

# Protonated base pairs explain the ambiguous pairing properties of $O^6$ -methylguanine

(hydrogen bonding/DNA structure/protonated cytosine/mutagenesis)

LOREN DEAN WILLIAMS AND BARBARA RAMSAY SHAW\*

Department of Chemistry, Paul M. Gross Chemical Laboratory, Duke University, Durham, NC 27706

Communicated by John A. Schellman, November 24, 1986

**ABSTRACT** The base-pairing interactions of promutagenic  $O^6$ -methylguanine ( $O^6$ -MeGua) with cytosine and thymine in deuterated chloroform were investigated by  $^1\text{H}$  NMR spectroscopy. Nucleosides were derivatized at hydroxyl positions with triisopropylsilyl groups to obtain solubility in nonaqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. We were able to observe hydrogen-bonding interactions between nucleic acid bases in a solvent of low dielectric constant, a condition that approximates the hydrophobic interior of the DNA helix.  $O^6$ -MeGua was observed to form a hydrogen-bonded mismatch with thymine. Whereas  $O^6$ -MeGua did not form hydrogen bonds with cytosine (via usual, wobble, or unusual tautomeric structures), it did form a 1:1 hydrogen-bonded complex with protonated cytosine. The pairing of unprotonated cytosine in chloroform is thus consistent with the known preference of  $O^6$ -MeGua for thymine over cytosine in polymerase reactions. In contrast, the pairing of protonated cytosine is consistent with the greater stability of oligonucleotide duplexes containing cytosine- $O^6$ -MeGua as compared with thymine- $O^6$ -MeGua base pairs [Gaffney, B. L., Markey, L. A. & Jones, R. A. (1984) *Biochemistry* 23, 5686-5691]. Our observation that cytosine must be protonated in order to pair with  $O^6$ -MeGua suggests that the cytosine- $O^6$ -MeGua base pair in DNA is stabilized by protonation of cytosine. Through this mechanism, methylation at the  $O^6$  position of guanine in double-stranded DNA could promote cross-strand deamination of cytosine (or 5-methylcytosine) to produce uracil (or thymine).

As first observed by Loveless (1),  $O^6$  guanine adducts in DNA are particularly important in induction of carcinogenesis (reviewed in ref. 2). The mutagenic potential of guanine alkylation at the  $O^6$  position appears to be determined both by its effects on the hydrogen-bonding properties of guanine and by the persistence of the lesion (3). *In vitro* studies have shown that  $O^6$ -methylguanine ( $O^6$ -MeGua) codes for thymine (or uracil) instead of cytosine during both DNA replication and transcription processes (4-8). In DNA synthesis, dTTP is incorporated at least an order of magnitude more frequently than dCTP opposite  $O^6$ -MeGua (4). Conversely,  $O^6$ -Me-dGTP is incorporated preferentially opposite thymine rather than cytosine by the Klenow fragment of *Escherichia coli* DNA polymerase I (5, 6). In transcription, UTP is incorporated preferentially over CTP opposite  $O^6$ -MeGua by RNA polymerase (7, 8).

The relative contributions of the four  $O^6$ -MeGua base pairs (with guanine, adenine, cytosine, or thymine) to oligonucleotide stability are in contradiction with the coding preferences of  $O^6$ -MeGua. Jones and coworkers (9) have reported that in oligodeoxynucleotide dodecamers the  $O^6$ -MeGua-thymine base pair has a greater destabilizing effect on duplex

stability than does any other  $O^6$ -MeGua base pair, whereas the  $O^6$ -MeGua-cytosine base pair is the least destabilizing. Thus, although  $O^6$ -MeGua interacts preferentially with thymine in the polymerase complex, it forms the least stable base pair with thymine in oligonucleotides. Such results imply either that base pairing is not as important as generally assumed in replication or that factors which stabilize  $O^6$ -MeGua base pairing in the polymerase complex differ significantly from those in the DNA helix. In two recent reports, Patel *et al.* (10, 11) deduced by  $^1\text{H}$  NMR that both the  $O^6$ -MeGua-thymine and the  $O^6$ -MeGua-cytosine base pairs are stacked in dodecamers. These authors were not able to establish direct evidence regarding the specific hydrogen-bonding schemes of  $O^6$ -MeGua with either cytosine or thymine. To understand the pairing properties of  $O^6$ -MeGua in the absence of stacking effects and to reconcile the discrepant effects of  $O^6$  methylation of guanine on the structure and biosynthesis of DNA, we have investigated the base-pairing interaction of  $O^6$ -MeGua with thymine and cytosine in  $\text{C}^2\text{HCl}_3$  with  $^1\text{H}$  NMR.

Numerous spectroscopic and crystallographic studies have shown that the hydrogen-bonding specificity expressed by nucleic acid bases in DNA is manifested between nucleic acid monomers in nonaqueous solvents (12-16). To date, however, the poor solubilities of nucleosides and bases in appropriate solvents have impeded physicochemical investigations of the base-pairing properties of the bases in solvents where solute-solute hydrogen bonding can be maximized. Here we report that the 2',3'-isopropylidene-5'-triisopropylsilyl derivatives of cytidine (compound 1) and  $O^6$ -methylguanosine (compound 2) and the 3',5'-bis(triisopropylsilyl) derivative of thymidine (compound 3) shown in Fig. 1 have properties that are suitable for spectroscopic investigation of the base-pairing properties of nucleic acid bases in solution. These nucleoside derivatives are soluble in chloroform at concentrations on the order of 50 mM at temperatures approaching  $-40^\circ\text{C}$ , thus allowing direct evaluation of the effects of methylation on DNA base pairing, as described in the experiments reported below. The purpose of the triisopropylsilyl and isopropylidene groups is to obtain solubility of the nucleosides and to prevent the ribose hydroxyls from forming hydrogen bonds.

## MATERIALS AND METHODS

**Synthesis.** Guanosine and cytidine (American Bionetics, Emeryville, CA) were converted in 10- to 20-g lots to the 2',3'-isopropylidene nucleosides (17). The 2',3'-isopropylidene nucleosides were added to a 1.2-fold molar excess of triisopropylsilyl chloride (Aldrich) in the presence of excess imidazole in dimethylformamide to form the 2',3'-isopropylidene-5'-triisopropylsilyl nucleosides (18). Thymidine was converted to 3',5'-bis(triisopropylsilyl)thymidine

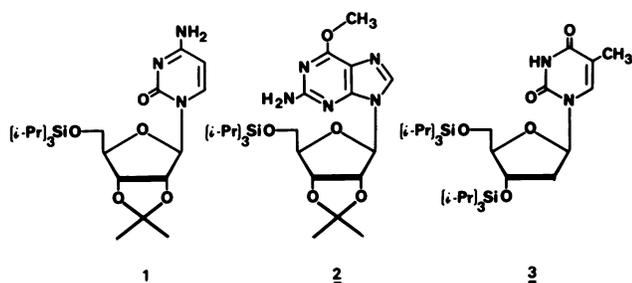


FIG. 1. Modified nucleosides (compounds 1–3) used in this study.

by the addition of 2.4 equivalents of triisopropylsilyl chloride, using a procedure otherwise similar to that used for the other nucleosides. All 2',3'-isopropylidene-5'-triisopropylsilyl nucleosides were purified by flash chromatography, followed by recrystallization from  $\text{CH}_2\text{Cl}_2$ /hexane. The products were pure as determined by thin-layer chromatography, high-performance liquid chromatography, elemental analysis, and  $^1\text{H}$  NMR.

**2',3'-O-Isopropylidene-5'-Triisopropylsilylcytidine (Compound 1).**  $^1\text{H}$  NMR ( $\text{C}^2\text{HCl}_3$ ,  $25^\circ\text{C}$ ),  $\delta$  1.006–1.172 (m, triisopropyl, 18H); 1.323 (s,  $\text{CH}_3$ , 3H); 1.552 (s,  $\text{CH}_3$ , 3H); 3.830–4.022 (m,  $5'\text{-CH}_2$ , 2H); 4.251–4.278 (m,  $4'$  H, 1H); 4.743–4.826 (m,  $3'$  H and  $2'$  H, 2H); 5.589, 5.619 (d, 5H, 1H); 5.907, 5.913 (d,  $1'$  H, 1H); 7.699, 7.728 (d, 6H, 1H).  $\text{UV}_{\text{max}}$  ( $\text{CH}_2\text{Cl}_2$ ) 275 nm. IR (KBr)  $1655\text{ cm}^{-1}$  (s). Elemental analysis calculated: C, 57.37; H, 8.48; N, 9.56. Found: C, 57.20; H, 8.34; N, 9.31.

The perchlorate salt of 2',3'-O-isopropylidene-5'-triisopropylsilylcytidine was obtained by addition of the appropriate volume of standardized perchloric acid (Mallinckrodt) to a stock solution of 2',3'-O-isopropylidene-5'-triisopropylsilylcytidine.

**2',3'-O-Isopropylidene-5'-Triisopropylsilyl- $O^6$ -Methylguanosine (Compound 2).** The conversion of 2',3'-isopropylidene-5'-triisopropylsilylguanosine to the  $O^6$ -methyl derivative was effected by procedures reported elsewhere, describing  $O^6$  substitution of a different ribose derivative of guanosine (19).  $^1\text{H}$  NMR ( $\text{C}^2\text{HCl}_3$ ,  $25^\circ\text{C}$ ),  $\delta$  1.006–1.060 (m, triisopropyl, 18H); 1.377 (s,  $\text{CH}_3$ , 3H); 1.605 (s,  $\text{CH}_3$ , 3H); 3.838–3.927 (m,  $5'\text{-CH}_2$ , 2H); 4.045 (s,  $O^6\text{-CH}_3$ , 3H); 4.308–4.337 (m,  $4'$  H, 1H); 4.84 (broad s,  $2\text{NH}_2$ , 2H); 4.982–5.019 (m,  $3'$  H, 1H); 5.176–5.211 (m,  $2'$  H, 1H); 6.045, 6.055 (d,  $1'$  H, 1H); 7.810 (s, 8H, 1H).  $\text{UV}_{\text{max}}$  ( $\text{CH}_2\text{Cl}_2$ ) 253 and 279 nm. Elemental analysis calculated: C, 55.95; H, 7.96; N, 14.19. Found: C, 56.01; H, 7.98; N, 14.12.

**3',5'-Bis(triisopropylsilyl)thymidine (Compound 3).**  $^1\text{H}$  NMR ( $\text{C}^2\text{HCl}_3$ ,  $25^\circ\text{C}$ ),  $\delta$  1.037–1.370 (m, triisopropyl, 36H); 1.884, 1.879 (d,  $\text{CH}_3$ , 3H); 2.008–2.269 (m,  $2'$  and  $2''$  H, 2H); 3.831–3.962 (m,  $5'\text{-CH}_2$ , 2H); 3.983–4.010 (m,  $3'$  H, 1H); 4.582–4.604 (m,  $4'$  H, 1H); 6.305–6.361 (m,  $1'$  H, 1H); 7.431, 7.436 (d, 6H, 1H); 8.30 (broad s, 3H, 1H). Elemental analysis calculated: C, 60.61; H, 9.81; N, 5.05. Found: C, 60.51; H, 9.67; N, 5.00.

**Experimental Procedures.** Nucleoside concentrations were established by weight and confirmed by UV absorbance using experimentally determined values of  $\epsilon_{281} = 5.5 \times 10^3$  for 2',3'-isopropylidene-5'-triisopropylsilylcytidine and  $\epsilon_{281} = 1.1 \times 10^4$  for 2',3'-isopropylidene-5'-triisopropylsilyl- $O^6$ -methylguanosine in methylene chloride. During mixing experiments, relative nucleoside concentrations were reconfirmed by NMR integrated peak intensities. Perchloric acid (70%; Mallinckrodt) was standardized and used without further purification. Trifluoromethanesulfonic acid (Aldrich) was standardized and stored under argon.  $\text{C}^2\text{HCl}_3$  (99.8% enriched; Stohler Isotope Chemicals, Waltham, MA) was distilled over  $\text{P}_2\text{O}_5$  and stored over 4-Å molecular sieves

(Fisher) in the dark under argon. Flash chromatography was performed with 230- to 400-mesh silica gel (Merck).  $^1\text{H}$  NMR spectra were obtained on Bruker WM 250 and Varian XL-300 NMR spectrometers. Typically, 48 transients were accumulated over 4000–5000 Hz, using 16K, double-precision (32-bit) data points. Chemical shifts are in reference to the  $\text{CHCl}_3$  (7.243 ppm) impurity contained in the deuterated chloroform. The probe temperatures were calibrated with the chemical shifts of methanol (20).

## RESULTS

**$O^6$ -MeGua Forms a Base Pair with Thymine.** Previous spectroscopic attempts to examine base pairing of  $O^6$ -MeGua have met with little success (21) primarily because  $O^6$ -MeGua interacts weakly, if at all, with the four normal bases, and because the experimental conditions were not optimal for examining hydrogen bonding. It is demonstrated here that  $O^6$ -MeGua and thymine (derivatized as shown in Fig. 1) form a non-Watson-Crick-type base pair in nonpolar solution, stabilized by two hydrogen bonds. When 25 mM  $O^6$ -MeGua (as compound 2) is combined with 25 mM thymine (as 3) in  $\text{C}^2\text{HCl}_3$ ,  $^1\text{H}$  resonances of both bases shift downfield in a manner indicative of formation of two hydrogen bonds between the two bases; at  $-5^\circ\text{C}$ , the  $\text{H}^3$  resonance of thymine is deshielded by 0.24 ppm and the  $\text{N}^2$  amino protons of  $O^6$ -MeGua are deshielded by 0.07 ppm (Fig. 2). Qualitatively similar shifts of decreasing magnitude are observed with increasing temperature (data not shown). The stoichiometry of the interaction can be determined from a continuous-variation plot of chemical shift versus mol fraction (Fig. 3). The change in chemical shift of the  $\text{H}^3$  resonance of thymine as a function of mol fraction shows a discontinuity centered around 0.5 mol fraction, indicating a 1:1 interaction of thymine with  $O^6$ -MeGua.

The two hydrogen bond donor sites are readily identified as the  $\text{H}^3$  of thymine and one of the  $\text{N}^2$  amino protons of  $O^6$ -MeGua. Space-filling models show that approach of the  $\text{H}^3$  of thymine to the  $\text{N}^3$  lone pair of  $O^6$ -MeGua is prohibited by steric clash of the non-hydrogen-bonded carbonyl oxygen of thymine with the ribose moiety at the  $\text{N}^9$  position of  $O^6$ -MeGua. Thus, the  $\text{N}^1$  lone pair of  $O^6$ -MeGua is the only hydrogen bond acceptor site that is in proper relative proximity to the  $\text{N}^2$  amino hydrogen bond donor site and available for hydrogen bond formation with the  $\text{H}^3$  proton of thymine. Two types of  $O^6$ -MeGua-thymine dimers fit these constraints:

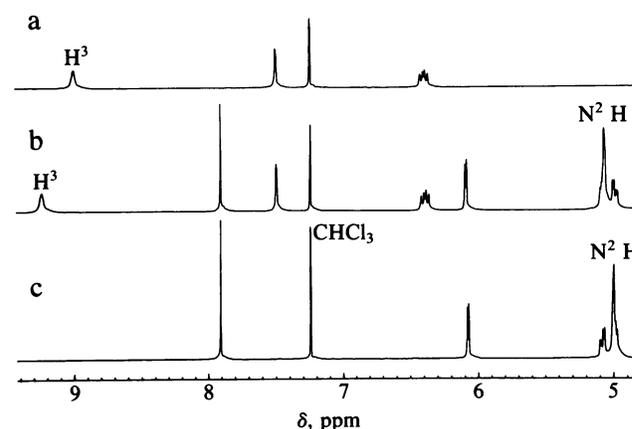


FIG. 2. Partial 250-MHz  $^1\text{H}$  spectra of compound 3 (spectrum a), compound 3 plus compound 2 (spectrum b), and compound 2 (spectrum c). Each nucleoside was 25 mM in deuterated chloroform with a probe temperature of  $-5^\circ\text{C}$ . The nonexchangeable proton assignments are  $\delta$  6.36–6.43 ( $1'$ ) and 7.52 ( $\text{H}^6$ ) in spectrum a and 4.97–5.02 ( $3'$ ), 5.06–5.11 ( $2'$ ), and 7.90 ( $\text{H}^8$ ) in spectrum c.

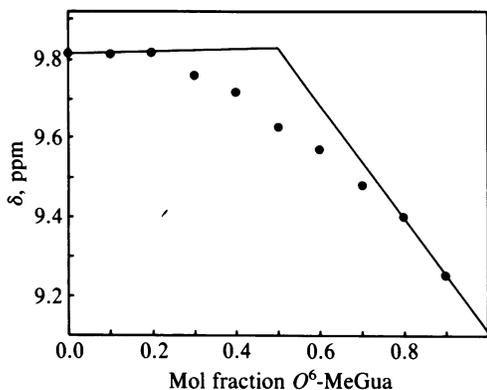


FIG. 3. Mixing curve (mol fraction of the  $O^6$ -MeGua nucleoside 2 vs. chemical shift) of the  $H^3$  resonance of the thymine nucleoside 3 at  $-40^\circ\text{C}$ . The total concentration of the two nucleosides was held constant at 25 mM for each data point.

the normal non-Watson-Crick  $O^6$ -MeGua-thymine base pair (Fig. 4a) and the reverse non-Watson-Crick  $O^6$ -MeGua-thymine base pair (Fig. 4b). Because the  $O^2$  and  $O^4$  of thymine compete as hydrogen bond acceptor sites on the monomer level (22, 23), two hydrogen-bonded structures (as shown in Fig. 4 a and b) would be expected to coexist in significant populations in solution.

**$O^6$ -MeGua Does Not Form a Base Pair with Normal Cytosine.** Solution IR spectroscopic studies indicated that  $O^6$ -MeGua and cytosine do not readily associate to form hydrogen-bonded dimers at room temperature (21). We have extended those studies by employing a combination of relatively low temperature ( $-20^\circ\text{C}$ ) and high concentration (50 mM) in chloroform, a solvent of low dielectric constant ( $= 4.8$ ), in an attempt to coax  $O^6$ -MeGua and cytosine into base-pair formation. Even under these conditions we find no evidence for hydrogen bonding upon mixing of the two bases

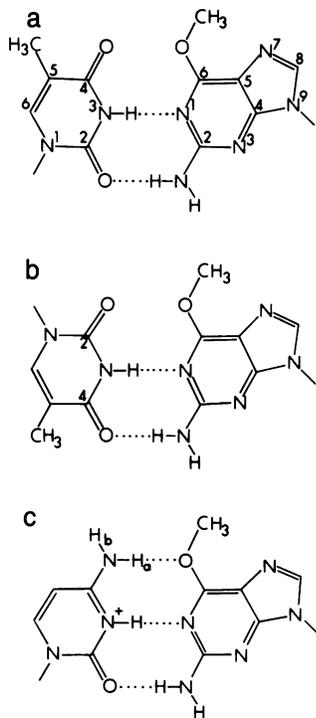


FIG. 4. Likely base pairing schemes of  $O^6$ -MeGua with thymine and with protonated cytosine ( $\text{CytH}^+$ ). (a) Non-Watson-Crick thymine- $O^6$ -MeGua base pair. (b) Reverse non-Watson-Crick thymine- $O^6$ -MeGua base pair. (c)  $\text{CytH}^+$ - $O^6$ -MeGua base pair.

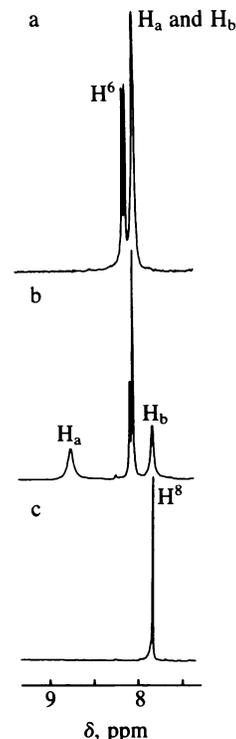


FIG. 5. Partial 250-MHz  $^1\text{H}$  spectra of protonated compound 1 (perchlorate salt) (spectrum a), compound 2 plus the perchlorate salt of 1 (spectrum b), and compound 2 (spectrum c). Each species was 25 mM with a probe temperature of  $25^\circ\text{C}$ .

(data not shown). Although it has been suggested that cytosine could be stabilized in its unfavorable imino tautomer (5, 24) by forming hydrogen bonds with  $O^6$ -MeGua, our data do not support such a hypothesis.

**Protonation of Cytosine Permits Base Pairing with  $O^6$ -MeGua.** Although no observable interaction occurs between  $O^6$ -MeGua and cytosine, we have considered the possibility that, when protonated, cytosine could hydrogen-bond with  $O^6$ -MeGua as depicted in Fig. 4c. When 25 mM  $O^6$ -MeGua is mixed with 25 mM protonated cytosine ( $\text{CytH}^+$ ) in  $\text{C}^2\text{HCl}_3$ , the resulting  $^1\text{H}$  NMR spectrum shows changes in chemical shifts as compared with the spectra of  $O^6$ -MeGua alone and  $\text{CytH}^+$  alone (Fig. 5). At room temperature, dimer formation shifts the signal of one of the  $\text{CytH}^+$   $\text{N}^4$  amino protons (presumably  $\text{H}_a$ ) 0.71 ppm downfield, while the other  $\text{N}^4$  amino proton ( $\text{H}_b$ ) signal moves 0.20 ppm upfield. [The chemical shifts described here are those observed with  $\text{ClO}_4^-$  as the anion; qualitatively similar changes in chemical shifts occur with other anions (e.g.,  $\text{CF}_3\text{SO}_3^-$ ; data not shown).] Other protons not directly involved in hydrogen bonding show more subtle changes upon mixing of  $O^6$ -MeGua with  $\text{CytH}^+$  (see below). Resonances corresponding to protons (other than the  $\text{N}^4$  amino protons of  $\text{CytH}^+$ ) directly involved in hydrogen-bond interactions are exchange-broadened and unobservable in these experiments.

Again, the stoichiometry of interaction has been derived from continuous-variation plots. Changes in chemical shifts as a function of mol fraction  $O^6$ -MeGua or  $\text{CytH}^+$  (Fig. 6) indicate that there is a 1:1 interaction of  $O^6$ -MeGua with  $\text{CytH}^+$ . Least-squares analyses of the two linear regions of each plot (for  $\text{H}_a$  and  $\text{H}_b$  protons of cytosine;  $\text{H}^8$  and  $O^6$ -methyl protons of guanine, in Fig. 6) yield two lines intersecting close to mol fraction 0.5, as expected for the 1:1 interaction



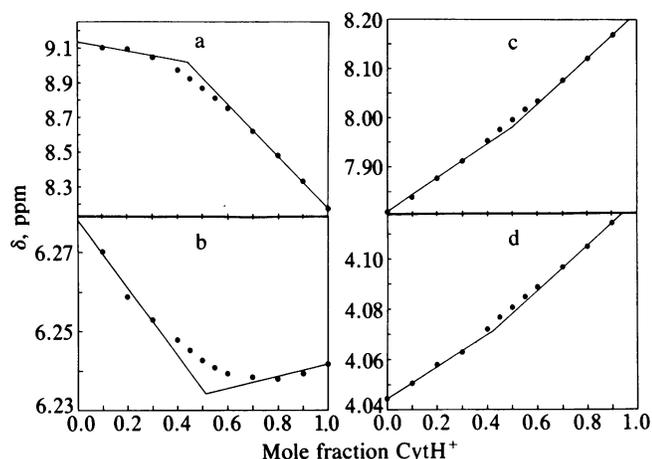


FIG. 6. Mixing curves of four proton resonances of the perchlorate salt of compound 1 (CytH<sup>+</sup>) and compound 2 (O<sup>6</sup>-MeGua) at 25°C: (a) H<sub>a</sub> resonance of CytH<sup>+</sup>. (b) H<sub>b</sub> resonance of CytH<sup>+</sup>. (c) H<sup>8</sup> resonance of O<sup>6</sup>-MeGua. (d) O<sup>6</sup>-methyl resonance of O<sup>6</sup>-MeGua. The total concentration of the two nucleosides was held constant at 25 mM for each data point.

The possibility that the interaction of O<sup>6</sup>-MeGua with CytH<sup>+</sup> is merely proton transfer from CytH<sup>+</sup> to O<sup>6</sup>-MeGua merits comment. The observed changes in chemical shifts are not consistent with simple proton transfer from CytH<sup>+</sup> to O<sup>6</sup>-MeGua. For example, addition of 1.0 equivalent of perchloric acid to 25 mM O<sup>6</sup>-MeGua resulted in a 0.13-ppm downfield shift of the sugar H<sup>1</sup> of O<sup>6</sup>-MeGua (data not shown), whereas the chemical shift of this hydrogen was invariant upon addition of 1.0 equivalent of CytH<sup>+</sup> ClO<sub>4</sub><sup>-</sup>. Additional evidence arguing against simple proton transfer is the observation that the formation of the half-protonated cytosine dimer essentially excludes O<sup>6</sup>-MeGua from interaction with CytH<sup>+</sup> (data not shown). If proton transfer from CytH<sup>+</sup> to O<sup>6</sup>-MeGua were to occur, it would result in 2:1 stoichiometry:



The mixing curves in Fig. 6 clearly show this not to be the case; the data are inconsistent with proton transfer from CytH<sup>+</sup> to O<sup>6</sup>-MeGua as in Eq. 2 but consistent with a 1:1 interaction as in Eq. 1.

We conclude that, in a nonpolar environment, an ionized base pair of O<sup>6</sup>-MeGua with CytH<sup>+</sup> is stabilized by three hydrogen bonds as shown in Fig. 4c. Protonation of cytosine confers complementarity to the arrangement of three hydrogen bond donor and acceptor sites of O<sup>6</sup>-MeGua with three hydrogen bond donor and acceptor sites of cytosine. This interaction requires protonation of the cytosine N<sup>1</sup> position, the most basic site on the four DNA bases. Changes in chemical shifts upon mixing indicate that the protonation-dependent 1:1 interaction of O<sup>6</sup>-MeGua with cytosine involves a hydrogen bond between an acceptor site on O<sup>6</sup>-MeGua and only one of the N<sup>4</sup> amino protons of CytH<sup>+</sup>, in analogy with the hydrogen bonding within the normal Watson-Crick C-G base pair.

## DISCUSSION

Physicochemical investigations of the properties of nucleic acid monomers have provided considerable insight into structure-function relationships of nucleic acids in general (12-16). Some important details regarding DNA and RNA were first elucidated through the study of monomeric nucleic acid bases, nucleosides, and nucleotides. The Hoogsteen base pair formed

between adenine and thymine was first observed in crystals of bases (25) many years before crystals of tRNA showed that similar patterns of hydrogen bonding can exist in polymers (ref. 26; for review, see ref. 27). The present report outlines our investigations into the hydrogen-bonding interactions of O<sup>6</sup>-MeGua with the potentially complementary bases cytosine and thymine. The poor solubility properties of guanine, cytosine, and derivatives of both bases have long been an obstacle to such investigations. By protecting ribose hydroxyls with highly lipophilic, nonplanar, nonaromatic triisopropylsilyl groups, we have prepared derivatives of guanosine and cytidine that possess high solubility in chloroform. We have found triisopropylsilyl derivatives of nucleosides simple to prepare, easy to purify, and stable under our experimental conditions. They are ideal for NMR studies defining hydrogen-bonding interactions between bases.

To summarize, we have found with <sup>1</sup>H NMR that O<sup>6</sup>-MeGua is able to pair with thymine (as depicted in Figs. 4a and b) in a hydrophobic environment. With these studies we have been able to provide direct evidence for the nature of the hydrogen-bonding interactions in the O<sup>6</sup>-MeGua-thymine aberrant base pair. The hydrogen bond donors are H<sup>3</sup> of thymine and one of the N<sup>2</sup> amino protons of O<sup>6</sup>-MeGua.

Further, we have shown that O<sup>6</sup>-MeGua can pair in nonpolar solvents with cytosine, but only when the cytosine is protonated. The structure most consistent with the <sup>1</sup>H NMR data is depicted in Fig. 4c, where the protonated base pair is stabilized by three hydrogen bonds. We find no evidence for the existence of wobble base pairs or imino tautomers of either cytosine or O<sup>6</sup>-MeGua.

There are several reasons why a protonated CytH<sup>+</sup>·O<sup>6</sup>-MeGua base pair might be found in DNA, as it is among our model bases. First, in molecular-dynamic calculations the dielectric constant within the helix is estimated to be small (28, 29); i.e., the helix interior is a nonpolar environment. Hydrogen bonding between bases in the helix interior might then be modeled in solvents of low dielectric constant like chloroform (dielectric constant = 4.8). Second, precedent for the ionized CytH<sup>+</sup>·O<sup>6</sup>-MeGua base pair within the helix lies in the analogous Cyt·CytH<sup>+</sup> base pair, which has been shown to exist predominantly in the protonated state in oligonucleotides at pH 7.0 (30). The positive free energy of protonation is evidently offset by negative free energy contributions from stacking and hydrogen bonding, resulting in formation of the thermodynamically stable base pair. Chen (31) and Pulleyblank *et al.* (32) have also obtained evidence, albeit indirect, for protonated base pairs in the helix. Third, on the basis of acid-base theory, one would predict that the frequency of protonation of cytosine when opposite O<sup>6</sup>-MeGua would be significantly greater than for cytosine that is free or paired to normal guanine (35). The N<sup>1</sup> site of guanine is a hydrogen bond proton donor but becomes a hydrogen bond proton acceptor upon O<sup>6</sup> methylation of guanine. The resulting repulsive interaction between the N<sup>1</sup> lone pair of O<sup>6</sup>-MeGua and the N<sup>3</sup> lone pair of unprotonated cytosine would be transformed into an attractive interaction upon protonation of cytosine. As proposed by Sowers *et al.* (35), the cytosine residue would thus be protonated with an increased probability when paired with a guanine that has been methylated at the O<sup>6</sup> position. Conversely, the pK<sub>a</sub> of the protonated base pair will be higher than that of unpaired protonated cytosine. As a precedent, the pK<sub>a</sub> of protonated cytosine increases from about 4.5 to around 7.5 upon formation of the hemiprotonated cytosine dimer (30). Because protonated cytosine (with pK<sub>a</sub> of 4.5) is less acidic than protonated O<sup>6</sup>-MeGua (pK<sub>a</sub> = 2.3) (2), the proton is expected to reside primarily on cytosine in the CytH<sup>+</sup>·O<sup>6</sup>-MeGua base pair.

The finding that protonated cytosine forms a base pair with O<sup>6</sup>-MeGua allows us to make a number of predictions.

(i) It explains the apparent dilemma that thymine pairs most weakly with  $O^6$ -MeGua in DNA oligomers (9) yet interacts preferentially in polymerase complexes (4–8). In both our  $^1\text{H}$  NMR experiments and in polymerase-directed reactions (4–8), thymine pairs preferentially over cytosine with  $O^6$ -MeGua. Assuming that RNA and DNA polymerases incorporate the physiologically predominant (un-ionized) form of each base, the preferential incorporation of thymine opposite  $O^6$ -MeGua is consistent with the pairing properties of the unprotonated bases we have observed in  $\text{C}^2\text{HCl}_3$ . However, the cytosine- $O^6$ -MeGua base pair, once established in the duplex by either methylation of DNA or by replication/transcription, may then be stabilized by protonation. We expect these results may be applicable to other modified base pairs in DNA.

(ii) Therefore, it may be a general rule that certain mechanisms such as ionization can stabilize hydrogen bonding interactions between bases in DNA but *not* in polymerase complexes.

(iii) An ionized base pair would be expected to alter DNA helix conformation, hydration, and dynamics.

(iv) Protonated base pairs like our proposed  $\text{CytH}^+\cdot O^6$ -MeGua base pair may constitute a previously unrecognized pathway for replication-independent mutagenesis. The protonated cytosine monomer is known to deaminate to uracil with a half-life of 17 days at  $37^\circ\text{C}$  (33, 34); however, cytosine in double-stranded DNA is normally unprotonated and protected from deamination [having a half-life  $>15,000$  years (L. Frederico, B.R.S. and T. Kunkel, unpublished data)]. Any damage that would increase the probability of finding a protonated cytosine in DNA could lead to increased levels of cytosine deamination. Methylation at  $O^6$  of guanine on one strand would favor protonation of cytosine on the opposite strand (35) and thereby lead to increased rates of cytosine deamination. We call this process *cross-strand deamination*. Whether formed by *in situ* alkylation or incorporation during DNA replication, a cytosine- $O^6$ -MeGua base pair would be stabilized by protonation. Yet, because it is protonated, the cytosine partner would be deaminated with an increased rate. The longer an unrepaired  $O^6$ -MeGua base remains in the double helix, the greater the probability of C $\rightarrow$ U transition mutations on the opposite strand. Likewise, alkylation of a guanine paired to a 5-methylcytosine residue is expected to increase the frequency of protonation and cross-strand deamination, resulting in a thymine- $O^6$ -MeGua base pair. In both cases the cell would be faced with repairing a double lesion—e.g., either a uracil- $O^6$ -MeGua or a thymine- $O^6$ -MeGua base pair.

We are grateful to Dr. Raymond Tsang and Ms. Nidhi Gupta for helpful comments during the progress of this work and to Robert R. Lee for technical assistance. This work was supported by grants to B.R.S. from the Duke University Research Council and the American Cancer Society (Grant NP-486).

1. Loveless, A. (1969) *Nature (London)* **223**, 206–207.
2. Singer, B. & Grunberger, D. (1983) *Molecular Biology of*

- Mutagens and Carcinogens* (Plenum, New York).
3. Goth, R. & Rajewsky, M. F. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 639–643.
4. Snow, E. T., Foote, R. S. & Mitra, S. (1984) *J. Biol. Chem.* **259**, 8095–8100.
5. Toorchen, D. & Topal, M. D. (1983) *Carcinogenesis* **4**, 1591–1597.
6. Eadie, J. S., Conrad, M., Toorchen, D. & Topal, M. D. (1984) *Nature (London)* **308**, 201–203.
7. Gershman, L. L. & Ludlum, D. B. (1973) *Biochim. Biophys. Acta* **308**, 310–316.
8. Mehta, J. R. & Ludlum, D. B. (1978) *Biochim. Biophys. Acta* **562**, 770–778.
9. Gaffney, B. L., Markey, L. A. & Jones, R. A. (1984) *Biochemistry* **23**, 5686–5691.
10. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L. & Jones, R. A. (1986) *Biochemistry* **25**, 1027–1035.
11. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L. & Jones, R. A. (1986) *Biochemistry* **25**, 1036–1042.
12. Katz, L. & Penman S. (1966) *J. Mol. Biol.* **15**, 220–231.
13. Kyogoku, Y., Lord, R. C. & Rich, A. (1966) *Science* **154**, 518–520.
14. Pitha, J., Jones, R. N. & Pithova, P. (1966) *Can. J. Chem.* **44**, 1045–1050.
15. Newmark, R. A. & Cantor, C. R. (1968) *J. Am. Chem. Soc.* **90**, 5010–5017.
16. Ts'o, P. O. P. (1974) in *Basic Principles in Nucleic Acid Chemistry*, ed. Ts'o, P. O. P. (Academic, New York), Vol. 1, pp. 454–584.
17. Tomasz, J. (1978) in *Nucleic Acid Chemistry*, eds. Townsend, L. B. & Tipson, R. S. (Wiley, New York), Part 2, pp. 765–769.
18. Ogilvie, K. K., Thompson, E. A., Quilliam, M. A. & Westmore, J. B. (1974) *Tetrahedron Lett.* **33**, 2865–2868.
19. Gaffney, B. L. & Jones, R. A. (1982) *Tetrahedron Lett.* **23**, 2253–2256.
20. Van Geet, A. L. (1970) *Anal. Chem.* **42**, 679–690.
21. Psoda, A., Kierdaszuk, B., Pohorille, A., Geller, M., Kusmierk, J. T. & Shugar, D. (1981) *Int. J. Quantum Chem.* **20**, 543–554.
22. Watanabe, M., Sugeta, H., Iwahashi, H., Kyogoku, Y. & Kainosho, M. (1980) *Eur. J. Biochem.* **117**, 553–558.
23. Iwahashi, H., Sugeta, H. & Kyogoku, Y. (1982) *Biochemistry* **21**, 631–638.
24. Topal, M. D. & Fresco, J. R. (1976) *Nature (London)* **263**, 2285–2289.
25. Hoogsteen, K. (1963) *Acta Crystallogr.* **16**, 907–916.
26. Sussman, J. L., Holbrook, S. R., Warrant, R. W. & Kim, S.-H. (1978) *J. Mol. Biol.* **123**, 607–630.
27. Saenger, W. (1983) *Principles of Nucleic Acid Structure* (Springer, New York).
28. Levitt, M. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 251–262.
29. Singh, U. C., Weiner, S. J. & Kollman, P. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 755–759.
30. Gray, M. D. & Cui, T. (1984) *Nucleic Acids Res.* **12**, 7565–7580.
31. Chen, F.-M. (1984) *Biochemistry* **23**, 6159–6165.
32. Pulleyblank, D. E., Haniford, D. B. & Morgan, A. R. (1985) *Cell* **42**, 271–280.
33. Lawley, P. D. & Brookes, P. (1963) *Biochem. J.* **89**, 127–138.
34. Leutzinger, E. E., Miller, P. S. & Kan, L.-S. (1982) *Biochim. Biophys. Acta* **697**, 243–251.
35. Sowers, L. C., Shaw, B. R., Veigl, M. L. & Sedwick, D. R. (1987) *Mutat. Res.*, in press.