Thioguanine and Thiouracil: Hydrogen-Bonding and Stacking Properties

J. Šponer,*,†,‡,§ J. Leszczynski,§ and P. Hobza†

J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague, Czech Republic, Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic, and Department of Chemistry, Jackson State University, Jackson, Mississippi 39217

Received: June 23, 1997; In Final Form: September 24, 1997[⊗]

Base stacking and H-bonding properties of thioguanine and thiouracils were studied using an ab initio quantum chemical method with inclusion of electron correlation (second-order Møller−Plesset perturbational method). Hydrogen-bonded base pairs containing thiobases are only slightly less stable (up to 2 kcal/mol) than the unmodified base pairs. The N···S distances are larger by 0.4−0.7 Å compared to the N···O distances in the standard base pairs. The thio group enhances polarizability of the monomers and their dipole moments. Thus, in stacked complexes of thiobases, both dispersion attraction and electrostatic interactions are enhanced. Mutual contact of the sulfur atoms and their interaction with second-row elements lead to steric clashes destabilizing the stacking, though, in DNA, such clashes should be eliminated by rather small adjustments of the local DNA conformation. The thio group significantly destabilizes the hydration of the 6-position of thioguanine with respect to guanine. The first hydration shell in the major groove might be significantly altered by thioguanine.

I. Introduction

Since 1994, an extensive high-level quantum-chemical analysis (ab initio method with inclusion of electron correlation) has been carried out on neutral dimers of nucleic acid bases.[1−11] These studies provided a rather complete picture of base−base interactions which could not be obtained by any other experimental or theoretical procedure and were recently reviewed.[6,7,9] Here we analyze H-bonding and stacking properties of two chemically modified analogues of DNA bases: thioguanine and thiouracils.

Chemically modified bases, such as 6-thioguanine, 2-thiouracil, 4-thiouracil, and 2,4-dithiouracil (designated ^{6S}G, ^{2S}U, ^{4S}U, ^{2,4S}U, cf. Figure 1), are frequently studied for their numerous pharmacological, biochemical, and biological capabilities (see refs 12−21 and references therein). The thiobases have the same distribution of hydrogen donors and acceptors as the standard bases. However, the sulfur atom may induce changes in the properties of bases and their interactions. Tautomeric equlibria of thiobases and selenobases have been studied by theoretical chemists.[22−24] Thiobases were also studied by experimental techniques;[25−29] for a review see ref 29. Interactions of thiobases were analyzed in several older quantum-chemical studies, mostly of a semiempirical nature.[30−33] The present paper provides the first high-level ab initio characterization of H-bonding and stacking properties of thiobases. We believe that although the tautomeric equilibria are of interest, the modified biochemical activity of thiobases is due to their altered molecular interactions.

Thiobases influence the structure of DNA, though the picture of such changes is not known at the molecular level. The partial incorporation of deoxy-^{6S}G is effective in inhibiting the formation of G tetrads in guanine-rich oligodeoxyribonucleotides.[14] On the other hand, thioguanine does not destabilize the formation of G.GC triple helixes, which indicates that the H-bonding involving thioguanine should not differ from guanine-containing complexes dramatically.[14] Different interactions of metal cations with guanine and thioguanine are known.[21,33] The crystal structure of 6-thioguanine reveals clear similarities with guanine crystal despite that thioguanine crystallizes in its 7H-tautomer form.[15] The H-bond distances involving the sulfur atom are increased by about 0.4 Å.

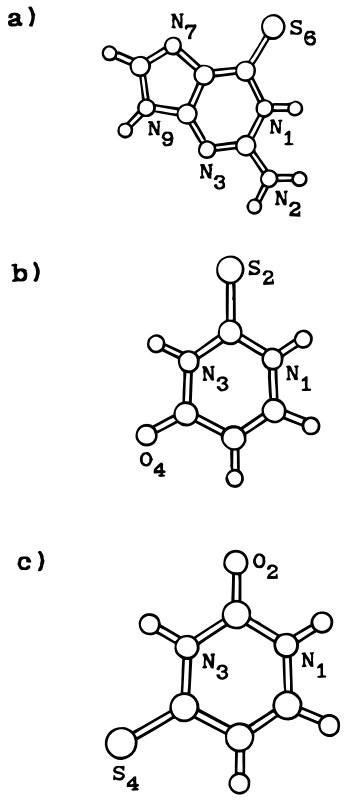

Figure 1. Molecular structure of (a) 6-thioguanine, (b) 2-thiouracil, and (c) 4-thiouracil.

II. Method

II.1. H-Bonded Base Pairs. Geometries of H-bonded base pairs were optimized within the Hartree−Fock (HF) approxima-

* Corresponding author, J. Heyrovský Institute of Physical Chemistry, Academy of Sciences, Prague. Fax: 420 2 858 2307. e-mail: sponer@indy.jh-inst.cas.cz

† Institute of Physical Chemistry.

‡ Institute of Biophysics.

§ Jackson State University.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S1089-5639(97)02040-9 CCC: $14.00

tion with the standard 6-31G* basis set using the gradient optimization procedure under C_s symmetry. The nature of planar optimized structures was determined by harmonic vibrational analysis. Nonplanar geometries were obtained for pairs where the planar optimized structure exhibited negative eigenvalues of the Hessian matrix.[2,5] Interaction energies were evaluated for planar optimized structures using the second-order Møller−Plesset perturbational method (MP2) with the 6-31G basis set augmented by diffuse d-polarization functions (with an exponent of 0.2 on the sulfur atom and 0.25 on the C, N, and O atoms; designated as 6-31G*(0.25)).[2,3] MP2/6-311G-(2df,p)//HF/6-31G** calculations were carried out to compare the interaction between a water molecule and 6-thioguanine and guanine.

II.2. Stacked base pairs. Base-stacking energies were evaluated at the MP2/6-31G*(0.25) level for a set of intermolecular geometries without relaxation of the monomers (single-point calculations). Geometries of monomers were optimized at the HF/6-31G* level within C_s symmetry.

One of the referees pointed out that the use of planar bases in base-stacking calculations ignores the amino group nonplanarity. Let us explain this point. The isolated nucleobases are nonplanar in their amino groups, and the amino group geometry is very sensitive to the intermolecular interactions.[5,7,34] To study these effects would require carrying out gradient optimizations. In contrast to the H-bonded pairs, the stacked pairs must be optimized at the MP2 level. We have recently carried out MP2 gradient optimization on several pyrimidine dimers (Hobza, P.; Šponer, J., unpublished data). Nevertheless, there are the following reasons to prefer the single-point technique in studies of base stacking.

(i) The gradient optimization of stacked base pairs revealed out-of-plane interactions involving the amino group hydrogen atoms and also some other deformations. On the other hand, the predicted stabilization energies basically agreed with estimates obtained by the usual single-point search. The reason is the mutual compensation of intramolecular and intermolecular contributions.

(ii) MP2 gradient optimization is exceptionally demanding and guanine dimer is still beyond our computer facilities.

(iii) Gradient optimization does not allow the characterization of the conformational space of stacked dimers. In contrast to H-bonded base pairs, the observed base-stacking arrangements are exceptionally variable and do not correspond to optimal structures predicted for isolated stacked dimers. In addition, no stable stacked structure was found for some dimers due to a transition to more stable H-bonded base pairs.

(iv) Gradient optimization is not corrected for the basis set superposition error (BSSE). The sum of BSSE from the HF and MP2 levels is rather large and can cause a deterioriation of the potential energy surface.

(v) The nucleobases are involved in H-bonded base pairs and surrounded by the adjacent base pairs in nucleic acids. These interactions eliminate most of the nonplanarities of monomers, so that the use of C_s symmetry for monomers is basically justified.

The standard counterpoise procedure has been applied in all energy calculations to eliminate the basis set superposition error. All orbitals of the ghost system were considered.[35] All calculations were done using the Gaussian94 set of programs.[36]

II.3. Reliability of Calculations. Benchmark calculations were recently reported for pyrimidine DNA base pairs and other H-bonded and stacked van der Waals clusters.[37] The calculations were carried out using the coupled cluster method with noniterative triple excitations (CCSD(T)) with a diffuse medium-sized polarized basis set and the MP2 level with larger basis sets (up to the aug-cc-pVTZ one). The MP2/6-31G*(0.25) base-stacking energies seem to be very close to the actual values, while the energies of H-bonded base pairs evaluated by the same method are probably underestimated by 1−2 kcal/mol.[37] Diffuse d-polarization functions are strongly required to cover the dispersion attraction for stacking interactions.[3,38,39] Recently, Cybulski et al.[40] reported reference calculations on polarizabilities of isolated bases and suggested that the MP2 calculations with medium-sized basis sets significantly underestimate base stacking. This conclusion is not correct for two reasons. First, Cybulski et al. did not consider medium-sized basis sets with diffuse polarization functions. Second, the MP2 theory itself overestimates the correlation stabilization for aromatic stacking (with respect to CCSD(T)).[37] Thus, the large basis sets recommended by Cybulski et al. would lead, at the MP2 level, to overestimation of aromatic stacking.

TABLE 1: Polarizability (α, au) and Dipole Moment (μ, D) of Selected Bases[a]

base	α	μ
6-thioguanine	103	7.8(8.5)
2,4-dithiouracil	98	4.9(5.7)
guanine	84	6.5(7.3)
4-thiouracil	78	4.8(5.6)
2-thiouracil	75	4.6(5.3)
inosine	72	4.7(5.6)
thymine	64	3.7(4.3)
cytosine	63	6.2(7.3)
uracil	56	3.9(4.7)

[a] Polarizability has been evaluated at the HF/6-31G*(0.25) level, the dipole moment at the MP2/6-31G*(0.25) level; the values in parentheses correspond to the HF dipole moment. For further data on electric properties of isolated nucleobases see refs 2, 3, 40−42.

III. Results and Discussion.

III.1. Isolated Bases. Table 1 compares the dipole moments and Hartree−Fock polarizabilities (relative values of polarizabilities should be considered) of thiobases and standard bases. The thiobases possess larger dipole moments than the standard bases, though the direction of the dipole moment is not changed (not shown). This means that the electrostatic dipole−dipole interaction in stacked and H-bonded complexes of thiobases will be enhanced. Due to the larger atomic radius of the sulfur atom, the exchange-repulsion is larger as well. This may cause steric problems in some configurations allowed for oxobases. Finally, thiobases are also characterized by increased vertical and total molecular polarizabilities. Therefore, increased intermolecular electron correlation stabilization is expected. The electronegativity of oxygen is larger than that of sulfur, which results in reduced polarity of the C═S bond.[32,33]

We have tried to rationalize the properties of thiobases using Mulliken population analysis. With the 6-31G*(0.25) basis set, the atomic charge on the sulfur atom is significantly more negative compared to the corresponding oxygen atom by ca. 0.3−0.4*e*. Accordingly, the carbon atom of the C═S bond is more positive with respect to the C═O bond by 0.3−0.6*e*; and the neighboring ring nitrogen atoms are more negative. However, when using the standard 6-31G* basis set, a fully opposite result was found. Both basis sets provide almost identical multipole moments of bases and potential-derived charges. Therefore, the electrostatic component of the interaction energy is not changed by the diffuse polarization functions, while the Mulliken populational analysis is strongly basis set dependent and cannot be used.

III.2. H-Bonding Properties of Thiobases. We analyzed interactions in six H-bonded base pairs: ^{2S}U·^{2S}U1, A·^{4S}U WC,

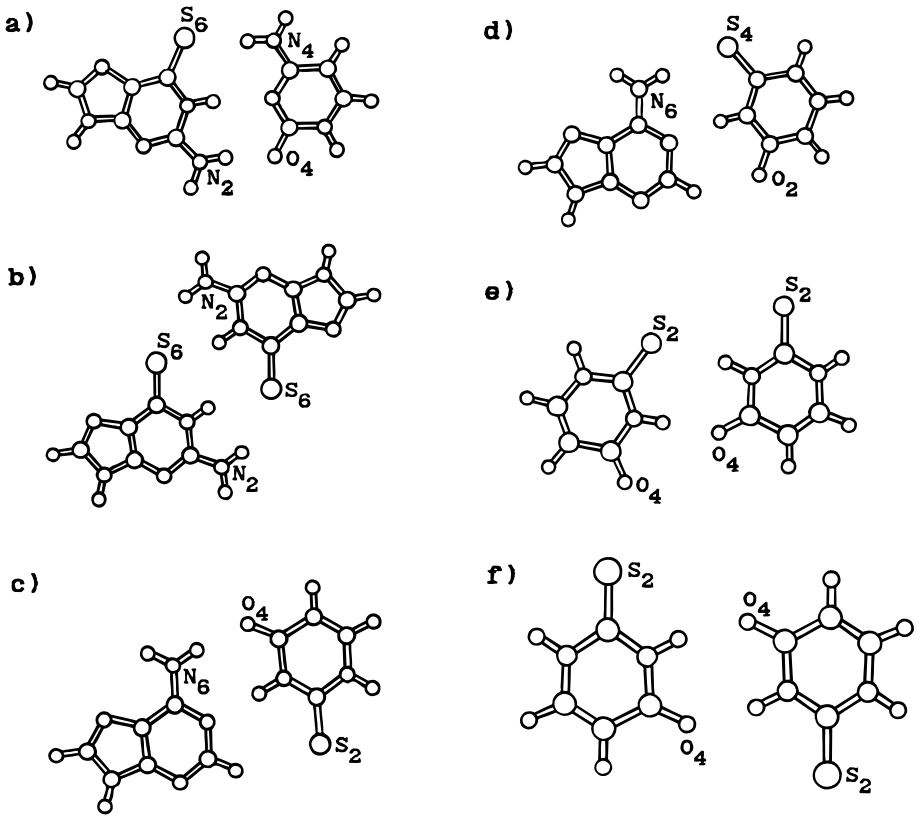

Figure 2. Molecular structure of H-bonded base pairs investigated in the present study: (a) ^{6S}G·C WC, (b) ^{6S}G·^{6S}G1, (c) A·^{2S}U WC, (d) A·^{4S}U WC, (e) ^{2S}U·^{2S}U1, (f) ^{2S}U·^{2S}U2.

TABLE 2: Interaction Characteristics of Planar H-Bonded Base Pairs Containing Thiobases and the Standard Base Pairs[2] (in Parentheses)[d]

base pair	ΔE^{HF}	ΔE^{COR}	ΔE^{DEF}	ΔE^{T}	ΔZPE	$\Delta H_0 0$	n_i	μ
^{6S}G·^{6S}G1	−19.3(−25.1)	−3.0(0.4)	2.4(3.2)	−19.9(−21.0)	1.0[a,b]	−18.9[b,c]	11,33,34	7.5[a](0.0)
^{6S}G·CWC	−23.1(−24.6)	−1.9(−1.2)	2.5(2.4)	−22.5(−23.4)	1.7(1.9)	−20.8(−21.5)	12,22,58	7.4(6.5)
A·^{4S}UWC	−8.4(−9.7)	−3.4(−2.7)	0.6(0.6)	−11.2(−11.8)	0.1(1.3)	−11.1(−10.5)	15,22,59	3.4(2.0)
A·^{2S}UWC	−9.6(−9.7)	−3.2(−2.7)	0.7(0.6)	−12.1(−11.8)	1.2(1.3)	−10.9(−10.5)	18,22,60	3.2(2.0)
^{2S}U·^{2S}U1	−6.9(−9.3)	−2.4(−1.3)	0.5(0.6)	−8.8(−10.0)	0.7(0.8)	−8.1(−9.2)	13,27,46	2.5(1.3)
^{2S}U·^{2S}U2	−8.7(−9.3)	−1.5(−1.3)	0.6(0.6)	−9.6(−9.9)	0.9(0.9)	−8.7(−9.1)	15,28,56	0.0(0.0)

[a] Evaluated for the nonplanar optimized structure. [b] ΔZPE for GG base pair cannot be evaluated from harmonic vibrational analysis because of strong anharmonicity.[43] [c] Estimated using interaction energy for planar pair, ΔZPE, and the HF/6-31G* energy difference between nonplanar and planar pair. [d] ΔE^{HF}, Hartree−Fock interaction energy; ΔE^{COR}, correlation contribution to the interaction energy; ΔE^{DEF}, deformation energy of bases with respect to optimized nonplanar monomers; $\Delta E^{T} = \Delta E^{HF} + \Delta E^{COR} + \Delta E^{DEF}$, total stabilization energy; ΔZPE, zero-point energy contribution; $\Delta H_0{}^0$, interaction enthalpy at 0 K evaluated within harmonic approximation; n_i, three lowest harmonic vibrational frequencies in cm^{-1}; μ, dipole moment (in D). The base pairs were optimized at the HF/6-31G* level under C_S symmetry; interaction energies were evaluated at the MP2/6-31G*(0.25) level; all data are in kcal/mol.

^{6S}G·C WC, ^{6S}G·^{6S}G1, A·^{2S}U WC, and ^{2S}U·^{2S}U2 (Figure 2). The designation of pairs (1,2,WC) is taken from the previous studies;[2] the superscript shows the position of the sulfur atoms. Table 2 shows the energy characteristics of the planar base pairs including the dipole moments and the three lowest intermolecular vibrational modes.[43] Table 3 summarizes the H-bond distances and some other interatomic distances, the so-called secondary interactions.[2,5,44] Comparison is made with the data obtained previously for the corresponding unmodified pairs at the MP2/6-31G*(0.25)//HF/6-31G** level.[2]

If the sulfur atom does not participate in the pairing (A·^{2S}U WC, ^{2S}U·^{2S}U2), the interaction characteristics and geometries of modified and standard base pairs are very similar. The small difference observed for the A·^{2S}U WC base pair with respect to the A·T WC one is due to the secondary interaction between the sulfur atom and the H2(C2) group of adenine. The pair is slightly opened toward the minor groove side, where the contact occurs and is about 0.5 kcal/mol more stable than its classical analogue. This is due to the increased dispersion attraction. When the sulfur atom participates in the hydrogen bond, the H-bond length increases by about 0.5−0.7 Å and the pairs are weaker than the parent structures by about 1−2 kcal/mol. The reduction in the Hartree−Fock intermolecular stabilization is compensated for by increasing dispersion stabilization.

Formation of the C═O···H−N hydrogen bond is accompanied by a nonnegligible prolongation of both C═O and N−H covalent bonds. (Formation of Y...H−X H-bonds is spectroscopically detected by a red shift of the X−H stretching frequency of the proton donor.) Similar changes of covalent bonds are observed also when a C═S···H−N bond is formed. However, there are differences with respect to C═O···H−N interactions. First, the prolongation of the N−H bond is reduced. Second, the prolongation of the C═S double bond is larger compared to the prolongation of C═O bonds. For example, the equilibrium N3−H3 distance of 2-thiouracil increases by 0.012 Å in N−H···O═C bonds of ^{2S}U·^{2S}U1 and ^{2S}U·^{2S}U2 complexes (from ca. 0.996 Å to ca. 1.008 Å), while this prolongation is only 0.010 Å for the N−H···S═C bond

TABLE 3: H-Bond Distances and Some Other Interatomic Distances (in Å) in Planar H-Bonded Base Pairs Containing Thiobases and the Parent Unmodified Pairs (Ref 2)

base pair	interaction X(H)···Y or H···Y	X···Y/XHY or H···Y, modified base pair	X···Y/XHY or H···Y, standard base pair
^{6S}G·^{6S}G1[a]	N1(H)···S6	3.57/163.3	2.87/178.1
	S6···H(N1)	3.57/163.3	2.87/178.1
	H2···S6	2.79	2.63
	S6···H2	2.79	2.63
^{6S}G·^{6S}G1[b]	N1(H)···S6	3.69/150.8	
	S6···H(N1)	3.69/150.8	
	H2···S6	2.74	
	S6···H2	2.74	
^{6S}G·CWC	N2(H)···O2	2.90/172.2	3.02/178.1
	N1(H)···N3	3.30/173.3	3.04/176.1
	S6···(H)N4	3.46/177.0	2.92/177.0
A·^{2S}UWC	N6(H)···O4	3.03/172.3	3.09/172.2
	N1···(H)N3	3.08/174.4	2.99/178.8
	H2···S2	3.11	2.96
A·^{4S}UWC	N6(H)···S4	3.66/173.7	3.09/172.2
	N1···(H)N3	3.08/170.3	2.99/178.8
	H2···O2	2.50	2.96
^{2S}U·^{2S}U1	N3(H)···O4	2.98/179.3	2.98/162.2
	S2···H(N3)	3.63/166.5	2.97/167.3
^{2S}U·^{2S}U2	O4···(H)N3	2.97/164.4	2.98/167.4
	N3(H)···O4	2.97/164.4	2.98/167.4

[a] Planar base pair. [b] Nonplanar base pair.

the in ^{2S}U·^{2S}U1 pair. On the other hand, the equilibrium C2═S2 distance (equilibrium value of 1.664 Å) is enhanced by about 0.0135 Å, while the C═O bonds are longer only by about 0.007−0.008 Å. The C4═S4 bond in A·^{4S}U WC base pair is longer by 0.011 Å, the C4═O4 bond in the A·^{2S}U WC base pair by 0.008 Å. Similar trends can be observed for the ^{6S}G·C WC and ^{6S}G·^{6S}G1 base pairs compared with the unmodified base pairs. The C═S bond length increases by 0.022 and 0.023 Å for ^{6S}G·C WC and ^{6S}G·^{6S}G1 base pairs, respectively; the corresponding values of C═O bond prolongations are 0.016 and 0.021 Å. The N3−H3 bond is prolonged by 0.012 Å in G·G1 and by 0.008 Å in ^{6S}G·^{6S}G1. Similar prolongations were found for the cytosine N4−H4 bonds, 0.014 Å in the GC WC base pair, and 0.011 Å in the ^{6S}G·C WC one. The N6−H6 distance is longer by 0.0045 Å (N6−H6...S4−C4, A·^{4S}U WC) and 0.006 Å (N6−H6...O4−C4, A·^{2S}U WC).

The weak (secondary) interaction between the C═X6 groups and the amino group in G·G1 types of base pairs is accompanied by a prolongation of the N2−H2 bond, by 0.002 Å in ^{6S}G·^{6S}G1 and by 0.004 Å in G·G1. Another weak secondary interaction is the interaction between the adenine C2−H2 group and the C2═X2 group of thiouracils on the minor groove side of the Watson−Crick adenine···thiouracil pairs. The C2═O2 distance in the A·^{4S}U WC base pair is increased by 0.002 Å, while the C2═S2 distance in the A·^{2S}U WC base pair is increased by about 0.0035 Å. However, the C−H bond is shorter in both pairs by about 0.002 Å, compared to the isolated base. Therefore, this interaction should not been considered a weak H-bond. The shortening of the C−H bonds is very likely due to the short-range repulsion between the hydrogen and oxygen (sulfur) atoms. It has been proposed that many C−H···O contacts are stabilized by the interaction between the two second-row elements, and the hydrogen itself is rather destabilizing.[45] The concept of attractive structure-making C−H···O hydrogen bonds is starting to be very popular in structural biology.[46−48] In many cases the C−H groups can form real weak H-bonds (typically sp and sp^2 carbons, or sp^3 carbons with some substituents; for a review see refs 47 and 48). Nevertheless, definitely not all close C−H···O contacts can be interpreted as H-bonds. This is the case of the presently considered C−H···O(S) contacts between bases and probably many weak interactions in biopolymers where dispersion attraction dominates.[9]

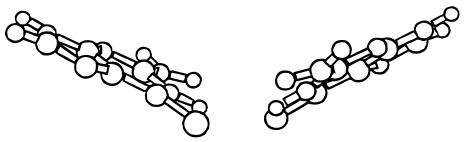

Figure 3. Nonplanar optimized (HF/6-31G*) geometry of the H-bonded ^{6S}G·^{6S}G1 base pair.

TABLE 4: Base Stacking in (2-Thiouracil)$_2$/(Uracil)$_2$ Complexes (Undisplaced Structures[a]) Evaluated at the MP2/6-31G*(0.25) Level[b]

twist (deg)	VSEP	ΔE^{HF}	ΔE^{COR}	ΔE^{MP2}
0	3.0	41.50/26.97	−24.24/−18.30	17.27/8.67
	3.2	25.77/16.60	−18.47/−13.91	7.30/2.70
	3.4	16.64/10.73	−14.09/−10.61	2.55/0.13
	3.6	11.29/7.34	−10.78/−8.11	0.51/−0.77
	3.8	8.08/5.35	−8.28/−6.25	−0.21/−0.91
	4.0	6.10/−	−6.41/−	−0.31/−
30	3.4	10.15/−	−12.27/−	−2.12/−
60	3.4	3.76/2.48	−10.21/−7.79	−6.45/−5.32
90	3.4	1.51/−	−9.59/−	−8.08/−
120	3.4	2.78/2.16	−10.06/−7.54	−7.28/−5.38
150	3.4	5.46/−	−10.89/−	−5.43/−
180	3.0	19.94/9.69	−19.26/−13.45	0.68/−3.76
	3.2	10.71/4.17	−14.54/−9.97	−3.83/−5.80
	3.3	7.74/2.51	−12.62/−8.56	−4.88/−6.05
	3.4	5.51/1.34	−10.94/−7.35	−5.43/−6.01
	3.5	3.85/0.55	−9.48/−6.31	−5.64/−5.76
	3.6	2.36/−0.05	−8.21/−5.41	−5.85/−5.46
	3.7	1.67/−	−7.11/−	−5.43/−
	3.8	0.99/−0.69	−6.15/−3.97	−5.16/−4.66

[a] Centers of mass stacked directly one above the other (cf. ref 3). [b] Data for (uracil)$_2$ are taken from ref 3. All energies are in kcal/mol. ΔE^{HF}, HF component of the interaction energy; ΔE^{COR}, correlation contribution to the interaction energy; ΔE^{MP2}, total MP2 interaction energy; VSEP (Å), vertical separation of bases.

The harmonic vibrational characteristics of modified and standard base pairs are similar. The only difference was found for the ^{6S}G·^{6S}G1 base pair, which is intrinsically nonplanar in contrast to the G·G1 base pair. We have optimized the nonplanar structure for the ^{6S}G·^{6S}G1 base pair; the energy difference (HF/6-31G*) between the planar and the nonplanar structure is less than 0.1 kcal/mol, although the base pair is buckled significantly (see Figure 3). The sulfur atoms interact with two hydrogen atoms, which makes the interaction rather different from the standard base pair. Let us recall that even the G·G1 base pair is known to be very flexible toward buckled and propeller-twisted structures, with the strongly anharmonic lowest buckle vibrational mode.[43] The G·G1 type of pairing is characterized by a delicate balance between contributions stabilizing the planar structure (primary H-bonds, attractive secondary interactions) and destabilizing it (amino group nonplanarity, repulsive secondary interactions).[2,5]

III.3. Stacking Properties of Thiobases. Tables 4 and 5 compare interactions in (6-thioguanine)$_2$ and (2-thiouracil)$_2$ stacked dimers with the parent dimers (guanine)$_2$ and (uracil)$_2$. The parallel undisplaced (2-thiouracil)$_2$ dimer is less stable than (uracil)$_2$ for vertical separation of monomers below 3.4 Å, due to the steric repulsion of the sulfur atoms. (Uracil)$_2$ and (2-thiouracil)$_2$ are of a similar stability for larger base···base separations. A similar trend was found for the parallel undisplaced (6-thioguanine)$_2$.

Also the antiparallel undisplaced (2-thiouracil)$_2$ is less stable than (uracil)$_2$. This is again due to the bulky sulfur atoms which interact with the second-row ring atoms. The optimal vertical

TABLE 5: Base Stacking in Undisplaced (6-Thioguanine)$_2$/(Guanine)$_2$ Complexes Evaluated at the MP2/6-31G*(0.25) Level[a]

twist (deg)	VSEP	ΔE^{HF}	ΔE^{COR}	ΔE^{MP2}
0	3.0	52.68/38.41	−34.27/−27.37	18.41/11.04
	3.2	33.01/23.78	−26.19/−20.80	6.82/2.97
	3.4	21.80/15.60	−20.06/−15.74	1.74/−0.14
	3.6	15.32/10.93	−15.42/−12.02	−0.10/−1.09
	3.8	11.47/8.20	−11.91/−9.24	−0.44/−1.04
30	3.4	13.24/9.68	−17.88/−14.19	−4.64/−4.51
60	3.4	6.59/4.83	−15.84/−12.52	−9.25/−7.69
90	3.4	4.35/1.74	−15.96/−11.91	−11.61/−10.17
120	3.4	3.95/1.27	−16.02/−11.85	−12.05/−10.58
150	3.4	3.23/−	15.24/−	−12.72/−
180	3.0	15.42/15.69	−25.87/−21.93	−10.45/−6.24
	3.1	10.35/−	−22.47/−	−12.12/−
	3.2	6.63/7.52	−19.50/−16.37	−12.87/−8.85
	3.3	3.91/5.01	−16.91/−14.13	−13.00/−9.12
	3.4	1.92/3.22	−14.65/−12.19	−12.72/−8.97
	3.6	−0.52/1.01	−10.98/−9.06	−11.49/−8.05
	3.8	−1.75/−	−8.22/−	−9.97/−
	4.0	−2.33/−0.68	−6.16/−5.02	−8.49/−5.70
	4.2	−2.56/−	−4.63/−	−7.19/−

[a] Data for (guanine)$_2$ are taken from ref 3. All energies are in kcal/mol. ΔE^{HF}, HF component of the interaction energy; ΔE^{COR}, correlation contribution to the interaction energy, ΔE^{MP2}, total MP2 interaction energy; VSEP, vertical separation of bases (Å).

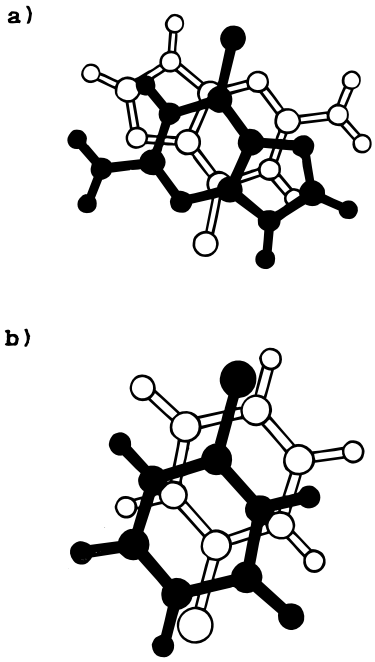

Figure 4. Antiparallel undisplaced (6-thioguanine)$_2$ and (2-thiouracil)$_2$ dimers.

separation of monomers is thus 3.6 Å for (2-thiouracil)$_2$ and 3.3 Å for (uracil)$_2$. This repulsion is eliminated by a mutual displacement of bases in such a direction that the sulfur atoms do not interact with rings. Figure 5 shows such a geometry with a displacement of 1.0 Å. Here, the optimal vertical separation of bases decreases to the usual value of 3.3 Å, and the dimer is remarkably stable, −7.0 kcal/mol.

Antiparallel undisplaced (6-thioguanine)$_2$ dimer differs from (2-thiouracil)$_2$ in that the sulfur atoms do not interact with the ring (Figure 4). Due to increased dipole−dipole and dispersion attractions, antiparallel undisplaced (6-thioguanine)$_2$ is much more stable than (guanine)$_2$.

It can be concluded that the stacking interactions involving thiobases do not differ from those for oxobases dramatically. The dispersion attraction is always enhanced; the electrostatic

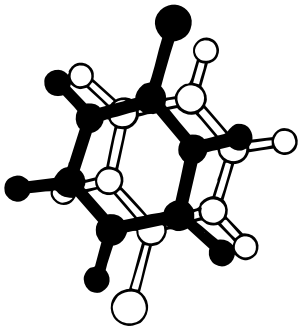

Figure 5. Displaced structure of antiparallel (2-thiouracil)$_2$ dimer without a direct contact of sulfur atom with the ring atoms.

contribution depends on the mutual orientation of bases. The bulky sulfur atom may cause destabilization due to steric clashes, though such steric contacts can be efficiently eliminated by a number of local DNA conformational variations.[49]

III.4. Empirical Potential Calculations. Base stacking in neutral dimers of bases is well reproduced by an empirical potential consisting of a Lennard-Jones potential combined with the standard atom-centered point charge Coulombic term.[3] The atomic charges must be derived from molecular electrostatic potential. In the previous papers,[3,9] the Coulombic term was combined with the scaled 6−9 Lifson-Hagler (6−9LH) Lennard-Jones empirical potential;[49,50] van der Waals interaction energies were scaled by a unique factor of 0.7[3,9] to match the absolute values of stacking energies provided by the MP2/6-31G*(0.25) procedure. Despite the overall agreement between the empirical potential and MP2 data, the potential significantly underestimates the energy difference between parallel and antiparallel undisplaced stacked homo-dimer structures for a vertical separation of bases below 3.4 Å.[3] The 6−9LH potential reproduced well the parallel dimers but overestimated the optimal separation of bases for antiparallel stacked dimers. Reduction of atomic radii would lead to a much better description of the antiparallel geometries; however, in this case short-range repulsion in the parallel dimers is not reproduced satisfactorily. In addition rather localized regions of an increased short-range repulsion were revealed for (cytosine)$_2$ and (adenine)$_2$ which also were not reproduced by the empirical potential.[3]

Here we use the same set of van der Waals parameters for N, C, O, and H atoms as before.[3,9] van der Waals parameters for the sulfur atom were estimated using the ab initio data for stacked (6-thioguanine)$_2$ and (2-thiouracil)$_2$. The equilibrium S···S distance should fall within 3.9−4.3 Å, while the well depth (energy minimum on the van der Waals S···S interaction curve) should be, when combined with scaled 6−9LH potential, 0.5−0.6 kcal/mol. The empirical potential calculations were carried out for the equilibrium S···S distance of 4.2 Å and the well depth of 0.55 kcal/mol if no other statement is made. The point atomic charges were derived by fitting to the molecular electric potential at the MP2/6-31G*(0.25) level.[3]

Figure 6 compares the MP2 ab initio and empirical potential data (twist dependence) for both (6-thioguanine)$_2$ and (2-thiouracil)$_2$. The agreement is not as good as for unmodified dimers (cf. Figure 3 in ref 3). Further, the potential is not accurate enough for compressed (vertical separation of bases of 3.0 Å) antiparallel (6-thioguanine)$_2$, where the MP2 procedure predicts a stacking energy of −10.5 kcal/mol while the potential gives a value of −3.2 kcal/mol. The agreement is not good also for extended (vertical separation of bases of 3.8 Å) antiparallel (6-thioguanine)$_2$; we obtained −8.5 and −11.0 kcal/mol for the MP2 and empirical potential, respectively. Similarly, the potential fails for compressed parallel (6-thioguanine)$_2$

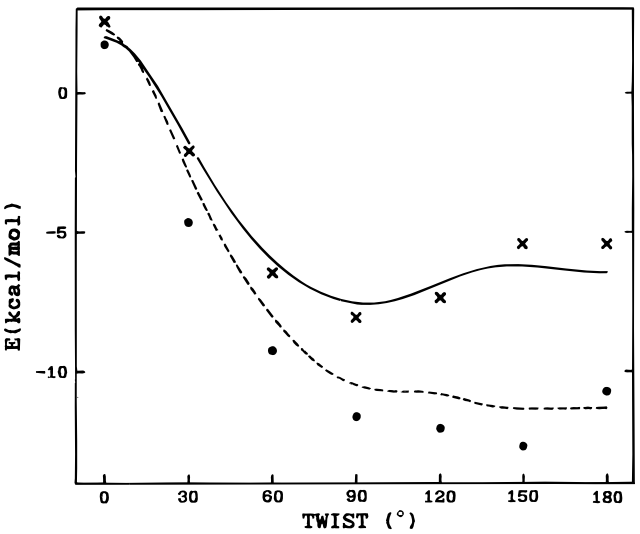

Figure 6. Dependence of stacking energy in (2-thiouracil)$_2$ and (2-thiouguanine)$_2$ on the twist angle. (2-thioguanine)$_2$: (solid circles) MP2/6-31G*(0.25) data, (dashed line) empirical potential data. (2-thiouracil)$_2$: (crosses) MP2/6-31G*(0.25) data, (solid line) empirical potential data. van der Waals radius of the sulfur atom was 2.1 Å; the depth of the S···S van der Waals potential energy curve was −0.55 kcal/mol.

(18.4 and 23.8 kcal/mol by the MP2 method and empirical potential, respectively). The MP2 and empirical potential procedures agree within 1.5 kcal/mol for compressed and extended parallel and antiparallel (2-thiouracil)$_2$.

Table 6 presents the stacking energy difference between parallel and antiparallel dimers obtained by the full MP2 procedure, by the electrostatic term of the potential, and by the whole empirical potential. The electrostatic term reproduces the MP2 prediction for large base−base separations. Upon compression of the dimer the energy difference evaluated by the MP2 method increases much faster than the empirical potential electrostatic energy, while the isotropic Lennard-Jones potential recovers only a small fraction of this difference. (The exceptions are the thioguanine and thiouracil dimers, due to overestimation of the S···S repulsion by our set of parameters.)

III.5. Interaction of Guanine and Thioguanine with a Water Molecule. The hydrogen-bonding and stacking properties of thiobases are rather similar to their parent oxomolecules. The small destabilization of hydrogen-bonded base pairs is not expected to influence the stability of the double helix. Recent experiments demonstrated that the double helix can incorporate even hydrophobic (non-hydrogen-bonded) base pairs.[51,52] Another possible source of the altered properties of thiobase-containing nucleic acids could be changes in the first hydration shell. We carried out a limited set of calculations on a water···nucleobase dimer (Table 7). First, we have optimized one water molecule on the major groove side of guanine and 6-thioguanine. The optimization resulted in a planar complex with the water molecule shared between the N7 and O6 atoms as indicated in Figure 7. In case of the H$_2$O···^{6S}G complex, the final geometry is similar, but the S6···O$_v$ distance is by 0.8 Å larger than the O6···O$_v$ one. It follows the expectation that the water···sulfur interaction is less favorable than the water···oxygen one. Despite this, the H$_2$O···^{6S}G complex is only 0.4 kcal/mol less stable than the H$_2$O···G complex. This could be due to the larger dipole moment of ^{6S}G with respect to G which can compensate for the reduced water···S6 interaction.

TABLE 6: Difference of Stacking Energy (kcal/mol) between Parallel and Antiparallel Stacked Dimers[a]

dimer	VSEP	E^{DIFMP2}	E^{DIFESP}	E^{DIFPOT}
(Cyt)$_2$	3.0	18.8	11.7	13.9
	3.2	13.5	10.1	10.8
	3.4	10.3	8.8	9.1
	3.6	8.2	7.8	7.9
	3.8	7.0	7.0	7.0
	4.0	6.0	6.3	6.2
	4.6	4.4	4.8	4.7
(Ura)$_2$	3.0	12.4	7.6	9.3
	3.4	6.1	5.4	5.5
	3.8	3.8	4.0	4.0
(Ade)$_2$	3.0	13.1	6.1	9.2
	3.4	5.2	3.8	4.3
	3.8	2.6	2.5	2.6
(Gua)$_2$	3.0	17.4	9.9	13.4
	3.4	8.8	7.4	7.8
(^{6S}Gua)$_2$	3.0	28.9	14.2	27.0(23.1)
	3.2	19.7	12.9	18.2(16.3)
	3.4	14.5	11.4	13.6(12.7)
	3.8	9.5	9.2	9.5(9.2)
(^{2S}Ura)$_2$	3.0	16.6	9.1	18.1(14.3)
	3.4	8.0	6.6	8.5(7.6)
	3.8	5.0	5.0	5.4(5.1)
Gua···Cyt[b]	3.0	13.9	9.8	10.8
	3.4	8.6	7.7	8.0
	3.8	6.3	6.3	6.4

[a] E^{DIFMP2}, MP2/6-31G*(0.25) data; E^{DIFESP}, Coulombic term with MP2/6-31G*(0.25) ESP derived charges; $E^{DIFPOT} = E^{DIFESP}$ plus scaled 6−9LH van der Waals potential; VSEP, vertical separation of bases (Å). In the case of sulfur-containing complexes, the values in parentheses were obtained with the equilibrium S···S van der Waals distance reduced to 4.0 Å. [b] The difference between the most stable and least stable undisplaced structure (cf. ref 3).

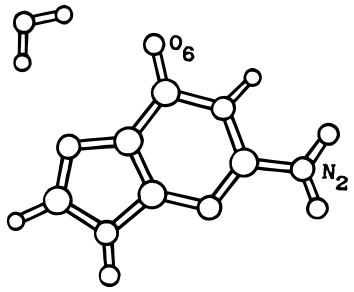

Figure 7. Guanine···H$_2$O complex.

However, in B-DNA there are two (probably bridged) water molecules on the major groove side of guanine.[53] One of them interacts with N7 but not with O6. The other water molecule hydrates the O6 atom with the O···O6C6 angle being around 135° and the O···O6C6C5 angle being about 0°.[53] We carried out additional constrained optimizations of the water···base

TABLE 7: Geometries and Energies of Selected H$_2$O···Guanine and H$_2$O···^{6S}Guanine Complexes: Distances in Angstroms, and Angles in Degrees. ΔE is the MP2/6-311G(2df,p)//HF/6-31G Interaction Energy (kcal/mol)**

H$_2$O···guanine					H$_2$O···^{6s}guanine				
C6O6···O$_v$	O6···O$_v$	C6···O$_v$	N7···O$_v$	ΔE	C6S6···O$_v$	S6···O$_v$	C6···O$_v$	N7···O$_v$	ΔE
113.1[a]	3.11	3.75	3.17	−7.0	96.7[a]	3.90	4.41	3.15	−6.5
135.0[b]	2.99	3.93	4.04	−5.7	135.0[b]	3.56	4.88	4.87	−3.1
170.0[b]	2.98	4.16	5.24	−4.9	170.0[b]	3.50	5.13	6.08	−1.9

[a] The angle has been optimized. [b] The angle has been fixed.

complex with fixed O···O(S)6C6 angles of 135° and 170°. The water oxygen atom was restricted to be coplanar with the base ring (O···O(S)6C6C5 angle of 0°). Table 7 demonstrates that when the water is restricted to interact with the O(S)6 position only, the H_2O···^{6S}G complex is much less stable than the H_2O···G one. The water···base distances (O_v···S(O)6, O_v···C6) are significantly larger for ^{6S}G compared with guanine. Further, while the orientation of water hydrogens remains essentially unchanged for the H_2O···G complex, in the restricted H_2O···^{6S}G complex both hydrogens are oriented toward the sulfur atom and are out-of-plane (one hydrogen atom is below the base, the other above). It is very likely that the B-DNA hydration pattern known for guanine will be altered by ^{6S}G.

III.6. Comparison with older theoretical studies. Our study represents the first high-level analysis of interactions of thiobases. Although we do not consider semiempirical methods as a reliable tool to study the base−base interactions,[7,54] it was encouraging to see that the present calculations basically confirmed some conclusions from the previous studies:[30−33] a small destabilization of H-bonded base pairs upon incorporation of the sulfur atom and increased polarity of thiobases with respect to oxobases. The differences between ^{6S}G···C and G···C base pairs are too small to destroy the normal biological functioning of nucleic acids.[33] Our calculations on the other hand indicate an alteration of the first hydration shell by the sulfur atom.

IV. Conclusions

H-bonded base pairs involving thiobases are almost as stable as the standard base pairs.

The stacking interactions of thiobases are influenced by three factors: (i) significantly enhanced dispersion attraction, (ii) enhanced dipole−dipole interaction due to the increased dipole moments of monomers, and (iii) moderate steric clashes between the sulfur atoms and between the sulfur atom and other second-row elements.

Solvation of the S6 position of 6-thioguanine is unfavorable. This indicates that the B-DNA hydration pattern with separately hydrated N7 and O6 positions of guanine will be perturbed by thioguanine.

Acknowledgment. This study was supported by Grant 203/97/0029 from the GA ČR, by Grant 204/97/K084 from the GA ČR, by ONR Grant N00014-95-1-0049, and by NIH Grant 5SO6GM08047-24. We would like to thank the Mississippi Center for Supercomputing Research for a generous allotment of computer time.

References and Notes

(1) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493.
(2) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965.
(3) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.
(4) Šponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 695.
(5) Šponer, J.; Florián, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 827.
(6) Šponer, J.; Hobza, P.; Leszczynski, J. In *Computational Chemistry. Reviews of Current Trends*; Leszczynski, J., Ed.; World Scientific Publisher: Singapore, 1996; p 185.
(7) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117.
(8) Florián, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010.
(9) Šponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. *Biophys. J.* **1997**, *73*, 76.
(10) Alhambra, C.; Luque, F. J.; Gago; F. Orozco, M. *J. Phys. Chem. B* **1997**, *101*, 3846.
(11) Brameld, K.; Dasgupta, S.; Goddard, W. A., III. *J. Phys. Chem. B* **1997**, *101*, 4851.
(12) Williams, A. M.; Jorangen, E. *Br. J. Cancer* **1961**, *15*, 342.
(13) Sartorelli, A. C.; Junga, I. G. *Cancer. Res.* **1958**, *18*, 938.
(14) Rao, T. S.; Durland, R. H.; Seth, D. M.; Myrick, M. A.; Bodepudi, V.; Revankar, G. *Biochemistry* **1995**, *34*, 765.
(15) Bugg, C. E.; Thewalt, U. *J. Am. Chem. Soc.* **1970**, *92*, 7441.
(16) Dzik, J. M; Bretner, M.; Kulikowski, T.; Golos, B.; Jarmula, A.; Poznanski, J.; Rode, W.; Shugar, D. *Biochim. Biophys. Acta* **1996**, *1293*, 1.
(17) Rogers, K. C.; Crescenzo, A. T.; Söll, D. *Biochimie* **1995**, *77*, 66.
(18) Rappaport, H. P *Biochemistry* **1993**, *32*, 3047.
(19) Kumar, R. J.; Davis, D. R. *J. Org. Chem.* **1995**, *60*, 7726.
(20) Sierputowska-Gracz, H.; Sochacka, E.; Malkiewicz, A.; Kuo, K.; Gehrke, C. W.; Agris, P. F. *J. Am. Chem. Soc.* **1987**, *109*, 7171.
(21) Dubler, E. In *Metal Ions in Biological Systems*; Sigel, A., Sigel, H., Eds.; Marcel Dekker: New York, 1996; Vol. 32, p 301.
(22) (a) Leszczynski, J. *J. Mol. Struct. (THEOCHEM)* **1994**, *311*, 37. (b) Stewart, M. J.; Leszczynski, J.; Rubin, Y. V.; Blagoi, Y. P. *J. Phys. Chem. A* **1997**, *101*, 4753.
(23) Alhambra, C.; Lugue, F.; Estelrich, J.; Orozco, M. *J. Org. Chem.* **1995**, *60*, 969.
(24) Leszczynski, J.; Šponer, J. *J. Mol. Struct. (THEOCHEM)* **1996**, *388*, 237.
(25) Lapinski, L.; Rostkowska, M.; Nowak, M.; Kwiatkowski, J. S.; Leszczynski, J. *Vibrat. Spectrosc.* **1996**, *13*, 29.
(26) Rostkowska, M.; Szczepaniak, K; Nowak, M. J.; Leszczynski, J.; KuBulat, K.; Person, W. B. *J. Am. Chem. Soc.* **1990**, *112*, 2147.
(27) Singh, K.; Yadav, R. A.; Yadav, J. S. *Spectrochim. Acta* **1991**, *47A*, 819.
(28) Rubin, Yu. V.; Bokovoi, V. A.; Blagoi, Yu. P. *J. Fluorescence* **1995**, *5*, 263.
(29) Nowak, M. J.; Lapinski, L.; Kwiatkowski, J. S.; Leszczynski, J. In *Computational Chemistry. Reviews of Current Trends. Vol. II;* Leszczynski, J., Ed.; World Scientific Publisher: Singapore, 1997; p 140.
(30) Chojnacki, H.; Sokalski, W. A. *J. Theor. Biol.* **1975**, *54*, 167.
(31) Geller, M.; Jaworski, A.; Pohorille, A. *Int. J. Quantum. Chem.* **1979**, *15*, 369.
(32) Aida, M.; Nagata, C.; Ohmine, I.; Morokuma, K. *J. Theor. Biol.* **1982**, *99*, 599.
(33) Lipinski, J. *J. Mol. Struct. (THEOCHEM)* **1989**, *201*, 87.
(34) Bludsky, O.; Šponer, J.; Leszczynski, J.;Špirko, V.; Hobza, P. *J. Chem. Phys.* **1996**, *108*, 11042.
(35) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, W. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, L. J.; Replogle, E. S.; Gomperts, R.; Martin, R. L; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian94;* Gaussian, Inc.: Pittsburgh, PA, 1995.
(37) Šponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263.
(38) Kroon-Batenburg, L. M. J.; van Duijneveldt, F. B. *J. Mol. Struct.* **1985**, *121*, 185.
(39) Hobza, P.; Selzle, H. L.; Schlag, E. W. *Chem. Rev.* **1994**, *116*, 3500.
(40) Johnson, R. C.; Power, T. D.; Holt, J. S.; Immaraporn, B.; Monat, J. E.; Sissoko, A.; Yanik, M. M.; Zagorodny, A. V.; Cybulski, M. S. *J. Phys. Chem.* **1996**, *100*, 18875.
(41) Bakalarski, G.; Grochowski, P.; Kwiatkowski, J. S.; Lesyng, B.; Leszczynski, J. *Chem. Phys.* **1996**, *204*, 301.
(42) Basch H.; Garemer, P. G.; Jasien, M.; Krauss, M.; Stevens, W. J. *Chem. Phys. Lett.* **1989**, *163*, 514.
(43) Špirko, V.; Šponer, J.; Hobza, P. *J. Chem. Phys.* **1997**, *106*, 1472.
(44) Pranata, J.; Wierschke, S. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.
(45) van Duijneveldt, F. B.; van Mourik, T. *J. Mol. Struct. (THEOCHEM)* **1995**, *341*, 63.
(46) Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12116.
(47) Desiraju, G. T. *Acc. Chem. Res.* **1991**, *24*, 290.
(48) Steiner, T. *Cryst. Rev.* **1996**, *6*, 1.
(49) Šponer, J.; Kypr J. *J. Biomol. Struct. Dyn.* **1993**, *11*, 27.
(50) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111.
(51) Schweitzer, B. A.; Kool, E. T. *J. Am. Chem. Soc.* **1994**, *117*, 1863.
(52) Moran, S.; Rex, X.-F.; Rumney, S.; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056.
(53) Schneider, B.; Berman, H. M. *Biophys. J.* **1995**, *69*, 2661.
(54) Hobza, P.; Kabeláč, M.; Mejzlík, P.; Šponer, J.; Vondrášek, J. *J. Comput. Chem.* **1997**, *18*, 1136.