

## Structure of Glycyl-L-phenylalanyl-L-phenylalanine Hemihydrate

BY GILLES PRÉCIGOUX, MICHEL COTRAIT AND SERGE GEOFFRE

Laboratoire de Cristallographie et de Physique Cristalline, UA 144 au CNRS, Université de Bordeaux I,  
351 cours de la Libération, 33405 Talence CEDEX, France

(Received 4 June 1985; accepted 27 September 1985)

**Abstract.**  $C_{20}H_{23}N_3O_4 \cdot 0.5H_2O$ ,  $M_r = 378.43$ , monoclinic,  $C2$ ,  $a = 19.983$  (8),  $b = 5.017$  (5),  $c = 20.476$  (9) Å,  $\beta = 110.36$  (2)°,  $V = 1924.6$  Å<sup>3</sup>,  $Z = 4$ ,  $D_m = 1.32$ ,  $D_x = 1.307$  g cm<sup>-3</sup>,  $\lambda(\text{Cu K}\alpha) = 1.5418$  Å,  $\bar{\mu} = 7.8$  cm<sup>-1</sup>,  $F(000) = 804$ ,  $T = 293$  K, final  $R = 0.049$  for 1142 independent reflections. The molecule exists in the crystal as a zwitterion with numerous intermolecular hydrogen bonds. The molecule partly assumes the conformation of a right-handed  $\alpha$  helix. In the  $c$  direction, molecules are hydrogen bonded around water molecules, located on the twofold axis, so giving infinite helicoidal coils.

**Introduction.** The glycyl-L-phenylalanyl-L-phenylalanine (Gly-L-Phe-L-Phe) tripeptide is part of the insulin B chain (23–25 fragment). Moreover, the L-Phe-L-Phe fragment is frequently found in neuropeptides such as substance P and somatostatin. Structural analysis of a sequence containing the L-Phe-L-Phe dipeptide can provide useful information for the understanding of preferential conformations and biological activity of linear peptides enclosing such a dipeptide.

**Experimental.** Synthesis according to Merrifield's (1963) solid-phase method. Automatic synthesizer apparatus. Protected amino acids from the Bachem Company. Chloromethylated polystyrene–1% divinylbenzene as solid phase. Amino acid chain elongation: 30% trifluoroacetic acid deprotection, neutralization by 10% diisopropylethylamide dissolved in dichloromethane and coupling with dicyclohexylcarbodiimide (Sheenan & Ness, 1955). Released from the resin by hydrogen fluoride in the presence of anisole. Purification by gel permeation chromatography (Sephadex G 15) and reverse-phase high-pressure liquid chromatography.

Colourless needles (0.1 × 0.2 × 0.5 mm) were grown by slow cooling of concentrated aqueous solutions.  $D_m$  determined by flotation; Enraf–Nonius CAD-4 diffractometer, graphite monochromator; unit-cell dimensions determined by a least-squares procedure on  $2\theta$  values of 19 reflections with  $\theta > 25^\circ$ ; no absorption or extinction correction applied; intensity data collected in one quarter of the Ewald sphere and  $\sin\theta/\lambda < 0.5$  Å<sup>-1</sup> ( $h: 0 \rightarrow 19$ ,  $k: 0 \rightarrow 4$ ,  $l: -20 \rightarrow 17$ ) with

$\omega-2\theta$  scan,  $\Delta\theta = (0.8 + 0.15 \text{tg}\theta)^\circ$ , detector aperture (1.3 + 1.0 tg $\theta$ ) mm; reflections 427 and 915 monitored during data collection with fluctuations <2%; 1212 independent reflections, 1061 with  $I > 2\sigma(I)$ ; structure solved by direct methods, MULTAN80 (Main, Fiske, Hull, Lessinger, Germain, Declercq & Woolfson, 1980). Except for the O of water, located on a subsequent Fourier map, all C, N and O atoms were identified on the best  $E$  map; structure refinement by least-squares techniques using first  $B_i$ , then  $\beta_{ij}$  anisotropic parameters on a CII-Mini 6/92 computer. H atoms, except those of the water molecule, were located from a difference Fourier synthesis; the refinement was resumed using isotropic temperature factors for the well positioned H atoms;  $\sum w(|F_o| - |F_c|)^2$  minimized with  $w = 1$  when  $|F_o| < p$  and  $\sqrt{w} = p/F_o$  when  $|F_o| > p$ , where  $p = |F_o(\text{max.})/10|^{1/2}$ ;  $(\Delta/\sigma)_{\text{max}} = 0.35$ ; final  $\Delta F$  excursions 0.35 and  $-0.2$  e Å<sup>-3</sup>; scattering factors from *International Tables for X-ray Crystallography* (1974) for C, N, O and from Stewart, Davidson & Simpson (1965) for H. Final  $R = 0.049$ ,  $wR = 0.055$ ,  $S = 1.28$  and  $w = 1/\sigma_f^2$ .

**Discussion.** The final atomic parameters for C, N and O atoms, labelled in accordance with the IUPAC–IUB Commission on Biochemical Nomenclature (1970), are listed in Table 1.\* A projection of the molecular packing is shown in Fig. 1.

The tripeptide exists as a zwitterion, with charged ammonium  $NH_3^+$  and carboxylate  $COO^-$  groups. It crystallizes with an additional water molecule, located on a twofold axis. Bond lengths and angles are given in Table 2; they agree reasonably well with those found for other peptides having a phenylalanyl residue (Benedetti, 1977).

The molecular conformation is defined by the following torsion angles (°) (e.s.d.'s  $\sim 2.5^\circ$ ): Gly:  $\psi = -163$ ,  $\omega = 183$ ; Phe<sub>1</sub>:  $\varphi = -126$ ,  $\psi = -56$ ,  $\omega = 171$ ,  $\chi_1 = -176$ ,  $\chi_2 = 86$ ; Phe<sub>2</sub>:  $\varphi = -78$ ,  $\psi = -20$ ,  $\chi_1 = -71$  and  $\chi_2 = 95$ .

\* Lists of structure factors, anisotropic thermal parameters and H-atom parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 42543 (10 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Both phenylalanyl residues assume a right-handed  $\alpha$ -helix conformation;  $\chi_2$  values are in the neighbourhood of  $90^\circ$ , as is the case for aromatic residues in peptides and proteins (Benedetti, Morelli, Nemethy & Scheraga, 1983). If the  $\chi_1$  value for Phe<sub>2</sub> corresponds to the most frequent rotamer, that for Phe<sub>1</sub> corresponds to the second most favourable rotamer. In such a conformation, phenylalanine side chains are approximately perpendicular ( $104.6^\circ$ ). At Gly level, the

Table 1. Atomic coordinates ( $\times 10^4$ ) with standard deviations and equivalent isotropic temperature factors,  $B_{eq}$ , for non-hydrogen atoms

$$B_{eq} = \frac{4}{3} \sum_i \sum_j \beta_{ij} a_i \cdot a_j$$

	x	y	z	$B_{eq}(\text{\AA}^2)$
<b>Glycyl</b>				
N	6785 (3)	927 (1)	365 (2)	3.4
C $^\alpha$	6445 (3)	982 (1)	884 (3)	3.6
C'	6515 (3)	749 (1)	1354 (3)	2.5
O	6961 (2)	576 (1)	1408 (2)	4.1
<b>Phenylalanyl-1</b>				
N	6067 (2)	749 (1)	1701 (2)	3.0
C $^\alpha$	6067 (3)	547 (1)	2212 (3)	5.6
C'	5323 (3)	425 (1)	2032 (3)	2.6
O	5238 (2)	180 (1)	1956 (2)	3.7
C $^\beta$	6327 (3)	666 (2)	2950 (3)	3.4
C $^\gamma$	6391 (3)	464 (1)	3516 (3)	3.3
C $^{\delta 1}$	5844 (4)	423 (2)	3762 (3)	4.8
C $^{\delta 2}$	7004 (3)	314 (2)	3795 (3)	4.3
C $^{\epsilon 1}$	5912 (4)	238 (2)	4278 (4)	5.7
C $^{\epsilon 2}$	7063 (4)	129 (2)	4310 (3)	5.0
C $^{\epsilon}$	6515 (4)	94 (2)	4549 (4)	5.6
<b>Phenylalanyl-2</b>				
N	4776 (2)	586 (1)	1960 (2)	2.6
C $^\alpha$	4048 (3)	491 (1)	1691 (3)	2.8
C'	3786 (3)	461 (2)	907 (3)	3.3
O'	4101 (2)	584 (1)	575 (2)	5.1
O''	3239 (2)	319 (1)	641 (2)	4.8
C $^\beta$	3563 (3)	679 (1)	1917 (3)	3.6
C $^\gamma$	3683 (3)	660 (1)	2681 (3)	3.2
C $^{\delta 1}$	3355 (4)	463 (2)	2917 (4)	4.8
C $^{\delta 2}$	4128 (4)	830 (2)	3152 (4)	5.0
C $^{\epsilon 1}$	3480 (5)	433 (2)	3629 (4)	6.5
C $^{\epsilon 2}$	4254 (4)	800 (2)	3861 (4)	6.3
C $^{\epsilon}$	3922 (4)	600 (2)	4090 (4)	5.8
W	5000	312 (2)	0000	7.6

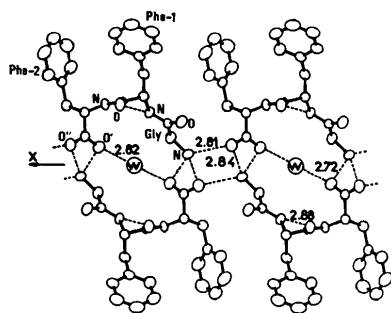


Fig. 1. Projection of the Gly-Phe-Phe structure. Hydrogen bonds are represented by dotted lines. The  $b$  axis points upward. O and N atoms are indicated.

$N-C_\alpha$  bond lies approximately in the Gly-Phe<sub>1</sub> peptide bond plane as observed in many other oligopeptides like Gly-Tyr (Cotrait & Bideau, 1974) or Gly-Gly-Phe-Leu (Prange & Pascard, 1979) in which Gly is the N-terminal residue.

All the N and O atoms, except the O(Gly) atom, are involved in the following hydrogen bonds:

	$d$ (Å)	$\alpha$ ( $^\circ$ )
$NH_3^+ \cdots COO^-$ (III)	2.81	159
$NH_3^+ \cdots COO^-$ (II + a)	2.72	166
$NH_3^+ \cdots COO^-$ (II + a + b)	2.84	168
$O(H) \cdots COO^-$ (II + a)	2.82	
$NH(Phe_1) \cdots O(Phe_2)$ (I + b)	2.88	148
$NH(Phe_2) \cdots O(Phe_1)$ (I + b)	3.12	158

where II =  $-x, y, z$  and III =  $\frac{1}{2} + x, \frac{1}{2} + y, z$ ,  $d$  is the  $N \cdots O$  distance and  $\alpha$  the  $N-H \cdots O$  angle (e.s.d.'s  $\sim 0.01 \text{ \AA}$  and  $1.0^\circ$ , respectively).

The packing of the peptides around water molecules (see Fig. 2) can be described as an endless helicoidal structure with a strand made by the succession of peptide backbones and strong hydrogen bonds between the  $NH_3^+$  and  $COO^-$  groups (2.72 Å). The hydrogen interactions between  $O(Phe_2)$ ,  $NH(Phe_1)$  and  $NH(Phe_2)$  contribute to the stability of the helix thus formed.

Table 2. Bond lengths (Å) and angles ( $^\circ$ ) with standard deviations

<b>Glycyl</b>			
N-C $^\alpha$	1.473 (8)	C'-O	1.221 (7)
C $^\alpha$ -C'	1.49 (1)	C'-N	1.321 (7)
N-C $^\alpha$ -C'	110.7 (6)	O-C'-N	124.3 (6)
C $^\alpha$ -C'-O	121.2 (6)	C'-N-C $^\alpha$	123.6 (6)
C $^\alpha$ -C'-N	114.6 (6)		
<b>Phenylalanyl-1</b>			
N-C $^\alpha$	1.458 (8)	C $^\gamma$ -C $^{\delta 1}$	1.37 (1)
C $^\alpha$ -C $^\beta$	1.530 (9)	C $^\gamma$ -C $^{\delta 2}$	1.38 (1)
C'-O	1.239 (7)	C $^{\delta 1}$ -C $^{\epsilon 1}$	1.38 (1)
C'-N	1.325 (8)	C $^{\delta 2}$ -C $^{\epsilon 2}$	1.38 (1)
C $^\alpha$ -C $^\beta$	1.54 (1)	C $^{\epsilon 1}$ -C $^{\epsilon}$	1.35 (1)
C $^\beta$ -C $^\gamma$	1.51 (1)	C $^{\epsilon 2}$ -C $^{\epsilon}$	1.36 (1)
N-C $^\alpha$ -C'	110.4 (5)	C $^\beta$ -C $^\gamma$ -C $^{\epsilon 2}$	121.2 (7)
C $^\alpha$ -C'-O	120.4 (6)	C $^{\epsilon 1}$ -C $^\gamma$ -C $^{\epsilon 2}$	118.3 (7)
C $^\alpha$ -C'-N	118.5 (6)	C $^\gamma$ -C $^{\delta 1}$ -C $^{\epsilon 1}$	120.3 (8)
O-C'-N	121.1 (6)	C $^\gamma$ -C $^{\delta 2}$ -C $^{\epsilon 2}$	120.7 (7)
C'-N-C $^\alpha$	121.7 (5)	C $^{\delta 1}$ -C $^{\epsilon 1}$ -C $^{\epsilon}$	120.9 (8)
N-C $^\alpha$ -C $^\beta$	110.6 (5)	C $^{\delta 2}$ -C $^{\epsilon 2}$ -C $^{\epsilon}$	119.9 (8)
C $^\alpha$ -C $^\beta$ -C $^\gamma$	113.9 (6)	C $^{\epsilon 1}$ -C $^{\epsilon}$ -C $^{\epsilon 2}$	119.9 (8)
C $^\beta$ -C $^\gamma$ -C $^{\epsilon 1}$	120.5 (6)	C $^\beta$ -C $^\alpha$ -C'	111.1 (5)
<b>Phenylalanyl-2</b>			
N-C $^\alpha$	1.443 (8)	C $^\gamma$ -C $^{\delta 1}$	1.36 (1)
C $^\alpha$ -C'	1.51 (1)	C $^\gamma$ -C $^{\delta 2}$	1.36 (1)
C'-O'	1.258 (9)	C $^{\delta 1}$ -C $^{\epsilon 1}$	1.40 (1)
C'-O''	1.240 (9)	C $^{\delta 2}$ -C $^{\epsilon 2}$	1.39 (1)
C $^\alpha$ -C $^\beta$	1.54 (1)	C $^{\epsilon 1}$ -C $^{\epsilon}$	1.34 (1)
C $^\beta$ -C $^\gamma$	1.50 (1)	C $^{\epsilon 2}$ -C $^{\epsilon}$	1.37 (1)
N-C $^\alpha$ -C'	111.6 (5)	C $^\beta$ -C $^\gamma$ -C $^{\epsilon 2}$	121.8 (7)
C $^\alpha$ -C'-O'	118.5 (6)	C $^{\epsilon 1}$ -C $^\gamma$ -C $^{\epsilon 2}$	118.6 (7)
C $^\alpha$ -C'-O''	116.3 (6)	C $^\gamma$ -C $^{\delta 1}$ -C $^{\epsilon 1}$	120.8 (8)
O'-C'-O''	125.1 (7)	C $^\gamma$ -C $^{\delta 2}$ -C $^{\epsilon 2}$	120.5 (8)
N-C $^\alpha$ -C $^\beta$	110.2 (5)	C $^{\delta 1}$ -C $^{\epsilon 1}$ -C $^{\epsilon}$	120.2 (9)
C $^\alpha$ -C $^\beta$ -C $^\gamma$	112.7 (6)	C $^{\delta 2}$ -C $^{\epsilon 2}$ -C $^{\epsilon}$	120.1 (8)
C $^\beta$ -C $^\gamma$ -C $^{\epsilon 1}$	119.5 (7)	C $^{\epsilon 1}$ -C $^{\epsilon}$ -C $^{\epsilon 2}$	119.8 (8)

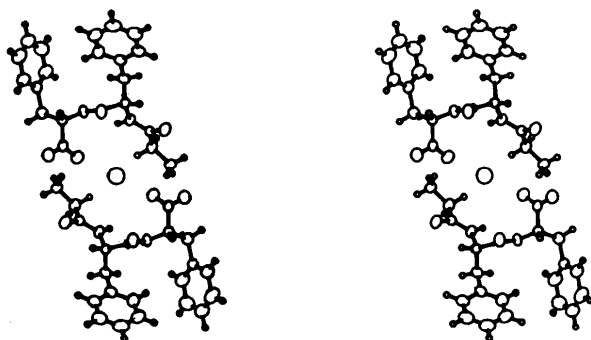


Fig. 2. Stereoview showing the packing of the peptides around the water molecules.

One water molecule is bonded to the peptide carboxyl groups related by the twofold axis and acts as proton donor for both hydrogen bonds.

In the crystallographic 'a' direction, these helices are related by a twofold axis and hydrogen bonded at the level of the COO<sup>-</sup> and NH<sub>3</sub><sup>+</sup> polar groups. In the 'c' direction, the stacking of the helices is only stabilized by

weak van der Waals interactions between atoms of the hydrophobic phenyl groups.

#### References

- BENEDETTI, E. (1977). *Peptides* (Proc. Fifth Am. Pept. Symp.), edited by GOODMAN & MEINHOFER, pp. 257–273. New York: John Wiley.
- BENEDETTI, E., MORELLI, G., NEMETHY, G. & SCHERAGA, H. A. (1983). *Int. J. Pept. Protein Res.* **22**, 1–15.
- COTRAIT, M. & BIDEAU, J. P. (1974). *Acta Cryst.* **B30**, 1024–1028. *International Tables for X-ray Crystallography* (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor D. Reidel, Dordrecht.)
- IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *Biochemistry*, **9**, 3471–3478.
- MAIN, P., FISKE, S. J., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J.-P. & WOOLFSON, M. M. (1980). *MULTAN80. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univs. of York, England, and Louvain, Belgium.
- MERRIFIELD, R. B. (1963). *J. Am. Chem. Soc.* **82**, 2149–2154.
- PRANGE, T. & PASCARD, C. (1979). *Acta Cryst.* **B35**, 1812–1819.
- SHEENAN, J. C. & NESS, G. P. (1955). *J. Am. Chem. Soc.* **77**, 1067–1068.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

*Acta Cryst.* (1986). **C42**, 317–319

### *rel*-(1*aR*,2*R*,4*aR*,7*S*,7*aR*,7*bR*)-Decahydro-7-hydroxy-1,1,7-trimethyl-4-methylene-1*H*-cycloprop[*e*]azulen-2-yl Benzoate

BY PREM C. JAIN\* AND WILLIAM H. WATSON†

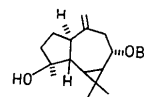
*FASTBIOS Laboratory, Department of Chemistry, Texas Christian University, Fort Worth, Texas 76129, USA*

(Received 29 May 1985; accepted 20 September 1985)

**Abstract.** C<sub>22</sub>H<sub>28</sub>O<sub>3</sub>, *M<sub>r</sub>* = 340.47, monoclinic, *P*2<sub>1</sub>, *a* = 10.957 (4), *b* = 9.170 (2), *c* = 9.804 (3) Å, β = 93.23 (3)°, *V* = 983.5 (5) Å<sup>3</sup>, *Z* = 2, *D<sub>x</sub>* = 1.150, *D<sub>m</sub>*(floatation) = 1.148 g cm<sup>-3</sup>, λ(Cu *Kα*) = 1.54178 Å, μ = 6.21 cm<sup>-1</sup>, *F*(000) = 368, *T* = 293 K, *R* = 0.035 for 1295 unique observed reflections. The compound is an aromadendrane-type of sesquiterpene isolated from *Ferulago antiochia* (Apiaceae). It contains a seven-membered ring in a chair conformation fused to a three-membered ring and to a five-membered ring in a half-chair conformation. The benzoate ester exhibits an extended conformation with the benzene ring planar to within 0.020 (11) Å.

**Introduction.** As part of a phytochemical screening of the genus *Ferula* or *Ferulago* (family Apiaceae) (Miski, Ulubelen & Mabry, 1983; Miski, Ulubelen, Mabry,

Vickovic, Watson & Holub, 1984), a newly described species *Ferulago antiochia* Saya & Miski (Saya & Miski, 1985) from Hatay province in the Southern region of Turkey has been investigated. In addition to several germacrane-type sesquiterpene esters, the petroleum ether extract of the roots of *F. antiochia* afforded a new aromadendrene ester (1). Spectral data indicated (1) to be 8β-benzoyloxyspathulenol; this structure has been confirmed by X-ray crystallography.



(1)

**Experimental.** A large colorless crystal was cleaved and air-ground to dimensions 0.48 × 0.49 × 0.36 mm, mounted along the *c* axis on a Syntex *P*2<sub>1</sub> diffractometer and data collected using a θ:2θ scan, 2θ<sub>max</sub>

\* On leave from Kurukshetra University, Kurukshetra, India.

† To whom correspondence should be addressed.