Hydrogen Bonding of Cysteine in the Solid Phase. Crystal and Molecular Structures of L-Cysteine Methyl Ester·HCl and L-Cysteine Ethyl Ester·HCl

C. H. Görbitz

Department of Chemistry, University of Oslo, P.O. Box 1033, Blindern N-0315 Oslo 3, Norway

Görbitz, C. H., 1989. Hydrogen Bonding of Cysteine in the Solid Phase. Crystal and Molecular Structures of L-Cysteine Methyl Ester·HCl and L-Cysteine Ethyl Ester·HCl. – Acta Chem Scand. 43: 871–875.

The hydrochloride salts of L-cysteine methyl ester (CM) and L-cysteine ethyl ester (CE) both crystallize in the orthorhombic space group $P2_12_12_1$. Unit cell parameters at 120 K for CM are $a = 5.128(1)$, $b = 11.779(2)$, $c = 13.498(2)$ Å; those for CE are $a = 5.149(1)$, $b = 8.085(1)$, $c = 21.302(3)$ Å. The structures were solved by direct methods and refined to conventional R factors 0.026 and 0.027 for CM and CE, respectively. The two molecules adopt remarkably similar conformations in the solid phase with the cysteine side chain +*gauche*. The thiol group of CE participates in a hydrogen bond as the donor with Cl^- as the acceptor. In CM, only a very weak interaction of this type is observed. Based on a survey of crystal structures involving cysteine, it is suggested that interactions of the types S–H···O, S–H···S and S–H···Cl^- are weakly attractive in the solid phase.

The term hydrogen bond for a complex X–H···Y is normally associated with the electronegative elements O, N, F and Cl. A statistical investigation of crystal structures has, however, revealed that even carbon displays a significant tendency to form short intermolecular contacts of the type C–H···O, C–H···N and C–H···Cl.[1] Sulfur is as electronegative as carbon and may exhibit similar properties. The role of the thiol group as a hydrogen bond donor and acceptor has been studied by a number of authors,[2–6] but in the solid state little evidence has been provided so far. Only a limited number of crystal structure determinations of compounds involving the amino acid residue cysteine are known,[7–15] and only in three has the thiol group hydrogen atom been located.

The thiol group of cysteine takes part in a variety of biochemical reactions.[16,17] The possible formation of weak hydrogen bonds at receptor sites is of considerable interest, as it might contribute to the biological response. The purpose of this work was to increase the number of high-precision X-ray studies of cysteine residues in order to extend the understanding of its hydrogen bonding properties as well as its conformational characteristics.

Experimental

Crystals of L-cysteine methyl ester·HCl (CM) were grown from a 1:1 water–methanol solution, crystals of L-cysteine ethyl ester·HCl (CE) from an aqueous solution. The data collection procedures are summarized in Table 1. Cell parameters were determined by a least-squares fit to the diffractometer settings for 25 general reflections. Standard deviations in the measured intensities were calculated as $\sigma I = [C_T+(0.02C_N)^2]^{1/2}$, where C_T is the total number of counts and C_N is the scan count minus the background count. The intensities were corrected for Lorentz and polarization effects, but not for absorption. Both structures were solved by MULTAN[18] and refined isotropically with subsequent introduction of all hydrogen atoms connected to carbon and nitrogen in theoretical positions. The thiol group hydrogen atoms were obtained from difference Fourier syntheses. All positional parameters, isotropic temperature factors for the hydrogen atoms, and anisotropic temperature factors for the non-hydrogen atoms

Table 1. Data collections.

Instrument	Nicolet P3	
Radiation	Graphite crystal Monochromated Mo K_α	
Scanning mode	$\theta/2\theta$	
Scan speed/° min^{-1}	3.0	
Scan range/°	$2\theta_{\alpha1}$ −0.8 to $2\theta_{\alpha1}$ +0.9	
Background count	For 35 % of scan time at scan limits	
Temperature/K	120	
2θ range/°	3.0–70.0	
	CM	CE
Crystal dimensions/mm	0.55×0.45×0.40	0.55×0.50×0.35
No. of refl. measured	2123	2310
No. of unique refl. $I>2.5\sigma I$	1969	2187

Table 2. Fractional coordinates with standard deviations and equivalent isotropic temperature factors, B_{eq}, for non-hydrogen atoms.

Atom	x	y	z	B_{eq}/Å²
L-Cys methyl ester·HCl				
Cl⁻	1.0703(1)	0.4189(1)	0.5258(1)	1.7
SG	0.8144(1)	0.4215(1)	0.7978(1)	1.9
O1	0.6002(2)	0.7330(1)	0.7709(1)	2.1
O2	0.8327(2)	0.6601(1)	0.6446(1)	2.2
N1	0.9384(3)	0.6610(1)	0.9082(1)	1.7
CA	1.0121(3)	0.6429(1)	0.8028(1)	1.4
CB	1.0806(3)	0.5195(1)	0.7836(1)	1.7
C1	0.7887(3)	0.6848(1)	0.7391(1)	1.6
CM	0.6273(4)	0.6924(2)	0.5764(1)	3.0
HS	0.747(7)	0.422(2)	0.718(2)	
HN11	0.917(5)	0.737(2)	0.922(2)	
HN12	1.056(5)	0.633(2)	0.948(2)	
HN13	0.792(4)	0.627(2)	0.922(2)	
HCA1	1.161(4)	0.693(2)	0.789(1)	
HCB1	1.148(4)	0.510(2)	0.711(1)	
HCB2	1.221(5)	0.499(2)	0.828(1)	
HCM1	0.471(5)	0.652(2)	0.604(2)	
HCM2	0.654(5)	0.659(2)	0.517(2)	
HCM3	0.618(6)	0.771(2)	0.569(2)	
L-Cys ethyl ester·HCl				
Cl⁻	1.3404(1)	0.1393(1)	0.3048(1)	1.6
SG	1.0561(1)	0.3514(1)	0.4448(1)	1.6
O1	1.0898(2)	0.6508(1)	0.3178(1)	1.7
O2	0.7989(2)	0.7239(1)	0.3925(1)	1.6
N1	0.8284(2)	0.3660(1)	0.2919(1)	1.3
CA	0.7370(2)	0.4677(1)	0.3456(1)	1.1
CB	0.7408(3)	0.3715(2)	0.4075(1)	1.3
C1	0.8994(2)	0.6241(1)	0.3493(1)	1.2
CE1	0.9344(3)	0.8800(2)	0.4038(1)	1.8
CE2	1.1743(3)	0.8537(2)	0.4427(1)	2.2
HS	1.165(6)	0.269(3)	0.404(1)	
HN11	0.711(5)	0.270(3)	0.286(1)	
HN12	0.995(5)	0.332(3)	0.296(1)	
HN13	0.816(5)	0.423(3)	0.257(1)	
HCA1	0.567(4)	0.503(2)	0.337(1)	
HCB1	0.675(5)	0.262(3)	0.403(1)	
HCB2	0.638(5)	0.428(3)	0.438(1)	
HE11	0.797(5)	0.948(3)	0.429(1)	
HE12	0.974(5)	0.935(3)	0.362(1)	
HE21	1.245(5)	0.956(3)	0.455(1)	
HE22	1.293(6)	0.773(4)	0.420(1)	
HE23	1.117(6)	0.800(3)	0.484(1)	

were refined by least-squares methods, giving $R = 0.026$ and $R_w = 0.032$ with goodness of fit $S = [\Sigma w\Delta^2/(m-n)]^{1/2} = 2.09$, and $R = 0.027$ and $R_w = 0.035$, with $S = 2.38$ for CM and CE, respectively. The final parameters are given in Table 2. Atomic scattering factors for free heavy atoms and spherically bonded hydrogen atoms were taken from Ref. 19.

Lists of structure factors and anisotropic thermal parameters are available from the author on request.

Crystal data. L-Cysteine methyl ester·HCl, $C_4H_{10}ClNO_2S$: orthorhombic, $a = 5.128(1)$, $b = 11.779(2)$, $c = 13.498(2)$ Å, $V = 815.3(2)$ Å³, $F_w = 171.6$, $Z = 4$, $F_{000} = 360$, space group $P2_12_12_1$, $D_C = 1.398$ g cm^{-3}.

L-Cysteine ethyl ester·HCl, $C_5H_{12}ClNO_2S$: orthorhombic, $a = 5.149(1)$, $b = 8.085(1)$, $c = 21.302(3)$ Å, $V = 886.8(2)$ Å³, $F_w = 185.7$, $Z = 4$, $F_{000} = 392$, space group $P2_12_12_1$, $D_C = 1.391$ g cm^{-3}.

Description and discussion

Views of both molecules with associated chloride ions showing atomic numbering, bond lengths and bond angles are presented in Figs. 1 and 2. All values compare favourably with the structures of other amino acids and peptides, and there are no important differences between corresponding parameters in the two molecules. The largest deviation is for the S–H bond in the thiol group, the value 1.13 Å obtained for CM seems to be an underestimation for this type of bond. O2–CM (1.449 Å in CM) is close to the expected value for methyl esters (1.448 Å),[20] but the bond O2–CE1 (1.462 Å in CE) is rather long compared with the mean value for O–C in other esters (1.452 Å).[20]

As can be seen from Figs. 1 and 2, the common part of CM and CE appears remarkably similar in the two crystal structures. This is confirmed by the torsion angles given in Table 3. An almost identical conformation is also observed for the cysteine ethyl ester·HCl–urea (1:1) complex.[10] Two features are of particular interest: (1) the torsion angle N1–CA–C1–O2 is very close to 180° in both structures, which means that the N1 atom is in the plane of the ester group. About 16 % of the amino acid and peptide esters in the Cambridge Structural Database (Jan. 1989 release)[21] have N–C$^\alpha$–C′–O″ torsion angles in the interval 180 ± 10°. An additional 5 % are in the interval 0 ± 10°. Accordingly, planar sytems are observed in 21 % of all structures. Substituting a nitrogen atom for O2, this torsion angle corresponds to ψ in peptides and proteins, which invariably adopts angles in the whole interval between −70 and 180°. For cysteine residues ψ or corresponding torsion angles are always close to 0 or 180°. There is no evident explanation for this planarity. (2) The single bond between C$^\alpha$ and C$^\beta$ essentially gives rise to three staggered conformations with

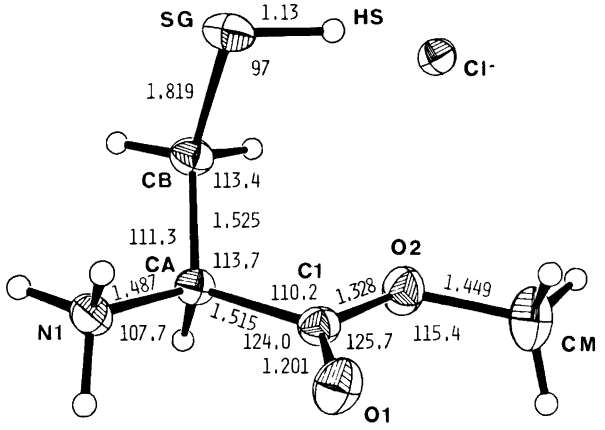

Fig. 1. View of L-cysteine methyl ester. The e.s.d.s are 0.002 Å and 0.1° for bond lengths and bond angles, respectively.

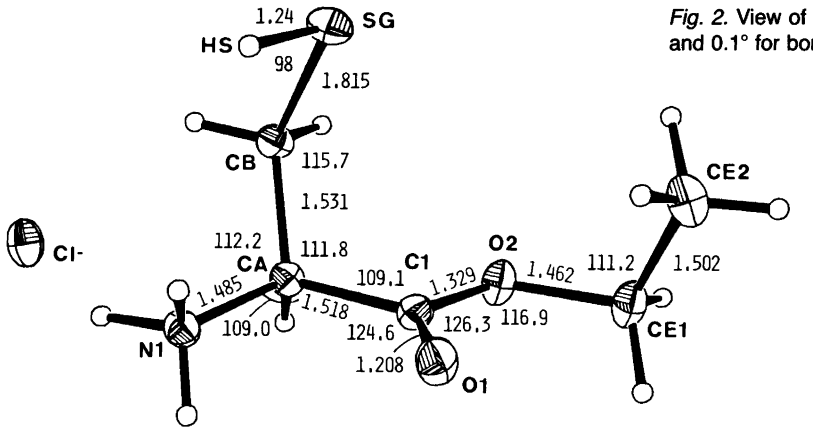

Fig. 2. View of L-cysteine ethyl ester. The e.s.d.'s are 0.002 Å and 0.1° for bond lengths and bond angles, respectively.

Table 3. Principal torsion angles (°) with standard deviations.

	L-Cys methyl ester·HCl	L-Cys ethyl ester·HCl
ψ_T (N1, CA, C1, O2)	−173.7(1)	172.8(1)
χ^1 (N1, CA, CB, SG)	66.0(1)	78.9(1)
χ^2 (CA, CB, SG, HS)	91.7(15)	−63.9(11)
θ_T (CA, C1, O2, CE1/CM)	177.8(1)	178.4(1)
θ' (C1, O2, CE1, CE2)	–	−76.9(2)

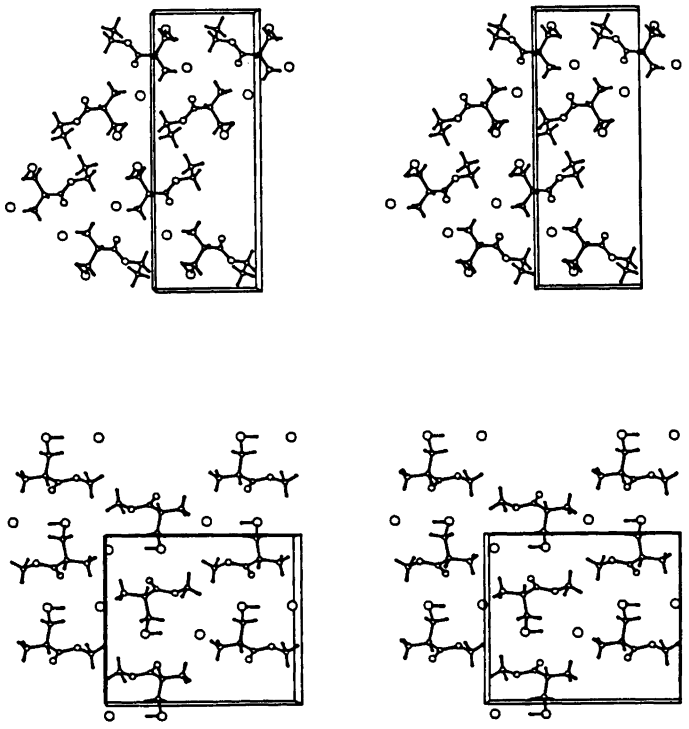

Fig. 3. Stereoscopic drawings of the crystal packing of L- cysteine methyl ester·HCl (top) and L-cysteine ethyl ester·HCl, both viewed along the *a*-axis.

Table 4. Hydrogen bond distances (Å) and angles (°).

D	H	A	D–H	D···A	H···A	D–H···A
L-Cys methyl ester·HCl						
SG	HS	Cl^-	1.13	3.899	3.08	130[a]
N1	HN11	Cl^-	0.92	3.165	2.26	169
N1	HN12	Cl^-	0.87	3.123	2.27	166
N1	HN13	Cl^-	0.87	3.195	2.39	154
CA1	HCA1	O1	0.99	3.223	2.31	153
L-Cys ethyl ester·HCl						
SG	HS	Cl^-	1.24	3.739	2.52	169
N1	HN11	Cl^-	0.99	3.122	2.22	152
N1	HN12	Cl^-	0.91	3.223	2.37	156
N1	HN13	Cl^-	0.88	3.144	2.33	153

[a]van der Waals' contact.

N–C^{α}–C^{β}–S^{γ} (χ^1) ≈ 60 (g^+), 180 (t) or −60° (g^-). Their relative frequencies in protein structures are 0.14, 0.25 and 0.61, respectively.[22] On the other hand, a statistical survey on peptide structures[23] gave the figures 0.50, 0.31 and 0.19, indicating a slight preference for g^+. When the cysteine residue is not substituted on the sulfur atom or is part of a cyclic system, the preference for g^+ is very strong with χ^1 in the interval 62–79°.[24] The two compounds described in the present paper confirm this statement. The conformational properties of cysteine will be described in detail elsewhere.[24]

The crystal packings with the unit cells are illustrated in Fig. 3. Three-dimensional hydrogen-bond networks link the molecules in both structures. Data for intermolecular hydrogen bonds are given in Table 4. The $-NH_3^+$ groups donate all their hydrogen atoms to the chloride ion in each structure. In addition, a short hydrogen bond-like interaction of the type C^{α}–H···O is seen in the structure of CM. More interest is focused on the role of the thiol groups in the hydrogen bonding. As can be seen from Table 4, the –SH group of CE donates its proton to form a moderately strong hydrogen bond to Cl^-, but does not accept any hydrogen bond. In the structure of CM a short contact exists between SG and O1 (3.211 Å), but the angle CB–SG···O1 (156°) is unfavourable for hydrogen bonding, and HS···O1 is 2.86 Å, beyond the sum of the van der Waals' radii. As only a van der Waals contact exists to Cl^- (Table 4), the thiol group is evidently not involved in any hydrogen bond-like interaction in the crystal structure considered.

As mentioned in the introduction, the lack of high-precision studies on cysteine precludes any firm conclusions regarding its hydrogen-bonding properties. A summary of crystal structures with cysteine residues is given in Table 5. In cases when hydrogen atoms were not localized in the original work, theoretical hydrogen atom positions were

Table 5. Observed and assumed hydrogen bonds in cysteine (Cys) structures with distances (Å) and angles (°).

Compound	Hydrogen bonds	C–D···A	D–H	D···A	H···A	D–H···A	Ref.
L-Cys-Gly NaI	None						7
L-Cys·HCl·H_2O	S–H···O	117	*1.25*[a]	3.43	*2.29*	*150*	8
or	S–H···Cl^-	151	*1.25*	3.50	*2.94*	*106*	
DL-Cys^- Tl^+	N–H···S	134	*0.93*	3.31	*2.46*	*147*	9
L-Cys ethyl ester·HCl urea complex	S–H···O	87	*1.25*	3.50	*2.28*	*164*	10
γ-L-Glu-L-Cys-Gly (Glutathione)	S–H···O	91	1.21	3.50	2.33	161	11
L-Cys (monoclinic)	S–H···O	96	*1.25*	3.48	*2.23*	*180*	12
or	S–H···S	106	*1.25*	3.68	*2.45*	*167*	
L-Cys (orthorhombic)[b,c]	S–H···O	101	1.25	3.39	2.40	134	13, 14
and	S–H···S	96	1.30	3.85	2.75	140	
N-Acetyl-L-Cys[b]	S–H···O	91	1.34	3.47	2.22	149	15
	N–H···S	–[d]	1.01	3.89	2.75	152	
L-Cys methyl ester·HCl	None						This work
L-Cys ethyl ester·HCl	S–H···Cl^-	92	1.24	3.74	2.52	169	This work

[a]Optimized values in italic types when the thiol hydrogen atom was not located, see the text. [b]Neutron diffraction values. [c]Thiol group disordered. [d]sp^2-nitrogen.

calculated: (1) hydrogen atom pointing in the direction of the acceptor; (2) bond angles C–S–H = 97°, C–N–H = 114°; (3) distances S–H = 1.25 Å, N–H = 0.93 Å. These are optimized geometries, and the data should therefore be treated with some caution. The true values for the distances H···A are likely to be slightly higher than indicated in the table, if the bond exists at all. The presence of an S–H···Cl^- hydrogen bond in the structure of L-cysteine · HCl · H_2O has been suggested,[8] but the data in Table 5 indicate that an S–H···O hydrogen bond is somewhat more favourable.

Realization and optimization of hydrogen bonds, e.g. in peptide structures, is of profound importance in establishing the build-up of a crystal. Naturally, one cannot expect the formation of hydrogen bonds of the types S–H···X and X–H···S to be the most important, but it seems from Table 5 that at least the former occurs quite frequently with O, S or Cl^- as the acceptor. This in itself suggests that the interaction is attractive, presumably through electrostatic stabilization. It is, however, not mandatory to have the –SH group involved in hydrogen bonds, as the structure of CM clearly demonstrates.

Conclusion

Available material on the amino acid cysteine is sparse, but indicates that short intermolecular contacts of the type S–H···X (X = O, Cl^-, S) are reasonably common. These interactions are then likely to be attractive rather than repulsive and of importance in determining the packing arrangement of the corresponding structure.

References

1. Taylor, R. and Kennard, O. *J. Am. Chem. Soc. 104* (1982) 5063.
2. Mikenda, W. *J. Mol. Struct. 147* (1986) 1.
3. Rao, C. N. R., Dwivedi, A. G., Randhawa, H. S., Ratajczak, H., Szczesniak, M. M., Romanowska, K. and Orville-Thomas, W. J. *J. Mol. Struct. 30* (1976) 271.
4. Westhaus, P. A. and Pohl, H. A. *J. Theor. Biol. 70* (1978) 157.
5. de Alencastro, R. B. and Sandorfy, C. *Can. J. Chem. 51* (1973) 1443.
6. Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y. and Winter, G. *Nature (London) 314* (1985) 235.
7. Dryer, H. B. *Acta Crystallogr. 4* (1951) 42.
8. Ayyar, R. R. *Z. Kristallogr. 126* (1968) 227.
9. Freeman, H. C. and Moore, C. J. *Acta Crystallogr., Sect. B 33* (1977) 2690.
10. Haas, D. J. *Acta Crystallogr. 19* (1965) 860.
11. Görbitz, C. H. *Acta Chem. Scand., Ser. B 41* (1987) 362.
12. Harding, M. M. and Long, H. A. *Acta Crystallogr., Sect. B 24* (1968) 1096.
13. Kerr, K. A. and Ashmore, J. P. *Acta Crystallogr., Sect. B 29* (1973) 2124.
14. Kerr, K. A., Ashmore, J. P. and Koetzle, T. F. *Acta Crystallogr., Sect. B 31* (1975) 2022.
15. Takusagawa, F., Koetzle, T. F., Kou, W. W. H. and Parthasarathy, R. *Acta Crystallogr., Sect. B 37* (1981) 1591.
16. Friedman, M. *The Chemistry and Biochemistry of Sulfhydryl Groups in Amino Acids, Peptides and Proteins,* Pergamon Press, New York 1973.
17. Jocelyn, P. C. *Biochemistry of the SH Groups,* Academic Press, New York 1972.
18. Germain, G., Main, P. and Woolfson, M. M. *Acta Crystallogr., Sect. A 27* (1971) 368.
19. *International Tables for X-Ray Crystallography,* Kynoch Press, Birmingham 1974, Vol. IV.
20. Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. and Taylor, R. *J. Chem. Soc., Perkin Trans. 2* (1987) S1.
21. Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R. and Watson, D. G. *Acta Crystallogr., Sect. B 35* (1979) 2331.
22. Ponder, J. W. and Richards, F. M. *J. Mol. Biol. 193* (1987) 775.
23. Benedetti, E., Morelli, G., Némethy, G. and Scheraga, H. A. *Int. J. Pept. Protein Res. 22* (1983) 1.
24. Görbitz, C. H. *To be published.*

Received March 10, 1989.