

The Crystal and Molecular Structure of Deoxyuridine

BY ASADUR RAHMAN AND H. R. WILSON

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

(Received 28 February 1972)

The crystal structure of 2'-deoxyuridine ($C_9H_{12}N_2O_5$) has been determined from data collected on a Hilger and Watts linear diffractometer. The crystals are monoclinic, space group $P2_1$, with cell dimensions $a=7.91 \pm 0.01$, $b=6.710 \pm 0.005$, $c=18.77 \pm 0.01$ Å and $\beta=96.6 \pm 0.1^\circ$. The structure was solved by Patterson interpretation methods. The final R value for 1611 observed reflexions is 0.084. There are two molecules in the asymmetric unit; in molecule I the dihedral angle between the base and the sugar is 53° , whilst in molecule II it is 47° . The glycosidic torsion angle φ_{CN} , is -26° in molecule I and -25° in molecule II. In molecule I, atom C(2') of the sugar ring is displaced 0.43 Å *endo* and atom C(3') 0.19 Å *exo* relative to the plane through atoms C(1'), O(1') and C(4'). In molecule II, atom C(2') is 0.32 Å *endo* and atom C(3') is 0.28 Å *exo*. The orientation of the C(5')-O(5') bond is *trans* to the C(4')-O(1') bond and *gauche* to C(4')-C(3') in both molecules.

Introduction

The structure determination of 2'-deoxyuridine (UdR) was undertaken as part of a series of structure determinations, being carried out in this laboratory, of nucleic acid components and their analogues. Although UdR is found in animal tissues it does not normally occur in DNA. An exception is the DNA of phage PBS2, a transducing bacteriophage for *Bacillus subtilis* (Takahashi, 1963; Takahashi & Marmur, 1963). The structure of this DNA appears to be the same as that of normal DNA (Langridge & Marmur, 1964). In ordinary cells UdR acts as a precursor for the thymidylic acid of DNA.

The structure of UdR is of interest for the comparison of its conformation with the conformations of related nucleosides, in particular, thymidine, which shows an unusual pucker of the sugar ring (Tollin, Wilson & Young, 1968; Young, Tollin & Wilson, 1969). A preliminary account of the conformational parameters of UdR has been given recently (Rahman & Wilson, 1971).

Experimental

Crystals of UdR ($C_9N_2O_5H_{12}$), whose chemical structure is shown in Fig. 1(a), were obtained from aqueous solutions. The crystals are elongated and they split very easily into thin plates. Most of the crystals gave diffraction patterns showing extensive streaking, and, after a survey of many crystals, two were selected as being suitable for data collection. The unit-cell dimensions were obtained from Weissenberg photographs using Cu $K\alpha$ radiation. The crystal system is monoclinic and the crystals are elongated along the b axis. The unit-cell dimensions are:

$$a=7.91 \pm 0.01; b=6.710 \pm 0.005; c=18.77 \pm 0.02 \text{ \AA}; \\ \beta=96.6 \pm 0.1^\circ, \\ \lambda \text{ Cu } K\alpha_1=1.54050 \text{ \AA}, \lambda \text{ Cu } K\alpha_2=1.54434 \text{ \AA}.$$

Systematic absences of the $0k0$ reflexions with k odd were observed. The space group therefore is either $P2_1$ or $P2_1/m$ but the optical activity of the nucleoside establishes the space group as $P2_1$. A calculation of the probable density of the crystals, based on the chemical formula, showed that there must be four molecules in the unit cell and hence two molecules in the asymmetric unit. The calculated density, assuming no molecules of water of crystallization, is 1.6106 g.cm^{-3} .

The intensities of 1611 significant reflexions, out of 2478 unique reflexions within the limiting sphere corresponding to the Cu $K\alpha$ radiations, were measured on a Hilger and Watts diffractometer using Mo $K\alpha$ radiation ($\lambda=0.71069$ Å). Two crystals were used for data collection, one mounted along the b axis and the other along the c axis. The one mounted along the b axis had a length of 0.50 mm and cross-section of 0.16×0.08 mm. The crystal mounted along the c axis had a length of 0.12 mm and cross-section of 0.14×0.06 mm. Four cycles of intensity measurements were made using balanced filters. The absorption coefficient for Mo $K\alpha$ radiation is 1.068 cm^{-1} . No absorption corrections were applied.

Structure determination

The orientations of the two pyrimidine bases in the asymmetric unit were determined by evaluating the $I(\theta, \varphi)$ function of Tollin & Cochran (1964) for θ and φ values ranging from 0 to 90° and 0 to 180° respectively. Sharpened Patterson coefficients and a value of 5 Å for the radius of the disc were used. The $I(\theta, \varphi)$ map, plotted on a Sanson-Flamsteed sinusoidal equal area projection, is shown in Fig. 2. There is only one significant peak in the map, at $\theta=88$ and $\varphi=0^\circ$ (θ and φ are spherical polar angles with respect to an orthogonal sets of axes defined by a^* , b , c), suggesting that the two bases are parallel or nearly parallel.

In order to find the orientations of the pyrimidine bases within these planes the sum rotation function $R(\theta_1, \theta_2, \theta_3)$ of Tollin & Cochran (1964) was computed using a program written by Dr A. R. I. Munns. This function was computed for the line θ_1 ranging from 0 to 180°, keeping θ_2 and θ_3 fixed. With the pyrimidine model chosen to lie in the a - b plane, $\theta_2 = -88^\circ$ and $\theta_3 = 90^\circ$. However, because of the approximate sixfold

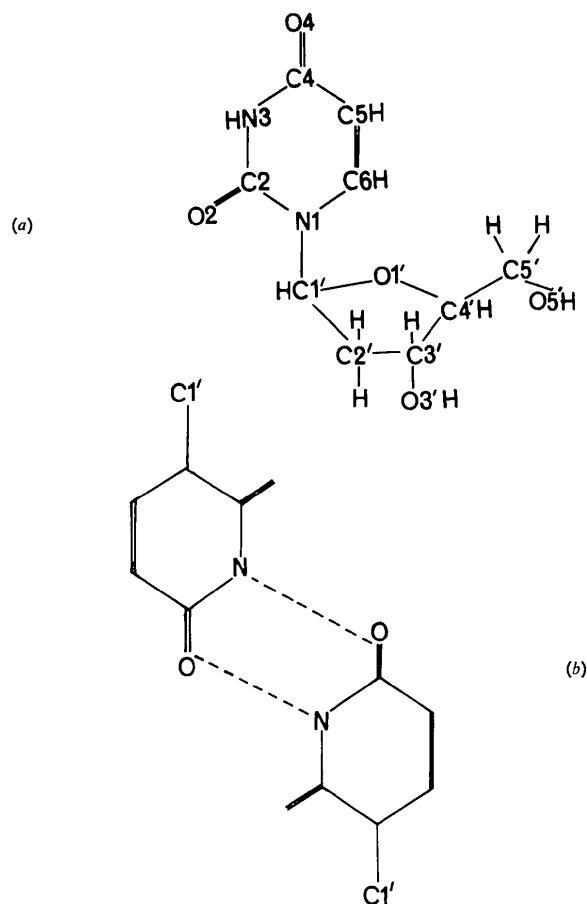


Fig. 1. (a) The 2'-deoxyuridine ($C_9H_{12}N_2O_5$) molecule and the atom numbering. (b) Model of a dimer pair of bases.

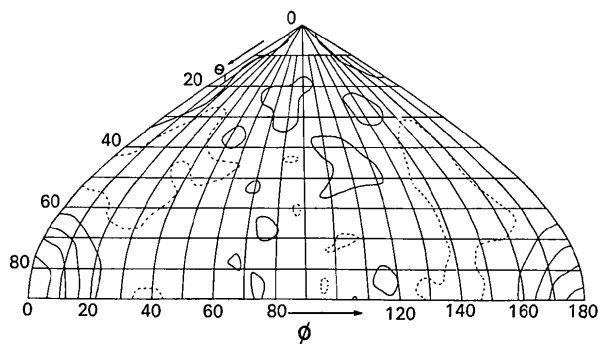


Fig. 2. The function $I(\theta, \phi)$ with $R = 5.0 \text{ \AA}$.

symmetry of the pyrimidine base used as a model, the orientations could not be determined unambiguously. In an attempt to resolve this ambiguity a sugar component was added to the model to be rotated. Normally sugar rings in nucleosides and nucleotides are puckered with either $C(3')$ *endo* or $C(2')$ *endo*, *i.e.* with either $C(3')$ or $C(2')$ displaced from the mean plane of the other four ring atoms and on the same side as $C(5')$. The angular rotation about the glycosidic bond is represented by φ_{CN} (Donohue & Trueblood, 1960) and it is normally found that φ_{CN} is significantly smaller for a $C(3')$ *endo* sugar than for a $C(2')$ *endo* sugar (Sundaralingam, 1969; Arnott & Hukins, 1969; Wilson, 1970). In the models to be rotated, a $C(3')$ *endo* sugar with $\varphi_{CN} = -13^\circ$ was used and the $C(2')$ *endo* sugar had $\varphi_{CN} = -58^\circ$. Each nucleoside model was used to calculate a line rotation for θ_1 . Also, as the $I(\theta, \varphi)$ map suggested that the two bases are parallel or nearly parallel, the possibility that they are coplanar and hydrogen bonded to form dimers was considered, and a similar line rotation was calculated for a model dimer of bases as illustrated in Fig. 1(b). The results of the line rotations for the $C(3')$ *endo* and $C(2')$ *endo* models indicate that the $C(2')$ *endo* model is preferable and shows the highest peak at $\theta_1 = 182.5^\circ$. This corresponds to a direction of the glycosidic bond approximately parallel to the z axis. Similarly, the line rotation for the model of dimer-bases showed the highest peak at 7.5° and at 187.5° , the two peaks being due to the twofold symmetry of the model. Figs. 3 and 4 show the sum rotation functions R for a single pyrimidine base and dimer base-pair, and for $C(3')$ *endo* and $C(2')$ *endo* sugars respectively.

To determine the translation (X_o, Z_o) of the molecule the $Q(X_o, Z_o)$ function of Tollin (1966) was calculated for both the $C(2')$ *endo* molecule and the dimer model, at orientation $(182.5, -88, 90^\circ)$ and $(7.5, -88, 90^\circ)$ respectively. The Q -function maps are

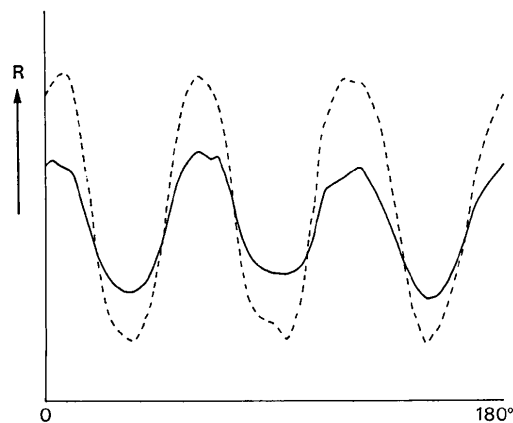


Fig. 3. The sum rotation function R for rotation of angle θ_1 . — indicates a single pyrimidine base, ---- a dimer base-pair.

shown in Figs. 5 and 6. The pseudo mm symmetry of the Q -function maps is due to the symmetry of the model. This is more marked in the case of the dimer model than in the case of the nucleoside model. False peaks are marked F and the highest peak is marked with a cross. The highest peak for the single molecule occurs at $(0.214a, 0.410c)$ and for the dimer model at $(0.247a, 0.149c)$. These positions are consistent and differ because of the different origin taken for the models. The position of the pyrimidine base in the single nucleoside model coincides with one of the bases of the dimer models.

Attempts were made to locate the second molecule from a Fourier synthesis using phases derived from the $C(2')$ *endo* nucleoside. This showed positive regions at positions consistent with a second base coplanar with the first and in a position similar to the second base in the dimer model. Although the position of the base was fairly clear there was no definite indication of the

positions of the sugar atoms of the second molecule.

Similar attempts to locate the atoms of the two sugars from a Fourier synthesis using phases derived from the dimer model of bases also failed.

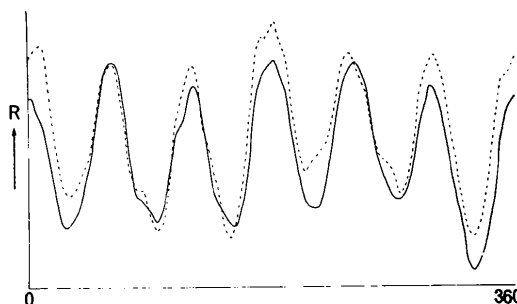


Fig. 4. The sum rotation function R for rotation of angle θ_1 . — indicates the model with $C(3')$ *endo* sugar, ---- the model with $C(2')$ *endo* sugar.

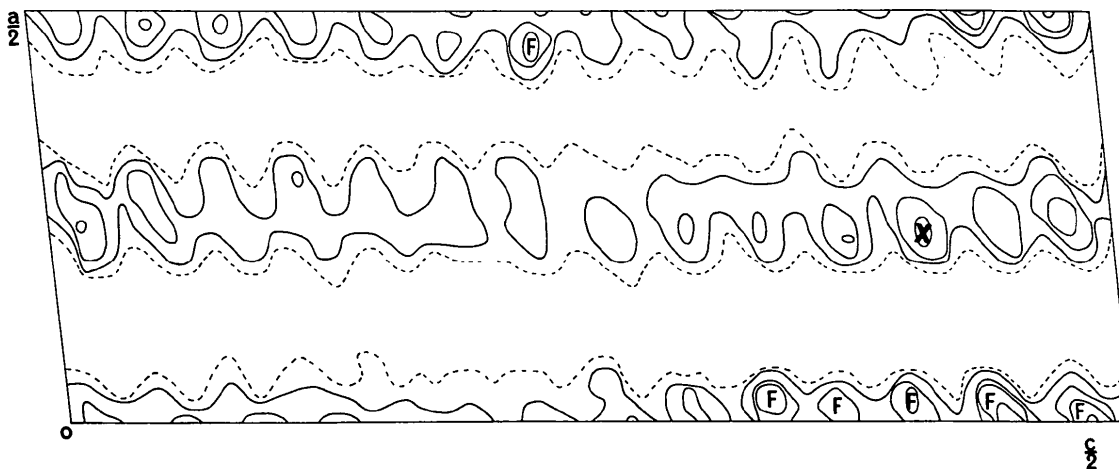


Fig. 5. The (X_0Z_0) function for a model with a $C(2')$ *endo* sugar. False peaks are indicated by F and the highest peak by X .

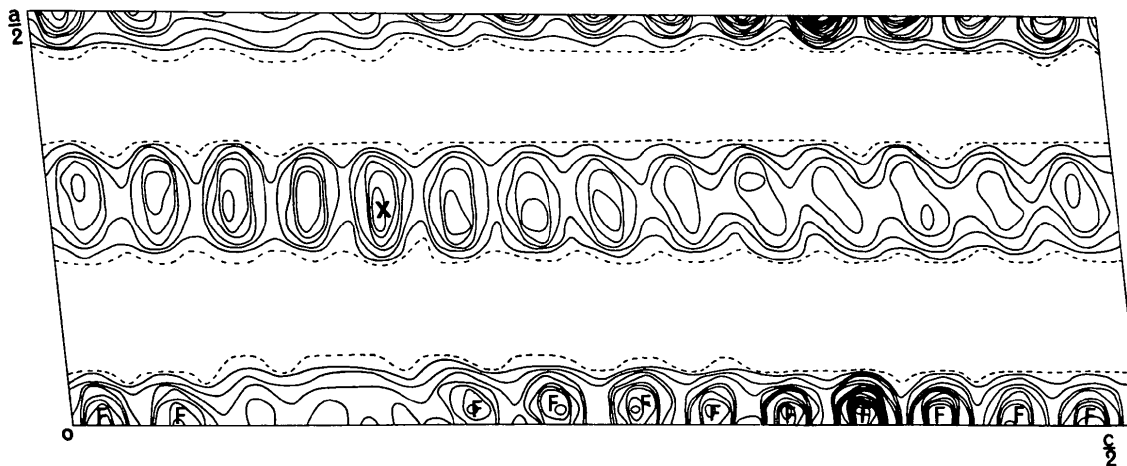


Fig. 6. The (X_0Z_0) function for a dimer model. The highest peak is marked X and false peaks are marked F .

Assuming that both molecules contained C(2') *endo* sugars and that the bases were hydrogen bonded in a coplanar arrangement, two C(2') *endo* sugars were attached to the dimer of bases and a structure factor calculation gave an R value of 0.48. A difference Fourier synthesis suggested that one molecule should be rotated through 180° about an axis along the glycosidic bond. Modifying atomic positions on the basis of the difference Fourier synthesis reduced the R value to 0.44. At this stage three cycles of least-squares refinement were carried out using isotropic temperature factors of 3 \AA^2 . The R value decreased to 0.32 but nine of the atoms in the sugar parts of the two molecules had high temperature factors. These atoms were removed and a Fourier synthesis was calculated using observed structure factors and phases derived from the other atoms. This Fourier synthesis revealed the positions of the removed atoms and a structure factor calculation gave an R value of 0.29. The arrangement of the bases was seen to be similar to that in the original dimer but with one molecule rotated through 180° about the glycosidic bond. The φ_{CN} values in the final structure were much smaller than those assumed in the theoretical models.

Refinement

The structure was refined by the method of least-squares using the block-diagonal approximation in a program written by Professor J. Trotter (Toronto University) and modified and adapted for use on an Elliott 4130 computer. Refinement was started with an isotropic temperature factor of 3 \AA^2 for all the atoms. Each reflexion was given a weight of unity except those having structure amplitudes greater than 100, in which case a weight of $(100/|F_o|)^2$ was applied. Six cycles of refinement brought the R value down to 0.13. At this stage a difference Fourier synthesis was calculated. It showed that there were no water molecules in the crystal but failed to reveal the positions of the hydrogen atoms clearly.

The positions of the hydrogen atoms, except those attached to the O(3') and O(5') atoms in each molecule, were calculated by using standard bond lengths and angles. Anisotropic temperature factors of the form

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

were introduced for the heavier atoms and three cycles of refinement were carried out. The hydrogen atoms were included in the structure factor calculations but their positional and thermal parameters, the latter being 3 \AA^2 , were not refined. The R value decreased to 0.091. At this stage another difference-Fourier synthesis was calculated using the phases of the non-hydrogen atoms. Of the hydrogen atoms attached to O(3') and O(5'), the one attached to O(5') in one molecule showed up clearest in the difference Fourier synthesis, and location of this hydrogen atom resolved the ambiguity in the hydrogen-bonding scheme involving

O(3') and O(5'). Two further cycles of anisotropic least-squares refinement were carried out, in which all the hydrogen atoms were included but not refined, and the final value of R was 0.084.

The maximum shifts in the positional parameter was 0.005 \AA , corresponding to 0.29σ , and the maximum thermal parameter shift was 0.34σ . The maximum standard deviations for bond-lengths and angles are 0.014 \AA and 1.2° .

Tables 1 and 2 show the final values of positional and thermal parameters; observed and calculated structure factors are listed in Table 3. The bond lengths and angles in the two molecules are shown in Table 4 and on Fig. 7.

Table 1. *Positional parameters of the non-hydrogen atoms with estimated standard deviations, ($\times 10^4$)*

Molecule I	x/a	y/b	z/c
N(1)	2100 (9)	-141 (13)	4932 (4)
C(2)	2188 (12)	-1807 (15)	4502 (5)
O(2)	2479 (10)	-3438 (11)	4752 (4)
N(3)	1936 (10)	-1427 (12)	3787 (4)
C(4)	1655 (11)	371 (16)	3448 (4)
O(4)	1399 (10)	470 (13)	2791 (4)
C(5)	1697 (14)	2014 (16)	3946 (5)
C(6)	1885 (12)	1733 (15)	4640 (5)
C(1')	2243 (11)	-458 (15)	5701 (4)
O(1')	1388 (8)	1111 (11)	6020 (3)
C(2')	4086 (12)	-359 (15)	6066 (5)
C(3')	3813 (12)	315 (16)	6818 (5)
O(3')	3313 (11)	-1369 (12)	7198 (4)
C(4')	2358 (13)	1757 (16)	6671 (5)
C(5')	2922 (13)	3890 (16)	6549 (5)
O(5')	3718 (10)	4642 (11)	7210 (4)
Molecule II			
N(1)	2712 (8)	4752 (12)	766 (4)
C(2)	2362 (10)	3126 (14)	1155 (4)
O(2)	2289 (7)	1470 (10)	923 (3)
N(3)	2078 (10)	3526 (11)	1849 (4)
C(4)	2068 (11)	5352 (16)	2164 (5)
O(4)	1748 (10)	5517 (12)	2790 (3)
C(5)	2397 (15)	7001 (16)	1717 (5)
C(6)	2706 (13)	6620 (16)	1041 (5)
C(1')	2962 (11)	4398 (14)	-4 (4)
O(1')	3943 (7)	5964 (10)	-227 (3)
C(2')	1336 (11)	4338 (14)	-500 (5)
C(3')	1899 (11)	5098 (14)	-1212 (5)
O(3')	2737 (9)	3485 (10)	-1515 (4)
C(4')	3235 (11)	6628 (16)	-927 (5)
C(5')	2485 (11)	8713 (15)	-844 (4)
O(5')	2119 (10)	9556 (12)	-1522 (4)

Results and discussion

The pyrimidine bases

Both pyrimidine bases are planar and the maximum deviation of a ring atom from the least-squares mean plane is 0.02 \AA for molecule I and 0.01 \AA for molecule II. The equation of the mean plane through the six atoms of the pyrimidine ring of molecule I is

$$0.9945x + 0.1045y + 0.0008z - 0.6150 = 0$$

and that for molecule II is

$$0.9480x - 0.0871y + 0.3060z - 2.0250 = 0$$

where x , y and z are coordinates in Å, with respect to axes a , b and c^* .

The displacements of all atoms of and attached to the bases are listed in Table 5.

The two pyrimidine bases are not strictly parallel, the dihedral angle between them being 21° . They are hydrogen bonded to each other through N(3) and O(4). The bond lengths and angles in the two pyrimidine rings are in good agreement with each other. The mean value of the C–N bond is 1.37 Å and the mean value of the C=O bond length is 1.21 Å. It is worth noting that the C=O bonds between C(2) and O(2) have lengths of 1.20 and 1.19 Å in molecules I and II respectively, whilst the value for the C(4)–O(4) bonds is 1.23 Å in both molecules. Similar differences between the two C=O bonds are also observed in thymidine (Young *et al.*, 1969), methyluridine (Hunt & Subramanian, 1969) and methylthymine (Hoogsteen, 1963*a, b*). The values for these C=O bonds are shown in Table 6.

The sugars

The sugar rings in both molecules are puckered so that C(2') is *endo*, *i.e.* C(2') is lying out of the mean

plane through the other sugar ring atoms on the same side of the plane as C(5'). This type of pucker is the most common in deoxynucleosides. The displacement of C(2') in molecule I is 0.58 Å and in molecule II it is 0.54 Å. The equations of the least-squares mean plane through C(1'), O(1'), C(3') and C(4') for molecules I and II are respectively

$$-0.6091x - 0.6354y + 0.4746z - 4.9300 = 0$$

and

$$-0.7178x + 0.6061y + 0.3426z - 0.0067 = 0.$$

If the displacements of the ring atoms are considered relative to the three-atom plane through C(1'), O(1') and C(4'), the pucker of the ring is C(2') *endo*–C(3') *exo*. For molecule I the *endo* displacement of C(2') is 0.43 Å and the *exo* displacement of C(3') is 0.19 Å. For molecule II the C(2') displacement is 0.32 Å and the C(3') displacement is 0.28 Å. The displacements from the sugar rings are listed in Table 7, and Fig. 8 shows views of the sugar rings along the C(1')O(1')C(4') plane. Fig. 9 shows views of the molecules along a line parallel to the glycosidic bond and passing through the pyrimidine plane.

The conformation of the C(5')–O(5') bond is *trans-gauche* in both molecules. The angles φ_{OO} and φ_{OC} , describing the orientation of C(5')–O(5') relative to

Table 2. Thermal parameters of the non-hydrogen atoms with estimated standard deviations ($\times 10^4 \text{ \AA}^2$)

Molecule I	B_{11}	B_{12}	B_{13}	B_{22}	B_{23}	B_{33}
N(1)	155 (12)	8 (28)	34 (9)	151 (19)	9 (11)	19 (2)
C(2)	172 (16)	20 (35)	25 (10)	144 (21)	36 (13)	14 (2)
O(2)	261 (14)	–109 (28)	9 (9)	128 (14)	13 (11)	26 (2)
N(3)	21 (14)	16 (29)	25 (9)	127 (17)	32 (11)	16 (2)
C(4)	147 (14)	8 (33)	35 (9)	159 (20)	13 (13)	15 (3)
O(4)	318 (16)	42 (34)	31 (9)	213 (19)	3 (11)	19 (2)
C(5)	232 (20)	–49 (38)	46 (12)	140 (25)	–20 (14)	20 (3)
C(6)	168 (16)	–16 (32)	53 (11)	106 (19)	8 (14)	26 (3)
C(1')	150 (14)	–8 (32)	5 (9)	144 (20)	9 (12)	13 (2)
O(1')	147 (10)	–46 (25)	15 (7)	229 (18)	22 (10)	20 (2)
C(2')	173 (16)	–88 (37)	25 (10)	180 (23)	–23 (14)	19 (2)
C(3')	158 (15)	–28 (36)	25 (11)	175 (23)	–22 (15)	27 (2)
O(3')	364 (18)	–89 (34)	51 (11)	191 (19)	–9 (12)	22 (3)
C(4')	191 (17)	23 (39)	22 (11)	171 (23)	11 (16)	22 (3)
C(5')	223 (19)	–2 (36)	32 (12)	141 (22)	–24 (14)	19 (2)
O(5')	319 (16)	72 (31)	21 (9)	146 (16)	4 (11)	25 (2)
Molecule II						
N(1)	123 (11)	–46 (26)	11 (8)	146 (18)	9 (11)	17 (2)
C(2)	96 (11)	16 (28)	18 (8)	115 (18)	7 (12)	15 (2)
O(2)	161 (10)	12 (23)	34 (6)	119 (13)	0 (9)	17 (2)
N(3)	161 (12)	19 (25)	27 (8)	92 (15)	7 (10)	16 (2)
C(4)	137 (14)	–24 (33)	7 (10)	144 (20)	–17 (14)	19 (3)
O(4)	325 (16)	19 (32)	48 (9)	171 (17)	–4 (10)	15 (2)
C(5)	260 (22)	20 (41)	27 (13)	156 (23)	4 (14)	19 (2)
C(6)	205 (18)	64 (37)	39 (12)	146 (20)	125 (15)	22 (3)
C(1')	146 (14)	–52 (31)	19 (9)	137 (20)	–17 (12)	10 (2)
O(1')	116 (9)	34 (22)	17 (6)	186 (16)	–5 (9)	13 (1)
C(2')	149 (14)	55 (31)	16 (9)	140 (19)	8 (11)	15 (2)
C(3')	130 (13)	42 (29)	10 (9)	112 (18)	10 (13)	19 (3)
O(3')	265 (13)	–5 (26)	43 (8)	127 (15)	7 (10)	25 (2)
C(4')	154 (15)	24 (34)	15 (9)	179 (21)	–5 (14)	13 (2)
C(5')	145 (14)	31 (32)	–12 (9)	165 (21)	26 (13)	14 (2)
O(5')	307 (15)	–14 (30)	19 (9)	173 (17)	–6 (10)	21 (2)

Table 3. Observed and calculated structure factors

H	K	L	F _{obs}	F _{calc}	H	K	L	F _{obs}	F _{calc}	H	K	L	F _{obs}	F _{calc}	H	K	L	F _{obs}	F _{calc}	
0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	
0	0	1	4.25	6.40	4	0	0	5.37	4.91	1	1	-3	20.95	22.28	5	1	-2	14.73	17.20	
0	0	2	12.78	13.09	4	0	7	10.32	11.69	1	1	-2	23.98	24.76	5	1	-1	15.07	15.17	
0	0	3	9.25	9.05	4	0	8	15.09	15.90	1	1	-1	18.09	18.09	5	1	0	17.26	17.26	
0	0	4	41.77	39.80	4	0	9	18.10	19.13	1	1	0	6.54	6.27	5	1	1	15.23	15.70	
0	0	5	51.74	48.95	4	0	10	19.30	19.50	1	1	1	7.94	7.47	5	1	2	12.98	12.49	
0	0	6	21.46	22.08	4	0	11	13.57	13.24	1	1	1	13.04	13.43	5	1	3	17.26	13.99	
0	0	7	13.13	13.33	4	0	12	18.79	21.40	1	1	1	3	17.84	18.09	5	1	4	14.87	12.92
0	0	8	37.10	33.94	4	0	13	5.99	1.87	1	1	4	33.56	34.81	5	1	5	10.35	9.51	
0	0	9	28.89	28.71	4	0	14	3.07	2.35	1	1	5	9.20	6.18	5	1	6	6.27	6.27	
0	0	10	32.26	32.06	5	0	-1	8.20	4.86	1	1	6	11.00	11.74	5	1	7	3.13	4.98	
0	0	11	9.41	4.91	5	0	-2	6.78	5.82	1	1	8	12.23	14.51	5	1	8	2.40	5.69	
0	0	12	9.05	7.88	5	0	-3	10.15	9.00	1	1	10	12.16	14.09	5	1	9	1.58	3.38	
0	0	13	24.24	24.67	5	0	-4	5.25	4.18	1	1	10	4.96	4.23	5	1	10	1.10	3.08	
0	0	14	18.10	17.46	5	0	-5	8.76	9.92	1	1	11	21.01	21.19	5	1	11	1.11	1.49	
0	0	15	9.05	7.88	5	0	-6	5.92	5.57	1	1	12	7.48	6.42	5	1	12	0.57	0.91	
0	0	16	24.24	24.67	5	0	-7	12.16	14.09	1	1	13	5.73	5.49	5	1	13	0.17	0.17	
0	0	17	9.05	7.88	5	0	-8	9.07	10.32	1	1	14	1.14	0.07	5	1	14	0.17	0.17	
0	0	18	9.05	7.88	5	0	-9	7.01	7.49	1	1	15	14.75	17.30	5	1	15	0.16	0.16	
0	0	19	9.42	7.20	5	0	-10	3.04	3.31	1	1	17	12.93	9.75	5	1	16	0.14	0.14	
0	0	20	20.85	20.85	5	0	-11	3.00	2.95	1	1	20	7.43	7.21	5	1	17	0.13	0.13	
0	0	21	9.05	7.88	5	0	-12	16.04	16.04	1	1	21	7.52	6.82	5	1	18	0.12	0.12	
0	0	22	9.05	7.88	5	0	-13	4.99	3.60	1	1	21	13.48	13.48	5	1	19	0.11	0.11	
0	0	23	9.05	7.88	5	0	-14	3.04	3.31	1	1	22	1.17	0.94	5	1	20	0.11	0.11	
0	0	24	9.05	7.88	5	0	-15	15.04	16.44	1	1	22	1.18	1.12	5	1	21	0.10	0.10	
0	0	25	9.05	7.88	5	0	-16	15.04	16.44	1	1	23	3.26	3.22	5	1	22	0.10	0.10	
0	0	26	9.05	7.88	5	0	-17	15.04	16.44	1	1	23	12.83	12.41	5	1	23	0.10	0.10	
0	0	27	9.05	7.88	5	0	-18	15.04	16.44	1	1	24	10.66	11.12	5	1	24	0.10	0.10	
0	0	28	9.05	7.88	5	0	-19	15.04	16.44	1	1	24	13.48	13.48	5	1	25	0.10	0.10	
0	0	29	9.05	7.88	5	0	-20	15.04	16.44	1	1	25	13.48	13.48	5	1	26	0.10	0.10	
0	0	30	9.05	7.88	5	0	-21	15.04	16.44	1	1	25	13.48	13.48	5	1	27	0.10	0.10	
0	0	31	9.05	7.88	5	0	-22	15.04	16.44	1	1	26	13.48	13.48	5	1	28	0.10	0.10	
0	0	32	9.05	7.88	5	0	-23	15.04	16.44	1	1	26	13.48	13.48	5	1	29	0.10	0.10	
0	0	33	9.05	7.88	5	0	-24	15.04	16.44	1	1	27	13.48	13.48	5	1	30	0.10	0.10	
0	0	34	9.05	7.88	5	0	-25	15.04	16.44	1	1	27	13.48	13.48	5	1	31	0.10	0.10	
0	0	35	9.05	7.88	5	0	-26	15.04	16.44	1	1	28	13.48	13.48	5	1	32	0.10	0.10	
0	0	36	9.05	7.88	5	0	-27	15.04	16.44	1	1	28	13.48	13.48	5	1	33	0.10	0.10	
0	0	37	9.05	7.88	5	0	-28	15.04	16.44	1	1	29	13.48	13.48	5	1	34	0.10	0.10	
0	0	38	9.05	7.88	5	0	-29	15.04	16.44	1	1	29	13.48	13.48	5	1	35	0.10	0.10	
0	0	39	9.05	7.88	5	0	-30	15.04	16.44	1	1	30	13.48	13.48	5	1	36	0.10	0.10	
0	0	40	9.05	7.88	5	0	-31	15.04	16.44	1	1	30	13.48	13.48	5	1	37	0.10	0.10	
0	0	41	9.05	7.88	5	0	-32	15.04	16.44	1	1	31	13.48	13.48	5	1	38	0.10	0.10	
0	0	42	9.05	7.88	5	0	-33	15.04	16.44	1	1	31	13.48	13.48	5	1	39	0.10	0.10	
0	0	43	9.05	7.88	5	0	-34	15.04	16.44	1	1	32	13.48	13.48	5	1	40	0.10	0.10	
0	0	44	9.05	7.88	5	0	-35	15.04	16.44	1	1	32	13.48	13.48	5	1	41	0.10	0.10	
0	0	45	9.05	7.88	5	0	-36	15.04	16.44	1	1	33	13.48	13.48	5	1	42	0.10	0.10	
0	0	46	9.05	7.88	5	0	-37	15.04	16.44	1	1	33	13.48	13.48	5	1	43	0.10	0.10	
0	0	47	9.05	7.88	5	0	-38	15.04	16.44	1	1	34	13.48	13.48	5	1	44	0.10	0.10	
0	0	48	9.05	7.88	5	0	-39	15.04	16.44	1	1	34	13.48	13.48	5	1	45	0.10	0.10	
0	0	49	9.05	7.88	5	0	-40	15.04	16.44	1	1	35	13.48	13.48	5	1	46	0.10	0.10	
0	0	50	9.05	7.88	5	0	-41	15.04	16.44	1	1	35	13.48	13.48	5	1	47	0.10	0.10	
0	0	51	9.05	7.88	5	0	-42	15.04	16.44	1	1	36	13.48	13.48	5	1	48	0.10	0.10	
0	0	52	9.05	7.88	5	0	-43	15.04	16.44	1	1	36	13.48	13.48	5	1	49	0.10	0.10	
0	0	53	9.05	7.88	5	0	-44	15.04	16.44	1	1	37	13.48	13.48	5	1	50	0.10	0.10	
0	0	54	9.05	7.88	5	0	-45	15.04	16.44	1	1	37	13.48	13.48	5	1	51	0.10	0.10	
0	0	55	9.05	7.88	5	0	-46	15.04	16.44	1	1	38	13.48	13.48	5	1	52	0.10	0.10	
0	0	56	9.05	7.88	5	0	-47	15.04	16.44	1	1	38	13.48	13.48	5	1	53	0.10	0.10	
0	0	57	9.05	7.88	5	0	-48	15.04	16.44	1	1	39	13.48	13.48	5	1	54	0.10	0.10	
0	0	58	9.05	7.88	5	0	-49	15.04	16.44	1	1	39	13.48	13.48	5	1	55	0.10	0.10	
0	0	59	9.05	7.88	5	0	-50	15.04	16.44	1	1	40	13.48	13.48	5	1	56	0.10	0.10	
0	0	60	9.05	7.88	5	0	-51	15.04	16.44	1	1	40	13.48	13.48	5	1	57	0.10	0.10	
0	0	61	9.05	7.88	5	0	-52	15.04	16.44	1	1	41	13.48	13.48	5	1	58	0.10	0.10	
0	0	62	9.05	7.88	5	0	-53	15.04	16.44	1	1	41	13.48	13.48	5	1	59	0.10	0.10	
0	0	63	9.05	7.88	5	0	-54	15.04	16.44	1	1	42	13.48	13.48	5	1	60	0.10	0.10	
0	0	64	9.05	7.88	5	0	-55	15.04	16.44	1	1	42	13.48	13.48	5	1	61	0.10	0.10	
0	0	65	9.05	7.88	5	0	-56	15.04	16.44	1	1	43	13.48	13.48	5	1	62	0.10	0.10	
0	0	66	9.05	7.88	5	0	-57	15.04	16.44	1	1	43	13.48	13.48	5	1	63	0.10	0.10	
0	0	67	9.05	7.88	5	0	-58	15.04	16.44	1	1	44	13.48	13.48	5	1	64	0.10	0.10	
0	0	68	9.05	7.88	5	0	-59	15.04	16.44	1	1	44	13.48	13.48	5	1	65	0.10	0.10	
0	0	69	9.05	7.88	5	0	-60	15.04	16.44	1	1	45	13.48	13.48	5	1	66	0.10	0.10	
0	0	70	9.05	7.88	5	0	-61	15.04	16.44	1	1	45	13.48	13.48	5	1	67	0.10	0.10	
0	0	71	9.05	7.88	5	0	-62	15.04	16.44	1	1	46	13.48	13.48	5	1	68	0.10	0.10	
0	0	72	9.05	7.88	5	0	-63	15.04	16.44	1	1	46	13.48	13.48	5	1	69	0.10	0.10	
0	0	73	9.05	7.88	5	0	-64	15.04	16.44	1	1	47	13.48	13.48	5	1	70	0.10	0.10	
0	0	74	9.05	7.88	5	0	-65	15.04	16.44	1	1	47	13.48	13.48	5	1	71	0.10	0.10	
0	0	75	9.05	7.88	5	0	-66	15.04	16.44	1	1	48	13.48	13.48	5	1	72	0.10	0.10	
0	0	76	9.05	7.88	5	0	-67	15.04	16.44	1	1	48	13.48	13.48	5	1	73	0.10	0.10	
0	0	77	9.05	7.88																

Table 4. Bond lengths and angles in the two molecules

(a) Bond lengths	Molecule I	Molecule II
N(1)–C(2)	1.39 Å	1.36 Å
C(2)–O(2)	1.20	1.19
C(2)–N(3)	1.36	1.37
N(3)–C(4)	1.37	1.36
C(4)–O(4)	1.23	1.23
C(4)–C(5)	1.44	1.43
C(5)–C(6)	1.31	1.34
C(6)–N(1)	1.37	1.36
N(1)–C(1')	1.45	1.50
C(1')–O(1')	1.42	1.40
O(1')–C(4')	1.43	1.44
C(1')–C(2')	1.54	1.50
C(2')–C(3')	1.52	1.54
C(3')–O(3')	1.42	1.42
C(3')–C(4')	1.51	1.53
C(4')–C(5')	1.52	1.54
C(5')–O(5')	1.42	1.39
(b) Bond angles		
N(1)–C(2)–O(2)	121.8°	123.8°
C(2)–N(1)–C(6)	121.2	121.9
C(2)–N(1)–C(1')	117.3	116.6
N(1)–C(2)–N(3)	114.4	114.7
C(2)–N(3)–C(4)	128.5	126.8
N(3)–C(4)–O(4)	120.9	120.5
N(3)–C(4)–C(5)	112.4	115.4
O(4)–C(4)–C(5)	126.7	124.1
C(4)–C(5)–C(6)	121.8	118.2
C(5)–C(6)–N(1)	121.6	123.0
C(6)–N(1)–C(1')	121.5	121.3
N(1)–C(1')–O(1')	108.9	108.0
N(1)–C(1')–C(2')	113.5	113.8
O(1')–C(1')–C(2')	104.7	107.5
C(1')–C(2')–C(3')	101.4	102.5
C(2')–C(3')–C(4')	102.0	110.1
C(2')–C(3')–O(3')	107.9	106.8
O(3')–C(3')–C(4')	110.7	108.5
C(1')–O(1')–C(4')	110.5	109.6
C(3')–C(4')–C(5')	113.6	112.9
C(3')–C(4')–O(1')	106.6	107.1
O(1')–C(4')–C(5')	107.1	107.7
C(4')–C(5')–O(5')	108.1	108.7

Table 5. Deviations of atoms from the least-squares plane of the pyrimidine ring

	Molecule I	Molecule II
N(1)	–0.023 Å	–0.012 Å
C(2)	0.021	0.013
N(3)	0.001	–0.005
C(4)	–0.021	–0.005
C(5)	0.021	0.008
C(6)	0.001	0.001
O(2)	0.082	0.056
O(4)	–0.076	0.015
C(1')	–0.097	0.061

Table 6. A comparison between the C(2)–O(2) and C(4)–O(4) bonds in UdR and some related structures

Structure	C(2)–O(2)	C(4)–O(4)	e.s.d.	Reference
Thymidine	1.206 Å	1.230 Å	0.006 Å	Young, Tollin & Wilson (1969)
Methyluridine	1.196	1.223	0.006	Hunt & Subramanian (1969)
Methylthymine I	1.214	1.230	0.004	Hoogsteen (1963a)
Methylthymine II	1.207	1.230	0.005	Hoogsteen (1963b)
Uracil	1.215	1.245	0.002	Stewart & Jensen (1967)
Deoxyuridine I	1.203	1.228	0.011	
Deoxyuridine II	1.192	1.235	0.011	

the bonds C(4')–O(1') and C(4')–C(3') (Shefter & Trueblood, 1965) are given in Table 8 and are illustrated in Fig. 10. The same type of C(5')–O(5') conformation has been observed in 5-fluorodeoxyuridine (Harris & Macintyre, 196 and i 5-iodouridine Rahman & Wilson, 1970) although it is not the most common type of conformation in nucleosides and nucleotides. Table 9 gives a comparison of the deoxy-nucleosides related to UdR.

The bond lengths and angles (Table 4) of the sugars in the two molecules are in good agreement with those found in other nucleosides and nucleotides. As in other structures (Sundaralingam & Jensen, 1965) the C(4')–O(1') bond is slightly longer than the C(1')–O(1') bond. The three types of angles within the furanose ring, C–C–C, C–C–O and C–O–C, have average values of 101.7, 105.6 and 110.5° in molecule I and average values of 101.3, 107.3 and 109.6° in molecule II. Averaging for the two molecules, the three types of angles have values of 101.5, 106.4 and 110.1° respectively. The differences between these are significant and agree with those found in other nucleosides and nucleotides. As in other structures the exocyclic angles N(1)–C(1')–O(1') and C(5')–C(4')–O(1') are significantly smaller than the angles N(1)–C(1')–C(2') and C(5')–C(4')–C(3').

Table 7. Deviations of atoms of the sugar ring from the plane through atoms C(1')O(1')C(4')

	Molecule I	Molecule II
C(2')	–0.431 Å	–0.323 Å
C(3')	0.191	0.275
O(3')	1.568	1.680
C(5')	–1.328	–1.367
O(5')	–1.292	–1.297

Table 8. Summary of the conformational parameters of the molecules

	Molecule I	Molecule II
φ_{CN}	–26°	–25°
φ_{OO}	186	192
φ_{OC}	69	71
$\varphi_{O(1')-C(1')}$	–16.8	–13.0
$\varphi_{C(1')-C(2')}$	33.9	31.1
$\varphi_{C(2')-C(3')}$	–37.5	–35.3
$\varphi_{C(3')-C(4')}$	28.8	28.8
$\varphi_{C(4')-O(1')}$	–7.6	–10.9

orientation of the sugar and the base (Donohue & Trueblood, 1960) is -26° in molecule I and -25° in molecule II. These values are appreciably smaller than those normally observed in C(2') *endo* pyrimidine molecules and this is probably related to the low values of the dihedral angles between base and sugar rings. Although the φ_{CN} values are lower than normally found they lie within the 'allowed range' determined by Haschemeyer & Rich (1962) and based on van der Waals contacts. Also, a calculation of the potential energy as a function of φ_{CN} , based on non-bonded interactions (Wilson & Rahman, 1971) shows that the deepest minimum occurs within 5° of the observed φ_{CN} in both molecules.

The angle φ_{CN} , together with the angles φ_{OO} and φ_{OC} , and the displacements of the atoms in the sugar ring define the conformation of the molecule. An alternative description of the pucker in the sugar ring is in terms of the torsional angles around the ring. These are listed in Table 8.

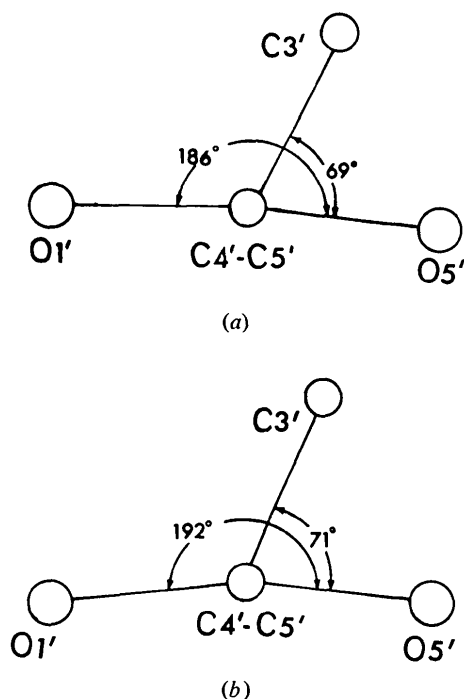


Fig. 10. The conformation of the C(4')-O(5') bond: (a) molecule I; (b) molecule II.

One of the reasons for studying deoxyuridine was to see whether it would have a C(3') *exo* pucker of the sugar ring like thymidine. This, however, is not the case. Both UdR molecules are similar to other uracil-containing deoxynucleosides in having C(2') *endo* puckers of the sugar. This is the case for fluorodeoxyuridine (Harris & McIntyre, 1964), bromodeoxyuridine (Iball, Morgan & Wilson, 1966), and iododeoxyuridine (Camerman & Trotter, 1965).

Hydrogen bonds and molecular packing

The arrangement of the molecules in the crystal and the hydrogen bonding scheme is shown in Fig. 11. O(4) in molecule I is accepting a hydrogen atom from N(3) in molecule II. The distance O(4)-N(3) is 2.8 Å and the angle C(4)-N(3)-O(4) is 112° . Similarly, O(4) in molecule II is accepting a hydrogen atom from N(3) in molecule I at $(x, y+1, z)$. The distance O(4)-N(3) $(x, y+1, z)$ is 2.77 Å and the angle C(4)-N(3)-O(4) is 114° . The bases form an infinite sheet parallel to the (100) plane, each sheet separated from the other by approximately 3.9 Å along the *a* axis. The hydrogen bonds are in a zigzag arrangement within the sheet. Crystals of UdR have a tendency to split into thin plates parallel to the (100) plane and this is no doubt due to the sheet-like structure.

In the sugar part, O(5') in molecule I is at a distance of 2.70 Å from O(3') in molecule I at $(x, y+1, z)$ and 2.71 Å from O(3') in molecule II at $(x, y, z-1)$. O(5') in molecule II is at a distance of 2.68 Å from O(3') in molecule II at $(x, y+1, z)$ and 2.75 Å from O(3') in molecule I at $(x, y+1, z+1)$. In the second difference Fourier synthesis mentioned above the hydrogen atom attached to O(5') in molecule I was clearly located at a distance of 1.00 Å from O(5'). The angle between the O(5')-O(5')H and O(5')-O(3') bonds in molecule I at $(x, y+1, z)$ is 11° , so that in the hydrogen bonding scheme O(5') in molecule I donates a hydrogen atom to O(3') in molecule II at $(x, y, z+1)$. The angles C(5')-O(5')-O(3') and C(3')-O(3')-O(5') are 107° and 109° respectively. Similarly, O(5') in molecule II donates a hydrogen atom to O(3') in molecule II at $(x, y+1, z)$ and accepts a hydrogen from O(3') in molecule I at $(x, y+1, z-1)$. The angles C(5')-O(5')-O(3') and C(3')-O(3')-O(5') are 112° and 113° respectively. Apart from the interatomic distances mentioned above, no other atoms are closer than the sum of their van der Waals radii. O(2) in both molecules does not take part in hydrogen bonding.

Table 9. A comparison of the conformational parameters of deoxynucleosides related to UdR

Nucleoside	φ_{CN}	Sugar pucker	C(5')-O(5') conformation	Reference
Fluorodeoxyuridine	-59°	C(2') <i>endo</i>	<i>trans-gauche</i>	Harris & Macintyre (1964)
Bromodeoxyuridine	-47°	C(2') <i>endo</i>	<i>gauche-trans</i>	Iball, Morgan & Wilson (1966)
Iododeoxyuridine	-63°	C(2') <i>endo</i>	<i>gauche-gauche</i>	Camerman & Trotter (1965)
Thymidine	-39°	C(3') <i>exo</i>	<i>gauche-trans</i>	Young, Tollin & Wilson (1969)
Deoxyuridine I	-28°	C(2') <i>endo</i>	<i>trans-gauche</i>	
Deoxyuridine II	-24°	C(2') <i>endo</i>	<i>trans-gauche</i>	

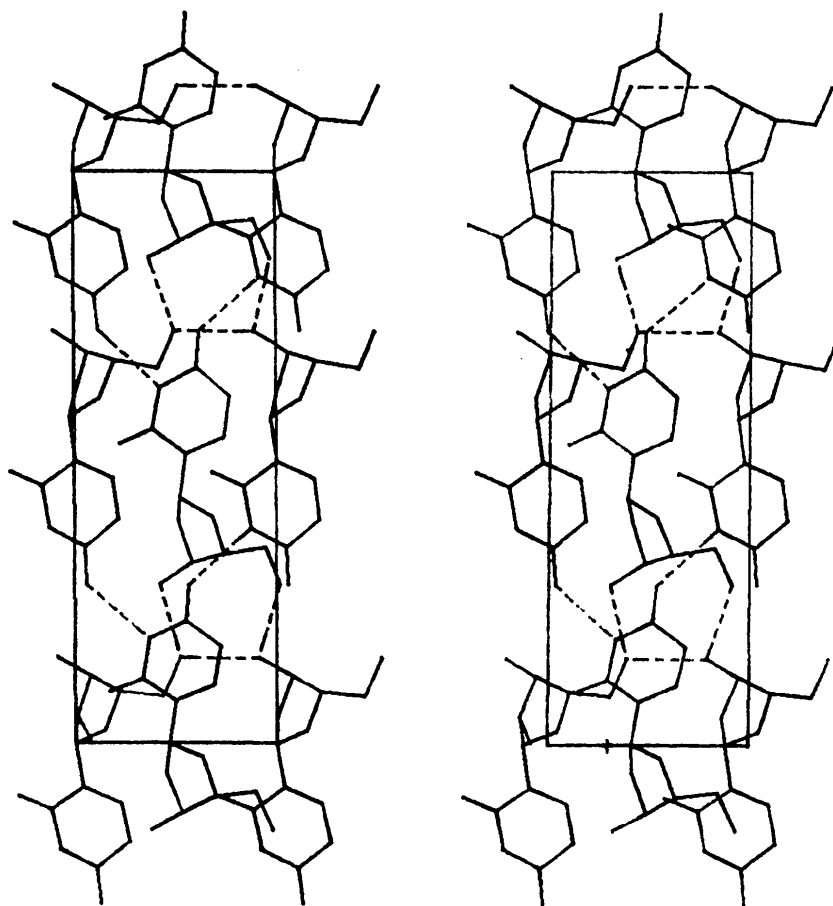


Fig. 11. Stereopair of the Udr structure viewed along the a axis. Broken lines represent hydrogen bonds.

We thank Dr P. Tollin for advice and discussion and Drs A.R.I. Munns and D. W. Young for discussion. We also thank Mr J. Low for assistance with the data collection and Miss J. Anderson for computational assistance. A Science Research Council grant for the purchase and maintenance of the diffractometer, and financial support by the Medical Research Council are gratefully acknowledged.

References

- ARNOTT, S. & HUKINS, D. W. L. (1969). *Nature, Lond.* **224**, 886.
- CAMERMAN, N. & TROTTER, J. (1965). *Acta Cryst.* **18**, 203.
- DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363.
- HARRIS, D. R. & MACINTYRE, W. M. (1964). *Biophys. J.* **4**, 203.
- HASCHEMEYER, A. E. V. & RICH, A. (1967). *J. Mol. Biol.* **27**, 369.
- HOOGSTEEN, K. (1963a). *Acta Cryst.* **16**, 28.
- HOOGSTEEN, K. (1963b). *Acta Cryst.* **16**, 907.
- HUNT, J. & SUBRAMANIAN, E. (1969). *Acta Cryst.* **B25**, 2144.
- IBALL, J., MORGAN, C. H. & WILSON, H. R. (1966). *Proc. Roy. Soc. A* **295**, 320.
- LANGRIDGE, R. & MARMUR, J. (1964). *Science*, **143**, 1450.
- RAHMAN, A. & WILSON, H. R. (1970). *Acta Cryst.* **B26**, 1765.
- RAHMAN, A. & WILSON, H. R. (1971). *Nature, Lond.* **232**, 333.
- SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067.
- STEWART, R. F. & JENSEN, L. H. (1967). *Acta Cryst.* **23**, 1102.
- SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821.
- SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 930.
- TAKAHASHI, I. (1963). *J. Genet. Microbiol.* **31**, 211.
- TAKAHASHI, I. & MARMUR, J. (1963). *Nature, Lond.* **197**, 794.
- TOLLIN, P. (1966). *Acta Cryst.* **21**, 613.
- TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1322.
- TOLLIN, P., WILSON, H. R. & YOUNG, D. W. (1968). *Nature, Lond.* **217**, 1148.
- WILSON, H. R. (1970). *Nature, Lond.* **225**, 545.
- WILSON, H. R. & RAHMAN, A. (1971). *J. Mol. Biol.* **56**, 129.
- YOUNG, D. W., TOLLIN, P. & WILSON, H. R. (1969). *Acta Cryst.* **B25**, 1423.