-324 (2)	3.8 (2)
C(17)	4456 (6)	1459 (7)	-765 (2)	3.5 (2)
C(18)	4343 (7)	2490 (8)	-1189 (2)	4.5 (2)
C(19)	4977 (8)	1918 (9)	-1644 (2)	4.8 (2)
C(20)	4197 (9)	999 (11)	-1934 (3)	6.4 (3)
C(21)	4779 (14)	482 (13)	-2362 (3)	9.4 (5)
C(22)	6136 (13)	890 (15)	-2487 (3)	11.1 (5)
C(23)	6863 (11)	1729 (17)	-2195 (4)	10.5 (5)
C(24)	6313 (9)	2239 (12)	-1776 (3)	7.0 (4)
C(25)	3594 (6)	182 (6)	-868 (2)	3.4 (2)
O(26)	2321 (4)	303 (5)	-927 (2)	4.5 (2)
N(27)	4223 (5)	-1027 (6)	-912 (2)	3.8 (2)
C(28)	3386 (7)	-2212 (7)	-1055 (2)	4.5 (2)
C(29)	4484 (10)	-3351 (9)	-124 (4)	7.5 (4)
C(30)	5702 (11)	-2897 (11)	-859 (5)	11.0 (5)
C(31)	5711 (7)	-1320 (8)	-843 (3)	5.0 (3)
C(32)	2319 (7)	-2609 (7)	-670 (3)	4.4 (3)
O(33)	1299 (6)	-3292 (6)	-862 (2)	6.6 (2)
O(34)	2441 (5)	-2357 (6)	-250 (2)	5.5 (2)
O(35)	4270 (6)	-908 (6)	324 (2)	7.1 (2)

## Discussion

### Bond lengths and angles

The bond lengths and the valency angles in the structure, given in Table 2, are by and large close to the respective standard values (Vijayan, 1976). A few bonds in the phenyl and the pyrrolidine rings, however, are abnormally short. All these bonds are associated with atoms having high thermal parameters. The shortening of the  $C^\beta-C^\gamma$  bond in the prolyl residue, a feature also observed in several other crystal structures (e.g. Kartha, Ashida & Kakudo, 1974; Karle, 1974), is of particular interest as it appears to be related to the conformational flexibility, albeit limited, of the pyrrolidine ring. Most of the conformational possibilities of the ring can be described as arising from the displacement of  $C^\beta$ ,  $C^\gamma$  or both from the plane defined by the rest of the atoms (Ashida & Kakudo, 1974). These displacements can assume a continuous range of values thus giving rise to the possibility of static disorder in a direction perpendicular to the plane of the ring. This disorder, along with thermal vibrations, could lead to highly anisotropic temperature factors for  $C^\beta$  and  $C^\gamma$ , and a consequent shortening of the  $C^\beta-C^\gamma$  bond. In the present structure, the  $B$  values along the principal axes of the thermal-vibration ellipsoid of  $C^\gamma$  are 20.2, 8.6 and 4.2 Å<sup>2</sup>; the corresponding values for  $C^\beta$  are 10.8, 8.3 and 3.3 Å<sup>2</sup>. In both cases, the major axis is nearly perpendicular to the mean plane of the five-membered ring, the angle between the major axis and the plane normal being 4.6 and 24.2° for  $C^\gamma$  and  $C^\beta$  respectively. The  $C^\beta-C^\gamma$  length corrected for the temperature factor is 1.499 and 1.582 Å, when the vibrations are approximated to 'riding motion' and 'uncorrelated motion' respectively (Busing & Levy, 1964). Although neither approximation corresponds to the real situation, the values do indicate that the shortening of the bond can indeed be explained in terms of the appropriate thermal parameters.

### Side-chain conformation

The prolyl residue assumes a  $C_\alpha-C^\gamma$ -*exo*- $C^\beta$ -*endo* conformation (Ashida & Kakudo, 1974) with  $C^\gamma$  and  $C^\beta$  deviating by -0.291 and 0.124 Å respectively from the plane defined by N,  $C^\alpha$  and  $C^\delta$ . The conformation of the phenylalanine side chain is defined by two dihedral angles  $\chi^1$  and  $\chi^{21}$  (IUPAC-IUB Commission on Biochemical Nomenclature, 1970). The observed value of 174.9° for  $\chi^1$  in the structure corresponds to the second most favourable conformation sterically with the phenyl ring *trans* to the NH group and *gauche* to the CO group. The observed value of  $\chi^{21}$  (81.4°) is in the neighbourhood of 90°, as is the case with most of the aromatic side chains in peptides and proteins (Benedetti, 1977; Bhat, Sasisekharan & Vijayan, 1979).

Table 2. Bond lengths (Å) and bond angles (°) involving non-hydrogen atoms

Standard deviations are given in parentheses. (See Fig. 1 for atom numbering.)

C(1)-C(2)	1.391 (14)	C(1)-C(6)	1.369 (13)
C(2)-C(3)	1.306 (17)	C(3)-C(4)	1.368 (16)
C(4)-C(5)	1.409 (14)	C(5)-C(6)	1.392 (14)
C(6)-C(7)	1.503 (12)	C(7)-O(8)	1.455 (11)
O(8)-C(9)	1.367 (8)	C(9)-O(10)	1.204 (9)
C(9)-N(11)	1.336 (9)	N(11)-C(12)	1.448 (8)
C(12)-C(13)	1.508 (11)	C(12)-C(14)	1.555 (9)
C(14)-O(15)	1.226 (7)	C(14)-N(16)	1.327 (8)
N(16)-C(17)	1.460 (8)	C(17)-C(18)	1.553 (9)
C(17)-C(25)	1.517 (9)	C(18)-C(19)	1.516 (10)
C(19)-C(20)	1.418 (12)	C(19)-C(24)	1.370 (11)
C(20)-C(21)	1.410 (13)	C(21)-C(22)	1.405 (18)
C(22)-C(23)	1.346 (18)	C(23)-C(24)	1.375 (14)
C(25)-O(26)	1.238 (7)	O(25)-N(27)	1.324 (8)
N(27)-C(28)	1.459 (9)	C(28)-C(29)	1.539 (12)
C(28)-C(32)	1.534 (10)	C(29)-C(30)	1.451 (16)
C(30)-C(31)	1.531 (13)	C(32)-O(33)	1.298 (9)
C(32)-O(34)	1.205 (8)		
C(2)-C(1)-C(6)	120 (1)	C(17)-C(18)-C(19)	112.0 (6)
C(1)-C(2)-C(3)	122 (1)	C(18)-C(19)-C(20)	119.9 (7)
C(2)-C(3)-C(4)	122 (1)	C(18)-C(19)-C(24)	121.1 (7)
C(3)-C(4)-C(5)	117 (1)	C(20)-C(19)-C(24)	118.9 (8)
C(4)-C(5)-C(6)	122 (1)	C(19)-C(20)-C(21)	119.9 (8)
C(1)-C(6)-C(5)	116.9 (9)	C(20)-C(21)-C(22)	118.5 (9)
C(1)-C(6)-C(7)	122.0 (8)	C(21)-C(22)-C(23)	120 (1)
C(5)-C(6)-C(7)	121.1 (8)	C(22)-C(23)-C(24)	122 (1)
C(6)-C(7)-O(8)	108.1 (8)	C(23)-C(24)-C(25)	120 (1)
C(7)-O(8)-C(9)	116.2 (6)	C(17)-C(25)-O(26)	119.0 (5)
O(8)-C(9)-O(10)	123.6 (7)	C(17)-C(25)-N(27)	119.5 (5)
O(8)-C(9)-N(11)	109.6 (6)	O(26)-C(25)-N(27)	121.4 (6)
O(10)-C(9)-N(11)	126.8 (7)	C(25)-N(27)-C(28)	118.2 (5)
C(9)-N(11)-C(12)	120.4 (6)	C(25)-N(27)-C(31)	127.0 (5)
N(11)-C(12)-C(13)	111.1 (6)	C(28)-N(27)-C(31)	114.7 (5)
N(11)-C(12)-C(14)	109.2 (5)	N(27)-C(28)-C(29)	102.9 (6)
C(13)-C(12)-C(14)	109.7 (5)	N(27)-C(28)-C(32)	111.8 (5)
C(12)-C(14)-O(15)	120.9 (6)	C(28)-C(29)-C(30)	105.7 (8)
C(12)-C(14)-N(16)	115.1 (6)	C(29)-C(30)-C(31)	108.9 (9)
O(15)-C(14)-N(16)	123.8 (6)	N(27)-C(31)-C(30)	100.6 (7)
C(14)-N(16)-C(17)	121.5 (5)	C(28)-C(32)-O(33)	110.0 (6)
N(16)-C(17)-C(18)	110.4 (5)	C(28)-C(32)-O(34)	124.6 (6)
N(16)-C(17)-C(25)	108.9 (5)	O(33)-C(32)-O(34)	125.3 (6)
C(18)-C(17)-C(25)	110.1 (5)		

### Backbone conformation and the intramolecular water bridge

A perspective view of the molecule is shown in Fig. 1. The lone water molecule in the structure, which bridges the two ends of the molecule through hydrogen bonds with the carbonyl O in the benzyloxycarbonyl group and one of the O atoms in the terminal carboxyl group, is also shown in the figure. The dihedral angles which define the main-chain conformation (IUPAC-IUB Commission on Biochemical Nomenclature, 1970) are listed in Table 3. All the peptide groups in the molecule are *trans* and nearly planar with 5.7° as the maximum value of  $\Delta\omega$ . The  $\phi$ ,  $\psi$  values of the prolyl residue may be considered as corresponding to those for the

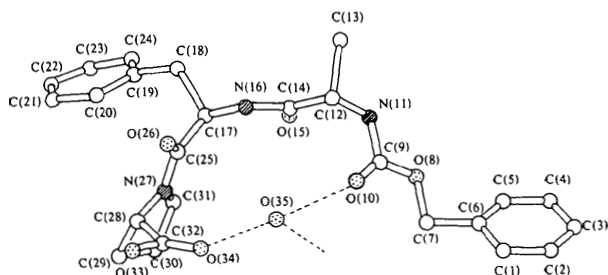


Fig. 1. A perspective view of the molecule. The internal water bridge is also shown.

Table 3. Main-chain torsion angles ( $^{\circ}$ )

E.s.d.'s are  $\text{ca } 0.6^{\circ}$ .

	$i = 1$	$i = 2$	$i = 3$
$\omega_i$	174.7	-174.3	
$\varphi_i$	-83.6	132.6	-66.9
$\psi_i$	147.0	-122.5	157.1*

\* This torsion angle is defined by the atoms N(27), C(28), C(32) and O(33).

polyproline II structure whereas those of the phenylalanyl residue are those appropriate for the  $\beta$  structure. In spite of its participation in an antiparallel sheet in the crystal structure (see below), the alanyl residue has  $\varphi$ ,  $\psi$  values substantially different from those for an ideal  $\beta$  structure, presumably on account of the distortions introduced by the formation of the internal water bridge.

Geometrical features of the hydration of peptides in crystal structures have been studied recently by Yang, Brown & Kopple (1979). This study indicates that intramolecular water bridges have been observed so far only in cyclic peptides [even in cyclic peptides, the water molecule appears to play an important role in determining peptide conformation only in antamanide (Karle, 1977)]. Thus the present structure constitutes the first example of an internal water bridge in a linear peptide. The role of the water molecule here appears to be the initiation of chain-reversal, which is often achieved in peptide chains through the formation of a  $\beta$ -turn stabilized by a 4 $\rightarrow$ 1 hydrogen bond.

Bound water molecules are believed to play an important role in maintaining the structural integrity of proteins, probably by contributing to the local stabilization of conformation (Finney, 1977). It has been noticed in protein structures that more water molecules bind to the main chain than to side chains. Although the main chain has an equal number of CO and NH groups, more water molecules are known to bind to the former than to the latter in proteins (Finney, 1977) as well as in small peptides (Yang *et al.*, 1979). Thus the main-chain CO groups are the most important

loci for water binding in polypeptide chains. In this context, the present structure provides a possible model, at atomic resolution, of a structurally important intramolecular water bridge involving two main-chain CO groups.

### Crystal structure and hydrogen bonding

The arrangement of the molecules in the crystal structure is shown in Fig. 2. The parameters of the hydrogen bonds, which along with van der Waals interactions stabilize the crystal structure, are given in Table 4. An interesting feature of the structure is the presence of an antiparallel hydrogen-bonded sheet involving the alanyl and the phenylalanyl residues. This sheet extends over the entire crystal along the direction of the hydrogen bonds. The only other crystal structure in which such an arrangement has been observed appears to be that of trialanine (Fawcett, Camerman & Camerman, 1975). The sheet accounts for two crystallographically independent N-H $\cdots$ O hydrogen bonds. In addition, there are three O-H $\cdots$ O hydrogen bonds in the structure, two of which are involved in the internal water bridge referred to earlier. In the third O-H $\cdots$ O hydrogen bond, the water molecule accepts a proton from the terminal carboxyl group of an adjacent peptide molecule.

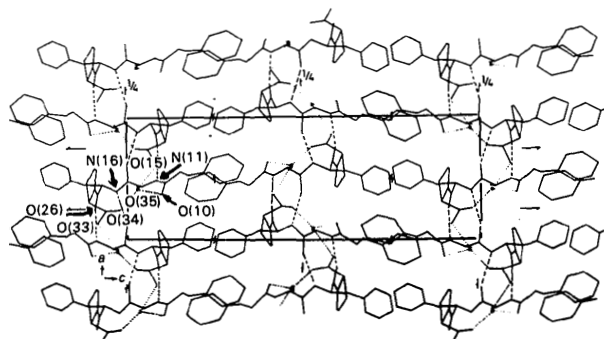


Fig. 2. Crystal structure as viewed along the  $b$  axis. The interchain N-H $\cdots$ O hydrogen bonds and the hydrogen bonds involving the water molecule are indicated by broken lines and dotted lines respectively. Only those atoms involved in hydrogen bonds are numbered.

Table 4. Hydrogen-bond parameters

The standard deviations are given in parentheses.

$D$	$A$	$D \cdots A$	$H-D \cdots A$
O(35) $\cdots$ O(10) <sup>a</sup>		2.765 (8) Å	25 (6) $^{\circ}$
N(11) $\cdots$ O'(26) <sup>b</sup>		2.900 (8)	13 (5)
N(16) $\cdots$ O'(15) <sup>c</sup>		2.944 (8)	14 (4)
O(35) $\cdots$ O(34) <sup>a</sup>		2.761 (8)	15 (6)
O(33) $\cdots$ O'(35) <sup>c</sup>		2.578 (9)	10 (6)

Symmetry code: (a)  $x, y, z$ ; (b)  $\frac{1}{2} + x, \frac{1}{2} - y, -z$ ; (c)  $\frac{1}{2} + x - 1, \frac{1}{2} - y, -z$ .



Another interesting feature of the crystal structure is the spatial segregation of non-polar and polar groups. In fact, they form alternating layers perpendicular to the *c* axis, one layer made up of the phenyl and the pyrrolidine rings and the other consisting largely of peptide units and water molecules.

The authors thank the University Grants Commission, India, and the Indian National Science Academy for financial assistance.

#### References

- ASHIDA, T. & KAKUDO, M. (1974). *Bull. Chem. Soc. Jpn.* **47**, 1129–1133.
- BENEDETTI, E. (1977). *Peptides, Proceedings of the Fifth American Peptide Symposium*, edited by M. GOODMAN & J. MEINHOFER, pp. 257–273. New York: John Wiley.
- BHAT, T. N., SASISEKHARAN, V. & VIJAYAN, M. (1979). *Int. J. Pept. Protein Res.* **13**, 170–184.
- BUSING, W. R. & LEVY, H. A. (1964) *Acta Cryst.* **17**, 142–146.
- CROMER, D. T. & WABER, J. T. (1965) *Acta Cryst.* **18**, 104–109.
- FAWCETT, J. K., CAMERMAN, N. & CAMERMAN, A. (1975). *Acta Cryst.* **B31**, 658–665.
- FINNEY, J. L. (1977). *Philos. Trans. R. Soc. London Ser. B.* **278**, 3–32.
- GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* **A27**, 368–376.
- HAMILTON, W. C. (1959). *Acta Cryst.* **12**, 609–610.
- IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *Biochemistry*, **9**, 3471–3479.
- KARLE, I. L. (1974). *Biochemistry*, **13**, 2155–2162.
- KARLE, I. L. (1977). *J. Am. Chem. Soc.* **99**, 5152–5157.
- KARTHA, G., ASHIDA, T. & KAKUDO, M. (1974). *Acta Cryst.* **B30**, 1861–1866.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.
- VIJAYAN, M. (1976). *Handbook of Biochemistry and Molecular Biology, Proteins*, Vol. 11, edited by G. D. FASMAN, pp. 742–759. Cleveland, Ohio: Chemical Rubber Co.
- YANG, C. H., BROWN, J. N. & KOPPLE, K. D. (1979). *Int. J. Pept. Protein Res.* **14**, 12–20.