

## The Crystal Structure and Absolute Configuration of 1-( $\beta$ -D-Arabinofuranosyl)cytosine Hydrochloride

BY JAMES S. SHERFINSKI AND RICHARD E. MARSH

Arthur Amos Noyes Laboratory of Chemical Physics,\*

California Institute of Technology, Pasadena, California 91109, U.S.A.

(Received 31 July 1972; accepted 4 October 1972)

Crystals of 1-( $\beta$ -D-arabinofuranosyl)cytosine hydrochloride,  $C_9H_{14}N_3O_5Cl$ , are monoclinic, space group  $P2_1$ , with cell dimensions  $a = 6.540$  (2),  $b = 15.260$  (2),  $c = 6.758$  (3) Å,  $\beta = 117.68$  (3)°. Intensity data for two quadrants of reciprocal space were collected using Cu  $K\alpha$  radiation and an automated diffractometer. The structure was solved by Patterson-Fourier techniques and refined by least squares to an  $R$  index of 0.022 for 2490 observed reflections. The absolute configuration of the sugar ring was confirmed utilizing the anomalous scattering by Cl, O, and N. The arabinose ring is puckered with C(2') *endo* and the conformation about the C(4')-C(5') bond is *gauche-trans*; the torsion angle about the glycosidic bond is  $+26.7^\circ$ . These conformational features as well as some anomalies in the bond distances and angles can be explained by the short O(2')...N(1) distance of 2.75 Å. The hydrogen-bonding scheme involving the cytosine ring and the chloride ion is similar to that found in several other salts of cytosine derivatives.

### Introduction

The compound 1-( $\beta$ -D-arabinofuranosyl)cytosine hydrochloride has been known for some time to be effective in inhibiting DNA synthesis. Evans, Musser, Mengel, Forsblad & Hunter (1961) demonstrated the antitumor capabilities of the compound against various carcinomas and leukemias. Its mode of action was somewhat elucidated by Chu & Fisher (1962), who showed that cytosine arabinoside inhibits the reduction of cytidylic acid to deoxycytidine. Recently Momparler (1969) found that there is some uptake of the 5'-triphosphate ester of cytosine arabinoside into *in vitro*-synthesized DNA and that this base is located primarily at the 3'-terminus. Since the polymer synthesis is 5' to 3', polymerization must cease upon incorporation of the arabinoside base. Therefore, since 1-( $\beta$ -D-arabinofuranosyl)cytosine can apparently act as a substrate for cytidylate reductase and DNA polymerase, we thought it important to compare its structural features with those already known for cytidylate nucleotides and nucleosides.

### Experimental

Crystals of 1-( $\beta$ -D-arabinofuranosyl)cytosine hydrochloride,  $(C_9H_{14}N_3O_5)^+ \cdot Cl^-$ , were kindly furnished by the Cancer Chemotherapy National Service Center. A small sample was dissolved in a minimum amount of distilled water at room temperature, and this solution was allowed to evaporate almost to dryness. Several

needle-shaped crystals elongated along  $c$  were isolated, although the majority of the sample precipitated in the form of a clear, colorless multicrystalline cake. One approximately triangular prism,  $0.37 \times 0.12 \times 0.15$  mm in size, was mounted along the needle axis. Oscillation, Weissenberg, and precession photographs showed Laue symmetry  $2/m$  and systematic absences  $0k0$  with  $k$  odd, indicating the space group  $P2_1$  or  $P2_1/m$ ; the former was chosen because of the presumed optical purity of the compound and the measured density (by flotation in a benzene/carbon tetrachloride solution), which indicated only two molecules in the cell. This same crystal was then mounted on a modified Daxex-automated General Electric XRD-5 quarter-circle diffractometer equipped with a copper X-ray tube and scintillation counter. Monochromatization was by means of a nickel filter and a pulse-height analyzer. Cell dimensions were calculated from a least-squares fit to the observed  $2\theta$  values for 10 high-angle reflections. Crystal data are listed in Table 1.

Table 1. Crystal data

$C_9H_{14}N_3O_5^+ \cdot Cl^-$	F.W. 279.7
Monoclinic	Space group $P2_1$
$a = 6.540$ (2) Å	$Z = 2$
$b = 15.260$ (2)	$F(000) = 292$
$c = 6.758$ (3)	$V = 597.3$ Å <sup>3</sup>
$\beta = 117.68$ (3)°	$D_x = 1.551$ g cm <sup>-3</sup>
$\lambda(\text{Cu } K\alpha) = 1.5418$ Å	$D_m = 1.555$ (2)
$\mu = 30.5$ cm <sup>-1</sup>	

Intensity data from both the  $hkl$ - $\bar{h}kl$  and  $h\bar{k}l$ - $\bar{h}\bar{k}l$  quadrants were collected using a  $\theta$ - $2\theta$  scan at a scan speed of  $1^\circ/\text{min}$  in  $2\theta$ . The scan width varied linearly from  $2.3^\circ$  at  $2\theta = 4^\circ$  to  $3.5^\circ$  at  $2\theta = 150^\circ$ . Backgrounds

\* Contribution No. 4572 from the Arthur Amos Noyes Laboratory of Chemical Physics. This investigation was supported in part by Public Health Service Research Grant No. GM 16966, from the National Institute of General Medical Sciences.

were counted for 20 sec at the beginning and end of each scan. Three check reflections  $\bar{2}, 10, 4, \bar{1}, 11, 2$  and 250 – were measured periodically and showed no significant change in intensity during the eight days of data collection. 2544 reflections with  $2\theta$  less than  $155^\circ$  were measured. The intensities were corrected for Lorentz and polarization effects and placed on an approximately absolute scale by a Wilson (1942) plot. Observational variances  $\sigma^2(F_o^2)$  included counting statistics for the scan and background plus an additional term  $(0.02s)^2$ , where  $s$  is the scan count. Structure factors smaller than  $3\sigma(F_o)$ , 50 in number, were assigned weights of zero. Four reflections (130,  $\bar{1}\bar{3}0$ , 011,  $0\bar{1}\bar{1}$ ) with counting rates apparently exceeding the linear range of the detector were also given zero weight; the  $3\bar{3}\bar{2}$  reflection was not measured. No correction was made for absorption. Intensity transmission factors for a random set of reflections varied from 0.58 to 0.64.

### Structure solution and refinement

The positions of the chloride ion and of two other atoms 4–5 Å away were assigned from a three-dimensional Patterson map calculated using one quadrant of data, and four structure-factor Fourier-map cycles revealed the remaining heavy atoms.

Refinement was by least-squares minimization of the quantity  $\sum w(F_o^2 - F_c^2/k^2)^2$ , where  $k$  is the scale factor and weights  $w$  were taken equal to  $\sigma^{-2}(F_o^2)$ . Form factors for C, N, O, and  $\text{Cl}^-$  were taken from *International Tables for X-ray Crystallography* (1962); those for H were from Stewart, Davidson & Simpson (1965). All computations were carried out under the CRYM system on an IBM 360/75 computer.

Initial refinement was based on one quadrant of data. After four cycles of isotropic refinement ( $R = 0.081$ ), a difference map indicated the positions of all the hydrogen atoms except those of the hydroxyl groups, appearing as peaks ranging from 0.2 to 0.5  $\text{e}\text{\AA}^{-3}$ . With the located hydrogen atoms contributing to  $F_c$ , the heavy atoms were then refined anisotropically for two cycles. A subsequent difference map clearly indicated the positions of the remaining hydrogen atoms. Further refinement, with a parameter list that included the coordinates of all 32 atoms, anisotropic temperature parameters for the 18 heavy atoms, isotropic temperature parameters for the 14 hydrogen atoms, a scale factor and an isotropic extinction parameter (Larson, 1967, equation 3), converged at an  $R$  index of 0.025.

The atomic form factors of O, N, and  $\text{Cl}^-$  were then corrected for the real and imaginary components of anomalous dispersion, using values for Cl from *International Tables for X-ray Crystallography* (1962) and for O and N from Hope (1969, private communication: for oxygen,  $\Delta f' = 0.049$ ,  $\Delta f'' = 0.032$ ; for nitrogen,

Table 2. *Enantiomorph determination*

	D	L
$R = \frac{\sum   F_c  -  F_o  }{\sum  F_o }$	0.023	0.032
$wR_2 = \frac{\sum w  F_c ^2 -  F_o ^2 ^2}{\sum w F_o ^4}$	0.003	0.006
Goodness-of-fit $\left( \frac{\sum w  F_o ^2 -  F_c ^2 }{(M-S)} \right)^{1/2}$	2.04	2.59
	$M = 1293$ observations $S = 219$ parameters	

Table 3. *Final heavy-atom parameters and their estimated standard deviations*

All parameters have been multiplied by  $10^5$ . The temperature factor is of the form

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

The final value of the extinction parameter  $g$  is  $(76.1 \pm 2.7) \times 10^{-6} \text{e}^{-2}$ .

	$x$	$y$	$z$	$b_{11}$	$b_{22}$	$b_{33}$	$b_{12}$	$b_{13}$	$b_{23}$
$\text{Cl}^-$	-26530 (6)	50000 (0)	-13406 (6)	2160 (10)	296 (1)	2714 (11)	-73 (7)	2685 (16)	-106 (7)
O(1')	25258 (17)	76721 (7)	20214 (19)	1852 (27)	238 (4)	2969 (33)	-33 (17)	3240 (51)	20 (18)
C(1')	40853 (21)	76350 (9)	11277 (22)	1401 (33)	256 (5)	1730 (34)	-64 (21)	1600 (56)	116 (21)
C(2')	47587 (24)	66651 (9)	11853 (24)	1768 (35)	255 (5)	2113 (40)	-92 (23)	2129 (62)	-164 (23)
O(2')	64450 (17)	64154 (7)	33326 (19)	1678 (27)	283 (4)	2866 (36)	172 (17)	1697 (51)	288 (19)
C(3')	25416 (25)	62108 (9)	8044 (24)	1943 (40)	275 (5)	1937 (37)	-275 (23)	1937 (63)	-106 (24)
O(3')	10332 (22)	61782 (9)	-15191 (19)	2617 (37)	512 (6)	1966 (32)	-1068 (24)	2120 (56)	-470 (22)
C(4')	15902 (21)	68121 (9)	19926 (23)	1386 (32)	258 (6)	1891 (38)	-107 (21)	1611 (57)	130 (22)
C(5')	21570 (27)	65302 (11)	43272 (27)	2178 (42)	359 (7)	2150 (43)	195 (27)	2423 (72)	332 (26)
O(5')	8753 (21)	70746 (8)	50562 (20)	2414 (34)	472 (6)	2235 (32)	212 (22)	2751 (57)	71 (22)
N(1)	60948 (18)	81903 (7)	25060 (19)	1596 (30)	235 (4)	1652 (29)	-100 (19)	1787 (51)	27 (18)
C(2)	73380 (23)	85045 (9)	14740 (23)	1702 (37)	278 (5)	1838 (38)	-87 (23)	1892 (63)	123 (23)
O(2)	68073 (20)	83654 (9)	-4678 (18)	2480 (37)	609 (7)	1776 (31)	-761 (25)	2342 (56)	-190 (23)
N(3)	92733 (20)	89906 (8)	28434 (20)	1802 (32)	286 (5)	1995 (35)	-267 (20)	2011 (56)	86 (20)
C(4)	100048 (24)	91468 (9)	50346 (24)	1827 (35)	213 (5)	2072 (38)	-44 (21)	1699 (61)	-17 (21)
N(4)	119094 (25)	95996 (10)	61689 (25)	2388 (41)	337 (6)	2652 (44)	-606 (26)	1740 (69)	-197 (25)
C(5)	86540 (27)	88171 (11)	60039 (25)	2436 (42)	291 (5)	1850 (40)	-153 (26)	2148 (69)	-283 (24)
C(6)	67449 (26)	83521 (9)	47093 (25)	2177 (44)	271 (6)	1954 (39)	-84 (24)	2486 (69)	-54 (23)

$\Delta f' = 0.032$ ,  $\Delta f'' = 0.019$ ). Refinements of both enantiomers were continued until parameter shifts were less than  $0.3\sigma$  for the heavy atoms and less than  $1.0\sigma$  for the hydrogen atoms. Results of the two refinements are summarized in Table 2. A comparison of intensity transmission factors (see *Experimental*) for representative  $hkl$  and corresponding  $h\bar{k}l$  reflections showed relative differences of no more than 4–5%. Moreover, the absorption effects on intensities are systematic throughout reciprocal space, while the Bijvoet differences are random. Accordingly, we feel that the differences in statistical sums are sufficient to establish the enantiomorph.

Coordinates and temperature parameters for the two refinements agreed within two standard deviations except for the  $y$  coordinates of the lighter atoms, which differed by an approximately constant amount of  $0.045 \text{ \AA}$ . A polar dispersion error of this magnitude was predicted by Cruickshank & McDonald (1967) as the result of refining the incorrect enantiomorph using the chlorine scattering factor corrected for anomalous scattering in Cu  $K\alpha$  radiation.

Final least-squares refinement of the D enantiomer was based on the full data set (two quadrants). Two matrices were collected, one containing the coordinates and anisotropic temperature coefficients of the heavy atoms, scale factor, and extinction parameter and the second containing the coordinates and isotropic temperature parameters of the hydrogen atoms. After two cycles, the largest parameter shifts associated with the two matrices were  $0.4\sigma$  and  $1.1\sigma$ , the  $R$  index was  $0.022$ , the weighted  $R$  was  $0.003$ , and the goodness-of-fit was  $1.83$ . The final parameters are listed in Tables 3 and 4; observed and calculated structure factors are given in Table 5.

### Discussion

The 1-( $\beta$ -D-arabinofuranosyl)cytosine cation is shown in Fig. 1. It differs from the cytidine cation only in

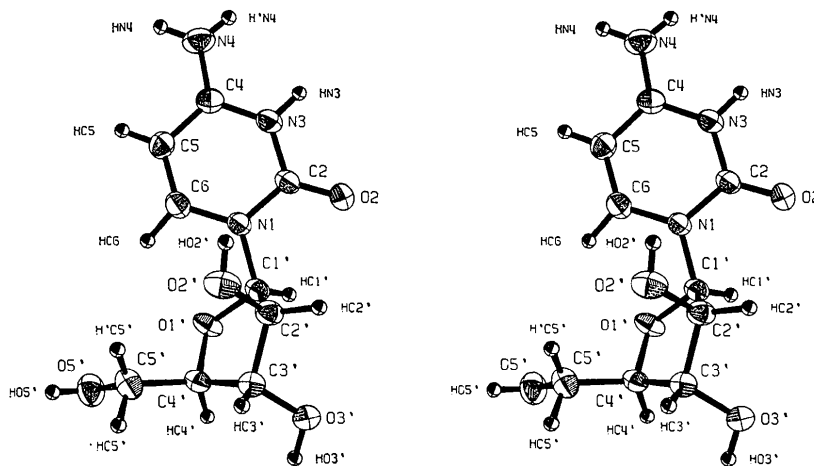


Fig. 1. A stereographic illustration (Johnson, 1965) of the 1-( $\beta$ -D-arabinofuranosyl)cytosine cation. Thermal ellipsoids are drawn at the 50% probability level. Hydrogen atoms are represented as spheres  $0.10 \text{ \AA}$  in radius.

Table 4. *Hydrogen atom parameters and their estimated standard deviations*

Positional parameters have been multiplied by  $10^3$ .

	$x$	$y$	$z$	$B(\text{\AA}^2)$
HC(1')	342 (3)	787 (1)	-24 (3)	2.8 (3)
HC(2')	518 (3)	650 (1)	-3 (3)	2.6 (3)
HO(2')	768 (3)	671 (1)	375 (3)	4.0 (4)
HC(3')	290 (3)	561 (1)	135 (3)	3.6 (4)
HO(3')	-16 (4)	591 (1)	-165 (3)	3.7 (4)
HC(4')	-6 (3)	682 (1)	101 (3)	3.8 (4)
HC(5')	160 (4)	593 (1)	439 (3)	3.9 (4)
H'C(5')	368 (3)	653 (1)	529 (3)	3.2 (4)
HO(5')	130 (5)	687 (2)	640 (4)	6.4 (6)
HN(3)	1011 (4)	920 (2)	221 (4)	4.4 (5)
HN(4)	1233 (4)	977 (2)	756 (4)	5.9 (6)
H'N(4)	1256 (4)	982 (2)	549 (4)	4.4 (5)
HC(5)	918 (4)	893 (2)	748 (4)	5.1 (5)
HC(6)	565 (3)	811 (1)	524 (3)	3.3 (4)

the configuration at the carbon atom C(2'). This difference – the interchange of an OH group and an H atom at C(2') – has rather important effects on the molecular conformation, for in 1-( $\beta$ -D-arabinofuranosyl)cytosine the OH group, O(2'), is in the position *cis* [relative to the C(1')–C(2') bond] to N(1) and hence is subject to close interactions with the pyrimidine ring. The following conformational features seem to be directly related to this change in configuration:

(1) The conformation about the glycosidic bond is *anti*, the torsion angle C(6)–N(1)–C(1')–O(1') being  $+26.7^\circ$ . A *syn* conformation would result in a short O(2)  $\cdots$  O(1') or O(2)  $\cdots$  O(2') contact, and other values for the torsion angle would result in a short O(1')  $\cdots$  H or O(2)  $\cdots$  H contact. A view along the glycosidic bond is shown in Fig. 2.

(2) The puckering of the sugar ring is C(2')–*endo*. This puckering helps to increase the O(2')  $\cdots$  N(1) separation.

(3) The conformation about the exocyclic bond C(4')–C(5') is *gauche-trans* (Fig. 3). The more common *gauche-gauche* conformation (see Sundaralingam

Table 5. Observed and calculated structure factors

Each set of four columns contains, from left to right, the  $k$  index,  $10|F_o|$ ,  $10|F_c|$ , and  $10(|F_o|^2 - |F_c|^2)/\sigma(|F_o|^2)$ . A reflection with (-) in the fourth column was given zero weight in the least-squares refinement.

-1 k 2	11 37 37 0	12 25 26 14	13 25 27 17	14 25 28 10	15 25 28 10	16 25 28 10	17 25 27 17	18 25 28 10	19 25 28 10	20 25 28 10	21 25 28 10	22 25 28 10	23 25 28 10	24 25 28 10	25 25 28 10	26 25 28 10	27 25 28 10	28 25 28 10	29 25 28 10	30 25 28 10	31 25 28 10	32 25 28 10	33 25 28 10	34 25 28 10	35 25 28 10	36 25 28 10	37 25 28 10	38 25 28 10	39 25 28 10	40 25 28 10	41 25 28 10	42 25 28 10	43 25 28 10	44 25 28 10	45 25 28 10	46 25 28 10	47 25 28 10	48 25 28 10	49 25 28 10	50 25 28 10	51 25 28 10	52 25 28 10	53 25 28 10	54 25 28 10	55 25 28 10	56 25 28 10	57 25 28 10	58 25 28 10	59 25 28 10	60 25 28 10	61 25 28 10	62 25 28 10	63 25 28 10	64 25 28 10	65 25 28 10	66 25 28 10	67 25 28 10	68 25 28 10	69 25 28 10	70 25 28 10	71 25 28 10	72 25 28 10	73 25 28 10	74 25 28 10	75 25 28 10	76 25 28 10	77 25 28 10	78 25 28 10	79 25 28 10	80 25 28 10	81 25 28 10	82 25 28 10	83 25 28 10	84 25 28 10	85 25 28 10	86 25 28 10	87 25 28 10	88 25 28 10	89 25 28 10	90 25 28 10	91 25 28 10	92 25 28 10	93 25 28 10	94 25 28 10	95 25 28 10	96 25 28 10	97 25 28 10	98 25 28 10	99 25 28 10	100 25 28 10
--------	------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	--------------

& Jensen, 1965b; Shefter & Trueblood, 1965) would result in a close O(5')...O(2') contact of about 2.5-2.6 Å.

A preliminary report on the structure of 5-bromouracil arabinoside (Tougaard, 1969) shows this similar nucleoside to be in the *anti* and *gauche-gauche* con-

formations, with the puckering of the arabinose ring quite different.

Values for the important torsion angles are given in Table 6. Bond distances and angles involving the heavy atoms are shown in Fig. 4. The largest deviations from standard values – the C(1')–C(2') distance, which is about 0.03 Å longer than in cytidylic acid (Sundaralingam & Jensen, 1965*a*; Bugg & Marsh, 1967) and 2'-deoxycytidine hydrochloride (Subramanian & Hunt, 1970), and the bond angles O(2')–C(2')–C(1') and C(2')–C(1')–N(1) – seem to reflect the uncomfortably close contact of 2.754 Å between N(1) and O(2'), as does the relatively large deviation of C(1') from the plane of the pyrimidine ring (see Table 7).

Table 6. *Torsion angles*

A positive angle corresponds to a right-handed screw (Klyne & Prelog, 1960). Estimated standard deviations are in the range 0.1–0.2°.

C(6)–N(1)–C(1')–O(1')	26.7°
O(1')–C(1')–C(2')–C(3')	25.2
C(1')–C(2')–C(3')–C(4')	–34.3
C(2')–C(3')–C(4')–O(1')	32.8
C(3')–C(4')–O(1')–C(1')	–4.6
C(4')–O(1')–C(1')–C(2')	–17.9
O(5')–C(5')–C(4')–C(3')	–171.5
O(5')–C(5')–C(4')–O(1')	68.7

Table 7. *Deviations from the plane of the pyrimidine ring*

The least-squares plane was passed through the six ring atoms, N(1)–C(6), all weighted equally. Its direction cosines relative to *a*, *b*, and *c* are –0.486, 0.839, and 0.007; the origin-to-plane distance is 8.566 Å.

	Dev. (Å)		Dev. (Å)		Dev. (Å)
N(1)	–0.006	C(1')	–0.085	HC(5)	–0.019
C(2)	–0.003	O(2)	–0.022	HC(6)	0.049
N(3)	0.011	N(4)	–0.030	Cl( <i>a</i> )	0.272
C(4)	–0.010	HN(3)	–0.005	Cl( <i>b</i> )	0.223
C(5)	0.002	HN(4)	0.062	O(2) ( <i>c</i> )	0.027
C(6)	0.006	H'N(4)	0.048		

- (*a*) at  $1-x, \frac{1}{2}+y, 1-z$   
 (*b*) at  $1-x, \frac{1}{2}+y, -z$   
 (*c*) at  $x, y, 1+z$

Distances and angles involving the hydrogen atoms are listed in Table 8. Formal standard deviations are approximately 0.02–0.03 Å and 1–2°; however, the spread among values expected to be equivalent suggests that these standard deviations should be approximately doubled.

A stereoscopic view of the structure is shown in Fig. 5. Of the three hydroxyl groups, two – O(2') and O(5') – donate hydrogen bonds to other hydroxyl groups; O(3') is hydrogen-bonded to a chloride ion. All three protons associated with nitrogen atoms form hydrogen bonds with chloride ions, resulting in an irregular fourfold coordination about the chloride ion (Fig. 6). Details of the hydrogen bonds are given in Table 9.

Table 8. *Bond distances and angles involving the hydrogen atoms*

Formal e.s.d.'s are about 0.02–0.03 Å and 1.5–2.0°.

HN(3)–N(3)	0.90 Å	HN(3)–N(3)–C(2)	116°
HN(4)–N(4)	0.89	HN(3)–N(3)–C(4)	119
H'N(4)–N(4)	0.83	HN(4)–N(4)–C(4)	121
HC(5)–C(5)	0.90	H'N(4)–N(4)–C(4)	119
HC(6)–C(6)	1.01	H'N(4)–N(4)–HN(4)	118
HC(1')–C(1')	0.90	HC(5)–C(5)–C(4)	116
HC(2')–C(2')	1.01	HC(5)–C(5)–C(6)	125
HC(3')–C(3')	0.98	HC(6)–C(6)–C(5)	124
HC(4')–C(4')	0.97	HC(6)–C(6)–N(1)	114
HC(5')–C(5')	1.00	HC(1')–C(1')–N(1)	106
H'C(5')–C(5')	0.90	HC(1')–C(1')–C(2')	114
HO(2')–O(2')	0.85	HC(1')–C(1')–O(1')	109
HO(3')–O(3')	0.85	HC(2')–C(2')–C(1')	114
HO(5')–O(5')	0.88	HC(2')–C(2')–C(3')	110
		HC(2')–C(2')–O(2')	113
		HC(3')–C(3')–C(2')	109
		HC(3')–C(3')–O(3')	108
		HC(3')–C(3')–C(4')	117
		HC(4')–C(4')–C(3')	104
		HC(4')–C(4')–O(1')	109
		HC(4')–C(4')–C(5')	112
		HC(5')–C(5')–C(4')	113
		HC(5')–C(5')–O(5')	104
		H'C(5')–C(5')–HC(5')	105
		H'C(5')–C(5')–O(5')	113
		H'C(5')–C(5')–C(4')	114
		HO(2')–O(2')–C(2')	112
		HO(3')–O(3')–C(3')	105
		HO(5')–O(5')–C(5')	102

The arrangement of cytosine rings and chloride ions (Fig. 7, Table 10) seems to be characteristic of hydrochloride salts of cytosine derivatives. Essentially the identical arrangement is found in crystals of 1-methylcytosine hydrochloride (Trus & Marsh, 1972) and deoxycytidine hydrochloride (Subramanian & Hunt,

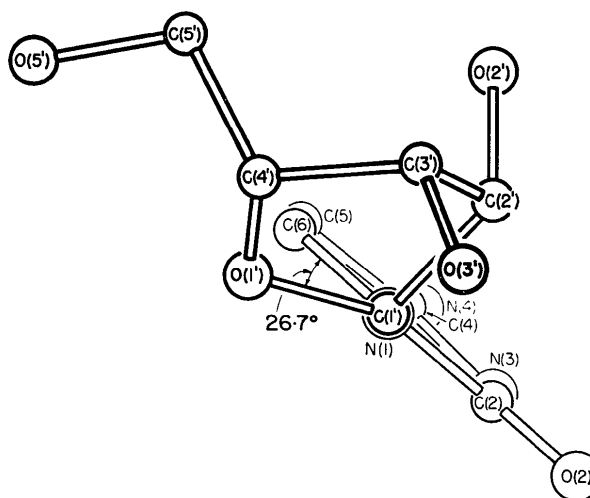


Fig. 2. A view down the C(1')–N(1) bond. The sign of the torsion angle is positive for a right-handed screw (Klyne & Prelog, 1960). This convention differs from that proposed by Donohue & Trueblood (1960) and Sundaralingam & Jensen (1965*a*) for glycosidic torsion angles, but is in accordance with most other torsion-angle sign conventions.

1970). [A similar relationship is found in the crystals of 1-methylcytosine hydrobromide (Bryan & Tomita, 1962), except that neighboring cytosine cations are

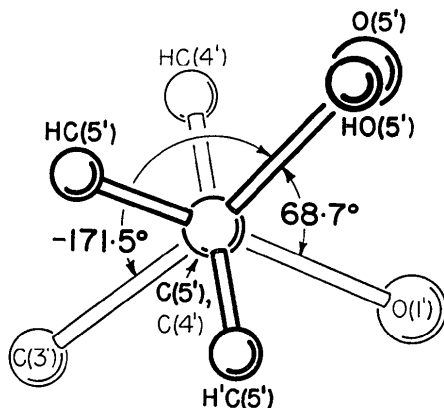


Fig. 3. A view down the C(5')-C(4') bond.

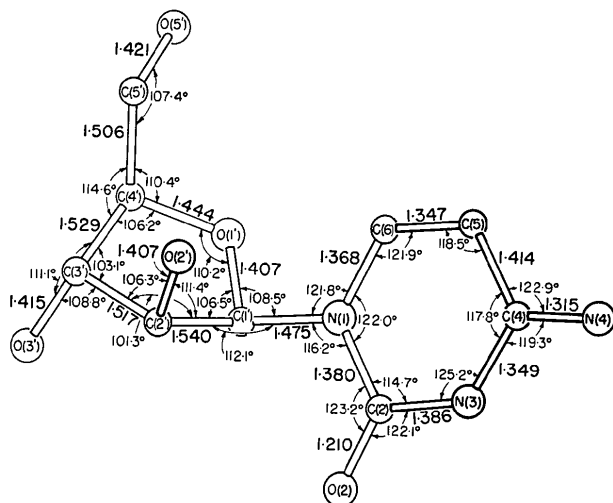


Fig. 4. Bond distances and angles in the 1-( $\beta$ -D-arabinofuranosyl)cytosine cation. Estimated standard deviations are about 0.002 Å and 0.2°.

Table 9. Distances and angles for the hydrogen bonds, D-H...A

D	H	A	Distance D-A	Distance H...A	Angle D-H...A
N(4)	HN(4)	Cl <sup>-</sup> (a)	3.354 Å	2.49 Å	165°
N(4)	H'N(4)	Cl <sup>-</sup> (b)	3.568	2.84	147
N(3)	HN(3)	Cl <sup>-</sup> (b)	3.221	2.35	165
O(3')	HO(3')	Cl <sup>-</sup>	3.054	2.22	166
O(5')	HO(5')	O(3') (c)	2.649	1.83	155
O(2')	HO(2')	O(5') (d)	2.762	1.93	165

(a) at  $1-x, \frac{1}{2}+y, 1-z$

(b) at  $1-x, \frac{1}{2}+y, -z$

(c) at  $x, y, 1+z$

(d) at  $1+x, y, z$

related by a  $2_1$  axis rather than by a translation.] This arrangement features, besides a continuous chain of NH-Cl hydrogen bonds, the interesting interactions C(5)-H...O(2) and C(6)-H...O(2). These same C-H...O contacts are also found in the crystal structures of the two forms of cytidylic acid (Sundaralingam & Jensen, 1965a; Bugg & Marsh, 1967), although the anionic moiety involved in the hydrogen bonding scheme is in those cases a phosphate group. All of these H...O distances are just in the range of the sum of van der Waals radii (Pauling, 1960).

Thermal ellipsoids for the cytosine cation and the chloride ion are shown in Figs. 1 and 6. The largest mean-square displacements are exhibited by atoms O(3') and O(2) and are directed approximately along the view vector in Fig. 1. The orientation of the chloride thermal ellipsoid seems to be primarily determined by its two shortest hydrogen bonds, involving HN(3) and HO(3').

*Note added in proof:* - The crystal structure of 1-( $\beta$ -D-arabinofuranosyl)-4-thiouracil monohydrate has recently been reported (Saenger, 1972). This molecule is also in the *anti* conformation, the glycosidic torsion angle being 36.0°. The C(4')-C(5') conformation, is *gauche-gauche*; however, any possibility of a close O(5')...O(2') contact is precluded by the furanose ring being C(3')-*endo*.

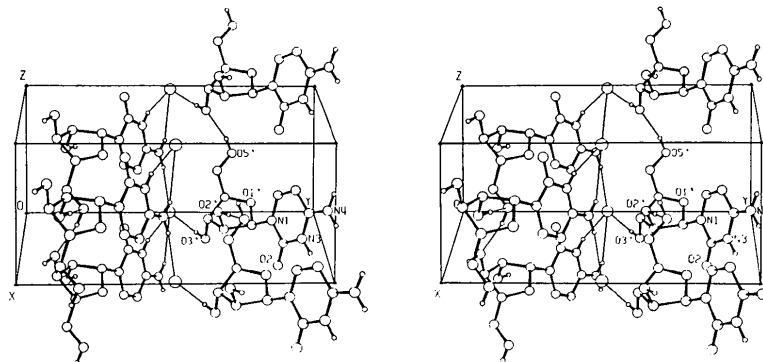


Fig. 5. A stereoscopic view (Johnson, 1965) showing the packing in 1-( $\beta$ -D-arabinofuranosyl)cytosine hydrochloride.

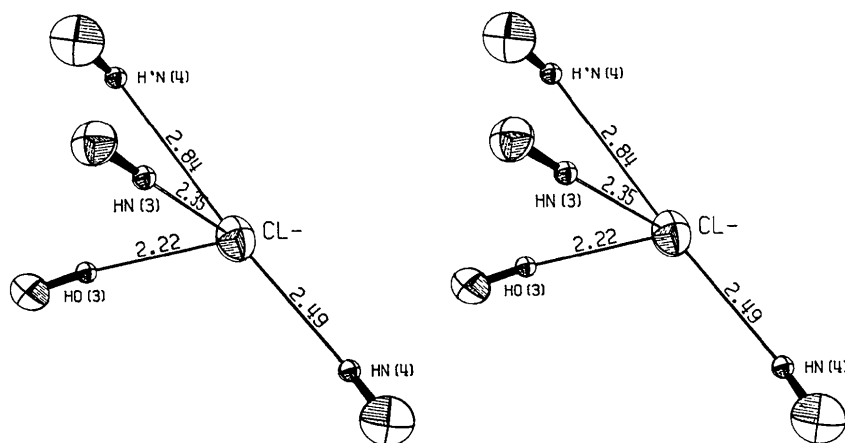


Fig. 6. A stereoscopic view (Johnson, 1965) showing the environment of the chloride ion. Thermal ellipsoids are drawn at the 50% probability level.

Table 10. *Interionic distances in various cytosine derivatives*

Letters refer to labeled distances in Fig. 7.

Distance	This study	MCHCl*†	2'-Deoxy*	Cyt- <i>m</i> *	Cyt- <i>o</i> *
<i>a</i>	2.84 Å	2.74 Å	3.05 Å		
<i>b</i>	3.568	3.49	3.689		
<i>c</i>	2.49	2.39	2.62		
<i>d</i>	3.354	3.30	3.415		
<i>e</i>	2.35	2.24	2.22		
<i>f</i>	3.221	3.10	3.131		
<i>g</i>	2.66	2.58	2.67	2.51 Å	2.84 Å
<i>h</i>	3.209	3.16	3.376	3.128	3.423
<i>i</i>	3.241	3.20	3.492	3.153	3.469
<i>j</i>	2.66	2.64	2.94	2.60	2.87
<i>k</i>	3.02	2.97	3.08		

\* MCHCl: 1-methylcytosine hydrochloride (Trus & Marsh, 1972). 2'-Deoxy: 2'-deoxycytidine hydrochloride (Subramanian & Hunt, 1970). Cyt-*m*: cytidylic acid *b* (monoclinic; Bugg & Marsh, 1967). Cyt-*o*: cytidylic acid *b* (orthorhombic; Sundaralingam & Jensen, 1965).

† Distances for the 1-methylcytosine hydrochloride structure are averages for the two independent molecules in the crystal.

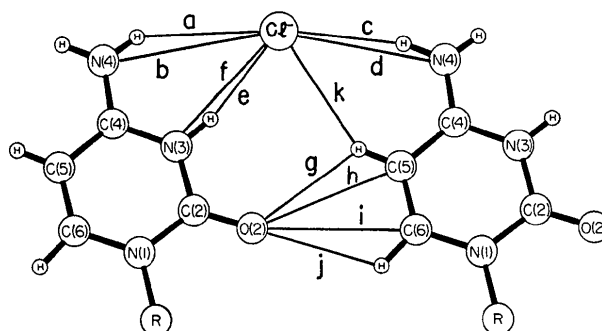


Fig. 7. The arrangement of cytosine rings and chloride ions. Values for the lettered distances, as found in crystals of several cytosine derivatives, are given in Table 10.

We thank Dr Sten Samson for excellent equipment, Miss Lillian Casler for many of the drawings, and Miss Allison Kimball for the typescript.

#### References

- BRYAN, R. F. & TOMITA, K. (1962). *Acta Cryst.* **15**, 1174–1178.
- BUGG, C. E. & MARSH, R. E. (1967). *J. Mol. Biol.* **25**, 67–82.
- CHU, M. Y. & FISHER, G. A. (1962). *Biochem. Pharmacol.* **11**, 423–430.
- CRUICKSHANK, D. W. J. & McDONALD, W. S. (1967). *Acta Cryst.* **23**, 9–11.
- DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363–371.
- EVANS, J. S., MUSSER, E. A., MENGEL, G. D., FORSLAD, K. R. & HUNTER, J. H. (1961). *Proc. Soc. Exp. Biol. Med.* **106**, 350–353.
- International Tables for X-ray Crystallography (1962). Vol. III. Birmingham: Kynoch Press.
- JOHNSON, C. K. (1965). *ORTEP*. ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KLYNE, W. & PRELOG, V. (1960). *Experientia*, **16**, 521–523.
- LARSON, A. C. (1967). *Acta Cryst.* **23**, 664–665.
- MOMPARLER, R. L. (1969). *Biochem. Biophys. Res. Commun.* **34**(4), 465–471.
- PAULING, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca: Cornell Univ. Press.
- SAENGER, M. (1972). *J. Amer. Chem. Soc.* **94**, 621–626.
- SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067–1077.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.
- SUBRAMANIAN, E. & HUNT, D. J. (1970). *Acta Cryst.* **B26**, 303–311.
- SUNDARALINGAM, M. & JENSEN, L. H. (1965a). *J. Mol. Biol.* **13**, 914–929.
- SUNDARALINGAM, M. & JENSEN, L. H. (1965b). *J. Mol. Biol.* **13**, 930–943.
- TOUGARD, P. (1969). *Biochem. Biophys. Res. Commun.* **37**(6), 961–964.
- TRUS, B. L. & MARSH, R. E. (1972). *Acta Cryst.* **B28**, 1834–1840.
- WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151–152.