

The valence density of urea at 123 K has been reported in a previous paper by Mullen & Hellner (1978). A comparison with the thiourea molecules described here shows that the peak height at N in urea ($3.75 \text{ e } \text{Å}^{-3}$) is very close to the values found for both thiourea molecules (3.62 and $3.69 \text{ e } \text{Å}^{-3}$). In contrast, the peak height at O in urea ($7.0 \text{ e } \text{Å}^{-3}$) is considerably higher than those at S in thiourea. The mean C—S bond population (3.03 e) may be compared with a value of 3.27 e for C—O in urea.

The C—N bond populations (e) compare very closely: 1.89 in urea; 1.85 and 1.92 in thiourea (Table 2). The N—H bond populations are lower in thiourea (1.15 e) than in urea (1.33 e), mean values being quoted in both cases. The populations of the residual charge at the protons show the opposite trend with mean values of 0.67 and 0.39 e for thiourea and urea respectively.

Lone-pair populations at S in thiourea are much lower (mean = 2.88 e) than at O in urea (4.16 e). The N lone pair also has a lower population in thiourea (1.64 e) than in urea (2.36 e).

It is intended in future papers in this series to publish experimental deformation densities and to compare these with quantum chemically derived densities. Additionally, form factors for bond charges obtained from the experimental model (Scheringer, Kutoglu, Mullen & Hellner, 1978) will be compared with those of Fritchie (1966) and Cromer & Larson (1974).

References

- CROMER, D. T. & LARSON, A. C. (1974). *J. Chem. Phys.* **60**, 176–184.
 DVORYANKINA, V. F. & VAINSHTEIN, B. K. (1960). *Kristallografiya*, **5**, 589–599.
 ELCOMBE, M. M. & TAYLOR, J. C. (1968). *Acta Cryst.* **A24**, 410–420.
 FINGER, L. (1972). A Fortran IV computer program for structure factor calculation and least-squares refinement of crystal structures. Carnegie Institution of Washington.
 FRITCHIE, C. J. (1966). *Acta Cryst.* **20**, 27–36.
 GOLDSMITH, G. J. & WHITE, J. G. (1959). *J. Chem. Phys.* **31**, 1175–1187.
 HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502–510.
 HELLNER, E. (1977). *Acta Cryst.* **B33**, 3813–3816.
International Tables for X-ray Crystallography (1952). Vol. 1. Birmingham: Kynoch Press.
International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press.
 JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.
 MULLEN, D. & HELLNER, E. (1977). *Acta Cryst.* **B33**, 3816–3823.
 MULLEN, D. & HELLNER, E. (1978). *Acta Cryst.* **B34**, 1624–1627.
 REES, B. (1976). *Acta Cryst.* **A32**, 483–488.
 SCHERINGER, C., KUTOGLU, A., MULLEN, D. & HELLNER, E. (1978). *Acta Cryst.* **A34**, 475–476.
 STEWART, R. F. (1976). *Acta Cryst.* **A32**, 565–574.
 TRUTER, M. R. (1967). *Acta Cryst.* **22**, 556–559.

Acta Cryst. (1978). **B34**, 2794–2798

The Crystal and Molecular Structure of 2',3',5'-Tri-*O*-acetyl-6-*O*-(mesitylenesulphonyl)guanosine

BY S. NEIDLE AND A. ACHARI

Department of Biophysics, University of London, King's College, 26–29 Drury Lane, London WC2B 5RL, England

G. M. SHELDRIK

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

AND C. B. REESE AND P. K. BRIDSON

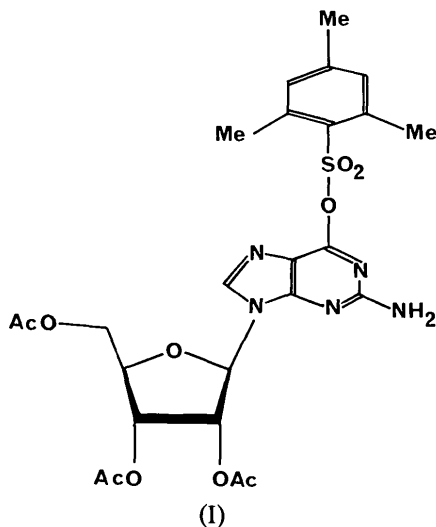
Department of Chemistry, University of London, King's College, Strand, London WC2R 2LS, England

(Received 22 March 1978; accepted 21 April 1978)

The structure of a 6-*O*-mesitylenesulphonyl derivative of 2',3',5'-tri-*O*-acetylguanosine, $\text{C}_{24}\text{H}_{29}\text{N}_5\text{O}_{10}\text{S}$, has been determined by X-ray diffraction. Crystals are monoclinic, $a = 26.370(4)$, $b = 8.200(2)$, $c = 17.991(3)$ Å, $\beta = 132.77(4)^\circ$. The solution of the structure in space group *C2* was not straightforward and is described in detail. Refinement converged at $R = 0.110$ for 1102 observed reflections. The guanine base displays some deviations from its usual geometry due to the loss of C(6)—O(6) double-bond character. The ribose sugar is C(2')-*endo* puckered.

Introduction

The title compound (I) was obtained by treatment of guanosine (fully protected by acetyl groups on the sugar hydroxyl positions), with mesitylenesulphonyl chloride (Bridson, Markiewicz & Reese, 1977). The proposed structure, with the O(6) atom of guanine rather than N(2) being the site of attack, was verified by chemical and spectroscopic means. The crystallographic analysis reported here provides definitive confirmation for this assignment. The O(6) oxygen is a favoured site for the binding of various mutagenic agents to guanine in nucleic acids.



Experimental

The title compound was crystallized from aqueous solution as elongated, extremely thin fibrous platelets. These showed parallel extinction under the polarizing microscope. Photographs revealed the crystals to be monoclinic. Systematic absences $hkl: h + k = 2n + 1$ and $0k0: k = 2n + 1$, together with the known optical activity of the compound, uniquely identified the space group as $C2$. Cell dimensions were obtained from diffractometer measurements.

Crystal data

$C_{24}H_{29}N_5O_{10}S$, $M_r = 579.0$, monoclinic plates, elongated along b . $a = 26.370$ (4), $b = 8.200$ (2), $c = 17.991$ (3) Å, $\beta = 132.77$ (4)°, $V = 2855.8$ Å³, $D_m = 1.33$ (by flotation), $D_c = 1.346$ g cm⁻³ for $Z = 4$. $F(000) = 1376$, $\mu = 14.9$ cm⁻¹ for Cu $K\alpha$ radiation ($\lambda_{K\alpha} = 1.54178$ Å). Space group $C2$ (No. 5, C_2^2).

Intensities were collected on an Enraf-Nonius CAD-4 automatic diffractometer, to a θ limit of 60° for Cu $K\alpha$ radiation. A total of 1597 reflections were

measured, of which only 1102 were considered to be observable, with $F > 2\sigma(F)$.

Structure solution and refinement

Initial attempts to solve the structure by conventional direct methods, heavy-atom location or search procedures, were all unsuccessful. An unconventional direct-methods approach, detailed below, eventually led to the correct solution.

Table 1. Final positional parameters ($\times 10^4$) for the non-hydrogen atoms, and isotropic (U) temperature factors (Å² $\times 10^3$) with estimated standard deviations in parentheses

	x	y	z	U
C(16)	3959 (7)	6529 (16)	9100 (9)	31 (5)
C(26)	3967 (7)	6090 (16)	8359 (9)	41 (5)
C(36)	4044 (7)	7285 (16)	7890 (9)	48 (6)
C(46)	4112 (7)	8919 (16)	8161 (9)	49 (6)
C(56)	4104 (7)	9358 (16)	8902 (9)	56 (6)
C(66)	4027 (7)	8163 (16)	9371 (9)	42 (6)
C(76)	3891 (13)	4321 (30)	8025 (17)	67 (7)
C(86)	4145 (13)	10238 (34)	7572 (17)	74 (7)
C(96)	4066 (12)	8878 (32)	10157 (15)	63 (7)
O(1')	2165 (6)	-3111 (18)	6795 (9)	39 (4)
C(1')	2016 (9)	-2352 (28)	7309 (14)	40 (5)
C(2')	1228 (8)	-2531 (23)	6633 (12)	20 (4)
O(2')	1055 (6)	-2370 (19)	7232 (9)	44 (4)
C(21)	447 (11)	-1834 (30)	6809 (17)	56 (6)
O(21)	7 (9)	-1345 (26)	5948 (13)	95 (6)
C(22)	313 (12)	-1755 (34)	7496 (17)	70 (7)
C(3')	1107 (9)	-4290 (23)	6203 (12)	28 (5)
O(3')	1291 (6)	-5442 (19)	6946 (9)	44 (4)
C(31)	777 (11)	-5957 (28)	6847 (15)	50 (6)
O(31)	166 (8)	-5462 (22)	6198 (11)	78 (5)
C(32)	1032 (12)	-7162 (34)	7708 (17)	70 (7)
C(4')	1666 (10)	-4391 (26)	6169 (13)	38 (5)
C(5')	1330 (10)	-4086 (29)	5080 (14)	47 (6)
O(5')	1877 (7)	-4013 (20)	5093 (9)	52 (4)
C(51)	2000 (9)	-2626 (29)	4857 (13)	42 (5)
O(51)	1616 (9)	-1452 (27)	4561 (12)	91 (6)
C(52)	2582 (12)	-2726 (34)	4914 (19)	71 (7)
N(1)	2282 (7)	2774 (21)	9516 (10)	29 (4)
C(2)	2676 (10)	1290 (27)	9572 (14)	38 (5)
N(2)	2689 (8)	1140 (23)	10337 (11)	51 (5)
N(3)	2449 (7)	32 (20)	8959 (10)	32 (4)
C(4)	2452 (8)	358 (22)	8245 (11)	21 (4)
C(5)	2637 (9)	1754 (24)	8084 (13)	29 (5)
C(6)	2865 (10)	2972 (27)	8769 (15)	42 (6)
O(6)	3102 (6)	4464 (17)	8760 (8)	38 (4)
N(7)	2533 (8)	1702 (23)	7212 (11)	39 (5)
C(8)	2272 (10)	288 (29)	6848 (14)	40 (5)
N(9)	2210 (7)	-636 (19)	7445 (10)	27 (4)
S	3885 (3)	5062	9708 (4)	*
O(16)	3883 (9)	5704 (23)	10429 (10)	*
O(26)	4282 (8)	3674 (25)	9964 (11)	*

* The anisotropic thermal parameters for these atoms are:

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
S	49 (3)	57 (4)	32 (3)	18 (3)	21 (3)	0 (4)
O(16)	146 (16)	87 (13)	43 (9)	-34 (9)	74 (11)	-49 (12)
O(26)	59 (11)	98 (15)	76 (12)	35 (12)	37 (10)	1 (13)

A large starting set was constructed containing some negative sign information. Four $k = 0$ reflections were assigned phase angles (φ) of 0° , and four more were assigned $\varphi = 180^\circ$, all on the basis of Σ_1 relations and seminvariant relations derived by combining triplet and quartet phase relationships (Sheldrick, 1976). Two trio relationships were added, each of one positive and two negative reflection quartets involving only $k = 0$ reflections. The total starting set included 18 $k = 0$ reflections and one $k = 1$ reflection (to fix the origin along b), plus six general reflections which were assigned the 'magic integer' phases $33x$, $49x$, $57x$, $61x$, $63x$ and $64x$ (Main, 1977). After weighted tangent-formula expansion to 250 of the largest E values, the figures of merit $NQEST$ (De Titta, Edmonds, Langs & Hauptman, 1975) and R_α (Roberts, Pettersen, Sheldrick, Isaacs & Kennard, 1973) were calculated. The 'best' solutions were expanded to 461 E values > 1.1 . One E map (ranked 25th on $NQEST$ and 2nd on R_α) revealed a fragment (O_2S-O-) attached to a six-membered aromatic ring, with a further atom attached *ortho* to the ring. Tangent expansion (Karle, 1968) was attempted, but did not develop the structure further.

Table 2. Final positional parameters for the hydrogen atoms

Each hydrogen atom was assigned a temperature factor (U) 0.005 Å² greater than the atom to which it is attached. The numbering scheme follows that of the non-hydrogen atom to which each hydrogen is attached.

	x	y	z
H(36)	405	694	732
H(56)	416	1062	911
H(761)	392	408	746
H(762)	340	391	775
H(763)	430	368	872
H(861)	416	980	702
H(862)	462	1082	818
H(863)	372	1108	721
H(961)	412	1016	1034
H(962)	454	825	1077
H(963)	365	837	1006
H(1')	227	-294	803
H(2')	94	-161	605
H(221)	-16	-110	712
H(222)	73	-119	822
H(223)	26	-300	762
H(3')	57	-443	551
H(321)	63	-769	765
H(322)	136	-641	838
H(323)	134	-812	776
H(4')	191	-557	647
H(5'1)	108	-290	483
H(5'2)	95	-501	457
H(521)	276	-154	491
H(522)	230	-329	418
H(523)	302	-347	549
H(21)	224	125	1025
H(22)	309	45	1101
H(8)	218	-48	628

Since it was suspected that the fragment was correctly oriented but wrongly positioned with respect to the space-group origin, a tangent formula expansion was performed in space group $C1$, with the data expanded to a full hemisphere. Two iterations revealed two (O_2S-O-) groups, each attached to an almost complete guanine ring, related by a twofold axis at $x = 0.178$, $z = 0.166$. Coordinates were calculated for space group $C2$ with the origin shifted to this position, and three cycles of tangent expansion with E values > 0.9 developed the complete structure.

The structure was refined by full-matrix least-squares techniques. In view of the paucity of observable reflection data, the benzene ring of the mesitylene group was treated as a rigid body, and only the S atoms and two O atoms were refined anisotropically. The positions of all hydrogen atoms were calculated, and their structure-factor contributions taken into account during refinement. The final R was 0.110, and R_w was 0.100, where $w = [0.8303/\sigma^2(F_o) + 0.005F_o^2]$.

All calculations were performed with programs written by GMS; diagrams were drawn with the aid of the program *PLUTO* (Motherwell, 1975).

Tables 1 and 2 list the final positional and thermal parameters.*

Discussion

Interatomic bond lengths and angles are given in Table 3; the numbering scheme used is shown in Fig. 1. Due

* A list of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33575 (6 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

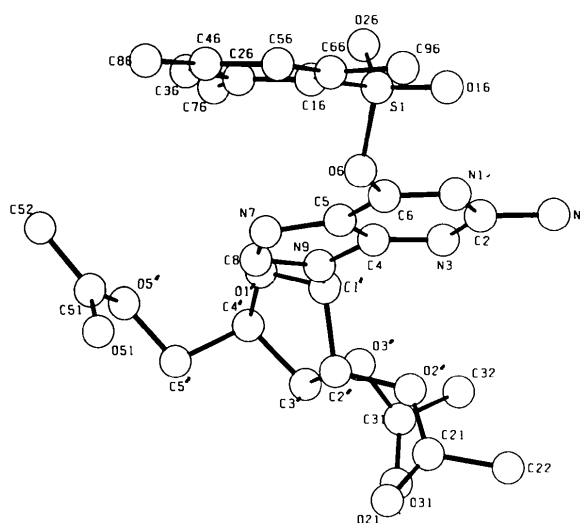


Fig. 1. A view of the molecule showing the numbering scheme used.

to the relatively poor quality of the crystal, which has resulted in an unusually small observable:variable ratio, the e.s.d.'s are not low.

The analysis has confirmed that O(6) of the guanine is the site of attachment of the sulphonyl group. The trimethylbenzene group is aligned approximately parallel to the plane of the guanine base (Fig. 1), although, as Fig. 2(a) shows, these two planar groups do not stack on each other at all. The two acetyl moieties at positions 2' and 3' on the ribose ring are also aligned roughly parallel to each other, presumably so as to minimize non-bonded repulsions.

The arrangement around the S atom shows some deviations from the tetrahedral; these are understandable in view of the steric bulk of some of its adjacent groups. Bond lengths involving the S atom are unexceptional within the limits of experimental error. The dimensions of the ribose group are likewise as expected, although there is perhaps an accentuation of the characteristic difference between the O(1')-C(1') and O(1')-C(4') bond lengths (Saenger, 1973).

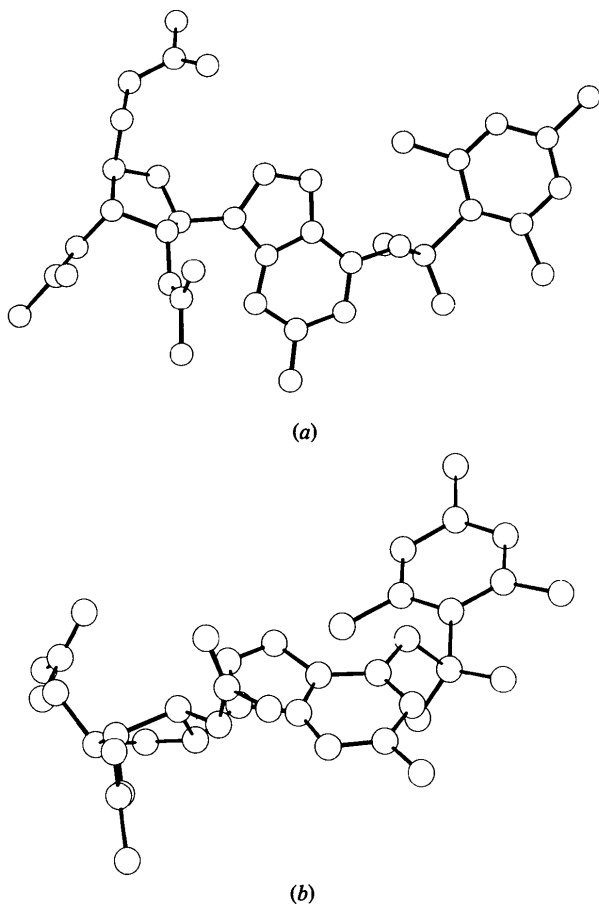


Fig. 2. (a) A view of the molecule projected onto the guanine plane. (b) A view looking down the C(4')-O(1')-C(1') sugar plane.

The guanine base is flat (Table 4). Its geometry is, however, appreciably different from the base in 9-ethyl-guanine (Destro, Kistenmacher & Marsh, 1974) and in guanosine dihydrate (Thewalt, Bugg & Marsh, 1970), as well as from the averaged values reported by Voet & Rich (1970). These differences can be ascribed to the O(6)-substitution, which has resulted in a lengthening

Table 3. Bond lengths (Å) and angles (°)

Estimated standard deviations for bond lengths are between 0.01 and 0.02 Å and 1 and 2° for bond angles. The bond lengths and angles in the benzene ring of the mesitylene group were constrained to 1.395 Å and 120° respectively.

O(1')-C(1)	1.37	N(1)-C(2)	1.36
O(1')-C(4)	1.45	N(1)-C(6)	1.32
C(1')-C(2')	1.55	C(2)-N(2)	1.36
C(1')-N(9)	1.46	C(2)-N(3)	1.32
C(2')-O(2')	1.43	N(3)-C(4)	1.32
C(2')-C(3')	1.56	C(4)-C(5)	1.35
O(2')-C(21)	1.30	C(4)-C(9)	1.38
C(21)-O(21)	1.21	C(5)-C(6)	1.37
C(21)-C(22)	1.50	C(5)-N(7)	1.40
C(3')-O(3')	1.43	C(6)-O(6)	1.38
C(3')-C(4')	1.52	N(7)-C(8)	1.28
O(3')-C(31)	1.31	C(8)-N(9)	1.41
C(31)-O(31)	1.25	O(6)-S	1.62
C(31)-C(32)	1.55	S-O(16)	1.40
C(4')-C(5')	1.53	S-O(26)	1.40
C(5')-O(5)	1.43	S-C(16)	1.72
O(5')-C(51)	1.33	C(26)-C(76)	1.53
C(51)-O(51)	1.23	C(46)-C(86)	1.56
C(51)-C(52)	1.47	C(66)-C(96)	1.47
C(1')-O(1')-C(4')	109	C(2)-N(1)-C(6)	116
O(1')-C(1')-C(2')	106	N(1)-C(2)-N(2)	115
O(1')-C(1')-N(9)	109	N(1)-C(2)-N(3)	128
C(2')-C(1')-N(9)	111	N(2)-C(2)-N(3)	117
C(1')-C(2')-C(3')	101	C(2)-N(3)-C(4)	111
C(1')-C(2')-O(2')	110	N(3)-C(4)-C(5)	129
C(3')-C(2')-O(2')	113	N(3)-C(4)-N(9)	126
C(2')-O(2')-C(21)	119	C(5)-C(4)-N(9)	105
O(2')-C(21)-O(21)	126	C(4)-C(5)-C(6)	115
O(2')-C(21)-C(22)	115	C(4)-C(5)-N(7)	113
O(21)-C(21)-C(22)	119	C(6)-C(5)-N(7)	132
C(2')-C(3')-C(4')	101	N(1)-C(6)-C(5)	121
C(2')-C(3')-O(3')	109	N(1)-C(6)-O(6)	115
C(4')-C(3')-O(3')	106	C(5)-C(6)-O(6)	124
C(3')-O(3')-C(31)	115	S-O(6)-C(6)	121
O(3')-C(31)-O(31)	125	C(5)-N(7)-C(8)	103
O(3')-C(31)-C(32)	111	N(7)-C(8)-N(9)	113
O(31)-C(31)-C(32)	123	C(4)-N(9)-C(8)	106
C(3')-C(4')-O(1')	109	C(4)-N(9)-C(1')	127
C(3')-C(4')-C(5')	108	C(8)-N(9)-C(1')	127
C(5')-C(4')-O(1')	108	O(6)-S-C(16)	99
C(4')-C(5')-O(5)	107	O(6)-S-O(16)	107
C(5')-O(5')-C(51)	120	O(6)-S-O(26)	104
O(5')-C(51)-O(51)	120	O(16)-S-C(16)	113
O(5')-C(51)-C(52)	114	O(16)-S-O(26)	118
O(51)-C(51)-C(52)	126	O(26)-S-C(16)	113
C(16)-C(26)-C(76)	122	C(36)-C(26)-C(76)	118
C(36)-C(46)-C(86)	119	C(36)-C(46)-C(86)	121
C(16)-C(66)-C(96)	129	C(56)-C(66)-C(96)	129
C(56)-C(66)-C(96)	111		

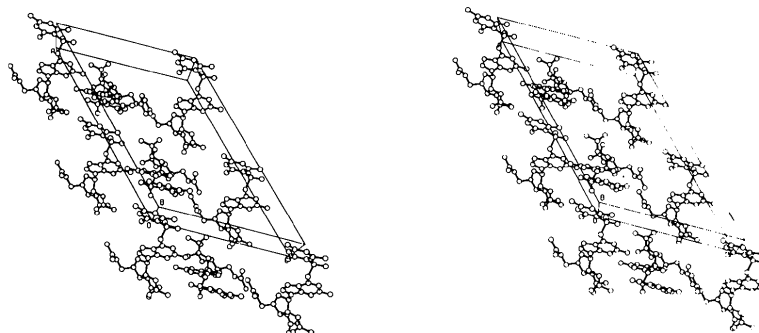


Fig. 3. A stereoview of the packing of molecules in the unit cell.

Table 4. Deviations (\AA) of atoms from various least-squares planes

An asterisk denotes atoms not included in the plane calculations.

Guanine ring		Ribose ring	
N(1)	-0.02	O(1')	-0.04
C(2)	0.00	C(1')	0.02
*N(2)	-0.05	*C(2')	-0.57
N(3)	0.03	C(3')	-0.02
C(4)	0.05		0.04
C(5)	0.03		
C(6)	0.00		
*O(6)	0.02		
N(7)	-0.02		
C(8)	-0.05		
N(9)	0.01		
*C(1')	0.15		

Table 5. Torsion angles ($^{\circ}$)

The torsion angle convention is that given by Sundaralingam (1969).

C(8)—N(9)—C(1')—O(1')	23.8	χ
C(4')—O(1')—C(1')—C(2')	-28.3	τ_0
O(1')—C(1')—C(2')—C(3')	37.7	τ_1
C(1')—C(2')—C(3')—C(4')	-31.3	τ_2
C(2')—C(3')—C(4')—O(1')	16.6	τ_3
C(3')—C(4')—O(1')—C(1')	7.2	τ_4
O(3')—C(3')—C(4')—C(5')	145.9	ψ'
C(3')—C(4')—C(5')—O(5')	173.9	ψ
O(1')—C(4')—C(5')—O(5')	56.2	

of the C(6)—O(6) bond from 1.228 to 1.38 \AA . This partial loss of carbonyl character has affected the delocalization in the purine such that, for example, the N(1)—C(6) bond has shortened from 1.402 \AA (Voet & Rich, 1970) to 1.32 \AA and the angles C(5)—C(6)—N(1) and C(6)—N(1)—C(2) have changed from 111 and 125 $^{\circ}$ (Voet & Rich, 1970), to 121 and 116 $^{\circ}$ respectively.

The ribose sugar shows a C(2')-endo pucker [Fig. 2(b) and Table 4]. The phase angle of pseudorotation (Altona & Sundaralingam, 1972), which expresses this puckering in a quantitative manner, has a value of

149.5 $^{\circ}$, which implies that the puckering is not far from being C(1')-exo, as may be seen in Fig. 2(a). The glycosidic angle (Table 5), is in the usual *anti* region (e.g. Neidle, Taylor & Robins, 1978). The exocyclic C(4')—C(5') bond has a *trans,gauche* conformation (Table 5).

The crystal structure is devoid of hydrogen bonds (Fig. 3), the molecules being held together by van der Waals forces alone; there are no unusually close contact distances.

We are grateful to the Cancer Research Campaign and the Lawson Tait Trust for support, and Dr M. B. Hursthouse for access to his data collection facilities.

References

- ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.
- BRIDSON, P. K., MARKIEWICZ, W. T. & REESE, C. B. (1977). *J. Chem. Soc. Chem. Commun.* pp. 791–792.
- DESTRO, R., KISTENMACHER, T. J. & MARSH, R. E. (1974). *Acta Cryst.* **B30**, 79–85.
- DE TITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 472–479.
- KARLE, J. (1968). *Acta Cryst.* **B24**, 182–186.
- MAIN, P. (1977). *Acta Cryst.* **A33**, 750–757.
- MOTHERWELL, W. D. S. (1975). Univ. of Cambridge. Unpublished.
- NEIDLE, S., TAYLOR, G. L. & ROBINS, A. B. (1978). *Acta Cryst.* **B34**, 1838–1841.
- ROBERTS, P. J., PETERSEN, R. C., SHELDRIK, G. M., ISAACS, N. W. & KENNARD, O. (1973). *J. Chem. Soc. Perkin Trans. 2*, pp. 1978–1984.
- SAENGER, W. (1973). *Angew. Chem. Int. Ed. Engl.* **12**, 591–601.
- SHELDRIK, G. M. (1976). Direct Methods Symposium, Buffalo, New York. To be published.
- SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.
- THEWALT, U., BUGG, C. E. & MARSH, R. E. (1970). *Acta Cryst.* **B26**, 1089–1101.
- VOET, D. & RICH, A. (1970). *Prog. Nucleic Acid Res. Mol. Biol.* **10**, 183–265.