Acta Cryst. (1995). **D51**, 48–59

Structure of Octreotide, a Somatostatin Analogue

BY EHMKE POHL, ANDREAS HEINE AND GEORGE M. SHELDRICK*

Institut für Anorganische Chemie der Universität Göttingen, Tammannstrasse 4, D-37077 Göttingen, Germany

ZBIGNIEW DAUTER AND KEITH S. WILSON

European Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

AND JÖRG KALLEN, WERNER HUBER AND PAUL J. PFÄFFLI

Preclinical Research Department, Sandoz Ltd, CH-4002 Basel, Switzerland

(*Received* 1 *July* 1993; *accepted* 27 *May* 1994)

Abstract

Octreotide, a synthetic somatostatin analogue, is an octapeptide with one disulfide bridge. Crystals of octreotide are orthorhombic, space group $P2_12_12_1$, a = 18.458 (5), b = 30.009 (7), c = 39.705 (27) Å, with three molecules of octapeptide, one ordered oxalate dianion and 52 water molecules in the asymmetric unit. Complete protonation of the NH_2 groups (as assumed in the refinement) would require three oxalate dianions in the asymmetric unit for charge neutrality; a chemical analysis indicated that four are present. In either case they are so disordered that they cannot be distinguished from the water molecules. The 18 951 unique reflections (R_{sym} = 0.026) used for structure solution and refinement were recorded with the EMBL imaging-plate scanner using synchrotron radiation. The structure was solved by Patterson interpretation, locating the three disulfide bridges, followed by tangent phase expansion and E-Fourier recycling. The anisotropic refinement against all F^2 data between 1.04 and 10.0 Å resolution by blocked restrained full-matrix least-squares techniques converged to a conventional R index based on F of 0.084 [$I > 2\sigma(I)$ and $10.0 > d > 1.04$ Å] and $wR2$, the weighted R-index on F^2, of 0.246 (for all data). One peptide molecule adopts a flat β-sheet structure; the other two possess different irregular backbone conformations, but are similar to each other. All three molecules have a distorted type II′ β-turn around the D-Trp–Lys region, but exhibit different side-chain conformations. The crystal structure is stabilized by a network of inter- and intramolecular hydrogen bonds.

Introduction

Somatostatin (somatotropin release-inhibiting factor, SRIF, Fig. 1), a cyclic tetradecapeptide, has been isolated from ovine hypothalamus (Brazeau *et al.*, 1973; Burgus, Ling, Butcher & Guillemin, 1973), but is also present in many other organs. It inhibits the release of many peptide hormones including glucagon, insulin, secretin and growth hormone (Gerick & Lorenzi, 1978). Furthermore somatostatin plays an important role in the central nervous system as a neurotransmitter and neuromodulator (Reichlin, 1982). Somatostatin has been employed in the treatment of acute gastrointestinal disorders, but its therapeutic use is limited by the very short *in vivo* half-life of 1 to 2 min (Reichlin, 1983). Extensive activity studies on somatostatin and its derivatives (reviewed by Veber, 1992) have shown that the fragment -Phe^7-Trp^8-Lys^9-Thr^{10}- is essential for biological activity, and that this activity may be increased by replacing L-Trp^8 by its enantiomer D-Trp^8 (Rivier, Brown & Vale, 1975).

The search for synthetic somatostatin agonists which would be more stable *in vivo* has concentrated on cyclic peptides containing the -Phe^7-(D)-Trp^8-Lys^9-Thr^{10}- fragment. Initial investigations involved carbacyclic and bicyclic peptides (Veber *et al.*, 1976; Veber *et al.*, 1979; Nutt, Veber & Saperstein, 1980). The cyclic hexapeptide c[Pro^6-Phe^7-(D)-Trp^8-Lys^9-Thr^{10}-Phe^{11}] (the superscript numbers refer to the sequence of somatostatin), that contains the key sequence and is more active than somatostatin itself, has been particularly intensively studied (Veber *et al.*, 1981). From NMR studies, Veber (1981) proposed a type II′ β-turn about D-Trp^8 and Lys^9, which would correspond to a type II turn in somatostatin; type II′ turns are favoured by a glycine or D-amino acid in the position occupied by the tryptophan. NMR studies by Kessler *et al.* (1983) also indicated a type II′ β-turn. Huang *et al.* (1992) and He, Huang, Raynor, Reisine & Goodman (1993), using 500 MHz two-dimensional NMR, binding studies and conformational analysis of methylated derivatives, showed that although the side chains are conformationally flexible, the biologically active conformation of this cyclic hexapeptide probably involves the *trans* rotamer of D-Trp^8 and the *gauche*$^-$ (g^-) rotamer of Lys^9.

* Author to whom correspondence should be addressed.

© 1995 International Union of Crystallography
Printed in Great Britain – all rights reserved

Possibly the most promising candidate for medical applications is the cyclic octapeptide octreotide (Fig. 1) (Bauer *et al.*, 1982). In octreotide the active tetrapeptide fragment is conformationally restricted by the disulfide bridge. The aromatic side chain of Phe1 in octreotide can occupy a similar region of space to Phe^6 in somatostatin, and it protects the disulfide bridge against enzymatic attack. In octreotide the terminal Thr-COOH group is reduced to an alcohol function, which should also be more stable to enzymatic degradation (Bauer *et al.*, 1982). Octreotide has a similar overall activity to somatostatin, but it is more selective in the inhibition of growth-hormone secretion as compared to the inhibition of insulin secretion. Octreotide is successfully used in the treatment of acromegaly, thyrotrophinomas, the carcinoid syndrome and various rare endocrine tumors (Lamberts, Uitterlinden, Verschoor, van Dongen & del Pozo, 1985; Reichlin, 1987; Battershill & Clissold, 1989; Pless, 1990; Koop, 1991). It is also being tested in the treatment of advanced breast cancer and prostate cancer (Manni, 1992).

Here we present the first X-ray crystal structure of a somatostatin analogue peptide. It provides a basis for further studies on the mechanism of the activity of octreotide and other somatostatin analogues.

Experimental

Crystallization of octreotide oxalate salt

1.14 g (1.0 mmol) of octreotide diacetate salt was dissolved in 5.25 ml (1.05 mmol) 0.2 *M* aqueous oxalic acid, concentrated under reduced pressure on a rotary evaporator, and the remaining substance dried at 323 K and 0.01 mbar. Dissolution in 5 ml H_2O and evaporation were repeated twice. The residue was dissolved in 2.5 ml H_2O, filtered and transferred to a flat-bottomed polystyrene flask. Upon standing overnight at 293 K the oxalate precipitated in the form of tiny crystals. In order to grow larger crystals this suspension was warmed twice a day to 303 K under careful rotation until about one third of the crystals had dissolved and were allowed to stand and grow again at 293 K. After 4 weeks crystals as long as 2 mm were obtained.

For analytical purposes a part of this crystal suspension was filtered off, washed with cold water and dried at 293 K and 0.01 mbar. Microanalytical determination of oxalate by ion chromatography indicated 105 μg of oxalate per mg of crystals corresponding to a 3:4 molar ratio of peptide to oxalic acid.

Elemental analysis of the dried material $C_{155}H_{206}N_{30}O_{46}S_6$ (3417.91), %found (calculated): C 54.4 (54.5), H 6.4 (6.1), N 12.4 (12.3), O 21.2 (21.5), S 5.6 (5.6).

X-ray data collection

The single crystal used for data collection (size 2.0 × 0.4 × 0.3 mm) was sealed in a Lindemann capillary with some mother liquor to prevent rapid decomposition by solvent loss. Data were collected at 285 K using synchrotron radiation from the EMBL beamline X31 on the DORIS storage ring at DESY in Hamburg. The EMBL imaging-plate scanner (220 mm diameter) was employed as a detector. The strategy of data collection was based on the rotation method developed for photographic film (Arndt & Wonacott, 1977). The wavelength used was 0.700 Å. Data were collected in three passes with the same crystal at high, medium and low resolution using different exposure times. High-resolution data were collected at a crystal-to-detector distance of 110 mm with 90° φ range and a φ increment of 2°, medium resolution at 220 mm and φ increment of 4°, low resolution at 350 mm and φ increment 8°. Multiple passes were needed as the stronger (low-resolution) reflections saturated the electronic scanner read-out and thus some pixels were unusable when the exposure time was long enough to obtain significant intensities for the weak (high-resolution) reflections. The images were subsequently processed using the program *DENZO* (Otwinowski, 1991). Only fully recorded reflections were accepted from each image, symmetry-equivalent reflections and Friedel opposites were not merged during data processing. No absorption or decay correction was applied.

Accurate cell dimensions were obtained from another crystal sealed in a capillary on a Stoe–Huber four-circle diffractometer with graphite-monochromated Mo $K\alpha$ radiation (λ = 0.71073 Å) at room temperature using 30 strong reflections centered at $+2\theta$, $+\omega$ and -2θ, $-\omega$ in the range $15° < 2\theta < 20°$. The relatively large standard deviation in the *c*-axis length is an indirect consequence of the weakness of the reflections used for cell determination, in particularly those with high $|l|$ values. Crystal data are summarized in Table 1.

H-Ala[1]-Gly[2]-Cys[3]-Lys[4]-Asn[5]-Phe[6]-Phe[7]-Trp[8]-Lys[9]-Thr[10]-Phe[11]-Thr[12]-Ser[13]-Cys[14]-OH

(*a*)

H-(D)-Phe- Cys-Phe-(D)-Trp-Lys-Thr-Cys-Thr(ol)

(*b*)

Fig. 1. The amino-acid sequences of (*a*) somatostatin and (*b*) octreotide.

Table 1. *Crystal data for octreotide*

Formula	$C_{149}H_{204}N_{30}O_{86}S_6$
Molecular mass (g mol^{-1})	3983.79
Crystal size (mm)	2 × 0.4 × 0.3
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
Temperature (K)	285
a (Å)	18.458 (5)
b (Å)	30.009 (7)
c (Å)	39.710 (27)
V (Å^3)	21993 (20)
Z	4
ρ_{calc} (Mg m^{-3})	1.203
μ (mm^{-1})	0.153
$F(000)$	8368
No. of reflections*	25806
Independent	18951
R_{int}	0.026
No. of data	18930
Restraints	3296
Parameters	2441
$wR2$ (all data)	0.246
$R1$ [$F > \sigma(F)$]	0.0843
$g1$, $g2$	0.1794, 8.3
Abs. structure par.	0.41 (10)
Max. (e Å^{-3})	0.351
Min. (e Å^{-3})	−0.590

* After data processing.

Structure solution

The octreotide structure was probably too large for direct methods when this work was performed in 1992. The critical factor influencing the probabilities of the phase relations is the number of atoms per (primitive) unit cell. With 21 993 Å^3 the cell is larger than for the proteins rubredoxin (19 184 Å^3) or crambin (16 838 Å^3) which have only recently been solved by direct methods (Sheldrick, Dauter, Wilson, Hope & Sieker, 1993; Hauptman, Weeks, Smith, Teeter & Miller, 1993). The FeS_4 cluster of rubredoxin plays an important role in enabling this structure to be solved by direct methods, and it appears that the six S atoms were also significant in the solution of crambin. The data for these two proteins extend to somewhat higher resolution (*ca* 0.9 Å) than for octreotide (1.04 Å). Despite exhaustive attempts, we were not successful in solving the octreotide structure by direct methods.

The structure was finally solved by the Patterson vector-superposition minimum-function approach using *SHELXS*92 (Sheldrick, 1992). First a sharpened Patterson function was calculated with coefficients $(E^3F)^{1/2}$, and the highest 20 peaks that were more than 6 Å from the origin were selected as superposition vectors. For each vector a superposition minimum function was calculated; in the ideal case this function should correspond to a double image of the structure, one image being the inverse of the other. To deconvolute this double image it is necessary to find an origin shift so that the space-group symmetry is fulfilled for one of the two images. Possible origin shifts were examined automatically by the program, and for each combination of superposition vector and origin shift a list of potential atomic sites was generated, assigning the potential atoms as S or as light atoms depending on the height of the contributing peaks in the superposition map. These solutions were then ranked in terms of the correlation coefficient between E_c and E_o defined by Fujinaga & Read (1987). The two solutions with the highest correlation coefficients both showed all three disulfide bridges. The following tangent phase expansion followed by iterative E-Fourier recycling (Sheldrick, 1982) revealed two almost complete peptide molecules and most of the third molecule.

Structure refinement

The structure was refined on F^2 with the program *SHELXL*93 (Sheldrick, 1993) using three overlapping blocks for the least-squares matrix. All non-H atoms were refined anisotropically. The H atoms were included in calculated positions with idealized geometry and refined using a riding model. All the NH_2 groups were considered to be protonated to NH_3^+. Rigid-bond restraints (Rollett, 1970; Hirshfeld, 1976; Trueblood & Dunitz, 1983) were applied as 'hard' restraints (e.s.d.'s 0.01 Å^2) with target values of zero for the differences of the anisotropic displacement components along the 1–2 and 1–3 vector directions. 'Similarity' restraints were applied to the anisotropic displacement parameters of atoms within 1.7 Å of each other to restrain the corresponding U_{ij} components to be equal with 'soft' e.s.d.'s of 0.05 Å^2 (and 0.10 Å^2 for terminal atoms). For the oxalate dianion the restraints applied on the anisotropic displacement parameters had to be tightened to e.s.d.'s of 0.002 Å^2 (rigid-bond restraints) and 0.01 Å^2 (similarity restraints). Chemically equivalent 1,2- and 1,3-distances in the three crystallographically independent molecules and in the two Thr and Phe residues within each molecule were restrained to be the same with e.s.d.'s of 0.03 Å. In addition the chiral volumes of all carbonyl C atoms were restrained to zero (with e.s.d.'s 0.1 Å^3) and the phenyl and tryptophan rings were restrained to be flat by restraining the volumes of individual atomic tetrahedra to be zero (with e.s.d.'s 0.1 Å^3). The weights for the restraints were calibrated with the R_{free} test suggested by Brünger (1992; 1993), and the above values (which are the program default values) proved to be reasonable.

'Anti-bumping' restraints were applied to prevent close contacts between isolated solvent molecules, and 'Babinet's principle' was employed to model diffuse solvent (Langridge *et al.*, 1960; Driessen *et al.*, 1989). The anisotropic displacement parameters of the water O atoms were restrained to be approximately isotropic (with e.s.d.'s 0.10 Å^2). All water positions were refined as fully occupied. In a test in which all the water occupancies were allowed to refine freely, subject to limiting values of 0.0 and 1.0, R_{free} decreased by only 0.09%, which we do not consider to be significant. In

this test 28 of the occupancies adopted the limiting value of 1.0, and the remainder refined to values in the range 0.47 to 1.0.

The refinement converged to a final $wR2 = [\Sigma w(F_o^2 - F_c^2)/\Sigma wF_o^4]^{1/2} = 0.2425$ for all 18 930 data between 1.04 and 10.0 Å resolution for a total of 2441 least-squares parameters. The weighting scheme was $w^{-1} = \sigma^2(F_o)^2 + (0.1794P)^2 + 8.30P$, where $P = (F_o^2 + 2F_c^2)/3$, and the goodness of fit $S = \{\Sigma[w(F_o^2 - F_c^2)^2]/(n - p)\}^{1/2}$ was 1.063. 3296 restraints were employed. The corresponding values of $wR2$ and S for the full resolution range (18 951 observations) were 0.246 and 1.063. The conventional R index based on F ($R1 = \Sigma\big||F_o| - |F_c|\big| / \Sigma|F_o|$) was 0.0843 for the 16 633 data with $I > 2\sigma(I)$ and 0.0953 for all 18 951 data. The known absolute configuration was assumed throughout the refinement; the Flack x parameter (Flack, 1983) refined to 0.4 (1); in view of the short synchrotron wavelength, it is not surprising that this value does not provide a satisfactory discrimination between enantiomorphs.

Results and discussion

The final atomic parameters for the non-H atoms are given in Table 2. The primary structure of the octapeptide molecule is shown in Fig. 2. Residue numbers 1–8 refer to molecule I, 9–16 to molecule II and 17–24 to molecule III. Stereoviews of the three crystallographically independent molecules are shown in Fig. 3. The 50% probability atomic displacement ellipsoids are illustrated in Fig 4. Selected mean bond lengths and angles are listed in Table 3, and important torsion angles are summarized in Table 4.

Molecular conformations

Molecule I forms a regular antiparallel β-pleated sheet with a type II′ β-turn spanning Trp4 and Lys5, stabilized by three intramolecular hydrogen bonds O1···HN8, O3···HN5 and O6···HN3. The hydrogen bond across the β-turn is rather long (O···HN = 3.56 Å). The overall shape of molecule I is flat with the Trp4 and Lys5 side chains pointing in one direction and Phe1 and Thr(ol)8 in the opposite direction. Molecules II and III also adopt a type II′ β-turn for the -Trp-Lys- unit, with more normal hydrogen-bond distances across the turn, but the tryptophan carbonyl group takes part in an additional hydrogen bond to the cysteine NH (O12···HN15 and O20···HN23 in molecules II and III, respectively), that gives rise to a quite different backbone conformation to that observed in I. Furthermore OG1 of the threonine side chain is within hydrogen-bonding distance of the phenylalanine NH (Phe11 in molecule II and Phe19 in molecule III). The overall shape of molecules II and III is more globular than that of molecule I but is still relatively flat. The geometries of the intramolecular hydrogen bonds in all three molecules are given in Table 5.

Table 2. *Atomic coordinates* ($\times 10^4$) *and equivalent isotropic displacement parameters* ($\text{Å}^2 \times 10^3$)

$$U_{eq} = (1/3)\Sigma_i \Sigma_j U_{ij} a_i^* a_j^* \mathbf{a}_i . \mathbf{a}_j .$$

	x	y	z	U_{eq}
O(1S)	2875 (5)	5831 (3)	470 (3)	202 (4)
O(2S)	1486 (4)	6551 (3)	3379 (3)	197 (4)
O(3S)	−38 (6)	6067 (4)	−2061 (3)	209 (4)
O(4S)	−2522 (6)	6025 (4)	−541 (2)	209 (4)
O(5S)	−1996 (9)	9132 (5)	251 (3)	261 (6)
O(6S)	−1663 (4)	4022 (3)	1652 (3)	194 (4)
O(7S)	1771 (9)	6250 (5)	4173 (3)	263 (6)
O(8S)	3745 (20)	−3360 (8)	679 (5)	438 (17)
O(9S)	−2077 (11)	6455 (12)	4103 (6)	473 (20)
O(10S)	1828 (12)	497 (7)	1595 (5)	334 (9)
O(11S)	−1204 (10)	6901 (10)	−2438 (6)	428 (16)
O(12S)	−580 (12)	3046 (5)	585 (5)	352 (11)
O(13S)	5305 (16)	4980 (13)	2865 (6)	468 (18)
O(14S)	−4690 (13)	7162 (10)	825 (4)	469 (19)
O(15S)	3125 (10)	6431 (5)	2456 (4)	293 (8)
O(16S)	−5060 (17)	7289 (13)	3299 (7)	476 (19)
O(17S)	7958 (12)	4065 (9)	2321 (7)	388 (12)
O(18S)	6181 (32)	7306 (17)	3104 (9)	683 (35)
O(19S)	−746 (30)	4768 (13)	−1129 (8)	647 (30)
O(20S)	−2426 (14)	6524 (13)	−1875 (8)	493 (19)
O(21S)	2083 (14)	−1026 (18)	1464 (11)	615 (28)
O(22S)	3955 (19)	3333 (10)	327 (9)	509 (21)
O(23S)	6553 (24)	1188 (18)	977 (12)	698 (35)
O(24S)	8142 (12)	3411 (10)	1069 (7)	396 (12)
O(25S)	−5588 (16)	6321 (9)	2608 (5)	399 (13)
O(26S)	6013 (21)	3802 (9)	−213 (7)	469 (18)
O(27S)	4992 (39)	810 (24)	816 (15)	802 (49)
O(28S)	−1117 (28)	3848 (17)	−566 (13)	665 (35)
O(29S)	1000 (21)	3323 (18)	−138 (13)	621 (29)
O(30S)	7055 (15)	2849 (12)	1789 (11)	502 (20)
O(31S)	1807 (16)	7697 (12)	1021 (11)	526 (21)
O(32S)	6414 (16)	4496 (14)	2473 (8)	537 (24)
O(33S)	5705 (19)	838 (17)	1760 (8)	659 (31)
O(34S)	6169 (15)	3628 (7)	450 (5)	373 (12)
O(35S)	2051 (11)	45 (17)	2079 (8)	614 (28)
O(36S)	5259 (19)	2881 (6)	−431 (3)	401 (16)
O(37S)	−1776 (27)	3032 (13)	−621 (7)	565 (26)
O(38S)	2404 (21)	−3186 (17)	451 (12)	670 (33)
O(39S)	2665 (21)	3183 (14)	−58 (7)	510 (21)
O(40S)	5420 (38)	2811 (15)	556 (9)	676 (33)
O(41S)	2147 (24)	7041 (16)	3758 (13)	664 (34)
O(42S)	7704 (18)	3645 (9)	207 (12)	534 (22)
O(43S)	1810 (23)	−2591 (10)	2247 (12)	557 (24)
O(44S)	7837 (14)	1829 (14)	2231 (7)	471 (18)
O(45S)	7335 (38)	2705 (18)	−198 (11)	789 (47)
O(46S)	2948 (22)	−2865 (12)	1252 (10)	520 (21)
O(47S)	2355 (9)	−66 (8)	764 (5)	335 (10)
O(48S)	7588 (18)	3153 (18)	2426 (11)	597 (28)
O(49S)	953 (20)	7099 (13)	4337 (7)	504 (21)
O(50S)	1084 (19)	−2598 (8)	1577 (11)	543 (22)
O(51S)	−2352 (34)	4239 (15)	−472 (13)	757 (40)
O(52S)	3296 (11)	−2762 (11)	1954 (6)	406 (14)
C(2X)	−553 (7)	5535 (4)	4092 (3)	190 (3)
C(1X)	−386 (6)	5086 (4)	4109 (3)	162 (3)
O(11X)	260 (4)	4960 (2)	4095 (2)	144 (2)
O(12X)	−898 (5)	4824 (3)	4143 (3)	193 (3)
O(21X)	−95 (5)	5823 (2)	4078 (2)	174 (3)
O(22X)	−1211 (6)	5648 (4)	4128 (3)	268 (4)
N1	−652 (6)	6643 (2)	3997 (2)	152 (3)
CA1	−312 (5)	6833 (2)	3698 (2)	102 (2)
C1	−540 (4)	6564 (2)	3390 (2)	86 (2)
O1	−1180 (3)	6539 (2)	3321 (2)	123 (2)
CB1	−522 (6)	7313 (2)	3674 (2)	115 (3)
CG1	−172 (5)	7566 (2)	3393 (2)	110 (2)
CD11	555 (5)	7647 (3)	3392 (2)	135 (3)
CE11	846 (6)	7889 (3)	3128 (3)	162 (4)
CZ1	473 (7)	8040 (3)	2883 (3)	152 (4)
CE21	−245 (6)	7966 (3)	2869 (2)	137 (3)
CD21	−572 (5)	7728 (2)	3125 (2)	116 (2)
N2	−24 (3)	6377 (2)	3219 (1)	84 (1)
CA2	−148 (4)	6104 (2)	2925 (1)	77 (2)
C2	87 (3)	5634 (2)	2984 (1)	76 (2)

Table 2 (*cont.*)

	x	y	z	U_{eq}		x	y	z	U_{eq}
O2	567 (3)	5545 (2)	3178 (1)	98 (1)	CA11	−3260 (3)	6190 (2)	971 (2)	81 (2)
CB2	274 (4)	6284 (2)	2615 (2)	95 (2)	C11	−2791 (3)	6306 (2)	673 (1)	76 (2)
SG2	128 (1)	6864 (1)	2528 (1)	104 (1)	O11	−2981 (3)	6589 (2)	465 (1)	102 (2)
N3	−236 (3)	5327 (2)	2795 (1)	78 (1)	CB11	−3886 (4)	5888 (3)	842 (2)	108 (2)
CA3	5 (4)	4872 (2)	2794 (1)	76 (2)	CG11	−3634 (3)	5517 (3)	623 (2)	100 (2)
C3	−198 (3)	4666 (2)	2455 (1)	70 (2)	CD111	−3328 (4)	5141 (3)	753 (2)	124 (3)
O3	−835 (2)	4629 (2)	2374 (1)	105 (2)	CE111	−3082 (5)	4805 (3)	547 (3)	149 (3)
CB3	−308 (4)	4603 (2)	3082 (2)	92 (2)	CZ11	−3142 (6)	4850 (4)	212 (3)	174 (5)
CG3	−5 (4)	4137 (2)	3100 (1)	85 (2)	CE211	−3431 (6)	5202 (4)	71 (2)	163 (4)
CD13	669 (4)	4061 (3)	3214 (2)	113 (2)	CD211	−3689 (5)	5553 (3)	276 (2)	126 (3)
CE13	958 (5)	3642 (3)	3235 (2)	136 (3)	N12	−2207 (3)	6066 (2)	633 (1)	70 (1)
CZ3	551 (6)	3295 (3)	3137 (2)	137 (3)	CA12	−1843 (4)	6041 (2)	305 (2)	76 (2)
CE23	−115 (6)	3356 (3)	3027 (2)	138 (3)	C12	−1517 (4)	6476 (2)	213 (2)	87 (2)
CD23	−403 (5)	3775 (3)	3005 (2)	114 (2)	O12	−1167 (3)	6700 (2)	418 (1)	101 (2)
N4	337 (2)	4540 (2)	2261 (1)	65 (1)	CB12	−1284 (4)	5675 (2)	318 (2)	86 (2)
CA4	216 (3)	4342 (2)	1935 (1)	62 (1)	CG12	−929 (4)	5587 (3)	−7 (2)	99 (2)
C4	45 (3)	4699 (2)	1683 (1)	60 (1)	CD112	−1133 (5)	5286 (3)	−243 (2)	154 (4)
O4	486 (2)	4994 (1)	1618 (1)	80 (1)	NE112	−684 (5)	5293 (3)	−515 (2)	179 (4)
CB4	900 (3)	4086 (2)	1836 (2)	73 (2)	CE212	−168 (6)	5599 (4)	−459 (2)	192 (5)
CG4	845 (3)	3816 (2)	1523 (2)	75 (2)	CD212	−319 (5)	5781 (3)	−143 (2)	146 (4)
CD14	248 (4)	3711 (2)	1327 (2)	89 (2)	CE312	153 (6)	6110 (4)	−22 (3)	226 (8)
NE14	437 (4)	3443 (2)	1073 (1)	99 (2)	CZ312	716 (7)	6235 (5)	−216 (4)	313 (11)
CE24	1154 (4)	3349 (2)	1092 (2)	89 (2)	CH212	844 (8)	6046 (6)	−524 (4)	331 (11)
CD24	1422 (3)	3572 (2)	1373 (2)	78 (2)	CZ212	412 (7)	5724 (5)	−659 (3)	299 (10)
CE34	2148 (4)	3529 (2)	1457 (2)	101 (2)	N13	−1605 (4)	6607 (2)	−107 (1)	114 (2)
CZ34	2581 (5)	3264 (2)	1256 (2)	135 (3)	CA13	−1286 (8)	6989 (3)	−241 (2)	177 (4)
CH24	2298 (6)	3053 (3)	987 (2)	136 (3)	C13	−1595 (9)	7416 (3)	−97 (2)	215 (5)
CZ24	1593 (5)	3079 (2)	889 (2)	120 (3)	O13	−1243 (10)	7766 (3)	−159 (3)	308 (8)
N5	−612 (2)	4688 (2)	1542 (1)	63 (1)	CB13	−1365 (12)	6995 (4)	−630 (2)	253 (8)
CA5	−846 (3)	4999 (2)	1287 (1)	64 (1)	CG13	−910 (13)	6657 (6)	−798 (3)	284 (9)
C5	−1434 (3)	5305 (2)	1408 (1)	64 (1)	CD13	−1111 (16)	6715 (10)	−1177 (3)	362 (12)
O5	−1778 (2)	5533 (2)	1200 (1)	87 (1)	CE13	−642 (16)	6575 (11)	−1395 (3)	366 (13)
CB5	−1064 (4)	4748 (2)	965 (2)	83 (2)	NZ13	−934 (7)	6695 (6)	−1750 (2)	247 (7)
CG5	−429 (4)	4600 (3)	758 (2)	99 (2)	N14	−2177 (7)	7396 (3)	86 (2)	213 (5)
CD5	−631 (5)	4312 (4)	467 (2)	142 (4)	CA14	−2508 (8)	7767 (4)	254 (3)	256 (5)
CE5	−19 (7)	4178 (6)	260 (3)	198 (6)	C14	−2064 (8)	7893 (3)	566 (2)	209 (6)
NZ5	−209 (9)	3828 (6)	11 (4)	297 (9)	O14	−2134 (7)	8267 (2)	694 (2)	211 (5)
N6	−1563 (3)	5332 (2)	1736 (1)	70 (1)	CB14	−3287 (8)	7711 (4)	300 (3)	275 (6)
CA6	−2083 (3)	5632 (2)	1879 (2)	73 (2)	OG114	−3478 (7)	7421 (4)	581 (3)	274 (5)
C6	−1751 (3)	5893 (2)	2161 (2)	78 (2)	CG214	−3764 (13)	7605 (7)	18 (5)	353 (12)
O6	−1239 (3)	5751 (2)	2330 (1)	102 (2)	N15	−1603 (6)	7600 (2)	696 (2)	152 (3)
CB6	−2747 (4)	5389 (3)	2000 (2)	102 (2)	CA15	−1188 (6)	7677 (3)	990 (2)	128 (3)
OG16	−2574 (4)	5104 (3)	2262 (2)	158 (3)	C15	−473 (7)	7907 (3)	906 (2)	158 (3)
CG26	−3131 (5)	5148 (3)	1740 (2)	126 (3)	O15	111 (6)	7699 (3)	916 (3)	204 (4)
N7	−2066 (3)	6283 (2)	2222 (1)	82 (1)	CB15	−1026 (5)	7257 (3)	1191 (2)	109 (2)
CA7	−1894 (3)	6547 (2)	2516 (2)	83 (2)	SG15	−1790 (1)	6960 (1)	1354 (1)	102 (1)
C7	−2612 (4)	6662 (2)	2685 (2)	92 (2)	N16	−518 (7)	8329 (2)	819 (2)	166 (3)
O7	−3149 (3)	6758 (2)	2521 (1)	111 (2)	CA16	94 (7)	8605 (3)	728 (2)	170 (4)
CB7	−1490 (4)	6970 (2)	2438 (2)	96 (2)	CT16	422 (8)	8801 (4)	1045 (3)	179 (5)
SG7	−685 (1)	6893 (1)	2187 (1)	100 (1)	O16	1088 (6)	9002 (3)	964 (3)	199 (4)
N8	−2598 (3)	6659 (2)	3022 (1)	103 (2)	CB16	−107 (9)	8942 (3)	461 (3)	196 (5)
CA8	−3207 (5)	6746 (3)	3239 (2)	126 (3)	OG116	−694 (6)	9198 (2)	591 (2)	184 (4)
CT8	−3014 (6)	7079 (4)	3501 (2)	139 (3)	CG216	−335 (12)	8739 (4)	140 (2)	252 (9)
O8	−3619 (7)	7214 (5)	3678 (3)	277 (6)	N17	1228 (4)	5545 (3)	3802 (2)	122 (2)
CB8	−3499 (8)	6324 (4)	3384 (4)	194 (5)	CA17	1733 (4)	5218 (3)	3656 (2)	101 (2)
OG18	−2948 (8)	6115 (4)	3585 (3)	262 (5)	C17	2178 (4)	5432 (3)	3379 (2)	100 (2)
CG28	−3728 (9)	6002 (5)	3135 (5)	240 (7)	O17	2561 (6)	5747 (3)	3439 (2)	192 (4)
N9	−5327 (4)	7253 (4)	2471 (3)	181 (4)	CB17	2227 (5)	5038 (3)	3937 (2)	117 (2)
CA9	−4558 (5)	7381 (3)	2407 (2)	125 (3)	CG17	2672 (4)	4662 (3)	3808 (2)	104 (2)
C9	−4444 (4)	7370 (3)	2031 (2)	118 (3)	CD117	3336 (4)	4741 (3)	3665 (2)	118 (2)
O9	−4876 (4)	7548 (3)	1850 (2)	169 (3)	CE117	3735 (5)	4381 (4)	3543 (2)	138 (3)
CB9	−4439 (5)	7848 (3)	2547 (2)	140 (3)	CZ17	3498 (5)	3968 (4)	3565 (2)	137 (3)
CG9	−3676 (5)	8006 (3)	2513 (2)	119 (2)	CE217	2848 (6)	3885 (3)	3710 (2)	144 (3)
CD19	−3176 (6)	7912 (3)	2757 (2)	148 (3)	CD217	2431 (5)	4226 (3)	3829 (2)	124 (3)
CE19	−2464 (6)	8063 (3)	2720 (3)	147 (3)	N18	2167 (3)	5256 (2)	3089 (1)	97 (2)
CZ9	−2267 (6)	8294 (3)	2446 (3)	154 (3)	CA18	2569 (4)	5401 (3)	2791 (2)	94 (2)
CE29	−2744 (6)	8388 (3)	2210 (2)	141 (3)	C18	2063 (3)	5402 (2)	2491 (2)	78 (2)
CD29	−3447 (5)	8241 (3)	2240 (2)	132 (3)	O18	1568 (3)	5132 (2)	2474 (1)	104 (2)
N10	−3893 (4)	7150 (2)	1916 (2)	112 (2)	CB18	3222 (4)	5106 (3)	2731 (2)	101 (2)
CA10	−3712 (5)	7141 (3)	1554 (2)	114 (2)	SG18	3009 (1)	4525 (1)	2699 (1)	104 (1)
C10	−3469 (4)	6672 (2)	1458 (2)	86 (2)	N19	2215 (3)	5690 (2)	2249 (1)	82 (1)
O10	−3187 (3)	6427 (2)	1661 (1)	90 (1)	CA19	1808 (3)	5711 (2)	1935 (1)	74 (2)
CB10	−3160 (5)	7483 (3)	1464 (2)	125 (2)	C19	2247 (3)	5526 (2)	1639 (2)	75 (2)
SG10	−2288 (2)	7411 (1)	1661 (1)	120 (1)	O19	2846 (2)	5670 (2)	1574 (1)	92 (1)
N11	−3555 (3)	6578 (2)	1136 (1)	93 (2)	CB19	1594 (4)	6189 (2)	1855 (2)	91 (2)

Table 2 (*cont.*)

	x	y	z	U_{eq}
CG19	1116 (3)	6219 (2)	1553 (2)	82 (2)
CD119	378 (4)	6170 (2)	1582 (2)	93 (2)
CE119	−70 (4)	6190 (2)	1303 (2)	109 (2)
CZ19	186 (5)	6260 (3)	1001 (2)	121 (3)
CE219	913 (6)	6312 (3)	958 (2)	152 (4)
CD219	1373 (4)	6291 (3)	1236 (2)	121 (3)
N20	1917 (3)	5230 (2)	1455 (1)	75 (1)
CA20	2180 (3)	5091 (2)	1123 (2)	77 (2)
C20	2945 (3)	4942 (2)	1137 (2)	78 (2)
O20	3164 (2)	4671 (2)	1348 (1)	87 (1)
CB20	1699 (4)	4711 (3)	997 (2)	113 (3)
CG20	1896 (6)	4522 (4)	673 (2)	155 (3)
CD120	2272 (6)	4131 (4)	584 (3)	181 (4)
NE120	2331 (6)	4075 (4)	266 (3)	226 (5)
CE220	1997 (5)	4432 (5)	96 (2)	209 (5)
CD220	1758 (5)	4680 (4)	361 (2)	177 (4)
CD320	1378 (5)	5074 (5)	277 (3)	214 (5)
CZ320	1297 (7)	5170 (6)	−48 (3)	268 (8)
CH220	1551 (7)	4914 (7)	−299 (4)	302 (9)
CZ220	1927 (7)	4512 (6)	−235 (3)	273 (8)
N21	3385 (3)	5108 (2)	897 (1)	81 (1)
CA21	4124 (3)	4960 (2)	858 (2)	84 (2)
C21	4587 (4)	5074 (2)	1162 (2)	85 (2)
O21	5158 (3)	4864 (2)	1207 (1)	107 (2)
CB21	4460 (5)	5137 (3)	536 (2)	117 (3)
CG21	4132 (6)	4967 (3)	215 (2)	125 (3)
CD21	4385 (9)	4510 (4)	138 (3)	189 (5)
CE21	4070 (13)	4327 (5)	−165 (4)	243 (8)
NZ21	4387 (14)	3874 (4)	−223 (3)	320 (11)
N22	4388 (3)	5389 (2)	1368 (1)	85 (1)
CA22	4787 (3)	5487 (2)	1671 (2)	88 (2)
C22	4788 (4)	5109 (3)	1921 (2)	99 (2)
O22	5235 (3)	5090 (3)	2152 (2)	142 (2)
CB22	4565 (4)	5916 (3)	1835 (2)	103 (2)
OG122	3901 (4)	5882 (3)	2006 (2)	155 (2)
CG222	4576 (6)	62767 (3)	1593 (3)	134 (3)
N23	4287 (3)	4799 (2)	1884 (1)	89 (2)
CA23	4287 (3)	4405 (2)	2094 (2)	95 (2)
C23	4889 (4)	4086 (3)	2015 (2)	107 (2)
O23	5063 (3)	3801 (2)	2228 (2)	137 (2)
CB23	3575 (4)	4157 (3)	2054 (2)	94 (2)
SG23	2782 (1)	4453 (1)	2198 (1)	95 (1)
N24	5186 (4)	4113 (3)	1709 (2)	125 (2)
CA24	5684 (6)	3791 (3)	1577 (3)	160 (3)
CT24	6449 (6)	3973 (6)	1572 (6)	272 (2)
O24	6947 (8)	3672 (8)	1467 (7)	408 (13)
CB24	5454 (9)	3654 (6)	1230 (4)	233 (6)
OG124	5698 (9)	4027 (5)	1015 (3)	300 (7)
CG224	4726 (10)	3552 (8)	1175 (6)	332 (12)

Table 3. *Mean values for bond lengths and angles* (Å, °)

C—N	1.317	CA—C—N	116.9
C=O	1.229	N—C—O	122.7
C—CA	1.511	C—N—CA	123.2
CA—N	1.448	CA—C—O	120.4
CA—CB	1.516	N—CA—C	110.7

Table 4. *Torsion angles* (°)

	φ	ψ	ω	χ_1	χ_2
D-Phe1		−120.3 (8)			
Cys2	−115.9 (7)	155.3 (6)	178.8 (6)		
Phe3	−153.4 (6)	115.8 (6)	171.6 (5)	−174.8 (6)	74.4 (8)
D-Trp4	79.0 (6)	−116.6 (5)	−179.8 (5)	173.3 (5)	−9.4 (10)
Lys5	−112.4 (6)	14.5 (7)	−177.6 (5)	−78.2 (7)	173.4 (8)
Thr6	−129.7 (6)	155.0 (5)	176.4 (5)	64.6 (8)	
Cys7	−128.6 (6)	142.2 (6)	170.0 (6)		
Thr8			−178.3 (7)		
D-Phe9		−129.6 (9)			
Cys10	−140.2 (8)	155.7 (7)	−176.4 (8)		
Phe11	−123.9 (7)	140.8 (6)	169.1 (7)	171.0 (6)	78.2 (9)
D-Trp12	67.6 (8)	−137.0 (7)	159.5 (5)	174.4 (6)	−92.2 (10)
Lys13	−68.4 (14)	−11.0 (14)	−175.0 (9)	−70.4 (19)	177.4 (15)
Thr14	−75.0 (13)	−19.0 (15)	176.3 (9)	79.3 (13)	
Cys15	−87.7 (13)	72.5 (10)	−177.6 (11)		
Thr16			−179.5 (8)		
D-Phe17		−124.0 (8)			
Cys18	−134.4 (8)	151.6 (6)	−178.2 (7)		
Phe19	−107.3 (7)	129.4 (6)	175.8 (6)	−174.5 (5)	86.0 (7)
D-Trp20	53.8 (8)	−133.0 (6)	165.8 (5)	−178.2 (8)	100.8 (12)
Lys21	−64.4 (8)	−20.7 (8)	−173.5 (5)	−65.9 (10)	−77.7 (12)
Thr22	−64.0 (8)	−18.1 (9)	175.6 (5)	−51.6 (8)	
Cys23	−71.3 (9)	−22.1 (10)	174.4 (6)		
Thr24			−170.0 (8)		
'Folded'					
Phe[7]*	−74	83		−54	97
Trp[8]	69	−139		178	−86
Lys[9]	−62	−35		−62	−179
Thr[10]	−84	75		58	
'Flat'					
Phe[7]	−142	167		−59	100
Trp[8]	88	−141		180	−87
Lys[9]	−66	−23		−62	−179
Thr[10]	−142	176		62	

* The superscript numbers refer to somatostatin.

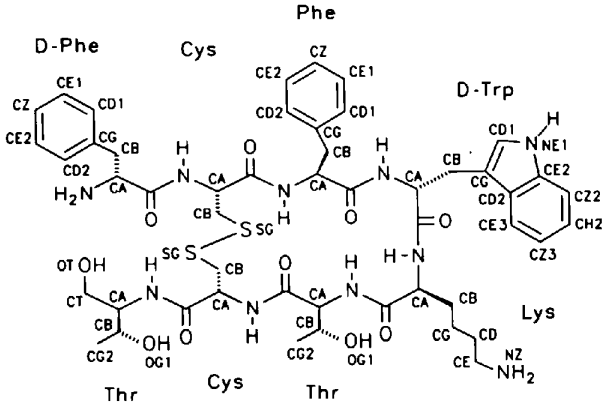


Fig. 2. Primary structure of octreotide.

The r.m.s. deviation in the least-squares superposition of the main-chain ring atoms of molecules I and III is 2.84 Å, whereas the corresponding value for molecules II and III is only 0.58 Å. The least-squares superposition of molecule II and III is shown in Fig. 5. The orientation of the terminal amino acids Phe and Thr(ol) is different from that in molecule I, and they are not stabilized by an intramolecular hydrogen bond. The peptide linkages in all three molecules are essentially planar except the bonds between Phe11-Trp12 (ω = 159.1°) and Phe19-Trp20 (ω = 165.5°). The mean deviations of the ω-angles from 180° are 3.3, 6.7 and 6.7° for molecules I, II and III, respectively. The bond distances and angles of all three molecules are in the normal range for small peptides.

Table 5. *Geometries of intramolecular hydrogen bonds*

	$D\cdots A$ (Å)	D—H$\cdots A$ (°)
O1···N8	2.895 (9)	172.7 (3)
O3···N6	3.559 (8)	167.3 (1)
O6···N3	2.909 (7)	161.4 (2)
O11···N14	3.214 (13)	149.4 (2)
O12···N15	3.026 (8)	165.0 (3)
O19···N22	3.078 (7)	136.5 (2)
O20···N23	2.996 (7)	160.5 (2)
OG114···N11	3.357 (14)	134.7 (3)
OG122···N19	3.311 (8)	130.5 (2)

Table 6. *Geometries of intermolecular hydrogen bonds*

	$D\cdots A$ (Å)	D—H$\cdots A$ (°)
O2···N17	2.762 (8)	145.0 (2)
O4···N20	2.811 (6)	161.0 (2)
O5···N12	2.871 (6)	171.1 (2)
O7···N10	3.007 (7)	161.6 (2)
O10···N7	3.071 (7)	162.7 (2)
O18···N4	3.004 (6)	140.5 (2)

Table 7. *Selected intermolecular hydrogen bonds involving the amino groups*

	$D\cdots A$ (Å)		$D\cdots A$ (Å)
N1···O21	2.69 (1)	N1···O22X	3.20 (2)
N1···O36S^{i}	2.78 (2)	N1···O9S	2.72 (2)
NZ5···O28S	2.84 (4)	NZ5···O29S	2.76 (4)
N9···O11S^{ii}	3.01 (2)	N9···O25S	2.89 (3)
NZ13···O11S	2.84 (3)	NZ13···O3S	2.79 (2)
N17···O7S	2.77 (1)	NZ13···O20S	2.85 (3)
N17···O21	2.80 (1)	N17···O11X	2.76 (1)
NZ21···O26S	3.01 (4)	NZ21···O21X^{iii}	3.20 (2)
NZ21···O22S	2.84 (3)	NZ21···O7S^{iii}	3.24 (3)

Symmetry codes: (i) $0.5 - x$, $1 - y$, $0.5 + z$; (ii) $-0.5 + x$, $1.5 - y$, $-z$; (iii) $0.5 - x$, $1 - y$, $-0.5 + z$.

Table 8. *Selected hydrogen bonds to water molecules*

	$D\cdots A$ (Å)		$D\cdots A$ (Å)
N2···O2S	2.91 (1)	O3···O17S^{iv}	2.81 (2)
NE14···O12S	2.95 (1)	N5···O6S	2.82 (1)
OG16···O32S^{iv}	2.74 (2)	OG18···O9S	2.80 (4)
N11···O14S	3.00 (2)	NE112···O19S	2.91 (2)
N13···O4S	2.98 (1)	O13···O8S^{v}	2.73 (2)
O14···O5S	3.14 (2)	OG114···O14S	2.60 (3)
O16···O21S^{vi}	2.71 (4)	OG116···O5S	2.76 (2)
O17···O2S	3.14 (2)	N19···O15S	2.91 (1)
NE120···O39S	3.03 (4)	N21···O1S	2.91 (1)
O22···O13S	2.85 (3)	OG122···O15S	2.82 (2)
O23···O3S^{vii}	2.85 (1)	O24···O6S^{viii}	2.87 (2)
O24···O24S	2.82 (3)	OG124···O34S	2.69 (2)

Symmetry codes: (iv) $-1 + x$, y, z; (v) $-0.5 + x$, $0.5 - y$, $-z$; (vi) x, $1 + y$, z; (vii) $0.5 - x$, $1 - y$, $0.5 + z$; (viii) $1 + x$, y, z.

Comparison with NMR studies of other somatostatin analogues

To the best of our knowledge, no NMR or modelling studies have been reported on octreotide itself, and no other crystal structures have been reported for somatostatin analogues. Although it made the structure solution much more difficult, we are perhaps fortunate that there are three crystallographically independent molecules of octreotide in the same crystal, because we can deduce that the three conformations must have similar energies. All three octreotide molecules exhibit the type II′ β-turn which has been predicted in all recent studies of biologically active peptides that contain the -Phe^7-(D)-Trp^8-Lys^9-Thr^{10}- sequence, confirming that this feature is indeed dominant. The crystal structure of octreotide reveals two alternative backbone conformations with different patterns of intramolecular hydrogen bonds. The differences are so substantial that we expect that interconversion would be slow on an NMR time scale.

For c[Pro^6-Phe^7-(D)-Trp^8-Lys^9-Thr^{10}-Phe^{11}], Kessler *et al.* (1983) proposed a ring conformation with a type II′ β-turn that at least qualitatively resembles our molecule I in the -Phe-(D)-Trp-Lys-Thr- region. For the corresponding -Gly^{11}-peptide they proposed an alternative conformation with two intramolecular hydrogen bonds which did not correspond to a type II′ β-turn. By exchanging partners in the two hydrogen bonds one could generate a conformation with a type II′ β-turn with the same hydrogen bonds and conformation as in molecules II and III reported here. We suspect that it would also be possible to account for the NMR data on the basis of such a structure. The two basic ring conformations found in the crystal would then correspond to the two conformations found by NMR in solution.

Huang *et al.* (1992) made detailed predictions for 'flat' and 'folded' conformations of the same cyclic hexapeptide c[Pro^6-Phe^7-(D)-Trp^8-Lys^9-Thr^{10}-Phe^{11}] which are compared with our results in Table 4. It will be seen that the φ and ψ torsion angles of the -(D)-Trp^8- and -Lys^9- residues for both their structures and our molecules I, II and III are all typical for type II′ β-turns. The differences in the tryptophan and lysine side-chain torsion angles observed in the crystal confirm the conformational flexibility which they report in solution. It is particularly interesting that molecule II (but not I or III) in the crystal possesses exactly the *trans*(D-Trp)/*gauche*$^-$(Lys) rotamer combination which they assign to the biologically active conformation; the torsion angles agree well with their predictions. It is not possible, however, to assign the conformations found in the crystal to either their 'flat' or their 'folded' conformation reliably because these conformations differ primarily in the region that has different amino acids to those present in octreotide.

Intermolecular hydrogen bonds and solvent structure

The crystal structure contains a network of intermolecular hydrogen bonds between the peptide molecules and to the solvent molecules. The intermolecular hydrogen bonds are shown in Fig. 6, where the phenylalanine and tryptophan ring systems were omitted for clarity. The Lys5-Thr6-Cys7-Thr8 strand

of molecule I forms an antiparallel β-pleated sheet with the (D)-Trp12-Phe11-Cys10-(D)-Phe9 strand of molecule II with hydrogen bonds O5···HN12, O11···HN6 and O7···HN10. Molecule I forms three hydrogen bonds to molecule III, O2···HN17, O18···HN4 and O4···HN20. The parameters of these hydrogen bonds are given in Table 6.

One oxalate dianion could be seen clearly in the difference electron-density map. This oxalate makes strong hydrogen bonds to the terminal amino groups of molecules I and III and a weak hydrogen bond to the Lys amino group of molecule II. Two out of the six amino groups show strong hydrogen bonding to the oxalate dianion, but all the amino N atoms are surrounded by at least three O atoms at typical hydrogen-bonding distances, so it is not clear which of the amino groups, if any, are not protonated. Even if all are protonated, two of the waters must be H_3O^+ ions if the chemical analysis is correct. In the refinement we assumed that all six amino groups are protonated to NH_3^+. The missing oxalates are presumably so highly disordered that they could not be distinguished from water molecules. Details of the intermolecular hydrogen bonds are given in Table 7.

A total of 52 water sites were found and refined. The positions were selected from the difference electron-density peaks when they showed distances between 2.6 and 3.2 Å to H—N, C═O or other water molecules, had no other short contacts, and gave rise to chemically reasonable angles involving hydrogen bonds. Thus, somewhat stricter criteria for the acceptance of hydrogen bonds were used than those proposed by Baker &

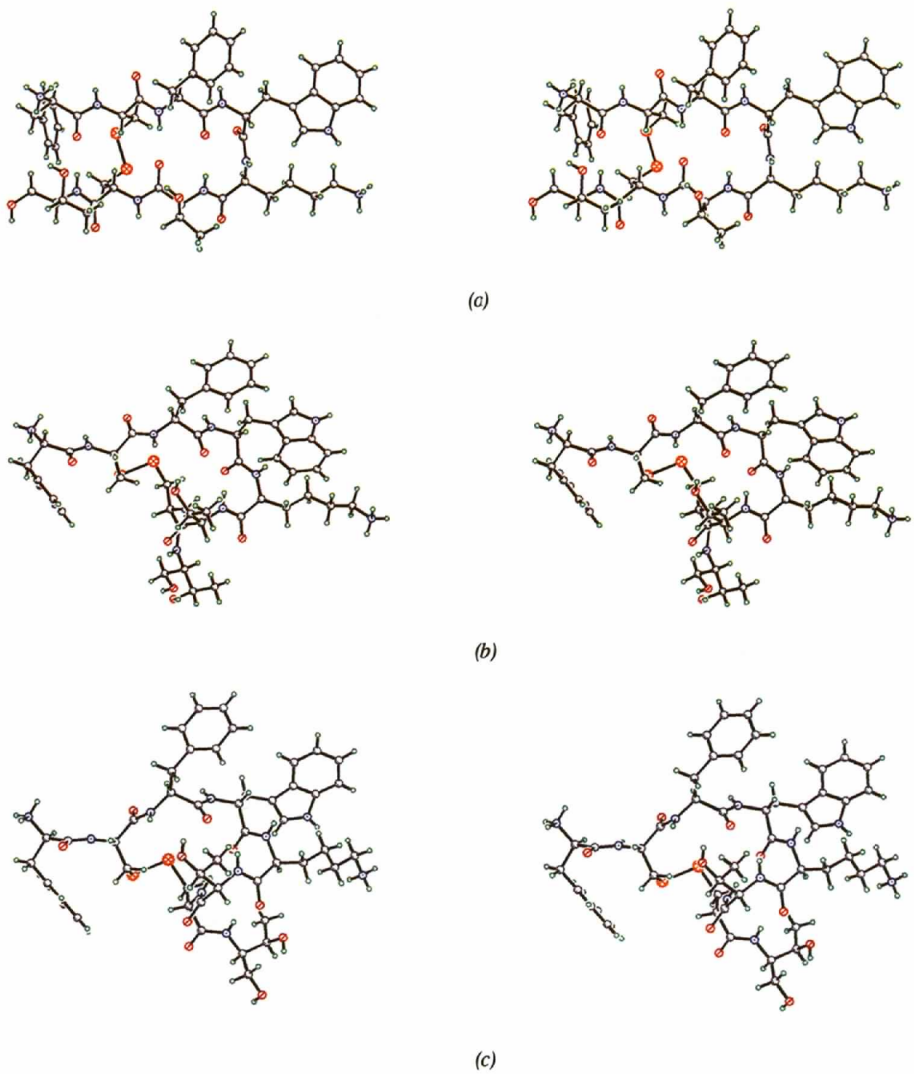


Fig. 3. Stereoview of the crystal structure of octreotide (*a*) molecule I, (*b*) molecule II and (*c*) molecule III.

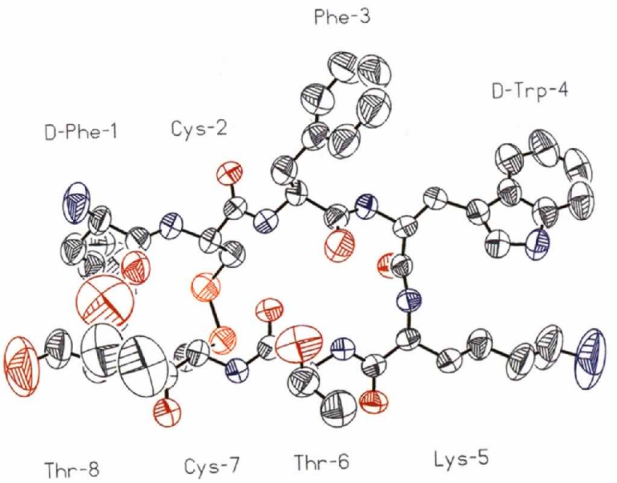
Hubbard (1984) for small proteins. Water positions were rejected when the U_{eq} value was higher than 1 Å^2 or when the position moved too close to other parts of the structure during refinement. 40 water molecules belong to the first hydration shell of the peptides, the remaining 12 form a second hydration shell with no hydrogen bonds to the peptides. All C═O and N—H groups not involved in intra- or intermolecular hydrogen bonds to the other peptide molecules form hydrogen bonds to surrounding water molecules. The C═O···O distances range from 2.7 to 3.2 Å, the N—H···O distances from 2.6 to 3.2 Å. The C═O···O angles are in the range 100–146°, the N—H···O angles are 100–160°. These values correspond well with those given by Jeffrey & Saenger (1991). The second water shell consists of the remaining 12 water molecules. Here the water is less well ordered and the U_{ij} values are significantly higher.

The solvent structure is shown in Fig. 7. The water molecules of the first hydration shell are represented as hatched circles and the water molecules of the second as open circles. The solvent network consist of three-, four-, five- and six-membered rings. Similar water structures were found in the high-resolution X-ray structures of crambin (Teeter, 1984) and rubredoxin (Dauter, Sieker & Wilson 1992).

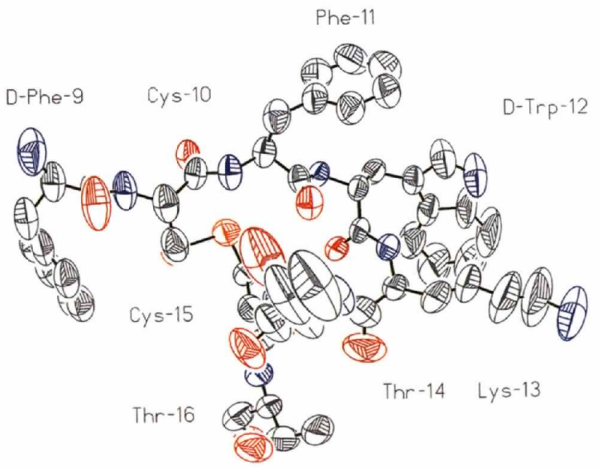
Concluding remarks

In its diffraction properties, octreotide behaves very much like a 'mini-protein' with 24 residues and about 20% water in the asymmetric unit; however, unlike the vast majority of proteins, it diffracts almost to atomic resolution. Although direct methods were unsuccessful, the data will doubtless provide a challenge for future developments in direct methods. The successful solution by automated Patterson interpretation to find the six S atoms, followed by partial structure expansion, clearly demonstrates the potential of such combined Patterson and statistical approaches. The use of 'non-crystallographic symmetry', in the form of restraints on chemically equivalent 1–2 and 1–3 distances, allowed the presence of three chemically identical but conformationally different molecules to be exploited effectively in the refinement.

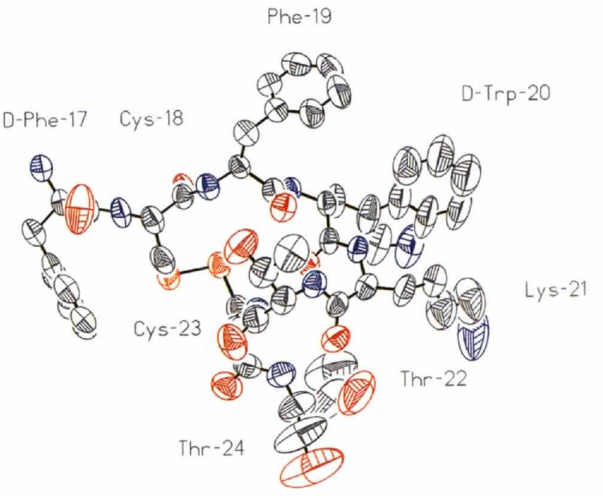
Two different backbone conformations were found for octreotide, both stabilized by intramolecular and



(a)



(b)



(c)

Fig. 4. 50% displacement ellipsoids of (*a*) molecule I, (*b*) molecule II and (*c*) molecule III.

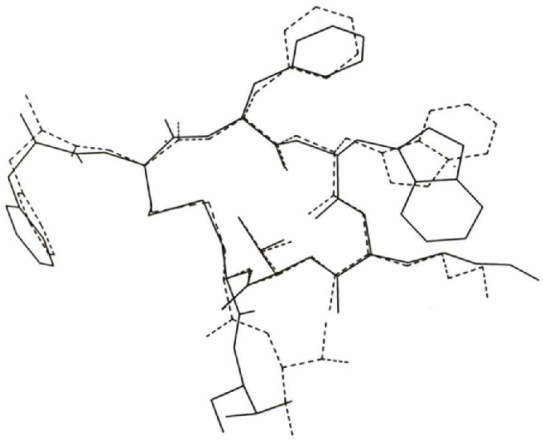


Fig. 5. Superposition of the 20 ring atoms of molecules II and III.

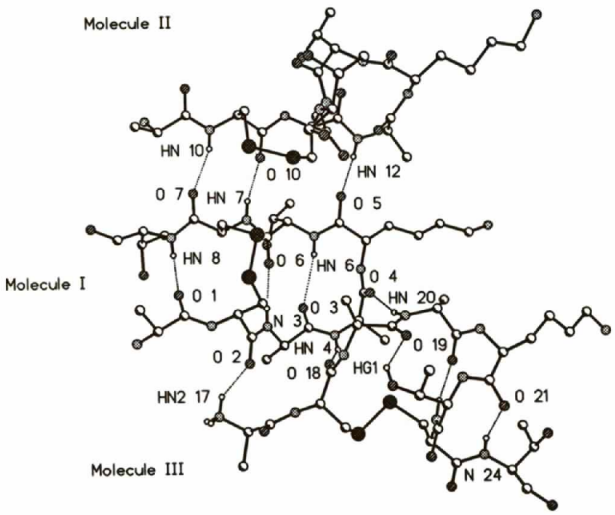


Fig. 6. Intramolecular hydrogen bonds of the three independent peptide molecules.

intermolecular hydrogen bonds. The oxalate dianion is not responsible for the different conformations since it binds strongly to the terminal amino groups of conformationally distinct molecules. One possible explanation for the presence of different conformations is that molecule I forms six intermolecular hydrogen bonds to the other peptide molecules, thus stabilizing an energetically favourable regular anti-parallel β-sheet structure. Molecules II and III form only three intermolecular hydrogen bonds to the other peptides, but they form more hydrogen bonds to the solvent; this conformation may be favoured in aqueous solution.

All three molecules of octreotide (Fig. 8) exhibit the expected type II′ β-turn involving the -(D)-Trp-Lys-unit, and one of the three molecules also adopts the *trans/gauche*$^-$ rotamer combination for these two amino acids predicted by Huang *et al.* for the biologically active conformation. Although conclusions about active-site geometries cannot be drawn from crystal structures alone, this structure illustrates the conformational flexibility of somatostatin analogues as well as establishing that the active conformation proposed by Huang *et al.* corresponds to one of the energy minima.

We gratefully acknowledge financial support by the Leibniz Program of the Deutsche Forschungsgemeinschaft.

References

ARNDT, U. W. & WONACOTT, A. J. (1977). Editors. *The Rotation Method in Crystallography*. Amsterdam: North Holland.

BAKER, E. N. & HUBBARD, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

BATTERSHILL, P. E. & CLISSOLD, S. P. (1989). *Drugs*, **38**, 658–702.

BAUER, W., BRINER, U., DOEPFNER, W., HALLER, R., HUGUENIN, R., MARBACH, P., PETCHER, T. J. & PLESS, J. (1982). *Life Sci.* **31**, 1133–1140.

BRAZEAU, P., VALE, W., BURGUS, R., LING, N., BUTCHER, M., RIVIER, J. & GUILLEMIN, R. (1973). *Science*, **179**, 77–79.

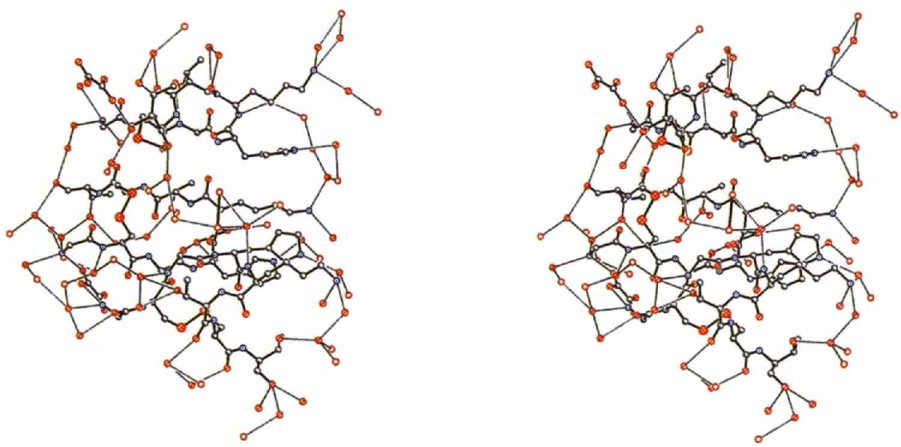


Fig. 7. Stereoview of the solvent structure in the crystal.

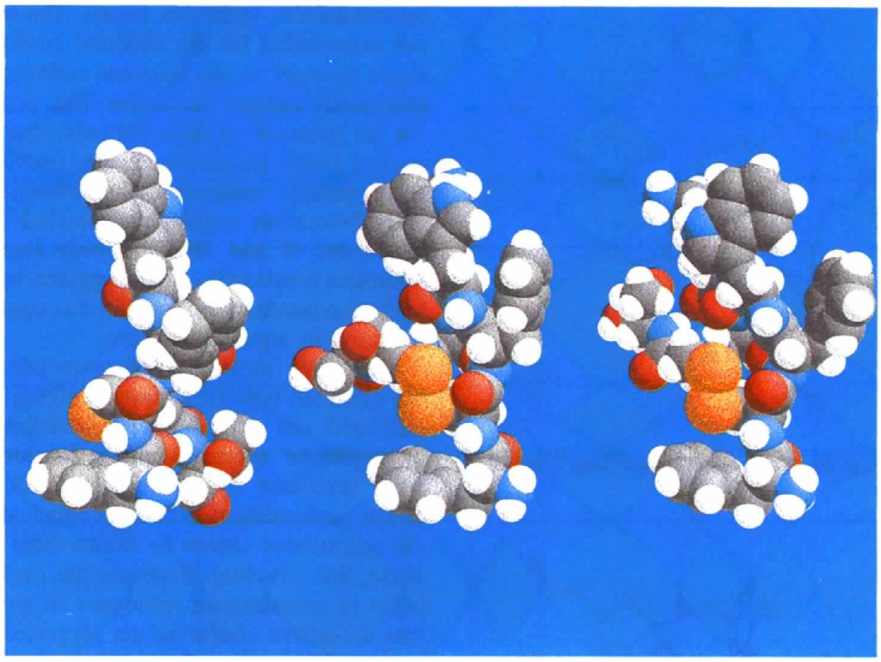


Fig. 8. Space-filling plots of the three independent peptide molecules.

BRÜNGER, A. (1992). *Nature (London)*, **355**, 472–475.

BRÜNGER, A. (1993). *Acta Cryst.* D**49**, 24–36.

BURGUS, R., LING, N., BUTCHER, M. & GUILLEMIN, R. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 684–688.

DAUTER, Z., SIEKER, L. C. & WILSON, K. S. (1992). *Acta Cryst.* B**48**, 42–59.

DRIESSEN, H., HANEEF, M. I. J., HARRIS, G. W., HOWLIN, B., KHAN, G. & MOSS, D. C. (1989). *J. Appl. Cryst.* **22**, 510–516.

FLACK, H.D. (1983). *Acta Cryst.* A**39**, 876–881.

FUJINAGA, M. & READ, R. J. (1987). *J. Appl. Cryst.* **20**, 517–521.

GERICK, J. E. & LORENZI, M. (1978). *Frontiers in Neuroendocrinology* Vol. 5, edited by W. F. GANONG & L. MARTINI, p. 265. New York: Raven Press.

HAUPTMANN, H. A., WEEKS, C. M., SMITH, G. D., TEETER, M. M. & MILLER, R. (1993). Am. Crystallogr. Assoc. Meet., Alberquerque. Poster.

HE, Y. B., HUANG, Z., RAYNOR, K., REISINE, K. & GOODMAN, M. (1993). *J. Am. Chem. Soc.* **115**, 8066–8072.

HIRSHFELD, F. L. (1976). *Acta Cryst.* A**32**, 239–244.

HUANG, Z., HE, Y. B., RAYNOR, K., TALLENT, M., REISINE, T. & GOODMAN, M. (1992). *J. Am. Chem. Soc.* **114**, 9390–9401.

JEFFREY, G. A. & SAENGER, W. (1991). *Hydrogen Bonding in Biological Structures*. Berlin: Springer-Verlag.

KESSLER, H., BERND, M., KOGLER, H., ZARBOCK, J., SØRENSEN, O. W., BODENHAUSEN, G. & ERNST, R. R. (1983). *J. Am. Chem. Soc.* **105**, 6944–6952.

KOOP, H. (1991). *Z. Gesamte. Inn. Med. Ihre Grenzgeb.* **46**, 302–305.

LAMBERTS, S. W. J., UITTERLINDEN, P., VERSCHOOR, L., VAN DONGEN, K. J. & DEL POZO, E. (1985). *N. Eng. J. Med.* **313**, 1576–1580.

LANGRIDGE, R., MARVIN, D. A., SEEDS, W. E., WILSON, H. R., HOOPER, C. P., WILKINS, M. H. F. & HAMILTON, L. D. (1960). *J. Mol. Biol.* **2**, 38–64.

MANNI, A. (1992). *Biotherapy*, **4**, 31-36.

NUTT, R. F., VEBER, D. F. & SAPERSTEIN, R. (1980). *J. Am. Chem. Soc.* **102**, 6539–6545.

OTWINOWSKI, Z. (1991). *DENZO. A Film Processing Program for Macromolecular Crystallography*. Yale Univ., New Haven, CT, USA.

PLESS, J. (1990). *Z. Gastroenterol.* (Suppl.) **28**, 4–7.

REICHLIN, S. (1982). *Molecular Genetic Neuroscience*, edited by F. O. SCHMITT, S. J. BIRD & F. E. BLOOM, pp. 359-372. New York: Raven Press.

REICHLIN, S. (1983). *N. Eng. J. Med.* **309**, 1495–1500; 1556–1563.

REICHLIN, S. (1987). Editor. *Proceedings of the 4th International Conference on Somatostatin*. New York: Plenum Press.

RIVIER, J. E., BROWN, M. & VALE, W. (1975). *Biochem. Biophys. Res. Commun.* **65**, 746-751.

ROLLET, J. S. (1970). *Crystallographic Computing*, edited by F. R. AHMED, S. R. HALL & C. P. HUBER, pp. 167–181. Copenhagen: Munksgaard.

SHELDRICK, G. M. (1982). *Crystallographic Computing*, edited by D. SAYRE, pp. 506–514. Oxford: Clarendon Press/Munksgaard.

SHELDRICK, G. M. (1990). *Acta Cryst.* A**46**, 467–473.

SHELDRICK, G. M. (1992). *Crytallographic Computing* 5, edited by D. MORAS, A. D. PODJARNY & J. C. THIERRY, pp. 145–157. Oxford: IUCr/Oxford Univ. Press.

SHELDRICK, G. M. (1993). *SHELXL93. Program for Crystal Stucture Refinement*, Univ. of Göttingen, Germany.

SHELDRICK, G. M., DAUTER, Z., WILSON, K. S., HOPE, H. & SIEKER, L. C. (1993). *Acta Cryst.* D**49**, 18–22.

TEETER, M. M. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 6014–6018.

TRUEBLOOD, K. N. & DUNITZ, J. P. (1983). *Acta Cryst.* B**39**, 120–133.

VEBER, D. F. (1981). *Peptides: Synthesis, Structure, Function*, edited by D. F. RICH & E. GROSS, pp. 685. Rockford, IL: Pierce Chemical Company.

VEBER, D. F. (1992). *Peptides, Chemistry and Biology, Proceedings of the Twelfth American Peptide Symposium*, edited by J. A. SMITH & J. E. RIVIER, pp. 3–14. Leiden: ESCOM.

VEBER, D. F., FREIDINGER, R. M., SCHWENK-PERLOW, D., PALEVEDA W. J. JR, HOLLY, F. W., STRACHAN, R. G., NUTT, R. F., ARISON, B. H., HOMNICK, C., RANDALL, W. C., GLITZER, M. S., SAPERSTEIN, R. & HIRSCHMANN, R. (1981). *Nature (London)*, **292**, 55–58.

VEBER, D. F., HOLLY, F. W., NUTT, R. F., BERGSTRAND, S. J., BRADY, S. F., HIRSCHMANN, R., GLITZER, M. S. & SAPERSTEIN, R. (1979). *Nature (London)*, **280**, 512–514.

VEBER, D. F., STRACHAN, R. G., BERGSTRAND, S. J., HOLLY, F. W., HOMNICK, C. F., HIRSCHMANN, R., TORCHIANA, M. L. & SAPERSTEIN, R. (1976). *J. Am. Chem Soc.* **98**, 2367–2369.