Crystal Structure of the Promutagen O^4-Methylthymidine: Importance of the Anti Conformation of the O(4) Methoxy Group and Possible Mispairing of O^4-Methylthymidine with Guanine†

Richard G. Brennan,‡,§ Danuta Pyżalska,‡ Wayne J. P. Blonski,‖ Frank E. Hruska,‖ and Muttaiya Sundaralingam*,‡

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin—Madison, Madison, Wisconsin 53706, and Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

Received July 9, 1985

ABSTRACT: O^4-Methylthymidine (O^4medT) is a promutagen. To correlate its biological properties to changes in the electronic, geometric, and conformational properties of the pyrimidine base resulting from the keto to enol shift arising from methylation, an X-ray study of O^4medT was undertaken. The crystal data are a = 4.950 (2) Å, b = 12.648 (1) Å, c = 19.305 (2) Å, space group $P2_12_12_1$, Z = 4, and R = 0.042. The D-deoxyribofuranosyl ring is puckered in the uncommon ${}_1T^2$ twist conformation with the phase angle of pseudorotation P = 133.8 (5)°. The amplitude of puckering τ_m = 31.4 (3)° shows that the ring is considerably flattened. The base is in the anti conformation [χ_{CN} = 40.6 (4)°], and the exocyclic C(4′)–C(5′) bond (Ψ) is gauche$^+$ [46.2 (5)°]. Methylation produces cytosine-like conjugation for the thymine base. The methoxy group takes the syn-periplanar conformation. Two types of mispairings with guanine are possible, and both require the anti conformation for the O(4) methoxy group. Semiempirical energy calculations have been carried out and reveal that the anti conformation can be energetically assumed in the double helix by widening the exocyclic angles C(5)–C(4)–O(4) and C(4)–C(5)–C(7) and the angle C(4)–O(4)–C(8) at the methoxy group. Such coordinated expansion relieves unfavorable interactions between the C(7) and C(8) methyl groups.

Lawley et al. (1973) showed that the carcinogenic agent *N*-methyl-*N*-nitrosourea can methylate the O(4) atom of thymidine. Subsequent studies revealed the promutagenic nature of this alkylated base. In particular, template activity experiments have shown that *Escherichia coli* DNA polymerase I was able to incorporate deoxyguanosine 5′-monophosphate (dGMP) into the complementary strands of alkylated poly(dA-dT) templates, which contained varying amounts of O^4-methylthymidine (O^4medT) (Abbott & Saffhill, 1977; Singer et al., 1983). A one to one correlation was found between the percent of dGMP incorporation and the percent of O^4medT in the template (Abbott & Saffhill, 1977). Additional studies demonstrated that O^4medT was able to be incorporated by *E. coli* DNA polymerase I and by rat spleen DNA polymerase α into strands complementary to templates of poly(dC-dG) but not poly(dA-dT) (Hall & Saffhill, 1983). It was also shown that O^4medT inhibited DNA synthesis only when it replaced deoxycytidine, which strongly implied that O^4medT bound to the polymerase/template complex as a deoxycytidine analogue.

In order to gain insight into the electronic and geometric perturbations resulting from O(4) methylation, that is, the keto to enol tautomeric shift, which might be the basis of alkylation-induced mutagenesis, we have carried out the X-ray structure determination of O^4medT. The structure and conformation of O^4medT are compared to the crystallographic results found for the O-alkylated pyrimidine O^4-methyluridine (Brenan et al., 1983). The base pairing between O^4medT and guanine is discussed in light of the crystal structure of O^4medT, semiempirical energy calculations, and relevant crystal and solution studies.

Experimental Procedures

Single crystals of O^4-methylthymidine ($C_{11}H_{16}N_2O_5$) were grown from an aqueous solution by slow evaporation. Oscillation and Weissenberg photographs revealed an orthorhombic lattice, and systematic absences indicated the space group $P2_12_12_1$. A crystal with dimensions 0.06 × 0.06 × 0.28 mm was chosen for data collection on an Enraf-Nonius CAD4 diffractometer using Ni-filtered Cu Kα radiation (λ = 1.5418 Å). The unit-cell parameters, refined by a least-squares algorithm using 25 centered reflections, were a = 4.950 (2) Å, b = 12.648 (1) Å, c = 19.305 (2) Å, V = 1208.7 (6) Å^3, and Z = 4. A total of 1242 independent reflections were collected to a 2θ maximum of 140° employing the ω–2θ scan mode. Of these reflections, 1087 had $I > 1.5\sigma(I)$ and were used in the structural analysis. An empirical ϕ curve correction for absorption and Lorentz and polarization corrections were applied to the intensities. Three standard reflections, monitored every 2 h of X-ray exposure, revealed negligible crystal decay; hence, no correction was applied.

The structure was solved by the direct methods program MULTAN (Main et al., 1974) employing 240 normalized structure factors (E's) with values ≥1.2. The E map calculated with the phase set, which had the best combined figure of merit, revealed all 18 non-hydrogen atoms. The initial R index was 0.184, and block-diagonal least-squares refinement of the non-hydrogen atomic positions and thermal parameters converged, for the anisotropic model, to an R index of 0.083. The atomic coordinates of the hydrogen atoms were obtained from two successive difference Fourier maps and were included in the final refinement with isotropic thermal parameters. The

† We thank the National Institutes of Health of the U.S. Public Health Service for Research Grant GM-17378 and the College of Agricultural and Life Sciences of the University of Wisconsin—Madison and NSERC of Canada (A6434) for their continued support.

‡ University of Wisconsin—Madison.

§ Present address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403-1229.

‖ University of Manitoba.

0006-2960/86/0425-1181$01.50/0

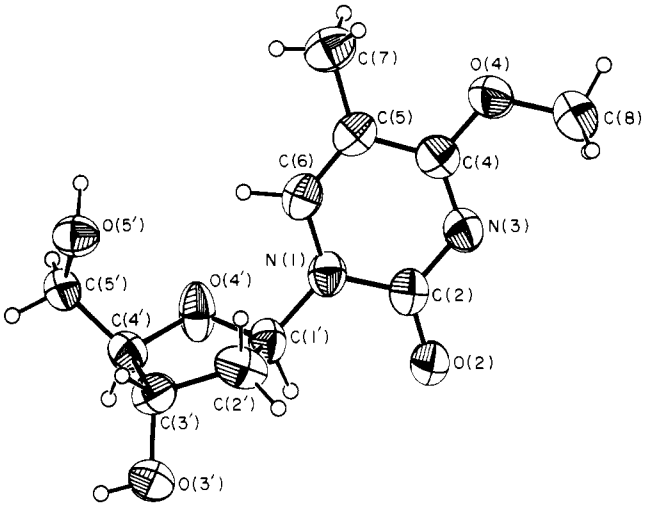

FIGURE 1: ORTEP drawing of O^4medT. The non-hydrogen atoms are represented by 50% probability ellipsoids while the hydrogen atoms are drawn as spheres of arbitrary size.

Table I: Atomic Parameters for O^4-Methylthymidine[a]

atom	x	y	z	B_{eq} or B
N(1)	−48 (6)	6722 (2)	2314 (1)	3.83 (7)
C(2)	−95 (8)	5982 (3)	1785 (2)	4.18 (9)
O(2)	−1594 (6)	5207 (2)	1832 (1)	5.09 (7)
N(3)	1553 (6)	6136 (2)	1233 (1)	4.31 (8)
C(4)	3155 (8)	6956 (3)	1218 (2)	3.98 (9)
O(4)	4825 (6)	7089 (2)	689 (1)	5.25 (7)
C(8)	4943 (13)	6282 (3)	169 (2)	7.31 (16)
C(5)	3284 (8)	7758 (2)	1740 (2)	3.85 (8)
C(7)	5136 (9)	8683 (3)	1687 (2)	5.14 (11)
C(6)	1626 (8)	7581 (3)	2281 (2)	4.00 (9)
C(1′)	−1714 (7)	6540 (3)	2938 (2)	3.94 (8)
C(2′)	−120 (8)	6122 (3)	3547 (2)	4.86 (10)
C(3′)	−1536 (8)	6583 (3)	4170 (2)	4.19 (9)
O(3′)	−3324 (7)	5811 (2)	4450 (1)	6.05 (9)
C(4′)	−3068 (8)	7533 (3)	3885 (2)	3.73 (8)
O(4′)	−2707 (6)	7525 (2)	3150 (1)	5.04 (7)
C(5′)	−2181 (8)	8593 (3)	4148 (2)	4.22 (9)
O(5′)	659 (5)	8717 (2)	4142 (1)	4.35 (6)
H(C1′)	−330 (8)	603 (2)	282 (2)	5.1 (9)
H1(C2′)	−17 (9)	538 (3)	350 (2)	6.0 (9)
H2(C2′)	173 (7)	643 (2)	353 (1)	3.7 (7)
H1(C3′)	−21 (7)	680 (2)	451 (2)	4.4 (8)
H(O3′)	−367 (9)	603 (2)	487 (2)	6.6 (11)
H(C4′)	−518 (8)	741 (2)	400 (1)	4.6 (8)
H1(C5′)	−309 (8)	919 (2)	390 (2)	5.2 (9)
H2(C5′)	−306 (8)	874 (3)	457 (2)	5.5 (9)
H(O5′)	107 (8)	929 (3)	384 (2)	7.0 (11)
H1(C8)	564 (9)	563 (3)	36 (2)	6.4 (10)
H2(C8)	612 (11)	655 (4)	−19 (2)	10.7 (14)
H3(C8)	295 (10)	608 (3)	−3 (2)	7.8 (12)
H1(C7)	472 (9)	912 (2)	129 (2)	5.7 (9)
H2(C7)	495 (8)	907 (2)	207 (2)	4.4 (8)
H3(C7)	703 (8)	843 (3)	164 (2)	5.9 (9)
H(C6)	137 (6)	806 (2)	267 (1)	3.0 (7)

[a] Fractional positional parameters are multiplied by 10^4 for non-hydrogen atoms and by 10^3 for hydrogen atoms. $B_{eq} = (4/3)\sum_i\sum_j B_{ij}a_ia_j$.

final R and R_w indices[1] for all observed reflections were 0.042 and 0.056, respectively. A modified counting statistics weighting scheme (Stout & Jensen, 1968) was used in which $w = 1/[\sigma^2F_0 + (0.01F_0)^2]$. Scattering factors for the oxygen, nitrogen, and carbon atoms were taken from Cromer and Waber (1965), and those for the hydrogen atoms were from Stewart et al. (1965).

Energy calculations using semiempirical potential functions were carried out to assess the effects of changing the exocyclic bond angles at C(4), C(5), and O(4) on the nonbonded and electrostatic interactions between the C(7) methyl and the C(8) methyl of the methoxy group. The total energy (V_{tot}) is expressed as the sum of the van der Waals interactions (V_{nb}), electrostatic interactions (V_{es}), and the torsional potential for rotation about a single bond (V_t). The energy expressions and the various constants used are identical with those used in calculations for O^4-methyluridine (Brennan et al., 1983).

[1] $R_w = [w(|F_o| - |F_c|)^2/\sum w|F_o|^2]^{1/2}$.

Table II: Bond Lengths (Å) and Angles (deg) for Non-Hydrogen Atoms

N(1)–C(2)	1.385 (4)	N(1)–C(2)–O(2)	119.5 (4)
C(2)–O(2)	1.233 (5)	N(1)–C(2)–N(3)	118.2 (4)
C(2)–N(3)	1.356 (5)	O(2)–C(2)–N(3)	122.3 (4)
N(3)–C(4)	1.306 (5)	C(2)–N(3)–C(4)	119.8 (4)
C(4)–O(4)	1.325 (5)	N(3)–C(4)–O(4)	119.8 (4)
O(4)–C(8)	1.433 (4)	N(3)–C(4)–C(5)	125.0 (4)
C(4)–C(5)	1.431 (5)	O(4)–C(4)–C(5)	115.2 (4)
C(5)–C(7)	1.490 (5)	C(4)–O(4)–C(8)	118.4 (4)
C(5)–C(6)	1.347 (6)	C(4)–C(5)–C(6)	113.6 (4)
C(6)–N(1)	1.368 (5)	C(4)–C(5)–C(7)	122.4 (4)
N(1)–C(1′)	1.478 (4)	C(6)–C(5)–C(7)	124.0 (4)
C(1′)–O(4′)	1.400 (5)	C(5)–C(6)–N(1)	122.5 (4)
C(1′)–C(2′)	1.511 (5)	C(6)–N(1)–C(2)	120.8 (4)
C(2′)–C(3′)	1.509 (6)	C(6)–N(1)–C(1′)	120.0 (3)
C(3′)–O(3′)	1.424 (5)	C(2)–N(1)–C(1′)	119.1 (3)
C(3′)–C(4′)	1.524 (5)	N(1)–C(1′)–O(4′)	107.2 (3)
C(4′)–O(4′)	1.430 (4)	N(1)–C(1′)–C(2′)	113.4 (4)
C(4′)–C(5′)	1.499 (5)	O(4′)–C(1′)–C(2′)	105.5 (4)
C(5′)–O(5′)	1.415 (5)	C(1′)–C(2′)–C(3′)	104.0 (4)
		C(2′)–C(3′)–O(3′)	109.0 (4)
		C(2′)–C(3′)–C(4′)	104.4 (4)
		O(3′)–C(3′)–C(4′)	111.6 (4)
		C(3′)–C(4′)–O(4′)	106.9 (4)
		C(3′)–C(4′)–C(5′)	115.9 (4)
		C(5′)–C(4′)–O(4′)	107.8 (4)
		C(4′)–O(4′)–C(1′)	109.9 (4)
		C(4′)–C(5′)–O(5′)	112.8 (4)

Table III: Selected Torsion Angles (deg)

C(4′)–O(4′)–C(1′)–C(2′) (τ_0)	−28.7 (5)
O(4′)–C(1′)–C(2′)–C(3′) (τ_1)	30.7 (5)
C(1′)–C(2′)–C(3′)–C(4′) (τ_2)	−21.1 (5)
C(2′)–C(3′)–C(4′)–O(4′) (τ_3)	5.0 (4)
C(3′)–C(4′)–O(4′)–C(1′) (τ_4)	14.9 (5)
O(4′)–C(1′)–N(1)–C(6) (χ_{CN})	40.6 (4)
C(3′)–C(4′)–C(5′)–O(5′) (Ψ)	46.2 (5)
O(4′)–C(4′)–C(5′)–O(5′)	−73.5 (5)
N(3)–C(4)–O(4)–C(8)	−4.2 (5)
C(5)–C(4)–O(4)–C(8)	175.9 (5)
O(4)–C(4)–C(5)–C(7)	0.8 (5)
N(3)–C(4)–C(5)–C(7)	−179.2 (6)

Table IV: Deviations (Å) of Atoms from the Least-Squares Plane of the Pyrimidine Ring[a]

*N(1)	−0.001 (3)	O(2)	−0.020 (4)
*C(2)	−0.001 (4)	C(7)	0.005 (6)
*N(3)	0.007 (4)	O(4)	−0.045 (5)
*C(4)	−0.010 (4)	C(8)	−0.155 (7)
*C(5)	0.007 (4)	C(1′)	−0.085 (5)
*C(6)	−0.002 (5)		

[a] Atoms defining the plane are marked with an asterisk. The equation of the plane is of the form $-0.718X + 0.521Y - 0.462Z - 2.38 = 0$, where X, Y, and Z are orthogonal coordinates in angstroms.

Results and Discussion

An ORTEP drawing showing the overall molecular conformation of the alkylated nucleoside and the atom labeling scheme is given in Figure 1. The fractional positional and isotropic thermal parameters are presented for all atoms in Table I. Geometric and conformational aspects of the structure are summarized in Tables II–V.

Table V: Geometry of the Hydrogen Bonds

D–H···A	D···A (Å)	D–H (Å)	H···A (Å)	D–H···A (deg)
O(5′)–H(O5′)···O(2) (i)[a]	2.702 (3)	0.95 (4)	1.76 (4)	170 (4)
O(3′)–H(O3′)···O(5′) (ii)[a]	2.828 (3)	0.87 (4)	1.96 (4)	171 (3)

[a] Symmetry code and translation: (i) $-x$, $0.5 + y$, $0.5 - z$; (ii) $-0.5 + x$, $1.5 - y$, $1 - z$.

Deoxyribose and Glycosyl Geometry and Conformation. The deoxyribofuranose ring is found in the C(1′) exo–C(2′) endo, ${}_1T^2$ twist, conformation. The phase angle (P) of pseudorotation (Altona & Sundaralingam, 1972) is 133.8 (5)°. It is outside the preferred range (144–190°) of S-type conformers by more than 10° (Sundaralingam, 1983). The amplitude of puckering (τ_m) is 31.4 (3)° and reveals that the deoxyribose ring is exceptionally flat. The C(1′) exo–C(2′) endo conformation found here is most uncommon, and a survey of 179 β-D-nucleosides (de Leeuw et al., 1980; Birnbaum et al., 1984) revealed that none assumes this sugar pucker. Furthermore, only three of the 179 nucleosides surveyed display smaller values for τ_m. The exocyclic torsion angle O(5′)–C(5′)–C(4′)–C(3′) (Ψ) assumes the preferred gauche$^+$ conformation [46.2 (5)°]. The conformation about the glycosyl bond is anti where the value of χ_{CN}, O(4′)–C(1′)–N(1)–C(6), is 40.6 (4)°. The N(1)–C(1′) bond length is 1.478 Å, and the O(4′)–C(1′)–N(1) bond angle is 107.2 (3)°.

O(4)–C(8) Methoxy Group Conformation and Base Planarity. The C(8) methyl group assumes the energetically favorable syn periplanar conformation [where the syn conformation is defined with respect to the N(3) atom]. In this conformation, two C(8) methyl group hydrogen atoms are 2.54 (3) and 2.71 (4) Å from the N(3) atom and are, in part, responsible for hindering N(3) from accepting a hydrogen bond. The N(3)–C(4)–O(4)–C(8) torsion angle is −4.2 (5)° and reveals the slight deviation of the methoxy group from the base plane (see also Table IV). In O^4-methyluridine the methoxy group is closer to the base plane; the corresponding N(3)–C(4)–O(4)–C(7) torsion angle is 0.4 (3)° for both independent molecules. The O-alkylated pyrimidine ring is flat with the average displacement of the ring atoms of 0.005 (4) Å from the least-squares plane (Table IV).

Hydrogen Bonding and Molecular Packing. A molecular packing diagram illustrating the hydrogen bonding of O^4medT is shown in Figure 2. Two hydrogen bonds are found; one between the O(3′) hydroxyl hydrogen atom and O(5′) hydroxyl oxygen atom and the other between the O(5′) hydroxyl hydrogen and O(2) carbonyl oxygen atoms (Table V). The latter of these bonds is the stronger where the O(5′) donor to O(2) acceptor distance is 2.702 (3) Å. O^4medT does not exhibit base–base stacking or, unlike O^4-methyluridine (Brennan et al., 1983), sugar–base stacking or very short sugar ring oxygen–base interactions. There are some alkyl–sugar and alkyl–base nonbonded contacts. The C(8) methyl group displays close contacts with a neighboring O(3′) hydroxyl oxygen atom [3.426 (5) Å] and with two O(4) methoxy oxygen atoms [3.582 (6) and 3.662 (6) Å] translated along the a axis. The C(7) methyl group makes close contacts with its neighboring O(3′) hydroxyl group [3.587 (5) Å] and with the O(4′) ribosyl ring oxygen atom [3.356 (4) Å].

Effects of Methylation on Pyrimidine Ring Geometry in Nucleosides. Methylation of the thymine O(4) oxygen atom produces the enol tautomeric form of the pyrimidine ring resulting in the deprotonation of the N(3) atom. This shift from the keto to enol tautomer perturbs the electronic and

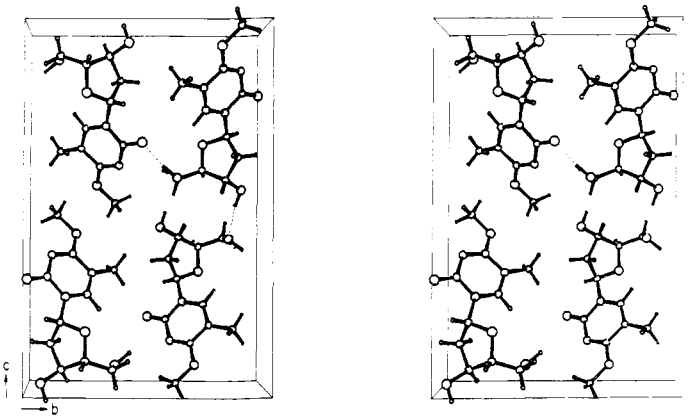

FIGURE 2: Stereo diagram of the crystal packing of O^4medT viewed down the a axis. The hydrogen bonds between H(O3′) and O(5′) and between H(O5′) and O(2) are depicted by dashed lines.

geometric properties of the O(4) alkylated base. The geometry of O^4medT is close to that of O^4-methyluridine (Brennan et al., 1983). The most important change resulting from O(4) methylation is formation of the N(3)–C(4) double bond [1.306 (5) Å] in the pyrimidine ring, the conjugation of which now more closely resembles that found in cytosine (Taylor & Kennard, 1982) than in thymine (M. Sundaralingam and Lin, unpublished results; Suck et al., 1974). However, since the methoxy group belongs to the conjugated system, the geometry around the C(4) atom in O^4medT is different from that in cytosine. For example, the internal angle at C(4) is almost the same in O^4medT and O^4-methyluridine [125.0 (4)°; 124.9 (3)° (molecule A) and 124.4 (3)° (molecule B), respectively], but more than 3° greater than that of cytosine [121.8 (2)°]. In O^4medT the angle at C(5) of 113.6 (4)° is smaller than that in cytosine [117 (1)°] and also O^4-methyluridine [116.3 (3)° (molecule A), 116.8 (3)° (molecule B)]. This could suggest greater sp^3 hybridization at the C(5) atom in O^4medT than in cytosine or O^4-methyluridine, but the values of the bnod lengths do not support this because they are comparable (within 2σ) for these compounds.

A comparison between the C(4) exocyclic bond angles of O^4medT, C(5)–C(4)–O(4) [115.2 (4)°] and N(3)–C(4)–O(4) [119.8 (4)°], with those of thymine, the respective values of which are 124.7 and 119.7°, reveals that the former bond angle contracts by more than 9° upon methylation while the latter is unchanged. This angular contraction is necessitated by the expansion of the internal C(4) angle (see above), thereby maintaining the overall trigonal character of the C(4) carbon atom. In contrast to this angular contraction, the C(5) exocyclic bond angle, C(4)–C(5)–C(7), widens to 122.4 (4)° in O^4medT as compared to the corresponding angle in thymine (119.2°).

O^4-Methylthymine–Guanine Base Pairing and the Anti Methoxy Group. A two hydrogen bond base pair between O^4meT and guanine (G) has been proposed by Abbott and Saffhill (1977) in which the N(3) and O(2) atoms of O^4meT act as hydrogen-bond acceptors from the N(1) and N(2) atoms of guanine, respectively (Figure 3). This pairing scheme requires that the O(4)–C(8) methoxy group takes the antiperiplanar conformation with respect to the N(3) atom. But, the crystal structure of O^4-methylthymidine reveals that the methoxy group shows an inherent preference for the syn-periplanar conformation as has been observed in the crystal structures of all O-alkylated nucleosides and nucleoside analogues (Brennan et al., 1983; Cook et al., 1980; Srikrishnan et al., 1983) (Figure 1). Also, it is seen that the proposed O^4meT-G base pair is unfavorable for the syn methoxy ori-

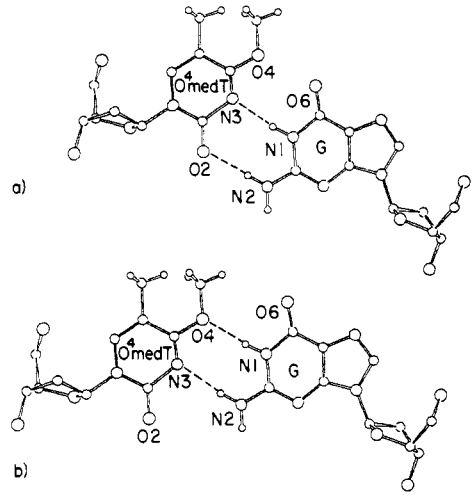

FIGURE 3: (a) Possible O⁴medT-G mispair in which the N(3) and O(2) atoms of O⁴medT accept hydrogen bonds from the N(1) and N(2) atoms of guanine. (b) Possible wobble pair in which the N(3) and O(4) atoms of O⁴medT accept hydrogen bonds from the N(1) and N(2) atoms of guanine. Only the hydrogen atoms of interest are shown.

entation because of (i) steric conflict between the C(8) methyl group and the guanine O(6) oxygen atom [where the O(6) to C(8) distance is ~1.6 Å] and (ii) less accessibility of the N(3) nitrogen for hydrogen bonding.

However, there is support for the presence of the anti-periplanar conformation for O⁴medT. Thermal stability studies (Singer et al., 1983) have shown that poly(dA-dT,O⁴medT) displays a cooperative melting curve and melting temperature (T_m) identical with those of poly(dA-dT). This suggests (Singer et al., 1983) that the thymine O(4) atom is engaged in hydrogen bonding to the N(6) amino group of adenine, which would require an anti-periplanar conformation for the O(4) alkyl group. In the crystal structure of the overcrowded molecule 1,5,N^4,N^4-tetramethylcytosine (Dattagupta et al., 1977), the $N(4)$ methyl groups were found to lie nearly in the plane of the pyrimidine ring. The steric compression between adjacent methyl groups was alleviated by an increase in the C(8)–N(4)–C(4), N(4)–C(4)–C(5), and C(4)–C(5)–C(8) bond angles by 5.8, 4.2, and 7.9°, respectively, from their standard sp^2 values of 120°. Similar distortions for the corresponding bond angles in O⁴medT would be expected to allow the O(4)–C(8) methoxy group to also take an anti-periplanar conformation. Shortening of the O(4)···O(6) distance by expansion of the O(4)–C(4)–C(5) angle can easily be relieved by propeller twisting of the base pair. Clearly, the conformation of the exocyclic methoxy group plays as important a role in O⁴meT-G base pairings as it does in the base pairing of other alkylated bases, e.g., those between N^6-methoxyadenosine and cytidine (Birnbaum et al., 1984; Stolarski et al., 1984) and N^4-methoxy(hydroxy)cytidine and adenosine (Psoda et al., 1981; Singer et al., 1984).

Semiempirical energy calculations were carried out to assess the effects of changing the values of the C(4)–O(4)–C(8), O(4)–C(4)–C(5), and C(4)–C(5)–C(7) bond angles on the nonbonded and electrostatic interactions between the C(7) and anti C(8) methyl groups. These calculations reveal that for the anti-periplanar conformation when the C(4)–O(4)–C(8), O(4)–C(4)–C(5), and C(4)–C(5)–C(7) bond angles are 118, 120, and 120°, respectively, there is close contact (~1.2 Å) between H(C8)...H(C7), resulting in high potential energy. However, when the above bond angles are widened to 125, 124, and 128°, respectively, the relative potential energy drops substantially and is only ~0.8 kcal greater than that calculated for the low energy syn conformer (−0.38 kcal mol^{-1}) of O⁴medT (Figure 4C). The plots in Figure 4 also show that to relieve the C(7)···C(8) unfavorable contacts all three bond angles undergo nearly equal expansion. The graphs show that many combinations of values for these angles can diminish the unfavorable methyl–methyl interactions in O⁴medT resulting in calculated potential energies of less than 5 kcal mol^{-1}, indicating that the anti conformation accompanied by concerted changes in the valency angles is energetically feasible.

The above findings allow us to predict that the anti conformer of O^4-methylthymine, and hence the proposed O⁴meT-G base pair, is possible. However, the rotation of the methoxy group from the syn to anti orientation will require additional energy. Studies by Engel and von Hippel (1978) have shown that *E. coli* DNA polymerase I is able to carry out the syn to anti conversion of the exocyclic methyl group

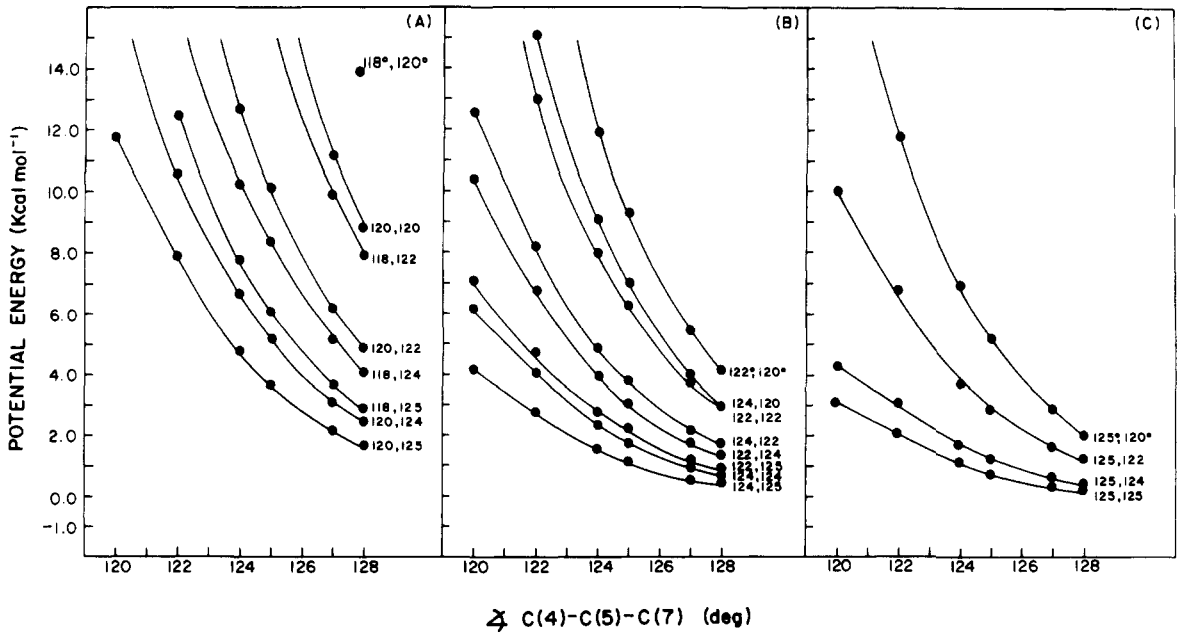

FIGURE 4: Potential energy as a function of the C(4)–C(5)–C(7) angle for various values of the C(4)–O(4)–C(8) angle (first value shown next to each curve) and O(4)–C(4)–C(5) angle (second value).

of N^6-methyladenosine, which displays a large rotational energy barrier (~12 kcal/mol). It is very likely then that *E. coli* DNA polymerase I and other DNA polymerases can similarly overcome the rotational energy barrier of O^4-methylthymine and rotate the O(4)–C(8) methoxy group into the anti position. This isomerization step, which is probably rate limiting, would explain the low incorporation of O^4meTMP into poly (dC-dG) (Hall & Saffhill, 1983).

Conclusions

The deoxyribofuranose ring of O^4-methylthymidine is very flat [τ_m = 31.4 (3)°] and takes the unusual $_1T^2$ twist, C(1′) exo–C(2′) endo conformation [P = 133.8 (5)°]. Apparently, there is no correlation between this conformation and the low value of τ_m (de Leeuw et al., 1980; Birnbaum et al., 1984). Methylation at O(4) results in the preferred syn conformation for the O(4)–C(8) methoxy group. Electronic and geometric changes in the pyrimidine ring accompanying methylation make the base very closely resemble cytidine. Formation of the O^4-methylthymine–guanine mispair requires the rotation of the O(4)–C(8) methoxy group into the anti-periplanar conformation. This conformation for the methoxy group would allow good stacking of the O^4medT-G mispair with the adjacent Watson–Crick base pairs in the DNA. Semiempirical energy calculations indicate that such a conformation is allowed if the C(8)–O(4)–C(4), O(4)–C(4)–C(5), and C(4)–C(5)–C(7) bond angles are widened. It appears then that the syn orientation of the methoxy group is favored for the unpaired monomer of O^4medT. However, once inserted into a double-stranded helix by DNA polymerases, the methoxy group is rotated from the syn to the anti orientation. The reason for this is that with the anti methoxy group O^4medT is able to participate in two hydrogen bonds and can stack well with adjacent base pairs in the helix. These favorable interactions apparently more than compensate the energy loss for the syn to anti rotation and the valency angle distortions. In addition to the proposed O^4meT-G mispair (Figure 3), an O^4meT-G "wobble" pair with N(1)-H···O(4) and N(2)-H···N(3) hydrogen bonding can be envisioned (Figure 3) where the N(3) and O(4) atoms act as hydrogen-bond acceptors. However, such wobble base pairing is considered unlikely since the O(4) methoxy oxygen atom shows little tendency for hydrogen bonding as exemplified by the present structure and the crystal structure of O^4-methyluridine.

Supplementary Material Available

Anisotropic thermal parameters and structure factor tables (7 pages). Ordering information is given on any current masthead page.

Registry No. O^4medT, 50591-13-4; guanine, 73-40-5.

References

Abbott, P. J., & Saffhill, R. (1977) *Nucleic Acids Res. 4*, 761–769.

Altona, C., & Sundaralingam, M. (1972) *J. Am. Chem. Soc. 94*, 8205–8212.

Birnbaum, G. I., Kierdaszuk, B., & Shugar, D. (1984) *Nucleic Acids Res. 12*, 2447–2460.

Brennan, R. G., Privé, G. G., Blonski, W. J. P., Hruska, F. E., & Sundaralingam, M. (1983) *J. Am. Chem. Soc. 105*, 7737–7742.

Cook, W. J., Gartland, G. L., & Bugg, C. E. (1980) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B36*, 2467–2470.

Cromer, D. T., & Waber, J. T. (1965) *Acta Crystallogr. 18*, 104–109.

Dattagupta, J. K., Saenger, W., Bolewska, K., & Kulakowska, I. (1977) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B33*, 85–89.

de Leeuw, H. P. M., Haasnoot, C. A. G., & Altona, C. (1980) *Isr. J. Chem. 20*, 108–126.

Engel, J. D., & von Hippel, P. H. (1978) *J. Biol. Chem. 253*, 935–939.

Hall, J. A., & Saffhill, R. (1983) *Nucleic Acids Res. 11*, 4185–4193.

Hruska, F. E., & Blonski, W. J. P. (1983) *Can. J. Chem. 60*, 3026–3032.

Johnson, C. K. (1976) *ORTEP II, Report ORNL-5138*, Oak Ridge National Laboratory, Oak Ridge, TN.

Lawley, P. D., Orr, D. R., Shah, S. A., Farmer, P. B., & Jarman, M. (1973) *Biochem. J. 135*, 193–201.

Lesyng, B., & Saenger, W. (1978) *Biochim. Biophys. Acta 544*, 215–224.

Main, P., Woolfson, M. M., Lessinger, L., Germain, G., & Declercq, J. P. (1974) *MULTAN. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*, Universities of York, York, England, and Louvain, Belgium.

Psoda, A., Kierdaszuk, B., Pohorille, A., Gelles, M., Kusmierek, F. T., & Shugar, D. (1981) *Int. J. Quantum. Chem. 20*, 543–554.

Singer, B., Sâgi, J., & Kuśmierek, J. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4884–4888.

Singer, B., Fraenkel-Conrat, H., Abbott, L. G., & Spengler, S. J. (1984) *Nucleic Acids Res. 12*, 4609–4619.

Srikrishnan, T., Parthasarthy, R., De, N. C., & Chheda, G. B. (1983) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B39*, 1441–1445.

Steward, R. F., Davidson, E. R., & Simpson, W. T. (1965) *J. Chem. Phys. 42*, 3175–3187.

Stolarski, R., Kierdaszuk, B., Hagberg, C.-E., & Shugar, D. (1984) *Biochemistry 23*, 2906–2913.

Stout, G. H., & Jensen, L. H. (1968) *X-ray Structure Determination, A Practical Guide*, pp 454–458, McMillan, New York.

Suck, D., Saenger, W., & Rodhe, W. (1974) *Biochim. Biophys. Acta 361*, 1–10.

Sundaralingam, M. (1983) *Conformation in Biology* (Srinivasan, R., & Sarma, R. H., Eds.) pp 191–225, Adenine Press, New York.

Taylor, R., & Kennard, O. (1982) *J. Am. Chem. Soc. 104*, 3209–3212.