Acta Cryst. (1980). B**36**, 878–885

The Structure of 4-Amino-1-[4-amino-2-oxo-1(2*H*)-pyrimidinyl]-1,4-dideoxy-*β*-D-glucopyranuronic Acid (*C*-Substance) Monohydrate: a Nucleoside Fragment of Gougerotin

By P. Swaminathan, J. McAlister and M. Sundaralingam*

Department of Biochemistry and College of Agricultural and Life Sciences, University of Wisconsin–Madison, 420 Henry Mall, Madison, Wisconsin 53706, *USA*

(*Received* 5 *September* 1979; *accepted* 20 *November* 1979)

Abstract

C-substance, the nucleoside fragment of the aminoacyl antibiotic gougerotin, crystallizes in space group $P2_1$ with cell dimensions $a = 10{\cdot}027$ (5), $b = 8{\cdot}606$ (3), $c = 14{\cdot}620$ (7) Å, $\beta = 98{\cdot}45$ (2)°; $Z = 4$. The structure was solved by direct methods and refined to an R index of 0·048 using 1987 intensities measured on a diffractometer. The two independent molecules are related by a pseudo twofold screw axis parallel to the z direction and have similar conformations, except for the orientation of the carboxyl group. The cytosine base is slightly nonplanar in both molecules with a shallow boat configuration. The pyranosyl-ring conformations are $^4C_1$ and the dispositions of the bases are *anti*, $\chi = 75{\cdot}0$ (6) and 71·7 (5)° for molecules A and B respectively. The two endocyclic C—O bonds of the pyranose rings are equal in both molecules with a shortened base–sugar glycosyl C(1′)—N(1) bond, a feature reminiscent of the anomeric effect observed in glucosides with an equatorial substituent at C(1′). The carboxyl groups are rotated differently about the exocyclic C(5′)—C(6′) bonds in the two molecules. In molecule B the carboxyl group is virtually in the plane of the ring O(5′) atom by virtue of a strong intramolecular hydrogen bond to the amino group, while in molecule A it is rotated into the plane of the ring C(4′) atom by a hydrogen bond through a water to the carbonyl oxygen O(2) of the base. Both molecules exist as zwitterions with a protonated amino group at the 4′ position and an ionized carboxyl group. The structure also exhibits two bifurcated hydrogen bonds and a short intermolecular C(6)···O(5′) contact (3·01 Å).

Introduction

C-substance {4-amino-1-[4-amino-2-oxo-1(2*H*)-pyrimidinyl]-1,4-dideoxy-*β*-D-glucopyranuronic acid} is a nucleoside fragment of the aminoglycosyl antibiotic gougerotin (Clark & Gunther, 1963; Clark & Chang, 1965; Sinohara & Skypeck, 1965). Gougerotin is composed of a pyranosyl nucleoside residue and an amino acid residue that are linked through an aminoacyl glycosyl linkage (Fig. 1*a*). Gougerotin inhibits protein synthesis by affecting the peptide-transfer reaction. The peptide synthetase accepts gougerotin in the place of the aminoacyl tRNA and thus is inhibited from completing the peptide-transfer step in the ribosome (Fox, Watanabe & Bloch, 1966). As part of a program of studies on the structures of protein synthesis inhbitors, we have undertaken the X-ray analysis of *C*-substance where the dipeptidyl chain at the 4′ position and the carboxamide at the 5′ position of gougerotin are replaced by an amino group and a carboxyl group respectively, Fig. 1(*b*). Our additional interest in *C*-substance and gougerotin is that they possess a pyranosyl sugar moiety instead of a furanosyl sugar and this has offered an opportunity of studying the conformational properties of the pyranosyl nucleosides.

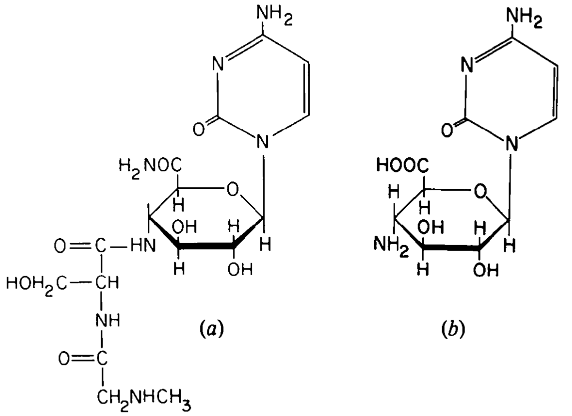


Fig. 1. Molecular formulae for (*a*) gougerotin and (*b*) *C*-substance.

Experimental

A sample of *C*-substance, kindly provided by Dr K. A. Watanabe, Sloan Kettering Research Institute, was

---

* To whom correspondence should be addressed.

0567-7408/80/040878-08$01.00

Table 1. *Crystal data and |E| statistics*

(*a*) Crystal data

Empirical formula	$C_{10}H_{14}N_4O_6.H_2O$	Z	4
Formula weight	304·26	V	1247·9 (1·7) Å^3
Space group	$P2_1$	D_c	1·619 Mg m^{-3}
a	10·027 (5) Å	λ(Cu $K\alpha$)	1·5418 Å
b	8·606 (3)	μ(Cu $K\alpha$)	1·206 mm^{-1}
c	14·620 (7)	Crystal size	0·15 × 0·10 × 0·40 mm
β	98·45 (2)°		

(*b*) Statistical distribution of |E| values and the starting set of phases

	Theoretical		Observed*	
$\|E\|$ statistics	centric	acentric	Debye	Wilson
$\langle \|E\|^2 \rangle$	1·00	1·00	1·00	1·00
$\langle \| \|E\|^2 - 1 \| \rangle$	0·97	0·74	0·84	0·85
$\langle \|E\| \rangle$	0·80	0·89	0·86	0·85

Starting set of phases

h	k	l	$\|E\|_{\text{Debye}}$	$\|E\|_{\text{Wilson}}$
3	5	−12	3·1†	3·1
8	0	−6	2·8§	3·0
4	7	−1	2·7‡	3·0†
4	6	2	2·7§	2·9§
4	8	−5	2·5†	2·4
3	4	−9	2·5§	2·6
7	1	−10	2·6	2·9‡
0	0	4	2·6	2·7§
4	6	0	2·1	2·2§
1	0	4	2·0	1·9†

* The values in the columns with the heading Debye or subscript Debye are calculated with molecular scattering factors, and those with the heading Wilson are calculated with atomic scattering factors.

† Origin-defining reflections.

‡ Enantiomorph-defining reflection.

§ Other reflections in the starting set.

crystallized in the monoclinic space group $P2_1$ by slow evaporation of its solution in a mixture of ethanol, propanol and butanol. The relevant crystal data for the compound are given in Table 1. Intensity data were collected on a needle-shaped crystal of approximate dimensions 0·15 × 0·10 × 0·40 mm to a 2θ limit of 127° using the θ–2θ scan mode at a speed of 2° min^{-1}, on a Picker FACS-I four-circle diffractometer using Ni-filtered Cu $K\alpha$ radiation. The scan range for each reflection was computed using the expression $\Delta\theta = 1{\cdot}5° + 2 \times 0{\cdot}14 \tan\theta$, where θ is the Bragg angle. Background counts were measured at both the extremities of the peak for 10 s. The intensities were corrected for differential absorption, by measuring the intensity variation of the 020 reflection at $\chi = 90°$, for different values of φ (North, Phillips & Mathews, 1968), and for the Lorentz and polarization effects. Out of 2099 reflections scanned, 1987 had $I > 1{\cdot}5\sigma(I)$, where $\sigma(I) = (I_{\text{scan}} + I_{\text{bkg}} + 0{\cdot}03I^2_{\text{scan}})^{1/2}$, and were considered in the structure analysis.

Structure determination and refinement

Initial attempts to solve the structure by the symbolic addition procedure and Patterson techniques using |E| values calculated with the atomic scattering factors (*International Tables for X-ray Crystallography*, 1962) were not successful. The statistical distribution of |E| values (Table 1*b*) lies midway between ideally centric and acentric distributions indicating that the structure has pseudosymmetry. The atomic scattering factors $\sum_j f_j^2$ used in calculating |E|'s assumes that the atoms are randomly distributed in the unit cell of the crystal, which is not strictly true in the case of highly symmetrical molecules. Hence, spherically averaged molecular scattering factors were used for calculating the |E|'s (Debye, 1915; Main, 1975; Swaminathan, 1975) by treating the molecule as a whole, or fragments of it, to be spherically symmetric and randomly distributed in the unit cell of the crystal. The expression for calculating the molecular scattering factor is

$$g_m^2 = \sum_i \sum_j \frac{f_i f_j \sin Kr_{ij}}{Kr_{ij}}, \tag{1}$$

where f_i, f_j are the atomic scattering factors, r_{ij} is the interatomic distance and $K = 4\pi \sin\theta/\lambda$. The normalized structure factor is given by $|E| = |F_o|/\sum_m [\varepsilon g_m^2]^{1/2}$ where m is the number of groups in the unit cell. In the present analysis, the molecule was treated as two rigid groups consisting of the atoms of the base and those of the sugar ring. The interatomic distances for the base and sugar moieties were taken from the mean values reported by Rao & Sundaralingam (1973) and Fries, Rao & Sundaralingam (1971), respectively. The molecular scattering factor was calculated using the program *NORMAL* (first part of the *MULTAN* program series, Main, Woolfson, Lessinger, Germain & Declercq, 1974). The Debye curve obtained with the molecular scattering factor agrees better with the least-squares straight line than does the regular Wilson plot obtained with atomic scattering factors (Fig. 2). The starting set of |E| values

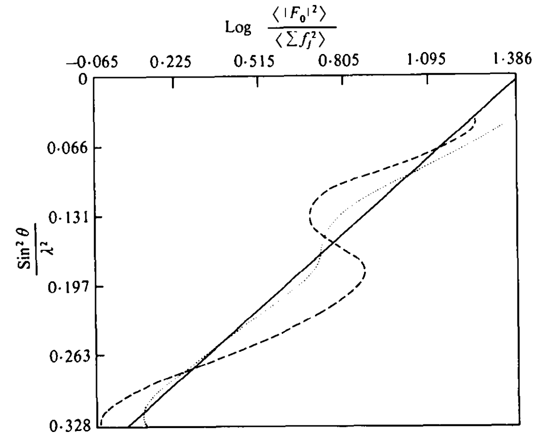


Fig. 2. Wilson curve using atomic scattering factors (dashed line) and Debye curve using the spherically averaged molecular scattering factors (dotted line). The agreement of the Debye curve with the least-squares fit (straight line) is significantly better than the Wilson curve.

Table 2. *Fractional positional parameters* ($\times 10^4$) *and equivalent isotropic temperature factors of the non-hydrogen atoms of C-substance*

	x	y	z	B_{eq}* (Å^2)
O(5')*A*	2401 (3)	4069 (5)	3207 (3)	1·87
C(1')*A*	2999 (5)	4478 (6)	4122 (4)	1·51
C(2')*A*	2025 (5)	5467 (6)	4581 (3)	1·54
O(2')*A*	2578 (4)	5775 (6)	5514 (3)	2·60
C(3')*A*	704 (5)	4574 (6)	4572 (3)	1·15
O(3')*A*	−250 (4)	5524 (5)	4941 (3)	1·66
C(4')*A*	162 (5)	4100 (6)	3589 (4)	1·36
N(4')*A*	−1112 (4)	3182 (6)	3560 (3)	2·14
C(5')*A*	1212 (5)	3147 (6)	3171 (3)	1·44
C(6')*A*	822 (5)	2777 (6)	2139 (4)	2·14
O(4')*A*	−342 (4)	3140 (6)	1759 (3)	2·70
O(6')*A*	1721 (4)	2154 (6)	1752 (3)	2·37
N(1)*A*	4249 (4)	5284 (6)	4010 (3)	1·51
C(2)*A*	5453 (5)	4405 (8)	4078 (4)	1·69
O(2)*A*	5418 (4)	3020 (6)	4270 (3)	2·00
N(3)*A*	6593 (5)	5139 (6)	3937 (3)	2·03
C(4)*A*	6549 (6)	6643 (8)	3686 (4)	2·02
N(4)*A*	7738 (5)	7317 (6)	3590 (4)	2·65
C(5)*A*	5330 (6)	7489 (8)	3507 (4)	2·10
C(6)*A*	4209 (6)	6783 (6)	3679 (4)	1·76
O(5')*B*	8262 (3)	2773 (5)	7959 (2)	1·87
C(1')*B*	7891 (5)	2581 (6)	8863 (3)	1·65
C(2')*B*	8979 (5)	1625 (6)	9453 (3)	1·43
O(2')*B*	8592 (4)	1370 (5)	10333 (2)	2·53
C(3')*B*	10340 (5)	2417 (6)	9513 (3)	1·59
O(3')*B*	11383 (4)	1442 (5)	9940 (3)	1·97
C(4')*B*	10640 (5)	2826 (6)	8543 (3)	1·91
N(4')*B*	11906 (4)	3734 (6)	8643 (3)	1·79
C(5')*B*	9444 (5)	3711 (6)	8002 (3)	1·50
C(6')*B*	9690 (5)	4220 (6)	7031 (4)	1·70
O(4')*B*	8900 (4)	3830 (6)	6338 (3)	3·49
O(6')*B*	10732 (4)	5033 (5)	7037 (3)	2·39
N(1)*B*	6584 (4)	1821 (6)	8763 (3)	1·64
C(2)*B*	5433 (5)	2637 (6)	8938 (4)	2·10
O(2)*B*	5514 (4)	4044 (5)	9102 (4)	3·63
N(3)*B*	4257 (4)	1829 (6)	8903 (3)	2·34
C(4)*B*	4214 (5)	315 (6)	8723 (4)	2·20
N(4)*B*	3052 (5)	−444 (6)	8744 (4)	3·28
C(5)*B*	5349 (6)	−507 (8)	8499 (4)	2·67
C(6)*B*	6499 (6)	289 (6)	8523 (4)	2·17
O(*W*1)	6316 (6)	1140 (8)	1839 (5)	4·78
O(*W*2)	4175 (5)	964 (6)	2894 (4)	5·72

* The equivalent isotropic temperature factor is calculated using the expression $B_{eq} = \frac{4}{3}\sum_i \sum_j \mathbf{a}_i \mathbf{a}_j \beta_{ij}$, where the $\mathbf{a}_i$'s are the unit-cell edges in direct space.

Table 3. *Fractional positional parameters* ($\times 10^3$) *and isotropic temperature factors of the hydrogen atoms of C-substance*

	x	y	z	B (Å^2)
H(1')*A*	317 (5)	346 (6)	442 (3)	2·46
H(2')*A*	181 (5)	645 (7)	421 (3)	1·30
H(3')*A*	88 (5)	360 (6)	491 (3)	1·69
H(4')*A*	−7 (5)	497 (7)	320 (3)	1·04
H(5')*A*	139 (5)	217 (6)	352 (3)	1·61
H(O2')*A*	322 (5)	651 (7)	552 (4)	4·59
H(O3')*A*	−93 (7)	509 (9)	485 (5)	10·55
H(N4')1*A*	−177 (7)	373 (10)	375 (5)	5·98
H(N4')2*A*	−140 (6)	293 (8)	292 (4)	5·22
H(N4')3*A*	−100 (6)	221 (8)	388 (4)	6·19
H(N4)1*A*	764 (6)	839 (8)	355 (4)	4·67
H(N4)2*A*	844 (7)	679 (9)	378 (5)	5·47
H(5)*A*	541 (5)	861 (7)	333 (4)	2·04
H(6)*A*	331 (5)	722 (7)	356 (4)	2·10
H(1')*B*	781 (6)	371 (8)	911 (4)	1·28
H(2')*B*	899 (5)	56 (6)	921 (3)	2·42
H(3')*B*	1030 (5)	339 (7)	985 (4)	1·59
H(4')*B*	1069 (5)	180 (7)	821 (3)	1·35
H(5')*B*	936 (5)	467 (7)	834 (4)	1·70
H(O2')*B*	900 (5)	178 (7)	1072 (4)	4·40
H(O3')*B*	1148 (6)	164 (8)	1046 (4)	4·03
H(N4')1*B*	1260 (7)	304 (9)	873 (5)	6·37
H(N4')2*B*	1187 (5)	449 (6)	906 (3)	3·97
H(N4')3*B*	1187 (5)	427 (7)	808 (4)	2·95
H(N4)1*B*	318 (6)	−149 (8)	873 (4)	3·92
H(N4)2*B*	252 (7)	5 (9)	908 (5)	6·19
H(5)*B*	532 (6)	−148 (8)	941 (4)	4·27
H(6)*B*	738 (6)	−14 (8)	833 (4)	2·38
H(*W*2)1	327 (8)	72 (12)	247 (6)	14·48
H(*W*2)2	392 (8)	152 (11)	346 (6)	12·38
H(*W*1)1	595 (4)	106 (6)	224 (3)	2·23
H(*W*1)2	565 (0)	38 (0)	144 (0)	2·25

picked by the *CONVERGENCE* routine of the program *MULTAN* (Main *et al.*, 1974) calculated using atomic and molecular scattering factors are listed in Table 1(*b*). Since the $|E|$'s calculated by the two methods are different, the order in which they appear in the phase propagation is different and hence the difference in the starting phases. While the set of starting phases based on the atomic scattering factors gave fragments of the structure at the wrong position, the starting set of phases picked using the $|E|$'s calculated with molecular scattering factors gave a 15-atom fragment of the molecule in the correct position. Tangent refinement based on this partial structure information followed by additional E maps yielded the positions of all the nonhydrogen atoms of the molecule.

Full-matrix least-squares refinement of the atoms with isotropic temperature factors for all the non-hydrogen atoms followed by difference Fourier maps revealed the positions of two water molecules and those of the H atoms with the exception of that attached to O(*W*1). The refinement was continued with anisotropic temperature factors for the nonhydrogen atoms and isotropic temperature factors for H atoms using Cruickshank, Pilling, Bujosa, Lovell & Truter's (1961) weighting scheme, given by $w = 1/(a + b|F_o| + c|F_o|^2)$, with $a = 2{\cdot}66$, $b = 1{\cdot}0$ and $c = 0{\cdot}0007$. During the refinement of the structure, y coordinates of different groups of connected atoms of molecule A and molecule B were kept fixed to avoid any bias by neglecting highly correlated parameters. The refinement converged at $R = 0{\cdot}048$, at which point the final shifts in the atomic parameters were all less than one tenth their standard deviations.

The least-squares refinements were carried out on both a Univac 1110 computer using the program *ORFLS* (Busing, Martin & Levy, 1962) and an in-house PDP 11/35 minicomputer using programs developed in these laboratories (Rao, McAlister & Merrit, unpublished results). The scattering factors used for C, N and O are those given by Cromer & Waber (1965) and for H atoms those given by Stewart, Davidson & Simpson (1965).

The fractional positional coordinates of the non-hydrogen and H atoms are given in Tables 2 and 3.*

Discussion

Both molecules of *C*-substance exist as zwitterions with an ionized carboxyl group and a protonated amino group at C(4′).

Bond lengths and angles

The bond lengths and angles of the non-hydrogen atoms are shown in Fig. 3. The mean standard deviation in the bond lengths is 0·006 Å and in angles 0·4°.

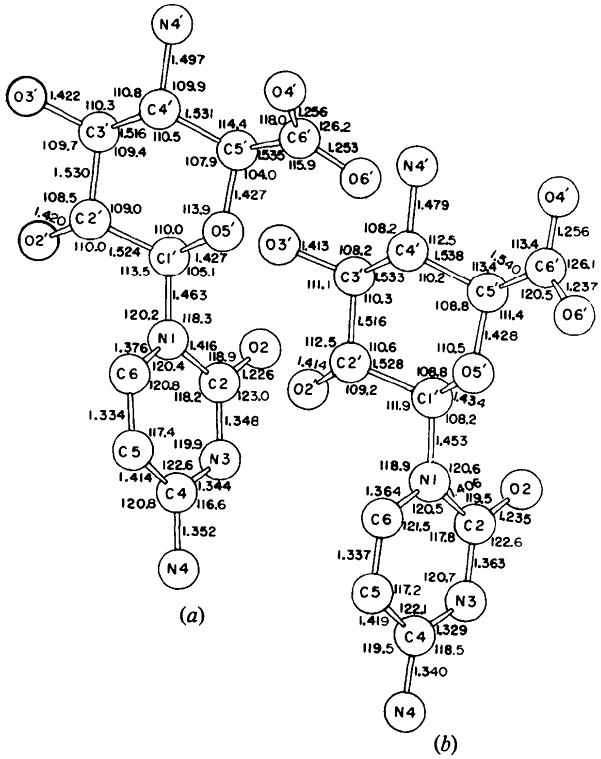


Fig. 3. Bond lengths (Å) and angles (°) observed in (*a*) molecule *A*, and (*b*) molecule *B*.

The sugar

The bond lengths and angles of the *β*-glucosyl moieties of the two molecules agree with each other within experimental error. The two endocyclic C—O bonds of the pyranose rings are nearly equal, with values of 1·427 (7), 1·427 (7) Å (molecule *A*) and 1·428 (6), 1·434 (6) Å (molecule *B*) for the C(5′)—O(5′) and O(5′)—C(1′) bonds, respectively, a feature which is reminiscent of the mean values of 1·441 and 1·426 Å found for these bond lengths in glucosides with an equatorial substituent at the anomeric carbon atom C(1′) (Berman, Chu & Jeffrey, 1967; Jeffrey, Pople, Binkley & Vishveshwara, 1978). The sugar—base C(1′)—N(1) glycosyl bond distances of 1·463 (7) (molecule *A*) and 1·453 (7) Å (molecule *B*) are shortened significantly compared to the mean value of 1·491 (4) Å (Lin, Sundaralingam & Arora, 1971) observed in the corresponding pyrimidine furanosyl nucleosides. Thus the equatorially linked heterocyclic nucleobase to the anomeric carbon atom C(1′) seems to produce the same electronic effect (anomeric effect) on the C—O—C—N bond sequence as that of an equatorial substituent on the C—O—C—O bond sequence in glucopyranosides.

The bond angles involving the ring oxygen atom O(5′) are 113·9 (4)° (molecule *A*) and 110·5 (3)° (molecule *B*). This angle varies from 107 to 114° in the known glucuronic acid structures (compounds *a*, *b*, *c* of Table 6). The exocyclic bond angles O(5′)—C(1′)—N(1) and O(5′)—C(5′)—C(6′) are significantly less than the mean tetrahedral value of 109·5°, in molecule *A*, while the exocyclic bond angles C(2′)—C(1′)—N(1) and C(4′)—C(5′)—C(6′) are greater than the tetrahedral value. The water bridge between the base and the carboxyl group in molecule *A* is, perhaps, instrumental in causing the observed distortions in the bond angles around C(1′) and C(5′). In molecule *B*, where there is no such water bridge, the angles around C(5′) and C(1′) are closer to the tetrahedral value.

The C—O bonds of the ionized carboxyl group in molecule *A* are nearly equal, as expected. However, in molecule *B*, the C(6′)—O(6′) bond is 0·019 Å less than the C(6′)—O(4′) bond. This may be attributed to the strong intramolecular hydrogen bonding between the carboxyl O(4′) and the amino N(4′) in molecule *B*, which has also resulted in an alteration in the bonding character by decreasing the bond angle C(5′)—C(6′)—O(4′) to 113·4 (5)° and increasing the bond angle C(5′)—C(6′)—O(6′) to 120·5 (4)°.

The base

The bond distances and angles of the cytosine bases compare well with the reported values for the neutral cytosine derivatives (Rao & Sundaralingam, 1973). The equations to the least-squares planes passing through the bases are given in Table 4. The pyrimidine

---

* Lists of structure factors and anisotropic thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 35013 (9 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 4. *Deviations of atoms from the least-squares planes passing through the base and the carboxyl groups*

Atoms used in fitting the least-squares planes are denoted by daggers. X, Y and Z refer to the orthogonal a, b and c^* axes.

Plane I (molecule *A*)

$$0\cdot0352X + 0\cdot3010Y + 0\cdot9530Z = 6\cdot9636$$

Plane II (molecule *B*)

$$0\cdot0999X - 0\cdot2268Y + 0\cdot9688Z = 12\cdot3704$$

	Molecule *A* (Å)	Molecule *B* (Å)
N(1)†	0·051 (7)	0·025 (7)
C(2)†	−0·041 (8)	0·009 (8)
O(2)	−0·138 (9)	−0·049 (9)
N(3)†	−0·004 (7)	−0·017 (7)
C(4)†	0·040 (8)	0·026 (8)
N(4)	0·125 (10)	0·087 (10)
C(5)†	−0·029 (8)	−0·009 (8)
C(6)†	−0·015 (8)	−0·016 (8)

Plane I (molecule *A*)

$$0\cdot3682X + 0\cdot9010Y - 0\cdot2292Z = 1\cdot5844$$

Plane II (molecule *B*)

$$0\cdot5553X - 0\cdot8279Y - 0\cdot0786Z = 0\cdot7500$$

	Molecule *A* (Å)	Molecule *B* (Å)
O(5′)	1·141 (8)	0·020 (8)
C(5′)†	0·002 (8)	−0·000 (7)
C(6′)†	−0·006 (7)	0·000 (7)
O(4′)†	0·002 (7)	−0·000 (7)
O(6′)†	0·002 (7)	−0·000 (7)

ring is not strictly planar, but has a shallow boat configuration, as is evident from the deviation of N(1) and C(4) on one side of the ring plane while C(2), N(3), C(5) and C(6) deviate on the opposite side. Such nonplanarity has previously been observed in cytosine base (Barker & Marsh, 1964), the nucleosides cytidine (Furberg, Petersen & Rømming, 1965) and deoxycytidine hydrochloride (Subramanian & Hunt, 1970), and the nucleotides cytidine 3′-monophosphate (Sundaralingam & Jensen, 1965) and deoxycytidine 5′-phosphate monohydrate (Viswamitra, Reddy, Lin & Sundaralingam, 1971). Apparently, the nonplanarity of the cytosine base is independent of whether the base is protonated or not.

Molecular conformation

An *ORTEP* (Johnson, 1965) drawing showing the atom numbering and the overall molecular conformation is given in Fig. 4.

Glycosyl conformation

The relative orientation of the base with respect to the sugar is given by the torsion angle (χ) O(5′)—C(1′)—N(1)—C(6) which is 75·0 (6)° in molecule *A* and 71·7 (5)° in molecule *B*. Thus, both molecules exhibit the *anti* conformation, which appears to be a preferred conformation for the pyranosyl as well as the furanosyl nucleosides.

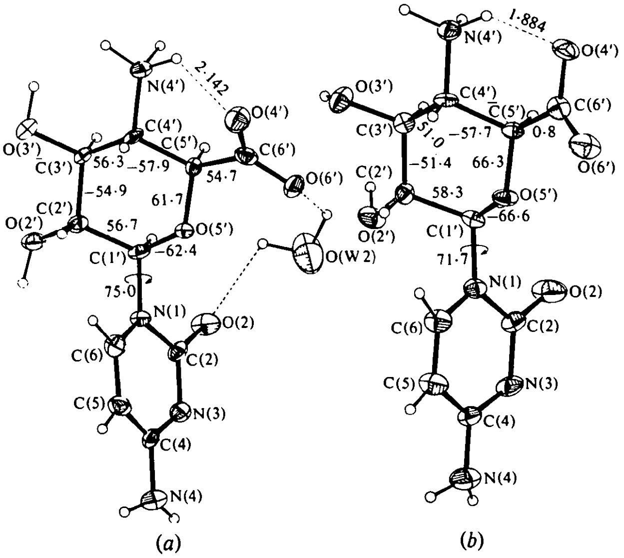


Fig. 4. *ORTEP* (Johnson, 1965) drawings of (*a*) molecule *A* and (*b*) molecule *B* showing the endocyclic torsion angles for the sugar and the torsion angle about the C(5′)—C(6′) bond and the glycosyl bond (°). The thermal ellipsoids represent 50% probability.

Pyranose-ring conformation

The pyranose ring is in the $^4C_1$ conformation in both molecules. The mean torsion angles about the endocyclic bonds of the pyranose moiety in molecule *A* and molecule *B* are 58·3 (5) and 58·6 (5)° respectively (Fig. 4). A comparison of the sugar-ring torsion angles of the two molecules shows that the individual endocyclic torsion angles show significant differences between the two molecules. These differences are possibly generated by the differences in the orientation of the carboxyl group and the presence of strong intramolecular hydrogen bonding in molecule *B*. Thus, the pyranose-ring conformation is flexible over a small range of values within its local-minimum-energy conformation and adapts itself to the intramolecular hydrogen bond. Cremer & Pople (1975) parameters for the two pyranose rings are, $\theta = 0\cdot9°$, $\varphi = 303\cdot6°$, $q_2 = 0\cdot009$ Å, $Q = q_3 = 0\cdot587$ Å (for molecule *A*); $\theta = 7\cdot1°$, $\varphi = 0\cdot7°$, $q_2 = 0\cdot074$ Å, and $Q = q_3 = 0\cdot595$ Å (for molecule *B*), where Q is the total amplitude of puckering, q_2, q_3, φ are the puckering coordinates and θ is the angle between Q and q's, such that, $q_2 = Q \sin \theta$, and $q_3 = Q \cos \theta$. The value of θ equal to 7·1° in molecule *B* indicates that it is quite distorted from the ideal chair conformation. In both molecules the greatest puckering occurs about the C(5′)–O(5′) and O(5′)–C(1′) bonds while the smallest is about the C(2′)–C(3′) and C(3′)–C(4′) bonds.

The torsion angles involving the exocyclic atoms are given in Table 5. In general, these angles are greater than the endocyclic torsion angles. The variation in the exocyclic torsion angles is greater in molecule *B* than in molecule *A*.

The orientations of the H atoms of the hydroxyl groups are given in Table 5. The hydroxyl hydrogen atoms H(O2′) and H(O3′) of molecule *A* and molecule *B* show rather large differences in their orientations with respect to the ring H atoms H(2′) and H(3′).

Conformation of the carboxyl group

As mentioned above, the carboxyl groups are twisted differently about the C(5′)–C(6′) bond in the two molecules. In Table 6 the conformations about the C(5′)–C(6′) bond in different uronic acid structures are listed. In molecule *B* and the calcium salt of D-glucuronic acid (Delucas, Bugg, Terzis & Rivest, 1975) the carboxyl group is nearly in the same plane as O(5′) while in the other compounds listed in Table 6, the carboxyl group is twisted from the plane containing O(5′).

Table 5. *Comparison of the exocyclic torsion angles* (°) *of the pyranose ring in molecule A and molecule B*

The mean standard deviation in the torsion angles is 0·5°.

	Molecule *A*	Molecule *B*
N(1)–C(1′)–C(2′)–O(2′)	−66·9	−58·0
O(2′)–C(2′)–C(3′)–C(4′)	−174·7	−173·8
O(2′)–C(2′)–C(3′)–O(3′)	64·2	66·2
O(3′)–C(3′)–C(4′)–C(5′)	177·1	172·7
O(3′)–C(3′)–C(4′)–N(4′)	−60·9	−64·0
N(4′)–C(4′)–C(5′)–C(6′)	64·4	57·0
N(4′)–C(4′)–C(5′)–O(5′)	179·6	−178·5
C(6′)–C(5′)–O(5′)–C(1′)	−176·4	−168·0
C(5′)–O(5′)–C(1′)–N(1)	175·0	171·7
H(O2′)–O(2′)–C(2′)–H(2′)	−46	132
H(O3′)–O(3′)–C(3′)–H(3′)	−66	28

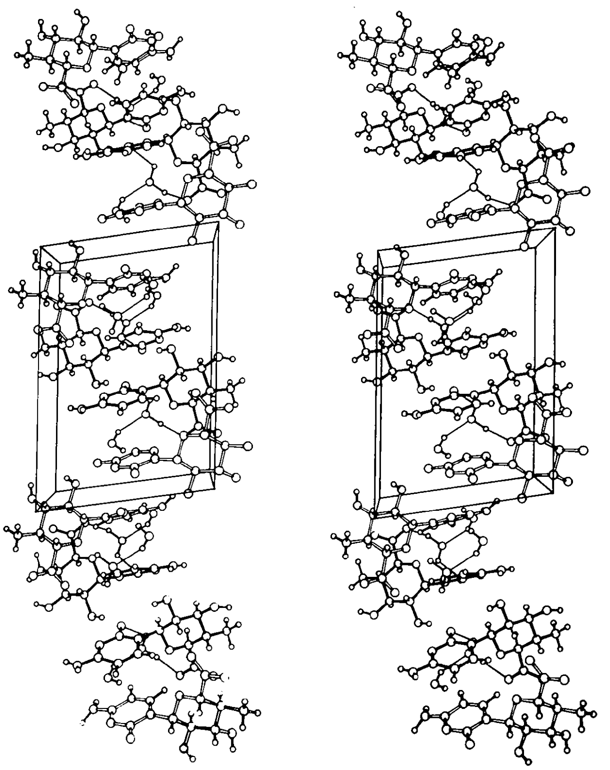


Fig. 5. Packing of the molecules viewed down the *b* axis. The spiralling of the molecule around the screw axis is seen to maximize base stacking.

Molecular packing, hydrogen bonding and base stacking

The molecules are packed in a helical arrangement with the helix axis parallel to the *z* direction (Fig. 5). The two molecules are related by an approximate twofold screw axis along the *z* direction. With reference to a Cartesian system of axes the direction cosines of the pseudo screw axis are 0·00, 0·10, 0·99 and it passes through the point 6·9, 4·8, 4·8 Å. The base planes are approximately perpendicular to the *z* axis. The water molecules sit in the groove of the helical arrangement of the molecules. The bases are stacked along the **c** direction with a separation of 3·7 Å. A view

Table 6. *Conformational angles* (°) *about the exocyclic* C(5′)–C(6′) *bond in different glucuronic acid structures*

	Molecule *A*	Molecule *B*	*a*	*b*	*c*
C(4′)–C(5′)–C(6′)–O(6′)	172·1	124·0	106·0	86·6	124·6
C(4′)–C(5′)–C(6′)–O(4′)	−6·0	−56·0	−73·7	−91·5	−57·3
O(5′)–C(5′)–C(6′)–O(6′)	54·7	0·8	−18·0	−32·7	2·9
O(5′)–C(5′)–C(6′)–O(4′)	−124·1	−179·1	162·2	149·2	−179·0

(*a*) Potassium β-D-glucuronate (Gurr, 1973).
(*b*) Aldotriuronic acid trihydrate (Moran & Richards, 1973).
(*c*) Calcium bromide salt of D-glucuronic acid (Delucas, Bugg, Terzis & Rivest, 1975).

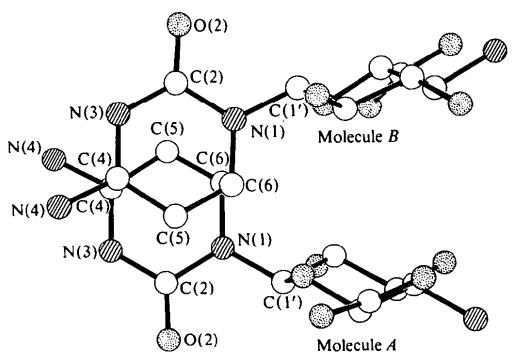
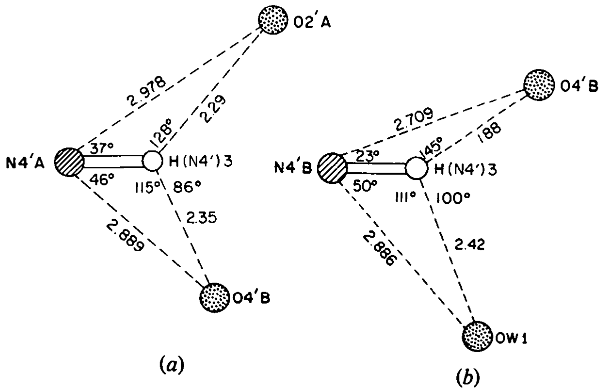


Fig. 6. Stacking of the bases of molecules *A* and *B* viewed normal to the base plane.



(*a*) (*b*)

Fig. 7. The bifurcated hydrogen bonds in (*a*) molecule *A* and (*b*) molecule *B*.

normal to the base plane of the molecule is given in Fig. 6. The angle between the base planes is 5·7°. Because of the steric hindrance caused by the water molecule, the overlap of the base planes is not complete.

The structure is stabilized by intramolecular, intermolecular and bifurcated hydrogen bonds. The hydrogen-bond distances and angles are listed in Table 7. There are 14 protons available for hydrogen bonding and all take part in hydrogen-bond interactions. The exocyclic hydroxyl groups act as both donors and acceptors. The water molecule *W*(2) donates both protons to molecule *A* and accepts one from water molecule *W*(1). The other proton of *W*(1) is donated to the symmetry-related molecule *B*. The intramolecular hydrogen bond in molecule *B* is between amino hydrogen H(N4′)3 and the carboxyl oxygen O(4′).

The intramolecular hydrogen bond in molecule *A* appears to be a weak interaction, as judged by the distances and angles involving the donor and acceptor atoms. There are two bifurcated hydrogen bonds observed in the structure (Fig. 7*a*,*b*). In the first, the amino hydrogen H(N4′)3 of molecule *A* is shared by the oxygens O(2′)*A* and O(4′)*B*. In the second the amino hydrogen H(N4′)3 of molecule *B* is shared between O(4′)*B* and O(*W*1). The geometries of these hydrogen bonds agree well with the values reported by

Table 7. *Hydrogen-bond distances* (Å) *and angles* (°) *in C-substance*

D–H···*A*	Symmetry code*	Translation for *A* (*x*,*y*,*z*)	*D*–H	*D*···*A*	H···*A*	∠*D*–H···*A*
O(2′)*A*–H(O2′)···O(2)*A*	2	(1,0,1)	0·91	2·772	1·88	171
O(3′)*A*–H(O3′)···O(6′)*B*	1	(−1,0,0)	0·77	2·745	2·46	104
N(4′)*A*–H(N4′)1···N(3)*A*	1	(−1,0,0)	0·88	2·968	2·09	170
N(4′)*A*–H(N4′)2···O(4′)*A*	1	(0,0,0)	0·96	2·850	2·14	129
N(4′)*A*–H(N4′)3···O(4′)*B*	2	(1,−1,1)	0·96	2·889	2·35	115
N(4′)*A*–H(N4′)3···O(2′)*A*	2	(0,−1,1)	0·96	2·978	2·29	128
N(4)*A*–H(N4)1···O(4′)*B*	2	(2,0,1)	0·93	3·012	2·41	122
N(4)*A*–H(N4)2···O(3′)*A*	1	(1,0,0)	0·85	3·030	2·27	150
O(2′)*B*–H(O2′)···O(4′)*A*	1	(1,0,1)	0·74	2·675	1·96	165
O(3′)*B*–H(O3′)···O(6′)*A*	1	(1,0,1)	0·76	2·692	1·93	179
N(4′)*B*–H(N4′)1···N(3)*B*	1	(1,0,0)	0·91	2·850	1·95	171
N(4′)*B*–H(N4′)2···O(2′)*B*	2	(2,0,2)	0·90	2·803	1·93	163
N(4′)*B*–H(N4′)3···O(4′)*B*	1	(0,0,0)	0·94	2·709	1·88	145
N(4′)*B*–H(N4′)3···O(*W*1)	2	(2,0,1)	0·94	2·886	2·42	111
N(4)*B*–H(N4)1···O(*W*1)	2	(1,−1,1)	0·91	3·151	2·28	159
N(4)*B*–H(N4)2···O(3′)*B*	1	(−1,0,0)	0·89	3·058	2·18	174
O(*W*2)–H(*W*2)1···O(6′)*A*	1	(0,0,0)	1·05	2·947	2·14	133
O(*W*2)–H(*W*2)2···O(2)*A*	1	(0,0,0)	1·02	2·827	2·19	119
O(*W*1)–H(*W*1)1···O(*W*2)	1	(0,0,0)	0·74	2·825	2·14	154
†O(*W*1)–H(*W*1)2···O(2)*B*	2	(1,−1,1)	1·05	2·788	1·75	173
Mean σ			(0·06)	(0·006)	(0·06)	(5)

* Symmetry codes are (1) *x*,*y*,*z*; (2) $-x, \frac{1}{2}+y, -z$.

† The hydrogen was fixed geometrically.

Jeffrey & Takagi (1978) for bifurcated hydrogen bonds. Apart from the two hydrogen-bonding neighbors O(3′)*B* and O(*W*1), N(4)*B* has a third neighbor O(4′)*A* at a distance of 2·970 Å. The H(N4)···O(4′)*A* distance is 2·88 Å and the angle N(4)*B*—H(N4)1···O(4′)*A* is 88°. The geometry suggests that the interaction between N(4)*B* and O(4′)*A* is probably not a hydrogen bond. A short intermolecular C—H···O contact is also observed in the structure. C(6) of molecule *B* is 3·01 Å from O(5′) of molecule *A*, the H(6)···O(5′) distance is 2·39 Å, and the angle C(6)—H(6)···O(5′)*A* is 124°.

We gratefully thank Dr K. A. Watanabe for kindly supplying a sample of *C*-substance. This work was supported by a grant (GM-17378) from the National Institutes of Health of the US Public Health Service. Continued support by the College of Agricultural and Life Sciences of the University of Wisconsin–Madison is hereby acknowledged.

References

BARKER, D. L. & MARSH, R. E. (1964). *Acta Cryst.* **17**, 1581–1587.

BERMAN, H. M., CHU, S. S. C. & JEFFREY, G. A. (1967). *Science*, **157**, 1576–1577.

BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. Report ORNL-TM-305. Oak Ridge National Laboratory, Tennessee.

CLARK, J. M. JR & CHANG, A. Y. (1965). *J. Biol. Chem.* **240**, 4734–4739.

CLARK, J. M. JR & GUNTHER, J. K. (1963). *Biochim. Biophys. Acta*, **76**, 636–638.

CREMER, D. & POPLE, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–111.

CRUICKSHANK, D. W. J., PILLING, P. E., BUJOSA, A., LOVELL, F. M. & TRUTER, M. R. (1961). *Computing Methods and the Phase Problems in X-ray Crystal Analysis*, p. 32. New York: Pergamon Press.

DEBYE, P. (1915). *Ann. Phys.* (*Leipzig*), **46**, 809–823.

DELUCAS, L., BUGG, C. E., TERZIS, A. & RIVEST, R. (1975). *Carbohydr. Res.* **41**, 19–29.

FOX, J. J., WATANABE, K. A. & BLOCH, A. (1966). *Prog. Nucleic Acid Res. Mol. Biol.* **5**, 251–313.

FRIES, D. C., RAO, S. T. & SUNDARALINGAM, M. (1971). *Acta Cryst.* B**27**, 994–1005.

FURBERG, S., PETERSEN, C. S. & RØMMING, C. (1965). *Acta Cryst.* **18**, 313–320.

GURR, G. E. (1973). *Acta Cryst.* **16**, 690–696.

International Tables for X-ray Crystallography (1962). Vol. III, p. 202. Birmingham: Kynoch Press.

JEFFREY, G. A., POPLE, J. A., BINKLEY, J. S. & VISHVESHWARA, S. (1978). *J. Am. Chem. Soc.* **100**, 373–379.

JEFFREY, G. A. & TAKAGI, S. (1978). *Acc. Chem. Res.* pp. 264–270.

JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.

LIN, G. H.-Y., SUNDARALINGAM, M. & ARORA, S. K. (1971). *J. Am. Chem. Soc.* **93**, 1235–1241.

MAIN, P. (1975). International Summer School on Crystallographic Computing, Prague, paper No. 48.

MAIN, P., WOOLFSON, M. M., LESSINGER, L., GERMAIN, G. & DECLERCQ, J. P. (1974). *MULTAN* 74. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univs. of York, England, and Louvain, Belgium.

MORAN, R. A. & RICHARDS, G. F. (1973). *Acta Cryst.* B**29**, 2770–2783.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. C. (1968). *Acta Cryst.* A**24**, 351–359.

RAO, S. T. & SUNDARALINGAM, M. (1973). *Synthetic Procedures in Nucleic Acid Chemistry*, Vol. II, edited by W. W. ZORBACH & S. TIPSON. New York: John Wiley.

SINOHARA, H. & SKYPECK, H. H. (1965). *Biochem. Biophys. Res. Commun.* **18**, 98–102.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

SUBRAMANIAN, E. & HUNT, D. J. (1970). *Acta Cryst.* B**26**, 303–311.

SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 914–929, 930–943.

SWAMINATHAN, P. (1975). *Acta Cryst.* B**31**, 1608–1613.

VISWAMITRA, M. A., REDDY, B. S., LIN, G. H.-Y. & SUNDARALINGAM, M. (1971). *J. Am. Chem. Soc.* **93**, 4565–4573.