Hydrogen-Bond Acid/Base Catalysis: A Density Functional Theory Study of Protonated Guanine-(Substituted) Cytosine Base Pairs as Models for Nucleophilic Attack on Mitomycin in DNA

J. J. Dannenberg* and Maria Tomasz

Contribution from the Department of Chemistry, City University of New York, Hunter College and The Graduate School, 695 Park Avenue, New York, New York 10021

Received August 30, 1999

Abstract: The mechanism of alkylation at the exocyclic nitrogen of guanine in G·C base pairs has been studied using density functional theory at the B3LYP/D95** level. Protonation of the amino group was used as a model for this reaction. The calculations indicate that the reaction is facilitated by a temporary transfer of the H-bonding hydrogen from the guanine amino position to the cytosine oxygen within the H-bond. Thus, the cytosine "loans" its basicity to the guanine within the H-bonded base pair. These calculations explain the previously observed dependence of guanine alkylation upon the substituent at the 5-position of cytosine. The generality of catalysis via the temporary transfer of a H-bonding hydrogen within an H-bond, hydrogen-bond, acid/base catalysis (HBA/BC), is discussed. This form of catalysis might be important in biochemistry, materials science, and the solid state.

Introduction

The study of a theoretical model for the alkylation of the amino group of guanine in DNA leads to both a more detailed understanding of the mechanism of this particular reaction, as well as, the manifestation of what may be a hitherto unrecognized form of acid/base catalysis that operates through stable hydrogen bonds (hydrogen-bond acid/base catalysis, HBA/BC). We shall first detail the investigation of this specific reaction. Then, we shall discuss the possible implications of HBA/BC to other problems of chemical interest.

In DNA the exocyclic amino group of guanine is a selective target for alkylation by a variety of cytotoxic and carcinogenic agents; among them, natural products with complex structures and reactive metabolites of polycyclic aromatic hydrocarbons. For maximum efficiency of alkylation the duplex structure of DNA is often required. Mitomycin C, an antitumor antibiotic is a prototypical example of such DNA-reactive agents. It alkylates and cross-links duplex DNA exclusively at the 2-amino group of guanines.[1] Recently, we suggested that the cytosine of the G·C base pair participates in the reaction between the activated mitomycin and the 2-amino group of guanine in DNA.[2,3] This suggestion was based on the observation that varying the 5-substituent of cytosine from CH_3 to H to F in duplex DNA affected the rates of the alkylation of the guanine by mitomycin. The rates of alkylation were observed to be in the order CH_3 > H > F. The data fit the Hammett linear free-energy ($\sigma-\rho$) relationship.[2,3] On this basis, we proposed that electronic effects of the cytosine-5 substituent were transmitted via G·C H-bonding to the guanine-amino group, thereby exerting an influence on the nucleophilic reactivity of this moiety (Figure 1). Similar rate enhancements upon 5-methyl substitution of cytosine were reported previously for other guanine N^2-alkylating agents such as the environmental carcinogen benzo-[*a*]pyrene diolepoxide (BPDE).[4,5] If our interpretation is correct, the modulation of the reactivity of guanine by its electronic interaction with cytosine via H-bonding may be a general property of duplex DNA.

In this paper, we report density functional theory (DFT) calculations on the protonation of the amino group of guanine within the G·C base pair. We consider this a reasonable model for nucleophilic attack by this group upon activated mitomycin **2** (Figure 1).[3,6] In this model the amino group is quaternized by a proton instead of an alkyl group. The model is justified in that nucleophilic attack by the amino group would give rise to an ammonium species at this center. Since the nucleophilic attack would presumably be endothermic, the transition state should resemble the ammonium species by the Hammond postulate.

While the purpose of this paper is not to simply calculate properties of the G·C base pair, this is clearly necessary for comparisons with the protonated species. Many ab initio calculations on the G·C base-pair have been previously reported.[7] Alkylated (in the same position, as well as others, as that considered here) bases and base pairs, some of which are thought to have mutagenic properties, have been the subjects of similar studies.[8] Ab initio reports of proton transfers in neutral base pairs[9] and the radical anions[10] have also appeared.

Methods

Molecular orbital calculations were performed using hybrid DFT methods at the B3LYP/D95(d,p) level. This method combines Becke's 3-parameter functional,[11] with the nonlocal correlation provided by the

(1) Tomasz, M. *Chem. Biol.* **1995**, *2*, 575−579.
(2) Johnson, W. S.; He, Q.-Y.; Tomasz, M. *Bioorg. Med. Chem.* **1995**, *3*, 851−860.
(3) Das, A.; Tang, K. S.; Gopalakrishnan, S.; Waring, M. J.; Tomasz, M. *Chem. Biol.* **1999**, *6*, 461−471.
(4) Dennisenko, M. F.; Chen, J. X.; Tang, M.-S.; Pfeifer, G. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3893−3898.
(5) Chen, J. X.; Zheng, Y.; West, M.; Tang, M.-S. *Cancer Res.* **1998**, *58*, 2070−2075.
(6) Suresh Kumar, G.; Lipmann, R.; Cummings, J.; Tomasz, M. *Biochemistry* **1997**, *36*, 14128−14136

10.1021/ja993145i CCC: $19.00
Published on Web 02/19/2000

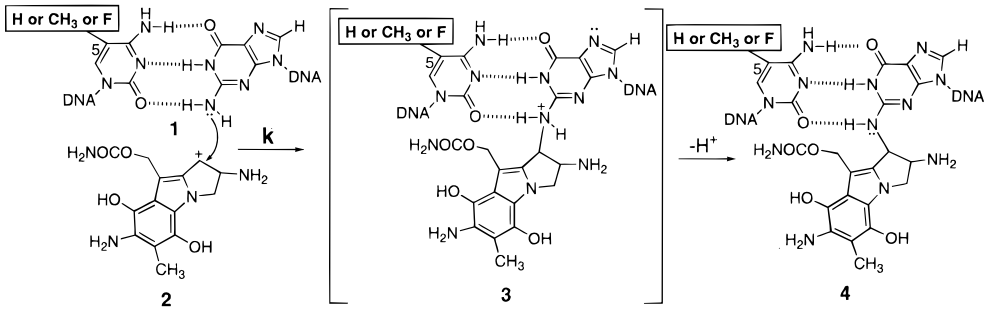


Figure 1. Presumed mechanism for alkylation of guanine by mitomycin.

correlation functional of Lee, Yang, and Parr.[12] This functional has led to better results[13,14] than the PW91LYP functional that we have used in the past.[15] We used the GAUSSIAN 98 suite of computer programs.[16] The geometries of all species were completely optimized with the constraint of the plane of the aromatic ring(s) taken as a symmetry plane. Calculation of vibrational force constants and the corresponding frequencies allowed us to obtain the enthalpy and free energies of the systems studied. The same vibrational calculations allowed us to verify the accuracy of the optimized minimum structures. All frequencies were found to be real with the exception of the those related to inversion about the NH_2's. Prior reports have noted the pyramidal structure of the amino group in guanine.[17] We have previously found that the planar structures of urea become minima after vibrational corrections.[18] For simplicity, we have assumed the same would occur in the cases studied in this report. The H-bonding energies have not been corrected for basis set superposition error (BSSE).[19] Only differences in the energies are important to the work presented in this paper. The BSSEs should approximately cancel in these differences. Furthermore, the CP correction is (a) controversial[20] and (b) usually added as a single-point correction without further optimization. This procedure does not find the correctly optimized structure.[21]

Table 1. B3LYP/D95(d,p) Protonation Energies in kcal/mol

species	energy	enthalpy (298 K)	free energy (298 K)
guanine	−200.9	−191.6	−192.3
cytosine	−239.5	−230.2	−230.4
fluorocytosine	−235.6	−226.3	−226.4
methylcytosine	−243.0	−233.7	−233.9
cytosine/guanine	−235.5	−227.9	−224.6
flurocytosine/guanine	−232.6	−225.1	−221.8
methylcytosine/guanine	−237.9	−230.2	−227.2

Table 2. Relative Protonation Energies at the Exocyclic Nitrogen of Guanine Calculated at the B3LYP/D95(d,p) Level (kcal/mol)

species	total energy	enthalpy (298k)	free energy (298 K)
cytosine/guanine	0	0	0
flurocytosine/guanine	2.8	2.8	2.7
methylcytosine/guanine	−2.4	−2.4	−2.6
guanine	34.5	36.2	32.2

Results

The protonation energies are collected in Tables 1−3 and illustrated in Figure 2. An expanded version of Table 1 which includes the total energies calculated for the various species, appears as Supporting Information. The properties and energies of the three base pairs are collected in Tables 4 and 5.

(7) (a) Clementi, E.; Mehl, J.; Von Niessen, W. *J. Chem. Phys.* **1971**, *54*, 508−20. (b) Del Bene, J. E. *THEOCHEM* **1985**, *25*, 201−12. (c) Hobza, P.; Sandorfy, C *J. Am. Chem. Soc.* **1987**, *109*, 1302−7. (d) Cieplak,; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 3734−9. (e) Aida, M. *J. Comput. Chem.* **1988**, *9*, 362−8. (f) Czerminski, R.; Kwiatkowski, J. S.; Person, W. B.; Szczepaniak, K. *J. Mol. Struct.* **1989**, *198*, 297−305. (g) Dive, G.; Dehareng, D.; Ghuysen, J. M. *Theor. Chim. Acta* **1993**, *85*, 409−21. (h) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493−9. (i) Santamaria, R.; Vazquez, A. *J. Comput. Chem.* **1994**, *15*, 981−96. (j) Jiang, S.-P.; Raghunathan, G.; Ting, K.-L.; Xuan, J. C.; Jernigan, R. L. *J. Biomol. Struct. Dyn.* **1994**, *12*, 367−82. (k) Santamaria, R.; Quiroz-Gutierrez, A.; Juarez, C. *THEOCHEM* **1995**, *357*, 161−70. (l) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965. (m) Brameld, K.; Dasgupta, S.; Goddard, W. A., III. *J. Phys. Chem. B.* **1997**, *101*, 4851−4859. (n) Florian, J.; Sponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884−892.

(8) (a) Pullman, A.; Armbruster, A. M. *Theor. Chim. Acta* **1979**, *50* 359−61. (b) Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 367−71. (c) Miertus, S.; Trebaticka, *J. Theor. Biol.* **1984**, *108*, 509−17. (d) Hobza, P.; Sandorfy, *Collect. Czech. Chem. Commun.* **1988**, *53*, 1943−5. (e) Stewart, E. L.; Foley, C. K.; Allinger, N. L.; Bowen, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 7282−6. (f) Bakalarski, G.; Grochowski, P.; Kwiatkowski, J. S.; Lesyng, B.; Leszczynski, J. *Chem. Phys.* **1996**, *204*, 301−11. (g) Broch, H.; Hamza, A.; Vasilescu, D. *Int. J. Quantum Chem.* **1996**, *60*, 8, Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology, **1996**, 21−40.

(9) (a) Kong, Y. Shik; J., Mu S.; Lowdin, P. O. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1987**, *14*, 189−209. (b) Florian, J.; Leszczynski, J.; Scheiner, S. *Mol. Phys.* **1995**, *84*, 469−80. (c) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010−17.

(10) (a) Colson, A. O.; Besler, B.; Close, D. M.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 661−8. (b) Colson, A. O.; Besler, B.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 9787−94. (c) Hutter, M.; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*, 7574−7577. (d) Bertran, J.; Oliva, A.; Rodriguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159−8167.

(11) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(12) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(13) Pudzianowski, A. T. *J. Phys. Chem.* **1996**, *100*, 4781.

(14) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Phys. Chem. A* **1999**, *103*, 1640.

(15) Sordo, T. L.; Dannenberg, J. J. *J. Org. Chem.* **1999**, *64*, 1922.

(16) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheesman, J. R.; Keth, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Lahan, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslwski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andre, J. L.; Replogle, E. S., Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Deefres, D. J.; Baker, J.; Stewar, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. ***Gaussian 94***; Gaussian, Inc.: Pittsburgh, PA, 1995.

(17) (a) Sponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **1996**, 827−833. (b) Sponer, J.; Hobza, P. *THEOCHEM* **1994**, 35−40. (c) Leszczynski, *J. Int. J. Quantum Chem., Quantum Biol. Symp.* **1992**, *19*, (*Proc. Int. Symp. Appl. Fund. Theory Probl. Biol. Pharmacol.* **1992**), 43−55

(18) Masunov, A.; Dannenberg, J. J. *J. Phys. Chem. A* **1999**, *103*, 178−84

(19) (a) Jansen, H. B.; Ross, P. *Chem. Phys. Lett.* **1969** *3*, 40. (b) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970** *19*, 553. (c) Meunier, A.; Levy, B.;. Berthier, G. *Theor. Chim. Acta* **29**, *49* 1973.

(20) (a) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1984**, *82*, 2418. (b) Frisch, M. J.;. Del Bene, J. E.; Binkley, J. S.;. Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 2279. (c) Szalewicz, K.; Cole, S. J.; Kolos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988,** *89*, 3662. (d) Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1993**, *97*, 2488. (e) van Lenthe, J. H.; van Duijneveldt-van Rijdt, J. G. C. M.; van Duijneveldt, F. B *Ab Initio Methods in Quantum Chemistry*; Lawley, K. P., Ed.; John Wiley & Sons: New York, 1987; Vol. II. (f) Gutowski,, M.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *J. Chem. Phys.* **1993**, *98*, 4728. (g) Cook, D. B.;. Sordo, J. A.; Sordo, T. L. *Int. J. Quantum Chem.* **1993**, *48*, 375. (h) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873.

(21) See for example: Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024 and references therein.

Table 3. Relative Protonation Energies at the Oxygen of (substituted) Cytosines Calculated at the B3LYP/D95(d,p) Level (kcal/mol)

species	total energy	enthalpy (298 K)	free energy (298 K)
cytosine	0	0	0
fluorocytosine	4.3	3.9	3.9
methylcytosine	−3.4	−3.5	−3.7

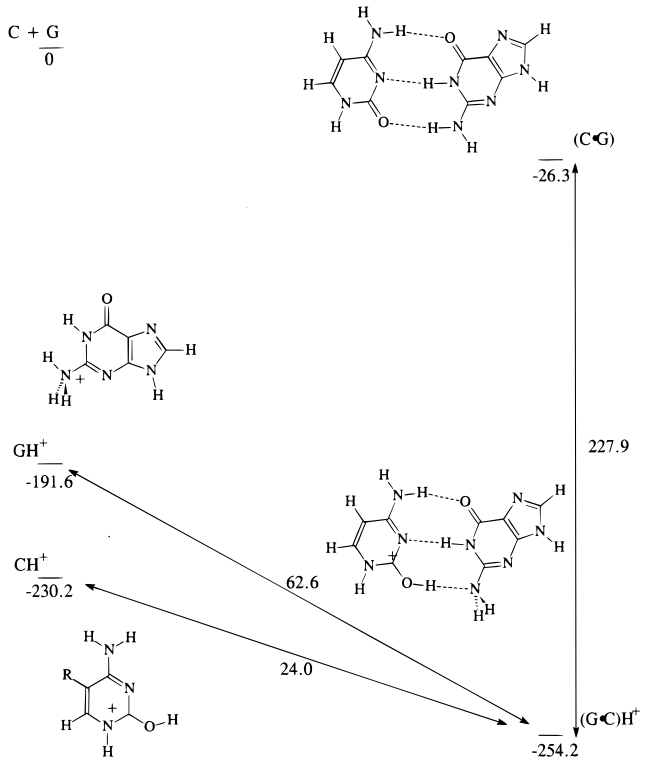


Figure 2. Relative enthalpies at 298 K for various species in kcal/mol. Zero is defined as the sum of the enthalpies of the isolated bases, G and C. The values associated with the double-headed arrows indicate differences in enthalpy.

As seen from Table 2, the enthalpy of protonation at the amino group of G in the G·C base pair is more favorable than protonation at the analogous position in monomeric guanine by 36.2 kcal/mol. The protonation at the amino group of guanine in the base pair leads immediately to a transfer of a proton from the guanine to the carbonyl of the cytosine to which it is H-bonded. We were completely unable to find a local minimum that corresponds to the protonated base pair prior to proton transfer. The resulting H-bonding complex can be thought of as one between a protonated cytosine and a distorted guanine (Figure 2; **7**). This striking result can be understood from the following: Protonation of monomeric guanine at 2-NH_2 leads to a quaternary nitrogen without a lone-pair of electrons to conjugate with the ring. As might, therefore, be expected, prior calculations indicate that guanine is best protonated elsewhere.[22] Consequently, guanine is an extremely weak base ($pK_a < 0.14$) at this position.[23] However, when participating in a G·C base pair, proton transfer from the N to the cytosine O allows the base pair to delocalize the charge throughout the π-system of the cytosine. As a result, the *cytosine of the G·C base pair replaces the guanine in the role of the base*. An analogous proton transfer within the G·C base pair, from guanine-N_1 to cytosine-N_3, was proposed for the Ag^+−guanine complex based on spectroscopic evidence.[24]

The basicities of the substituted cytosines considered here have been calculated (see Tables 1 and 2). The proton affinity of cytosine (at C═O) is substantially greater ($\Delta\Delta H = 38.6$ kcal/mol) than that of guanine (at the amino group). One might, therefore, suspect that nucleophilic attack at this position of cytosine might be favorable. However, much like the reaction at the oxygen of an enolate anion, such a nucleophilic attack might not lead to a stable product. *In the G·C base pair, the cytosine C═O can be used to mediate the nucleophilic substitution reaction by "loaning" its basicity to the guanine*. This substantially reduces the energy of the quaternary ammonium intermediate whose formation is likely to be the slow step in the overall process.

This "loaned" basicity can stabilize an intermediate in a manner that is different from, but somewhat analogous to, the strong low-barrier hydrogen bonds (LBHBs) that have been suggested to explain certain enzymatic reaction paths.[25] These have received much recent attention (and precipitated significant controversy). According to one of the suggestions that have been promulgated to explain the unusual stability of LBHBs, the H-bonding strength should decrease with increasing disparity between the pK_a's of the H-bonding species.[26] In the protonation of the G·C base pair, we propose that HBA/BC might be an alternative to the LBHB mechanism for catalysis through H-bonding.

Comparison of the results in Table 2 with those in Table 3 indicate that, although they follow the same qualitative order, the differences in the proton affinities at oxygen of the three cytosines are greater than those of the similarly substituted G·C base pairs. This apparent anomaly is due to the fact that the amino group of the guanine that becomes protonated (as it transfers another proton to the cytosine) is no longer conjugated to the π-system. Rather, its lone-pair electrons, which are involved in the H-bond, are in the plane of the σ-bonds. Were the protonated base pair to dissociate (into guanine and protonated cytosine), the amino group of guanine would rotate to restore the lost conjugation. Ab initio studies of protonation at other sites of guanine have been reported.[22,27] These indicate that guanine is preferentially protonated elsewhere. Previous calculations on protonation of cytosine suggest that protonation at O and at N_3 are energetically similar.[22] We have calculated the enthalpy of protonation at O to be more favorable than at N_3 by 0.2 kcal/mol using B3LYP/D95(d,p). At the highest level reported in the literature, MP4/6311++G(d,p)/MP2/6-31G(d), ΔH for protonation at O is 0.6 kcal/mol more favorable than protonation at N_3.[22c] The proton affinity of cytosine at this level was reported[22c] to be 227 kcal/mol, slightly less than our value of 230.2 and slightly more than experimental reports of 223.8 (for cytosine)[28] and 225.9 (for cytidine)[29] kcal/mol.

(22) (a) Lavery, R.; Pullman, A.; Pullman, B. *Theor. Chim. Acta* **1978**, *50*, 67−73. (b) Del Bene, J. *J. Phys. Chem.* **1983**, *87*, 367−71. (c) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811−6821. (d) Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. *J. Comput. Chem.* **1998**, *19*, 989−1000. (e) Gorb, L.; Leszczynski, J. *Int. J. Quantum Chem.* **1998**, *70*, 855−862.

(23) Shapiro, R. In *Progress in Nucleic Acid Reseaech and Molecular Biology*; Davidson, J. N., Cohn, W. E., Eds.; Academic Press: New York, 1968; Vol. 8, pp 73−112.

(24) Sowers, L. C.; Shaw, B. R.; Veigl, M. L.; Sedgwick, W. D. *Mutation Res.* **1987**, *177*, 201−8.

(25) For a minireview", see: Cleland, W. W.; Frey, P. A.; and Gerlt, J. A. *J. Biol. Chem.* **1998**, 273, 25529.

(26) (a) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1887. (b) Frey, P. A. *Science* **1995**, *269*. 104. Also see: (c) Zeegers-Huyskens, T. *J. Org. Chem.* **1999**, *64*, 4986. (d) Chen, J.; McAllister, M. A.; Lee. J. K.; Houk, K. N. *J. Org. Chem.* **1998**, *63*, 4611. (e) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163−70 and references therein.

(27) Del Bene, J. E. *THEOCHEM* **1985**, *25*, 201−12.

(28) Lias, S.; Liebmann, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695.

Table 4. Selected Parameters of H-bonding Base-pairs

Charges[a]

R	O	N_3	N (exo)[b]	H (hb)[c]	H[d]	F/Me[e]	C_2 (O)	C_4 (N)
				Cytosine				
H	−0.309	−0.123	−0.454	0.250	0.272		0.084	0.208
F	−0.303	−0.115	−0.455	0.272	0.276	−0.196	0.089	0.116
methyl	−0.313	−0.122	−0.463	0.262	0.270	−0.100	0.073	0.172
relative to H								
F	0.005	0.009	0.000	0.022	0.004	−0.196	0.005	−0.092
methyl	−0.004	0.002	−0.009	0.011	−0.002	0.000	−0.011	−0.037
				Cytosine in G·C Base Pair				
H	−0.384	−0.219	−0.450	0.344	0.246		0.159	0.231
F	−0.379	−0.216	−0.451	0.347	0.266	−0.186	0.160	0.136
methyl	−0.389	−0.215	−0.459	0.341	0.257	−0.083	0.150	0.114
relative to H								
F	0.005	0.003	0.000	0.003	0.020	−0.186	0.001	−0.095
methyl	−0.005	0.004	−0.008	−0.003	0.012	−0.083	−0.009	−0.118

	O	N_1	N (exo)[b]	H (hb)[c]	H[d]	C_2 (N)	C_6 (O)
			Guanine				
	−0.298	−0.223	−0.464	0.256	0.277	0.172	0.122
			Guanine in G·C Base-Pair				
H	−0.392	−0.223	−0.473	0.324	0.255	0.142	0.156
F	−0.392	−0.221	−0.473	0.323	0.256	0.144	0.158
methyl	−0.392	−0.215	−0.473	0.326	0.254	0.142	0.158
relative to H							
F	0.000	0.002	0.000	−0.001	0.001	0.002	0.002
methyl	0.000	−0.008	0.000	0.002	−0.001	0.000	0.002

H−Bond Distances (Å)[f]

	A			B			C		
	N−H..O	H..O	N−H	NH,N	H..N	N−H	NH..O	H..O	N−H
H	2.762	1.719	1.042	2.916	1.879	1.038	2.904	1.879	1.025
F	2.750	1.706	1.045	2.917	1.882	1.036	2.924	1.901	1.024
methyl	2.762	1.719	1.042	2.916	1.879	1.038	2.890	1.865	1.020
relative to H									
F	−0.012	−0.014	0.002	0.001	0.003	−0.002	0.021	0.022	−0.001
methyl	0.000	0.000	0.000	0.000	−0.001	0.000	−0.014	−0.014	−0.005

[a] From Mulliken populations. [b] Exocyclic nitrogen. [c] H-bonding hydrogen on NH_2. [d] Non-H-bonding hydrogen on NH_2. [e] Total charge on the substituent. [f] See Figure 3 for definitions of A, B, and C.

Table 5. Calculated H-Bonding Energies (B3LYP/D95**) for the Three Base Pairs (uncorrected for BSSE) in kcal/mol[a]

	ΔE	ΔE (ZPVE)[b]	ΔH	ΔG
H	−29.3	−26.9	−26.4	−15.7
F	−28.8	−26.4	−25.9	−15.1
methyl	−29.9	−27.4	−26.9	−16.0

[a] Enthalpy and free energy at 298 K. [b] Corrected for zero-point vibrational energies.

The effect of substitution on the charge distribution and H-bond lengths of the neutral base-pairs (Table 4) are noticeable, but not very large. In themselves, they do not provide a plausible explanation for the experimental observations discussed above. All of the changes in atomic charges (Mulliken populations) listed are greater for the isolated cytosine species than for the corresponding cytosine in the base pair except for the H-bonding hydrogen and the exocyclic nitrogen. Fluorine substitution slightly shortens H-bond A and lengthens H-bond C (see Figure 3). Methyl substitution has little effect upon H-bond A, but slightly shortens H-bond C. Neither substitution has any significant effect upon H-bond B. Methyl substitution strengthens the H-bond by about 0.5 kcal/mol, while fluorine substitution weakens it by about the same amount. The values listed in Table 5 are not corrected for BSSE; this error should cancel when differences are taken.

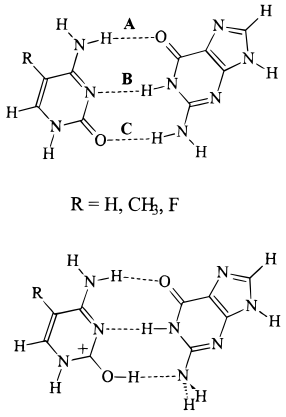


Figure 3. Protonation model for the alkylation of the G·C base pair. For the designations of the hydrogen bonds (A, B, C), refer to Table 4.

Although the precise H-bonding interactions of the G·C base pairs are not directly relevant to our purpose (only the differences are), we have included this information for completeness. The data in Tables 4 and 5 (which are uncorrected for BSSE) indicate that the present calculations predict some-

(29) Greco, F.; Liguori, A.; Sindona, G.; Uccella, N. *J. Am. Chem. Soc.* **1990**, *112*, 9092. Reference 22c incorrectly cites this value as the proton affinity of cytosine.

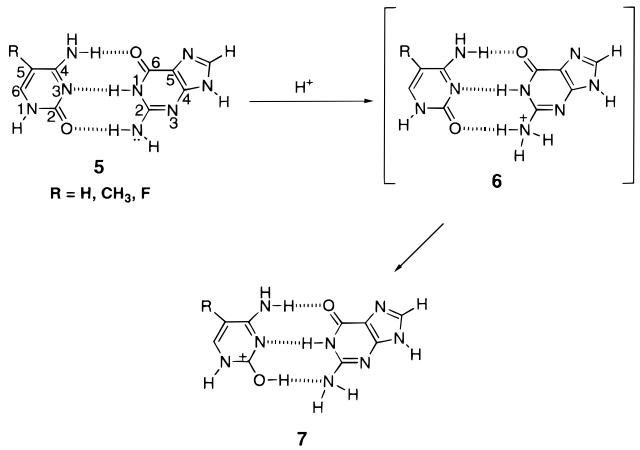


Figure 4. Protonation of the amino group of guanine in the G·C base-pair. Cytosine becomes the base.

what stronger interactions than the best (uncorrected) value reported by Goddard of 24.9 kcal/mol for LMP2/cc-pVTZ(-f).[7m] The H-bonding distances are also somewhat shorter than the reported crystallographic values.[30] To our knowledge, no appropriate gas-phase data is available. Calculated H-bonding interaction distances that have not been geometrically optimized on a CP-corrected surface are generally too short.[31] This effect is particularly noticeable for DFT calculations. For example, the O···O H-bonding distance of B3LYP/d95++** calculations on water dimer is too short before BSSE correction. It increases when optimized on the CP-corrected potential surface.[14] The cumulative BSSEs of the three H-bonds, together with the smaller basis set used in these larger systems, should accentuate the underestimation of the H-bonding distances in the present calculations.

Discussion

Mechanism of Alkylation of the Guanine 2-Amino Group. Protonation of the exocyclic amino group of guanine in the G·C Watson−Crick base pair is 36.2 kcal/mol[32] more favorable than the analogous protonation in guanine. The greater stability of the protonated guanine in the base pair form is due to the transfer of a proton from the guanine-$NH_3(+)$ group to the O^2 atom of the cytosine component of the base-pair. The resulting structure is a protonated G·C base-pair in which the O^2-protonated cytosine is now the H-bond donor and the guanine 2-NH_2 is the acceptor (see Figure 4). These findings provide a novel insight into the mechanism for the alkylation of the guanine 2-amino group. Furthermore, they suggest that analogous catalytic action might be provided by other stable H-bonding systems. In fact, the concept of HBA/BC might prove to be generally important.

Since the proton affinity of cytosine at C═O is much greater than that of guanine at the 2-amino group (by 38.6 kcal/mol), one might suspect that nucleophilic attack at this position of cytosine might be favorable. However, much like the reaction at the oxygen of the enolate anion, such a nucleophilic attack would lead to a relatively unstable product, that is, the "imino tautomer" of cytosine. In the G·C base pair, the cytosine C═O can mediate the nucleophilic substitution reaction at guanine-N^2 by "loaning" its basicity to the guanine by means of hydrogen-bond base catalysis (see discussion below). This substantially stabilizes the (putative) quaternary ammonium intermediate of the alkylation by mitomycin, **3**, whose formation is likely to be the slow step in the overall process. Figure 5 depicts the mechanism in general terms: Transfer of the H-bonded guanine-N^2 proton to the cytosine facilitates the formation of the positively charged alkylation intermediate **8** which is likely to be the rate-determining step. This mechanism serves as the likely explanation for the earlier experimental observation that the electronic properties of the cytosine substituents affect the rate of N^2 alkylation of the base-pair guanine.[2,3] The proposed mechanism further explains why the basicity of the cytosine affects the nucleophilicity of the G-NH_2 group. The cytosine, thus, assumes a catalytic role within the base-pair structure. Guanine-N^2 alkylation by ectenaiscidin was proposed to be catalyzed by proton transfer to a basic N in the ectenaiscidin molecule itself rather than to the base-pair cytosine as proposed in the present work.[33] Since single-stranded DNA can be alkylated at guanine-N^2 by mitomycin,[34] BPDE[35] and other agents, the catalytic proton transfer by cytosine may not be absolutely required for all such reactions. Quantitative experimental assessment of the effect of HBBC from comparison of the alkylation rates of single- and double-stranded DNA is not feasible, due to the numerous variables which can influence the reaction rates. However, the results of the mitomycin alkylation experiments, utilizing 5-substitution of cytosine as the experimental variable,[1−5] provide a qualitative demonstration of the existence of catalysis by cytosine in duplex DNA. The present calculations elucidate the mechanism.

There are clear differences between the isolated base pairs of a theoretical study and the corresponding base pairs in native or crystalline DNA. For example, the present theoretical study assumed the base pairs to be planar. This assumption is justified by the allusion to our previous studies of urea (where pyramidal amino groups became planar after appropriate corrections were made) and the fact that the only imaginary force constants calculated for the base pairs in this study involved the pyramidalization of the amino groups. Thus, the twisted base pairs seen in some crystal structures are not likely to be minima for the isolated base pairs.[36] The observed structures are likely due to the influence of other factors such as interaction with nearest neighbor molecules and molecular fragments such as neighboring base pairs, the DNA backbone, and water. Nevertheless, the calculated relative energies of protonation are in excellent agreement with the experimental observations. Furthermore, these differences are readily understood from hydrogen-bond base catalysis (discussed in detail below).

Hydrogen-Bond Acid/Base Catalysis (HBA/BC). The alkylation of the guanine in the G·C base pair is an example of a definable class of catalysis, which we shall call hydrogen-bond acid/base catalysis (HBA/BC). These catalytic reactions are illustrated schematically in Figure 6. We define HBA/BC as a reaction in which a proton that is involved in an H-bond is (a) first transferred from the H-bond donor to the H-bond acceptor

(30) Rosenberg, J. M.; Seeman, N. C.; Dat, R. O.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 145−67. Also see: Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, Berlin, Heidleberg, 1991; pp 259−66 and references therein.

(31) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024.

(32) Unless otherwise noted, all energetic values refer to enthalpies at 298 K.

(33) Moore, B. M., II; Seaman, F. C.; Wheelhouse, R. T.; Hurley, L. H. *J. Am. Chem. Soc.* **1998**, *120*, 2490−2491.

(34) Tomasz, M.; Lipman, R.; Verdine, G. L.; Nakanishi, K. *Biochemistry* **1986**, *25*, 4337−4344.

(35) Pirogov, N.; Shafirovich, A.; Kolbanovskiy, A.; Solntsev, K.; Courtney, S. A.; Amin, S.; Geacintov, N. E. *Chem. Res. Toxicol.* **1998**, *11*, 381−388

(36) Luisi, B.; Orozco, M.; Sponer, J.; Luque, F.; Shakked, Z. *J. Mol. Biol.* **1998**, *279*, 1123.

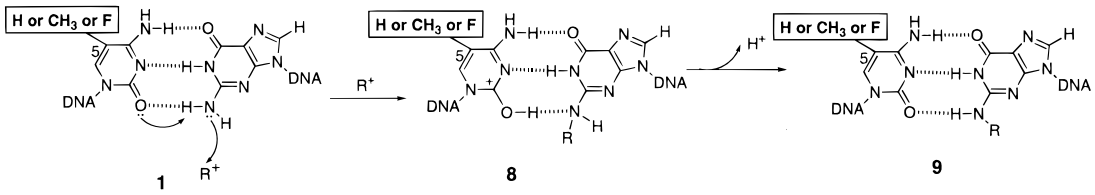


Figure 5. Mechanism for alkylation of guanine invoking hydrogen-bond base catalysis.

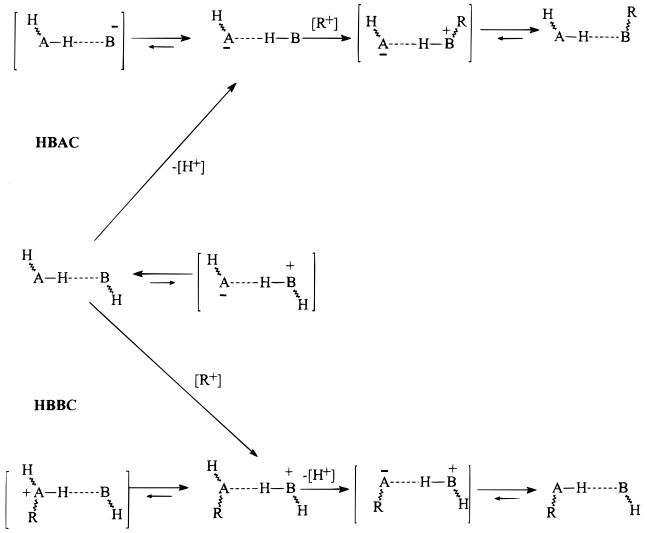
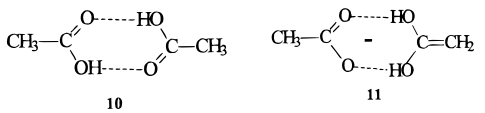


Figure 6. Simple scheme for hydrogen-bond acid catalysis (HBAC) and hydrogen-bond base catalysis (HBBC). The H's and/or R's are not necessarily meant to be on the H-bonding atoms. They can be in any appropriate position on A or B. The species in brackets are not minima on the potential energy surface that we have studied. However, they could be minima for other reactions involving HBA/BC mechanisms.

within the H-bond and (b) then transferred back to its original position in a later stage of the reaction.

The lower section of Figure 6 illustrates HBBC, the base-catalyzed process occurring in the alkylation of guanine in the G·C base pair. In this mechanism, the H-bonded complex **A**·**B** is attacked by an electrophile, R^+, at **A** (the H-bond donor) which transfers the H-bonding proton to **B**. **A** subsequently loses a different proton, while the H-bonding proton is returned to it by **B**. The H-bond remains intact throughout the process, although the donor/acceptor relationships are reversed twice.

The upper portion of Figure 6 outlines HBAC, the analogous scheme of acid-catalyzed alkylation of the H-bonded complex A·B. In this process, a proton is removed from the H-bond acceptor, **B**, while the H-bond donor transfers the H-bonding proton to **B**. This process produces an anion of the original H-bond donor which has become the *acceptor* of the new H-bond. **B** has essentially exchanged the H-bonding proton for the one it has lost. Thus, **B** has now become the H-bond *donor*. An electrophile, R^+, attacks **B**, which returns the H-bonding proton to its original position on **A**. In this mechanism, the original H-bond donor, **A**-H, loans its H-bonding proton to **B** when the latter is deprotonated. **A** recovers this proton when **B** reacts with electrophile.



An example of HBAC might be the dissociation of a proton from the methyl group of acetic acid dimer, **10**. We have suggested that the carbon acidity of acetic acid should be greatly enhanced in the dimer (relative to the monomer), due to a proton transfer in the dimer leading to the anion **11**.[37] One might immediately object to the suggestion that the acetic acid dimer would deprotonate at carbon in solution, as it could more easily deprotonate at oxygen, even at the cost of a hydrogen bond. Thus, although HBAC would substantially lower the activation barrier for reactions that might proceed via deprotonation at carbon, deprotonation at oxygen would still prevail, rendering these reactions difficult in solution. In the solid state, however, the situation might be quite different. Breaking a hydrogen bond in the solid state might be substantially more disruptive than breaking the analogous hydrogen bond in solution. Hydrogen bonds are generally more durable in the solid state than in solution. Many biochemical entities resemble the solid state in this regard. The $^{15}N\cdots^{15}N$, $^{13}C\cdots^{15}N$, and $^{1}H\cdots^{15}N$ coupling across hydrogen bonds in several such systems confirms the relatively long life of these interactions.[38−40]

These processes recall acid/base catalysis, which are generally classified as specific or general acid base catalysis. In specific acid/base catalysis, the reaction rate is first order in the lyonium or lyate ion concentration. In general acid/base catalysis, the rate depends on a sum of terms, each of which is first order in the concentration of a different acid or base. In HBA/BC, there is no kinetic dependence upon acid or base concentration as the acid and base units are associated by a hydrogen bond before the reaction. Thus, HBA/BC would be unperceived by kinetics if the concentration of H-bonded species were constant.

In the present study, we could not locate any local minima that correspond to the protonated base pairs prior to proton transfer. We therefore presume that proton transfer occurs as the guanine is attacked, without the intervention of any intermediate. In general, this might not always be true. In cases where proton transfer occurs after the initial attack, the activation barrier for this transfer might have kinetic consequences. This barrier (as well as the equilibria between the two H-bonded species) might be usefully studied using isotope effects.

We suggest that HBA/BC might be a previously unrecognized form of catalysis that might have eluded detection since there is no implied kinetic law involving acid or base. HBA/BC is more likely to be important in large assemblies of molecules, enzymes or in the solid state than in solution (where H-bonds are subject to dynamic exchanges). We suggest that these phenomena might be particularly important in biochemistry and in materials science.

As mentioned above, HBA/BC might be an alternative explanation to the low-barrier hydrogen bond (LBHB) suggestion used to explain certain enzymatic reactions.[25] Instead of a strong hydrogen bond providing the stabilization for the

(37) Dannenberg, J. J.; Paraskevas, L.-R.; Sharma, V. submitted for publication.

(38) Dingley, A. J.; Grzesiek, A. *J. Am. Chem. Soc.* **1998**, *120*, 8293; Cordier, F.; Grzesiek, A. *J. Am. Chem. Soc.* **1999**, *121*, 1601.

(39) Cornilescu, G.; Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949.

(40) Pervushin, K.; Ono, A.; Fernández, C.; Szyperski, T.; Kainosho, M.; Wüthrich. K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147−51.

enhanced acidity of certain hydrogens, this enhancement might be due to a loan of acidity from a H-bonded neighbor by means of HBAC. In this case, the apparent acidity of the proton abstracted would, in reality, be related to the acidity of the H-bonding neighbor that transfers the H-bonding proton (see Figure 6).

Experimental verification of HBA/BC might be accomplished by means of further experiments, similar to those discussed above, where structural modification of one of the H-bonding partners that leads to enhanced acidity (or basicity) has the appropriate effect on a reaction site located on a different H-bonding partner.

Conclusions

Substitution of the cytosine with methyl or fluorine in place of hydrogen affects the alkylation of the amino group of guanine by mitomycin and other agents in G•C base pairs of DNA.[2−5] This alkylation occurs as a result of hydrogen bond base catalysis (HBBC), a newly recognized variation of base catalysis. The experimental observations are due to a catalytic loan of basicity of the (substituted) cytosine to the guanine through one of the hydrogen bonds.

This newly recognized form of catalysis, together with its H-bond acid catalysis (HBAC) counterpart is likely to be quite general. They should be expected to occur preferentially in systems where the hydrogen bonds are relatively stable, such as the solid state, biological systems, and molecular aggregates which are important to materials science.

Acknowledgment. This work was supported by PSC-CUNY and NCI Grant CA28681. We thank Dr. Richard W. Franck for helpful discussions.

Supporting Information Available: Expanded Table 1 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA993145I