J. Phys. Chem. A **2003,** *107,* 6441−6443

Cooperative Hydrogen-Bonding in Adenine−Thymine and Guanine−Cytosine Base Pairs. Density Functional Theory and Møller−Plesset Molecular Orbital Study

Amparo Asensio, Nadya Kobko, and J. J. Dannenberg*

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

Received: February 24, 2003; In Final Form: June 3, 2003

The cooperative contributions to the H-bonding interaction energies of the adenine−thymine and guanine−cytosine base pairs have been evaluated using molecular orbital theory. The energies of the individual bonds in each base pair were ascertained by using models structures that keep one H-bond at a time intact by rotating one base with respect to the other about the axis of each H-bond to form structures with the bases perpendicular to each other. The energies of the individual H-bonds calculated in this way are compared with those of the planar base pairs. Optimized geometries were obtained using ab initio molecular orbital theory with electron correlation (MP2/D95**) and density functional theory (B3LYP/D95**). The cooperative contributions are of similar magnitude for each base pair. However, since the A•T overall interaction is weaker, the cooperative interaction provides 31% of its stability versus only 12−16% for G•C. The relatively smaller cooperative contribution to G•C is due to the difficulty of forming three optimal H-bonds between two rigid molecules. Structural modifications that might strengthen one H-bond tend to weaken another. As a result the central H-bond of G•C appears to be compressed by the attractive interaction in the two outer H-bonds. To the extent that these observations can be generalized, they should be important to the design of materials that utilize H-bonding motifs for self-assembly.

Since the initial report by Clementi,[1] the hydrogen bonds in the purine/pyrimidine base pairs of DNA have been studied many times using molecular orbital methods.[2−11] The aggregate energies of the hydrogen bonds in these base pairs have been carefully evaluated and are well documented. The cooperativity in these interactions has been addressed.[7] However, the extent to which hydrogen bond cooperativity contributes to the stability of these base pairs has not previously been quantitatively investigated. In this paper, we evaluate the energies of the individual H bonds in the two base pairs of normal DNA. We compare these energies with the total H-bonding interaction of each base pair to evaluate the cooperative contributions to the overall stabilization of the pairs. While the interaction energies presented here have been calculated at a reasonably high level, this study aims to characterize the cooperative interactions in the base pairs rather than to perform a state of the art determination of the interaction energies, which would require larger basis sets and optimization on a potential energy surface corrected for BSSE. While such calculations are possible,[12] they would require an enormous amount of computer time for H-bonding systems such as G•C whose geometry converges extremely slowly.

Hydrogen-bond cooperativity has been extensively studied for H-bonds that form chains, such as those formed by the amidic functions in peptides and proteins,[13−21] and in molecular crystals, such as urea,[22−24] acetic acid,[25] and the enol of cyclohexane-1,3-dione.[26] These systems generally have only one H-bond between pairs of molecules. Thus, the cooperativity could be studied by determining the interaction energies as molecular aggregates increase in size, or by doing calculations on infinite periodic systems.[24,27]

The DNA base pairs contain multiple H-bonds between one pair of molecules.To evaluate the individual H-bonds in the DNA base pairs, one needs to be able to evaluate the energy of a particular hydrogen bond in the absence of the others. To do this, we have initially optimized the geometries of the normal base pairs to obtain the interaction energies of each pair. To obtain an estimate of the energies of the individual H-bonds in each pair, we rotated one of the bases with respect to the other about the axis of each hydrogen bond, in turn, so that the planes of the individual bases become perpendicular to each other. We, then, optimized the geometries of these structures with the constraints that (a) the planes of the bases remain perpendicular to each other and (b) that the angle of the hydrogen bond in which the hydrogen atom is central remains fixed. All other geometric parameters were allowed to vary freely. We take the difference between the sum of the individual H bond energies and the total interaction of the base pairs in their normal coplanar geometries as the cooperative part of the H-bonding interactions. Due to the constraints described above, the structures that contain the single H-bonds are not true minima on a potential energy surface (PES). Thus, the vibrational analyses that are necessary for the evaluation of enthalpies and entropies could not be properly performed. The cooperative contribution to the energies of interactions would be equivalent to the enthalpies if the vibrational energies of the planar and twisted base pairs be the same. Clearly, the model we use for evaluating the cooperativity does not directly address the entropic contribution.

Methods

Density functional theory (DFT) and second-order Møller−Plesset (MP2) ab initio molecular orbital (MO) calculations were

---

* Author to whom correspondence should be addressed. E-mail: jdannenberg@gc.cuny.edu.

10.1021/jp0344646 CCC: $25.00

TABLE 1: Energies for H-Bonds in Base Pairs[a]

	MP2/D95**		B3LYP/D95**	
	no correction	CP-corrected	no correction	CP-corrected
A·T				
interaction	−17.58	−11.69	−14.70	−12.56
NH···O (A)	−8.78	−4.68	−5.90	−4.59
N···H (B)	−5.98	−3.38	−4.79	−3.85
CH···O (C)	−2.29	−0.06	−1.01	−0.16
cooperativity	−0.53	−3.57	−3.00	−3.95
G·C				
interaction	−31.60	−24.50	−29.33	−26.65
O···H (A)	−6.56	−3.55	−5.39	−4.22
N···H (B)	−12.53	−7.92	−8.78	−7.24
NH···O (C)	−13.04	−9.98	−12.45	−10.87
cooperativity	0.53	−3.05	−2.70	−4.32

[a] The interaction energy (kcal/mol) followed by the energies of each H-bond individually and the cooperativity (total interaction less the sum of the individual bonds).

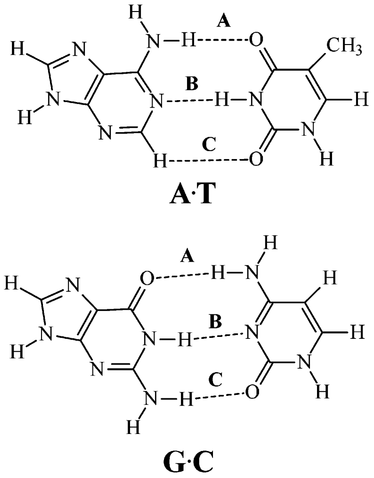

Figure 1. Base pairs with the H-bonds labeled.

performed using the GAUSSIAN 98 suite of computer programs[28] on our parallelized cluster Pentium 3, Pentium 4, and AMD Athlon computers using LINDA. All calculations used the D95(d,p) basis set. The hybrid DFT calculations used the B3LYP functional. This method combines Becke's 3-parameter functional,[29] with the nonlocal correlation provided by the correlation functional of Lee, Yang, and Parr.[30] The geometries of all species were completely optimized with the following constraints: (1) each of the bases was kept planar, (2) in the perpendicular species, each planar base was kept perpendicular to the other with the X−H···Y angle fixed at its value in the optimized planar base pair (X and Y refer to the heavy atoms in the H-bond). The number of compute nodes used for each calculation varied with the sizes of the systems studied. The vibrational frequencies were calculated for the planar structures, using the normal harmonic approximations employed in the GAUSSIAN 98 program, to verify the stationary points and to calculate the enthalpies of the various species. All frequencies were real except for some very low frequency imaginary vibrations that involved pyramidalization of the NH_2 groups of the individual bases. Vibrational analyses were not carried out on the perpendicular species as they are not true minima on the PES. The single-point *a posteriori* counterpoise corrections (CP) were calculated using the procedure incorporated in GAUSSIAN 98. Optimization on the CP-corrected PES values (CP−OPT)[12] were not completed due to the excessive CPU time required for the MP2 calculations.

Results and Discussion

The energetic results are collected in Table 1. The A·T and G·C base pairs with the individual H-bonds labeled to facilitate the following discussion are depicted in Figure 1.

Adenine·Thymine. The A·T base pair has two traditional H-bonds and one weak C−H···O interaction. After BSSE correction, the interaction energies are calculated to be −11.7 and −12.6 kcal/mol by MP2 and DFT, respectively. The MP2 calculations have a much greater CP correction than the DFT calculations. This implies that had the CP−OPT geometry been calculated, it should differ from the normal optimized geometry more for the MP2 than for the DFT optimized geometry. Thus, the MP2 geometry is farther from the minimum on the CP-corrected surface than that for DFT. This explains the weaker interaction calculated using MP2. The sums of the magnitudes of interactions for the individual H-bonds are 3.57 and 3.95 kcal/mol less than the interaction energy of the planar base pair for MP2 and DFT, respectively. Thus the cooperative contribution to the H-bonding interaction is 31% of the total interaction for each MO method. Closer examination of the individual H-bonds in the perpendicular base pairs shows that each of the traditional H-bonds is longer than in the planar structure. This result agrees with the observation that shorter H-bonds are generally stronger. On the other hand, the C−H···O interaction shortens considerably upon twisting the planar structure about the C−H···O axis. Thus, the C−H···O interaction must be weakened in the planar base pair. This interaction is sacrificed so that the two other stronger interactions can be optimized. The C−H···O interaction is negligible, even in the perpendicular structure. One can safely conclude this last interaction makes no significant contribution to the stability of the planar base pair, where the C−H···O interaction is much longer. Previous studies found no evidence for a C−H···O stabilizing interaction in A·T, despite finding such interactions stabilizing in uracil dimers.[7,11]

Guanine·Cytosine. The energetic analysis for the G·C base pair is more complex. Once again, the CP-corrected DFT interaction of −26.65 kcal/mol is slightly stronger than for MP2 (−24.50). The cooperative contributions of 3.05 and 4.32 kcal/mol for MP2 and DFT, respectively, are similar in magnitude to those calculated for A·T (see above). However, this corresponds to only 12−16% of the total interaction of the three H-bonds. Examination of the individual H-bonds in the perpendicular structures shows that only one of them, the C═O···H interaction (H-bond **A** in Figure 1), lengthens upon twisting about its H-bond axis for the MP2 calculations, while both this and the central NBH···N interaction (H-bond **B**) lengthen in the DFT optimization. For both calculational methods, the third H-bond shortens. The clear implication is that the molecular geometries do not allow all three H-bonds to simultaneously achieve their optimal interactions. A system with three H-bonds between rigid monomers cannot adjust its intermolecular geometry to simultaneously optimize all three interactions. Formation of an optimal H-bond **A**, for example, requires some geometrical compromise which may lead to less than optimal formation of the other two H-bonds, etc. For a system with only two H-bonds, small adjustments in the angle of one of the H-bonds can lead to a more stable H-bonding distance of the other without appreciably affecting the first H-bonding distance. However, with three H-bonds, such an adjustment of an angle might lead to a more stable interaction of one of the other two H-bonds, but not necessarily both. Thus, the total cooperativity for the three H-bond system (G·C) remains similar in magnitude (but lower in terms of the fraction of the overall interaction) to that for the two H-bond system (A·T).

TABLE 2: H-Bond Distances (Å) for the Planar and Twisted Base Pairs

	MP2/D95**			B3LYP/D95**		
	planar	twisted	difference	planar	twisted	difference
A•T						
NH···O (A)	1.915	1.997	0.082	1.793	1.9225	0.130
N···H (B)	1.774	1.863	0.089	1.887	1.991	0.104
CH···O (C)	2.741	2.336	−0.405	2.834	2.4055	−0.429
G•C						
O···H (A)	1.742	1.880	0.138	1.719	1.8805	0.162
N···H (B)	1.890	1.882	−0.008	1.879	1.963	0.084
NH···O (C)	1.883	1.843	−0.040	1.881	1.8305	−0.050

A more detailed comparison of the H-bonds in the two base pairs (Table 2) indicates that H-bond **A** in G•C has the shortest interaction distance in the planar structure and lengthens considerably more than any of the other H-bonds in either base pair when the bases are held perpendicular to each other. One might be tempted to interpret this as an indication of a large cooperative interaction involving this H-bond. However, the observation that both of the other H-bonds contract when the bases are perpendicular suggests that H-bond **A** is compressed in the planar base pair. This compression, which allows the other two H-bonds to enhance their interactions, relaxes in the perpendicular structure that preserves H-bond **A**. Variation of the structure of G•C by tilting (while keeping the base pair planar) about the central H-bond (**B** in Figure 1) would simultaneously shorten (strengthen) **A** and lengthen (weaken) **C**, or vice versa. Thus, small tilts of this nature would have a much more significant effect upon the structure than upon the energy of the base pair.

Bickelhaupt et al. have noted that the best calculated values for A•T are in better agreement with experiment than those for G•C. They have attributed the poorer agreement of the G•C H-bond lengths to the experimental environment. The inclusion of environmental effects in their calculations improved the agreement with experiment.[10] Another possible explanation might be that BSSE affects the geometry of the G•C pair more than A•T. The combined BSSE of the three H-bonds should exert a larger nonphysical attraction between G and C than between A and T, while the ability of the G•C pair to tilt about H-bond **B** would provide it with an additional means of relaxing under this artifactual force. We have previously shown the H-bonding distances on counterpoise-corrected surfaces to be very sensitive to small changes in the corrected interaction energies.[31]

One should note that the observations that both base pairs have substantial cooperative interactions differ from the predictions that might be expected from the secondary electrostatic interactions described originally by Jorgensen.[32] However, we had previously noted that the interaction energy of acetic acid dimer is more than twice that expected for one H-bond,[33] in disaccord with these predictions. More recently, Leszczynski has also concluded that secondary electrostatic interactions can be deceptive based upon a comprehensive study of 17 multiply H-bonded complexes.[34]

Conclusions

The cooperative contribution to each of the A•T and G•C pairs is similar in magnitude, but provides a much larger fraction of the overall stability of A•T, which has the weaker overall interaction.

The relatively small fraction of cooperativity in the G•C pair appears to be due to the inability of two rigid molecules to form three H-bonds with close to optimal geometries for each individual H-bond. A generalization of this observations suggests that similar problems should be inherent in other systems with more than two H-bonds between inflexible molecules. Thus, the entropic advantage of multiple H-bonds between rigid (rather than flexible) molecules can be partially offset by enthalpic disadvantage of less than optimal individual H-bonding structures.

Acknowledgment. This work was supported in part by grants from the National Institutes of Health (S06GM60654), and from PSC-CUNY. A.A. is most indebted to the Minesterio de Educación, Cultura y Deporte (Spain) for a postdoctoral fellowship.

References and Notes

(1) Clementi, E.; Mehl, J.; Von Niessen, W. *J. Chem. Phys.* **1971**, *54*, 508.
(2) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493.
(3) Sponer, J.; Hobza, P.; Leszczynski, J. *Comput. Chem.* **1996**, *1*, 185.
(4) Sponer, J.; Hobza, P. *Chem. Phys.* **1996**, *204*, 365.
(5) Brameld, K. A.; Goddard, W. A., III. *J. Am. Chem. Soc.* **1999**, *121*, 985.
(6) Brameld, K.; Dasgupta, S.; Goddard, W. A., III. *J. Phys. Chem. B* **1997**, *101*, 4851.
(7) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *Chemistry− A European Journal* **1999**, *5*, 3581.
(8) Sponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1999**, *103*, 2528.
(9) Dannenberg, J. J.; Tomasz, M. *J. Am. Chem. Soc.* **2000**, *122*, 2062.
(10) Guerra, C. F.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117.
(11) Hobza, P.; Sponer, J.; Cubero, E.; Orozco, M. *J. Phys. Chem. B* **2000**, *104*, 6286.
(12) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024.
(13) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746.
(14) Kobko, N.; Paraskevas, L.; del Rio, E.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2001**, *123*, 4348.
(15) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2003**, *125*, 8124.
(16) Suhai, S. *Int. J. Quantum Chem.* **1994**, *52*, 395.
(17) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273.
(18) Sheridan, R. P.; Lee, R. H.; Peters, N.; Allen, L. C. *Biopolymers* **1979**, *18*, 2451.
(19) Van Duijnen, P. T.; Thole, B. T. *Biopolymers* **1982**, *21*, 1748.
(20) Zhao, Y.-L.; Wu, Y.-D. *J. Am. Chem. Soc.* **2002**, *124*, 1570.
(21) Wu, Y.-D.; Zhao, Y.-L. *J. Am. Chem. Soc.* **2001**, *123*, 5313.
(22) Belosludov, R. V.; Li, Z.-Q.; Kawazoe, Y. *Mol. Eng.* **1999**, *8*, 105.
(23) Masunov, A.; Dannenberg, J. J. *J. Phys. Chem. B* **2000**, *104*, 806.
(24) Dovesi, R.; Causa, M.; Orlando, R.; Roetti, C.; Saunders, V. R. *J. Chem. Phys.* **1990**, *92*, 7402.
(25) Turi, L.; Dannenberg, J. J. *J. Am. Chem. Soc.* **1994**, *116*, 8714.
(26) Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1992**, *96*, 5819.
(27) Suhai, S. *J. Phys. Chem.* **1996**, *100*, 3950.
(28) 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.; Salvador, P.; Dannenberg, J. J.; 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*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 1998.
(29) Becke, A. D. *J. Chem Phys.* **1993**, *98*, 5648.
(30) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(31) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Phys. Chem. A* **1999**, *103*, 1640.
(32) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.
(33) Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1993**, *97*, 12197.
(34) Lukin, O.; Leszczynski, J. *J. Phys.. Chem. A* **2002**, *106*, 6775.