

The Crystal Structure of L-Cysteic Acid Monohydrate

BY WAYNE A. HENDRICKSON* AND J. KARLE

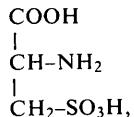
Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D.C. 20390, U.S.A.

(Received 6 February 1970)

The structure of L-cysteic acid monohydrate ($C_3H_7NO_5S \cdot H_2O$) has been determined by an X-ray diffraction analysis of a crystal in space group $P2_{1}2_{1}2_{1}$ having 4 molecules in a unit cell of dimensions $a = 6.927 \pm 1$, $b = 19.027 \pm 3$, $c = 5.305 \pm 1$ Å; density $D_x = 1.77$, $D_{\text{meas}} = 1.75 \text{ g.cm}^{-3}$. The intensity data were measured on an automatic diffractometer, phases were determined by the symbolic addition procedure and the structure was refined to an R of 6.7% by the method of least squares. The dimensions and conformation of the molecule are comparable to those of L-cysteine and taurine. The water of hydration is in the orientation determined by proton magnetic resonance. The molecules are packed in layers in the crystal and are held together by hydrogen bonding and ionic interactions.

Introduction

L-Cysteic acid, with structural formula



is a naturally occurring amino acid, although not one which is normally incorporated into proteins. It does occur in the part of a sheep's fleece exposed to light and weather and is an intermediate in the metabolic degradation of L-cysteine. It is also commonly used in the analysis of the amino-acid composition of proteins through the chromatographic identification of oxidized cysteine and cystine (Greenstein & Winitz, 1961).

Preliminary crystallographic work on cysteic acid, including proton magnetic resonance measurements on the proton-proton vectors of the water of hydration, has been reported (El Saffar, Hendrickson & Koski, 1969). The structures of the closely related compounds L-cysteine and taurine, decarboxylated cysteic acid, have also been reported (Ramachandra Ayyar, 1968; Harding & Long, 1968; Sutherland & Young, 1963; Okaya, 1966).

Experimental

A crystal from the same batch as those described by El Saffar *et al.* (1969) was mounted along the c axis for data collection on a Picker automated four-circle diffractometer. Measurements of 2θ values were fitted by a least-squares procedure to yield the lattice parameters and associated standard deviations which are listed in Table 1 together with other physical data. These cell constants differ from those based on film data (El Saffar *et al.*, 1969) by $3-5\sigma$. All the accessible intensity data ($\theta_{\text{max}} = 66.25^\circ$) were measured by $\theta-2\theta$ scans of $2^\circ + 2\theta(\alpha_2) - 2\theta(\alpha_1)$ at 2 deg.min^{-1} with 10

sec background counts before and after each peak. Lorentz and polarization corrections were applied to the net counts and normalized structure factors $|E|$ and structure factors $|F|$ were derived on an approximately absolute scale from a K -curve (Karle & Hauptman, 1953) of the data. No correction for absorption is included in the results cited below. The effect of an absorption correction was not greater than $\frac{1}{3}\sigma$ in bond lengths and angles nor greater than $\frac{1}{2}\sigma$ in thermal parameters, excepting B_{33} values which were decreased by 0.18 Å² on the average.

Table 1. Physical data

Molecular formula	$C_3H_7NO_5S \cdot H_2O$
Molecular weight	187.17
Habit	Prismatic c
Crystal size	$0.08 \times 0.11 \times 0.21$ mm
Space group	$P2_{1}2_{1}2_{1}$
Z	4 molecules/unit cell
a	6.927 ± 0.001 Å
b	19.027 ± 0.003
c	5.305 ± 0.001
V	699.20 Å ³
Density, flotation*	1.75 g.cm ⁻³
Density, calculated	1.77
Radiation	$Cu K\alpha$, 1.5418 Å; Ni-filtered
Linear absorption coefficient, μ	40.2 cm ⁻¹
Temperature	22°C, ambient
Number of independent reflections	741

* El Saffar, Hendrickson & Koski (1969).

Structure analysis

The phases for the structure factor amplitudes were determined by an application of the symbolic addition procedure for non-centrosymmetric crystals (Karle & Karle, 1966). The reflections which were assigned phases to specify an origin and enantiomorph are listed in Table 2 along with those given symbolic assignments. Other phase indications were obtained by hand from these by using the relationship

$$\varphi_h = \langle \varphi_k + \varphi_{h-k} \rangle_{k_r}$$

with interacting reflections restricted to those with

* A Resident Postdoctoral Research Associate of the National Research Council.

$|E| > 1.75$. The recurring indications that $b = a$ and $c = a + \pi/2$ left two possible phase-sets depending on the value of a . Each of these sets of 48 phases was then gradually expanded with the tangent formula (Karle & Hauptman, 1956; program of Brenner & Gum, 1968) to generate phases for the 275 reflections with $|E| > 1.0$. The seven major peaks in an E -map synthesized from the set following from $a = -\pi/2$ were interpretable as atoms of the structure. Phases calculated from this partial structure were applied to reflections with $|E| > 1.4$ for which $|F|$ calculated at least 0.25 of the observed value. These reflections were then entered into the tangent formula for refinement of the partial structure phases and extension to all reflections with $|E| > 1.0$ (Karle, 1968). The four remaining non-hydrogen atoms were revealed in the succeeding E -map. The structure resulting from this phase determination proved to be opposite to the stereoconfiguration known for L-amino acids. The enantiomorphous structure was produced by inversion through the (001) plane and as a consequence is referred to an origin different from that specified by the phases in Table 2.

Table 2. Reflections used to initiate the phase determination

	h	ϕ_h	$ E_h $
	0 2 5	0	2.32
	7 0 3	$+\pi/2$	2.08
	2 3 0	0	2.07
	0 11 1	$+\pi/2$	2.46
	0 1 6	$a (\pm \pi/2)$	2.41
	2 2 4	b	2.17
	3 1 1	c	2.09

The structure was refined by full-matrix least-squares minimization of $\sum w(|F_o| - |F_c|)^2$ using the Busing, Martin & Levy (1962) program with scattering factors as listed in *International Tables for X-ray Crystallography* (1962). No allowance was made in the refinement for the small effect of anomalous dispersion. After three cycles of isotropic refinement using unit weights, the R value was reduced to 0.105. Inspection of the $(|F_o| - |F_c|)$ discrepancies at this point showed systematic deviations attributable to extinction for reflections with $|F_o| > 80$ and atypically large discrepancies for reflections with $h = 7, 8$ and $k \leq 4$. The validity of the measurements for this latter class of reflections was already suspect since reflections 740, 741, 820 and 821 had skewed peak profiles extending outside the scan range and had been eliminated from the data set. Tests indicated that neither stray radiation, additional crystal fragments nor instrumental failure was responsible for the discrepant observations. Although still unexplained, the 23 reflections with $h \geq 7$ and $k \leq 4$ were given zero weight in ensuing cycles of refinement as were the 7 reflections with $|F_o| > 80$. Four of the hydrogen atoms were located with confidence in a difference map following a cycle of anisotropic refinement and the other five were found in a second difference map computed after including these as constant contributors in another cycle of refinement. Hydrogen atoms were given the isotropic temperature factors of the atoms to which they were attached and positional parameters were varied in subsequent refinement. A plot of $(\Delta F)^2$ vs $|F_o|$ was used to establish the final weighting scheme: $w = (6 - |F_o|)^{-1}$, $|F_o| < 5$; $w = 1$, $5 \leq$

Table 3. Final atomic parameters†

	x	y	z	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
S	0.8527	0.3906	-0.0298	1.27	1.91	1.40	-0.06	0.13	-0.01
O(1)	0.3893	0.3127	0.1788	2.85	2.46	1.78	-0.28	0.34	0.27
O(2)	0.3257	0.4194	0.3346	2.55	2.68	1.53	0.13	0.67	-0.24
O(3)	0.8844	0.4663	-0.0547	3.16	2.35	2.48	-0.50	-0.31	-0.09
O(4)	0.8199	0.3726	0.2328	3.17	4.01	1.51	-0.69	-0.24	0.90
O(5)	0.9982	0.3501	-0.1569	1.92	4.44	3.86	1.04	0.39	-1.13
N	0.4787	0.4868	-0.0567	2.07	2.55	2.93	0.78	0.48	1.06
C(1)	0.3838	0.3802	0.1690	0.93	2.53	1.95	0.14	-0.36	0.43
C(2)	0.4536	0.4089	-0.0802	1.44	2.74	1.87	0.05	-0.08	0.64
C(3)	0.6323	0.3736	-0.1935	2.20	3.03	1.13	-0.20	-0.06	-0.18
O(<i>W</i>)	0.2624	0.2653	0.5909	4.43	2.77	3.21	0.50	1.82	0.41
H(1, C2)	0.337	0.401	-0.161						
H(2, C3)	0.623	0.320	-0.218						
H(3, C3)	0.654	0.384	-0.402						
H(4, N)	0.364	0.507	-0.031						
H(5, N)	0.525	0.500	0.050						
H(6, N)	0.574	0.507	-0.183						
H(7, O1)	0.333	0.299	0.288						
H(8, <i>W</i>)	0.308	0.224	0.585						
H(9, <i>W</i>)	0.187	0.283	0.590						
Standard deviations									
S	0.0002	0.0001	0.0003	0.06	0.05	0.05	0.04	0.05	0.04
O(1)-O(5)	0.0007	0.0002	0.0009	0.22	0.19	0.20	0.17	0.18	0.16
N, C, O(<i>W</i>)	0.0009	0.0003	0.0011	0.25	0.23	0.24	0.20	0.21	0.20
H	0.014	0.004	0.018						

† Positional parameters are in fractions of a cell edge and temperature factors are of the form $\exp[-\frac{1}{4}(h^2a^*{}^2B_{11} + k^2b^*{}^2B_{22} + l^2c^*{}^2B_{33} + 2hka^*b^*B_{12} + 2hla^*c^*B_{13} + 2klb^*c^*B_{23})]$.

Table 4. Observed and calculated structure factors

Successive columns contain k , $|F_{\text{obs}}| \times 10$, $|F_{\text{calc}}| \times 10$ and $\varphi_{\text{calc}}(\text{cycles}) \times 100$. The $|F_{\text{calc}}|$ contain a factor of 1.194 to bring them to the empirical absolute scale of the $|F_{\text{obs}}|$. Reflections given zero weight in the refinement are designated by asterisks.

$|F_o| \leq 15$; $w = (0.1|F_o| - 0.5)^{-1}$, $|F_o| > 15$; and $w=0$ for 30 reflections specified above. Two final cycles brought the refinement to convergence with a weighted R of 0.067 and a standard R of 0.050 excluding reflections of zero weight. The resulting atomic parameters are listed in Table 3 and the list of observed and calculated structure factors is given in Table 4.

Discussion

The molecular structure of L-cysteic acid is depicted in a stereo-drawing, Fig. 1, and holds no surprises. The molecule exists in the crystal as a zwitterion with the amino group extracting a hydrogen atom from the more acidic sulfonate group rather than the carboxyl group. This is evident from the disposition of hydrogen atoms and is confirmed by the dimensions of the two acid groups. Bond lengths and angles in cysteic acid (Table 5) are in close agreement with comparable dimensions in an averaged amino acid (Marsh & Donohue, 1967) and in the accurately determined structure of taurine (Okaya, 1966). Lengths agree to within two least-squares estimated standard deviations and angles compare to within three standard deviations except for a 3.8° difference in the C(1)-C(2)-C(3) angles. As in taurine the sulfonate S-O bonds are slightly non-equivalent, the shortest differing by about 4 standard deviations from the longest. Unlike taurine, the asymmetry here, if real, is not correlated with hydrogen bond involvement. It may, however, be significant that in both molecules the shortest S-O bond is to the most

remote oxygen atom and the longest bond is to the oxygen atom in close intramolecular contact with the nitrogen atom (2.837 \AA in this case). The conformation adopted by the molecule is described by the torsion angles in Table 5 and is illustrated in Fig. 1. As shown in Fig. 2, this combines the configuration usually taken by cysteine and its derivatives (Lakshminarayanan, Sasisekharan & Ramachandran, 1967; Ramachandra Ayyar, 1968; Harding & Long, 1968) with a sulfonate group oriented in the 60° position about C_x-C_β rather than at -60° as in taurine (Sutherland & Young, 1963; Okaya, 1966).

Although the hydrogen positions of the water in the crystal were not accurately established, the orientation found for the water molecule agrees well with that determined by proton magnetic resonance. The H(8, W)-H(9, W) vector makes an angle of $37 \pm 6^\circ$ with [010] and $89 \pm 6^\circ$ with [001] as compared with the respective values of $36 \pm 1^\circ$ and $87 \pm 1^\circ$ measured by El Saffar *et al.* (1969). Since the O-H bonds of the water molecule are short, owing to the displacement of electron clouds from proton positions as well as errors, the observed H-H distance of 1.40 ± 0.15 Å is considerably less than the more accurate value of 1.59 ± 0.01 Å determined by proton magnetic resonance.

As is illustrated by Fig. 3, the molecules are packed in layers parallel to (010) in the crystal. Hydrogen bonds through the amino hydrogen atoms and ionic interactions between amino and sulfonate groups hold together molecular bilayers which in turn are fixed to one another by hydrogen bonds involving the water

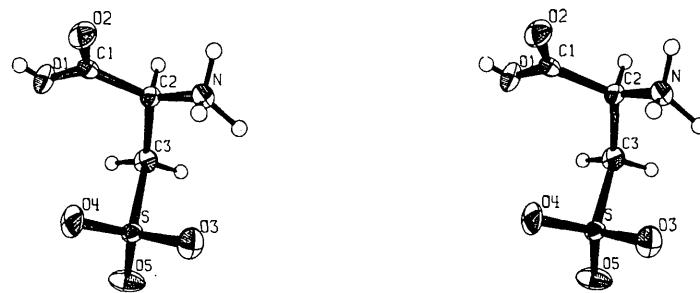


Fig. 1. Stereodrawing of the L-cysteic acid molecule. The Figure was drawn by computer with a program due to Johnson (1965).

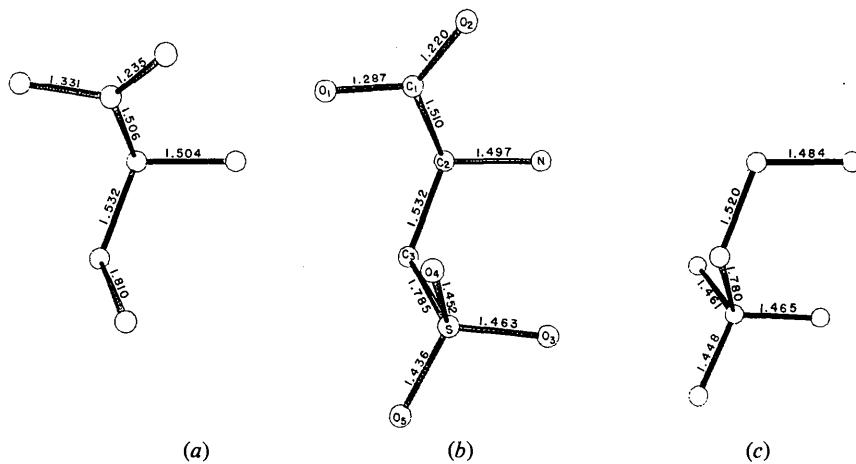


Fig. 2. Comparison of (a) L-cysteine, (b) L-cysteic acid and (c) taurine. The configuration and C–O bond lengths of L-cysteine are taken from Ramachandra Ayyar (1968); other L-cysteine bond lengths are averaged from three molecules (Ramachandra Ayyar, 1968; Harding & Long, 1968). The configurations and bond lengths of taurine and L-cysteic acid are taken respectively from Okaya (1966) and from this work. The molecules were drawn by computer (Johnson, 1965) with the N–C(2)–C(3) plane of each in the same orientation.

Table 5. Molecular dimensions

Bond lengths ($\sigma=0.005-8$)		Bond angles ($\sigma=0.3-5$)	
C(1)–O(1)	1.287 Å	C(2)–C(1)–O(1)	112.7°
C(1)–O(2)	1.220	C(2)–C(1)–O(2)	121.0
C(2)–N	1.497	O(1)–C(1)–O(2)	126.2
C(2)–C(1)	1.510	C(1)–C(2)–C(3)	116.3
C(2)–C(3)	1.532	C(1)–C(2)–N	108.8
S–C(3)	1.785	C(3)–C(2)–N	111.9
S–O(3)	1.463	C(2)–C(3)–S	114.9
S–O(4)	1.452	C(3)–S–O(3)	105.2
S–O(5)	1.436	C(3)–S–O(4)	106.9
		C(3)–S–O(5)	105.9
		O(3)–S–O(5)	110.0
		O(4)–S–O(5)	115.7
		O(5)–S–O(3)	112.3
Average lengths to hydrogen atoms		Torsion angles*	
C–H	1.03	ψ_1	190.0
N–H	0.87	ψ_2	12.6
O–H	0.73	χ_1	55.4
		χ_2	292.2

* The torsion angles follow the conventions defined by Lakshminarayanan, Sasisekharan & Ramachandran (1967) with χ_2 referring to O(3).

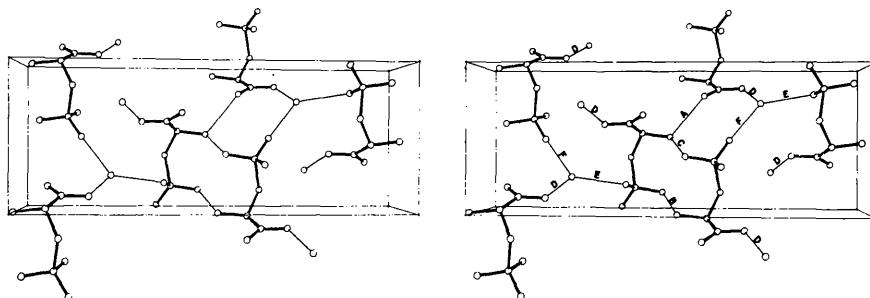


Fig. 3. Stereographic packing diagram. The unit cell is oriented with **b** across, **a** down and **c** out of the page. Hydrogen bonds are designated by the letters assigned in Table 6. The Figure was drawn by the computer program of Johnson (1965).

Table 6. *Hydrogen bond lengths*

Bond	Symmetry operation on acceptor atom			Length	$X\text{-H}\cdots\text{O}$ angle
<i>A</i> N-H \cdots O(2)	$\frac{1}{2}-x$	$1-y$	$-\frac{1}{2}+z$	2.823 Å	144.1°
<i>B</i> N-H \cdots O(3)	$1\frac{1}{2}-x$	$1-y$	$-\frac{1}{2}+z$	2.946	167.0
<i>C</i> N-H \cdots O(3')	$1\frac{1}{2}-x$	$1-y$	$\frac{1}{2}+z$	2.965	144.6
<i>D</i> O(1) \cdots H \cdots O(<i>W</i>)	<i>x</i>	<i>y</i>	<i>z</i>	2.523	164.3
<i>E</i> O(<i>W</i>) \cdots H \cdots O(4)	$-\frac{1}{2}+x$	$\frac{1}{2}-y$	$1-z$	2.813	144.8
<i>F</i> O(<i>W</i>) \cdots H \cdots O(5)	$-1+x$	<i>y</i>	$1+z$	2.783	142.2
Configuration about hydrogen donors					
(C-N). <i>A</i>	C(2) \cdots N $\cdots\cdots$ O(2)			121.5°	
(C-N). <i>B</i>	C(2) \cdots N $\cdots\cdots$ O(3)			105.3	
(C-N). <i>C</i>	C(2) \cdots N $\cdots\cdots$ O(3')			114.2	
<i>A.B</i>	O(2) \cdots N $\cdots\cdots$ O(3)			97.7	
<i>A.C</i>	O(2) \cdots N $\cdots\cdots$ O(3')			103.4	
<i>B.C</i>	O(3) \cdots N $\cdots\cdots$ O(3')			127.7	
(C-O). <i>D</i>	C(1) \cdots O(1) \cdots O(<i>W</i>)			111.2	
<i>D.E</i>	O(1) \cdots <i>W</i> \cdots O(4)			124.9	
<i>D.F</i>	O(1) \cdots <i>W</i> \cdots O(5)			116.0	
<i>E.F</i>	O(4) \cdots <i>W</i> \cdots O(5)			118.3	

molecule. All oxygen atoms and the nitrogen atom participate in hydrogen bonding (Table 6). This layer structure is no doubt responsible for the cleavage plane parallel to (010) noted by El Saffar *et al.* (1969).

References

- BRENNER, S. A. & GUM, P. H. (1968). NRL Report 6697, U.S. Naval Research Laboratory, Washington, D.C.
- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). ORFLS, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.
- EL SAFFAR, Z. M., HENDRICKSON, W. A. & KOSKI, W. S. (1969). *Acta Cryst.* **B25**, 160.
- GREENSTEIN, J. P. & WINITZ, M. (1961). *Chemistry of the Amino Acids*. London, New York: John Wiley.
- HARDING, M. M. & LONG, H. A. (1968). *Acta Cryst.* **B24**, 1096.
- International Tables for X-ray Crystallography (1962). Vol. III. Birmingham: Kynoch Press.
- JOHNSON, C. K. (1965). ORTEP, ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.
- KARLE, J. (1968). *Acta Cryst.* **B24**, 182.
- KARLE, J. & HAUPTMAN, H. (1953). *Acta Cryst.* **6**, 473.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
- LAKSHMINARAYANAN, A. V., SASISEKHARAN, V. & RAMACHANDRAN, G. N. (1967). In *Conformation of Biopolymers*, Ed. G. N. RAMACHANDRAN, Vol. 1, p. 61. London, New York: Academic Press.
- MARSH, R. E. & DONOHUE, J. (1967). *Advanc. Protein Chem.* **22**, 235.
- OKAYA, Y. (1966). *Acta Cryst.* **21**, 726.
- RAMACHANDRA AYYAR, R. (1968). *Z. Kristallogr.* **126**, 227.
- SUTHERLAND, H. H. & YOUNG, D. W. (1963). *Acta Cryst.* **16**, 897.