418 *J. Phys. Chem. A* **2003,** *107,* 418−426

An Analysis of the Interactions between Nucleic Acid Bases: Hydrogen-Bonded Base Pairs

Rafał R. Toczyłowski and Sławomir M. Cybulski*

Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056

Received: August 5, 2002; In Final Form: November 12, 2002

Thirty-two nucleic acid hydrogen-bonded base pairs have been examined using second-order Møller−Plesset perturbation theory (MP2) and the 6-31G*(0.25) and modified aug-cc-pVDZ basis sets. Complexes of C_s symmetry have been constructed from rigid monomers which allowed us to use a compact parametrization of their geometry based on the center of mass separation and Euler angles. The dependence of the interaction energy and its components on small geometrical modifications around the minima have been examined. The electrostatic and the exchange energies have been found to be the most important components of the overall interaction energy, although the dispersion and the induction energies also play important roles. The exchange energy, while typically not the largest in magnitude at the minima, is the most anisotropic component for rotations of the monomers in the plane of the complex. The analysis of the electron-correlated components reveals that the effect of the attractive dispersion energy is to a large degree canceled out by the repulsive correlation correction to the exchange energy.

1. Introduction

The number of publications concerned with modeling of nucleic acids or their fragments continues to grow at an impressive rate. Recent activities in this area have been summarized in several reviews whose authors discuss how both experimentalists and theorists attempt to understand the behavior of nucleic acids.[1−4] Perhaps the most important aspect of modeling studies is the choice of the empirical force-field. According to the developers of the AMBER 4.1 force-field "we have reached the limit for accurately representing biomolecular systems with an effective two-body additive potential employing quantum mechanically derived atom centered charges",[5] so it is appropriate to consider the development of a new generation of force-fields. It is not yet clear what functional forms they will employ but it is certain that they will be based on parametrization of quantum chemical calculations. Ab initio methods should be particularly useful for developing new potentials for nonbonding interactions because of the large number of parameters required and the scarcity of relevant experimental results. In addition, the availability of separate components of the interaction energy may be beneficial in the model potential development.

Most of ab initio studies of nucleic acid bases (NABs) published so far have been authored by Hobza and co-workers. They have been summarized in a recent review.[6] The focus of their work was on the optimization of geometries of NAB pairs and the calculations of interaction energies at the minima. They used the self-consistent field (SCF) method and the standard 6-31G**[7,8] basis set for optimizations, but the interaction energies were obtained from the second-order Møller−Plesset perturbation theory calculations (MP2) in which the 6-31G*(0.25) basis set was used.[9] The latter was obtained by replacing a polarization d function with the exponent of 0.8 in the standard 6-31G*[7,8] basis set with a more diffuse d function with the exponent of 0.25. Harmonic vibrational frequencies were also calculated to ensure that the true minima were found. Unfortunately, our requests for the optimized SCF geometries remained unanswered so it was impossible for us to verify the results of Hobza et al.[6] Recently, Šponer and Hobza[10] reported the results of MP2 optimizations of four H-bonded configurations. They did not correct their results for the basis set superposition error (BSSE) which caused the intermolecular distances to be artificially too short. In the present paper an attempt is made to eliminate this problem.

Despite large number of publications devoted to interactions of NABs Hobza et al. have not addressed several key issues. The most important problem is that their supermolecule results suffer from the usual problem of the lack of insight into the nature of the interactions. They tested the sensitivity of their MP2 results to the basis set size and found "a steep improvement" associated with an increase of the size of the basis set.[10] However, they performed their analysis for only a few base pairs and did not explain the reasons which were responsible for the basis set effects that they found. By focusing on single optimized geometries of the H-bonded base pairs they did not address the issue of the dynamical nature of intermolecular interactions. Even though the H-bonded interactions involving NABs are rather strong, the complexes are certainly not going to be rigid and fixed to their optimized geometries, but will undergo various kinds of motions which will not be properly understood until accurate potential energy surfaces (PESs) are developed.

Due to the number of atoms and electrons in nucleic acid bases it is still impractical to study their intermolecular complexes by employing large basis sets and sophisticated electron-correlated methods like, for example, coupled-cluster approach. Such calculations can be performed for a limited number of points and are reliable, but are too time-consuming to be used for developing potential energy surfaces. This is the main reason in most ab initio studies of nucleic acid base pairs the cheapest electron-correlated approach, second-order Møller−Plesset perturbation theory (MP2), has been used. There were also attempts to use density functional theory (DFT) to study NAB pairs, but as Šponer and Hobza[10] observed "this method cannot be used for reference calculations even for H-bonded systems, for reasons well documented in the literature".[10] Recently, a more sophisticated DFT studies[11,12] have appeared,

10.1021/jp021805r CCC: $25.00

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{pmatrix} \cos\chi & -\sin\chi & 0 \\ \sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\phi\cos\theta\cos\chi - \sin\phi\sin\chi & -\sin\phi\cos\theta\cos\chi - \cos\phi\sin\chi & \sin\theta\cos\chi \\ \cos\phi\cos\theta\sin\chi + \sin\phi\cos\chi & -\sin\phi\cos\theta\sin\chi + \cos\phi\cos\chi & \sin\theta\sin\chi \\ -\cos\phi\sin\theta & \sin\phi\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

but questions have been raised about the correctness of the empirical treatment of the dispersion energy in these hybrid (DFT + empirical dispersion) approaches.[13] It has been shown[13] that in the vicinity of the global minimum for the H-bonded thymine-adenine base pair the dispersion energy obtained at the MP2 level of theory with even small basis sets is larger in magnitude by 100−400% than the empirical dispersion energy added to the DFT results.[11]

We treat this work as an exploratory investigation of potential energy surfaces of more than 30 nucleic acid base pairs in the immediate vicinity of the H-bonded minima. We present the decomposition of the supermolecule interaction energy into physically meaningful components: electrostatic, exchange, induction and dispersion energies. To the best of our knowledge, such an analysis has not been presented so far for nucleic acid base pairs. To make our results easily reproducible, we devised a compact notation that without any additional information should allow any interested researcher to reproduce our supermolecule results.

In the following section we present our methodology. Next follows a presentation of the results. We conclude with a brief summary and suggestions for future studies.

2. Methods

2.1. Description of the Geometry of the Base Pairs. The nucleic acid bases are denoted by one or two letter abbreviations. We consider the four DNA bases: adenine (A), cytosine (C), guanine (G), and thymine (T), as well as fluorouracil (FU), hypoxanthine (HX), and uracil (U). To describe the complexes, we follow the abbreviations used by Hobza et al.[6] and where necessary we use numbers. We do it to distinguish between different configurations of the same complex and not to introduce any kind of ordering. For the most common configurations we use the usual abbreviations: Watson−Crick (WC), reversed Watson−Crick (RWC), Hoogsteen (H), and reversed Hoogsteen (RH). Most of the complexes studied are constructed from the four DNA bases: A, C, G, and T. The additional bases are each used in one complex. FU and U as a replacement of T in the T−A−WC base pair, and HX instead of G in the C−G−WC complex.

The geometry of each of the seven nucleic acid bases were optimized at the MP2 level of theory using the 6-31G(d,p) basis set, imposing the C_s point group symmetry in each case. The geometries are very similar to those reported by Johnson et al.[14] with the largest structural differences occurring for the bases containing the amine group: adenine, cytosine, and guanine. Once the geometry of each base was optimized, the intramolecular parameters were frozen and the C_s symmetry was kept in all subsequent calculations of nucleic acid base pairs. This approach enabled us to propose a simple and compact parametrization of the geometries of the hydrogen-bonded nucleic acid base pairs based on the center of mass separation and Euler angles. For each base the principal axes orientation (PAO) was chosen as the reference geometry, that is the geometry for which all of the Euler angles are equal to zero. The moment of inertia tensor for the PAO is diagonal, with the X, Y, and Z eigenvalues arranged in ascending order. The reference geometries of all seven nucleic acid bases are available as Supporting Information. The PAO coordinates, (X, Y, Z), can be transformed into coordinates in the complex, (x, y, z), using the above transformation (eq 1).

In the present case, because of the C_s symmetry of our complexes, only five parameters are needed to fully characterize the geometry: the center of mass separation, R, and four Euler angles (two for each base), ϕ_A, θ_A, ϕ_B, and θ_B. The remaining two Euler angles, χ_A and χ_B, are both 0°. As can be seen in eq 1, ϕ and θ are the angles of rotation of a molecule around the Z and Y axes, respectively. Moreover, in our case, θ_A and θ_B may only be equal to 0° or 180°. The angles refer to the orientation in which the centers of mass of the two bases A and B forming a complex A−B are both on the x axis with $x_{\mathrm{com(B)}} > x_{\mathrm{com(A)}}$, where "com(X)" refers to the center of mass of the appropriate base. This is shown in Figure 1. The optimized values of R, ϕ_A, θ_A, ϕ_B, and θ_B for each base pair examined are listed in Table 1. We note, however, that the Euler angles are not unique for complexes of C_s symmetry and various equivalent sets of parameters are possible. For example, using the Euler angles $\phi_X' = \phi_X - 180°$, $\theta_X' = \theta_X - 180°$, and $\chi_X' = \chi_X$ (X = A and B) instead of ϕ_X, θ_X, χ_X gives an equivalent configuration.

The optimal geometries were found by performing single point calculations at the MP2 level of theory with the 6-31G*(0.25) basis set. All calculations were performed with GAUSSIAN 94[15] and TRURL 98[16] packages. The center of mass separation, R, was varied by 0.05 Å and the Euler angles, ϕ_A and ϕ_B, by 5°. In each case the interaction energy was corrected for the basis set superposition error (BSSE) by using the counterpoise procedure of Boys and Bernardi.[17] We decided not to try to find the minima more precisely because of the limitations of both the method and the basis set. If we were able to use a more sophisticated ab initio method or a larger basis set the optimal parameters could have easily changed by more than 0.05 Å and 5°.

Subsequently, we found the interaction energies for each minimum with the modified aug-cc-pVDZ basis set which we will refer to as m-avdz. The m-avdz basis set is a result of the decontraction of the aug-cc-pVDZ[18−20] basis set according to the algorithm proposed by Davidson.[21] The results obtained with the modified basis set are virtually identical to those obtained with the original aug-cc-pVDZ basis set but can be achieved with savings in execution time for programs which do not take advantage of using the same primitive functions for several contracted functions. We stress that because of the small size of the 6-31G*(0.25) basis set and the accompanying underestimation of the dispersion contribution, the intermolecular distances given in Table 1 are too long. For several complexes, such as C−G−WC, T−A−WC, and T−A−H, for which we performed additional calculations with the m-avdz basis set, the optimal intermolecular distances were by 0.05−0.10 Å shorter than those in Table 1.

The just described approach can be contrasted with the standard unconstrained optimization in which all inter- as well as intramolecular parameters are optimized. At the end of optimization one obtains the geometry which, considering the

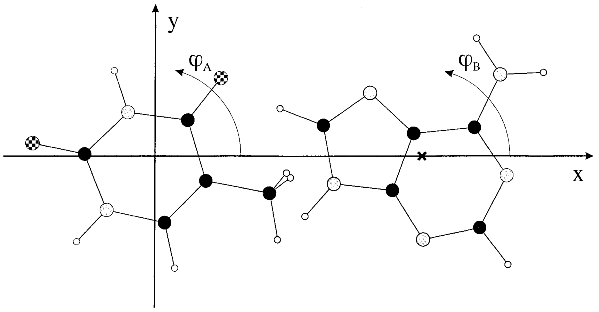

Figure 1. The initial orientation for a base pair prior to applying the in-plane rotations (ϕ_A and ϕ_B) that result in structures listed in Table 1. Rotations by the θ angles are not shown. The center of mass of the first base is at the origin and the center of mass of the second base lies on the x axis at the indicated point.

TABLE 1: Intermolecular Parameters[a]

base pair	R (Å)	ϕ_A (deg)	θ_A (deg)	ϕ_B (deg)	θ_B (deg)
C−G−WC	5.75	95	180	35	0
G−G−1	6.45	50	180	230	180
C−HX	5.90	95	180	0	180
T−G−3	6.60	330	180	235	180
G−G−3	6.70	325	0	210	180
C−C−1	5.50	115	180	295	180
T−G−1	6.15	90	180	235	180
A−G−1	6.75	175	180	220	180
T−G−2	6.40	225	0	240	180
C−A−1	6.25	285	0	340	180
T−A−H	5.80	255	0	255	180
T−A−RH	5.85	240	0	70	0
G−G−5	7.05	290	0	20	0
C−G−1	6.30	115	180	130	180
C−A−2	5.80	285	0	95	0
C−G−RWC	6.60	55	180	195	180
FU−A	6.20	75	180	0	180
T−A−WC	6.15	75	180	0	180
U−A	5.90	95	180	5	180
A−G−3	6.40	260	0	225	180
T−A−RWC	6.15	65	180	185	0
A−A−1	7.10	345	0	170	0
A−G−4	6.95	340	0	120	180
C−T−2	5.40	275	0	255	180
A−A−2	6.65	170	180	90	0
T−T−1	5.95	225	0	280	180
T−T−2	5.65	270	0	90	0
T−T−3	6.20	220	0	40	0
C−T−1	5.40	90	180	245	180
A−G−2	6.60	100	180	120	180
A−A−3	6.30	265	0	85	0
G−G−4	7.10	310	180	130	180

[a] Euler angles ϕ_A or θ_A always refer to the base given first.

methods and basis sets that could currently be used in automatic optimizations, will not be much more reliable than the geometry we found. It should also be kept in mind that the interaction energy of van der Waals complexes of large molecules is only a tiny portion of the overall energy and that somewhat different optimized geometries can be obtained depending on the cutoff criteria. An additional problem which is not easy to circumvent is the difficulty of saturating the dispersion energy which all basis sets customarily used in automatic optimizations severely underestimate. All these factors influenced our decision to search for only approximate minima. We also note that while still dictated by the available resources the optimization of all geometrical parameters of complexes even larger than nucleic acid base pairs became routine. On the other hand, the exploration of larger portions and particularly parametrization of multidimensional potential energy surfaces is still far from routine and in many cases it is prohibitively expensive. For example, for the thymine-adenine complex there are 84 degrees of freedom if both intra- and intermolecular parameters are considered. If no new efficient methods of parametrization are introduced soon it seems very likely that in the foreseeable future the rigid-monomer calculations will be the method of choice of exploration of multidimensional potential energy surfaces outside of the immediate vicinity of the minimum region.

We verified for a few complexes that the quantitative geometrical differences between our C_s symmetry structures and fully optimized geometries were comparable in magnitude to differences caused by using different basis sets which we described earlier. For example, for the C−G−WC complex, the absolute differences between our optimized geometry and the one found in the MP2 gradient optimization with the 6-31G-(d,p) basis set and without the removal of BSSE do not exceed 0.07 Å for the bond lengths and 0.17 Å for the intermolecular distances. In general, we expect these differences not to exceed 0.1−0.2 Å. The difference between the interaction energies for the two geometries is somewhat more significant. For the C−G−WC complex the results obtained with the 6-31G*(0.25) basis set differed by 2.3 kcal/mol. The differences for other base pairs are, in general, expected to be smaller although slightly larger difference for the strongly interacting bases would not be very surprising.

2.2. Analysis of ΔE^{SCF} and $\Delta E^{(2)}$. As mentioned earlier, despite a large number of publications devoted to interactions of nucleic acid bases, the physical origin of these interactions has not been properly explained. Intermolecular Møller−Plesset perturbation theory[22] allows us to understand the supermolecule interaction energies in terms of physically meaningful components: electrostatic, exchange, induction, and dispersion.[23,24] The MP2 interaction energy ΔE^{MP2} is the sum of the SCF interaction energy ΔE^{SCF} and the second-order correlation correction $\Delta E^{(2)}$.

$$\Delta E^{MP2} = \Delta E^{SCF} + \Delta E^{(2)} \tag{2}$$

ΔE^{SCF} can be decomposed into electrostatic, exchange, and deformation contributions.

$$\Delta E^{SCF} = \epsilon_{es}^{(10)} + \epsilon_{exch}^{HL} + \Delta E_{def}^{SCF} \tag{3}$$

The sum of $\epsilon_{es}^{(10)}$ and ϵ_{exch}^{HL} is also known as the Heitler−London (HL) energy.

$$\Delta E^{HL} = \epsilon_{es}^{(10)} + \epsilon_{exch}^{HL} \tag{4}$$

The deformation energy contains the second- and higher order induction energies with exchange effects taken into account. For large intermolecular separations, ΔE_{def}^{SCF} converges to the second-order induction energy $\epsilon_{ind,r}^{(20)}$.

The second-order correlation correction to the interaction energy $\Delta E^{(2)}$ contains the second-order correlation correction to the electrostatic energy $\epsilon_{es,r}^{(12)}$, the second-order dispersion energy $\epsilon_{disp}^{(20)}$, the second-order correlation correction to the SCF deformation $\Delta E_{def}^{(2)}$, and the second-order correlation correction to the HL exchange $\Delta E_{exch}^{(2)}$.

$$\Delta E^{(2)} = \epsilon_{es,r}^{(12)} + \epsilon_{disp}^{(20)} + \Delta E_{def}^{(2)} + \Delta E_{exch}^{(2)} \tag{5}$$

All partitioning calculations were carried out using the TRURL 98[16] package.

3. Results and Discussion

Thirty-two H-bonded complexes of 13 different nucleic acid base pairs which were examined are shown in Figure 2. The

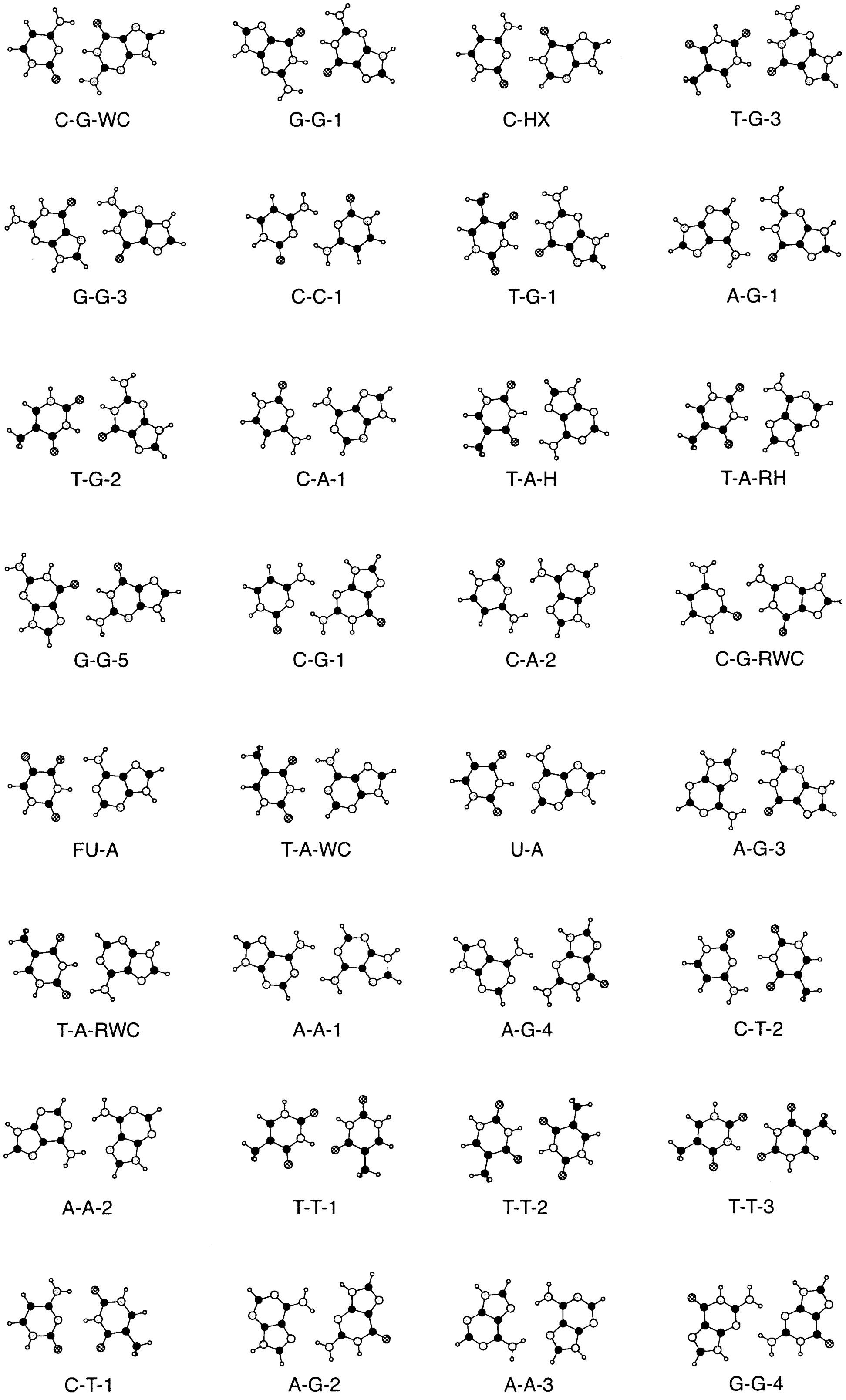

Figure 2. Optimized geometries of the 32 base pairs.

interaction energies for 26 of them can be found in the review by Hobza and Šponer.[6] The new base pairs are C−G−RWC, C−HX, FU−A, G−G−5, T−G−3, and U−A. The complexes shown in Figure 2 are ordered according to the decreasing magnitude of the interaction energies found with the m-avdz basis set. We note that in all tables presented in this work we keep the same ordering.

For each complex, at least seven points in the vicinity of the minimum were examined. The total MP2/6-31G*(0.25) interaction energies at these points are listed in Table 2. As can be seen, a change of intermolecular distance by 0.05 Å does not affect the interaction energy very much. None of the deviations from the energy at the minimum is greater than 0.2 kcal/mol, which amounts to no more than 1.4% of ΔE^{MP2}. This indicates

TABLE 2: ΔE^{MP2} (in kcal/mol) for the Minima (min) and Several Neighboring Configurations Obtained with the 6-31G*(0.25) Basis Set

		ΔR		$\Delta\phi_{\mathrm{A}}$		$\Delta\phi_{\mathrm{B}}$	
base pair	min	−0.05 Å	+0.05 Å	−5°	+5°	−5°	+5°
C−G−WC	−23.5	−23.4	−23.4	−22.3	−21.6	−20.7	−22.7
G−G−1	−21.6	−21.4	−21.5	−20.0	−21.0	−21.2	−21.0
C−HX	−18.1	−18.1	−17.9	−17.6	−16.7	−17.1	−17.3
T−G−3	−17.8	−17.6	−17.8	−17.0	−16.4	−16.9	−16.7
G−G−3	−17.1	−17.1	−17.0	−16.1	−16.9	−16.5	−15.9
C−C−1	−17.1	−17.1	−17.1	−16.8	−15.4	−16.8	−15.4
T−G−1	−14.1	−14.1	−14.0	−13.4	−13.1	−13.7	−13.0
A−G−1	−13.7	−13.6	−13.6	−11.5	−13.6	−12.4	−13.2
T−G−2	−13.3	−13.3	−13.2	−12.9	−11.4	−13.1	−12.2
C−A−1	−13.4	−13.4	−13.3	−12.6	−12.6	−12.0	−12.8
T−A−H	−12.7	−12.7	−12.6	−12.7	−11.7	−12.3	−11.3
T−A−RH	−12.6	−12.4	−12.6	−11.9	−11.7	−12.2	−10.9
G−G−5	−13.3	−13.3	−13.2	−13.2	−12.7	−12.8	−12.8
C−G−1	−12.9	−12.8	−12.9	−12.7	−11.5	−12.6	−9.6
C−A−2	−12.5	−12.4	−12.5	−12.3	−10.7	−11.1	−12.0
C−G−RWC	−12.6	−12.5	−12.6	−12.2	−12.2	−12.4	−11.6
FU−A	−12.3	−12.2	−12.2	−12.0	−11.0	−10.7	−11.7
T−A−WC	−12.1	−12.0	−12.0	−11.1	−11.6	−10.5	−11.6
U−A	−12.1	−12.0	−12.0	−11.2	−11.5	−11.1	−11.1
A−G−3	−12.0	−11.9	−12.0	−11.2	−10.2	−11.4	−11.2
T−A−RWC	−11.6	−11.5	−11.6	−10.2	−11.2	−9.6	−11.5
A−A−1	−10.7	−10.7	−10.6	−9.3	−10.0	−10.5	−9.3
A−G−4	−10.6	−10.5	−10.5	−9.9	−9.4	−9.3	−9.2
C−T−2	−10.6	−10.5	−10.5	−9.8	−9.3	−9.8	−9.8
A−A−2	−10.1	−10.0	−10.1	−9.2	−9.3	−9.2	−8.8
T−T−1	−9.9	−9.8	−9.9	−9.0	−9.2	−9.8	−8.6
T−T−2	−9.8	−9.7	−9.8	−8.4	−9.8	−8.4	−9.8
T−T−3	−9.8	−9.7	−9.8	−8.8	−9.0	−8.8	−9.0
C−T−1	−9.7	−9.7	−9.7	−9.0	−8.5	−9.6	−8.0
A−G−2	−9.1	−9.0	−9.1	−9.0	−7.2	−8.6	−7.1
A−A−3	−8.7	−8.7	−8.7	−7.5	−8.5	−7.5	−8.5
G−G−4	−8.8	−8.7	−8.7	−8.5	−6.1	−8.5	−6.1

that it will be very easy for the interacting bases to change somewhat the intermolecular separation without greatly affecting the interaction energy. Future calculations should explore the interaction in a wider radial interval.

On the other hand, the interaction energies are quite sensitive to small changes in the angular parameters, although usually the four rotations listed in Table 2 vary in impact that they have on the interaction energy. For example, for the T−T−2 complex a change of ϕ_{B} by +5° does not change the interaction energy, but a change of the same angle by −5° causes a drop in the magnitude of the interaction energy by 1.4 kcal/mol. The largest deviations from the energy at the minimum caused by a rotation of one of the bases by ±5° in the molecular plane are observed for the following three complexes: C−G−1 (3.3 kcal/mol), C−G−WC (2.8 kcal/mol), and G−G−4 (2.7 kcal/mol).

The components of ΔE^{MP2} found with the 6-31G*(0.25) basis set are presented in Table 3. The largest component of the interaction energy in terms of its magnitude is the first-order electrostatic energy, $\epsilon_{\mathrm{es}}^{(10)}$. The second in importance is $\epsilon_{\mathrm{exch}}^{\mathrm{HL}}$ which in one case (G−G−4) exceeds the magnitude of $\epsilon_{\mathrm{es}}^{(10)}$. $\Delta E_{\mathrm{def}}^{\mathrm{SCF}}$ and $\epsilon_{\mathrm{disp}}^{(20)}$ are also large indicating that the overall interaction energy is a result of complex interplay of various factors. We also note the importance of the correlated exchange energy which we can only deduce indirectly from a comparison of $\Delta E^{(2)}$ with the sum of $\epsilon_{\mathrm{es,r}}^{(12)}$ and $\epsilon_{\mathrm{disp}}^{(20)}$. The remaining two terms in eq 5, $\Delta E_{\mathrm{def}}^{(2)}$ and $\Delta E_{\mathrm{exch}}^{(2)}$, must be fairly large in magnitude and repulsive to cancel out, to a large degree, the effect of the attractive $\epsilon_{\mathrm{disp}}^{(20)}$. We verified that $\epsilon_{\mathrm{ind,r}}^{(22)}$, which is the leading component of $\Delta E_{\mathrm{def}}^{(2)}$, is small and therefore only $\Delta E_{\mathrm{exch}}^{(2)}$ could be responsible for the large repulsive contribution to $\Delta E^{(2)}$.

The results in Table 3 show that in general the inclusion of the electron-correlation is essential to the quantitative description of the H-bonded nucleic acid bases. $\Delta E^{(2)}$ constitutes more than 20% of ΔE^{MP2} for 19 out of 32 base pairs considered. As we will show later, $\Delta E^{(2)}$ plays an even more important role as the size of the basis set is increased. The tabulated values of $\Delta E^{(2)}$ vary from 0.1 to −4.2 kcal/mol. Such a large range of values indicates that the impact of the electron correlation on the interaction energy is not uniform for all bases. For some complexes, e.g., G−G−1 and T−G−3, $\Delta E^{(2)}$ is only a small fraction of ΔE^{MP2}, and for some other base pairs, e.g., A−A−2 and A−A−3, it amounts to more than 40% of ΔE^{MP2}. We note that for the G−G−1 and T−G−3 complexes $\Delta E^{(2)}$ was found to have a small positive value. This will no longer be the case if larger basis sets are used.

The importance of the electron correlation increases for shorter distances since the overall interaction energy becomes smaller in magnitude, but the correlation correction becomes larger. For the C−G−WC complex, for example, the addition of $\Delta E^{(2)}$ to ΔE^{SCF} causes a 6.3% increase of the strength of the interaction at $R = 5.75$ Å, but 8.5% at $R = 5.60$ Å. For T−A−WC, the effect of $\Delta E^{(2)}$ is even more significant: 28.5% at $R = 6.15$ Å and 37.4% at $R = 6.00$ Å. For longer distances, the electron correlation gradually loses its importance, but it should be stressed that the distances at which it can be neglected are far away from the minima and that in all studies of the NAB pairs electron correlation should be taken into account. When analyzing the influence of rotations on $\Delta E^{(2)}$, we found that the trends vary from complex to complex. For example, for the T−A−WC base pair, all of the in-plane rotations of the bases by ±5° lead to configurations for which $\Delta E^{(2)}$ is more important than at the minimum. However, the opposite is true for the C−G−WC complex, when cytosine is rotated by −5° or guanine is rotated by +5° in the plane of the complex. Even though the total MP2 interaction energies in the vicinity of the minima do not change much with the intermolecular separation, the components undergo more dramatic changes.

The examination of the influence of 5° rotations on the components of the interaction energy can be used as a measure of the anisotropy of areas in the vicinity of the H-bonded minima. In Table 4 we present the results for four selected, representative complexes: C−G−WC and G−G−1 (the two most stable), T−A−WC (midrange), and G−G−4 (the least stable). Depending on whether ϕ_{A} or ϕ_{B} is varied, different changes are observed with respect to the values at the appropriate minimum. Table 4 contains only the largest differences in the component energies caused by one of the ±5° rotations. As can be seen, the exchange energy is affected by the rotations much more than the largest in magnitude, electrostatic energy. The results in Table 4 also show that the SCF deformation energy is strongly affected and that the dispersion energy is, in general, the least anisotropic component.

To test the quality of the results obtained with the 6-31G*(0.25) basis set, we carried out additional calculations using a larger basis set. Table 5 contains the MP2 interaction energy and its components found with the m-avdz basis set for the same geometries as in Table 3. In this case $\Delta E^{(2)}$ constitutes more than 30% of ΔE^{MP2} for 18 out of 32 base pairs considered. A comparison of data in Tables 3 and 5 shows the deficiencies of the 6-31G*(0.25) basis set which are manifested not so much in the values of ΔE^{MP2}, but in the values of the components. The values of ΔE^{MP2} found with the 6-31G*(0.25) basis set are underestimated by 2−8% (no more that 1.2 kcal/mol) with respect to those found with the m-avdz basis set. The only

TABLE 3: ΔE^{MP2} and Its Components (in kcal/mol) for the Indicated Configurations Obtained with the 6-31G*(0.25) Basis Set

base pair	$\epsilon_{\mathrm{es}}^{(10)}$	$\epsilon_{\mathrm{exch}}^{\mathrm{HL}}$	$\epsilon_{\mathrm{ind,r}}^{(20)}$	$\Delta E_{\mathrm{def}}^{\mathrm{SCF}}$	ΔE^{SCF}	$\epsilon_{\mathrm{es,r}}^{(12)}$	$\epsilon_{\mathrm{disp}}^{(20)}$	$\Delta E^{(2)}$	ΔE^{MP2}
C−G−WC	−34.4	24.1	−12.9	−11.9	−22.1	1.4	−9.0	−1.4	−23.5
G−G−1	−32.5	21.7	−12.5	−10.9	−21.6	2.2	−8.3	0.0	−21.6
C−HX	−27.0	18.8	−10.2	−9.1	−17.3	1.5	−7.3	−0.8	−18.1
T−G−3	−28.3	20.0	−10.7	−9.6	−17.9	2.1	−7.0	0.1	−17.8
G−G−3	−23.2	15.1	−8.2	−7.4	−15.5	0.9	−6.9	−1.6	−17.1
C−C−1	−25.7	20.5	−10.3	−9.0	−14.2	−0.2	−7.6	−2.9	−17.1
T−G−1	−21.8	16.2	−8.5	−7.7	−13.3	1.0	−6.4	−0.8	−14.1
A−G−1	−18.7	14.8	−7.7	−6.5	−10.3	−0.7	−6.6	−3.3	−13.7
T−G−2	−21.2	16.0	−8.2	−7.4	−12.6	1.1	−6.3	−0.8	−13.3
C−A−1	−19.0	15.8	−7.6	−6.4	−9.7	−1.1	−6.4	−3.7	−13.4
T−A−H	−19.9	16.0	−7.2	−5.9	−9.8	−0.2	−6.8	−2.9	−12.7
T−A−RH	−21.1	18.1	−8.0	−6.6	−9.6	−0.2	−7.2	−3.0	−12.6
G−G−5	−15.7	8.5	−4.6	−4.4	−11.6	0.6	−4.8	−1.7	−13.3
C−G−1	−21.2	18.5	−8.7	−7.2	−9.9	−0.4	−6.9	−3.0	−12.9
C−A−2	−18.5	16.2	−7.4	−6.3	−8.5	−1.1	−6.8	−4.0	−12.5
C−G−RWC	−16.3	10.1	−5.5	−5.1	−11.3	0.6	−5.0	−1.3	−12.6
FU−A	−19.5	15.9	−7.6	−6.1	−9.8	−0.1	−6.5	−2.5	−12.3
T−A−WC	−20.1	17.1	−7.9	−6.4	−9.4	−0.2	−6.8	−2.7	−12.1
U−A	−20.7	18.0	−8.5	−6.7	−9.4	−0.2	−7.1	−2.7	−12.1
A−G−3	−16.8	14.4	−7.0	−6.2	−8.6	−0.5	−6.5	−3.4	−12.0
T−A−RWC	−19.9	17.6	−8.2	−6.5	−8.8	−0.2	−7.0	−2.8	−11.6
A−A−1	−15.8	13.8	−6.1	−4.9	−6.8	−1.3	−5.8	−3.8	−10.7
A−G−4	−16.7	15.2	−6.8	−5.3	−6.9	−1.1	−6.3	−3.7	−10.6
C−T−2	−14.5	12.4	−6.5	−5.6	−7.7	−0.6	−5.8	−2.8	−10.6
A−A−2	−15.6	14.5	−6.0	−4.8	−5.9	−1.3	−6.2	−4.2	−10.1
T−T−1	−17.1	13.9	−6.4	−5.5	−8.6	0.5	−5.7	−1.3	−9.9
T−T−2	−16.7	14.4	−6.8	−5.8	−8.1	0.0	−5.8	−1.8	−9.8
T−T−3	−16.6	13.5	−6.1	−5.3	−8.4	0.5	−5.5	−1.4	−9.8
C−T−1	−13.3	12.5	−6.6	−5.6	−6.4	−1.0	−5.9	−3.3	−9.7
A−G−2	−14.4	13.6	−5.7	−4.4	−5.3	−1.0	−6.1	−3.8	−9.1
A−A−3	−11.7	10.4	−4.2	−3.4	−4.7	−1.1	−5.6	−4.1	−8.7
G−G−4	−15.0	15.3	−6.8	−5.3	−5.0	−1.2	−6.3	−3.8	−8.8

TABLE 4: The Maximum Change in the Magnitudes of the Interaction Energy Components Caused by One of the ϕ_{A} or ϕ_{B} Rotations by ±5°

base pair	$\epsilon_{\mathrm{es}}^{(10)}$ (kcal/mol)	%	$\epsilon_{\mathrm{exch}}^{\mathrm{HL}}$ (kcal/mol)	%	$\Delta E_{\mathrm{def}}^{\mathrm{SCF}}$ (kcal/mol)	%	$\epsilon_{\mathrm{disp}}^{(20)}$ (kcal/mol)	%
C−G−WC	1.5	4.5	5.1	21.3	1.4	11.8	0.7	8.4
G−G−1	4.4	13.5	7.2	32.8	2.8	25.2	0.8	9.9
T−A−WC	2.1	10.6	3.4	19.9	0.9	14.4	0.6	7.9
G−G−4	3.4	22.7	7.7	50.5	1.8	34.5	1.6	25.4

exception among the 32 complexes is the least attractive structure, G−G−4, for which the 6-31G*(0.25) basis set overestimates the interaction energy by 0.3%. Two of the component energies, $\epsilon_{\mathrm{exch}}^{\mathrm{HL}}$ and $\Delta E_{\mathrm{def}}^{\mathrm{SCF}}$ are not significantly affected by the change of the basis set. However, $\epsilon_{\mathrm{es}}^{(10)}$ and $\epsilon_{\mathrm{disp}}^{(20)}$ are more sensitive to the basis set size. When the 6-31G*(0.25) basis set is used $\epsilon_{\mathrm{es}}^{(10)}$ is overestimated by 0.5−1.4 kcal/mol (3−7%), but $\epsilon_{\mathrm{disp}}^{(20)}$ is underestimated by 1.0−2.0 kcal/mol (14−16%) compared to the results obtained with the m-avdz basis set. These two effects to a large degree cancel each other out so that ΔE^{MP2} appears to be insensitive to the increase of the basis set size. Unfortunately, different components depend in different ways on the intermolecular separation and orientation and even a fairly small change of geometry can result in not only significant differences between components, but also the overall interaction energies obtained with different basis sets. In ref 13 the results for the T−A−WC complex are used to illustrate this problem. A change of the intermolecular separation from 6.15 to 5.85 Å increases the difference between ΔE^{MP2} values obtained with the 6-31G*(0.25) and m-avdz basis sets from 6% to more than 21%.[13]

In their review Hobza and Šponer[6] wrote that "stabilization of H-bonded NA base pairs is of electrostatic origin, therefore, rather reliable characteristics are obtained already at the HF level with medium-sized polarized basis sets of atomic orbitals." A comparison of ΔE^{SCF} values obtained with the 6-31G*(0.25) basis set and ΔE^{MP2} values obtained with the m-avdz basis set shows that for the 32 complexes considered in this work the former underestimates the latter by an average error of 25% and therefore, in our opinion, cannot be considered to be "rather reliable characteristics".

Table 6 contains lengths of hydrogen bonds and angles that give a measure of distortions of these bonds from linearity for all 32 complexes. The distances between the electronegative atoms, X_1 and X_2, involved in an H-bond are denoted by r. Asterisks indicate that oxygen is one of the electronegative atoms. α is equal to $180° - \beta$, where β is the $X_1{-}H{-}X_2$ hydrogen bond angle. It gives the distortion of an H-bond from linearity, with $\alpha = 0°$ indicating a linear H-bond. C−G−WC is the only complex with three hydrogen bonds, but the results for G−G−1 show that it is possible to form nearly as strong an interaction using what is typically described as two hydrogen bonds.[6] Unfortunately, the explanation of this phenomenon is not as simple as the one proposed by Hobza and Šponer.[6] It is not true that the G−G−1 base pair is so stable due to the interaction of the dipole moments. They are antiparallel which is not a particularly favorable orientation. Furthermore, we note the presence of two longer and highly nonlinear NH···O hydrogen bonds that involve amine groups. They certainly make the interaction energy somewhat more attractive. Hobza and

TABLE 5: ΔE^{MP2} and Its Components (in kcal/mol) for the Indicated Configurations Obtained with the m-avdz Basis Set

base pair	$\epsilon_{es}^{(10)}$	ϵ_{exch}^{HL}	$\epsilon_{ind,r}^{(20)}$	ΔE_{def}^{SCF}	ΔE^{SCF}	$\epsilon_{es,r}^{(12)}$	$\epsilon_{disp}^{(20)}$	$\Delta E^{(2)}$	ΔE^{MP2}
C−G−WC	−33.1	23.9	−12.9	−12.0	−21.2	1.8	−10.6	−3.3	−24.5
G−G−1	−31.1	21.6	−12.5	−11.0	−20.6	2.3	−9.8	−2.2	−22.8
C−HX	−26.1	18.6	−10.2	−9.2	−16.6	1.8	−8.5	−2.4	−19.0
T−G−3	−27.1	19.8	−10.7	−9.8	−17.1	2.0	−8.3	−2.0	−19.0
G−G−3	−22.6	15.0	−8.4	−7.6	−15.1	1.1	−8.0	−3.0	−18.1
C−C−1	−25.1	20.3	−10.4	−9.2	−14.0	0.5	−8.8	−4.0	−18.0
T−G−1	−20.6	16.1	−8.6	−7.8	−12.3	1.0	−7.6	−2.5	−14.8
A−G−1	−17.7	14.8	−7.8	−6.7	−9.6	−0.2	−7.7	−4.4	−14.1
T−G−2	−20.1	16.0	−8.3	−7.5	−11.6	1.1	−7.4	−2.5	−14.1
C−A−1	−18.5	15.6	−7.6	−6.5	−9.4	−0.3	−7.5	−4.4	−13.8
T−A−H	−19.4	15.9	−7.3	−6.1	−9.5	0.2	−7.9	−4.1	−13.6
T−A−RH	−20.6	18.0	−8.1	−6.8	−9.4	0.2	−8.4	−4.3	−13.6
G−G−5	−14.9	8.5	−4.7	−4.5	−11.0	0.8	−5.6	−2.5	−13.5
C−G−1	−20.4	18.3	−8.8	−7.3	−9.4	0.3	−8.0	−3.9	−13.3
C−A−2	−18.0	16.1	−7.5	−6.4	−8.3	−0.3	−8.0	−4.8	−13.1
C−G−RWC	−15.5	10.1	−5.6	−5.2	−10.6	0.8	−5.9	−2.3	−12.9
FU−A	−18.8	15.8	−7.6	−6.2	−9.3	0.4	−7.5	−3.6	−12.9
T−A−WC	−19.4	17.0	−8.0	−6.6	−9.0	0.3	−7.9	−3.9	−12.9
U−A	−20.1	17.9	−8.5	−6.8	−9.0	0.3	−8.2	−3.9	−12.9
A−G−3	−16.0	14.3	−7.1	−6.4	−8.0	0.0	−7.7	−4.5	−12.5
T−A−RWC	−19.4	17.5	−8.3	−6.6	−8.5	0.3	−8.2	−4.0	−12.5
A−A−1	−15.3	13.7	−6.2	−5.0	−6.6	−0.6	−6.8	−4.4	−11.0
A−G−4	−16.0	15.1	−6.8	−5.5	−6.4	−0.3	−7.3	−4.4	−10.8
C−T−2	−13.8	12.5	−6.6	−5.7	−7.0	−0.2	−6.8	−3.7	−10.8
A−A−2	−15.1	14.4	−6.1	−5.0	−5.7	−0.6	−7.3	−4.8	−10.5
T−T−1	−16.1	13.9	−6.5	−5.6	−7.8	0.5	−6.7	−2.7	−10.5
T−T−2	−15.7	14.4	−6.8	−5.9	−7.2	0.2	−6.9	−3.2	−10.4
T−T−3	−15.7	13.5	−6.2	−5.4	−7.6	0.5	−6.5	−2.8	−10.4
C−T−1	−12.7	12.6	−6.7	−5.7	−5.8	−0.5	−6.9	−4.2	−10.0
A−G−2	−13.8	13.5	−5.8	−4.6	−4.8	−0.3	−7.0	−4.5	−9.3
A−A−3	−11.2	10.4	−4.3	−3.6	−4.4	−0.5	−6.5	−4.6	−9.0
G−G−4	−14.1	15.2	−6.9	−5.4	−4.3	−0.4	−7.2	−4.4	−8.8

TABLE 6: Lengths and Distortions from Linearity of Hydrogen Bonds for Structures Given in Table 1[a]

base pair	r_1 [Å]	r_2 [Å]	α_1 [deg]	α_2 [deg]
C−G−WC[b]	*2.96	*3.02	4.2	1.0
G−G−1	*2.92	*2.92	13.6	13.6
C−HX	*2.97	3.01	2.4	4.4
T−G−3	*2.88	*2.95	2.0	4.3
G−G−3	2.98	*3.19	9.4	12.4
C−C−1	3.03	3.03	1.5	1.5
T−G−1	*2.96	*2.96	0.4	5.2
A−G−1	*2.96	3.21	10.2	6.4
T−G−2	*2.91	*2.99	3.4	12.7
C−A−1	3.05	3.19	1.4	1.7
T−A−H	2.97	*3.13	7.8	19.0
T−A−RH	2.95	*3.04	2.8	18.8
G−G−5	*2.96	3.12	5.0	26.3
C−G−1	2.97	3.17	5.8	0.6
C−A−2	3.03	3.12	4.0	16.3
C−G−RWC	*2.96	3.33	11.5	8.5
FU−A	*3.00	3.07	3.6	5.5
T−A−WC	*3.00	3.04	2.5	2.4
U−A	2.97	*3.10	0.3	6.0
A−G−3	*2.89	3.19	25.8	4.0
T−A−RWC	3.00	*3.07	2.3	0.0
A−A−1	3.07	3.27	2.6	4.6
A−G−4	3.09	3.15	1.0	3.4
C−T−2	*2.95	3.25	0.7	13.6
A−A−2	3.10	3.11	21.0	2.0
T−T−1	*2.91	*3.02	12.3	14.2
T−T−2	*2.98	*2.98	1.8	1.8
T−T−3	*2.97	*2.97	11.0	11.0
C−T−1	*3.00	3.20	3.8	7.6
A−G−2	3.10	3.11	11.1	25.6
A−A−3	3.16	3.16	18.1	18.1
G−G−4	3.11	3.11	7.6	7.6

[a] The bonds are ordered from the shortest to the longest. The lengths given in the table will be obtained when geometries from Table 1 are used. Asterisks denote hydrogen bonds with oxygen atoms involved. [b] $r_3 = 3.07$ Å, $\alpha_3 = 2.0°$.

Šponer also wrote that “the positive (destabilizing) value of ΔE^{COR}” ($\Delta E^{(2)}$ in our notation) “found for the GG1 pair is due to the fact that the repulsive contribution caused by reduction of the dipole moments of monomers due to electron correlation is larger in absolute value than the dispersion attraction".[10] This again is not true as even when the smaller, 6-31G*(0.25), basis set is used, the dispersion energy, $\epsilon_{\mathrm{disp}}^{(20)}$, is almost four times greater in magnitude than the second-order correlation correction to the electrostatic energy, $\epsilon_{\mathrm{es,r}}^{(12)}$ which takes into account the influence of intramolecular correlation effects on multipole moments. As we already mentioned, the repulsive correlated exchange energy cancels out most of the attractive dispersion energy.

There is no clear relationship between the length, linearity and number of hydrogen bonds and the magnitude of the interaction energy. More stable complexes tend to contain shorter and more linear hydrogen bonds, but for example, the T−T−2 complex contains two almost linear hydrogen bonds, yet it has essentially the same interaction energy as the interaction energies of the T−T−1 and T−T−3 complexes in which the H-bonds deviate from linearity by more than 13°.

Several complexes were chosen to examine the effect of functional groups on the H-bonds. A comparison of C−G−WC and C−HX indicates that breaking of one of hydrogen bonds influences the interaction energy significantly. In the case of C−HX where one of the amine groups is missing, only two hydrogen bonds can be formed and this causes a decrease of the magnitude of ΔE^{MP2} by 5.5 kcal/mol compared to C−G−WC. The magnitude of all component energies also decreases. On the other hand, when modified functional groups do not participate in H-bonds the interaction energies and their components are not significantly affected as the results for T−A−WC, FU−A, and U−A illustrate. Substitution of the methyl group in thymine by either fluorine or hydrogen atoms causes only minor changes in the interaction energies and the components, especially those found with the m-avdz basis set. The geometrical differences between these three complexes is a side effect of the definition of the intermolecular parameters, which depend on the center of mass separation and principal axes orientations. But the geometries of hydrogen bonds are very similar for the three complexes. The parameters characterizing the hydrogen bonds are in good agreement with the results of Šponer et al.,[9] although there are differences caused by different optimization procedures.

4. Summary and Conclusions

A picture of hydrogen bonding interactions between nucleic acid bases that emerges from our calculations is more complex than what has been presented in the literature heretofore. Typically,[6] the role of electrostatic energy was emphasized, which is not incorrect since this in most cases is the most important component of the interaction energy. However, as we have shown, other components are also very important. We doubt that simple Lennard-Jones type exchange−dispersion potentials[5] can properly reproduce the sum of exchange and dispersion energies, and the deformation (induction) energy is too important to be completely omitted from a force field. A very important role is also played by the correlation correction to the exchange energy $\Delta E_{\mathrm{exch}}^{(2)}$, which was found to be repulsive and which canceled out a major portion of the attractive dispersion energy resulting in small values of the supermolecule $\Delta E^{(2)}$ contribution.

It is difficult to gauge the quality of our results. If we applied the MP2 approach and a basis set of aug-cc-pVDZ size to study smaller complexes the results would only be of mediocre quality as many studies, including our own, amply illustrate.[25,26] But for complexes of nucleic acid bases, electron-correlated calculations with basis sets larger than m-avdz are too time-consuming for a large number of configurations. So at present we are unable to examine the effect of an additional increase of basis set size on the MP2 results. The differences between the 6-31G*(0.25) and m-avdz basis set seem to indicate that the effect could be fairly significant. This is confirmed by the recent MP2 result of Šponer and Hobza[10] obtained for the U−U−4 configuration of the uracil dimer with the aug-cc-pVTZ basis set. They found that a basis set upgrade from 6-31G*(0.25) to aug-cc-pVDZ changes the interaction energy more than if the aug-cc-pVDZ is replaced by the aug-cc-pVTZ basis set.

We did try to gauge the importance of higher orders of MP theory using fourth-order Møller−Plesset perturbation theory (MP4) and the 6-31G*(0.25) basis set to examine two complexes: C−G−WC and T−A−WC. In both cases the MP4 results were lower in energy than their MP2 counterparts, but by only 0.3 kcal/mol (1.2%) for C−G−WC and 0.2 kcal/mol (1.9%) for T−A−WC. Interestingly, also the MP3 values were very close to the MP4 (and MP2) results. However, the MP4(SD) and MP4(SDQ) interaction energies were smaller in magnitude and only the inclusion of triple excitations made the MP4 results similar to the MP2 ones. Šponer and Hobza[27] also found, for the uracil and the cytosine dimers, that the MP2 results were similar to the coupled-cluster, CCSD(T), interaction energies obtained with the cc-pVDZ basis set. Of course, it would be premature to make sweeping generalizations about the quality of the MP2 approach based on the results of only a few calculations, but so far the MP2 approach has not been found to be unacceptable. Certainly, more higher level calculations are needed.

In essentially every study of nucleic acid bases or base pairs or even larger clusters the authors emphasize the role that they play as carriers of genetic information. The most interesting phenomena, replication and transcription, are of dynamical nature yet the available ab initio studies have been primarily concerned with the structural and static aspects of the interaction. In our opinion a much more important contribution of ab initio calculations would be in developing new potentials which could be used in a more reliable treatment of dynamical phenomena. A brute force method based on fitting the results of thousands of ab initio calculations will be very expensive for systems as large as nucleic acid bases. It is premature to say that the approach based on separate fitting of energy components would be a better alternative, but there is no doubt that it could be useful for testing the suitability of various approximate functions to reliably model various component energies. There have been attempts based on a similar idea to develop intermolecular potentials but for much smaller systems,[28] so at this time it is not possible to predict how well this will work for nucleic acid bases pairs. In the future we plan to investigate out-of-plane and stacked configurations in addition to wider areas of potential energy surfaces for hydrogen bonded interactions. We also plan to test the reliability of the most popular force fields that are currently available.

Acknowledgment. This work was supported by the National Institutes of Health (Grant No. GM59586-01A1).

Supporting Information Available: Geometries of the seven bases in the principal axes orientations. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Auffinger, P.; Westhof, E. *Curr. Opin. Struct. Biol.* **1998**, *8*, 227.
(2) Lafontaine, I.; Lavery, R. *Curr. Opin. Struct. Biol.* **1999**, *9*, 170.

(3) Beveridge, D. L.; McConnell, K. J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 182.

(4) MacKerell, A. D., Jr.; Nilsson, L. *Comput. Biochem. Biophys.* **2001**, 441.

(5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(6) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247.

(7) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(8) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(9) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965.

(10) Šponer, J.; Hobza, P. *J. Phys. Chem. A* **2000**, *104*, 4592.

(11) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.

(12) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515.

(13) Cybulski, S. M.; Bledson, T. M.; Toczyłowski R. R. *J. Chem. Phys.* **2002**, *116*, 11039.

(14) Johnson, R. C.; Power, T. D.; Holt, J. S.; Immaraporn, B.; Monat, J. E.; Sissoko, A. A.; Yanik, M. M.; Zagorodny, A. V.; Cybulski, S. M. *J. Phys. Chem.* **1996**, *100*, 18875.

(15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheesman, J. R.; Keith, T. A.; 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 B.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(16) Cybulski, S. M. Trurl 98 package: Oxford, OH, 1998.

(17) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(18) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(19) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(20) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975.

(21) Davidson, E. R. *Chem. Phys. Lett.* **1996**, *260*, 514.

(22) Szalewicz, K.; Jeziorski, B. In *Molecular Interactions: From van der Waals to Strongly Bound Complexes*; Scheiner, S., Ed.; Wiley: Chichester, 1997; Chapter 1.

(23) Chalasiński, G.; Szczęśniak, M. M. *Mol. Phys.* **1988**, *63*, 205.

(24) Cybulski, S. M.; Chalasiński, G.; Moszynski, R. *J. Chem. Phys.* **1990**, *92*, 4357.

(25) Fernandez, B.; Koch, H. *J. Chem. Phys.* **1998**, *109*, 10255.

(26) Cybulski, S. M.; Toczyłowski R. R. *J. Chem. Phys.* **1999**, *111*, 10520.

(27) Šponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263.

(28) Mooij, W. T. M.; van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Eijck, B. P. *J. Phys. Chem. A* **1999**, *103*, 9872.