

## The Crystal and Molecular Structure of 2'-Deoxycytidine

BY D. W. YOUNG AND H. R. WILSON

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

(Received 4 November 1974; accepted 7 November 1974)

$C_9H_{13}N_3O_4$  is triclinic, space group  $P1$ , with  $a=7.285$ ,  $b=6.886$ ,  $c=11.074$  Å,  $\alpha=104^\circ 19'$ ,  $\beta=84^\circ 53'$ ,  $\gamma=72^\circ 26'$ ,  $Z=2$ . The structure was solved by Patterson methods. The final  $R$  for 1274 counter reflexions is 0.054. The two molecules in the cell are linked through two hydrogen bonds between the cytosine bases. The glycosyl torsion angle,  $O(4')-C(1')-N(1)-C(2)$ , is  $201.2^\circ$  for molecule I and  $222.2^\circ$  for molecule II. The sugar ring pucker relative to the least-squares plane through the five-atom sugar ring is  $C(3')-endo-C(2')-exo$  in molecule I and  $C(2')-endo-C(3')-exo$  in molecule II. In both molecules the orientation of  $C(5')-O(5')$  is *gauche* relative to both  $C(4')-O(4')$  and  $C(4')-C(3')$ .

### Introduction

In our conformational studies of nucleic acid components we have determined the crystal structure of 2'-deoxycytidine (CDR). It is of particular interest to compare the conformational parameters of CDR (Fig. 1) with those of deoxycytidine hydrochloride (Subramanian & Hunt, 1970) and deoxycytidine 5'-phosphate (Viswamitra, Reddy, Lin & Sundaralingam, 1971), in both of which the glycosyl torsion angles have values at the extreme end of the observed range. Comparison of the conformation with those of cytidine (Furberg, Petersen & Rømming, 1965) and cytidine 3'-phosphate (Sundaralingam & Jensen, 1965; Bugg & Marsh, 1967) is also of interest.

### Experimental

Crystals were obtained from aqueous solutions. Weissenberg and precession photographs showed that the crystals were triclinic and the asymmetry of the nucleoside establishes the space group  $P1$ . The cell dimensions are:  $a=7.285$  (6),  $b=6.866$  (6),  $c=11.074$  (9) Å,  $\alpha=104^\circ 19'$  (5)',  $\beta=84^\circ 53'$  (5)',  $\gamma=72^\circ 26'$  (5)', [ $\lambda(\text{Mo } K\alpha)=0.71069$  Å].

The calculated density for  $Z=2$  is  $1.560$  g cm $^{-3}$ .

The intensities of 1274 significant out of 2122 unique reflexions within a sphere corresponding to the limiting sphere for Cu  $K\alpha$  radiation were measured on a Hilger and Watts linear diffractometer with Mo  $K\alpha$  radiation and balanced filters. Three crystals, mounted about **a**, **b** and **c**, were used. The crystals were very small and no absorption corrections were applied.

### Structure determination

The orientations of the pyrimidine bases were determined by evaluating the  $I(\theta, \varphi)$  function of Tollin & Cochran (1964), for  $\theta$  and  $\varphi$  values ranging from  $0^\circ$  to  $180^\circ$ . The largest 300 sharpened Patterson coefficients were used with a disc radius of 4 Å. The two largest peaks in the map occur at  $\theta=41^\circ$ ,  $\varphi=165^\circ$  and

$\theta=139^\circ$ ,  $\varphi=13^\circ$  (Fig. 2). Since the second set of angles is equivalent to  $\theta=41^\circ$ ,  $\varphi=193^\circ$ , this suggests that the two bases are nearly parallel to each other and possibly nearly coplanar. The intermediate orientation,  $\theta=41^\circ$ ,  $\varphi=180^\circ$ , was taken for the mean direction of the normals to the pyrimidine planes, and an  $I(\theta_1, \theta_2, \theta_3)$  function (Munns, 1971) was calculated to determine the azimuthal orientations of the pyrimidine rings. This function showed three peaks instead of the possible six, suggesting that the two pyrimidine rings either had the same azimuthal angles or angles differing by  $180^\circ$ . This observation suggested the possibility of dimer association of the rings as is often found in related structures (Voet & Rich, 1970). Structure factors calculated with one cytosine base in the orientation indicated by the largest peak in the  $I(\theta_1, \theta_2, \theta_3)$  map gave an  $R$  of 0.50. The calculated phases were used to obtain an electron density map which contained peaks consistent with a second pyrimidine ring forming a dimer with the first. Several other peaks were also observed which were consistent with their belonging to the deoxyribose sugars. The original cytosine base and three peaks thought to be part of its sugar attachment gave an  $R$  of 0.45 and the electron density map thus phased revealed all the atoms of the second molecule. This molecule and the initial cytosine base gave an  $R$  of 0.40 and the electron density map phased with these atoms revealed all the remaining atoms. With an overall isotropic temperature factor of  $3.0$  Å $^2$  the value of  $R$  was 0.28.

### Refinement

The structure was refined by block-diagonal least-squares calculations. All observed structure factors were assigned unit weights and the non-hydrogen atom coordinates and individual isotropic temperature factors were refined. After six cycles the value of  $R$  had dropped to 0.097. A difference map calculated at this stage contained peaks, most of which were interpretable as hydrogen atoms.

With fixed hydrogen atom coordinates and isotropic

temperature factors of  $3.0 \text{ \AA}^2$ , the positional and isotropic thermal parameters of the non-hydrogen atoms were refined. The weighting scheme used at this stage was:

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

where  $a$  and  $b$  were set as 12.0 and 16.0 respectively.

Anisotropic temperature factors were of the form:

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

After six cycles the value of  $R$  was 0.054\* and in the last cycle the root-mean-square shifts in the positional and thermal parameters of the non-hydrogen atoms were  $0.08\sigma$  and  $0.23\sigma$  respectively.

Prior to the final two cycles a difference synthesis was computed and there were peaks in positions consistent with the location of all the hydrogen atoms except those attached to C(5') of molecule II. The peak heights at the hydrogen positions varied between

\* A copy of this final data set with the observed and calculated structure amplitudes has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 30768 (10 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

0.51 and  $0.28 \text{ e \AA}^{-3}$ . There were three other peaks in the difference maps of height greater than  $0.28 \text{ e \AA}^{-3}$ , the largest being of height  $0.36 \text{ e \AA}^{-3}$ .

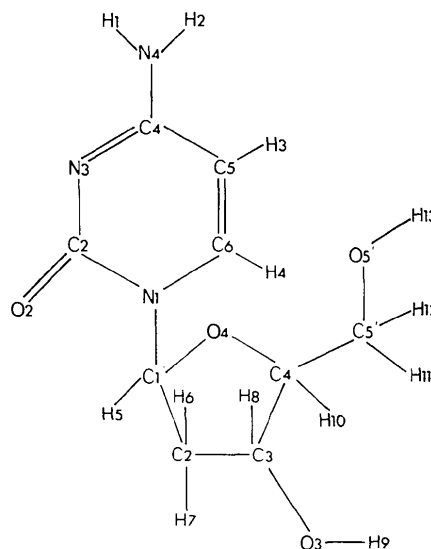


Fig. 1. The CDR molecule and the atom numbering.

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

| Molecule I  |            |            |           |          |           |           |          |          |          |
|-------------|------------|------------|-----------|----------|-----------|-----------|----------|----------|----------|
|             | $x/a$      | $y/b$      | $z/c$     | $B_{11}$ | $B_{12}$  | $B_{13}$  | $B_{22}$ | $B_{23}$ | $B_{33}$ |
| O(2)        | 6942 (8)   | 14651 (7)  | 8986 (5)  | 198 (12) | -172 (17) | -137 (13) | 134 (11) | 88 (12)  | 96 (5)   |
| C(2)        | 6800 (10)  | 12855 (10) | 8793 (7)  | 142 (15) | -124 (23) | -31 (16)  | 133 (15) | 69 (15)  | 54 (6)   |
| N(3)        | 8046 (9)   | 11281 (9)  | 9172 (6)  | 128 (12) | -92 (20)  | -68 (13)  | 126 (12) | 72 (13)  | 66 (5)   |
| C(4)        | 7853 (11)  | 9360 (10)  | 8928 (7)  | 163 (16) | -63 (25)  | -74 (16)  | 105 (15) | 56 (15)  | 54 (6)   |
| N(4)        | 9005 (10)  | 7917 (9)   | 9362 (6)  | 192 (14) | -121 (22) | -121 (15) | 132 (13) | 97 (14)  | 83 (6)   |
| C(5)        | 6378 (11)  | 8894 (11)  | 8244 (7)  | 172 (17) | -110 (26) | -51 (18)  | 115 (15) | 44 (17)  | 72 (7)   |
| C(6)        | 5154 (10)  | 10477 (11) | 7912 (7)  | 124 (14) | -104 (24) | -67 (16)  | 145 (16) | 49 (16)  | 58 (6)   |
| N(1)        | 5315 (8)   | 12461 (8)  | 8192 (5)  | 126 (12) | -95 (19)  | -64 (13)  | 115 (12) | 63 (13)  | 55 (5)   |
| C(1')       | 3848 (11)  | 14279 (11) | 7928 (7)  | 137 (16) | -19 (24)  | -68 (17)  | 128 (15) | 72 (16)  | 67 (6)   |
| C(2')       | 4535 (12)  | 14691 (12) | 6691 (7)  | 175 (16) | -119 (26) | -99 (17)  | 161 (16) | 76 (17)  | 65 (6)   |
| C(3')       | 3658 (11)  | 13376 (11) | 5729 (7)  | 168 (17) | -97 (26)  | -50 (17)  | 147 (16) | 55 (16)  | 56 (6)   |
| O(3')       | 3343 (9)   | 14098 (9)  | 4644 (5)  | 310 (16) | -262 (25) | -117 (15) | 334 (16) | 168 (13) | 67 (5)   |
| C(4')       | 1726 (11)  | 13623 (12) | 6483 (7)  | 118 (15) | 5 (28)    | -73 (16)  | 186 (18) | 61 (18)  | 65 (6)   |
| O(4')       | 2142 (7)   | 13726 (8)  | 7758 (5)  | 127 (10) | -80 (19)  | -61 (11)  | 192 (12) | 83 (11)  | 60 (4)   |
| C(5')       | 868 (12)   | 11891 (14) | 6046 (8)  | 171 (18) | -169 (33) | -140 (20) | 314 (24) | 91 (22)  | 80 (8)   |
| O(5')       | 2114 (9)   | 9818 (9)   | 5977 (6)  | 220 (13) | -231 (20) | -64 (16)  | 213 (13) | 71 (15)  | 107 (6)  |
| Molecule II |            |            |           |          |           |           |          |          |          |
|             | $x/a$      | $y/b$      | $z/c$     | $B_{11}$ | $B_{12}$  | $B_{13}$  | $B_{22}$ | $B_{23}$ | $B_{33}$ |
| O(2)        | 12335 (8)  | 6182 (7)   | 11446 (5) | 164 (10) | -181 (16) | -108 (12) | 146 (10) | 129 (11) | 104 (5)  |
| C(2)        | 12818 (10) | 7750 (11)  | 11401 (7) | 106 (14) | -93 (25)  | -35 (16)  | 171 (16) | 82 (16)  | 56 (6)   |
| N(3)        | 11917 (8)  | 9156 (9)   | 10806 (6) | 120 (12) | -73 (20)  | -66 (13)  | 131 (12) | 67 (13)  | 62 (5)   |
| C(4)        | 12458 (10) | 10847 (10) | 10809 (6) | 137 (15) | -102 (23) | -15 (16)  | 108 (14) | 48 (15)  | 48 (6)   |
| N(4)        | 11479 (9)  | 12198 (9)  | 10221 (6) | 117 (12) | -73 (21)  | -80 (14)  | 171 (14) | 112 (14) | 79 (6)   |
| C(5)        | 14032 (11) | 11180 (11) | 11387 (8) | 130 (15) | -143 (24) | -59 (18)  | 155 (16) | 78 (17)  | 83 (7)   |
| C(6)        | 14964 (10) | 9768 (11)  | 11969 (7) | 103 (14) | -107 (23) | -34 (17)  | 161 (16) | 85 (16)  | 72 (7)   |
| N(1)        | 14330 (8)  | 8119 (8)   | 12037 (6) | 131 (12) | -104 (19) | -46 (13)  | 128 (12) | 88 (12)  | 57 (5)   |
| C(1')       | 15245 (10) | 6639 (10)  | 12742 (7) | 142 (15) | -110 (25) | -57 (17)  | 126 (15) | 64 (16)  | 56 (6)   |
| C(2')       | 17315 (11) | 5095 (12)  | 12098 (7) | 145 (16) | -63 (28)  | -59 (18)  | 182 (18) | 42 (18)  | 60 (6)   |
| C(3')       | 18186 (11) | 4639 (11)  | 13246 (7) | 166 (17) | -105 (26) | 5 (18)    | 152 (16) | 64 (16)  | 57 (7)   |
| O(3')       | 17441 (9)  | 3163 (9)   | 13675 (6) | 232 (14) | -98 (20)  | -18 (15)  | 159 (12) | 104 (12) | 93 (5)   |
| C(4')       | 17349 (10) | 6772 (11)  | 14250 (7) | 123 (15) | -75 (25)  | -38 (18)  | 146 (16) | 73 (17)  | 65 (7)   |
| O(4')       | 15505 (7)  | 7870 (8)   | 13921 (4) | 139 (10) | -74 (18)  | -69 (11)  | 164 (11) | 38 (11)  | 49 (4)   |
| C(5')       | 18639 (12) | 8156 (13)  | 14396 (8) | 195 (18) | -141 (29) | -99 (19)  | 197 (19) | 42 (20)  | 78 (7)   |
| O(5')       | 19042 (8)  | 8761 (8)   | 13283 (6) | 178 (13) | -140 (20) | 0 (13)    | 209 (13) | 115 (14) | 122 (6)  |

The peaks initially interpreted as H(11) (II) and H(12) (II) were too close together, but the calculated positions of these two hydrogens were only about 0.1 Å away from each peak and still in regions of high electron density. For these two hydrogens the calculated coordinates were used in the final refinement cycles.

Table 2. *Positional parameters of the hydrogen atoms*

| Molecule I |            |            |            |
|------------|------------|------------|------------|
|            | <i>x/a</i> | <i>y/b</i> | <i>z/c</i> |
| H(1)       | 0.9669     | 0.7989     | 1.0024     |
| H(2)       | 0.8573     | 0.6753     | 0.9321     |
| H(3)       | 0.6376     | 0.7552     | 0.7925     |
| H(4)       | 0.4012     | 1.0346     | 0.7540     |
| H(5)       | 0.3266     | 1.5796     | 0.8753     |
| H(6)       | 0.5909     | 1.4008     | 0.6438     |
| H(7)       | 0.3759     | 1.6733     | 0.7010     |
| H(8)       | 0.4290     | 1.2089     | 0.5551     |
| H(9)       | 0.4588     | 1.3344     | 0.4133     |
| H(10)      | 0.0556     | 1.5354     | 0.6683     |
| H(11)      | -0.0392    | 1.2018     | 0.6668     |
| H(12)      | 0.0784     | 1.1904     | 0.5196     |
| H(13)      | 0.3370     | 0.9023     | 0.5392     |

## Molecule II

|       | <i>x/a</i> | <i>y/b</i> | <i>z/c</i> |
|-------|------------|------------|------------|
| H(1)  | 1.0545     | 1.2181     | 0.9811     |
| H(2)  | 1.1332     | 1.3782     | 1.0840     |
| H(3)  | 1.4846     | 1.2280     | 1.1251     |
| H(4)  | 1.6512     | 0.9731     | 1.2292     |
| H(5)  | 1.4375     | 0.5929     | 1.2893     |
| H(6)  | 1.8229     | 0.5994     | 1.1442     |
| H(7)  | 1.7503     | 0.3964     | 1.1426     |
| H(8)  | 1.9588     | 0.4050     | 1.2752     |
| H(9)  | 1.8326     | 0.1984     | 1.3341     |
| H(10) | 1.6970     | 0.6715     | 1.5174     |
| H(11) | 1.7987     | 0.9506     | 1.5164     |
| H(12) | 1.9955     | 0.7342     | 1.4633     |
| H(13) | 2.0033     | 0.7814     | 1.2808     |

## Results and discussion

The values of the positional and thermal parameters of the non-hydrogen atoms are given in Table 1. The positional parameters of the hydrogen atoms are given in Table 2. The bond lengths and angles involving non-hydrogen atoms are given in Table 3. The average estimated standard deviation in bond lengths is 0.01 Å, and in the bond angles, 0.7°.

*The cytosine bases*

The least-squares plane through the ring atoms of molecule I is:

$$-0.4711x - 0.1208y + 0.8738z - 4.1602 = 0$$

and that for molecule II is:

$$-0.5060x + 0.1709y + 0.8454z - 5.4500 = 0$$

where *x*, *y* and *z* are coordinates, in Å, with respect to the axial set  $O(\mathbf{b} \times \mathbf{c}^*, \mathbf{b}, \mathbf{c}^*)$ .

The polar angles,  $\theta$  and  $\varphi$ , defining the normals to these planes with respect to the axial set  $O(\mathbf{a}^*, \mathbf{b}, \mathbf{a}^* \times \mathbf{b})$  are  $\theta = 39.1^\circ$ ,  $\varphi = 191.3^\circ$ , and  $\theta = 42.1^\circ$ ,  $\varphi = 165.0^\circ$ , which are in good agreement with those found in the  $I(\theta, \varphi)$  function.

Both cytosine rings are essentially planar although that of molecule II is less strictly so. The deviations of some of the attached atoms from the least-squares planes of the cytosine rings are quite large (Table 4), although similar deviations are observed in related structures.

Cytosine bases in crystal structures occur in both the protonated and unprotonated forms. In deoxycytidine hydrochloride (Subramanian & Hunt, 1970), deoxycytidine 5'-phosphate (Viswamitra *et al.*, 1971) and both forms of cytidine monophosphate (Sundara-

Table 3. *Bond lengths (Å) and angles (°) for the non-hydrogen atoms*

|                 | Molecule I | Molecule II |                   | Molecule I | Molecule II |
|-----------------|------------|-------------|-------------------|------------|-------------|
| O(2)-C(2)       | 1.239      | 1.243       | N(1)-C(1')        | 1.491      | 1.472       |
| C(2)-N(3)       | 1.369      | 1.345       | C(1')-C(2')       | 1.537      | 1.542       |
| C(2)-N(1)       | 1.380      | 1.412       | C(1')-O(4')       | 1.420      | 1.432       |
| N(3)-C(4)       | 1.336      | 1.337       | C(2')-C(3')       | 1.535      | 1.529       |
| C(4)-N(4)       | 1.314      | 1.334       | C(3')-O(3')       | 1.421      | 1.442       |
| C(4)-C(5)       | 1.437      | 1.412       | C(3')-C(4')       | 1.510      | 1.515       |
| C(5)-C(6)       | 1.337      | 1.346       | C(4')-O(4')       | 1.460      | 1.456       |
| C(6)-N(1)       | 1.368      | 1.364       | C(4')-C(5')       | 1.506      | 1.519       |
|                 |            |             | C(5')-O(5')       | 1.421      | 1.429       |
|                 | Molecule I | Molecule II |                   | Molecule I | Molecule II |
| O(2)-C(2)-N(3)  | 122.1      | 122.8       | N(1)-C(1')-C(2')  | 112.6      | 114.5       |
| O(2)-C(2)-N(1)  | 118.3      | 118.8       | N(1)-C(1')-O(4')  | 107.6      | 108.0       |
| N(3)-C(2)-N(1)  | 119.6      | 118.4       | C(2')-C(1')-O(4') | 106.2      | 104.8       |
| C(2)-N(3)-C(4)  | 120.0      | 120.6       | C(1')-C(2')-C(3') | 102.0      | 101.0       |
| N(3)-C(4)-N(4)  | 118.5      | 117.5       | C(2')-C(3')-O(3') | 114.3      | 110.1       |
| N(3)-C(4)-C(5)  | 121.0      | 121.6       | C(2')-C(3')-C(4') | 103.1      | 103.0       |
| N(4)-C(4)-C(5)  | 120.4      | 120.9       | O(3')-C(3')-C(4') | 110.1      | 107.6       |
| C(4)-C(5)-C(6)  | 117.6      | 118.3       | C(3')-C(4')-O(4') | 103.7      | 105.8       |
| C(5)-C(6)-N(1)  | 121.5      | 120.0       | C(3')-C(4')-C(5') | 116.9      | 116.0       |
| C(2)-N(1)-C(6)  | 120.3      | 120.7       | O(4')-C(4')-C(5') | 109.9      | 110.3       |
| C(2)-N(1)-C(1') | 118.1      | 118.3       | C(1')-O(4')-C(4') | 111.4      | 110.2       |
| C(6)-N(1)-C(1') | 121.6      | 120.9       | C(4')-C(5')-O(5') | 114.5      | 113.1       |

Table 4. Deviations of atoms from the least-squares planes of the pyrimidine ring atoms

The atoms marked with an asterisk here and in Table 5 were used to calculate the mean planes.

|       | Molecule I | Molecule II |
|-------|------------|-------------|
| O(2)  | -0.051 Å   | -0.013 Å    |
| C(2)* | -0.018     | -0.014      |
| N(3)* | 0.000      | -0.012      |
| C(4)* | +0.016     | +0.020      |
| N(4)  | +0.105     | +0.045      |
| C(5)* | -0.016     | -0.002      |
| C(6)* | -0.002     | -0.024      |
| N(1)* | +0.019     | +0.032      |
| C(1)  | +0.158     | +0.100      |

lingam & Jensen, 1965; Bugg & Marsh, 1967) the cytosine bases are protonated. In cytosine itself (Barker & Marsh, 1964), cytosine monohydrate (Jeffrey & Kinoshita, 1963) and cytidine (Furberg, Petersen & Rømming, 1965) the cytosine bases are unprotonated. Both bases are also unprotonated in the present structure, and the bond lengths and angles are in good agreement with those in the other unprotonated cytosine structures. They differ significantly from a number of those in the protonated cytosine bases as has been discussed in detail by Viswamitra *et al.* (1971).

#### The sugars

The sugar ring in molecule I is puckered so that C(3') is *endo*, *i.e.* lying out of the mean plane through the other sugar ring atoms and on the same side as C(5'). Described relative to the least-squares plane of the five-atom ring the pucker is C(3')-*endo*-C(2')-*exo* and can be represented as  ${}^3T_2$  where  $T$  represents the twist of the ring. In molecule II the sugar ring is puckered so that C(2') is *endo*, or C(2')-*endo*-C(3')-*exo* relative to the five-atom plane, and can be represented as  ${}^2T_3$ . These two types of pucker are those most commonly found in nucleosides and nucleotides although it is unusual to find the two types in the same crystal, and CDR is the first example among deoxynucleosides and nucleotides. There are two crystal structures containing ribose sugars that have similar differences in pucker, namely 5-iodouridine (Rahman & Wilson, 1970) and adenosine triphosphate (Kennard *et al.*, 1972). These observations show that the sugar pucker is affected by its environment in the crystal. Theoretical calculations show that the energy difference between the two conformations is relatively small (Saran, Perahia & Pullman, 1973).

Table 5. Deviations of atoms from the least-squares plane of the five-atom sugar rings

|        | Molecule I | Molecule II |
|--------|------------|-------------|
| C(1')* | -0.093 Å   | -0.176 Å    |
| C(2')* | +0.196     | +0.239      |
| C(3')* | -0.230     | -0.214      |
| O(3')  | +0.408     | +0.619      |
| C(4')* | +0.173     | +0.106      |
| O(4')* | -0.046     | +0.045      |
| C(5')  | -0.574     | +1.439      |
| O(5')  | -1.983     | +2.560      |

Table 6. The torsion angles around the sugar rings

|                         | Molecule I | Molecule II |
|-------------------------|------------|-------------|
| C(2')-C(1')-O(4')-C(4') | 3.2°       | 340.4°      |
| C(1')-O(4')-C(4')-C(3') | 19.7       | 355.3       |
| O(4')-C(4')-C(3')-C(2') | 325.7      | 27.2        |
| C(4')-C(3')-C(2')-C(1') | 35.8       | 322.4       |
| C(3')-C(2')-C(1')-O(4') | 335.6      | 35.4        |

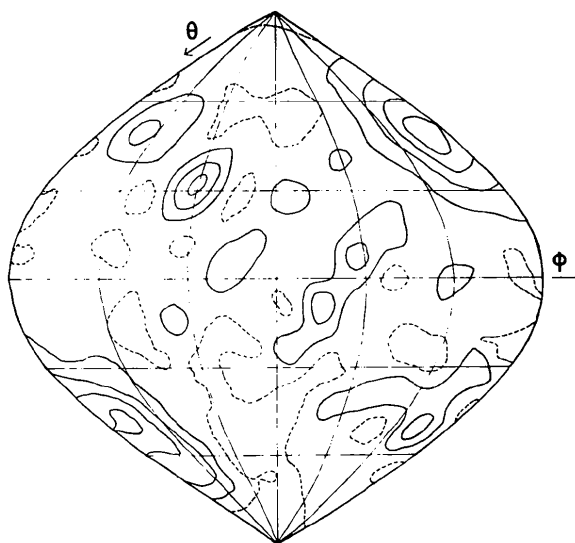
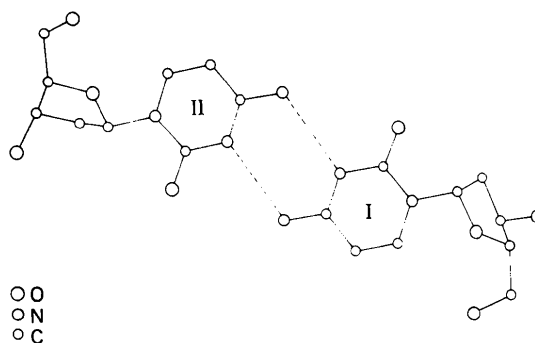
Fig. 2. The  $I(\theta, \phi)$  map. The ranges for both  $\theta$  and  $\phi$  are 0-180°.

Fig. 3. A view of the two molecules along the normal to the mean plane through the two pyrimidine rings.

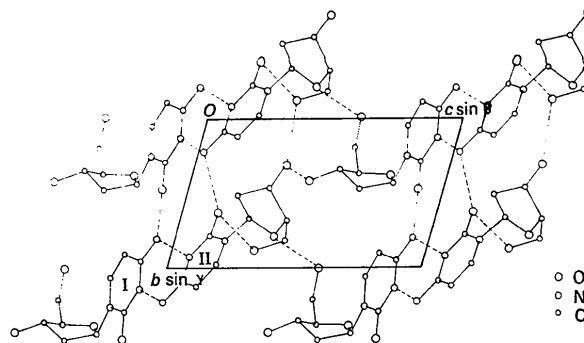
Fig. 4. A projection of the structure viewed along  $a$ . Hydrogen bonds are shown as broken lines.

Table 7. *A summary of the conformational parameters in CDR and related structures* $\chi_{CN}$  is defined for the sequence O(4') C(1') N(1) C(2),  $\psi$  for C(3') C(4') C(5') O(5').

| Compound                             | $\chi_{CN}$ | Sugar pucker | $\psi$ | C(5')-O(5')<br>conformation | Reference                                     |
|--------------------------------------|-------------|--------------|--------|-----------------------------|---|
| Cytidine                             | 197.2°      | ${}^3T_2$    | 47.1°  | <i>gg</i>                   | Furberg, Petersen & Rømming (1965)            |
| Cytidine 3'-phosphate (orthorhombic) | 226.1       | ${}^2T_3$    | 43.9   | <i>gg</i>                   | Sundaralingam & Jensen (1965)                 |
| Cytidine 3'-phosphate (monoclinic)   | 214.3       | ${}^2T_3$    | 45.5   | <i>gg</i>                   | Bugg & Marsh (1967)                           |
| Deoxycytidine hydrochloride          | 187.5       | ${}^3T_2$    | 46.2   | <i>gg</i>                   | Subramanian & Hunt (1970)                     |
| Deoxycytidine 5'-phosphate           | 182.1       | ${}^3T^4$    | 57.3   | <i>gg</i>                   | Viswamitra, Reddy, Lin & Sundaralingam (1971) |
| Deoxycytidine molecule I             | 201.2       | ${}^3T_2$    | 56.7   | <i>gg</i>                   | Present work                                  |
| Deoxycytidine molecule II            | 222.2       | ${}^2T_3$    | 62.5   | <i>gg</i>                   |   |

Table 8. *Hydrogen-bond lengths and angles*

| X-H...Y                |                   | X...Y  | X-H...Y |
|------------------------|-------------------|--------|---------|
| N(4)I—H(1)···N(3)II    | ( x, y, z)        | 2.97 Å | 143.1°  |
| N(4)II—H(1)···N(3)I    | ( x, y, z)        | 3.02   | 157.8   |
| N(4)I—H(2)···O(2)I     | ( x, -1+y, z)     | 3.03   | 162.9   |
| N(4)II—H(2)···O(2)II   | ( x, 1+y, z)      | 3.02   | 152.2   |
| O(3')I—H(9)···O(3')II  | (-1+x, 1+y, -1+z) | 2.90   | 148.3   |
| O(3')II—H(9)···O(5')II | ( x, -1+y, z)     | 2.80   | 140.6   |
| O(5')I—H(13)···O(4')II | (-1+x, y, -1+z)   | 2.93   | 164.2   |
| O(5')II—H(13)···O(2)II | ( 1+x, y, z)      | 2.80   | 164.3   |

The equations of the least-squares mean plane through the five-atom rings for molecules I and II respectively are:

$$-0.5425x + 0.8236y + 0.1657z - 6.2183 = 0$$

and

$$0.4111x + 0.8420y - 0.3494z - 4.5575 = 0.$$

Displacements of atoms from these planes are listed in Table 5 and the torsion angles in Table 6.

The conformation of C(5')-O(5') is *gauche* relative to both C(4')-O(4') and C(4')-C(3').

#### Conformations of the molecules

The dihedral angle between the base and the sugar ring is 72.5° in molecule I and 68.9° in molecule II. The torsion angle describing the rotation of the sugar ring about the glycosyl bond, O(4')-C(1')-N(1)-C(2), is 201.2° in molecule I and 222.2° in molecule II. These values are quite normal and fall in the range of values observed among nucleosides and nucleotides, where nucleosides with C(2')-*endo* puckering of the sugar have a larger torsion angle than those with the C(3')-*endo* pucker. Both molecules are in the *anti* conformation.

The conformational parameters of the two CDR molecules are listed in Table 7 together with those of related nucleosides and nucleotides.

#### Hydrogen bonds and molecular packing

The arrangement of the molecules in the crystal and the hydrogen-bonding scheme is shown in Fig. 4. The two molecules in the cell are hydrogen bonded together through N(4) and N(3). This is also illustrated in Fig. 3. There is an approximate twofold axis relating the two molecules, and there is an angle of 17° between the planes of the bases. A similar angle is found in cytosine (Barker & Marsh, 1964).

Hydrogen-bond lengths and angles are listed in Table 8. As in the other unprotonated cytosine structures the O(2) atoms are involved in hydrogen bonding. O(4') of molecule II participates in hydrogen bond formation but that of molecule I does not. All available hydrogen atoms participate in hydrogen bonding.

We acknowledge the assistance of the staff of the Computing Laboratory, University of Dundee. We thank Mr John Low for assistance with the data collection, Mr Mark Ritchie for technical assistance and Dr P. Tollin for discussions.

#### References

- BARKER, D. L. & MARSH, R. E. (1964). *Acta Cryst.* **17**, 1581-1587.
- BUGG, C. E. & MARSH, R. E. (1967). *J. Mol. Biol.* **25**, 67-82.
- FURBERG, S., PETERSEN, C. S. & RØMMING, C. (1965). *Acta Cryst.* **18**, 313-320.
- JEFFREY, G. A. & KINOSHITA, Y. (1963). *Acta Cryst.* **16**, 20-28.
- KENNARD, O., ISAACS, N. W., MOTHERWELL, W. D. S., COPPOLA, J. C., WAMPLER, D. L., LARSON, A. C. & WATSON, D. G. (1972). *Proc. Roy. Soc. A* **325**, 401-436.
- MUNNS, A. R. I. (1971). Ph.D. Thesis, Univ. of Dundee.
- RAHMAN, A. & WILSON, H. R. (1970). *Acta Cryst.* **B26**, 1765-1775.
- SARAN, A., PERAHIA, D. & PULLMAN, B. (1973). *Theoret. Chim. Acta*, **30**, 31-44.
- SUBRAMANIAN, E. & HUNT, D. J. (1970). *Acta Cryst.* **B26**, 303-311.
- SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 914-929.
- TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1322-1324.
- VISWAMITRA, M. A., REDDY, B. S., LIN, G. H. Y. & SUNDARALINGAM, M. (1971). *J. Amer. Chem. Soc.* **93**, 4565-4573.
- VOET, D. & RICH, A. (1970). *Prog. Nucl. Acid Res. Mol. Biol.* **10**, 183-265.