

Metahybrid Density Functional Theory and Correlated *ab Initio* Studies on Microhydrated Adenine–Thymine Base Pairs

Ahmed Dkhissi, ^{*,†,‡} and Ralf Blossey[†]

Interdisciplinary Research Institute, USTL, USR 3078 CNRS, c/o IEMN UMR 8520 CNRS, Avenue Poincaré BP 60069, Cite Scientifique, Villeneuve d'Ascq F-59652 France, and Laboratoire d'Analyse et d'Architecture des Systèmes - CNRS, 7 Av du Colonel Roche 31077 Toulouse Cedex 04, France

Received: April 10, 2008; Revised Manuscript Received: May 09, 2008

Monohydrated and dihydrated adenine–thymine base pairs are characterized using metahybrid density functional theory and correlated *ab initio* approaches. The motivation of this work is twofold. First, the high-level geometries and interaction energies computed for different complexes serve as a reference for the testing of recently developed density functional theory (DFT) with respect to its ability to correctly describe the balance between the electrostatic and the dispersion contributions that bind these complexes. Second, these studies of nucleic acid base pairs are important for finding binding sites of water molecules around bases and for a better understanding of the influence of the solvent on the stability of the structure of DNA duplexes.

Introduction

H-bonding and stacking interactions contribute to the stability and conformational variability of nucleic acids.^{1,2} A proper description of these two different types of noncovalent interactions helps to understand the basic principles governing the formation of the 3D nucleic acid architectures, nanoarchitectures of DNA,³ and so forth. Contrary to hydrogen bonding, the stacking interactions between nucleic acid bases are difficult to study experimentally. The energy information cannot be directly obtained because of the flatness of the potential energy landscape. Solution thermodynamics experiments do not give the structure–energy relationship at the atomic level, and the energy data are largely influenced by other contributions such as hydration and so forth. In principle, theoretical quantum mechanical models are very useful in this regard because they can directly provide the intrinsic strength and geometrical preferences of different types of DNA base pairs, and they allow us to assign reliable values of energy to each base–base configuration. However, calculations should be performed at an adequate theoretical level.

Each DNA structure is the result of a balance of several interactions of the individual DNA building blocks. There are at least three interactions studied in the literature: the stacking of base pairs stabilized by the London dispersion, hydrogen bonding stabilized by the electrostatic interactions, and the interactions with ions and water molecules. Indeed, in the following, we list some studies related to these interactions:

(1) Many theoretical studies, by a broad palette of methods (HF, MP2, CCSD(T), and density functional theory (DFT)), have been devoted to the study of hydration of isolated adenine and thymine bases.^{4–11} The most stable hydration site of the canonical form of adenine was found between the N₃ and H₉ atoms (for the standard atom labeling of the bases). In the most stable structure of dihydrated adenine, the water molecules interact and contact the H₉ and N₃ atoms of adenine.

(2) Further, the stability of H-bonded and stacked base pairs has been largely studied in recent years.^{12–20}

(3) The interaction of metal ions with nucleic acid (NA) bases has been recently studied with correlated *ab initio* methods.²¹

We return now to the solvation of base pairs. From experimental data, it is very well-known that in a nonpolar solvent and also in the gas phase the base pairs prefer a planar hydrogen bonded arrangement, whereas in bulk water, the stacked configurations are preferred.^{22–24} These experimental data have been confirmed by theoretical studies^{25,26} using free-energy/molecular dynamics empirical potential and high *ab initio* studies.

The hydration of both H-bonded and stacked base pairs has been the subject of only a few experimental and theoretical studies.^{27–32}

The first aim of this paper is to characterize mono- and dihydrated complexes of adenine–thymine base pairs using density functional theory and MP2 methods. The dominant intermolecular interactions of adenine–thymine pairs are hydrogen-bonding and π -stacking interactions. It is now generally recognized that the latter interaction can be of equal importance as hydrogen-bonding interactions.³³ However, the difficulty of calculating dispersive interactions using quantum mechanical methods is well-known. High-level *ab initio* methods, at least at the level of MP2 with large basis sets, are required with the associated computational expense. Density functional theory methods, although reducing the computational expense, often fail to adequately describe dispersive interactions.^{34,35} However, the more recently developed functionals of Truhlar and co-workers^{36–38} and DFT-D^{39–41} are showing more promise. So, our second aim of the present work is to reevaluate the credibility of DFT/MPWB1K on both stacking and hydrogen-bonding interactions by comparing the DFT results with MP2 and high *ab initio* advanced *ab initio* calculations. In this context, Kabelac et al.³¹ published recently high *ab initio* calculations combining MP2 with very large basis sets in order to predict a credible interaction energy for adenine–thymine base pairs in the presence of water molecules. These calculations are very interesting and serve as a reference for testing of recently

* To whom correspondence should be addressed.

[†] Interdisciplinary Research Institute.

[‡] Laboratoire d'Analyse et d'Architecture des Systèmes - CNRS.

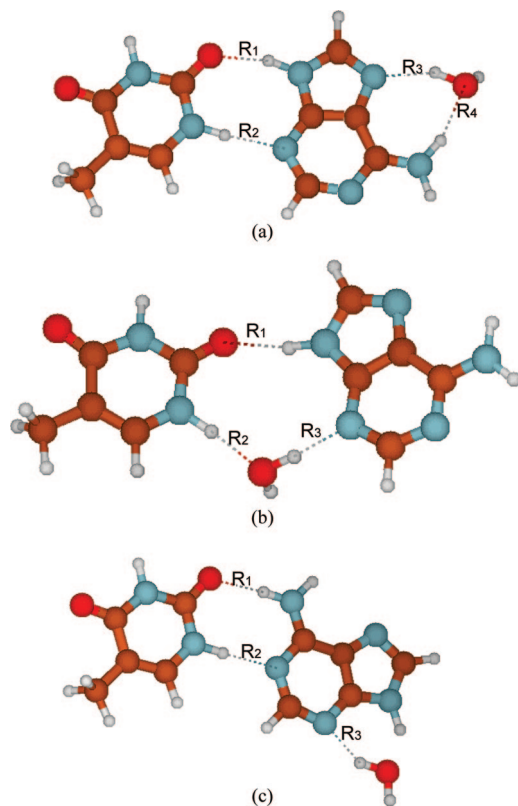


Figure 1. H-bond structures of monohydrated adenine–thymine base pairs. (a) Structure of AT1wHB1; (b) structure of AT1wHB2; (c) structure of AT1wHB3.

developed density functional theories in predicting both stacking and hydrogen-bonding interactions.

Methodology

In our recent study, dedicated to the performance of hybrid metadensity functionals in predicting accurately the stacking and hydrogen-bonding interactions,⁴² we have compared the predictive abilities of different methods in reproducing the known energetic and structural properties of different complexes. The purpose of the investigation was to select the most effective functional for future DFT studies for these interactions. As expected, the study demonstrated that the MPWB1K method appears to be a reliable tool to study the two noncovalent interactions: stacking and hydrogen bonding. Therefore, the latter functional is selected for this study of adenine–thymine in the presence of water molecule complexes.

For the molecular orbital expansion, we have used the 6–31+G** basis set. It has been demonstrated that this basis set is sufficient to predict the main intermolecular properties.^{42–44}

To compare the DFT and correlated ab initio methods, MP2 calculations were performed with the same basis set.

The structures of AT1w and AT2w complexes are fully optimized without any symmetry constraints, and we did not freeze the coordinate of the monomers (rigid monomers) as was done in similar studies.^{12,45,46}

Results and Discussion

In the recent work done by Kabelac et al.,³¹ the potential energy surface (PES) of monohydrated and dihydrated adenine–thymine base pairs was investigated using the MD/Q (molecular dynamics/quenching) method with the Cornell et al.⁴⁷ empirical force field. The PES of monohydrated AT base pair (AT-1w) was

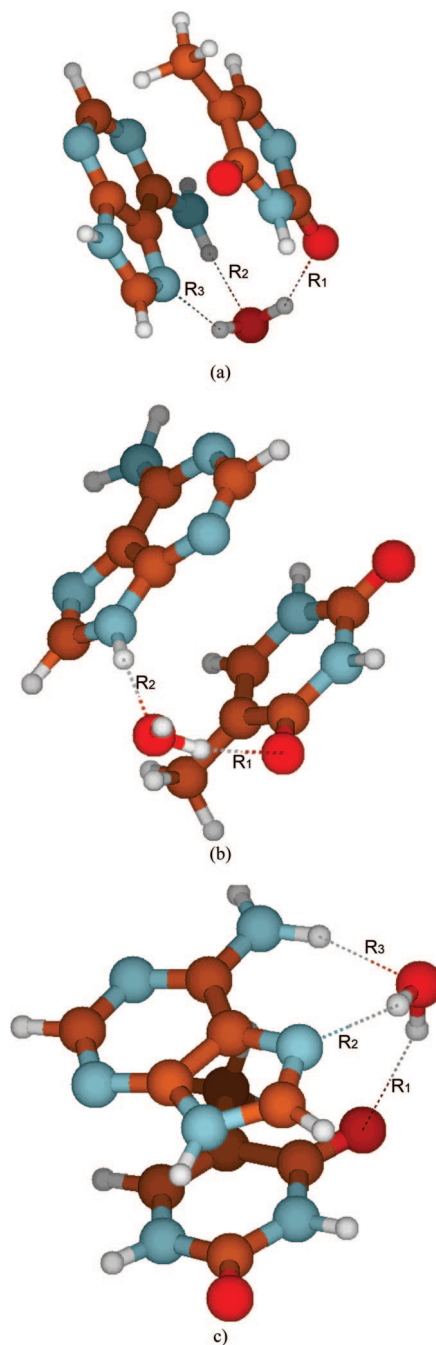


Figure 2. Stacking structures of monohydrated adenine–thymine base pairs. (a) Structure of AT1wS1, (b) structure of AT1wS2, and (c) structure of AT1wS3.

mapped by 337 structures while for dihydrated AT base pair (AT-2w) more than 1740 different structures were localized. In this work, we have studied 12 structures of AT-1w and 12 of AT-2w corresponding to the most stable structures. These structures have water molecules in the most energetically preferred positions. Most of these structures were hydrogen-bonded (HB) or stacked (S). Both HB and S structures can exist in the planar or slightly nonplanar arrangement.

Monohydrated Adenine–Thymine Base Pairs. The structures of the complexes studied in this work are shown in Figure 1.

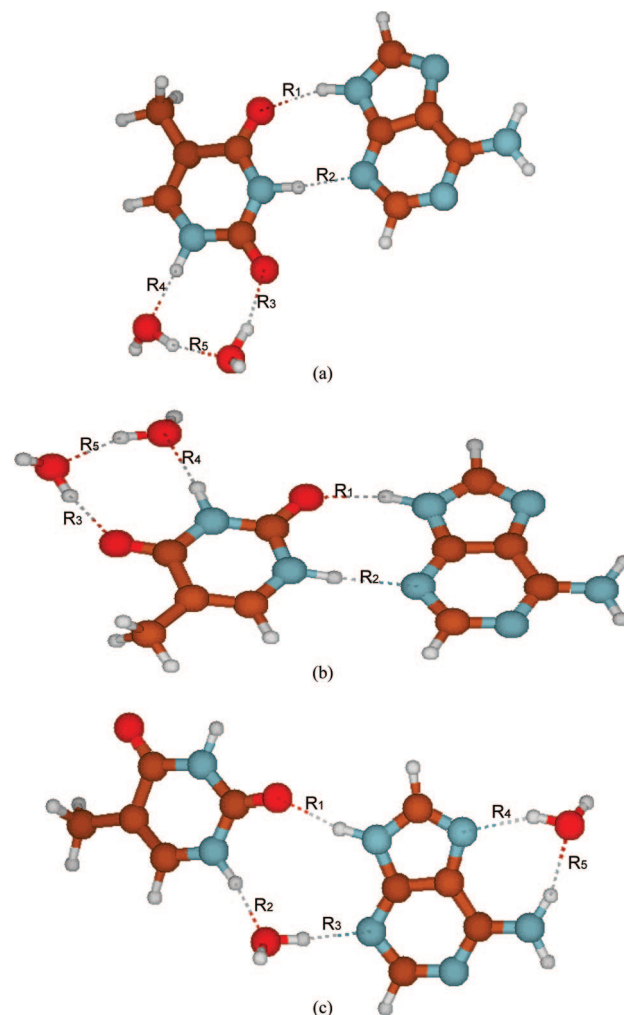
Hydrogen-Bonded Structures. The three structures studied are named AT1wHB1, AT1wHB2, and AT1wHB3. In the first and the third structures, the bases interact together with two classical, strong hydrogen bonds (preserving Watson–Crick

TABLE 1: Selected Structural Data (Distances in Å and Angles in Degrees) for Monohydrated Adenine–Thymine Base Pairs Calculated with DFT and MP2 Methods

	DFT	ab initio ^a
AT1wHB1		
interaction energy	16.20	28.05
interaction R1		
R	2.803	2.759
α	163.9	166.3
interaction R2		
R	2.905	2.871
α	176.7	176.9
interaction R3		
R	2.815	2.833
α	154.0	155.6
interaction R4		
R	2.904	2.872
α	162.3	164.6
AT1wHB2		
interaction energy	14.97	26.23
interaction R1		
R	2.814	2.759
α	168.4	163.3
interaction R2		
R	2.785	2.730
α	175.0	174.1
interaction R3		
R	2.775	2.773
α	174.7	169.0
AT1wHB3		
interaction energy	13.17	25.31
interaction R1		
R	2.895	2.881
α	175.3	175.9
interaction R2		
R	2.916	2.872
α	178.3	178.8
interaction R3		
R	2.823	2.854
α	140.7	139.5
AT1wS1		
interaction energy	12.58	23.71
interaction R1		
R	2.926	2.907
α	151.4	152.3
interaction R2		
R	2.865	2.830
α	161.7	163.8
interaction R3		
R	2.902	2.881
α	145.6	146.8
AT1wS2		
interaction energy	11.91	23.38
interaction R1		
R	2.770	2.788
α	164.0	164.7
interaction R2		
R	2.849	2.750
α	162.6	160.5
AT1wS3		
interaction energy	11.83	23.12
interaction R1		
R	2.937	2.934
α	151.6	152.2
interaction R2		
R	2.901	2.900
α	145.6	145.8
interaction R3		
R	2.877	2.878
α	160.9	161.0

^a The interaction energies reported in the table are from ref 31.

structure). The water molecules usually lie in the plane of the bases. In the second structure (AT1wHB2), the water molecule breaks the H-bond interaction N–H···H and is inserted there. AT1whb1 and AT1whb2 are stabilized with four hydrogen bonds (N–H···O or N–H···N) while the third structure is stabilized with only three hydrogen bonds. The DFT calculations indicate that the most stable complex is AT1Whb1, which is in good agreement with the high ab initio calculations. What is

**Figure 3.** H-bond structures of dihydrated adenine–thymine base pairs. (a) Structure of AT2wHB1, (b) structure of AT2wHB2, and (c) structure of AT2wHB3.

interesting is that the order of stability between the three structures obtained with DFT is the same as those obtained with the ab initio methods. Further, the energetic difference between each two structures calculated with DFT method is in good agreement with that obtained from RI-MP2/TZVPP//RI-MP2/cc-pvDZ approaches.

The DFT predicted geometries are usually similar to geometries obtained with MP2. Table 1 contains the selected structural properties for monohydrated adenine–thymine base pair complexes. The performance of MPWB1K is illustrated, in this table, in comparison with MP2.

The intermolecular distances R(N–N) or R(N–O) for N–H···N and N–H···O H-bonds, respectively, calculated with DFT are in excellent agreement with those obtained with MP2. The mean deviations for the predicted intermolecular distances (in comparison with MP2) are less than 0.06 Å. This range is the same as has been obtained with the popular B3LYP functional in similar complexes.^{42–44} This result is very interesting and confirms the credibility of MPWB1K in predicting the structural properties of H-bond interactions.

Stacking Structures. In these structures, water molecules are usually oriented perpendicularly to the planes of the bases helping to join them together. These structures are stabilized with two interaction types: stacking interactions between the bases and several hydrogen bonds. In each complex, the water molecule behaves as a proton donor and as a proton acceptor.

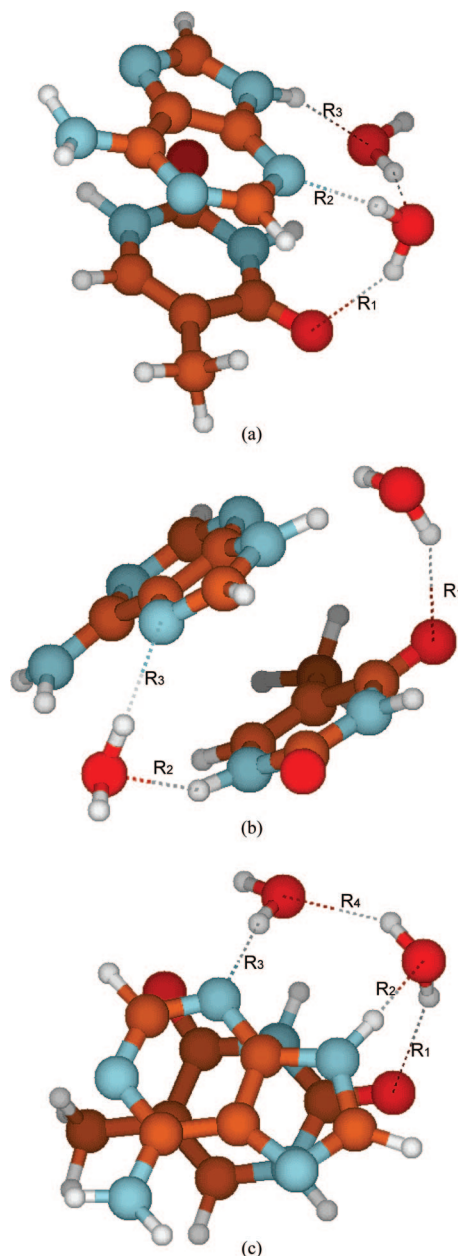


Figure 4. Stacking structures of dihydrated adenine–thymine base pairs. (a) Structure of AT2wS1, (b) structure of AT2wS2, and (c) structure of AT2wS3.

The DFT calculated intermolecular distances between the bases are 3.39, 3.27, and 3.42 Å for AT1wS1, AT1wS2, and AT1wS3, respectively. These results are in reasonable agreement with those obtained with MP2 (3.22, 3.09, and 3.31 Å). For hydrogen-bonding interactions, the prediction obtained with DFT is in better agreement with MP2.

Comparing the HB and S structures for AT-1w, the most stable structure corresponds to the HB and to the fourth local minimum, and it is 3.6 kcal/mol less stable than the global HB minimum. This finding is again in excellent agreement with MP2 results (see Table 1).

Dihydrated Adenine–Thymine Base Pairs. The structures of the complexes studied in this section are shown in Figures 2–4.

Hydrogen-Bonded Structures. Two possible arrangements exist in the presence of two water molecules. The first is a motif with two water molecules joined together in one side of thymine (AT2wHB1 and AT2wHB2) preserving the Watson–Crick

TABLE 2: Selected Structural data (Distances in Å and Angles in Degrees) for Dihydrated Adenine–Thymine Base Pairs Calculated with DFT and MP2 Methods

	DFT	MP2 ^a
AT2wHB1		
interaction energy (ΔE)	21.47	36.56
interaction R1		
R	2.820	2.804
α	159.8	162.2
interaction R2		
R	2.922	2.852
α	174.0	176.2
interaction R3		
R	2.751	2.764
α	167.3	167.3
interaction R4		
R	2.830	2.772
α	177.1	177.8
interaction R5		
R	2.696	2.693
α	159.6	162.6
AT2wHB2		
ΔE	20.94	36.09
interaction R1		
R	2.803	2.761
α	163.0	165.5
interaction R2		
R	2.928	2.886
α	176.5	177.2
interaction R3		
R	2.761	2.782
α	168.5	168.2
interaction R4		
R	2.829	2.786
α	173.6	176.1
interaction R5		
R	2.713	2.716
α	158.3	163.7
AT2wHB3		
ΔE	21.28	36.04
interaction R1		
R	2.815	2.803
α	169.3	168.0
interaction R2		
R	2.774	2.776
α	173.1	173.5
interaction R3		
R	2.763	2.759
α	175.6	174.6
interaction R4		
R	2.809	2.811
α	154.2	154.1
interaction R5		
R	2.908	2.910
α	162.3	162.3
AT2wS1		
ΔE	22.04	37.20
interaction R1		
R	2.841	2.841
α	149.1	149.1
interaction R2		
R	2.892	2.892
α	159.0	159.0
interaction R3		
R	2.862	2.862
α	158.3	158.3
interaction R4		
R	2.645	2.645
α	163.5	163.5
AT2wS2		
ΔE	21.17	36.40
interaction R1		
R	2.896	2.876
α	152.0	152.2
interaction R2		
R	2.902	2.827
α	150.1	154.1
interaction R3		
R	2.774	2.769
α	150.4	151.9
AT2wS3		
ΔE	21.59	36.22
interaction R1		
R	2.841	2.869
α	146.9	149.1
interaction R2		
R	2.787	2.739
α	163.0	161.5
interaction R3		
R	2.880	2.876
α	165.1	162.0
interaction R4		
R	2.879	2.886
α		
	145.4	149.2

^a Interaction energies are reported from ref 30.

structure. In the second motif, the two water molecules lie separated on the opposite sides of adenine giving rise to the

AT2wHB3 where one water molecule breaks the strong interaction between the bases and is inserted there as in the case of monohydrated AT base pairs. All these structures are stabilized with several hydrogen-bond interactions: three perfectly linear hydrogen bonds (R1–R3) and two distorted hydrogen bonds (R4–R5). The DFT calculations indicate that these structures have similar stability, which is in good agreement with the results obtained with the correlated ab initio method.

The selected structural properties calculated with DFT and MP2 for dihydrated AT base pairs are summarized in Table 2. As noted before, the results from DFT and MP2 agree well.

Stacked Structures. The dihydrated S structures contain two types of such water bridges: one contains two separate water molecules that lie on opposite side of the bases, while in the second they interact with each other. These structures are named AT2wS1, AT2wS2, and AT2wS3. These complexes are stabilized with two noncovalent bonds: stacking interactions between adenine and thymine and hydrogen bonding due to the presence of water molecules. The DFT calculations indicate that the most stable structure in stacking configurations is AT2wS1, and there are only very small differences in energies between the different structures. These results are again in good agreement with the prediction obtained with MP2.

The intermolecular distance between the bases calculated with DFT are 3.44, 3.35, and 3.45 Å for AT2wS1, AT2wS2, and AT2wS3, respectively. These results are again in reasonable agreement with those obtained with MP2 (3.44, 3.20, and 3.22 Å).

For all S structures, the thymine base is rotated by approximately 225° with respect to the adenine molecule. This rotation is in good agreement with those obtained with MP2 (229°). Such an orientation of stacked bases has been already observed for adenine–uracil.⁴⁸

Comparing the HB and S structures for dihydrated adenine–thymine, the presence of two water molecules leads to larger stability of S structures. The global minimum is AT2wS1. However, this complex is slightly more stable than AT2wHB1 (only 0.5 kcal/mol difference).

Conclusions

The most stable structures for mono- and dihydrated adenine–thymine complexes, obtained by Kabelac et al.,³¹ were fully optimized with DFT/6–31+G** and MP2/6–31+G** approaches. In the gas phase, we obtained that the HB structures are more stable than the S structures by more than 90%. The presence of one water molecule in the AT base pair reduced the difference between the most stable HB and S structures by 4 kcal/mol. Also, the addition of the second water molecule leads to equivalent stability of S and HB structures. As has been demonstrated by Kabelac et al.,³¹ our results confirm that the preference for S structures of nucleic acids base pairs in an aqueous environment is due to hydrophilic interactions of a rather small amount of water molecules with the base pairs and is not due to a hydrophobic interaction between a large bulk of solvent and base pair as is generally believed.

Further, the meta-hybrid density functional DFT/MPWB1K is very promising for studies of larger biomolecules as DNA duplex or at least to serve as development of new force fields that are computationally efficient and capable of modeling large biomolecules.

References and Notes

- (1) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911.

- (2) Bommarito, S.; Peyret, N.; SantaLucia, J. *J. Nucleic Acids Res.* **2000**, *28*, 1929.
- (3) Feldkamp, H.; Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2006**, *45*, 2.
- (4) van Mourik, T.; Price, S. L.; Clary, D. C. *J. Phys. Chem. A* **1999**, *103*, 1611.
- (5) van Mourik, T.; Benoit, D. M.; Price, S. L.; Clary, D. C. *Phys. Chem. Chem. Phys.* **2000**, *2*, 1281.
- (6) van Mourik, T. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2886.
- (7) Kryachko, E.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2001**, *105*, 3379.
- (8) Aamouche, A.; Berthier, G.; Cadioli, B.; Gallinella, E.; Ghomi, M. *THEOCHEM* **1998**, *426*, 307.
- (9) Ghomi, M.; Aamouche, A.; Cadioli, B.; Berthier, G.; Grajcar, L.; Baron, M. H. *J. Mol. Struct.* **1997**, *410*, 323.
- (10) Aleman, C. *Chem. Phys. Lett.* **1999**, *302*, 461.
- (11) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **1998**, *102*, 6010.
- (12) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247–3276.
- (13) Sponer, J.; Hobza, P. *Coll. Czech. Chem. C* **2003**, *68*, 2231–2282.
- (14) Shukla, M. K.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 11338.
- (15) Haranczyk, M.; Gutowski, M. *J. Am. Chem. Soc.* **2005**, *127*, 699.
- (16) MacNaughton, J.; Moewes, A.; Kurmaev, E. Z. *J. Phys. Chem. B* **2005**, *109*, 7749.
- (17) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
- (18) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
- (19) Perun, S.; Sobolewski, A. L.; Domcke, W. *Chem. Phys.* **2005**, *313*, 107.
- (20) Blancafort, L.; Cohen, B.; Hare, P. M.; Kohler, B. *J. Phys. Chem. A* **2005**, *109*, 4431.
- (21) Kabelac, M.; Hobza, P. *J. Phys. Chem. B* **2006**, *110*, 14515.
- (22) Yanson, B. T.; Sukhodub, I. K. *Biopolymers* **1979**, *18*, 1149–1156.
- (23) Schweizer, M. P.; Broom, A. D.; Tso, P. O. P.; Hollis, D. P. *J. Am. Chem. Soc.* **1968**, *90*, 1042.
- (24) Binford, J.; Holloway, D. M. *J. Mol. Biol.* **1968**, *31*, 91–97.
- (25) Cieplak, P.; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 3734–3739.
- (26) Miller, J. L.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 8587–8594.
- (27) Sivanesan, D.; Sumathi, I.; Welsh, W. J. *Chem. Phys. Lett.* **2003**, *367*, 351.
- (28) Sivanesan, D.; Babu, K.; Gadre, S. R.; Subramanian, V.; Ramasami, T. *J. Phys. Chem. A* **2000**, *104*, 10887.
- (29) Kabelac, M.; Hobza, P. *Chem. Eur. J.* **2001**, *7*, 2067.
- (30) Kabelac, M.; Sherer, E. C.; Cramer, C. J.; Hobza, P. *Chem. Eur. J.* **2007**, *13*, 2067.
- (31) Kabelac, M.; Zendlova, L.; Recha, D.; Hobza, P. *J. Phys. Chem. B* **2005**, *109*, 12206.
- (32) Zendlova, L.; Hobza, P.; Kabelac, M. *Chem. Phys. Chem.* **2006**, *7*, 439.
- (33) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
- (34) Kohn, W.; Meir, Y.; Makarov, D. E. *Phys. Rev. Lett.* **1998**, *80*, 4153.
- (35) Mourik, T. V.; Gdanitz, R. J. *J. Chem. Phys.* **2002**, *116*, 9620.
- (36) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 415.
- (37) Zhao, Y.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2701.
- (38) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415.
- (39) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
- (40) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287.
- (41) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.
- (42) Dkhissi, A.; Blossey, R. *Chem. Phys. Lett.* **2007**, *439*, 35.
- (43) Dkhissi, A.; Adamowicz, L.; Maes, G. J. *Phys. Chem. A* **2000**, *104*, 2112.
- (44) Dkhissi, A.; Adamowicz, L.; Maes, J. *Phys. Chem. A* **2000**, *104*, 5625.
- (45) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
- (46) Sponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *276*, 263.
- (47) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (48) Florian, J.; Sponer, J.; Warshel, A. J. *Phys. Chem. B* **1999**, *103*, 884.