Acta Cryst. (1979). B35, 110–114

The Crystal and Molecular Structure of *tert*-Butyloxycarbonyl-L-prolyl-L-prolylglycinamide

BY ISAO TANAKA AND TAMAICHI ASHIDA

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

AND YASUTSUGU SHIMONISHI AND MASAO KAKUDO

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

(*Received* 2 *June* 1978; *accepted* 8 *August* 1978)

Abstract

The crystal structure of the title compound has been determined by X-ray methods. The space group is $P2_1$ with $a = 10{\cdot}243$ (1), $b = 9{\cdot}941$ (4), $c = 10{\cdot}550$ (1) Å, $\beta = 114{\cdot}92$ (1)° and $Z = 2$. The structure was solved by a direct method and the final R value was 0·041. The molecule consists of the two well-known conformations, a polyproline II type structure at the N-terminal half of the molecule and a β-turn structure at the C-terminal half with amide group as a H donor. A comparison of the main-chain structures with several linear oligopeptides which fold into a β-turn is made.

Introduction

Several oligopeptides, such as *p*-bromobenzyloxycarbonyl(Z)-Gly-Pro-Leu-Gly-OH (Ueki, Ashida, Kakudo, Sasada & Katsube, 1969), Z(*o*-Br)-Gly-Pro-Leu-Gly-Pro-OH (Ueki, Bando, Ashida & Kakudo, 1971) and Z-Gly-Pro-Leu-Gly-Pro-OH (Bando, Tanaka, Ashida & Kakudo, 1973) have been investigated by X-ray methods in connection with the substrate specificity of the enzymatic reaction by collagenase, and were shown to have the same compact conformation, *i.e.* a β-turn structure by the convention of Venkatachalam (1968). Furthermore, recently *S*-benzyl-Cys-Pro-Leu-Gly-NH_2 (Rudko & Low, 1975) and *tert*-butyloxycarbonyl(Boc)-Pro-Leu-Gly-OH (Ashida, Tanaka, Shimonishi & Kakudo, 1977) were also shown to fold into a β-turn structure.

Most of the linear peptides folded into a β-turn structure with the 4 → 1 intramolecular hydrogen bond in the crystalline state contain a proline residue at the second position of the turn. Protein structure analyses at atomic resolution have shown a tendency for proline to play an important role at the second position when the peptide folds into a β-turn (Crawford, Lipscomb & Schellman, 1973). These facts stimulated us to investigate more precisely the β-turn structure containing proline in this position.

Boc-Pro-Pro-Gly-NH_2 was synthesized and its structure investigated by X-ray methods. Two different types of folding for this peptide were expected, the first that of the collagen-like structure, the other that of a β-turn structure with the amide group as the donor in the intramolecular hydrogen bond.

Experimental

tert-Butyloxycarbonyl-L-prolyl-L-prolylglycinamide was prepared by the amination of the *p*-nitrophenyl ester previously obtained by the dehydration reaction of *tert*-butyloxycarbonyl-L-prolyl-L-prolylglycine with *p*-nitrophenol, and crystallized from an ethanol solution.

Crystal data

$C_{17}H_{28}N_4O_5$, $M_r = 404{\cdot}48$; m.p. = 489–490 K; $[\alpha]_D^{24} = -122{\cdot}2°$ (c 1·101, AcOH), space group $P2_1$, $a = 10{\cdot}243$ (1), $b = 9{\cdot}941$ (4), $c = 10{\cdot}550$ (1) Å, $\beta = 114{\cdot}92$ (1)°, $D_m = 1{\cdot}27$, $D_c = 1{\cdot}29$ Mg m^{-3} (for $Z = 2$).

A 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 values of high-angle reflections. The intensity data were collected on a computer-controlled Hilger & Watts four-circle diffractometer using Ni-filtered Cu $K\alpha$ radiation. The θ–2θ step-scanning method was adopted. The symmetrical A setting for the reflections with $0 < 2\theta < 100°$ and fixed χ setting (Arndt & Willis, 1966) 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.

0567-7408/79/010110-05$01.00

Structure determination

The structure was solved by the direct method using the program *MULTAN* (Germain, Main & Woolfson, 1971). All the non-hydrogen atoms were located on the *E* map. The structure was refined by the block-diagonal least-squares method (*HBLS* V: Ashida, 1973). The function minimized was $\sum w(|F_o| - |F_c|)^2$, where weights w were unity for all the non-zero reflections and zero for zero reflections. All the H atoms found on the ΔF synthesis map were included in further refinement. At the final stage, an extinction correction was applied to a few seriously affected reflections. The final R value was 0·041. The final positional parameters are listed in Tables 1 and 2.* All the atomic scattering factors were taken from *International Tables for X-ray Crystallography* (1974).

Discussion

The crystal structure projected along the *b* axis is given in Fig. 1. Fig. 2 illustrates an *ORTEP* (Johnson, 1965) drawing of the molecule with thermal ellipsoids enclosing 50% probability. The bond lengths and angles, and the torsion angles are shown in Figs. 3 and 4.

The most striking feature concerning the main-chain conformation is that the subsequent prolyl residues take different conformations, namely an extended (polyproline II type) conformation for the first prolyl residue Pro(1) ($\varphi = -60{\cdot}9$, $\psi = 156{\cdot}3°$) and a folded (β-turn type) conformation for the second prolyl residue Pro(2) ($\varphi = -64{\cdot}9$, $\psi = -23{\cdot}0°$). In molecules having consecutive prolyl residues, it is impossible for the first

* Lists of structure factors and thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33873 (12 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 1. *The atomic positional parameters* ($\times 10^4$) *with their e.s.d.'s in parentheses*

	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 2. *Hydrogen positional parameters* ($\times 10^3$) *with their e.s.d.'s in parentheses*

	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)

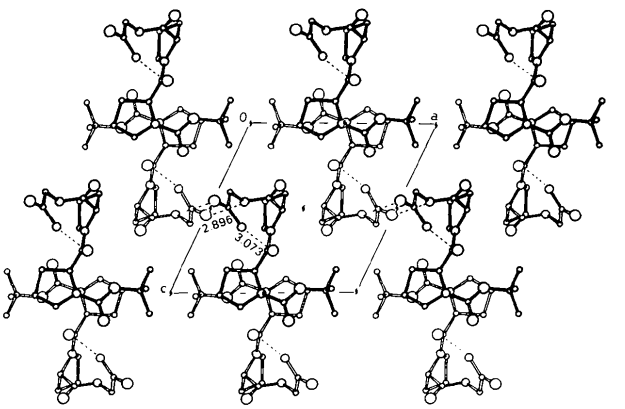

Fig. 1. The crystal structure viewed along the *b* axis. Both intra- (3·073 Å) and intermolecular (2·896 Å) hydrogen bonds are shown with dashed lines.

to take the folded conformation because of the steric hindrance of the second pyrrolidine ring. This is consistent with the extended conformation of the Aoc-Pro_3-OH (Kartha, Ashida & Kakudo, 1974), Boc-Pro_4-OBzl (Matsuzaki, 1974) and polyproline II (Sasisekharan, 1959). On the other hand the folded conformation for the second prolyl residue in the present peptide is not only free from steric hindrance but also largely stabilized by the 4 → 1 hydrogen bond between the first prolyl and the C-terminal amide group.

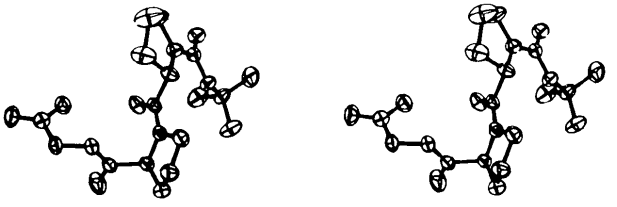

Fig. 2. Stereodrawing of the molecule with thermal ellipsoids drawn to enclose 50% probability.

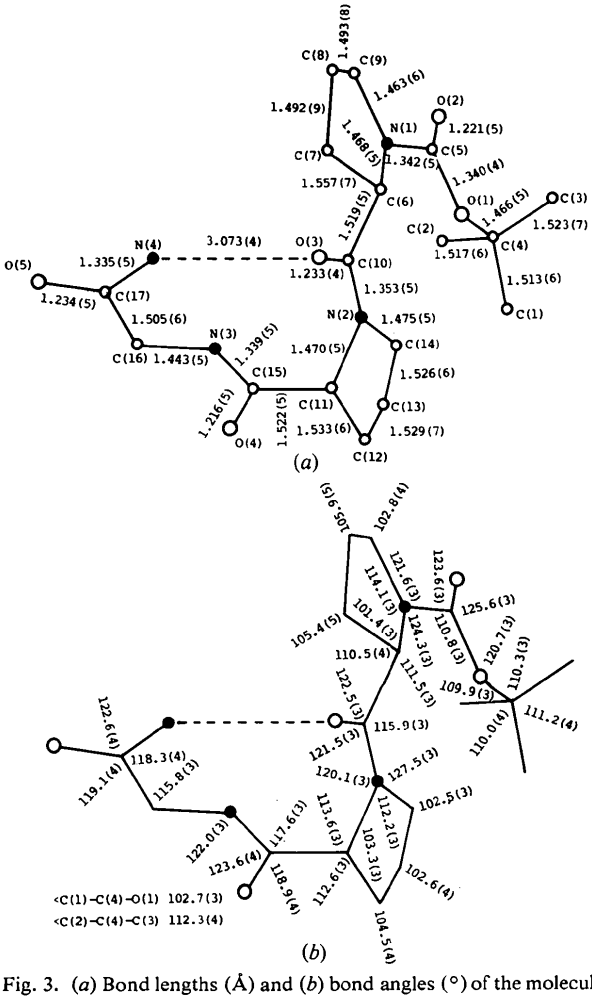

Fig. 3. (*a*) Bond lengths (Å) and (*b*) bond angles (°) of the molecule.

The NC$^{\alpha}$C′ angles, 111·5° for Pro(1) and 113·6° for Pro(2), are slightly but obviously different. A widening of the latter angle is caused by the 4 → 1 hydrogen-bond formation as pointed out by Ashida *et al.* (1977). Following the notation of the pyrrolidine ring conformation (Ashida & Kakudo, 1974), the rings are C_s–C^{γ}-*exo* for Pro(1) and C_s–C^{γ}-*endo* for Pro(2). As *A* (C^{γ}-*exo*) and *B* (C^{γ}-*endo*) conformations (Balasubramanian, Lakshminarayanan, Sabesan, Tegoni, Venkatesan & Ramachandran, 1971) for pyrrolidine rings are found in both main-chain conformations, polyproline II type and β-turn type [see for example Ashida *et al.* (1977)], a close relation between the main-chain conformation and the pyrrolidine-ring conformation seems unlikely to exist.

There are two features of the β-turn of this molecule. The first is that the Pro-Gly sequence for the second and third positions of the turn takes the type I folding, not type II. As suggested by Venkatachalam (1968) both I and II types are permissible for the peptides having a glycyl residue at the third position and type II folding has been found in the tripeptide H-Pro-Leu-Gly-NH_2 (Reed & Johnson, 1973), as well as in many proteins. Venkatachalam (1968) expected the Pro-Gly sequence to be accommodated in the type II folding, which may be less favorable than the type I because of the steric repulsion between the pyrrolidine ring and the peptide plane. The second feature is the ability of the C-terminal amide group to make the β-turn. This kind of β-turn has commonly been observed in crystal structures such as *S*-benzyl-Cys-Pro-Leu-Gly-NH_2 (Rudko & Low, 1975) and H-Pro-Leu-Gly-NH_2 (Reed & Johnson, 1973), suggesting that it is one of the favorite conformations.

The torsion angles (φ,ψ) at the second and third positions in the linear oligopeptides folded into the β-turn structure are listed in Table 3 and are plotted on the (φ,ψ) map in Fig. 5. It is natural that the rotational flexibility of the second position is smaller than that of the third position because the former positions are

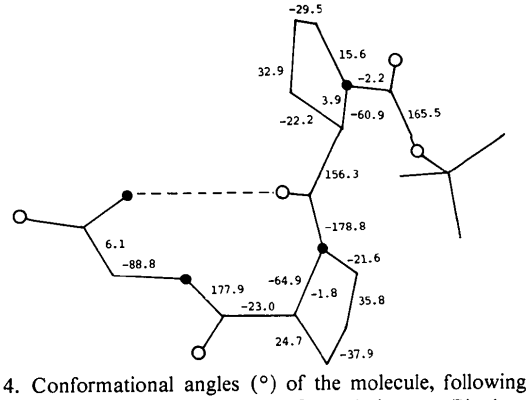

Fig. 4. Conformational angles (°) of the molecule, following the convention of the IUPAC–IUB Commission on Biochemical Nomenclature (1970).

Table 3. *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		6·1
H-Pro-Leu-Gly-NH_2	−61·2	127·8		–

* Some of the torsion angles for these molecules were incorrectly given in the original paper (Rudko & Low, 1975). The angles recalculated from the published atomic coordinates are given here.

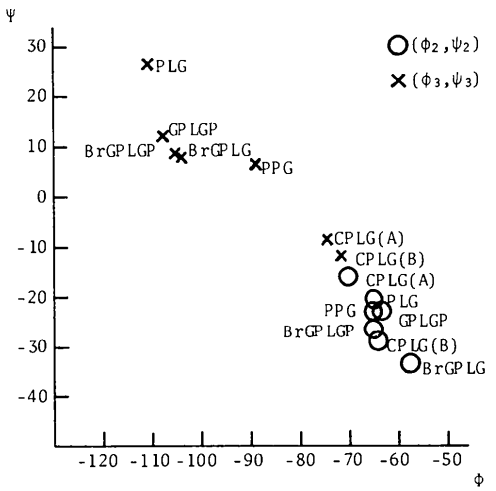

Fig. 5. Ramachandran plot of the main-chain torsion angles of the β-turn. Open circles refer to the angles (φ_2,ψ_2), the torsion angles for the second residue, and crosses refer to the angles (φ_3,ψ_3), the torsion angles for the third residue. PPG(Boc-Pro-Pro-Gly-NH_2), PLG(Boc-Pro-Leu-Gly-OH), GPLGP(Z-Gly-Pro-Leu-Gly-Pro-OH), BrGPLGP[Z(*o*-Br)-Gly-Pro-Leu-Gly-Pro-OH], BrGPLG [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).

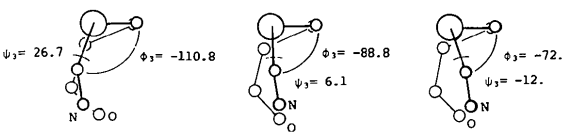

Fig. 6. Newman projection of the β-turn nine-membered ring viewed along the bond C_α—N of the third residue. The molecules are (left to right) Boc-Pro-Leu-Gly-OH, Boc-Pro-Pro-Gly-NH_2 and *S*-Bzl-Cys-Pro-Leu-Gly-NH_2 (*B*).

mostly occupied by proline. Therefore, the following discussion about the β-turn structure is mainly concerned with the third position.

The variety of the torsion angles may result from the crystal-packing requirement, but also may reflect the intrinsic nature of each molecule. 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 probably show the typical β-turn structure for the Pro-*X* system (*X*: any acyl group). Boc-Pro-Leu-Gly-OH and Boc-Pro-Gly-NH_2 are special cases, because the former accommodates the Boc group at the first position and the latter the NH_2 group at the fourth position. 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 O atom of the prolyl residue and the C-terminal amide group. Actually, in *S*-benzyl-Cys-Pro-Leu-Gly-NH_2 (*B*), the subsequent 4 → 1 hydrogen bonds are formed, so the conformation may be referred to the 3_{10} helix rather than the β-turn.

Fig. 5 also illustrates that there is a relation between φ and ψ rotation, which can be explained qualitatively by the Newman projections of three typical molecules in Fig. 6. Although the β-turn structure is considered to be rigid there is still some flexibility without the 4 → 1 hydrogen bond being disturbed.

The authors would like to express their sincere thanks to Professor J. Tanaka of Nagoya University for kindly permitting them to use the diffractometer. Part of the cost of this research was covered by a Scientific Research Grant from the Ministry of Education, for which the authors' thanks are due.

References

ARNDT, U. W. & WILLIS, B. T. M. (1966). *Single Crystal Diffractometry*. Cambridge Univ. Press.

ASHIDA, T. (1973). *UNICS*–Osaka, pp. 55–61. The Computation Center, Osaka Univ.

ASHIDA, T. & KAKUDO, M. (1974). *Bull. Chem. Soc. Jpn*, **47**, 1129–1133.

ASHIDA, T., TANAKA, I., SHIMONISHI, Y. & KAKUDO, M. (1977). *Acta Cryst.* B**33**, 3054–3059.

BALASUBRAMANIAN, R., LAKSHMINARAYANAN, A. V., SABESAN, M. N., TEGONI, G., VENKATESAN, K. & RAMACHANDRAN, G. N. (1971). *Int. J. Protein Res.* **3**, 25.

BANDO, S., TANAKA, N., ASHIDA, T. & KAKUDO, M. (1973). Ann. Meet. Crystallogr. Soc. Japan, Osaka.

CRAWFORD, J. L., LIPSCOMB, W. N. & SCHELLMAN, C. G. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 538–542.
GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.
International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press.
IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *J. Mol. Biol.* **52**, 1–17.
JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.
KARTHA, G., ASHIDA, T. & KAKUDO, M. (1974). *Acta Cryst.* B**30**, 1861–1866.
MATSUZAKI, T. (1974). *Acta Cryst.* B**30**, 1029–1036.
REED, L. L. & JOHNSON, P. L. (1973). *J. Am. Chem. Soc.* **95**, 7523–7524.
RUDKO, A. D. & LOW, B. W. (1975). *Acta Cryst.* B**31**, 713–725.
SASISEKHARAN, V. (1959). *Acta Cryst.* **12**, 897–903.
UEKI, T., ASHIDA, T., KAKUDO, M., SASADA, Y. & KATSUBE, Y. (1969). *Acta Cryst.* B**25**, 1840–1849.
UEKI, T., BANDO, S., ASHIDA, T. & KAKUDO, M. (1971). *Acta Cryst.* B**27**, 2219–2231.
VENKATACHALAM, C. M. (1968). *Biopolymers*, **6**, 1425–1436.

Acta Cryst. (1979). B**35**, 114–118

Crystal Structure of the Neurotensin Tetrapeptide L-Pro-L-Tyr-L-Ile-L-Leu

BY MICHEL COTRAIT, SERGE GEOFFRE, MICHEL HOSPITAL AND GILLES PRECIGOUX

Laboratoire de Cristallographie et de Physique Cristalline associé au CNRS, Université de Bordeaux I, 351 *cours de la Libération*, 33405 *Talence, France*

(*Received* 26 *May* 1978; *accepted* 27 *September* 1978)

Abstract

The crystal structure of a neurotensin fragment, Pro-Tyr-Ile-Leu ($C_{26}H_{40}N_4O_6$), an important biological peptide, has been determined. It crystallizes in space group $P2_1$ with cell parameters $a = 9{\cdot}840\ (4)$, $b = 26{\cdot}750\ (9)$, $c = 5{\cdot}305\ (5)$ Å, $\beta = 92{\cdot}72\ (3)°$ and $Z = 2$. The final R value is 0·062. The molecular conformation of the main chain is defined by the following (φ,ψ) angles: (−54, 169), (−71, −48), (−127, 134) and (−139°, 147°), respectively, for Pro, Tyr, Ile and Leu. There is no intramolecular hydrogen bond; the hydrogen-bond network is rather loose but there are some van der Waals interactions between hydrophobic side chains.

Introduction

Neurotensin is a biologically important peptide recently isolated from bovine hypothalami. Its amino-acid sequence is *p*Glu-Leu-Tyr-Glu-Asn-Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu-OH. It produces hypotension, increases vascular permeability, induces a pain sensation and affects the contractibility of various non-vascular smooth muscles, like kinins. Neurotensin also possesses properties which are not shared by the kinins, like a rapid hyperglycemia (Carraway & Leeman, 1975). It was recently shown that it binds very specifically to synaptic membranes from rat brain (Kitabgi, Carraway, Van Rietschoten, Granier, Morgat, Menez, Leeman & Freychet, 1977).

Partial sequences of neurotensin were synthetized and their biological effects tested. It was shown that the biological action of this peptide resides essentially in its carboxylic terminal group and that the smallest peptide with a still quite important activity is the pentapeptide Arg-Pro-Tyr-Ile-Leu; NH_2-terminal partial sequences of neurotensin as large as the (1–10) decapeptide were found to be ineffectual (Carraway & Leeman, 1976).

Experimental

Synthesis and purification

The tetrapeptide was synthetized according to the Merrifield phase procedure with an automatic peptide synthetizer built in our laboratory. Only Boc amino acids were used; cleavage from the resin by hydrogen fluoride yielded the crude peptide which was purified by chromatography Bio-gel P2 using 0·1 *N* acetic acid as eluant. The fraction corresponding to the major peak of absorption was lyophilized and used to obtain crystals by slow evaporation of a 0·1 *N* acetic acid solution.

Crystal data

The crystals belong to the monoclinic space group $P2_1$. The cell constants were obtained manually from the measurements of ω, χ, and φ Eulerian angles for 10 reflections with a Siemens four-circle diffractometer. The refined parameters are: $a = 9{\cdot}840\ (\pm 0{\cdot}004)$, $b =$

0567-7408/79/010114-05$01.00