

## The Crystal and Molecular Structure of Thymidine

BY D. W. YOUNG, P. TOLLIN AND H. R. WILSON

*Carnegie Laboratory of Physics, University of Dundee, Scotland*

(Received 8 August 1968)

The crystal structure of thymidine ( $C_{10}N_2O_5H_{14}$ ) has been determined from data collected on a Hilger and Watts linear diffractometer and a Wooster four-circle diffractometer. Thymidine is orthorhombic, space group  $P2_12_12_1$ , with the following cell dimensions:  $a = 4.860 \pm 0.005$ ,  $b = 13.91 \pm 0.01$  and  $c = 16.32 \pm 0.01$  Å. The structure was solved by a Patterson interpretation method and the positional and thermal parameters were refined by the method of least-squares. The carbon, nitrogen and oxygen atom thermal parameters were refined anisotropically. The final  $R$  value for the 1230 observed reflexions is 0.066 and the average standard deviations in the bond lengths and angles are about 0.006 Å and  $0.4^\circ$  respectively.

The pyrimidine ring in thymidine is planar and atom C(1') is also coplanar with the ring. The dihedral angle between the base and sugar planes is  $74^\circ$ , with the glycosidic torsion angle,  $\varphi_{CN} = -40^\circ$ . Atom C(3') of the deoxyribose ring is displaced by 0.57 Å from the plane of the remaining ring atoms and on the opposite side to atom C(5'). The orientation of the C(5')-O(5') bond is *gauche* to C(4')-O(1') and *trans* to C(4')-C(3').

### Introduction

The determination of the structure of thymidine was undertaken as part of a series of structure determinations of nucleic acid components. When this structure determination was started thymidine was the only nucleoside which had been studied by electron spin resonance of irradiated single crystals (Pruden, Snipes & Gordy, 1965). Radiation effects on thymine make an important contribution to the spectra obtained in electron spin resonance experiments on irradiated deoxyribose nucleic acid and it therefore seemed important to have a knowledge of the structure of the undamaged thymidine molecule (Gordy, Pruden &

Snipes, 1965). In addition, the approximate orientation of the pyrimidine ring in the crystal has been predicted from the electron spin resonance studies and it was hoped to confirm this result. A preliminary account of the conformation of the molecule and a comparison with the electron spin resonance results has been given recently (Tollin, Wilson & Young, 1968).

### Experimental

Crystals of thymidine ( $C_{10}N_2O_5H_{14}$ ), whose structural formula appears in Fig. 1 along with the numbering system used in this paper, were obtained by evaporation from aqueous solution. The cell dimensions were obtained from Weissenberg and precession photographs with Cu  $K\alpha$  radiation. The crystals are orthorhombic with unit-cell dimensions:

$$a = 4.860 \pm 0.005, \quad b = 13.91 \pm 0.01, \quad c = 16.32 \pm 0.01 \text{ \AA}$$

$$(\lambda \text{ Cu } K\alpha_1 = 1.54050 \text{ \AA}, \lambda \text{ Cu } K\alpha_2 = 1.54434 \text{ \AA}).$$

The systematic absences,  $h00$  with  $h = 2n + 1$ ,  $0k0$  with  $k = 2n + 1$  and  $00l$  with  $l = 2n + 1$  uniquely determine the space group as  $P2_12_12_1$ . There are four molecules in the unit cell and these results are in good agreement with those of Furberg & Hordvik (1956).

In all 1230 reflexions were measured on a Hilger and Watts linear diffractometer, molybdenum  $K$  radiation and balanced filters being used. The data were collected out to a  $\sin\theta$  value corresponding to the radius of the limiting sphere for copper  $K\alpha$  radiation. A number of reflexions close to the centre of each layer, which cannot be measured with any accuracy on the diffractometer because of the uneven background observed when the reflexion is close to the direct beam, were measured by use of a Wooster four-circle diffractometer.

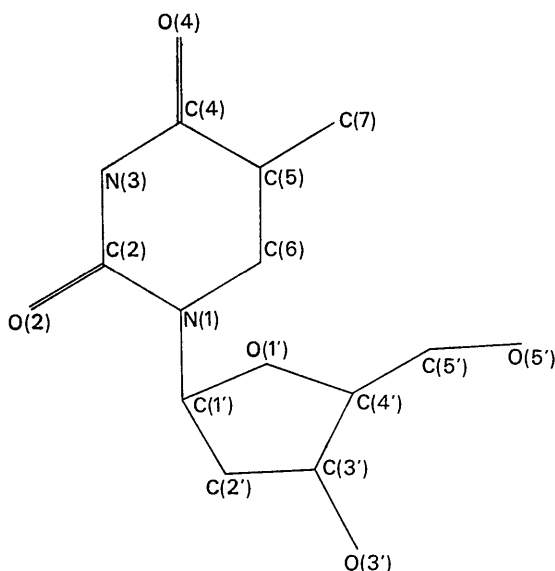


Fig. 1. The thymidine molecule and the atom numbering.

### Structure determination

The function  $I(\theta, \varphi)$  (Tollin & Cochran, 1964) was computed with the 100 largest sharpened Patterson coefficients,  $|F_s(h)|^2$ , calculated by the modification function defined by Lipson & Cochran (1957) with a value of  $p=29.0$ . A value of  $3.4 \text{ \AA}$  was used for the radius  $R$  of the disc and the results obtained, plotted on a Sanson-Flamsteed projection, are shown in Fig. 2. The largest peak on this map indicates that the normal to the pyrimidine has a direction defined by  $\theta=90^\circ$  and  $\varphi=43.5^\circ$ , where  $\theta$  and  $\varphi$  are the spherical polar angles with respect to the axial set  $a, b, c$ . The section through the origin of sharpened Patterson function, computed in the orientation given by those  $\theta$  and  $\varphi$  values is shown in Fig. 3.

The theoretical vector set for the pyrimidine residue with attachments and including the atom C(1'), was superimposed on the Patterson section and rotated until the best fit was obtained. Because the angle  $\theta=90^\circ$  this plane contains the crystallographic  $c$  axis and hence the Patterson section has  $mm$  symmetry. The Patterson section, therefore, contains the vector sets of two of the pyrimidine residues related by a mirror plane and this has to be taken into account when the theoretical vector set is fitted to the Patterson function.

At this stage the angular coordinates defining the orientation of the pyrimidine ring are known and it remains to find the position of this ring in the unit cell. The  $Q$  functions defined by Tollin (1966) were used to determine this position. Fig. 4 shows the three  $Q$  functions  $Q(X_0, Y_0)$ ,  $Q(X_0, Z_0)$  and  $Q(Y_0, Z_0)$  which define the position of the pyrimidine residue with respect to the  $2_1$ -axes parallel to  $c, b$  and  $a$  respectively. The largest peaks on these maps are mutually consistent and indicate a vector of  $x=0.127, y=0.361, z=0.216$  which must be added to the relative fractional coordinates of the pyrimidine ring to give its position in the unit cell.

Only for the map of  $Q(Y_0, Z_0)$  are there atoms having relative fractional  $x$  coordinates separated by  $\frac{1}{2}$ , and these give rise to false peaks which can be predicted *a priori* as has been noted previously (Tollin, 1966). The positions of these false peaks are indicated on the  $Q(Y_0, Z_0)$  map. The pyrimidine ring along with atom C(1') is approximately centrosymmetric and, therefore, the  $Q$  functions ought to possess approximate centres of symmetry, which are evident in Fig. 4.

The 280 largest three-dimensional structure factors were calculated by use of the coordinates of the atoms of the planar group only and the conventional  $R$  value

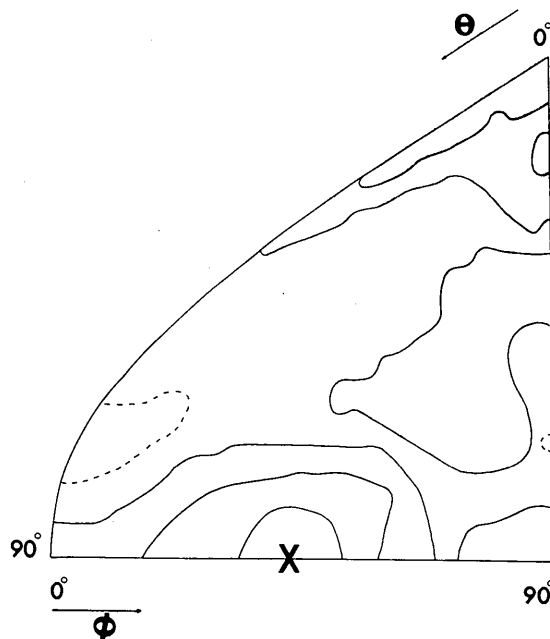


Fig. 2. The  $I(\theta, \varphi)$  map with  $R=3.4 \text{ \AA}$ . The largest peak is marked X.

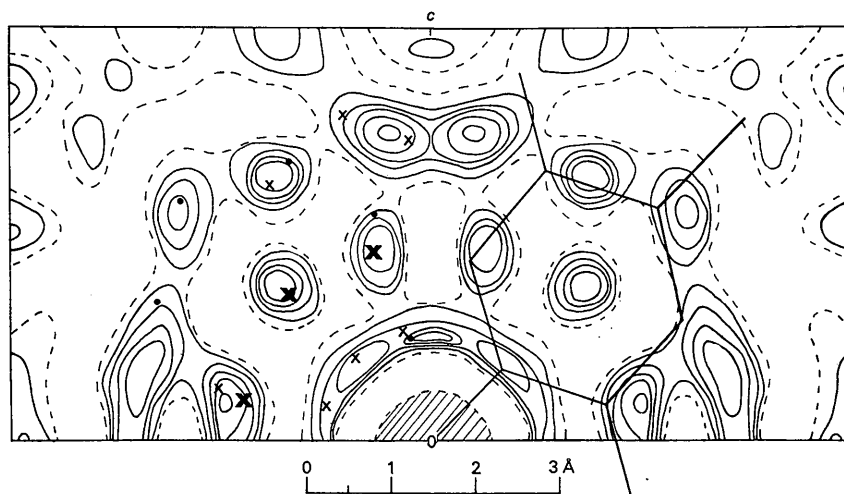


Fig. 3. The Patterson section through the origin, in the orientation given by the largest peak in the  $I(\theta, \varphi)$  function. Contours are drawn at equal arbitrary intervals except at the origin. Superimposed is the pyrimidine ring and its vector set. Peaks marked . are weight 1, x weight 2 and x are weight 4.

defined by  $R = \sum |F_o| - |F_c| / \sum |F_o|$ , was 0.44. The phases from these structure factors along with the magnitudes of the observed structure factors were used to calculate a three-dimensional electron density map and the peaks corresponding to the atoms of the sugar group were clearly visible and approximately half the height of those of the atoms of the pyrimidine nucleus.

The positions were extracted from the map and the  $R$  value for the 280 largest three-dimensional structure factors was 0.32.

### Refinement

The structure was refined by a least-squares refinement procedure, using the same 280 largest structure factors and an overall isotropic temperature factor to an  $R$  value of 0.19. The introduction of individual isotropic temperature parameters and refinement with all the

reflexion data reduced the  $R$  value to 0.13. The least-squares refinement up to this stage was performed on the STANTEC ZEBRA computer using the program of J.C. Schoone (Rogers, 1961). The positions of the hydrogen atoms H(N3), H(C6), H(Cl'), H(C2'), H2(C2'), H(C3), H(C4'), H(C5') and H2(C5') were calculated assuming normal bond lengths and angles. Keeping the coordinates of these atoms fixed, the structure was refined by least-squares using anisotropic thermal parameters for the heavy atoms, to an  $R$  value of 0.09 in two cycles. This refinement was performed on the IBM 1620 II computer using the *SFLS* program of Dr G.A. Mair. The program uses the block-diagonal approximation and the coefficients were weighted by

$$\omega = 1 / \{1 + (KF_o - b)^2 / a^2\},$$

where  $a = 16$  and  $b = 10$ . The anisotropic temperature factors had the form

$$\exp[-(B_{11}h^2 + B_{22}k^2 + B_{33}l^2 + B_{12}hk + B_{23}kl + B_{13}hl)].$$

An electron density map and a difference Fourier synthesis were calculated. The five largest peaks in the difference Fourier, with peak heights between 0.42 and 0.55 e.Å<sup>-3</sup>, were interpreted as the locations of the five missing hydrogen atoms. The next highest peak had a height of 0.32 e.Å<sup>-3</sup>. The positions of the hydrogen atoms H(O3') and H(O5') confirmed the hydrogen bond scheme suggested by the calculated intermolecular distances. The electron density map was entirely consistent with this interpretation of the difference Fourier synthesis. With all the hydrogen atoms included in the structure factor calculation, but refining only the heavy atom parameters, least-squares refinement proceeded to a final  $R$  value of 0.066.

Including the unobserved reflexions with structure factors of zero value led to an  $R$  value of 0.091. At the final stage the maximum positional parameter shift was 0.0016 Å corresponding to 0.30  $\sigma$  and the maximum thermal parameter shift was 0.26  $\sigma$ . Fourteen sets of structure factor calculations were performed, in each of which one of the hydrogen atoms was omitted and for each a difference Fourier synthesis was performed in the region of the omitted hydrogen atom. Fig. 5 is a composite of these difference density calculations. The final values of the parameters of the heavy atoms with their estimated standard deviations are given in Tables 1 and 2, and the parameters for the hydrogen atoms in Table 3. The observed and calculated structure factors are listed in Table 4. Table 5 lists the intramolecular bond lengths and angles, and Table 6 gives the intermolecular contacts.

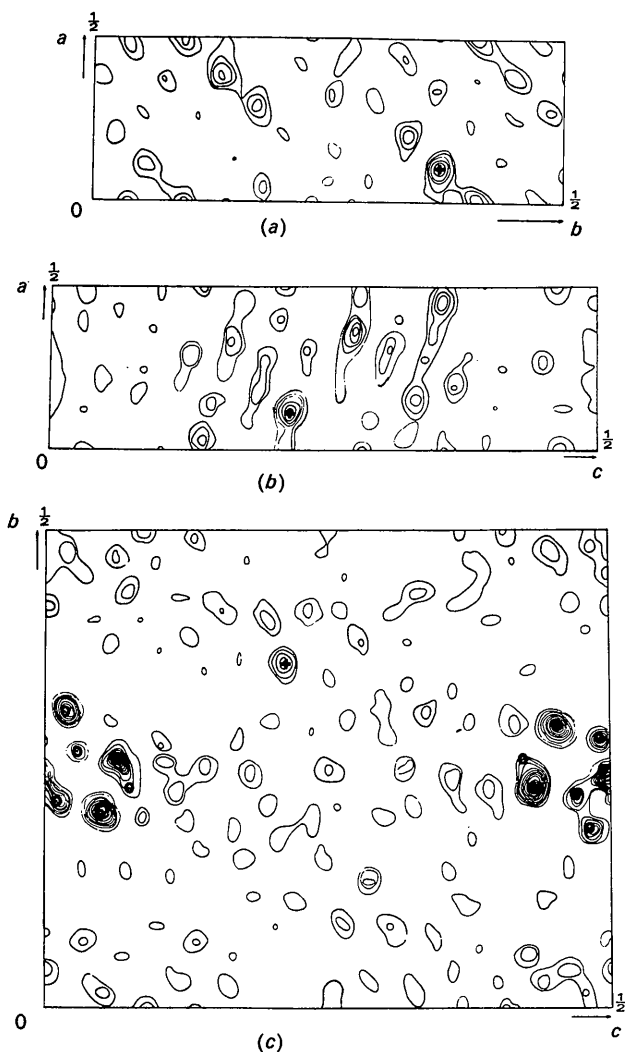


Fig. 4. The  $Q$  function maps, each contoured at arbitrary equal intervals. (a) is  $Q(X_o, Y_o)$ , (b) is  $Q(X_o, Z_o)$  and (c) is  $Q(Y_o, Z_o)$ . The false peaks in (c) are indicated by  $\odot$ ; the correct locating peaks are marked +.

Table 1. Positional parameters of the non-hydrogen atoms with estimated standard deviations, in fractional coordinates,  $\times 10^3$

	$x/a$	$y/b$	$z/c$
N(1)	0.0244 (8)	0.4122 (3)	0.2758 (2)
C(2)	0.0974 (11)	0.4329 (3)	0.1957 (3)
N(3)	-0.0448 (9)	0.3797 (3)	0.1382 (2)

Table 1 (cont.)

	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>
C(4)	-0.2460 (10)	0.3119 (3)	0.1523 (3)
C(5)	-0.3086 (9)	0.2928 (3)	0.2378 (2)
C(6)	-0.1703 (10)	0.3439 (3)	0.2942 (2)
O(2)	0.2687 (9)	0.4914 (3)	0.1764 (2)
O(4)	-0.3575 (8)	0.2716 (2)	0.0940 (2)
C(7)	-0.5195 (12)	0.2176 (4)	0.2589 (3)
C(1')	0.1715 (10)	0.4657 (3)	0.3408 (2)
C(2')	0.0091 (10)	0.5478 (3)	0.3779 (3)
C(3')	0.1095 (10)	0.5490 (3)	0.4663 (3)
C(4')	0.1453 (9)	0.4419 (3)	0.4845 (2)
C(5')	-0.1158 (9)	0.3942 (2)	0.5148 (3)
O(1')	0.2307 (7)	0.4003 (2)	0.4064 (2)
O(3')	0.3767 (7)	0.5911 (2)	0.4725 (2)
O(5')	-0.0648 (7)	0.2935 (2)	0.5195 (2)

Table 3. The atomic parameters for the hydrogen atoms

	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>	<i>B</i>
H(N3)	0.0104	0.3832	0.0820	3.50 Å <sup>2</sup>
H(C6)	-0.2105	0.3356	0.3514	3.50
H1(C7)	-0.5937	0.2320	0.3152	3.50
H2(C7)	-0.4406	0.1536	0.2594	3.50
H3(C7)	-0.6811	0.2191	0.2208	3.50
H(C1')	0.3340	0.4917	0.3151	3.50
H1(C2')	0.0475	0.6110	0.3527	3.50
H2(C2')	-0.1882	0.5307	0.3823	3.50
H(C3')	-0.0174	0.5857	0.5008	3.50
H(C4')	0.2905	0.4290	0.5261	3.50
H1(C5')	-0.1584	0.4192	0.5705	3.50
H2(C5')	-0.2833	0.4099	0.4786	3.50
H(O3')	0.3718	0.6510	0.4480	3.50
H(O5')	-0.2183	0.2671	0.5124	3.50

Table 2. Thermal parameters of the non-hydrogen atoms with their estimated standard deviations in brackets

The anisotropic temperature factors and their e.s.d. have been multiplied by 10<sup>5</sup>.

	<i>B</i> <sub>11</sub>	<i>B</i> <sub>12</sub>	<i>B</i> <sub>13</sub>	<i>B</i> <sub>22</sub>	<i>B</i> <sub>23</sub>	<i>B</i> <sub>33</sub>
N(1)	2799 (163)	151 (109)	58 (81)	337 (18)	15 (24)	146 (10)
C(2)	3052 (210)	208 (134)	-49 (107)	366 (22)	-13 (33)	227 (15)
N(3)	3758 (187)	84 (117)	165 (88)	385 (18)	-37 (26)	148 (11)
C(4)	2676 (180)	57 (117)	169 (98)	250 (18)	-99 (29)	205 (13)
C(5)	2284 (185)	-3 (120)	43 (92)	294 (19)	-31 (30)	191 (13)
C(6)	2802 (200)	336 (125)	-209 (91)	359 (21)	15 (30)	139 (12)
O(2)	5954 (197)	1955 (111)	-347 (94)	591 (19)	26 (28)	218 (11)
O(4)	5685 (198)	267 (112)	539 (95)	340 (15)	-140 (23)	240 (11)
C(7)	4047 (244)	473 (157)	-129 (120)	403 (24)	-3 (37)	284 (17)
C(1')	2794 (191)	352 (118)	198 (95)	293 (18)	70 (28)	156 (13)
C(2')	2969 (203)	-105 (121)	649 (96)	257 (18)	35 (30)	234 (14)
C(3')	2887 (200)	25 (125)	238 (97)	251 (18)	-2 (29)	192 (13)
C(4')	1846 (158)	47 (113)	33 (87)	271 (17)	-72 (27)	171 (13)
C(5')	1809 (170)	54 (125)	-288 (102)	331 (20)	-26 (34)	295 (16)
O(1')	3278 (136)	-479 (89)	-90 (70)	273 (13)	-26 (21)	161 (8)
O(3')	3094 (139)	323 (91)	359 (70)	287 (13)	74 (21)	187 (9)
O(5')	2624 (140)	411 (94)	-384 (97)	310 (15)	72 (31)	592 (18)

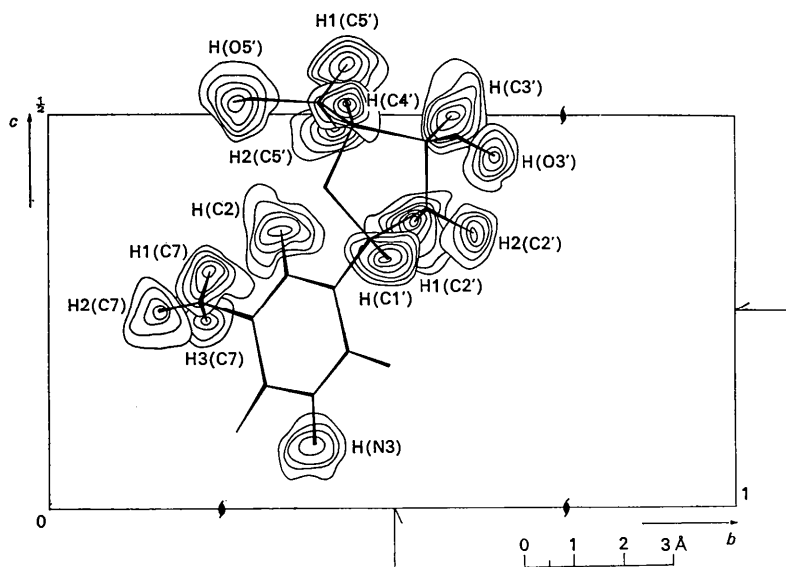


Fig. 5. Composite  $\Delta F$  synthesis showing the electron density in sections through the hydrogen atoms. Contours are at intervals of 0.1 e.Å<sup>-3</sup> with the first contour at 0.1 e.Å<sup>-3</sup>.

Table 4. The observed and calculated structure factors

Both the observed and calculated structure factors have been multiplied by 100. The unobserved reflexions are indicated by an asterisk.

Table with columns for h, k, l, Fobs, Fcalc, and multiple rows of numerical data representing structure factors for various reflections.



atom in the ring. In thymidine, the C(1') atom of the sugar is attached to N(1) so the same reasoning applies. The percentage contributions of the resonance structures proposed by Hoogsteen (1963a) for 1-methylthymine should thus be in good agreement with the

bond lengths observed in thymidine. This is certainly the case, even for the C(4)–C(5) bond, which has a calculated value of 1.441 Å (Hoogsteen, 1963a).

The bond lengths in thymine have been derived from molecular orbital calculation by Pullman & Pullman

Table 6. *Intermolecular contacts between atoms less than 3.5 Å apart (excluding hydrogen atoms)*

The distances marked \* are the hydrogen bonding contacts.

<i>i</i>	<i>j</i>	Equipoint of atom <i>j</i>	<i>d<sub>ij</sub></i> Å
O(2)	C(4)	1 + <i>x</i> , <i>y</i> , <i>z</i>	3.457 Å
O(3')	C(2')	1 + <i>x</i> , <i>y</i> , <i>z</i>	3.492
C(3')	O(4)	– <i>x</i> , $\frac{1}{2} + y$ , $\frac{1}{2} - z$	3.465
O(3')	O(4)	– <i>x</i> , $\frac{1}{2} + y$ , $\frac{1}{2} - z$	*2.737
C(4')	O(2)	$\frac{1}{2} + x$ , 1 – <i>y</i> , $\frac{1}{2} + z$	3.293
O(3')	N(3)	$\frac{1}{2} - x$ , 1 – <i>y</i> , $\frac{1}{2} + z$	*2.854
O(4)	C(3')	$\frac{1}{2} - x$ , 1 – <i>y</i> , $-\frac{1}{2} + z$	3.474
C(5')	O(5')	$\frac{1}{2} + x$ , $\frac{1}{2} - y$ , 1 – <i>z</i>	3.448
O(5')	O(5')	$\frac{1}{2} + x$ , $\frac{1}{2} - y$ , 1 – <i>z</i>	*2.788
O(5')	O(1')	$\frac{1}{2} + x$ , $\frac{1}{2} - y$ , 1 – <i>z</i>	3.117

Table 7. *A comparison between the bond lengths of thymine as determined for 1-methylthymine, thymine monohydrate and thymidine*

	NMT(1)	NMT(2)	TMH	TdS	M.O.
N(1)–C(2)	1.379	1.376	1.355	1.385	1.38
C(2)–N(3)	1.379	1.378	1.361	1.381	1.38
N(3)–C(4)	1.375	1.377	1.391	1.378	1.38
N(1)–C(6)	1.383	1.382	1.382	1.374	1.36
C(4)–C(5)	1.432	1.422	1.447	1.453	1.45
C(5)–C(6)	1.346	1.333	1.349	1.343	1.38
C(5)–C(7)	1.497	1.510	1.503	1.504	1.50
C(2)–O(2)	1.214	1.207	1.234	1.206	1.24
C(4)–O(4)	1.237	1.238	1.231	1.230	1.24

NMT(1): 1-Methyl thymine (Hoogsteen, 1963a);  $\sigma(L) \approx 0.004$  Å.

NMT(2): 1-Methyl thymine, 9-methyl adenine complex (Hoogsteen, 1963b);  $\sigma(L) \approx 0.005$  Å.

TMH: Thymine monohydrate (Gerdil, 1961);  $\sigma(L) \approx 0.004$  Å.

TdS: Present investigation;  $\sigma(L) \approx 0.006$  Å.

M.O.: Molecular orbital calculations (Pullman & Pullman, 1963).

$\sigma(L)$  represents the average estimated standard deviations of the bond lengths in the various structure determinations.

Table 8. *A comparison between the bond angles of thymine as determined for 1-methylthymine, thymine monohydrate and thymidine*

	NMT(1)	NMT(2)	TMH	TdS
N(1)–C(2)–N(3)	115.4	115.0	115.2	113.7
C(2)–N(3)–C(4)	126.3	126.3	126.3	127.5
N(3)–C(4)–C(5)	116.1	115.7	115.6	115.8
C(4)–C(5)–C(6)	118.3	119.0	118.2	117.2
C(5)–C(6)–N(1)	123.3	123.2	121.8	124.1
C(6)–N(1)–C(2)	120.6	120.4	122.8	121.8
C(1')–N(1)–C(6)	121.2	121.1		121.6
C(1')–N(1)–C(2)	118.2	119.4		116.6
N(1)–C(2)–O(2)	123.3	124.3	122.7	124.3
N(3)–C(2)–O(2)	121.3	120.7	122.1	122.0
N(3)–C(4)–O(4)	120.0	119.9	118.3	119.7
C(5)–C(4)–O(4)	123.9	124.4	126.1	124.5
C(7)–C(5)–C(4)	119.3	117.3	119.0	119.3
C(7)–C(5)–C(6)	122.4	123.6	122.8	123.5

NMT(1): See Table 7  $\sigma(\phi) \approx 0.2^\circ$ .

NMT(2): See Table 7  $\sigma(\phi) \approx 0.2^\circ$ .

TMH: See Table 7  $\sigma(\phi) \approx 0.2^\circ$ .

TdS: See Table 7  $\sigma(\phi) \approx 0.4^\circ$ .

$\sigma(\phi)$  represents the average estimated standard deviations of the bond angles in the various structures.

(1963). These values are also listed in Table 7, and comparison shows that the agreement between them and those found in thymidine is good, except for the bond C(2)–O(2).

Although the agreement between the bond lengths found in thymidine and 1-methylthymine and thymine monohydrate is good, quite a few of the angles differ by four times the standard deviation, which is significant.

The dimensions of the thymine part of thymidine are in good agreement with those predicted by Spencer (1959) from the known crystal data at that time and which were used in constructing the molecular models of the *B*, *C* and *A* forms of deoxyribonucleic acid (Langridge, Marvin, Seeds, Wilson, Hooper, Wilkins & Hamilton, 1960; Marvin, Spencer, Wilkins & Hamilton, 1961; Fuller, Wilkins, Wilson & Hamilton, 1965).

The glycosidic bond length of 1.480 Å is in agreement with that found in other nucleosides and nucleotides.

#### The deoxyribose ring

Sugar rings in nucleosides and nucleotides are usually puckered with either C(2') or C(3') out of the plane formed by the other four atoms (Spencer, 1959). The amount of displacement is about 0.5 Å and the displaced atom may either be *endo*, *i.e.* lying on the same side of the sugar plane as atom C(5') and the ring of the base, or *exo*, lying on the opposite side of the sugar plane to C(5') and the base. The most common conformation of the furanose ring is with C(2') or C(3') *endo*, and it has been suggested (Jardetsky, 1960) that in solution C(2') is *endo* in purine nucleosides and C(3') is *endo* in pyrimidine nucleosides. In the crystalline state, however, this is not the case and it appears that the hydrogen bonding and packing arrangement determines the sugar pucker because the same nucleoside may have a different type of pucker in different environments (Haschemeyer & Sobell, 1965; Iball, Morgan & Wilson, 1968).

In thymidine the conformation is C(3') *exo*, with a displacement of 0.567 Å from the least-squares plane through the other four ring atoms, which is given by the equation

$$0.8661x + 0.4876y + 0.1099z = 4.4544.$$

Thymidine is the only pyrimidine nucleoside or nucleotide so far studied that has this conformation. The only other nucleoside which has C(3') *exo* is deoxyadenosine (Watson, Sutor & Tollin, 1965) where the displacement of the atom from the least-squares plane is 0.552 Å. In calcium thymidylate the sugar is puckered with C(3') *endo*, and in this conformation the atom O(3') is brought close to the least-squares plane. When C(3') is *exo* however, O(3) is about 2 Å from the least-squares plane. In thymidine, atom O(3') is 1.988 Å from the plane, which may be compared with 1.970 Å in deoxyadenosine.

Another variable parameter in sugars is the conformation of the C(5')–O(5') bond with respect to the

furanose ring, and this can be defined in terms of rotation about the C(4')–C(5') bond. Shefter & Trueblood (1965) have designated the angle between the projected C(5')–O(5') and C(4')–O(1') bonds as  $\varphi_{OO}$ , and the angle between the projected C(5')–O(5') and C(4')–C(3') bonds as  $\varphi_{OC}$ . In thymidine  $\varphi_{OO} = 56.1^\circ$  and  $\varphi_{OC} = 172.8$ , so that C(5')–O(5') is *gauche* to C(4')–O(1') and *trans* to C(4')–C(3'). In the majority of nucleosides and nucleotides that have so far been studied, and in all the 5'-nucleotides, the conformation is such that C(5')–O(5') is *gauche* to both C(4')–O(1') and C(4')–C(3'). Deoxyadenosine, however, does have the *gauche-trans* conformation, so that this feature, as well as the C(3') *exo* pucker, is the same as in thymidine. The only other reported *gauche-trans* conformations are in 5-bromodeoxyuridine (Iball, Morgan & Wilson, 1966) where the sugar ring is puckered with C(2') *endo*, and in adenosine 3'-phosphate (Sundaralingam, 1966) where the pucker is C(3') *endo*.

The bond lengths and angles of the sugar in thymidine are listed in Table 5. Sundaralingam & Jensen (1965) have discussed in detail the various conformations assumed by the sugar residues in nucleosides and nucleotides and have also compared the bond lengths and angles in a number of well refined structures. The bond lengths found in thymidine are in good agreement with the mean values calculated by Sundaralingam & Jensen. As in other structures, the C(4')–O(1') bond in thymidine is significantly longer than the C(1')–O(1') bond, the difference in this case being 0.026 Å. The length of the C(3')–O(3') bond is 1.428 Å which is close to the normal value for a C–O bond and may be compared with the value 1.422 Å in deoxyadenosine. These values are about 0.03 Å greater than the average value for the same bond when C(3') is *endo* (Sundaralingam & Jensen, 1965).

The three types of angles within the furanose ring C–C–C, C–C–O and C–O–C have average values of

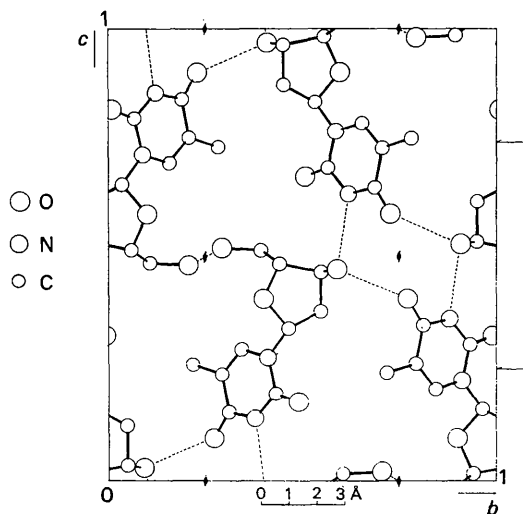


Fig. 6. The contents of a unit cell viewed down the *a* axis. The intermolecular hydrogen bonds are shown as broken lines.



102.4°, 105.5° and 110.1°. The differences between these are significant and agree with those found in other nucleosides and nucleotides (Sundaralingam & Jensen, 1965). Angle C(2')-C(3')-C(4') is only 0.6° greater than angle C(1')-C(2')-C(3') in the present structure. This is less than the average difference between the angle involving the out-of-plane atom and that involving the in-plane atom, and in deoxyadenosine the difference is 1.8°. The external angles involving the *exo* C(3') atom in thymidine differ from those found in deoxyadenosine. In thymidine, C(2')-C(3')-O(3') is greater than C(4')-C(3')-O(3') by 4.9°, whereas in deoxyadenosine C(4')-C(3')-O(3') is the greater by 6.4°.

In agreement with other structures (Sundaralingam & Jensen, 1965) the external angles N(3)-C(1')-O(1') and C(5')-C(4')-O(1') of 108.3° and 110.5°, respectively, are significantly smaller than the angles N(3)-C(1')-C(2') and C(3')-C(4')-C(5') of 114.4° and 113.2° respectively.

#### Conformation of the molecule

The dihedral angle between the thymine base and the best plane of the sugar ring is 74.3°. This value is similar to those found in other nucleosides and nucleotides. The orientation of the sugar ring relative to the thymine base can be specified in terms of rotation about the glycosidic bond (Donohue & Trueblood, 1960). Donohue & Trueblood define the torsion angle  $\varphi_{CN}$  as the angle formed by the trace of the plane of the base with the projection of C(1')-O(1') bond of the furanose ring when viewed along the C-N glycosidic bond. Like all the other pyrimidine nucleosides and nucleotides which have been studied, the sugar in thymidine is in the *anti* conformation with  $\varphi_{CN} = -40.0^\circ$ .

This may be compared with  $\varphi_{CN} = -48^\circ$  in calcium thymidylate. Both values fall within the allowed range of torsion angles determined by Haschemeyer & Rich (1967), based on accepted intermolecular van der Waals radii.

The angle  $\varphi_{CN}$ , together with the angles  $\varphi_{OO}$  and  $\varphi_{OC}$ , and the displacement of atom C(3') completely

Table 9. A comparison between the conformational parameters in calcium thymidylate and thymidine

	Calcium thymidylate	Thymidine
$\Delta 1$	0.05 Å	0.004 Å
$\varphi_{CN}$	48°	40°
$\varphi_{OO}$	63°	56°
$\varphi_{OC}$	57°	173°
$\psi$	75°	74°
$\Delta 3$	0.53 Å	0.57 Å
Conformation	<i>endo</i>	<i>exo</i>

$\psi$  is the angle between the plane of the deoxyribose ring and the plane of the pyrimidine nucleus.

$\Delta 3$  is the magnitude of the displacement of atom C(3') from the plane of other deoxyribose ring atoms.

$\Delta 1$  is the magnitude of the displacement of atom C(1') from the plane of the pyrimidine nucleus.

define the conformation of the molecule. A summary of the conformation parameters in thymidine and calcium thymidylate are listed in Table 9.

#### Molecular packing and hydrogen bonds

The arrangement of the thymidine molecules in the unit cell is shown in Fig. 6. Hydrogen bonds are represented by dashed lines.

The thymine bases form hydrogen bonds with atoms O(3') of the sugars of other molecules. There is a hydrogen bond between atom O(3') in the reference molecule at  $x, y, z$  and atom O(4) in the molecule at  $-x, \frac{1}{2} + y, \frac{1}{2} - z$ , with an OH...O length of 2.74 Å. In this case, atom O(3') is donating the hydrogen atom. The same O(3') atom also accepts a hydrogen atom from N(3) in the molecule at  $\frac{1}{2} - x, 1 - y, \frac{1}{2} + z$ , and the N-H...O length is 2.85 Å. A third hydrogen bond is formed between atom O(5') of the reference molecule and atom O(5') of the molecule at  $-\frac{1}{2} + x, \frac{1}{2} - y, 1 - z$ , the O(5') atom on the reference molecule donating the hydrogen. The same O(5') atom also accepts a hydrogen atom from the O(5') atom on the molecule at  $\frac{1}{2} + x, \frac{1}{2} - y, 1 - z$ . These O-H...O bonds are of length 2.79 Å and they form an infinite chain of bonds about the twofold screw axis.

#### Hydrogen atoms

The covalent bond lengths involving the hydrogen atoms, as determined from the difference Fourier synthesis, all lie in the range 0.85–1.03 Å. As mentioned earlier, all the hydrogen atoms were located in the difference Fourier synthesis. This was an unexpected result because Pruden, Snipes & Gordy (1965) conclude from their electron spin resonance studies of irradiated single crystals of thymidine that the methyl group is rotating about the C(5)-C(7) bond. The fact that the methyl hydrogen atoms are clearly resolved in the difference Fourier shows that in the undamaged molecules in the crystal the methyl group is not rotating.

We wish to acknowledge the assistance given in this investigation by the staff of the Computing Departments of the Universities of Dundee and of St Andrews: also the use of the program of Mr P. Adamson of St Andrews which produced Fig. 6. Thanks are also expressed to Mrs A. Brydie and Miss E. Green for technical assistance. We also wish to acknowledge the financial assistance from the Science Research Council towards the purchase and maintenance of the Hilger and Watts linear diffractometer.

#### References

- DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363.  
 FULLER, W., WILKINS, M. H. F., WILSON, H. R. & HAMILTON, L. D. (1965). *J. Mol. Biol.* **12**, 60.  
 FURBERG, S. & HORDVIK, A. (1956). *Acta Chem. Scand.* **10**, 135.  
 GERDIL, R. (1961). *Acta Cryst.* **14**, 333.

- GORDY, W., PRUDEN, B. & SNIPES, W. (1965). *Proc. U. S. Nat. Acad. Sci.* **53**, 751.
- HARRIS, D. R. & MACINTYRE, W. M. (1964). *Biophys. J.* **4**, 203.
- HASCHEMEYER, A. E. V. & RICH, A. (1967). *J. Mol. Biol.* **27**, 369.
- HASCHEMEYER, A. E. V. & SOBELL, H. M. (1965). *Acta Cryst.* **18**, 525.
- HOOGSTEEN, K. (1963a). *Acta Cryst.* **16**, 28.
- HOOGSTEEN, K. (1963b). *Acta Cryst.* **16**, 907.
- IBALL, J., MORGAN, C. H. & WILSON, H. R. (1966). *Proc. Roy. Soc. A* **295**, 320.
- IBALL, J., MORGAN, C. H. & WILSON, H. R. (1968). *Proc. Roy. Soc. A* **302**, 225.
- JARDETSKY, C. D. (1960). *J. Amer. Chem. Soc.* **82**, 229.
- LANGRIDGE, R., MARVIN, D. A., SEEDS, W. E., WILSON, H. R., HOOPER, C. W., WILKINS, M. H. F. & HAMILTON, L. D. (1960). *J. Mol. Biol.* **2**, 38.
- LIPSON, H. & COCHRAN, W. (1957). *The Determination of Crystal Structures*, p. 170. London: Bell.
- MARVIN, D. A., SPENCER, M., WILKINS, M. H. F. & HAMILTON, L. D. (1961). *J. Mol. Biol.* **3**, 19.
- PRUDEN, B., SNIPES, W. & GORDY, W. (1965). *Proc. U. S. Nat. Acad. Sci.* **53**, 917.
- PULLMAN, B. & PULLMAN, A. (1963). *Quantum Biochemistry*, p. 153. London: Interscience.
- ROGERS, D. (1961). In *Computing Methods and the Phase Problem in X-ray Crystal Analysis*. Oxford: Pergamon Press.
- SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067.
- SPENCER, M. (1959). *Acta Cryst.* **12**, 59.
- SUNDARALINGAM, M. (1966). *Acta Cryst.* **21**, 495.
- SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 930.
- TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1332.
- TOLLIN, P. (1966). *Acta Cryst.* **21**, 613.
- TOLLIN, P., WILSON, H. R. & YOUNG, D. W. (1968). *Nature, Lond.* **217**, 1148.
- TRUEBLOOD, K. N., HORN, P. & LUZZATI, V. (1961). *Acta Cryst.* **14**, 965.
- WATSON, D. C., SUTOR, D. J. & TOLLIN, P. (1965). *Acta Cryst.* **19**, 111.

*Acta Cryst.* (1969). B25, 1432

## The Crystal Structure of 1,3,5-Trithiane: A Refinement\*

By G. VALLE, V. BUSETTI AND M. MAMMI

*Centri di Strutturistica e di Chimica delle Macromolecole del CNR, Istituto di Chimica Organica dell'Università, 35100 Padova, Italy*

AND G. CARAZZOLO

*Centro Ricerche Resine, Montecatini-Edison, 21053 Castellanza (Varese), Italy*

(Received 5 August 1968)

The crystal structure of 1,3,5-trithiane (orthorhombic, space group  $Pmn2_1$ ,  $a=7.668$ ,  $b=7.003$ ,  $c=5.285$  Å;  $Z=2$ ) has been refined by the full-matrix least-squares method with photographic three-dimensional data. Final values of the molecular parameters are: C-S =  $1.814 \pm 0.009$  Å, C-S-C =  $98.9 \pm 0.6^\circ$ , S-C-S =  $114.7 \pm 0.7^\circ$ .  $R$  index, 6.9%. The crystal packing is discussed also in connexion with the observed anisotropic thermal vibrations.

### Introduction

A knowledge of the structure of trithiane was a basic requirement for the investigation of the solid-state polymerization of trithiane (Lando & Stannett, 1964; Carazzolo & Mammi, 1964) and for the correlation of the crystallographic features of the polymer and the monomer (Mammi, Carazzolo & Valle, 1965; Mammi, Valle & Carazzolo, 1968 and references therein). The structure, first determined by Moerman & Wiebenga (1937) and later confirmed by Valle, Carazzolo &

Mammi (1965), has now been refined by three-dimensional data.

### Experimental

The single crystals of trithiane used in the analysis were obtained by sublimation at  $180^\circ\text{C}$  in the form of long transparent prisms elongated along the  $c$  axis and with nearly square cross-sections. Cell parameters were determined on Weissenberg and precession photographs, with Cu  $K\alpha$  and Mo  $K\alpha$  radiations.

### Crystal data

Trithiane,  $\text{C}_3\text{H}_6\text{S}_3$ , F.W. 138.3, m.p.  $215^\circ\text{C}$ . Orthorhombic, space group  $Pmn2_1-C_{2v}^7$ .  $a=7.668 \pm 0.005$ ,  $b=7.003 \pm 0.005$ ,  $c=5.285 \pm 0.005$  Å,

\* Work carried out with CNR aid. A report has been given at the VIIth International Congress of I.U.Cr., Moscow, 1966, paper 8-66.