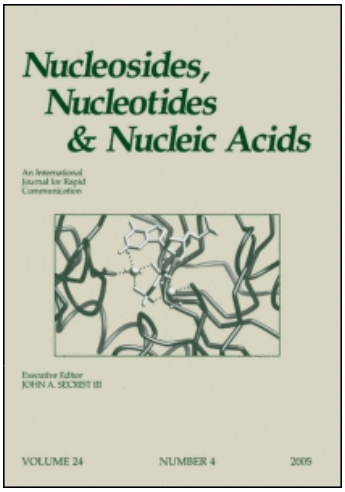


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

CRYSTAL STRUCTURE OF 1,3,5-TRIMETHYL-N⁴-HYDROXYCYTOSINE, AND ITS RELEVANCE TO THE MECHANISM OF HYDROXYLAMINE MUTAGENESIS

Karin Bjåmer Birnbaum[a]; Borys Kierdaszuk[b]; David Shugar[bc]
[a] National Research Council, Ottawa, Ontario, Canada KIA OR6 [b] Department of Biophysics, Institute of Experimental Physics University of Warsaw, Warszawa, PL, Poland [c] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawq, PL, Poland

To cite this Article Birnbaum, Karin Bjåmer , Kierdaszuk, Borys and Shugar, David(1996) 'CRYSTAL STRUCTURE OF 1,3,5-TRIMETHYL-N⁴-HYDROXYCYTOSINE, AND ITS RELEVANCE TO THE MECHANISM OF HYDROXYLAMINE MUTAGENESIS', Nucleosides, Nucleotides and Nucleic Acids, 15: 11, 1805 — 1819
To link to this Article: DOI: 10.1080/07328319608002734
URL: http://dx.doi.org/10.1080/07328319608002734

NUCLEOSIDES & NUCLEOTIDES, 15(11&12), 1805-1819 (1996)

CRYSTAL STRUCTURE OF 1,3,5-TRIMETHYL-N^4-HYDROXYCYTOSINE, AND ITS RELEVANCE TO THE MECHANISM OF HYDROXYLAMINE MUTAGENESIS

Karin Bjåmer Birnbaum [a], Borys Kierdaszuk [b] and David Shugar[b,c] *

[a] National Research Council, Ottawa, Ontario, Canada K1A OR6; [b] Department of Biophysics, Institute of Experimental Physics, University of Warsaw, 93 Żwirki i Wigury Street, PL-02089 Warszawa, Poland; [c] Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 5[a] Pawinskiego Street, PL-02106 Warszawa, Poland

ABBREVIATIONS: Cyt, cytosine; Ura, uracil; Ade, adenine; Gua, guanine; Cyd, cytidine; Urd, uridine Thd, thymidine; Thy, thymine; Ado, adenosine; OH^4Cyt, N^4-hydroxycytosine; 1-Me-OH^4Cyt, 1-methyl-N^4-hydroxycytosine; 1,3-dMe-OH^4Cyt, 1,3-dimethyl-N^4-hydroxycytosine; 1,N^4-dMe-OH^4Cyt, 1,N^4-dimethyl-N^4-hydroxycytosine; 1,3,5-tMe-OH^4Cyt, 1,3,5-trimethyl-N^4-hydroxycytosine; 1,5,N^4,N^4-tetraMe-Cyt, 1,5,N^4,N^4-tetramethylcytosine; OMe^4Cyt, N^4-methoxycytosine; OMe^4Cyd, N^4-methoxycytydine; OH^4Cyd, N^4-hydroxycytydine

ABSTRACT: 1,3,5-Trimethyl-N^4-hydroxycytosine, an analogue of the promutagenic N^4-hydroxycytosine and 5-methyl-N^4-hydroxycytosine nucleosides, crystallizes in the monoclinic space group P $2_1/n$ with cell dimensions at -147 °C: a = 7.1481(7), b = 9.2565(5), c = 13.3086(12) Å, β = 97.90(2)°, V = 872.24(13) Å^3, ρ_c = 1.426 Mg m^{-3}, Z = 4, F(000) = 401.39, μ = 0.91 mm^{-1}, λ(Cu) = 1.54056 Å, 2θ(max) = 139.3°. The crystal structure has been solved by X-ray difraction and refined to R = 3.7 % for 1457 reflections. Notwithstanding the steric hindrance imposed by methyl groups at both N(3) and C(5), the exocyclic N^4-OH group is located essentially in the plane of the ring, giving rise to an "overcrowded" molecule, like that of 1,5,N^4,N^4-tetramethylcytosine. The conformational parameters have also been compared with those of a number of related and previously reported N(1)-substituted cytosines. In the present compound the N^4-OH rotamer is in the anti conformation relative to the ring N(3), hence similar to that of one of the rotamers in N(1)-substituted N^4-hydroxycytosine, which permits normal Watson-Crick base pairing of the latter, relevant to the mechanism of hydroxylamine mutagenesis.

*Correspondence to this author: Fax:(+4822)220248; E-mail: shugar@asp.biogeo.uw.edu.pl

INTRODUCTION

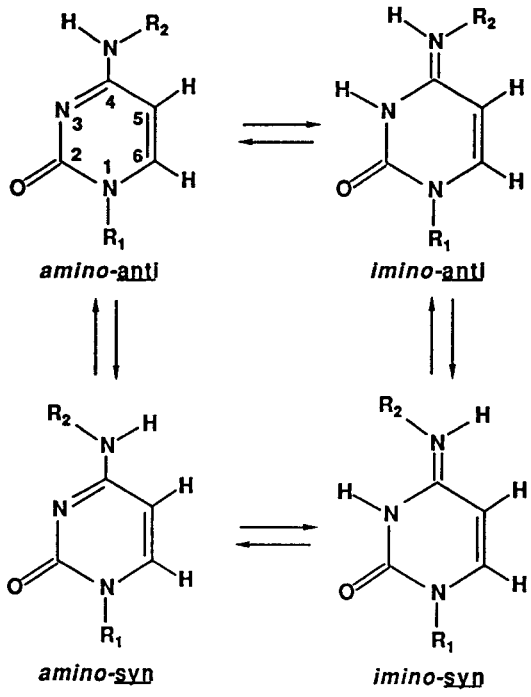
The N^4-hydroxy and N^4-methoxy derivatives of cytidine (Scheme 1), products of reaction of the latter with the known potent mutagens hydroxylamine (NH_2OH) and methoxyamine (NH_2OCH_3), are promutagens in a variety of prokaryotic and eukaryotic systems.[1] They exhibit dual functionality in that they may base pair like Cyd or Urd,[2,3] and have been shown to induce Cyd → Thd(Urd) transitions. Their promutagenic properties have also been documented in in vitro systems during replication and transcription with RNA and DNA polymerases.[2,3]

The dual functionality of N(1)-substituted OH^4Cyt and OMe^4Cyt is related to their existance in two tautomeric forms, amino ⇌ imino (see Scheme 1).[4] The imino form, which mimics Ura or Thy, is predominant in aqueous[5] and low-polar aprotic[6-8] media, and in the gas phase.[9] Only the imino forms are observed in the solid state structures of 1,5-dMe-OH^4Cyt[10] and protonated 1-Me-OH^4Cyt, as the HCl salt.[11]

The tautomeric equilibrium of OMe^4Cyt has been shown to influence base pairing with potentially complementary bases, e.g. the imino form of OMe^4Cyt forms planar associates (base pairs) only with Ade.[7,8] Furthermore, because of planarity of the molecule, there are two possible rotamers of the exocyclic N^4-OH (or N^4-OMe) group about the C(4)-N^4 bond, denoted as syn and anti relative to the ring N(3) atom (Scheme 1), and this also plays a key role in the nature of the base pairing with potentially complementary nucleosides.[4,8] A spectroscopic study of oligonucleotide duplexes showed both Watson-Crick and wobble base-pairing of OMe^4Cyt-Gua, with OMe^4Cyt in the imino form and the methoxy group both in the syn and anti conformations[12]. The preferred syn conformation of the methoxy group had a destabilizing effect on an octamer of an OMe^4Cyt-Ade duplex[13] in equilibrium between Watson-Crick and wobble configurations at low temperatures, with an anti-oriented methoxy group. The melting transition was accompanied by isomerisation to the syn conformation.[13]

X-ray diffraction[10] and spectroscopic[7,9] studies have shown that 1,5-dMe-OH^4Cyt and the corrsponding N^4-methoxy congener prefer the form imino-syn (see Scheme 1) as a result of steric hindrance by the C(5)-methyl. In the former, the syn rotamer, apparently stabilized by an intramolecular hydrogen bond, N(3)-H···O^4. OMe^4Cyt, in the crystal of a self-complementary oligonucleotide, where it forms a wobble base-pair with Gua,[14] and protonated 1-Me-OH^4Cyt,[11] both prefer the imino-syn conformation in the solid state, despite the absence of steric hindrance at C(5).

With the foregoing in mind, it appeared of interest to examine the conformation of the exocyclic N^4-OH group in the presence of steric hindrance at both the anti and syn positions, i.e. in 1,3,5-tMe-OH^4Cyt, and to compare this with published data for 1,5-dMe-OH^4Cyt,[10] 1-Me-OH^4Cyt hydrochloride,[11] the average of twelve N(3)-



Scheme 1

Tautomeric amino-imino and rotameric syn-anti equilibria of the exocyclic N^4-OH group in 1-Me-OH^{4}Cyt (R_1=CH_3; R_2=OH) and OH^{4}Cyd (R_1= ribose; R_2=OH) or exocyclic N^4-OMe group in 1-Me-OMe^{4}Cyt (R_1=CH_3; R_2=OCH_3) and OMe^{4}Cyd (R_1= ribose; R_2=OCH_3).

protonated, N(1)-substituted cytosine derivatives,[15] the average of eleven neutral N(1)-substituted cytosine derivatives,[15] and 1,5,N^4,N^4-tetraMe-Cyt.[16] The overall findings further contribute to our understanding of the dual functionality of OH^{4}Cyt.

EXPERIMENTAL

1,3,5-tMe-OH^{4}Cyt, m.p. 124 ^{o}C, was synthesized according to procedures described previously for 1,5-dMe-OH^{4}Cyt and 1,3-dMe-OH^{4}Cyt.[5,17] Its structure and purity were confirmed chromatographically and by ^{1}H NMR spectra and pH-dependent UV absorption spectra.

1,3,5-tMe-OH^{4}Cyt crystallized as colorless prisms from aqueous ethanol. A crystal, cut to size, 0.28 x 0.08 x 0.12 mm, was mounted and the cell dimensions determined

from 24 reflections with 2θ angles in the range 100 - 140°. The low-temperature (-147 °C) crystal data are: $O_2N_3C_7H_{11}\cdot H_2O$, F.W. = 187.20, monoclinic, P $2_1/n$, a = 7.1481(7), b = 9.2565(5), c = 13.3086(12) Å, β = 97.90(2)°, V = 872.24(13) Å^3, ρ_c = 1.426 Mg m^{-3}, Z = 4, F(000) = 401.39, μ = 0.91 mm^{-1}, λ(Cu) = 1.54056 Å, 2θ(max) = 139.3°. No absorption correction was considered necessary due to the low μ value.

The intensity data were collected at low temperature on a Nonius diffractometer, using the θ/2θ scan mode. 1458 unique reflections had intensities greater than 2.5σ and were considered observed. The structure was solved on the NRCVAX system[18] with the symbolic addition method and refined with full-matrix least-squares to:

R = 0.037(R' = 0.037) for the observed reflections
R = 0.041(R' = 0.041) for all reflections

with 26 atoms, 171 parameters (including the secondary extinction coefficient) and 1457 observed out of 1623 reflections up to $2\theta_{max}$ of 139.3°, where $R = \Sigma \left| |F_o| - |F_c| \right| / \Sigma |F_o|$ and $R' = \Sigma w\Delta^2/\Sigma wF_o^2$. Weights based on counting-statistics were used: $w = 1/(\sigma^2 F_o + kF_o^2)$ with a k value of 0.0004. No shift was larger than 0.001σ after the last cycle of refinement.

The final difference Fourier map shows peaks similar to those reported for cytosine monohydrate,[19] the deepest hole being -0.26 e/Å^3 and the highest peak 0.33 e/Å^3. Most of the peaks are located in the middle of bonds and can be attributed to bonding electrons, the largest peak representing the bonding density of the C(4)-C(5) bond. In addition there is a fairly round peak at a distance of 1.00 Å from N^4 with total density of the order of 0.4-0.5 electron. However, this could neither be a partial hydrogen atom nor the lone-pair electron on N^4 since the angles (C(4)-N^4-Peak 163° and O^4-N^4-Peak 64°) deviate too much from accepted values.

RESULTS AND DISCUSSION

Solid state structure of 1,3,5-tMe-Cyt

Despite the presence of steric hindrance at both the anti and syn positions, it is notable that the N^4-hydroxy group is anti to the N(3) atom of the cytosine ring. The existence of such an anti conformation in the presence of a methyl substituent at the C(5) ring atom, which is the case here, has been doubted in the past.

The final atomic coordinates and thermal parameters, listed in Table 1, were employed to calculate the bond lengths and bond angles (Fig. 1) and the 3D-structure

TABLE 1. Atomic coordinates x, y, z and Biso for 1,3,5-tMe-OH^4Cyt. The e.s.d's refer to the last digit(s).

Atom	x	y	z	Biso*
N(1)	0.17925 (19)	0.68781 (13)	0.98418 (9)	1.56(5)
C(2)	0.16216 (21)	0.59532 (15)	0.90390 (11)	1.46(5)
O^2	0.09970 (16)	0.63715 (12)	0.81818 (8)	1.97(5)
N(3)	0.21608 (18)	0.45440 (13)	0.92528 (9)	1.56(5)
C(4)	0.29635 (21)	0.40226 (16)	1.02133 (11)	1.44(5)
N^4	0.35111 (19)	0.26959 (14)	1.02049 (9)	1.82(5)
O^4	0.42600 (19)	0.21964 (13)	1.11799 (8)	2.58(5)
C(5)	0.30462 (21)	0.50619 (16)	1.10464 (11)	1.48(6)
C(6)	0.24713 (22)	0.64122 (17)	1.08106 (11)	1.66(6)
C(7)	0.1205 (3)	0.83840 (16)	0.96509 (13)	2.05(7)
C(8)	0.1903 (3)	0.35216 (18)	0.84018 (12)	2.34(7)
C(9)	0.3704 (3)	0.46965 (17)	1.21435 (12)	2.18(7)
W	0.48533 (20)	0.06184 (12)	0.88366 (9)	2.57(5)
H(4)	0.452 (3)	0.199 (3)	1.1012 (17)	4.0 (5)
H(6)	0.255 (3)	0.7164 (20)	1.1337 (15)	2.3 (4)
H(71)	-0.20 (3)	0.8456 (20)	0.9445 (15)	2.7 (4)
H(72)	0.184 (3)	0.8796 (24)	0.9123 (17)	3.5 (5)
H(73)	0.152 (3)	0.886 (3)	1.0315 (18)	4.0 (5)
H(81)	0.313 (3)	0.336 (20)	0.8154 (16)	2.7 (4)
H(82)	0.146 (3)	0.2581 (21)	0.8653 (13)	2.1 (4)
H(83)	0.091 (4)	0.393 (3)	0.7873 (20)	5.3 (6)
H(91)	0.364 (3)	0.5651 (21)	1.2548 (17)	3.2 (4)
H(92)	0.284 (3)	0.3927 (22)	1.2396 (15)	2.9 (4)
H(93)	0.502 (3)	0.4360 (20)	1.2303 (14)	2.4 (4)
HW1	0.433 (4)	0.127 (3)	0.9238 (19)	4.0 (5)
HW2	0.453 (5)	0.087 (3)	0.819 (3)	7.5 (9)

* Biso is the mean of the principal axes of the thermal ellipsoid.

and conformation, which is depicted by the stereoscopic diagram (Fig. 2). As can be seen from the bond lengths, the structure is, as expected, in the imino form, with a short C(4)-N^4 bond (1.290(2) Å) and a considerably longer C(4)-N(3) bond (1.412(2) Å), consistent with double and almost single bonds, respectively. Application of the empirical formula[20] $1/d^2 = 0.316 + 0.144p$ for calculating the double-bond character (p) from the C-N bond length (d) gave the values shown in Table 2. The p value for the C(4)-N^4 bond points to its double-bond character, while the N(3)-C(4) bond, which is the closest to a single bond, still retains some degree of double-bond character. The sp^2 character of the C(4) and N(3) atoms is well preserved, best at C(4), where the deviation

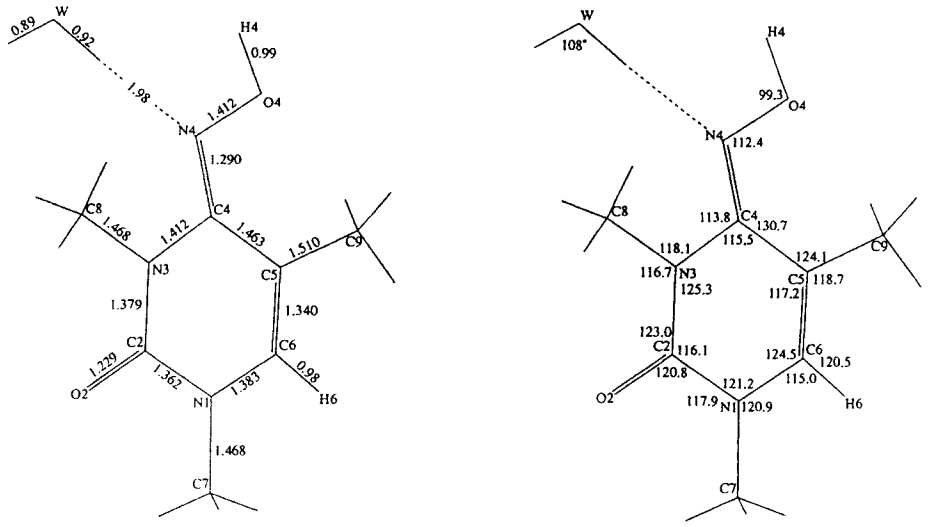


FIG. 1. Final bond lengths and angles. The e.s.d.'s are 0.0016-0.0021 Å and 0.12-0.14°, respectively; and about 10-fold larger when H atoms are involved. The closest van der Waals distances from the methyl hydrogens to the exocyclic N^4, O^4 and O^2 are: N^4···H(82) 2.37(2) Å, O^4···H(93) 2.51(2) Å, O^2···H(83) 2.30(3) Å, O^2···H(72) 2.60(2) Å.

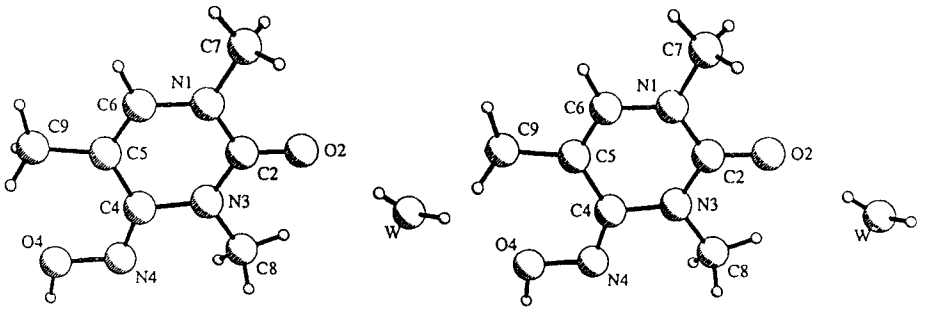


FIG. 2. Stereoscopic view of the solid state structure of 1,3,5-tMe-OH^{4}Cyt.

TABLE 2. Calculated double-bond character (p)* from C-N bond lengths (d). [20]

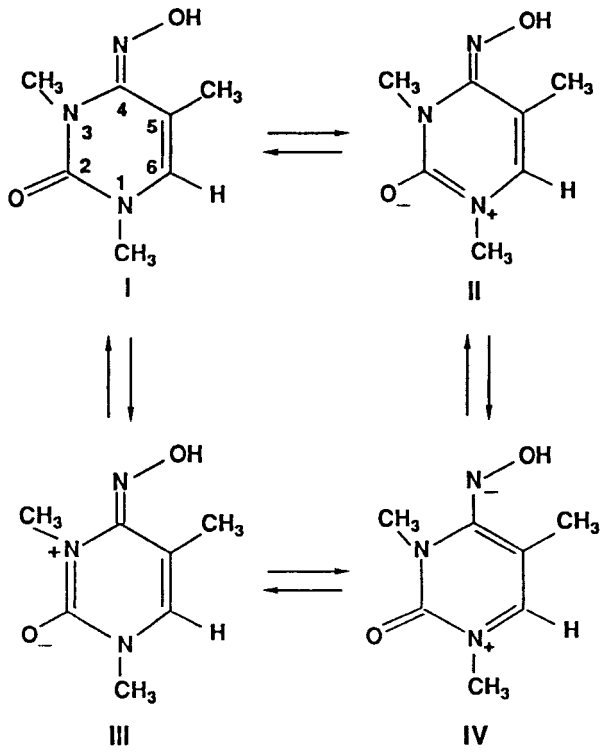
	Bond				
	C(4)-N^4	N(3)-C(4)	N(1)-C(6)	C(2)-N(3)	C(2)-N(1)
d (Å)	1.290	1.412	1.383	1.379	1.362
p	1.979	1.289	1.436	1.457	1.549

* $1/d^2 = 0.316 + 0.144p$.

of 1σ (-0.002 Å) from the plane through the three atoms to which it is bonded, is not significant, and somewhat less so at N(3), which deviates by 5σ (-0.008 Å) from planar. The double-bond character for the other three C-N bonds, viz. N(1)-C(6), C(2)-N(3), and C(2)-N(1), is midway between single and double bonds. However, the C(2)-N(1) bond is significantly shorter, by 0.019 Å (13σ) from the average of the other two (1.3807(15) Å), which are equal in value within experimental errors.

The bond lengths of the molecule, with their degree of double-bond character, indicate that, in addition to the imino form (I, Scheme 2), there are three other canonical forms (II, III and IV, Scheme 2), all anti, that contribute to the ground state of the molecule, with a larger contribution from II than from III and IV, since the C(2)-N(1) bond is shorter than C(2)-N(3) and N(1)-C(6).

Despite the presence of a methyl group at both the N(3) and C(5) positions, the molecule is essentially planar with a small angle of 3.8 (1)° between the plane of C(4)-N^4-O^4 and the cytosine ring plane (Table 3), manifested also by the small C(5)-C(4)-N^4-O^4 torsion angle of 1.8°. The difference in the value of these two angles is due to deviations from planarity of the cytosine ring, which are significant. The largest deviations from the ring plane occur for the exocyclic substituents, and to a lesser degree in the cytosine ring itself, where the deviations are smaller, the two largest ones being C(4) and N(3), 0.031(2) and -0.022(2) Å, respectively. These deviations are due either to the single-bond character of the N(3)-C(4) bond, even though the sp^2 character of the C(4) and N(3) atoms is fairly well preserved as discussed above, or to overcrowding from the C(8) and C(9) methyl atoms. Deviations from planarity in most structures of this type are due to overcrowding.[21] The methyl atoms C(8) and C(9) are



Scheme 2

Canonical forms of imino-anti 1,3,5-tMe-OH^{4}Cyt.

displaced to the same side (by -0.089(3) Å and -0.077(3) Å) of the plane as the N(3) and C(5) ring atoms to which they are attached, and this also results in deviations from planarity of the torsion angles about the N(3)-C(4) and C(4)-C(5) bonds by up to 6 and 5°, respectively. There are distortions of the exocyclic angles at C(4) and C(5) as manifested by the opening up of the angles N^4-C(4)-C(5) (130.7°) and C(4)-C(5)-C(9) (124.0°) and a consequent closing of the adjacent ones. This can probably be attributed more to packing forces resulting from formation of the hydrogen bond from H(W1) to N^4 than to steric hindrance from the C(9) methyl group. The O^4···H(93) distance (2.51(2) Å) is longer than the N^4···H(82) distance (2.37(2) Å) even though the van der Waals radius for O is about 0.1 Å shorter than for N.

TABLE 3. Weighted least-squares planes [a].

Plane 1 (Cytosine ring)		Plane 2 (N^4-hydroxy group)	
Atoms	Distance (Å)	Atoms	Distance (Å)
N(1)	0.011(2)	C(4)	-0.0001(21)
C(2)	0.007(2)	N^4	0.0001(19)
N(3)	-0.022(2)	O^4	0.0001(19)
C(4)	0.031(2)	H^4	-0.05 (3)
C(5)	-0.012(2)	N(3)	-0.036 (3)*
C(6)	-0.012(2)	C(5)	0.035 (4)*
O^2	0.012(2)*		
N^4	0.114(2)*		
O^4	0.129(3)*		
C(7)	0.019(3)*		
C(8)	-0.089(3)*		
C(9)	-0.077(3)*		
χ^2	618.1	χ^2	3.024

[a] Equations of the planes:

1. 6.8354(12)X + 2.220(5)Y - 3.956(8)Z = 1.152(8) Å.
2. 6.7257(24)X + 2.798(14)Y - 3.733(25)Z = 0.69(3).

Dihedral angle between planes: 3.78(11) °.

* Atoms excluded from calculation of the plane.

The same pattern of the exocyclic angles is not seen at N(3), where they are both less than 120°, but the opening of the N(3)-C(2)-O^2 angle is probably due to the steric effect of the C(8) methyl group, the O^2···H(83) distance being short (2.30(3) Å) while O^2···H(72) is longer (2.60(2) Å). Otherwise, there is very little strain in the lower half of the ring as indicated by the exocyclic atoms O^2 and C(7) being close to the ring plane.

Comparison with other related structures

A comparison of the bond lengths in the present structure with those in 1,5-dMe-OH^4Cyt,[10] 1-Me-OH^4Cyt hydrochloride,[11] the average of twelve N(3)-protonated, N(1)-substituted cytosine derivatives[15] (R-N(1)-CytH^+), the average of eleven neutral N(1)-substituted cytosine derivatives[15] (R-N(1)-Cyt), and 1,5,N^4,N^4-tetramethylcyto-

sine[16] (1,5,N^4,N^4-tetraMe-Cyt) is shown in Table 4. Compounds 1, 2, 3, and 4 occur in the imino form while 5 and 6 are in the amino form. This is manifested in the lengths of the N(3)-C(4) and C(4)=N^4 bonds, the former possessing more single-bond character, and the latter more double-bond character in the imino form, while the reverse holds for the amino form. As may be seen from the table, the N(3)-C(4) bond is the one that is most sensitive to variations in structure, the difference between the longest and the shortest being as much as 0.073 Å. The values decrease gradually when moving from compound 1 towards compound 6. Conversely, the length of C(4)=N^4 gradually increases from compound 1 to 6, but the range of values for this bond is narrower, only 0.036 Å. The gradual increase and decrease in the lengths of these two bonds indicate that, in addition to the imino and amino forms, other canonical species, some of them mentioned above, also play a role in all these structures. The type of canonical forms present would depend on the substituents on the ring. Packing forces such as hydrogen bonding can affect the bond lengths and angles. It should be noted that the protonated compounds (3, 4) have bond lengths more or less midway between the nonprotonated imino and amino forms, and could have been placed in a group by themselves.

Most of the bond lengths in the present structure correspond very closely to those of 1,5-dMe-OH^{4}Cyt (compound 2, Table 4). However, some of the bonds observed here are significantly longer, the largest difference being in the N(3)-C(4) bond (0.025 Å, about 8σ), the other deviations in the ring being for the adjacent bonds, C(4)-C(5), and C(2)-N(3) (0.017 Å, and 0.012 Å, respectively), another indication that a larger degree of canonical form II might be present here.

As may be seen from Table 4, all the ring bond angles are consistent with those in other N(3)-substituted cytosine derivatives with an N(3)-C(4) single bond (the imino form): the C(2)-N(3)-C(4) angle (125.3°) at N(3) being about 5° larger than for N(3)-unsubstituted cytosine derivatives (the amino form), where it is close to 120°, while the adjacent angles, N(1)-C(2)-N(3) (116.1°) and N(3)-C(4)-C(5) (115.5°), are about 4° less than 120°. There is much less difference between the imino and amino forms in the other three ring angles, the C(4)-C(5)-C(6) angles being close to 117°, the C(5)-C(6)-N(1) angles exceeding 120° by 3° on the average, and the C(6)-N(1)-C(2) angles being about 121°.

It is of considerable interest that, like 1,5,N^4,N^4-tetraMe-Cyt,[16] 1,3,5-tMe-OH^{4}Cyt is an "overcrowded" molecule which deviates only slightly (3.8°) from planarity. In the former molecule alleviation of the steric strain is achieved by opening up of the angles

TABLE 4: Comparison of the bond lengths (Å) and the ring bond angles (°) in the present structure with those in various cytosine and N^4-hydroxycytosine derivatives.

	Imino				Amino	
	This work (1)	1,5-dMe -OH^4Cyt (2)	1-Me -OH^4Cyt (3)	R-N(1) -CytH+ (4)	R-N(1)-Cyt (5)	1,5,N^4,N^4- -tetraMe-Cyt (6)*
Bond						
N(1)-C(2)	1.362	1.362	1.369	1.377	1.399	1.389
N(1)-C(6)	1.383	1.381	1.367	1.361	1.361	1.373
C(2)-N(3)	1.379	1.367	1.376	1.381	1.350	1.348
C(2)=O^2	1.229	1.222	1.222	1.215	1.241	1.240
N(3)-C(4)	1.412	1.387	1.354	1.352	1.335	1.335
C(4)=N^4	1.290	1.288	1.302	1.313	1.329	1.350
C(4)-C(5)	1.463	1.446	1.419	1.414	1.429	1.449
C(5)=C(6)	1.340	1.330	1.336	1.343	1.343	1.332
N^4-O^4	1.412	1.416	1.384			
N(1)-C(7)	1.468	1.461	1.470	1.483	1.474	1.458
N(3)-C(8)	1.468					
C(5)-C(9)	1.510	1.492				1.524
Angle						
N(1)-C(2)-N(3)	116.1	115.6	116.0	115.1	119.1	118.4
C(2)-N(3)-C(4)	125.3	125.7	124.4	124.9	120.6	121.9
N(3)-C(4)-C(5)	115.5	116.2	117.9	117.6	121.1	121.1
C(4)-C(5)-C(6)	117.2	117.1	117.8	118.5	117.7	115.6
C(5)-C(6)-N(1)	124.5	124.2	122.9	122.0	121.1	123.2
C(6)-N(1)-C(2)	121.2	121.0	120.9	121.9	120.3	120.7

* These are average values for the two independent molecules in this structure.

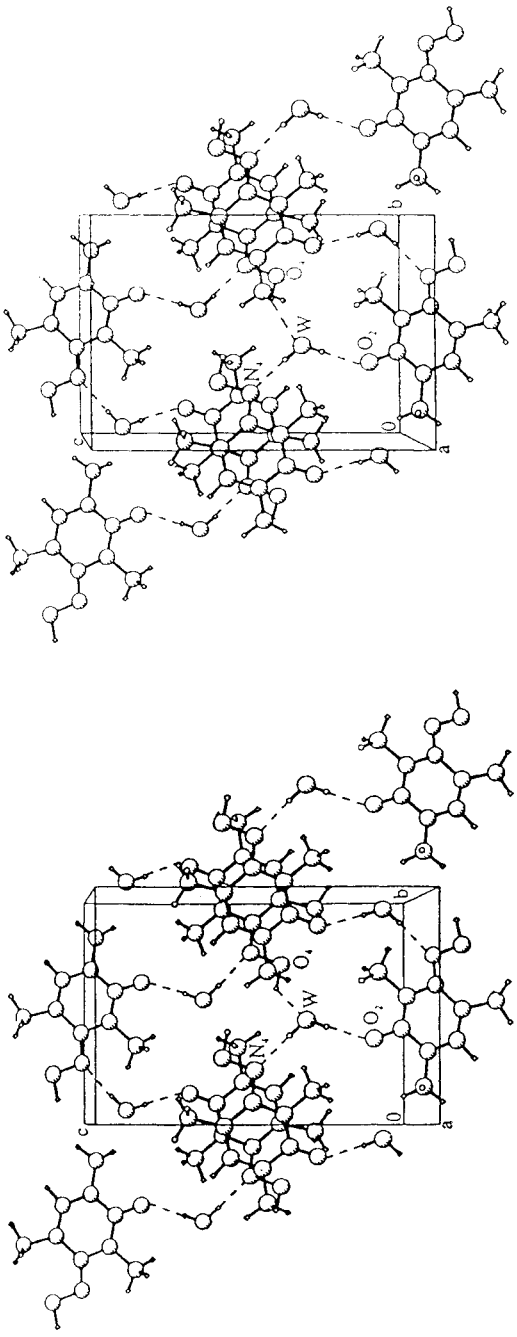


FIG. 3. Stereoscopic view of the packing and hydrogen bonding of 1,3,5-tMe-OH^{4}Cyt in the solid state.

TABLE 5. Hydrogen-bond distances (Å) and angles (°).

Donor atom	Acceptor atom	Position of acceptor atom			Distance D···A	Distance H···A	Angle (°)
O^4	W	1-x	-y	2-z	2.682(2)	1.73(3)	O^4-H^4···W 160.3 (20)
W	O^2	0.5-x	y+0.5	1.5-z	2.761(2)	1.87(3)	W-HW2···O^2 176 (3)
W	N^4	x	y	z	2.901(2)	1.98(3)	W-HW1···N^4 172.9 (22)

at C(4), C(5), and N^4 (N^4-C(4)-C(5) 124.2°, C(4)-C(5)-C(9) 127.9°, C(4)-N^4-C(8) 125.7°) and a deviation from planarity of 4.0° between the heterocycle and the dimethylamino group. In the one investigated here similar values of the exocyclic angles (N^4-C(4)-C(5) 130.7° and C(4)-C(5)-C(9) 124.1°) are observed, as well as a similar deviation from planarity (3.8°). However, the opening up of the angles at C(4) and C(5) cannot be attributed to steric hindrance alone, but also to packing forces from hydrogen bonds, as mentioned earlier. Another example of an increase in these angles from 120° in the absence of steric hindrance is 1,3-dMe-Ura[22] with no methyl at C(5), but there is still opening of the adjacent angles (O^4-C(4)-C(5) 125.5°) and C(4)-C(5)-H(5) 123.4°). It is not clear whether the deviations in the latter structure are due to packing forces involving the C(4)-O^4 carbonyl groups, which overlap with adjacent molecules and are attributed to dipole-induced dipole forces. In 1,5-dMe-OH^4Cyt[10] on the other hand, there is also no steric hindrance between the C(5) methyl group and the N^4-hydroxyl; since the latter is syn to N(3), there is no distortion of the exocyclic angles at C(4) and C(5).

Finally, it should be noted that the solid state structure has been reported for a self-complementary oligodeoxynucleotide containing an OMe^4C residue in the imino form as the syn rotamer in each chain.[14] Although the resolution was low, the C(4)-N^4 bond lengths were estimated as 1.28 Å and 1.32 Å, hence close to the corresponding bond lengths in the imino forms of the compounds listed in Table 4.

Hydrogen bonding and packing

As may be seen from the packing diagram (Fig. 3), and the hydrogen bond distances and angles listed in Table 5, the molecules stack head-to-tail in the [100] direction with approximately equal intermolecular distances from one molecule to the one below and above it, respectively. The N^4-O^4 bond stacks above the N(1) atom

with the shortest intermolecular contact to the molecule below being C(8)···C(6') 3.432(3) Å, and to the one above N(1)···N(4") 3.389(2) Å. With the absence of a hydrogen atom at N(3), neither intramolecular nor hydrogen bonding to a symmetry-related molecule, and consequent base-pairing, is possible. Consequently, all the hydrogen bonds involve water molecules and link the cytosine molecules in three dimensions to the water channels running between the cytosine stacks in the [100] direction; the hydrogen of the hydroxyl group is donated to the water molecule which, in turn, donates its hydrogen atoms to the exocyclic atoms N^4 and O^2 of two symmetry-related molecules.

Hydroxylamine mutagenesis of T-even bacteriophages

Relevant to the foregoing is the fact that hydroxylamine is highly mutagenic against the T-even bacteriophages,[23] the DNA of which contains, in place of cytosine, 5-hydroxymethylcytosine and/or glucosylated 5-hydroxymethylcytosine. It would clearly be desirable to establish, in this case, the extent to which a C(5)-hydroxymethyl substituent affects the syn-anti equilibrium of the exocyclic N^4-OH, most simply with the aid of a model compound such as 1-methyl-N^4-hydroxy-5-hydroxymethylcytosine.

Acknowledgements

Publication No. 39098 of the National Research Council of Canada. We are indebted to Dr. Anna Wróblewska (Medical Academy, Łódź, Poland) for synthesis of 1,3,5-tMe-OH^{4}Cyt. We wish to express our thanks to Drs. Eric Gabe and Gary Enright for helping with data collection and the use of the NRCVAX computer programs. The research of B.K. and D.S. was supported by an International Research Scholar's award (HHMI 75195-543401) of the Howard Hughes Medical Institute, and, in part, by the Polish Ministry of High Education (BST-501/BF).

REFERENCES

1. Budowsky, E.I. *Prog. Nucleic Acid. Res. Mol. Biol.* 1976, **16**, 125-188.
2. Flavell, R.A.; Sabo, D.L.; Bandle, E.F.; Weissmann, C. *J. Biol. Chem.* 1974, **89**, 255-272.
3. Singer, B.; Spengler, S. *Biochemistry* 1981. **20**, 1127-1132.
4. Shugar, D.; Kierdaszuk, B. *Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci.* 1985, **8**, 657-668.
5. Brown, D.M.; Hewlins, M.J.E.; Schell. P. *J. Chem. Soc.* 1968, **C**, 1925-1929.
6. Brown, D.M.; Hewlins, M.J.E. *Nature* 1969, **221**, 656-657.
7. Kierdaszuk, B.; Shugar, D. *Biophys. Chem.* 1983, **17**, 285-295.

8. Kierdaszuk, B.; Stolarski, R.; Shugar, D. *Eur. J. Biochem.* 1983, **130**, 559-564.
9. Kulinska, K.; Psoda, A.; Shugar, D. *Acta Biochim. Polon.* 1980, **27**, 57-65.
10. Shugar, D.; Huber, C.P.; Birnbaum, G.I. *Biochim. Biophys. Acta* 1976, **447**, 274-284.
11. Birnbaum, G.I.; Kulikowski, T.; Shugar, D. *Can. J. Biochem.* 1979, **57**, 308-313.
12. Nedderman, A.N.R.; Stone, M.J.; Williams, D.H.; Kong Thoo Lin, P.; Brown, D.M. *J. Mol. Biol.* 1993, **230**, 1068-1076.
13. Stone, M.J.; Nedderman, A.N.R.; Williams, D.H.; Kong Thoo Lin, P.; Brown, D.M. *J. Mol. Biol.* 1991, **222**, 711-723.
14. Van Meervelt, L.; Moore, M.H.; Kong Thoo Lin, P.; Brown, D.M.; Kennard, O. *J. Mol. Biol.* 1990, **216**, 773-781.
15. Wang, A.H.-J.; Barrio, J.R.; Paul, I.C. *J. Am. Chem. Soc.* 1976, **98**, 7401-7408.
16. Dattagupta, J.K.; Saenger, W.; Bolewska, K.; Kulakowska, I. *Acta Cryst.* 1977, **B33**, 85-89.
17. Janion, C.; Shugar, D. *Biochem. Biophys. Res. Commun.* 1965, **18**, 617-622.
18. Gabe, E.J.; Le Page, Y.; Charland,J.-P.; Lee, F.L.; White, P.S. *J. Appl. Cryst.* 1989, **22**, 384-387.
19. Eisenstein, M. *Acta Cryst.* 1988, **B44**, 412-426.
20. Staab, H.A. *Einführung in die theoretische organische Chemie, (Weinheim: Verlag Chemie)*, 1960, p. 197.
21. Voet, D.; Rich, A. *Progr. Nucleic Acid Res. Mol. Biol.* 1970, **10**, 183-265.
22. Banerjee, A.; Dattagupta, J.K.; Saenger, W.; Rabczenko, A. *Acta Cryst.* 1977, **B33**, 90-94.
23. Champe, S.P.; Benzer, S. *Proc. Natl. Acad. Sci. U.S.* 1962, **48**, 532-546.

Received April 17, 1996
Accepted September 11, 1996