

metastable. So far we can only say that no transition point has been found.

References

- ALLMANN, R. (1967). *Acta Cryst.* **22**, 432–433.
 BERG, J. M. VAN DEN & RUTTEN-KEULEMANS, E. W. M. (1963). *An Algol Program for the Calculation of Cell Dimensions from Angles*.
 DOYLE, P. A. & TURNER, P. S. (1968). *Acta Cryst.* **A24**, 390–397.
 KRYGER, L., RASMUSSEN, S. E. & DANIELSEN, J. (1972). *Acta Chem. Scand.* **26**, 2339–2348.
 HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502–510.
International Tables for X-ray Crystallography (1962). Vol. III. Birmingham: Kynoch Press.
 MOORE, F. H. (1963). *Acta Cryst.* **16**, 1169–1175.
 RAMACHANDRAN, G. N. & SRINIVASAN, R. (1959). *Acta Cryst.* **12**, 410–411.
 SCHOUWSTRA, Y. (1972). *Acta Cryst.* **B28**, 2217–2221.
 SCHOUWSTRA, Y. (1973). *Acta Cryst.* **B29**, 1–4.
 WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151–152.

Acta Cryst. (1973). **B29**, 1641

The Crystal and Molecular Structure of Uracil- β -D-arabinofuranoside

BY P. TOLLIN, H. R. WILSON AND D. W. YOUNG

Carnegie Laboratory of Physics, University of Dundee, Dundee, Scotland

(Received 20 March 1973; accepted 30 March 1973)

The crystal structure of uracil- β -D-arabinofuranoside ($C_9H_{12}N_2O_6$) has been determined from data collected on a Hilger and Watts linear diffractometer. The crystals are orthorhombic, space group $P2_12_12_1$, with cell dimensions: $a = 6.810$ (5), $b = 6.870$ (5), $c = 20.98$ (1) Å. The structure was solved by Patterson interpretation methods. The final R for 1164 observed reflexions is 0.058. The glycosidic torsion angle defined for the sequence of atoms O(1')–C(1')–N(1)–C(6) is 34.0° . The sugar ring pucker is C(2') *endo* and, relative to the least-squares plane through the five-atom sugar ring, it is C(2') *endo*–C(1') *exo*. The orientation of the C(5')–O(5') bond is *gauche* to both the C(4')–O(1') and C(4')–C(3') bonds. There is an intramolecular hydrogen bond between atoms O(2') and O(5'), in which O(2') is the donor.

Introduction

The structure determination of uracil- β -D-arabinofuranoside (ara U) was undertaken as part of a series of structure determinations of nucleosides and nucleotides. Ara U was first isolated from sponges of the species *Cryptotethya crypta* (Bergmann & Feeney, 1950, 1951), and for this reason is also referred to as spongouridine. We were particularly interested in comparing the conformation of ara U with those of related ribose and deoxyribose nucleosides. A preliminary account of the conformational parameters of ara U has been given recently (Tollin, Wilson & Young, 1973).

Experimental

Crystals of ara U ($C_9H_{12}N_2O_6$) (Fig. 1) were obtained by evaporation from aqueous solutions. The unit-cell dimensions, obtained from Weissenberg photographs with Cu $K\alpha$ radiation, are

$$a = 6.810 \text{ (5)}, \quad b = 6.870 \text{ (5)}, \quad c = 20.98 \text{ (1) \AA}$$

$$(\lambda \text{ Cu } K\alpha_1 = 1.54050 \text{ \AA}, \quad \lambda \text{ Cu } K\alpha_2 = 1.54434 \text{ \AA})$$

and the space group is $P2_12_12_1$. The calculated density, with one molecule per asymmetric unit, is 1.652 g cm^{-3} .

The intensities of 1165 unique reflexions were measured on a Hilger and Watts linear diffractometer with Mo K radiation and balanced filters. The data were collected to a $\sin\theta$ value corresponding to the radius of the limiting sphere for Cu $K\alpha$ radiation. Two crystals were used, one mounted along **a** and the other along **b**. Both were about 0.5 mm long and 0.15×0.15 mm in cross section. No absorption corrections were applied.

Structure determination

The structure amplitudes were sharpened and used to calculate the $I(\theta\phi)$ function (Tollin & Cochran, 1964) in order to determine the orientation of the plane of the base. This function showed two large peaks of comparable height, one at $(90^\circ, 0^\circ)$ of height 154 on an arbitrary scale, the other at $(90^\circ, 90^\circ)$ of height 138. In order to distinguish between these two possibilities, and to determine the azimuthal angle which defines the orientation of the base in its plane, the function $I(\theta_1\theta_2\theta_3)$ was calculated (Munns, 1971; Tollin & Munns, to be published). This function rotates a model of the base in its plane until the best fit with the Patterson function is obtained. The peak at $(90^\circ, 90^\circ)$ was shown to be the correct one, the height of the peak in the $I(\theta_1\theta_2\theta_3)$ function being 364. The largest peak in

the function for the direction ($90^\circ, 0^\circ$) was only 224. Once the orientation of the base had been determined its position in the cell was found from the Q -functions (Tollin, 1966). Since the space group is $P2_12_12_1$ there are three possible Q functions $Q(x_0y_0)$, $Q(y_0z_0)$ and $Q(x_0z_0)$, not mutually independent, defining the position $(x_0y_0z_0)$ of the base. Because of the special orientation of the base in the (010) plane the Q -functions showed extra symmetry and the largest peaks from the three maps were inconsistent. However by adding the three maps together the largest peak was found at $x = \frac{1}{3}$, $y = 0$, $z = 0$. This corresponded to the largest peak in $Q(y_0z_0)$ and $Q(x_0y_0)$ and one of the larger peaks in $Q(x_0z_0)$. The largest peak in this last function had a height of 57 while the correct peak was seventh in order of magnitude and of height 50. A more detailed discussion of ways of combining Q -functions will be given elsewhere (Tollin & Munns, to be published). Structure factors calculated using the base alone gave an R of 0.43. The calculated phases were used to obtain an electron density map from which the positions of the remaining atoms were found.

Refinement

The structure was refined by the method of least squares with the block-diagonal approximation. A program written by Professor J. Trotter, modified for use on an Elliott 4130 computer, was used.

With unit weights for all the observed structure factors the non-hydrogen atom coordinates and individual isotropic temperature factors were refined until R had dropped to 0.12. It had been noted that two low-order reflexions, 020 and 110, calculated consistently too large (the 020 reflexion particularly so) and it was decided to remove this structure amplitude from further cycles of refinement. The observed value of the 020 structure factor was 133 but the calculated value was 188. Since the 020 reflexion makes a full contribution to the peak at $(90^\circ, 90^\circ)$ in the $I(\theta\phi)$ function such a reduction in the 020 reflexion can explain why the peak at $(90^\circ, 90^\circ)$ is lower than that at $(90^\circ, 0^\circ)$. Increase of the $|F|^2$ contribution from the 020 reflexion by a factor of two makes the peak at $(90^\circ, 90^\circ)$ the largest in the $I(\theta\phi)$ function.

The locations of the twelve hydrogen atoms were calculated with normal bond lengths and angles and with the possible hydrogen bond scheme suggested by close contacts of non-hydrogen atoms. With fixed hydrogen atom locations and isotropic temperature factors of 3.0 \AA^2 , the positional and isotropic thermal parameters of the non-hydrogen atoms were further refined until R was 0.095. The weighting scheme used at this stage of the refinement was

$$w = 1/[1 + \{(|F_o| - a)/b\}^2]$$

where a and b are constants chosen at this stage as 12.0 and 16.0 respectively.

The least-squares refinement was continued with

individual anisotropic temperature factors of the form

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

for the non-hydrogen atoms. The same weighting scheme was retained. After two such cycles of refinement R had dropped to 0.064. Twelve sets of structure factor calculations were performed in each of which one of the hydrogen atoms was omitted and for each a difference Fourier synthesis was computed in the region of the omitted hydrogen atom. Fig. 4 is a composite of these difference electron density calculations. Each hydrogen atom was clearly visible as a maximum with a peak height between 0.4 and 0.7 $e \text{ \AA}^{-3}$ and in a position very close to that expected.

Two further cycles of least-squares refinement were performed in which the coordinates and isotropic temperature factors of the hydrogen atoms were also allowed to vary. In the final stages the root mean square positional parameter shift was 0.15σ and the root mean square shift in the thermal parameters was 0.42σ where σ is the estimated standard deviation.

The final values of the parameters of the heavy atoms with their estimated standard deviations are given in Tables 1 and 3, and the parameters of the hydrogen atoms in Table 2. The observed and calculated structure factors are listed in Table 4. The final R for 1164 non-zero observed reflexions used in the anisotropic least-squares refinement was 0.058. R for all reflexions inside the copper $K\alpha$ limiting sphere, including the unobservable reflexions which were given

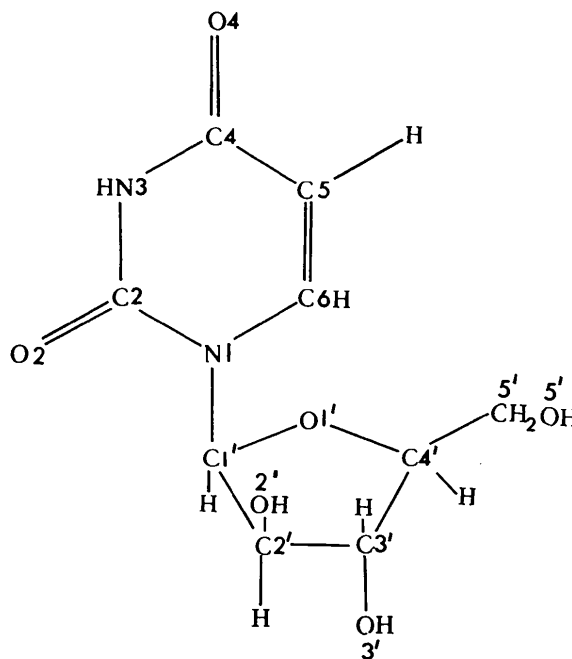


Fig. 1. The ara U ($C_9H_{12}N_2O_6$) molecule and the atom numbering.

an amplitude of 0.7 of the significant structure amplitude at the appropriate region of reciprocal space, was 0.077.

Table 1. *Fractional coordinates of the non-hydrogen atoms, with estimated standard deviations*

	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>
N(1)	0.3376 (5)	-0.0110 (6)	0.2953 (2)
C(2)	0.5143 (6)	-0.0016 (7)	0.2645 (2)
O(2)	0.6724 (5)	-0.0093 (6)	0.2914 (1)
N(3)	0.5029 (5)	0.0131 (6)	0.1993 (2)
C(4)	0.3338 (7)	0.0365 (7)	0.1627 (2)
O(4)	0.3482 (5)	0.0537 (5)	0.1051 (1)
C(5)	0.1564 (7)	0.0348 (7)	0.1990 (2)
C(6)	0.1628 (6)	0.0087 (7)	0.2624 (2)
C(1')	0.3402 (6)	-0.0327 (7)	0.3650 (2)
O(1')	0.1717 (5)	-0.1385 (5)	0.3817 (1)
C(2')	0.3269 (6)	0.1602 (7)	0.4018 (2)
O(2')	0.2054 (5)	0.2969 (5)	0.3706 (2)
C(3')	0.2434 (6)	0.0887 (7)	0.4647 (2)
O(3')	0.4044 (5)	0.0188 (5)	0.5011 (1)
C(4')	0.1057 (7)	-0.0778 (6)	0.4450 (2)
C(5')	-0.1081 (7)	-0.0300 (7)	0.4421 (2)
O(5')	-0.1459 (5)	0.1318 (5)	0.4007 (3)

Table 2. *Positional parameters of the hydrogen atoms, with estimated standard deviations*

	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>
H(1)	0.6186 (8)	-0.0143 (8)	0.1764 (2)
H(2)	0.0303 (9)	0.0481 (10)	0.1782 (3)
H(3)	0.0353 (10)	0.0157 (10)	0.2866 (3)
H(4)	0.4636 (8)	-0.1082 (8)	0.3782 (2)
H(5)	0.0940 (10)	0.2813 (9)	0.3749 (3)
H(6)	0.4597 (9)	0.2156 (9)	0.4075 (3)
H(7)	0.3678 (8)	-0.0121 (8)	0.5384 (2)
H(8)	0.1633 (8)	0.1880 (8)	0.4897 (2)
H(9)	0.1288 (9)	-0.1831 (9)	0.4675 (3)
H(10)	-0.1980 (11)	0.0880 (11)	0.3710 (3)
H(11)	-0.1531 (7)	0.0111 (8)	0.4831 (2)
H(12)	-0.1922 (8)	-0.1525 (8)	0.4290 (2)

Discussion

The pyrimidine base

The pyrimidine base is essentially planar, the maximum deviation of a ring from the least-squares plane

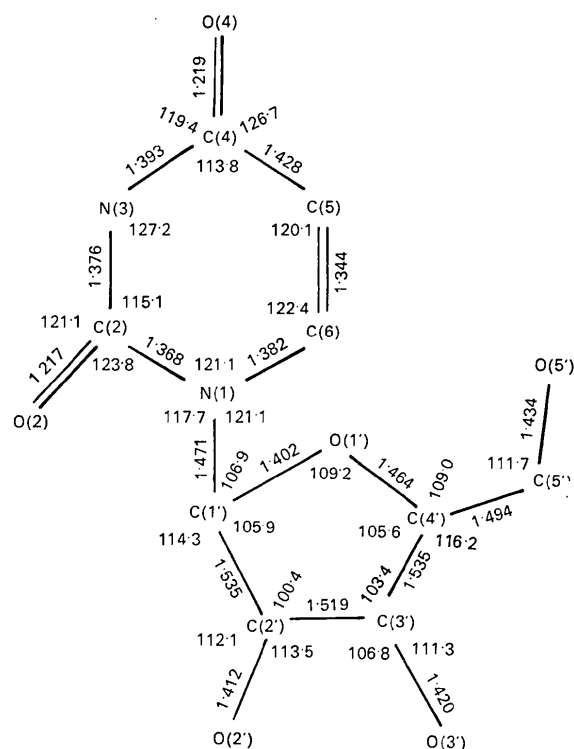


Fig. 2. Bond lengths and angles.

being 0.031 Å (Table 6). The equation of the least-squares plane through the six-membered ring is

$$0.0446x + 0.9928y + 0.1110z + 0.7300 = 0$$

where *x*, *y* and *z* are in Å.

Atom C(1') of the sugar ring and atom O(4) are coplanar with the six-membered ring, but atom O(2) is displaced 0.089 Å from the plane.

The pyrimidine is in the diketo form and the bond lengths and angles, which are shown in Fig. 2 and listed in Table 5, are similar to those in related ribose and deoxyribose nucleosides.

Table 3. *Thermal parameters of the non-hydrogen atoms with their estimated standard deviations ($\times 10^5$)*

	B_{11}	B_{22}	B_{33}	B_{12}	B_{23}	B_{13}
N(1)	798 (63)	1562 (81)	77 (6)	430 (166)	-13 (42)	-8 (38)
C(2)	909 (76)	1157 (85)	92 (8)	459 (185)	-113 (49)	-29 (48)
O(2)	880 (58)	2576 (91)	106 (6)	179 (167)	-67 (43)	-93 (35)
N(3)	884 (65)	1540 (82)	74 (6)	307 (160)	-83 (43)	52 (39)
C(4)	1148 (82)	1116 (86)	103 (7)	375 (175)	-69 (46)	-158 (51)
O(4)	1729 (74)	1977 (79)	66 (5)	94 (158)	-36 (37)	-115 (39)
C(5)	866 (81)	1487 (98)	131 (9)	446 (183)	-127 (55)	-105 (53)
C(6)	683 (71)	1647 (99)	135 (9)	84 (196)	-79 (56)	-66 (48)
C(1')	932 (74)	1297 (90)	89 (7)	369 (169)	-13 (48)	27 (51)
O(1')	1360 (63)	1250 (62)	91 (5)	-726 (132)	-140 (33)	183 (36)
C(2')	771 (76)	1124 (84)	80 (7)	-87 (169)	19 (46)	-12 (49)
O(2')	977 (61)	1168 (61)	142 (6)	119 (115)	171 (37)	35 (37)
C(3')	706 (75)	1112 (88)	104 (8)	62 (155)	1 (47)	79 (46)
O(3')	1042 (57)	1930 (78)	76 (5)	13 (141)	126 (40)	-93 (35)
C(4')	1171 (87)	868 (80)	87 (7)	-226 (165)	76 (47)	6 (49)
C(5')	1014 (86)	1466 (105)	132 (9)	-532 (184)	125 (58)	-6 (52)
O(5')	1195 (68)	1639 (74)	181 (8)	-99 (147)	117 (44)	-118 (43)

Table 4. Observed and calculated structure factors

M	K	L	F _o B	F _c AL	M	K	L	F _o B	F _c AL	M	K	L	F _o B	F _c AL	M	K	L	F _o B	F _c AL	M	K	L	F _o B	F _c AL	
0	2	2	25.07	25.91	0	7	7	6.94	6.18	1	5	12	18.15	18.04	2	4	1	14.79	14.29	2	23	10	24	0.98	0.4
0	4	4	38.76	39.82	0	7	7	17.58	16.98	2	4	2	38.27	38.82	3	2	4	15.23	15.26	3	2	24	0.98	0.4	
0	0	6	27.32	27.64	0	7	7	4.98	4.74	2	4	2	4	3.12	3.12	3	2	4	11.75	11.12	3	2	4	1.25	1.12
0	0	8	18.42	18.19	0	7	7	11.55	11.29	1	5	13	15.15	14.94	2	4	2	15.98	15.83	3	2	4	1.25	1.12	
0	0	10	9.95	9.81	0	7	7	2.32	2.32	1	5	17	8.17	8.04	2	4	2	27.92	27.17	3	3	3	21.01	20.17	
0	0	12	14.22	14.22	0	7	7	6.89	6.89	1	5	19	10.40	10.35	2	4	2	10.40	10.35	3	3	3	10.40	10.35	
0	0	14	12.08	12.00	0	7	7	8.09	8.09	1	5	20	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	16	20.76	21.47	0	7	7	14.74	14.74	1	5	21	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	18	6.03	5.95	0	7	7	0.81	0.81	1	5	22	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	20	6.44	6.44	0	7	7	6.89	6.89	1	5	23	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	22	12.08	12.00	0	7	7	8.09	8.09	1	5	24	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	24	20.76	21.47	0	7	7	14.74	14.74	1	5	25	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	26	6.03	5.95	0	7	7	0.81	0.81	1	5	26	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	28	6.44	6.44	0	7	7	6.89	6.89	1	5	27	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	30	12.08	12.00	0	7	7	8.09	8.09	1	5	28	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	32	20.76	21.47	0	7	7	14.74	14.74	1	5	29	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	34	6.03	5.95	0	7	7	0.81	0.81	1	5	30	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	36	6.44	6.44	0	7	7	6.89	6.89	1	5	31	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	38	12.08	12.00	0	7	7	8.09	8.09	1	5	32	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	40	20.76	21.47	0	7	7	14.74	14.74	1	5	33	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	42	6.03	5.95	0	7	7	0.81	0.81	1	5	34	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	44	6.44	6.44	0	7	7	6.89	6.89	1	5	35	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	46	12.08	12.00	0	7	7	8.09	8.09	1	5	36	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	48	20.76	21.47	0	7	7	14.74	14.74	1	5	37	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	50	6.03	5.95	0	7	7	0.81	0.81	1	5	38	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	52	6.44	6.44	0	7	7	6.89	6.89	1	5	39	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	54	12.08	12.00	0	7	7	8.09	8.09	1	5	40	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	56	20.76	21.47	0	7	7	14.74	14.74	1	5	41	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	58	6.03	5.95	0	7	7	0.81	0.81	1	5	42	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	60	6.44	6.44	0	7	7	6.89	6.89	1	5	43	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	62	12.08	12.00	0	7	7	8.09	8.09	1	5	44	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	64	20.76	21.47	0	7	7	14.74	14.74	1	5	45	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	66	6.03	5.95	0	7	7	0.81	0.81	1	5	46	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	68	6.44	6.44	0	7	7	6.89	6.89	1	5	47	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	70	12.08	12.00	0	7	7	8.09	8.09	1	5	48	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	72	20.76	21.47	0	7	7	14.74	14.74	1	5	49	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	74	6.03	5.95	0	7	7	0.81	0.81	1	5	50	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	76	6.44	6.44	0	7	7	6.89	6.89	1	5	51	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	78	12.08	12.00	0	7	7	8.09	8.09	1	5	52	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	80	20.76	21.47	0	7	7	14.74	14.74	1	5	53	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	82	6.03	5.95	0	7	7	0.81	0.81	1	5	54	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	84	6.44	6.44	0	7	7	6.89	6.89	1	5	55	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	86	12.08	12.00	0	7	7	8.09	8.09	1	5	56	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	88	20.76	21.47	0	7	7	14.74	14.74	1	5	57	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	90	6.03	5.95	0	7	7	0.81	0.81	1	5	58	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	92	6.44	6.44	0	7	7	6.89	6.89	1	5	59	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	94	12.08	12.00	0	7	7	8.09	8.09	1	5	60	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	96	20.76	21.47	0	7	7	14.74	14.74	1	5	61	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	98	6.03	5.95	0	7	7	0.81	0.81	1	5	62	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	100	6.44	6.44	0	7	7	6.89	6.89	1	5	63	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	102	12.08	12.00	0	7	7	8.09	8.09	1	5	64	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	104	20.76	21.47	0	7	7	14.74	14.74	1	5	65	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	106	6.03	5.95	0	7	7	0.81	0.81	1	5	66	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	108	6.44	6.44	0	7	7	6.89	6.89	1	5	67	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	110	12.08	12.00	0	7	7	8.09	8.09	1	5	68	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	112	20.76	21.47	0	7	7	14.74	14.74	1	5	69	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	114	6.03	5.95	0	7	7	0.81	0.81	1	5	70	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	116	6.44	6.44	0	7	7	6.89	6.89	1	5	71	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	118	12.08	12.00	0	7	7	8.09	8.09	1	5	72	10.32	10.32	2	4	2	10.32	10.32	3	3	3	10.32	10.32	
0	0	120	20.76	21.47	0	7	7	14.74	14.74	1	5	73	16.46	16.46	2	4	2	16.46	16.46	3	3	3	16.46	16.46	
0	0	122	6.03	5.95	0	7	7	0.81	0.81	1	5	74	7.96	7.96	2	4	2	7.96	7.96	3	3	3	7.96	7.96	
0	0	124	6.44	6.44	0	7	7	6.89	6.89	1	5	75	9.52	9.52	2	4	2	9.52	9.52	3	3	3	9.52	9.52	
0	0	126	12.08	12.00	0	7	7	8.09	8.09	1	5	76													

Table 5. Bond lengths and angles

The mean e.s.d. for bond lengths is 0.006 Å and for angles 0.35°.

Bond lengths(Å)		Bond angles(°)	
N(1)-C(2)	1.368	N(1)-C(2)-O(2)	123.8
C(2)-O(2)	1.217	C(2)-N(1)-C(6)	121.1
C(2)-N(3)	1.376	C(2)-N(1)-C(1')	117.7
N(3)-C(4)	1.393	N(1)-C(2)-N(3)	115.1
C(4)-O(4)	1.219	C(2)-N(3)-C(4)	127.2
C(4)-C(5)	1.428	N(3)-C(4)-O(4)	119.4
C(5)-C(6)	1.344	N(3)-C(4)-C(5)	113.8
C(6)-N(1)	1.382	O(4)-C(4)-C(5)	126.7
N(1)-C(1')	1.471	C(4)-C(5)-C(6)	120.1
C(1')-O(1')	1.402	C(5)-C(6)-N(1)	122.4
O(1')-C(4')	1.464	C(6)-N(1)-C(1')	121.1
C(1')-C(2')	1.535	N(1)-C(1')-O(1')	106.9
C(2')-C(3')	1.519	N(1)-C(1')-C(2')	114.3
C(3')-O(3')	1.420	O(1')-C(1')-C(2')	105.9
C(2')-O(2')	1.412	C(1')-C(2')-C(3')	100.4
C(3')-C(4')	1.535	C(2')-C(3')-C(4')	103.6
C(4')-C(5')	1.494	C(2')-C(3')-O(3')	106.8
C(5')-O(5')	1.434	O(3')-C(3')-C(4')	111.3
		C(1')-O(1')-C(4')	109.2
		C(3')-C(4')-C(5')	116.2
		C(3')-C(4')-O(1')	105.6
		O(1')-C(4')-C(5')	109.0
		C(4')-C(5')-O(5')	111.7
		C(1')-C(2')-O(2')	112.1
		O(2')-C(2')-C(3')	113.5

Table 6. Deviations (in Å) of atoms from the least-squares plane of the pyrimidine ring

N(1)	-0.015
C(2)	0.031
O(2)	0.089
N(3)	-0.024
C(4)	-0.001
O(4)	-0.014
C(5)	0.018
C(6)	-0.010
C(1')	0.001

The sugar ring

The most common type of pucker of sugar rings in nucleosides and nucleotides is with either atom C(2')

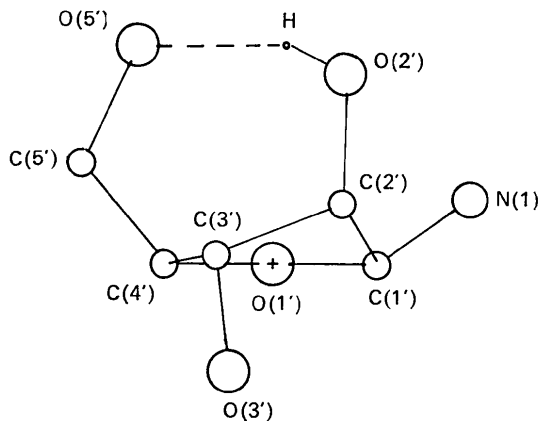


Fig. 3. A view of the sugar ring along the plane containing atoms C(1'), O(1') and C(4').

or atom C(3') displaced from the plane of the other four atoms on the same side as atom C(5'), when the pucker is called C(2') or C(3') *endo*, respectively. In ara U the pucker is C(2') *endo* (Fig. 3), and the displacement is 0.591 Å. The pucker thus differs from that observed in 1-β-D-arabinofuranosyl-4-thiouracil where the pucker is C(3') *endo* (Saenger, 1972). Described relative to the least-squares plane through the five ring atoms the pucker is C(2') *endo*-C(1') *exo* and may be written as ²T₁, where T indicates the twist conformation of the ring. In this respect the sugar pucker is similar to that in the related deoxynucleosides bromo-deoxyuridine (Iball, Morgan & Wilson, 1966), iodo-deoxyuridine (Camerman & Trotter, 1965) and chloro-deoxyuridine (Young & Morris, 1973). The equation of the five-atom plane in ara U is

$$0.7264x - 0.5351y + 0.4314z + 4.9006 = 0;$$

the deviations of atoms from this plane are listed in Table 7.

Table 7. Deviations (in Å) of atoms from the least-squares plane of the sugar-ring atoms

C(1')	0.206
O(1')	-0.088
C(2')	-0.236
C(3')	0.183
C(4')	-0.064
C(5')	-1.324
O(2')	-1.622
O(3')	1.566
O(5')	-2.480

The torsion angles around the sugar ring are listed in Table 8 and are similar to those in other nucleosides with similar puckering of the sugar ring. The bond lengths and angles are also similar to those in other nucleosides and are listed in Table 5.

Table 8. The torsion angles around the sugar ring

O(1')-C(1')	-27.5°
C(1')-C(2')	+38.1
C(2')-C(3')	-34.0
C(3')-C(4')	+19.4
C(4')-O(1')	+5.0

The conformation of the C(5')-O(5') bond relative to the furanose ring is *gauche* to both the C(4')-O(1') and C(4')-C(3') bonds. This *gauche-gauche* conformation is that most commonly found in nucleosides and nucleotides. In ara U the torsion angle for the sequence of atoms O(5') C(5') C(4') C(3') is 55.8°.

Conformation of the molecule

The dihedral angle between the mean planes of the pyrimidine base and the sugar ring is 63.2°. The conformation of the sugar ring relative to the base can be described in terms of a torsion angle about the glyco-

sic bond (Donohue & Trueblood, 1960). This angle, χ , when defined for the sequence of atoms O(1') C(1') N(1) C(6) is 34.0° and the nucleoside is in the *anti* conformation. A comparison of the main characteristics with those of related ribose and deoxyribose nucleosides is made in Table 9.

Molecular packing and hydrogen bonding

The arrangement of the molecules in the unit-cell is shown in Fig. 5. They are arranged with the pyrimidine bases approximately perpendicular to **b** and the bases are stacked about 3.4 Å apart, with a similar overlap of bases to that observed in related structures (Bugg, Thomas, Sundaralingam & Rao, 1971).

The hydrogen bonding scheme is also indicated in Fig. 5. There is a hydrogen bond between O(2) and O(5') of the molecule at $(1+x, y, z)$ in which O(5') is donating and where the O(5')-O(2) distance is 2.781 Å. O(4) is hydrogen bonded to O(3') of the molecule at $(\frac{1}{2}-x, -y, -\frac{1}{2}+z)$; O(3') is donating and the O(3')-O(4) distance is 2.822 Å. N(3) is hydrogen bonded to O(2') of the molecule at $(1-x, \frac{1}{2}+y, \frac{1}{2}-z)$; N(3) is donating and the N(3)-O(2') distance is 2.881 Å. These are the only intermolecular hydrogen bonds. However, there is an intramolecular hydrogen bond between O(2') and O(5') with O(2') donating, and where the O(2')-O(5') distance is 2.723 Å. This intramolecular bond is made possible because of the C(2') *endo* pucker of the sugar, and would not be possible if the pucker were C(3') *endo*. The formation of this intramolecular hydrogen bond in the crystalline state probably ac-

counts for the difference between this conformation and that predominating in solution. Proton magnetic resonance studies of ara U in solution (Hruska, 1972) indicate that there is a destabilization of the *gauche-gauche* conformation about the C(4')-C(5') bond and a more marked C(3') *endo* pucker compared with uridine in solution. The present results, however, show that the interaction between the 2'-hydroxyl and the 5'-hydroxyl is one of attraction when the latter is in the *gauche-gauche* conformation, and hence situated over the sugar ring. It may be that the predominance of other conformations in solution is due to the availability of water molecules for hydrogen bonding.

We thank Mr J. Low for assistance with the data collection and Miss Ruth Brand for computational assistance. A Science Research Council grant for the

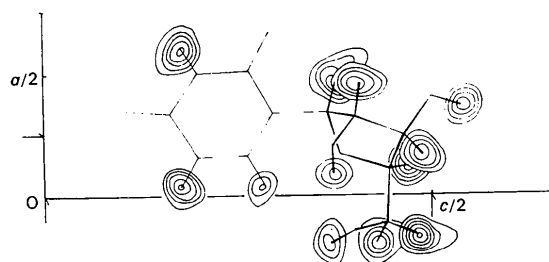


Fig. 4. A composite difference Fourier synthesis showing the hydrogen atoms. The view is along **b**, as in Fig. 5

Table 9. A comparison of the conformational parameters of ribose and deoxyribose nucleosides closely related to ara U

The angle χ_{CN} is defined for the sequence of atoms O(1') C(1') N(1) C(6).				
Deoxyribose nucleosides	χ_{CN}	Pucker	C(5')-O(5')	Reference
Deoxyuridine	28°			
Mol I	28	C(2') <i>endo</i>	<i>t-g</i>	Rahman & Wilson (1972)
Mol 2	24	C(2') <i>endo</i>	<i>t-g</i>	
Fluorodeoxyuridine	59	C(2') <i>endo</i>	<i>t-g</i>	Harris & Macintyre (1964)
Chlorodeoxyuridine	41.1	C(2') <i>endo</i>	<i>g-t</i>	Young & Morris (1973)
Bromodeoxyuridine	47	C(2') <i>endo</i>	<i>g-t</i>	Iball, Morgan & Wilson (1966)
Iododeoxyuridine	63	C(2') <i>endo</i>	<i>g-g</i>	Camerman & Trotter (1965)
Thymidine	39	C(3') <i>exo</i>	<i>g-t</i>	Yong, Tollin & Wilson (1969)
Ribose nucleosides				
Uridine	16.8	C(3') <i>endo</i>	<i>g-g</i>	Green, Shiono, Rosenstein & Abraham (1971)
Mol I				
Mol II	23.8	C(3') <i>endo</i>	<i>g-g</i>	
Chlorouridine	51.4	C(2') <i>endo</i>		Hawkinson & Coulter (1971)
Bromouridine	51	C(2') <i>endo</i>	<i>g-g</i>	Iball, Morgan & Wilson & (1966)
(+ DMSU)	62.2	C(2') <i>endo</i>	<i>g-g</i>	Iball, Morgan & Wilson (1968)
(+ adenosine)	20.0	C(3') <i>endo</i>	<i>g-g</i>	Haschemeyer & Sobell (1965)
Iodourine				
Mol I	11.8	C(3') <i>endo</i>	<i>g-g</i>	Rahman & Wilson (1970)
Mol II	55.8	C(2') <i>endo</i>	<i>t-g</i>	
Methyluridine	29.4	C(3') <i>endo</i>	<i>g-g</i>	Hunt & Subramanian (1969)
Arabinose nucleoside				
Uracil arabinofuranoside	34.0	C(2') <i>endo</i>	<i>g-g</i>	Present work

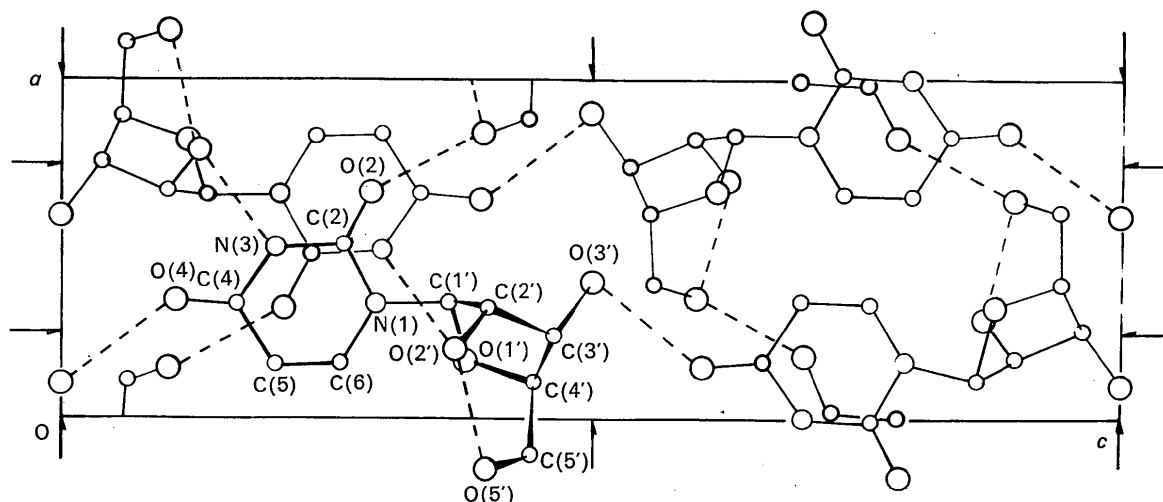


Fig. 5. A projection of the structure viewed along *b*. Hydrogen bonds are shown as broken lines.

purchase and maintenance of the diffractometer is gratefully acknowledged.

References

- BERGMANN, W. & FEENEY, R. J. (1950). *J. Amer. Chem. Soc.* **72**, 2809–2810.
- BERGMANN, W. & FEENEY, R. J. (1951). *J. Org. Chem.* **16**, 981–987.
- BUGG, C. E., THOMAS, J. M., RAO, S. T. & SUNDARALINGAM, M. (1971). *Biopolymers*, **10**, 175–219.
- CAMERMAN, N. & TROTTER, J. (1969). *Acta Cryst.* **18**, 203–211.
- DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363–371.
- GREEN, E. A., SHIONO, R., ROSENSTEIN, R. D. & ABRAHAM, D. K. (1971). *Chem. Commun.* pp. 53–54.
- HARRIS, D. R. & MACINTYRE, W. M. (1964). *Biophys. J.* **4**, 203–225.
- HASCHEMEYER, A. E. V. & SOBELL, H. M. (1965). *Acta Cryst.* **18**, 525–532.
- HAWKINSON, S. W. & COULTER, C. (1971). *Acta Cryst.* **B24**, 34–42.
- HRUSKA, F. E. (1972). In *Conformation of Biological Molecules and Polymers*. The Fifth Jerusalem Symposium (The Israel Academy of Sciences and Humanities).
- HUNT, D. J. & SUBRAMANIAN, E. (1969). *Acta Cryst.* **B25**, 2144–2152.
- IBALL, J., MORGAN, C. H. & WILSON, H. R. (1966). *Proc. Roy. Soc. A* **295**, 320–333.
- IBALL, J., MORGAN, C. H. & WILSON, H. R. (1968). *Proc. Roy. Soc. A* **302**, 225–236.
- MUNNS, A. R. I. (1971). Ph. D. Thesis. Univ. of Dundee.
- RAHMAN, A. & WILSON, H. R. (1970). *Acta Cryst.* **B26**, 1765–1775.
- RAHMAN, A. & WILSON, H. R. (1972). *Acta Cryst.* **B28**, 2260–2270.
- SAENGER, W. (1972). *J. Amer. Chem. Soc.* **94**, 621–626.
- TOLLIN, P. (1966). *Acta Cryst.* **21**, 613–614.
- TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1322–1324.
- TOLLIN, P., WILSON, H. R. & YOUNG, D. W. (1973). *Nature, Lond.* **242**, 49–50.
- YOUNG, D. W. & MORRIS, E. M. (1973). *Acta Cryst.* **B29**, 1259–1264.
- YOUNG, D. W., TOLLIN, P. & WILSON, H. R. (1969). *Acta Cryst.* **B25**, 1423–1432.