

Crystal Structure of Thiocytosine*

BY SVEN FURBERG† AND LYLE H. JENSEN

Department of Biological Structure, University of Washington, Seattle, Washington, U.S.A.

(Received 31 July 1969)

Crystals of thiocytosine are monoclinic, space group $C2$, with unit-cell dimensions $a=19.598$, $b=4.094$, $c=13.527$ Å, $\beta=97.54^\circ$. The crystal structure has been solved using Nb-filtered Mo radiation collected on a manual four-circle diffractometer. 3603 reflections were recorded with measurable intensities. The structure based on low-angle reflections ($\sin \theta/\lambda < 0.64$) has been compared with that derived from high-angle reflections ($\sin \theta/\lambda > 0.64$) alone. The two structures differ significantly only in the length of the double bond, C(4)–C(5), which is found to be about 0.01 Å longer in the high-angle structure. This difference may be due to the effect of the valence electrons in the double bond. The final R indices were 0.018 for the low-angle reflections, 0.030 for the high-angle ones, and 0.025 for all reflections. The standard deviations in bond lengths involving nonhydrogen atoms lie in the range 0.0016–0.0024 Å for the low-angle structure and 0.0010–0.0019 Å for the high-angle one. Two crystallographically non-equivalent molecules are linked in pairs in the crystal by N–H...N and N–H...S bonds. Very small, but probably significant, differences are observed between the two crystallographically independent molecules.

As a part of a program of investigation of the structure of nucleic acid components and related compounds we have determined the structure of thiocytosine. Thiopyrimidines are of considerable biological interest and have been observed to be present in soluble ribonucleic acid. The structures of two thiopyrimidines, 2-thiouracil and 2,4-dithiouracil, have been reported (Tsernoglou, 1966; Shefter & Mautner, 1967), and the structure of cytosine is well known from analyses on cytosine (Baker & Marsh, 1964), cytosine monohydrate (Jeffrey & Kinoshita, 1963) and cytidine (Furberg, Petersen & Rømming, 1965).

Thiocytosine forms high quality crystals which are not noticeably damaged by X-rays and in which thermal motion is moderate. We have, therefore, used this opportunity to study the influence on the structure analysis of including high-angle reflections (Stewart & Jensen, 1967).

Experimental

A sample of thiocytosine supplied by Sigma Chemical Co. contained nicely developed prismatic crystals elongated along b , with (001) as the predominant face. Weissenberg photographs showed the crystals to have a monoclinic, C -centered lattice, the only systematic absences occurring for $h+k=2n+1$. The space group is $C2$, the other two possibilities, Cm and $C2/m$, being ruled out by space considerations and by the intensities of the 020 and 040 reflections. There are eight molecules in the unit cell (calculated density 1.570 g.cm⁻³) and thus two molecules in the asymmetric unit.

The unit-cell dimensions were determined from the 2θ values for 15 reflections measured on a four-circle

diffractometer. Both $+2\theta$ and -2θ were measured for each reflection. The values were read directly to 0.01° and an estimate was made of the third decimal place. The unit-cell parameters derived from a least-squares fit to the data are as follows: $a=19.598 \pm 0.0017$, $b=4.0943 \pm 0.0003$, $c=13.527 \pm 0.0007$ Å; $\beta=97.542 \pm 0.014^\circ$, based on Mo $K\alpha=0.71069$ Å. The uncertainties are estimated standard deviations derived from the inverse matrix of the refinement.

All intensity measurements were carried out on a single crystal of dimensions $0.27 \times 0.33 \times 0.37$ mm. The corners of the crystal were rounded on a 'liquid lathe' (Jensen & Mabis, 1966) with water as solvent. The size of the crystal is well within the intensity plateau of the beam (focal spot width, 0.8 mm; take-off angle, 3°; 1.0 mm aperture). Measurements were made by the $\omega/2\theta$ scan technique at a rate of 2° minute⁻¹. Scan times were variable as computed from the tangent relation of Alexander & Smith (1964), and backgrounds were measured for half the scan time at each limit of the scan range. All 2648 independent reflections with 2θ less than 70° were first measured; of these, only 132 had intensities less than a 2σ cut-off limit. Later the measurements were extended to $2\theta=90^\circ$ ($\sin \theta/\lambda=1.0$) and 1090 reflections were recorded with measurable intensities, using for these high-angle data a cut-off limit of 4σ and measuring each reflection twice.

Three reflections were chosen as standards and were measured at the beginning and end of each four-hour shift. A scale factor based on the mean value of the standard reflections bracketing each shift was applied to the data for that shift and served to put all data on a common scale. The measurements of the strongest reflections were corrected for coincidence losses which were determined both by use of filters and also by remeasuring the reflections with reduced current on the X-ray tube. No corrections were made for

* Presented at the Tucson A.C.A. meeting, February, 1968.

† Permanent address: Department of Chemistry, University of Oslo, Blindern, Norway.

Table 1. Observed and calculated structure factors
Each group of three columns contains respectively h, 10|F_o|, and 10|F_c|.

Table with multiple columns of numerical data, organized into groups labeled H,0,0 through H,2,27. Each group contains three columns: h, 10|F_o|, and 10|F_c|.

Table 1 (cont.)

H	h	k	l	Intensity	Classification
H 3,16	16	0	0	30	Observed
H 3,15	15	0	0	30	Observed
H 3,14	14	0	0	30	Observed
H 3,13	13	0	0	30	Observed
H 3,12	12	0	0	30	Observed
H 3,11	11	0	0	30	Observed
H 3,10	10	0	0	30	Observed
H 3,9	9	0	0	30	Observed
H 3,8	8	0	0	30	Observed
H 3,7	7	0	0	30	Observed
H 3,6	6	0	0	30	Observed
H 3,5	5	0	0	30	Observed
H 3,4	4	0	0	30	Observed
H 3,3	3	0	0	30	Observed
H 3,2	2	0	0	30	Observed
H 3,1	1	0	0	30	Observed
H 3,0	0	0	0	30	Observed
H 4,16	16	1	0	30	Observed
H 4,15	15	1	0	30	Observed
H 4,14	14	1	0	30	Observed
H 4,13	13	1	0	30	Observed
H 4,12	12	1	0	30	Observed
H 4,11	11	1	0	30	Observed
H 4,10	10	1	0	30	Observed
H 4,9	9	1	0	30	Observed
H 4,8	8	1	0	30	Observed
H 4,7	7	1	0	30	Observed
H 4,6	6	1	0	30	Observed
H 4,5	5	1	0	30	Observed
H 4,4	4	1	0	30	Observed
H 4,3	3	1	0	30	Observed
H 4,2	2	1	0	30	Observed
H 4,1	1	1	0	30	Observed
H 4,0	0	1	0	30	Observed
H 5,16	16	0	1	30	Observed
H 5,15	15	0	1	30	Observed
H 5,14	14	0	1	30	Observed
H 5,13	13	0	1	30	Observed
H 5,12	12	0	1	30	Observed
H 5,11	11	0	1	30	Observed
H 5,10	10	0	1	30	Observed
H 5,9	9	0	1	30	Observed
H 5,8	8	0	1	30	Observed
H 5,7	7	0	1	30	Observed
H 5,6	6	0	1	30	Observed
H 5,5	5	0	1	30	Observed
H 5,4	4	0	1	30	Observed
H 5,3	3	0	1	30	Observed
H 5,2	2	0	1	30	Observed
H 5,1	1	0	1	30	Observed
H 5,0	0	0	1	30	Observed
H 6,16	16	0	0	30	Observed
H 6,15	15	0	0	30	Observed
H 6,14	14	0	0	30	Observed
H 6,13	13	0	0	30	Observed
H 6,12	12	0	0	30	Observed
H 6,11	11	0	0	30	Observed
H 6,10	10	0	0	30	Observed
H 6,9	9	0	0	30	Observed
H 6,8	8	0	0	30	Observed
H 6,7	7	0	0	30	Observed
H 6,6	6	0	0	30	Observed
H 6,5	5	0	0	30	Observed
H 6,4	4	0	0	30	Observed
H 6,3	3	0	0	30	Observed
H 6,2	2	0	0	30	Observed
H 6,1	1	0	0	30	Observed
H 6,0	0	0	0	30	Observed

absorption since intensity errors from this source are less than 1% in the present case ($\mu = 6.2 \text{ cm}^{-1}$). Secondary extinction was subsequently corrected for by comparing observed and calculated structure factors.

Structure analysis

In the structure analysis, the reflections were divided into three sets. Set I consists of 1347 reflections with $\sin \theta / \lambda < 0.64$ of which 32 were coded 'unobserved', *i.e.* intensities $< 2\sigma$. Set I comprises the reflections within the limit of the Cu sphere. Set II comprises 1298 reflections in the range $0.64 < \sin \theta / \lambda < 0.81$ ($54^\circ < 2\theta < 70^\circ$) of which 100 were coded 'unobserved'. Finally, set III

comprises the 1090 high-angle reflections in the region $0.81 < \sin \theta / \lambda < 1.0$ recorded as described above.

The structure was solved and the preliminary refinement carried out using data sets I and II. The Patterson synthesis showed clearly the positions of the two sulfur atoms. Positions of the carbon and nitrogen atoms were derived from the electron density projection on (010) and approximately known bond lengths and angles. The postulated structure had an *R* index of 0.25, which was reduced to 0.027 by full-matrix least-squares refinement of the nonhydrogen atoms, the hydrogen atoms being kept in fixed positions derived from a ΔF synthesis. Secondary extinction was then approximately corrected for by multiplying F_o by $1 + kI_o$

Table 2. Parameters for nonhydrogen atoms in molecules A and B from refinement of data set II + III

Molecule	x	y	z	Standard deviations are given in parentheses.					
				B ₁₁	B ₂₂	B ₃₃	B ₁₂	B ₁₃	B ₂₃
C(2)	0.08468 (4)	0.17746 (32)	0.24076 (6)	1.62 (2)	2.28 (3)	1.68 (2)	-0.04 (2)	0.21 (2)	0.07 (2)
B	0.33934 (4)	0.31605 (31)	0.25406 (6)	1.58 (2)	2.29 (3)	1.77 (2)	-0.02 (2)	0.30 (1)	0.20 (2)
C(4)	0.13406 (4)	0.46321 (35)	0.37997 (6)	1.74 (2)	2.66 (2)	1.71 (2)	-0.08 (2)	0.22 (1)	-0.08 (2)
B	0.38563 (4)	0.09017 (31)	0.12068 (7)	1.59 (2)	2.53 (3)	1.75 (2)	0.05 (2)	0.24 (1)	0.10 (2)
C(5)	0.06918 (5)	0.50292 (54)	0.41522 (7)	1.94 (2)	3.78 (5)	2.12 (2)	-0.26 (2)	0.55 (2)	-0.51 (4)
B	0.32144 (5)	-0.05569 (41)	0.08274 (9)	1.77 (2)	3.33 (5)	2.32 (3)	-0.10 (2)	0.14 (2)	-0.50 (3)
A	0.01419 (5)	0.36251 (51)	0.36033 (9)	1.79 (2)	4.18 (6)	2.31 (3)	-0.39 (3)	0.65 (2)	-0.54 (3)
B	0.26788 (4)	0.00219 (47)	0.13415 (8)	1.56 (2)	3.04 (4)	2.47 (2)	-0.20 (3)	0.05 (1)	-0.26 (3)
A	0.02227 (4)	0.19802 (37)	0.27554 (7)	1.62 (2)	3.31 (4)	2.03 (2)	-0.26 (2)	0.31 (1)	-0.33 (2)
N(1)	0.27659 (4)	0.18902 (33)	0.21752 (7)	1.50 (2)	2.76 (4)	2.20 (2)	-0.11 (2)	0.33 (2)	-0.06 (2)
B	0.14072 (4)	0.30492 (31)	0.29466 (6)	1.57 (2)	2.61 (3)	1.79 (2)	-0.06 (2)	0.16 (1)	-0.13 (2)
N(3)	0.39347 (4)	0.27277 (31)	0.20403 (6)	1.56 (2)	2.66 (3)	1.79 (2)	-0.07 (2)	0.33 (1)	0.00 (1)
B	0.19045 (5)	0.58513 (48)	0.43318 (9)	1.89 (2)	4.34 (6)	2.43 (4)	-0.34 (3)	0.11 (2)	-0.99 (3)
N(4)	0.44028 (5)	0.04278 (45)	0.07343 (7)	1.95 (2)	3.88 (5)	2.17 (2)	-0.12 (3)	0.68 (2)	-0.43 (2)
B	0.08729 (1)	-0.01271 (0)	0.12951 (2)	1.91 (1)	2.89 (1)	2.06 (1)	-0.23 (1)	0.38 (0)	-0.45 (1)
S(2)	0.34745 (1)	0.52968 (13)	0.36290 (2)	2.10 (1)	2.95 (1)	2.08 (1)	-0.43 (1)	0.10 (0)	-0.31 (1)

for the reflections (about 20) with I_o greater than 30,000 dekaaccounts. For all other reflections the effect was considered negligible (less than 0.5% in F) and the angular dependence (Zachariasen, 1965) is also negligible.

The hydrogen positions were next refined, first with fixed thermal parameters equal to those of the heavy atoms to which they are attached, then with a variable isotropic thermal parameter. Both procedures gave essentially the same hydrogen positions and the same R index.

Least-squares refinements were carried out for the following groups of reflections: (1) set I, (2) set II, (3) set III, (4) set I + II, (5) set II + III, (6) set I + II + III. In Table 1 are listed the structure factors, F_o and F_c , where the F_c are based on parameters from the refinement for data set I + II + III.

In Table 2 the parameters for nonhydrogen atoms resulting from the refinement of data set II + III are tabulated. Although these results are slightly less precise than those from refining the more extensive set I + II + III, we regard them as more accurate for the reason discussed below.

In Table 3 the parameters for the hydrogen atoms are listed. These were determined from data set I + II since the contributions of hydrogen to the high-angle reflections of set III are negligible.

All refinements were carried out by the least-squares method using the full matrix of the normal equations. The weighting scheme applied was based on the assumption that the standard deviation, $\sigma(I)$, of the relative intensities is given by

$$\sigma(I) = [C_T + (0.01 C_N)^2]^{1/2},$$

where C_T is the total number of counts (peak plus backgrounds) and C_N is the net count (peak minus backgrounds).

The scattering factors used were those of Dawson (1960) for sulfur, of Berghuis, Haanappel, Potters, Loopstra, MacGillavry & Veenendaal, (1955) for carbon and nitrogen and of Stewart, Davidson & Simpson (1965) for hydrogen.

Discussion

Accuracy

Table 4 summarizes some data and results from the refinements. Listed for each of the six data sets are the number of observed and 'unobserved' reflections, the scale factor, the conventional residual $R = \sum |F_o| - |F_c| / \sum |F_o|$, the weighted residual $R_w = \sum w |F_o| - |F_c| / \sum w |F_o|$ and the standard deviation of an observation of unit weight, $\sum w \Delta F^2 / (m - n)$, where m is the number of observations and n is the number of parameters, sometimes also termed 'goodness of fit'. In set III the contribution of the hydrogen atoms was neglected.

The value of R for data set I, the data to the limit of Cu $K\alpha$ radiation, is remarkably low, 0.018. This is due, in part, to the size of the crystal used, which was large enough to insure relatively low counting error, and to

the use of Mo $K\alpha$ radiation. Moreover, the presence of a heavy atom such as S and the fact that the structure is noncentrosymmetric (part of the error presumably being absorbed in the phases) also tend toward lower R values.

It will be seen that for data sets I, II and III, R values increase from one set to the next, *i.e.* R increases with increasing $\sin \theta/\lambda$. This stems from the fact that reflections become progressively weaker and counting errors relatively greater with increasing $\sin \theta/\lambda$.

In general, it is to be expected that R_w would be less than R since in refinement it is $\sum w|\Delta F|^2$ that is minimized. Indeed, this is found to be true in Table 4 for any data set not including I and the converse for any set including I. This suggests that $w=1/\sigma^2$ does not properly weight the ΔF values for the low-angle data.

A similar effect can be seen for the 'goodness of fit', a quantity which should be unity if the weighting scheme is correct, if no systematic errors are present and if the model is a proper one. In Table 4 this is found to be the case, within the limit of experimental error, for any data set not including I. But for any set including I, the values greatly exceed unity. This again suggests that $w=1/\sigma^2$ does not properly weight the low-angle ΔF values. Nevertheless, the same weighting function was used for all data sets, and it appears to weight properly sets not including I.

These observations indicating that the weighting function was unsatisfactory for low-angle reflections can be explained in terms of deficiencies in the model used to calculate structure factors. Since the quantity $\sum w(|F_o| - |F_c|)^2$ occurs in the expression for the goodness of fit, it is evident that its value depends not only on w and F_o but also on F_c , *i.e.* on the model.

Spherical atoms have been assumed and these have been modified by anisotropic thermal parameters. But actual bonded atoms have aspherical electron clouds and neglecting this by using spherical symmetrical scattering factors will affect primarily the low-order reflections. This is evident, for example, in the scattering factors for C and C^{4+} which are essentially equal for $0.5 < \sin \theta/\lambda$. It follows, then, that a model based on spherical atoms is not proper for low-angle data but is more nearly so for high-angle data.

If the model is implicated in the trend of the 'goodness of fit' as suggested, then it should follow that structures containing spherical ions or heavy atoms should give values more nearly equal to unity. The value of 1.7 found for arginine hydrochloride monohydrate (Dow, Jensen, Mazumdar, Srinivasan & Ramachandran, 1970) with an appreciable fraction of the electrons associated with the chloride ion supports this view.

In the case of uracil (Stewart & Jensen, 1967), a structure without a heavy atom, a value of 2.34 was found for the 'goodness of fit'. Since appreciable high-angle data were included in the uracil refinement, this value is to be compared with the 2.13 for the thiocytosine data set I+II. As expected, the compound containing a heavy atom has the lower value.

In comparing values for the 'goodness of fit', it should be noted however, that this quantity is sensitive to other effects such as crystal size (Hope, 1969), the radiation used and the size of the unit cell. Indeed, any factor which causes random errors to dominate will tend to yield values more nearly equal to unity, provided, of course, that the weighting function is a proper one.

Table 3. Parameters for hydrogen atoms in molecules A and B from refinement of data set I+II

Standard deviations are given in parentheses.

	Molecule	x	y	z	B
H(1)	A	-0.0110 (9)	0.1180 (48)	0.2475 (11)	3.0 (4)
	B	0.2399 (9)	0.2196 (49)	0.2448 (11)	2.9 (4)
H(4)	A	0.2346 (8)	0.5800 (51)	0.4132 (11)	3.1 (3)
	B	0.4826 (8)	0.1610 (49)	0.0917 (11)	3.5 (3)
H(4')	A	0.1871 (8)	0.6943 (49)	0.4904 (12)	3.7 (4)
	B	0.4345 (8)	-0.0776 (49)	0.0173 (11)	3.1 (3)
H(5)	A	0.0651 (8)	0.6173 (49)	0.4729 (12)	3.4 (3)
	B	0.3160 (8)	-0.1919 (50)	0.0261 (12)	3.9 (3)
H(6)	A	-0.0328 (8)	0.3743 (47)	0.3756 (11)	3.2 (3)
	B	0.2225 (8)	-0.0916 (49)	0.1201 (10)	3.0 (3)

Table 4. Number of observed and unobserved reflections and summary data of refinements

Data set	I	II	III	I+II	II+III	I+II+III
No. observed reflections	1315	1198	1090	2513	2288	3603
No. 'unobserved*' reflections	32	100	†	132	100	132
Scale factor	0.994	1.007	1.048	0.997	1.018	1.014
R	0.018	0.024	0.036	0.023	0.030	0.025
R_w	0.023	0.021	0.033	0.026	0.026	0.028
Goodness of fit	2.52	1.02	1.09	2.13	1.05	1.92

* Reflections less than 2σ recorded as 'unobserved'.

† Only reflections greater than 4σ were recorded for data set III.

By altering the weighting scheme, it is possible to normalize the 'goodness of fit' for the various ranges of $\sin \theta/\lambda$. The weighting function we have used, however, was chosen with reference to the reproducibility of the standard reflections and should account satisfactorily for *random* errors. To alter the weighting so as to normalize the 'goodness of fit' could decrease, but not eliminate, the effects of model deficiency, but it would tend to obscure the very effects whose cause we should seek to eliminate.

In general, bond lengths of corresponding bonds in molecules *A* and *B* as listed in Table 5 agree well, the largest difference being 0.007 Å for the C(2)–N(3) bond for data set I. If only random errors operate, the quantity (root-mean-square δ_{A-B})/ $\sqrt{2}$ should be comparable with r.m.s. σ . It is found to be somewhat less than r.m.s. σ for data sets not including I and somewhat more for those including I, *i.e.* agreement between corresponding bonds in the two independent molecules is somewhat better than expected for high-angle data and conversely for any set including low-angle data.

The high-angle structure (set II+III) is appreciably better than any of the others in the sense that the r.m.s.

of the differences between corresponding bonds in the two independent molecules is smallest. On the other hand, the r.m.s. σ for the high-angle structure is slightly higher than that for the refinement of the full set of data. The difference is probably not significant, however, in spite of the fact that the full set of data contains 1315 more observed reflections than the high-angle set. The additional data should have led to an increase of precision of about 20%. The fact that the precision increased only slightly is consistent with the deficiency of the model in the low-angle region. The scattering factors are quite accurate, however, in the high-angle region where the contribution of the valence electrons for the C, N and S atoms are negligible. It is for this reason that we regard the nonhydrogen atom parameters derived from data set II+III as most accurate and have tabulated final parameters for these atoms in Table 2 only for this set.

The bond lengths derived from refinements of different sets of data agree within the limits of error, (Table 5), apart from the double bond, C(5)–C(6), which is somewhat shorter in the low-angle structure (set I) than in the high-angle structure (sets II+III, and

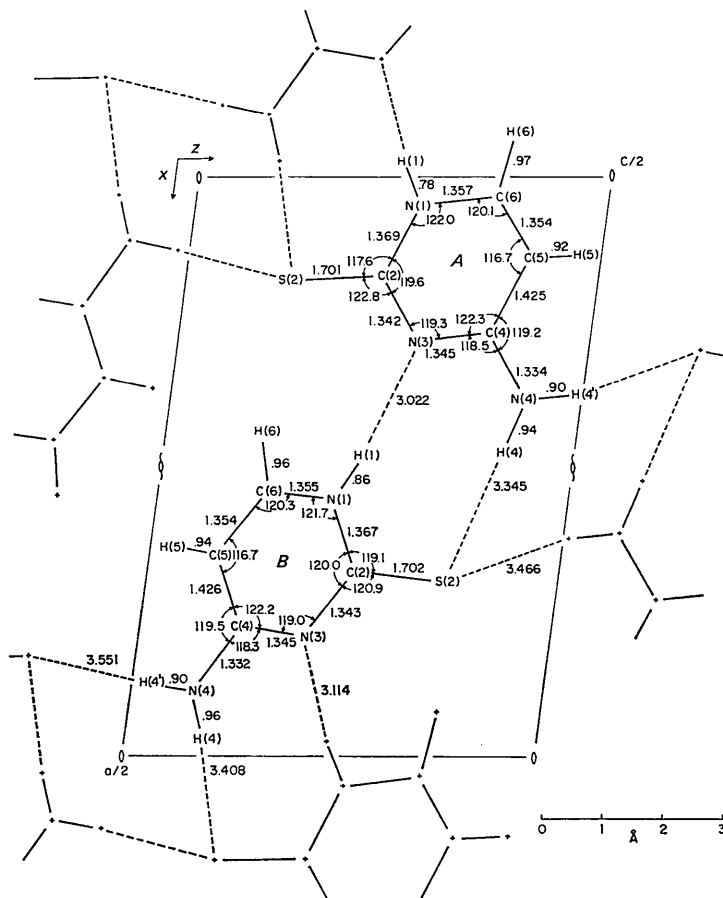


Fig. 1. Structure as viewed along [010] with bond lengths and angles designated. Hydrogen bond lengths are distances between donor and acceptor atoms.

also set II and set III). The difference is 0.015 Å for molecule *A* and 0.011 Å for molecule *B*, more than five times the estimated standard deviations. This possibly stems from the fact that the bond C(5)–C(6) has more double-bond character than the other bonds in the molecule, the bonding electrons tending to shift slightly the center of gravity of the electron clouds of atoms C(5) and C(6) in the observed direction. However, a similar effect is not observed for the C–S and C–N bonds, which also have double-bond character, although not to the same extent as C(5)–C(6).

The final difference map based on data set I gives peaks of heights 0.1 to 0.2 e.Å⁻³ in bonds between nonhydrogen atoms, whereas that based on set II+III gives no peaks greater than about 0.04 e.Å⁻³. This again shows the influence of bonding electrons.

The e.s.d.'s lie in the range 0.0010 to 0.0019 Å in the bond lengths and 0.07 to 0.12° in the bond angles for bonds between nonhydrogen atoms. For bonds involving hydrogen atoms, the corresponding ranges are 0.016–0.018 Å and 1.0–1.4°. No corrections for thermal effects were made.

Molecular structure

The bond lengths and bond angles calculated from the

coordinates of Tables 2 and 3 are given in Fig. 1 and in the column headed II+III of Table 5.

The two crystallographically independent molecules have very nearly the same structure in the crystal. Thus the largest difference in the length of corresponding bonds between nonhydrogen atoms is found to be only 0.002 Å, indicating identity of the bonds at this level of precision. Most of the angles in *A* and *B* are also closely similar, the only significant difference being in the angles at C(2), the angle S(2)–C(2)–N(3) being 0.9° larger and S(2)–C(2)–N(1) 1.5° smaller in molecule *A* than in molecule *B*. This is probably due to intermolecular forces, since there is a short contact of 2.69 Å between the sulfur atom in molecule *A* and hydrogen atom H(6) in molecule *B* (Fig. 1). The corresponding distance between the sulfur atom in molecule *B* and H(6) in molecule *A* is considerably longer.

The bonds to the hydrogen atoms are essentially equal in length in the two molecules except N(1)–H(1), which is found to be 0.78 Å in molecule *A* and 0.86 Å in molecule *B*. This difference, as well as the difference of about 0.05 Å between the two N–H bonds in each amino group, is probably significant, and related to the different strengths of the hydrogen bonds in which they take part. As commonly found in X-ray analysis,

Table 5. Bond lengths involving nonhydrogen atoms for molecules *A* and *B*

Results for each of six data sets. Bond lengths in Å units, standard deviations in parentheses in units of 10⁻⁴ Å.

Molecule		Data set					I+II+III
		I	I+II	II	II+III	III	
N(1)–C(2)	<i>A</i>	1.372 (22)	1.372 (16)	1.369 (18)	1.369 (12)	1.368 (23)	1.372 (13)
	<i>B</i>	1.373 (21)	1.372 (15)	1.367 (17)	1.367 (12)	1.365 (23)	1.371 (13)
C(2)–N(3)	<i>A</i>	1.340 (20)	1.341 (15)	1.342 (16)	1.342 (12)	1.348 (23)	1.342 (13)
	<i>B</i>	1.347 (22)	1.347 (16)	1.343 (17)	1.343 (12)	1.344 (21)	1.347 (13)
N(3)–C(4)	<i>A</i>	1.349 (21)	1.348 (16)	1.345 (18)	1.345 (14)	1.344 (27)	1.347 (13)
	<i>B</i>	1.347 (21)	1.346 (16)	1.347 (18)	1.345 (14)	1.342 (28)	1.345 (14)
C(4)–C(5)	<i>A</i>	1.424 (22)	1.423 (16)	1.424 (19)	1.425 (13)	1.424 (27)	1.423 (14)
	<i>B</i>	1.420 (22)	1.421 (16)	1.424 (19)	1.426 (14)	1.428 (29)	1.421 (13)
C(5)–C(6)	<i>A</i>	1.339 (24)	1.343 (18)	1.353 (23)	1.354 (18)	1.357 (36)	1.343 (15)
	<i>B</i>	1.343 (23)	1.347 (17)	1.352 (21)	1.354 (15)	1.355 (28)	1.347 (14)
C(6)–N(1)	<i>A</i>	1.351 (24)	1.353 (18)	1.359 (24)	1.357 (19)	1.350 (38)	1.353 (15)
	<i>B</i>	1.353 (24)	1.353 (18)	1.357 (22)	1.355 (18)	1.354 (36)	1.353 (15)
S(2)–C(2)	<i>A</i>	1.702 (16)	1.700 (12)	1.700 (13)	1.701 (10)	1.701 (20)	1.700 (10)
	<i>B</i>	1.697 (17)	1.697 (12)	1.700 (13)	1.702 (10)	1.704 (21)	1.697 (11)
C(4)–N(4)	<i>A</i>	1.332 (21)	1.335 (16)	1.337 (20)	1.334 (15)	1.333 (31)	1.335 (14)
	<i>B</i>	1.332 (21)	1.333 (15)	1.332 (18)	1.332 (13)	1.331 (26)	1.333 (12)

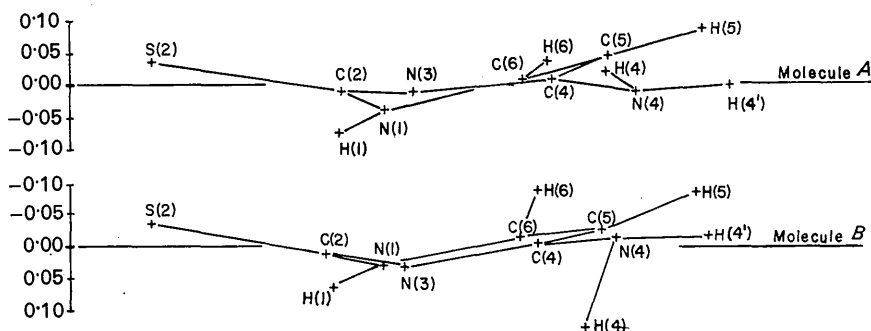


Fig. 2. Molecules *A* and *B* viewed along the least-squares plane for each molecule.

the bonds to hydrogen atoms are roughly 0.1 Å shorter than the internuclear distances. The effect is especially large for the N(1)–H(1) bond, indicating a greater displacement of the center of gravity of the bonding electrons in this bond than in the N–H bonds of the amino group. The same trend is found in a number of other recent crystal structure analysis.

Both molecules are nearly planar, the distances from the nonhydrogen atoms to the least-squares plane through the six ring atoms lying in the range up to 0.043 Å (Table 6). The largest of these deviations are, however, significant. Furthermore, the two molecules deviate from a planar structure in essentially the same manner, as is evident from Fig. 2. Also the hydrogen

atoms lie nearly in the least-squares plane, the largest deviation being 0.13 Å [atom H(4) of molecule *B*]. The amino group of molecule *A* is planar and coplanar with the pyrimidine ring, whereas that of molecule *B* is slightly distorted.

A comparison of thiocytosine with cytosine reveals significant differences only in the region of atom C(2) to which the sulfur atom is attached. The C–S bond is 1.701 Å and has apparently less double-bond character than the C–O bond in cytosine, which is about 1.24 Å. For adjacent ring bonds C(2)–N(3) and N(1)–C(2), the values 1.374 and 1.364 Å were found in cytosine, slightly larger than the thiocytosine values of 1.368 and 1.342 Å respectively. These changes may be interpreted in terms of flow of bonding electrons from the neighboring C(2)–N bonds [especially N(1)–C(2)] to the external double bond on substitution of the sulfur atom by oxygen.

The bond C(5)–C(6) is found to be 1.354 Å in both molecules *A* and *B* slightly but significantly longer than in the low-angle structure of this and other pyrimidine derivatives, for which values very close to a normal double bond (1.34 Å) are reported. From a theoretical point of view, the bond should be expected to have less double-bond character, in agreement with the result of the analysis of the high-angle data.

Comparison of thermal parameters for corresponding atoms in the two molecules, *A* and *B* (Table 2) shows that the patterns of thermal motion in the molecules are remarkably similar. It is also evident that the largest component of motion is approximately normal to the plane of the ring and is appreciably greater for atoms N(4), C(5) and C(6), particularly in molecule *A*. The thermal ellipsoids are plotted in Fig. 3.

Molecular packing

The arrangement of the molecules in the crystal is shown in Fig. 1. The basic unit is a dimer of the two crystallographically independent molecules, which are held together in a very nearly planar arrangement by the two strongest hydrogen bonds of their kind in the structure, N(1')–H(1')...N(3) of 3.022 Å and N(4)–H(4)...S(2') of 3.345 Å. An additional short contact H(6')...S(2) of 2.69 Å between these two molecules probably represents a repulsion, the angles at C(2) being slightly distorted, as described above. (The normal van der Waals distance between sulfur and hydrogen atoms is probably about 2.9 Å). The dimers are linked to neighboring dimers by weaker hydrogen bonds. In the *x* direction, there are bonds N(1)–H(1)...N(3') of 3.114 Å and N(4')–H(4')...S(2) of 3.408 Å, analogous to those in the dimer but weaker. Actually there are infinite, roughly planar ribbons of doubly linked molecules extending in the *x* direction. The ribbon plane forms an angle of about 30° with the *xz* plane. Neighboring ribbons are linked together by even weaker N–H...S bonds of lengths 3.466 and 3.551 Å. All the N–H...S bonds take part in bonding molecules in spirals up twofold axes and twofold screw

Table 6. Deviation of atoms from least-squares planes through nonhydrogen atoms in molecules *A* and *B*

	Molecule <i>A</i> deviation from l.s. plane	Molecule <i>B</i> deviation from l.s. plane
N(1)	–0.043 Å	0.029 Å
C(2)	–0.011	0.011
N(3)	–0.014	0.033
C(4)	0.006	0.004
C(5)	0.041	–0.023
C(6)	0.003	–0.009
S(2)	0.032	–0.034
N(4)	–0.014	–0.010
H(1)	–0.076	0.063
H(4)	0.016	0.124
H(4')	–0.004	–0.015
H(5)	0.084	–0.081
H(6)	0.034	–0.089

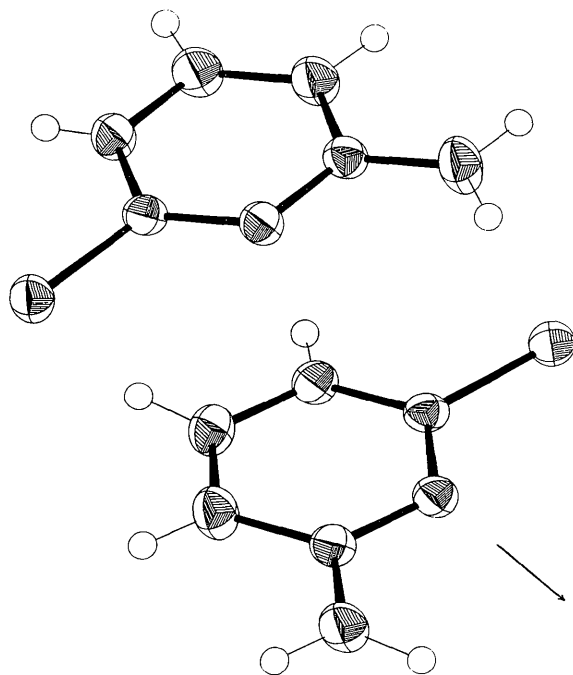


Fig. 3. Thermal ellipsoids plotted at the 50% level as viewed along [150] (indicated by arrow).

axes parallel to **b**. The arrangement is essentially the same as that found in crystals of cytosine, but more complicated due to the existence of two types of molecules *A* and *B*. The hierarchy of structure in crystals of thiocytosine is interesting: molecules–dimers–ribbons–bundles or ribbons, each step being associated with progressively weaker bonds.

The hydrogen bonds conform to the general stereochemistry of the molecule, being nearly linear, lying approximately in molecular planes, and forming angles of about 120° with adjacent C–N bonds. For example the angle between the two N–H...S bonds of the amino group is 123·5° in molecule *A* and 124·3° in *B*. In general, the shortest hydrogen bonds are associated with the longest N–H distances, as is to be expected. The shortest intermolecular distance between hydrogen atoms is 2·46 Å (amino hydrogen atoms) and between nitrogen and sulfur atoms, 3·28 Å. The latter contact is, however, not a hydrogen bond, but rather a van der Waals contact, since there is no hydrogen atom between these two atoms.

Support under Grant AM 3288 from the National Institutes of Health is gratefully acknowledged. We thank Dr J. M. Stewart for the *X-ray* 63 *Crystal Structure*

Calculating System which was used for most of the computational work in this study and to Dr C. K. Johnson for the *ORTEP* program.

References

- ALEXANDER, L. E. & SMITH, G. S. (1964). *Acta Cryst.* **17**, 1195.
 BARKER, D. L. & MARSH, R. (1964). *Acta Cryst.* **17**, 1581.
 BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.
 DAWSON, B. (1960). *Acta Cryst.* **13**, 403.
 DOW, J., JENSEN, L. H., MAZUMDAR, S. K., SRINIVASAN, R. & RAMACHANDRAN, G. N. (1969). To be published.
 FURBERG, S., PETERSEN, C. S. & RØMMING, C. (1965). *Acta Cryst.* **18**, 313.
 HOPE, H. (1969). Private communication.
 JEFFREY, G. A. & KINOSHITA, Y. (1963). *Acta Cryst.* **16**, 20.
 JENSEN, L. H. & MABIS, A. J. (1966). *Acta Cryst.* **21**, 770.
 SHEFTER, E. & MAUTNER, H. G. (1967). *J. Amer. Chem. Soc.* **89**, 1249.
 STEWART, R. F., DAVIDSON, E. & SIMPSON, W. (1965). *J. Chem. Phys.* **42**, 3175.
 STEWART, R. F. & JENSEN, L. H. (1967). *Acta Cryst.* **23**, 1102.
 TSEBNOGLOU, D. (1966). Thesis, Yale Univ.
 ZACHARIASEN, W. H. (1965). *Acta Cryst.* **18**, 710.

Acta Cryst. (1970). B26, 1268

The Crystal Structure of 2'-Hydroxymethanesulfonanilide

BY HAROLD P. KLUG

Mellon Institute, Carnegie-Mellon University, Pittsburgh, Pa. 15213, U.S.A.

(Received 19 June 1969)

The simple compound, 2'-hydroxymethanesulfonanilide (HMSA), HO–C₆H₄–NH–SO₂CH₃, is related to a series of new compounds possessing blood-pressure-control activity. This biological activity is possibly related to the hydrogen atom of the phenyl –N–H portion of these compounds. Furthermore, the infrared study of this compound had revealed an interesting hydrogen bonding pattern in the crystal. A detailed study of HMSA has been carried out to determine the stereochemistry around the nitrogen atom and to verify the hydrogen bonding. The crystals of HMSA are orthorhombic, space group *P*2₁2₁2₁, with *a* = 5·5913 ± 0·004, *b* = 10·8101 ± 0·0009, *c* = 13·6305 ± 0·0012 Å. The unit cell contains four molecules. The structure was solved using three-dimensional intensity data obtained with an automatic single-crystal diffractometer. Phasing was done by the Hall–Maslen method, and refinement was by least-squares methods with anisotropic temperature factors. The hydrogen atoms were revealed by a three-dimensional ΔF difference synthesis. HMSA molecules through hydrogen bonding form infinite chains parallel to the *a* axis. Each HMSA molecule also has an intramolecular hydrogen bond between the OH group and its nitrogen atom. The amine hydrogen atom sticks out alone on one side of the benzene ring while the whole methylsulfone group is on the opposite side. The thermal motions of the molecule have been analyzed.

Introduction

The incorporation of the alkylsulfonamido group into the benzene ring of phenethanolamines unexpectedly leads to a series of compounds with interesting biological activity (Larsen & Lish, 1964). Many of these new compounds are strikingly similar in action to cer-

tain well-known phenolic phenethanolamines, e.g. phenylephrine and epinephrine. As an approach to an understanding of their biological activity a study of the crystal structure and stereochemistry of methanesulfonanilide (MSA), C₆H₅–NH–SO₂CH₃, was made (Klug, 1968). The related compound, 2'-hydroxymethanesulfonanilide (HMSA), HO–C₆H₄–NH–SO₂CH₃, also had