Acta Cryst. (1975). B31, 713

The Crystal Structure of *S*-Benzyl-L-cysteinyl-L-prolyl-L-leucylglycinamide and its Selenium Analog

BY ANDREW D. RUDKO AND BARBARA W. LOW

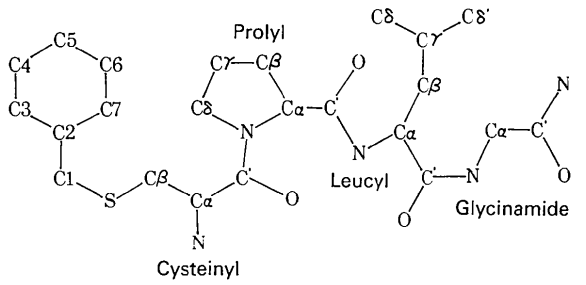
Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, N.Y. 10032, *U.S.A.*

(*Received* 29 *April* 1974; *accepted* 11 *October* 1974)

The crystal structure of the protected C-terminal tetrapeptide fragment of oxytocin, *S*-benzyl-L-cysteinyl-L-prolyl-L-leucylglycinamide, has been determined, as well as that of the analogous peptide in which sulfur is replaced by selenium. The crystals of both compounds are monoclinic, space group $P2_1$. Lattice constants are $a=13{\cdot}65$, $b=8{\cdot}97$, $c=20{\cdot}27$ Å, $\beta=93{\cdot}33°$ for the sulfur peptide and $a=13{\cdot}71$, $b=9{\cdot}07$, $c=20{\cdot}31$Å, $\beta=94{\cdot}00°$ for the nearly isomorphous selenium peptide. The selenium analog structure was solved by the heavy-atom method. The atomic parameters so derived were used as a starting point for the refinement of both structures. The least-squares refinement was not straightforward; the selenium peptide would refine to an R value of only 0·21, while the sulfur peptide was finally refined to an R value of 0·085. There are two molecules in the asymmetric unit with very similar conformations and orientations, related by a translation of approximately $\frac{1}{4}b+\frac{1}{2}c$. The peptides are in a compact '3_{10} bend' conformation with an intramolecular hydrogen bond between the carbonyl oxygen of the cysteinyl residue and the amide nitrogen of the leucyl residue. In one of the molecules in the asymmetric unit there is a second hydrogen bond between the carbonyl oxygen of the prolyl residue and the C-terminal amide nitrogen.

Introduction

S-Benzyl-L-cysteinyl-L-prolyl-L-leucylglycinamide is the protected C-terminal fragment of the peptide hormone oxytocin. The intact hormone has a twenty-atom ring formed by the first six residues and closed by a disulfide bond, cys(1)–cys(6), with a 'tail' part which starts at the Cα of cys(6). The peptide described here includes that tail part as well as the sulfur and β carbon of the cys(6) residue.



The crystal structure analyses of this peptide ($C_{23}H_{35}O_4N_5S$) and of its virtually isomorphous selenium analog ($C_{23}H_{35}O_4N_5Se$) were undertaken for three reasons. We were primarily interested in the relationship between the precise conformation of the peptide in isolation and in the complete hormone. We were interested also in the general aspects of relationships between amino acid sequence and peptide conformation. The development of an empirical method for the prediction of α-helical regions in proteins (Low, Lovell & Rudko, 1968) suggested that protein conformation was largely determined by near-neighbor interactions. The results of this study provide some comment concerning the relevance of near-neighbor interactions in determining the conformations of short-chain peptide fragments.

Few structural studies which provide opportunities for detailed comparison of the conformations, bond lengths, and bond angles of sulfur and selenium analog structures have been made. These analyses provide an opportunity for such comparative evaluation.

A preliminary account of the principal features of the structure has been published previously (Rudko, Lovell & Low, 1971).

Experimental

The sulfur peptide crystals were prepared by dissolving the synthetic material in absolute ethanol and adding diethyl ether to slight turbidity. After standing for six days at 4 °C large tablets appeared lying on (100). The crystal used for data collection was an approximately rectangular parallelepiped $0{\cdot}5\times0{\cdot}2\times0{\cdot}2$ mm.

Both sulfur and selenium peptides could be crystallized by dissolving the compound in water and permitting the solution to evaporate slowly over a period of four days at 4 °C. This latter method gave the best crystals of the selenium analog. The crystal used for data collection was grown by this procedure; it was approximately $0{\cdot}4\times0{\cdot}15\times0{\cdot}13$ mm.

The space-group determination presented some problems. Precession photographs of both peptides were wholly compatible with monoclinic symmetry and a primitive lattice. The space group appeared initially to be $P2$ with some $0k0$ odd reflections weak but observable. The Patterson functions were however clearly compatible in the monoclinic system with the space

group $P2_1$ alone. Relative intensities of the $0,2n+1,0$ reflections varied considerably from crystal to crystal in both sulfur and selenium peptides. The strongest of these reflections, an 010 reflection in the sulfur peptide, corresponds to an absolute F value of 5·6 e. These reflections were not Renninger reflections, as could be clearly demonstrated by measuring the intensities of the 010 and 030 reflections while changing the azimuthal angle, *i.e.* rotating about the normal to the reflecting planes.

An alternative possibility would be a space group $P1$ (pseudo $P2_1$). A study of 332 sulfur peptide reflections on the diffractometer was made and showed no significant differences between reflections hkl and $\bar{h}k\bar{l}$ related by the twofold axis. Within the limits of experimental accuracy the structure is therefore not triclinic ($P1$) but monoclinic ($P2_1$). Accordingly, we have based our analysis on space group $P2_1$; further comments on the possible origin of the $0k0$ reflections with k odd are provided elsewhere.

The lattice parameters (Table 1) were calculated by least-squares refinement from precise diffractometer measurements of high-angle reflections. Crystal densities were determined by the modified gradient tube method. The molecular weight of the sulfur peptide ($C_{23}H_{35}O_4N_5S$) is 477·6 and that of the selenium is 524·5. The experimentally determined weights of the unit-cell contents are 1969 ± 15 and 2094 ± 15 Daltons respectively. Thus there are four molecules in each unit cell, two molecules in the asymmetric unit and no solvent.

Table 1. *Crystal data for sulfur and selenium peptides*

	S-peptide	Se-peptide
a(Å)	13·65 ± 0·01*	13·71 ± 0·03*
b(Å)	8·97 ± 0·01	9·07 ± 0·02
c(Å)	20·27 ± 0·01	20·31 ± 0·03
β	93·33 ± 0·08°	94·00 ± 0·08°
Unit-cell volume (Å^3)	2477·7	2519·4
Density (g cm^{-3})†	1·32 ± 0·01	1·38 ± 0·01
Density (calc.) (g cm^{-3})	1·28	1·38
Weight of unit-cell contents (Daltons)	1969 ± 15	2094 ± 15
Space group	$P2_1$	$P2_1$

* The errors cited for the unit-cell dimensions are root mean square deviations from the mean.

† Gradient columns of water-saturated brombenzene/xylene (Low & Richards, 1952).

For the sulfur peptide, peak-height X-ray diffraction intensities were measured on a Picker manual diffractometer using Mo $K\alpha$ ($\lambda = 0{\cdot}7107$ Å) radiation. The 4995 unique reflections with 2θ less than 55° were measured; 3470 were considered significantly above background: $I > 3\sigma(I)$. For the selenium peptide, intensities were measured by θ–2θ scans on a General Electric automatic diffractometer with Cu $K\alpha$ ($\lambda = 1{\cdot}5418$ Å) radiation. All reflections with 2θ less than 150° were measured: a total of 5415 reflections of which 3974 were significant.

The limit of data collection for both peptides corresponds to spacings of 0·8 Å. Absorption corrections were not applied. Absorption should be negligible for the sulfur peptide for Mo $K\alpha$ radiation ($\mu = 1{\cdot}5$ cm^{-1}) but may have some effect on the selenium analog data ($\mu = 24{\cdot}4$ cm^{-1} for Cu $K\alpha$ radiation). The data were placed on an approximately absolute scale by the use of Wilson (1950) plots.

Structure determination

The selenium peptide structure was essentially solved by the heavy-atom technique. The positions of the selenium atoms could readily be determined by examination of the three-dimensional Patterson function calculated with approximately absolute $|F_{hkl}|^2$. Although there are two selenium atoms in the asymmetric unit they are not in unrelated general positions and provide only centrosymmetric phase information. The electron density distribution calculated with the selenium phases contains the structure and its superimposed mirror image. An unweighted Fourier map was impossible to interpret. When the terms in the Fourier summation were weighted by Sim's (1961) methods the interpretation was straightforward. False (noise) peaks disappeared and it became easy to distinguish between the two centrosymmetrically related images. All atoms (except hydrogens) were located in this map.

The cell dimensions and general intensity distribution of sulfur peptide and selenium analog are very similar, indicating that the two compounds crystallize in broadly isomorphous forms. Accordingly, the rough model derived from the selenium-phased electron density map was used as a starting point for the refinement of both structures.

The R values after the first complete structure-factor calculation were 0·36 for the selenium peptide and 0·43 for the sulfur peptide. Each structure was first refined by a three-dimensional ΔF synthesis. The corrected parameters resulted in R values of 0·30 for the selenium structure and 0·39 for the sulfur structure.

Further refinement of both structures was by least-squares calculations. Because of the large number of parameters, the structures were divided into sections and refined by a block-matrix approximation, *i.e.*, only terms corresponding to interactions between parameters of atoms within the same group are considered. At first the structures were divided into three groups: (*a*) the sulfur or selenium atoms, (*b*) the oxygen and nitrogen atoms, and (*c*) the carbon atoms. In later cycles sections (*a*) and (*b*) were combined. Anisotropic temperature parameters were used for both sulfur and selenium atoms, isotropic temperature parameters were used for all other atoms. Unit weights were employed initially, later refinement used weights $w = 1/(\sigma_{hkl} + aF_{hkl})$, where a is a parameter chosen so that the average weighted ΔF^2 is approximately constant for all ranges of F. All scattering factors were taken from *International Tables for X-ray Crystallography* (1962).

Table 2. *Atomic parameters*

(*a*) Fractional coordinates and isotropic thermal parameters

		x	*y*	*z*	*B*(Å²)
Se peptide					
Molecule *A*					
	Se	0·2310 (2)	0·0000 (0)	0·3900 (1)	
Benzyl	C(1)	0·1611 (16)	−0·0030 (32)	0·2990 (11)	4·48 (0·46)
	C(2)	0·1489 (18)	−0·1510 (30)	0·2733 (12)	4·14 (0·49)
	C(3)	0·0820 (22)	−0·2391 (41)	0·3005 (14)	6·33 (0·69)
	C(4)	0·0776 (28)	−0·3901 (49)	0·2750 (18)	8·53 (0·92)
	C(5)	0·1444 (31)	−0·4307 (52)	0·2308 (20)	9·06 (1·03)
	C(6)	0·2029 (24)	−0·3525 (42)	0·2062 (16)	6·66 (0·77)
	C(7)	0·2164 (24)	−0·2034 (41)	0·2264 (16)	6·73 (0·77)
Cysteinyl	N	0·5149 (17)	0·0904 (31)	0·3402 (11)	6·30 (0·56)
	Cβ	0·3682 (15)	−0·0203 (27)	0·3655 (10)	3·47 (0·40)
	Cα	0·4088 (18)	0·1235 (31)	0·3525 (12)	4·74 (0·51)
	C′	0·4092 (16)	0·2107 (26)	0·4177 (10)	3·64 (0·43)
	O	0·4390 (11)	0·1581 (19)	0·4724 (17)	4·50 (0·33)
Prolyl	N	0·3717 (15)	0·3445 (24)	0·4139 (9)	4·60 (0·42)
	Cα	0·3760 (18)	0·4412 (31)	0·4721 (12)	4·90 (0·53)
	Cβ	0·3457 (22)	0·6002 (40)	0·4443 (15)	6·55 (0·70)
	Cγ	0·2975 (26)	0·5716 (46)	0·3823 (18)	8·36 (0·90)
	Cδ	0·3253 (18)	0·4182 (31)	0·3536 (12)	4·65 (0·52)
	C′	0·3112 (17)	0·3906 (30)	0·5240 (11)	4·47 (0·49)
	O	0·3271 (15)	0·4320 (27)	0·5816 (10)	7·57 (0·54)
Leucyl	N	0·2372 (12)	0·3070 (20)	0·5044 (8)	3·28 (0·32)
	Cα	0·1734 (15)	0·2311 (29)	0·5543 (9)	3·63 (0·40)
	Cβ	0·0849 (18)	0·1818 (29)	0·5242 (12)	4·49 (0·50)
	Cγ	0·0082 (16)	0·2895 (25)	0·4894 (10)	3·63 (0·42)
	Cδ	0·0262 (23)	0·3224 (40)	0·4176 (15)	6·67 (0·73)
	Cδ′	−0·0965 (25)	0·2379 (48)	0·4850 (16)	8·28 (0·81)
	C′	0·2221 (16)	0·1300 (29)	0·5999 (11)	3·97 (0·45)
	O	0·1882 (14)	0·0840 (25)	0·6469 (9)	6·52 (0·46)
Glycyl	N	0·3143 (15)	0·0706 (26)	0·5818 (9)	4·72 (0·43)
	Cα	0·3621 (21)	−0·0311 (37)	0·6256 (14)	6·20 (0·65)
	C′	0·4254 (17)	0·0371 (29)	0·6850 (11)	4·18 (0·48)
	O	0·4562 (13)	−0·0515 (23)	0·7276 (9)	6·20 (0·43)
	N	0·4433 (18)	0·1700 (32)	0·6866 (12)	5·89 (0·51)
Molecule *B*					
	Se	0·2324 (2)	0·2419 (4)	0·8707 (1)	
Benzyl	C(1)	0·1575 (18)	0·2464 (35)	0·7871 (11)	4·98 (0·51)
	C(2)	0·1519 (23)	0·4335 (39)	0·7644 (15)	6·22 (0·71)
	C(3)	0·0794 (24)	0·5284 (44)	0·7971 (16)	6·88 (0·76)
	C(4)	0·0873 (40)	0·6509 (74)	0·7753 (26)	13·19 (1·65)
	C(5)	0·1559 (28)	0·7144 (51)	0·7226 (18)	8·52 (0·96)
	C(6)	0·2068 (27)	0·6160 (49)	0·6832 (18)	7·74 (0·92)
	C(7)	0·1920 (27)	0·4702 (48)	0·7170 (18)	7·91 (0·90)
Cysteinyl	N	0·5204 (14)	0·3187 (24)	0·8220 (9)	4·37 (0·41)
	Cβ	0·3648 (15)	0·2259 (28)	0·8439 (10)	3·59 (0·40)
	Cα	0·4145 (18)	0·3591 (31)	0·8336 (11)	4·44 (0·49)
	C′	0·4187 (16)	0·4485 (27)	0·9025 (11)	3·96 (0·45)
	O	0·4384 (12)	0·3803 (21)	0·9575 (8)	4·76 (0·36)
Prolyl	N	0·3900 (13)	0·5842 (23)	0·9044 (9)	3·90 (0·37)
	Cα	0·3932 (16)	0·6703 (27)	0·9651 (11)	3·84 (0·44)
	Cβ	0·3660 (20)	0·8313 (25)	0·9410 (13)	5·39 (0·59)
	Cγ	0·3124 (18)	0·8117 (31)	0·8762 (12)	4·82 (0·52)
	Cδ	0·3536 (15)	0·6745 (25)	0·8446 (10)	3·29 (0·39)
	C′	0·3213 (16)	0·6051 (28)	1·0100 (11)	4·06 (0·45)
	O	0·3404 (11)	0·6334 (20)	1·0712 (7)	4·62 (0·34)
Leucyl	N	0·2465 (13)	0·5505 (23)	0·9856 (9)	4·06 (0·38)
	Cα	0·1649 (16)	0·4794 (29)	1·0279 (10)	3·95 (0·44)
	Cβ	0·0779 (17)	0·4557 (30)	0·9815 (11)	4·32 (0·50)
	Cγ	−0·0172 (20)	0·4002 (34)	1·0149 (13)	5·54 (0·60)
	Cδ	−0·0428 (26)	0·5056 (50)	1·0674 (17)	8·25 (0·88)
	Cδ′	−0·0976 (27)	0·3954 (47)	0·9610 (18)	8·41 (0·93)
	C′	0·2019 (17)	0·3535 (29)	1·0651 (11)	4·03 (0·45)
	O	0·1579 (12)	0·3010 (21)	1·1089 (8)	5·47 (0·39)

Table 2 (*cont.*)

Glycyl	N	0·2910 (13)	0·2832 (22)	1·0490 (8)	4·02 (0·39)
	Cα	0·3236 (16)	0·1610 (28)	1·0891 (11)	4·15 (0·46)
	C′	0·3820 (19)	0·2205 (35)	1·1574 (13)	5·33 (0·57)
	O	0·4025 (13)	0·1279 (23)	1·2004 (9)	5·44 (0·38)
	N	0·4036 (16)	0·3602 (30)	1·1644 (11)	5·66 (0·50)
S peptide					
Molecule *A*					
	S	0·2285 (2)	0·0000 (0)	0·3838 (1)	
Benzyl	C(1)	0·1670 (12)	0·0062 (18)	0·3004 (7)	
	C(2)	0·1501 (9)	−0·1486 (15)	0·2749 (5)	
	C(3)	0·0745 (9)	−0·2334 (17)	0·3010 (5)	
	C(4)	0·0592 (10)	−0·3842 (19)	0·2784 (6)	
	C(5)	0·1223 (13)	−0·4408 (21)	0·2334 (7)	
	C(6)	0·1960 (12)	−0·3525 (22)	0·2070 (3)	
	C(7)	0·2097 (10)	−0·2031 (20)	0·2276 (5)	
Cysteinyl	N	0·5129 (7)	0·0643 (13)	0·3381 (5)	4·55 (0·23)
	Cβ	0·3550 (9)	−0·0393 (16)	0·3641 (6)	4·42 (0·28)
	Cα	0·4096 (7)	0·1026 (13)	0·3516 (5)	2·98 (0·20)
	C′	0·4081 (7)	0·1995 (12)	0·4142 (5)	2·73 (0·20)
	O	0·4396 (5)	0·1473 (9)	0·4693 (3)	3·24 (0·15)
Prolyl	N	0·3731 (6)	0·3384 (10)	0·4099 (4)	2·68 (0·16)
	Cα	0·3785 (8)	0·4343 (12)	0·4689 (5)	2·90 (0·20)
	Cβ	0·3517 (8)	0·5905 (15)	0·4405 (5)	3·76 (0·24)
	Cγ	0·2843 (9)	0·5545 (16)	0·3773 (6)	4·58 (0·28)
	Cδ	0·3234 (8)	0·4086 (13)	0·3504 (5)	3·48 (0·23)
	C′	0·3105 (8)	0·3866 (14)	0·5199 (5)	3·70 (0·23)
	O	0·3231 (6)	0·4320 (11)	0·5782 (4)	5·17 (0·20)
Leucyl	N	0·2369 (5)	0·2915 (9)	0·5034 (3)	2·32 (0·14)
	Cα	0·1735 (7)	0·2347 (13)	0·5549 (4)	2·86 (0·19)
	Cβ	0·0771 (8)	0·1622 (13)	0·5239 (5)	3·22 (0·21)
	Cγ	0·0065 (8)	0·2772 (13)	0·4889 (5)	3·45 (0·22)
	Cδ	0·0301 (10)	0·3047 (17)	0·4157 (6)	5·36 (0·22)
	Cδ′	−0·0969 (10)	0·2145 (17)	0·4910 (6)	5·37 (0·32)
	C′	0·2263 (7)	0·1179 (12)	0·5996 (5)	2·87 (0·19)
	O	0·1879 (5)	0·0751 (10)	0·6485 (4)	4·25 (0·17)
Glycyl	N	0·3144 (6)	0·0693 (10)	0·5816 (4)	3·26 (0·18)
	Cα	0·3691 (9)	−0·0426 (14)	0·6227 (5)	3·79 (0·23)
	C′	0·4273 (9)	0·0264 (14)	0·6816 (5)	3·61 (0·23)
	O	0·4570 (6)	−0·0689 (10)	0·7253 (4)	4·57 (0·18)
	N	0·4440 (8)	0·1683 (12)	0·6854 (5)	4·35 (0·22)
Molecule *B*					
	S	0·2356 (2)	0·2494 (4)	0·8701 (1)	
Benzyl	C(1)	0·1578 (11)	0·2641 (18)	0·7944 (6)	
	C(2)	0·1497 (10)	0·4289 (17)	0·7689 (5)	
	C(3)	0·0858 (11)	0·5238 (21)	0·7997 (6)	
	C(4)	0·0825 (12)	0·6710 (24)	0·7751 (9)	
	C(5)	0·1416 (14)	0·7148 (23)	0·7206 (9)	
	C(6)	0·2010 (12)	0·6100 (24)	0·6945 (7)	
	C(7)	0·2061 (11)	0·4685 (20)	0·7157 (6)	
Cysteinyl	N	0·5200 (6)	0·3153 (11)	0·8198 (4)	3·47 (0·19)
	Cβ	0·3581 (8)	0·2136 (14)	0·8417 (5)	3·73 (0·24)
	Cα	0·4158 (7)	0·3560 (12)	0·8349 (4)	2·68 (0·19)
	C′	0·4167 (7)	0·4420 (12)	0·9006 (4)	2·54 (0·18)
	O	0·4410 (5)	0·3763 (9)	0·9542 (3)	3·22 (0·14)
Prolyl	N	0·3910 (5)	0·5851 (9)	0·9008 (3)	2·29 (0·15)
	Cα	0·3943 (8)	0·6701 (12)	0·9628 (5)	2·91 (0·20)
	Cβ	0·3703 (9)	0·8308 (14)	0·9414 (5)	3·90 (0·25)
	Cγ	0·3084 (8)	0·8122 (14)	0·8752 (5)	3·89 (0·24)
	Cδ	0·3534 (8)	0·6745 (13)	0·8424 (5)	3·21 (0·21)
	C′	0·3231 (7)	0·6119 (12)	1·0114 (5)	2·92 (0·20)
	O	0·3385 (5)	0·6286 (9)	1·0716 (3)	3·82 (0·16)
Leucyl	N	0·2412 (6)	0·5427 (9)	0·9856 (3)	2·54 (0·15)
	Cα	0·1646 (8)	0·4827 (12)	1·0285 (5)	3·00 (0·20)
	Cβ	0·0737 (8)	0·4526 (13)	0·9800 (5)	2·98 (0·20)
	Cγ	−0·0208 (9)	0·4012 (14)	1·0165 (5)	3·75 (0·24)
	Cδ	−0·0492 (10)	0·5069 (18)	1·0686 (6)	5·32 (0·31)

Table 2 (*cont.*)

	Cδ′	−0·1010 (10)	0·3891 (18)	0·9609 (7)	5·47 (0·33)
	C′	0·2002 (7)	0·3450 (12)	1·0660 (4)	2·71 (0·19)
	O	0·1552 (6)	0·3000 (10)	1·1113 (3)	4·28 (0·18)
Glycyl	N	0·2852 (6)	0·2769 (10)	1·0484 (4)	3·08 (0·17)
	Cα	0·3281 (8)	0·1577 (13)	1·0907 (5)	3·54 (0·23)
	C′	0·3814 (8)	0·2145 (13)	1·1545 (5)	3·30 (0·22)
	O	0·4006 (6)	0·1216 (10)	1·1984 (4)	4·54 (0·18)
	N	0·4057 (7)	0·3588 (12)	1·1592 (5)	4·11 (0·21)

(*b*) Anisotropic thermal parameters in Å², expressed as $T=\exp[-\frac{1}{4}(B_{11}h^2a^{*2}+B_{22}k^2b^{*2}+B_{33}l^2c^{*2}+2B_{12}hka^*b^*+2B_{13}hla^*c^*+2B_{23}klb^*c^*)]$

		B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
Se peptide							
Molecule *A*							
	Se	4·82 (0·15)	5·37 (0·17)	4·92 (0·19)	−1·01 (0·13)	0·24 (0·13)	−1·89 (0·15)
Molecule *B*							
	Se	3·82 (0·13)	4·74 (0·14)	3·56 (0·15)	−0·96 (0·13)	0·57 (0·11)	−0·44 (0·14)
S peptide							
Molecule *A*							
	S	4·78 (0·19)	5·31 (0·20)	4·90 (0·16)	−1·09 (0·16)	0·46 (0·14)	−1·90 (0·15)
Benzyl	C(1)	7·26 (0·99)	4·69 (0·77)	6·89 (0·86)	1·24 (0·74)	−2·03 (0·73)	0·29 (0·71)
	C(2)	3·69 (0·61)	4·49 (0·63)	3·41 (0·50)	0·20 (0·53)	−0·46 (0·44)	−0·33 (0·48)
	C(3)	3·33 (0·61)	6·05 (0·79)	3·76 (0·53)	0·17 (0·59)	−1·79 (0·47)	−0·11 (0·56)
	C(4)	4·58 (0·73)	6·34 (0·87)	4·45 (0·63)	0·06 (0·69)	−1·07 (0·55)	0·16 (0·64)
	C(5)	6·72 (0·97)	7·16 (1·04)	5·09 (0·74)	0·70 (0·86)	−1·93 (0·69)	0·98 (0·75)
	C(6)	7·27 (1·02)	8·10 (1·10)	3·26 (0·59)	2·21 (0·89)	−1·88 (0·62)	−2·08 (0·69)
	C(7)	5·42 (0·80)	8·44 (1·04)	2·59 (0·51)	1·18 (0·75)	−0·52 (0·51)	−0·36 (0·62)
Molecule *B*							
	S	3·71 (0·15)	4·63 (0·16)	3·36 (0·11)	−1·18 (0·14)	0·71 (0·11)	−0·50 (0·12)
Benzyl	C(1)	6·47 (0·87)	4·96 (0·77)	5·08 (0·65)	−1·74 (0·69)	2·28 (0·62)	0·20 (0·62)
	C(2)	4·88 (0·72)	5·56 (0·77)	3·31 (0·52)	−0·75 (0·63)	−2·31 (0·49)	−0·07 (0·54)
	C(3)	4·74 (0·80)	7·16 (0·95)	4·20 (0·63)	1·33 (0·73)	−0·53 (0·56)	−0·07 (0·67)
	C(4)	5·59 (0·95)	8·69 (1·27)	7·65 (1·03)	1·55 (0·90)	−3·32 (0·82)	−1·83 (0·98)
	C(5)	7·68 (1·11)	7·13 (1·15)	6·30 (0·92)	−0·06 (0·98)	−2·78 (0·84)	1·88 (0·89)
	C(6)	6·26 (0·98)	8·39 (1·18)	4·90 (0·75)	−1·08 (0·92)	−0·82 (0·68)	1·32 (0·81)
	C(7)	5·47 (0·82)	7·38 (0·99)	3·34 (0·57)	−1·99 (0·74)	−0·27 (0·52)	0·48 (0·63)

Anomalous dispersion corrections for selenium had no significant effects and were not used in the final refinement. For most calculations the X-RAY 67 programs (Stewart, 1967) were used on an IBM 360/91 computer.

After 11 cycles the *R* value of the sulfur peptide was reduced from 0·39 to 0·21. No further refinement was possible by this method. Most of the parameters converged at chemically reasonable values except for obvious problems with the benzene-ring atoms. Their positional parameters oscillated and the thermal parameters tended to grow very large.

After many attempts at further refinement including the use of partial data sets and various weighting functions a change was made in the blocking scheme for the sulfur peptide. The molecule was divided into three groups of atoms: (1) *S*-benzyl-cysteinyl-prolyl, (2) leucyl, and (3) glycinamide. The effect of this change was surprisingly radical. In eight cycles the *R* value went to 0·105 and all parameters, including those of the benzene ring atoms, refined to reasonable values. Anisotropic thermal parameters were then applied to the seven carbon atoms of each benzyl group. After three cycles the *R* value was 0·096. 38 of the 68 hydrogen atoms were located in a difference map. They were included in the structure but not refined. After six more cycles of refinement the final *R* value was 0·085. In the last cycle the maximum parameter shift was less than 0·3 standard deviations.

The refinement of the selenium peptide initially followed a similar pattern to that of the sulfur peptide. After 14 cycles the *R* value was reduced from 0·30 to 0·22. No further refinement was possible. As with the sulfur peptide, the parameters converged at chemically reasonable values, apart from the problems described with the benzene-ring atoms. The subsequent procedure of changing the blocking scheme, when employed with the selenium analog structure, had no effect whatsoever. After eight cycles the *R* value remained 0·215. A discussion of diverse efforts to resolve the problems encountered in refinement and our conclusions concerning their nature are given in the Appendix.

Results

Final parameters for both structures are given in Table 2.* The structures of the sulfur and selenium

* A table of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 30729 (83 pp., 2 microfiches). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

peptides are nearly identical. Unless otherwise noted, the structure discussed will be that of the sulfur peptide; it is known to a greater degree of precision. All conclusions apply equally well to both structures.

The Patterson maps of both peptides contain pseudo-origins at approximately $0,\frac{1}{4},\frac{1}{2}$. This suggested, and the structure determination confirmed, that the two molecules (*A* and *B*) in the asymmetric unit are in almost identical orientations with similar conformations and are related by a translational vector $0,b/4,c/2$ (see Fig. 1). The atoms of the backbone chain portion of molecule *A* (including the proline ring, the α-carbon and sulfur of cysteine but excluding the terminal amide group) are related to corresponding atoms of molecule *B* by an average translation of $0{\cdot}003a$, $0{\cdot}238b$, and $0{\cdot}484c$ with standard deviations of $0{\cdot}017a$, $0{\cdot}018b$, and $0{\cdot}011c$ respectively.

Viewed down the *y* axis as in Figs. 2 and 3, the molecule can be seen to have a hydrophobic side and a hydrophilic side. On the hydrophobic side are the benzyl group and leucyl side chain. The free amino group and terminal amide group lie on the hydrophilic side. The peptide molecules are packed with hydrophobic sides near the plane at $x=0$ and their hydrophilic sides near the plane at $x=\frac{1}{2}$. The crystal structure is thus divided into alternating regions of hydrophilic and hydrophobic interactions.

Intermolecular interactions – hydrophilic

The hydrophilic interactions involve the free amino groups, terminal amides, and prolyl carbonyls. They are shown in Fig. 3. Both molecules *A* and *B* form four hydrogen bonds to four different molecules: two equivalent (symmetry-related), and two non-equivalent. The amino group of molecule *A* is hydrogen bonded to the prolyl carbonyl group of an equivalent molecule related by the operation $1-x,y+\frac{1}{2},1-z$. The prolyl carbonyl of molecule *A* is in turn hydrogen bonded to the free amino group of another equivalent molecule related by $1-x,y-\frac{1}{2},1-z$. The amide carbonyl of molecule *B* is bonded to the free amino group of an equivalent molecule related by $1-x,y-\frac{1}{2},2-z$, and the free amino group of molecule *B* is in turn bonded to the amide carbonyl of the molecule related by the operation $1-x,y+\frac{1}{2},2-z$.

The free amino group of molecule *B* also forms a hydrogen bond to the terminal amide nitrogen of a nonequivalent molecule (*A*) related by the approximate translation $0,-0{\cdot}24,-0{\cdot}48$. The other hydrogen bond between nonequivalent molecules joins the amide carbonyl group of molecule *A* with the amide nitrogen of a molecule (*B*) related by the translation $0,0{\cdot}24,0{\cdot}48$ followed by the operation $1-x,y-\frac{1}{2},2-z$.

The most striking feature of the hydrogen-bond network is the end-to-end hydrogen bonding of the terminal amide groups, forming an endless chain extending in the *y* direction. As shown in Fig. 3, the amide nitrogen of a molecule *A* is hydrogen bonded to the free amino group of a molecule *B*, which in turn is bonded to the carbonyl of the terminal amide of another *B* molecule and the nitrogen of this molecule is bonded to the amide carbonyl of another molecule *A*. The nitrogen of that amide group is bonded to the free amino group of a type *B* molecule *et seq.* Hydrogen-bond distances and angles are given in Table 3.

Intermolecular interactions – hydrophobic

Hydrophobic interactions are of two kinds. The leucyl hydrocarbon side chains lie with their β-carbons very close to twofold screw axes; the leucyl side chain of molecule *A* lies on the axis at $x=0,z=\frac{1}{2}$, and the side chain of molecule *B* lies on the axis at $x=0,z=1$. The hydrocarbon groups are thus stacked one above

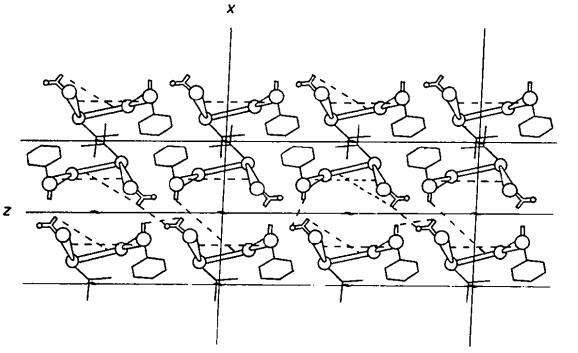


Fig. 1. Schematic representation of the molecular packing as viewed down the *b* axis.

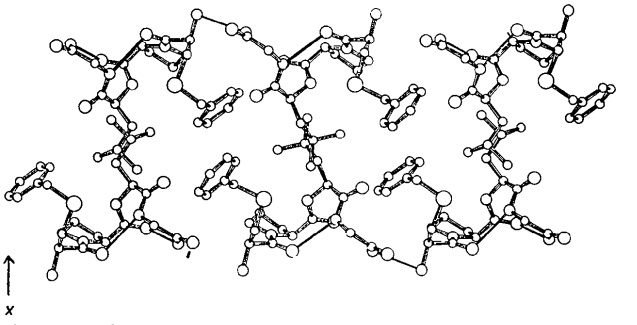


Fig. 2. Molecular packing viewed down the *b* axis showing the hydrophobic interactions.

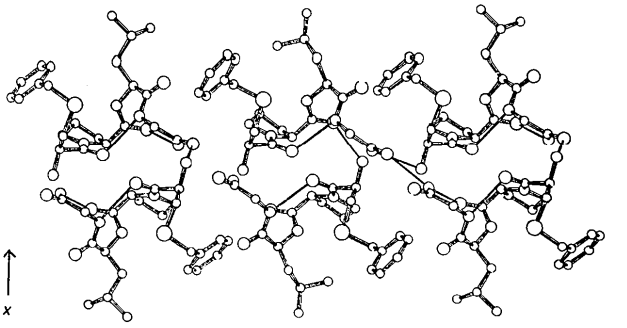


Fig. 3. Molecular packing viewed down the *b* axis showing the hydrophilic interactions.

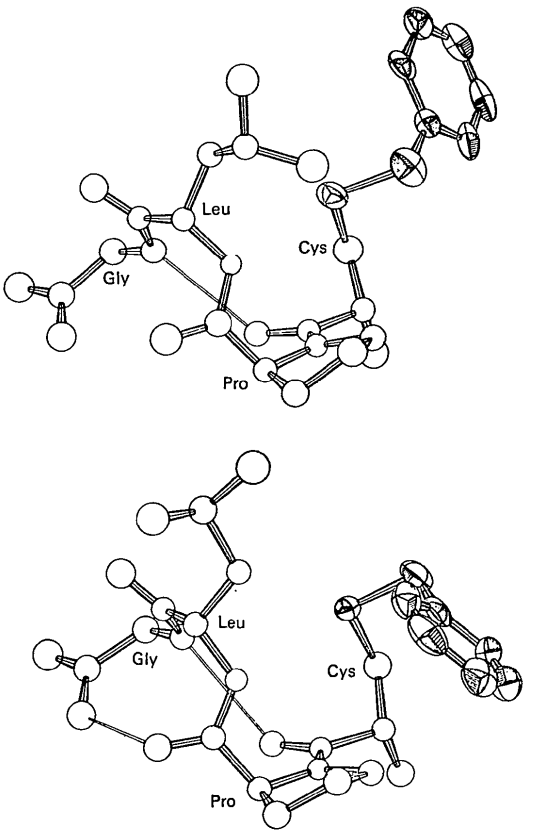


Fig. 4. Thermal ellipsoid plots (50% probability level) of the two molecules of the sulfur peptide in the asymmetric unit. Anisotropic temperature parameters were obtained only for the sulfur atoms and benzyl groups.

the other in the y direction; symmetry-related groups alternating at intervals of $b/2$ (4·48 Å).

The benzene rings are arranged in pairs about the pseudocenters at $0, -0{\cdot}38, 0{\cdot}25$, and $0, 0{\cdot}38, 0{\cdot}75$, as shown in Fig. 2, each benzene being paired with one of a nonequivalent molecule. In each pair the planes of the rings are parallel and 4·42 Å apart. If one ring is projected on the plane of the other there is approximately one third overlap. As Fig. 1 indicates, this arrangement results in two continuous channels of paired benzene rings at $x=0, z=0{\cdot}25$ and $x=0, z=0{\cdot}75$ extending in the y direction, each pair of benzene rings being stacked directly above a symmetry-related pair.

Each molecule is involved in hydrophobic interactions with three others: two equivalent and one nonequivalent. These are different molecules from the four to which each is hydrogen bonded. In total, therefore, each molecule interacts with seven neighboring molecules.

Peptide conformation

The backbone conformations of the two molecules in the asymmetric units are very similar, as shown for the sulfur peptide in Fig. 4 and for the selenium peptide in Fig. 5. Torsion angles about single bonds in the peptide backbone are given in Table 4. Each molecule is folded back on itself forming a compact, tightly twisted conformation with the carbonyl oxygen of the cysteinyl residue hydrogen bonded to the glycyl amino nitrogen forming a ten-membered ring (–OC–Pro–Leu–NH–), see Fig. 5.

In molecule *B* a second intramolecular hydrogen bond joins the prolyl carbonyl oxygen to the nitrogen

Table 3. *Hydrogen-bond distances and angles*

	Distance (Å)*	Angle (C=O···N)†
Sulfur analog		
Intramolecular bonds		
O(Cys)–NH(Gly) Mol. *A*	3·02 (2)	125 (1)°
O(Cys)–NH(Gly) Mol. *B*	3·08 (2)	121 (1)
O(Pro)–NH(Amide) Mol. *B*	3·13 (2)	130 (1)
Intermolecular bonds		
O(Pro, Mol. *A*)–NH(Cys, Mol. *A*)	2·99 (2)	136 (1)
O(Gly, Mol. *A*)–NH(Amide, Mol. *B*)	3·00 (2)	148 (1)
N(Cys, Mol. *B*)–NH(Amide, Mol. *A*)	3·16 (2)	
O(Gly, Mol. *B*)–NH(Cys, Mol. *B*)	3·01 (2)	127 (1)
Selenium analog		
Intramolecular bonds		
O(Cys)–NH(Gly) Mol. *A*	3·00 (4)	126 (2)
O(Cys)–NH(Gly) Mol. *B*	2·97 (4)	125 (2)
O(Pro)–NH(Amide) Mol. *B*	3·20 (4)	117 (2)
Intermolecular bonds		
O(Pro, Mol. *A*)–NH(Cys, Mol. *A*)	2·97 (4)	137 (3)
O(Gly, Mol. *A*)–NH(Amide, Mol. *B*)	2·92 (4)	151 (2)
N(Cys, Mol. *B*)–NH(Amide, Mol. *A*)	3·18 (5)	
O(Gly, Mol. *B*)–NH(Cys, Mol. *B*)	3·04 (4)	126 (3)

* Estimated standard deviation × 100 in parentheses.
† Estimated standard deviation in parentheses.

of the terminal amide forming a second ten-membered ring (OC–Leu–Gly–NH_2). Although the conformation of molecule *A* is quite similar, the prolyl carbonyl oxygen is too far from the terminal amide nitrogen (3·57 Å) to permit hydrogen-bond formation. Both atoms are involved in intermolecular hydrogen bonding. Apart from the presence of a second intramolecular hydrogen bond in molecule *B* the principal *A*/*B* differences involve side-chain conformations.

Side-chain conformations

The most striking difference between molecules *A* and *B* is in the positions of the benzyl groups. The torsion angles about the sulfur–benzyl bonds differ by 180°; thus the benzene ring in molecule *A* is bent towards the peptide, in molecule *B* it extends outwards. The benzene rings are, as expected, closely planar (see Table 5) with standard deviations of 0·017 Å (molecule *A*) and 0·010 Å (molecule *B*).

The conformation about the Cα–Cβ bond of the cysteinyl residue is the same in molecules *A* and *B* with the sulfur atom *trans* to the amino nitrogen. This differs from conformations reported for similar structures. For example, in the cysteinyl residue of glutathione (Wright, 1958), hexagonal L-cystine (Oughton & Harrison, 1959), L-cystine hydrobromide (Peterson, Steinrauf & Jensen, 1960) and in *N*,*N*′-diglycylcystine (Yakel & Hughes, 1954) the sulfur is always *trans* to the α-hydrogen.

Table 4. *Dihedral angles*

Following the conventions suggested by the IUPAC–IUB Commission on Biochemical Nomenclature (1970).

Peptide backbone angles

		Sulfur analog		Selenium analog		Z-Gly-Pro-Leu-Gly[a]	Z-Gly-Pro-Leu-Gly-Pro[b]
		Molecule *A*	Molecule *B*	Molecule *A*	Molecule *B*		
Cysteinyl	ψ	62°	67°	63°	65°		
	ω	4	3	0	6		
Prolyl	φ	−70	−64	−65	−71	−58	−65
Prolyl	ψ	−16	−28	−34	−22	−33	−27
	ω	4	1	−1	8		
Leucyl	φ	−105	−108	−116	−118	−104	−105
Leucyl	ψ	9	12	15	21	8	8
	ω	0	−6	−1	−1		
Glycinamide	φ	−97	−104	−97	−98		
Glycinamide	ψ	14	18	7	11		

Side-chain angles (sulfur analog)

Residue	Angle	Molecule *A*	Molecule *B*
Cysteinyl	N—Cα–Cβ–S	178°	173°
	Cα–Cβ–S—C	87	270
Leucyl	N—Cα–Cβ–Cγ	292	175
	Cα–Cβ–Cγ–Cδ	273	305

(*a*) Ueki *et al.* (1969). (*b*) Ueki *et al.* (1971); angles converted to newer convention.

Table 5. *Best planes*

(*a*) Equations†

Plane		χ^2	*A*	*B*	*C*	*D*
Molecule *A*						
I	Benzene ring	6·56	0·603	−0·304	0·738	5·505
II	Prolyl ring	151·36	0·945	0·184	−0·269	2·606
III	Peptide group Cys–Pro	23·72	0·931	0·321	−0·175	3·865
IV	Peptide group Pro–Leu	17·24	0·604	−0·764	0·224	1·865
V	Peptide group Leu–Gly	0·35	0·422	0·733	0·533	8·300
VI	Amide group $GlyNH_2$	131·58	−0·804	−0·039	0·592	4·022
Molecule *B*						
VII	Benzene ring	1·88	0·712	9·251	0·655	11·964
VIII	Prolyl ring	38·34	0·967	0·192	−0·166	2·082
IX	Peptide group Cys–Pro	8·74	0·962	0·252	−0·108	3·494
X	Peptide group Leu–Gly	115·36	0·468	0·616	0·633	16·324
XI	Peptide group Pro–Leu	0·91	−0·480	0·876	−0·047	2·368
XII	Amide group $GlyNH_2$	0·05	0·900	−0·189	−0·391	−6·059

Table 5 (*cont.*)

(*b*) Displacements (Å) of atoms from the planes

I		II		III	
C(2)	0·017	Cα	0·082	Cα(Cys)	0·021
C(3)	0·003	Cβ	0·048	C′	0·010
C(4)	0·022	Cγ	0·488*	O	0·006
C(5)	0·019	Cδ	0·053	N	0·032
C(6)	0·002	N	0·088	Cα(Pro)	0·027
C(7)	0·020				
IV		**V**		**VI**	
Cα(Pro)	0·021	Cα(Leu)	0·002	Cα(Gly)	0·003
C′	0·017	C′	0·000	C′	0·089
O	0·002	O	0·001	O	0·016
N	0·025	N	0·005	N	−0·102
Cα(Leu)	0·023	Cα(Gly)	0·004		
VII		**VIII**		**IX**	
C(2)	0·010	Cα	0·042	Cα(Cys)	0·014
C(3)	0·002	Cβ	0·024	C′	0·010
C(4)	0·011	Cγ	0·547*	O	0·002
C(5)	0·007	Cδ	0·026	N	0·017
C(6)	0·006	N	0·044	Cα(Pro)	0·016
C(7)	0·014				
X		**XI**		**XII**	
Cα(Pro)	0·005	Cα(Leu)	0·050	Cα(Gly)	0·000
C′	0·005	C′	0·020	C′	0·002
O	0·000	O	0·016	O	0·001
N	0·005	N	0·076	N	0·001
Cα(Leu)	0·005	Cα(Gly)	0·063		

* These atoms not included in best planes calculation.

† $Ax+By+Cz=D$ is the equation of the plane in orthogonal space (y is parallel to crystal b axis, z parallel to c, and x perpendicular to y and z; x, y, and z are in Å). $\chi^2=\sum_n(\Delta_n/\sigma_n)^2$ where Δ_n is the displacement of the nth atom from the mean plane and σ_n is the standard deviation in the location of the nth atom, calculated for a direction normal to the plane.

As observed elsewhere, four atoms of the proline ring, Cα, Cβ, Cδ, and N, are approximately coplanar (see Table 5) while Cγ is displaced about 0·5 Å from the plane (0·49 Å in molecule *A*, 0·55 Å in molecule *B*) and toward the carboxyl group as in copper proline (Mathieson & Welsh, 1952) and opposite to that in tosylprolylhydroxyproline (Fridrichsons & Mathieson, 1967) and *p*-bromocarbobenzoxy-Gly-L-Pro-L-Leu-Gly (Ueki, Ashida, Kakudo, Sasada & Katsube, 1969). In L-Leu-Pro-Gly (Leung & Marsh, 1958) it is found in both positions at approximately half occupancy. In *o*-bromocarbobenzoxy-Gly-L-Pro-L-Leu-Gly-L-Pro (Ueki, Bando, Ashida & Kakudo, 1971) the γ-carbons are (*a*) opposite the carboxyl in Pro(2) and (*b*) on the same side in Pro(5).

The conformations of the leucyl side chains of molecules *A* and *B* are different. In molecule *B* the β-carbon is very nearly *trans* to the nitrogen and the α-carbon is *trans* to one of the δ-carbons [*cf.* orthorhombic L-leucine hydrobromide (Subramanian, 1967)]. In molecule *A* the β-carbon is *trans* to the carbonyl carbon [*cf.* leucyl residues in L-Leu-L-Pro-Gly (Leung & Marsh, 1958) and *p*-bromocarbobenzoxy-Gly-L-Pro-L-Leu-Gly (Ueki *et al.*, 1969)]. The β–γ bond is, however, twisted by about 30° from the normal staggered conformation towards a conformation in which the hydrogen on the δ-carbon is eclipsed by the α-carbon.

Bond lengths and bond angles for the sulfur and selenium analogs are given in Tables 6 and 7 respectively. None differs significantly from accepted values. The peptide and amide groups are effectively planar, the greatest deviations being in the third peptide bond (Leu–Gly) of molecule *B* and the terminal amide group of molecule *A*, with root mean square deviations from the mean plane of 0·056 and 0·078 Å respectively. None of the other groups has standard deviations greater than 0·024 Å.

Table 6. *Bond lengths* (Å)

Numbers in parentheses are estimated standard deviations ×100.

	Sulfur analog		Selenium analog	
	Molecule *A*	Molecule *B*	Molecule *A*	Molecule *B*
Benzyl group				
S(Se)–C(1)	1·85 (2)	1·82 (1)	1·98 (4)	1·92 (4)
C(1) –C(2)	1·51 (2)	1·56 (2)	1·40 (5)	1·76 (6)
C(2) –C(3)	1·40 (2)	1·42 (2)	1·41 (5)	1·63 (7)
C(3) –C(4)	1·45 (2)	1·41 (3)	1·48 (5)	1·41 (9)
C(4) –C(5)	1·38 (2)	1·44 (3)	1·36 (8)	1·38 (9)
C(5) –C(6)	1·42 (2)	1·36 (3)	1·27 (7)	1·39 (7)
C(6) –C(7)	1·43 (2)	1·34 (2)	1·42 (7)	1·42 (8)
C(7) –C(2)	1·39 (2)	1·42 (2)	1·44 (6)	1·25 (7)
Cysteinyl residue				
Cα—N	1·49 (2)	1·50 (1)	1·52 (3)	1·52 (3)
Cα—Cβ	1·51 (2)	1·51 (2)	1·56 (3)	1·47 (4)
Cβ—S(Se)	1·83 (1)	1·82 (1)	1·99 (2)	1·94 (2)
Cα—C′	1·53 (1)	1·55 (1)	1·43 (4)	1·62 (3)
C′ —O	1·25 (1)	1·26 (1)	1·24 (3)	1·29 (3)
C′ —N	1·34 (1)	1·33 (1)	1·31 (3)	1·29 (3)
Prolyl residue				
Cα—N	1·47 (1)	1·47 (1)	1·45 (3)	1·44 (3)
Cα—Cβ	1·55 (2)	1·53 (2)	1·61 (5)	1·58 (4)
Cβ—Cγ	1·55 (2)	1·55 (2)	1·41 (5)	1·48 (4)
Cγ—Cδ	1·54 (2)	1·55 (2)	1·54 (5)	1·53 (4)
Cδ—N	1·49 (1)	1·49 (1)	1·51 (3)	1·52 (3)
Cα—C′	1·49 (2)	1·52 (2)	1·52 (4)	1·49 (3)
C′ —O	1·26 (1)	1·25 (1)	1·24 (3)	1·29 (3)
C′ —N	1·34 (1)	1·35 (1)	1·31 (3)	1·22 (3)
Leucyl residue				
Cα—N	1·49 (1)	1·51 (1)	1·53 (3)	1·59 (3)
Cα—Cβ	1·55 (2)	1·56 (1)	1·39 (3)	1·48 (3)
Cβ—Cα	1·56 (2)	1·59 (2)	1·57 (3)	1·61 (4)
Cγ—Cδ	1·56 (2)	1·49 (2)	1·53 (4)	1·48 (5)
Cγ—Cδ′	1·52 (2)	1·53 (2)	1·52 (4)	1·50 (5)
Cα—C′	1·54 (2)	1·51 (2)	1·45 (3)	1·44 (3)
C′ —O	1·21 (1)	1·21 (1)	1·18 (3)	1·19 (3)
C′ —N	1·36 (1)	1·37 (1)	1·44 (3)	1·44 (3)
Glycinamide				
Cα—N	1·47 (2)	1·48 (2)	1·41 (4)	1·42 (3)
Cα—C′	1·53 (2)	1·52 (2)	1·56 (4)	1·65 (4)
C′ —O	1·26 (2)	1·23 (1)	1·24 (4)	1·23 (4)
C′ —N	1·31 (2)	1·35 (2)	1·22 (4)	1·32 (4)

Discussion

The peptide backbone conformation is remarkably similar to those of the related portion of *p*-bromo-

Table 7. *Bond angles* (°)

	Sulfur analog		Selenium analog	
	Molecule *A*	Molecule *B*	Molecule *A*	Molecule *B*
Benzyl group				
S(Se)–C(1)–C(2)	109 (1)	114 (1)	114 (2)	106 (2)
C(1) –C(2)–C(3)	117 (1)	118 (1)	117 (3)	117 (4)
C(2) –C(3)–C(4)	118 (1)	114 (1)	118 (3)	109 (4)
C(3) –C(4)–C(5)	119 (1)	122 (2)	120 (4)	127 (6)
C(4) –C(5)–C(6)	121 (2)	119 (2)	121 (5)	120 (5)
C(5) –C(6)–C(7)	121 (1)	123 (2)	126 (5)	115 (4)
C(6) –C(7)–C(2)	117 (1)	119 (1)	116 (4)	129 (5)
C(7) –C(2)–C(3)	124 (1)	123 (1)	120 (3)	117 (4)
C(7) –C(2)–C(1)	119 (1)	119 (1)	123 (3)	118 (4)
Cysteinyl residue				
N—Cα—Cβ	108 (1)	110 (1)	105 (2)	108 (2)
Cα–Cβ—S(Se)	111 (1)	114 (1)	111 (2)	116 (2)
Cβ–S(Se)–Cl	101 (1)	101 (1)	100 (1)	102 (1)
Cβ–Cα—C′	108 (1)	108 (1)	106 (2)	107 (2)
N—Cα—C′	109 (1)	110 (1)	106 (2)	106 (2)
Cα–C′—O	121 (1)	120 (1)	123 (2)	120 (2)
Cα–C′—N	119 (1)	120 (1)	116 (2)	121 (2)
O—C′—N	120 (1)	120 (1)	120 (2)	119 (2)
Prolyl residue				
C′–N –Cα	120 (1)	120 (1)	121 (2)	124 (2)
N—Cα–Cβ	102 (1)	105 (1)	107 (2)	105 (2)
Cα–Cβ–Cγ	104 (1)	103 (1)	103 (3)	103 (2)
Cβ–Cγ–Cδ	105 (1)	105 (1)	115 (3)	107 (2)
Cγ–Cδ–N	102 (1)	101 (1)	101 (2)	102 (2)
Cδ–N—Cα	114 (1)	113 (1)	112 (2)	112 (2)
Cβ–Cα–C′	113 (1)	112 (1)	109 (2)	105 (2)
N—Cα–C′	114 (1)	113 (1)	113 (2)	109 (2)
Cα–C′–O	120 (1)	122 (1)	120 (2)	115 (2)
Cα–C′–N	120 (1)	117 (1)	116 (2)	119 (2)
O—C′–N	120 (1)	121 (1)	124 (2)	125 (2)
Leucyl residue				
C′–N—Cα	120 (1)	121 (1)	121 (2)	124 (2)
N—Cα–Cβ	112 (1)	104 (1)	112 (2)	106 (2)
Cα–Cβ–Cγ	113 (1)	112 (1)	120 (2)	115 (2)
Cβ–Cγ–Cδ	113 (1)	115 (1)	114 (2)	110 (2)
Cβ–Cγ–Cδ′	108 (1)	103 (1)	115 (2)	106 (3)
Cδ–Cγ–Cδ′	109 (1)	112 (1)	102 (2)	110 (2)
Cβ–Cγ–C′	108 (1)	113 (1)	115 (2)	117 (2)
N—Cα–C′	111 (1)	112 (1)	115 (2)	112 (2)
Cα–C′–O	119 (1)	119 (1)	124 (2)	122 (2)
Cα–C′–N	117 (1)	118 (1)	117 (2)	120 (2)
O—C′–N	124 (1)	122 (1)	118 (2)	117 (2)

carbobenzoxy-glycyl-L-prolyl-L-leucyl-glycine (Ueki *et al.*, 1969) and *o*-bromocarbobenzoxy-glycyl-L-prolyl-L-leucyl-glycyl-L-proline (Ueki *et al.*, 1971). These two peptides differ from the tetrapeptide described here only in the N-terminal and C-terminal regions. All three peptides contain a common fragment: –HN–CO–Pro–Leu–NH–CH_2–CO–. In all the structures the amino nitrogen of the glycyl(4) residue is hydrogen bonded to a carbonyl group three residues back along the chain. Comparative dihedral angles for the peptide backbones of these peptides are given in Table 4.

Venkatachalam (1968) has shown that for a system of three linked peptide units there are three general types of conformations for which 4–1 hydrogen bonding of this kind is possible. The conformation defined by Venkatachalam as type 1, with dihedral angles φ_2, ψ_2, φ_3, and ψ_3 equal approximately to $-60°$, $-30°$, $-90°$, and 0° respectively, is clearly the one adopted by these three peptides. This type of conformation has been referred to as a β-twist (Urry & Ohnishi, 1970) a β-fold (Geddes, Parker, Atkins & Beighton, 1968), a 3_{10}-bend (Dickerson *et al.*, 1971), or a U-fold (Ueki *et al.*, 1971). It has been found in cyclohexaglycyl (Karle & Karle, 1963), lysozyme (Blake, Johnson, Mair, North, Phillips & Sarma, 1967), and other peptides and proteins.

The second hydrogen-bonded ring in molecule *B* is descriptively similar to the first but it does not have the type I conformation; the dihedral angles do not fall within the range defined by Venkatachalam. This type of twist conformation can occur only when the nitrogen is in a terminal acid amide.

Conclusions

The peptide –HN–CH–CO–Pro–Leu–NH–CH_2–CO– has been shown to occur in identical conformations in four different solid-state environments: the two molecules in these asymmetric units and the two carbobenzoxy peptides (Ueki *et al.*, 1969, 1971). This conformational stability illustrates the importance of local interactions between near-neighbor peptide units to peptide conformation, suggesting that certain peptide sequences may have an intrinsic tendency to adopt a 3_{10}-bend conformation.

Lewis, Momany & Scheraga (1971) have pointed out that such sequences in proteins would have an important directing effect on the folding of the polypeptide chain. They have devised an empirically-based scheme for predicting the occurrence of bends in proteins. The success of this method is further evidence that the 3_{10}-bend conformation is primarily a product of local forces.

Table 8. *Correlation coefficients*

Molecules *A* and *B*–benzyl carbon 1

	X	*Y*	*Z*	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
X	−0·25	−0·00	0·04	−0·01	−0·01	0·01	−0·02	−0·05	0·04
Y	−0·07	0·32	0·01	0·00	0·04	0·03	−0·02	−0·01	0·04
Z	0·05	−0·01	−0·21	−0·03	−0·05	−0·04	−0·05	−0·01	−0·02
U_{11}	0·04	−0·02	0·00	−0·32	0·06	0·07	0·00	0·02	0·00
U_{22}	−0·01	0·05	0·00	0·07	−0·39	0·07	0·12	−0·01	0·02
U_{33}	−0·03	0·00	0·05	0·09	0·08	−0·16	−0·04	0·04	0·00
U_{12}	−0·05	−0·04	0·02	−0·06	0·00	0·01	0·18	0·03	−0·07
U_{13}	0·07	−0·06	0·01	0·04	−0·01	0·04	0·01	−0·21	−0·02
U_{23}	0·04	−0·05	−0·03	0·03	−0·01	−0·01	−0·04	−0·05	0·22

The inherent tendency to adopt a particular conformation does not necessarily mean that the same sequence must adopt the same conformation under all conditions. N.m.r. studies by Hruby, Brewster & Glasel (1971) on the *S*-benzyl-oxytocin-C-terminal tetrapeptide in $(CD_3)_2SO$ show that the peptide adopts a different conformation in this solvent. Their results suggest that the prolyl residue exists as both *cis* and *trans* isomers in approximately equal amounts.

As the crystal structure of the oxytocin molecule has not yet been determined, we do not know what the C-terminal 'tail' region conformation is *in situ*. There is some information about conformation of oxytocin in solution. The n.m.r. studies of oxytocin [in $(CD_3)_2SO$ solution] by Johnson, Schwartz & Walter (1969) give results consistent with the hypothesis that the C-terminal portion adopts a conformation similar to that of the crystalline tetrapeptide. On the basis of these results Urry & Walter (1971) have predicted a 3_{10}-bend conformation for the C-terminal portion of oxytocin. Conformational energy calculations by Kotelchuk, Scheraga & Walter (1972), starting with the oxytocin structure proposed by Urry & Walter (1971), indicate that the 3_{10}-bend conformation of the 'tail' portion is a stable, low-energy conformation in water.

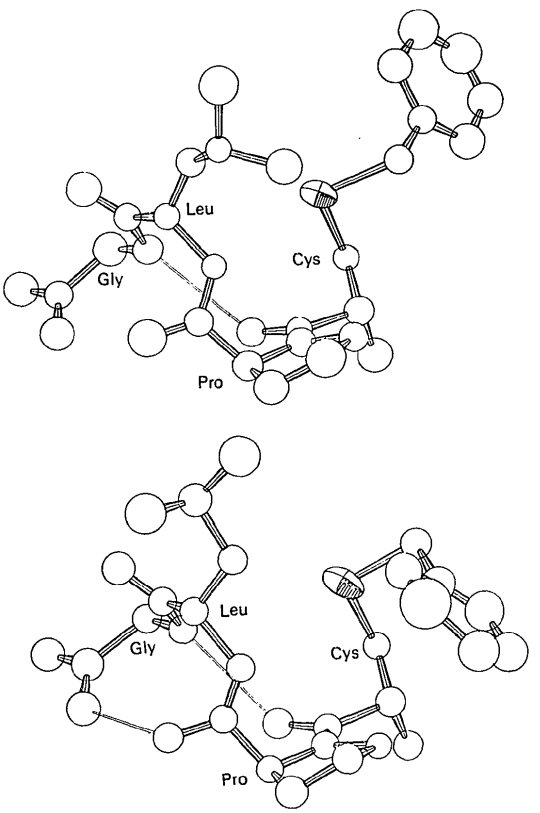


Fig. 5. Thermal ellipsoid plots (50% probability level) of the two molecules of the selenium peptide in the asymmetric unit. Anisotropic temperature factors were obtained only for the selenium atoms.

N.m.r. studies of oxytocin in aqueous solution (Feeney, Roberts, Rockey & Burgen, 1971) have given contradictory results. No detailed conformational information could be derived from this work as it was not possible to definitely assign the peptide NH proton resonances. It was evident that none of the temperature coefficients of the chemical shifts were low enough to suggest hydrogen bonding. However, on the basis of diffusion studies in water, Craig (1964) concluded that the C-terminal 'tail' of the oxytocin molecule is in a very compact, almost spherical, conformation. This is clearly consistent with the structure we have found for the isolated tetrapeptide.

We thank Professor Roderick Walter who provided both synthetic peptides. The intensity data for the sulfur peptide were collected by Dr F. M. Lovell, who carried out preliminary studies on this peptide. We have appreciated and found helpful the critical comments of an anonymous referee, both in agreement and in disagreement. This work was supported by National Institutes of Health Grant AM 01320 and National Science Foundation Grant GB 7272. One of us (A.D.R.) was supported by a Public Health Service Traineeship. We are grateful to the Columbia University Computing Center for the use of their computing facilities.

APPENDIX

Refinement

As discussed earlier, the refinements of the sulfur and selenium analog structures both stopped initially at abnormally high *R* values and only the sulfur analog could finally be refined.

The fact that one of the structures did finally refine suggests that the problems which had inhibited refinement did not lie in the data and were not, therefore, the result of systematic errors in data collection or data handling, or the consequence of gross disorder in the crystal. Disorder may, indeed must, exist in some crystals as indicated by the presence of forbidden $0k0$ reflections for odd values of k, but its effect on the data must be small.

The refinement problems could formally have resulted from errors in the structure model, but this is evidently not the case here. The unusual nature of this crystal structure (two molecules in the asymmetric unit with similar conformations, related by a simple translation) makes it possible to imagine several distinctly different structures which would have similar structure-factor distributions. For example, (*a*) keep the seleniums as they are but interchange the remainder of the two molecules in the asymmetric unit by translating all the light atoms of molecule *A* by 0,0·24,0·48 and all the light atoms of molecule *B* by 0, −0·24, −0·48; (*b*) keep both the seleniums and the benzyl groups as they are and interchange the rest of the structures; (*c*) keep the seleniums and the peptides and interchange the

benzyl groups. When the refinement would not proceed, all these alternatives were considered. The R values are all high [for (*a*) and (*b*) comparable to the R value with selenium alone] and ΔF syntheses show all the shifted atoms to be in incorrect positions (Rudko, 1972). We also calculated ΔF syntheses omitting sections of the structure, *e.g.*, the benzyl groups, the seleniums plus the benzyl groups. We hoped to discover whether these sections of the structure were incorrect or possibly disordered. These maps look as one would expect, with peaks at the positions of the omitted atoms and no very significant peaks elsewhere, but the benzyl groups did come up somewhat weaker and more diffuse than one might expect. The final difference map of the selenium peptide did not suggest any errors in the structure. The principal peaks are satellites about the Se position.

Thus, the problem inhibiting refinement of the selenium structure appears to be of the same kind as the problem which made refinement of the sulfur structure so difficult. Final refinement of the sulfur peptide structure from $R=0{\cdot}2$ to $R=0{\cdot}1$ involved no parameter shifts greater than would be expected in this stage of structure refinement. The principal changes were in the thermal parameters of the sulfur atoms and the positional and thermal parameters of benzene-ring carbon atoms.

If the structural model is essentially correct and the data are satisfactory but the model does not refine, then the problem would appear to lie in the least-squares refinement procedure itself. One limitation of the method of least squares is the problem of 'inverse overlap', the interaction between parameters which arises when a centrosymmetric group of atoms is present in a noncentrosymmetric crystal. This problem has been examined by Srinivasan (1961) and Rae & Maslen (1963). They concluded that accurate refinement of the affected parameters may not be possible by any method, not even by full-matrix least-squares calculations.

An example of the effect of inverse overlap is the structure of guanidinium aluminum sulfate hexahydrate (Geller & Katz, 1962). Good data were available and the general nature of the structure was apparent, but satisfactory refinement was never obtained. Lingafelter, Orioli, Schein & Stewart (1966) re-investigated the problem and found that the difficulties in refinement could be explained in terms of the strong interactions between parameters of atoms related by a pseudocenter at $\frac{5}{6},\frac{1}{6},0{\cdot}54$. They concluded that inverse overlap was responsible for the failure to refine.

An analogous situation exists in the tetrapeptide structure described here. A partial pseudocenter at $0,0{\cdot}38,0{\cdot}75$ relates the selenium (or sulfur) and benzyl group of molecule A to the corresponding portion of molecule B (as the atomic positions are given here the center relates part of molecule A to the corresponding part of a molecule related to B by the screw axis at $x=0,z=1$). Inverse overlap arising from this pseudocenter also affects the atoms of the proline ring and the terminal amide group to a lesser degree. As the fraction of the total electrons in the cell involved in the pseudocentrosymmetry is greater for the selenium analog, the effect should be stronger for the selenium than for the sulfur analog. Thus, the effect of inverse overlap could explain the difficulty in refining both peptides as well as the erratic behavior of the benzene-ring parameters. Examination of the correlation matrix reveals covariance terms as large as 0·4 between corresponding parameters of atoms related by the pseudocenter. Examples of these types of correlations are given in Table 8.

Disorder

We have established that the observed $0k0$ (k odd) reflections are not Renninger reflections nor do they imply $P2$ symmetry. It appears, therefore, that disorder effects involving local deviations from $P2_1$ symmetry may exist.

The rather unusual packing of the peptide molecules in the crystal suggests possible ways in which 'faults' may occur in the crystal with local deviations from $P2_1$ symmetry but little distortion of the lattice (lattice distortion would result in diffuse reflections). The crystal structure (see Fig. 1) consists of alternating sheets of type A and type B molecules, with chemical environments of the types A and B different in spite of relatively few conformational differences.

One can imagine a 'fault' consisting of a translation of part of the structure by $b/2$ in such a way that a sheet of type A molecules becomes related to the next sheet of type B molecules by the translation $-b/4+c/2$. Effectively, this means that a sheet of molecules is on one side (towards $z=0$) in the normal chemical environment of type A molecules and on the other side (towards $z=1$) in the normal chemical environment of type B molecules. As there are not very many differences between A and B it is not difficult to imagine either molecule or an intermediate form fitting in such an environment. Such a molecule would have type A packing interactions at the N-terminal end and type B packing interactions at the C-terminal end. The opposite type of 'fault' may, of course, also occur: a sheet of molecules with the environment of B at the N-terminal end and of A at the C-terminal end.

'Faults' of this kind, in which part of the structure is dislocated by $b/2$, would have little overall effect on the observed structure factors if they occurred infrequently. As the $P2_1$ symmetry is broken at the 'fault', weak $0k0$ reflections for odd values of k would be observed.

References

BLAKE, C. C. F., JOHNSON, L. N., MAIR, G. A., NORTH, A. C. T., PHILLIPS, D. C. & SARMA, V. R. (1967). *Proc. Roy. Soc.* B**167**, 378–388.

CRAIG, L. (1964). *Science*, **144**, 1093–1099.

DICKERSON, R. E., TAKANO, T., EISENBERG, D., KALLAI, O. B., SAMSON, L., COOPER, A. & MARGOLIASH, E. (1971). *J. Biol. Chem.* **246**, 1511–1533.
FEENEY, J., ROBERTS, G. C. K., ROCKEY, J. H. & BURGEN, A. S. V. (1971). *Nature New Biol.* **232**, 108–110.
FRIDRICHSONS, J. & MATHIESON, A. MCL. (1967). *Acta Cryst.* **15**, 569–577.
GEDDES, A. J., PARKER, K. D., ATKINS, D. T. & BEIGHTON, E. (1968). *J. Mol. Biol.* **32**, 343–358.
GELLER, S. & KATZ, H. (1962). *Bell Syst. Tech.* **41**, 425–439.
HRUBY, V. J., BREWSTER, A. I. & GLASEL, J. A. (1971). *Proc. Natl. Acad. Sci. U.S.* **68**, 450–453.
International Tables for X-ray Crystallography (1962). Vol. III, pp. 210–212. Birmingham: Kynoch Press.
IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *Biochemistry*, **9**, 3471.
JOHNSON, L. F., SCHWARTZ, I. L. & WALTER, R. (1969). *Proc. Natl. Acad. Sci. U.S.* **64**, 1269–1275.
KARLE, I. L. & KARLE, J. (1963). *Acta Cryst.* **16**, 969–975.
KOTELCHUCK, D., SCHERAGA, A. & WALTER, R. (1972), *Proc. Natl. Acad. Sci. U. S.* **69**, 3629–3633.
LEUNG, Y. C. & MARSH, R. E. (1958). *Acta Cryst.* **11**, 17–31.
LEWIS, P. N., MOMANY, F. A. & SCHERAGA, H. A. (1971). *Proc. Natl. Acad. Sci. U.S.* **68**, 2293–2297.
LINGAFELTER, E. C., ORIOLI, P. L., SCHEIN, J. B. & STEWART, J. B. (1966). *Acta Cryst.* **20** 451–455.
LOW, B. W., LOVELL, F. M. & RUDKO, A. D. (1967). *Proc. Natl. Acad. Sci. U.S.* **60**, 1519–1526.
LOW, B. W. & RICHARDS, F. M. (1952). *J. Amer. Chem. Soc.* **74**, 1660.
MATHIESON, A. MCL. & WELSH, H. K. (1952). *Acta Cryst.* **5**, 599–604.
OUGHTON, B. & HARRISON, P. M. (1959). *Acta Cryst.* **12**, 396–404.
PETERSON, J., STEINRAUF, L. K. & JENSEN, L. H. (1960). *Acta Cryst.* **13**, 104–109.
RAE, A. I. M. & MASLEN, E. N. (1963). *Acta Cryst.* **16**, 703–704.
RUDKO, A. D. (1972). Ph.D. Dissertation, Columbia Univ.
RUDKO, A. D., LOVELL, F. M. & LOW, B. W. (1971). *Nature New Biol.* **232**, 18–19.
SIM, G. A. (1961). In *Computing Methods and the Phase Problem*, edited by R. PEPINSKY, J. M. ROBERTSON & J. C. SPEAKMAN, pp. 227–235. Oxford: Pergamon Press.
SRINIVASAN, R. (1961). *Acta Cryst.* **14**, 1163–1167.
STEWART, J. M. (1967). *X-RAY 67 Programs for X-ray Crystallography*. Technical Report 67–58, Computer Center, Univ. of Maryland.
SUBRAMANIAN, E. (1967). *Acta Cryst.* **22**, 910–917.
UEKI, T., ASHIDA, T., KAKUDO, M., SASADA, Y. & KATSUBE, Y. (1969). *Acta Cryst.* B**25**, 1840–1849.
UEKI, T., BANDO, S., ASHIDA, T. & KAKUDO, M. (1971). *Acta Cryst.* B**27**, 2219–2231.
URRY, D. W. & OHNISHI, M. (1970). In *Spectroscopic Approaches to Biomolecular Conformation*, edited by D. W. URRY, pp. 263–299. Chicago: Amer. Med. Assoc. Press.
URRY, D. W. & WALTER, R. (1971). *Proc. Natl. Acad. Sci. U.S.* **68**, 956–958.
VENKATACHALAM, C. M. (1968). *Biopolymers*, **6**, 1425–1436.
WILSON, A. J. C. (1950). *Acta Cryst.* **3**, 397–398.
WRIGHT, W. B. (1958). *Acta Cryst.* **11**, 632–642.
YAKEL, H. L. & HUGHES, E. W. (1954). *Acta Cryst.* **7**, 291–297.

Acta Cryst. (1975). B**31**, 725

Structure Cristalline de $Sr(CH_3COO)(CH_3COS).4H_2O$

PAR M. M. BOREL ET M. LEDÉSERT

Laboratoire de Chimie Minérale B et Laboratoire de Cristallographie-Minéralogie, Groupe de Cristallographie et Chimie du Solide (E.R.A. n° 305), U.E.R. des Sciences, Université, 14032 Caen Cedex, France

(*Reçu le* 23 *juillet* 1974, *accepté le* 14 *octobre* 1974)

The crystals of strontium ethanoate thioethanoate are monoclinic, space group $P2_1/c$ with four molecules in a unit cell of dimensions $a=12{\cdot}72$, $b=7{\cdot}09$, $c=12{\cdot}97$ Å, $\beta=111{\cdot}13$ (38)°. The structure was solved by the heavy-atom method from 871 intensities collected from layers $h0l$ to $h5l$ on a Weissenberg camera. The final $R=0{\cdot}062$. The coordination of the strontium ion is nine: one sulphur atom and eight oxygen atoms.

Dans le cadre d'une étude comparative thermochimique et structurale des composés oxygénés et sulfurés, un certain nombre de thioéthanoates et dithioéthanoates ont été préparés (Borel & Ledésert, 1973; Bernard, Borel & Ledésert, 1973).

La méthode de préparation utilisée nous a conduit pour le calcium et le strontium à deux composés originaux, l'éthanoate thioéthanoate de calcium trihydraté $Ca(CH_3COO)(CH_3COS).3H_2O$ et l'éthanoate thioéthanoate de strontium tétrahydraté $Sr(CH_3COO)(CH_3COS).4H_2O$.

La mise en évidence (Bernard & Borel, 1972) de ces composés originaux nous a amené à entreprendre les déterminations deleurs structures cristallines en commençant par $Sr(CH_3COO)(CH_3COS).4H_2O$.