

## The Crystal Structure of 5-Ethyl-6-methyluracil\*

BY GEORGE N. REEKE, JR.† AND RICHARD E. MARSH

*Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California, U.S.A.*

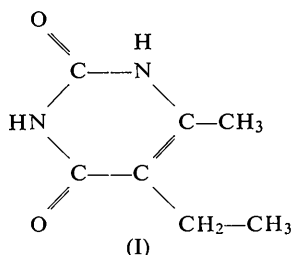
(Received 2 July 1965)

The crystal structure of 5-ethyl-6-methyluracil has been determined and refined by three-dimensional least-squares methods. The final *R* index for 1369 observable reflections is 0.080, and the standard deviations in the positions of the C, N, and O atoms are in the range 0.002–0.005 Å.

The crystals are triclinic, space group  $P\bar{1}$ , with  $a = 7.478$ ,  $b = 11.803$ ,  $c = 4.776$  Å,  $\alpha = 99.39$ ,  $\beta = 107.65$ ,  $\gamma = 92.14^\circ$ ; there are two molecules per unit cell. The bond distances and angles are very close to those in thymine monohydrate. The pyrimidine ring is significantly non-planar, being buckled about the C(2)–C(5) line apparently to relieve strain between the substituent groups. Two pairs of N–H  $\cdots$  O hydrogen bonds, 2.78 and 2.82 Å in length and arranged about centers of symmetry, hold the molecules together to form chains along the [102] direction.

### Introduction

Current interest in the structures of the nucleic acids and related compounds led one of us (GNR) to select a pyrimidine derivative for his first attempt at an X-ray crystal analysis. A large number of organic compounds prepared in the laboratories of the late Professor Treat B. Johnson have been made available to us by Professor James English of Yale University. From among these compounds, 5-ethyl-6-methyluracil (I) was selected for study.



Preliminary crystallographic investigation on this compound was carried out by Gerdil (1960).

### Experimental

The material was available in crystalline form. Typical crystals were white needles about 1 mm in length and 0.1 mm in diameter; many were poorly formed or bent. Recrystallization from water at room temperature yielded similar crystals. Preliminary oscillation and Weissenberg photographs indicated that the crystals are triclinic with cell dimensions close to those reported by Gerdil (1960) for his modification '(a)'. They also showed that most of the crystals were twins, with the twin plane parallel to the needle axis *c*.

\* Contribution No. 3269 from the Gates and Crellin Laboratories of Chemistry. This work was supported, in part, by a grant (G-21608) from the National Science Foundation, U.S.A.

† Present address: Department of Chemistry, Harvard University, Cambridge, Massachusetts, U.S.A.

Two different crystals were used in this investigation. One was mounted about its needle (*c*) axis; the second was cleaved into a roughly equi-dimensional rhomb, mounted and oriented about an axis which turned out to be *b*, and subsequently remounted and oriented about [0 $\bar{1}$ 1]. Unit-cell dimensions were obtained from a least-squares treatment of measurements made on zero-layer Weissenberg photographs about these three axes; the photographs were prepared in a special camera with the film held in the asymmetric position. The resulting cell dimensions, together with those reported by Gerdil (1960), are given in Table 1. The density calculated on the basis of two molecules per unit cell is 1.297 g.cm<sup>-3</sup>; the approximate value observed by flotation in dioxane–chloroform mixtures is 1.25 g.cm<sup>-3</sup>. The space group  $P\bar{1}$  was assumed.

Table 1. *Unit-cell dimensions and their estimated standard deviations (in parentheses)*

	This investigation	Gerdil (1960)
<i>a</i>	7.478(1) Å	7.45 Å
<i>b</i>	11.803(1)	11.7
<i>c</i>	4.776(1)	4.83
$\alpha$	99.39(2)°	99.91°
$\beta$	107.65(1)	108
$\gamma$	92.14(1)	92.5

Intensities were estimated visually on multiple-film equi-inclination Weissenberg photographs of layer lines 0–4 about *c* and of the zero layer lines about the *b* and [0 $\bar{1}$ 1] axes. The processing of these intensities – including the evaluation of film and pack factors, correction for Lorentz and polarization effects, correlation of the measurements from the different photographs, and statistical evaluation of the standard deviations in  $F^2$  – was carried out by means of a series of programs operating under the CRYRM system on an IBM 7094 computer (Duchamp, 1964).

Of a total of about 1850 reflections within the effective sphere of copper radiation, 1583 were recorded; of these, 1369 were strong enough to be observed.

### Determination and refinement of the structure

In view of the short  $c$  axis, the initial structure determination was based on the  $hk0$  reflections. Sharpened and unsharpened Patterson projections indicated the orientation of the pyrimidine ring, and hydrogen-bond and packing considerations soon led to a reasonable trial structure. Two-dimensional refinement by Fourier and least-squares methods, including individual isotropic temperature factors for the heavy atoms, led to a final  $R$  index of 0.17 for 187  $hk0$  reflections. Three-dimensional refinement was then initiated.

Refinement was by full-matrix least squares, using the triclinic sub-program of the CRYRM system for the IBM 7094 computer (Duchamp, 1964). The quantity minimized was  $\sum w(F_o^2 - F_c^2/K^2)^2$ , where  $K$  is the scale parameter; atomic form factors were taken from *International Tables for X-ray Crystallography* (1962). Initial  $z$  coordinates were assigned assuming a planar molecule lying in the  $\bar{2}11$  plane, which was represented by a reflection having approximately twice the intensity of any other; after a few refinement cycles the  $z$  coordinate of C(10) was reassigned on the basis of reasonable values for the C(9)–C(10) distance and the C(5)–C(9)–C(10) angle.

During the early refinement cycles the atoms were assigned isotropic temperature factors and the weighting scheme was that proposed by Hughes (1941), modified to reflect the expected uncertainties in  $F_o^2$  rather than in  $F_o$ . When the  $R$  index levelled off at 0.19, anisotropic temperature parameters were introduced, and three least-squares cycles reduced  $R$  to 0.12. Difference Fourier maps were then calculated in the planes of the pyrimidine ring and of the methylene and two methyl groups, from which preliminary coordinates of the hydrogen atoms were derived.

Refinement was concluded with a number of full-matrix least-squares cycles in which 130 parameters – coordinates of all 21 atoms, anisotropic temperature parameters of the 11 heavy atoms, and a scaling factor – were adjusted; the hydrogen atoms were assigned constant isotropic temperature factors with  $B=4.0$ . Weights were assigned according to the expression

$$w = (0.0875 + 0.1326F_o + 0.0730F_o^2)^{-2},$$

which was derived during the data-reduction calculations (see *Experimental*) as the best quadratic representation of the reciprocals of the observational variances,  $[\sigma(F_o^2)]^{-2}$ . Unobserved reflections were included only if the calculated value of  $F$  exceeded the observational threshold. A few strong reflections were given zero weight because of presumed extinction effects. During the last refinement cycle, no heavy-atom parameter shifted as much as 1/10 of its standard deviation; the maximum shift in a hydrogen-atom coordinate was approximately  $0.7\sigma$ .

The final  $R$  index for 1368 reflections of non-zero weight is 0.080, and the 'goodness of fit'

$$\left( \frac{\sum w(F_o^2 - F_c^2)^2}{m-s} \right)^{\frac{1}{2}},$$

is 1.8. This latter value is significantly larger than the expectation value 1.0, apparently indicating systematic errors in the data. We do not know the source of these errors.

Table 2. *The final parameters and their standard deviations (in parentheses)*

All values have been multiplied by  $10^4$ . The temperature factor is of the form  $\exp[-(b_{11}h^2 + b_{22}k^2 + b_{33}l^2 + b_{12}hk + b_{13}hl + b_{23}kl)]$ . The hydrogen atoms were assigned isotropic temperature factors with  $B=4.0$ .

	x	y	z	$b_{11}$	$b_{22}$	$b_{33}$	$b_{12}$	$b_{13}$	$b_{23}$
N(1)	44.67(3)	1086(2)	761.7(5)	164(4)	72(2)	360(12)	-2(4)	8(10)	50(6)
C(2)	3086(3)	021.7(2)	6181(5)	168(4)	70(2)	343(12)	13(4)	62(10)	48(6)
N(3)	1839(3)	0409(2)	3592(4)	171(4)	67(1)	370(12)	-12(4)	3(10)	44(6)
C(4)	1822(3)	1414(2)	2459(5)	166(4)	67(2)	374(13)	16(4)	78(11)	46(6)
C(5)	3267(3)	2537(2)	4200(5)	176(5)	66(2)	448(14)	9(4)	119(12)	38(7)
C(6)	4584(3)	2119(2)	6666(5)	168(4)	70(2)	423(13)	0(4)	96(11)	26(7)
O(7)	2986(2)	-0687(1)	7110(4)	207(4)	75(1)	429(11)	-18(3)	-16(9)	95(5)
O(8)	0650(2)	1480(1)	0059(4)	209(4)	75(1)	428(10)	-6(3)	-52(9)	94(5)
C(9)	3171(4)	3459(2)	3081(7)	226(6)	72(2)	549(17)	-14(5)	112(14)	95(8)
C(10)	1766(6)	4196(3)	4018(12)	343(10)	81(3)	1119(32)	91(8)	393(27)	172(14)
C(11)	6280(4)	2914(3)	8575(8)	198(6)	87(2)	628(20)	-4(6)	9(15)	42(9)
H(12)	660(5)	283(3)	1073(10)						
H(13)	605(6)	367(4)	892(9)						
H(14)	717(6)	275(3)	802(9)						
H(15)	441(6)	382(3)	370(9)						
H(16)	296(5)	335(3)	085(10)						
H(17)	232(6)	444(3)	627(10)						
H(18)	150(6)	480(4)	291(9)						
H(19)	050(6)	364(3)	331(9)						
H(20)	529(6)	095(3)	918(11)						
H(21)	096(6)	-019(4)	243(9)						

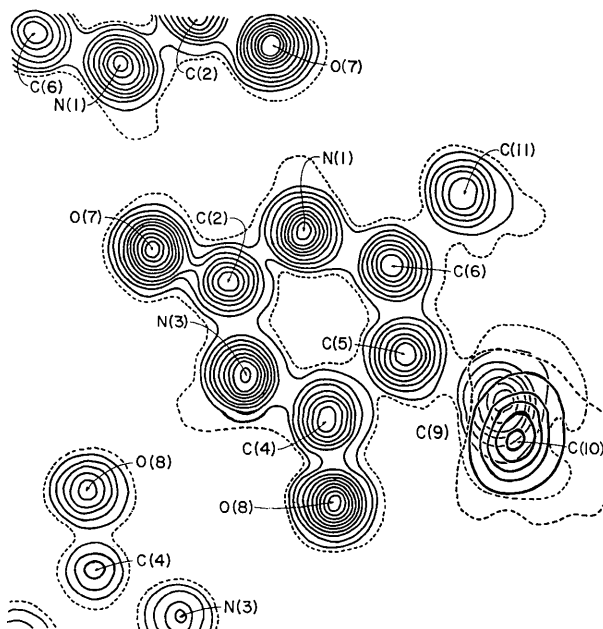


Fig. 1. The electron density, evaluated at the conclusion of the refinement, in the plane of the uracil ring and in a parallel plane passing through C(10). Contours are at intervals of  $1 \text{ e.}\text{\AA}^{-3}$  beginning with  $1 \text{ e.}\text{\AA}^{-3}$  (dashed).

Table 3. Observed and calculated structure factors

Within each group are values of h, 10F<sub>o</sub>, and 10F<sub>c</sub>. Minus signs preceding values of 10F<sub>o</sub> should be read 'less than'; asterisks indicate reflections omitted from the final least-squares cycles.

Table with multiple columns containing numerical data for structure factors. The table is organized into groups, with each group containing values for h, 10F<sub>o</sub>, and 10F<sub>c</sub>. The data is presented in a grid-like format across the page.

The final parameters of the heavy atoms are listed in Table 2; observed and calculated structure factors are listed in Table 3. In Fig. 1 is shown the electron density, evaluated at the conclusion of the refinement, in the plane of the uracil ring and in a parallel plane passing through C(10). Difference syntheses, for which the contributions of the hydrogen atoms were omitted from  $F_o$ , were evaluated in the planes of the uracil ring, of the methylene hydrogen atoms, and of the two sets of methyl hydrogen atoms; they are shown in Fig. 2.

### Discussion

#### The geometry of the molecule

The standard deviations in the atomic coordinates (Table 2) correspond to positional uncertainties of about 0.002 Å for atoms N(1)–O(8), 0.003 Å for C(9) and C(11), and about 0.005 Å for C(10). Standard deviations in distances between heavy atoms, then, are in of the neighborhood of 0.005 Å. In view of uncertainties in the pattern of thermal motions, this value should probably be increased to about 0.01 Å for bonds involving non-ring atoms. Standard deviations in the bond angles are about 1°.

The bond distances and angles involving the heavy atoms are shown in Fig. 3. Values in parentheses include estimated corrections for the effects of thermal motion; these corrections were calculated on the basis of a riding model (Busing & Levy, 1964), the atom with the larger temperature motions being assumed to ride on the atom with the smaller motions.

Distances and angles involving the hydrogen atoms are given in Table 4.

The dimensions of the thymine portion of the molecule (atoms 1–9) agree very closely with those found by Gerdil (1961) in thymine monohydrate. The maximum difference in bond distances between the two investigations, using values uncorrected for thermal motion, is 0.010 Å and the average difference is 0.005 Å – well within the experimental uncertainties. Moreover, the anisotropic thermal parameters reported by Gerdil (his Table 5) suggest that, had he estimated corrections to the bond distances, they would have been very similar to the corrections we have estimated for 5-ethyl-6-methyluracil.

Deviations of various atoms from the mean plane of the six-membered ring are given in Table 5. The deviations of O(8), C(9), and C(11) are relatively large and of

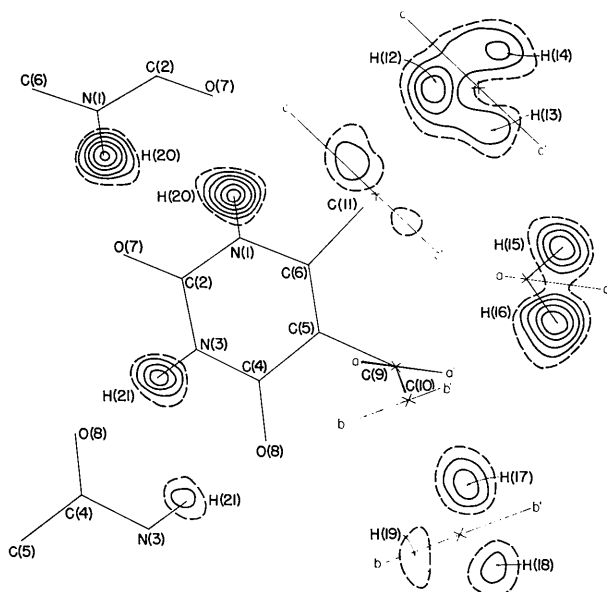


Fig. 2. The ' $F_o - F_c$ ' electron density evaluated in the planes of the uracil ring and of the hydrogen atoms of the methyl and methylene groups. Values of  $F_c$  were calculated from the final parameters (Table 2) but with the contributions of the hydrogen atoms omitted. Contours are at intervals of 0.1 e.Å<sup>-3</sup> beginning with 1 e.Å<sup>-3</sup> (dashed).

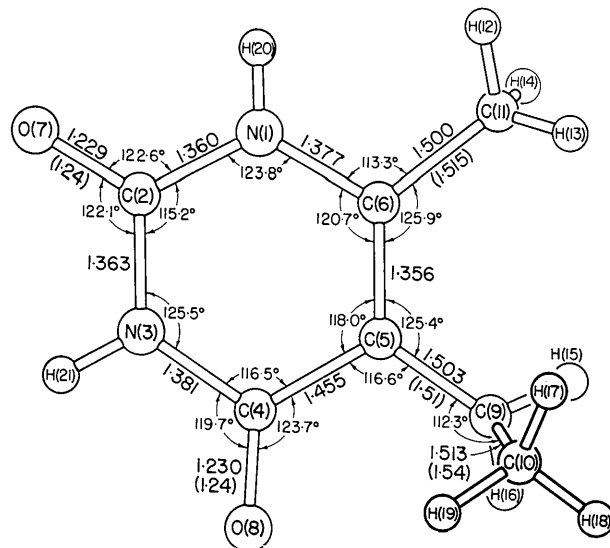


Fig. 3. Bond distances and angles. Values in parentheses include an approximate correction for the effects of thermal motion.

Table 4. Bond distances and angles involving hydrogen atoms

N(1)–H(20)	0.86 Å	C(2)–N(1)–H(20)	115°	C(6)–N(1)–H(20)	121°
N(3)–H(21)	0.93	C(2)–N(3)–H(21)	118	C(4)–N(3)–H(21)	117
C(9)–H(15)	0.94	C(5)–C(9)–H(15)	107	C(10)–C(9)–H(15)	115
C(9)–H(16)	1.02	C(5)–C(9)–H(16)	113	C(10)–C(9)–H(16)	112
		H(15)–C(9)–H(16)	97		
C(10)–H(17)	1.01	C(9)–C(10)–H(17)	105	H(19)–C(10)–H(17)	115
C(10)–H(18)	0.95	C(9)–C(10)–H(18)	111	H(17)–C(10)–H(18)	116
C(10)–H(19)	1.05	C(9)–C(10)–H(19)	104	H(18)–C(10)–H(19)	106
C(11)–H(12)	1.01	C(6)–C(11)–H(12)	111	H(14)–C(11)–H(12)	109
C(11)–H(13)	0.91	C(6)–C(11)–H(13)	114	H(12)–C(11)–H(13)	94
C(11)–H(14)	0.81	C(6)–C(11)–H(14)	109	H(13)–C(11)–H(14)	118

alternating sense, suggesting considerable crowding of the side groups. The crowding is relieved not by a simple bending of the C(4)–O(8), C(5)–C(9), and C(6)–C(11) bonds out of the plane of the ring atoms but rather by a folding of the ring itself along the C(2)–C(5) line. As a result of this folding, the ring is significantly non-pla-

nar but the three bonds about each of the atoms C(2), C(4), C(5), and C(6) remain approximately coplanar. The amount of the folding – that is, the dihedral angle between the mean planes of C(2), N(3), C(4), and C(5) and of N(1), C(2), C(5), and C(6) – is  $3^\circ$ .

Table 5. Deviations from the least-squares plane of the six-membered ring

The direction cosines of the plane normal relative to the crystallographic axes are:  $q_a = -0.7999$ ;  $q_b = 0.3071$ ;  $q_c = 0.6924$ . The origin-to-plane distance is  $0.250 \text{ \AA}$ .

Atom	Deviation	Atom	Deviation
N(1)	-0.010 Å	C(10)	1.543 Å
C(2)	0.026	H(12)	0.37
N(3)	-0.014	H(13)	0.42
C(4)	-0.014	H(14)	-0.90
C(5)	0.031	H(15)	-0.28
C(6)	-0.020	H(16)	-0.53
O(7)	0.066	H(20)	-0.04
O(8)	-0.083	H(21)	-0.10
C(9)	0.125	O(7')*	0.065
C(11)	-0.115	O(8')*	-0.416

\* Atoms O(7') and O(8') are hydrogen-bond acceptors of neighboring molecules.

#### Packing of the molecules

The structure viewed along the  $c$  axis is shown in Fig. 4. Molecules are held together by N–H $\cdots$ O hydrogen bonds to form chains along the [102] direction. These hydrogen bonds, which occur in pairs about centers of symmetry, are of length  $2.82 \text{ \AA}$  [N(3)–O(8)] and  $2.78 \text{ \AA}$  [N(1)–O(7)]. In thymine monohydrate (Gerdil, 1961) the two N $\cdots$ O distances are significantly longer –  $2.84 \text{ \AA}$ ; moreover, one oxygen atom [O(7)] is the acceptor for both N–H $\cdots$ O bonds, the other oxygen atom [O(8)] accepting a hydrogen bond from the water molecule of crystallization.

Other intermolecular distances are given in Table 6. The hydrogen atoms of the methyl groups form interesting interlocking arrays along the plane  $y = \frac{1}{2}$  (see Fig. 4); presumably as a result of this interlocking, there appears to be little or no rotation of the methyl groups (see Fig. 2).

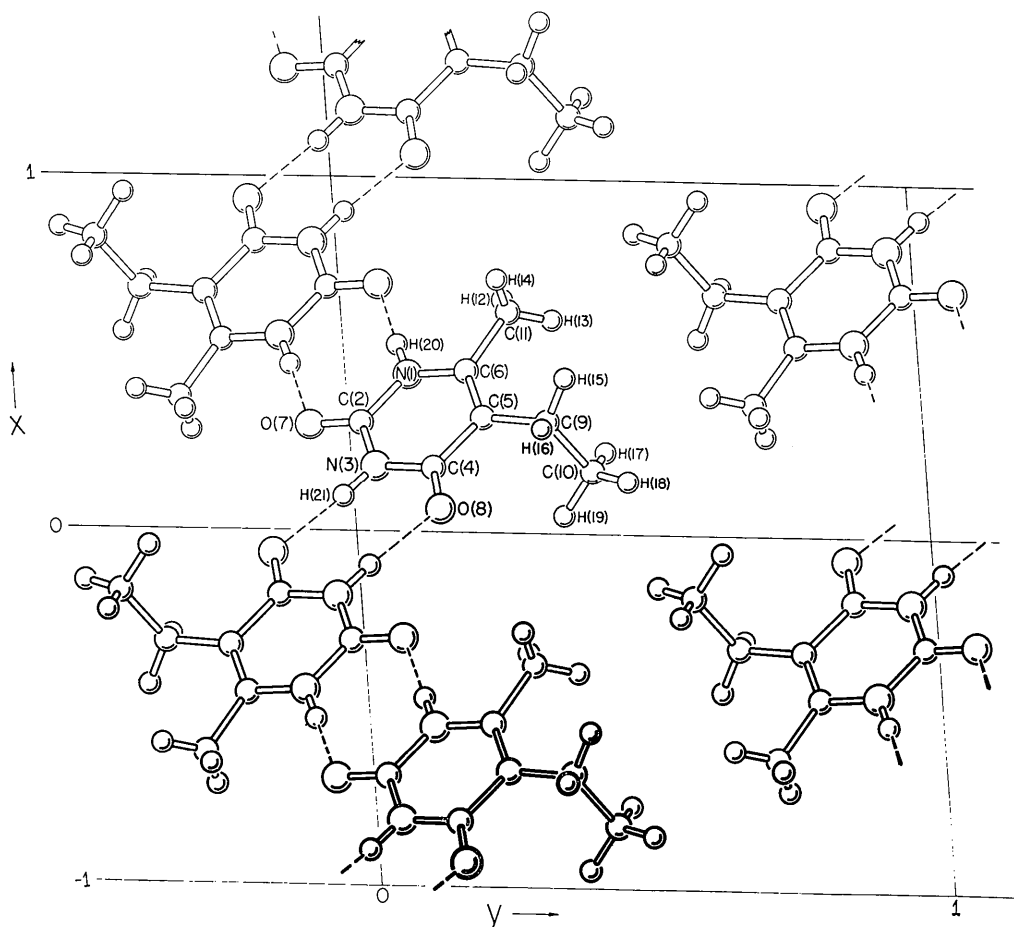


Fig. 4. The structure viewed along the  $c$  axis.

Table 6. *The shortest intermolecular distances*

All distances less than 3.6 Å between pairs of heavy atoms and less than 3.0 Å when a hydrogen atom is involved are included.

From	to	in molecule at		
C(2)	O(8)	$x,$	$y, 1+z$	3.21
N(1)	O(8)			3.42
N(1)	C(4)			3.46
O(7)	O(8)			3.50
C(2)	C(4)			3.51
O(7)	N(3)			3.52
N(3)	O(8)			3.52
C(6)	H(16)			2.87
H(17)	H(16)			2.64
H(12)	H(15)			2.66
H(13)	H(16)			2.76
H(12)	H(16)			2.83
H(13)	H(15)			2.88
H(13)	H(17)	$1-x,$	$1-y, 2-z$	2.86
H(15)	H(15)	$1-x,$	$1-y, 1-z$	2.85
H(13)	H(18)			2.94
H(19)	H(12)	$x-1,$	$y, z-1$	2.86
H(19)	H(14)			2.98
N(1)	C(2)	$1-x,$	$-y, 1-z$	3.21
O(7)	C(6)			3.26
O(7)	N(1)			3.36
N(1)	O(7)			3.36
C(2)	C(2)			3.41
N(3)	N(1)			3.48
N(1)	N(1)			3.57
C(10)	C(10)	$-x,$	$1-y, 1-z$	3.57

#### The temperature factors

The magnitudes and direction cosines, relative to the crystallographic axes, of the ellipsoids of thermal motion implied by the anisotropic temperature factors of Table 2 are given in Table 7. The anisotropies of the ring atoms appear to be due primarily to a translational or lattice vibration with maximum amplitude perpendicular to the plane of the ring and minimum along the [102] direction – that is, parallel to the hydrogen-bonded chains. The non-ring atoms, O(7) to C(11), have additional out-of-plane amplitudes and C(10) is also librating approximately perpendicular to the C(9)–C(10) bond.

Table 7. *The magnitudes and direction cosines, relative to the crystallographic axes, of the temperature factor ellipsoids*

	Axis $i$	$B_i$	$q_i^a$	$q_i^b$	$q_i^c$
N(1)	1	4.51	-0.846	0.294	0.646
	2	3.85	-0.197	-0.955	0.402
	3	2.37	0.496	0.048	0.649
C(2)	1	3.98	-0.928	0.007	0.634
	2	3.83	0.027	-1.000	0.160
	3	2.55	0.371	-0.010	0.756
N(3)	1	4.74	-0.864	0.247	0.650
	2	3.62	-0.096	-0.953	0.451
	3	2.39	0.493	0.173	0.612
C(4)	1	3.95	-0.856	-0.242	0.717
	2	3.62	0.262	-0.970	0.004
	3	2.76	0.445	0.021	0.697
C(5)	1	4.14	-0.700	-0.290	0.867
	2	3.75	0.525	-0.857	0.136
	3	3.18	0.480	0.427	0.480
C(6)	1	4.12	-0.444	-0.649	0.812
	2	3.97	-0.732	0.682	0.282
	3	2.92	0.516	0.337	0.511
O(7)	1	5.99	-0.838	0.336	0.625
	2	3.69	-0.291	-0.940	0.349
	3	2.76	0.462	-0.065	0.698
O(8)	1	6.22	-0.855	0.222	0.678
	2	3.88	-0.225	-0.965	0.225
	3	2.58	0.467	-0.138	0.700
C(9)	1	5.71	-0.823	0.321	0.656
	2	4.01	0.396	-0.338	0.741
	3	3.54	-0.408	-0.885	0.140
C(10)	1	9.01	-0.156	0.006	0.975
	2	7.07	-0.943	-0.261	0.199
	3	3.86	0.293	-0.965	0.121
C(11)	1	6.62	-0.730	0.195	0.813
	2	5.30	0.194	-0.885	0.504
	3	2.97	0.655	0.423	0.293

The work described here was carried out as an undergraduate research project (GNR).

#### References

- BUSING, W. R. & LEVY, R. A. (1964). *Acta Cryst.* **17**, 142.  
 DUCHAMP, D. J. (1964). *Program and Abstracts*, Amer. Cryst. Assoc. Meeting, Bozeman, Montana; paper B-14.  
 Gerdil, R. (1960). *Acta Cryst.* **13**, 165.  
 Gerdil, R. (1961). *Acta Cryst.* **14**, 333.  
 HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.  
*International Tables for X-Ray Crystallography* (1962). Vol. III, p. 202. Birmingham: Kynoch Press.