Biochemistry (Moscow), Vol. 70, No. 3, 2005, pp. 366-376. Translated from Biokhimiya, Vol. 70, No. 3, 2005, pp. 447-458.
Original Russian Text Copyright © 2005 by Monajjemi, Chahkandi, Zare, Amiri.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM04-122, December 19, 2004.

Study of the Hydrogen Bond in Different Orientations of Adenine–Thymine Base Pairs: an *ab initio* Study

M. Monajjemi[1]*, B. Chahkandi[1,2], K. Zare[1], and A. Amiri[3]

[1]*Science and Research Campus, Islamic Azad University, Hesarak, Poonak,*
P.O. Box 14515-775, Tehran, Iran; E-mail: m_monajjemi@yahoo.com
[2]*Department of Chemistry, Shahrood Campus, Islamic Azad University, Shahrood, Iran*
[3]*Department of Chemistry, Central Tehran Campus, Islamic Azad University, Tehran, Iran*

Received May 15, 2004
Revision received September 13, 2004

Abstract—In order to gain deeper insight into structure, charge distribution, and energies of A–T base pairs, we have performed quantum chemical *ab initio* and density functional calculations at the HF (Hartree–Fock) and B3LYP levels with *3-21G**, *6-31G**, *6-31G***, and *6-31++G*** basis sets. The calculated donor–acceptor atom distances in the Watson–Crick A–T base pair are in good agreement with the experimental mean values obtained from an analysis of 21 high resolution DNA structures. In addition, for further correction of interaction energies between adenine and thymine, the basis set superposition error (BSSE) associated with the hydrogen bond energy has been computed via the counterpoise method using the individual bases as fragments. In the Watson–Crick A–T base pair there is a good agreement between theory and experimental results. The distances for (N2...H23–N19), (N8–H13...O24), and (C1...O18) are 2.84, 2.94, and 3.63 Å, respectively, at B3LYP/*6-31G*** level, which is in good agreement with experimental results (2.82, 2.98, and 3.52 Å). Interaction energy of the Watson–Crick A–T base pair is –13.90 and –10.24 kcal/mol at B3LYP/*6-31G*** and HF/*6-31G*** levels, respectively. The interaction energy of model (9) A–T base pair is larger than others, –18.28 and –17.26 kcal/mol, and for model (2) is the smallest value, –13.53 and –13.03 kcal/mol, at B3LYP/*6-31G*** and B3LYP/*6-31++G*** levels, respectively. The computed B3LYP/*6-31G*** bond enthalpies for Watson–Crick A–T pairs of –14.4 kcal/mol agree well with the experimental results of –12.1 kcal/mol deviating by as little as –2.3 kcal/mol. The BSSE of some cases is large (9.85 kcal/mol) and some is quite small (0.6 kcal/mol).

Key words: Watson–Crick base pairs, *ab initio*, B3LYP, hydrogen bonding

The interaction between nucleotide bases is an important element in the structure of DNA sequences. These interactions involve hydrogen bonding between base pairs [1-4]. Apart from providing water with physical properties that make it the ideal medium for many life processes to take place in, it is responsible for various types of self-organization and molecular recognition, such as the folding of proteins. In 1953, Watson and Crick [4] using data from other laboratories built a model of DNA. This model involves pairing the nucleic base adenine with thymine and cytosine with guanine. They showed that the fundamental unit of DNA involves the helical intertwining of two chains held together by hydrogen bonds through the pairings of A–T and C–G. Adenine and guanine are purines and thymine and cytosine are pyrimidines.

In the past decade, *ab initio* and density functional theory (DFT) quantum chemical studies [5-16] have appeared on the geometry, energy, and other aspects of the hydrogen bonds that hold A–T and G–C pairs together. The adequacy of DFT for hydrogen-bonded systems has received much attention lately [17-33]; this is known from the investigations of Sim et al. [17]. They found that DFT results are of comparable quality to those from correlated *ab initio* methods. Others [9, 14-16] have shown that this is also true for the corresponding structures; minor but signification deviations from experimental values were obtained with both the DFT and *ab initio* methods. Guerra and others [34-37] have shown that these structural deviations are a result of intermolecular interactions of the base pairs with the environment in the crystal.

Maranon and coworkers [38] obtained the geometry of A–T base pair in DNA molecule (see Fig. 1, A–T (1))

Abbreviations: DFT) density functional theory; HF) Hartree–Fock approximation; B3LYP (Beck3-LYP) and BP86) methods of density functional theory; BSSE) basis set superposition error.
* To whom correspondence should be addressed.

in molecular dynamic simulations (MD) using the Gaussian 94 package [39]. The geometry of the A–T base pair was optimized with the restriction that both molecules can only move over the same plane. The *ab initio* calculations were carried out using the HF/*3-21G** level. The results of this calculation showed that the lengths of hydrogen were d(N2...H23–N19) = 0.2796 nm and d(N8–H13...O24) = 0.2988 nm [38]. Maranon and coworkers [38] studied MD simulation of the double hydrogen bonds of the A–T pair in water. They simulated each hydrogen bond of the base pair by two covalent bonds and used an adequate value of the Van der Waals parameters in order to give a good simulation of the hydrogen bond.

Also, structural geometries and energies of A–T pair were carried out by HF/*6-31G*** and B3LYP/*6-31G*** base by Meyer and Suhnel [14].

In the present paper, we address a different point. Both density functional and *ab initio* methods satisfactorily reproduce experimental A–T hydrogen bond distance and interaction energies in different states of A–T base pairs. We try to calculate geometries, energies, enthalpies, and atomic charges of the hydrogen bonds in the Watson–Crick DNA base pair and other different orientations of A–T base pairs (see Fig. 1) with *ab initio* and DFT quantum chemical studies.

Density functional quantum chemical calculations have recently provided a relatively consistent picture on base pair interaction energies and geometries; we have performed calculations of this type for the A–T base pairs. This can lead to more detailed information on structure, charge distribution, and energetics of the base pair as compared to the simple iso-steric concept. The enthalpies of A–T base pairs in nine different orientations are calculated at B3LYP/*6-31G**, B3LYP/*6-31G***, and HF/*6-31G*** levels with frequency calculations at these levels.

METHODS OF INVESTIGATION

Ab initio calculation in the HF/*3-21G**, HF/*6-31G**, HF/*6-31G***, and DFT studies at the B3LYP/*3-21G**, B3LYP/*6-31G**, B3LYP/*6-31G***, and B3LYP/*6-31++G*** levels were carried out to determine the complex structures with the Gaussian 98 program [40]. The interaction energies were corrected for the basis set superposition error (BSSE) by the standard counterpoise method (CP) approach [41].

Ab initio interaction energies were evaluate using HF/*3-21G**, HF/*6-31G**, and HF/*6-31G*** optimized geometries and density functional theory B3LYP/*3-21G**, B3LYP/*6-31G**, B3LYP/*6-31G***, and B3LYP/*6-31++G*** levels. The counterpoise correction method was calculated according to the following equation:

$$\Delta E_{BSSE} = E(AB.r_C)^{AB} - E(A, r_e) - E(B, r_e) + E^{DEF}, \quad (1)$$

where

$$E^{DEF} = \left[E(A, r_e) - E(A, r_C)^{AB}\right] + \left[E(B, r_e) - E(B, r_C)^{AB}\right].$$

Here, the deformation energy E^{DEF} is defined as the energy difference between the geometry optimized monomers and the structures of the monomers adopted in the complex; the label r_c is used to indicate the geometry of the product complex AB while r_e indicates the geometry of the separate reactants.

Mean values of intermolecular donor–acceptor atom distance of A–T pairs in the Watson–Crick model have been determined from 21 DNA structures with resolution better than 1.5 Å using the program HBExplore [42] for comparison between experimental and calculated data.

The frequencies of A–T in the Watson–Crick and eight different orientations were calculated at B3LYP/*6-31G***, B3LYP/*6-31G**, and HF/*6-31G*** levels. Bond enthalpies (ΔH_{298}) were obtained from frequency calculations at 298 K and 1 atm assuming an ideal gas:

$$\Delta H_{298} = \sum (E_0 + H_{Corr})_{Product} - \sum (E_0 + H_{Corr})_{reactants}, \quad (2)$$

where

$$H_{Corr} = E_{total} + k_B T$$

and

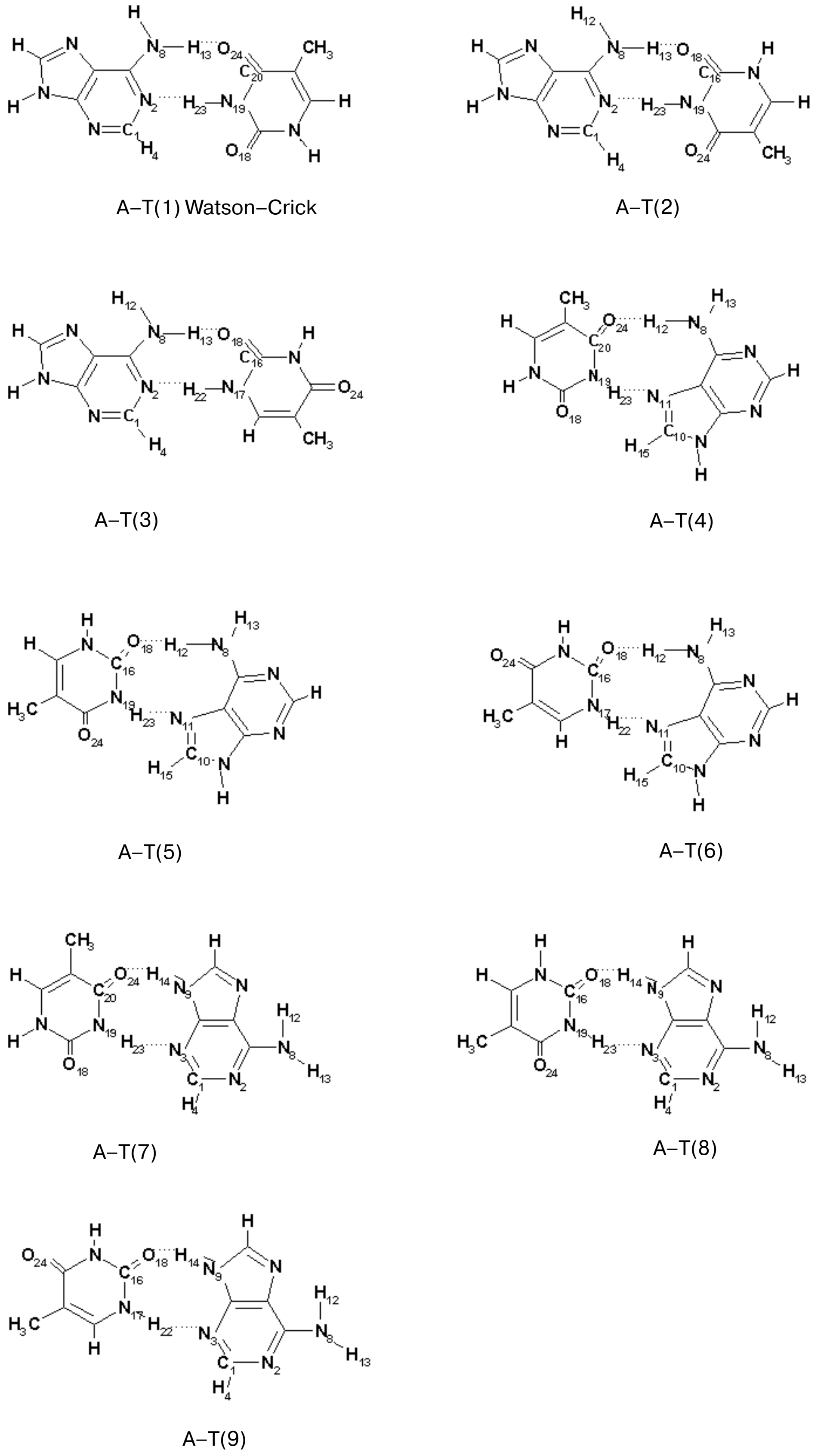
$$E_{total} = E_{trans} + E_{rot} + E_{vib} + E_{elec}.$$

Here E_0 is total electronic energy of monomers and dimer and H_{Corr} is the thermal correction for enthalpy. Also, E_{trans}, E_{rot}, E_{vib}, and E_{elec} are the translational, rotational, zero point vibrational, and electronic energies based on our frequency calculations. The BSSE associated with the hydrogen bond energy has been computed by the counterpoise method using the individual bases as fragments.

RESULTS AND DISCUSSION

The results obtained give a clear picture of the differences between different orientations of A–T base pairs. Nevertheless, we discuss briefly a methodological aspect of the calculations first and then proceed with a discussion of the basic results without reiterating statements on the finer details of approximations and methods. The results of our B3LYP and HF study on the formation of the adenine–thymine complexes are summarized and compared with literature in Tables 1 and 2.

The computed B3LYP/*6-31G*** bond enthalpies for Watson–Crick A–T pairs of –14.4 kcal/mol agree well with the experimental results of –12.1 kcal/mol [43], deviating by as little as –2.3 kcal/mol (see Table 2). The



A–T(1) Watson–Crick

A–T(2)

A–T(3)

A–T(4)

A–T(5)

A–T(6)

A–T(7)

A–T(8)

A–T(9)

Fig. 1. Watson–Crick geometry of A–T base pair (1) and other eight different orientations of A–T.

Table 1. Geometries of A–T base pairs in nine different orientations in gas phase obtained from various levels of theory

Base pairs	HF/ /3–21G*	HF/ /6–31G*	HF/ /6–31G **	B3LYP/ /3–21G*	B3LYP/ /6–31G*	B3LYP/ /6–31G**	B3LYP/ /6–31++G**
1	2	3	4	5	6	7	8
A–T(1)							
r(H13...O24)	1.971	2.075	2.087	1.822	1.928	1.921	1.920
r(N2...H23)	1.747	1.979	1.976	1.611	1.831	1.799	1.828
r(H4...O18)	2.574	2.981	2.963	2.497	2.849	2.804	2.863
r(N8...O24)	2.973	3.071	3.081	2.856	2.948	2.940	2.939
r(N2 ...N19)	2.779	2.994	2.990	2.690	2.876	2.848	2.875
r(C1 ...O18)	3.419	3.787	3.775	3.363	3.672	3.633	3.684
∠(N8H13O24)	172.1	172.9	172.9	174.2	174.5	174.6	174.1
∠(H13O24C20)	124.3	126.4	126.3	122.8	124.5	124.5	125.8
∠(N2H23N19)	178.1	178.0	178.8	178.1	179.7	179.7	179.4
∠(H23N2C1)	115.2	117.0	116.6	116.3	117.8	117.5	117.3
A–T(2)							
r(H13...O18)	1.972	2.087	2.092	1.839	1.949	1.944	1.943
r(N2...H23)	1.745	1.985	1.978	1.606	1.834	1.796	1.818
r(H4...O24)	2.571	2.981	2.964	2.473	2.833	2.776	2.825
r(N8...O18)	2.973	3.082	3.086	2.870	2.967	2.961	2.960
r(N2...N19)	2.778	2.998	2.991	2.686	2.878	2.844	2.866
r(C1 ...O24)	3.419	3.970	3.777	3.346	3.660	3.610	3.651
∠(N8H13O18)	171.7	172.7	172.7	173.4	174.2	174.1	173.9
∠(H13O18C16)	123.1	124.8	124.7	121.3	122.4	122.1	122.0
∠(N2H23N19)	178.0	177.8	177.0	178.7	178.8	179.1	178.8
∠(H23N2C1)	114.3	116.1	115.9	115.3	116.8	116.5	116.5
A–T(3)							
r(H13...O18)	1.846	2.001	2.011	1.727	1.877	1.869	1.876
r(N2...H22)	1.881	2.039	2.047	1.732	1.896	1.875	1.887
r(N8...O18)	2.855	3.001	3.009	2.766	2.900	2.891	2.898
r(N2...N17)	2.904	3.048	3.054	2.795	2.934	2.914	2.926
∠(N8H13O18)	175.6	175.8	176.4	176.7	176.9	177.1	176.6
∠(H13O18C16)	123.8	123.9	123.5	122.2	122.1	121.0	122.7
∠(N2H22N17)	177.2	177.5	177.8	178.0	178.6	178.7	178.6
∠(H22N2C1)	119.5	120.4	120.6	120.6	121.2	121.1	120.6
A–T(4)							
r(H12...O24)	2.018	2.148	2.148	1.871	1.959	1.964	1.965
r(N11...H23)	1.726	1.954	1.950	1.599	1.804	1.779	1.791
r(N8...O24)	3.01	3.136	3.135	2.895	2.970	2.973	2.974
r(N11...N19)	2.751	2.964	2.959	2.668	2.844	2.821	2.836
r(H15...O18)	2.518	2.915	2.907	2.432	2.802	2.756	2.791
r(C10 ...O18)	3.245	3.594	3.588	3.185	3.492	3.453	3.483
∠(N8H12O24)	172.6	170.0	170.0	171.8	170.3	170.1	170.1
∠(H12O24C20)	130.8	132.0	132.0	129.9	131.7	131.2	131.9
∠(N11H23N19)	172.6	175.1	175.2	172.7	174.8	174.6	174.9
∠(H23N11C10)	117.7	118.6	118.5	118.3	119.7	119.5	119.5
A–T(5)							
r(H12...O18)	2.008	2.155	2.155	1.882	1.987	1.988	1.985
r(N11...H23)	1.720	1.951	1.946	1.595	1.804	1.767	1.801
r(N8...O18)	3.008	3.144	3.143	2.907	2.996	2.996	2.993

Table 1. (Contd.)

1	2	3	4	5	6	7	8
r(N11...N19)	2.747	2.961	2.956	2.665	2.844	2.811	2.845
r(H15...O24)	2.522	2.906	2.900	2.413	2.777	2.716	2.798
r(C10 ...O24)	3.247	3.590	3.586	3.171	3.473	3.421	3.492
∠(N8H12O18)	172.9	170.4	170.4	172.3	170.2	170.2	170.4
∠(H12O18C16)	129.5	130.6	130.5	128.1	129.0	128.5	129.9
∠(N11H23N19)	173.3	176.5	176.5	173.3	175.9	175.3	176.3
∠(H23N11C10)	117.3	117.6	117.6	117.7	118.7	118.6	118.4
A–T(6)							
r(H12...O18)	1.850	2.023	2.025	1.740	1.888	1.883	1.888
r(N11...H22)	1.888	2.072	2.070	1.737	1.906	1.891	1.989
r(N8...O18)	2.849	3.007	3.006	2.765	2.895	2.888	2.894
r(N11...N17)	2.908	3.079	3.076	2.796	2.940	2.925	2.933
∠(N8H12O18)	170.8	167.4	167.2	170.2	168.5	168.1	168.2
∠(H12O18C16)	132.2	130.9	130.8	130.3	129.9	129.6	129.0
∠(N11H22N17)	177.8	178.1	177.0	176.9	176.5	176.4	176.5
∠(H22N11C10)	122.3	122.8	122.9	122.0	123.5	123.5	123.3
A–T(7)							
r(H23...N3)	1.800	2.021	2.011	1.661	1.864	1.827	1.865
r(O24...H14)	1.822	1.995	1.996	1.712	1.854	1.838	1.850
r(N9...O24)	2.798	2.955	2.953	2.733	2.852	2.834	2.841
r(N19 ...N3)	2.829	3.032	3.020	2.733	2.905	2.871	2.908
r(O18 ...H4)	3.288	3.594	3.756	3.162	3.439	3.391	3.464
r(C1 ...O18)	3.917	4.197	4.184	3.813	4.056	4.014	4.084
∠(N9H14O24)	160.6	159.3	158.9	164.5	163.2	162.5	161.2
∠(N19H23N3)	176.1	175.1	174.9	177.2	176.8	176.7	175.8
∠(H14O24C20)	133.1	134.8	135.0	129.2	131.1	131.6	133.9
∠(H23N3C1)	128.1	129.6	129.2	129.2	130.9	130.4	129.6
A–T(8)							
r(H23...N3)	1.797	2.019	2.016	1.654	1.871	1.837	1.860
r(O18...H14)	1.829	2.002	2.002	1.730	1.879	1.864	1.872
r(N9...O18)	2.805	2.958	2.956	2.745	2.874	2.857	2.858
r(N19 ...N3)	2.826	3.027	3.024	2.725	2.911	2.879	2.902
r(O24 ...H4)	3.276	3.586	3.581	3.128	3.429	3.386	3.437
r(C1 ...O24)	3.903	4.196	4.192	3.788	4.049	4.012	4.064
∠(N9H14O18)	160.9	158.4	158.4	163.5	162.4	161.9	160.9
∠(N19H23N3)	175.9	173.7	173.8	176.4	175.7	175.7	174.5
∠(H14O18C16)	130.8	133.7	133.7	127.8	129.2	129.4	131.6
∠(H23N3C1)	127.0	128.3	128.2	128.1	129.9	129.5	128.4
A–T(9)							
r(H22...N3)	1.878	2.056	2.048	1.731	1.901	1.870	1.880
r(O18...H14)	1.773	1.950	1.948	1.668	1.822	1.803	1.811
r(N9...O18)	2.763	2.925	2.921	2.695	2.830	2.810	2.817
r(N17 ...N3)	2.900	3.063	3.055	2.793	2.937	2.908	2.918
∠(N9H14O18)	164.1	162.0	162.7	166.0	165.7	165.1	164.6
∠(N17H22N3)	175.9	176.7	176.6	175.0	177.3	177.4	177.3
∠(H14O18C16)	129.7	130.1	130.3	127.7	127.5	127.8	128.5
∠(H22N3C1)	131.7	133.5	133.3	131.9	133.9	133.7	133.4

Table 2. Calculated binding energy and enthalpy of bonds at 298 K in nine A–T base pairs at various levels (in kcal/mol)

Level	ΔE	ΔE_{BSSE}	ΔH_{298}	$\overline{\Delta H}_{298}$
1	2	3	4	5
A–T(1)				−12.75
HF/3–21G*	−22.0	−14.05		
HF/6–31G*	−11.73	−10.02		
HF/6–31G **	−11.81	−10.24	−10.08	
B3LYP/3–21G*	−28.89	−19.32		
B3LYP/6–31G*	−16.20	−13.47	−14.7	
B3LYP/6–31G**	−16.44	−13.90	−14.4	
B3LYP/6–31++G**	−12.79	−13.35		
BP86/TZ2P 12	−13.0	−12.3	−11.8	
A–T(2)				−12.78
HF/3–21G*	−21.98	−14.02		
HF/6–31G*	−11.65	−9.96		
HF/6–31G **	−11.73	−10.19	−10.03	
B3LYP/3–21G*	−28.46	−18.86		
B3LYP/6–31G*	−15.76	−13.06	−14.3	
B3LYP/6–31G**	−15.98	−13.54	−14.0	
B3LYP/6–31++G**	−12.45	−13.03		
A–T(3)				−13.97
HF/3–21G*	−21.60	−15.86		
HF/6–31G*	−12.69	−11.69		
HF/6–31G **	−12.70	−11.86	−10.90	
B3LYP/3–21G*	−28.08	−21.04		
B3LYP/6–31G*	−17.01	−15.33	−16.0	
B3LYP/6–31G**	−17.13	−15.64	−15.0	
B3LYP/6–31++G **	−14.11	−14.80		
A–T(4)				−13.60
HF/3–21G*	−23.06	−15.10		
HF/6–31G*	−12.36	−10.66		
HF/6–31G **	−12.41	−10.90	−10.69	
B3LYP/3–21G*	−29.99	−20.14		
B3LYP/6–31G*	−16.76	−14.14	−15.2	
B3LYP/6–31G**	−16.97	−14.46	−14.9	
B3LYP/6–31++G**	−13.54	−14.19		
A–T(5)				−13.46
HF/3–21G*	−23.11	−15.18		
HF/6–31G*	−12.33	−10.64		
HF/6–31G **	−12.39	−10.88	−10.67	
B3LYP/3–21G*	−29.89	−20.02		
B3LYP/6–31G*	−16.52	−13.88	−15.0	
B3LYP/6–31G**	−16.76	−14.28	−14.7	
B3LYP/6–31++G**	−13.35	−13.92		
A–T(6)				−12.81
HF/3–21G*	−20.51	−15.08		
HF/6–31G*	−11.51	−10.66		
HF/6–31G **	−11.49	−10.83	−9.72	
B3LYP/3–21G*	−27.29	−20.33		
B3LYP/6–31G*	−15.92	−14.36	−14.9	
B3LYP/6–31G**	−15.97	−14.50	−13.8	
B3LYP/6–31++G**	−13.04	−13.82		

Table 2. (Contd.)

1	2	3	4	5
A–T(7)				−14.80
HF/3–21G*	−23.05	16.85		
HF/6–31G*	−13.19	−11.90		
HF/6–31G **	−13.31	−12.11	−11.69	
B3LYP/3–21G*	−29.16	−21.72		
B3LYP/6–31G*	−17.27	−15.25	−16.5	
B3LYP/6–31G **	−17.52	−15.65	−16.2	
B3LYP/6–31++G**	−14.48	−15.15		
A–T(8)				−14.43
HF/3–21G*	−23.0	16.75		
HF/6–31G*	−13.08	−11.79		
HF/6–31G **	−13.20	−12.04	−11.59	
B3LYP/3–21G*	−28.57	−21.11		
B3LYP/6–31G*	−16.78	−14.77	−16.0	
B3LYP/6–31G**	−17.00	−15.10	−15.7	
B3LYP/6–31++G **	−14.08	−14.71		
A–T(9)				−16.79
HF/3–21G*	−25.20	−19.40		
HF/6–31G*	−15.34	−14.32		
HF/6–31G **	−15.41	−14.54	−13.67	
B3LYP/3–21G*	−31.10	−24.38		
B3LYP/6–31G*	−19.39	−17.91	−18.5	
B3LYP/6–31G**	−19.59	−18.29	−18.2	
B3LYP/6–31++G**	−16.79	−17.26		

Note: ΔE and ΔE_{BSSE} are the bond energy at 0 K without and with correction for the BSSE, respectively. ΔH_{298} is the bond enthalpy at 298 K. ΔH_{exp} is experimental ΔH from spectrometry data [43] with corrections for A–T according to Brameld et al. [13]. For A–T (1) model $\Delta H_{exp} = -12.1$ kcal/mol. We assume that the ΔH_{exp} of the other forms of A–T (*n*) ($n \neq 1$) is near the $\overline{\Delta H}_{298}$ (average of theoretical ΔH_{298}) value that is obtained from theoretical calculations.

BSSEs of some cases are large (9.85 kcal/mol) and some are quite small (0.6 kcal/mol). Comparison of the hydrogen bond enthalpies (ΔH_{298}) from *ab initio* and DFT calculations is shown in Fig. 2.

The distance for (N2...H23–N19), (N8–H13...O24), and (C1...O18) for the Watson–Crick A–T base pair are 2.84, 2.94, and 3.63 Å, respectively, at B3LYP/*6-31G*** level, in good agreement with experimental results (2.82, 2.98, and 3.52 Å) [42]. It turns out that the DFT (B3LYP/*6-31G***) distances are in excellent agreement with the experimental data, whereas HF distances are somewhat too long (Table 1). For other models of A–T base pair result at different levels are summarized in Table 1.

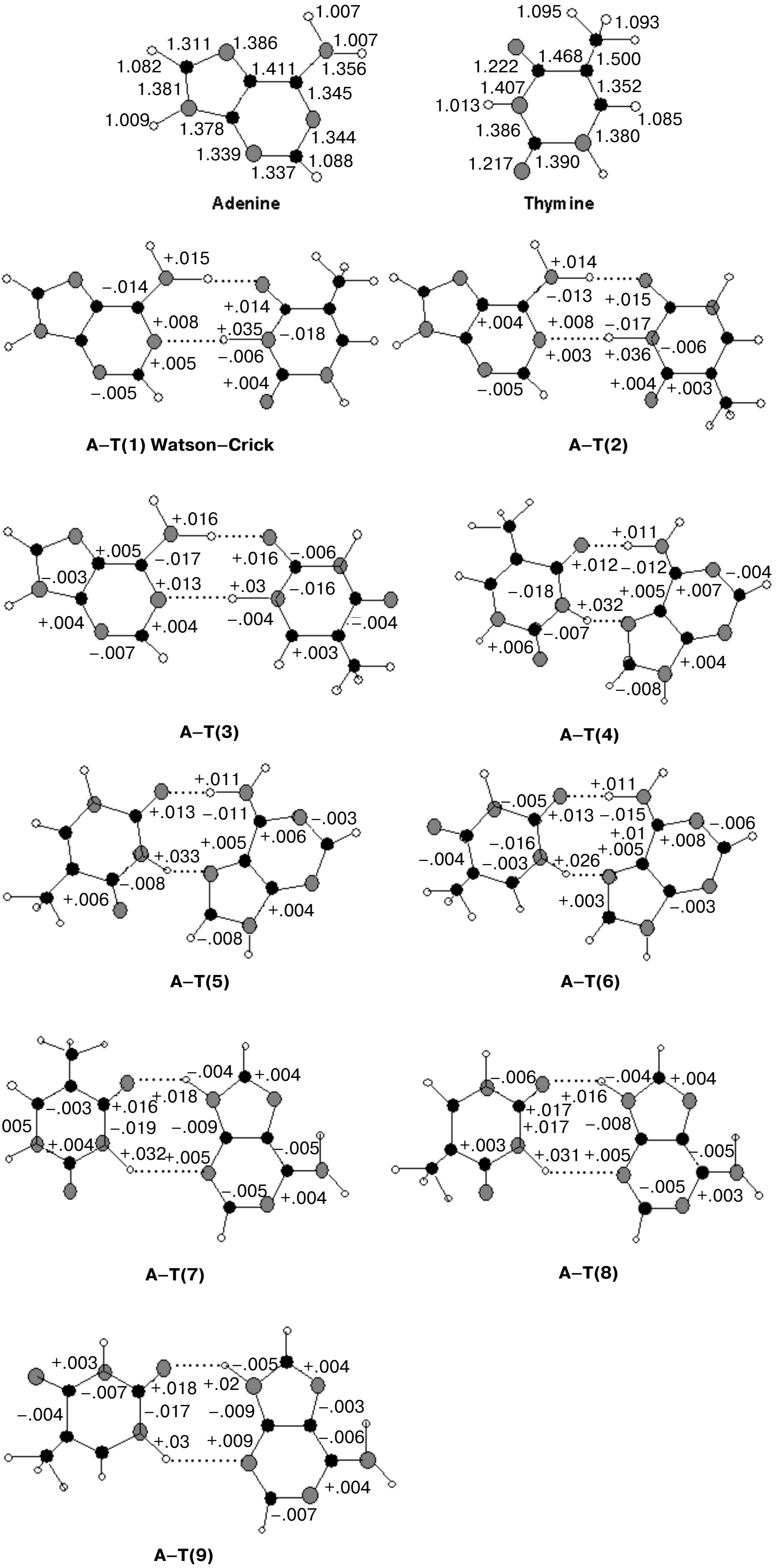
Therefore, we have to conclude that, at least for the hydrogen-bonded systems studied in this work, density functional calculations give similar or even better results than the conventional *ab initio* studies.

The interaction energy of model (9) A–T base pair is larger than others (−18.28 and −17.26 kcal/mol) and for model (2) is the smallest value (−13.53 and −13.03 kcal/mol) at B3LYP/*6-31G*** and B3LYP/*6-31++G*** levels, respectively. Comparison of the hydrogen bond energies, ΔE_{BSSE} and ΔE, from *ab initio* and DFT calculations is shown in Figs. 3 and 4.

The HF approach yields distances that are up to 0.15 Å longer than DFT values. In general, the intermolecular distances determined with density functional theory are somewhat shorter compared to HF optimizations.

The deformation of the bases (changes in bond lengths larger than 0.003 Å) caused by the formation of the hydrogen bonds is shown in Fig. 5. All the N–H bonds that participate in hydrogen bonding expand by 0.011-0.036 Å. The largest elongations are found for the N19–H23 of thymine in the A–T (model (2)) orientation and the smallest are found for N8–H12 of adenine. The C=O distances of oxygen atoms involved in hydrogen bonding increase by some 0.018 Å.

Atomic charges of A–T base pairs are shown in Fig. 6. The small interaction energy between A and T is also reflected in the electron distribution at the atoms involved in hydrogen bonding and the corresponding charges (Fig. 6). It turns out to be the case, as can be seen from Fig. 6



Adenine **Thymine**

A–T(1) Watson–Crick **A–T(2)**

A–T(3) **A–T(4)**

A–T(5) **A–T(6)**

A–T(7) **A–T(8)**

A–T(9)

Fig. 2. Comparison of the hydrogen bond enthalpies (ΔH_{298}, kcal/mol) of the different orientations of A–T base pair obtained from *ab initio* and DFT calculations.

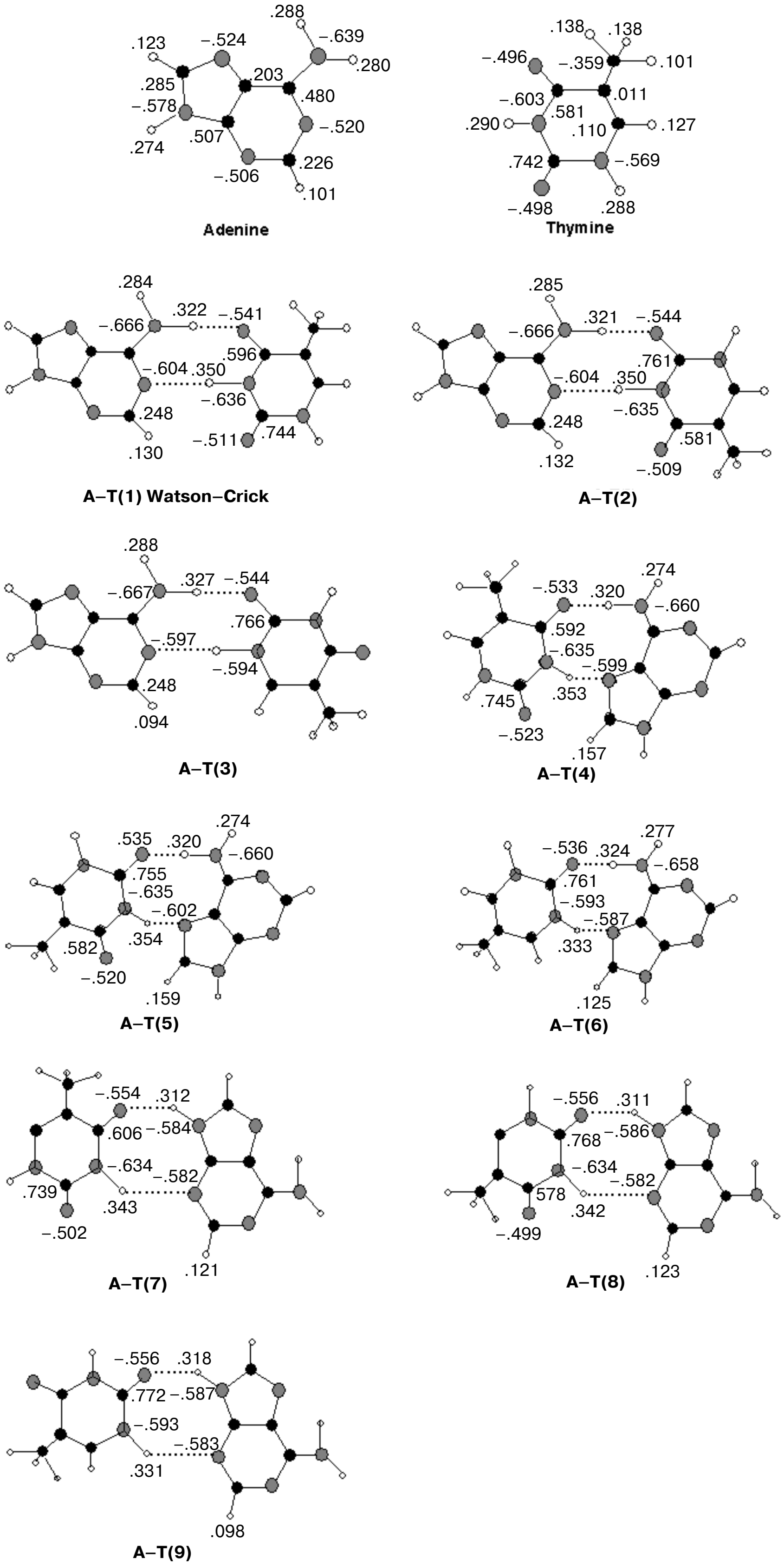


Fig. 3. Comparison of the hydrogen bond energies (ΔE_{BSSE}, kcal/mol) for different orientations of A–T base pair obtained from *ab initio* and DFT calculations.

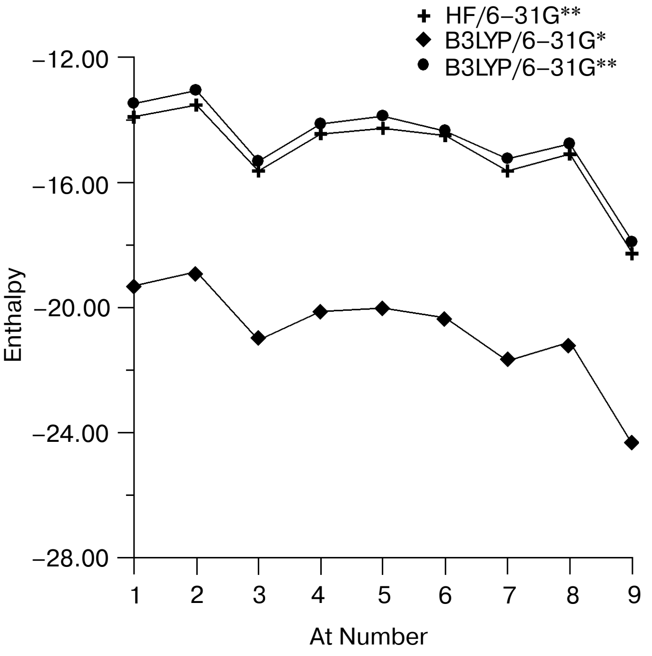


Fig. 4. Comparison of the hydrogen bond energies (ΔE, kcal/mol) for different orientations of A–T base pair obtained from *ab initio* and DFT calculations.

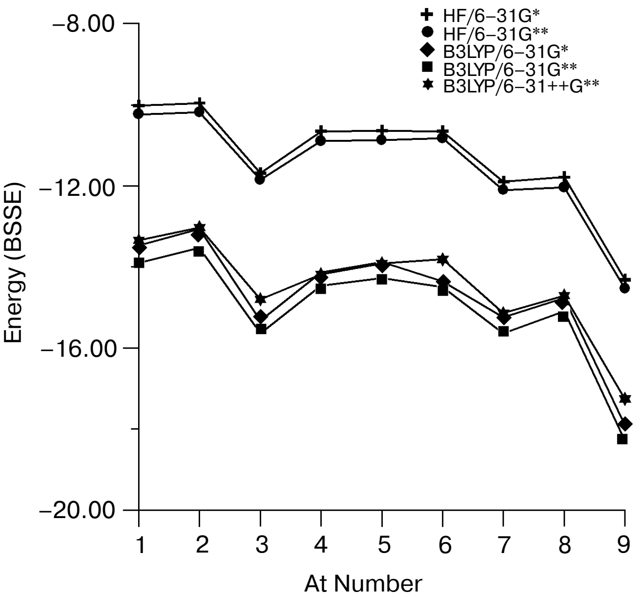


Fig. 5. Deformation (Å) of the individual bases caused by hydrogen bonding in the base pairs, from B3LYP/*6-31G*** optimizations without any symmetry constraint (only changes in bond length ≥0.003 Å are given).

which displays the atomic charges for the separate, noninteracting bases, that all proton acceptor atoms have negative charges, whereas the corresponding protons they face are all positively charged.

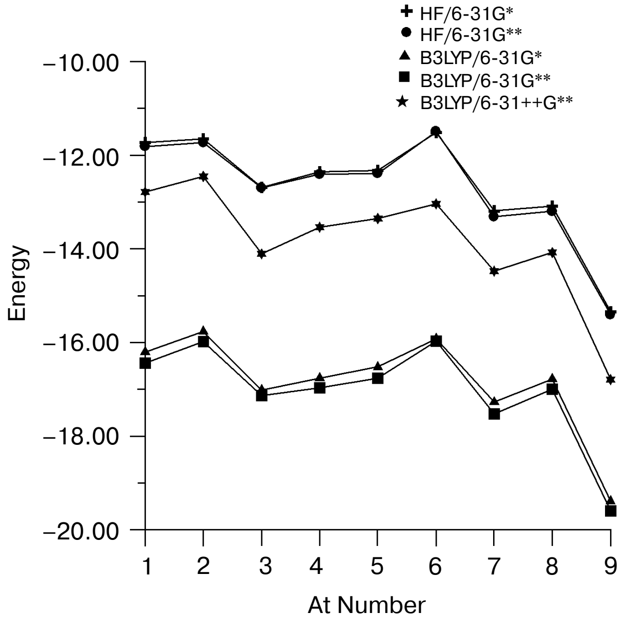


Fig. 6. Selected atomic charges of monomers and different orientation A–T base pairs from a Mulliken population analysis (B3LYP/*6-31G***).

The results obtained give the intermolecular (H14...O18) distance at the A–T (9) 1.811 Å in B3LYP/*6-31G*** level and show that with reduced interaction energies in different models of A–T the intermolecular (H...O) distance increases (Table 2).

We have unraveled a hitherto unresolved discrepancy between theoretical and experimental hydrogen bond lengths and energies in Watson–Crick and other different orientations. Our results clearly show that A–T base pairs select different geometries that vary in energy and structures. This difference is the reason for various hydrogen bonds between O, N, and H atoms. In general the results obtained in B3LYP/*6-31G*** are in good agreement with experimental values. It turns out that the B3LYP/*6-31G*** distances are in excellent agreement with the experimental data, whereas HF distances are somewhat too long.

REFERENCES

1. Jeffrey, G. A., and Saenger, W. (1991) *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Berlin-New York-Heidelberg.
2. Jeffrey, G. A. (1997) *An Introduction to Hydrogen Bonding*, Chap. 10, Oxford University Press, New York-Oxford.
3. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York-Berlin-Heidelberg-Tokyo.
4. Watson, J. D., and Crick, F. H. C. (1953) *Nature*, **171**, 737-738.

5. Lewis, J. P., and Sankey, O. F. (1995) *Biophys. J.*, **69**, 1068.
6. Kong, Y. S., Jhon, M. S., and Lowdin, P. O. (1987) *Int. J. Quantum Chem. Symp. QB*, **14**, 189.
7. Nagata, C., and Aida, M. (1988) *J. Mol. Struct.*, **179**, 451-466.
8. Gould, I. R., and Kollman, P. A. (1994) *J. Am. Chem. Soc.*, **116**, 2493-2499.
9. Sponer, J., Leszczynski, J., and Hobza, P. (1996) *J. Phys. Chem.*, **100**, 1965-1974.
10. Sponer, J., Leszczynski, J., and Hobza, P. (1996) *Biomol. Struct. Dyn.*, **14**, 117-135.
11. Sponer, J., Hobza, P., and Leszczynski, J. (1996) in *Computational Chemistry. Reviews of Current Trends* (Leszczynski, J., ed.) World Scientific Publisher, Singapore, pp. 185-218.
12. Hutter, M., and Clark, T. (1996) *J. Am. Chem. Soc.*, **118**, 7574-7577.
13. Brameld, K., Dasgupta, S., and Goddard, W. A. (1997) *J. Phys. Chem. B*, **101**, 4851-4859.
14. Meyer, M., and Suhnel, J. (1997) *J. Biomol. Struct. Dyn.*, **15**, 619-624.
15. Santamaria, R., and Vazquez, A. (1994) *J. Comp. Chem.*, **15**, 981-996.
16. Bertran, J., Olivia, A., Rodriguez Santiago, L., and Sodupe, M. (1998) *J. Am. Chem. Soc.*, **120**, 8159-8167.
17. Sim, F., St.-Amant, A., Papai, I., and Salahub, D. R. (1992) *J. Am. Chem. Soc.*, **114**, 4391-4400.
18. Guo, H., Sirois, S., Proynov, E. I., and Salahub, D. R. (1997) in *Theoretical Treatment of Hydrogen Bonding* (Hadzi, D., ed.) Wiley, New York.
19. Sirois, S., Proynov, E. I., Nguyen, D. T., and Salahub, D. R. (1997) *J. Chem. Phys.*, **107**, 6770-6781.
20. Rablen, P. R., Lockman, J. W., and Jorgenson, W. L. (1998) *J. Phys. Chem.*, **102**, 3782-3797.
21. Kim, K., and Jordan, K. D. (1994) *J. Phys. Chem.*, **98**, 10089-10094.
22. Novoa, J. J., and Sosa, C. (1995) *J. Phys. Chem.*, **99**, 15837-15845.
23. Latajka, Z., and Bouteiller, Y. (1994) *J. Chem. Phys.*, **101**, 9793-9799.
24. Del Bene, J. E., Person, W. B., and Szczepaniak, K. (1995) *J. Phys. Chem.*, **99**, 10705-10707.
25. Florian, J., and Johnson, B. G. (1995) *J. Phys. Chem.*, **99**, 5899-5908.
26. Combariza, J. E., and Kestner, N. R. (1995) *J. Phys. Chem.*, **99**, 2717-2723.
27. Civalleri, B., Garrone, E., and Ugliengo, P. (1997) *J. Mol. Struct.*, **419**, 227-238.
28. Lozynski, M., Rusinska Roszak, D., and Mack, H. G. (1998) *J. Phys. Chem.*, **102**, 2899.
29. Chandra, A. K., and Nguyen, M. (1998) *J. Chem. Phys.*, **223**, 299-306.
30. Paizs, B., and Suhai, S. (1998) *J. Comp. Chem.*, **19**, 575-584.
31. McAllister, M. A. (1998) *J. Mol. Struct.*, **427**, 39-53.
32. Pan, Y. P., and McAllister, M. A. (1998) *J. Mol. Struct.*, **427**, 221-227.
33. Gonzalez, L., Mo, O., and Yanez, M. (1997) *J. Comp. Chem.*, **18**, 1124-1135.
34. Guerra, C. F., and Bickelhaupt, F. M. (1999) *Angew. Chem.*, **111**, 3120-3122.
35. Guerra, C. F., and Bickelhaupt, F. M. (1999) *Angew. Chem. Int. Ed.*, **38**, 2942-2945.
36. Guerra, C. F., Bickelhaupt, F. M., Snijders, J. G., and Baerends, E. J. (2000) *J. Am. Chem. Soc.*, 4117.
37. Guerra, C. F., Bickelhaupt, F. M., and Snijders, J. G. (1999) *Chem. Eur. J.*, **5**, 3581-3594.
38. Maranon, J., Fantoni, A. F., and Grigera, J. R. (1999) *J. Mol. Liq.*, **79**, 177-186.
39. Gaussian 94 Package, Gaussian, Inc., Pittsburgh, PA 1506, USA.
40. Gaussian 98 Package, Gaussian, Inc., Pittsburgh, PA 1506, USA.
41. Boys, S. F., and Bernardi, F. (1970) *Mol. Phys.*, **19**, 553-566.
42. Lindauer, K., Bendic, C., and Suhnel, J. (1996) *Compute. Appl. Biosci.*, **12**, 281-289.
43. Yanson, I. K., Teplitsky, A. B., and Sukhodub, L. F. (1979) *Biopolymers*, **18**, 1149-1170.