

## The Crystal and Molecular Structure of 2,2'-Anhydro-1- $\beta$ -D-arabino-furanosyl Uracil

BY L. T. J. DELBAERE AND M. N. G. JAMES

*Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2E1, Canada*

(Received 7 May 1973; accepted 3 July 1973)

The compound 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil crystallizes in space group  $P2_12_12_1$  with  $Z=8$ ; the unit-cell dimensions are  $a=13.687$  (2),  $b=18.241$  (3) and  $c=7.439$  (1) Å. Intensity data for reflexions with  $3^\circ \leq 2\theta \leq 60^\circ$  were collected on a four-circle diffractometer with Mo  $K\alpha$  radiation and a graphite monochromator. 2593 reflexions had intensities  $>3\sigma$ . The structure was solved by direct methods and refined (anisotropically) to a final  $R$  index of 0.033. The conformation of both crystallographically independent molecules is  $C(2')$ -*exo*,  $C(3')$ -*exo* with displacements from the  $C(1')$ -O(1')-C(4') plane of 0.366 (2) and 0.675 (2) Å, respectively, for one molecule, and 0.237 (2) and 0.688 (2) Å, respectively, for the other. The glycosidic torsion angles O(1')-C(1')-N(1)-C(6) are  $294.5$  and  $290.0^\circ$  for the molecules. The N(3) atom of the uracil ring of only one of the molecules accepts a hydrogen bond and in the final difference electron-density map there is a shift of electron density from the N(3)-C(2) and N(3)-C(4) bonds of this molecule, into the  $sp^2$  orbital of N(3), and is directed towards the hydrogen-bond donor.

In order to demonstrate a Karplus (1959) type relationship between  $^{13}\text{C}$ -vicinal- $^1\text{H}$  coupling constants and the corresponding torsional angles in solution n.m.r. spectroscopy, Lemieux, Nagabhushan & Paul (1972) examined the spectra of uridine and other related compounds that had the 2 position of the uracil ring enriched with  $^{13}\text{C}$ . They were able to show that an approximate Karplus relation indeed exists, but the torsional angles they used were only roughly determined from molecular models. Among the compounds that were examined by these workers were two that contained rigid fused polycyclic ring systems, 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil and 2,5'-anhydro-2',3'-isopropylidene cyclouridine. The rigidity of these two molecules due to the ring fusions permits a high degree of confidence in comparing the conformations in solution and in the crystalline state; these two molecules were therefore selected as candidates to place the empirical Karplus relationship on firm experimental foundations by obtaining the torsional angles as accurately as possible. The fact that the asymmetric unit of the crystals of 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil, the subject of the present paper, contains two crystallographically independent molecules provides a measure of the conformational flexibility (or lack thereof) of this molecule in differing molecular environments. A brief description of the torsional angles for the two aforementioned cyclic nucleosides has appeared (Delbaere, James & Lemieux, 1973).

### Experimental

Crystals of the title compound were kindly supplied by Professor R. U. Lemieux of the Chemistry Department, University of Alberta. The crystals were colourless prisms having the predominant forms  $\{110\}$  and  $\{001\}$ . The following crystal data were obtained: chemical formula,  $\text{C}_9\text{H}_{10}\text{N}_2\text{O}_5$ ; molecular weight, 226.2 dalton; systematic absences,  $h00$  ( $h=2n+1$ ),  $0k0$  ( $k=2n+1$ ),  $00l$  ( $l=2n+1$ ); space group,  $P2_12_12_1$ ; unit-cell dimen-

sions,  $a=13.687$  (2),  $b=18.241$  (3),  $c=7.439$  (1) Å,  $V=1857.2$  Å<sup>3</sup>; measured density (by flotation in  $\text{C}_2\text{H}_4\text{Br}_2$  and  $\text{C}_6\text{H}_5\text{Br}$ ),  $1.56$  g cm<sup>-3</sup>; calculated density,  $1.61$  g cm<sup>-3</sup>; unit-cell content  $Z=8$ , absorption coefficient,  $\mu(\text{Mo } K\alpha)=1.4$  cm<sup>-1</sup>;  $F(000)=944$  e.

A single crystal of 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil measuring  $0.28 \times 0.29 \times 0.30$  mm was used for the data collection on a Picker FACS-1 diffractometer. The radiation used was Mo  $K\alpha$  ( $\lambda=0.71069$ ) and the tube output was monochromated by a highly oriented graphite crystal. The intensity data were measured with a  $2\theta$  scan of  $1.5^\circ$  and a scan speed of  $2^\circ/\text{min}$ . Two 10 s stationary-crystal stationary-counter background measurements were made on either side of the peak position. The data collected were within the range  $3^\circ \leq 2\theta \leq 60^\circ$ . 2593 out of 3086 reflexions measured (84%) had net counts greater than  $3\sigma(I)$ . Observational weights for the reflexions were derived from the formula  $1/w=2F_o/[T+(0.04I)^2+t^2B]^{1/2}$  where  $I$  is the net peak count,  $T$  the total peak count,  $B$  the total background count, and  $t$  is the ratio of the peak-scan time to the total background counting time. No absorption corrections were made to the intensity data as the crystal was small and equidimensional.

### Structure determination and refinement

A Wilson (1942) plot was used to convert the observed structure amplitudes to  $E$ 's; the overall isotropic temperature factor is  $2.33$  Å<sup>2</sup>. Initial symbolic addition by hand (Karle & Karle, 1966) led to the following starting set of phases:

$h$	$k$	$l$	$E$	Phase	
12	0	1	4.5	$90^\circ$	} origin defining reflexions
13	2	0	3.5	90	
0	1	2	3.3	90	
8	13	0	3.5	0	— defines enantiomorph
10	0	6	2.6	180	— from $\Sigma_1$ -type indications
13	3	3	3.2	$a$	

Tangent-formula refinement on the phases of the 278 reflexions with  $E > 1.6$  for the four solutions with values of  $a = 0, 90, 180$  and  $270^\circ$  gave  $R_{\text{Karle}}$  values (Karle & Karle, 1966) of 0.32, 0.34, 0.16 and 0.33 respectively. An  $E$  map for the  $a = 180^\circ$  solution ( $R_{\text{Karle}} = 0.16$ ) contained all of the non-hydrogen atoms of both molecules in the asymmetric unit. An  $R$  index

$[R = (\sum |F_o| - |F_c|) / (\sum |F_o|)]$  of 0.326 was obtained for the structure-factor calculation based on these atoms. The coordinates of the atoms were changed from those of the  $E$  map to correspond to the correct expected enantiomorph (D-arabinose). Least-squares refinement was carried out in three stages. Firstly, three cycles of full-matrix (isotropic) refinement gave an  $R$  index of

Table 1. Atomic parameters and their estimated standard deviations (*e.s.d.*)

(a) Molecule A

(i) Non-hydrogen atoms (parameters  $\times 10^5$ ).

Temperature factors are of the form  $T = \exp -(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)$ .

	$x/a$	$y/b$	$z/c$	$\beta_{11}$	$\beta_{22}$	$\beta_{33}$	$\beta_{12}$	$\beta_{13}$	$\beta_{23}$
O(2)	27227 (9)	29000 (7)	-212 (20)	229 (6)	142 (3)	1567 (29)	6 (3)	21 (11)	101 (8)
O(4)	59398 (10)	31241 (9)	12315 (27)	269 (7)	304 (5)	2509 (42)	48 (5)	-159 (15)	-180 (13)
O(1')	18785 (9)	44472 (6)	14362 (19)	288 (6)	118 (3)	1483 (27)	9 (4)	144 (11)	-6 (8)
O(3')	1814 (9)	34545 (7)	1985 (19)	200 (5)	187 (4)	1385 (26)	-7 (4)	-133 (11)	29 (8)
O(5')	23059 (11)	40720 (8)	50805 (23)	429 (8)	237 (4)	1680 (34)	60 (5)	-338 (14)	-236 (10)
N(1)	33634 (10)	40072 (8)	1613 (23)	215 (7)	141 (4)	1281 (31)	-1 (4)	57 (12)	6 (9)
N(3)	43510 (11)	29554 (9)	5026 (24)	228 (7)	175 (4)	1391 (33)	32 (4)	23 (13)	-69 (10)
C(2)	35296 (12)	32797 (9)	1593 (26)	219 (7)	145 (4)	1070 (31)	6 (5)	80 (13)	-58 (10)
C(4)	51443 (13)	34069 (11)	9080 (29)	248 (8)	230 (6)	1288 (36)	14 (6)	12 (15)	-105 (12)
C(5)	49954 (14)	42022 (11)	9074 (33)	286 (9)	201 (5)	1719 (44)	-66 (6)	33 (17)	-92 (14)
C(6)	41111 (14)	44825 (10)	5590 (29)	294 (9)	149 (5)	1465 (38)	-37 (5)	68 (16)	-17 (12)
C(1')	23277 (12)	41817 (9)	-1221 (27)	220 (7)	143 (4)	1216 (34)	7 (5)	31 (14)	80 (11)
C(2')	19125 (12)	34087 (10)	-5060 (26)	233 (8)	173 (5)	1039 (30)	4 (5)	-37 (14)	-13 (10)
C(3')	11235 (12)	32958 (9)	9312 (25)	194 (7)	145 (4)	977 (30)	3 (5)	-91 (12)	37 (10)
C(4')	13869 (12)	38555 (9)	23822 (26)	189 (7)	137 (4)	1132 (33)	2 (4)	20 (13)	10 (10)
C(5')	20064 (14)	35237 (10)	38580 (27)	339 (9)	164 (5)	1078 (32)	7 (6)	-124 (16)	-45 (11)

(ii) Hydrogen atoms (atomic coordinates  $\times 10^3$ )

	$x/a$	$y/b$	$z/c$	$B_{\text{iso}}(\text{\AA}^2)$		$x/a$	$y/b$	$z/c$	$B_{\text{iso}}(\text{\AA}^2)$
H(O3')	-7 (1)	307 (1)	2 (3)	2.1	H(2')	166 (1)	334 (1)	-173 (3)	2.1
H(O5')	287 (2)	400 (1)	550 (3)	3.1	H(3')	114 (1)	281 (1)	134 (3)	1.7
H(5)	554 (1)	452 (1)	114 (3)	2.7	H(4')	81 (1)	405 (1)	294 (3)	1.8
H(6)	393 (1)	501 (1)	53 (3)	2.4	H(5')	259 (1)	327 (1)	335 (3)	2.3
H(1')	227 (1)	454 (1)	-93 (3)	2.1	H(5'')	163 (1)	316 (1)	445 (3)	2.3

(b) Molecule B

(i) Non-hydrogen atoms (parameters  $\times 10^5$ ).

Temperature factors are of the form  $T = \exp -(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)$ .

	$x/a$	$y/b$	$z/c$	$\beta_{11}$	$\beta_{22}$	$\beta_{33}$	$\beta_{12}$	$\beta_{13}$	$\beta_{23}$
O(2)	72876 (9)	32254 (7)	71693 (20)	246 (6)	167 (4)	1368 (26)	-9 (4)	-29 (11)	169 (8)
O(4)	40277 (11)	34070 (9)	60040 (30)	288 (7)	283 (5)	3160 (49)	-44 (5)	-237 (17)	198 (14)
O(1')	81275 (9)	45558 (7)	48095 (17)	335 (7)	145 (3)	845 (21)	8 (4)	47 (11)	38 (7)
O(3')	97878 (10)	39332 (9)	66789 (21)	276 (7)	351 (5)	1387 (30)	-82 (5)	-216 (12)	194 (11)
O(5')	77135 (10)	37817 (8)	15996 (19)	421 (8)	240 (4)	959 (26)	-65 (5)	-177 (12)	69 (9)
N(1)	66609 (10)	43018 (8)	64089 (22)	249 (7)	141 (4)	1093 (28)	-4 (4)	30 (12)	-2 (9)
N(3)	56429 (11)	32536 (9)	65686 (25)	264 (7)	179 (5)	1502 (34)	-20 (5)	-29 (13)	146 (11)
C(2)	64743 (13)	35830 (10)	66978 (25)	275 (8)	153 (5)	897 (30)	9 (5)	46 (14)	72 (10)
C(4)	48501 (14)	36852 (11)	60905 (31)	290 (9)	212 (5)	1382 (38)	16 (6)	-79 (17)	58 (12)
C(5)	50148 (14)	44606 (11)	57136 (31)	304 (10)	195 (5)	1529 (41)	67 (6)	-35 (16)	48 (13)
C(6)	59054 (14)	47510 (10)	58800 (28)	355 (10)	144 (4)	1210 (35)	41 (5)	47 (16)	16 (11)
C(1')	77003 (12)	44681 (9)	65286 (24)	268 (8)	146 (4)	809 (30)	-14 (5)	35 (13)	-47 (10)
C(2')	80936 (13)	37455 (10)	73240 (24)	275 (8)	194 (5)	733 (27)	-32 (5)	-40 (13)	49 (10)
C(3')	89410 (13)	35488 (10)	60901 (26)	224 (8)	188 (5)	1014 (31)	4 (5)	-79 (14)	61 (11)
C(4')	86421 (12)	38909 (9)	42999 (25)	219 (8)	167 (5)	843 (28)	14 (5)	17 (12)	10 (10)
C(5')	80447 (14)	33901 (10)	31244 (25)	384 (10)	174 (5)	813 (29)	15 (6)	-42 (15)	-34 (10)

(ii) Hydrogen atoms (atomic coordinates  $\times 10^3$ )

	$x/a$	$y/b$	$z/c$	$B_{\text{iso}}(\text{\AA}^2)$		$x/a$	$y/b$	$z/c$	$B_{\text{iso}}(\text{\AA}^2)$
H(O3')	997 (1)	374 (1)	784 (3)	3.0	H(2')	829 (1)	378 (1)	856 (3)	2.1
H(O5')	717 (1)	359 (1)	150 (3)	2.8	H(3')	904 (1)	304 (1)	604 (3)	2.1
H(5)	446 (1)	473 (1)	534 (3)	2.6	H(4')	919 (1)	406 (1)	361 (3)	1.9
H(6)	610 (1)	523 (1)	578 (3)	2.3	H(5')	750 (1)	319 (1)	375 (3)	2.3
H(1')	778 (1)	490 (1)	714 (4)	1.9	H(5'')	849 (1)	302 (1)	272 (3)	2.3

0.098. Then a difference electron-density map revealed the positions of the hydrogen atoms, the subsequent inclusion of which produced a residual  $R$  of 0.089. The hydrogen atoms were assigned the temperature factors of the atoms to which they were bonded. Block-diagonal least-squares refinement on all of the parameters (except the hydrogen-atom temperature factors) with anisotropic temperature factors for the non-hydrogen atoms converged to a final  $R$  index of 0.033 (weighted  $R=0.044$ ). Two intense low-angle reflexions (012 and 040) appeared to suffer from secondary extinction and were not included in the refinement. The maximum ratio of parameter shift to e.s.d. in the last cycle of refinement was 0.3 for the non-hydrogen atoms and 0.6 for the hydrogens. Observational weights were employed (see *Experimental* section) during the final 3 cycles of least-squares refinement. The scattering factors used for carbon, nitrogen and oxygen were

those of Cromer & Mann (1968) with the anomalous dispersion corrections of Cromer & Liberman (1970); the form-factor curve for hydrogen was that of Stewart, Davidson & Simpson (1965). Computations were carried out on an IBM 360/67 computer using the N.R.C. crystallographic programs (Ahmed, Hall, Pippy & Huber, 1966) and a locally modified version of *ORFLS* (Busing, Martin & Levy, 1962). The observed and final calculated structure amplitudes with phase angles have been filed with 'The Depository of Unpublished Data, National Science Library, National Research Council of Canada, Ottawa, Canada', from which a copy of these data may be obtained at a nominal cost.\* The final atomic parameters are listed in Table 1.

\* Copies of these data have also been deposited with the National Lending Library, England (Supplementary Publication No. SUP 30172). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13, White Friars CH1 1NZ, England.

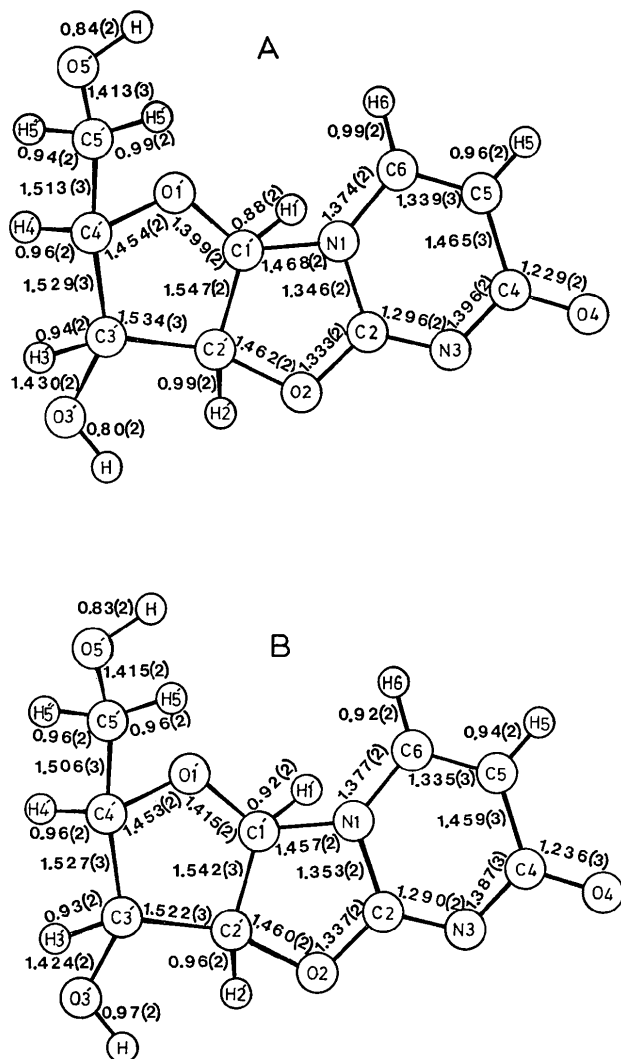


Fig. 1. Bond distances found for molecules *A* and *B* in the crystal structure of 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil.

Table 2. Least-squares planes

(a) Arabinose

	Distance ( $\text{\AA}$ ) from plane			
	Molecule <i>A</i>		Molecule <i>B</i>	
	I	II	III	IV
C(1')	0.0*	0.024 (2)*	0.0*	-0.049 (2)*
C(2')	-0.366 (2)	-0.023 (2)*	0.237 (2)	0.029 (2)*
C(3')	-0.675 (2)	0.014 (2)*	0.688 (2)	0.506 (2)
C(4')	0.0*	-0.423 (2)	0.0*	-0.031 (2)*
O(1')	0.0*	-0.016 (1)*	0.0*	0.051 (1)*
C(5')	1.388 (2)	-1.907 (2)	-1.375 (2)	-1.416 (2)
r.m.s. $\Delta$	-	0.020	-	0.041

Equations of the planes

	$\chi^2$
I. $0.8351X - 0.1822Y + 0.5190Z - 1.2237 = 0$	-
II. $-0.6914X + 0.4295Y - 0.5810Z - 1.1022 = 0$	579.0
III. $0.8340X + 0.3514Y + 0.4253Z - 13.7193 = 0$	-
IV. $0.8267X + 0.4374Y + 0.3539Z - 14.0458 = 0$	3188.7

(b) Uracil ring

	Distance ( $\text{\AA}$ ) from plane	
	Molecule <i>A</i>	Molecule <i>B</i>
	V	VI
N(1)	0.002 (2)*	0.011 (2)*
C(2)	0.003 (2)*	-0.004 (2)*
N(3)	-0.002 (2)*	-0.008 (2)*
C(4)	-0.003 (2)*	0.014 (2)*
C(5)	0.008 (2)*	-0.007 (2)*
C(6)	-0.007 (2)*	-0.006 (2)*
O(2)	-0.019 (2)	-0.009 (2)
O(4)	0.001 (2)	0.057 (2)
C(1')	-0.117 (2)	-0.105 (2)
C(2')	-0.033 (2)	0.092 (2)
H(5)	0.04 (2)	-0.04 (2)
H(6)	0.00 (2)	0.05 (2)
r.m.s. $\Delta$	0.005	0.009

Equations of the planes:

	$\chi^2$
V. $0.2374X + 0.0407Y - 0.9705Z - 1.2728 = 0$	28.7
VI. $-0.1861X + 0.2057Y + 0.9608Z - 4.4870 = 0$	128.3

\* Indicates atoms which are included in calculating least-squares plane.  $X$ ,  $Y$  and  $Z$  refer to the crystallographic coordinate system ( $\text{\AA}$  units).

### Results and discussion

The bond lengths and angles determined by this X-ray analysis are shown in Figs. 1 and 2 respectively. The chemically equivalent bond distances are markedly similar in both molecules with the largest variation, involving non-hydrogen atoms, occurring for the C(1')-O(1') bond; this distance is 1.399 (2) Å in molecule *A* and 1.415 (2) Å in molecule *B*. The chemically equivalent bond angles between non-hydrogen atoms in both independent molecules agree to within 2°, except for the 3.7° difference in the C(4')-C(3')-O(3') bond angle reflecting the adjustment of O(3') to different hydrogen-bonding schemes in the two molecules. The dimensions of the two molecules are similar; the conformations are also expected to be similar with the largest changes probably due to differences in the O(3') environment. Fig. 3 shows the torsion angles around the arabinose ring as Newman projections. The greatest variation in torsion angle, in comparing both molecules, occurs along the C(3')-C(2') bond with the largest deviation (14.2°), not unexpectedly, being in the O(3')-C(3')-C(2')-C(1') torsion angle. A stereo drawing of the two molecules in the asymmetric unit is depicted in Fig. 4. The relevant least-squares planes through the arabinose and uracil ring atoms are in Table 2. Table 3 contains the intermolecular non-bonded contacts.

#### Arabinose-ring conformation

Several dimensions of the arabinose rings in the present structure vary significantly from the average values (Arnott & Hukins, 1972) found in crystal

Table 3. Intermolecular non-bonded contacts

Molecule <i>A</i> ...Molecule <i>A</i>			Molecule <i>A</i> ...Molecule <i>B</i>		
C(3')	O(4')	3.059 (2) Å	O(2)	O(2 <sup>iv</sup> )	3.113 (2) Å
O(1')	O(5' <sup>iii</sup> )	3.092 (2)	C(2)	O(4')	3.174 (3)
C(2')	O(4')	3.143 (2)	C(5')	O(4)	3.201 (3)
O(2)	O(4')	3.165 (2)	O(2)	C(5' <sup>i</sup> )	3.229 (2)
C(1')	O(5' <sup>iii</sup> )	3.228 (2)	O(2)	O(5' <sup>i</sup> )	3.237 (2)
O(3')	O(4')	3.240 (2)	O(4)	C(5')	3.243 (2)
C(6)	O(5' <sup>iii</sup> )	3.292 (2)	C(5')	O(2' <sup>iv</sup> )	3.303 (2)
O(3')	C(4')	3.494 (2)	C(4)	N(3 <sup>v</sup> )	3.311 (3)
C(3')	N(3')	3.497 (2)	O(1')	C(5' <sup>iii</sup> )	3.313 (2)
H(6)	O(5' <sup>iii</sup> )	2.41 (2)	C(6' <sup>vi</sup> )	O(3')	3.364 (2)
			N(1)	O(4')	3.404 (3)
			C(3' <sup>iv</sup> )	N(3)	3.447 (2)
			O(2)	O(4')	3.460 (3)
			C(5' <sup>vi</sup> )	O(3')	3.462 (3)
			N(3)	N(3 <sup>v</sup> )	3.462 (3)
			N(3)	O(4')	3.475 (3)
			C(3' <sup>iv</sup> )	O(2)	3.498 (2)
			O(4)	N(3 <sup>v</sup> )	3.500 (3)
			O(3' <sup>ii</sup> )	C(3')	3.500 (2)
			H(O3' <sup>iii</sup> )	H(O3')	2.03 (3)
			H(4' <sup>vii</sup> )	H(4')	2.27 (3)

#### Symmetry codes

- |   |   |
|---|---|
| i. $-\frac{1}{2}+x, \frac{1}{2}-y, -z$    | v. $x, y, -1+z$                         |
| ii. $1+x, y, 1+z$                         | vi. $\frac{3}{2}-x, 1-y, \frac{1}{2}+z$ |
| iii. $\frac{1}{2}-x, 1-y, -\frac{1}{2}+z$ | vii. $1+x, y, z$                        |
| iv. $-\frac{1}{2}+x, \frac{1}{2}-y, 1-z$  |   |

structures containing ribofuranose rings. The C(1')-C(2') bond distances of 1.547 (2) and 1.542 (3) Å in molecules *A* and *B* respectively are longer than the average value of 1.525 Å, quoted by Arnott, and the C(1')-C(2')-C(3') angles of 104.6 (1)° in *A* and 103.7 (1)° in *B* are also larger than the average value of 101.4°. However, these values are similar to those of 1.550 (3) Å and 103.6 (2)° which were found for these parameters in the crystal structure of 2,5'-anhydro-2',3'-isopropylidene cyclouridine (Delbaere & James, 1973).

The conformation of the arabinose ring of each molecule (*A* and *B*) is C(2')-*exo*, C(3')-*exo* (Table 2) having planes I and III with displacements of atoms C(2') and C(3') from the C(4')-O(1')-C(1') plane of 0.366 (2) and 0.675 (2) Å respectively for molecule *A*, and 0.237 (2) and 0.688 (2) Å respectively for molecule *B*. The deoxyribose ring in deoxycytidine 5'-phosphate monohydrate (Viswamitra, Reddy, Lin & Sundaralin-

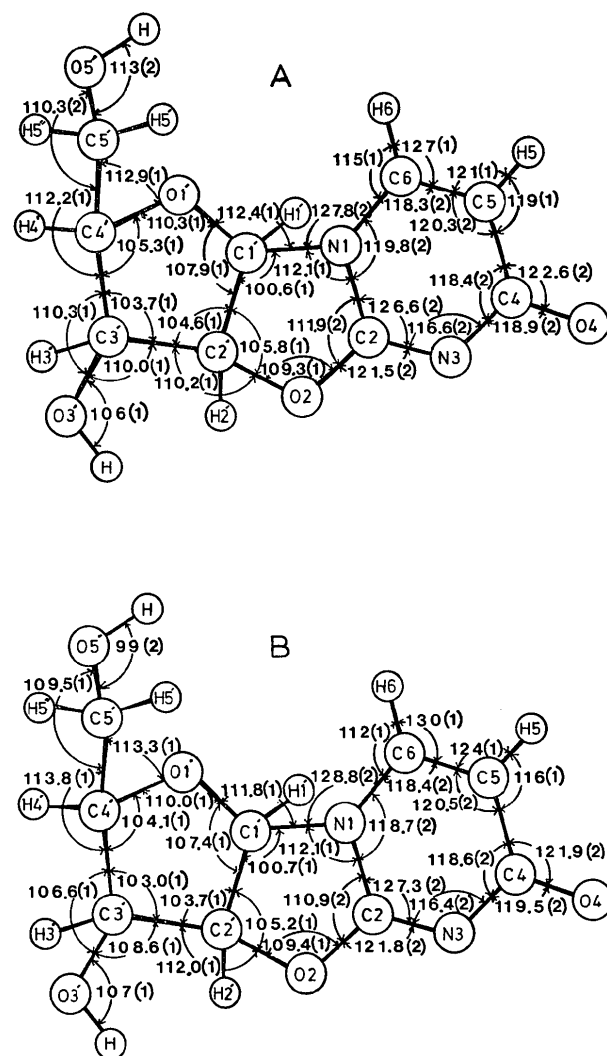


Fig. 2. Bond angles found for molecules *A* and *B* in the crystal structure of 2,2'-anhydro-1- $\beta$ -D-arabino-furanosyl uracil.

gam, 1971) has a similar ring puckering with displacements of 0.222 and 0.647 Å of C(2') and C(3') respectively from the C(4')-O(1')-C(1') plane.

The best four-atom planes of the arabinose rings are

C(1')-C(2')-C(3')-O(1') for molecule *A* and C(1')-C(2')-C(4')-O(1') for molecule *B* but neither of these are strictly planar. Note the large values of  $\chi^2$  for planes II and IV in Table 2. Although there are some varia-

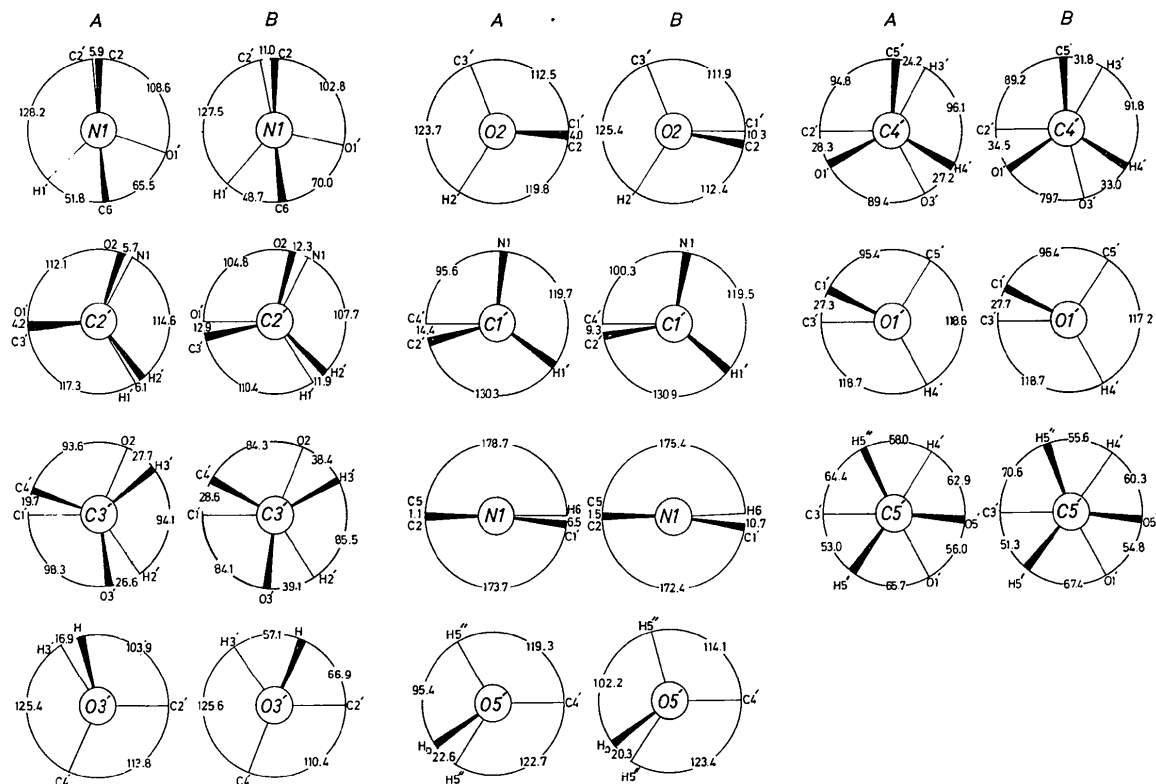


Fig. 3. Newman projections of the relevant torsion angles involving the arabinose-ring atoms and the uracil-ring atoms.

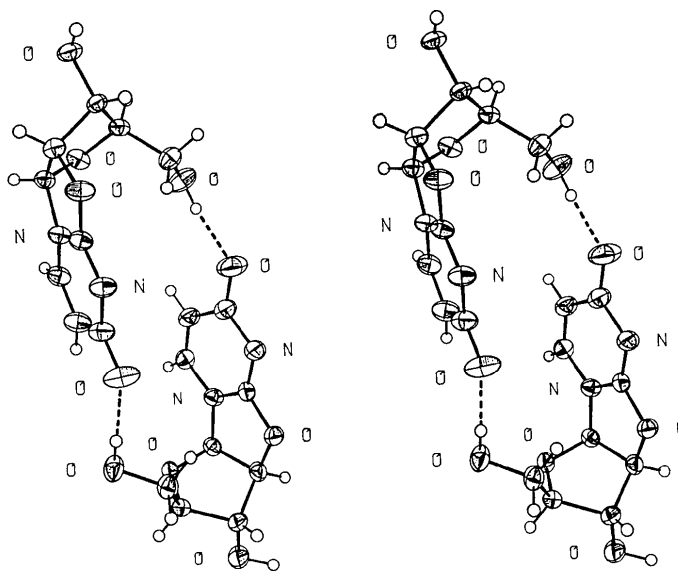


Fig. 4. Stereoscopic drawing of the two molecules in the asymmetric unit in the crystal structure of 2,2'-anhydro-1- $\beta$ -D-arabinofuranosyl uracil. The thermal ellipsoids include 50% probability and the hydrogen atoms were assigned 0.7 Å<sup>2</sup> temperature factors. The program ORTEP (Johnson, 1965) was used for drawing Figs. 4, 5 and 7.

tions in the torsion angles due to different molecular environments, the overall conformation is the same for both arabinose rings, thus confirming the initial postulate that the molecule is fairly rigid.

The conformation along the C(5')-C(4') bond is *gauche-trans* for *A* and *B* (Fig. 3), *i.e.*, the O(5')-C(5') bond is *gauche* to the C(4')-O(1') bond and *trans* to the C(4')-C(3') bond. The angles involved, as defined by Shefter & Trueblood (1965), are  $\varphi_{OO} = 56.0^\circ$  for *A* and  $54.8^\circ$  for *B* and  $\varphi_{OC} = 174.7^\circ$  for *A* and  $173.5^\circ$  for *B*. This conformation allows the formation of the two hydrogen bonds which link the two crystallographically independent molecules (see Fig. 4).

#### Arabinose-uracil linkages and the uracil ring

The glycosidic torsion angle,  $\chi_{CN}$  (Sundaralingam, 1969), O(1')-C(1')-N(1)-C(6) is  $294.5$  and  $290.0^\circ$  for molecules *A* and *B* respectively. These values differ from those usually occurring in nucleosides and nucleotides (Sundaralingam, 1969, and Arnott & Hukins, 1972). However, they agree well with the  $\chi_{CN} = 299^\circ$  obtained for 2,2'-anhydro-1- $\beta$ -D-*arabino*-furanosyl cytosine hydrochloride (Sundaralingam, 1973) which has a similar fused-ring system.

The glycosidic bond length, C(1')-N(1), is  $1.468$  (2) Å in *A* and  $1.457$  (2) Å in molecule *B* which may be compared to the average value of  $1.476$  Å for ribofuranose glycosidic linkages (Voet & Rich, 1970). The formation of the 2,2'-anhydro linkage of the base with the sugar ring, causes the uracil ring to shift from the *keto keto* to the *enol keto* tautomeric form. The C(2)-N(3) bond shortens by  $0.09$  Å and the C(2)-O(2) bond length increases by  $0.11$  Å from the average values obtained for these bonds in crystal structures containing uracil residues (Voet & Rich, 1970). There is considerable multiple-bond character in the C(2)-O(2) bond as seen in the lengths of  $1.333$  (2) Å in molecule *A* and  $1.337$  (2) Å in *B*, since they are intermediate in length between a single and a double bond; O(2) is within  $0.02$  Å of the least-squares plane (Table 2, plane V) through the uracil ring in molecule *A* and within  $0.01$  Å

of the similar plane (plane VI) in molecule *B* which allows overlap of the non-bonded electrons of the oxygen atom with the  $\pi$ -electron density of the ring. In 2'-deoxy-5-diazo-6(*H*)-O<sup>6</sup>,5'-cyclouridine (Abraham, Cochran & Rosenstein, 1971), the C(6)-O(5') bond is  $1.415$  Å, *i.e.*  $0.08$  Å longer, where C(6) has a tetrahedral configuration. The C(2')-O(2) bond,  $1.462$  (2) and  $1.460$  (2) Å in molecules *A* and *B*, is  $0.04$  Å longer than the  $1.417$  Å average for this bond in ribofuranose rings (Arnott & Hukins, 1972) but it is identical to the C(5')-O(2) distance of  $1.458$  (3) Å found in 2,5'-anhydro-2',3'-isopropylidene cyclouridine (Delbaere & James, 1973).

The formation of both the glycosidic bond, C(1')-N(1), and the 2,2'-anhydro linkage between the uracil ring and the arabinose moiety results in a five-membered ring, C(1')-N(1)-C(2)-O(2)-C(2'), being fused to the six-membered uracil ring. This causes all of the ring bond angles, of the atoms involved in this five-membered ring, to decrease from their average values (Arnott & Hukins, 1972; Voet & Rich, 1970) found in previously determined structures containing ribofuranose and uracil residues as follows: angle N(1)-C(1')-C(2') decreases by  $13^\circ$ , N(1)-C(2)-O(2) by  $11^\circ$ , C(1')-C(2')-O(2) by  $7^\circ$  and C(1')-N(1)-C(2) by  $5^\circ$ . The other bond angles about N(1), C(2) and N(3) of the uracil ring adjust accordingly to maintain the near planarity of each of the uracil rings (Table 2, planes V and VI), *i.e.* angle C(1')-N(1)-C(6) increases by  $7^\circ$ , N(1)-C(2)-N(3) increases by  $11^\circ$  and C(2)-N(3)-C(4) decreases by  $10^\circ$ .

#### Hydrogen bonding

Two hydrogen bonds, each between atom O(5') of one molecule and atom O(4) of the base of the other molecule, form the dimer found in the asymmetric unit (Fig. 4). The distance between O(5') of molecule *A* and O(4) of molecule *B* is  $2.738$  (2) Å with an H(O5') $\cdots$ O(4) distance of  $1.96$  (2) Å and an angle O(5')-H $\cdots$ O(4) of  $155$  (2)°. The dimensions of the second hydrogen bond are: molecule *B* O(5') $\cdots$ O(4) of molecule *A*,

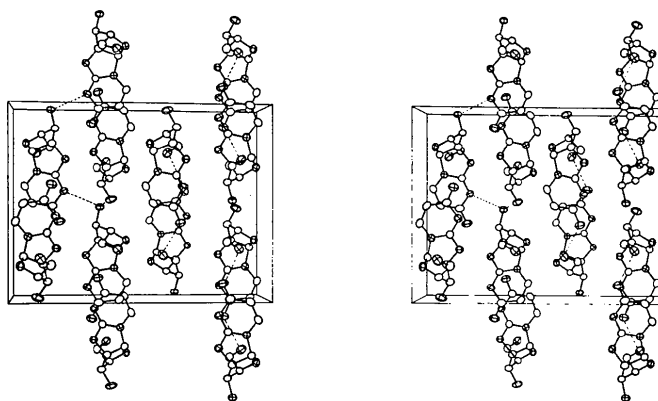


Fig. 5. Stereo packing diagram projected along the *c* axis of the unit cell. The origin of the unit cell is in the bottom right-hand corner with the *a* axis vertical and the *b* axis horizontal.

distance 2.722 (2) Å, H(O5')...O(4) distance 1.90 (2) Å with the angle O(5')-H...O(4) being 178 (2)°.

A hydrogen bond occurs between atom O(3') of molecule *A* and atom N(3) of a symmetry-equivalent

molecule *A*, N(3<sup>1</sup>), where the superscript refers to the symmetry code in Table 3. The O(3')...N(3<sup>1</sup>) distance is 2.860 (2) Å, H(O3')...N(3<sup>1</sup>) distance is 2.06 (2) Å and angle O(3')-H...N(3<sup>1</sup>) is 177 (2)°.

Atom O(3') of molecule *B* donates a hydrogen bond to atom O(3') of the symmetry equivalent of molecule *A* with symmetry code *ii* (Table 3). The dimensions of this hydrogen bond are molecule *B* O(3')...O(3'<sup>ii</sup>) of molecule *A*, distance 2.812 (2) Å, H(O3')...O(3'<sup>ii</sup>) distance 1.85 (2) Å with an angle O(3')-H...O(3'<sup>ii</sup>) of 172 (2)°. The hydrogen bonding in the unit cell may be seen in Fig. 5 which is a stereo view of a projection down the *c* axis of the unit cell.

Hanson, Sieker & Jensen (1972, 1973) have shown that in the accurate X-ray structure determination of sucrose, the difference Fourier map calculated from all of the data with the atomic parameters from the refinement on only the high-order data, had the same general features as the X-ray neutron difference map with the residual densities reduced and more diffuse. Their study also indicated that there were distortions in the electron density arising from the formation of a hydrogen bond.

A separate least-squares refinement of the structure 2,2'-anhydro-1-β-D-arabino-furanosyl uracil, in which only the reflexion data with  $\sin \theta/\lambda \geq 0.6$  were used in the refinement, produced a subsequent difference map which contained an enhancement of the bonding and lone-pair electron-density features of the final difference electron-density map; this suggests that although these latter features are not quantitatively determined, they are significant. Fig. 6 shows the sections of the final difference electron-density map through the atoms and neighbouring atoms of each of the two uracil rings (*A* and *B*); the skeletal outlines of these atoms are superimposed on the sections. The e.s.d. of the electron density in this map calculated by the formula of Cruickshank (1967) is 0.03 e Å<sup>-3</sup>. The lowest contour drawn is +0.09 e Å<sup>-3</sup> and the contour interval is 0.02 e Å<sup>-3</sup>. Negative contours are denoted as broken curves start-

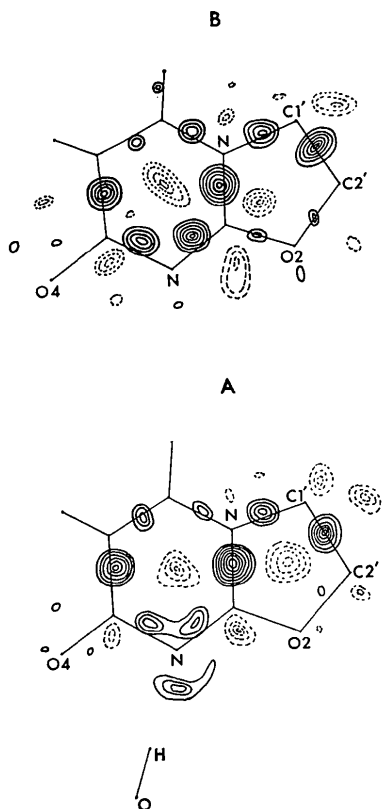


Fig. 6. Portions of the final difference map through the uracil ring and neighbouring atoms for molecules *B* and *A* respectively. The features shown are greater than three times the e.s.d. of the electron density (0.03 e Å<sup>-3</sup>). The first positive contour is at 0.09 e Å<sup>-3</sup> (first negative at -0.09 e Å<sup>-3</sup>) with a contour interval of 0.02 e Å<sup>-3</sup>. Broken curves denote negative contours.

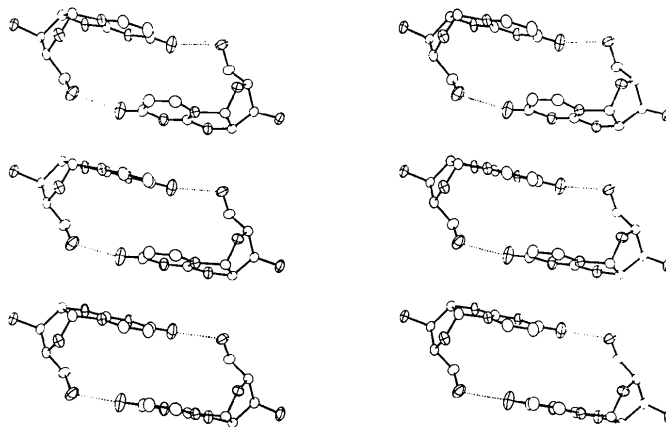


Fig. 7. Stereo view of a projection along the *b* axis of the unit cell; molecules *A* and *B* and the two pairs of equivalent molecules at (1+z) and (-1+z) are shown.

ing at  $-0.09 \text{ e } \text{\AA}^{-3}$  and the same contour interval. Bonding densities are apparent and are similar in *A* and *B* with significant exceptions. The striking feature when comparing the two sections is the difference between the bonding and lone-pair electron-density distribution around atom N(3). As described above, this atom in molecule *A* accepts a hydrogen bond from atom O(3') of a neighbouring molecule whereas that in *B* does not. As a result, electron density in the N(3)–C(4) and N(3)–C(2) bonds of molecule *A* is drawn into the  $sp^2$  orbital of N(3) directed towards the hydrogen atom of the hydrogen bond donor. This work thus provides evidence of a shift in bonding-electron density from a bonding orbital to a non-bonding orbital following the formation of a hydrogen bond.

#### *Intermolecular contacts and molecular packing*

The intermolecular non-bonded contacts, of 3.5 Å and less between non-hydrogen atoms, are listed in Table 3. Base–base interactions exist between molecule *A* and the molecule *B* one unit cell below (and between the base of molecule *B* and the base of the molecule *A* in the unit cell above), symmetry code *v*, Table 3. The angle between the least-squares planes of the uracil rings, planes V and VI, is  $14.5^\circ$ . The base–base interaction is demonstrated in Fig. 7 which is a stereo diagram of the projection down the *b* axis of the unit cell illustrating only molecules *A* and *B* and the two pairs of equivalent molecules above ( $1+z$ ) and below ( $-1+z$ ). The base–base contacts are similar to those occurring in many uracil derivatives (Bugg, Thomas, Sundaralingam & Rao, 1971) and involve the overlap of the electronegative carbonyl–oxygen atom O(4') of molecule *B* with the base of molecule *A*. The closest base–base contact is 3.174 (3) Å which occurs between C(2) of molecule *A* and O(4') of molecule *B*, Table 3.

The O(3')–H...N(3<sup>i</sup>) hydrogen-bond formation causes the closest non-bonding intermolecular contact, 3.059 (2) Å, between C(3') of molecule *A* and O(4<sup>i</sup>) of an equivalent molecule *A*, Table 3. As a consequence of the hydrogen-bonding network and the base–base interactions, there are much closer intermolecular contacts in the crystal structure of 2,2'-anhydro-1- $\beta$ -D-*arabino*-furanosyl uracil than there are in the crystal structure of 2,5'-anhydro-2',3'-isopropylidene cyclo-uridine (Delbaere & James, 1973) (closest intermolecular contact is 3.239 (3) Å) in which base–base interactions are present but no hydrogen bonding is possible. These closer contacts are reflected in the greater

density ( $1.61 \text{ g cm}^{-3}$ ) of the former compound in comparison to that ( $1.47 \text{ g cm}^{-3}$ ) of the latter.

The authors wish to thank Professor R. U. Lemieux for the gift of the crystals. This work was generously supported by the Medical Research Council of Canada grant MA-3406 to M.N.G.J.

#### References

- ABRAHAM, D. J., COCHRAN, T. G. & ROSENSTEIN, R. D. (1971). *J. Amer. Chem. Soc.* **93**, 6280–6281.
- AHMED, F. R., HALL, S. R., PIPPY, M. E. & HUBER, C. P. (1966). N.R.C. Crystallographic Programs for the IBM/360 System, *World List of Crystallographic Computer Programs*, 2nd ed., Appendix, p. 52.
- ARNOTT, S. & HUKINS, D. W. L. (1972). *Biochem. J.* **130**, 453–465.
- BUGG, C. E., THOMAS, J. M., SUNDARALINGAM, M. & RAO, S. T. (1971). *Biopolymers*, **10**, 175–219.
- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. Report ORNL-TM-305, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- CROMER, D. T. & LIBERMAN, D. (1970). *J. Chem. Phys.* **53**, 1891–1898.
- CROMER, D. T. & MANN, J. B. (1968). *Acta Cryst.* **A24**, 321–324.
- CRUICKSHANK, D. W. J. (1967). In *International Tables for X-ray Crystallography*, Vol. II, p. 331, edited by J. S. Kasper & K. Lonsdale. Birmingham: Kynoch Press.
- DELBAERE, L. T. J. & JAMES, M. N. G. (1973). Unpublished results.
- DELBAERE, L. T. J., JAMES, M. N. G. & LEMIEUX, R. U. (1973). *J. Amer. Chem. Soc.* In the press.
- HANSON, J. C., SIEKER, L. C. & JENSEN, L. H. (1972). *Trans. Amer. Cryst. Assoc.* **8**, 133–138.
- HANSON, J. C., SIEKER, L. C. & JENSEN, L. H. (1973). *Acta Cryst.* **B29**, 797–808.
- JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
- KARPLUS, M. (1959). *J. Chem. Phys.* **30**, 11–15.
- LEMIEUX, R. U., NAGABHUSHAN, T. L. & PAUL, B. (1972). *Canad. J. Chem.* **50**, 773–776.
- SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067–1077.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3177.
- SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.
- SUNDARALINGAM, M. (1973). Private communication.
- VISWAMITRA, M. A., REDDY, B. S., LIN, G. H. & SUNDARALINGAM, M. (1971). *J. Amer. Chem. Soc.* **93**, 4565–4573.
- VOET, D. & RICH, A. (1970). *Prog. Nucl. Res. Mol. Biol.* **10**, 183–265.
- WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151–152.