Interaction Energies of Hydrogen-Bonded Formamide Dimer, Formamidine Dimer, and Selected DNA Base Pairs Obtained with Large Basis Sets of Atomic Orbitals

Jiří Šponer* and Pavel Hobza

J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic,
Dolejškova 3, 182 23 Prague, Czech Republic

Received: December 17, 1999; In Final Form: February 24, 2000

Interaction energies of selected H-bonded nucleic acids base pairs, formamide dimer, and formamidine dimer have been evaluated using the second-order Møller−Plesset method (MP2) with extended basis sets of atomic orbitals. Basis sets with two sets of d-polarization functions on each non-hydrogen atom give the same interaction energies as medium-sized 6-31G** and 6-31G*(0.25) basis sets. This is due to a compensation of errors in the evaluation of the Hartree−Fock and correlation interaction energies. With a further increase of the size of the basis set, a steep improvement (increase in absolute value) of the interaction energy appears. It amounts to 0.6−1.1 kcal/mol with the cc-pVTZ basis set. The aug-cc-pVQZ and cc-pV5Z basis sets improve the interaction energy of model complexes by additional ca. 1.3 kcal/mol while these data appear to be very close to the basis set limit. In contrast to the MP2 method, Becke3LYP density functional theory method does not show any systematic improvement of the interaction energies with the increase of the basis set size. The previously published reference values for interaction energies of H-bonded base pairs obtained with medium-sized basis sets of atomic orbitals are likely to be underestimated (in absolute value) by 2.0−2.5 kcal/mol while relative stabilities of base pairs are correctly reproduced. Nevertheless, with the present computer facilities we cannot yet investigate in detail the influence of the quality of the optimized geometries of the dimers on their stabilization energies. It is expected, however, that this contribution is considerably smaller compared to the primary basis set effects evaluated in this study.

Introduction

Interactions of nucleic acid bases significantly influence the structure, dynamics, and function of nucleic acids. Therefore, studies of interactions of nucleobases are important to understand their role in nucleic acids[1−11] and to parametrize accurate molecular mechanical force fields for molecular modeling of nucleic acids.[12] Accurate gas phase data are required since it is very difficult to extract the intrinsic base−base interactions from condensed phase and/or crystal data.[13] In the absence of a sufficient amount of gas phase experimental data, the basic tool to study interactions of nucleic acid bases is the ab initio quantum-chemical method. Inclusion of electron correlation is necessary to achieve reasonable accuracy.[1,10] The base pairing should be studied by standard variational supermolecular post-Hartree−Fock method. The perturbation method in the form of symmetry-adapted perturbation theory (SAPT),[14a] which is very successful for studies of interactions of small rigid systems,[14b] is not well suited for base pairs requiring to optimize all inter- and intramolecular coordinates. Local methods of inclusion of electron correlation effetcs[14c] are still far from routine applications. We advice considerable caution and a careful verification when applying existing codes to molecular clusters since actual MP2 correlation interaction energy might be underestimated by the local MP2 method substantially.[14d] Density functional theory (DFT) presently does not capture the intersystem electron correlation effects properly[15] and brings no advantage over the HF approximation for base stacking[2b] and related interactions,[15] although work is in progress to improve the applicability of DFT for studies of molecular clusters.[15c−f]

Our presently available reference values for H-bonding[1a] and stacking interactions[1b,c] of nucleobases were evaluated using second-order Møller−Plesset (MP2) perturbational theory with medium-sized basis sets of atomic orbitals having a set of diffuse d-polarization functions on all second-row elements. The structures of the H-bonded base pairs were obtained using gradient optimization carried out within the Hartree−Fock (HF) approximation. While the interaction energies were corrected for an artifact known as the basis set superposition error (BSSE), the optimization procedure did not include this correction. Thus, the available reference values for H-bonding of bases can be spoiled by several inaccuracies: (i) The optimized geometries can be inaccurate due to the neglect of electron correlation, BSSE, and small size of the basis sets used for the optimizations. (ii) The subsequent evaluation of interaction energies is affected by the use of still rather small basis sets.

Recently, several attempts have been made to improve the quality of calculations of DNA base pairs and to estimate quality of the older data.

Several base pairs have been reoptimized at the MP2 level using medium-sized basis sets of atomic orbitals.[1a,7] The calculations indicate that optimizations carried out at the HF level introduce no significant error into the evaluation of the stability of the base pairs. Further, the MP2-optimized geometries are not necessarily more accurate than the Hartree−Fock ones, since at the MP2 level the basis set superposition error increases substantially. With current computer resources, counterpoise (CP)-corrected gradient optimization techniques[16] can-

---

* Corresponding author. E-mail: sponer@indy.jh-inst.cas.cz. Fax: 420 2 858 2307.

10.1021/jp9943880 CCC: $19.00

not be carried out for base pairs with inclusion of electron correlation effects and with large basis sets. Nevertheless, CP-corrected gradient optimization has been reported for AT, TC, and GA base pairs at the HF/6-31G** level.[10e] The CP-corrected structures provided almost identical stabilization energies as obtained after the standard optimization. Also, the geometries of base pairs were not substantially affected, except of the GA base pair. This base pair is intrinsically nonplanar and the interbase plane angle was moderately changed by the CP procedure.[10e]

The importance of higher-order electron correlation effects was investigated. It was shown that the MP2 procedure provides almost identical results as coupled cluster method with inclusion of noniterative triple electron excitations [CCSD(T)] for cytosine−cytosine and uracil−uracil H-bonded base pairs.[8] Similar results were obtained for other H-bonded systems.[8,17−19] It is due to compensation of errors and there is no guarantee that one can rely on the compensation over the whole spectrum of H-bonds. Nevertheless, it is reasonable to expect that the MP2 procedure is sufficiently accurate for H-bonded base pairs. Interestingly, for aromatic stacking clusters the MP2 method overestimates the stabilization.[8,20]

Finally, the influence of the size of basis set on the stabilization energies was considered. Local MP2 (LMP2) method was applied with the cc-pVTZ(-f) basis set (the notation -f means that f functions were omitted).[9] The calculations have shown for several base pairs an agreement (within 1 kcal/mol) with the previous data. However, several recent studies on small H-bonded complexes revealed that inclusion of higher-order polarization functions leads to a substantial improvement of the correlation part of the interaction energies.[8,19] In addition, it still remains to prove the accuracy of the LMP2 method compared to the standard MP2 procedure.[14d]

In this paper we present a new reevaluation of interaction energies of selected DNA base pairs, using correlation consistent cc-pVTZ basis set and several other large basis sets. For model systems, systematic calculations are performed up to the MP2/aug-cc-pVQZ and MP2/cc-V5Z levels. The calculations reveal that the previous reference studies underestimated (in absolute values) the interaction energies of base pairs while even the cc-pVTZ basis set is not sufficient to reach complete convergence. We estimate that medium-sized basis sets underestimate the interaction energies of base pairs in absolute value by 2.0−2.5 kcal/mol. The effect appears to be systematic and relative stability of base pairs reported before should not be affected. Size of the basis set used in the single-point interaction energy evaluations is the most important factor influencing the accuracy of calculations of gas phase dimerization of H-bonded base pairs.

Method

All calculations were done using the Gaussian94 code.[21] The structures were optimized at HF or MP2 levels of theory, as indicated below. The interaction energies were evaluated using the MP2 method and frozen core approximation. A variety of Pople's[21] and Dunning's correlation[22] consistent basis sets were used (Table 1).

The *interaction energy* in this paper is defined as the energy difference between the total electronic energy of the dimer and the electronic energies of the monomers separated into infinity without allowing their geometries to relax (i.e., assuming the same intramolecular geometries of the monomers as within the complex). All interaction energies were a posteriori corrected for the basis set superposition error using the full counterpoise procedure. We have separately calculated the *deformation energies* of the monomers upon the formation of the complex, ΔE^{DEF}. The deformation energy is the difference between electronic energies of the two monomers assuming they have the same geometry as within the dimer and electronic energies of fully relaxed isolated monomers. ΔE^{DEF} is repulsive contribution and should be added to the interaction energy in order to obtain the *complexation energy* of the dimer. Let us note that some authors include the deformation energy as a part of the interaction energy; the interaction energies are then equivalent to complexation energies in the present paper. The reader should always check whether in a paper the interaction energies include the deformation energies or not.

TABLE 1: Basis Set of Atomic Orbitals Used in the Present Paper[a]

basis set	second-row elements	hydrogen atoms	HF energy	MP2 energy
6-31G*(0.25)	3s2p1d	2s	−168.877 558	−0.377 016
6-31G**	3s2p1d	2s1p	−168.937 361	−0.483 611
6-311G(2d,p)	4s3p2d	2s1p	−168.982 524	−0.547 127
6-311G(2df,p)	4s3p2d1f	2s1p	−168.988 713	−0.591 936
6-311+G(2df,pd)	5s4p3d1f	3s1p1d	−168.995 588	−0.601 557
6-311++G(3df,dp)	5s4p3d1f	3s1p1d	−168.998 216	−0.608 336
aug-cc-pVDZ	4s3p2d	3s2p	−168.962 103	−0.512 153
cc-pVTZ	4s3p2d1f	3s2p1d	−168.999 146	−0.605 891
aug-cc-pVTZ	5s4p3d2f	4s3p2d	−169.002 773	−0.617 703
cc-pVQZ	5s4p3d2f1g	4s3p2d1f	−169.012 434	−0.649 713
aug-cc-pVQZ	6s5p4d3f2g	5s4p3d2f	−169.013 560	−0.655 312
cc-pV5Z	6s5p4d3f2g1h	5s4p3d2f1g	−169.015 815	−0.667 298

[a] The first two columns show the number of contracted basis functions. the last two columns show total HF and MP2 electronic energies (au) of formamide evaluated with these basis sets of atomic orbitals.

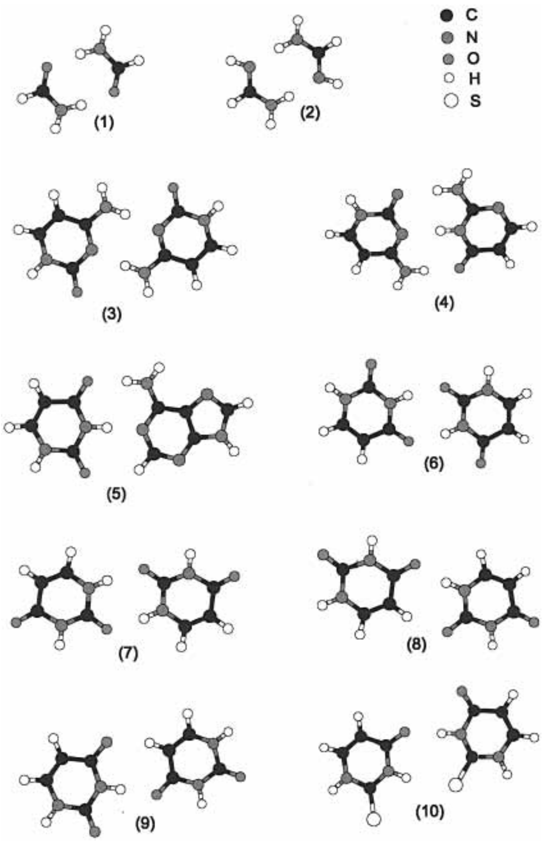

Figure 1. (1) FaFa dimer, (2) FiFi dimer, (3) CC base pair, (4) iCC WC base pair, (5) AU WC base pair, (6) UU1 base pair, (7) UU4 base pair, (8) UU7 base pair, (9) UU C base pair, (10) thUthU base pair.

The following systems have been studied (Figure 1): adenine−uracil Watson−Crick base pair (AU), isocytosine−cytosine

TABLE 2: Interaction Energies (kcal/mol) of Selected H-Bonded Dimers (See Figure 1)[a]

	6-31G*(0.25)	6-31G**	6-311G(2d,p)	6-311G(2df,p)	aug-cc-pVDZ	ΔE^{DEF}
FaFa[b]	−12.63/−12.81	−12.98/−12.67	−13.32/−11.93	−13.72/−11.89	−14.09/−12.16	1.73[b]
FiFi[b]	−12.30/−9.82	−14.32/−11.93	−14.05/−10.12	−14.62/−10.25	−14.56/−10.17	1.33[b]
CC[c]	−19.69/−16.31	−20.22/−17.67	−20.36/−16.17	−20.84/−16.20	−20.82/−16.06	2.54[c]
iCC[d]	−25.74/−24.21	−26.22/−25.17	−26.00/−23.36	−26.46/−23.42	−26.70/−23.24	2.31[d]
AU[c]	−13.39/−10.43	−13.77/−11.26	−14.19/−10.06	−14.65/−10.13	−15.03/−10.27	1.69[c]
UU1[d]	−10.46/−9.22	−10.17/−9.09	−10.18/−8.20	−10.45/−8.30	−11.03/−8.49	0.62[d]
UU4[d]	−15.87/−15.75	−15.97/−15.72	−15.99/−14.78	−16.28/−14.87	−16.85/−15.00	1.09[d]
UU 7[c]	−11.25/−11.03	−11.30/−11.45	−11.65/−10.49	−11.93/−10.50	−12.49/−10.65	0.77[c]
UU C[c]	−8.12/−6.97	−8.13/−7.15	−8.30/−6.41	−8.54/−6.41	−9.11/−6.61	0.56[c]
thUthU[d]	−9.27/−6.85	−8.49/−6.81	−8.75/−5.99	−9.00/−6.03	−9.80/−6.28	0.53[d]

[a] The first value in each column is the MP2 interaction energy, the second number is the HF component of the interaction energy. The last column (ΔE^{DEF}) provides deformation energies of the monomers based on the gradiently optimized structures (see Method). [b] MP2/DZ(d,p) optimized geometry. [c] MP2/6-31G** optimized geometry. [d] HF/6-31G** optimized geometry.

Watson−Crick base pair (iCC), cytosine−cytosine base pair (CC), four uracil−uracil base pairs (UU1, UU4, UU7, UU C), thiouracil−thiouracil base pair (thUthU), formamide dimer (FaFa), and formamidine dimer (FiFi). The iCC base pair is a smaller model complex for the standard guanine−cytosine base pair. The four uracil dimers show a wide variety of hydrogen bonds including two cases with a C−H···O bond pattern. The designation of the uracil base pairs is taken from our previous molecular dynamics study.[4b] The thiouracil dimer was included for a comparison since it has larger dispersion contribution. Formamide and formamidine dimers have been studied in the past and represent important model complexes having many features similar to base pairs.[16,23] The following systems were studied assuming C_i symmetry: CC, UU4, FaFa, and FiFi. Calculations for UU7, UU C, AU, and CC are based on MP2/6-31G** optimized geometries; FaFa and FiFi were optimized at the MP2/DZ(d,p) level. iCC base pair has been optimized at the MP2/6-31G** and HF/6-31G** levels. The remaining structures were optimized at the HF/6-31G** level. The effect of the size of basis set on the interaction energy is very similar for HF- and MP2-optimized geometries. Geometries of all systems can be obtained from the authors upon request.

Results

We first compared interaction energies obtained using medium-sized 6-31G** and 6-31G*(0.25) basis sets (Table 2). The 6-31G*(0.25) basis set has momentum-optimized exponents of the d-functions.[24] We have used the 6-31G*(0.25) basis set in our older studies in order to improve the inclusion of the dispersion attraction.[1] Further studies have revealed that while the use of the 6-31G*(0.25) basis set is essential for base stacking studies the diffuse polarization functions do not provide a significant improvement for the H-bonding. For H-bonded base pairs both basis sets provide similar results, as confirmed also in Table 2.

Then, we have considered the larger 6-311G(2d,p) basis set (Table 2). This basis set could be still applied for routine evaluations of base pair interaction energies. This improvement of the flexibility of the basis set has no influence on the stabilization energies. However, closer inspection of the data shows that there has been a considerable improvement of the electron correlation interaction energy. However, it has been canceled by a substantial opposing change in the HF interaction energy. In the next column of Table 2 we present data evaluated with 6-311G(2df,p) basis set. These are the first calculations on base pairs including f-polarization functions. Adding the f functions has already no marked effect on the HF interaction energy and in all cases the MP2 interaction energy is improved (i.e., larger in absolute value). Further considerable improvement of the correlation interaction energy is achieved using the aug-cc-pVDZ basis set. Although this basis set does not contain f-functions, it possesses two sets of d-polarization functions on second row elements and two sets of p-polarization functions on hydrogens. The basis set is very diffuse and has a large number of primitive functions; both factors improve the interaction energies.

In the next round of calculations we have increased the size of the basis set to the cc-pVTZ one (Table 3). This basis set contains one set of f- and two sets of d-functions on the second row elements similar to the 6-311G(2df,p) one, but it has a larger number of primitive functions. Further, d-polarization functions and a second set of p-polarization functions are augmented on the hydrogen atoms. This improvement has only marginal effect on the HF interaction energy, which appears to be basically converged. Because of a pronounced increase in the electron correlation stabilization the total interaction energy has improved compared to the MP2/6-311G(2d,p) level by 0.6−1.1 kcal/mol. The smallest improvement has been found for the two base pairs having C−H···O contact.

We have evaluated the interaction energy of the UU4 base pair with the aug-cc-pVTZ basis set, and this calculation demonstrated a further increase of the stabilization by ca. 0.6 kcal/mol, in line with the trend observed for FaFa and FiFi dimers. We have also tested two extended Pople's type basis sets [6-311+G(2df,pd) and 6-311++G(3df,pd)] for selected systems, and both basis sets show a good performance.

Then we have extended the calculations for the formamidine and formamide dimers by systematically applying larger and larger correlation-consistent basis sets up to the cc-pV5Z one. The calculations show further improvements in the interaction energies. It is interesting to note that MP2/aug-cc-pVTZ and MP2/cc-pVQZ results differ only marginally. However, the subsequent extension to the diffuse aug-cc-pVQZ basis set brings additional 0.3 kcal/mol of stabilization (Table 3). It is a clear indication that diffuse functions contribute to the dispersion energy in the case of H-bonded systems. The MP2/aug-cc-pVQZ calculations improved the interaction energy in the dimers by ca. 2.4 (FaFa) and 2.2 (FiFi) kcal/mol compared to calculations with 6-311G(2d,p) basis set. The largest cc-pV5Z basis set provided the same interaction energy as the aug-cc-pVQZ one. On the basis of literature data published recently for formic acid dimer[18] showing similar trends, we can assume that the MP2 basis set limit for the FaFa and FiFi dimers is by about 0.5 and 0.2 kcal/mol above (in absolute value) the MP2/cc-pVQZ and MP2/cc-pV5Z interaction energies, respectively.[18,25] We expect that underestimation (in absolute value) of the interaction energies with respect to the basis set limit would be 1.4−1.6

TABLE 3: Interaction Energies (kcal/mol) of Selected H-Bonded Dimers (See Figure 1)[a]

	6-311G(2d,p)	6-311+G(2df,pd)	cc-pVTZ	6-311++G(3df,pd)	aug-cc-pVTZ	cc-pVQZ	aug-cc-pVQZ	cc-pV5Z	ΔE^{DEF}
FaFa[b]	−13.32	−14.18	−14.41	−14.68	−15.25	−15.42	−15.75	−15.77	1.73[b]
	−11.93	−12.01	−12.16	−12.20	−12.29	−12.33	−12.34	−12.35	
FiFi[b]	−14.10		−15.17		−15.88	−15.97	−16.32	−16.32	1.33[b]
	−10.17		−10.39		−10.43	−10.43	−10.46	−10.46	
CC[c]	−20.36	−21.18	−21.36	−21.40					2.59[c]
	−16.17	−16.20	−16.35	−16.08					
iCC[d]	−26.00		−26.94						2.31[d]
	−23.36		−23.46						
iCC[c]	−28.47								4.34[c]
	−25.12								
AU[c]	−14.19		−15.19	−15.59					1.69[c]
	−10.06		−10.22	−10.29					
UU1[d]	−10.18		−10.79						0.62[d]
	−8.20		−8.37						
UU4[d]	−15.99	−16.77	−16.79	−17.18	−17.61				1.09[d]
	−14.78	−14.95	−14.99	−15.04	−15.07				
UU7[c]	−11.65		−12.36						0.77[c]
	−10.49		−10.57						
UU C[c]	−8.30		−8.93						0.56[c]
	−6.41		−6.47						
thUthU[d]	−9.00		−9.58	−10.06					0.53[d]
	−6.03		−6.27	−6.23					

[a] The first value in each column is the MP2 interaction energy, and the second number is the HF component of the interaction energy. [b] MP2/DZ(d,p) optimized geometry. [c] MP2/6-31G** optimized geometry. [d] HF/6-31G** optimized geometry.

kcal/mol and 2.4−2.7 kcal/mol for the cc-pVTZ and 6-311G(2d,p) basis sets, respectively. Concerning the base pairs it seems that the basis set effects are slightly less pronounced compared with the two model systems. Therefore, we predict that older calculations on H-bonded base pairs underestimated the interaction energies in absolute values by about 2.0−2.5 kcal/mol. To make this estimate we utilize the observation that the 6-311G(2d,p) basis set gives very similar results for base pairs as the 6-31G*(0.25) and 6-31G** basis sets typically used in earlier studies of base pairing.

We have also carried out some calculations with the smallest correlation consistent cc-pVDZ basis set; however, this basis set is apparently too small to show the proper systematic trends. For example, for the FaFa dimer it clearly gives the smallest stabilization among all basis sets used, interaction energy of −11.7 kcal/mol, while for the FiFi dimer a rather reasonable estimate of −13.1 kcal/mol was obtained. The cc-pVDZ basis set was shown to strongly exaggerate pyramidalization of amino groups of isolated nucleobases.[26] The aug-cc-pVDZ basis set (Table 2), second in the series of correlation consistent basis sets, is known to provide very good estimates of dispersion attraction for stacked clusters (mainly when applying the CCSD(T) method). Also for H-bonded dimers, this basis set provides reliable stabilization energies mostly comparable with the cc-pVTZ basis set (cf. the Tables 2 and 3).

The last column of Tables 2 and 3 shows the deformation energies of the monomers upon formation of the dimers, calculated using the respective gradiently optimized structures (see method). This repulsive contribution should be added to the interaction energies reported in Tables 2 and 3 in order to obtain the complexation energies. Note that the deformation energies increase with the strength of the base pairing. The deformation energies for CC, iCC, and AU base pairs are calculated with respect to fully optimized (nonplanar) monomers having nonplanar amino groups.

All values presented in Tables 2 and 3 are corrected for the basis set superposition error. We do not present the uncorrected data in the tables since the uncorrected interaction energies have no physical meaning. Basis set superposition error is a purely mathematical artifact caused by the finite size of the basis set. Let us nevertheless provide a few values proving that the counterpoise procedure has to be applied even for large basis sets of atomic orbitals. The magnitude of the BSSE for the UU4 base pair using 6-31G*(0.25), aug-cc-pVDZ, cc-pVTZ, and aug-cc-pVTZ basis set is 6.28, 2.81, 1.91, and 1.65 kcal/mol. The BSSE remains as large as 0.46 kcal/mol for the MP2/cc-V5Z calculations carried out for FaFa and FiFi dimers. The effect of the basis set size on the interaction energies would be clearly reversed if BSSE is not eliminated. The magnitude of the BSSE is significant even for large basis sets of atomic orbitals, and especially for smaller basis sets the CP-corrected results are considerably closer to the actual values than the CP-uncorrected ones.

Let us make a few comments on the HF interaction energies. As shown above, 6-31G** and 6-31G*(0.25) basis sets somewhat exaggerate the HF stabilization, compared with say the 6-311G(2d,f) basis set. Interestingly, the use of the largest basis sets introduces an additional small correction in the opposite direction, i.e., somewhat back toward the HF/6-31G** values. Nevertheless, the HF interaction energies are much less sensitive to the size of the basis set than the correlation interaction energies. Another interesting point concerns the FiFi dimer. For this complex, the 6-31G** basis set provides substantially better interaction energy than the 6-31G*(0.25) basis set, in contrast to all other complexes. Further, in contrast to other H-bonded systems, FiFi shows a considerable difference between MP2 and CCSD(T) interaction energies.[17] It on the one hand indicates that FiFi is a less realistic model for base pairs than FaFa. On the other hand, it shows that caution is still necessary when extrapolating from one system to another unless a very high quality of calculations is reached.

We have repeated the calculations for formamide dimer using the Becke3LYP density functional method. These calculations show no apparent improvement of the interaction energy with increasing the size of the basis set (−14.26 and −14.38 kcal/mol for cc-pVTZ and cc-pVQZ basis sets, respectively). Thus, this method cannot be used for reference calculations even for H-bonded systems, for reasons well documented in the literature.[15]

Concluding Remarks

We have carried out new reference calculations of the interaction energies of selected base pairs, formamide dimer, and formamidine dimer. In the study we addressed primarily the accuracy of the single-point interaction energy evaluations, leaving the question of the accuracy of the geometries to future studies. Nevertheless, the dependence of the interaction energy on the size of the basis is the major factor influencing the accuracy of the predictions.

The calculations show that improving the previous data obtained with basis sets such as 6-31G** or 6-31G*(0.25) by basis sets with two d-polarization functions on each non-hydrogen atom does not change the results. It reflects a certain compensation of errors between HF and correlation interaction energies. With a further increase of the basis set the HF interaction energy becomes saturated and a fast improvement of the interaction energy appears, being driven by the electron correlation component. It amounts to ca. 1 kcal/mol (in absolute value) with the cc-pVTZ basis set for all base pairs and the model complexes and reaches a value above 2 kcal/mol when MP2/aug-cc-pVQZ and MP2/cc-pV5Z calculations are applied for the model complexes. The aug-cc-pVQZ and cc-pV5Z basis sets are close to the basis set limit. Therefore, the previously published reference values for interaction energies of H-bonded DNA base pairs are underestimated in absolute value by ca. 2.5 kcal/mol. Relative differences in stabilities of various base pairs are properly predicted with the medium-sized basis sets.

With the present computer facilities we could not investigate in detail the influence of the quality of the optimized geometries of the dimers on their stabilization energies. It is expected, however, that this contribution is considerably smaller compared to the primary basis set effects evaluated in this study. We have nevertheless estimated the influence of the inclusion of electron correlation effects during the optimization. The data for iCC base pair (Table 3) show that the single-point calculations carried out using the MP2/6-31G** optimized geometry provide significantly better interaction energy compared to the data evaluated for the HF-optimized geometry. However, in contrast to the basis set dependence of the interaction energy discussed in detail above, this improvement of interaction energy concerns mainly the HF component of the interaction energy (see Table 3). Further, the improvement of the interaction energy achieved by using the MP2-optimized geometry is basically canceled when the deformation energies of monomers are considered to calculate the complexation energies.[1a,7] This can be explained in the following way. The inclusion of electron correlation effects increases the deformability of the monomers. Thus, the monomers show larger adjustments of their geometries (such as prolongation of the $X-H$ bonds participating in the $X-H\cdots Y$ hydrogen bonds) upon dimerization at the MP2 level. This obviously leads to better HF interaction energies, however, at the expense of larger deformation energies of the monomers.

Let us finally comment on the other type of nucleobase interactions, namely the base stacking. Current reference values of base stacking energies are based on MP2 data with medium-sized basis sets of atomic orbitals with diffuse (momentum-optimized) exponents of the d-polarization functions.[1a,c,7] Use of the diffuse polarization functions is required to obtain meaningful values of the dispersion energy stabilizing the stacked clusters. Although stacking energies converge faster with the size of the basis set than H-bonding energies,[8] one should assume that the basis sets used in previous studies are still rather small. With current computer resources we would be able to apply much larger basis sets. Nevertheless, we did not attempt any such reinvestigation for the following reason: Reference calculations on model complexes clearly demonstrated that the MP2 procedure overestimates (in absolute values) the aromatic stacking energies compared to CCSD(T) values.[8] This imbalance of the MP2 procedure compensates for the size of the basis set, and MP2 base stacking energies obtained with diffuse medium-sized basis sets are likely to be close to the actual (unknown) values. In contrast to H-bonding, to further reduce the uncertainty in the stacking evaluations would require a simultaneous substantial increase in the quality of the basis set and electron correlation method. This has not been feasible with the current computer facilities while MP2 calculations with large basis sets or CCSD(T) calculations with medium-sized basis sets are likely to provide less accurate values than those currently available. We will address the accuracy of the base stacking calculations in the near future.

Acknowledgment. This study was supported by a grant A4040903 from IGA AS CR and by grant 203/00/0633 from the GA ČR. Our special thanks go to the Supercomputer Center, Brno.

References and Notes

(1) (a) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem*. **1996**, *100*, 5590. (b) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem*. **1996**, *100*, 1965. (c) Šponer, J.; Gabb, H. A.; Leszczynski, J.; Hobza, P. *Biophys. J*. **1997**, *73*, 76.

(2) (a) Hobza, P.; Šponer, J.; Polasek, M. *J. Am. Chem. Soc*. **1995**, *117*, 792. (b) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Comput. Chem.* **1996**, *12*, 841. (c) Šponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *13*, 695. (d) Alhambra, C.; Luque, F. J.; Gago, F.; Orozco, M. *J. Phys. Chem. B* **1997**, *101*, 3846. (e) Saito, I.; Nakamura, T.; Nakatani, K.; Yoshioka, Y.; Yamaguchi, K.; Sugiama, H. *J. Am. Chem. Soc*. **1998**, *120*, 12686.

(3) (a) Florian, J.; Leszczynski, J. *J. Biomol. Struct. Dyn*. **1995**, *12*, 1055. (b) Florian, J. Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010. (c) Šponer, J.; Burda, J. V.; Mejzlik, P.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1997**, *14*, 613. (d) Bertran, J.; Oliva, A.; Rodriquez-Santiago, L.; Sodupe, M. *J. Am. Chem. Soc*. **1998**, *120*, 8159. (e) Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. B* **1998**, *102*, 9109. (f) Zhanpeisov, N. U.; Leszczynski, J. *Int. J. Quantum, Chem*. **1998**, *69*, 37 (g) Gadre, S. R.; Pundlik, S. S.; Limaye, A. C.; Rendel, A. P. *Chem. Commun.* **1998**, 573. (h) Cysewski, P. *J. Chem. Soc., Faraday Trans*. **1998**, *94*, 3117. (i) Kawahara, S.-I.; Wada, T.; Kawauchi, S.; Uchimaru, T.; Sekine, M. *J. Phys. Chem. A* **1999**, *103*, 8516. (j) Barsky, D.; Kool, E. T.; Colvin, M. E. *J. Biomol. Struct. Dyn*. **1999**, *16*, 1119. (k) Colson, A.-O.; Sevilla M. D. *Int. J. Radiat. Biol*. **1995**, *67*, 627. (l) Shiskhin, O. V.; Šponer, J.; Hobza, P. *J. Mol. Struct.* **1999**, *477*, 15. (m) Raimondi, M.; Famulari, A.; Gianinetti, E. *Int. J. Quantum Chem*. **1999**, *74*, 259.

(4) (a) Florian, J.; Šponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884. (b) Kratochvil, M.; Engkvist, O.; Šponer, J.; Jungwirth, P.; Hobza, P. *J. Phys. Chem*. A **1998**, *102*, 6921. (c) Pohorille, A.; Burt, S. K.; MacElroy, R. D. *J. Am. Chem. Soc.* **1984**, *10*, 402. (d) Danilov, V. I.; Tolokh, I. S. *J. Biomol. Struct. Dyn.* **1984**, *2*, 119. (e) Cieplak, P.; Kollman, P. A. *J. Am. Chem. Soc*. **1988**, *110*, 3334. (f) Friedman, R. A.; Honig, B. *Biophys. J.* **1995**, *69*, 1528. (g) Arora, N.; Jayram, B. *J. Phys. Chem.* B **1998**, *102*, 6139. (h) Danilov, V. I.; Zheltovsky, N. V.; Slyusarchuk, O. N.; Poltev, V. I.; Alderfer, J. L. *J. Biomol. Struct. Dyn.* **1998**, *15*, 69. Subramanian, V.; Sivanesan, D.; Ramasami, T. *Chem. Phys. Let*. **1998**, *290*, 189.

(5) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1997**, *101*, 9489.

(6) (a) Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1997**, *101*, 9670. (b) Šponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951. (c) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, P. Hobza, *J. Phys. Chem. B* **1999**, *103*, 2528. (d) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J. Hobza, P. *J. Biomol. Struct. Dyn.* **1998**, *16*, 139. (e) Šponer, J.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1999**, *17*, 61. (f) Gresh, N.; Šponer, J. *J. Phys. Chem. B* **1999**, *103*, 11415.

(7) Hobza, P.; Šponer, J. *Chem. Phys. Lett.* **1998**, *288*, 7.

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

(9) Brameld, K.; Dasgupta, S.; Goddard, W. A., III *J. Phys. Chem. B* **1997**, *101*, 4851.

(10) (a) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117 (b) Šponer, J.; Leszczynski, J.; Hobza, P. In *Computational Chemistry. Reviews of Current Trends;* Leszczynski, J., Ed.; World Scientific

Publisher: Singapore, 1996; p 185. (c) Šponer, J.; Hobza, P. *Encycl. Comput. Chem.* **1998**, 777. (d) Šponer, J.; Leszczynski, J.; Hobza, P. In *Theoretical Computational Chemistry. Vol. 8, Computational Molecular Biology*; Leszczynski, J., Ed.; Elsevier: New York, 1999; p 85. (e) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247

(11) (a) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149. (b) Desfrancois, C.; Abdoul-Carime, H.; Schulz, C. P.; Schermann, J.-P. *Science* **1995**, *269*, 1707 (c) Schnier, P. D.; Klassen, J. S.; Stritmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc*. **1998**, *120*, 9605.

(12) (a) Hobza, P.; Kabelác, M.; Šponer, J.; Mejzlik, P.; Vondrášek, J. *J. Comput. Chem.* **1997**, *18*, 1136. (b) Cieplak, P. *Encycl. Comput. Chem.* **1998**, 1922.

(13) (a) Kuechler, E.; Derkosch, J. Z. *Naturforsch. B: Anorg. Chem., Org. Chem.* **1966**, *21b*, 209. (b) Pitha, J.; Jones, R. N.; Pithova, P. *Can. J. Chem*. **1966**, *44*, 1045. (c) Ts'o, P. O. P.; Melvin, I. S.; Olson, A. C. *J. Am. Chem. Soc*. **1963**, *85*, 1289. (d) Gray, D. M. *Biopolymers* **1997**, *42*, 783 and references therein. (e) SantaLucia, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460. (f) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993. (g) Sartorius, J.; Schneider, H.-J. *J. Chem. Soc., Perkin Trans.* **1997**, *2*, 2319 and references therein. (h) Turner, D. H.; Freier, S. M.; Sugimoto, N. *Annu. Rev. Biophys. Biophys. Chem.* **1998**, *17*, 167. (i) Mizutani, M.; Kubo, I.; Jitsukawa, K.; Masuda, H.; Einaga, H. *Inorg. Chem.* **1999**, *38*, 420 (j) Sigel, R. K. O.; Lippert, B. *Chem. Commun*. **1999**, 2167. (k) Nakano, S.-I.; Fujimoto, M.; Hara, H.; Sugimoto, N. *Nucleic Acid. Res.* **1999**, *27*, 2957.

(14) (a) Jeziorski, B.; Moszczynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887. (b) Jeziorski, B.; Szalewicz, K. *Encycl. Comput. Chem.* **1999**, 1376. (c) Saebo, S.; Pulay, P. *J. Chem Phys*. **1987**, *86*, 914 (d) Head-Gordon, M. 5th International Conference "Computers in Chemistry", Szklarska Poreba, Poland, June 1999.

(15) (a) Kristyan, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175. (b) Hobza, P.; Šponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *17*, 1315. (c) Wesolowski, T. A.; Parisel, O.; Ellinger, Y.; Weber, J. *J. Phys. Chem. A* **1997**, *101*, 7818. (d) Wesolowski, T. A.; Ellinger, Y.; Weber, J. *J. Chem. Phys.* **1998**, *108*, 6078. (e). Perez-Jorda, J. M.; San-Fabian, E.; Perez-Jimenez, A. J. *J. Chem. Phys*. **1999**, *110*, 1916 and references therein. (f) Kurita, N.; Araki, M.; Nakano, K.; Kobayashi, K. *Chem. Phys. Lett.* **1999**, *313*, 693.

(16) (a) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys*. **1996**, *105*, 11024. (b) Hobza, P.; Bludsky, O.; Suhai, S. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3073. (c) Hobza, P.; Havlas, Z. *Theor. Chem. Acc.* **1998**, *99*, 372. (d) Famulari, A.; Gianinetti, E.; Raimondi, M.; Sironi, M.; Vandoni, I. *Theor. Chem. Acc.* **1998**, *99*, 358. (e) Famulari, A.; Gianinetti, E.; Raimondi, M.; Sironi, M. *Int. J. Quantum Chem.* **1998**, *69*, 151.

(17) Hobza, P.; Šponer, J. *J. Mol. Struct. (THEOCHEM)* **1996**, *388*, 115.

(18) (a) Harrison, R. J.; Bartlett, R. J. *Int. J. Quantum Chem.* QCS **1986**, *20*, 437. (b) Novoa, J. J.; Sosa, C. *J. Phys. Chem.* **1995**, *99*, 15837.

(19) Tsuzuki, S.; Uchimaru, T.; Matsumara, K.; Mikami, M.; Tanabe, K. *J. Chem. Phys.* **1999**, *110*, 11906.

(20) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18970.

(21) *Gaussian*; 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, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 1995.

(22) (a) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358 (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(23) (a) Colominas, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem. A* **1999**, *103*, 6200. (b) Hobza, P.; Muelder, F.; Sandorfy, C. *J. Am. Chem. Soc.* **1982**, *104*, 926. (c) Neuheuser, T.; Hess, B. A.; Reutel, C.; Weber, E. *J. Phys. Chem*. **1994**, *98*, 6469. (d) Dixon, D. A.; Dobbs, K. D. *J. Phys. Chem.* **1994**, *98*, 13435. (e) Florian, J.; Johnson, B. G. *J. Phys. Chem.* **1995**, *99*, 5899. (f) Suhai, S. *J. Chem. Phys.* **1995**, *103*, 7030. (g) Ben-Tal, N.; Sitkoff, D.; Topol, I. A.; Yang, A.-S.; Burt, S. K.; Honig, B. *J. Phys. Chem.* **1997**, *101*, 450. (h) Colominas, C.; Teixido, J.; Cemeli, J.; Luque, F. J.; Orozco, M. *J. Phys. Chem*. **1997**, *102*, 22269. (i) Aleman, C.; Vega, M. C.; Tabernaro, L.; Bell, J. *J. Phys. Chem.* **1996**, *100*, 11480.

(24) Kroon-Batenburg, L. M. J.; van Duijneveldt, F. B. *J. Mol. Struct.* **1985**, *121*, 185.

(25) (a) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104. (b) Feyreisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993.

(26) Bludský, O.; Šponer, J.; Leszczynski, J.; Špirko, V.; Hobza, P. *J. Chem. Phys.* **1996**, *105*, 11042.