

Crystal Structure of 5,6-Dihydro-1-methyl-4-thiouracil

By M. J. E. Hewlins, Department of Chemistry, University College, Cardiff

The X-ray crystal structure of the title compound was solved by the heavy-atom method from diffractometer data and refined by least-squares methods to R 0.107 for 1476 reflections. The crystals are monoclinic, space-group $P2_1/c$, $Z = 8$, with $a = 8.086$, $b = 19.57$, $c = 8.864$ Å, and $\beta = 100.53^\circ$. The two independent molecules in the asymmetric unit both have the ketothione form, and a half-chair conformation. They are linked by two hydrogen-bonds between N(3) and O(2), the molecules being twisted at 33° to one another.

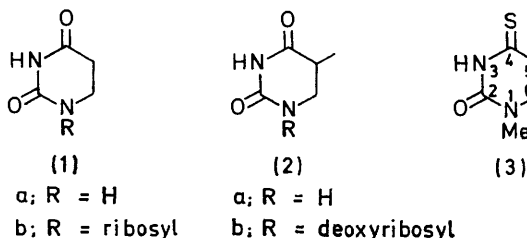
DIHYDROURACIL (1a) is one of the rarer bases of t-RNA,¹ and dihydro-derivatives of the normal nucleic acid pyrimidines are believed to be involved in certain

mutagenic processes.² Since the nonplanarity of the dihydropyrimidines can be expected to influence the secondary structure of polynucleotides, there has been

¹ For a review, see H. G. Zachau, *Angew. Chem. Internat. Edn.*, 1969, **8**, 711.

² B. Singer and H. Fraenkel-Conrat, *Progr. Nucleic Acid Res.*, 1969, **9**, 1.

recent interest in the conformations of these compounds. Crystal structures have been reported for dihydrouracil (1a),³ dihydrothymine (2a),⁴ and corresponding nucleosides [(1b) and (2b)],^{5,6} and the conformation of dihydrouracils in solution has been studied by n.m.r.⁷ The crystal structure of 5,6-dihydro-1-methyl-4-thiouracil (3) has now been determined in order to provide



further information about the geometry and hydrogen-bonding of dihydropyrimidines. Interest in thio-derivatives of the RNA bases has developed since the discovery of several thiopyrimidines in t-RNA.¹ Crystallographic information on sulphur-containing bases has been summarised recently.⁸

EXPERIMENTAL

The compound (3) was prepared as described in ref. 9 as pale yellow prismatic crystals, m.p. 194–196 °C (lit.⁹ 193–195) from methanol.

Crystal Data.—C₅H₈N₂OS, *M* = 144.1, Monoclinic, *a* = 8.086(10), *b* = 19.57(2), *c* = 8.864(10) Å, β = 100.53(10)°, *U* = 1379 Å³, *D_m* = 1.38 (by flotation), *Z* = 8, *D_c* = 1.39, *F*(000) = 608. Space-group *P*2₁/*c* (*C*_{2h}⁵, No. 14) from systematic absences. Cu-*K*_α radiation, λ = 1.5418 Å, μ(Cu-*K*_α) = 34.0 cm⁻¹. The asymmetric unit contains two independent molecules.

Crystallographic Measurements.—The space-group and approximate lattice constants were determined from Weissenberg and rotation photographs, and accurate lattice constants were then obtained from measurements with a Stoe two-circle Weissenberg diffractometer. Relative intensities for layers *h*0—12*l* were measured on the diffractometer by the θ—2θ scan technique by use of a crystal mounted about *b*. Four octants were measured and the intensities of equivalent reflections were averaged. Of 1476 independent reflections with 2θ < 120°, 1096 were considered to be 'observed'. Structure factors were calculated in the usual way after application of Lorentz and polarisation corrections. No absorption correction was made.

Structure Analysis.—Approximate scale and overall temperature factors were found by Wilson's method. The structure was solved by the heavy-atom procedure. There was some initial difficulty in interpreting the Patter-

son function owing to pseudo-translational symmetry, but eventually all non-hydrogen atoms were located. *R* was reduced from 0.33 to 0.18 by least-squares refinement with isotropic temperature factors, unit weights, and data limited to sin θ/λ < 0.4.

The remaining data out to sin θ/λ 0.56 were now included, and refinement continued with anisotropic temperature factors with the weighting scheme √*w* = 40/*F_o* for *F_o* > 40, and √*w* = 1 otherwise. The two independent molecules were refined separately, because of limitations of computer storage. When *R* was 0.123 a difference synthesis revealed 14 of the 16 hydrogen atoms with peak heights in the range 0.3–0.6 eÅ⁻³; the hydrogen atoms attached to nitrogen were not apparent. The 14 hydrogen atoms were included in the two final least-squares cycles, in which the non-hydrogen atoms were refined anisotropically and the hydrogen atoms isotropically. The final value of *R* was 0.107. The observed and calculated structure factors are listed in Supplementary Publication No. SUP 20279 (11 pp., 1 microfiche).*

Calculations were carried out with the 'X-Ray '63' system of programs on the S.R.C. Atlas computer at Chilton, and the 'X-Ray '67' system on the ICL 4/70 computer at University College, Cardiff. The full-matrix program ORFLS was used for all least-squares calculations, refinement being based on *F*. Scattering factors were taken from ref. 10a, and are for neutral atoms. 'Unobserved' reflections were included in all calculations except the least-squares sums.

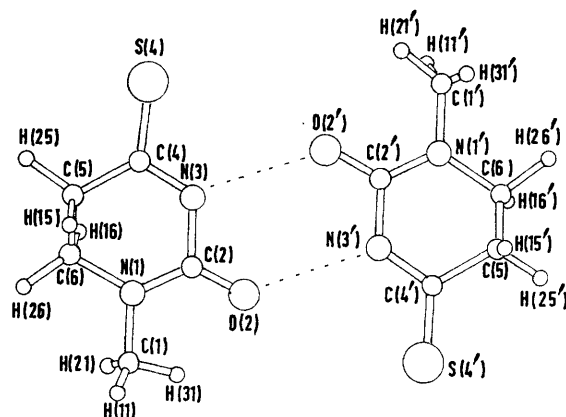


FIGURE 1 The two molecules of the asymmetric unit projected onto the least-squares plane through all non-hydrogen atoms

RESULTS

The final positional and thermal parameters are shown in Tables 1 and 2, together with the estimated standard deviations taken from the least-squares calculations. Table 3 shows interatomic distances and angles. Table 4 shows least-squares planes through the two molecules,

⁷ A. R. Katritzky, M. R. Nesbit, B. J. Kurtev, M. Lyapova, and I. G. Pojarlieff, *Tetrahedron*, 1969, **25**, 3807.

⁸ (a) G. H.-Y. Lin and M. Sundaralingam, *Acta Cryst.*, 1971, **B27**, 961; (b) W. Saenger and D. Suck, *Acta Cryst.*, 1971, **B27**, 1178; (c) G. H.-Y. Lin, M. Sundaralingam, and S. K. Arora, *J. Amer. Chem. Soc.*, 1971, **93**, 1235.

⁹ V. Skaric, B. Gaspert, I. Jerkunica, and D. Skaric, *Croat. Chem. Acta*, 1965, **37**, 199.

¹⁰ 'International Tables for X-Ray Crystallography,' vol. 3, Kynoch Press, Birmingham, 1962, (a) p. 202; (b) p. 276.

* For details of Supplementary Publications see Notice to Authors No. 7 in *J. Chem. Soc. (A)*, 1970, Issue No. 20.

³ D. C. Rohrer and M. Sundaralingam, *Acta Cryst.*, 1970, **B26**, 546.

⁴ S. Furberg and L. H. Jensen, *J. Amer. Chem. Soc.*, 1968, **90**, 470.

⁵ (a) M. Sundaralingam, S. T. Rao, and J. Abola, *Science*, 1971, **172**, 725; (b) D. Suck, W. Saenger, and K. Zechmeister, *FEBS Letters*, 1971, **12**, 257.

⁶ J. Konnert, I. L. Karle, and J. Karle, *Acta Cryst.*, 1970, **B26**, 770.

calculated by use of all the non-hydrogen atoms except C(5) and C(6). All the intermolecular contacts were calculated and none of these was unreasonable. Numbering of the atoms is indicated in Figure 1 which shows the asymmetric unit projected on to the least-squares plane through all the non-hydrogen atoms. The last digit in the number given to each hydrogen atom is the number of the carbon atom bonded to it. The packing in the unit

TABLE 1

Final atomic fractional co-ordinates with estimated standard deviations in parentheses. For hydrogen atoms the isotropic temperature parameter is also given

Molecule (1)	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> × 10
S(4)	0.6256(3)	-0.0115(1)	0.7968(2)	
O(2)	0.8202(6)	0.0874(3)	0.3556(5)	
N(1)	0.8157(7)	-0.0285(5)	0.3558(7)	
N(3)	0.7217(7)	0.0345(5)	0.5495(6)	
C(1)	0.8664(12)	-0.0323(6)	0.2028(9)	
C(2)	0.7909(8)	0.0326(7)	0.4095(9)	
C(4)	0.7021(8)	-0.0211(5)	0.6369(8)	
C(5)	0.7593(10)	-0.0879(5)	0.5854(9)	
C(6)	0.7495(11)	-0.0921(5)	0.4127(9)	
H(11)	0.987(16)	-0.044(7)	0.213(4)	93(35)
H(21)	0.807(13)	-0.053(7)	0.146(12)	60(28)
H(31)	0.855(16)	0.018(7)	0.159(15)	95(36)
H(15)	0.898(15)	-0.092(7)	0.616(14)	73(33)
H(25)	0.661(13)	-0.138(6)	0.610(12)	60(28)
H(16)	0.615(13)	-0.085(6)	0.368(13)	59(28)
H(26)	0.844(12)	-0.140(5)	0.397(11)	45(25)

Molecule (2)	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> × 10
S(4')	1.0488(2)	0.2677(1)	0.3489(2)	
O(2')	0.6170(6)	0.1692(3)	0.6001(2)	
N(1')	0.6048(8)	0.2845(5)	0.5800(7)	
N(3')	0.7981(7)	0.2220(4)	0.4704(6)	
C(1')	0.4551(11)	0.2880(6)	0.6547(10)	
C(2')	0.6641(10)	0.2246(7)	0.5544(9)	
C(4')	0.8884(9)	0.2771(5)	0.4396(8)	
C(5')	0.8394(11)	0.3438(5)	0.5013(9)	
C(6')	0.6549(11)	0.3477(5)	0.5069(10)	
H(11')	0.335(12)	0.269(5)	0.611(1)	36(23)
H(21')	0.436(13)	0.246(6)	0.733(12)	52(26)
H(31')	0.468(14)	0.313(7)	0.694(13)	73(30)
H(15')	0.936(13)	0.351(6)	0.583(13)	71(31)
H(25')	0.857(11)	0.389(5)	0.433(11)	33(22)
H(16')	0.565(10)	0.355(5)	0.418(10)	26(20)
H(26')	0.626(11)	0.397(5)	0.583(11)	38(23)

TABLE 2

Thermal parameters a ($\text{\AA}^2 \times 10^3$) with estimated standard deviations

Atom	U_{11}	U_{22}	U_{33}	$2U_{12}$	$2U_{23}$	$2U_{13}$
S(4)	69(1)	52(2)	37(1)	-6(2)	5(2)	42(2)
S(4')	53(1)	55(2)	54(1)	-15(2)	2(2)	42(2)
O(2)	68(3)	24(5)	41(3)	-6(5)	5(5)	39(5)
O(2')	60(3)	39(5)	54(3)	-8(6)	6(5)	48(5)
N(1)	48(4)	25(8)	42(4)	1(7)	-7(8)	26(6)
N(3)	40(3)	40(6)	28(3)	-1(6)	6(7)	21(5)
N(1')	55(4)	24(8)	40(4)	15(8)	-4(7)	25(6)
N(3')	41(3)	44(7)	33(3)	-17(7)	1(6)	26(5)
C(1)	75(6)	92(10)	41(5)	-4(11)	-26(10)	40(9)
C(2)	37(4)	16(11)	41(5)	6(8)	-4(11)	11(6)
C(4)	30(4)	52(9)	37(4)	-6(8)	2(9)	18(6)
C(5)	57(5)	31(9)	52(5)	14(9)	6(9)	43(8)
C(6)	62(5)	27(9)	53(5)	-19(9)	-12(9)	36(8)
C(1')	51(5)	85(10)	62(6)	19(10)	1(11)	47(9)
C(2')	47(5)	29(12)	35(4)	-3(10)	0(10)	3(7)
C(4')	41(4)	36(8)	31(4)	9(8)	-1(8)	4(6)
C(5')	69(5)	35(9)	45(5)	-20(10)	-6(9)	25(8)
C(6')	71(5)	25(9)	49(5)	6(10)	14(9)	21(8)

^a In the form $\exp[-2\pi^2(h^2a^{*2}U_{11} + k^2b^{*2}U_{22} + l^2c^{*2}U_{33} + 2hka^*b^*U_{12} + 2kib^*c^*U_{23} + 2hla^*c^*U_{13})]$.

cell is illustrated in Figure 2, which shows the structure viewed down *c*.

There are no significant differences in bond lengths or angles between the two independent molecules. Both are in a half-chair conformation, the atoms C(5) and C(6)

TABLE 3

Interatomic distances (\AA) and angles ($^\circ$) with estimated standard deviations

	Molecule (1)	Molecule (2)	
(a) Distances			
S(4)-C(4)	1.658(8)	1.657(8)	
O(2)-C(2)	1.215(14)	1.242(14)	
C(1)-N(1)	1.488(11)	1.484(12)	
N(1)-C(2)	1.315(16)	1.302(16)	
C(2)-N(3)	1.452(10)	1.424(11)	
N(3)-C(4)	1.361(11)	1.358(11)	
C(4)-C(5)	1.487(14)	1.497(14)	
C(5)-C(6)	1.520(12)	1.504(13)	
C(6)-N(1)	1.480(11)	1.486(13)	
C(1)-H(11)	0.99(13)	1.04(9)	
C(1)-H(21)	0.75(11)	1.10(11)	
C(1)-H(31)	1.05(14)	0.60(13)	
C(5)-H(15)	1.11(12)	0.98(10)	
C(5)-H(25)	1.30(11)	1.10(10)	
C(6)-H(16)	1.09(10)	0.98(8)	
C(6)-H(26)	1.24(12)	1.22(10)	
N(3)-O(2')	2.840(10)	N(3')-O(2)	2.829(10)
(b) Angles			
C(5)-C(4)-N(3)	116.9(7)	115.6(7)	
C(4)-N(3)-C(2)	124.7(8)	124.4(8)	
N(3)-C(2)-N(1)	116.1(9)	117.6(10)	
C(2)-N(1)-C(6)	123.6(7)	122.3(8)	
N(1)-C(6)-C(5)	110.0(7)	108.9(8)	
C(6)-C(5)-C(4)	113.2(8)	112.7(8)	
S(4)-C(4)-N(3)	119.8(7)	120.4(7)	
S(4)-C(4)-C(5)	123.2(7)	123.9(7)	
O(2)-C(2)-N(3)	116.5(10)	116.5(10)	
O(2)-C(2)-N(1)	127.3(8)	125.8(8)	
C(1)-N(1)-C(2)	117.5(9)	118.4(9)	
C(1)-N(1)-C(6)	116.4(8)	117.8(8)	

C-C-H and N-C-H angles were in the range 96–129°, with estimated standard deviations of 4–11°.

TABLE 4

Equations of least-squares planes, and, in square brackets, deviations (\AA) of atoms from the planes

Molecule 1: $0.94463X - 0.01586Y + 0.32774Z = +6.73167$
 [S(4) -0.02, O(2) 0.04, N(1) 0.04, N(3) -0.08, C(1) -0.05, C(2) 0.00, C(4) 0.06, C(5) * 0.33, C(6) * -0.24]

Molecule (2): $0.61049X + 0.05753Y + 0.78993Z = +6.60080$
 [S(4') 0.01, O(2') 0.07, N(1') 0.01, N(3') -0.12, C(1') -0.01, C(2') -0.02, C(4') 0.06, C(5') * 0.39, C(6') * -0.25]

Equations are in the form $lX + mY + nZ = p$, where *X*, *Y*, and *Z* are orthogonal co-ordinates defined by $X = ax \sin \beta$, $Y = by$, $Z = cz + ax \cos \beta$.

* Atoms not used in determining plane.

being above and below the plane formed by the other non-hydrogen atoms.

Although the hydrogen atoms attached to nitrogen were not located, it is clear from the bond lengths that the molecules are in the expected ketothione form; the C=S distance (1.658 \AA) is typical of pyrimidine thiones,^{8c} and the C=O distance (mean 1.229 \AA) is typical of amides and ureas.

The two independent molecules are linked by two hydrogen bonds between N(3) and O(2); the distances

(2.840 and 2.829 Å) are typical of N-H...O bonds.¹¹ There is a twist of 33° between the planes of the two hydrogen-bonded molecules.

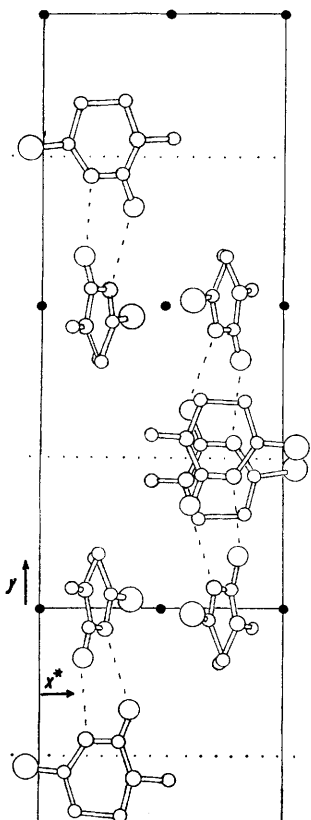


FIGURE 2 The unit cell viewed down c

DISCUSSION

The molecular geometry of the compound (3) is generally similar to the geometries of dihydrouracil (1a),³ dihydrothymine (2a),⁴ and dihydrothymidine (2b).⁶ One significant difference is that the N(3)-C(2) distance (mean 1.438 Å) is greater than the value of 1.395, 1.383, and 1.387 Å in the related compounds [(1a), (2a), and (2b)]. Electron distribution in the conjugated part of the molecule has been altered by the replacement of oxygen by sulphur at C(4).

The C(5)-C(6) bond length (mean 1.512 Å) is shorter than the usual C(sp³)-C(sp³) distance,^{10b} but its value is similar to those reported for the other dihydropyrimidines [(1a) 1.507, (2a) 1.516, and (2b) 1.500 Å]. Thermal motion has not been allowed for in the bond lengths and may be responsible for some of the apparent shortening.

The deviations of C(5) and C(6) from the plane formed by the other non-hydrogen atoms are also typical of the dihydropyrimidines. The extent of puckering in the

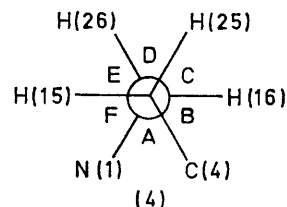
rings is indicated by the C(4)-C(5)-C(6)-N(1) torsion angles: 45.6 and 48.4°, cf. 45.4 (1a), 55.6 (2a), and 52° (1b).^{5b} Table 5 lists all the torsion angles at the

TABLE 5

Torsion angles (°) with estimated standard deviations ^a			
Angle	Molecule (1)	Molecule (2)	Dihydrouracil ^b
C(4)-C(5)-C(6)-N(1)	+45.6(0.8)	+48.4(0.8)	+45.4
C(4)-C(5)-C(6)-H(16)	-59(6)	-76(8)	-73
H(25)-C(5)-C(6)-H(16)	+60(9)	+49(9)	+38
H(25)-C(5)-C(6)-H(26)	-78(9)	-68(9)	-73
H(15)-C(5)-C(6)-H(26)	+46(9)	+42(9)	+55
H(15)-C(5)-C(6)-N(1)	-71(6)	-77(6)	-74

^a Calculated by the approximate formula of R. F. Bryan and J. D. Dunitz, *Helv. Chem. Acta*, 1960, **43**, 3. ^b Ref. 3.

C(5)-C(6) bond, with the corresponding values for dihydrouracil, taken from ref. 3. Although the estimated standard deviations in the present work are large for the angles involving hydrogen atoms, the general conclusion is that the angle D in (4) is *ca.* 70°,



while angles C and E are considerably smaller. This is consistent with the coupling constants found for dihydrouracils,⁷ which are approximately J_{ee} 2, J_{ae} 4, and J_{aa} 12 Hz.

It is worthwhile to compare the hydrogen-bonding in dihydropyrimidines with that in the fully saturated compounds. Voet and Rich¹¹ have pointed out that uracil derivatives, with the exception of some 5-halogeno-compounds, tend to form hydrogen-bonds from O(4) in complexes with themselves or with purines. For example, 1-methyluracil¹² and 1-methylthymine¹³ both form cyclic hydrogen-bonded pairs linked between N(3) and O(4), while O(2) is free. In the present compound the hydrogen-bonded pair is linked through N(3) and O(2). This less usual arrangement may be attributable either to the saturation of the 5,6-double bond, or to the presence of sulphur at C(4) in place of oxygen. There is some justification for both these suggestions. Firstly, N-H...S bonds are weaker and less likely than N-H...O bonds in situations where both are possible;¹⁴ an illustration of this is the structure of the 1-methyl-4-thiouracil:9-methyladenine complex¹⁵ in which the adenine is bonded to N(3) and O(2) of the pyrimidine, whereas in similar complexes formed by 1-methyluracil and 1-methylthymine, N(3) and O(4) are involved.¹⁶ On the other hand, in both

¹¹ D. Voet and A. Rich, *Progr. Nucleic Acid Res.*, 1970, **10**, 183.

¹² D. W. Green, F. S. Mathews, and A. Rich, *J. Biol. Chem.*, 1962, **237**, 3573.

¹³ K. Hoogsteen, *Acta Cryst.*, 1963, **16**, 28.

¹⁴ J. Donohue, *J. Mol. Biol.*, 1969, **45**, 231.

¹⁵ W. Saenger and D. Suck, *Nature*, 1970, **227**, 1046; *J. Mol. Biol.*, 1971, **60**, 87.

¹⁶ F. S. Mathews and A. Rich, *J. Mol. Biol.*, 1964, **8**, 89; K. Hoogsteen, *Acta Cryst.*, 1963, **16**, 907.

dihydrouracil and dihydrothymine O(4) is free from hydrogen-bonding while N(1), N(3), and O(2) are involved in making infinite hydrogen-bonded sheets. It seems probable that when the 5,6-bond is saturated, O(2) is a relatively better hydrogen-bond acceptor than when the 5,6-bond is unsaturated.

In the present structure the two hydrogen-bonded molecules are twisted at 33° to one another (see Figure 2). In the numerous hydrogen-bonded pyrimidine-pyrimidine and pyrimidine-purine pairs reviewed by Voet and Rich,¹¹ the angle between the two molecules is always small, the largest being 18° . The present result suggests that dihydropyrimidines are less specific in geometrical requirements for effective interaction than are the fully unsaturated compounds. In dihydro-

uracil and dihydrothymine the hydrogen-bonding is across centres of symmetry, so planarity is imposed. It would be of considerable interest to know the structure of a complex between a dihydrouracil derivative and an alkyladenine; there is i.r. evidence for such interaction in solution.¹⁷

I thank the S.R.C. and the staff of the Atlas Computer Centre for computing facilities, and Drs. W. Jones and R. Hine for implementing 'X-Ray '67' on the ICL 4/70 computer at University College, Cardiff.

[1/1791 Received, September 29th, 1971]

¹⁷ Y. Kyogoku, R. C. Lord, and A. Rich, *Proc. Nat. Acad. Sci. U.S.A.*, 1967, **57**, 250.