

The Crystal Structure of *N*-Acetylneuraminic Acid Methyl Ester Monohydrate

BY A. M. O'CONNELL*

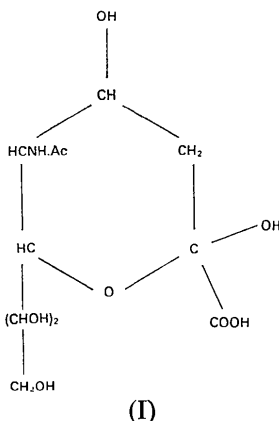
Department of Physics, University of Western Australia, Nedlands, Australia 6009,
and Crystallography Group, University of Göteborg, Medicinaregatan 9, Göteborg, Sweden

(Received 10 October 1972; accepted 19 April 1973)

The crystal structure of the methyl ester of *N*-acetylneuraminic acid monohydrate has been determined from three-dimensional X-ray diffraction data. The crystals are orthorhombic, space group $P2_12_12_1$, with unit-cell constants $a=7.954 \pm 1$, $b=11.675 \pm 4$, $c=16.974 \pm 2$ Å. The structure was solved using the tangent formula and refined by block-diagonal least-squares techniques to $R=0.042$. The three bulky substituent groups, the *N*-acetyl, the glycerol and the methyl ester groups, are all equatorially bonded to the pyranose ring while the hydroxyl group at the anomeric carbon atom C(2) is in the axial position. An extensive system of seven inter- and one intramolecular hydrogen bonds holds the molecules quite rigidly in the crystal lattice, resulting in an average overall isotropic temperature factor coefficient of only 2.7 Å².

Introduction

Neuraminic acid, $C_9H_{17}O_8N$, is the basic unsubstituted unit common to a group of naturally occurring compounds. These compounds, which are *N*-acylated and also sometimes *O*-substituted are known by the group name of sialic acid (Blix, Gottschalk & Klenk, 1957). Sialic acid was first isolated by Blix (1936) by heating bovine submaxillary mucoprotein in water. Interest in this group of compounds developed when it was realized that certain mucoproteins of epithelial secretions are biologically active and that this activity is lost upon treatment with influenza virus or an exoenzyme from *Vibrio Cholerae* resulting in the release of sialic acid from the mucoprotein. In a series of experiments Hirst (1942*a,b*), Burnet (1948) and Gottschalk & Lind (1949*a,b*) demonstrated the role of sialic acid in inhibiting the haemagglutinating power of heat-inactivated influenza virus. During the early 1950's a number of papers on the constitution of sialic acids and related compounds appeared in the literature, the correct structure of *N*-acetylneuraminic acid (I) being proposed by Gottschalk (1955). A review of this work and of the chemical and biological significance of sialic acids is given by Gottschalk (1960).



The present investigation was undertaken in order to provide an independent check on the proposed structure and also to provide a detailed knowledge of the molecular dimensions, stereochemistry and intermolecular packing.

Experimental

Crystals of the methyl ester of *N*-acetylneuraminic acid prepared from human brain gangliosides were kindly provided by Dr L. Svennerholm. Thin-layer chromatography indicated small amounts of the parent acid in the crystals. This was confirmed by comparison of the infrared spectra of the crystals with the spectra reported for the pure acid and its methyl ester (Fischmeister, 1958). Subsequent refinement of the methyl group population in the least-squares procedure showed that the crystals contained 86% of the ester and 14% of the acid (e.s.d.=1%). The ability of various sialic acids to form solid solutions without deforming the crystal lattice has been reported by Abrahamsson, Fischmeister & Svennerholm (1961).

Crystal data

N-Acetyl neuraminic acid methyl ester monohydrate

($C_{12}H_{21}O_9N \cdot H_2O$).

Orthorhombic

$a=7.954 \pm 1$, $b=11.675 \pm 4$,

$c=16.974 \pm 2$ Å ($\lambda=0.71069$).

$U=1576.1$ Å³.

$Z=4$.

Space group $P2_12_12_1$.

Absent spectra: $h00$, h odd; $0k0$, k odd; $00l$, l odd.

$\mu=1.36$ cm⁻¹ for Mo $K\alpha$ radiation.

The cell constants and reflexion data were measured on a Picker FACS1 diffractometer using a crystal

* Present address: C.S.I.R.O., Division of Land Resources Management, Floreat Park, Western Australia.

having dimensions $0.28 \times 0.49 \times 0.44$ mm and Zr-filtered Mo $K\alpha$ radiation. Intensities were measured by the $\omega/2\theta$ scan technique with a scan speed of 1° per min and a scan range at low Bragg angles of 1.2° . This range was continuously incremented to account for the increased α_1 - α_2 dispersion at higher angles. Stationary background counts of 20 sec duration were made at the beginning and end of each scan range. A total of 2081 reflexions having $2\theta \leq 55^\circ$ were recorded. All of these reflexions were included in the subsequent least-squares refinement. Terms having negative $|F_o|^2$ values were set to zero when calculating Fourier and difference maps and R values based on F_o .

The reflexion data were corrected for Lorentz-polarization but not for extinction or absorption ($\mu r_{\max} = 0.10$).

Table 1. *Initial phase assignments*

h	k	l	φ	$ E $	} Origin and enantiomorph Σ_1
3	13	0	1.5π	2.45	
0	9	8	0.5π	2.43	
1	12	0	1.5π	2.14	
2	0	1	1.5π	1.96	
0	2	0	0.0	2.06	

Solution

The phase determination was carried out by means of the tangent formula (Karle & Karle, 1966) using a system of programs written by Hall (1968). Four phases were chosen to define the origin and enantiomorph (Karle & Hauptman, 1956) and in addition the phase ϕ_{020} was evaluated using the Σ_1 relations. This list of

Table 2. *Atomic coordinates and hydrogen atom isotropic thermal parameters*

Positional e.s.d.'s for C, N and O are multiplied by 10^5 and for H by 10^4 . Hydrogen B value e.s.d.'s are multiplied by 10

	x/a	y/b	z/c	B
C(1)	-0.20842 (31)	-0.06514 (20)	0.48322 (12)	
C(2)	-0.13593 (29)	0.01337 (19)	0.54780 (12)	
C(3)	0.05452 (28)	0.02135 (20)	0.54757 (12)	
C(4)	0.12075 (27)	0.08305 (19)	0.62056 (13)	
C(5)	0.04464 (27)	0.03125 (18)	0.69476 (12)	
C(6)	-0.14804 (26)	0.01863 (17)	0.68948 (12)	
C(7)	-0.21793 (27)	-0.05184 (17)	0.75836 (12)	
C(8)	-0.40486 (28)	-0.08130 (19)	0.74855 (13)	
C(9)	-0.46288 (35)	-0.17001 (23)	0.80751 (19)	
C(10)	0.13571 (28)	0.05986 (19)	0.83278 (12)	
C(11)	0.19754 (36)	0.14591 (24)	0.89211 (14)	
C(12)	-0.17019 (55)	-0.23007 (31)	0.40406 (21)	
N(13)	0.09322 (23)	0.10132 (15)	0.76288 (10)	
O(14)	-0.10978 (24)	-0.15209 (15)	0.46574 (11)	
O(15)	-0.34033 (27)	-0.04702 (19)	0.45235 (13)	
O(16)	-0.19268 (19)	-0.04036 (12)	0.61807 (8)	
O(17)	-0.12481 (20)	-0.15614 (13)	0.76859 (9)	
O(18)	-0.49784 (23)	0.02122 (15)	0.76031 (11)	
O(19)	-0.63386 (26)	-0.20150 (17)	0.79553 (15)	
O(20)	-0.20190 (21)	0.12404 (13)	0.53880 (10)	
O(21)	0.29897 (20)	0.06817 (16)	0.62666 (9)	
O(22)	0.12733 (20)	-0.04466 (13)	0.84927 (9)	
O(23)	0.45144 (23)	0.15941 (15)	0.49397 (10)	
H(31)	0.1041 (31)	-0.0513 (20)	0.5434 (13)	2.4 (5)
H(32)	0.0956 (32)	0.0542 (21)	0.4974 (13)	2.5 (5)
H(41)	0.0909 (28)	0.1661 (18)	0.6187 (12)	1.5 (11)
H(51)	0.0946 (27)	-0.0415 (18)	0.7007 (12)	1.2 (4)
H(61)	-0.1937 (31)	0.0885 (20)	0.6913 (14)	2.5 (5)
H(71)	-0.2127 (29)	-0.0072 (19)	0.8101 (13)	1.8 (4)
H(81)	-0.4243 (31)	-0.1104 (22)	0.6989 (14)	2.6 (5)
H(91)	-0.4462 (36)	-0.1367 (24)	0.8657 (16)	3.8 (6)
H(92)	-0.3930 (38)	-0.2393 (25)	0.8037 (17)	4.5 (7)
H(111)	0.1784 (37)	0.2282 (25)	0.8710 (16)	4.1 (6)
H(112)	0.1521 (46)	0.1359 (29)	0.9410 (20)	6.1 (9)
H(113)	0.3099 (47)	0.1290 (30)	0.9064 (20)	6.0 (9)
H(121)	-0.1038 (48)	-0.3029 (31)	0.3994 (21)	8.7 (11)
H(122)	-0.1713 (47)	-0.1888 (29)	0.3525 (19)	7.9 (9)
H(125)	-0.2915 (48)	-0.2488 (31)	0.4143 (20)	8.3 (10)
H(131)	0.1081 (30)	0.1715 (20)	0.7565 (14)	2.3 (5)
H(171)	-0.0466 (33)	-0.1346 (22)	0.7984 (14)	2.9 (5)
H(181)	-0.5472 (35)	0.0364 (23)	0.7212 (15)	3.6 (6)
H(191)	-0.6981 (56)	-0.1462 (36)	0.8162 (26)	9.2 (12)
H(201)	-0.3084 (37)	0.1211 (24)	0.5226 (16)	3.7 (6)
H(211)	0.3391 (42)	0.0902 (27)	0.5871 (19)	5.2 (8)
H(231)	0.4346 (40)	0.1191 (27)	0.4616 (18)	4.8 (7)
H(232)	0.4129 (42)	0.2150 (27)	0.4870 (18)	5.0 (8)

five known phases (Table 1) was extended and refined using the tangent formula. Initially only reflexions with $|E| \geq 1.8$ were used in the tangent calculations.

Terms were rejected if their calculated $|E|$ values were less than 0.5 or if the phases calculated in successive refinement cycles tended to oscillate. After a number

Table 3. Thermal parameters (with *e.s.d.*'s $\times 10^5$) in the form $\exp[-(b_{11}h^2 + b_{22}k^2 + b_{33}l^2 + 2b_{12}hk + 2b_{13}hl + 2b_{23}kl)]$

	b_{11}	b_{22}	b_{33}	b_{12}	b_{13}	b_{23}
C(1)	0.01112 (37)	0.00485 (16)	0.00147 (7)	-0.00039 (24)	-0.00043 (14)	-0.00008 (9)
C(2)	0.00915 (32)	0.00405 (15)	0.00147 (6)	0.00017 (20)	0.00009 (15)	-0.00003 (9)
C(3)	0.00834 (32)	0.00510 (16)	0.00142 (6)	0.00020 (21)	-0.00001 (13)	0.00013 (9)
C(4)	0.00723 (30)	0.00440 (15)	0.00176 (6)	-0.00040 (19)	-0.00007 (14)	0.00017 (9)
C(5)	0.00825 (33)	0.00341 (14)	0.00133 (6)	0.00007 (19)	-0.00017 (13)	-0.00013 (8)
C(6)	0.00743 (32)	0.00335 (13)	0.00145 (6)	-0.00012 (18)	-0.00016 (12)	-0.00004 (8)
C(7)	0.00796 (31)	0.00340 (14)	0.00154 (6)	-0.00011 (19)	0.00015 (12)	-0.00011 (8)
C(8)	0.00846 (34)	0.00435 (16)	0.00194 (7)	-0.00047 (20)	0.00047 (13)	-0.00030 (9)
C(9)	0.01080 (42)	0.00558 (20)	0.00463 (12)	-0.00019 (26)	0.00207 (20)	0.00129 (14)
C(10)	0.00660 (30)	0.00477 (16)	0.00180 (7)	0.00025 (20)	0.00004 (13)	-0.00011 (9)
C(11)	0.01468 (45)	0.00679 (21)	0.00202 (8)	-0.00150 (28)	-0.00164 (17)	-0.00022 (11)
C(12)	0.02408 (89)	0.00671 (30)	0.00375 (15)	-0.00097 (41)	0.00163 (28)	-0.00242 (16)
N(13)	0.00916 (29)	0.00333 (12)	0.00169 (6)	-0.00068 (16)	-0.00041 (11)	-0.00007 (7)
O(14)	0.01457 (32)	0.00577 (14)	0.00311 (7)	0.00064 (19)	-0.00065 (13)	-0.00149 (9)
O(15)	0.01679 (39)	0.00985 (20)	0.00435 (9)	0.00283 (25)	-0.00435 (16)	-0.00252 (12)
O(16)	0.00969 (23)	0.00408 (10)	0.00141 (4)	-0.00122 (14)	-0.00018 (10)	-0.00009 (6)
O(17)	0.00926 (24)	0.00331 (10)	0.00257 (6)	0.00001 (14)	-0.00044 (11)	0.00044 (7)
O(18)	0.01054 (26)	0.00583 (13)	0.00305 (7)	0.00186 (17)	-0.00027 (11)	-0.00042 (8)
O(19)	0.01317 (34)	0.00586 (15)	0.00714 (12)	-0.00181 (20)	0.00431 (18)	-0.00044 (12)
O(20)	0.00971 (25)	0.00403 (11)	0.00254 (6)	0.00050 (15)	-0.00076 (11)	0.00023 (7)
O(21)	0.00711 (23)	0.00878 (16)	0.00199 (5)	-0.00085 (18)	-0.00012 (11)	0.00093 (9)
O(22)	0.00931 (24)	0.00468 (11)	0.00184 (5)	0.00035 (15)	0.00011 (10)	0.00036 (7)
O(23)	0.01274 (31)	0.00532 (12)	0.00236 (6)	0.00068 (18)	0.00038 (11)	0.00003 (8)

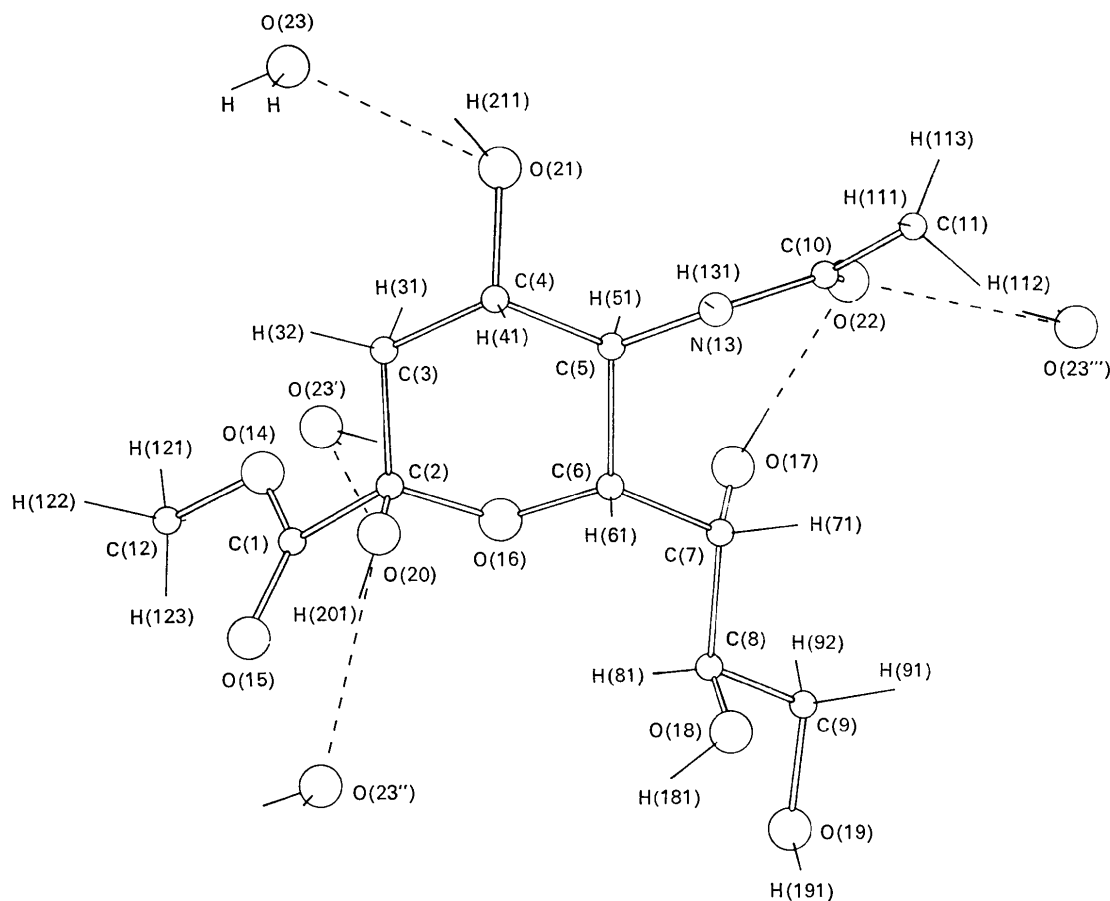


Fig. 1. Molecular stereochemistry and atom numbering. Four equivalent water molecules are shown.

of cycles during which the $|E|$ threshold was reduced to 1.5 the phases of 220 reflexions were determined.

It was found necessary to generate several sets of phases using different sets of origin and enantiomorph defining reflexions in order to obtain a self-consistent set of final phases. The 'correctness' of a set of phases was judged on the degree of convergence obtained and on the uniformity of the distribution of calculated $|E|$ values. The set of phases generated from the starting reflexions listed in Table 1 showed good phase convergence and only one calculated $|E|$ value was less than 0.5. A Fourier E map based on this set clearly showed 16 atoms of the molecule. The positions of the remaining 7 atoms were determined from difference syntheses.

Refinement

The structure was refined using block-diagonal least-squares techniques to minimize the quantity $\sum w(F_o^2 - F_c^2)^2$. Weights were initially based on the counting statistics according to Evans (1961), but during the

later stages of the refinement somewhat better convergence was obtained using the function

$$w = \frac{1.0}{(F_o^2 + 50)} \quad \text{for } F_o^2 < 200$$

$$w = \frac{250}{(F_o^2 + 50)^2} \quad \text{for } F_o^2 > 200.$$

For hydrogen the scattering curve of Stewart, Davidson & Simpson (1965) was used and for carbon, nitrogen and oxygen the values listed in *International Tables for X-ray Crystallography* (1962) were assumed.

Anisotropic thermal parameters were introduced at $R = 0.15$ and at $R = 0.085$ a difference synthesis showed the positions of all hydrogen atoms with peak heights in the range 0.3–0.6 $e \text{ \AA}^{-3}$. Further refinement of non-hydrogen positional and thermal parameters and the population of the C(12) methyl group reduced R to 0.059. The hydrogen positional and isotropic temperature factor coefficients were then included in the refinement and the R index decreased to 0.043. During

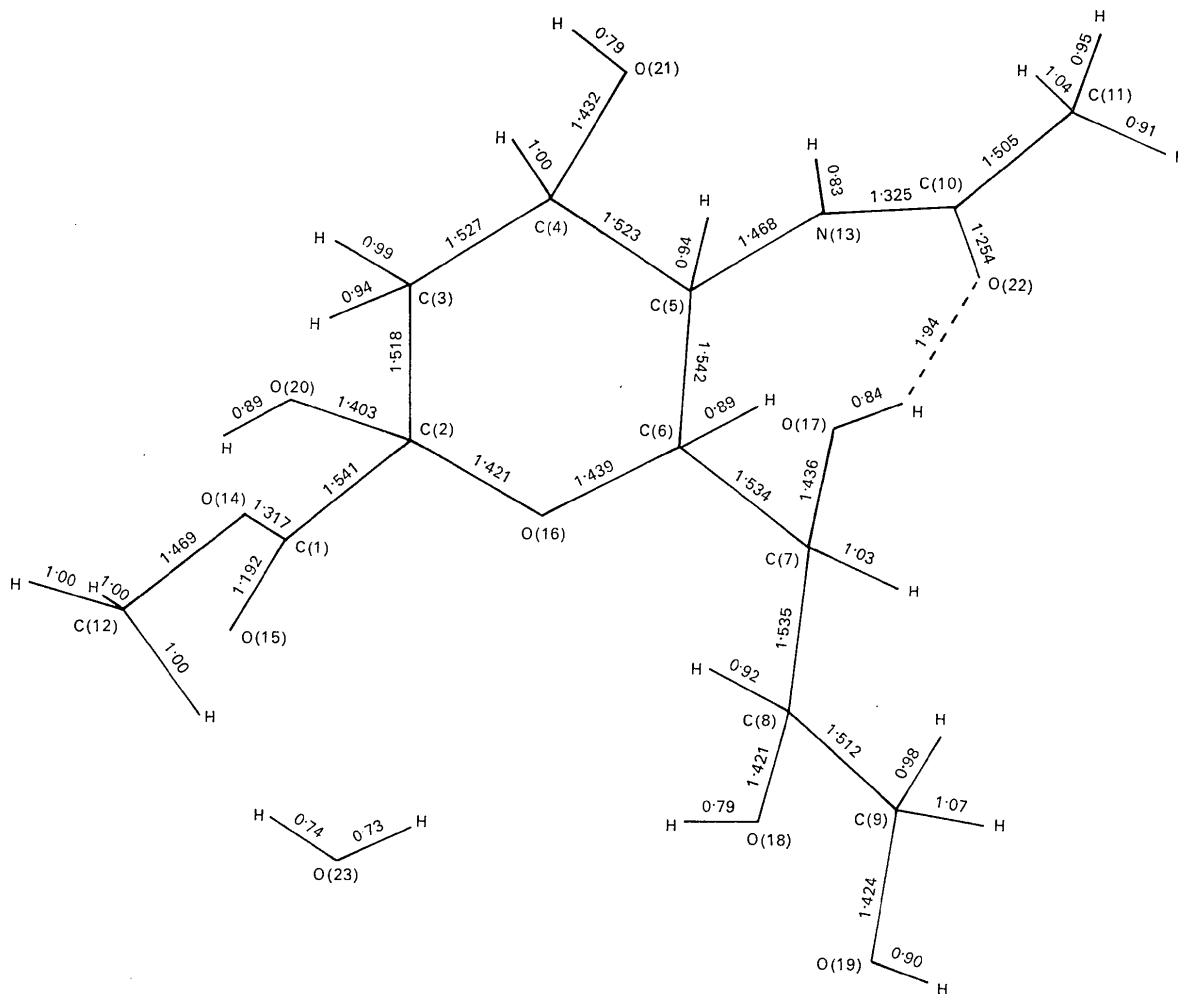


Fig. 2. Intramolecular bond lengths. The e.s.d.'s in the C-X and H-X (X=C, N, O) bonds are 0.003 and 0.03 Å respectively.

the final cycles the parameters of the methyl hydrogen atoms at C(12), which showed rather large oscillations in the least squares, were fixed at values deduced from a difference map. A hydrogen atom having population 0.14 was included at its expected position to account for the partial replacement of the ester by its parent acid. The final value for the population of the C(12) methyl group was 0.86 ± 0.01 .

Positional parameters are given in Table 2 and the thermal parameters in Table 3.

Discussion

The stereochemistry of *N*-acetylneuraminic acid methyl ester is shown diagrammatically in Fig. 1. The pyranose fragment has the normal chair conformation and dimensions that agree generally with those for other sugars (Ramachandran, Ramakrishnan & Sasisekharan, 1963). The three bulky substituents, namely the acetylamino, the glycerol and the methyl ester groups, are bonded equatorially to the ring, as is the hydroxyl group at C(4). The hydroxyl group at the anomeric carbon atom C(2) is in the axial position. The equations of the best planes through the acetylamino, the glycerol and the methyl ester groups and the deviations of individual atoms from the planes are given in Table 4. The ester group is planar within experimental error while the other two groups are approximately planar. In the acetylamino case even the hydrogen atom at N(13) lies within the plane of the other atoms so that the nitrogen atom is truly trigonal. A similar configuration exists in *N*-acetyl glycine (Donohue & Marsh, 1962).

Table 4. *Least-squares planes through the acetylamino (I), the glycerol (II) and the methyl ester (III) groups and the deviations from these planes*

(I)	$0.947X - 0.122Y - 0.297Z + 3.246 = 0$		
(II)	$-0.300X + 0.801Y + 0.518Z - 6.665 = 0$		
(III)	$-0.455X - 0.555Y + 0.697Z - 6.889 = 0$		
Deviations			
	I	II	III
C(5)	0.037 Å	C(6) -0.074 Å	C(1) 0.004 Å
N(13)	-0.039	C(7) 0.038	C(2) -0.003
C(10)	-0.012	C(8) 0.124	C(12) -0.003
C(11)	0.032	C(9) -0.049	O(14) 0.003
O(22)	-0.009	O(19) -0.039	O(15) -0.001
H(131)	-0.009		

The bond lengths are shown in Fig. 2 and the bond angles are listed in Table 5. The mean e.s.d.'s in the C-X and H-X (X=C, N, O) bond lengths are 0.003 and 0.03 Å, respectively. The dimensions of the acetylamino group are in general agreement with those in similar structures, e.g. in triacetyl sphingosine (O'Connell & Pascher, 1969); *N*-acetyl glycine (Donohue & Marsh, 1962); and *N*-acetyl- α -D-glucosamine (Johnson, 1966). The five C-OH type bonds have an average length of 1.421 Å. The axial C-OH bond at the

anomeric carbon atom C(2) is 0.02–0.03 Å shorter than the four other bonds of similar type. In a number of structure determinations of monosaccharides the length of the C(1)-O(1)H bond at the anomeric carbon atom has been found to be consistently shorter than bonds of similar type in the molecules (Kim & Jeffrey, 1967). In the presence of a glycosidic link, or if O(1) carries a methyl group in place of hydrogen, the C-OH bond tends to assume a more normal value and differences appear in the ring C-O lengths, the bond adjacent to the anomeric carbon atom being shorter; e.g. sucrose (Brown & Levy, 1963) and methyl- α -D-glycopyranoside (Berman & Kim, 1968). In the present structure the C-OH bond at C(2) is shorter by 6σ than similar bonds in the molecule and in addition the ring C-O distances differ by 6σ , the C(2)-O(16) distance being the shorter. The latter effect may be due to the additional methyl ester grouping at C(2), although it should be noted that the ring C-O distances are highly correlated since they are linked by the common O(16) oxygen atom.

Table 5. *Bond angles and (in parentheses) their e.s.d.'s*

C(2)-C(1)-O(14)	113.3 (0.2)°
C(2)-C(1)-O(15)	122.5 (0.2)
O(14)-C(1)-O(15)	124.2 (0.2)
C(1)-C(2)-C(3)	114.1 (0.2)
C(1)-C(2)-O(16)	102.5 (0.2)
C(1)-C(2)-O(20)	109.3 (0.2)
C(3)-C(2)-O(16)	110.3 (0.2)
C(3)-C(2)-O(20)	108.5 (0.2)
O(16)-C(2)-O(20)	112.3 (0.2)
C(2)-C(3)-C(4)	111.8 (0.2)
C(3)-C(4)-C(5)	110.3 (0.2)
C(3)-C(4)-O(21)	110.1 (0.2)
C(4)-C(5)-C(6)	112.6 (0.2)
C(4)-C(5)-N(13)	109.0 (0.2)
C(5)-C(6)-C(7)	111.5 (0.2)
C(5)-C(6)-O(16)	109.9 (0.2)
C(7)-C(6)-O(16)	107.2 (0.2)
C(6)-C(7)-C(8)	112.9 (0.2)
C(6)-C(7)-O(17)	111.1 (0.2)
C(8)-C(7)-O(17)	108.8 (0.2)
C(7)-C(8)-C(9)	112.2 (0.2)
C(7)-C(8)-O(18)	107.5 (0.2)
C(9)-C(8)-O(18)	109.0 (0.2)
C(8)-C(9)-O(19)	111.9 (0.2)
C(11)-C(10)-N(13)	116.0 (0.2)
C(11)-C(10)-O(22)	121.2 (0.2)
N(13)-C(10)-O(22)	122.8 (0.2)
C(5)-N(13)-C(10)	124.7 (0.2)
C(1)-O(14)-C(12)	116.3 (0.2)
C(2)-O(16)-C(6)	114.7 (0.2)
*C-O-H	107.2 (3.0)
*X-C-H	110.0 (2.5)
*H-C-H	107.0 (5.3)
*C-N-H	117.4 (1.5)

* Mean values (X=C, O or N).

Within the pyranose ring the C(5)-C(6) distance is 0.02 Å longer than the three similar C-C bonds. The carbon valence angles vary from 110 to 114°. The mean C-H and O-H distances are 0.98 ± 0.05 Å and

$0.84 \pm 0.05 \text{ \AA}$. In the neutron diffraction studies of sucrose (Brown & Levy, 1963) and α -D-glucose (Brown & Levy, 1965) the mean C-H and O-H distances are 1.10 and 0.97 \AA . Thus the present results indicate an apparent shortening of 0.12–0.13 \AA in the X-ray case due to polarization of the H-X bonds.

An analysis of the thermal motion in terms of the ellipsoids of vibration and the angles their axes make with the cell axes is given in Table 6. The vibrations are small and generally isotropic. The mean overall isotropic thermal coefficient is 2.7 \AA^2 which indicates the stabilizing effect of the hydrogen-bond network. Within the pyranose ring the mean isotropic temperature coefficient is 2.0 \AA^2 , while even the water oxygen has the remarkably small value of 2.9 \AA^2 . The only markedly anisotropic motion is towards the end of the glycerol chain. Both C(9) and O(19) have their maximum amplitudes perpendicular to the covalent bond and in the latter case to the O(9)–H(7)···O(19) hydrogen bond. The hydrogen atom temperature factors are physically reasonable and on the average are 1.4 times greater than the equivalent isotropic temperature factor of the parent atoms.

Table 6. *Magnitudes and directions with respect to the crystallographic axes of the principal axes of the vibrational ellipsoids*

	B_i	a	b	c
	$2.92 (10) \text{ \AA}^2$	$27 (12)^\circ$	$116 (14)^\circ$	$98 (4)^\circ$
C(1)	2.59 (10) 1.64 (8)	66 (13) 78 (4)	26 (13) 85 (4)	100 (5) 14 (4)
C(2)	2.35 (9) 2.18 (9) 1.69 (7)	24 (24) 113 (24) 95 (7)	67 (24) 24 (24) 87 (8)	87 (7) 95 (8) 6 (8)
C(3)	2.10 (8) 2.80 (9) 1.64 (7)	6 (7) 84 (6) 89 (9)	96 (6) 8 (5) 95 (4)	92 (9) 85 (4) 6 (5)
C(4)	1.80 (8) 2.48 (9) 1.98 (8)	14 (7) 103 (6) 87 (22)	77 (8) 22 (8) 107 (10)	89 (21) 73 (8) 17 (9)
C(5)	2.11 (8) 1.88 (8) 1.50 (7)	15 (15) 102 (17) 82 (7)	80 (17) 18 (13) 75 (10)	101 (7) 103 (10) 17 (8)
C(6)	1.93 (8) 1.83 (8) 1.63 (8)	26 (23) 77 (35) 69 (13)	109 (37) 23 (33) 78 (18)	106 (17) 108 (21) 25 (15)
C(7)	2.06 (8) 1.87 (8) 1.72 (8)	28 (18) 64 (20) 99 (15)	108 (18) 35 (23) 61 (22)	70 (12) 112 (23) 31 (18)
C(8)	1.93 (9) 2.11 (9) 2.71 (10)	43 (16) 61 (19) 62 (6)	94 (17) 42 (7) 132 (7)	132 (13) 63 (17) 54 (6)
C(9)	2.07 (11) 2.99 (12) 6.05 (16)	40 (5) 125 (5) 73 (2)	60 (5) 36 (5) 72 (2)	115 (2) 95 (3) 25 (2)
C(10)	1.66 (8) 2.62 (9) 2.07 (8)	8 (7) 85 (5) 85 (10)	96 (5) 10 (7) 82 (8)	95 (10) 99 (7) 10 (8)

Table 6 (cont.)

C(11)	4.42 (14) 3.51 (12) 1.81 (9)	38 (4) 66 (5) 62 (2)	123 (6) 36 (6) 78 (3)	107 (3) 115 (3) 31 (3)
C(12)	6.94 (24) 2.02 (12) 5.12 (20)	42 (5) 94 (2) 132 (5)	114 (3) 41 (2) 122 (3)	57 (4) 49 (2) 58 (4)
N(13)	2.49 (8) 1.64 (7) 1.95 (7)	27 (4) 64 (4) 83 (7)	108 (5) 40 (9) 124 (10)	109 (6) 63 (9) 35 (9)
O(14)	3.52 (9) 2.16 (7) 4.74 (10)	22 (3) 88 (3) 112 (3)	105 (3) 40 (2) 126 (2)	75 (3) 50 (2) 44 (2)
O(15)	2.14 (10) 3.94 (11) 8.55 (13)	47 (2) 123 (2) 120 (1)	78 (2) 38 (1) 125 (1)	45 (1) 74 (2) 49 (1)
O(16)	2.81 (7) 1.92 (6) 1.57 (6)	38 (3) 56 (4) 76 (5)	128 (3) 42 (4) 74 (6)	92 (3) 112 (8) 22 (8)
O(17)	2.28 (7) 1.70 (6) 3.13 (7)	18 (4) 97 (5) 106 (4)	79 (5) 18 (4) 76 (2)	76 (4) 107 (2) 22 (3)
O(18)	2.12 (7) 3.81 (9) 3.43 (8)	38 (2) 70 (6) 121 (5)	125 (2) 45 (6) 114 (7)	77 (3) 52 (9) 41 (9)
O(19)	2.15 (8) 3.39 (9) 9.22 (16)	32 (2) 113 (3) 68 (1)	64 (3) 27 (3) 96 (1)	107 (1) 76 (1) 23 (1)
O(20)	2.46 (7) 1.96 (7) 3.17 (8)	49 (5) 125 (5) 119 (4)	47 (5) 44 (5) 86 (4)	72 (4) 112 (3) 29 (3)
O(21)	1.77 (6) 5.02 (9) 2.09 (6)	7 (5) 96 (1) 87 (10)	84 (3) 16 (1) 105 (2)	92 (10) 75 (1) 15 (2)
O(22)	2.31 (6) 2.74 (7) 1.97 (6)	20 (7) 70 (7) 89 (9)	107 (8) 32 (5) 117 (4)	99 (9) 65 (4) 27 (5)
O(23)	3.42 (8) 2.80 (8) 2.63 (7)	31 (5) 106 (9) 115 (7)	64 (6) 33 (13) 71 (17)	74 (5) 118 (18) 32 (16)

The intermolecular packing is characterized by an extensive three-dimensional system of hydrogen bonds, seven inter- and one intramolecular involving all available hydrogen atoms. The bonds vary in length from 2.75 to 2.89 \AA (Fig. 3) and none of the hydrogen atoms differ from linearity by more than 17° . The dominating feature of the hydrogen-bond network is the role played by the water molecule. It is involved in four intermolecular bonds, twice as a donor and twice as an acceptor. Since the z coordinate of O(23) differs from 0.5 by only 0.06 \AA the space-group translations imply that all water molecules lie close to $z=0$ or $z=0.5$. Thus molecules of *N*-acetylneuraminic acid methyl ester are sandwiched between sheets of water molecules parallel to the (001) plane (Fig. 4). Within these sheets the O(23)–H(231)···O(20) and O(20)–H(201)···O(23) bonds form infinite chains parallel to a .

The coordination of the intra- and intermolecularly bonded hydrogen atoms about the water molecules is approximately tetrahedral [Fig. 5(a)]. A fifth short contact involving the water molecule [O(23)···O(15) = 3.01 Å] is probably not a hydrogen bond since the angle O(23)–H(231)···O(19) (= 111°) differs considerably from the expected near linear configuration and the O···H distance (= 2.64 Å) is long for a hydrogen bond. The existence of this bond would also imply that H(231) is involved in a bifurcated system between O(23)···O(15) and O(23)···O(22).

An interesting feature of the hydrogen-bond system is that the acetyl oxygen atom O(22) is involved in three acceptor bonds, two of which are inter- and one intramolecular. The coordination at O(22) can be described as a distorted tetrahedron [Fig. 5(b)]. The mean C(10)–O(22)···H angle is $121 \pm 6^\circ$ and the mean H···O(22)···H angle $95 \pm 5^\circ$. That O(22) is bound tightly in the crystal structure is illustrated by its small average isotropic temperature factor coefficient ($B = 2.3 \text{ \AA}^2$) even though it occupies a terminal position in the molecule.

The author is indebted to Dr Bo Nilsson for making the infrared and thin-layer chromatography measurements. Financial support was obtained from the Australian Research Grants Committee, the Swedish Medical Research Council, the Swedish National Science Research Council and the U.S. Public Health Service.

References

- ABRAHAMSSON, S., FISCHMEISTER, I. & SVENNERHOLM, L. (1961). *Ark. Kem.* **18**, 435–440.
 BERMAN, M. H. & KIM, S. H. (1968). *Acta Cryst.* **B24**, 897–904.
 BLIX, G. (1936). *Z. Phys. Chem.* **240**, 43.
 BLIX, G., GOTTSCHALK, A. & KLENK, E. (1957). *Nature, Lond.* **179**, 1088.
 BROWN, G. M. & LEVY, H. A. (1963). *Science*, **141**, 921–923.
 BROWN, G. M. & LEVY, H. A. (1965). *Science*, **147**, 1038–1039.
 BURNET, F. M. (1948). *Aust. J. Exp. Biol. Med. Sci.* **26**, 371.

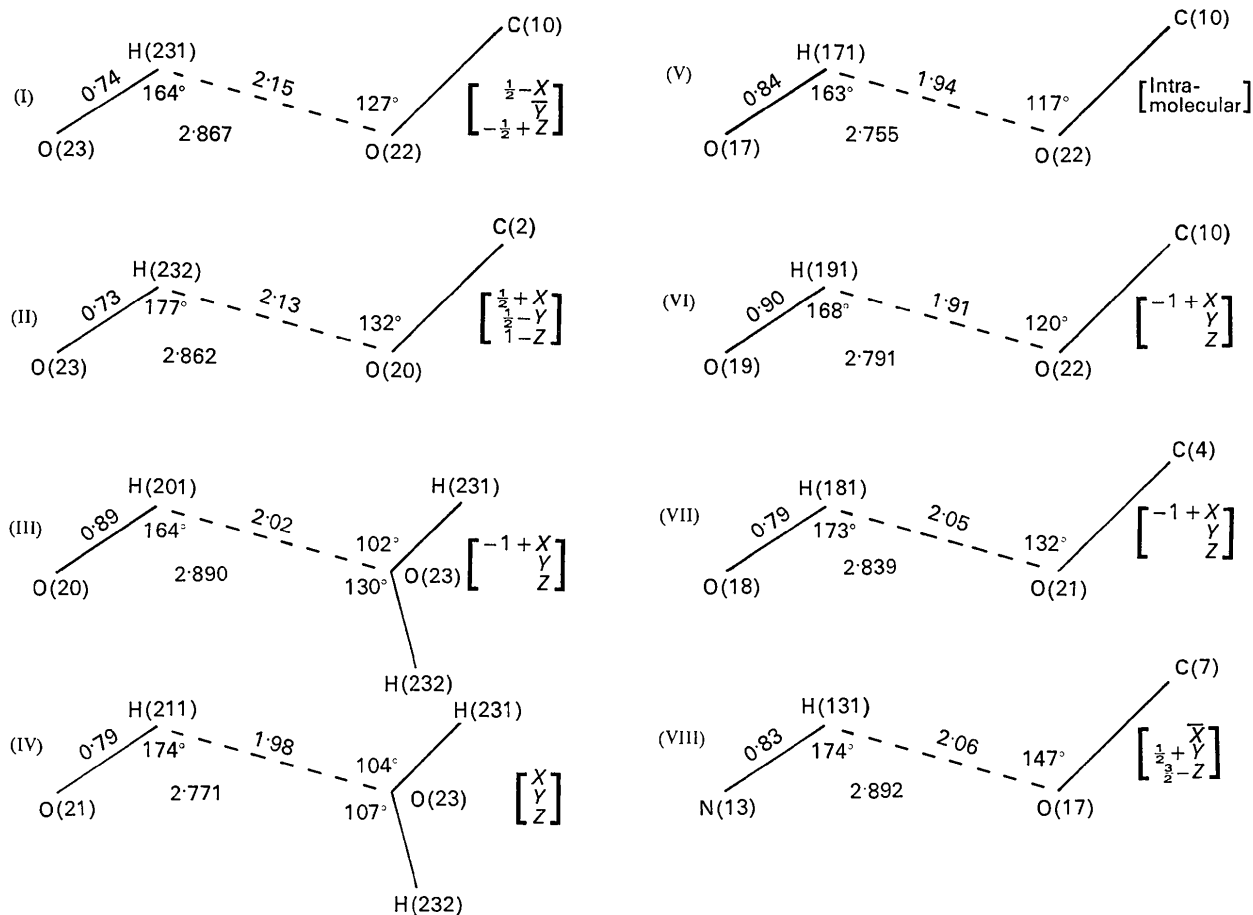


Fig. 3. Distances (Å) and angles ($^\circ$) associated with the eight hydrogen bonds. The equivalent positions of the acceptor atoms are given in square brackets.

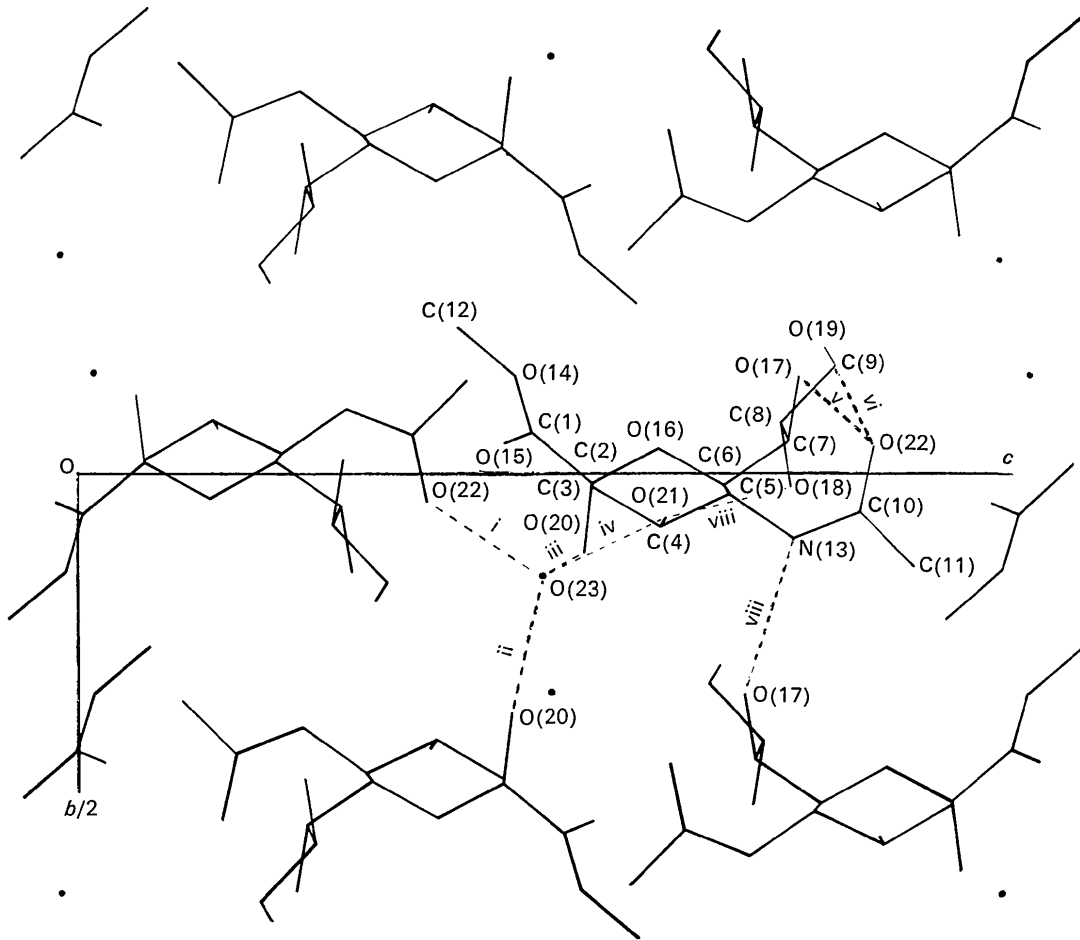


Fig. 4. Intermolecular packing viewed down the *a* axis. Hydrogen bonds are indicated by dotted lines.

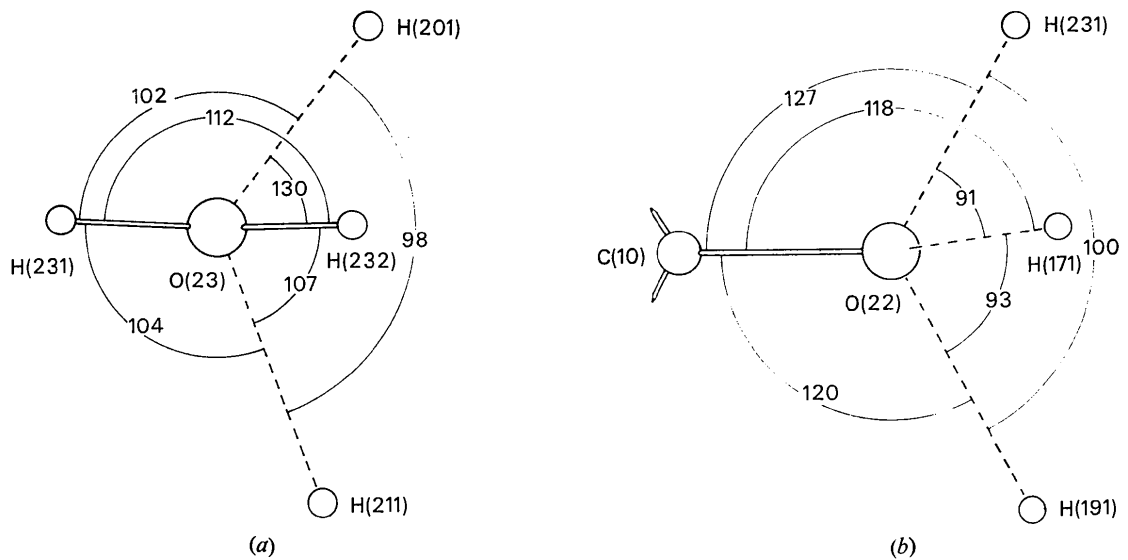


Fig. 5. (a) Coordination angles at the water molecule. (b) Coordination angles at O(22).

- DONOHUE, J. & MARSH, R. E. (1962). *Acta Cryst.* **15**, 941–945.
- EVANS, H. T. (1961). *Acta Cryst.* **14**, 689.
- FISCHMEISTER, I. (1958). *Ark. Kem.* **13**, 247–258.
- GOTTSCHALK, A. (1955). *Nature, Lond.* **176**, 881–882.
- GOTTSCHALK, A. (1960). *The Chemistry and Biology of Sialic Acids and Related Substances*. Cambridge Univ. Press.
- GOTTSCHALK, A. & LIND, P. E. (1949a). *Brit. J. Exp. Path.* **30**, 85.
- GOTTSCHALK, A. & LIND, P. E. (1949b). *Nature, Lond.* **164**, 232–233.
- HALL, S. R. (1968). *Direct Phasing Methods*. Univ. of Western Australia Computer Program UWAS-17.
- HIRST, G. K. (1942a). *J. Exp. Med.* **75**, 49.
- HIRST, G. K. (1942b). *J. Exp. Med.* **76**, 195.
- International Tables for X-ray Crystallography* (1962). Vol. III, p. 202. Birmingham: Kynoch Press.
- JOHNSON, L. N. (1966). *Acta Cryst.* **21**, 885–891.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **22**, 849–868.
- KIM, S. H. & JEFFREY, G. A. (1967). *Acta Cryst.* **22**, 537–545.
- O'CONNELL, A. M. & PASCHER, I. (1969). *Acta Cryst.* **B25**, 2553–2561.
- RAMACHANDRAN, G. N., RAMAKRISHNAN C., & SASISEKHARAN, V. (1963). *Aspects of Protein Structure*, p. 121. New York & London: Academic Press.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

Acta Cryst. (1973). **B29**, 2328

Crystal and Molecular Structure of 2-Thio-1-(β -D-ribofuranosyl)-3H-benzimidazole

BY P. PRUSINER AND M. SUNDARALINGAM*

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706, U.S.A.

(Received 25 May 1973; accepted 28 May 1973)

2-Thio-1-(β -D-ribofuranosyl)-3H-benzimidazole crystallizes as the monohydrate ($C_{12}H_{14}N_2O_4S \cdot H_2O$) in the orthorhombic space group $P2_12_12_1$ with $a = 21.726$ (8), $b = 9.006$ (3), and $c = 6.875$ (2) Å, $Z = 4$. The structure has been determined from 1194 X-ray intensities measured on a Picker four-circle automated diffractometer with Cu $K\alpha$ radiation. The structure was solved by the heavy-atom method and refined by least-squares calculations to a final R of 0.04. The nucleoside adopts the *syn* conformation with a glycosyl torsion angle χ_{CN} of 249.8° , and the mode of puckering of the ribose ring is $C(2')\text{-endo-C}(1')\text{-exo}$ (2T_1). These conformational features are consistent with the other known 8-substituted purine nucleosides. The conformation about the $C(4')\text{-C}(5')$ bond is *gauche-trans*. All potential sites on the sugar and base are involved in hydrogen bonding and the sulfur atom is found to accept two hydrogen bonds, one from the water molecule and the other from an adjacent $O(5')$ hydroxyl group. There is no interbase hydrogen bonding. The bases form a stacked column along alternating screw axes running parallel to c . The interplanar base separation is 3.42 Å and the shortest interbase contact is 3.61 Å involving the $C(4)$ and $C(8)$ atoms of adjacent bases. The stacking pattern observed here differs markedly from those reported for the 8-halogenated nucleosides as well as the majority of the known thio purine crystal structures.

Introduction

2-Thio-(β -D-ribofuranosyl)-3H-benzimidazole (TRB) belongs to a class of nucleoside analogs containing a benzimidazole base. The compound could exist in one of two tautomeric structures shown in Fig. 1. The benzimidazole base is also of biological interest since 5,6-dimethyl-1-(α -D-ribofuranosyl)-benzimidazole occurs as a fragment of vitamin B_{12} . The purpose of this study was to obtain information on the molecular dimensions and conformation and to compare it with the structure of the 2-chloro-1-(β -D-ribofuranosyl)-benzimidazole (CRB) which was recently completed in this laboratory (Sprang & Sundaralingam, 1973). In addition

it was of interest to determine the effects of the thio substituent on the base stacking, the hydrogen bonding and the overall conformation of the structure and to compare them with other sulfur-containing nucleic acid base analogs.

Experimental

A crystal with dimensions approximately 0.2 mm in all directions was chosen for the X-ray structure analysis. The crystal was mounted about c . The systematic absences: $h00, h = 2n + 1$; $0k0, k = 2n + 1$; $00l, l = 2n + 1$, show that the crystal belongs to the space group $P2_12_12_1$. The cell constants were determined from a least-squares refinement of the 2θ , ω and χ angles of 12 medium-angle reflections measured on our Picker

* To whom to address correspondence.