Mutual Relationship between Stacking and Hydrogen Bonding in DNA. Theoretical Study of Guanine−Cytosine, Guanine−5-methylcytosine, and Their Dimers

Carles Acosta-Silva, Vicenç Branchadell, Joan Bertran,* and Antoni Oliva*

Departament de Química, Universitat Autònoma de Barcelona, Bellaterra 08193 Spain

Received: April 28, 2010; Revised Manuscript Received: June 21, 2010

The mutual relationship between stacking and hydrogen-bonding and the possible influence of stacking in the different behavior of cytosine (C) and 5-methylcytosine (C′) in DNA have been studied through complete DFT optimization of different structures of G-C and G-C′ dimers (i.e., G-C/C-G and G-C′/C′-G), using four different functionals. Our results show that stacking leads to an increase of the $O_6\cdots H-N_4$ hydrogen bond length and to a simultaneous decrease of the $N_2-H\cdots O_2$ one, in such a way that both lengths approach each other and, in some cases, an inversion occurs. These results suggest that stacking can be a factor to explain the disparity between theory and experiment on the relative strength of the two lateral hydrogen bonds. Regarding the difference between cytosine and 5-methylcytosine, we have shown that methylation enhances the stacking interactions, mainly due to the increase of polarizability. Methylation also favors the existence of slid structures which can produce local distortions of DNA.

Introduction

Aromatic base stacking and hydrogen bonds were usually assumed to be independent interactions in nucleic acids.[1] However, thermodynamic studies have tentatively suggested that hydrogen bonds in DNA are coupled with stacking interactions.[2] Geerlings et al.[3,4] have studied the variation of the molecular electrostatic potential (MEP) in the vicinity of the hydrogen donor and of the hydrogen acceptor atoms when stacking interactions are present, and they have shown that the *intra*strand $\pi-\pi$ stacking interaction enhances the *inter*strand hydrogen-bond accepting capacity of DNA/RNA bases.[3]

In a previous study,[5] we carried out a complete DFT optimization of different structures of A-U and A-T dimers (i.e., A-U/U-A and A-T/T-A), using different functionals developed by the group of Truhlar,[6] to confirm the mutual relationship between stacking and hydrogen bonding. The system in which the connection was more evident was the A-U/U-A, corresponding to stacked Watson−Crick base pairing with an inversion center. In this case, the simultaneous effect of stacking on $N-H\cdots O$ and $N\cdots H-N$ hydrogen bonds leads to a reversal of the lengths of both hydrogen bonds, in such a way that the $N-H\cdots O$ hydrogen bond is the shortest one after stacking. A second aim of our work was to study the different behavior of DNA and RNA as a consequence of the replacement of a thymine base by an uracil one. We found that the A-U/U-A and A-T/T-A Watson−Crick structures are quite different when one looks at geometry, energy, charge transfer, electron density topology, or electrostatic potential. We showed that this peculiar behavior may be interpreted through the formation of a stabilizing CH/π interaction as described previously by us for an analogous system.[7]

For the guanine−cytosine (G-C) Watson−Crick base pairing (Figure 1), there is some discrepancy between the theoretical and experimental results about the relative length of the different hydrogen bonds. Thus, while theoretical results[8−11] show that the O6···N4 distance is shorter than the N2···O2 one, the experimental crystallographic values[12] present the opposite order. According to Baerends and Bickelhaupt,[13−16] the source of the divergence turns out to be the molecular environment (water, counterions, ...) of the base pairs in the crystal experimentally studied. Our previous results in the A-T and A-U systems may suggest that stacking can be another factor to explain the disparity between theory and experiment.

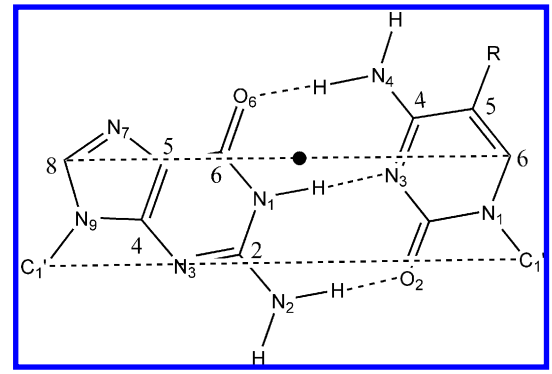

Figure 1. Scheme of a base pair (G-C′ if R = CH_3 and G-C if R = H). The dashed line C_6-C_8 represents the long base-pair axis, which is roughly parallel to the $C_1'-C_1'$ line, where C_1' stands for the sugar carbon atoms bonded to the bases. The twist angle is defined as the rotation around the midpoint of the C_6-C_8 axis (denoted by a dot).

The different behavior when a thymine base is replaced by a uracil one, can now be studied by methylation of the cytosine base to obtain the 5-methlcytosine (C′). DNA methylation in higher eukaryotes occurs at cytosines in CpG dinucleotides by the transfer of a methyl group from S-adenosylmethionine to the C5 position of cytosine catalyzed by DNA methyltransferases.[17−20] Methylation adds extra information to the DNA that is not encoded in the sequence. This so-called epigenetic information has many important biological functions. In eukaryotes, DNA methylation contributes to the control of gene expression, the protection of the genome against selfish DNA, maintenance of genome integrity, parental imprinting, X-chromosome inactivation in mammals, and regulation of development.[18,21] In addition, hypo- and hypermethylation reactions of DNA have been suggested to be involved in the initiation and progression steps of human cancers.[22−26] Deamination of 5-methylcytosine in cellular DNA, leading to the formation of

10.1021/jp103850h © 2010 American Chemical Society
Published on Web 07/21/2010

thymine, is more favorable than the corresponding deamination of cytosine to uracil, thus increasing the potential mutagenic effects.[22,27−30] The mutational damage of deamination of 5-methylcytosine is augmented by the lower repair efficiency of T/G mismatches (arising from deamination of 5-methylcytosine) as compared to U/G mismatches (arising from deamination of unmethylated cytosine). Uracil is a nonnatural base in DNA and is efficiently recognized and excised by the repair enzyme uracil-DNA glycosylase.[18] As a consequence, CpGs occur at one-fifth of their expected frequency in mammalian genomes, giving that the successive methylation and deamination of cytosine lead to TpG and CpA.[31,32]

Several reasons have been invoked to interpret the different behavior between cytosine and 5-methylcytosine in DNA. A first reason is the fact that the introduction of a methyl group in the C5 position of cytosine leads to an increase of the one-electron oxidation relative to the guanine−cytosine base pair.[33,34] A second interpretation suggests that the reactivity changes are based on an electronic effect that is transferred from the substituent through the cytosine carbonyl to the guanine amino group, in such a way that the rate of alkylation of the guanine by mitomycin is increased.[32,35−37] A third group of authors affirm that the 5-methyl group affects the local structure of DNA.[38−42] In fact, the 5-methyl group of cytosine in the major groove is found to be essential for the enzyme−DNA interaction.[38−40] In this work we will focus our attention on the enhancement of the stacking produced by the presence of the methyl group and the subsequent local distortions as, for instance, the change in the strength of the hydrogen bond between the carbonyl group of cytosine and the amino group of guanine.

Many papers have been devoted to the study of the stacking or hydrogen bonding in nucleic acids base pairs.[43−46] The main difficulty of this kind of study arises from the fact that hydrogen bonding originates mainly in electrostatic interactions while the stacking bonding is mainly due to the dispersion energy.[47,48] So, a high-level calculation in which both kinds of interactions are well represented is needed. This explains the very limited number of works published on the mutual relationship between both phenomena, since the size of the systems makes the calculations prohibitive. In one of them, Schütz et al.[49] have analyzed, using the symmetry adapted perturbation theory (SAPT), the different energy contributions in the stacking of two base pairs and they have shown that the dispersion energy is the main stabilizing term. Cooper et al.,[50] using the vdW-DF method, have mapped out the stacking interactions between acid base pairs as a function of the distance (rise) and the twist angle (twist) between them. However, the fact of keeping fixed the geometry of each one of the base pairs impedes the study of the coupling between hydrogen bonding and stacking. Recently, Deyà et al.[51] have studied the interplay between hydrogen bonding and stacking in Z-DNA through a model in which two cytosines are stacked while each one has three hydrogen bonds with a very simplified structure of guanine.

The object of the present work is to achieve a deeper insight on the mutual relationship of stacking and hydrogen-bonding with the aim of analyzing the possible influence of stacking in the different behavior of cytosine (C) and 5-methylcytosine (C′) in DNA base pairs. To reach this goal, we will carry out a complete DFT optimization of different structures of G-C and G-C′ dimers (i.e., G-C/C-G and G-C′/C′-G).

Computational Methods

The standard method for introducing dispersion energy in theoretical studies is to combine MP2 theory in the complete basis set (CBS) limit with a ΔCCSD(T) correction computed in a smaller basis to estimate the CBS CCSD(T) results.[44−46,52] However, this approach is not feasible for large systems of biological interest, such as the ones studied in this paper. One option is to use methods based on density functional theory (DFT), but classical functionals are not appropriate to treat systems in which dispersion forces are crucial, since they are too repulsive at long range.[53−55]

Among the most suitable methods, it is worth mentioning the building up of new functionals such as those proposed by Truhlar and co-workers.[56] In particular, M05-2X[56,57] and M06-2X[56,58] are highly parametrized metahybrid methods that are very adequate for the study of noncovalent interactions, since the equilibrium structures of noncovalent complexes are dominated by medium range exchange and correlation energies. Different authors have emphasized the good performance of the M06-2X method[59−63] although some others have shown some limitations.[64,65]

Other very suitable methods are the so-called DFT-D, in which an empirical correction to account for van der Waals interactions is included in DFT calculations. The empirical correction consists of damped pairwise interatomic potentials of the form C_6/r^6.[66] One of these methods is the B97-D, proposed by Grimme,[67,68] using the quite flexible power series expansion of Becke[69] as a basis. The same correction has also been introduced in B3LYP, leading to the B3LYP-D method.[67,70,71] Grimme[72] has recently proposed a new benchmark to test the accuracy of a series of DFT based methods, including the four ones used in the present work: M05-2X, M06-2X, B97-D, and B3LYP-D.

Following the recommendation of Truhlar et al.,[73] the augmented polarized valence double-ζ set 6-31+G(d,p)[74] has been used. The calculations with the M05-2X, M06-2X, B97-D, and B3LYP-D functionals have been performed with the Gaussian 03[75] and Gaussian 09[76] packages. For B3LYP-D calculations, the Grimme dispersion term and gradients have been programmed in an external driver. All minimum energy structures have been characterized through the calculation of the harmonic vibrational frequencies to verify that all frequencies are real. Full Natural Bond Orbital Analysis has been carried out with NBO Version 3.[77] The interaction energy between the fragments has been decomposed using the ADF program[78,79] into contributions associated with orbital and electrostatic terms following a Morokuma-type energy decomposition method.[80] These calculations have been done using an uncontracted polarized triple-ζ basis set of Slater type orbitals (TZP).

In some structures the presence of bond critical points[81] has been analyzed using the XAIM program.[82] To analyze the relative strength of hydrogen bonds in structures with multiple hydrogen bonds, several strategies have been proposed.[83−88] Our analysis will be based on the hydrogen bond lengths, the bond critical point electron density, the natural charge on the H atom and the σ^*(N−H) population.[88]

Results and Discussion

As stated in the Introduction, we carried out complete DFT optimization of different structures of G-C/C-G and G-C′/C′-G Watson−Crick dimers (where C′ states for 5-methylcytosine) using different functionals. We will first discuss the results obtained with the M06-2X functional.[56,58] Initial configurations were constructed in such a way that the planes of the two pairs of bases are parallel to each other at a distance of about 3.2 Å (this parameter, *R*, corresponds to the rise coordinate of Olson et al.[89]). Three different starting points for the optimization

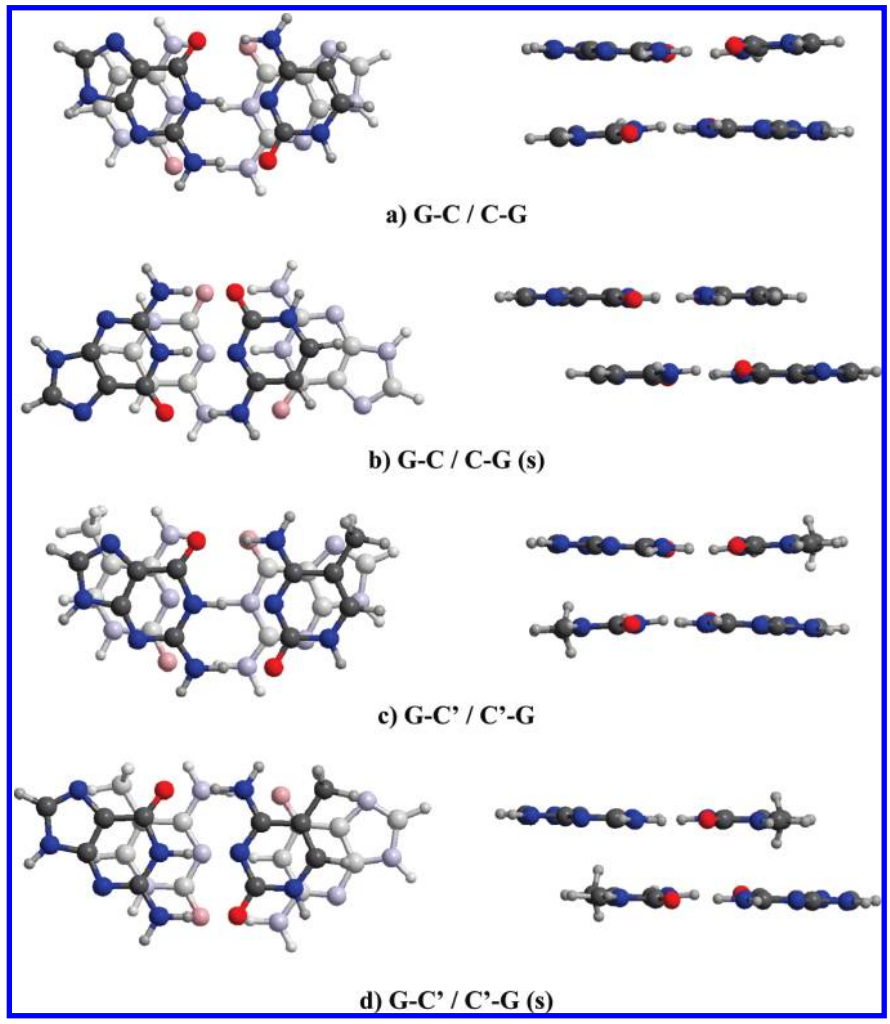

Figure 2. Totally optimized structures of the G-C/C-G and G-C′/C′-G dimers. The letter "s" in parentheses stands for "slid" structures. In each case, a lateral vision (right) and a vision from above (left) are presented.

procedure have been built: two of them correspond to the arrangements found in the DNA structures (where the four sugars are oriented to the same side), while in the third one an inversion center has been imposed (this leads each pair of two sugars to be oriented to a different side). The two arrangements corresponding to the DNA structures differ in the value of the twist coordinate of Olson et al.[89] This coordinate can be defined from the structure shown in Figure 1. The dashed line joining the C8 atom of guanine to the C6 atom of the pyrimidine (cytosine or 5-methylcytosine) is the long base-pair axis and the twist coordinate, which we will call θ, is the rotation of one base pair around the center of its C6−C8 axis. We have considered two initial values for this angle: $\theta_0 = 0°$ and $\theta_0 = 36°$.

We will successively discuss the geometry parameters of optimized structures, their energy and other aspects such as charge transfer and electrostatic potentials.

Geometry Parameters. The structures of all dimers optimized at the M06-2X/6-31+G(d,p) level of calculation are shown in Figures 2 and 3. In all cases, a lateral vision and a vision from above are presented.

The lateral vision shows that the planarity of the base pairs and the parallelism between them are slightly broken. To study the mean distance between the two base pairs, we have defined the *xy* plane by the C_6 atom of the pyrimidine base, and the C_8 and N_1 atoms of guanine. The mean distance is then defined as the difference between the mean *z* values of the atoms of the upper base pair and the one of the atoms of the lower base pair.

For the G-C/C-G structure, the optimization from the two different starting points (0 and 36°) lead to the same structure, in which the twist angle is about 18°. As the optimized base pairs are not strictly planar and parallel, the twist coordinate of the optimized structures is defined as the angle θ between the projections on the *xy* plane of the C_6-C_8 axis of each base pair. This twist angle is in good agreement with the results obtained by Cooper et al.[50] keeping unchanged the geometries of both base pairs and the parallelism between them. When 5-methylcytosine is used instead of cytosine, two different optimized structures are obtained. Both of them show again a twist angle of about 18°, but in the one starting from $\theta_0 = 36°$, one base pair is slid around 3 Å with respect to the other one (see Figure 2). When this slid structure is reoptimized after removing the methyl groups, a slid structure for G-C/C-G is also obtained. It is worth mentioning that the slid structures keep better the planarity of the base pairs and that one hydrogen atom of the methyl group of the G-C′/C′-G structure points to the N atom of the π five-member ring of guanine.

The values of the rise (R) and twist (θ) parameters, along with the hydrogen bond lengths for the pairs of bases, the values of electron density at the X···H bond critical points, the natural charge on the H atom and the σ^*(N−H) population are presented

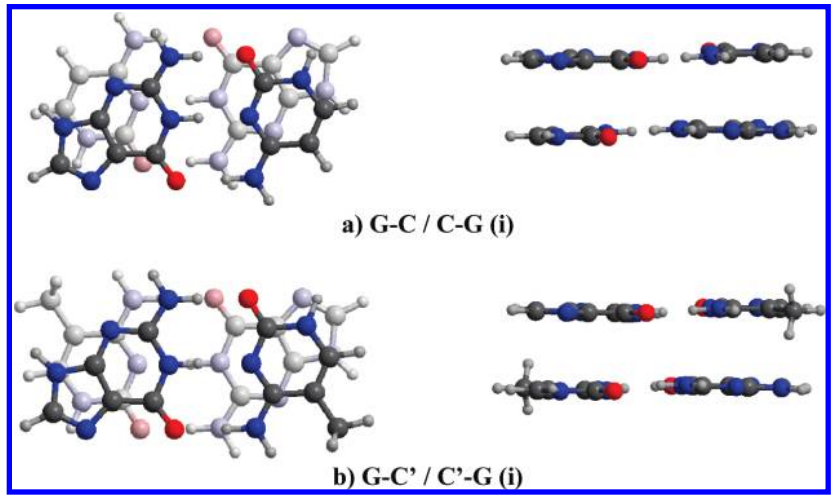

Figure 3. Totally optimized structures of the G-C/C-G (i) and G-C′/C′-G (i) dimers. In both structures, an inversion center has been imposed. In each case, a lateral vision (right) and a vision from above (left) are presented.

TABLE 1: Hydrogen Bond Lengths, Electron Density at the X···H Bond Critical Points ($\rho_{bcpX\cdots H}$), Natural Charge on the H Atom (q_H), σ^*_{N-H} Population, and Twist (θ) and Rise (R) Parameters[a]

structures	G···C/C′	distance[b]	d(X···H)[c]	$\rho_{bcpX\cdots H}$	q_H	σ^*_{N-H}	θ	R
	$O_6\cdots N_4$	2.80 (2.90)	1.769	0.0362	0.4865	0.0541		
G-C	$N_1\cdots N_3$	2.94 (2.92)	1.909	0.0321	0.4775	0.0616		
	$N_2\cdots O_2$	2.92 (2.84)	1.904	0.0268	0.4718	0.0343		
	$O_6\cdots N_4$	2.87	1.844	0.0303	0.4823	0.0459		
G-C/C-G	$N_1\cdots N_3$	2.93	1.895	0.0333	0.4799	0.0674	18.41	3.00
	$N_2\cdots O_2$	2.88	1.863	0.0295	0.4739	0.0404		
	$O_6\cdots N_4$	2.87	1.847	0.0303	0.4756	0.0450		
G-C/C-G (s)	$N_1\cdots N_3$	2.91	1.876	0.0348	0.4822	0.0715	18.24	3.02
	$N_2\cdots O_2$	2.84	1.817	0.0325	0.4807	0.0442		
	$O_6\cdots N_4$	2.89	1.863	0.0293	0.4830	0.0461		
G-C/C-G (i)	$N_1\cdots N_3$	2.93	1.899	0.0330	0.4785	0.0668	0	2.77
	$N_2\cdots O_2$	2.83	1.803	0.0336	0.4772	0.0460		
	$O_6\cdots N_4$	2.80	1.766	0.0366	0.4868	0.0549		
G-C′	$N_1\cdots N_3$	2.94	1.905	0.0325	0.4776	0.0625		
	$N_2\cdots O_2$	2.91	1.891	0.0276	0.4728	0.0356		
	$O_6\cdots N_4$	2.86	1.840	0.0307	0.4826	0.0467		
G-C′/C′ -G	$N_1\cdots N_3$	2.93	1.894	0.0334	0.4797	0.0680	16.32	3.02
	$N_2\cdots O_2$	2.87	1.851	0.0303	0.4736	0.0418		
	$O_6\cdots N_4$	2.87	1.846	0.0303	0.4763	0.0446		
G-C′/C′-G (s)	$N_1\cdots N_3$	2.91	1.873	0.0350	0.4831	0.0726	18.11	3.03
	$N_2\cdots O_2$	2.83	1.809	0.0331	0.4821	0.0452		
	$O_6\cdots N_4$	2.87	1.837	0.0309	0.4837	0.0477		
G-C′/C′-G (i)	$N_1\cdots N_3$	2.93	1.892	0.0335	0.4796	0.0681	0	2.79
	$N_2\cdots O_2$	2.83	1.814	0.0327	0.4799	0.0443		

[a] Distances are in Å and angles in degrees. [b] In parentheses, experimental values (ref 12). [c] X being O_6 of guanine, N_3 and O_2 of cytosine, respectively.

in Table 1. The values for the G-C and G-C′ base pairs are also included for comparison.

The rise parameter has a value of about 3 Å for the four dimer structures having a twist angle close to 18°. In the two structures in which an inversion center has been kept, the twist angle is 0° and the rise parameter is about 2.8 Å.

Regarding to the lengths of the hydrogen bonds, several interesting trends can be observed from the values shown in Table 1. For the G-C pair of bases (without stacking) the shortest hydrogen bond is the one between O_6 and N_4 while the experimental results[12] show that the shortest hydrogen bond is the other lateral one (the one between N_2 of guanine and O_2 of cytosine). Our result agrees with other density functional calculations[16] and with the benchmark results of Sponer et al.[90] The relative length of the two lateral hydrogen bonds is well correlated with the electron density at the X···H bond critical point, the positive natural charge on the H atom and the population of the antibonding N−H orbital. The introduction of a methyl group in cytosine (G-C′ pair of bases) leads to a slight decrease of the X···H distances, this decrease being more important for the hydrogen bond between the N atom of guanine and the O atom of the pyrimidine.

If one looks at the results corresponding to the dimers (G-C/C-G and G-C′/C′-G), one can observe the important influence of stacking on hydrogen bonds. For the central $N_1-H\cdots N_3$ hydrogen bonds, the stacking produces a quite slight increase of their strength. The effect of stacking is more important for the two lateral hydrogen bonds, in such a way that the strength of the $O_6\cdots H-N_4$ hydrogen bond decreases while the strength of the $N_2-H\cdots O_2$ one increases. This opposite behavior leads in some cases (the slid dimers and the ones with an inversion center) to the inversion of the lengths of the two lateral hydrogen bonds. The variation in the lengths of the three hydrogen bonds due to the stacking corresponds to a sort of rotation of one base

TABLE 2: Total Energy (in Hartree), Formation Energy and Interaction Energies (in kcal mol^{-1}) for the Different Base Pairs and Dimers

structures	E_{tot}	ΔE_f	E_{int}	E_{HB}	E_S	E_{MB}
G-C	−937.211833	−28.44	−31.59	−31.59		
G-C/C-G	−1874.461767	−23.91	−89.02	−62.60	−26.16	−0.26
G-C/C-G (s)	−1874.461788	−23.92	−89.05	−63.01	−25.20	−0.84
G-C/C-G (i)	−1874.456892	−20.85	−84.79	−62.08	−22.38	−0.33
G-C′	−976.5118458	−28.98	−32.26	−32.26		
G-C′/C′-G	−1953.064724	−25.75	−92.04	−63.84	−28.51	0.31
G-C′/C′-G (s)	−1953.069687	−28.86	−95.33	−64.37	−30.51	−0.44
G-C′/C′-G (i)	−1953.057901	−21.47	−86.74	−63.59	−23.24	0.09

with respect to the other, a movement that we will analyze in the next section. The change of the hydrogen bond lengths is also confirmed through the data of ρ and of the natural population analysis, since one can observe that the charge of the hydrogen atom is inverted in the slid dimers but not in the ones with an inversion center. The behavior of the σ^*(N−H) population presents more complex trends. This inversion of the lengths of the two lateral hydrogen bonds suggests that stacking can be a factor to be considered in the discrepancy between theoretical and experimental results in the G-C base pair.

Energies. The energy values obtained for the optimized structures are presented in Table 2. In this table, the first column corresponds to the total energy of the systems in atomic units. The second column presents the energy of formation, ΔE_f, of each structure from the separated units (the two optimized bases for G-C and G-C′, and the optimized pairs of bases for the dimers). The next column contains the interaction energy, E_{int}, calculated by subtracting from the total energy the energy of the different base units (2 for base pairs and 4 for the dimers) with the geometry they have in the complex. In the three remaining columns an analysis of the interaction energy is done by dividing it into different two-body and many-body interactions, all of them calculated with the geometries in the total system. E_{HB} is the two-body hydrogen bonding interaction in each base pair of the system, calculated as the difference between the energy of the base pair and the sum of the energies of the two bases. E_S contains the other four two-body interactions between the four bases: two intrastrand and two interstrand stacking interactions, the interstrand terms being usually repulsive. Finally, E_{MB} stands for the multibody interaction energy which has been calculated by subtracting the two-body terms (E_{HB} and E_S) from the total interaction energy, E_{int}.

In Table 2 one can observe that the energies of formation of G-C and G-C′ are very similar, that of G-C′ being 0.5 kcal mol^{-1} more stabilizing than the one for G-C. This is in good agreement with the strengthening of the hydrogen bonds shown in the analysis of geometries. BSSE is quite small, its values being 1.01 and 1.03 kcal mol^{-1}, respectively. For G-C, the energy of formation (−27.43 kcal mol^{-1}) is in good agreement with the benchmark value of −28.2 kcal mol^{-1} obtained by Hobza et al.[91] One can also observe that the interaction energies are, in both cases, more negative than the energies of formation, as it was to be expected given that the deformation energy is missing in E_{int}.

Regarding the G-C/C-G and G-C′/C′-G dimers, the two first columns show that the structures in which an inversion center has been imposed are the less stable ones. For the other two structures, both are almost degenerate in the case of G-C/C-G while the slid structure is more stable by 3.11 kcal mol^{-1} in the case of G-C′/C′-G. It is also worthwhile mentioning that the G-C′/C′-G structures are always more stable than the corresponding G-C/C-G ones. In all cases, the energies of formation of the dimers (a measure of the stacking interactions) are quite similar to the energies of formation of the base pairs (a measure of hydrogen bond interactions). In contrast with the base pairs, the interaction energy of the dimers is very different from the formation energy, since, as stated before, the interaction energy is referred to the sum of the energies of the four bases. However, both energies follow the same trends.

Let us now discuss the partition of the interaction energy in two-body and many-body terms. For the pairs of bases there is only one term, the one corresponding to the two-body hydrogen bonding interaction, E_{HB}. For the dimers, the E_{HB} values correspond to the addition of the hydrogen bond interactions in the upper and lower pairs of bases. So, E_{HB} values for dimers are about twice the corresponding term for the base pairs. The slid dimers, which are the most stable ones, have also the biggest E_{HB} interaction values.

Regarding the E_S interaction terms, they are more negative than the energy of formation, as they do not contain the energy for the geometry distortion. As mentioned before, the interstrand terms are usually repulsive, but they are attractive for the structures with an inversion center. This peculiar behavior can be explained through the electrostatic interactions between both cytosines and both guanines due to their high and opposite dipolar moments. This fact suggests that electrostatic terms also play an important role in the stacking of guanine−cytosine base pairs. Finally, the many-body interaction terms are small (less than 1 kcal mol^{-1}) when compared with the two-body ones and can be cooperative ($E_{MB} < 0$) or anticooperative ($E_{MB} > 0$). Thus, the total interaction is mainly due to two-body interactions.

It is worth discussing the effect of replacing cytosine by 5-methylcytosine. As mentioned before, Table 2 shows that the G-C′/C′-G dimers have greater values of the formation energy and of the total interaction energy, the E_S term, which is due to the stacking, being the most important one for understanding the difference in E_{int}. The main reason for the different behavior between 5-methylcytosine and cytosine is the increase of polarizability associated with the presence of a methyl group in 5-methylcytosine,[37] this leading to an increase of the dispersion energy and, consequently, to enhanced stacking interactions. A second reason may be the interaction of methyl with the π system of the adjacent basis as discussed by Cooper et al.[50] in the analysis of the difference of behavior between A-T/T-A and A-U/U-A, suggesting that a weak hydrogen bond is formed between a hydrogen atom of the methyl group and a nitrogen atom of adenine. This argument is also valid for the dimers of the G-C pair of bases, since, as mentioned before, one hydrogen atom of the methyl group of G-C′/C′-G points to the N atom of the π five-membered ring of guanine. As a matter of fact, we have found a bond critical point between these two atoms, its electron density being 0.0082, quite smaller than the densities shown in Table 1, thus indicating that it corresponds to a weak hydrogen bond.

Let us now make an energy analysis of the three movements mentioned in the geometry section: the one characterized by

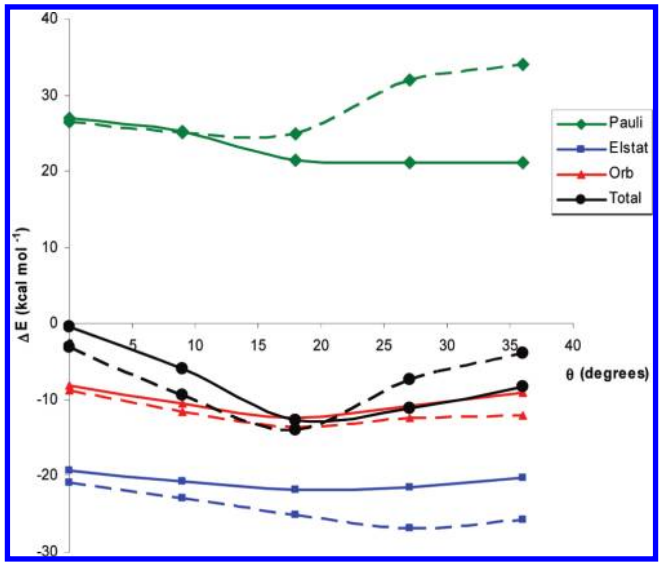

Figure 4. Variation of the interaction energy and of its decomposition contributions for the G-C/C-G (solid lines) and G-C′/C′-G (dashed lines) systems: with the twist angle (from 0° to 360°).

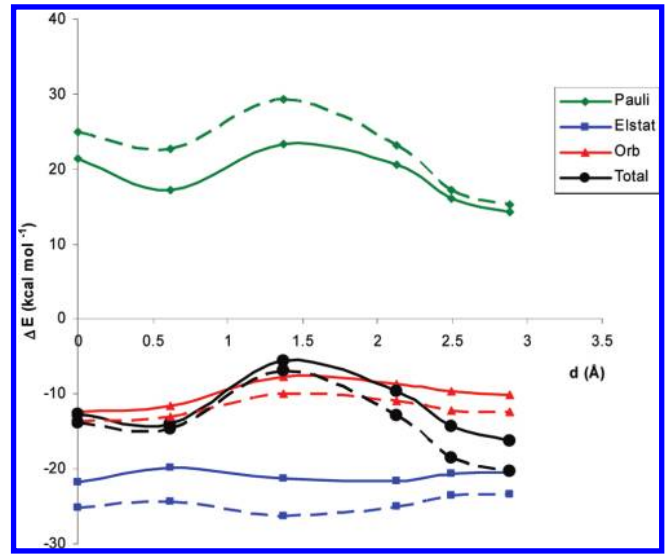

Figure 5. Variation of the interaction energy and of its decomposition contributions for the G-C/C-G (solid lines) and G-C′/C′-G (dashed lines) systems with the slide distance (from 0 to 3 Å).

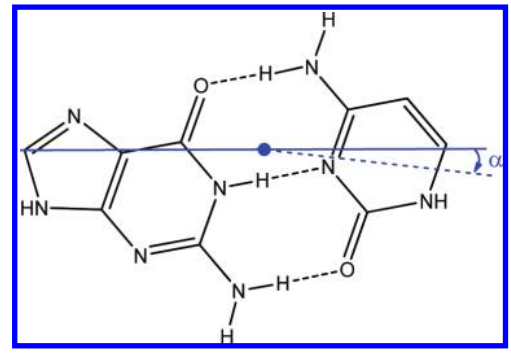

Figure 6. Schematic representation of the third movement for the G-C/C-G dimer. See text for details.

the twist angle, the slide of one base pair with respect to the other and the pseudorotation in each base pair due to the variation of the three hydrogen bond lengths. As mentioned in the Computational Methods section, this analysis has been carried out with the ADF program,[78,79] through a decomposition of the interaction energy between the rigid fragments into three contributions:

$$\Delta E_{\text{int}} = \Delta E_{\text{elst}} + \Delta E_{\text{Pauli}} + \Delta E_{\text{orb}} \quad (1)$$

where the electrostatic term ΔE_{elst} corresponds to the classical electrostatic interaction between the unperturbed charge distributions of the rigid fragments, ΔE_{Pauli} comprises the destabilizing interactions between occupied orbitals and is responsible for any steric repulsion, and the orbital interaction contribution, ΔE_{orb}, accounts for charge transfer, polarization and dispersion terms.

To study the twist movement, we have placed the two pairs of bases in parallel planes at a distance of 3 Å and we have changed the twist angle between 0 and 36°. Figure 4 shows the variation of the interaction energy and of its decomposition contributions for the G-C/C-G (solid lines) and G-C′/C′-G (dashed lines) systems. If one looks at the total interaction energies (black lines), one can observe that both systems present a minimum, the one for the methylated system being slightly displaced to smaller values of the twist angle, in good agreement with the values shown in Table 1. The existence of these minima is due to the behavior of the three terms of the energy decomposition. For the G-C/C-G system, the Pauli contribution (green solid line) decreases as far as the twist angle increases, but the other two energy terms show a minimum at an intermediate angle. In the methylated system, the Pauli term (green dashed line) shows already the presence of a minimum, while its value notably increases when the twist angle approaches 36°. This different behavior of the Pauli term is due to the repulsion when the methyl group approaches the other base pair. The curves for the electrostatic and orbital interaction terms run below for the G-C′/C′-G system, since it has a greater dipolar moment and a greater polarizability.

The slide movement has been analyzed by placing the two bases in parallel planes at a distance of 3 Å, with a twist angle of 18° (which is approximately the value it has in the minima) and sliding one base pair with respect to the other. Figure 5 shows the variation of the interaction energy and of its decomposition contributions for the G-C/C-G (solid lines) and G-C′/C′-G (dashed lines) systems with the slide distance (from 0 to 3 Å). The curves corresponding to the total interaction energy (in black) show the presence of two minima, one for a small value of the slide distance and the other for a slide of about 3 Å. Both minima are, of course, separated by an energy barrier at a slide of about 1.5 Å. The existence of these two minima agrees with the two optimized structures shown in Figure 2. The fact that the minimum of the slid structure of the G-C′/C′-G dimer is in a deeper well with respect to the energy barrier is mainly due to the diminution of the Pauli repulsion term through sliding.

The third movement is more difficult to explain and is schematically presented in Figure 6 for the G-C/C-G dimer. In this figure only one base pair is shown, the other being below, at a distance of 3 Å and with a value of 18° of the twist angle. Then the bases on the right (cytosine in the upper base pair and guanine in the lower base pair) are rotated an angle α respect to the axis perpendicular to the middle point of the C_6-C_8 axis. This partial rotation leads to a decrease of the hydrogen bond lengths $N_2-H\cdots O_2$ and $N_1-H\cdots N_3$ and to an increase of the $O_6\cdots H-N_4$ one. It also leads to a slight loss of linearity of the hydrogen bonds, which increases as far as α is increased. For this reason, only small values of α (up to 1°) have been used. The energy partition analysis of this third movement for the

TABLE 3: Energy Partition Analysis in the Base Pair G-C and the Dimer G-C/C-G for Two Values of the Angle α Defined in Figure 6[a]

	G-C			G-C/C-G		
	α = 0°	α = 1°	ΔΔ*E*	α = 0°	α = 1°	ΔΔ*E*
ΔE_{Pauli}	27.94	29.08	1.14	57.74	60.17	2.43
ΔE_{elstat}	−42.46	−43.20	−0.74	−81.46	−83.12	−1.66
ΔE_{orb}	−17.28	−17.67	−0.39	−33.34	−34.80	−1.46
ΔE_{int}	−31.80	−31.79	0.01	−57.06	−57.75	−0.69

[a] Energies in kcal mol^{-1}.

TABLE 4: Charges on Each Base for the Optimized Structures

structures	Q_G	$Q_{C/C'}$
G-C	−0.0306	0.0306
G-C/C-G	−0.0383	0.0383
G-C/C-G (s)	−0.0551	0.0551
G-C/C-G (i)	−0.0437	0.0437
G-C′	−0.0322	0.0322
G-C′/C′ -G	−0.0378	0.0378
G-C′/C′-G (s)	−0.0548	0.0548
G-C′/C′-G (i)	−0.0367	0.0367

G-C base pair and for the G-C/C-G dimer is presented in Table 3. One can observe that while the rotation barely increases the energy of the base pair, it stabilizes noticeably the dimer by 0.69 kcal mol^{-1}. The stabilization of the dimer is mainly due to the orbital interaction term that is clearly more stabilizing than in the base pair (one would expect the values for the dimer to be about twice the values of the base pair, which is what happens for the Pauli and electrostatic contributions, but the orbital interaction term of the dimer is about four times that of the G-C base pair). This can be confirmed through the values of the σ^*(N−H) population in G-C and G-C/C-G (see Table 1). We will come back to this point in the next section when discussing the values of the charge transfer.

Charge Transfer and Electrostatic Potentials. We have carried out the natural population analysis[77] in all the studied structures. This has permitted us to calculate the charge transfer between the different bases. Table 4 presents the charges on each base for the Watson−Crick base pairs and for the corresponding dimers. In the base pairs G-C and G-C′, one can observe that charge transfer is always produced from the pyrimidinic base to guanine. This can be explained in the following way. On one hand, it is well-known that the ionization potential of guanine is lower than those of the pyrimidinic bases,[92] while the electron affinity is greater for the pyrimidinic bases,[93] in such a way that π electron transfer is expected to occur from guanine to cytosine and 5-methylcytosine. On the other hand, there is also a σ electron transfer through hydrogen bonding, from the lone pair of the acceptor atom to an antibonding orbital of the N−H bond, this charge transfer being more important as far as the hydrogen bond is stronger. Given that the hydrogen bond $O_6\cdots H-N_4$ contributes to a σ charge transfer from guanine to cytosine but $N_2-H\cdots O_2$ and $N_1-H\cdots N_3$ lead to a σ charge transfer on the opposite sense, one can expect that the global σ charge transfer is produced from the pyrimidinic base to guanine. So, the sense of the charge transfer will depend on the relative values of the σ and π contributions. The values of the charges in Table 4 show that the σ charge transfer is the predominant contribution.

For the dimers, Table 4 shows that the charge transfer from the pyrimidinic base to guanine increases. This can be explained through the above-mentioned strengthening of the hydrogen bonds $N_2-H\cdots O_2$ and $N_1-H\cdots N_3$ (which contribute to the σ charge transfer from cytosine to guanine) and to the corresponding weakening of the hydrogen bond $O_6\cdots H-N_4$ (which contributes in the opposite sense). The increase of charge transfer is especially important for the slid structures given that the change in hydrogen bonds is greater in this case (see Table 1). Finally, the fact that the charge transfer shown by the structures with an inversion center is smaller than that corresponding to the slid structures can be explained through the decrease of the distance R between the planes of both dimers, thus implying an increase of the π contribution from guanine to cytosine.

The interplay between stacking and hydrogen bonding can be interpreted following the strategy proposed by Geerlings et al.,[3,4] in which the molecular electrostatic potential (MEP) is evaluated in the vicinity of the hydrogen bond. Geerlings states that the hydrogen bonding capacity can be evaluated as the minimum of the MEP in the direction of the hydrogen bond. In fact, he suggests that this minimum is at 1.25 Å of the hydrogen acceptor atom. Table 5 presents the electrostatic potentials at a distance of 1.25 Å from the hydrogen acceptor atom O_6 of guanine and N_3 and O_2 atoms of cytosine or 5-methylcytosine. In each case, two situations have been considered: the corresponding base taken alone or stacked with the other base pair, the geometry being also the one found in the dimer. As one would expect, the electrostatic potentials of the stacked N_3 and O_2 atoms of cytosine or 5-methylcytosine increases in all cases with respect to the ones of the isolated basis. Nevertheless, the electrostatic potential of the stacked O_6 atom of guanine does not always decrease as was to be expected, although this behavior is compensated by the greater increase of the electrostatic potentials in the other two atoms. This analysis of the electrostatic potentials suggests, in good agreement with the results shown in the preceding section, that the electrostatic contribution is not the most important term to explain the effect of stacking on hydrogen bonding.

Comparison with Other Functionals. We will now present the results obtained with the other functionals mentioned in the Computational Methods Section: M05-2X, B97-D, and B3LYP-D.

All geometrical parameters of the optimized structures are available as Supporting Information. In all cases, we have found the six minima of the dimers obtained with the M06-2X functional. However, the G-C′/C′-G (i) structure obtained with the B97-D functional is slid with respect to the ones calculated with the other functionals. This slid structure is shown in Figure 7a. Starting from this structure and removing the two methyl

TABLE 5: Values of the MEP (in au) at 1.25 Å from the H Acceptor Atoms of Guanine and Cytosine or 5-Methylcytosine

	GC/CG	GC/CG(s)	GC/CG(i)	GC′/C′G	GC′/C′G(s)	GC′/C′G(i)
O_6 (G) alone	−0.0857	−0.0861	−0.0849	−0.0858	−0.0866	−0.0851
O_6 (G) stacked	−0.0976	−0.0794	−0.0926	−0.0981	−0.0794	−0.0924
N_3 (C/C′) alone	−0.1015	−0.1012	−0.1012	−0.1034	−0.1029	−0.1032
N_3 (C/C′) stacked	−0.1152	−0.1077	−0.1143	−0.1172	−0.1103	−0.1163
O_2 (C/C′) alone	−0.1231	−0.1232	−0.1228	−0.1264	−0.1265	−0.1257
O_2 (C/C′) stacked	−0.1333	−0.1292	−0.1366	−0.1364	−0.1334	−0.1404

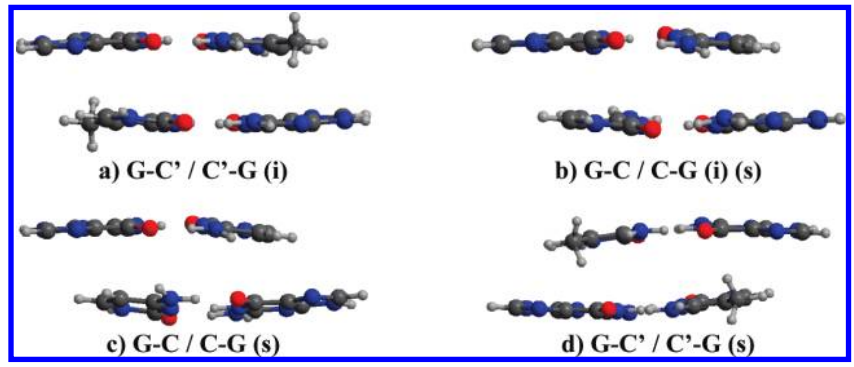

Figure 7. Lateral vision of totally optimized structures of some G-C/C-G and G-C′/C′-G dimers with the B97-D functional.

TABLE 6: Lengths (in Å) of the Three Hydrogen Bonds of the Pair Bases and of the Dimers Calculated with the Four Functionals

structures	G···C/C′	M06-2X	M05-2X	B97-D[a]	B3LYP-D
	$O_6\cdots N_4$	2.80	2.82	2.76	2.75
G-C	$N_1\cdots N_3$	2.94	2.96	2.90	2.89
	$N_2\cdots O_2$	2.92	2.94	2.91	2.88
	$O_6\cdots N_4$	2.87	2.90	2.81	2.80
G-C/C-G	$N_1\cdots N_3$	2.93	2.96	2.89	2.88
	$N_2\cdots O_2$	2.88	2.90	2.86	2.84
	$O_6\cdots N_4$	2.87	2.90	2.83	2.82
G-C/C-G (s)	$N_1\cdots N_3$	2.91	2.95	2.88	2.87
	$N_2\cdots O_2$	2.84	2.86	2.82	2.81
	$O_6\cdots N_4$	2.89	2.89	2.81 (2.82)	2.80
G-C/C-G (i)	$N_1\cdots N_3$	2.93	2.96	2.88 (2.86)	2.87
	$N_2\cdots O_2$	2.83	2.87	2.83 (2.82)	2.80
	$O_6\cdots N_4$	2.80	2.83	2.76	2.75
G-C′	$N_1\cdots N_3$	2.94	2.97	2.91	2.90
	$N_2\cdots O_2$	2.91	2.93	2.90	2.88
	$O_6\cdots N_4$	2.86	2.89	2.81	2.80
G-C′/C′ -G	$N_1\cdots N_3$	2.93	2.96	2.89	2.87
	$N_2\cdots O_2$	2.87	2.89	2.84	2.82
	$O_6\cdots N_4$	2.87	2.90	2.82	2.81
G-C′/C′-G (s)	$N_1\cdots N_3$	2.91	2.94	2.88	2.87
	$N_2\cdots O_2$	2.83	2.86	2.83	2.81
	$O_6\cdots N_4$	2.87	2.88	2.81	2.80
G-C′/C′-G (i)	$N_1\cdots N_3$	2.93	2.96	2.87	2.86
	$N_2\cdots O_2$	2.83	2.86	2.80	2.79

[a] In parentheses, values for a slid structure.

groups, the optimization leads to a second minimum for the G-C/C-G (i) structure, which is shown in Figure 7b. One can observe that the slid structure of G-C/C-G (i) presents a loss of planarity of the two base pairs. The loss of planarity is more important in the B97-D slid structures for G-C/C-G and G-C′/C′-G, as can be observed in Figure 7c,d. A similar loss of planarity is also found for the slid structures obtained with the B3LYP-D functional.

Table 6 presents the lengths of the three hydrogen bonds for the two base pairs (G-C and G-C′) and for the six dimers obtained with the four functionals we have used. As already mentioned for the M06-2X functional, the $O_6\cdots H-N_4$ hydrogen bond is, in all cases, the shortest one in the base pairs, while stacking tends to enlarge this bond length and to decrease the $N_2-H\cdots O_2$ hydrogen bond, in such a way that both lengths approach each other and, in some cases, an inversion of the strength of the two lateral hydrogen bonds is produced. The comparison between the corresponding methylated and nonmethylated structures confirms that the introduction of a methyl group in cytosine tends to slightly decrease the length of the $N_2-H\cdots O_2$ hydrogen bond. We can conclude that these two trends, already mentioned for M06-2X, are not an artifact of this functional and seem to correspond to a real behavior.

TABLE 7: Energies of Formation (in kcal mol^{-1}) Obtained with the Different Functionals

structures	M06-2X	M05-2X	B97-D	B3LYP-D
G-C	−28.44	−28.46	−28.68	−31.63
G-C/C-G	−23.91	−17.06	−20.93	−22.87
G-C/C-G (s)	−23.92	−18.71	−22.37	−23.90
G-C/C-G (i)	−20.85	−13.45	−17.50 (−25.15)	−19.58
G-C′	−28.98	−28.95	−29.15	−32.17
G-C′/C′-G	−25.75	−18.49	−23.23	−25.06
G-C′/C′-G (s)	−28.86	−22.49	−27.11	−29.21
G-C′/C′-G (i)	−21.47	−14.21	−26.31	−25.80

[a] In parentheses, value for a slid structure.

Table 7 presents the energies of formation of the base pairs and of the dimers. One can observe that all functionals show that G-C′ is about 0.5 kcal mol^{-1} more stable than G-C. It is also worthwhile mentioning that the energies of formation calculated with the B3LYP-D functional are noticeably more negative than the ones obtained with the other three functionals and the benchmark value of Hobza et al.[91] Regarding the values obtained for the dimers, one can observe that the M05-2X functional clearly underestimates the stacking, while the other three functionals show quite similar results. Table 7 also confirms that methylation provokes a noticeable increase of stacking. Finally, sliding stabilizes the dimers, this stabilization being again quite more important in the methylated dimers. This suggests that the methylated structures present a local distortion, as has been proposed by some authors.[38−42]

Sherrill et al.[64] argued that M05-2X and M06-2X are successful in the study of noncovalent interactions since both methods reproduce quite well medium range interactions. They affirm, however, that they have problems in the study of large range interactions, as, for instance, hydrogen-bonded and interstrand base-pair stacking, and that these interactions are better described by the addition of Grimme's empirical dispersion terms. To discuss these arguments, we have calculated the interstrand interactions with M06-2X and B97-D functionals. With the first functional, the interstrand interaction for G-C/C-G, G-C/C-G (s), G-C′/C′-G, and G-C′/C′-G (s) are always repulsive (6.54, 5.36, 6.98, and 5.99 kcal mol^{-1}, respectively), while when one uses the B97-D functional, the corresponding values are less repulsive (5.01, 1.95, 4.61, and 4.64 kcal mol^{-1}, respectively). So, the effect of the introduction of the term in R^{-6} is small.

To compare the M06-2X and the B3LYP-D functionals, we have carried out the decomposition of the interaction energy between the two pairs of bases in G-C/C-G. The two pairs of bases have been placed in two parallel planes, at a distance of 3 Å and with a value of 18° for the twist angle. The results are presented in Table 8. One can observe that the electrostatic and

TABLE 8: Decomposition of the Interaction Energy between the Two Pairs of Bases in G-C/C-G[a]

	M06-2X	B3LYP-D	$\Delta\Delta E$
ΔE_{Pauli}	21.45	52.46	31.01
ΔE_{elstat}	−21.77	−23.14	−1.37
ΔE_{orb}	−12.33	−10.15	2.18
ΔE_{disp}		−34.04	−34.04
ΔE_{int}	−12.65	−14.86	−2.21

[a] Energies in kcal mol^{-1}.

TABLE 9: Influence of the Counterpoise (CP) Correction in the Decomposition of the Total M06-2X Interaction Energy for the G-C/C-G Dimer[a]

	no CP	CP
ΔE_{tot}	−89.02	−82.20
$\Sigma\Delta^2 E_{HB}$	−62.60	−60.44
$\Sigma\Delta^2 E_S$	−26.16	−21.67
$\Sigma\Delta^2 E$	−88.76	−82.11
$\Sigma\Delta^3 E$	0.20	0.56
$\Sigma\Delta^4 E$	−0.46	−0.65
$\Sigma\Delta^{3,4} E$	−0.26	−0.09

[a] Energies in kcal mol^{-1}.

orbital terms are quite similar with both functionals. On the contrary, the B3LYP-D method overestimates the Pauli repulsion term, which is 31.01 kcal mol^{-1} greater than the one calculated by M06-2X. This overestimation is nearly compensated by the R^{-6} dispersion term, in such a way that the total interaction energies are very similar in both cases.

All the results presented above with the different functionals have been done without taking into account the BSSE corrections. Some authors[70] argue that BSSE corrections are already included in the parametrization of DFT-D methods. However, we decided to check the influence of these corrections in interaction energies and geometries of the G-C/C-G dimer calculated with the M06-2X functional.

Table 9 presents the influence of the counterpoise (CP) correction in the decomposition of the total M06-2X interaction energy for the G-C/C-G dimer. According to the methodology used by Hobza et al.,[94] the three- and four-body terms have been directly calculated. The CP correction for the pair interaction corresponding to hydrogen bonds is 2.16 kcal mol^{-1} in good agreement with the correction of about 1 kcal mol^{-1} for the hydrogen bond interactions in one base pair. The CP correction for the pair terms corresponding to the four stacking interactions in the dimer (two intrastrand and two interstrand) is, however, 4.49 kcal mol^{-1}, thus showing that BSSE affects more stacking than hydrogen bonding interactions.

Regarding the many-body terms, one can observe that three-body terms are repulsive while four-body terms are attractive, in good agreement with the results obtained by Hobza et al.[94] with all the methods they used. One can also observe that the CP correction makes the three-body term 0.36 kcal mol^{-1} more repulsive while the four-body term becomes 0.19 kcal mol^{-1} more attractive. Since both terms tend to compensate each other, the influence of the CP correction in the whole many-body terms is of only 0.17 kcal mol^{-1}, in such a way that these nonadditive terms change from −0.26 to −0.09 kcal mol^{-1}. It is worth mentioning that while dispersion energy is obviously additive in empirically corrected DFT schemes (such as B97-D and B3LYP-D), there is some nonadditivity in functionals like M06-2X, where there is not a separate term for dispersion energy.

The next step has been to analyze the influence of BSSE in the geometries of the pair of bases and of the dimers. We have found that the geometries barely change when CP correction is introduced. For instance, the lengths of the three hydrogen bonds in the base pair and in the dimers have been increased by about 0.01−0.02 Å, in such a way that the differences between the lengths in base pairs and dimers do not change. For the dimers, the rise parameter, R, increases only by about 0.05 Å and the twist angle, θ, opens by only about 0.2°.

Conclusions

To study the mutual relationship of stacking and hydrogen-bonding and the possible influence of stacking in the different behavior of cytosine (C) and 5-methylcytosine (C′) in DNA, we carried out complete DFT optimization of different structures of G-C and G-C′ dimers (i.e., G-C/C-G and G-C′/C′-G), using different functionals.

All results show that stacking leads to an increase of the $O_6\cdots H-N_4$ hydrogen bond length and to a simultaneous decrease of the $N_2-H\cdots O_2$ one, in such a way that both lengths approach each other and, in some cases, an inversion occurs. These results suggest that stacking can be a factor to explain the disparity between theory and experiment on the relative strength of the two lateral hydrogen bonds.

The geometries of the dimers may be analyzed in terms of three different movements. The first one corresponds to the value of the twist angle (see Figure 1), which is about 18° in all cases. It has to be emphasized, however, that we have not taken into account the lateral chains present in DNA. The second movement is the slide of one base pair with respect to the other one. The third movement is a sort of rotation of one base respect to the other one and it is related to the change of the hydrogen bond lengths due to stacking. The three movements have been analyzed through a decomposition of the interaction energy in several terms (Pauli repulsion, electrostatic, and orbitalic). The relative importance of these terms in each movement has been discussed.

A second aim of our work was to study the different behavior of cytosine and 5-methylcytosine in DNA. On one side the base pair G-C′ is about 0.5 kcal mol^{-1} more stable than the nonmethylated G-C. More important is the fact that methylation noticeably enhances the stacking interactions, mainly due to the increase of polarizability associated with the presence of a methyl group in 5-methylcytosine,[37] this leading to a higher dispersion energy and, consequently, to stronger stacking interactions. Methylation also favors the existence of slid structures which can produce local distortions in DNA.

This work complements our previous study[5] of the dimers coming from the adenine−thymine (A-T) and adenine−uracil (A-U) base pairs. In spite of the similarity of those systems with the ones treated in this paper, two differences deserve to be mentioned. On one side, the dimers A-T/T-A and A-U/U-A present a second minimum for a twist angle of about 39°. This second minimum was especially stable in A-T/T-A as a result of the existence of a CH/π interaction between the methyl group of thymine and the π system of adenine. The second difference is directly related to the mutual relationship between stacking and hydrogen bonding. In all cases, the stacking interaction leads to a decrease of the hydrogen bond interaction, but while in A-T/T-A and A-U/U-A, the decrease is greater, the greater the stacking interaction, the opposite occurs in G-C/C-G and G-C′/C′-G: the decrease of hydrogen bond interaction is smaller, the higher the stacking interaction (see Table 2). This peculiar behavior is probably due to the fact that the base pairs studied in this paper have three hydrogen bonds, while for A-T and A-U, there are only two "normal" hydrogen bonds since the

third one (C−H···O) is very weak. In the latter systems, both hydrogen bond interactions can be simultaneously maximized in such a way that the stronger the distortion due to stacking, the weaker the hydrogen bond interactions. On the contrary, the existence of three hydrogen bonds in G-C and G-C′ prevents the simultaneous maximization of all of them due to the rigidity of the two bases and a compromise has to be reached. Different compromises are possible and so the higher distortion produced by an increase of the stacking does not necessarily imply a higher decrease of the hydrogen bond interaction.

Acknowledgment. Financial support from Ministerio de Ciencia e Innovación (through the grants CTQ2007-61704 and CTQ2008-06381), Generalitat de Catalunya (through the grants SGR2009-733S and XRQTC) and allowance of computer resources from the CESCA supercomputing center are gratefully acknowledged.

Supporting Information Available: Geometrical parameters of all the optimized structures. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Manalo, M. N.; Pérez, L. M.; LiWang, A. *J. Am. Chem. Soc.* **2007**, *129*, 11298–11299.
(2) Moody, E. M.; Bevilacqua, P. C. *J. Am. Chem. Soc.* **2003**, *125*, 16285–16293.
(3) Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P. *Nucleic Acids Res.* **2005**, *33*, 1779–1789.
(4) Vanommeslaeghe, K.; Mignon, P.; Loverix, S.; Tourwé, D.; Geerlings, P. *J. Chem. Theory Comput.* **2006**, *2*, 1444–1452.
(5) Gil, A.; Branchadell, V.; Bertran, J.; Oliva, A. *J. Phys. Chem. B* **2009**, *113*, 4907–4914.
(6) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(7) Gil, A.; Branchadell, V.; Bertran, J.; Oliva, A. *J. Phys. Chem. B* **2007**, *111*, 9372–9379.
(8) Spooner, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965–1974.
(9) Hutter, M.; Clark, T. *J. Am. Chem. Soc.* **1996**, *118*, 7574–7577.
(10) Bertran, J.; Oliva, A.; Rodríguez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc.* **1998**, *120*, 8159–8167.
(11) Cerón-Carrasco, J. P.; Requena, A.; Zúñiga, J.; Michaux, C.; Perpète, E. A.; Jacquemin, D. *J. Phys. Chem. A* **2009**, *113*, 10549–10556.
(12) Girard, E.; Prangé, T.; Dhaussy, A. C.; Migianu-Griffoni, E.; Lecouvey, M.; Chervin, J. C.; Mezouar, M.; Kahn, R.; Fourme, R. *Nucleic Acids Res.* **2007**, *35*, 4800–4808.
(13) Fonseca Guerra, C.; Bickelhaupt, F. M. *Angew. Chem., Int. Ed.* **1999**, *38*, 2942–2945.
(14) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117–4128.
(15) van der Wijst, T.; Fonseca Guerra, C.; Swart, M.; Bickelhaupt, F. M. *Chem. Phys. Lett.* **2006**, *426*, 415–421.
(16) Fonseca Guerra, C.; van der Wijst, T.; Poater, J.; Swart, M.; Bickelhaupt, F. M. *Theor. Chem. Acc.* **2010**, *125*, 245–252.
(17) Stains, C. I.; Furman, J. L.; Segal, D. J.; Ghosh, I. *J. Am. Chem. Soc.* **2006**, *128*, 9761–9765.
(18) Jeltsch, A. *ChemBioChem* **2002**, *3*, 274–293.
(19) Peräkylä, M. *J. Am. Chem. Soc.* **1988**, *120*, 12895–12902.
(20) Forde, G. K.; Kedzierski, P.; Sokalski, A.; Forde, A. E.; Hill, G. A.; Lesczynski, J. *J. Phys. Chem. A* **2006**, *110*, 2308–2313.
(21) Kangaspeska, S.; Stried, B.; Métivier, R.; Polycarpou-Schwarz, M.; Ibberson, D.; Carmouche, R. P.; Benes, V.; Gannon, F.; Reid, G. *Nature* **2008**, *452*, 112–116.
(22) Labet, V.; Morell, C.; Cadet, J.; Eriksson, L. A.; Grand, A. *J. Phys. Chem. A* **2009**, *113*, 2524–2533.
(23) Petrovich, M.; Veprintsev, D. B. *J. Mol. Biol.* **2009**, *386*, 72–80.
(24) Esteller, M.; Corn, P. G.; Baylin, S. B.; Herman, J. G. *Cancer Res.* **2001**, *61*, 3225–3229.
(25) Costello, J. F.; Frühwald, M. C.; Smiraglia, D. J.; Rush, L. J.; Robertson, G. P.; Gao, X.; Wright, F. A.; Feramisco, J. D.; Peltomäki, P.; Lang, J. C.; Schuller, D. E.; Yu, L.; Bloomfield, C. D.; Caligiuri, M. A.; Yates, A.; Nishikawa, R.; Su Huang, H.-J.; Petrelli, N. J.; Zhang, X.; O'Dorisio, M. S.; Held, W. H.; Cavenee, W. K.; Christoph Plass, C. *Nat. Genet.* **2000**, *25*, 132–138.
(26) Lal, G.; Padmanabha, L.; Provenzano, M.; Fitzgerald, M.; Weydert, J.; Domann, F. M. *Cancer Lett.* **2008**, *267*, 165–174.
(27) Dong-Hyung, L.; Pfeifer, G. P. *J. Biol. Chem.* **2003**, *278*, 10314–10321.
(28) Almatarneh, M. H.; Flinn, C. G.; Poirier, R. A.; Sokalski, W. A. *J. Phys. Chem. A* **2006**, *110*, 8227–8234.
(29) Vu, B.; Cannistraro, V. J.; Sun, L.; Taylor, J. S. *Biochemistry* **2006**, *45*, 9327–9335.
(30) Cannistraro, V. J.; Taylor, J. S. *J. Mol. Biol.* **2009**, *392*, 1145–1157.
(31) Simmen, M. W. *Genomics* **2008**, *92*, 33–40.
(32) Das, A.; Tang, K. S.; Gopalkrishnan, S.; Waring, M. J.; Tomasz, M. *Chem. Biol.* **1999**, *6*, 461–471.
(33) Kawai, K.; Wata, Y.; Hara, M.; Tojo, S.; Majima, T. *J. Am. Chem. Soc.* **2002**, *124*, 3586–3590.
(34) Kanvah, S.; Schuster, G. B. *J. Am. Chem. Soc.* **2004**, *126*, 7341–7344.
(35) Moser, A.; Guza, R.; Tretyakova, N.; York, D. M. *Theor. Chem. Acc.* **2009**, *122*, 179–188.
(36) Dannenberg, J.; Tomasz, M. *J. Am. Chem. Soc.* **2000**, *122*, 2062–2068.
(37) Tretyakova, N.; Guza, R.; Matter, B. *Nucleic Acids Symp. Ser.* **2008**, *52*, 49–50.
(38) Norberg, J.; Vihinen, M. *J. Mol. Struct. Theochem* **2001**, *546*, 51–62.
(39) Rauch, C.; Trieb, M.; Wellenzohn, B.; Loferer, M.; Voegele, A.; Wibowo, F. R.; Liedl, K. R. *J. Am. Chem. Soc.* **2003**, *125*, 14990–14991.
(40) Hodges-Garcia, Y.; Hagerman, P. J. *Biochemistry* **1992**, *31*, 7595–7599.
(41) Wibowo, F. R.; Trieb, M.; Rauch, C.; Wellenzohn, B.; Liedl, K. R. *J. Phys. Chem. B* **2005**, *109*, 557–564.
(42) Pearlman, D. A.; Kollman, P. A. *Biopolymers* **1990**, *29*, 1193–1209.
(43) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247–3276.
(44) Dabkowska, I.; Gonzalez, H. V.; Jurecka, P.; Hobza, P. *J. Phys. Chem. A* **2005**, *109*, 1131–1136.
(45) Dabkowska, I.; Jurecka, P.; Hobza, P. *J. Chem. Phys.* **2005**, *122*, 204322−1–9.
(46) Sponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595–2610.
(47) Cybulski, H.; Sadlej, J. *J. Chem. Theory Comput.* **2008**, *4*, 892–897.
(48) Riley, K. E.; Pitonák, M.; Cerny, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66–80.
(49) Fiethen, A.; Jansen, G.; Hesselmann, A.; Schütz, M. *J. Am. Chem. Soc.* **2008**, *130*, 1802–1803.
(50) Cooper, V. R.; Thonhauser, T.; Puzder, A.; Schröder, E.; Lundqvist, B. I.; Langreth, D. C. *J. Am. Chem. Soc.* **2008**, *130*, 1304–1308.
(51) Escudero, D.; Estarellas, C.; Frontera, A.; Quiñonero, D.; Deyà, P. M. *Chem. Phys. Lett.* **2010**, *485*, 221–225.
(52) Pitonák, M.; Janowski, T.; Neogrády, P.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1761–1766.
(53) Cybulski, S. M.; Seversen, C. E. *J. Chem. Phys.* **2005**, *122*, 014117−1–9.
(54) DiLabio, G. A. *Chem. Phys. Lett.* **2008**, *455*, 348–353.
(55) Swart, M.; van der Wijst, T.; Fonseca Guerra, C.; Bickelhaupt, F. M. *J. Mol. Model.* **2007**, *13*, 1245–1257.
(56) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(57) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
(58) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
(59) Hargis, J. C.; Schaefer, H. F., III; Houk, K. N.; Wheeler, S. E. *J. Phys. Chem A* **2010**, *114*, 2038–2044.
(60) Gu, J.; Wang, J.; Leszcynski, J.; Xie, Y.; Schaefer, H. F., III. *Chem. Phys. Lett.* **2008**, *459*, 164–166.
(61) Gu, J.; Wang, J.; Leszcynski, J:; Xie, Y.; Schaefer, H. F., III. *Chem. Phys. Lett.* **2009**, *473*, 209–210.
(62) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146–10159.
(63) Steinmann, S. N.; Csonka, G.; Corminbouef, C. *J. Chem. Theory Comput.* **2009**, *5*, 2950–2958.
(64) Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. *J. Chem. Theory Comput.* **2008**, *4*, 1996–2000.
(65) Valdes, H.; Spiwok, V.; Rezac, J.; Reha, D.; Abo-Riziq, A. G.; de Vries, M. S.; Hobza, P. *Chem.−Eur. J.* **2008**, *14*, 4886–4898.
(66) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
(67) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
(68) Antony, J.; Brüske, B.; Grimme, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8440–8447.
(69) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554.
(70) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287–5293.
(71) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.

(72) Korth, M.; Grimme, S. *J. Chem. Theory Comput.* **2009**, *5*, 993–1003.

(73) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.

(74) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Quantum Theory*; Wiley: New York, 1986.

(75) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03*, Revision E.01; Gaussian, Inc.: Wallingford, CT, 2004.

(76) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02, Gaussian, Inc.: Wallingford, CT, 2009.

(77) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.

(78) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931–967.

(79) ADF2009.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.

(80) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1975**, *10*, 325–340.

(81) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Clarendon: Oxford, U.K., 1990.

(82) Xaim-1.0. Program developed by José Carlos Ortiz and Carles Bo, Universitat Rovira i Virgili, Tarragona, Spain. http://www.quimica.urv.es/XAIM/.

(83) Asensio, A.; Kobko, N.; Dannenberg, J. J. *J. Phys. Chem. A* **2003**, *107*, 6441–6443.

(84) Grunenberg, J. *J. Am. Chem. Soc.* **2004**, *126*, 16310–16311.

(85) Matta, C. F.; Castillo, N.; Boyd, R. J. *J. Phys. Chem. B* **2006**, *110*, 563–578.

(86) Espinosa, E.; Molins, E. *J. Chem. Phys.* **2000**, *113*, 5686–5694.

(87) Dong, H.; Hua, W.; Li, S. *J. Phys. Chem. A* **2007**, *111*, 2941–2945.

(88) Ebrahimi, A.; Habibi Khorassani, S. M.; Delarami, H. *Chem. Phys.* **2009**, *365*, 18–23.

(89) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickersont, R. E.; Gerstein, M.; Harvey, S. C.; Heinermann, U.; Lu, X.-J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C.-S.; Westhof, E.; Wolberger, C.; Berman, H. M. *J. Mol. Biol.* **2001**, *313*, 229–237.

(90) Sponer, J.; Jureka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142–10151.

(91) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(92) Close, D. M. *J. Phys. Chem. A* **2004**, *108*, 10376–10379.

(93) Li, X.; Cai, Z.; Sevilla, M. D. *J. Phys. Chem. A* **2002**, *106*, 1596–1603.

(94) Pitonák, M.; Neogrády, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1369–1378.

JP103850H