Stabilization Energies of the Hydrogen-Bonded and Stacked Structures of Nucleic Acid Base Pairs in the Crystal Geometries of CG, AT, and AC DNA Steps and in the NMR Geometry of the 5′-d(GCGAAGC)-3′ Hairpin: Complete Basis Set Calculations at the MP2 and CCSD(T) Levels

Iwona Dąbkowska,[†,‡] Haydee Valdes Gonzalez,[†] Petr Jurečka,[†] and Pavel Hobza*[,†]

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic, and Department of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland

Received: July 22, 2004; In Final Form: September 27, 2004

Stabilization energies of the H-bonded and stacked structures of a DNA base pair were studied in the crystal structures of adenine−thymine, cytosine−guanine, and adenine−cytosine steps as well as in the 5′-d(GCGAAGC)-3′ hairpin (utilizing the NMR geometry). Stabilization energies were determined as the sum of the complete basis set (CBS) limit of MP2 stabilization energies and the $\Delta E^{CCSD(T)} - \Delta E^{MP2}$ correction term evaluated with the 6-31G*(0.25) basis set. The CBS limit was determined by a two-point extrapolation using the aug-cc-pVXZ basis sets for X = D and T. While the H-bonding energies are comparable to those of base pairs in a crystal and a vacuum, the stacking energies are considerably smaller in a crystal. Despite this, the stacking is still important and accounts for a significant part of the overall stabilization. It contributes equally to the stability of DNA as does H-bonding for AT-rich DNAs, while in the case of GC-rich DNAs it forms about one-third of the total stabilization. Interstrand stacking reaches surprisingly large values, well comparable to the intrastrand ones, and thus contributes significantly to the overall stabilization. The hairpin structure is characterized by significant stacking, and both guanine···cytosine pairs possess stacking energies larger than 11.5 kcal/mol. A high portion of stabilization in the studied hairpin comes from stacking (similar to that found for AT-rich DNAs) despite the fact that it contains two GC Watson−Crick pairs having very large H-bonding stabilization. The DFT/B3LYP/6-31G** method yields satisfactory values of interaction energies for H-bonded structures, while it fails completely for stacking.

Introduction

Intermolecular interactions among nucleic acid (NA) bases play an important role in assembly of various architectures of DNA and RNA, and by far the most dominant among them are hydrogen-bonding (H-bonding) and stacking interactions.[1] The relative importance of these contributions, playing a key role in determining the structure and dynamics not only of DNA and RNA but also of many complexes of these biomacromolecules with drugs, was not known for a long time. It was generally accepted that H-bonding is much stronger and thus contributes dominantly to the stability of DNA and RNA, but any rational support for this suggestion was absent. This idea was probably based on a different origin of H-bonding and stacking (electrostatic and dispersion interactions) and on a belief that the former interactions should systematically be more stable. Experimental justification of the relative role of H-bonding and stacking is not easy. There exist several studies[2−4] determining the relative strength of both contributions, but these experiments concern the change of the free energy, and furthermore, they all include solvation and desolvation contributions. Thus, for example, one experiment[2,3] indicated that both contributions in RNA are comparable and contribute to the total stability by (only) about 1 kcal/mol (the stabilization energy of H-bonded base pairs lies between 15 and 30 kcal/mol). On the other hand, the relative importance of both contributions can be elucidated by using quantum chemical calculations. Evidently, the calculations should be performed at a very high theoretical level excluding the traditional problems connected with extended molecular clusters such as the size of the basis set or the portion of correlation energy covered.

In our previous paper[5] on this issue, we reported on our investigation of the gas-phase interaction of adenine and thymine, and guanine and cytosine, and also their 9- and 1-methyl analogues. Calculations revealed that final stabilization energies were very large, much larger than published up to now.[6,7] For the methyl analogues the following stabilization energies (kcal/mol) were obtained: mA···mT Hoogsteen, 16.3; mA···mT stacked, 13.1; mG···mC Watson−Crick, 28.8; mG···mC stacked, 18.0. On the basis of comparison with experimental data,[8] it was, however, concluded that they represent the lower boundary of the true stabilization energies. While the H-bonded energies were already close to the true values, the stacked energies were estimated to be still too low by about 10%. In our paper[5] we concluded that surprisingly large values of stacked energies can change the current view of the importance of the specific H-bonding interactions and nonspecific stacking interactions in DNA. It is, however, clear that the geometries of gas-phase pairs differ from the crystal geometries. While the H-bonded pairs in the gas phase and

* To whom correspondence should be addressed. E-mail: hobza@uochb.cas.cz.

† Academy of Sciences of the Czech Republic and Center for Complex Molecular Systems and Biomolecules.

‡ University of Gdańsk.

10.1021/jp046738a CCC: $30.25
Published on Web 01/26/2005

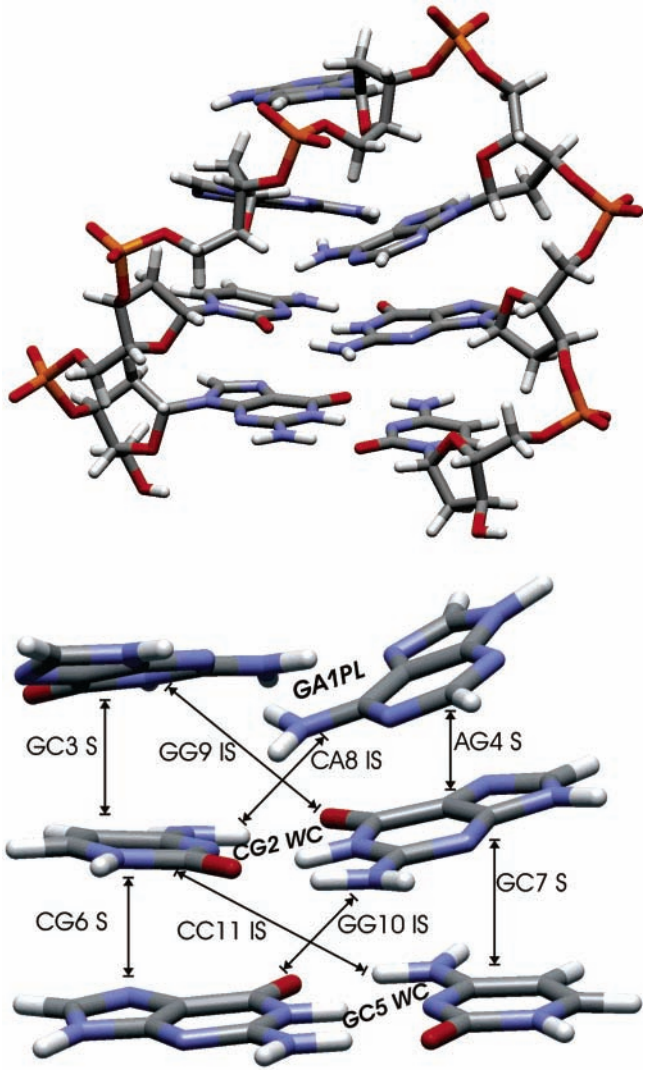

Figure 1. Structures of the (a, top) full 5′-d(GCGAAGC)-3′ hairpin and (b, bottom) investigated nucleic acid base pairs in the geometry of the hairpin structure.

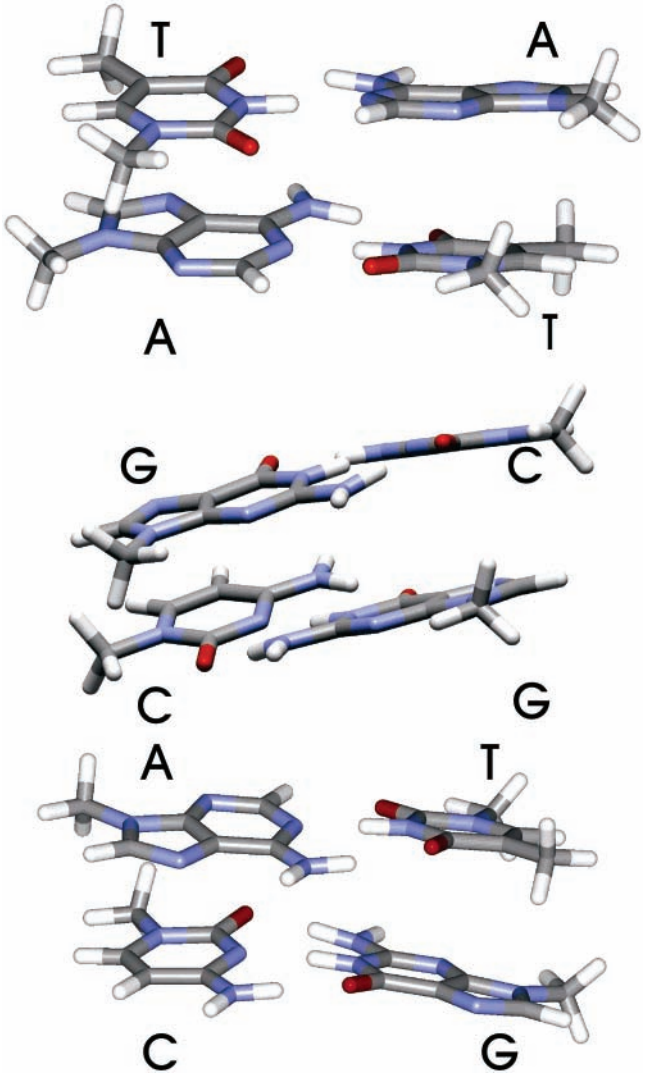

Figure 2. Methylated nucleic acid base pairs in the geometries of crystal B-DNA: (a, top) AT step, (b, middle) CG step, (c, bottom) AC step (A, adenine; T, thymine; G, guanine; C, cytosine).

crystal show high similarity, different geometries are expected for stacked pairs.

The aim of the present work is to study H-bonding and stacking interactions in experimental geometry. We will investigate the adenine−thymine, cytosine−guanine, and adenine−cytosine steps in the B-DNA crystal geometries[9,10] and also the 5′-d(GCGAAGC)-3′ hairpin, the geometry of which was determined from the NMR measurement.[11] We will include the hairpin structure by intention, since it is expected that stacking interactions here play a more important role than in a B-DNA architecture. The knowledge of H-bonding and stacking energies will help us to understand the origin and nature of stabilization in DNA, RNA, and other nucleic acid architectures.

Methods

Structures. The geometries of the CG and AC steps were taken from the Dickerson decamer[9] with PDB ID 5DNB, 5′-d(CCA**ACG**TTGG)-3′ (the pairs considered are in bold), and the geometry of the AT step was taken from a PDB ID 1ENN crystal,[10] 5′-(GCGA**AT**TCG)-3′. The geometry of the hairpin originated from NMR measurements[11] (PDB ID 1PQT, sequence 5′-(**GCG**A**AGC**)-3′). In subsequent calculations we considered only atoms of nucleic acid bases; phosphate groups were disregarded and sugar units replaced by methyl groups in the case of B-DNA steps, and just by hydrogen in the case of the hairpin. The geometry of the hairpin contains all atoms (but the cutting points) including hydrogen atoms (final refinement in the original paper was accomplished using the AMBER empirical potential), while crystal structures do not contain hydrogen atoms. The positions of heavy atoms in the latter cases were considered rigid, and the positions of hydrogen atoms at bases as well as at methyl groups were determined by the gradient optimization (B3LYP/6-31G**) of the respective H-bonded base pair. Figure 1 shows the full geometry of the hairpin (a) as well as of investigated combinations of base pairs (b), and in Figure 2 we present the geometries of AT, CG, and AC steps.

The interaction energy was always evaluated for one isolated base pair only, and the rest of the step was neglected. The total interaction energy was then computed as a sum of the pairwise interaction energies, and many-body terms were neglected. Error resulting from this approximation should be small compared to the total interaction energy; in our previous work[12] the largest three-body term among the AG, AT, and GC steps (in slightly different geometries) was found to be 1.18 kcal/mol for the GC step. Nevertheless, to estimate the role of the many-body interactions, the three- and four-body terms were calculated for the CG step, in which their contribution is expected to be the

largest. Calculations were performed at the RI-MP2/aug-cc-pVDZ level and corrected for the basis set superposition error (BSSE) by the counterpoise technique.[13]

Complete Basis Set Limit of the MP2 Stabilization Energies. The HF interaction energy converges with respect to the one-electron basis set already for relatively small basis sets, while the correlation interaction energy converges unacceptably slowly to its complete basis set (CBS) limit. To correct the computed results for basis set incompleteness error, the extrapolation scheme of Helgaker and co-workers[14] has been employed:

$$E_X^{\mathrm{HF}} = E_{\mathrm{CBS}}^{\mathrm{HF}} + A\exp(-\alpha X) \quad \text{and}$$

$$E_X^{\mathrm{corr}} = E_{\mathrm{CBS}}^{\mathrm{corr}} + BX^{-3}$$

where E_X and E_{CBS} are energies for the basis set with the largest angular momentum X and for the complete basis set, respectively, and α is the parameter fitted in the original work. In the present study (as in the previous one) we used augmented Dunning basis sets[15] rather than nonaugmented sets, which we believe reduce the extrapolation error remarkably. The BSSE correction[13] and frozen-core approximation were applied throughout this study. Extrapolation was also applied to the BSSE correction term.

All calculations were carried out using the TURBOMOLE 5.6 program suite,[16] the aug-cc-pVDZ and aug-cc-pVTZ basis sets, and standard (default) auxiliary basis sets. Recently, we explored[12] the applicability of the RI-MP2 method for nucleic acid base pairs and larger DNA fragments and have demonstrated that the method is capable of accurate description of H-bonded and stacked DNA base interactions. The results obtained with the RI-MP2 method differ only marginally from those evaluated with the exact MP2 method, while the time saving is as large as 1 order of magnitude.

Correction for Higher Order Correlation Effects. The difference between CCSD(T) and MP2 interaction energies ($\Delta E^{\mathrm{CCSD(T)}} - \Delta E^{\mathrm{MP2}}$) is known to depend only negligibly on the basis set size,[17] which allows an approximation of the CBS CCSD(T) interaction energy as

$$\Delta E^{\mathrm{CCSD(T)}} = \Delta E_{\mathrm{CBS}}^{\mathrm{MP2}} + (\Delta E^{\mathrm{CCSD(T)}} - \Delta E^{\mathrm{MP2}})|_{\mathrm{small\ basis\ set}}$$

Our previous studies[17,18] show that the 6-31G(0.25) basis set provides a satisfactory value of the CCSD(T)−MP2 difference for the DNA basis, and this basis set was systematically used for all CCSD(T) calculations.

DFT Calculations. The DFT calculations were performed with the B3LYP functional using the 6-31G** basis set, as this combination of the functional and basis set is commonly used in the literature.

Codes. Except CCSD(T) calculations where we used the MOLPRO 2002.6 suite of programs,[19] all calculations were performed using the TURBOMOLE 5.6[16] and GAUSSIAN 03 program packages.[20]

Results and Discussion

Adenine−Thymine step. Table 1 shows interaction energies for all base pairs in the AT step. A Watson−Crick arrangement possessing two H-bonds is considerably more stable than all three remaining stacking structures. Passing from the aug-cc-pVDZ to the aug-cc-pVTZ basis set enlarges the stabilization energy by more than 1 kcal/mol, and extrapolation to the CBS limit yields yet another nonnegligible stabilization energy increase (~0.4 kcal/mol). The MP2−CCSD(T) interaction energy difference is negligible, and this fully agrees with previous results. The final H-bonding energy agrees very well with the gas-phase results and is slightly larger. The difference can be explained by considering the deformation energy in the gas-phase calculations, which reduces the value of the stabilization energy (regarding the crystal DNA step, the deformation energy of the monomers was not taken into account). Among three stacked structures, the intrastrand one possesses the largest stabilization of more than 9 kcal/mol. Passing to the aug-cc-pVTZ and CBS limit levels brings additional stabilization, while inclusion of the higher correlation energy contributions yields destabilization. This destabilization is again in accord with data known from gas-phase DNA base pairs. The stabilization of the intrastrand AT pair is considerably smaller (by about 4 kcal/mol) than that of the optimized gas-phase structure and is due to different geometries in the crystal and gas-phase optimized dimer. The final stabilization energy of the stacked pair is about half that of the Watson−Crick pair. Both interstrand stacked pairs are characterized by small interaction energies (below 1 kcal/mol), and while the AA pair brings small stabilization, the TT pair provides destabilization.

TABLE 1: Interaction Energies (kcal/mol) of DNA Base Pairs in the Adenine−Thymine, Cytosine−Guanine, and Adenine−Cytosine Steps Evaluated with Various Basis Sets[a]

	RIMP2/aDZ[b]	RIMP2/aTZ[b]	D → T[c]	MP2 → CCSD(T)[d]	ΔE[e]	B3LYP[f]
Adenine−Thymine Step						
AT WC	−14.8	−15.9	−16.4	0.0	−16.4	−13.4
AT S	−9.2	−9.8	−10.1	2.0	−8.1	3.6
AA IS	−0.6	−0.8	−0.9	0.2	−0.7	2.1
TT IS	0.9	0.8	0.8	0.2	1.0	1.6
Cytosine−Guanine Step						
GC WC	−32.7	−34.6	−35.4	−0.4	−35.8	
GC S	−8.1	−8.3	−8.3	0.4	−7.9	
GG IS	−4.9	−5.4	−5.6	1.1	−4.5	
CC IS	1.4	1.3	1.3	0.1	1.4	
Adenine−Cytosine Step						
AT WC	−16.6	−17.7	−18.2	−0.2	−18.4	−15.1
GC WC	−32.7	−34.6	−35.4	−0.4	−35.8	−32.7
AC S	−7.7	−8.1	−8.3	1.6	−6.7	4.6
TG S	−7.0	−7.6	−7.9	1.7	−6.2	4.1
AG IS	−4.7	−4.9	−5.0	0.2	−4.8	−1.7
TC IS	−0.2	−0.2	−0.2	0.1	−0.1	0.4

[a] WC, S, and IS are acronyms for Watson−Crick, stacked, and interstrand stacked structures. [b] aDZ and aTZ mean aug-cc-pVDZ and aug-cc-pVTZ basis sets. [c] Extrapolation to the CBS limit using aug-cc-pVDZ and aug-cc-pVTZ energies. [d] Difference between CCSD(T) and MP2 interaction energies determined with the 6-31g* (0.25) basis set. [e] Total interaction energy evaluated as a sum of the CBS RI-MP2 interaction energy and the difference between the CCSD(T) and MP2 interaction energies. [f] 6-31G** basis set.

Cytosine−Guanine Step. Table 1 presents the interaction energies for all the base pairs, and the WC H-bonded pair is now by far the most stable. The final stabilization energy is higher than 35 kcal/mol. Passing from aug-cc-pVDZ to aug-cc-pVTZ and to the CBS description brings about 2 and 1 kcal/mol, respectively. The CCSD(T) correction is slightly negative (though stabilizing). While the H-bonding energies of the crystal and gas-phase optimized AT dimers were similar, in the present case a rather large difference of more than 5 kcal/mol was found (the gas-phase dimer is less stable). The difference is due to considerably larger deformation energies of guanine in comparison with adenine. The stacked CG pair is comparably as stable as the AT pair and again significantly less stable (by about 10 kcal/mol) than the respective pair in the gas-phase structure. The GG interstrand pair brings nonnegligible stabilization which is much larger than that found in the AT step. The overall

TABLE 2: Interaction Energies (kcal/mol) of DNA Base Pairs in the Hairpin Evaluated with Various Basis Sets[a]

	RIMP2/ aDZ[b]	RIMP2/ aTZ[b]	D → T[c]	MP2→ CCSD(T)[d]	ΔE[e]	B3LYP[f]
GA1 HB	−10.2	−11.4	−11.9	−0.6	−11.3	−11.4
CG2 WC	−27.7	−29.5	−30.3	0.4	−30.7	−32.1
GC5 WC	−28.4	−30.2	−31.0	0.4	−31.4	−32.7
GC3 S	−8.3	−8.9	−9.2	−1.5	−7.7	−1.0
AG4 S	−7.8	−8.5	−8.8	−2.3	−6.5	5.4
CG6 S	−12.9	−13.5	−13.8	−1.4	−12.4	−5.0
GC7 S	−11.8	−12.4	−12.7	−1.1	−11.6	−4.6
CA8 IS	−2.7	−2.9	−3.1	−0.1	−3.0	−2.2
GG9 IS	−4.9	−5.0	−5.1	0.1	−5.2	−4.3
GG10 IS	0.4	0.1	0.0	−0.8	0.8	2.7
CC11 IS	2.9	2.9	2.9	−0.2	3.1	3.3

[a] HB, WC, S, and IS are acronyms for hydrogen-bonded, Watson−Crick, stacked, and interstrand stacked structures. [b] aDZ and aTZ mean aug-cc-pVDZ and aug-cc-pVTZ basis sets. [c] Extrapolation to the CBS limit using aug-cc-pVDZ and aug-cc-pVTZ energies. [d] Difference between CCSD(T) and MP2 interaction energies determined with the 6-31g* (0.25) basis set. [e] Total interaction energy evaluated as a sum of the CBS RI-MP2 interaction energy and the difference between CCSD(T) and MP2 interaction energies. [f] 6-31G** basis set.

stabilization from H-bonding equals 71.6 kcal/mol, while that from stacking is only 18.8 kcal/mol. The three- and four-body effects in the CG step are more pronounced (equaling about 1 kcal/mol) and are of repulsive origin.

Adenine−Cytosine Step. The geometry of the CG pair is identical to that of the CG step, and thus, the interaction energies are also identical. On the other hand, the AT WC geometry is different from that in the AT step, what is reflected in higher H-bonding energies (all energies are 1.8 kcal/mol higher). The CCSD(T) correction term is slightly negative (similarly to the case of a GC WC pair). Both stacking energies are slightly smaller than those of the AT and CG pairs, and again, the CCSD(T) correction is relatively large and repulsive. The AG interstrand stacking is comparable to that of GG, while the TC interstrand stacking is negligible.

Both H-bonding energies together provide 54 kcal/mol, while all four stacks contribute 17.8 kcal/mol. Clearly, the present step represents a transition between the CG and AT steps, and there is no remarkable deviation.

Hairpin. In the case of the hairpin we have two GC WC H-bonded pairs and one GA mispair. Table 2 shows that the H-bonding energy of the mispair is smaller than that of WC structures but it is not negligible. The GC WC stabilization energies are similar, and both are smaller than that of the GC WC pair in the CG step. Apparently, the hairpin structure is less organized, which can be attributed to the presence of a mispair. Extrapolation to the CBS limit is important in all three cases and brings for the WC structures more than 2.5 kcal/mol. The CCSD(T) correction term is systematically small and similar in value to those of previously studied GC and AT WC structures.

Among eight stacked structures, the intrastrand interactions are considerably stronger than the interstrand ones. Further, the CG6 and GC7 structures have stabilization higher than 12.5 kcal/mol, which is about 3 kcal/mol stronger than in the case of the most favorable stacked pairs in the AT and CG steps of the regular (Dickerson) double helix. Evidently, with respect to a regular coil the hairpin structure loses some stabilization from the H-bonded pairs but gains additional stability due to increased stacking. Among four interstrand stacked pairs, the largest stabilization was found for the GG10 pair and only the CC11 pair brings destabilization. The CCSD(T) correction term is again repulsive for all the stacked structures except one (GG9), where it is, however, negligible.

Summarizing all H-bonding, we obtain a stabilization of 73.4 kcal/mol, while from stacking we have 42.5 kcal/mol.

Performance of the DFT Method. Tables 1 and 2 show stabilization energies obtained with the popular DFT method using the B3LYP functional and 6-31G** basis set. In the case of H-bonded complexes DFT yields reasonable stabilization energies which for AT WC and GA complexes are comparable to aug-cc-pVDZ results and for both GC WC pairs even with aug-cc-pVTZ ones. On the other hand, in the case of stacking DFT fails completely; AT S at the MP2/aug-cc-pVDZ level is stabilized by about 9 kcal/mol, while DFT yields a repulsion of about 4 kcal/mol. This means the total error is as large as about 13 kcal/mol, which is 150% of the most accurate value. Very similar results were obtained for all remaining stacked pairs. DFT underestimates the stacking attraction in GC3, AG4, CG6, and GC7 by about 7, 13, 8, and 7 kcal/mol. DFT gives stacked energies comparable to that of the MP2 method only if they are small or repulsive (the data presented concern the hairpin structure, but the same holds for AT and CG steps). The finding that DFT underestimates the stacking energy is definitely not new and does not represent any surprise, and we were among the first to point out problems of describing stacking by the DFT method already in 1995.[21] Presenting these data explicitly, we stress again that up to now no functional being able to describe the stacking in DNA base pairs exists. This failure not only is certainly limited to stacking of DNA bases but concerns any interaction governed by the London dispersion energy. Because of the importance of the London dispersion energy in biological systems, we must expand our warning by stating that the use of the present DFT methods in solving problems of the structure, dynamics, energetics, and function of biomacromolecules is not justified and the results obtained should be taken with care.

Many-Body Terms. On the basis of a suggestion of a reviewer, we have also calculated three- and four-body terms for the CG step and the three-body term of the CGT sequence of bases in the CGT step geometry. The dominant contribution to the many-body effects came from the HF calculation (more than 96%), and the effect of correlation (and therefore also dispersion) is negligible. This suggests that many-body contributions to the interaction energy could be evaluated at a lower (noncorrelated) level of theory than the pairwise interactions. Two three-body terms in the (centrosymmetric) CG step amount to 2.3 kcal/mol (CG/C) and 0.2 kcal/mol (CG/G), and the four-body term is −2.3 kcal/mol. Although these values are not negligible, they partially cancel (three- and four-body terms have opposite signs), and the sum of pairwise interaction energies represents the total interaction energy relatively well. However, if accurate results are required, many-body interactions should be included. In the case of the CGT stacked trimer, the three-body term due to stacking amounts to 0.05 kcal/mol and is therefore negligible.

Relative Importance of Stacking and H-Bonding in Different DNA Architectures. Extrapolating results from CG and AT steps to the polyCG and polyAT duplexes, we found that the dispersion energy contributes about one-third (31%) to the total stabilization in the former duplex and half (50%) in the latter. (Numbers in parentheses represent the percent of the vertical stacking of one base pair—two bases—with respect to the total stabilization due to this base pair, computed as a sum of two vertical stacking contributions and one H-bonding contribution.) This means that in the polyAT duplexes the

H-bonding and stacking contribute similarly to the stability of DNA. In the case of the hairpin GC step, deformation of the structure with respect to the purely double helical one brings weakening of hydrogen bonds and also increased overlap of vertical stacking, which results in a larger stacking contribution to the total stability (about 40% if polyGC in the geometry of the hairpin was considered). We can thus conclude by stating that stacking plays an important role especially in AT-rich DNAs and also in the hairpin architectures.

Comparison with Experiment. Throughout the present study only interaction energies in vacuo were considered. In the gas phase, passing from energy to Gibbs energy is connected with substantial destabilization of the complex due to association entropy. As we have shown previously[22] at room temperature (298 K) the $-T\Delta S$ term amounts to 12.2 kcal/mol for the GC WC base pair and 11.1 kcal/mol for the AT WC base pair. The large energy (enthalpy) stabilization is thus significantly reduced when entropy is taken into account.

Furthermore, when passing from the gas phase to a water environment, molecules are stabilized by interaction with water molecules. According to Kollman,[23] the Gibbs hydration energies of 9-methyladenine, 1-methylthymine, 9-methylguanine, and 1-methylcytosine are −12.0, −12.4, −22.4 and −18.4 kcal/mol, respectively. Upon incorporation of the base into the NA backbone, a substantial part (a very rough estimate could be about 80%) of ΔG_{HYDR} is lost, which means that hydration destabilizes the double helix with respect to the uncoiled chain. If we add the stabilizing enthalpy contributions per base pair (e.g., in the AT step the summation of all the stabilization interactions per base pair gives −32.0 kcal/mol) with the destabilizing contribution of lost hydration (80% of 12.0 for Ade and 12.4 for Thy gives 19.5 kcal/mol), we will get a −12.8 kcal/mol stabilization per pair. The difference between this number and the typical excess stabilization Gibbs energies of RNA pairs measured by Turner and co-workers[3] (ranging between −1.6 and −3.4 kcal/mol) is most likely due to loss of conformational entropy and other effects connected with double-helix formation. Unfortunately, as long as the precise magnitudes of hydration and entropy effects are unknown, no better comparison with experimental Gibbs energy values can be given.

Regarding the role of H-bonding and stacking in the stabilization of the double helix, we have found that in the case of AT-rich sequences both interactions contribute almost equally and in the CG-rich sequences stacking makes about one-third of the stabilization. Although these conclusions apply to the situation in vacuo, they are in line with the results of the dangling end experiments of Turner and co-workers,[3] who measured the stabilization Gibbs energy due to H-bonding and stacking in RNA in solution. A possible explanation is that the hydration affects both kinds of interactions in similar ways.

Conclusions

(i) For obtaining reasonable relative values of H-bonding and stacking energies using the experimental geometry, the extrapolation to the complete basis set is essential.

(ii) The CCSD(T) correction term should be systematically included. Similarly to the case of gas-phase geometries, it is small (negligible) for H-bonded structures and large and repulsive for stacked structures.

(iii) The DFT/B3LYP/6-31G** method yields satisfactory values of H-bonded energies, while it fails completely for stacking.

(iv) While H-bonding energies are comparable for base pairs in crystal geometries and in in vacuo optimized geometries, the stacking energies in the crystal are considerably lower.

(v) Stacking contributes equally to the stability of DNA as H-bonding does for AT-rich DNAs, while for GC-rich DNAs it forms about one-third of the total stabilization.

(vi) Interstrand stacking reaches surprisingly large values, well comparable to the intrastrand ones, and thus contributes significantly to the overall stabilization.

(vii) The hairpin structure studied is characterized by a high portion of stacking stabilization (similar to that found for AT-rich DNAs), despite the fact that it contains two GC WC pairs having very large H-bonding stabilization.

Acknowledgment. This work was part of the research project Z4 0550506 and was supported by a grant 203/05/0009 from the Grant Agency of the Czech Republic. I.D. was sponsored by the International Visegrad Fund under Contract No. S-004-2003 and Polish State Committee for Scientific Research (KBN), Grant 4 T09A 012 24. H.V.G. holds an EMBO scholarship (Grant ALTF 580-2003). This research was performed in part using the Molecular Science Computing Facility (MSCF) in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory. Pacific Northwest is operated for the Department of Energy by Battelle.

References and Notes

(1) Müller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143.
(2) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473.
(3) Freier, S. M.; Sugimoto, N.; Sinclair, A.; Neilson, T.; Kierzek, R.; Caruthers, M. H.; Turner, D. H. *Biochemistry* **1986**, *25*, 3214.
(4) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182.
(5) Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608.
(6) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493.
(7) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.
(8) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.
(9) Prive, G. G.; Yanagi, K.; Dickerson, R. E. *J. Mol. Biol.* **1991**, *217*, 177.
(10) Soler-Lopez, M.; Malinina, L.; Subirana, J. A. *J. Biol. Chem.* **2000**, *275*, 23034.
(11) Žídek, L.; Wu, H.; Feigon, J.; Sklenář, U. *J. Biomol. NMR* **2001**, *21*, 153.
(12) Jurečka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578.
(13) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(14) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
(15) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(16) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
(17) Jurečka, P.; Hobza, P. *Chem. Phys. Lett.* **2002**, *365*, 89.
(18) Hobza, P.; Šponer, J. *J. Am. Chem. Soc.* **2002**, *124*, 11802.
(19) MOLPRO, a package of ab initio programs designed by H.-J. Werner and P. J. Knowles version 2002.1: R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and H.-J. Werner.
(20) Gaussian 03, Revision A.1: Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich,

S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A., Gaussian, Inc., Pittsburgh, PA, 2003.

(21) Hobza, P.; Šponer, J., Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315.

(22) Hobza, P.; Šponer, J. *Chem. Phys. Lett.* **1996**, *261*, 379.

(23) Miller, J. L.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 8587.