Acta Cryst. (1982). B38, 1758–1764

The Structure of the Hydrochloride Salt of Trp-Met-Asp-Phe-NH_2.CH_3OH.-0·5$C_2H_5OC_2H_5$, the C-terminal Tetrapeptide Amide of Gastrin*

BY WILLIAM B. T. CRUSE, ERNST EGERT, M. A. VISWAMITRA† AND OLGA KENNARD

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

(*Received* 14 *September* 1981; *accepted* 4 *January* 1982)

Abstract

$C_{29}H_{37}N_6O_6S^+.Cl^-.CH_4O.0{\cdot}5C_4H_{10}O$, $M_r = 702{\cdot}3$, crystallizes in space group $P2_1$ with $a = 14{\cdot}391$ (3), $b = 23{\cdot}757$ (4), $c = 11{\cdot}362$ (3) Å, $\beta = 103{\cdot}48$ (2)°, $U = 3777{\cdot}5$ (7) Å^3, $Z = 4$, $D_c = 1{\cdot}235$ Mg m^{-3}, μ(Cu $K\alpha$) = 1·74 mm^{-1}. The structure was refined to $R = 0{\cdot}082$ for 3143 reflections. The two peptide molecules in the asymmetric unit have different conformations: one is rather extended, whereas the other has the phenyl group bent back towards the backbone enabling the formation of an intramolecular hydrogen bond between the terminal amide group and the Asp side chain. The H-bonded peptide dimer is the repeat unit of an infinite antiparallel β-pleated sheet extending along **c**. The individual β-sheets are connected by a dense network of H bonds with participation of the Cl^- ions. The packing of the side chains, which extend mainly along **a**, is stabilized by both components of the solvent mixture: the methanol molecules accept H bonds from Trp and Asp, while diethyl ether fills a large pocket enclosed by the hydrophobic peptide residues. Although the observed peptide conformations probably do not resemble the 'active' form, the structure demonstrates the ability of linear peptides to crystallize as β-sheets provided the solvent molecules have the appropriate size, shape and polarity to solve the considerable packing problems.

Introduction

The gastrin family of peptide hormones comprises the linear polypeptides gastrin and cholecystokinin (pancreozymin), found in the gut, and the decapeptide caerulein, isolated from the skin of frogs. All are characterized by a common C-terminal tetrapeptide amide sequence Trp-Met-Asp-Phe-NH_2 which by itself possesses the whole range of the biological activity of gastrin but at a reduced level (Tracy & Gregory, 1964). Gastrin, the most studied member of the trio, is best known for its ability to stimulate gastric-acid secretion, although it affects nearly every major secretory, absorptive and smooth-muscle activity of the digestive tract.

In an effort to relate the primary structure with hormone activity, Morley (1968) synthesized a large number of analogues of the C-terminal tetrapeptide amide. His work showed that a penultimate Asp residue and a terminal amide group were essential for gastric-secretory activity and that the three remaining residues could be modified as long as they retained their general size, shape and hydrophobic character. Although no direct conformational description of the gastrin receptor exists, schematic models (Morley, 1968) have provided a basis for the development of the 'anti-gastrin' compound *N*-butoxycarbonyl-Gly-Trp-Met-Gly-NH_2. Efforts have been made to interpret the nature of the receptor by conformational analysis of Trp-Met-Asp-Phe-NH_2 using NMR spectra from dimethyl sulphoxide solution (Feeney, Roberts, Brown, Burgen & Gregory, 1972) and MO calculations (Kier & George, 1972). The importance of this peptide and the scarcity of detailed structural information for linear peptides prompted the study of its crystal structure.

Experimental

Trp-Met-Asp-Phe-NH_2 was obtained from Sigma as the hydrochloride salt and used without further purification. Crystals were grown by the vapour diffusion of diethyl ether into a methanolic peptide solution containing 0·005 *M* HCl. A microcrystalline precipitate appeared on the test-tube walls near the meniscus within 12 h; large elongated crescent-shaped crystals formed away from the meniscus within a further 3 days. They had to be mounted within 10 days since their prolonged exposure to the mother liquor resulted in visible physical deterioration accompanied by the precipitation of oily globules from solution. The

* Trp-Met-Asp-Phe-NH_2 = L-tryptophanyl-L-methionyl-L-aspartyl-L-phenylalanylamide.

† Permanent address: Indian Institute of Science, Bangalore, India.

0567-7408/82/061758-07$01.00

crystals decomposed in the atmosphere. Consequently, they were sealed in glass capillaries with considerable difficulty using protein-crystal-mounting techniques with one end of the capillary containing mother liquor. Attempts to grow crystals with less volatile ethers or other alcohols were unsatisfactory.

Intensity data were collected with a crystal 0·2 × 0·4 × 0·7 mm on a Syntex $P2_1$ diffractometer using graphite-monochromated Cu $K\alpha$ radiation and θ–2θ scans. Of 3633 unique reflections with $0 < 2\theta < 96°$, 3143 had $F > 3\sigma(F)$ and were treated as observed. Lorentz and polarization corrections were applied; no corrections were made for absorption. The asymmetric unit contains a peptide hydrochloride dimer together with one ether and two methanol molecules – altogether 95 non-hydrogen atoms.

Structure determination and refinement

The positions of the two Cl^- ions obtained from sharpened Patterson functions served as input partial structure for the difference-structure phase-refinement program *DIRDIF* (van den Hark, Prick & Beurskens, 1976). The resulting *E* map showed several chain fragments from which 11 atoms were accepted, although a later examination revealed that the 25 highest peaks all represented correct atomic positions. The structure was expanded to 51 atoms (the two peptide backbones and the Asp side chains) by several cycles of Karle tangent refinement. Successive Fourier syntheses determined another 24 atomic locations, but several ring atoms within the Trp and Phe residues (especially in the 'unprimed' molecule) gave rise to unacceptable bond lengths and angles or did not show up at all. Consequently, the four six-membered rings were idealized and treated as rigid bodies throughout the refinement, which was carried out with *SHELX*76 (Sheldrick, 1976).

After the first least-squares cycles, three solvent molecules ($2CH_3OH$ + $1C_4H_{10}O$) could be located from difference maps. H atoms were placed at their calculated positions (r_{X-H} = 0·98 Å) with temperature factors fixed at the values of the atoms to which they were attached. The ether atoms were refined with an occupancy factor of 0·5 to accommodate their partial disorder, and some low-angle reflections were omitted. Blocked full-matrix refinement was then continued with all non-hydrogen atoms anisotropic except for the solvent molecules, the atoms within rigid groups and also C(23) and C(23′), which showed very large temperature factors. H atoms were allowed to ride on their pivot atom. Finally, the restrictions imposed on the Trp and Phe residues of the 'primed' peptide molecule were released, while those for the 'unprimed' were retained since otherwise unacceptable geometries

Table 1. *Fractional coordinates* (×10⁴) *and equivalent isotropic temperature factors* (Å² ×10³), *with e.s.d.'s in parentheses*

Atoms refined isotropically are designated with an asterisk. $U_{eq} = \frac{1}{3}\sum_i \sum_j U_{ij}\, a_i^* a_j^* \mathbf{a}_i . \mathbf{a}_j$.

	x	*y*	*z*	U_{eq}/U (Å²)
Cl(1)	−1643 (2)	4241 (1)	92 (2)	59 (1)
Cl(2)	1217 (3)	3262 (2)	2791 (3)	105 (2)
N(1)	497 (8)	4089 (4)	9525 (9)	78 (5)
C(1*A*)	521 (9)	4697 (5)	9312 (11)	72 (6)
C(1)	−157 (9)	4851 (6)	8095 (11)	72 (6)
O(1)	−351 (7)	4484 (4)	7293 (8)	90 (4)
N(2)	−474 (7)	5377 (4)	7964 (9)	70 (5)
C(2*A*)	−945 (9)	5613 (5)	6828 (11)	75 (5)
C(2)	−486 (9)	6186 (6)	6732 (13)	73 (6)
O(2)	−13 (7)	6427 (4)	7643 (8)	87 (4)
N(3)	−660 (7)	6416 (4)	5612 (9)	72 (4)
C(3*A*)	−255 (9)	6955 (5)	5464 (11)	73 (6)
C(3)	−912 (9)	7250 (6)	4391 (12)	73 (6)
O(3)	−998 (8)	7039 (5)	3339 (9)	112 (5)
N(4)	−1355 (7)	7710 (4)	4581 (9)	68 (4)
C(4*A*)	−1891 (9)	8038 (6)	3598 (11)	79 (6)
C(4)	−1282 (10)	8426 (6)	3142 (14)	82 (7)
O(4)	−1439 (7)	8525 (4)	2004 (9)	94 (4)
N(5)	−633 (10)	8698 (6)	3883 (11)	132 (7)
C(11)	1549 (9)	4846 (6)	9276 (12)	83 (6)
C(12)	2236 (13)	4776 (9)	10385 (17)	114 (9)
C(13)	2791 (18)	4311 (14)	10821 (29)	213 (18)
N(6)	3354 (19)	4355 (21)	11937 (21)	313 (28)
C(14)	3076 (15)	4845 (7)	12189 (20)	253 (30)
C(15)	3508 (15)	5109 (7)	13275 (20)	395 (25)*
C(16)	3307 (15)	5673 (7)	13460 (20)	192 (9)
C(17)	2674 (15)	5971 (7)	12559 (20)	281 (16)*
C(18)	2242 (15)	5707 (7)	11473 (20)	178 (9)*
C(19)	2443 (15)	5144 (7)	11288 (20)	151 (8)*
C(21)	−2015 (10)	5698 (7)	6754 (15)	116 (8)
C(22)	−2584 (16)	5132 (12)	6729 (23)	211 (16)
S	−3813 (5)	5241 (5)	6255 (11)	295 (7)
C(23)	−4230 (34)	4633 (18)	6660 (42)	468 (33)*
C(31)	722 (10)	6895 (6)	5251 (12)	93 (6)
C(32)	1224 (11)	7454 (7)	5198 (15)	95 (8)
O(5)	1738 (11)	7433 (5)	4407 (13)	175 (8)
O(6)	1150 (7)	7855 (5)	5734 (9)	96 (5)
C(41)	−2680 (11)	8344 (8)	4053 (13)	124 (8)
C(42)	−3303 (11)	8762 (8)	3141 (15)	165 (12)
C(43)	−3214 (11)	9335 (8)	2922 (15)	203 (10)*
C(44)	−3847 (11)	9594 (8)	1957 (15)	196 (9)*
C(45)	−4571 (11)	9280 (8)	1212 (15)	285 (15)*
C(46)	−4661 (11)	8708 (8)	1431 (15)	531 (38)*
C(47)	−4027 (11)	8448 (8)	2396 (15)	322 (19)*
N(1′)	949 (8)	8127 (4)	−1649 (9)	85 (5)
C(1*A*′)	1067 (9)	7569 (6)	−1011 (11)	71 (5)
C(1′)	407 (10)	7544 (6)	−177 (13)	71 (6)
O(1′)	296 (7)	7942 (4)	454 (8)	92 (4)
N(2′)	−15 (8)	7055 (5)	−158 (8)	71 (4)
C(2*A*′)	−637 (10)	6897 (6)	701 (12)	86 (6)
C(2′)	−476 (10)	6274 (5)	998 (13)	73 (6)
O(2′)	−232 (8)	5956 (4)	260 (9)	101 (5)
N(3′)	−682 (7)	6102 (4)	1997 (9)	70 (5)
C(3*A*′)	−639 (9)	5516 (6)	2279 (12)	79 (6)
C(3′)	−1458 (10)	5397 (6)	2916 (12)	72 (6)
O(3′)	−1537 (6)	5635 (4)	3806 (8)	88 (4)
N(4′)	−2046 (7)	4968 (5)	2383 (9)	69 (4)
C(4*A*′)	−2864 (9)	4796 (6)	2830 (12)	74 (6)
C(4′)	−2687 (14)	4265 (8)	3600 (14)	101 (8)

Table 1 (*cont.*)

	x	*y*	*z*	U_{eq}/U (Å^2)
O(4′)	−3330 (9)	4054 (6)	3976 (11)	140 (7)
N(5′)	−1820 (10)	4041 (5)	3832 (13)	112 (7)
C(11′)	2058 (9)	7494 (7)	−353 (12)	91 (6)
C(12′)	2721 (14)	7500 (10)	−1158 (17)	117 (9)
C(13′)	3297 (16)	7936 (12)	−1389 (20)	156 (12)
N(6′)	3837 (16)	7774 (15)	−2211 (23)	196 (14)
C(14′)	3558 (21)	7226 (17)	−2566 (30)	198 (18)
C(15′)	3903 (19)	6807 (20)	−3294 (25)	265 (24)
C(16′)	3514 (42)	6315 (20)	−3466 (49)	334 (37)
C(17′)	2882 (34)	6119 (23)	−2734 (34)	301 (32)
C(18′)	2608 (16)	6502 (12)	−2025 (19)	147 (12)
C(19′)	2924 (15)	7046 (11)	−1804 (20)	123 (10)
C(21′)	−1731 (11)	6967 (7)	80 (15)	106 (7)
C(22′)	−2003 (12)	7565 (9)	−252 (15)	137 (10)
S′	−3294 (4)	7599 (4)	−1023 (6)	184 (4)
C(23′)	−3652 (25)	7585 (18)	255 (33)	360 (20)*
C(31′)	344 (9)	5350 (6)	3137 (13)	98 (7)
C(32′)	461 (12)	4745 (8)	3270 (19)	125 (10)
O(5′)	1076 (12)	4510 (6)	2807 (16)	202 (10)
O(6′)	33 (10)	4441 (6)	3762 (16)	209 (10)
C(41′)	−3707 (11)	4708 (7)	1728 (14)	103 (8)
C(42′)	−4051 (13)	5196 (8)	1023 (21)	106 (9)
C(43′)	−3672 (13)	5393 (11)	173 (20)	135 (11)
C(44′)	−4059 (20)	5875 (16)	−494 (26)	188 (17)
C(45′)	−4873 (32)	6159 (12)	−121 (35)	217 (21)
C(46′)	−5231 (20)	5867 (18)	668 (31)	184 (18)
C(47′)	−4815 (19)	5435 (10)	1237 (20)	142 (12)
C(1*M*)	5499 (21)	8522 (14)	5782 (24)	304 (17)*
O(1*M*)	5086 (13)	8584 (8)	6789 (17)	233 (7)*
C(2*M*)	2187 (20)	8517 (12)	2954 (25)	274 (15)*
O(2*M*)	2678 (9)	8312 (6)	4113 (13)	172 (5)*
C(1*E*)	4053 (29)	7313 (18)	2710 (36)	167 (16)*
C(2*E*)	4977 (44)	7092 (27)	3267 (57)	233 (27)*
O(3*E*)	5423 (35)	7128 (19)	4198 (43)	228 (17)*
C(4*E*)	6138 (47)	6816 (30)	4323 (57)	256 (28)*
C(5*E*)	6620 (35)	6578 (22)	3494 (44)	198 (20)*

were found. Because of the high thermal motion of some side chains and the solvent molecules, unit weights were used throughout. The refinement converged at $R = 0{\cdot}082$. There were no peaks higher than 0·35 e Å^{-3} in the final difference map. Parameters are listed in Table 1.*

Results

The conformation of the peptides

Fig. 1 illustrates the conformations of the two independent peptide molecules. The bond lengths and angles (Tables 2 and 3) within the peptide backbones are in the expected range and agree with calculated

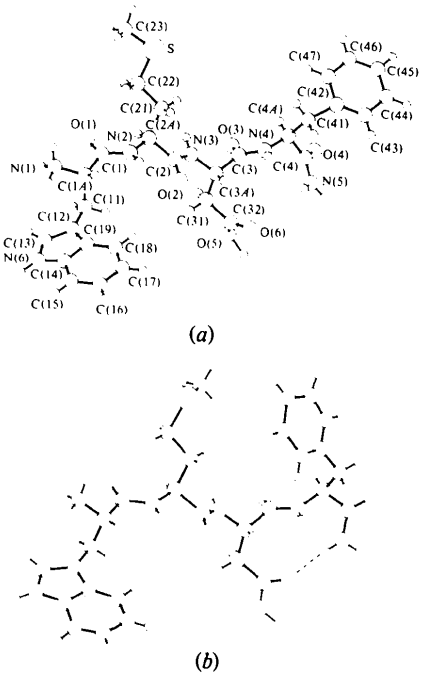

Fig. 1. The conformations of the two peptide molecules with the atom-numbering scheme (shown for the 'unprimed' peptide). The intramolecular hydrogen bond N(5′)—H···O(6′) is indicated by a broken line.

average values (Corey & Pauling, 1953; Marsh & Donohue, 1967). The geometries of the side chains show larger discrepancies consistent with the substantial thermal motions which tend to increase with the distance of the atoms from the main chain.

The dihedral angles φ, ψ and ω (Table 4) are characteristic of a β-pleated sheet (Miyazawa, 1961; Arnott, Dover & Elliott, 1967). The α-C atoms of Trp and Asp are separated by 6·85 and 6·93 Å, respectively, which agrees well with the peptide-chain-repeat distance of 6·89 Å in the anti-parallel β-sheet formed by poly(alanine) (Arnott *et al.*, 1967). The main chain is regular from N(1) to C(3) but is twisted at the C-terminal in both peptides. Comparison of the angle pairs φ_2/ψ_2 and φ_3/ψ_3 indicates a marked conformational change from the favoured right-hand (Chothia, 1973; Raghavendra & Sasisekharan, 1979) to a left-hand twist. The conformations of the side chains, which are situated alternately above and below the backbone, are similar to those observed in other crystal structures (Bhat, Sasisekharan & Vijayan, 1979). Five χ^1 values are close to −60°, while three residues (2 Asp and 1 Phe) have their γ-C atom *trans* to the peptide N atom. The conformations of the four aromatic side chains are characterized by dihedral angles χ^2 in the commonly found range near 90° ($82° < \chi^2 < 100°$) in agreement with other Phe-containing peptide structures (van den Veen & Low, 1972; Cotrait

* Lists of structure factors, anisotropic thermal parameters and H-atom coordinates have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 36713 (18 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 2. *Bond lengths* (Å) *with e.s.d.'s in parentheses*

N(1)–C(1*A*)	1·47 (2)	N(1′)–C(1*A*′)	1·50 (2)
C(1*A*)–C(1)	1·54 (2)	C(1*A*′)–C(1′)	1·49 (2)
C(1*A*)–C(11)	1·53 (2)	C(1*A*′)–C(11′)	1·46 (2)
C(1)–O(1)	1·24 (2)	C(1′)–O(1′)	1·22 (2)
C(1)–N(2)	1·33 (2)	C(1′)–N(2′)	1·31 (2)
N(2)–C(2*A*)	1·43 (2)	N(2′)–C(2*A*′)	1·52 (2)
C(2*A*)–C(2)	1·53 (2)	C(2*A*′)–C(2′)	1·52 (2)
C(2*A*)–C(21)	1·54 (2)	C(2*A*′)–C(21′)	1·58 (2)
C(2)–O(2)	1·24 (2)	C(2′)–O(2′)	1·24 (2)
C(2)–N(3)	1·35 (2)	C(2′)–N(3′)	1·30 (2)
N(3)–C(3*A*)	1·43 (2)	N(3′)–C(3*A*′)	1·43 (2)
C(3*A*)–C(3)	1·53 (2)	C(3*A*′)–C(3′)	1·55 (2)
C(3*A*)–C(31)	1·49 (2)	C(3*A*′)–C(31′)	1·57 (2)
C(3)–O(3)	1·27 (2)	C(3′)–O(3′)	1·19 (2)
C(3)–N(4)	1·31 (2)	C(3′)–N(4′)	1·37 (2)
N(4)–C(4*A*)	1·43 (2)	N(4′)–C(4*A*′)	1·45 (2)
C(4*A*)–C(4)	1·45 (2)	C(4*A*′)–C(4′)	1·52 (2)
C(4*A*)–C(41)	1·54 (2)	C(4*A*′)–C(41′)	1·54 (2)
C(4)–O(4)	1·28 (2)	C(4′)–O(4′)	1·21 (2)
C(4)–N(5)	1·28 (2)	C(4′)–N(5′)	1·33 (2)
C(11)–C(12)	1·42 (2)	C(11)–C(12′)	1·47 (3)
C(12)–C(13)	1·39 (4)	C(12′)–C(13′)	1·39 (4)
C(12)–C(19)	1·33 (3)	C(12′)–C(19′)	1·37 (4)
C(13)–N(6)	1·34 (4)	C(13′)–N(6′)	1·40 (4)
N(6)–C(14)	1·28 (5)	N(6′)–C(14′)	1·39 (5)
C(14)–C(15)	1·39	C(14′)–C(15′)	1·45 (6)
C(14)–C(19)	1·39	C(14′)–C(19′)	1·46 (4)
C(15)–C(16)	1·39	C(15′)–C(16′)	1·29 (7)
C(16)–C(17)	1·39	C(16′)–C(17′)	1·44 (8)
C(17)–C(18)	1·39	C(17′)–C(18′)	1·33 (6)
C(18)–C(19)	1·39	C(18′)–C(19′)	1·37 (4)
C(21)–C(22)	1·57 (3)	C(21′)–C(22′)	1·50 (3)
C(22)–S	1·74 (2)	C(22′)–S′	1·86 (2)
S–C(23)	1·67 (5)	S′–C(23′)	1·65 (4)
C(31)–C(32)	1·52 (2)	C(31′)–C(32′)	1·45 (2)
C(32)–O(5)	1·29 (3)	C(32′)–O(5′)	1·26 (3)
C(32)–O(6)	1·15 (2)	C(32′)–O(6′)	1·17 (3)
C(41)–C(42)	1·56 (2)	C(41′)–C(42′)	1·43 (3)
C(42)–C(43)	1·39	C(42′)–C(43′)	1·30 (3)
C(42)–C(47)	1·39	C(42′)–C(47′)	1·31 (4)
C(43)–C(44)	1·39	C(43′)–C(44′)	1·41 (4)
C(44)–C(45)	1·39	C(44′)–C(45′)	1·50 (6)
C(45)–C(46)	1·39	C(45′)–C(46′)	1·33 (6)
C(46)–C(47)	1·39	C(46′)–C(47′)	1·28 (4)
C(1*M*)–O(1*M*)	1·41 (4)	C(1*E*)–C(2*E*)	1·43 (7)
C(2*M*)–O(2*M*)	1·43 (3)	C(2*E*)–O(3*E*)	1·10 (7)
		O(3*E*)–C(4*E*)	1·25 (8)
		C(4*E*)–C(5*E*)	1·41 (9)

Table 3. *Bond angles* (°) *with e.s.d.'s in parentheses*

N(1)–C(1*A*)–C(1)	110 (1)	N(1′)–C(1*A*′)–C(1′)	109 (1)
N(1)–C(1*A*)–C(11)	107 (1)	N(1′)–C(1*A*′)–C(11′)	110 (1)
C(1)–C(1*A*)–C(11)	110 (1)	C(1′)–C(1*A*′)–C(11′)	111 (1)
C(1*A*)–C(1)–O(1)	118 (1)	C(1*A*′)–C(1′)–O(1′)	122 (1)
C(1*A*)–C(1)–N(2)	117 (1)	C(1*A*′)–C(1′)–N(2′)	114 (1)
O(1)–C(1)–N(2)	124 (1)	O(1′)–C(1′)–N(2′)	123 (1)
C(1)–N(2)–C(2*A*)	123 (1)	C(1′)–N(2′)–C(2*A*′)	125 (1)
N(2)–C(2*A*)–C(2)	107 (1)	N(2′)–C(2*A*′)–C(2′)	108 (1)
N(2)–C(2*A*)–C(21)	111 (1)	N(2′)–C(2*A*′)–C(21′)	111 (1)
C(2)–C(2*A*)–C(21)	109 (1)	C(2′)–C(2*A*′)–C(21′)	107 (1)
C(2*A*)–C(2)–O(2)	121 (1)	C(2*A*′)–C(2′)–O(2′)	120 (1)
C(2*A*)–C(2)–N(3)	116 (1)	C(2*A*′)–C(2′)–N(3′)	116 (1)
O(2)–C(2)–N(3)	123 (1)	O(2′)–C(2′)–N(3′)	124 (1)
C(2)–N(3)–C(3*A*)	119 (1)	C(2′)–N(3′)–C(3*A*′)	120 (1)
N(3)–C(3*A*)–C(3)	108 (1)	N(3′)–C(3*A*′)–C(3′)	106 (1)
N(3)–C(3*A*)–C(31)	111 (1)	N(3′)–C(3*A*′)–C(31′)	112 (1)
C(3)–C(3*A*)–C(31)	110 (1)	C(3′)–C(3*A*′)–C(31′)	110 (1)
C(3*A*)–C(3)–O(3)	118 (1)	C(3*A*′)–C(3′)–O(3′)	123 (1)
C(3*A*)–C(3)–N(4)	119 (1)	C(3*A*′)–C(3′)–N(4′)	113 (1)
O(3)–C(3)–N(4)	123 (1)	O(3′)–C(3′)–N(4′)	124 (1)
C(3)–N(4)–C(4*A*)	121 (1)	C(3′)–N(4′)–C(4*A*′)	122 (1)
N(4)–C(4*A*)–C(4)	111 (1)	N(4′)–C(4*A*′)–C(4′)	113 (1)
N(4)–C(4*A*)–C(41)	107 (1)	N(4′)–C(4*A*′)–C(41′)	108 (1)
C(4)–C(4*A*)–C(41)	112 (1)	C(4′)–C(4*A*′)–C(41′)	110 (1)
C(4*A*)–C(4)–O(4)	120 (1)	C(4*A*′)–C(4′)–O(4′)	120 (2)
C(4*A*)–C(4)–N(5)	120 (1)	C(4*A*′)–C(4′)–N(5′)	118 (2)
O(4)–C(4)–N(5)	120 (2)	O(4′)–C(4′)–N(5′)	121 (2)
C(1*A*)–C(11)–C(12)	115 (1)	C(1*A*′)–C(11′)–C(12′)	112 (1)
C(11)–C(12)–C(13)	129 (2)	C(11′)–C(12′)–C(13′)	129 (2)
C(11)–C(12)–C(19)	126 (2)	C(11′)–C(12′)–C(19′)	125 (2)
C(13)–C(12)–C(19)	104 (2)	C(13′)–C(12′)–C(19′)	106 (2)
C(12)–C(13)–N(6)	116 (3)	C(12′)–C(13′)–N(6′)	112 (2)
C(13)–N(6)–C(14)	98 (3)	C(13′)–N(6′)–C(14′)	106 (3)
N(6)–C(14)–C(15)	120 (2)	N(6′)–C(14′)–C(15′)	134 (3)
N(6)–C(14)–C(19)	119 (2)	N(6′)–C(14′)–C(19′)	106 (3)
C(15)–C(14)–C(19)	120	C(15′)–C(14′)–C(19′)	119 (3)
C(14)–C(15)–C(16)	120	C(14′)–C(15′)–C(16′)	121 (4)
C(15)–C(16)–C(17)	120	C(15′)–C(16′)–C(17′)	121 (5)
C(16)–C(17)–C(18)	120	C(16′)–C(17′)–C(18′)	116 (4)
C(17)–C(18)–C(19)	120	C(17′)–C(18′)–C(19′)	128 (3)
C(12)–C(19)–C(14)	103 (2)	C(12′)–C(19′)–C(14′)	109 (3)
C(12)–C(19)–C(18)	137 (2)	C(12′)–C(19′)–C(18′)	137 (2)
C(14)–C(19)–C(18)	120	C(14′)–C(19′)–C(18′)	113 (3)
C(2*A*)–C(21)–C(22)	114 (1)	C(2*A*′)–C(21′)–C(22′)	113 (1)
C(21)–C(22)–S	111 (2)	C(21′)–C(22′)–S′	109 (1)
C(22)–S–C(23)	101 (2)	C(22′)–S′–C(23′)	94 (1)
C(3*A*)–C(31)–C(32)	113 (1)	C(3*A*′)–C(31′)–C(32′)	112 (1)
C(31)–C(32)–O(5)	110 (1)	C(31′)–C(32′)–O(5′)	118 (2)
C(31)–C(32)–O(6)	127 (2)	C(31′)–C(32′)–O(6′)	127 (2)
O(5)–C(32)–O(6)	123 (2)	O(5′)–C(32′)–O(6′)	115 (2)
C(4*A*)–C(41)–C(42)	116 (1)	C(4*A*′)–C(41′)–C(42′)	117 (1)
C(41)–C(42)–C(43)	133 (1)	C(41′)–C(42′)–C(43′)	124 (2)
C(41)–C(42)–C(47)	107 (2)	C(41′)–C(42′)–C(47′)	116 (2)
C(43)–C(42)–C(47)	120	C(43′)–C(42′)–C(47′)	120 (2)
C(42)–C(43)–C(44)	120	C(42′)–C(43′)–C(44′)	120 (2)
C(43)–C(44)–C(45)	120	C(43′)–C(44′)–C(45′)	117 (3)
C(44)–C(45)–C(46)	120	C(44′)–C(45′)–C(46′)	114 (3)
C(45)–C(46)–C(47)	120	C(45′)–C(46′)–C(47′)	123 (5)
C(42)–C(47)–C(46)	120	C(42′)–C(47′)–C(46′)	124 (3)
C(1*E*)–C(2*E*)–O(3*E*)	132 (7)	O(3*E*)–C(4*E*)–C(5*E*)	133 (6)
C(2*E*)–O(3*E*)–C(4*E*)	110 (6)		

& Barrans, 1974) and solution studies (Becker, Bleich, Day, Freer, Glasel & Visintainer, 1979). This results in an orientation where the two indole rings are almost parallel to the peptide group linking Trp and Met (interplanar angles: 4 and 2°). The methionine side chains have the S atoms in the preferred *trans* arrangement (Stenkamp & Jensen, 1975; Chen & Parthasarathy, 1977), but differ in the conformation of the terminal methyl group, which is *trans* in the 'unprimed' and *gauche* in the 'primed' molecule. Although C(23) and C(23′) showed very high temperature factors, there was no indication of alternative positions for these atoms during the refinement.

The main conformational differences between the two peptides occur at the C-termini reflected by the dissimilarities in the dihedral angles ψ_4, $\chi_3^{2,1}$ and χ_4^1. The 'unprimed' molecule has an extended conformation with a *trans* Phe residue, whereas the other peptide contains an intramolecular H bond N(5′)–H···O(6′) between the terminal amide group and the Asp side chain with the phenyl ring bent back towards the backbone. Model-building studies supported by van der

Table 4. *Important dihedral angles* (°)

The dihedral angles are designated according to the conventions of the IUPAC–IUB Commission on Biochemical Nomenclature (1970). E.s.d.'s are 2° for φ, ψ and ω and 3° for χ.

	'Unprimed' molecule	'Primed' molecule		'Unprimed' molecule	'Primed' molecule
ψ_1	157	140	χ_1^1	−66	−61
ω_1	167	173	$\chi_1^{2,1}$	91	100
φ_2	−133	−145	χ_2^1	−67	−64
ψ_2	164	158	χ_2^2	−165	178
ω_2	180	174	χ_2^3	−166	81
φ_3	−151	−144	χ_3^1	−175	−171
ψ_3	115	127	$\chi_3^{2,1}$	33	−68
ω_3	173	−179	χ_4^1	176	−65
φ_4	−85	−101	χ_4^2	87	82
ψ_4	−42	4			

Waals energy calculations using the program *EENY* (Motherwell & Isaacs, 1972) suggest a correlation between the formation of the internal H bond and the change from a *trans* to a *gauche* Phe residue, which is necessary to relieve steric hindrance between the carbonyl O(4) and the phenyl ring.

The β-pleated sheet

The peptide backbones are situated approximately in the *bc* plane lying side by side in a head-to-tail fashion. The two chains are, however, not collinear but inclined at an angle of about 30° (Fig. 2). Translocation of this dimeric species along **c** results in the formation of an antiparallel β-pleated sheet (Fig. 3) held together by four H bonds between the atom pairs N(2)/O(2) and N(3)/O(3) of both peptides (Table 5). The twofold screw axis extends this 'one-dimensional' structure into a herring-bone pattern where the individual β-sheets are connected by an extensive H-bonding network. The 'unprimed' peptide forms an infinite zigzag chain along **b** with adjacent members linked together by a pair of N–H···O bonds between the N- and C-terminal groups. Symmetry-related 'primed' peptides do not interact directly with each other but instead are bridged by Cl(1) and O(6) supported by N(1)–H···O(1′). The Cl^- ions connect N(1) and N(4) within each dimer by accepting four H bonds and thus contribute to the stabilization of the β-pleated sheet. They may also be responsible for the considerable twist of the peptide backbones from N(4) to the C-terminal.

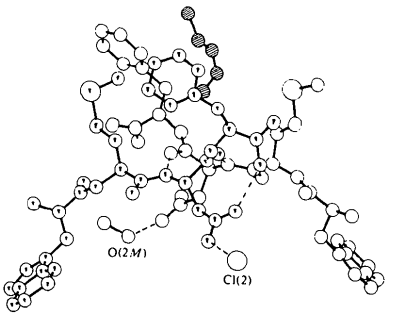

Fig. 2. The conformation of the peptide dimer. H bonds are indicated by broken lines. The atoms of the 'primed' peptide molecule are marked by a '▼' and the ether molecule is hatched. H atoms are omitted.

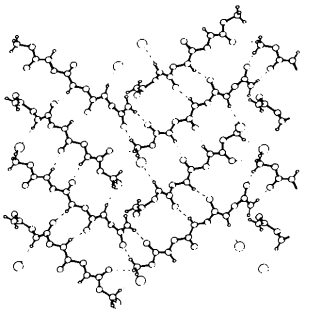

Fig. 3. A view along **a*** with **b** horizontal and **c** vertical, showing the β-pleated sheets. Side chains and H atoms attached to carbon are omitted. H bonds are dashed.

Table 5. *Hydrogen-bond distances* (Å) *with e.s.d.'s in parentheses*

N(1)–H···O(1′)	2·96 (3)	N(1′)–H···Cl(2)	3·10 (3)
N(1)–H′···O(4)	2·78 (3)	N(1′)–H′···O(6)	3·12 (4)
N(1)–H″···Cl(1)	3·31 (3)	N(1′)–H″···Cl(1)	3·21 (3)
N(2)–H···O(2′)	2·90 (3)	N(2′)–H···O(2)	2·91 (3)
N(3)–H···O(3′)	2·84 (3)	N(3′)–H···O(3)	2·80 (3)
N(4)–H···Cl(2)	3·23 (3)	N(4′)–H···Cl(1)	3·29 (3)
N(5)–H···O(6′)	3·15 (4)	N(5′)–H···O(6′)	2·85 (4)
N(5)–H′···O(1)	2·86 (3)	N(5′)–H′···O(6)	2·98 (4)
N(6)–H···O(1*M*)	3·00 (4)	N(6′)–H···O(1*M*)	3·03 (4)
O(5)–H···O(2*M*)	2·55 (4)	O(5′)–H···Cl(2)	2·97 (4)
O(1*M*)–H···O(4′)	2·71 (3)	O(2*M*)–H···O(4′)	2·79 (3)

The packing of the side chains

Since the peptide backbones are contained in the *bc* plane, the side chains extend mainly along **a**. Within each β-sheet, the Trp and Asp residues, which can participate in H bonding, lie on opposite sides from the non-polar Met and Phe side chains. Adjacent carboxylic groups point away from each other with their acidic H atoms bonded to Cl(2) and the methanol oxygen O(2*M*), respectively (Fig. 2). There are three important features concerning side-chain interactions (Fig. 4): (1) all polar *X*–H groups donate H bonds; (2) no stacking of aromatic rings occurs; (3) the non-polar side chains form an elongated 'hydrophobic pocket' which is filled by an ether molecule.

The overall H-bonding pattern (Table 5) is remarkably symmetrical. The amide group in the 'primed' peptide forms four H bonds; it donates towards both

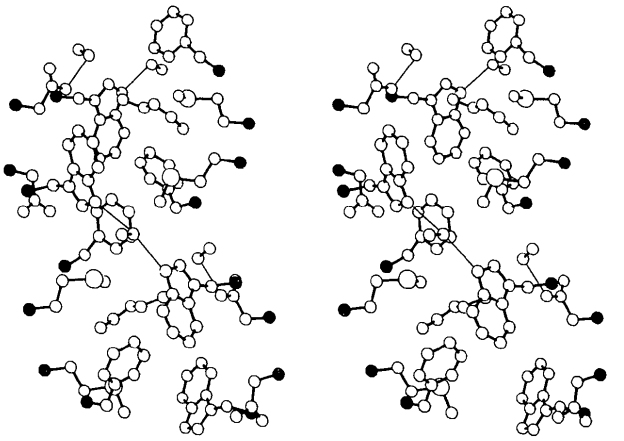

Fig. 4. Stereoscopic view along **c*** with **a** horizontal and **b** vertical, showing the packing of the side chains. The main chains are omitted except for the α-carbon atoms (black circles). H bonds are indicated by thin lines.

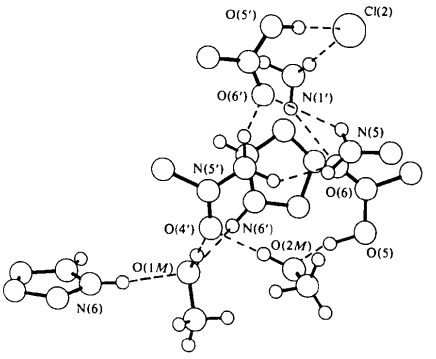

Fig. 5. H bonds (broken lines) formed by side chains and solvent molecules. O, N and Cl atoms are labelled.

Asp carbonyl O atoms and accepts from both methanol molecules. Furthermore, the N—H groups of the Trp residues are bonded to the same methanol and O(6′) interacts with both N(5) and N(5′). Altogether the methanol molecules participate in five H bonds thus helping to stabilize the packing of the polar side chains of the peptides (Fig. 5).

The Met and Phe residues of a peptide dimer form an elongated cavity capped by two indole rings (Figs. 2 and 4). It has the ideal size for one molecule of diethyl ether but is too narrow to permit longer or branched-chain ethers. The presence of the volatile diethyl ether within the crystals is consistent with their quick decomposition when removed from the mother liquor. The partial loss of diethyl ether as indicated from electron-density maps would also explain the very high temperature factors of the atoms surrounding the cavity.

Discussion

The conformations observed for Trp-Met-Asp-Phe-NH_2 do not agree with the results of extended Hückel MO calculations (Kier & George, 1972) which predict that the two aromatic side chains are situated on the same side of the peptide backbone with Met and Asp on the opposite side. NMR studies (Feeney *et al.*, 1972) suggest an extended coil conformation and rule out the formation of an intramolecular H bond. However, the crystal structure is dominated by intermolecular H bonding and thus describes the conformation of a peptide interacting with itself rather than the situation in solution or at the receptor site. Morley (1968) has shown that replacement of Met by *N*-methylnorleucine does not impair the activity, although this analogue is unable to form a similar β-ribbon due to methylation at N(3). Therefore, the conformation found in the present crystal structure is unlikely to represent the biologically active form of the tetrapeptide amide but may be relevant to states in which gastrin is transported or stored.

The antiparallel β-pleated sheet extending over the entire crystal is similar to that formed by Ala-Ala-Ala (Fawcett, Camerman & Camerman, 1975). This tripeptide crystallizes as a neutral dimer with water molecules instead of Cl^- ions as links between successive β-ribbons. The connection of β-sheet-like segments by anions is also observed with other crystal structures (Tokuma, Ashida & Kakudo, 1969; Rao & Parthasarathy, 1973) and seems to be a characteristic way of packing the individual ribbons together, because it maximizes the number of H bonds in addition to interstrand H bonding. However, the formation of a β-sheet sandwich also requires an efficient burial of the side chains between the backbone layers. In this respect, Trp-Met-Asp-Phe-NH_2 seems unlikely to show such an arrangement because the side chains differ considerably in size, shape and hydrophobic/hydrophilic character. Indeed, crystal growth was slowest perpendicular to the β-sheet, whereas in Ala-Ala-Ala, which permits an efficient close packing of side chains, this direction corresponds to maximum crystal growth. The packing problems due to the rather different side chains are alleviated by the two components of the solvent: methanol interacts mainly with polar groups *via* H bonding and the non-polar ether molecule, with its considerable flexibility, is able to fill the large pocket enclosed by the hydrophobic side chains. Both kinds of solvent molecules thus play important roles in the design of the whole crystal structure.

The present structure suggests that antiparallel β-ribbon formation, with its simple way of maximizing the number of interpeptide H bonds, represents a very general type of structure. Whereas the conformations of cyclic peptides are limited by the constraints due to ring closure, linear peptides are quite flexible and can

adopt a multitude of energetically favourable conformations. In crystal structures, however, packing interactions and intermolecular H bonding may dominate all other considerations. Many linear peptides will probably crystallize as regular β-sheets, provided the packing problems which arise from the specific amino acid sequence can be solved by solvent molecules of the appropriate size, shape and polarity. Therefore, this structure shows one major difficulty in utilizing X-ray analysis to elucidate the 'active' conformation of linear peptides. It also stresses the care needed in selecting the right peptide modification if one is to relate crystal structure to biological function.

We thank the Medical Research Council for financial support and a Visiting Professorship (MAV), the Science Research Council for the diffractometer, the Deutsche Forschungsgemeinschaft for a Postdoctoral Fellowship (EE) and Professor G. M. Sheldrick for creating an extended version of *SHELX*76.

References

ARNOTT, S., DOVER, S. D. & ELLIOTT, A. (1967). *J. Mol. Biol.* **30**, 201–208.
BECKER, E. L., BLEICH, H. E., DAY, A. R., FREER, R. J., GLASEL, J. A. & VISINTAINER, J. (1979). *Biochemistry*, **18**, 4656–4668.
BHAT, T. N., SASISEKHARAN, V. & VIJAYAN, M. (1979). *Int. J. Pept. Protein Res.* **13**, 170–184.
CHEN, C. & PARTHASARATHY, R. (1977). *Acta Cryst.* B**33**, 3332–3336.
CHOTHIA, C. (1973). *J. Mol. Biol.* **75**, 295–302.
COREY, R. B. & PAULING, L. (1953). *Proc. R. Soc. London Ser. B*, **141**, 10–20.
COTRAIT, M. & BARRANS, Y. (1974). *Acta Cryst.* B**30**, 1018–1023.
FAWCETT, J. K., CAMERMAN, N. & CAMERMAN, A. (1975). *Acta Cryst.* B**31**, 658–665.
FEENEY, J., ROBERTS, G. C. K., BROWN, J. P., BURGEN, A. S. V. & GREGORY, H. (1972). *J. Chem. Soc. Perkin Trans.* 2, pp. 601–604.
HARK, T. E. M. VAN DEN, PRICK, P. & BEURSKENS, P. T. (1976). *Acta Cryst.* A**32**, 816–821.
IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *J. Mol. Biol.* **52**, 1–17.
KIER, L. B. & GEORGE, J. M. (1972). *J. Med. Chem.* **15**, 384–386.
MARSH, R. E. & DONOHUE, J. (1967). *Adv. Protein Chem.* **22**, 235–256.
MIYAZAWA, T. (1961). *J. Polym. Sci.* **55**, 215–231.
MORLEY, J. S. (1968). *Proc. R. Soc. London Ser. B*, **170**, 97–111.
MOTHERWELL, W. D. S. & ISAACS, N. W. (1972). *J. Mol. Biol.* **71**, 231–241.
RAGHAVENDRA, K. & SASISEKHARAN, V. (1979). *Int. J. Pept. Protein Res.* **14**, 326–338.
RAO, S. N. & PARTHASARATHY, R. (1973). *Acta Cryst.* B**29**, 2379–2388.
SHELDRICK, G. M. (1976). *SHELX*76. Program for crystal structure determination. Univ. of Cambridge, England.
STENKAMP, R. E. & JENSEN, L. H. (1975). *Acta Cryst.* B**31**, 857–861.
TOKUMA, Y., ASHIDA, T. & KAKUDO, M. (1969). *Acta Cryst.* B**25**, 1367–1373.
TRACY, H. J. & GREGORY, R. A. (1964). *Nature (London)*, **204**, 935–938.
VEEN, J. M. VAN DER & LOW, B. W. (1972). *Acta Cryst.* B**28**, 3548–3559.

Acta Cryst. (1982). B**38**, 1764–1768

Etude Conformationnelle de la Diphényl-2,4 Δ-1-Pyrrolinedicarboxylate-5,5 de Diéthyle

PAR FRANÇOIS THÉOBALD

Laboratoire de Chimie Physique, UER des Sciences Exactes et Naturelles, 25030 Besançon CEDEX, France

ET MOHAMED BIROUK ET JEAN-FRANÇOIS ROBERT*

Laboratoire de Chimie Pharmaceutique, UER des Sciences Médicales et Pharmaceutiques, 25030 Besançon CEDEX, France

(*Reçu le* 28 *avril* 1981, *accepté le* 15 *février* 1982)

Abstract

The crystal structure of diethyl 2,4-diphenyl-1-pyrroline-5,5-dicarboxylate, $C_{22}H_{23}NO_4$, has been established in order to ascertain results from a previous NMR study. Triclinic, $P\bar{1}$, $a = 8{\cdot}654$ (2), $b = 14{\cdot}012$ (2), $c = 8{\cdot}472$ (3) Å, $\alpha = 98{\cdot}97$ (2), $\beta = 102{\cdot}89$ (2), $\gamma = 102{\cdot}35$ (1)°, $Z = 2$, $V = 955{\cdot}5$ Å^3, $d_x = 1{\cdot}260$ Mg m^{-3}, μ(Cu $K\alpha$) = 0·625 mm^{-1}. The final R is 0·064 for 3998 independent reflexions. The angle the 1-pyrroline plane makes with the plane of the φ_2 phenyl group is 8 (1)° whereas the angle with φ_4 is 65 (1)°, *i.e.* an angle

* L'auteur auquel toute correspondance doit être adressée.

0567-7408/82/061764-05$01.00