J|A|C|S
ARTICLES
Published on Web 07/31/2002

Electron Affinity of the Guanine−Cytosine Base Pair and Structural Perturbations upon Anion Formation

Nancy A. Richardson,[†] Steven S. Wesolowski,[‡] and Henry F. Schaefer, III*,[†]

Contribution from the Center for Computational Quantum Chemistry, University of Georgia, Athens, Georgia 30602-8107, and Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Received January 3, 2002. Revised Manuscript Received May 23, 2002

Abstract: The adiabatic electron affinity (AEA) for the Watson−Crick guanine−cytosine (GC) DNA base pair is predicted using a range of density functional methods with double- and triple-ζ plus polarization plus diffuse (DZP++ and TZ2P++) basis sets in an effort to bracket the true electron affinity. The methods used have been calibrated against a comprehensive tabulation of experimental electron affinities (*Chem. Rev.* **2002**, *102*, 231). Optimized structures for GC and the GC anion are compared to the neutral and anionic forms of the individual bases as well as Rich's 1976 X-ray structure for sodium guanylyl-3′,5′-cytidine nonahydrate, GpC·9H_2O. Structural distortions and natural population (NPA) charge distributions of the GC anion indicate that the unpaired electron is localized primarily on the cytosine moiety. Unlike treatments using second-order perturbation theory (MP2), density functional theory consistently predicts a substantial *positive* adiabatic electron affinity for the GC pair (e.g., TZ2P++/B3LYP: +0.48 eV). The stabilization of C^- via three hydrogen bonds to guanine is sufficient to facilitate adiabatic binding of an electron to GC and is also consistent with the positive experimental electron affinities obtained by photoelectron spectroscopy of cytosine anions incrementally microsolvated with water molecules. The pairing (dissociation) energy for GC^- (35.6 kcal/mol) is determined with inclusion of electron correlation and shows the anion to have greater thermodynamic stability; the pairing energy for neutral GC (TZ2P++/B3LYP 23.9 kcal/mol) compares favorably to previous MP2/6-31G* (23.4 kcal/mol) results and a debated experiment (21.0 kcal/mol).

I. Introduction

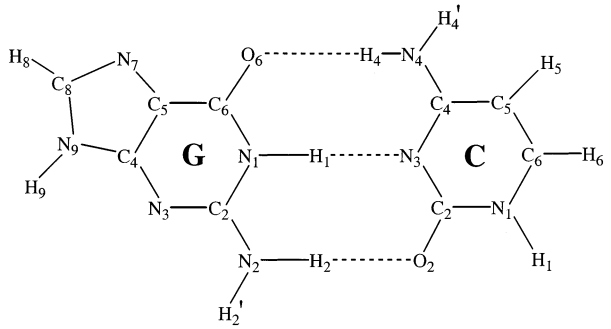
High-energy radiation damage to DNA is suspected to proceed from the formation of transient charged radicals within the strand. Specifically, electron trapping within nucleobase sites is believed to play a key role in DNA damage and repair since the cascade of reactions leading to mutations likely stems from acquisition of excess charge.[1−4] Furthermore, accurate electron affinities for nucleic acid base pairs and the distribution of excess electron sites are essential for prediction of electron- and hole transfer in DNA.[5−11]

Individual nucleic acid bases (NABs) are observable in the gas phase, although precise experimental determination of their electron affinities is difficult.[12−17] Theoretical work at varying levels of sophistication has complemented these studies.[18−24] The combination of elegant spectroscopic studies and increasingly rigorous computational efforts has provided insights into the nature of electron attachment to the individual NABs. Despite overwhelming evidence that conventional covalent anions of the nucleobases exist in solution and the solid state, only the thymine (T) and uracil (U) anions have strong support for long-lived existence in the gas phase. ("Covalent" electron affinities are to be distinguished from so-called "dipole-bound"

* To whom correspondence should be addressed. E-mail: hfsiii@uga.edu.
‡ Yale University.
† University of Georgia.

(1) Steenken, S.; Telo, J. P.; Novais, H. M.; Candeias, L. P. *J. Am. Chem. Soc.* **1992**, *114*, 4701.
(2) Colson, A. O.; Sevilla, M. D. *Int. J. Radiat. Biol.* **1995**, *67*, 627.
(3) Desfrançois, C.; Abdoul-Carime, H.; Schermann, J. P. *J. Chem. Phys.* **1996**, *104*, 7792.
(4) Huels, M. A.; Hahndorf, I.; Illenberger, E.; Sanche, L. *J. Chem. Phys.* **1998**, *108*, 1309.
(5) Cai, Z.; Sevilla, M. D. *J. Phys. Chem. B* **2000**, *104*, 6942.
(6) Messer, A.; Carpenter, K.; Forzley, K.; Buchanan, J.; Yang, S.; Razskazovskii, Y.; Cai, Z.; Sevilla, M. D. *J. Phys. Chem. B* **2000**, *104*, 1128.
(7) Cai, Z.; Gu, Z.; Sevilla, M. D. *J. Phys. Chem. B* **2000**, *104*, 10406.
(8) Giese, B.; Spichty, M. *ChemPhysChem* **2000**, *1*, 195.
(9) Giese, B.; Spichty, M.; Wessely, S. *Pure and Appl. Chem.* **2001**, *73*, 449.
(10) Berlin, Y. A.; Burin, A. L.; Ratner, M. A. *J. Am. Chem. Soc.* **2001**, *123*, 260.
(11) Bixon, M.; Jortner, J. *J. Phys. Chem. A.* **2001**, *105*, 10322.
(12) Chen E. C. M.; Chen, E. S. D.; Wentworth, W. E. *Biochem. Biophys. Res. Commun.* **1990**, *171*, 97.
(13) Wiley: J. R.; Robinson, J. M.; Ehdaie, S.; Chen, E. C.; Chen, E. S.; Wentworth, W. E. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 841.
(14) Hendricks, J. H.; Lyapustina, S. A.; de Clercq, H. L.; Bowen, K. H. *J. Chem. Phys.* **1996**, *104*, 7788.
(15) Desfrançois C.; Perquet, V.; Bouteiller, Y.; Schermann, J. P. *J. Phys. Chem. A* **1998**, *102*, 1274.
(16) Schiedt, J.; Weinkauf, R.; Neumark, D. M.; Schlag, E. W. *Chem. Phys.* **1998**, *239*, 511.
(17) Chen, E. C. M.; Chen, E. S. D. *J. Phys. Chem. B* **2000**, *104*, 7835.
(18) Oyler, N. A.; Adamowicz, L. *J. Phys. Chem.* **1993**, *97*, 11122.
(19) Sevilla, M. D.; Besler, B.; Colson, A. *J. Phys. Chem.* **1995**, *99*, 1060.
(20) Desfrançois C.; Abdoul-Carime, H.; Carles, S.; Periquet, V.; Schermann, J. P.; Smith, D. M. A.; Adamowicz, L. *J. Chem. Phys.* **1999**, *110*, 11876.
(21) Smith, D. M. A.; Jalbout, A. F.; Smets, J.; Adamowicz, L. *Chem. Phys.* **2000**, *260*, 45.
(22) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *Chem. Phys. Lett.* **2000**, *322*, 129.
(23) Wesolowski, S. S.; Leininger, M. L.; Pentchev, P. N.; Schaefer, H. F. *J. Am. Chem. Soc.* **2001**, *123*, 4023.
(24) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **2001**, *105*, 8782.

Scheme 1



electron affinities.[25] The dipole-bound electron affinity refers to the energy difference between a neutral molecule and an anion characterized by a very diffuse molecular orbital containing the additional electron which weakly binds via dipole interactions. The additional electron in conventional "covalent" anions fills the lowest unoccupied valence molecular orbital. The additional electron in the covalent anion causes significant structural changes, whereas the additional electron in the dipole-bound anion is too distant to influence the structure.) The best estimates for the (covalent) adiabatic electron affinities (AEAs) of T and U are 0.16 and 0.19 eV, respectively.[23] On the other hand, the AEAs of isolated guanine (G) and cytosine (C) are very close to zero ($<\pm 0.1$ eV) and may not be bound at all. In fact, Li, Cai, and Sevilla estimate an AEA of -0.75 eV for guanine.[26] Several theoretical studies[27−32] identify geometry changes that lead to accommodation of excess charge and may justify differences in stability. In general, the amino groups slightly pyramidalize, and the rings tend to pucker in the anions.

While the electronic properties and geometric changes of the isolated neutral and anionic nucleobases represent a fascinating challenge to both spectroscopists and theorists, the incorporation of additional components of the DNA strand clearly alters these properties. No experimental results exist for the electron affinity of the GC pair (Scheme 1). However, spectroscopic determination of electron affinities for microsolvated nucleobases shows that hydrated nucleobases are markedly different from their isolated counterparts. These studies provide an important set of experimental data for systems which may mimic some of the properties of GC base pairs. The groups of Bowen,[33] Schlag,[16,34] and Desfrancois[3,20,35] have obtained photoelectron spectra for incrementally solvated uracil, thymine, and cytosine. Their pioneering experiments clearly demonstrate that although naked nucleobase covalent anions are not directly observed, microsolvation with even a single water molecule provides sufficient stabilization to facilitate electron binding. For example, the cytosine−water complex has an electron affinity of ca. 0.3 eV.[16] Successive microsolvation further increases the stabilization of these anion complexes. The largest complex investigated, $C^{\bullet}(H_2O)_5$, has an electron affinity of more than 1.0 eV. In this experiment, the authors note that upon microsolvation the cytosine anion has spectra characteristic of a covalent rather than a dipole-bound anion.

Given the positive electron affinities for solvated cytosine complexes, the electron affinity of GC might be expected to be positive as well. However, previous theoretical efforts have failed to reach a consensus concerning the correct sign of the electron affinity of the GC pair.[27,29,31,36,37] The seminal contribution by Colson, Besler, and Sevilla using HF/3-21+G(d) suggested a negative AEA of -0.75 eV.[27] Inclusion of electron correlation via MP2 by Adamowicz and co-workers yields a value of -0.056 eV, while their prediction using the B3LYP density functional method is $+0.391$ eV. However, these authors repose little confidence in the latter result: "...the parametrization of the DFT/B3LYP method was not developed to provide accurate EA values and thus [their DFT] result should be considered as an approximation with a rather uncertain trust bracket." Recently, Li, Cai, and Sevilla[37] also determined a positive AEA (0.49 eV) for GC using B3LYP with a 6-31+G-(d) basis.

While B3LYP is widely considered to be one of the most reliable density functionals, systematic trends of AEAs predicted with a series of density functionals have been noted for over 100 chemical systems.[38] These trends have suggested a bracketing technique which provides a practical means for assessing electron affinities of molecules for which no experimental data exist. In this study we predict the adiabatic electron affinity of the GC pair using five different density functional combinations in an effort to bracket the true AEA and establish the correct sign. In addition, natural population atomic (NPA) charges for the GC pair are computed to monitor the location of excess charge. This systematic bracketing procedure also provides a unique platform for comparing structures, vibrational frequencies, and pairing (dissociation) energies (obtained with a variety of density functionals) to the robust body of theoretical and experimental work addressing these properties.

II. Theoretical Methods

Absolute energies, optimized geometries, and natural charges were determined for the hydrogen-bonded Watson−Crick base pair guanine-cytosine (GC). Five generalized gradient approximation (GGA) exchange-correlation density functionals were used. These are denoted B3LYP, B3P86, BHLYP, BLYP, and BP86 and are combinations of one of Becke's exchange functionals: the three-parameter HF/DFT hybrid exchange functional (B3),[39] a modified half-and-half HF/DFT hybrid method (BH)[40] as implemented in GAUSSIAN 94, or the pure DFT exchange functional (B)[41] with the dynamical correlation functional of Lee, Yang, and Parr (LYP)[42] or that of Perdew (P86).[43,44] The GAUSSIAN 94 system of DFT programs[45] was used for all results.

(25) Simons, J.; Jordan, K. D. *Chem. Rev.* **1987**, *87*, 535.
(26) Li, X.; Cai, Z.; Sevilla, M. D. *J. Phys. Chem. A* **2002**, *106*, 1596.
(27) Çolson, A. O.; Besler, B.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 9787.
(28) Šponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161.
(29) Al-Jihad, I.; Smets, J.; Adamowicz, L. *J. Phys. Chem. A* **2000**, *104*, 2994.
(30) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117.
(31) Smets, J.; Jalbout, A. F.; Adamowicz, L. *Chem. Phys. Lett.* **2001**, *342*, 342.
(32) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Mol. Struct.* **2001**, *573*, 43.
(33) Hendricks, J. H.; Lyapustina, S. A.; de Clercq, H. L.; Bowen, K. H. *J. Chem. Phys.* **1998**, *108*, 8.
(34) Smets, J.; Smith, D. M. A.; Elkadi, Y.; Adamowicz, L. *J. Phys. Chem. A* **1997**, *101*, 9152.
(35) Periquet, V.; Moreau, A.; Carles, S.; Schermann, J. P.; Desfrançois, C. *J. Electron Spectrosc. Relat. Phenom.* **2000**, *106*, 141.
(36) Saettel, N. J.; Wiest, O. *J. Am. Chem. Soc.* **2001**, *123*, 2693.
(37) Li, X.; Cai, Z.; Sevilla, M. D. *J. Phys. Chem. B* **2001**, *105*, 10115.
(38) Rienstra-Kiracofe, J. C.; Tschumper, G. S.; Schaefer, H. F.; Nandi, S.; Ellison, G. B. *Chem. Rev.* **2002**, *102*, 231.
(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(40) Becke, A. D. *J. Chem. Phys.* **1992**, *98*, 1372.
(41) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(42) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785.
(43) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(44) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.

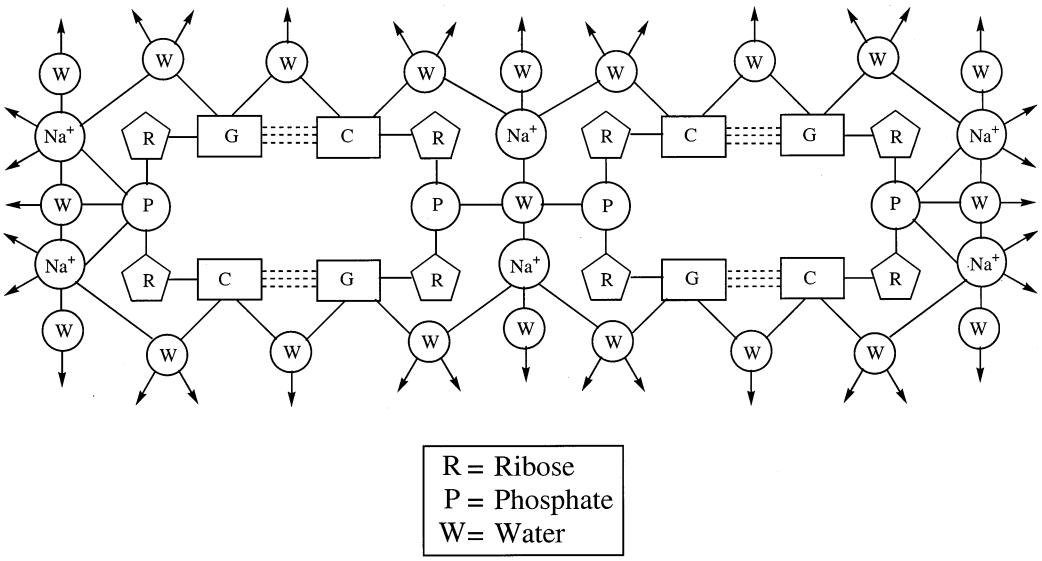


Figure 1. Schematic representation of the crystal structure of guanylyl-3′,5′-cytidine nonahydrate (GpC). Connectivity among some water molecules is omitted for clarity.

Double-ζ quality basis sets with polarization and diffuse functions (denoted DZP++) were used throughout for optimizations and vibrational frequency analyses. The DZP++ basis sets were constructed by augmenting the Huzinaga−Dunning[46,47] set of contracted double-ζ Gaussian functions with one set of p-type polarization functions for each H atom and one set of five d-type polarization functions for each C, N, and O atom ($\alpha_p(H) = 0.75$, $\alpha_d(C) = 0.75$, $\alpha_d(N) = 0.80$, $\alpha_d(O) = 0.85$). To complete the DZP++ basis, one even-tempered s diffuse function was added to each H atom while sets of even-tempered s and p diffuse functions were centered on each heavy atom. The even-tempered orbital exponents were determined according to the prescription of Lee and Schaefer:[48]

$$\alpha_{\text{diffuse}} = \frac{1}{2}\left(\frac{\alpha_1}{\alpha_2} + \frac{\alpha_2}{\alpha_3}\right)\alpha_1 \qquad (1)$$

where α_1, α_2, and α_3 are the three smallest Gaussian orbital exponents of the s- or p-type primitive functions for a given atom ($\alpha_1 < \alpha_2 < \alpha_3$). The final DZP++ set contains six functions per H atom (5s1p/3s1p) and 19 functions per C, N, or O atom (10s6p1d/5s3p1d), yielding a total of 421 contracted functions for the GC pair. This basis has the tactical advantage that it has previously been used in successful studies[38] for a wide range of electron affinities.

In addition, single-point energies at the DZP++ optimized geometries were computed using a triple-ζ quality basis set (TZ2P++). This basis was formed from the Huzinaga−Dunning[46,47] sp sets augmented with two sets of polarization functions (two sets of five d-type functions on C, N, and O, and two sets of p functions on H). The exponents for the polarization functions are $\alpha_p(H) = 1.50, 0.375$, $\alpha_d(C) = 1.50, 0.375$, $\alpha_d(N) = 1.60, 0.40$, and $\alpha_d(O) = 1.70, 0.425$. Even-tempered diffuse s- and p-type functions were added in a fashion analogous to that for the DZP++ set. The final TZ2P++ set contains 10 functions per H atom (6s2p/4s2p) and 28 functions per C, N, or O atom (11s7p2d/6s4p2d), yielding a total of 632 contracted functions for the GC pair.

Both the neutral and anion stationary points were optimized via analytic gradients until the residual RMS gradient was less than 10^{-4} hartree/bohr. Numerical integration was performed using the GAUSSIAN 94[45] default grid consisting of 75 radial shells with 302 angular points per shell.

Valence adiabatic electron affinities were computed as the difference between the absolute energies of the neutral and anion species at their respective optimized geometries.

$$\text{AEA} = E_{\text{neut}} - E_{\text{anion}} \qquad (2)$$

All molecular orbital plots were constructed with the TZ2P++ basis using the MOLDEN software package[49] and utilized the appropriate B3LYP/DZP++ optimized structures.

Natural population atomic (NPA) charges were determined at the B3LYP/DZP++ level using the natural bond order (NBO) analysis of Reed and Weinhold.[50−53]

III. Results

A. Geometry. While experimental gas-phase data are available for individual NABs, analogous experiments for the Watson−Crick pairs present major challenges. Gas-phase GC has been difficult to observe,[54] though an IR spectrum was obtained with resonance enhanced multiphoton ionization (REMPI).[55] In contrast, numerous solid-phase crystal structures of larger DNA fragments are well-known,[56] as well as sodium guanylyl-3′,5′-cytidine nonahydrate (GpC)[57] and sodium adenylyl-3′,5′-uridine hexahydrate (ApU).[58] These latter two structures represent the smallest DNA or RNA fragments containing GC and AU pairs, respectively. Despite their biochemical simplicity, these fragments have a relatively complex environment including deoxyribose and phosphate units as well as sodium ions and nine water molecules per unit as schematically represented in Figure 1. This work has served as the primary

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision c.3; Gaussian, Inc.: Pittsburgh, PA, 1995.
(46) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.
(47) Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823.
(48) Lee, T. J.; Schaefer, H. F. *J. Chem. Phys.* **1985**, *83*, 1784.
(49) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Design* **2000**, *14*, 123.
(50) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
(51) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 1736.
(52) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.
(53) Reed, A. E.; Schleyer, P. R. *J. Am. Chem. Soc.* **1990**, *112*, 1434.
(54) Desfrançois, C.; Abdoul-Carime, H.; Schulz, C. P.; Schermann, J. P. *Science* **1995**, *269*, 1707.
(55) Nir, E.; Kleinermanns, K.; de Vries, M. S. *Nature* **2000**, *408*, 949.

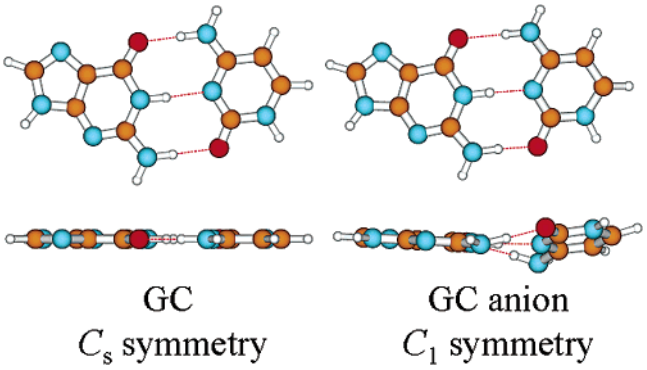


Figure 2. Side views of GC. The neutral strutures are of C_s symmetry; however, the anion deviates visibly from planarity. The energy difference between the C_s and C_1 structures for the anion is 1.5 kcal/mol.

comparison for theoretical studies of bases and base pairs.[30,37,59−62] The challenge is that the environment of the crystal and that of the isolated gas phase molecules modeled with theory are significantly different. Baerends and co-workers concluded that these differences were due largely to charge effects.[30] They were able to reproduce the geometric parameters by including a selection of ions and water molecules (but not corresponding to the locations in the GpC crystal.)

Although base pairs also present a greater challenge to theory than do the isolated NABs, much success has been realized in the studies of geometries (especially hydrogen-bond lengths),[30,59,60,63−70] pairing energetics,[32,63,70−75] and vibrational analyses.[76,77] Careful analysis of these quantities for both the neutral and the anion provides some rationale for conclusions about the AEA and the location of the charge. The effect of added charge in the Watson−Crick GC base pair is most evident in the geometry changes. Figure 2 shows the remarkable differences between the neutral GC and the anionic form. Specific changes in relevant dihedral angles around the amino groups and the pyramidalization of the nitrogens are detailed in Table 1. (The neutral GC has dihedral angles of 0° and planar nitrogens.) The optimized geometry of neutral GC is planar (C_s symmetry.) This large structural perturbation indicates that the anion is covalent and not dipole-bound.

Table 1. Dihedral and Pyramidalization Angles (in degrees) of Selected Nitrogens in G, G$^-$, C, C$^-$, GC, and GC$^-$ Using the DZP++[a] Basis Set

	N3C2N2H2′			N1C2N2H2			N2 pyramidalization		
	G	G$^-$	GC$^-$	G	G$^-$	GC$^-$	G	G$^-$	GC$^-$
BHLYP	12.5	16.2	8.1	−32.3	−19.8	−15.7	345.5	351.1	355.9
BLYP	13.7	19.4	11.5	−36.0	−30.6	−18.5	342.6	340.6	353.8
B3LYP	13.0	18.7	8.5	−34.2	−27.1	−17.4	344.2	344.6	355.5
BP86	13.6	19.7	10.8	−36.6	−31.0	−19.7	342.1	340.1	353.4
B3P86	12.9	19.2	9.0	−34.3	−27.3	−17.7	344.2	343.9	354.9
MP2/ 6-311G(2df,p)	13.3			−39.6			339.6		

	C5C6N4H4′			N3C4N4H4			N4 pyramidalization		
	C	C$^-$	GC$^-$	C	C$^-$	GC$^-$	C	C$^-$	GC$^-$
BHLYP	−14.1	19.4	28.1	8.9	−38.4	−25.8	356.4	331.0	339.7
BLYP	−19.4	23.5	26.6	11.9	−29.6	−25.7	353.3	335.0	341.4
B3LYP	−16.4	19.5	28.2	10.1	−39.2	−26.3	355.1	329.3	338.6
BP86	−19.6	24.3	29.4	12.1	−28.9	−26.0	353.1	334.3	339.3
B3P86	−15.9	18.9	29.2	9.9	−40.6	−25.2	355.4	328.8	339.7
MP2/ 6-311G(2df,p)[b]	−21.4			12.6			351.9		

[a] Isolated cytosine more closely resembles the structure of cytosine within the base pair anion, GC$^-$, while isolated guanine more closely resembles the structure of guanine within the neutral pair. Note that in the neutral GC all the dihedral angles are 0° and the sum of the angles around both amino nitrogens is 360°. [b] Reference 67.

Although the C_s symmetry structure is higher in energy (<0.01) than the C_1 structure, the difference appears to be due to numerical errors in the DFT procedures, not an actual difference in structure. The C_s harmonic vibrational frequencies are all real and essentially identical to those for C_1 symmetry. Also, neutral GC shows negligible deviation from planarity in the nitrogens of the amino groups.

In contrast, the optimized GC anion has a considerably puckered pyrimidine ring. In accord with this visible geometric difference, the C_s structure of the GC anion lies 1.5 kcal/mol higher in energy than the optimized C_1 structure. The dihedral angle N3C2N2H2′ of GC$^-$ more closely resembles N3C2N2H2 of neutral G than it resembles that of G$^-$; in fact N3C2N2H2′ is flatter in GC$^-$ than in either G or G$^-$ alone. The same is true for the pyramidalization of N2: less pyramidalization occurs in GC$^-$. For the dihedral angle N1C2N2H2 the GC$^-$ value lies slightly lower than the value for either G or G$^-$. Pairing flattens out both the anion and neutral G. For C the effects are quite different. The dihedral angles and extent of pyramidalization of N4 in GC$^-$ resemble C$^-$ to a much greater extent than C. The sign of the dihedral angles reverses from the neutral to the anion. In GC$^-$ both C5C6N4H4′ and N3C4N4H4 are of the same sign as C$^-$. The C5C6N4H4 dihedral angle becomes even larger in GC$^-$, while N3C4N4H4 is slightly less negative than in C$^-$. The pyramidalization of N4 is less than in C$^-$ but greater than in C. The changes appear to be similar for all the functionals, with BHLYP giving a slightly smaller N1C2N2H2 dihedral angle for G$^-$. These angular changes suggest that the excess charge is largely located on C in GC$^-$.

In addition to the angular changes, the variations in bond length show a similar, although more subtle, pattern. Figure 3 shows the bond lengths of the individual neutral and anionic G and C. Previous MP2/6-31G* work[28,78] for the neutral bases compares favorably to our neutral geometries. Figure 4 shows

(56) Kennard, O.; Hunter, W. N. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 1254.
(57) Rosenberg, J. M.; Seeman, N. C.; Day, R. O.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 145.
(58) Seeman, N. C.; Rosenberg, J. M.; Suddath, F. L.; Kim, J.; Rich, A. *J. Mol. Biol.* **1976**, *104*, 109.
(59) Brameld, K.; Dasgupta, S.; Goddard, W. A. *J. Phys. Chem. B* **1997**, *101*, 4851.
(60) Guerra, C. F.; Bickelhaupt, F. M. *Angew. Chem., Int. Ed.* **1999**, *38*, 2942.
(61) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247.
(62) Desfrançois, C.; Carles, S.; Schermann, J. P. *Chem. Rev.* **2000**, *100*, 3943.
(63) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493.
(64) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965.
(65) Hutter, M.; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*, 7574.
(66) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *Chem. Eur. J.* **1999**, *5*, 3581.
(67) Šponer, J.; Florian, H.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **1996**, *13*, 827.
(68) Smith, D. M. A.; Smets, J.; Adamowicz, L. *J. Phys. Chem. A* **1999**, *103*, 5784.
(69) Bertran, J.; Oliva, A.; Rodriguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159.
(70) Raimondi, M.; Famulari, A.; Gianietti, E. *Int. J. Quantum Chem.* **1999**, *74*, 259.
(71) Dey, M.; Moritz, F.; Grotemeyer, J.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 9211.
(72) Šponer, J.; Hobza, P. *Chem. Phys. Lett.* **1996**, *261*, 379.
(73) Kawahara, S.; Uchimaru, T. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2869.
(74) Kabelac, M.; Hobza, P. *J. Phys. Chem. B* **2001**, *105*, 5804.
(75) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
(76) Špirko, V.; Šponer, J.; Hobza, P. *J. Chem. Phys.* **1997**, *106*, 1472.
(77) Santamaria, R.; Charro, E.; Zacarias, A.; Castro, M. *J. Comput. Chem.* **1999**, *20*, 511.

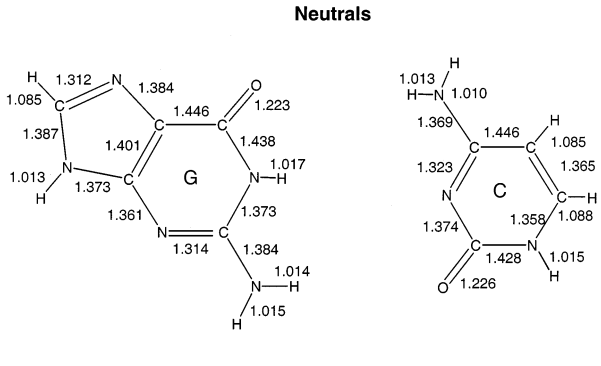


Figure 3. B3LYP/DZP++ bond lengths for the isolated anion and neutral G and C. Distances are reported in angstroms.

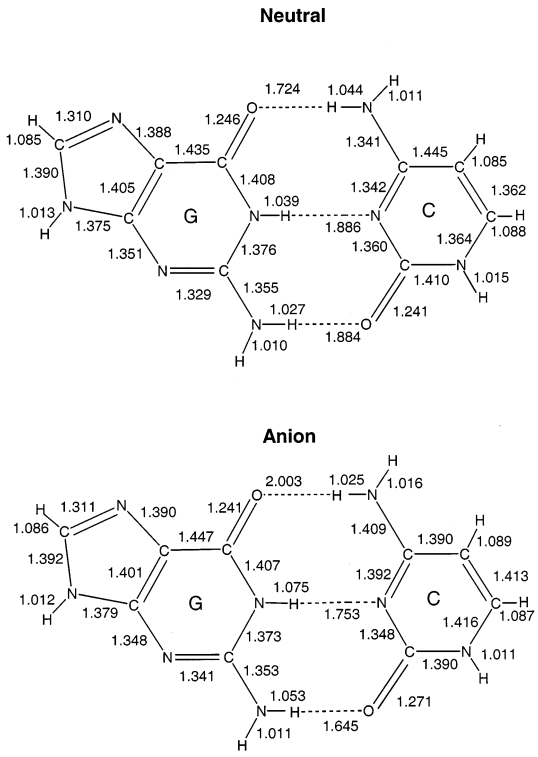


Figure 4. B3LYP/DZP++ bond lengths for anion and neutral GC pairs. Distances are reported in angstroms.

all the bond lengths of the neutral and anionic GC pair. Finally, Figure 5 illustrates the significant (>0.010 Å) bond length changes that occur between the neutral and anionic GC pair. Few changes occur for G either by itself or in the pair. In contrast, C shows increases in most of the bond lengths, indicating addition of electron density to antibonding orbitals.

(78) Bludsky, O.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Chem. Phys.* **1996**, *105*, 24.

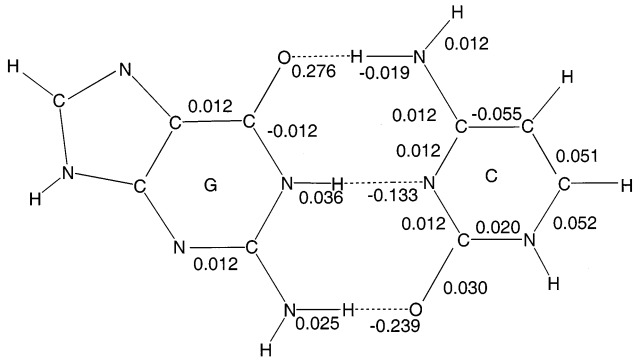


Figure 5. Significant (>0.010 Å) changes in GC base pair bond lengths (anion − neutral). Differences are reported in angstroms.

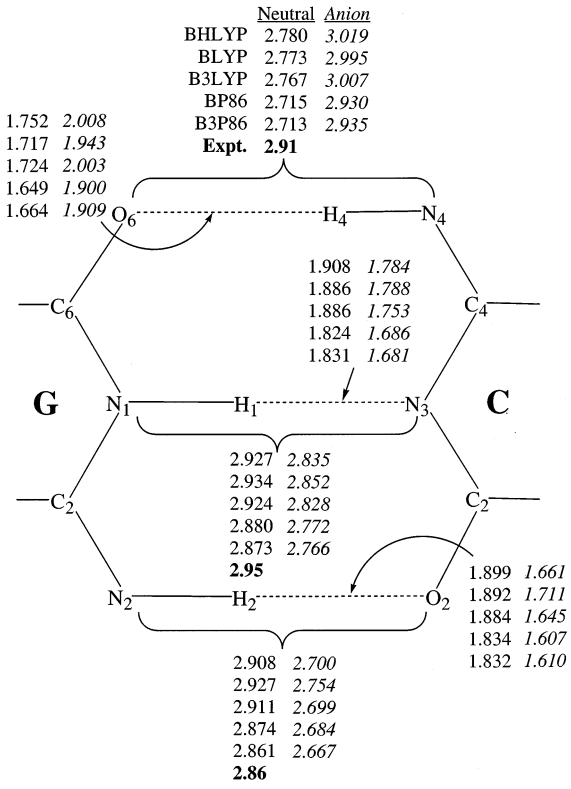


Figure 6. Covalent heavy atom−hydrogen, hydrogen bond heavy atom−hydrogen, and heavy atom bond lengths of the guanine−cytosine anion base pair with five DFT functionals using the DZP++ basis set. (Comparison is made to the experimental distances[57] determined by crystallography and shown in Figure 1.)

Only the C4−C5 bond decreases slightly, as there is a slight increase in double bond character. A similar shortening also occurs for N3−C2 in isolated C; however, this bond lengthens in the pair due to N3 participation in a hydrogen bond with atom H1 of G. The bond changes for C in GC which differ from those in C alone are also related to hydrogen-bonding interactions.

The changes in bond lengths for the atoms involved in hydrogen bonds between G and C also give an indication of the location of excess charge. As shown in Figure 6, the N4−H4 bond distance on C decreases, while the N1−H1 and N2−H2 bonds on G both increase. This correspondence is in accord with the increase in pyramidalization of N4 on C and the decrease in pyramidalization of N2 on G. Traditional hybridization arguments can be made: when a nitrogen gains charge, it becomes more nearly sp^3 hybridized and more pyramidal; when

Table 2. Harmonic Vibrational Frequencies (in cm^{-1}) Corresponding to the Six "New" Vibrational Modes Introduced upon GC and GC^- Pairing[a]

	ω_1	ω_2	ω_3	ω_4	ω_6	ω_7
BHLYP/DZP++	27, 29	38, 37	78, 70	91, 97	130, 112	141, 150
BLYP/DZP++	24, 27	40, 34	78, 73	94, 95	130, 112	133, 155
B3LYP/DZP++	27, 28	40, 35	80, 97	93, 97	129, 112	135, 140
BP86/DZP++	21, 27	38, 36	76, 72	98, 101	124, 115	140, 136
B3P86/DZP++	26, 28	38, 37	79, 71	97, 101	133, 117	143, 144
BP86/6-311G[b] (GC)	29	44	85	105	145	150
B3PW(91)/6-311G[c] (GC)	30	45	87	103	144	148
expt (GC)[c]	−	−	65.7	81.8	114.4	119.5

[a] Entries given as pairs correspond to the neutral and anion GC values, respectively. A thorough description of these vibrational modes is provided in reference 77. [b] Reference 77. [c] Reference 55.

Table 3. Pairing or Dissociation Energies (in kcal/mol) for the Neutral and Anionic G, C, GC[a]

	$GC \rightarrow G + C$	$GC^- \rightarrow G + C^-$	$GC^- \rightarrow G + e^- + C$
BHLYP/DZP++	−28.3 (−26.4)	−40.5 (−39.4)	−44.4 (−44.5)
BLYP/DZP++	−25.6 (−23.9)	−36.8 (−36.1)	−33.8 (−35.8)
B3LYP/DZP++	−27.2 (−25.4)	−39.4 (−38.4)	−40.6 (−40.8)
BP86/DZP++	−27.5 (−26.1)	−39.9 (−39.4)	−39.5 (−39.8)
B3P86/DZP++	−29.1 (−27.5)	−42.2 (−41.5)	−45.8 (−45.8)
TZ2P++/B3LYP	−24.8 (−23.9)	−36.6 (−35.6)	−33.3 (−33.4)
HF/6-31G(d)[b]	−23.02	−34.99	
MP2/6-31G* [c]	−23.8		
expt[d]	−21.0		

[a] Zero-point corrected pairing energies are given in parentheses. [b] Reference 27. [c] Reference 64. [d] From field-ionization mass spectroscopy, reference 79.

nitrogen loses charge, it becomes more sp^2 hybridized and more planar. These changes in turn explain why the corresponding hydrogen-bond length changes have an opposite pattern: hydrogen-bond distances increase if they are adjacent to decreasing N−H bonds. The hydrogen bond between O6 of G and H4 of C is weakened and lengthened by the removal of some electron density, while the hydrogen bonds between atom H1 of G and atom N3 of C and between atom H2 of G and atom O2 of C are strengthened and shortened by the increase of electron density.

A final geometrical relationship to consider is the pattern among the heavy atom only distances. Figure 6 also shows the heavy atom distances for all five functionals and the experimental values[57] for GpC. The distance from O6 of G to N4 of C increases overall, while the N1 of G to N3 of C and the N2 of G to O2 of C distances decrease. Interestingly, the O6 of G to N3 of C distance for the anionic GC more closely agrees with the experimental distance of 2.91 Å determined by Rich and co-workers.[57] While GC is formally neutral in the crystal structure, the phosphates being close to the GC pair and the sodium counterions being further distant may support the idea that in this crystal the O6−N4 interaction is partly governed by some negative charge distribution on GC. Indeed, Baerends et al. showed that the hydrogen-bond distances of GC found in the GpC crystal may be reproduced by appending positive sodium ions and water molecules in certain locations around GC (but not corresponding to locations in the GpC crystal).[30] The N1−N3 distance and the N2−O2 distance are successively less affected by what may be the partial negative charge distributed to GC. This is likely due to their closer proximity to atoms N3 of G and N1 of C, which are connected to ribose which in turn connects to the phosphate.

B. Vibrational Frequencies and Pairing Energies. The changes in geometry and, as shown in Table 2, smaller variations in vibrational frequencies occur as GC accepts a negative charge. However, since the changes in vibrational frequencies are relatively small for the modes related to GC interaction, the overall stability of the molecule is not expected to decrease. Vibrational analysis shows a decrease of about 15 cm^{-1} in the harmonic vibrational frequency of the anion mode associated primarily with the antisymmetric stretching vibration between G and C. The asymmetric out-of-plane bending mode also decreases slightly. These two modes are the only ones to be significantly affected by the addition of an electron to neutral G−C. The inclusion of the diffuse functions in the DZP++ basis set reduces the magnitude of the vibrational frequencies compared to those determined with the 6-311G basis set.[77] The recent experimental work using REMPI has also determined frequencies for these low-energy vibrational modes.[55] Our theoretical harmonic vibrational frequencies still exceed those determined by experiment by about 20 cm^{-1} since the low-energy modes are likely to be very anharmonic. The inclusion of anharmonicity in the determination of the intermolecular vibrational modes at the Hartree−Fock level by Spirko, Sponer, and Hobza provides closer agreement with the experimental fundamentals.[76]

In addition to frequency changes which occur upon pairing, the energetics provide a quantitative measure of stability. Yanson, Teplitsky, and Sukhodub provide experimental estimates for the pairing energy for a number of base pairs[79] including GC, estimated at 21.0 kcal/mol. The early HF/6-31G(d) results of Colson, Besler, and Sevilla were intriguingly close to this value through a cancellation of errors.[27] Table 3 provides the pairing energies of GC and GC^- for all five functionals. While a pairing energy for $GC^- \rightarrow G + C^-$ is the most favorable, we also report $GC^- \rightarrow G + e^- + C$, where the energy for the righthand side is that of G + C neutral. We have not considered BSSE corrections in this work, but they are generally less for DFT than MP2.[61] Also, our use of a larger basis set with diffuse functions on both the heavy atoms and the hydrogens reduces the need for BSSE corrections. This thermodynamic analysis indicates that B3LYP/TZ2P++//B3LYP/DZP++ gives a stability of 23.9 kcal/mol for neutral GC and 32.0 kcal/mol for anionic GC. The neutral compares favorably to the MP2/6-31G* results of Šponer, Leszczynski, and Hobza who found the neutral GC to have a 23.8 kcal/mol pairing energy,[64] and their most recent estimate based on higher level studies of model systems, which adds 2.5 to 3.0 kcal/mol to their earlier estimate.[80] The early work[27] on the anion was again fortuitously close to our DFT result. The most recent theoretical studies[80] repose little confidence in the 1979 experiments.[79]

C. Bracketed Electron Affinity. Table 4 shows the electron affinities of G, C, and GC for all five DFT functionals. Even the nonzero-point corrected BHLYP functional, which is well-known to underestimate electron affinities, indicates that an electron binds to GC. These five functionals have been shown to provide an accurate bracket for electron affinities in many molecules with BHLYP underestimating (providing a lower bound) and B3P86 overestimating (providing an upper bound).[38]

(79) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.
(80) Šponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2002**. In press.

Table 4. Zero-Point Vibrationally Corrected Adiabatic Electron Affinities (in eV) for Isolated G, C, and the GC Base Pair Using the DZP++ Basis Set[a]

functional	G[a]	C	GC
BHLYP	−0.36 (−0.42)	−0.14 (−0.25)	0.42 (0.28)
BLYP	−0.01 (−0.10)	−0.01 (−0.13)	0.51 (0.35)
B3LYP	−0.07 (−0.14)	0.30 (−0.90)	0.60 (0.44)
BP86	0.11 (0.02)	0.13 (0.01)	0.71 (0.54)
B3P86	0.36 (0.27)	0.54 (0.42)	1.15 (0.99)
B3LYP/TZ2P++	0.07 (<0.01)	−0.02 (−0.02)	0.48 (0.37)

[a] Uncorrected electron affinities are provided in parentheses. B3LYP/TZ2P++ single-point electron affinities are provided for comparison and employ the corresponding DZP++ zero-point corrections. [b] These values reflect the AEA for the dipole-bound anion.

These values suggest that the electron affinity of GC is at least 0.28 eV and at most 1.15 eV. The three most reliable functionals (BLYP, B3LYP, and BP86) suggest a narrowing of this range from 0.50 to 0.71 eV. In an attempt to estimate the AEA even more closely the zero-point corrected single-point TZ2P++/B3LYP//B3LYP/DZP++ value of 0.48 eV was determined. This also agrees well with that of Sevilla et al. who determined the B3LYP/6-31+G(d) AEA to be 0.49 eV,[37] showing that, while diffuse functions are important, basis-set size itself does not change these AEAs. Li, Cai, and Sevilla did determine that the effects of additional diffuse functions could be used to determine the dipole-bound anion as well; however, the dipole-bound anion is closer to the energy of the neutral and thus produces a smaller AEA.[26]

Thus. the B3LYP/6-31++G** electron affinity of 0.39 eV reported by Smets, Jalbout, and Adamowicz[31] is much more reasonable than their MP2/6-31++G**//HF/6-31++G**(6d) value of −0.06 eV. One explanation for the poorer performance of the MP2 wave function points to the effects of spin polarization,[22,69] resulting in spin contamination which may cause an overestimation of the total energy for the anion. This leads to values of the electron affinity that are too small.

D. Further Rationalization of GC's Positive AEA: A Comparison to Experimental Solvation Results and Theoretical NPA Charge Analysis. Since significant structural changes are evident upon the addition of an electron, the excess electron must occupy a valence orbital close to the anion, and not a diffuse orbital of GC. In Figure 7 the B3LYP/TZ2P++//B3LYP/DZP++ orbital plots for the SOMO of the G and C anions are shown. Also shown are the GC HOMO and the GC anion SOMO. Consistent with the results of Sevilla et al., G^- is dipole-bound while C^- is valence-bound.[26] From the SOMO it is also evident that in GC^- the electron density from the added electron is located primarily on the cytosine, and the GC anion is thus a valence-bound anion.

The sizable AEA of GC is somewhat surprising in light of the small electron affinities of the individual bases G and C. These EAs were shown to oscillate between small positive and somewhat negative values in previous work.[22,23,26] The increased anionic stability for the combination of the two bases shows that the interactions of the molecule must stabilize the excess charge. However, the experimental results of Schiedt and co-workers[16] show that when cytosine is microsolvated by just one water molecule, the photoelectron spectrum indicates a conventional anion, with an electron affinity near 0.30 eV. A comparison of solvated electron affinities of $C\cdot(H_2O)_n$ with selected AEA values for GC is shown in Figure 8. From our

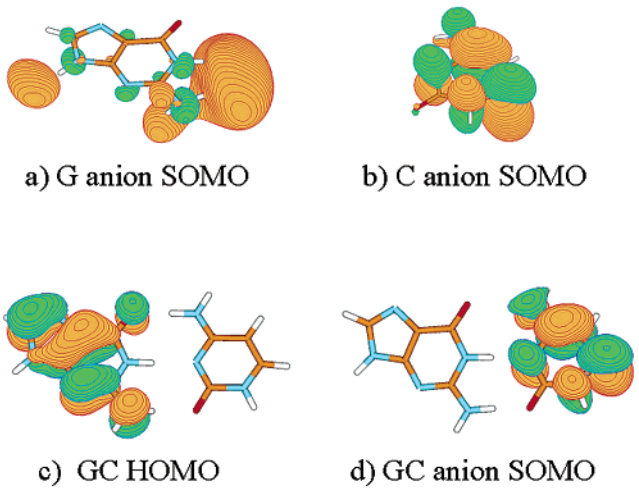


Figure 7. Qualitative comparison of several relevant TZ2P++ molecular orbitals involved in GC anion formation. The SOMO of the G anion clearly reveals dipole-bound character, while the C anion SOMO displays valence character. The HOMO of neutral GC supports electron density on the guanine. However, the GC anion SOMO is strikingly similar to the free C anion SOMO and suggests that the unpaired electron in the GC anion lies predominantly on the cytosine.

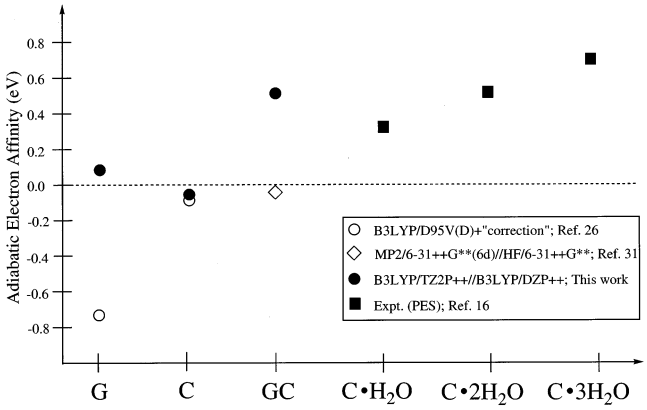


Figure 8. Comparison of the adiabatic electron affinity (AEA) of microsolvated C and the AEAs for G, C, and GC using MP2 and B3LYP; the positive AEAs of G reflect those of the dipole-bound state.

geometrical analysis (above) we show that in the anionic GC dimer C resembles C^-, while G resembles neutral G. By comparison to the experimental microsolvation data, we find that G solvates C by about as much as two water molecules. This effect supports the increase in electron affinity with respect to the isolated base.

Consideration of the location of the excess charge in GC^- shows that the flow is toward cytosine. Figure 9 shows the NPA charges for individual neutral and anionic G and C, while Figure 10 shows the charges for the neutral and the anionic GC pair. The largest changes not related to simply shifting charge within the same ring occur for C5 and C6 on cytosine. Also, the sum of all the charge on G and all the charge on C shows a significant shift of electron density toward C in the anion. The neutral G takes a net −0.36 charge, while C has a +0.36 charge, but when the pair has an overall −1 charge, C accounts for −0.92 e^-, and G carries −0.08 e^-. The change in hydrogen-bond lengths from the neutral to the anion explains why G bears negative charge in the neutral, but not in the anion: in the neutral the O6−(H4)N4 distance is much shorter and G is accepting a greater amount of electron density, while the donor hydrogen

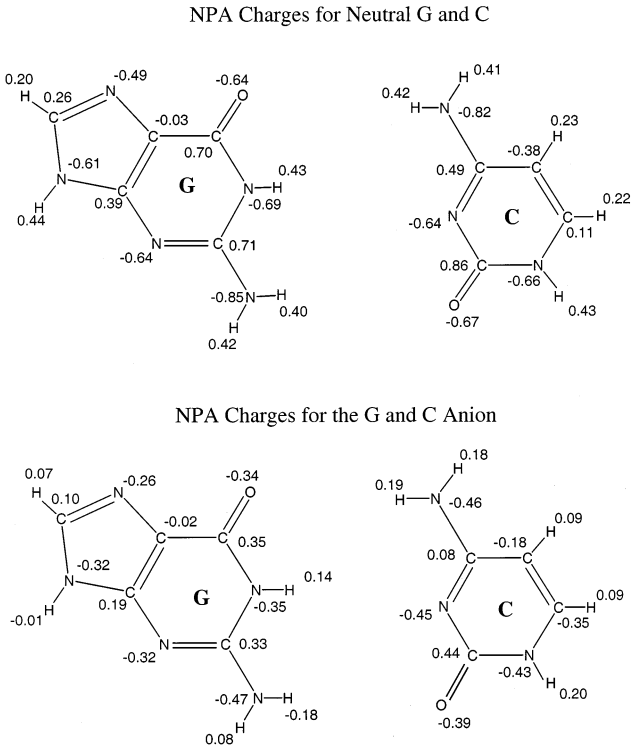


Figure 9. Neutral and anion natural population (NPA) charges on each atom in G and in C.

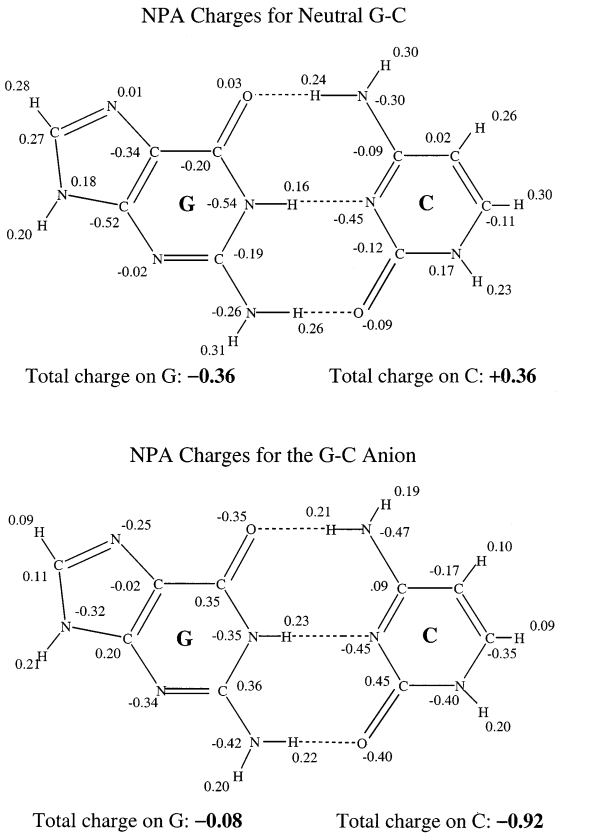


Figure 10. Neutral and anion natural population (NPA) charges on each atom in the GC base pair.

bonds of G are longer. In the anion the situation is reversed, with the G acceptor bond much longer and the G donor bonds shorter. The bearing of negative charge for G in the neutral results in greater distortions of G than G^- alone.

As shown in Figure 7, the electron density in the SOMO of the G anion is located outside of the ring structure. In the base pair G cannot accommodate the excess charge in a diffuse orbital as it can in the isolated NAB. Thus, electrons forced into close proximity with G and forming valence-bound anions are highly unstable. Within the anionic base pair, C is stablized by a substantial amount of negative charge. As shown in Figure 7 the charge is accommodated in an antibonding orbital of the C anion ring in a fashion similar to that observed in the anion of the pair. These charge distribution effects also serve to explain why the dihedral angles and pyramidalization in G and C change, as observed when the neutral pair becomes a negative ion. In the anionic pair G has flatter dihedrals and less pyramidalization than does the isolated G, while C has larger dihedrals and more pyramidalization than does anionic C. Taken together the structures, charges, and SOMOs of the isolated C anion and cytosine within the GC anion are strikingly similar and support the notion that the unpaired electron is largely on the cytosine in GC^-.

IV. Conclusions

The bracketing procedure with DFT functionals and analogy to microsolvation provides good evidence that the AEA of GC is positive, with the best value being around 0.48 eV. The geometric parameters of G within GC^- are most like those of neutral G rather than like anionic G (with angles flattening out more than for neutral G). The question of the true electron affinity of isolated G awaits a much higher-level theoretical approach which can account for the mixing of valence and dipole-bound anions. The geometric parameters of C are more like those of anionic C than neutral C (with the angles even more pyramidal than those in anionic C).

The geometry changes and charge distribution changes support the conclusion that excess charge resides on C in GC^-. The changes in hydrogen-bond lengths between G and C can be traced to the charge distributions within G and C. As the complex becomes more negative, the O6−N4 bond lengthens and becomes very similar to the length in the GpC crystal in which some negative charge is likely delocalized into the G−C fragment, lengthening the O6−N4 bond from what it would be in the isolated neutral G−C molecule. The formation of a stable covalent anion is largely due to the solvation effect of the G. Since the electron affinity of GC is determined to be positive, the observation of isolated GC^- should be possible.

Acknowledgment. We sincerely thank Drs. Jerzy Leszczynski, Jira Šponer, and Pavel Hobza for several insightful discussions. We also thank Dr. Matthew Leininger for seminal contributions to the individual nucleobase work as well as for a number of stimulating discussions. This work was supported in part by the U.S. National Science Foundation, Grant CHE-9815397.

JA020009W