Acta Cryst. (1977). B33, 1349–1356

The Crystal and Molecular Structure of 2′-*O*-Methylcytidine

BY B. HINGERTY,* P. J. BOND† AND R. LANGRIDGE

Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, New Jersey 08540, *USA*

AND F. ROTTMAN

Biochemistry Department, Michigan State University, Michigan 48823, *USA*

(*Received* 8 *September* 1976; *accepted* 15 *October* 1976)

A conformation for 2′-*O*-methylcytidine ($C_{10}H_{15}N_3O_5$) has been established by X-ray diffraction analysis. The compound was synthesized by methylation of cytidine with diazomethane. It crystallizes in space group $P2_12_12_1$ with $a = 8{\cdot}742\ (5)$, $b = 8{\cdot}673\ (4)$, $c = 31{\cdot}537\ (15)$ Å, $Z = 8$. The structure was solved by direct methods and refined to an R of 6·9% for 2488 observed reflections. The two independent molecules have similar conformations with C(2′)*endo*–C(3′)*exo* puckered furanose rings oriented *anti* to their bases. The orientation of O(5′) about C(4′)–C(5′) is *gauche–gauche* in each molecule. The methyl group in molecule 1 is disordered with two possible staggers: α[C(1′)–C(2′)–O(2′)–CO(2′)] = 95 and 156°. In molecule 2 only the $\alpha = 168°$ stagger was observed. The structure has implications for ribonucleic acid conformations in the nucleus and cytoplasm since 2′-*O*-methylnucleotides are present, albeit to a minor extent, in tRNA's, mRNA's, rRNA's and HnRNA's.

Introduction

Modified nucleotides, methylated in the 2′-hydroxyl position of the β-D-ribose, occur in both nuclear and cytoplasmic ribonucleic acids (Barrel & Clark, 1974). *C*(2′)-*O*-Methylcytidilic acid is found at the 5′-end of the anticodon loop of several tRNA sequences (Yosiha, 1973; Dube, Marcker, Clark & Gory, 1968; Raj-Bhandary, Chang & Stuart, 1967; Dudock, Katy, Taylor & Holley, 1969; Hersh, 1970). There is evidence that 2′-*O*-methylation might be a factor in the cellular mechanism for processing ribonucleic acid molecules.

Novikoff hepatoma poly(*A*) containing mRNA molecules have one or two 2′-*O*-methylnucleotides linked by a 5′–5′ pyrophosphate bond to a 5′-terminal 7-methylguanosine (Desrosiers, Friderici & Rottman, 1975). The 2′-*O*-methylation can occur with any of the four common nucleotides and the structure for most eucaryotic mRNA 5′-termini may be represented as $m^7G^5'ppp5'(Nmp)_{1\text{ or }2}Np$. 2′-*O*-methylated constituents are present in the blocked 5′-terminal structures of heterogeneous nuclear RNA of mouse *L* cells (Perry, Kelley, Friderici & Rottman, 1975), several viral mRNA's (Wei & Moss, 1975; Furuichi, Morgan, Muthukrishnan & Shatkin, 1975; Moyer, Abraham, Adler & Banerjee, 1975) and low molecular weight nuclear RNA's (Reddy, Ro-Choi, Henning & Busch, 1974). However, Newcastle disease virus mRNA (Colonno & Stone, 1976) and the tobacco mosaic virus genome (Zimmern, 1975; Keith & Fraenkel-Conrat, 1975) lack 2′-*O*-methylnucleotides in their 5′-end sequences.

The resolution and quality of X-ray data from crystals of large nucleic acid molecules is not usually good enough for atoms to be identified in electron density maps. Thus, it is from high-resolution single-crystal analyses of their components, like 2′-*O*-methylcytidine, that conformation is revealed in detail and useful model-building information is gained.

Experimental

Cytidine (Miles Laboratories) was methylated by treatment with diazomethane and purified on Dowex 1(OH^-) columns (Rottman & Heinlein, 1968). Crystals were obtained after eluting the column with 30% aqueous methanol and concentrating the chromatographically homogeneous fraction to about 0·5 *M* at 37°C under a vacuum for 24 h. The large crystals of OMC which formed suggested tetragonal symmetry. They were square based pyramids, although the base was somewhat extended indicating the underlying orthorhombic symmetry.

* Author to whom correspondence should be addressed at MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

† Present address: Fritz-Haber-Institut der Max-Planck-Gesellschaft, Teilinstitut für Strukturforschung, 1 Berlin 33-Dahlem, Germany (BRD).

A crystal was mounted on a Picker four-circle X-ray diffractometer FACS-1 system with Zr-filtered Mo $K\alpha$ radiation ($\lambda = 0{\cdot}70926$ Å for α_1) and the data were collected to $2\theta < 60°$. 3986 reflections were collected from the unique portion of the reciprocal lattice. The reflections were measured with a $\theta/2\theta$ scan or three-circle setting. A 2θ range of 1·2° was used plus a dispersion correction at higher resolution to correct for the separation of the $K\alpha_1$ and $K\alpha_2$ lines. The scan speed was 1° min^{-1} with high and low background counts of 20 s each.

12 well-centered reflections were obtained and the cell constants and orientation matrix refined by least squares (Busing & Levy, 1967). In order to obtain more accurate cell dimensions only reflections with $2\theta > 33°$ were used. This enabled the centering to be done on $K\alpha_1$ only. The results of the refinement are given in Table 1. The Lorentz, polarization and background effects were corrected for. The standard deviation in the intensities was determined from the counting statistics with a 2% allowance for beam instability. All reflections for which $I/\sigma < 3$ were considered unobserved, so that only 2488 reflections were used.

Initial Cu $K\alpha$ precession photographs indicated either an orthorhombic or tetragonal space group. In order to resolve the ambiguity precession photographs were taken with Mo $K\alpha$ radiation. These showed that the correct symmetry was orthorhombic and the space group $P2_12_12_1$. The observed density from a flotation gradient was 1·450 g cm^{-3}. This indicated two molecules of OMC per asymmetric unit (as was expected from the pseudosymmetry), with water molecules absent. The actual structure determination showed no evidence of water molecules and the calculated density was 1·428 g cm^{-3}.

In order to collect a high-resolution data set, Mo $K\alpha$ radiation was used to measure reflections at $2\theta > 50°$. A crystal 0·35 × 0·40 × 0·30 mm was used to obtain greater intensity at the highest resolution. The linear absorption coefficient was so small (1·25 cm^{-1}) that this size was considered acceptable. Statistical tests were performed on the data to investigate the pseudosymmetry. Firstly, an R was calculated between hkl and khl reflections to determine the extent of the pseudosymmetry. This resulted in an R of 22% compared with about 5% for symmetrically equivalent reflections. Secondly, the statistical distribution of the normalized structure factors (or E's) showed no hypercentricity due to pseudosymmetry (Stout & Jensen, 1968).

Table 1. *Crystal data*

Space group $P2_12_12_1$	Formula $C_{10}H_{15}N_3O_5$
$a = 8{\cdot}742\,(5)$ Å	$Z = 8$
$b = 8{\cdot}673\,(4)$	$D_o = 1{\cdot}450$ g cm^{-3}
$c = 31{\cdot}537\,(15)$	$D_c = 1{\cdot}428$
Crystal size 0·35 × 0·40 × 0·30 mm	
$\mu = 1{\cdot}25$ cm^{-1}	

Solution of the structure

Direct methods were then applied with *MULTAN* (Germain, Main & Woolfson, 1970). The correct set of phases was obtained with 236 E's $> 1{\cdot}7$ by restricting the number of unique phase relations to 1650 with a probability of 0·75. The set with the highest figure of merit (1·06) lowest residual (20·0) and relative minimum in Ψ_0 was considered correct. An E map indicated all 36 non-hydrogen atoms. The two molecules are related by a set of pseudo equivalent points given in Table 2 in addition to the $P2_12_12_1$ symmetry. This involved a translation along x, y and z of $\frac{1}{4}$ and an approximate 90° rotation about z.

The structure was refined with the program *SFLS* (supplied by G. Reeke). Full-matrix isotropic refinement reduced R to 16%. A difference map was calculated to determine if any water molecules were present. No large peaks could be detected except one peak of approximately 1 e Å^{-3} fairly close to the methyl C of molecule 1.

The structure was next refined anisotropically and this reduced R to 12% by blocking the matrices into four groups each containing a ring of the structure. The position of the methyl C of molecule 1 was highly elongated in one direction and because a peak of 1 e Å^{-3} was still observed in the difference map this methyl group was considered to be disordered. To determine the H positions a difference map was calculated. All the geometrically fixed H atoms could be selected, reducing R to 9%. The H atoms not fixed by geometry were found in the next difference synthesis which reduced R to 6·9%.

Although the absorption coefficient was small it was decided to apply this correction, because the crystal (max 0·4 mm) was large. The program used was *ACACA* (Wuensch & Prewitt, 1965) to which a Gaussian integration table was appended with a grid of 4 × 4 × 4. A secondary extinction parameter was also included in the refinement with a value of $(1{\cdot}2 \pm 0{\cdot}2) \times 10^{-6}$ and yielded a small but observable improvement in the residuals.

Table 2. $P2_12_12_1$ *equivalent points, pseudo-equivalent points*

$P2_12_12_1$	(1)	x	y	z
	(2)	$\frac{1}{2}-x$	$-y$	$\frac{1}{2}+z$
	(3)	$\frac{1}{2}+x$	$\frac{1}{2}-y$	$-z$
	(4)	$-x$	$\frac{1}{2}+y$	$\frac{1}{2}-z$
Pseudo-equivalent points				
		x	y	z
		$\frac{1}{4}+y$	$\frac{1}{4}-x$	$\frac{1}{4}+z$

Table 3. *OMC coordinates and temperature factors*

The coordinates are fractional ($\times 10^4$; for H $\times 10^3$) and the temperature factors are in the form:

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

The b_{ij} are multiplied by 10^5 with the standard deviation corresponding to the last position. Isotropic temperature factors are given as $\exp[-B(\sin\theta/\lambda)^2]$. CO(2′)*A* and CO(2′)*B* have occupancies of 0·66 and 0·34 respectively.

Molecule OMC1

	x	*y*	*z*	b_{11}	b_{22}	b_{33}	b_{12}	b_{13}	b_{23}
N(4)	−491 (6)	−287 (6)	896 (1)	208 (9)	173 (8)	107 (5)	−77 (16)	−24 (3)	−19 (3)
C(4)	225 (5)	573 (5)	1201 (1)	125 (6)	120 (6)	91 (5)	−39 (11)	1 (3)	−9 (3)
N(3)	1079 (4)	−240 (4)	1477 (1)	141 (5)	101 (5)	94 (4)	1 (10)	−11 (2)	−5 (2)
C(2)	1890 (5)	564 (4)	1782 (1)	102 (5)	80 (5)	113 (5)	2 (10)	3 (2)	1 (2)
N(1)	1715 (4)	2148 (4)	1809 (1)	97 (4)	89 (4)	86 (4)	1 (8)	−11 (2)	7 (2)
C(6)	796 (5)	2923 (5)	1527 (1)	125 (6)	62 (6)	97 (5)	23 (11)	−9 (2)	0 (2)
C(5)	53 (5)	2166 (5)	1223 (1)	127 (6)	115 (7)	98 (5)	3 (12)	−15 (3)	8 (3)
O(2)	2715 (4)	−151 (3)	2035 (1)	156 (4)	115 (4)	148 (4)	18 (8)	−36 (2)	13 (2)
C(1′)	2426 (5)	2991 (4)	2159 (1)	115 (5)	77 (5)	70 (4)	−24 (10)	7 (2)	7 (2)
C(2′)	3726 (4)	4093 (5)	2040 (1)	84 (5)	108 (6)	74 (4)	0 (10)	−1 (2)	7 (2)
O(2′)	5139 (4)	3328 (5)	2016 (1)	118 (4)	296 (8)	148 (4)	165 (11)	−11 (2)	−22 (3)
CO(2′)*A*	5592 (10)	2821 (10)	1604 (3)	5·6 (0·3)					
CO(2′)*B*	6332 (21)	3676 (21)	1814 (5)	5·6 (0·5)					
C(3′)	3646 (5)	5247 (5)	2402 (1)	113 (6)	137 (7)	84 (4)	−108 (12)	1 (2)	12 (3)
O(3′)	4243 (5)	4580 (5)	2782 (1)	153 (6)	224 (7)	89 (4)	−136 (12)	−24 (2)	18 (2)
C(4′)	1921 (5)	5383 (5)	2472 (1)	105 (5)	135 (6)	78 (4)	−45 (12)	3 (2)	−10 (3)
C(5′)	1068 (6)	6668 (6)	2241 (2)	157 (8)	140 (8)	120 (6)	46 (15)	0 (3)	−17 (3)
O(5′)	1310 (4)	6582 (3)	1793 (1)	189 (5)	113 (4)	109 (4)	20 (9)	−7 (2)	2 (2)
O(1′)	1299 (3)	3909 (3)	2353 (1)	97 (3)	120 (4)	102 (3)	−36 (7)	12 (1)	−5 (2)

	x	*y*	*z*	*B*		*x*	*y*	*z*	*B*
H1N(4)	−119 (8)	21 (8)	84 (2)	9 (2)	HC(3′)	423 (5)	632 (4)	236 (1)	4 (1)
H2N(4)	−8 (5)	−168 (6)	87 (1)	6 (1)	HO(3′)	500 (6)	454 (8)	277 (2)	7 (2)
HC(6)	77 (4)	380 (4)	158 (1)	1 (1)	HC(4′)	171 (5)	540 (5)	280 (1)	4 (1)
HC(5)	−59 (5)	282 (4)	105 (1)	4 (1)	H1C(5′)	−13 (8)	660 (7)	231 (2)	11 (2)
HC(1′)	280 (3)	228 (3)	235 (1)	1 (1)	H2C(5′)	172 (5)	775 (5)	235 (1)	5 (1)
HC(2′)	349 (3)	462 (3)	178 (1)	1 (1)	HO(5′)	142 (7)	785 (7)	162 (2)	9 (2)

Molecule OMC2

	x	*y*	*z*	b_{11}	b_{22}	b_{33}	b_{12}	b_{13}	b_{23}
N(4)	2317 (7)	2600 (8)	3344 (2)	129 (7)	285 (11)	111 (5)	40 (16)	−8 (3)	50 (3)
C(4)	3116 (5)	2030 (5)	3664 (1)	93 (5)	168 (7)	86 (5)	52 (11)	−4 (2)	17 (3)
N(3)	2380 (3)	1182 (4)	3956 (1)	75 (4)	167 (6)	98 (4)	17 (10)	0 (2)	16 (2)
C(2)	3176 (4)	509 (5)	4278 (1)	66 (4)	133 (6)	100 (4)	20 (10)	8 (2)	7 (3)
N(1)	4756 (3)	783 (4)	4305 (1)	57 (3)	136 (5)	85 (3)	16 (8)	4 (2)	17 (2)
C(6)	5496 (5)	1662 (5)	4006 (1)	66 (5)	131 (6)	96 (5)	6 (10)	2 (2)	16 (3)
C(5)	4744 (5)	2292 (6)	3684 (1)	91 (5)	165 (8)	87 (5)	−9 (11)	10 (2)	28 (3)
O(2)	2551 (3)	−318 (4)	4545 (1)	84 (3)	189 (5)	131 (3)	−28 (8)	13 (2)	40 (2)
C(1′)	5620 (4)	143 (4)	4672 (1)	72 (4)	109 (6)	73 (4)	35 (10)	6 (2)	19 (2)
C(2′)	6666 (4)	−1208 (5)	4564 (1)	72 (4)	103 (5)	67 (4)	25 (9)	7 (2)	2 (2)
O(2′)	5795 (3)	−2564 (3)	4573 (1)	127 (4)	125 (4)	124 (4)	−48 (8)	25 (2)	−1 (2)
CO(2′)	6495 (8)	−3844 (8)	4403 (2)	224 (11)	179 (11)	191 (10)	−107 (21)	60 (5)	−5 (5)
C(3′)	7892 (5)	−1020 (5)	4907 (1)	91 (5)	141 (7)	83 (4)	80 (11)	5 (2)	3 (3)
O(3′)	7307 (4)	−1584 (4)	5306 (1)	175 (5)	190 (6)	71 (3)	129 (11)	11 (2)	26 (2)
C(4′)	8066 (5)	681 (5)	4947 (1)	100 (5)	160 (7)	75 (4)	40 (12)	−8 (2)	−14 (3)
C(5′)	9297 (5)	1481 (7)	4686 (2)	86 (5)	185 (9)	117 (6)	−86 (13)	−7 (3)	−25 (4)
O(5′)	9194 (3)	1075 (5)	4246 (1)	95 (3)	217 (6)	96 (3)	−30 (9)	3 (2)	−1 (2)
O(1′)	6579 (3)	1330 (3)	4823 (1)	85 (3)	138 (4)	102 (3)	84 (7)	−4 (1)	−19 (2)

	x	*y*	*z*	*B*		*x*	*y*	*z*	*B*
H1N(4)	287 (5)	327 (5)	315 (1)	4 (1)	H3*M*	687 (8)	−338 (9)	417 (2)	10 (2)
H2N(4)	165 (9)	276 (10)	334 (3)	11 (3)	HC(3′)	877 (4)	−164 (4)	485 (1)	2 (1)
HC(6)	655 (5)	179 (5)	403 (1)	4 (1)	HO(3′)	733 (7)	−247 (6)	532 (2)	8 (2)
HC(5)	520 (5)	296 (5)	350 (1)	4 (1)	HC(4′)	836 (4)	111 (5)	523 (1)	4 (1)
HC(1′)	495 (4)	−51 (4)	487 (1)	3 (1)	H1C(5′)	937 (5)	257 (5)	477 (1)	3 (1)
HC(2′)	720 (5)	−115 (5)	429 (1)	4 (1)	H2C(5′)	1029 (6)	160 (6)	482 (1)	6 (1)
H1*M*	561 (6)	−322 (6)	410 (2)	6 (1)	HO(5′)	955 (11)	30 (10)	435 (3)	16 (4)
H2*M*	566 (7)	−465 (7)	443 (2)	10 (2)					

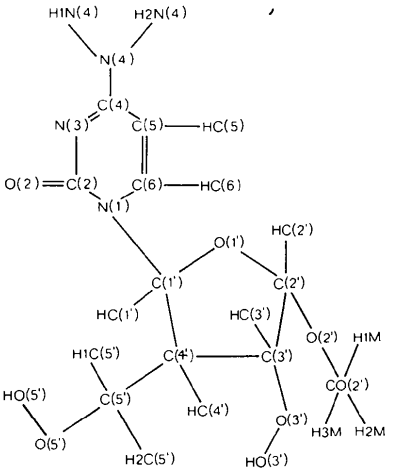

Fig. 1. Labeling nomenclature for OMC.

To resolve the disorder of the methyl group of molecule 1, a difference map was calculated after the entire group was deleted from the structure factor calculation. The map showed two peaks of 2 and 1 e $Å^{-3}$ separated by about 1·2 Å. This could be explained by allowing the methyl group to occupy two staggered positions, both permitted stereochemically. The previous average methyl-group position eclipsed the H of C(2′) in the ribose. When the disorder was included in the refinement the population factors for the two sites were constrained to be p and $1 - p$. The population distributions were found to be 66 and 34% and the two positions then became staggered and displayed reasonable geometry.

The final R for the observed reflections is 6·9%; for all the data it is 8·1%.* The final shifts were all less than the standard deviations. The coordinates are given in Table 3, while Fig. 1 explains the labeling system.

Results

(*a*) *OMC*1 *and OMC*2 *are similar but not identical*

Because of the presence of pseudotetragonal symmetry it was expected that the two molecules might be

* A list of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 32228 (39 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

Table 4. *Conformational-angle definitions for OMC*

Angle	Atoms
(Sundaralingam, 1973)	
χ	O(1′)–C(1′)–N(1′)–C(6)
ρ	C(5′)–C(4′)–C(3′)–O(3′)
ψ	O(5′)–C(5′)–C(4′)–C(3′)
τ (0)	C(4′)–O(1′)–C(1′)–C(2′)
τ (1)	O(1′)–C(1′)–C(2′)–C(3′)
τ (2)	C(1′)–C(2′)–C(3′)–C(4′)
τ (3)	C(2′)–C(3′)–C(4′)–O(1′)
τ (4)	C(3′)–C(4′)–O(1′)–C(1′)
α	C(1′)–C(2′)–O(2′)–CO(2′)
α^1	C(3′)–O(2′)–O(2′)–CO(2′)
(Arnott, 1970)	
χ	C(2′)–C(1′)–N(1)–C(2)
σ	C(5′)–C(4′)–C(3′)–O(3′)
ϵ	O(5′)–C(5′)–C(4′)–C(3′)
τ's	same as above

Table 5. *Observed conformational and sugar torsion angles for OMC*

Standard deviations are given in parentheses.

(*a*) Observed conformational angles

	χ (Arnott, 1970)	ρ, σ	ψ, ϵ	α	α^1
OMC1	112·9 (4)	151·5 (4)	55·9 (5)	95·0 (5)[A]	−151·6 (5)
				156·0 (10)[B]	−90·7 (11)
OMC2	107·9 (4)	151·1 (4)	51·6 (6)	167·9 (4)	−79·6 (5)

(*b*) Sugar torsion angles for OMC

	τ (0)	τ (1)	τ (2)	τ (3)	τ (4)	P
OMC1	−15·7 (4)	33·8 (4)	−37·8 (4)	29·9 (4)	−9·0 (4)	174·8
OMC2	−18·1 (4)	34·0 (3)	−37·1 (4)	27·7 (4)	−6·2 (4)	170·2

(*c*) Deviations in Å of C(2′), C(3′) and C(5′) from the plane defined by O(1′), C(4′) and C(1′)

	OMC1	OMC2
C(2′)	−0·404	−0·456
C(3′)	+0·227	+0·154
C(5′)	−1·253	−1·239

(*A*) 66% occupancy.
(*B*) 34% occupancy.

very similar. This turned out to be the case as can be seen from Tables 4 and 5 where the conformational angles of the two molecules are listed. Both bases are in the *anti* orientation and their χ angles are within the range for other structures possessing C(2′)*endo* sugar conformations (Sundaralingam, 1973). Both ribose moieties have the C(2′)*endo*–C(3′)*exo* twist conformation and a pseudorotation parameter P like that of other structures in the same class (Altona & Sundaralingam, 1972). The deviations of C(3′), C(2′) and C(5′) from the plane defined by C(1′)–O(1′)–C(4′) are given in Table 5. Both O(5′) atoms are *gauche–gauche* with respect to O(1′) and C(3′). Both methyl groups adopt staggered configurations. The $\alpha = 95$ and 156° conformers of the methyl group are observed with a statistical distribution in molecule 1; only the $\alpha = 168°$ conformer is observed in molecule 2.

(*b*) *Bond distances and angles*

A summary of the bond distances and angles is given in Figs. 2 and 3. None of the bond lengths deviate from standard values. For clarity, ribose H angles are not given in the figure but are listed in Table 6. (A number of H atomic and methyl-group positions involved anomalous bond lengths and angles. However, since disorder was present and since the H atoms were located in difference maps we considered these results to be presentable.)

A half-normal probability plot was computed (De Camp, 1973; Abrahams & Keve, 1971) to detect any significant conformational differences between the two molecules (Fig. 4*b*,*d*). With only non-hydrogen and non-methyl bond lengths, bond angles and conformational angles only small differences were noticeable. When the analysis was repeated with all interatomic distances less than 4·65 Å, the differences were more pronounced with a maximum ΔP (the number of standard deviations separating two observations) of 38·4 for the O(2)(cytosine) to O(2′)(ribose) distance. The large ΔP's were from O(2′), O(2) and to a lesser extent N(4), C(4) of the base. This can be traced to the angle χ which differs by 5° and significantly affects O(2′), O(2) interatomic distances. Other differences of 9·9, 8·4, 8·2 and 8·1 for C(2′)–O(1′), O(3′)–O(1′), C(1′)–O(2′), and O(2′)–C(3′) may be due to the 4·6° difference in pseudorotation parameter P. The differences from N(4) and C(4) may be due to slightly modified

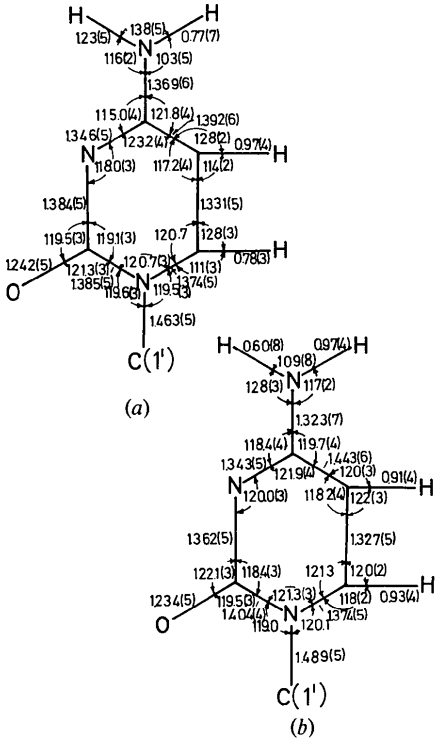

Fig. 2. Bond lengths and angles for cytosine moieties in (*a*) OMC1 and (*b*) OMC2 with standard deviations given in parentheses estimated from least squares.

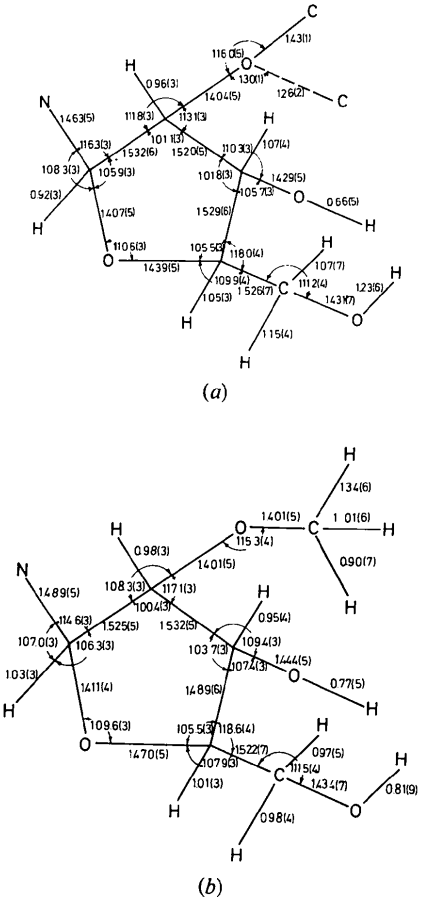

Fig. 3. Bond lengths and angles for ribose moieties in (*a*) OMC1 and (*b*) OMC2 with standard deviations given in parentheses estimated from least squares.

Table 6. *Hydrogen ribose bond angles*

Standard deviations are given in parentheses.

	OMC1	OMC2
HC(1′)–C(1′)–N(1)	108 (2)	113 (2)
HC(1′)–C(1′)–O(1′)	110 (2)	122 (2)
HC(1′)–C(1′)–C(2′)	108 (2)	93 (2)
HC(2′)–C(2′)–C(3′)	109 (2)	106 (2)
HC(2′)–C(2′)–C(1′)	110 (2)	116 (2)
HC(2′)–C(2′)–O(2′)	112 (2)	109 (2)
HC(3′)–C(3′)–C(2′)	117 (2)	112 (2)
HC(3′)–C(3′)–O(3′)	106 (2)	105 (2)
HC(3′)–C(3′)–C(4′)	115 (2)	119 (2)
HO(3′)–O(3′)–C(3′)	110 (5)	112 (4)
H1C(5′)–C(5′)–H2C(5′)	118 (4)	74 (4)
H1C(5′)–C(5′)–O(5′)	110 (3)	120 (2)
H1C(5′)–C(5′)–O(4′)	110 (3)	110 (2)
H2C(5′)–C(5′)–C(4′)	102 (2)	116 (3)
H2C(5′)–C(5′)–O(5′)	105 (2)	121 (3)
HO(5′)–O(5′)–C(5′)	114 (3)	78 (7)
HC(4′)–C(4′)–O(1′)	102 (2)	109 (2)
HC(4′)–C(4′)–C(3′)	108 (2)	118 (2)
HC(4′)–C(4′)–C(5′)	112 (2)	98 (2)
H1*M*–CO(2′)–H2*M*	*	86 (4)
H1*M*–CO(2′)–H3*M*		57 (5)
H2*M*–CO(2′)–H3*M*		130 (6)
H1*M*–CO(2′)–O(2′)		72 (2)
H2*M*–CO(2′)–O(2′)		102 (4)
H3*M*–CO(2′)–O(2′)		97 (5)

* OMC1 does not have values for these hydrogen atoms since the electron density was too difficult to interpret because of disorder.

Table 7. *Close interatomic contacts*

Position: $P2_12_12_1$ equivalent points from Table 2 with unit-cell translations along each axis as indicated. Molecule: refers to contact molecule as OMC1 or OMC2. Angle: apex angle of hydrogen bond (A–H $\cdots$ B).

	Contact	Position	Molecule	Distance (Å)	Angle (°)
(*a*) OMC1 close interatomic contacts					
N(4)–H	O(3′)	(2) 0,0,–1	2	2·94	143
N(4)–H	O(5′)	(4) 1,–1,0	2	3·38	172
N(3)	H–O(5′)	(1) 0,–1,0	1	2·94	162
O(2)	H–O(3′)	(4) 1,–1,0	1	2·73	157
O(3′)	H–N(4)	(1) 0,0,0	2	2·99	173
O(5′)	H–N(4)	(4) 0,0,0	2	3·32	144
O(5′)–H	N(3)	(1) 0,1,0	1	2·94	162
O(3′)	O(2)	(4) 1,0,0	1	2·73	157
O(5′) (intra)	H–C(6)	(1) 0,0,0	1	3·31	167
(*b*) OMC2 close intermolecular contacts					
N(4)–H	O(3′)	(1) 0,0,0	1	2·99	173
N(4)–H	O(5′)	(4) 0,–1,0	1	3·32	144
O(2)	H–O(5′)	(1) –1,0,0	2	3·31	128
O(2)	H–O(3′)	(3) –1,–1,1	2	2·74	170
O(3′)	H–N(4)	(2) 0,0,0	1	2·94	143
O(5′)	H–N(4)	(4) 1,0,0	1	3·38	172
O(5′)–H	O(2)	(1) 1,0,0	2	3·31	128
O(3′)–H	O(2)	(3) 0,–1,1	2	2·74	170
O(5′) (intra)	H–C(6)	(1) 0,0,0	2	3·36	156

intermolecular hydrogen bonding. For molecule 1 the N(4)–H $\cdots$ O(3′) angle is 143 and N(4)–H $\cdots$ O(5′) 172°, while molecule 2 is reversed geometrically with N(4)–H $\cdots$ O(3′) 173 and N(4)–H $\cdots$ O(5′) 144°. To estimate the reliability of the standard deviations a final analysis was done with only bond lengths and angles and O(2′), O(2), N(4) and C(4) removed. The slope of this curve was 2·3 and maximum ΔP was 4·7, indicating that the standard deviations are underestimated by a factor of two.

(*c*) *All intermolecular hydrogen bonds are between bases and sugars*

One of the striking features of the crystal structure is the absence of any base–base hydrogen bonds. A list of hydrogen bonds is given in Table 7. All hydrogen donors and acceptors, except O(1′) and O(2′), are involved in hydrogen bonds. We speculate that O(2′) is not involved since it is blocked by the presence of the methyl group and this may be significant in native RNA interactions. The packing (Fig. 4*a*,*c*) indicates that each molecule shares eight intermolecular hydrogen bonds or four unique hydrogen bonds per molecule. Although some of these bonds are quite long, the observation of H atoms adds to their credibility.

(*d*) *Intramolecular hydrogen bonding*

The only intramolecular hydrogen bond that may be present is between C(6)–H and O(5′). Although this bond is quite long (3·3 Å) it is nearly linear. A close van der Waals contact between C(6)–H and O(1′) can be detected in both molecules but the angle C(6)–H $\cdots$ O(1′) is only 100°.

(*e*) *Little base stacking*

There is little perceptible base stacking at about the 5 Å level (Fig. 4*a*,*c*). The molecules are involved in so many base-sugar interactions that base stacking does not appear to be energetically favored.

(*f*) *The base planes of the asymmetric unit make an angle of* 68°

The best mean-square planes defined by the six atoms of the cytosine rings are given in Table 8 with the observed deviations from planarity. The amino group appears to deviate significantly from the plane of the ring. From the pseudosymmetry the two molecules are related by a translation of $\frac{1}{4}$ of the unit cell along each axis and a rotation of about 90° about **c**. This is only approximate and the true rotation of the base planes is 68°. From this relation the bases are approximately 20° from being perpendicular to each other but they are physically separated by ribose moieties.

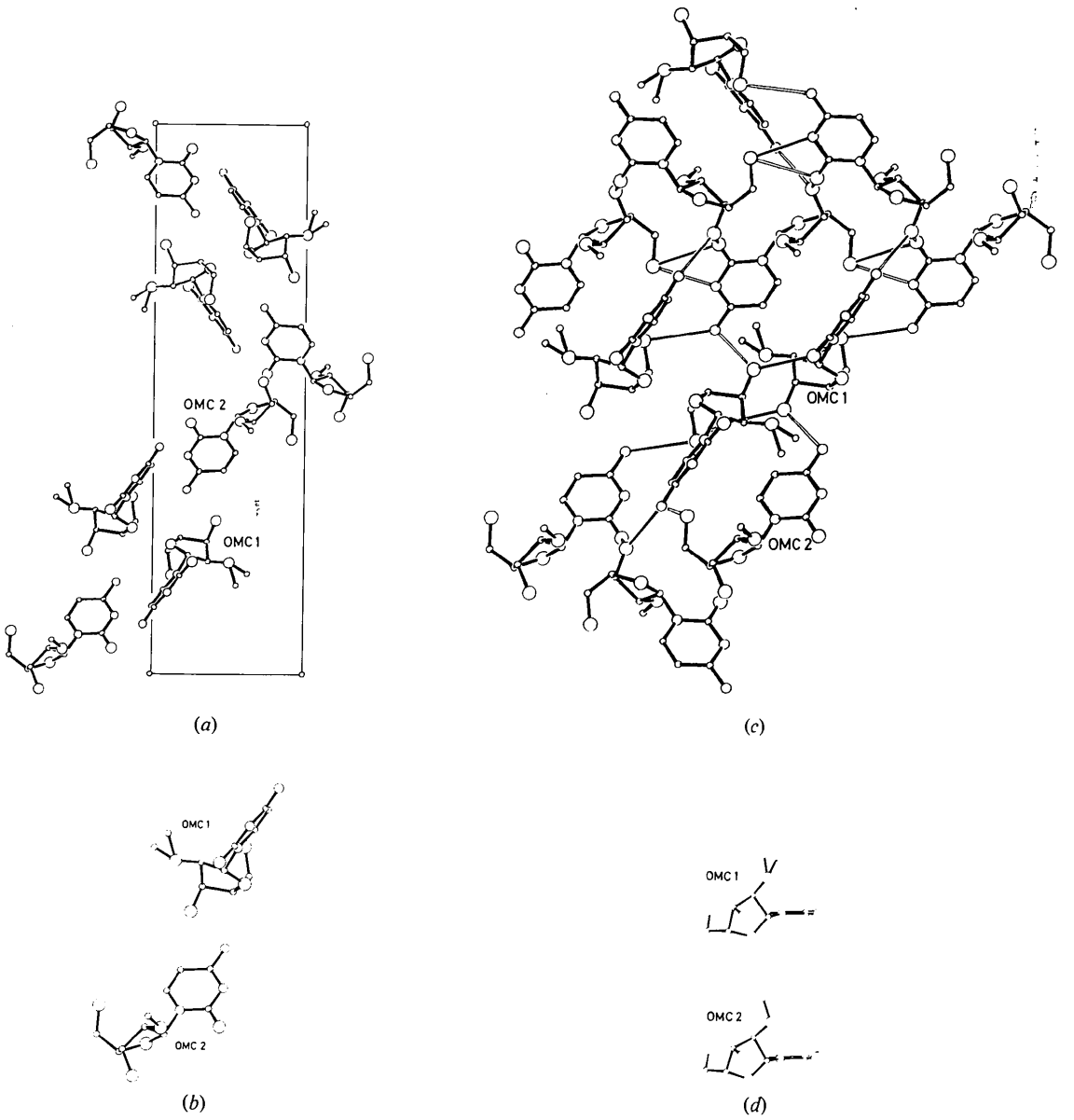

Fig. 4. (*a*) The eight molecules of the unit cell with the borders of the *a* and *c* axes shown. Some minor base parallelism can be seen but is at over 5 Å. (*b*) Unique contents of the asymmetric unit with disorder of OMC1 methyl shown. Hydrogens are deleted for clarity. (*c*) Hydrogen-bonding scheme with all contacts less than 3·3 Å for O···N and O···O contacts only. (*d*) OMC1 and OMC2 shown in the same orientation indicating structural similarity.

Discussion

The crystal structure of cytidine (Furberg, Petersen & Rømming, 1965) revealed the molecule in a C(3′)*endo* puckered state while our study has shown 2′-*O*-methylcytidine to be in the C(2′)*endo* puckered state as is 3′-*O*-methylarabinofuranosylcytosine (Birnbaum, Darzynkiewicz & Shugar, 1975). But the latter has O(5′) *gauche–trans* about C(4′)–C(5′) compared with *gauche–gauche* for this study. This and the analysis of 2′-*O*-methyladenosine (Sundaralingam & Prusiner, 1973) demonstrate that ribose methylation does not prevent the furanose ring from adopting any of the commonly observed sugar puckerings. Many stabilizing interactions observed in recent tRNA structural analyses (Ladner, Jack, Robertus, Brown, Rhodes, Clark & Klug, 1975) involved hydrogen bonds from 2′-hydroxyl groups. The donor potential of this group is removed upon methylation, which may thus impose significant tertiary structural limitations.

Table 8. *Least-squares planes of cytosine in crystal structure of OMC*

The best planes are determined by the six ring atoms of OMC [O(2), N(4), C(1') not included].

(*a*) OMC1 $0{\cdot}78001x' + 0{\cdot}11088y' - 0{\cdot}61586z' + 2{\cdot}13462 = 0$.

Atom deviations from plane in Å

N(4)	+0·031	C(2)	+0·016	C(5)	+0·004
C(4)	+0·011	N(1)	−0·002	O(2)	+0·020
N(3)	−0·020	C(6)	−0·008	C(1')	−0·118

(*b*) OMC2 $0{\cdot}15980x' - 0{\cdot}87171y' - 0{\cdot}55389z' + 7{\cdot}40373 = 0$.

Atom deviations from plane in Å

N(4)	+0·043	C(2)	+0·014	C(5)	+0·007
C(4)	+0·001	N(1)	−0·007	O(2)	+0·047
N(3)	−0·011	C(6)	−0·004	C(1')	−0·073

This research has been supported by USPHS, NIH grants RR-00578, GM-16539 and NSF grant GB-28021. Helpful advice was received from D. Appleman, W. Saenger and E. Subramanian. We thank the Max Planck Institut für Experimentelle Medizin in Göttingen, Germany (BRD) for computer time and technical assistance. We acknowledge the MRC Laboratory of Molecular Biology, Cambridge, England for secretarial and photographic help. BH would also like to thank the National Science Foundation for a NATO postdoctoral grant and the National Institute of Health for a National Cancer Institute grant IF22 CA02210-01.

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* A**27**, 157–165.

ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Amer. Chem. Soc.* **94**, 8205–8212.

ARNOTT, S. (1970). *Progr. Biophys. Mol. Biol.* **21**, 265–319.

BARREL, B. & CLARK, B. (1974). *Handbook of Nucleic Acid Sequences*. Oxford: Joyson-Bruvers.

BIRNBAUM, G. I., DARZYNKIEWICZ, E. & SHUGAR, D. (1975). *J. Amer. Chem. Soc.* **97**, 5904–5908.

BUSING, W. R. & LEVY, H. A. (1967). *Acta Cryst.* **22**, 457–464.

COLONNO, R. J. & STONE, H. (1967). *Nature, Lond.* **261**, 611–613.

DE CAMP, W. H. (1973). *Acta Cryst.* A**29**, 148–150.

DESROSIERS, R. C., FRIDERICI, K. H. & ROTTMAN, F. M. (1975). *Biochemistry*, **14**, 4367–4374.

DUBE, S. K., MARCKER, K. A., CLARK, B. F. C. & GORY, S. (1968). *Nature, Lond.* **218**, 232–233.

DUDOCK, B. S., KATY, G., TAYLOR, E. K. & HOLLEY, R. W. (1969). *Proc. Natl. Acad. Sci. US*, **62**, 941–945.

FURBERG, S., PETERSEN, C. S. & RØMMING, C. (1965). *Acta Cryst.* **18**, 313–320.

FURUICHI, Y., MORGAN, M., MUTHUKRISHNAN, S. & SHATKIN, A. J. (1975). *Proc. Natl. Acad. Sci. US*, **72**, 362–366.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* B**26**, 274–285.

HERSH, D. (1970). *Nature, Lond.* **228**, 57.

KEITH, J. & FRAENKEL-CONRAT, H. (1975). *FEBS Lett.* **57**, 31–33.

LADNER, J. E., JACK, A., ROBERTUS, J. D., BROWN, R. S., RHODES, D., CLARK, B. F. C. & KLUG, A. (1975). *Proc. Natl. Acad. Sci. US*, **72**, 4414–4418.

MOYER, S. A., ABRAHAM, G., ADLER, R. & BANERJEE, A. K. (1975). *Cell*, **5**, 59–67.

PERRY, R. P., KELLEY, D. E., FRIDERICI, K. H. & ROTTMAN, F. M. (1975). *Cell*, **6**, 13–19.

RAJBHANDARY, U. L., CHANG, S. H. & STUART, A. (1967). *Proc. Natl. Acad. Sci. US*, **57**, 751–758.

REDDY, R., RO-CHOI, T. S., HENNING, D. & BUSCH, H. (1974). *J. Biol. Chem.* **249**, 6486–6494.

ROTTMAN, F. & HEINLEIN, K. (1968). *Biochemistry*, **7**, 2634–2641.

STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination*, p. 321. New York: Macmillan.

SUNDARALINGAM, M. (1973). *Jerusalem Symposium on Quantum Chemistry and Biochemistry* V, Jerusalem. Edited by E. BERGMANN & B. PULLMAN, pp. 417–455. New York: Academic Press.

SUNDARALINGAM, M. & PRUSINER, P. (1973). *Amer. Cryst. Assoc. Abs.* Q3, p. 194.

WEI, C. M. & MOSS, B. (1975). *Proc. Natl. Acad. Sci. US*, **72**, 318–322.

WUENSCH, B. J. & PREWITT, C. T. (1965). *Z. Kristallogr.* **122**, 24–59.

YOSIHA, H. (1973). *Biochem. Biophys. Res. Commun.* **50**, 779–784.

ZIMMERN, D. (1975). *Nucleic Acids Res.* **2**, 1189–1201.