Comprehensive Study of the Effects of Methylation on Tautomeric Equilibria of Nucleic Acid Bases

Gareth K. Forde,[†,§] Aviane E. Forde,[†] Glake Hill,[†] Adrian Ford,[†] Alejandro Nazario,[‡] and Jerzy Leszczynski*[,†]

Computational Center for Molecular Structure and Interactions, Jackson State University, Jackson, Mississippi 39217, Department of Biology, Universidad Metropolitana, Rio Piedras, Puerto Rico 00928, and School of Medicine, University of Minnesota, Minneapolis, Minnesota 55455

Received: November 14, 2005; In Final Form: March 31, 2006

Minor tautomers of nucleic acid bases can result by intramolecular proton transfer. These rare tautomers could be stabilized through the addition of methyl groups to DNA bases. A comprehensive theoretical study of tautomers of methylated derivatives of guanine, adenine, cytosine, thymine, and uracil was performed. Molecular geometries of all tautomers were obtained at the density functional theory and MP2 levels with the 6-31G(d,p) basis set, and single-point calculations were performed at the CCSD(T)/6-311G(d,p) level. Tautomers obtained by protonation at the preferred protonation site for methylated isolated bases were compared to their nonmethylated counterparts. The effects of methylation on the relative stabilities of nucleic acid base tautomers are also studied and discussed in this work. The results suggest that some sites on the bases may not be mutagenic and may even stabilize the canonical Watson−Crick form. The results also indicate that a number of methylation sites can stabilize the tautomers, suggesting possible mechanisms for mutagenic changes.

1. Introduction

The role of tautomerization in genetic expression has been the subject of numerous studies, both experimental and theoretical.[1−13] A number of these studies have determined that the production of these tautomers can lead to mispairs and eventual mutations.[14−18] These mutations are the precursors to many molecular-based diseases, including cancer. A greater understanding of this phenomenon is crucial to understanding the mechanisms leading to gene expressions and cellular changes.

The interactions between neighboring bases in nucleic acids are largely responsible for their observed structures.[19] These interactions usually fall into one of two categories, hydrogen bonding and base stacking. Base stacking, until recently, had been very difficult to analyze by high-level theoretical ab initio methods because of the large dispersion components of the interaction energy.[20,21] Hydrogen bonding, however, has long been investigated by ab initio methods.[22] The interactions of adjacent bases, which are mediated by hydrogen bonding, are responsible for base complimentarity. This phenomenon results in the double-stranded character of DNA described over half a century ago.[23] The mutagenic potential of the tautomerization of nucleic acid bases (NABs) lies in their ability to form noncomplimentary base pairs resulting in spontaneous mutations. Because of this discovery, tautomerism in NABs has been the subject of many investigations.[24,25] Both the experimental and the theoretical evidence for tautomeric equilibria in NABs is well documented.[22]

Previous investigations on the tautomers of NABs have included evidence of stabilization of the rare tautomers by substituent groups.[26,27] In these studies the investigated substituents included heavy metals, bromine, and methyl groups. The focus of the present work centers on methylation changes in NABs.

The recent works on the methyl substituents mainly evaluated cytosine methylation at the C(5) position.[28−30] It should be noted that the C(5) cytosine atom is one of many well-established promutagenic methylation sites. Moreover, the endogenous methylation of cytosine at C5 is essential to the normal regulation of gene expression. However, the exogenous methylation of the more electronegative sites in NABs such as N7, N3, and O6 in the purines and O2, O4, and N3 in the pyrimidines is also known to be responsible for mutagenic events. It is known that methylation induces several classes of base pair substitutions and frame shift mutations. Unfortunately, to date, no investigation has been devoted solely to comprehensive studies of the major promutagenic methyl derivatives of the tautomers of both the pyrimidines and the purines.

The role that purines play in methylation has been determined to be important in the methylation cascade. The most reactive sites in NABs have been determined to exist at the N7 and N3 sites of guanine and the N3 site of adenine in S_N2 reactions.[31] These are the most favored sites for S_N2 alkylating agents. Contrarily, the S_N1 alkyating compounds favor exocyclic base oxygens and oxygens in the phosphodiester backbone.

Experimental studies confirm that the N7 site is the preferred site of guanine methylation.[32] The importance of this fact is, however, still debated. Some studies suggest that since N7 does not directly participate in hydrogen bonding of the base pair, it is only weakly mutagenic if at all.[32] Others suggest that the N7 methylated species can undergo ring opening that blocks any DNA replication.[33−35] This seems to be supported by the fact

* Author to whom correspondence should be addressed. Phone: (601)-979-3723. Fax: (601)979-3982. E-mail: jerzy@ccmsi.us.
[†] Jackson State University.
[‡] Universidad Metropolitana.
[§] University of Minnesota.

10.1021/jp056574d CCC: $33.50
Published on Web 07/15/2006

TABLE 1: Comparison of the Protonation Sites of Guanine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}	$\Delta H^{CCSD(T)}$
			Nonmethylated Species		
N9H	0.06	0.07	0.0	0.01	0.00
	−542.448221	−542.481203	−540.922858	−540.956002	−541.2216896
N7H	0.00	0.00	0.01	0.00	0.18
	−542.448315	−542.481312	−540.922837	−540.956036	−541.2214057
			N3 Methylated Species		
N9H	14.16	13.90	14.08	14.37	
	−582.081802	−582.117411	−580.420047	−580.455696	
N7H	0.00	0.00	0.00	0.00	
	−582.104364	−582.139562	−580.442483	−580.478603	

[a] Relative energies are in kcal/mol; total energies are in au; see Figure 1.

TABLE 2: Comparison of the N9H Protonated Keto and Enol Tautomers of Guanine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}	$\Delta H^{CCSD(T)}$
			Nonmethylated Species		
N9H keto	0.00	0.00	0.0	0.0	0.00
	−542.448221	−542.481203	−540.922858	−540.956002	−541.2216896
N9H enol	0.86	0.97	1.00	1.18	0.28
	−542.446851	−542.479651	−540.921243	−540.954120	−541.2212396
			N7 Methylated Species		
N9H keto	0.00	0.00	0.00	0.00	0.00
	−582.115603	−582.151689	−580.451096	−580.487318	−580.777839
N9H enol	3.84	4.38	3.90	4.33	3.29
	−582.10948	−582.144708	−580.444879	−580.480413	−580.772589
			N3 Methylated Species		
N9H keto	7.25	7.17	5.50	5.17	
	−582.081802	−582.117411	−580.420047	−580.455696	
N9H enol	0.00	0.00	0.00	0.00	
	−582.09336	−582.128839	−580.428805	−580.463949	

[a] Relative energies are in kcal/mol; total Energies are in au; see Figure 2.

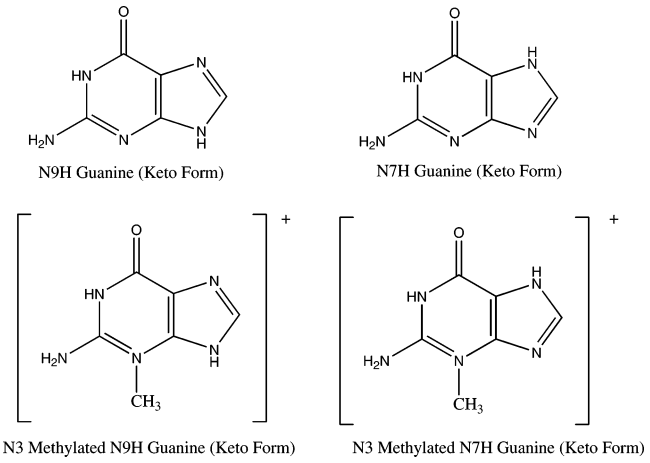

Figure 1. Molecular structures of N9H and N7H guanine, nonmethylated and methylated (Table 1).

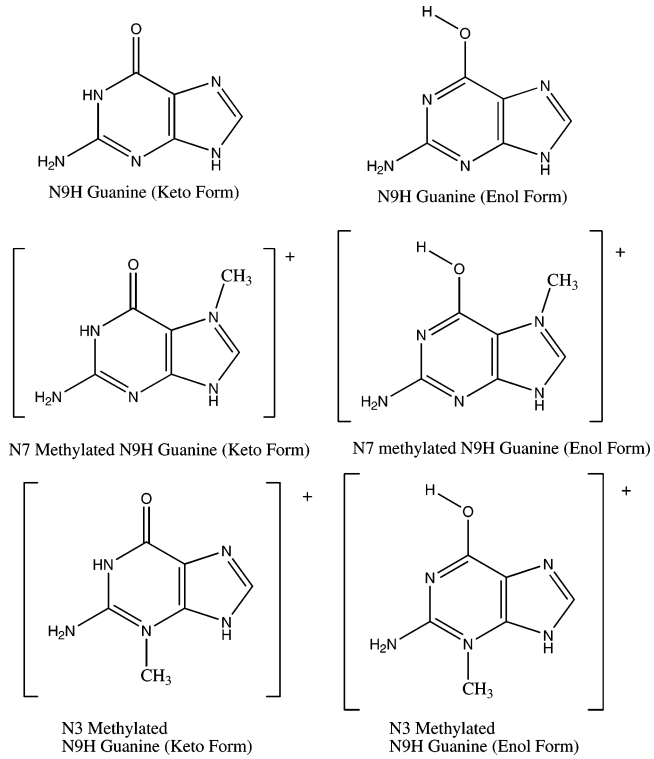

Figure 2. Molecular structures of N9H guanine, keto and enol tautomers, nonmethylated and methylated (Table 2).

that smokers experience about twice as much N7 methylation as nonsmokers.[31] It is widely believed that tobacco smoke, containing a methylating agent, causes a number of cancers.

Two sites of protonation are available in isolated guanine bases, the N9 and the N7. Both of these are important to the classical and nonclassical binding of guanine. Normally the deoxyribose sugar from the DNA backbone is attached at the N9 site. However, reverse and nonclassical binding may cause interactions with the N7 site. The N3 site protonation is virtually nonexistent except in the isolated gas phase studies. Previous studies have shown that it is highly unfavorable when compared to the N9H and N7H species.[36]

Adenine methylation has been suggested to be as important as guanine methylation. The N3 site of adenine is thought to be the second most favored site of methylation, second to the sites of guanine.[31,37] Some studies have suggested that this N3 methylated site can cause major consequences and can be the most responsible factor for toxicity by a methylating agent.[37]

Pyrimidine methylation has been the subject of a great deal of attention, especially cytosine methylation. Cytosine methylation is thought to be important in eukaryotic cellular development,[38] viral latancy,[39] and regulation of transcription during

TABLE 3: Comparison of the Protonation Sites of Adenine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}	$\Delta H^{CCSD(T)}$
			Nonmethylated Species		
N9H	0.00	0.00	0.00	0.00	
	−467.219176	−467.250944	−465.877645	−465.909271	
N7H	8.11	8.19	7.67	7.63	
	−467.206256	−467.237892	−465.865423	−465.897116	
N3H	8.57	8.30	9.01	8.85	
	−467.205523	−467.237713	−465.863293	−465.895170	
			N1 Methylated Species		
N9H	0.00	0.00	0.00	0.00	
	−506.874780	−506.908206	−505.396591	−505.430347	
N7H	10.87	10.63	12.04	12.41	
	−506.857452	−506.891266	−505.377398	−505.410999	
N3H	24.68	24.51	24.46	24.66	
	−506.835450	−506.869154	−505.357603	−505.391046	
			N3 Methylated Species		
N9H	2.16	2.41	1.78	2.14	2.05
	−506.874357	−506.908077	−505.394317	−505.428249	−505.676738
N7H	0.00	0.00	0.00	0.00	0.00
	−506.877802	−506.911914	−505.397161	−505.431661	−505.680010
			N7 Methylated Species		
N9H	6.83	6.56	6.28	6.43	
	−506.864713	−506.898810	−505.383963	−505.417730	
N3H	0.00	0.00	0.00	0.00	
	−506.875597	−506.909270	−505.393973	−505.427973	

[a] Relative energies are in kcal/mol; total Energies are in au; see Figure 3.

TABLE 4: Comparison of the N9H Protonated Tautomers of Adenine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}
		Nonmethylated Species		
N9H amino	0.00	0.00	0.00	0.00
	−467.219176	−467.250944	−465.877645	−465.909271
N9H imino	12.51	12.68	12.40	12.33
	−467.199232	−467.230740	−465.857878	−465.889626
		N3 Methylated Species		
N9H amino	0.00	0.00	0.00	0.00
	−506.874357	−506.908077	−505.394317	−505.428249
N9H imino	23.80	23.77	22.57	22.43
	−506.836426	−506.870189	−505.358360	−505.392502
		N7 Methylated Species		
N9H amino	0.00	0.00	0.00	0.00
	−506.864713	−506.89881	−505.383963	−505.417730
N9H imino	10.30	10.63	9.57	9.55
	−506.848299	−506.881871	−505.368708	−505.402510

[a] Relative energies are in kcal/mol; total Energies are in au; see Figure 4.

cellular differentiation.[40,41] Hyper- or hypomethylation has been a target in the initiation and progression of human tumors.[42,43] Contrary to purines, thymine and uracil methylation has not been reported so extensively. However, there are some studies that determined that thymine and uracil methylation can be a common occurrence. O4 thymine methylation occurs more regularly than O6 guanine methylation and is more mutagenic in bacteria.[31] This is probably due to a lack of the adequate repair enzyme for this thymine methylation. Yet there has been little evidence of thymine or uracil methylation importance in human cell lines. Nevertheless, for completeness, all DNA and RNA bases were included in this study.

There is an obvious interest in the evaluation of the combined effects of methylation and tautomerism. In this investigation, high-level ab initio techniques were used to determine the effect of methylation at different possible sites on the relative stability of rare tautomers of both the purines (adenine and guanine) and the pyrimidines (cytosine, uracil, and adenine).

2. Computational Details

In this work, the tautomeric equilibria of canonical nucleic acid bases and their methyl derivatives (Figure 8) were studied using the Gaussian 98 program.[44] All of the geometries considered in this study were fully optimized without symmetry constraints with density functional theory (DFT) and Moller−Plesset (MP2) methods using the 6-31G(d,p) basis set. The applied DFT method used the B3LYP hybrid functional (a parametrized combination of Becke's exchange functional,[45] the Lee, Yang, and Parr correlation functional, and the exact exchange). Vibrational frequency calculations were performed to verify that for all species the obtained geometries correspond to local minima. The obtained geometries were used as the inputs for the full MP2/6-31G(d,p)-level optimization and vibrational frequency calculations. The single-point calculations were carried out with the coupled cluster methods (CCSD(T)) methods and the 6-311G(d,p) basis set using the geometries obtained with MP2 optimization for the pyrimidines and their methyl derivatives. For the purines, because the coupled cluster single-point calculations are computationally demanding, the CCSD(T)/6-311G(d,p)//MP2/6-31G(d,p)-level evaluations of relative stabilities were only performed when the difference between a low-lying tautomer and the global minimum tautomer at the MP2/6-31G(d,p) level is less than 5 kcal/mol.

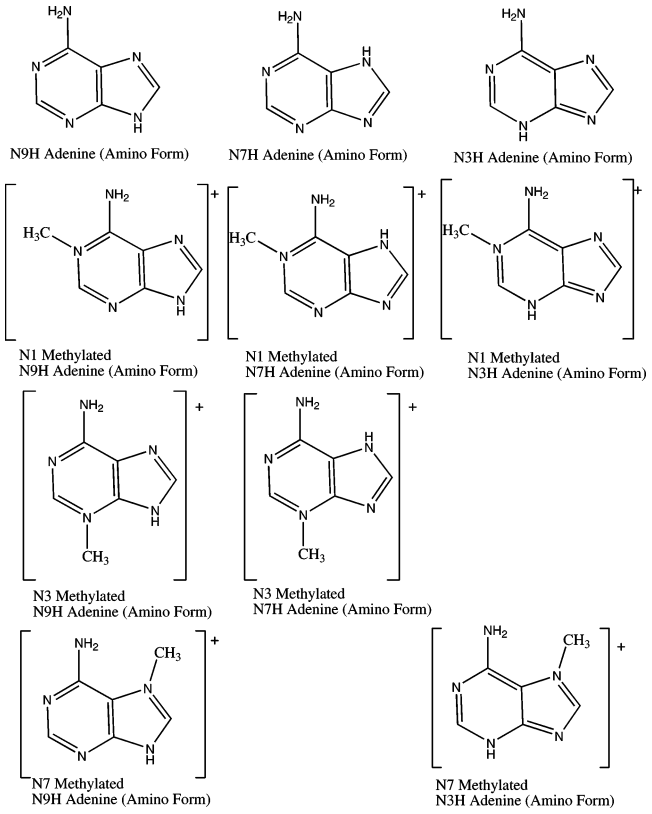

Figure 3. Molecular structures of N9H, N7H, and N3H adenine, nonmethylated and methylated (Table 3).

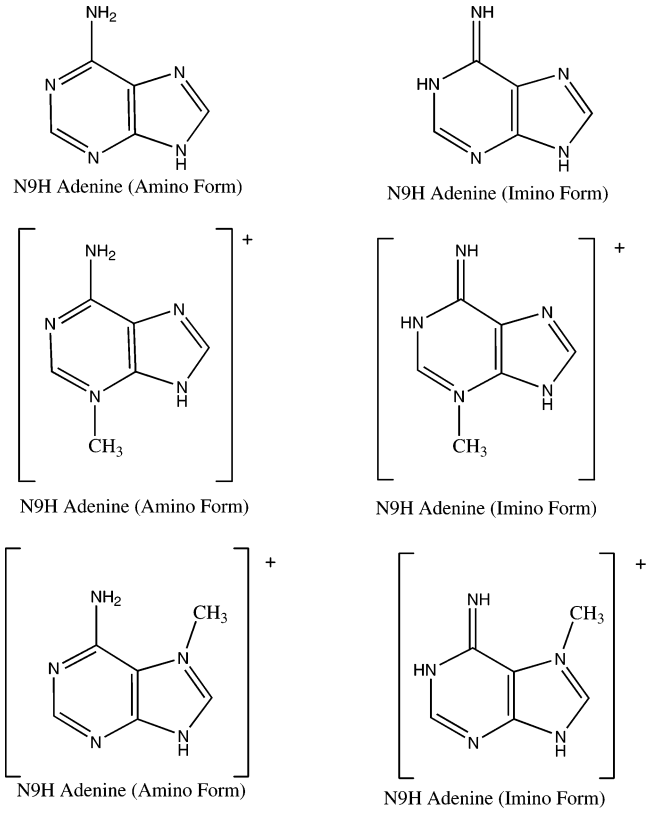

Figure 4. Molecular structures of N9H adenine, amino and imino tautomers, nonmethylated and methylated (Table 4).

3. Results

A. Purine Methylation. In guanine nucleic acid base, the N9H and the N7H protonated structures are similar in energy (Table 1). This is consistent with previous experimental[24] and theoretical studies.[36] The order of stability is debatable, with DFT predicting the N7H tautomer to be the most stable (0.06 and 0.07 kcal/mol), MP2 predicting equal energies, and CCSD-(T) favoring the N9H form. The QCISD(T) level of theory predicts the N7H tautomer to be the most stable, but with a very insignificant energy difference.[36] Regardless of the applied method, the two tautomers are essentially equal in predicted energy. When methylated at the N3 position, the N7H position is the energetically predominant form (between 13.9 and 14.4 kcal/mol difference). This seems to suggest that the N3 methylation favors a reverse or nonclassical interaction in DNA or RNA.

O6 exocyclic methylation favors the N9H conformation, e.g., the conformation that is normally detected in DNA. This does not mean that the O6 conformation is not responsible for changes in the DNA. The experimental studies concluded that the O6 methylation could induce a G to A transition in DNA replication.[46−51] The O6 methylated guanine can also form weak pairing with thymine, although this is weaker than the Watson−Crick GC base pair.[52]

Table 2 compares the keto and enol tautomers of N9H protonated guanine when methylated at the N7 and N3 sites. These tautomers are compared to the unmethylated species. There is very little difference in stability of the canonical N9H keto form and the N9H enol forms, with the keto form being calculated to be at most 1.2 kcal/mol lower in energy at the MP2 level and at least 0.3 kcal/mol lower at the CCSD(T) level. Upon N7 methylation, the order of stability remains the same for the nonmethylated species. In fact, the predicted energy difference clearly indicates that the N9H tautomer is the most stable form. Methylation of the N7 site seems to further stabilize the canonical form, casting doubt on whether this site actually causes any real changes in DNA replication.

N3 methylation causes significant changes in the order of stability of the N9H keto and enol forms. N3 methylation results in additional stability of the N9H enol form over the N9H keto tautomer. This methylation can have detrimental effects on the canonical base pairing and could promote point mutations with other nonclassical bases or tautomers. An intramolecular proton transfer could cause a disruption in the canonical three hydrogen bonds pattern that the GC base pair normally exhibits, creating a weaker interaction.

Adenine has similar responses to methylation as guanine. Table 3 compares the sites of protonation in adenine, the N9, N7, and N3 positions. N9 is the site of the sugar−adenine interaction and therefore will not be protonated in the DNA. However, isolated bases and nonclassical adenine forms are important to RNA; the formation of different triads and tetrads could be detected by spectroscopy. Therefore all three sites of protonation are displayed in this table. Without methylation, both DFT and MP2 methods reveal a moderate relative stability for the N9H conformation. The N7 and N3 protonated bases are higher in energy, and they are therefore not predicted to exist in any significant amount.

N1 methylation also favors the same order of stability with an increasing gap between the three species. The N3H form is significantly less stable relative to the N9H species, when methylated at the N1 site, versus unmethylated species. N3 methylation and N7 methylation, however, cause a reverse in the order of stability. The N7H protonated species is slightly more stable than the N9H protonated species in N3 methylation. The N3H protonated form is significantly more stable than the N9H protonated species in N7 methylation.

TABLE 5: Comparison of the N7 Protonated Tautomers of Adenine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}
		Nonmethylated Species		
N7H amino	0.00	0.00	0.00	0.00
	−467.206256	−467.237892	−465.865423	−465.897116
N7H imino	8.88	8.92	9.63	9.47
	−467.192094	−467.223681	−465.850070	−465.882028
		N3 Methylated Species		
N7 amino	0.00	0.00	0.00	0.00
	−506.877802	−506.911914	−505.397161	−505.431661
N7 imino	18.73	18.89	17.09	17.26
	−506.847953	−506.881805	−505.369922	−505.404154

[a] Relative energies are in kcal/mol; total energies are in au; see Figure 5.

TABLE 6: Comparison of the Amino, Hydroxy, and Imino Tautomers of Cytosine, Nonmethylated and Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}	$\Delta H^{CCSD(T)}$
			Nonmethylated Species		
2-oxo	0.00	0.00	1.59	1.38	1.93
	−394.842786	−394.873479	−393.708872	−393.739509	−393.932475
2-hydroxy	0.16	0.37	0.00	0.00	0.00
	−394.842524	−394.872885	−393.711413	−393.741714	−393.935561
4-imino	1.83	1.91	2.45	2.09	1.73
	−394.839863	−394.870440	−393.707508	−393.738388	−393.932796
			C5 Methylated Species		
2-oxo	0.22	0.03	2.27	2.08	1.65
	−434.135295	−434.167796	−432.867493	−432.899941	−433.119988
2-hydroxy	0.00	0.00	0.00	0.00	0.00
	−434.135645	−434.167842	−432.871103	−432.903263	−433.122622
4-imino	3.53	3.65	4.97	4.84	3.71
	−434.130025	−434.162024	−432.863190	−432.895553	−433.116714
			N3 Methylated Species		
2-oxo	0.00	0.00	0.00	0.00	0.00
	−434.505030	−434.537297	−433.234925	−433.267300	−433.484774
2-hydroxy	9.22	9.15	8.90	8.72	8.47
	−434.490333	−434.522711	−433.220737	−433.253397	−433.471265

[a] Relative energies are in kcal/mol; total Energies are in au; see Figure 6.

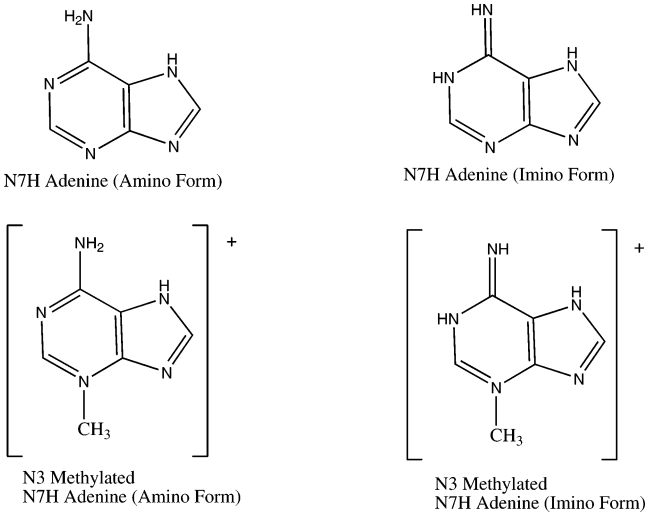

Figure 5. Molecular structures of N7H, amino and imino tautomers, nonmethylated and methylated (Table 5).

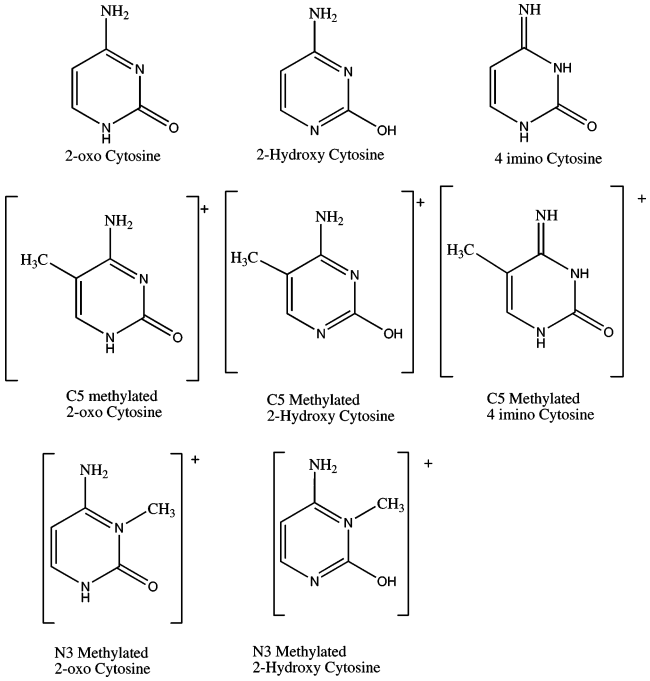

Figure 6. Molecular structures of 2-oxo, 2-hydroxy, and 4-imino cytosine, nonmethylated and methylated (Table 6).

Since the N9H and N7H protonated bases are the most prevalent in nature, we focus the investigation of methylation effects on tautomerization on these forms. Table 4 compares the tautomers of the N9H protonated species, methylated and nonmethylated. For the unmethylated NAB tautomers, the canonical amino form is significantly more stable, around 12.5 kcal/mol. In N3 methylation, this gap almost doubles with the imino form becoming significantly less stable. This is somewhat contradictory to experimental results. However, our study concludes that N3 methylation does not change the order of stability. This suggests that further studies are warranted to evaluate the N3 adenine methylation effects. N7 methylation results in the same order of stability with a slight decrease in

TABLE 7: Comparison of the Keto and Enol Tautomers of Thymine, Nonmethylated and O2 Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}
		Nonmethylated Species		
keto tautomer	0.0	0.0	0.0	0.0
	−454.033775	−454.066310	−452.622240	−452.771765
enol tautomer	12.90	13.02	12.94	13.30
	−454.013221	−454.045569	−452.601622	−452.750566
		O2 Methylated Species		
keto tautomer	8.90	8.79	8.07	8.33
	−493.643257	−493.678056	−492.049584	−492.256055
enol tautomer	0.00	0.00	0.00	0.00
	−493.657437	−493.692068	−492.062451	−492.242767

[a] Relative energies are in kcal/mol; total energies are in au; see Figure 7.

TABLE 8: Comparison of the Keto and Enol Tautomers of Uracil, Nonmethylated and O2 Methylated[a]

	ΔH^{DFT}	ΔG^{DFT}	ΔH^{MP2}	ΔG^{MP2}
		Nonmethylated Species		
keto tautomer	0.0	0.0	0.0	0.0
	−414.738605	−414.769029	−413.578147	−413.608741
enol tautomer	11.96	12.05	12.33	12.46
	−414.719539	−414.749832	−413.558495	−413.588880
		O2 Methylated Species		
keto tautomer	8.46	8.37	7.68	7.49
	−454.336085	−454.377824	−453.044386	−453.077186
enol tautomer	0.0	0.0	0.0	0.0
	−454.349740	−454.391156	−453.056619	−453.089129

[a] Relative energies are in kcal/mol; total Energies are in a.u; see Figure 8.

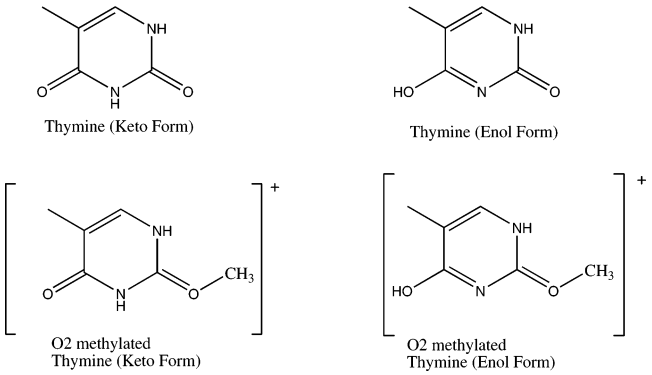

Figure 7. Molecular structures of thymine, keto and enol tautomers, nonmethylated and methylated (Table 7).

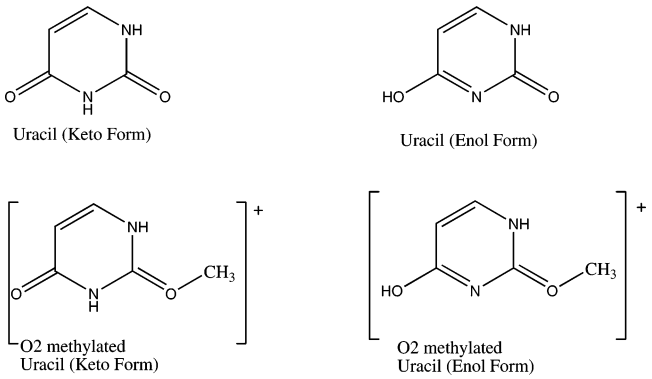

Figure 8. Molecular structures of uracil, keto and enol tautomers, nonmethylated and methylated (Table 8).

the gap between the two forms. This change is not very large and therefore suggested not to be of much consequence. Table 5 shows that even in N7 protonated species, the N3 methylation only increases the gap between the two tautomers by about 10 kcal/mol and causes no change in the order of stability. The inertness of the tautomeric changes in adenine questions the importance of adenine methylation and suggests that the effect that adenine methylation displays may not be as important as previously thought.

B. Pyrimidine Methylation. The unique characteristic of cytosine tautomers is the low relative energy of the three lowest tautomers (Table 6). These three tautomers are separated by no more than 4 kcal/mol at both the B3LYP and MP2 levels of theory. There are differences in the order of stabilities between the DFT and the MP2 levels of calculations. This discrepancy has been previously reported.[36] Density functional theory suggests that the 2-oxo form is the lowest while MP2 favors the 2-hydroxy form. The higher-level CCSD(T) method is in agreement with the MP2 level, suggesting that DFT may be problematic in suggesting the order of stability in cytosine tautomers. Also the CCSD(T) results obtained with a larger basis set, reported in earlier theoretical works, confirm the MP2 relative order of stability.[53,54] Even at higher levels, the gaps are small and suggest that both forms can exist in vivo.

Upon methylation at the C5 site, all three methods predict that the hydroxy form is the lowest in energy (Table 6), although once again DFT underestimates the MP2 and CCSD(T) results and suggests virtually equal energies. The relative order of stability predicted by the DFT calculations at the higher levels of theory is unchanged. The hydroxy form is slightly more stabilized upon C5 methylation. The stability of the imino form is unchanged.

O2 methylation is predicted to create changes in the energy differences. The hydroxy form is barred from being formed due to the methylation on the oxygen site. The relative energy difference of the imino tautomer widens to almost 8 times the energy gap compared to the relative stability of the parent form. O2 methylation not only blocks the formation of the hydroxy site but energetically makes the imino form unfavorable. N3 methylation also stabilizes the oxo form over the hydroxy tautomer. The obtained results suggest that the methylation could provide a mechanism for the promotion of the oxo form over other tautomeric structures.

Thymine (Table 7) and uracil (Table 8) have similar energetic characteristics. In both bases the canonical keto form is significantly stabilized over the enol form. This suggests that the enol tautomer is the predominant form that is present within the DNA and RNA structures. Upon O2 methylation, these trends are reversed. The enol form is overwhelmingly more stable than the keto form. This methylation causes a definite opportunity for point mutations. Our results are consistent with the findings that without the proper repair enzymes, thymine methylation can be detrimental to bacteria.[55]

Both uracil and thymine methylation stabilize the forms that could contribute toward the shift of the proton that is donated in base pairing to the oxygen atom that is also instrumental in base pairing. This creates a definite opportunity for these two bases to undergo nonclassical binding. Because the donating proton is now on another atom, and the nitrogen site that is normally attached to this proton becomes free, therefore facilitating the reverse base pairing. Ribonucleic acid, which displays many more deviations of the nucleic acid bases from their canonical forms, could possibly stabilize rare tautomers through uracil methylation. Again, the reverse in the stability trends revealed in our investigation suggests that further studies should be performed on the influence of thymine and uracil methylation.

4. Conclusion

A systematic high-level theoretical study of the five nucleic acid bases was performed using theoretical techniques to predict and compare the relative energies of methylated and unmethylated tautomers. These predictions show that methylation can influence the order of tautomer stability in these nucleic acid bases. The following conclusions summarize predicted trends:

(a) In guanine, N3 methylation reverses the relative energies of protonated bases, causing the N7 protonation site to be more stable than the N9H protonated species. O6 methylation merely stabilizes the canonical order of stability.

(b) In guanine, N3 methylation stabilizes the rare enol form in the N9H protonated species while N7 stabilizes the keto form.

(c) In adenine, N3 and N7 methylation stabilizes rare protonation sites while only N1 promotes the N9H canonical protonated base.

(d) In adenine, methylation in both the N9H and the N7H protonated species does not affect the order of relative stability.

(e) In cytosine, only C5 methylation favors the 2-hydroxy form. O2 and N3 methylation favor the 2-oxo Watson−Crick tautomer.

(f) O2 methylation in both uracil and thymine favor the enol form over the keto form.

This predicted stability could be an important factor in the development of point mutations, mispairing, and subsequent tumor formation. These postmethylation changes can also stabilize the canonical forms of the bases, making nonclassical pairing unlikely.

Acknowledgment. This work has been supported by the National Science Foundation (CREST Grant No. HRD-0318519), the National Institutes of Health (SCORE Grant No. SO6 GMO08047), and the Office of Naval Research (Grant No. N00014-03-1-0116). G.H. acknowledges support of the Chemical Materials and Computational Modeling Project funded by the Department of Defense through the U. S. Army Engineer Research and Development Center (Vicksburg, MS; Contract No. W912HZ-04-2-0002).

References and Notes

(1) Markova, N.; Enchev, V.; Timtcheva, I. *J. Phys. Chem. A* **2005**, *109*, 1981.
(2) Gorb, L.; Shishkin, O.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **2005**, *22*, 441.
(3) Podolyan, Y.; Gorb, L.; Leszczynski, J. *Int. J. Mol. Sci.* **2003**, *4*, 410.
(4) Balu, N.; Gamcsik, M. P.; Colvin, M. E.; Colvin, O. M.; Dolan, M. E.; Ludeman, S. M. *Chem. Res. Toxicol.* **2002**, *15*, 380.
(5) Cysewski, P. *J. Mol. Struct. (THEOCHEM)* **1999**, *466*, 49.
(6) Cysewski, P. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 3117.
(7) Cysewski, P.; Jeziorek, D.; Olinski, R. *J. Mol. Struct. (THEOCHEM)* **1996**, *369*, 93.
(8) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010.
(9) Florian, J.; Hrouda, V.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 1457.
(10) Poltev, V. I.; Gonzalez, E. H.; Teplukhin, A. V. *Mol. Biol.* **1995**, *29*, 213.
(11) Sigel, H. *Pure Appl. Chem.* **2004**, *76*, 1869.
(12) Blas, J. R.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 154.
(13) Grebneva, H. A. *J. Mol. Struct.* **2003**, *645*, 133.
(14) Legler, P. M.; Massiah, M. A.; Mildvan, A. S. *Biochemistry* **2002**, *41*, 10834.
(15) Venkateswarlu, D.; Leszczynski, J. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 373.
(16) Stark, A. A.; Rastetter, W. H. *Environ. Mol. Mutagen.* **1996**, *28*, 284.
(17) Leszczynski, J. *Chem. Phys. Lett.* **1991**, *181*, 123.
(18) Leszczynski, J.; Lammertsma, K. *J. Phys. Chem.* **1991**, *95*, 3128.
(19) Hill, G.; Forde, G.; Hill, N.; Lester, W. A.; Sokalski, W. A.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *381*, 729.
(20) Luo, R.; Gilson, M. K. *Biophys. J.* **1999**, *76*, A53.
(21) Hobza, P.; Kabelac, M.; Sponer, J.; Mejzlik, P.; Vondrasek, J. *J. Comput. Chem.* **1997**, *18*, 1136.
(22) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61*, 3.
(23) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
(24) Mons, M.; Dimicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *J. Phys. Chem. A* **2002**, *106*, 5088.
(25) Leszczynski, J. *J. Phys. Chem. A* **1998**, *102*, 2357.
(26) Sponer, J.; Sponer, J. E.; Gorb, L.; Leszczynski, J.; Lippert, B. *J. Phys. Chem. A* **1999**, *103*, 11406.
(27) Rueda, M.; Luque, F. J.; Lopez, J. M.; Orozco, M. *J. Phys. Chem. A* **2001**, *105*, 6575.
(28) Tian, S. X.; Xu, K. Z. *Int. J. Quantum Chem.* **2002**, *89*, 106.
(29) Sambrano, J. R.; de Souza, A. R.; Queralt, J. J.; Olive, M.; Andres, J. *Chem. Phys.* **2001**, *264*, 333.
(30) Sambrano, J. R.; de Souza, A. R.; Queralt, J. J.; Andres, J. *Chem. Phys. Lett.* **2000**, *317*, 437.
(31) Marnett, L. J.; Burcham, P. C. *Chem. Res. Toxicol.* **1993**, *6*, 771.
(32) Saffhill, R.; Margison, G. P.; Oconnor, P. J. *Biochim. Biophys. Acta* **1985**, *823*, 111.
(33) Boiteux, S.; Laval, J. *Biochem. Biophys. Res. Commun.* **1983**, *110*, 552.
(34) Oconnor, T. R.; Boiteux, S.; Laval, J. *Nucleic Acids Res.* **1988**, *16*, 5879.
(35) Tudek, B.; Boiteux, S.; Laval, J. *Nucleic Acids Res.* **1992**, *20*, 3079.
(36) Piacenza, M.; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83.
(37) Lindahl, T.; Sedgwick, B.; Sekiguchi, M.; Nakabeppu, Y. *Annu. Rev. Biochem.* **1988**, *57*, 133.
(38) Groudine, M.; Eisenman, R.; Weintraub, H. *Nature* **1981**, *292*, 311.
(39) Doerfler, W. *Biol. Chem. Hoppe−Seyler* **1991**, *372*, 557.
(40) Cedar, H. *Cell* **1988**, *53*, 3.
(41) Hergersberg, M. *Experientia* **1991**, *47*, 1171.
(42) Laird, P. W.; Jaenisch, R. *Hum. Mol. Genet.* **1994**, *3*, 1487.
(43) Jones, P. A.; Rideout, W. M.; Shen, J. C.; Spruck, C. H.; Tsai, Y. C. *Bioessays* **1992**, *14*, 33.
(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(45) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(46) Singer, B.; Chavez, F.; Goodman, M. F.; Essigmann, J. M.; Dosanjh, M. K. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8271.
(47) Singer, B.; Dosanjh, M. K. *Mutat. Res.* **1990**, *233*, 45.
(48) Loechler, E. L.; Green, C. L.; Essigmann, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6271.
(49) Ellison, K. S.; Dogliotti, E.; Connors, T. D.; Basu, A. K.; Essigmann, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8620.
(50) Pauly, G. T.; Hughes, S. H.; Moschel, R. C. *Biochemistry* **1991**, *30*, 11700.
(51) Pletsa, V.; Gentil, A.; Margot, A.; Armier, J.; Kyrtopoulos, S. A.; Sarasin, A. *Nucleic Acids Res.* **1992**, *20*, 4897.
(52) Price, S. L.; Locelso, F.; Treichel, J. A.; Goodfellow, J. M.; Umrania, Y. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 3407.
(53) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.
(54) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slavicek, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192.
(55) Mathieu, O.; Picard, G.; Tourmente, S. *Chromosome Res.* **2002**, *10*, 455.