First Principles Calculations of the pK_a Values and Tautomers of Isoguanine and Xanthine

Katherine Noyes Rogstad,[†] Yun Hee Jang,[‡,§] Lawrence C. Sowers,[†,‡] and William A. Goddard, III*[,‡]

Department of Biochemistry and Microbiology, Loma Linda University School of Medicine, Loma Linda, California 92350, Materials and Process Simulation Center, Beckman Institute (139-74), California Institute of Technology, Pasadena, California 91125, and School of Chemistry, Seoul National University, Seoul 151-747, Korea

Received April 3, 2003

The accurate replication of DNA requires the formation of complementary hydrogen bonds between a template base and the base moiety of an incoming deoxynucleotide-5′-triphosphate. Recent structural studies suggest that some DNA polymerases contribute additional constraints by interrogating the minor groove face of the incoming and template bases. Therefore, the hydrogen bond-donating or -accepting properties of the base pairing as well as minor groove faces of the bases could be important determinants of correct base selection. In this paper, we investigate two purines that could arise by endogenous damage of the normal DNA bases: isoguanine (which can be generated by the oxidation of adenine) and xanthine (which can be generated by the deamination of guanine). In both cases, the potential exists for the placement of a proton in the N3 position, converting the N3 position from a hydrogen bond acceptor to a donor. In this paper, we use first principles quantum mechanical methods (density functional theory using the B3LYP functional and the 6-31G++G**basis set) to predict the ionization and tautomeric equilibria of both isoguanine and xanthine in the gas phase and aqueous solution. For isoguanine, we find that the N1H and N3H neutral tautomeric forms are about equally populated in aqueous solution, while the enol tauotomers are predominant in the gas phase. In contrast, we find that xanthine displays essentially no tautomeric shifts in aqueous solution but is nearly equally populated by both an anionic and a neutral form at physiological pH. To obtain these results, we carried out an extensive examination of the tautomeric and ionic configurations for both xanthine and isoguanine in solution and in the gas phase. The potential hydrogen-bonding characteristics of these damaged purines may be used to test predictions of the important components of base selection by different DNA polymerases during DNA replication.

Introduction

The DNA of all organisms must be replicated with high fidelity to maintain the integrity of the genome. In higher organisms, overall fidelity is attributed to a combination of processes including correct selection of the incoming deoxynucleoside-5′-triphosphate by the DNA polymerase, preferential extension of correctly paired primer−template termini, and postreplication mismatch repair (*1*−*6*). The complexity of this task is increased substantially in metabolically active cells in which the DNA bases are continuously damaged by oxidation and hydrolysis (*7*−*11*). Repair systems do exist to remove damaged bases prior to DNA replication (*12*). However, if replication occurs prior to repair, proper selection of an incoming deoxynucleoside-5′-triphosphate may no longer be possible.

Among the multitude of damaged DNA bases that occur daily in essentially all living cells are isoG[1] (also known as 2-hydroxyadenine; Figure 1) and X (Figure 2). isoG is derived from the oxidation of adenine, whereas X is derived from both the hydrolytic and the oxidative deamination of guanine (*11*, *13*, *14*), although the oxidative nitrous acid-mediated pathway (*11*) is likely to be more important due to the slow rates of hydrolysis of purines (*15*). Common to both purines is that the chemical damage event could place a proton at the N3 position of the purine that is normally a hydrogen bond acceptor. Recent structural studies of DNA polymerases, in particular DNA pol β, have suggested that interactions between the polymerase and the purine N3 position can augment proper base selection by ensuring the correct alignment of the bases in the polymerase active site (*16*−*25*).

The damaged purines, isoG and X, could exist in several potential tautomeric or ionized forms (*26*−*32*). Previous studies have indicated the importance of both ionization and tautomerization in the generation of base

* To whom correspondence should be addressed. E-mail: wag@wag.caltech.edu.
[†] Loma Linda University School of Medicine.
[‡] California Institute of Technology.
[§] Seoul National University.

[1] Abbreviations: isoG, isoguanine; X, xanthine; QM, quantum mechanics; DFT, density functional theory; KF, Klenow fragment of DNA polymerase I from *E. coli*; Taq polymerase, DNA polymerase from *Thermus aquaticus*; isodGTP, 2′-deoxyisoguanosine triphosphate; isoC, isocytidine; HIV1-RT, human immunodeficiency virus 1 reverse transcriptase; κ, 2,6-diaminopyrimidine; dκTP, the deoxynucleoside triphosphate analogue of κ.

10.1021/tx034068e CCC: $25.00
Published on Web 10/25/2003

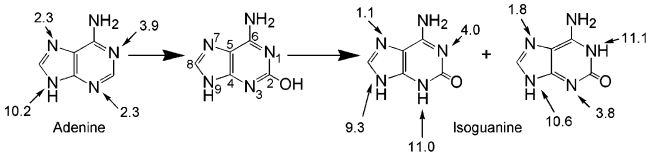


Figure 1. Oxidation of adenine to form the two predominant neutral tautomers of isoG. The calculated microscopic pK_a values of each site are indicated.

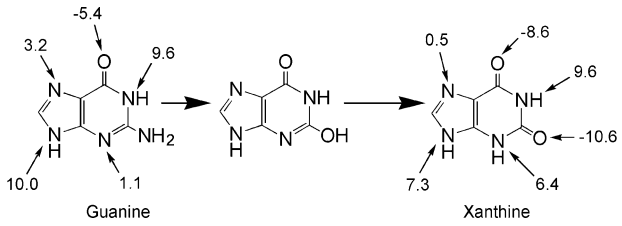


Figure 2. Deamination of guanine to form X. The calculated microscopic pK_a values of each site are indicated.

mispairs (*33*−*38*). In this study, the tautomeric and ionization equilibria of both X and isoG have been studied systematically using first principles QM methods. Such QM studies provide information that is very difficult to extract from experiment, allowing us to enhance substantially our understanding of potential equilibrium structures that may contribute to the fidelity of DNA replication and to explain why some modifications decrease replication fidelity.

Recently, we used first principles QM (B3LYP/6-31G++G**) to study the pK_a values and tautomeric equilibria of a number of oxidized and other DNA bases (*39*−*42*). Our methods have been validated by excellent agreement with available experimental data. Furthermore, we have been able to calculate some values that are experimentally inaccessible, making the QM-based approach extremely useful in the study of biologically relevant chemical compounds. Here, we show that isoG exists as two neutral, keto tautomers (N1H and N3H forms), whereas X exists as a diketo neutral tautomer and a singly ionized form deprotonated at the N3 position. The site specific pK_a values of isoG and X as well as their composite pK_a values are presented, where we find an excellent correlation with available experimental data. The interactions of isoG and X derivatives with DNA polymerase β and other polymerases are discussed in light of the enzyme's interactions with the N3 position of the minor groove of the DNA bases.

Computational Methods

The relative free energies of 88 tautomers of isoG and 108 tautomers of X (neutral, cationic, anionic, doubly anionic, and 9-methyl derivatives) were examined in both the gas phase and the aqueous phase using DFT calculations at the B3LYP/6-31++G** level of theory. The effect of solvation by water was taken into account using the Poisson−Boltzmann (PB) continuum solvation model. This same methodology was used in our previous work (*41*, *42*) for the calculation of the relative energies of tautomers and pK_a values of guanine and 8-oxoguanine in aqueous solution; hence, we will not discuss the details further here.

Free Energy in Solution. The standard free energy of a species in water (ΔG^0_{aq}) can be written as the sum of the gas phase standard free energy (ΔG^0_g) and the standard free energy of solvation in water (ΔG^0_{solv}):

$$\Delta G^0_{aq} = \Delta G^0_g + \Delta G^0_{solv} \quad (1)$$

The gas phase standard free energy of each species (ΔG^0_g) is obtained by

$$\Delta G^0_g = E_{0\,K} + \mathrm{ZPE} + \Delta\Delta G_{0\rightarrow 298\,K} \quad (2)$$

The total energy of the molecule at 0 K ($E_{0\,K}$) is calculated at the optimum geometry from QM. The zero-point energy (ZPE) and the Gibbs free energy change from 0 to 298 K ($\Delta\Delta G_{0\rightarrow 298\,K}$) are calculated from the vibrational frequencies obtained using QM. The translational and rotational free energy contributions are also calculated for the gas phase. All QM calculations used the DFT (B3LYP) (*45*−*49*) implemented in Jaguar v4.0 quantum chemistry software (*43*, *44*). Because frequency calculations are much more time consuming, the calculation was done in two steps. The 6-31G** basis set was used to optimize the geometry and to calculate the vibrational frequencies. Then, the 6-31++G** basis set was used for a final geometry optimization starting with the optimum 6-31G** geometry.

The standard free energy of solvation in water (ΔG^0_{solv}) is calculated using the continuum solvation approach (*50*−*52*) by solving the PB equation numerically (*53*). The solute is described as a low dielectric cavity ($\epsilon_{solute} = 1$) immersed in a high dielectric continuum of solvent ($\epsilon_{H_2O} = 80$ for water) (*54*). The solute/solvent boundary is described by the surface of closest approach as a sphere of radius 1.4 Å (probe radius for water) rolled over the van der Waals (vdW) envelope of the solute. The atomic radii used to build the vdW envelope of the solute were taken from our previous work on guanine (*42*). These are reduced by 6% from the values of Marten and co-workers (*52*): 1.88 (sp^2 C), 1.79 (sp^3 C), 1.46 (sp^2 O), 1.41 (sp^2 N), 1.18 (H attached to sp^2 C), and 1.08 Å (other H's). These calculations were carried out using Jaguar v4.0 (*43*, *44*) at the B3LYP/6-31++G** level, and the geometry was reoptimized in solution.

Summarizing, the calculation of each species was done in three steps: (i) step 1, B3LYP/6-31G**(g) preliminary geometry optimization and frequency calculation; (ii) step 2, B3LYP/6-31++G**(g) for the final geometry optimization; and (iii) step 3, B3LYP/6-31++G**(aq) for solution phase geometry optimization.

The final standard free energy of each species in water is expressed as

$$\Delta G^0_{aq} = \mathrm{ZPE}^{6-31G^{**}} + \Delta\Delta G_{0\rightarrow 298\,K}{}^{6-31G^{**}} + E_{0\,K,g}{}^{6-31++G^{**}} + \Delta G^0_{solv}{}^{6-31++G^{**}} \quad (3)$$

pK_a Calculation. The pK_a of an acid HA is given by (*55*, *56*):

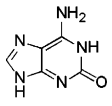
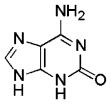
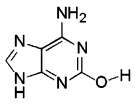
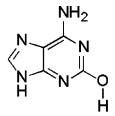
$$\mathrm{p}K_a = \frac{1}{2.303RT}\Delta G^0_{deprot,aq} \quad (4)$$

where R is the gas constant and T is the temperature. The standard free energy of deprotonation of HA in water ($\Delta G^0_{deprot,aq}$) is defined as

$$\Delta G^0_{deprot,aq} = \Delta G^0_{aq}(A^-) + \Delta G^0_{aq}(H^+) - \Delta G^0_{aq}(HA) \quad (5)$$

The standard free energies of HA and A^- in water

Table 1. Relative Free Energies (ΔΔ*G* in kcal/mol) and Populations (*f*) of Tautomers of Neutral isoG in the Gas Phase and Aqueous Solution

	 **1**	 **2**	 **3**	 **4**
$\Delta\Delta G^g$	5.5	4.9	0.4	0.0
f^g	6×10^{-5}	2×10^{-4}	0.34	0.66
$\Delta\Delta G^{aq}$	0.0	0.2	6.7	6.8
f^{aq}	0.51	0.38	6×10^{-6}	5×10^{-6}

[ΔG^0_{aq}(HA) and ΔG^0_{aq}(A^-)] were calculated as described in the previous section, Free Energy in Solution. For a proton, the standard free energy in water [ΔG^0_{aq}(H^+)] was estimated according to eq 1 from its gas phase free energy taken from literature (*55*, *56*)

$$\Delta G^0_g(H^+) = 2.5\,RT - T\Delta S^\circ = 1.48 - 7.76 = -6.28 \text{ kcal/mol} \quad (6)$$

The standard free energy of solvation in water [ΔG^0_{solv}(H^+)] is chosen as −263.47 kcal/mol, which gave the best match between calculated pK_a values of guanine and experimental ones in our previous work (*42*). The experimental value remains uncertain (−254 to −264 kcal/mol) (*55*, *57*−*59*).

pK_a Calculation in the Presence of Multiple Tautomers. In the presence of multiple tautomers, the overall macroscopic pK_a value is estimated as (*42*)

$$pK_a = pK_a^{ij} - \log f_i + \log f_j' \quad (7)$$

from a site specific value (pK_a^{ij}) corresponding to deprotonation from the *i*-th tautomer to the *j*-th tautomer at the deprotonated state, which is calculated according to eqs 4 and 5. Here, f_i is the relative population of the *i*-th tautomer before deprotonation and f_j' is the relative population of the *j*-th tautomer after deprotonation. These relative populations are calculated from the Boltzmann distribution based on the relative free energies of those tautomers.

Results

Isoguanine. Upon the oxidation of adenine to isoG, a hydroxyl substituent is generated at the 2-position (Figure 1). In this work, we examined the many possible tautomeric forms of this compound. The most stable form in aqueous solution is the 6-amino, 2-keto configuration with a proton added at the N1 position (N1H tautomer, **1**; Table 1). The N1 position is normally a hydrogen-bonding site. Therefore, this oxidative modification would perturb formation of a normal Watson−Crick A−T base pair. The N3H tautomer **2** is of essentially identical energy, making the N1H and N3H tautomers similarly populated in aqueous solution. The N3 position is examined by DNA pol β during DNA replication (*16*−*21*, *23*, *24*); therefore, formation of this tautomeric form could interfere with the purine−polymerase interaction.

The presence of the 2-keto substituent of isoG also introduces an additional acidic proton. In the parent purine, adenine, the most acidic proton among the N6 and N9 protons is the N9 proton with a calculated site specific pK_a of 10.2 (Figure 1). In isoG, among the N1/N3, N6, and N9 protons, the N9 proton is similarly the most acidic with a calculated pK_a of 9.3 or 10.6 (for the N3H and N1H tautomers, respectively; Figure 1). The N1/N3 proton could be ionized in more basic solution with a pK_a of 11.

The most basic site among the N1, N3, and N7 sites of adenine is N1, which accepts a proton with a pK_a of 3.9. The N1/N3 positions of isoG are similarly the most basic with a calculated pK_a values of 4.0 and 3.8 (for the N3H and N1H tautomers, respectively).

The most substantial effect due to the conversion of adenine to isoG is the presence of an additional proton that could interfere with base pair formation (N1H) or minor groove interaction with the polymerase (N3H). In no case did we find evidence that ionization of isoG would substantially shift its tautomeric equilibrium.

Xanthine. The deamination of guanine with the formation of X replaces the 2-amino group with a hydroxyl function (Figure 2). As has been observed previously with guanine and 8-oxoguanine (*41*, *42*), the most stable configurations of X in water are the keto configurations, **x1** and **x6**, (N9H and N7H tautomers, respectively; Table 2). The conversion from keto to enol carries an energetic penalty of at least 10 kcal/mol, making any of the enol tautomeric forms highly unlikely under physiological conditions. The tautomeric form in which the N9 proton shifts to N7 (N7H tautomer; **x6**) is slightly more stable, even though the energy difference is very small. Small energy differences between the N7H and the N9H tautomers have been seen previously for the purines, guanine and 8-oxoguanine (*41*, *42*).

The most basic site of the parent purine, guanine, is the N7 position when a proton is held at the N9 position (Figure 2). However, the pK_a drops dramatically from 3.2 (N7 of guanine) to less than 1 (N7 of X) upon the deamination of guanine (Figure 2).

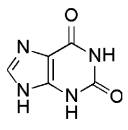
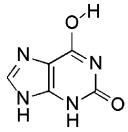
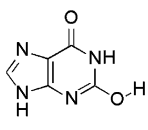
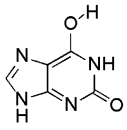
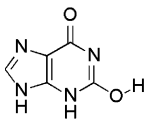
The most acidic proton of guanine is the N1 proton with a pK_a of 9.6, followed by the N9 proton with a pK_a of 10.0 (Figure 2). X acquires an additional acidic proton relative to guanine. The most acidic proton among the N1, N3, and N9 protons of X is the N3 proton with a calculated pK_a of 6.4. The N9 proton is also acidic, with a pK_a of 7.3. Upon deamination of guanine to X, the acidity of the N1 proton (which is the most acidic proton of guanine) remains constant at ∼9.6. In contrast, the N9 proton becomes substantially more acidic; the pK_a drops from 10.0 to 7.3.

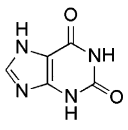
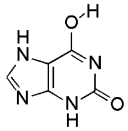
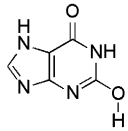
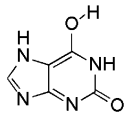
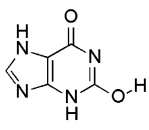
The most abundant configuration of X in aqueous solution is the species ionized (that is, deprotonated) at N3 (**x11**; Table 3). We examined whether the N1- and N3-ionized forms might be of similar energy so that they would be similarly populated. Above, in the case of isoG, the N1H and N3H neutral tautomers (**1** and **2**) were of essentially identical energy. In contrast, for the case of X, ionization at the N1 position leading to **x12** costs 4.3 kcal/mol more than ionization at the N3 position leading to **x11**, making N3 ionization the predominant form at physiological pH.

Discussion

Over the years, several determinants of DNA replication fidelity and sources of mispairing have been discussed, including the rare tautomer model (*60*−*62*), base ionization (*33*, *34*, *36*, *63*), thermodynamic discrimination (*4*, *64*, *65*), and geometric selection (*5*, *6*, *23*). Most

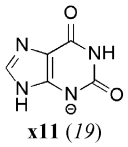
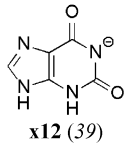
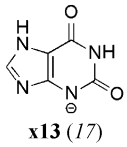
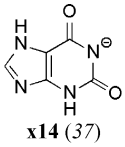
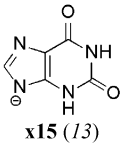
Table 2. Relative Free Energies (ΔΔ*G* in kcal/mol) and Populations (*f*) of Tautomers of Neutral X in the Gas Phase and Aqueous Solution[a]

	 **x1** (*139*)	 **x2** (*36ri9*)	 **x3** (*12u9*)	 **x4** (*16ri9*)	 **x5** (*2u39*)
$\Delta\Delta G^{g}$	9.1	21.3	18.7	25.8	29.2
f^{g}	2×10^{-7}	2×10^{-16}	2×10^{-14}	1×10^{-19}	4×10^{-22}
$\Delta\Delta G^{aq}$	0.6	12.7	10.9	14.6	14.7
f^{aq}	0.258	4×10^{-10}	7×10^{-9}	2×10^{-11}	1×10^{-11}

	 **x6** (*137*)	 **x7** (*36ri7*)	 **x8** (*12d7*)	 **x9** (*16ri7*)	 **x10** (*2u37*)
$\Delta\Delta G^{g}$	0.0	16.1	11.5	33.0	17.1
f^{g}	1.0	2×10^{-12}	4×10^{-9}	6×10^{-25}	3×10^{-13}
$\Delta\Delta G^{aq}$	0.0	12.3	10.4	16.1	13.3
f^{aq}	0.741	8×10^{-10}	2×10^{-8}	1×10^{-12}	1×10^{-10}

[a] The structure numbers italicized in parentheses refer to the naming scheme used in the Supporting Information.

Table 3. Relative Free Energies (ΔΔ*G* in kcal/mol) and Populations (*f*) of Tautomers of Ionized X in the Gas Phase and Aqueous Solution[a]

	 **x11** (*19*)	 **x12** (*39*)	 **x13** (*17*)	 **x14** (*37*)	 **x15** (*13*)
$\Delta\Delta G^{g}$	5.7	27.3	9.3	16.0	0.0
f^{g}	7×10^{-5}	1×10^{-20}	2×10^{-7}	2×10^{-12}	1.00
$\Delta\Delta G^{aq}$	0.0	4.3	1.2	3.2	1.1
f^{aq}	0.78	5×10^{-4}	0.097	0.003	0.12

[a] The structure numbers italicized in parentheses refer to the naming scheme used in the Supporting Information.

recently, it has been reported that DNA pol β interrogates the minor groove of the incoming base and its template partner (*16*−*21*, *23*). Because all four normal DNA bases have a hydrogen bond acceptor at identical positions in the minor groove where amino acid contact is made (*66*), pol β could reject as a substrate any structures deviant at the N3 position of purines or the O2 position of pyrimidines. The tautomeric equilibria and pK_a values of isoG and X, which deviate from the normal bases at the N3 position, are presented and discussed in light of their interactions with various polymerases.

Isoguanine. Our calculations show that in aqueous solution isoG exists in equilibrium between two tautomers (N1H and N3H tautomers, **1** and **2**) having populations of 51% **1** and 38% **2**, whereas the enol forms (**3** and **4**) predominate in the gas phase. These results are in agreement with the existence of an N3H tautomer as previously proposed (*67*−*69*). The predominance of enol forms in the gas phase is not surprising as previous studies have observed increasing proportions of enol isoG in increasingly nonpolar solvents (*27*, *30*).

The pK_a values of 2′-deoxyisoguanosine have been determined as 3.45 and 9.80 (*30*), for 9-methylisoguanine as 3.85 and 9.9 (*27*), and for isoG as 4.51 and 8.99 (*70*, *71*), comparable to our calculated macroscopic pK_a values of 3.9 (pK_{a1}; cationic to neutral) and 9.6 (pK_{a2}; neutral to anionic) (Table 4). These values render isoG as predominantly neutral at physiological pH.

Table 4. Calculated and Experimental Macroscopic pK_a Values of Adenine, Guanine, isoG, and X

compd	pK_a (expt)	pK_a (calcd)	ref
adenine	pK_{a1} < 0	pK_{a1} = −1.5	*71*, *91* (expt)
	pK_{a2} = 4.22	pK_{a2} = 4.2	
	pK_{a3} = 9.8	pK_{a3} = 10.2	
guanine	pK_{a1} = 3.3	pK_{a1} = 3.2	*70*, *71* (expt), *42* (calcd)
	pK_{a2} = 9.2	pK_{a2} = 9.4	
	pK_{a3} = 12.3	pK_{a3} = 12.6	
isoG	pK_{a1} = 4.5	pK_{a1} = 3.9	*70*, *71* (expt)
	pK_{a2} = 9.0	pK_{a2} = 9.6	
X	pK_{a1} ∼ 0.8	pK_{a1} = −0.1	*71*, *81*, *82* (expt)
	pK_{a2} = 7.44	pK_{a2} = 6.9	
	pK_{a3} = 11.12	pK_{a3} = 11.6	

Several papers regarding the interaction of isoG with varying polymerases have appeared, in which isoG derivatives are the incoming triphosphate or the templating base (*13*, *72*−*75*). DNA pol β, the KF, and mammalian pol α insert both dTTP and dCTP opposite isoG, although dTTP is the preferred substrate (*13*, *73*, *74*). In addition, Taq polymerase can insert dTTP opposite isoG (*75*), although we are unaware of any study measuring isodGTP incorporation opposite T by Taq polymerase. The KF will insert isodGTP opposite isoC (its synthetic Watson−Crick pairing partner) and T but

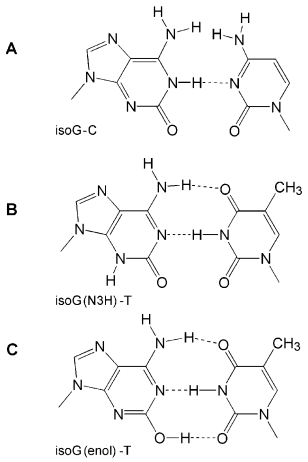


Figure 3. Proposed base-pairing geometries for isoG−pyrimidine matches. (A) isoG (N1H tautomer)−C with one hydrogen-bonding interaction coupled with a possible unfavorable repulsion of the NH_2 hydrogens. (B) isoG (N3H tautomer)−T, the configuration for polymerases that do not interrogate the minor groove. (C) isoG (enol tautomer)−T, the configuration explaining the behavior of pol β.

will not insert isodGTP opposite C (*13*, *72*). Taq polymerase can insert isodCTP opposite isoG (*75*). In contrast, pol α inserts isodGTP opposite both C and T (*13*). Of these four polymerases, ternary crystal structures have been solved for both pol β (*17*) and Taq polymerase (*76*). Pol β is known to have amino acid contacts in the minor groove of both the templating base (Arg 283) and the incoming dNTP (Asn 279) (*17*). In contrast, Taq polymerase makes specific contacts to the minor groove of the most recently formed base pair in both the primer and the template (Arg 573, Gln 754) but makes no minor groove contacts to the incoming dNTP or its templating base (*76*). Although no crystallographic data exist for the protein−DNA−nucleotide interactions of KF or pol α, a site specific amino acid substitution on KF implicates contacts with the primer terminus as being important for fidelity but not contacts with either the templating base or the incoming nucleotide (*22*). Furthermore, a study on a polymerase in the pol α family suggests that pol α probably operates similarly to KF (*77*). Thus, Taq polymerase, KF, and pol α are not expected to reject aberrant base pairs at the insertion step by interrogation of the minor groove, although they may prevent extension past the mispair due to the polymerase contact at the primer terminus.

isoG can pair with cytosine in a pseudo-Watson−Crick geometry (Figure 3a), forming one central hydrogen bond. The amino groups at the C6 positions may clash sterically, but this repulsion could be alleviated by out-of-plane rotation of the amino groups rather than a base pair shift to wobble geometry. This tautomer of isoG does not have an N3 proton, which is consistent with pol β's acceptance of this pairing. Furthermore, this isoG−C pair, while acceptable to the four mentioned polymerases, is not as preferred as the isoG−T pair (*74*), perhaps due to the steric destabilization of the amino groups.

We show that isoG exists in equilibrium between the N1H and N3H tautomers (**1** and **2**) in almost equal proportion, so that we can explain isoG−T pairings by the N3H tautomer depicted in Figure 3b. Both KF and pol α accept this base pair whether isoG is the templating base or the incoming nucleotide (*13*, *72*−*74*), and we suspect that the same is true for Taq polymerase, although only dTTP incorporation opposite isoG has been measured (*75*). However, this N3H−tautomeric configuration is not expected to be acceptable to pol β, given the presence of an N3 proton. Because pol β has the ability to incorporate dTTP opposite isoG (*74*), we must propose an alternative base-pairing structure that possesses Watson−Crick geometry. On the basis of our calculations, the most plausible possibility is that isoG shifts to an enol tautomer in the active site of pol β.

In our previous calculations on guanine (*42*), we observed that enol tautomers were exceedingly stable in the gas phase. This stability was due to in part to electrostatic repulsion of neighboring amino and N1 protons that destabilized the keto form coupled with dipole moment effects. In aqueous solution, water molecules were able to shield the protons, resulting in stabilization of the keto form. This observation is consistent both with experimental data on isoG, showing increased proportions of enol tautomers in increasingly nonpolar solvents (*27*, *30*), and with our gas phase calculations implicating a predominance of the enol forms. The dielectric constant of the active site of pol β is not known, although it is reasonable to expect that its value would lie somewhere between $\epsilon = 1$ (gas phase) and $\epsilon = 80$ (water). In fact, a model for polymerase fidelity was proposed based on the desolvation of the DNA bases in the active site of a DNA polymerase such that the thermodynamic differences between right and wrong base pairings would be amplified (*64*). Furthermore, the ternary crystal structure of T7 DNA polymerase shows an active site that very tightly fits around the nascent base pair (*78*). Perhaps a desolvated active site could approximate the properties of a nonpolar solvent and increase the relative proportion of the enol forms of isoG present. Thus, to explain the observation that pol β can insert dTTP opposite isoG and in fact prefers this pairing to dCTP−isoG (*74*), we must invoke the enol tautomer, which has no perturbations at the N3 position. Enol tautomers of isoG have been implicated previously to explain its base-pairing behavior (*13*, *72*, *73*, *75*, *79*, *80*).

Xanthine. Our calculations on X show one predominant diketo neutral tautomer. However, we calculate the pK_a value for X to be 6.9 (neutral-to-anionic pK_{a2}; Table 4), meaning that X is half ionized at physiological pH. This pK_a value agrees with the experimental value for the free base X, which is 7.44 (*71*, *81*, *82*). Interestingly, the experimental pK_a value of the nucleoside analogue of X, xanthosine, is substantially lower at 5.7 (*29*). This trend is unique among the other bases, whose pK_a values are comparable in both the free base and the nucleoside forms.

Another interesting result is that we calculate the pK_a of the N9 position of X to be 7.3, whereas the pK_a of the N9 position of guanine is substantially higher at 10.0 (Figure 2). Because the polarizability of the glycosidic bond can contribute to depurination in DNA, this result suggests that under neutral conditions deoxyxanthosine would be more labile to depurination than would deoxyguanosine. This conclusion is consistent with two recent experimental studies demonstrating that deoxyxanthosine is somewhat more labile to depurination than is deoxyguanosine at pH 7 (*83*, *84*).

Previous studies have examined the coding properties of X analogues. HIV1-RT is the most accepting of dXTP, as it will insert both dTTP and dCTP opposite this base in DNA (*83*). Additionally, HIV1-RT will insert dXTP

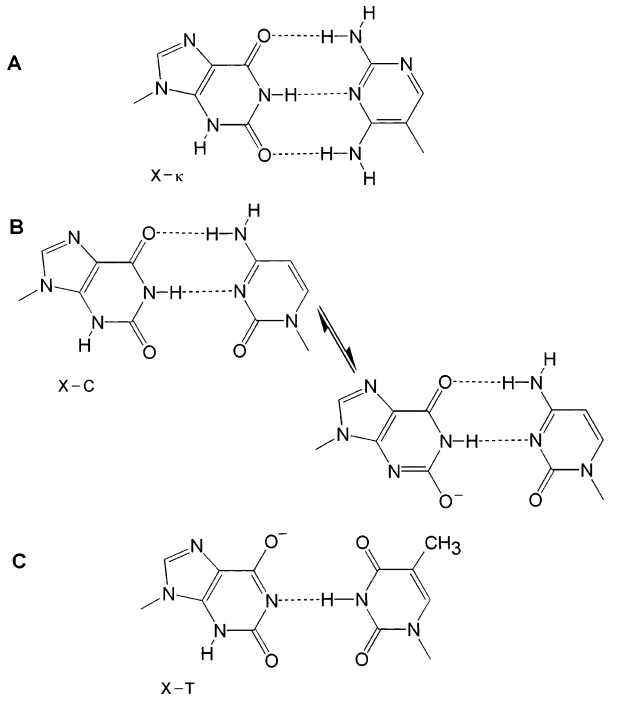


Figure 4. Proposed base-pairing geometries for X−pyrimidine matches. (A) X (neutral tautomer)−κ in the Watson−Crick geometry with three hydrogen bonds, which should be accepted by polymerases that do not interrogate the minor groove. (B) X−C, a pseudo-Watson−Crick base pair in equilibrium between the neutral and the ionized forms of X. (C) X−T, a pseudo-Watson−Crick base pair with one hydrogen bond. X must adopt a more rare N1-ionized tautomer in order for this geometric configuration to occur.

opposite κ and dκTP opposite X (*85*). Pol α is somewhat less accepting, as it will insert dTTP, dCTP, and dκTP (*86*, *87*) opposite X, although no extension occurs for dκTP (*88*). Interestingly, the KF will insert dXTP opposite κ (*89*) but not dκTP opposite X (*88*). Piccirilli and co-workers showed that dXTP opposite T is rejected, but a low level of dXTP−C is observed (*89*). The more recent study by Wuenschell and co-workers showed that KF inserts dXTP opposite both T and C with a preference for the latter (*83*). To our knowledge, no normal DNA bases have been paired opposite X by pol β, and furthermore, Horlacher and co-workers showed that pol β accepts neither dXTP opposite κ nor dκTP opposite X, even though these are apparent Watson−Crick complements (*88*).

In Figure 4, we suggest some possible tautomeric and ionized forms of X to explain the observed polymerase behavior in light of protein−DNA contacts made by these enzymes. As mentioned previously, KF, Taq polymerase, and pol α most likely do not monitor the minor groove of the incoming dNTP and its template partner whereas pol β does. The ternary crystal structure of HIV1-RT (*90*) shows protein−DNA contacts in the minor groove of several DNA bases, but there are none for the incoming dNTP and its template partner. A ternary crystal structure also exists for T7 DNA polymerase, showing similar hydrogen-bonding patterns to KF, Taq polymerase, and HIV1-RT (that is, minor groove contacts to the primer terminus and its corresponding template base, but no minor groove hydrogen bonds to the incoming dNTP or its base-pairing partner) (*78*). While we are unaware of any data on the behavior of T7 DNA polymerase with either isoG or X derivatives, we would expect this polymerase to accept incoming dNTPs that possess an N3 proton. Thus, of the polymerases mentioned, we would only expect pol β to reject bases containing a proton at the N3 position.

Figure 4a shows the X−κ base pair. The N3 proton prevents pol β from accepting this base pair as a substrate. Figure 4b shows the probable X−C configuration, in equilibrium between the neutral and the ionized forms of X. KF, pol α, and HIV1-RT accept the X−C base pair to some degree (*83*, *86*, *89*). We would expect that pol β also accepts this base pair, as it adopts Watson−Crick geometry, and in the ionized form has no N3 proton. The ability of pol α and HIV1-RT to insert dTTP opposite X is less obvious. There are no possible Watson−Crick pairings of any predominant neutral or ionized tautomer with T. Therefore, either a neutral enol tautomer or an N1-ionized tautomer (since pol α and HIV1-RT do not interrogate the minor groove at the insertion step) may be involved. The potential enol form **x3** (Table 2) occurs at a frequency of 7×10^{-9} in aqueous phase and 2×10^{-14} in gas phase (2×10^{-8} and 8×10^{-8}, respectively, among the tautomers with a proton kept at the N9 position only), thus ruling out this tautomer as an explanation. The ionized form presented in Figure 4c (equivalent to **x12**; Table 3) costs 3.6 kcal/mol from **x1** in aqueous solution at pH 7 (see Table S10 for details), rather than 10.3 kcal/mol for the previously mentioned enol tautomer **x3** relative to **x1**. While this makes the N1-ionized tautomer **x12** significantly more likely to occur than the enol tautomer **x3**, it is 4.3 kcal/mol higher than the N3-ionized form (**x11**) and is still ∼1000 times less likely to occur than the N3-ionized form **x11**. We note that guanine and X have essentially equal site specific pK_a values of 9.6 at the N1 position (Figure 2). Goodman and co-workers previously noted that a G−T mispair can be forced at high pH, ionizing G or T, and permitting a base pair approximating Watson−Crick geometry (*36*). An analogous situation may occur here, such that low levels of dTTP are inserted opposite X, which would adopt the N1-ionized tautomer **x12** depicted in Figure 4c at a frequency of ∼1/1000 (given a pK_a value of 9.6). We would not expect pol β to accept a T−X base pair, as this form of X would have an obligatory N3 proton in order to produce Watson−Crick geometry.

Conclusion

Our calculations on isoG and X show that many DNA bases, both normal and endogenous DNA damage products, exist as a composite of several structures in equilibrium. The theoretical methods presented here provide a powerful means to determine these tautomeric and ionized equilibria, allowing us to make conclusions about the mispairing potential of many DNA damage products. We show that isoG exists as a composite of two neutral forms in aqueous solution but can shift to enol forms in gas phase conditions. X, on the other hand, is a composite between one neutral, keto form and one ionized form. Both molecules have an equilibrium structure with an N3 proton, which affects their interactions with DNA polymerases such as pol β that interrogate the minor groove at the insertion step.

Acknowledgment. This work was supported in part by the National Institutes of Health GM41336 (L.C.S.) and CA85779 (L.C.S. and W.A.G.) as well as the BK21

program of Korea (Y.H.J.). In addition, the facilities of the MSC used for this project were supported by ARO-DURIP, ONR-DURIP, NSF-MRI, and IBM (SUR grant).

Supporting Information Available: Relative free energies and populations of all neutral and ionized tautomers calculated as well as additional data regarding pK_a values and gas phase dipole moments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Loeb, L. A., and Kunkel, T. A. (1982) Fidelity of DNA synthesis. *Annu. Rev. Biochem.* **52**, 429−457.
(2) Echols, H., and Goodman, M. F. (1991) Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.* **60**, 477−511.
(3) Kunkel, T. A. (1992) DNA replication fidelity. *J. Biol. Chem.* **267**, 18251−18254.
(4) Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol.* **28**, 83−126.
(5) Goodman, M. F. (1997) Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10493−10495.
(6) Diederichsen, U. (1998) Selectivity of DNA replication: the importance of base-pair geometry over hydrogen bonding. *Angew. Chem., Int. Ed.* **37**, 1655−1657.
(7) Ames, B. N. (1983) Dietary carcinogens and anticarcinogens: oxygen radicals and degenerative diseases. *Science* **221**, 1256−1264.
(8) Imlay, J. A., and Linn, S. (1988) DNA damage and oxygen radical toxicity. *Science* **240**, 1302−1309.
(9) Loeb, L. A. (1989) Endogenous carcinogenesis: molecular oncology into the twenty-first century. *Cancer Res.* **49**, 5489−5496.
(10) Mullaart, E., Lohman, P. H., Berends, F., and Vijg, J. (1990) DNA damage metabolism and aging. *Mutat. Res.* **237**, 189−210.
(11) Suzuki, T., Ide, H., Yamada, M., Endo, N., Kanaori, K., Tajima, K., Morii, T., and Makino, K. (2000) Formation of 2′-deoxyoxanosine from 2′-deoxyguanosine and nitrous acid: mechanism and intermediates. *Nucleic Acids Res.* **28**, 544−551.
(12) Hoeijmakers, J. H. J. (2001) Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366−374.
(13) Kamiya, H., and Kasai, H. (1995) Formation of 2-hydroxydeoxyadenosine triphosphate, an oxidatively damaged nucleotide, and its incorporation by DNA polymerases. *J. Biol. Chem.* **270**, 19446−19450.
(14) de Rojas-Walker, T., Tamir, S., Ji, H., Wishnok, J. S., and Tannenbaum, S. R. (1995) Nitric-oxide induces oxidative damage in addition to deamination in macrophage DNA. *Chem. Res. Toxicol.* **8**, 473−477.
(15) Karran, P., and Lindahl, T. (1980) Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry* **19**, 6005−6011.
(16) Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994) Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer and ddCTP. *Science* **264**, 1891−1903.
(17) Beard, W. A., Osheroff, W. P., Prasad, R., Sawaya, M. R., Jaju, M., Wood, T. G., Kraut, J., Kunkel, T. A., and Wilson, S. H. (1996) Enzyme-DNA interactions required for efficient nucleotide incorporation and discrimination in human DNA polymerase β. *J. Biol. Chem.* **271**, 12141−12144.
(18) Morales, J. C., and Kool, E. T. (1999) Minor groove interactions between polymerase and DNA: more essential to replication than Watson−Crick hydrogen bonds? *J. Am. Chem. Soc.* **121**, 2323−2324.
(19) Osheroff, W. P., Beard, W. A., Wilson, S. H., and Kunkel, T. A. (1999) Base substitution specificity of DNA polymerase β depends on interactions in the DNA minor groove. *J. Biol. Chem.* **274**, 20749−20752.
(20) Osheroff, W. P., Beard, W. A., Yin, S., Wilson, S. H., and Kunkel, T. A. (2000) Minor groove interactions at the DNA polymerase β active site modulate single-base deletion error rates. *J. Biol. Chem.* **275**, 28033−28038.
(21) Morales, J. C., and Kool, E. T. (2000) Functional hydrogen-bonding map of the minor groove binding tracks of six DNA polymerases. *Biochemistry* **39**, 12979−12988.
(22) Spratt, T. E. (2001) Identification of hydrogen bonds between *Escherichia coli* DNA polymerase I (Klenow fragment) and the minor groove of DNA by amino acid substitution of the polymerase and atomic substitution of the DNA. *Biochemistry* **40**, 2647−2652.
(23) Kool, E. T. (2001) Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1−22.
(24) Shah, A. M., Conn, D. A., Li, S. X., Capaldi, A., Jager, J., and Sweasy, J. B. (2001) A DNA polymerase β mutator mutant with reduced nucleotide discrimination and increased protein stability. *Biochemistry* **40**, 11372−11381.
(25) Beard, W. A., Shock, D. D., Yang, X.-P., DeLauder, S. F., and Wilson, S. H. (2002) Loss of DNA polymerase β stacking interactions with templating purines, but not pyrimidines, alters catalytic efficiency and fidelity. *J. Biol. Chem.* **277**, 8235−8242.
(26) Subramanian, E., and Marsh, R. E. (1971) The crystal structure of isoguanine sulfate monohydrate. *Acta Crystallogr.* **B27**, 753−759.
(27) Sepiol, J., Kazimierczuk, Z., and Shugar, D. (1976) Tautomerism of isoguanine and solvent-induced keto−enol equilibrium. *Z. Naturforsch.* **31c**, 361−370.
(28) Banerjee, A., Saenger, W., Lesyng, B., Kazimierczuk, Z., and Shugar, D. (1978) 9-methylisoguanine hydrochloride dihydrate: crystal structure and charge densities. *Acta Crystallogr.* **B34**, 2472−2477.
(29) Roy, K. B., and Miles, H. T. (1983) Tautomerism and ionization of xanthosine. *Nucleosides Nucleotides* **2**, 231−242.
(30) Seela, F., Wei, C., and Kazimierczuk, Z. (1995) Substituent reactivity and tautomerism of isoguanine and related nucleosides. *Helv. Chim. Acta* **78**, 1843−1854.
(31) Roberts, C., Bandaru, R., and Switzer, C. (1997) Theoretical and experimental study of isoguanine and isocytosine: base pairing in an expanded genetic system. *J. Am. Chem. Soc.* **119**, 4640−4649.
(32) Robinson, H., Gao, Y. G., Bauer, C., Roberts, C., Switzer, C., and Wang, A. H. (1998) 2′-Deoxyisoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry* **37**, 10897−10905.
(33) Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., and Fazakerly, G. V. (1988) Equilibrium between a wobble and ionized base pair formed between fluorouracil and guanine in DNA as studied by proton and fluorine NMR. *J. Biol. Chem.* **263**, 14794−14801.
(34) Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B., and Fazakerley, G. V. (1989) Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions. *J. Mol. Biol.* **205**, 437−447.
(35) Fazakerley, G. V., Gdaniec, Z., and Sowers, L. C. (1993) Base-pair induced shifts in the tautomeric equilibrium of a modified DNA base. *J. Mol. Biol.* **230**, 6−10.
(36) Yu, H., Eritja, R., Bloom, L. B., and Goodman, M. F. (1993) Ionization of bromouracil and fluorouracil stimulates base mispairing frequencies with guanine. *J. Biol. Chem.* **268**, 15935−15943.
(37) Suen, W., Spiro, T. G., Sowers, L. C., and Fresco, J. R. (1999) Identification by UV resonance Raman spectroscopy of an imino tautomer of 5-hydroxy-2′-deoxycyditine, a powerful base analogue transition mutagen with a much higher unfavored tautomer frequency than that of the natural residue 2′-deoxycytidine. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4500−4505.
(38) Sowers, L. C., Boulard, Y., and Fazakerley, G. V. (2000) Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry* **39**, 7613−7620.
(39) La Francois, C. J., Jang, Y. H., Cagin, T., Goddard, W. A., III, and Sowers, L. C. (2000) Conformation and proton configuration of pyrimidine deoxynucleoside oxidation damage products in water. *Chem. Res. Toxicol.* **13**, 462−470.
(40) Jang, Y. H., Sowers, L. C., Cagin, T., and Goddard, W. A., III (2001) First principles calculation of pK_a values for 5-substituted uracils. *J. Phys. Chem. A* **105**, 274−280.
(41) Jang, Y. H., Goddard, W. A., III, Noyes, K. T., Sowers, L. C., Hwang, S., and Chung, D. S. (2002) First principles calculations of the tautomers and pK_a values of 8-oxoguanine: implications for mutagenicity and repair. *Chem. Res. Toxicol.* **15**, 1023−1035.
(42) Jang, Y. H., Goddard, W. A., III, Noyes, K. T., Sowers, L. C., Hwang, S., and Chung, D. S. (2003) pK_a values of guanine in water: density functional theory calculations combined with Poisson−Boltzmann continuum-solvation model. *J. Phys. Chem. B* **107**, 344−357.
(43) *Jaguar 4.0* (2000) Schrodinger Inc., Portland, OR.
(44) Greeley, B. H., Russo, T. V., Mainz, D. T., Friesner, R. A., Langlois, J.-M., Goddard, W. A., III, Donnelly, R. E., Jr., and Ringnalda, M. N. (1994) New pseudospectral algorithms for electronic structure calculations: length scale separation and analytical two-electron integral corrections. *J. Chem. Phys.* **101**, 4028−4041.

(45) Slater, J. C. (1974) *Quantum Theory of Molecules and Solids, Vol. 4. The Self-Consistent Field for Molecules and Solids*, McGraw-Hill, New York.
(46) Vosko, S. H., Wilk, L., and Nusair, M. (1980) Accurate spin-dependent electron liquid correlation energies for local spin-density calculations—a critical analysis. *Can. J. Phys.* **58**, 1200−1211.
(47) Lee, C. T., Yang, W. T., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **37**, 785−789.
(48) Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic-behavior. *Phys. Rev. A* **38**, 3098−3100.
(49) Miehlich, B., Savin, A., Stoll, H., and Preuss, H. (1989) Results obtained with the correlation-energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **157**, 200−206.
(50) Tannor, D. J., Marten, B., Murphy, R., Friesner, R. A., Sitkoff, D., Nicholls, A., Ringnalda, M., Goddard, W. A., III, and Honig, B. (1994) Accurate first principles calculation of molecular charge-distributions and solvation energies from ab initio quantum mechanics and continuum dielectric theory. *J. Am. Chem. Soc.* **116**, 11875−11882.
(51) Honig, B., and Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science* **268**, 1144−1149.
(52) Marten, B., Kim, K., Cortis, C., Friesner, R. A., Murphy, R. B., Ringnalda, M. N., Sitkoff, D., and Honig, B. (1996) New model for calculation of solvation free energies: correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. *J. Phys. Chem.* **100**, 11775−11788.
(53) Nicholls, A., and Honig, B. (1991) A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson−Boltzmann equation. *J. Comput. Chem.* **12**, 435−445.
(54) Archer, D. G., and Wang, P. M. (1990) The dielectric-constant of water and Debye−Huckel limiting law slopes. *J. Phys. Chem. Ref. Data* **19**, 371−411.
(55) Lim, C., Bashford, D., and Karplus, M. (1991) Absolute pK_a calculations with continuum dielectric methods. *J. Phys. Chem.* **95**, 5610−5620.
(56) Topol, I. A., Tawa, G. J., Burt, S. K., and Rashin, A. A. (1997) Calculation of absolute and relative acidities of substituted imidazoles in aqueous solvent. *J. Phys. Chem. A* **101**, 10075−10081.
(57) Marcus, Y. (1985) *Ion Solvation*, p 105, John Wiley and Sons Ltd., New York.
(58) Reiss, H., and Heller, A. (1985) The absolute potential of the standard hydrogen electrode—a new estimate. *J. Phys. Chem.* **89**, 4207−4213.
(59) Tissandier, M. D., Cowen, K. A., Feng, W. Y., Gundlach, E., Cohen, M. H., Earhart, A. D., Coe, J. V., and Tuttle, T. R. (1998) The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster ion solvation data. *J. Phys. Chem. A* **102**, 7787−7794.
(60) Watson, J. D., and Crick, F. H. C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737−738.
(61) Watson, J. D., and Crick, F. H. C. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964−967.
(62) Topal, M. D., and Fresco, J. R. (1976) Complementary base pairing and the origin of substitution mutations. *Nature* **263**, 285−289.
(63) Sowers, L. C., Shaw, B. R., Veigl, M. L., and Sedwick, W. D. (1987) DNA base modification: ionized base pairs and mutagenesis. *Mutat. Res.* **177**, 201−218.
(64) Petruska, J., Sowers, L. C., and Goodman, M. F. (1986) Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1559−1562.
(65) Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr. (1988) Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6252−6256.
(66) Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 804−808.
(67) Pitsch, S., Krishamurthy, R., Bolli, M., Wendebord, S., Holzner, A., Minton, M., Lesueur, C., Schlonvogt, I., Jaun, B., and Eschenmoser, A. (1995) Pyranosyl-RNA ('p-RNA'): Base-pairing selectivity and potential to replicate. *Helv. Chim. Acta* **78**, 1621−1635.
(68) Eschenmoser, A. (1993) Hexose nucleic acids. *Pure Appl. Chem.* **65**, 1179−1188.
(69) Krishnamurthy, R., Pitsch, S., Minton, M., Miculka, C., Windhab, N., and Eschenmoser, A. (1996) Pyranosyl-RNA: base pairing between homochiral oligonucleotide strands of opposite sense of chirality. *Angew. Chem., Int. Ed. Engl.* **35**, 1537−1541.
(70) Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M. (1986) *Data for Biochemical Research*, 3rd ed., Oxford University Press, Oxford.
(71) Albert, A., and Brown, D. J. (1954) Purine Studies. Part I. Stability to acid and alkali. Solubility. Ionization. Comparison with pteridines. *J. Chem. Soc.*, 2060−2071.
(72) Switzer, C., Moroney, S. E., and Benner, S. A. (1989) Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **111**, 8322−8323.
(73) Kamiya, H., Ueda, T., Ohgi, T., Matsukage, A., and Kasai, H. (1995) Misincorporation of dAMP opposite 2-hydroxyadenine, an oxidative form of adenine. *Nucleic Acids Res.* **23**, 761−766.
(74) Kamiya, H., and Kasai, H. (1996) Effects of sequence contexts on misincorporation of nucleotides opposite 2-hydroxyadenine. *FEBS Lett.* **391**, 113−116.
(75) Maciejewska, A. M., Lichota, K. D., and Kusmierek, J. T. (2003) Neighbouring bases in template influence base-pairing of isoguanine. *Biochem. J.* **369**, 611−618.
(76) Li, Y., Korolev, S., and Waksman, G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.* **17**, 7514−7525.
(77) Franklin, M. C., Wang, J. M., and Steitz, T. A. (2001) Structure of the replicating complex of a pol α family DNA polymerase. *Cell* **105**, 657−667.
(78) Doublie, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 angstrom resolution. *Nature* **391**, 251−258.
(79) Switzer, C. Y., Moroney, S. E., and Benner, S. A. (1993) Enzymatic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry* **32**, 10489−10496.
(80) Chaput, J. C., and Switzer, C. (2000) Nonenzymatic transcription of an isoG-isoC base pair. *J. Am. Chem. Soc.* **122**, 12866−12867.
(81) Bergmann, F., and Dikstein, S. (1955) The relationship between spectral shifts and structural changes in uric acids and related compounds. *J. Am. Chem. Soc.* **77**, 691−696.
(82) Ogston, A. G. (1935) The constitution of the purine nucleosides. Part III. Potentiometric determination of the dissociation constants of methylated xanthines. *J. Chem. Soc.* 1376−1379.
(83) Wuenschell, G. E., O'Connor, T. R., and Termini, J. (2003) Stability, miscoding potential, and repair of 2′-deoxyxanthosine in DNA: implications for nitric-oxide induced mutagenesis. *Biochemistry* **42**, 3608−3616.
(84) Vongchampa, V., Dong, M., Gingipalli, L., and Dedon, P. (2003) Stability of 2′-deoxyxanthosine in DNA. *Nucleic Acids Res.* **31**, 1045−1051.
(85) Lutz, M. J., Held, H. A., Hottiger, M., Hubscher, U., and Benner, S. A. (1996) Differential discrimination of DNA polymerases for variants of the nonstandard nucleobase pair between xanthosine and 2,4-diaminopyrimidine, two components of an expanded genetic alphabet. *Nucleic Acids Res.* **24**, 1308−1313.
(86) Eritja, R., Horowitz, D. M., Walker, P. A., Ziehler-Martin, J. P., Boosalis, M. S., Goodman, M. F., Itakura, K., and Kaplan, B. E. (1986) Synthesis and properties of oligonucleotides containing 2′-deoxynebularine and 2′-deoxyxanthosine. *Nucleic Acids Res.* **14**, 8135−8153.
(87) Kamiya, H., Shimizu, M., Suzuki, M., Inoue, H., and Ohtsuka, E. (1992) Mutation induced by deoxyxanthosine in codon 12 of a synthetic c-Ha-*ras* gene. *Nucleosides Nucleotides* **11**, 247−260.
(88) Horlacher, J., Hottiger, M., Podust, V. N., Hubscher, U., and Benner, S. A. (1995) Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6329−6333.
(89) Piccirilli, J. A., Krauch, T., Moroney, S. E., and Benner, S. A. (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **343**, 33−37.
(90) Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669−1675.
(91) Taylor, H. F. W. (1948) The dissociation constants of benziminazole and certain purine derivatives. *J. Chem. Soc.* 765−766.

TX034068E