which the H-2 and H-5 ENDOR signals in Figure 9 emanate. If g_x is perpendicular to this imidazole plane, then the dipolar contribution to coupling would be about -1.9 mHz (as opposed to a positive coupling if g_x is parallel to the plane). The calculated magnitudes of the hyperfine coupling to H-2 and H-5 would be 2.8 and 2.2 MHz, to be compared with the experimentally measured 2.6 and 1.8 MHz. In MbIm, the exogenous imidazole lies so that the normal to its plane makes an angle of about 24° to the g_x direction. Thus, for $(TPP)Fe(Im)_2^+$ and MbIm, there is evidence that g_x is close to the normal to an imidazole plane so that if it is π bonding to the imidazole that removes the degeneracy of the d_{xz} and d_{yz} orbitals, then it would be π back-bonding to the d_{xz} orbital that lowers the energy of that orbital with respect to the d_{yz} orbital.

In proteins where the concentration of the paramagnetic centers is highly limited and many proton resonances overlap, NMR sensitivity for H-2 and H-5 is low. Thus, ENDOR seems the method of choice for observing strongly hyperfine-coupled imidazole protons, particularly for high molecular weight proteins like cytochrome *a*. Since protein is thought to modulate heme electronic behavior through the histidine-to-Fe bond, the ENDOR spectra like those of Figure 10 may give important clues to this modulation through its different effect in different proteins. We believe that there is considerable structural information in the overall behavior of imidazole proton ENDOR as a function of *g* values. We look for the overall angle-selected ENDOR technique to give more complete insight into the precise physical causes for differences in imidazole spectra between proteins.

The protons on C-4 and N-3 (called H-4 and H-3) are both about 5.2 Å from the heme iron, and they lie away from the heme normal by about 15°. Contact interactions are, respectively, -0.4 and -0.25 MHz for H-3 and H-4.[9] We have computed first-order hyperfine coupling magnitudes for H-3 and H-4 that are in the range 1.1-1.3 MHz at g_z, 0.9-1.1 MHz at g_y, and 0.7-0.9 MHz at g_x. Unfortunately this means that the ENDOR from these protons generally overlaps with the ENDOR from porphyrin protons. However, there should be no confusion of ENDOR signals between H-4 and H-5 because the H-5 protons are much more strongly hyperfine coupled than H-4; it is for this reason that we have assigned the strongly coupled proton ENDOR in Figure 8d to H-5 even though H-4 is also protonated. The most obvious case where the ENDOR from exchangeable proton N-3 could be assigned was with MbIm for Figure 11 spectra taken near g_y. Taking g_y as approximately perpendicular to the Fe-proton vector of H-3, we computed its dipolar coupling of -0.7 MHz. On adding the H-3 contact interaction[9] to the dipolar contribution, we predicted for H-3 a coupling whose magnitude was 1.1 MHz, to be compared with 1.2 MHz measured by ENDOR. We are not sure why this exchangeable H-3 proton should be so clearly distinguishable by ENDOR only in MbIm; perhaps the exogenous imidazole to which it belongs is rigidly oriented through bonding to the endogenous distal histidine.

Acknowledgment. This work was supported by Grant AM-17884 from the National Institute of Arthritis, Metabolism and Digestive Diseases, Grant GM 35103 from the National Institute of General Medical Sciences, and Grant RR07122 from the Biomedical Research Support Grant Program, Division of Research Resources, NIH. We are grateful to J. L. De Jesus-Collazo and C. Fan for performing calculations related to proton ENDOR interpretation, to H. Taylor for technical assistance in performing the ENDOR measurements, and to Prof. R. Kreilick, University of Rochester, and co-workers for discussions on the angle-selected ENDOR method.

Carcinogenic Alkylation of Nucleic Acid Bases. Structure and Conformation of O^4-Ethyl-2′-deoxythymidine in the Solid State and in Solution[1]

George I. Birnbaum,*[2a] Krishan L. Sadana,[2b] Wayne J. P. Blonski,[2b] and Frank E. Hruska*[2b]

Contribution from the Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6, and the Department of Chemistry, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2. *Received August 14, 1985*

Abstract: O^4-Ethyl-2′-deoxythymidine (e^4dT) crystallizes in the monoclinic space group $P2_1$, and the cell dimensions are a = 5.079 (1) Å, b = 15.054 (1) Å, c = 8.467 (1) Å, β = 94.07 (1)°. X-ray intensity data were measured with a diffractometer, and the structure was solved by direct methods. Least-squares refinement, which included all hydrogen atoms, converged at R = 0.030 for 1365 observed reflections. The *O*-ethyl group is coplanar with the pyrimidine ring, the methylene carbon atom being syn to N3. It is shown that O⁴-alkylation causes significant changes in the geometry of the ring which can be attributed to an altered electronic structure. The conformation about the glycosidic bond is anti with χ_{CN} = 22.8°. The deoxyribose ring adopts the unusual C3′ endo/C2′ exo twist pucker, and the gauche$^+$ rotamer of the CH_2OH side chain is stabilized by an intramolecular C6—H···O5′ hydrogen bond. Proton NMR data for e^4dT and e^4dU reveal the usual preference for the C2′ endo sugar pucker and a conformer distribution for the C4′-C5′ bond which is expected for 2′-deoxyribosides. Comments are made on the relevance of the structure to base mispairing of *O*-alkyl pyrimidines and their enzymatic repair.

Alkylating agents are known to be mutagens and carcinogens by virtue of their reactions with various nucleophilic sites of DNA.[3] Following early interest in N7-alkylation of guanine,[4,5] considerable attention shifted to O-alkylation of the nucleic acid bases when Loveless[6] prepared O^6-alkyl guanosine and proposed that it might form atypical base pairs. The pyrimidine bases can be alkylated at either the O2 or O4 positions,[3] but there remains uncertainty as to the behavior of the altered bases during transcription and replication. In the case of O^4-alkyl thymine, Abbott and Saffhill[7] suggested the incorporation of dGMP, using randomly methylated poly(dA-dT) as a template. Singer et al.[8] prepared O^4-methyldeoxythymidine (m^4dT) which was then incorporated into a poly(dA-dT) template. Mispairing of m^4dT with guanosine was proven unambiguously by copying the synthesized polymers with DNA polymerase I. Singer et al.[8] also showed that the misincorporation of m^4dT in place of dT does not distort the helical structure of DNA. Consequently, it is important to determine exactly the electronic, hydrogen bonding, and conformational properties of nucleosides with O-alkylated bases. So far, only O^4-methyluridine (m^4U)[9,10] and 6-methoxypurine riboside (O^6-

(1) Issued as NRCC No. 25419.
(2) (a) National Research Council of Canada. (b) University of Manitoba.
(3) Singer, B.; Kušmierek, J. T. *Annu. Rev. Biochem.* **1982**, *51*, 655-693.
(4) Brookes, P.; Lawley, P. D. *Biochem. J.* **1961**, *80*, 496-503.
(5) Fraenkel-Conrat, H. *Biochim. Biophys. Acta* **1961**, *49*, 169-180.
(6) Loveless, A. *Nature (London)* **1969**, *223*, 206-207.
(7) Abbott, P. J.; Saffhill, R. *Nucleic Acids Res.* **1977**, *4*, 761-769.
(8) Singer, B.; Sägi, J.; Kušmierek, J. T. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 4884-4888.

0002-7863/86/1508-1671$01.50/0

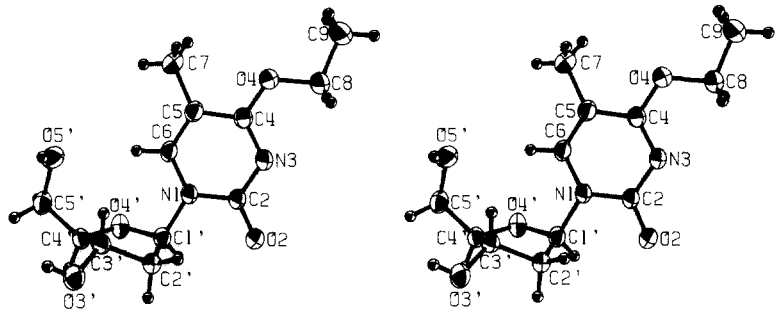

Figure 1. Stereoscopic view of e⁴dT; the ellipsoids correspond to 50% probability.

Table I. Final Atomic Parameters and Their Standard Deviations[a]

atom	x	y	z	U_{eq}/U
N(1)	4292 (3)	70466	11690 (16)	304
C(2)	5825 (3)	78006 (10)	10070 (19)	326
O(2)	5438 (3)	82710 (10)	−1788 (18)	471
N(3)	7696 (3)	80051 (10)	21967 (18)	338
C(4)	8147 (3)	74396 (11)	33705 (19)	322
O(4)	10000 (3)	76120 (9)	45200 (15)	399
C(5)	6745 (3)	66287 (11)	35392 (19)	354
C(6)	4776 (3)	64772 (10)	24034 (19)	325
C(7)	7408 (5)	60027 (16)	48841 (26)	535
C(8)	11443 (4)	84444 (13)	44643 (25)	458
C(9)	13466 (5)	84389 (16)	58411 (27)	523
C(1′)	2192 (3)	68767 (10)	−1195 (20)	321
C(2′)	3305 (3)	64228 (11)	−15383 (19)	339
C(3′)	2868 (3)	54434 (11)	−11933 (18)	308
O(3′)	2794 (3)	49402 (9)	−25995 (16)	470
C(4′)	246 (3)	54709 (11)	−4179 (19)	320
O(4′)	302 (2)	62966 (8)	4506 (16)	371
C(5′)	−301 (3)	47105 (12)	6803 (24)	392
O(5′)	1808 (3)	45296 (10)	18181 (16)	427
H(6)	373 (5)	596 (2)	241 (3)	18 (5)
H(71)	758 (8)	630 (3)	590 (5)	52 (10)
H(72)	606 (8)	662 (3)	501 (5)	56 (10)
H(73)	900 (8)	574 (3)	480 (5)	57 (10)
H(81)	1029 (6)	893 (2)	453 (4)	33 (8)
H(82)	1222 (6)	845 (2)	346 (3)	28 (6)
H(91)	1453 (8)	786 (3)	583 (5)	50 (9)
H(92)	1459 (7)	893 (3)	572 (4)	42 (8)
H(93)	1262 (7)	839 (3)	669 (4)	42 (9)
H(1′)	138 (4)	740 (1)	−43 (2)	5 (4)
H(2′)	510 (5)	661 (2)	−166 (3)	16 (5)
H(2″)	216 (5)	657 (2)	−248 (3)	24 (6)
H(3′)	414 (4)	525 (2)	−47 (3)	11 (5)
H(O3′)	265 (7)	444 (3)	−241 (4)	36 (8)
H(4′)	−111 (5)	547 (2)	−126 (3)	15 (5)
H(5′)	−66 (6)	422 (2)	−6 (4)	29 (6)
H(5″)	−181 (5)	483 (2)	126 (3)	22 (6)
H(O5′)	270 (6)	419 (2)	140 (4)	32 (8)

[a] The x coordinates and U_{eq} values of the non-hydrogen atoms were multiplied by 10^4; the y and z coordinates were multiplied by 10^5. All hydrogen atom parameters were multiplied by 10^3.

methylinosine; m⁶I)[11] have been examined by X-ray diffraction studies. In the present work we report our X-ray analysis of the first deoxyriboside with an O-alkylated base, O^4-ethyl-2′-deoxythymidine (e⁴dT). The X-ray data are compared with those of m⁴U and with the aqueous solution structure of e⁴dT and e⁴dU determined by ¹H NMR spectroscopy.

Experimental Section

To obtain e⁴dT, the procedure of Divakar and Reese[12] was used to prepare the 4-(1,2,4-triazol-1-yl) intermediate of 2′-deoxythymidine (dT), 3′,5′-protected by acetyl groups. This intermediate was treated with sodium ethoxide, and the e⁴dT was easily separated from the triazole byproduct and other impurities by flash chromatography. UV data of the product compared favorably with those of Kusmierek and Singer.[13] e⁴dU was prepared in an identical manner from 2′-deoxyuridine (dU).

Table II. ¹H Chemical Shifts (δ) and Coupling Constants (J) for e⁴dT and e⁴dU in Aqueous Solution[a,b]

δ	e⁴dT	e⁴dU	^3J	e⁴dT	e⁴dU
6	7.89	8.13	1′2′	6.6	6.5
5	1.99	6.22	1′2″	6.5	6.5
1′	6.28	6.26	1′3′	0.3	0.4
2′	2.32	2.33	2′2″	−14.1	−14.1
2″	2.49	2.53	2′3′	6.7	6.5
3′	4.46	4.44	2″3′	4.3	4.3
4′	4.09	4.11	3′4′	4.0	4.0
5′	3.87	3.87	4′5′	3.8	3.6
5″	3.79	3.78	4′5″	5.1	5.3
8[c]	4.41	4.39	5′5″	−12.6	−12.5
9[c]	1.39	1.37	56	1.1	7.5
			89[c]	7.1	7.1

[a] δ (in ppm) from TSP, J in Hz. [b] Data at 300 MHz, 298 K. [c] Methylene (8) and methyl (9) hydrogens of the O^4-ethyl group.

O^4-Ethyl-2′-deoxythymidine, $C_{12}H_{18}N_2O_5$, was crystallized from water to give colorless prisms. Precession photographs indicated the monoclinic space group $P2_1$. A crystal fragment measuring 0.10 × 0.20 × 0.25 mm was mounted on an Enraf-Nonius CAD-4 diffractometer; it provided the following data: a = 5.079 (1) Å, b = 15.054 (1) Å, c = 8.467 (1) Å, β = 94.07 (1)°, V = 645.7 Å³, ρ_c = 1.39 g cm⁻³, Z = 2 (21 °C; Cu Kα_1, λ = 1.54056 Å); $F(000)$ = 288, μ(Cu Kα) = 8.7 cm⁻¹.

Cell dimensions were determined from angular settings of 22 high-order reflections. X-ray intensities were measured with Ni-filtered Cu Kα radiation, using $\omega/2\theta$ scans with variable scan ranges and speeds. There were 1401 unique reflections accessible to the diffractometer ($2\theta \leq 152°$) of which 36 with $I < 3\sigma(I)$ were considered unobserved. The intensities were corrected for Lorentz and polarization factors; absorption corrections were considered unnecessary. The structure was determined by direct methods with the aid of the computer program MULTAN78.[14] Of the 40 starting sets subjected to tangent refinement, the solution with the highest combined figure of merit yielded an E map on which all 19 non-hydrogen atoms were located. Atomic parameters were refined by block-diagonal least squares. All hydrogen atoms were found on difference Fourier maps and refined with isotropic temperature parameters. The scattering factors were taken from the "International Tables for X-ray Crystallography"[15] and the oxygen curve was corrected for anomalous dispersion. Throughout the refinement the function $\sum w(|F_o| - |F_c|)^2$ was minimized, and a factor of 0.8 was applied to all shifts. The following weighting scheme was used during the final stages: $w = w_1 \cdot w_2$, where $w_1 = 1$ for $|F_o| \leq 5.0$, $w_1 = 5.0/|F_o|$ for $|F_o| > 5.0$, $w_2 = \sin^2\theta/0.8$ for $\sin^2\theta < 0.8$, and $w_2 = 1$ for $\sin^2\theta \geq 0.8$. This scheme made the average values of $w(\Delta F^2)$ independent of $|F_o|$ and $\sin^2\theta$. After the final cycle the average parameter shift equalled 0.1σ, and the largest one equalled 0.6σ. The conventional residual index R is 0.030, and the weighted index R' is 0.034 for 1365 reflections. A final difference Fourier map showed no significant features. The atomic coordinates are listed in Table I; lists of anisotropic temperature parameters and of observed and calculated structure factors are available (see paragraph at the end of the paper).

¹H NMR spectra of e⁴dT and e⁴dU were obtained at 300.13 MHz on a Bruker AM300 spectrometer. Aqueous solutions (4 mg/mL; pH 7.2; 305 K) were prepared and purified of paramagnetic ions as described

(9) Brennan, R. G.; Privé, G. G.; Blonski, W. J. P.; Hruska, F. E.; Sundaralingam, M. *J. Am. Chem. Soc.* **1983**, *105*, 7737–7742.
(10) Brennan, R. G.; Privé, G. G.; Blonski, W. J. P.; Hruska, F. E.; Sundaralingam, M. *J. Biomol. Struct. Dyn.* **1983**, *1*, 939–945.
(11) Cook, W. J.; Gartland, G. L.; Bugg, C. E. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1980**, *B36*, 2467–2470.
(12) Divakar, K. J.; Reese, C. B. *J. Chem. Soc., Perkin Trans. 1* **1982**, 1171–1176.
(13) Kušmierek, J. T.; Singer, B. *Nucleic Acids Res.* **1976**, *3*, 989–1000.
(14) Main, P.; Hull, S. E.; Lessinger, L.; Germain, G.; Declercq, J.-P.; Woolfson, M. M. MULTAN78, University of York, England, and University of Louvain, Belgium, 1978.
(15) "International Tables for X-ray Crystallography"; Ibers, J. A., Hamilton, W. C., Eds.; Kynoch Press: Birmingham, England, 1974; Vol. IV.

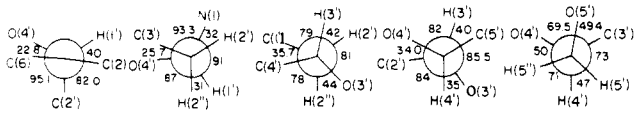

Figure 2. Newman projections along (left to right) N1–C1′, C1′–C2′, C2′–C3′, C3′–C4′, and C4′–C5′.

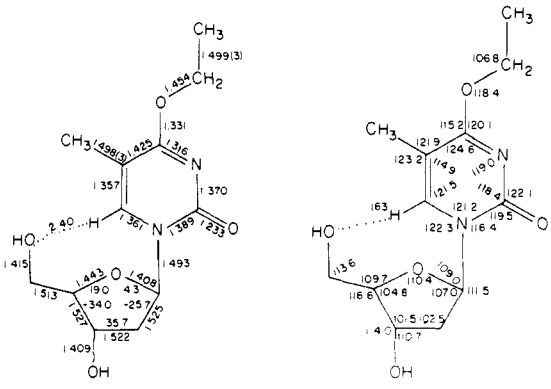

Figure 3. (left) Bond distances (Å) and endocyclic torsion angles (deg). Unless otherwise indicated, their estimated standard deviations (esd's) are 0.002 Å and 0.16°, respectively. (right) Bond angles (deg); their esd's are 0.12–0.16°.

earlier.[16,17] TSP was used as the internal reference, spectral assignments followed Allore et al.,[17] and the analysis was carried out with LAME.[18] The chemical shifts (δ) and coupling constants (J) are given in Table II.

Results and Discussion

A stereoscopic view of e⁴dT, showing the overall conformation in the crystal and the atomic numbering scheme, is presented in Figure 1. Details of the conformation can be seen in the Newman projections (Figure 2) and in Figure 3 which also shows the bond lengths and bond angles.

Pyrimidine Moiety. The pyrimidine ring is not exactly planar; C2 and C5 form the prows of a shallow boat and their displacements from the mean plane are 0.041 (2) and 0.026 (2) Å, respectively. In molecule A of m⁴U these two atoms were found to deviate by 0.037 (3) and 0.031 (3) Å, respectively, from the least-squares plane.[9] This was attributed to an unusually short intermolecular O4′···C2 contact of 2.977 Å. Yet, in e⁴dT the displacement of C2 is larger, even though the shortest intermolecular contact (also to O4′) is much weaker, the distance being 3.267 Å. In fact, significant deviations from planarity are by no means unusual in pyrimidine bases. A survey carried out several years ago[19] revealed that the average torsion angles are in the range 1.4 (2)–3.1 (3)°, rather than 0°.

The ethyl group and the pyrimidine ring are essentially coplanar with an N3–C4–O4–C8 torsion angle of 1.7 (2)°, comparable to the corresponding torsion angle in m⁴U.[9] This syn-periplanar conformation is preferred on steric grounds since it avoids the unacceptably close contacts between the C8 hydrogen atoms and the 5-methyl group which arise in the anti-periplanar arrangement. An analogous conformation was found in the structure of m⁶I.[11] Perpendicular orientations also seem unlikely in view of the calculations on 2-methoxypyridine.[20] The C4–O4–C8–C9 torsion angle is anti-periplanar (−177.2 (2)°) and directs the C9 methyl group away from the pyrimidine ring.

In order to assess the effect of O⁴-alkylation on the stereochemistry and electron distribution of the aglycon, we compared its geometry in e⁴dT with that of 5-methyluracil. Although "standard" geometries of nucleic acid–base residues have been published,[19] it would not be appropriate to use the uracil values because they do not take into account the effects of substitution

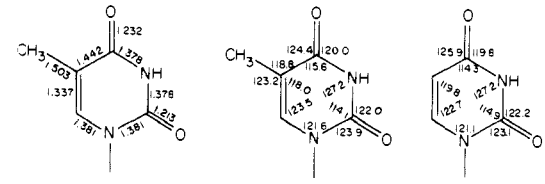

Figure 4. Average bond lengths (Å) and bond angles (deg) in 5-methyl substituted and unsubstituted uracil residues. Note the differences in bond angles at C4, C5, and C6.

Table III. Distortions of Bond Angles (in deg) Caused by O⁴-Alkylation[a]

bond angle	e⁴dT−⟨T⟩	m⁴U−⟨U⟩	e⁴dT−m⁵C	m⁴U−⟨C⟩
C6–N1–C2	−0.4	−0.6	0.5	−0.1
N1–C2–N3	4.7	4.3	−0.6	0.3
C2–N3–C4	−8.2	−8.1	−1.5	−0.9
N3–C4–C5	9.0	10.1	2.6	2.6
C4–C5–C6	−3.1	−3.0	−1.2	−0.8
C5–C6–N1	−2.0	−1.6	−0.1	0.1
N1–C2–O2	−4.4	−5.5	0.2	−1.6
O2–C2–N3	0.1	1.0	0.4	1.4
N3–C4–O(N)4	0.1	0.7	2.4	2.6
C5–C4–O(N)4	−9.2	−9.8	−5.1	−4.2
C4–C5–C7	3.1		0.3	
C6–C5–C7	0.0		0.8	

[a] ⟨T⟩ and ⟨U⟩ correspond to averages calculated from crystal structures with T and U residues, respectively. The values of m⁵C and ⟨C⟩ are taken from ref 22 and 19, respectively.

at C5.[21,22] We, therefore, selected those residues without substituents at C5 as well as those with a methyl substituent at C5. The calculated averages (Figure 4) show significant differences in bond angles, while average bond lengths in unsubstituted uracil (not shown) differ by ≤0.005 Å from those in 5-methyluracil. Using the results of the structure analyses of e⁴dT and m⁴U and the average geometry of thymine and uracil residues, we can now show (Table III) that O⁴-alkylation causes significant angular deformations of the pyrimidine moiety and that the effects are very similar in both systems. Not surprisingly, the largest changes are observed at C4. In contrast to other uracil and thymine residues, the O4–C4–C5 angle is much smaller than N3–C4–O4.

It has already been noted[9] that most bond lengths in m⁴U resemble those in cytidine more than those in uridine. The third column in Table III shows that the bond angles in e⁴dT are also fairly similar to those in the 5-methylcytosine residue of 5-methylarabinosylcytosine (m⁵araC).[22] The largest deviations, for reasons mentioned above, involve the N3–C4–O(N)4 and O(N)4–C4–C5 angles. The similarity between e⁴dT and cytidine, created by O⁴-alkylation of dT, results in an O2=C2—N3=C4—O4 fragment which was earlier interpreted as being "strongly conjugated".[9] We believe, however, that the delocalization of electrons from O4 to O2 is not as strong as it is from N4 to O2 in a cytosine residue. The evidence is provided by the bond lengths in this fragment. N3–C4 is 0.020 Å shorter than in m⁵araC while C2–N3 is 0.017 Å longer. Furthermore, C4–O4 is as long as C4–N4 in m⁵araC even though C–O bonds are generally shorter than C–N bonds. The decreased delocalization in e⁴dT is not surprising when one considers that OR is a weaker base than NH_2. We conclude, therefore, that the electron density at O2 is lower than in the case of cytidine, thus decreasing that atom's capacity of acting as a hydrogen bond acceptor in a Watson–Crick base pair with guanosine.

Deoxyribose Moiety. The bond lengths and bond angles of the sugar moiety (Figure 3) are normal. The furanose ring adopts an *N*-type pucker[23] with a phase angle of pseudorotation (P) of 11.8° ($^3T_2$) and a maximum amplitude of pucker (τ_m) of 36.5°. This is unusual for deoxynucleosides, most of which adopt an

(16) Hruska, F. E.; Blonski, W. J. P. *Can. J. Chem.* **1982**, *60*, 3026–3032.
(17) Allore, B. D.; Queen, A.; Blonski, W. J.; Hruska, F. E. *Can. J. Chem.* **1983**, *61*, 2397–2402.
(18) Haigh, C. W.; Williams, J. M. *J. Mol. Spectrosc.* **1969**, *32*, 398–406.
(19) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 3209–3212.
(20) Blonski, W. J. P.; Hruska, F. E.; Wildman, T. A. *Org. Magn. Reson.* **1984**, *22*, 505–509.
(21) Birnbaum, G. I.; Hruska, F. E.; Niemczura, W. P. *J. Am. Chem. Soc.* **1980**, *102*, 5586–5590.
(22) Birnbaum, G. I.; Gentry, G. A. *J. Am. Chem. Soc.* **1983**, *105*, 5398–5402.
(23) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205–8212.

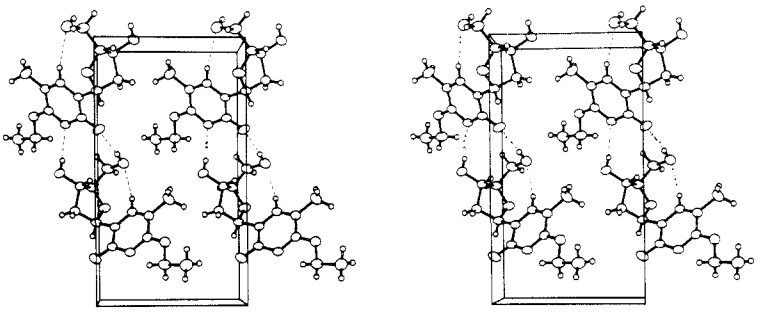

Figure 5. Stereoscopic view of the molecular packing in the crystal. Some hydrogen bonds are indicated by dotted lines. The directions of the axes are x, ⊙; y, ↓; z, →.

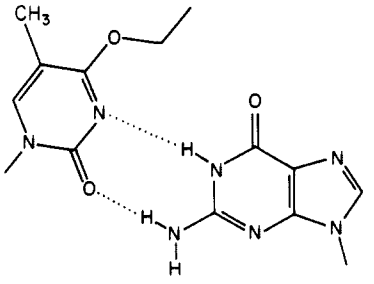

Figure 6. Proposed scheme of base pairing between e⁴dT and guanosine. Note the O^4-ethyl group in the observed syn-periplanar orientation.

S-type pucker as a result of the gauche effect.[24] The conformation about C4′–C5′ is gauche$^+$ in e⁴dT as well as in the seven other *N*-type pyrimidine deoxyribosides which we found in the current Cambridge Crystallographic Database.[25] In contrast, crystalline *S*-type deoxynucleosides show a slight preference for the trans over the gauche$^+$ conformation about C4′–C5′.[25,26] This is in line with the observed tendency in solution toward gauche$^+$ in *N*-type pyrimidine nucleosides.[27] However, it is interesting to note that in the photodimerized TpT (cyanoethyl ester), a molecule with a highly perturbed conformation, the pT fragment was found to have a type *N* pucker and a trans conformation about C4′–C5′.[28] In the present structure, the gauche$^+$ conformation is stabilized by an intramolecular C6—H···O5′ hydrogen bond (see below).

Glycosyl Bond. The *N*-glycosyl bond length (1.493 Å) is not unusual. The conformation about this bond is anti with χ_{CN}(C6–N1–C1′–O4′) = 22.8°. This torsion angle is well within the range expected for nucleosides with an *N*-type pucker of the furanose ring. A correlation between χ_{CN} and sugar ring pucker is well-known for ribofuranosides,[26] but, owing to paucity of experimental data, it is less well established for deoxyribofuranosides. From the eight molecules of pyrimidine deoxyribofuranosides with *N*-type pucker[25] we calculate $\langle\chi_{CN}\rangle$ = 26°. This value is not significantly different from the average χ_{CN} (19°) calculated for ribofuranosides.[26]

Hydrogen Bonding and Packing. In addition to the intramolecular hydrogen bond mentioned above, there are two additional protons which can, and do, participate in hydrogen bonds. The network can be represented schematically as follows:

C6—H···O5′—H···O2 O3′—H···N3

The distances and angles are given in Table IV. As commonly observed in X-ray analyses, the C–H and O–H bonds appear shorter than their real values. By extending the covalent bond lengths to their nominal values (1.09 and 0.97 Å, respectively) one obtains corrected H···*A* distances. The H6···O5′ distance of 2.26 Å is similar to those previously observed.[29,30] The O5′—H···O2 bond is fairly strong, as one would expect on the basis of the "cooperative effect".[31] The hydrogen bond to N3 is particularly interesting since it demonstrates that this atom is accessible to certain properly oriented proton donors even when the O^4-R group is oriented syn-periplanar to N3–C4. No hydrogen bond to N3 was noted in crystalline m⁴U,[9] though N1 of m⁶I serves as a proton acceptor for the 5′-hydroxyl group of a second molecule.[11] The biological significance of this hydrogen bond is discussed below.

Table IV. Distances (Å) and Angles (deg) for Hydrogen Bonds

	C6—H···O5′	O5′—H···O2	O3′—H···N3
acceptor (*A*) at	x, y, z	$1-x$, $-1/2+y, \bar{z}$	$1-x$, $-1/2+y, \bar{z}$
D···*A*	3.318 (2)	2.784 (2)	2.946 (2)
H···*A*	2.40 (3)	2.00 (3)	2.18 (4)
H···A_{corr}	2.26 (3)	1.82 (3)	1.98 (4)
D—H···*A*	164 (2)	173 (3)	173 (4)

The packing diagram (Figure 5) reveals the absence of base stacking. The hydrogen bonds join the molecules into endless ribbons parallel to y.

NMR Analysis. The magnitudes of the corresponding J values for the furanose ring (Table II) are similar to those of the O^4-methylated analogues and of the parent nucleosides dT and dU.[17] This indicates that O^4-alkylation of a deoxypyrimidine has little influence on the deoxyribose conformation in aqueous solution. Estimates based on J(3′4′) reveal a slight bias (60%) toward the C2′ endo pucker which is not unusual for deoxypyrimidines in solution but contrasts with the preference for the *N*-type pucker in crystalline e⁴dT. From the J(4′5′) and J(4′5″) data we estimate[32] the following populations for the C4′–C5′ conformation of e⁴dT and e⁴dU: gauche$^+$ (47%), trans (38%), and gauche$^-$ (15%). This distribution is comparable to that noted for dT, dU, and their O^4-methylated analogues.[17] This suggests that the C6—H···O5′ hydrogen bond found in crystalline e⁴dT plays no exceptional role in stabilizing the gauche$^+$ rotamer of e⁴dT in an aqueous solution.

Comparison of the sugar proton chemical shifts with the data of Allore et al.[17] suggests that the anti–syn distribution about the C1′–N1 bond is comparable to that of the parent molecules, dT and dU. The appearance of a five-bond coupling between H5 and H1′ in the spectrum of e⁴dU is supportive of the anti conformation.[33]

(24) Birnbaum, G. I.; Shugar, D. In "Topics in Nucleic Acid Structure", Part 3; Neidle, S., Ed.; Macmillan: London, in press.
(25) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1979**, *B35*, 2331–2339.
(26) de Leeuw, H. P. M.; Haasnoot, C. A. G.; Altona, C. *Isr. J. Chem.* **1980**, *20*, 108–126.
(27) Hruska, F. E. *Jerusalem Symp. Quantum Chem. Biochem.* **1973**, *5*, 345–366.
(28) Cadet, J.; Voituriez, Z.; Hruska, F. E.; Grand, A. *Biopolymers* **1985**, *24*, 897–903.
(29) Birnbaum, G. I.; Cygler, M.; Watanabe, K. A.; Fox, J. J. *J. Am. Chem. Soc.* **1982**, *104*, 7626–7630.
(30) Birnbaum, G. I.; Blonski, W. J. P.; Hruska, F. E. *Can. J. Chem.* **1983**, *61*, 2299–2304.
(31) Jeffrey, G. A.; Mitra, J. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1983**, *B39*, 469–480.
(32) Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Recl. Trav. Chim. Pays-Bas* **1979**, *98*, 576–577.
(33) Hruska, F. E. *Can. J. Chem.* **1971**, *49*, 2111–2118.

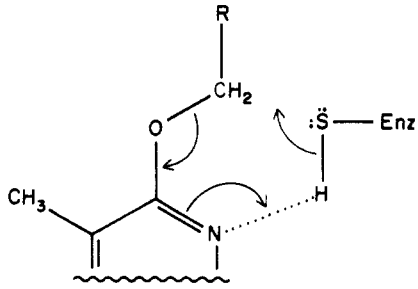

Figure 7. Proposed mechanism of the transfer of the O^4-alkyl group to the repair enzyme.

The ethyl groups of e^4dT and e^4dU display A_3X_2 patterns with $J(A,X) = 7.1$ Hz. That the methylene protons are isochronous means that they are symmetrically placed relative to the pyrimidine and in particular to the N3 ring atom. This situation obtains in the conformation revealed in the crystal state, namely the syn-periplanar and anti-periplanar orientations about the C4-O4 and O4-C8 bonds, respectively.

The $J(5,6)$ coupling constant in e^4dU is 7.5 Hz, in line with observations[16] that it is typically 0.5 Hz smaller in pyrimidines with "cytosine-like" electronic structures than in those with a 2,4-diketo structure. Following the literature,[34] Hruska and Blonski[16] attributed this decrease to a decrease in the π-bond order of the C5-C6 bond. However, as shown in Table III, the bond angles in "cytosine-like" structures (e^4dT, m^4U, and m^5araC) are significantly different from those in 2,4-diketo structures (U and dT). These differences, particularly at C5 and C6, would affect the $J(5,6)$ coupling and complicate the π bond order interpretation.

Biological Implications. In order to explain the misincorporation of guanine in replicative and transcriptive systems following DNA alkylation, Abbott and Saffhill[7] postulated a G--m^4dT base mispair involving N1(G)—H···N3(T) and N2(G)—H···O2(T) hydrogen bonds. However, their scheme shows the O-CH_3 bond in m^4dT anti-periplanar to N3-C4. This places the methyl group into a position which is unacceptably close to the C5-methyl group. Brennan et al.[9] proposed a G--m^4U base pair with the same Watson-Crick hydrogen bonds but with the N3-C4-O4-CH_3 torsion angle rotated 110° from the syn-periplanar conformation.

Alternatively, we can envisage a base pair (Figure 6) in which the alkoxy group in dT remains in its favored syn-periplanar conformation and at an acceptable distance from O6(G) (~3.5 Å). Here, only the N2—H···O2 hydrogen bond would have the normal strength, whereas the N1—H···N3 bond would be substantially weakened. On the other hand, this scheme does not suffer from the destabilizing rotation of the O-alkyl group which would increase the relative energy by ~6 kcal/mol,[20] thus counteracting the stabilization gained by a hydrogen bond. No estimates have been made of the binding strength of a G--O^4-alkyl-dT base pair, but we predict that it will be shown to be weaker than might be expected for a base pair joined by two normal hydrogen bonds.

Hydrogen bonding to N3 could also play an important role in the enzymatic repair of O^4-alkylated pyrimidine residues of DNA in prokaryotes.[35,36] The process appears to proceed in the same way as the repair of O^6-methylated guanine residues by a methyl transferase,[37] i.e., transfer of the alkyl group to a cysteinyl residue, regeneration of the parent base, and deactivation of the transferase (suicide repair[38]). Details of the mechanism we envisage are shown in Figure 7. This mechanism would represent an example of general base catalysis in which the base (N3) and the attacked center reside in the same molecule. The concerted transfer of electrons is facilitated by the syn-periplanar orientation of the O^4-alkyl group which has been found in the crystal structures of e^4dT and m^4U.

This scheme can be extended to the repair of O^6-alkylated guanines. Furthermore, hydrogen bonding at N3 would probably also facilitate the repair of O^2-alkylated pyrimidines by a DNA glycosylase.[36] It may also be significant that the O^4-alkyl base bears a structural and electronic resemblance to the 5-methylcytosine base which can play an important role in the long-term inactivation of eukaryotic genes.[39]

Acknowledgment. K.L.S. and F.E.H. acknowledge NSERC for operating support (Grants A1549 and A6434). We appreciate helpful comments of W. Fritz and Drs. N. R. Hunter and J. L. Charlton. Apart from MULTAN78,[14] all crystallographic computations were carried out with programs written by Ahmed et al.[40] Figures 1 and 5 were drawn with the ORTEP program of Johnson.[41]

Supplementary Material Available: Tables of anisotropic temperature parameters of the non-hydrogen atoms and of observed and calculated structure amplitudes (7 pages). Ordering information is given on any current masthead page.

(34) Veillard, A. *J. Chim. Phys. Phys.-Chim. Biol.* **1962**, 1056-1066.
(35) McCarthy, T. V.; Karran, P.; Lindahl, T. *EMBO J.* **1984**, *3*, 545-550.
(36) Ahmmed, Z.; Laval, J. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 1-8.
(37) Olsson, M.; Lindahl, T. *J. Biol. Chem.* **1980**, *255*, 10569-10571.
(38) Walsh, C. T. *Annu. Rev. Biochem.* **1984**, *53*, 493-535.
(39) Doerfler, W. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 919-931.
(40) Ahmed, F. R.; Hall, S. R.; Pippy, M. E.; Huber, C. P. *J. Appl. Crystallogr.* **1973**, *6*, 309-346.
(41) Johnson, C. K. ORTEP II Report ORNL-5138; Oak Ridge National Laboratory: Oak Ridge, TN, 1976.

New Enzymatic Synthesis of 2′-Deoxynucleoside-2′,2′-d_2 and the Determination of Sugar Ring Flexibility by Solid-State Deuterium NMR

Siddhartha Roy,*† Yukio Hiyama,‡ Dennis A. Torchia,‡ and Jack S. Cohen†

Contribution from the Clinical Pharmacology Branch, National Cancer Institute, and Mineralized Tissue Research Branch, National Institute of Dental Research, NIH, Bethesda, Maryland 20205. *Received August 26, 1985*

Abstract: An enzymatic synthesis of nucleosides selectively deuterated in the deoxyribose ring is described. The method is illustrated with the synthesis of 2′-deoxyguanosine-2′,2′-d_2 and thymidine-2′,2′-d_2. The availability of such specifically labeled nucleosides enables the motional characteristics of the deoxyribose ring to be investigated with deuterium NMR spectroscopy in solution and the solid state.

Rapid molecular motion in macromolecules results from interconversion between conformational states of comparable energy with relatively low activation energy barriers. Information about motion can provide understanding of the accessible conformational states of a macromolecule. It is possible that such states could be used for a variety of biological functions that a protein or nucleic acid performs.

Molecular motions in macromolecules have been investigated in detail[1] and motions of various frequencies and amplitudes have

† National Cancer Institute.
‡ National Institute of Dental Research.