Base-Base and Deoxyribose-Base Stacking Interactions in B-DNA and Z-DNA: A Quantum-Chemical Study

J. Šponer,*# H. A. Gabb,§ J. Leszczynski,¶ and P. Hobza*

*J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, 182 23 Prague 8, Czech Republic; #Institute of Biophysics, Academy of Sciences of the Czech Republic, 612 65 Brno, Czech Republic; §Imperial Cancer Research Fund, London WC2A 3PX, England; and ¶Department of Chemistry, Jackson State University, Jackson, Mississippi 39217 USA

ABSTRACT Base-stacking interactions in canonical and crystal B-DNA and in Z-DNA steps are studied using the ab initio quantum-chemical method with inclusion of electron correlation. The stacking energies in canonical B-DNA base-pair steps vary from −9.5 kcal/mol (GG) to −13.2 kcal/mol (GC). The many-body nonadditivity term, although rather small in absolute value, influences the sequence dependence of stacking energy. The base-stacking energies calculated for CGC and a hypothetical TAT sequence in Z-configuration are similar to those in B-DNA. Comparison with older quantum-chemical studies shows that they do not provide even a qualitatively correct description of base stacking. We also evaluate the base-(deoxy)ribose stacking geometry that occurs in Z-DNA and in nucleotides linked by 2′,5′-phosphodiester bonds. Although the molecular orbital analysis does not rule out the charge-transfer n-π^* interaction of the sugar O4′ with the aromatic base, the base-sugar contact is stabilized by dispersion energy similar to that of stacked bases. The stabilization amounts to almost 4 kcal/mol and is thus comparable to that afforded by normal base-base stacking. This enhancement of the total stacking interaction could contribute to the propensity of short $d(CG)_n$ sequences to adopt the Z-conformation.

INTRODUCTION

Stacking interactions represent one of the important sources of DNA stability and conformational variability (Calladine and Drew, 1986; Hunter, 1993; Šponer and Kypr, 1991b, 1993a,b; Gorin et al., 1995; Olson and Zhurkin, 1996). Base-stacking interactions are usually studied by means of empirical potentials. Empirical potentials allow a thorough search over a wide range of conformations. However, the information obtained is limited in several ways: 1) the current empirical potentials are pair-additive and do not cover many-body contributions; 2) both the van der Waals parameters and atomic charges used by various force fields differ considerably (Šponer and Kypr, 1991a, 1993b; Rudnicki and Lesyng, 1994; Hobza et al., 1996a, 1997); 3) all empirical potentials are based on the same, very simple analytical expression for the intermolecular interactions, which may not be sufficient to cover all physical contributions (some contributions can even be completely absent).

The uncertainties of empirical potential treatment may be eliminated by using nonempirical (ab initio) quantum-chemical techniques. A proper quantum-chemical evaluation of aromatic stacking is a considerably more difficult task (Hobza et al., 1994a,b, 1995; Smith and Jaffe, 1996; Šponer et al., 1996a,b,d,e)). Stacked complexes are stabilized by the dispersion attraction, which originates in the intermolecular correlation of electron motion. Lower level quantum chemical methods (e.g., semiempirical approaches, Hartree-Fock (HF) ab initio calculations, density functional theory (DFT)), which can be satisfactorily used to characterize hydrogen bonding in bases, fail to evaluate base stacking (Šponer et al., 1996b,d; Hobza et al., 1997). We must especially underline the limited accuracy of all DFT-based ab initio approaches, although they can be used for relatively large systems (Hutter et al., 1996; Carloni and Andreoni, 1996).

Recent improvements in computer hardware and software allowed us to characterize base stacking by using the second-order Møller-Plesset (MP2) perturbation theory, which covers a significant portion of the electron correlation effects (Hobza et al., 1995; Šponer et al., 1996a,b,d,e). The MP2 method, with a diffuse-polarized basis set of atomic orbitals, is a low-cost theoretical tool that is free of empirical parameters and covers most important contributions to the interaction energy. The electrostatic interaction, which is the structure-making term, has converged at this level, as demonstrated by calculations of electrostatic properties of isolated nucleic acid bases with large basis sets of atomic orbitals (Šponer et al., 1996a,b). The dispersion attraction is also properly covered, which is evidenced for smaller aromatic stacking complexes (benzene dimer, aminopyrimidine dimers, etc.). For these complexes, both large basis sets of atomic orbitals and the very accurate coupled cluster method with noniterative triple electron excitations (CCSD(T)) were extensively tested (Hobza et al., 1996b; Šponer and Hobza, 1997). Finally, MP2 calculations with medium-sized diffuse-polarized basis sets of atomic orbitals agree with available gas-phase experimental data on aromatic stacking complexes (benzene···Ar, Krause and Neusser, 1993; Brupbacher et al., 1994; benzene dimer, Arunam and Gutowsky, 1993; Henson et al., 1992). It should be noted that no such experiments are available for base stack-

Received for publication 12 August 1996 and in final form 11 April 1997.

Address reprint requests to Dr. J. Šponer, J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague 8, Czech Republic. Tel.: 420-2-6605-2056; Fax: 420-2-858-2307; E-mail: sponer@indy.jh-inst.cas.cz.

0006-3495/97/07/76/12 $2.00

ing, because nucleic acid bases form hydrogen-bonded complexes in the gas phase.

For nearly a decade, the quantum-mechanical calculations by Aida (Aida, 1988; Aida and Nagata, 1982, 1986) represented reference values for base stacking. These studies were made within the HF approximation with the 4-31G basis set and second-order perturbational evaluation of nonexpanded dispersion energy. The present MP2 calculations with the 6-31G*(0.25) diffuse-polarized basis set represent an improvement, concerning mainly the following items. The 4-31G basis set was too small to provide accurate values of electrostatic energies (polarization functions are necessary) and strongly underestimated the dispersion attraction (diffuse polarization functions are inevitable). Furthermore, the reduction of the dipole moments and dipole-dipole interactions due to the electron correlation could not be considered. Finally, the previous reference values were not corrected for the so-called basis set superposition error.

Recent MP2 studies on base-base interactions sought mainly to characterize in a basic way the conformational space of neutral and protonated base dimers. Special attention is given to the nature of stacking (Hobza et al., 1995; Šponer et al., 1996a,b,d,e), hydrogen bond interactions (Gould and Kollman, 1994; Šponer and Hobza, 1994b,c; Hobza et al., 1995; Florián and Leszczynski, 1995, 1996; Šponer et al., 1996c,d,e), and the ability of various empirical potentials and semiempirical and density functional theory techniques to reproduce the ab initio data (Hobza et al., 1996b, 1997; Šponer et al., 1996a–e). The calculations demonstrated that a properly parameterized empirical potential provides a better description of the nucleic acid-base interactions than any semiempirical, density functional theory or Hartree-Fock ab initio method. Here we extend our previous calculations and evaluate the base-stacking energies in 10 canonical B-DNA base-pair and Z-DNA steps, and compare them with stacking evaluated for crystal B-DNA geometries. Base-pair step stacking energy is calculated as a sum of four individual base-base interactions evaluated by the MP2 method and the many-body correction. The many-body correction covers nonadditivity of interactions and has been estimated at the HF level. This energy contribution is not included in available force fields and has not been studied in previous quantum-chemical studies.

We also study the origin and strength of the purine-(deoxy)ribose stacking interaction. This type of interaction rarely occurs in nucleic acid helices. It does, however, occur in Z-DNA (Wang et al., 1979), RNA tetraloops (Cheong et al., 1990), and various dinucleotides containing 2′,5′-phosphodiester linkages (Krishnan and Seshadri, 1993, 1994). A detailed survey of these contacts has been made recently (Egli and Gessner, 1995). This interaction has not been studied quantum chemically until now, although its presence in Z-DNA has evoked speculation about its origin and importance (Egli and Gessner, 1995). One of the aims of the present study is to clarify whether this interaction exhibits any unusual interaction that could imply a failure of the standard empirical potentials. Z-DNA is likely to exist in vivo (Jaworski et al., 1987; Rhamouni and Wells, 1989; Lukomski and Wells, 1994; Müller et al., 1996) and perform some biological function (Rich, 1994; Wölfl et al., 1996). The presence of Z-DNA binding proteins, for example, indicates that aspects of the Z-conformation are being targeted (Herbert et al., 1993). So, the (*syn*)-purine-deoxyribose stacking geometry is a potentially important protein recognition site. However, no detailed structures exist for protein bound to DNA in the Z conformation at the present time. Despite advances in elucidating Z-DNA biology, the mechanism of the B- to Z-DNA transition itself remains largely unsolved. Structural models for the transition exist (Olson et al., 1983; Harvey, 1983; Saenger and Heinemann, 1989), but none are conclusive and none consider the possible role of the purine-sugar attraction. Also potentially important are the 2′,5′-linked dinucleotides that appear in interferon-treated cells (Kerr and Brown, 1978). Such dinucleotides are known to adopt the (*syn*)-purine-sugar stacking geometry (Krishnan and Seshadri, 1993, 1994) and could be acting as effectors in cellular processes induced by interferon. Proper consideration of the base-sugar stacking geometry could improve structure-based drug design or even suggest a new class of biologically active compounds.

METHOD

Evaluation of interaction energies

The interaction energies were calculated using the second-order Møller-Plesset (MP2) perturbational method. A standard split-valence 6-31G basis set of atomic orbitals, augmented by a set of diffuse d-polarization functions with an exponent of 0.25 added to the second row elements (designated 6-31G*(0.25)), was used (Kroon-Batenburg and van Duijneveldt, 1985; Hobza et al., 1986). The use of modified exponents of polarization functions is quite essential for covering the space between the interacting monomers, to include the dispersion attraction. Fortunately, modification of the exponents of polarization functions does not significantly influence the other contributions (Šponer et al., 1996b). The use of the 6-31G* basis set with standard d-polarization functions (exponent of 0.8) would underestimate base stacking by almost 50% for approximately the same computational requirements.

Although the MP2 method covers a significant portion of electron correlation effects, higher order contributions to electron correlation not included in the MP2 approximation may still be important. Such calculations are now available for a number of smaller, symmetrical systems. A reduction in correlation stabilization energy (mainly dispersion energy) is encountered when passing from MP2 to higher-level calculations (MP4(SDTQ) and CCSD(T) methods) for all stacked complexes containing an aromatic molecule (1–2 kcal/mol per dimer). On the other hand, extension of the basis set increases the dispersion attraction (Hobza et al., 1994a, 1996; Smith and Jaffe, 1996; Šponer and Hobza,

1997), so that both effects compensate. Thus the currently reported MP2/6-31G*(0.25) values of base-stacking energies are expected to be close to the actual values.

The interaction energy of a dimer of nucleobases, E^{int}, consists of two components: the Hartree-Fock (HF) and correlation interaction energies. E^{int} is evaluated as a difference between the energy of the complex and the sum of the monomer energies (supermolecule approach). Both the HF and the correlated components of the interaction energy are corrected for the basis set superposition error.

The stacking interaction energy of the base-pair step is evaluated as a sum of four MP2 base-base contributions (Fig. 1). In addition, many-body effects are estimated by evaluating the difference (many-body correction) between base stacking calculated as the interaction of two base pairs and as the sum of four base-base interactions. These calculations can only be made at the HF/6-31G*(0.25) level. The HF approximation covers first-order exchange and induction (deformation) nonadditivities only. We are not able to estimate the dispersion nonadditivities, which could certainly affect the total nonadditivities. The evaluation of higher order nonadditivities, including the dispersion nonadditivity, would require at least the MP3 level of calculations (Chalasinski and Szczesniak, 1994). Such calculations are not currently feasible.

Ab initio calculations are performed using the Gaussian94 program suite (Frisch et al., 1994). The scf = tight option must be used to evaluate the many-body correction, because the standard single-point cutoffs for the HF procedure are insufficient for the present system.

Geometries of stacked base-pair steps

The base stacking energies were calculated for several sets of geometries. First, starting from the B-DNA fiber diffraction geometry (Arnott et al., 1980), the 10 canonical dinucleotide base steps were built and minimized using the JUMNA program and the FLEX nucleic acid force field (Lavery et al., 1986; Lavery, 1988). Using the optimization, we tried to account for the possible backbone contributions. The initial B-DNA geometry is characterized by the following parameters: rise 3.38 Å, helical twist 36°, propeller twist 0°, tilt 0°, inclination 0°, base pair roll 0°, and buckling 0°. During minimization, mononucleotide symmetry is imposed on the structure, and base-pair rise, tilt, inclination, propeller twist, and buckle (as defined by the Cambridge conventions; Dickerson et al., 1989) are held fixed to maintain B-form helical geometry. In addition, ApT, GpC, CpG, and TpA Z-DNA steps were built in a similar fashion, but with important differences. Starting from the original $d(CG)_3$ crystal structure (Wang et al., 1979), the central nucleotides were replaced with the desired sequence, using the Curves program (Lavery and Sklenar, 1988), and minimized by using JUMNA without symmetry constraints but with the same parameters as above locked at their starting values. We built the Z-DNA steps in this fashion to obtain a "standard" Z configuration.

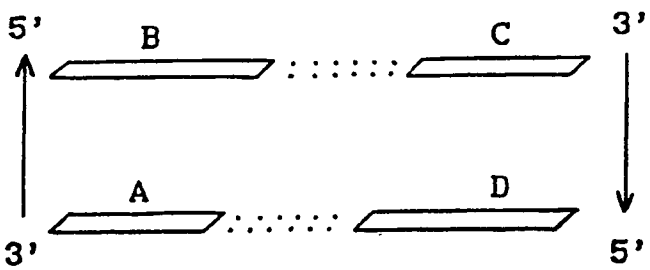

FIGURE 1 Stacking of a base-pair step. AD···BC is being approximated as a sum of four base-base contributions (A···B, C···D, A···C, B···D) evaluated at the MP2/6–31G*(0.25) level, plus the many-body correction (AD···BC interaction energy minus the base-base contributions) evaluated at the HF/6–31G*(0.25) level.

During the review process, one of the referees criticized the use of force-field preoptimizated structures in subsequent ab initio calculations. Potential energy optimization substantially relaxes the DNA fiber diffraction structure (Lavery and Hartmann, 1994). So to make our results more easily reproducible by other investigators and to eliminate any force-field dependence in our results, a second set of B-DNA geometries was prepared. We used optimized geometries of AT and GC Watson-Crick base pairs obtained at the HF/6-31G** level (Šponer et al., 1996c). We calculated the centers of the two C6–C8 base pair axes and prepared base-pair steps where all parameters were 0, except rise (3.38 Å) and helical twist (36°). The C6–C8 axis centers were stacked directly above one another; i.e., displacements were zero. No further adjustment has been made. These geometries can be easily reproduced and used to test other computational techniques. Throughout this paper, the undisplaced standard B-DNA geometries will be referred to as set I, and those produced by JUMNA will be referred to as set II. Fig. 2 shows the stacking of the GG steps in these two arrangements. For all other steps the set II geometry is very close to that for the GG step. The two sets of geometries differ in two parameters: geometry set II has twist 32°-34°, and the Y-displacement is −2.5 Å to −3.5 Å.

We did not consider any further geometries for the Z-DNA steps. The constrained JUMNA optimization did not introduce any visible changes in the geometry. There are no experimental geometries for the hypothetical AT and TA Z-DNA steps, so we must rely on model geometries.

The present calculations on canonical B-DNA structures were carried out with propeller twist of zero for the following reason. Propeller twist influences stacking in a sequence-dependent way. Thus the use of an average propeller twist would improve the energy in some steps, whereas in some others it would cause artificial destabilization due to steric clashes (Šponer and Kypr, 1990, 1993a). To estimate the influence of propeller twist and other local conformational variations, the present data will be compared with calculations using geometries from high-quality B-DNA decamer structures: all five independent steps of the $d(CCAACGTTGG)_2$ B-DNA decamer at 1.5-Å resolution and

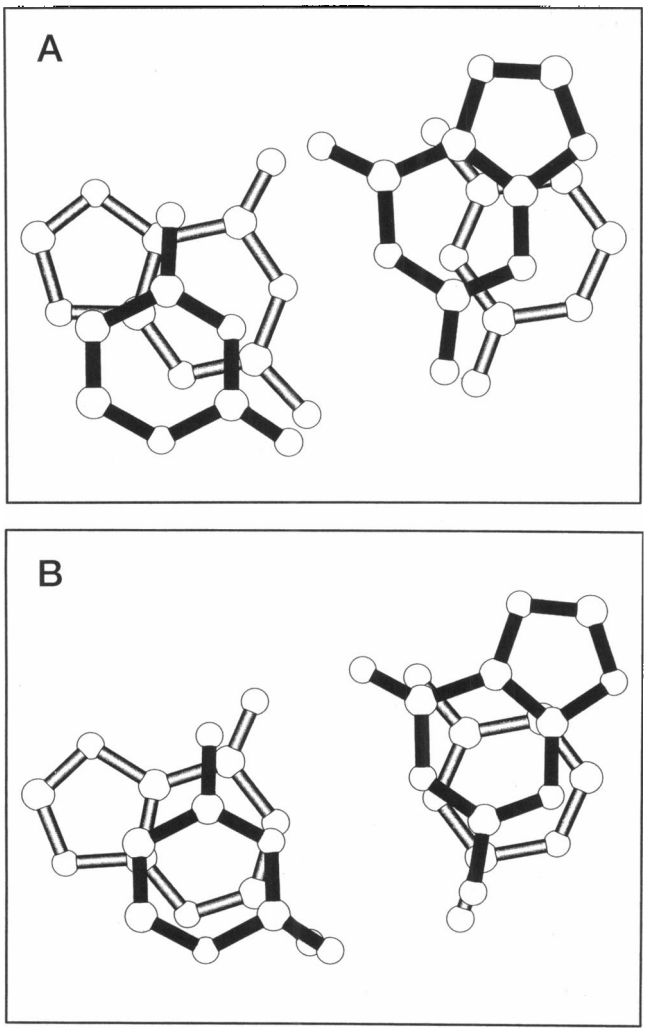

FIGURE 2 B-DNA GG base pair step. (*A*) Geometry I (undisplaced); (*B*) geometry II (displaced, processed by JUMNA).)

the central GC step from the B-DNA decamer d(CCAGGCCTGG)$_2$ solved to a resolution of 1.6 Å. The base-base MP2/6-31G*(0.25) interaction energies are taken from an earlier study (Šponer et al., 1996a), but with the many-body term recalculated at the HF/6-31G*(0.25) level.

In quantum-chemical calculations only atoms of bases were considered and the glycosidic bonds are replaced by hydrogen atoms. Furthermore, "force-field" base geometries (set II) are replaced by planar MP2-optimized geometries of bases (Šponer and Hobza, 1994a), as in our previous study (Šponer et al., 1996a). Use of nonoptimized geometries of bases could affect the dipole moments and electrostatic interaction energies.

Analysis of the sugar-base stacking

The initial geometry of the sugar-base complex was obtained by starting from a crystal structure (Gessner et al., 1989). The phosphate is replaced by a hydrogen, and the C5′ atom is replaced by either a hydrogen or an optimized methyl group. The geometry of ribose is relaxed (i.e., optimized with all dihedral angles involving crystallographically observable atoms frozen according to the respective nucleotide/dinucleotide structure) at the ab initio HF/6-31G** level. The optimized position of the Mg^{2+} cation with respect to the guanine has been obtained at the HF/6-31G* level. The metal ion is placed between the N7 and O6 atoms of guanine (Burda et al., 1996). To measure the spatial orientation of the furanose oxygen with respect to the pyrimidine ring of the guanine base, a simple vector algebraic algorithm is used (Gabb et al., 1996). First, calculate the mean plane and the geometric center of the pyrimidine ring of the purine base. Next, extend a cylinder normal to this mean plane with its origin at the geometric center. The radius of the cylinder is equal to the distance between the ring geometric center and the midpoint of the C4-C5 bond of the purine. Let **S** be the normal vector extending from the ring geometric center. The vector **V** represents the distance from the ring geometric center to the furanose oxygen. Two parameters, z and r, describe the position of the furanose oxygen with respect to the pyrimidine ring of the purine base. These two parameters are calculated using the following equations:

$$\phi = \cos^{-1}[(\mathbf{V} \cdot \mathbf{S})/(|\mathbf{V}| \cdot |\mathbf{S}|)]$$

where $\mathbf{V} \cdot \mathbf{S}$ is the dot product of the two vectors and

$$z = |\mathbf{V}|\cos(\phi)$$

$$r = |\mathbf{V}|\sin(\phi)$$

So z is the vertical distance of the oxygen over the base and r is the lateral distance of the oxygen from the center of the base ring.

RESULTS AND DISCUSSION

Base stacking

Interstrand and intrastrand stacking

Table 1 shows the total stacking energies and components for the base-pair steps: the canonical undisplaced B-DNA geometry (set I), geometries that were relaxed by the JUMNA program starting from the fiber DNA (set II, in parentheses), Z-DNA structures, and six B-DNA and two Z-DNA steps from crystals. Thus our present analysis is based on a comparison of four different sets of geometries, covering a large sample of structures.

The total stacking energy in B-DNA ranges from −9.5 kcal/mol (GG) to −13.2 kcal/mol (GC) for standard geometry set I, from −8.3 kcal/mol (GG) to −15.8 kcal/mol (GC) for geometry set II, and from −8.9 kcal/mol (GG) to −12.6 kcal/mol (GT) for the crystal data. The stabilization is dominated by the intrastrand contribution to the interaction energy, which ranges (for set I) from −18.1 kcal/mol (GC) to −4.6 kcal/mol (GG). The rather unstable intrastrand

TABLE 1 Stacking energies for base-pair steps in B and Z conformation

Base-pair step	E_{intra}	E_{inter}	E_{MB}	E_T
AA*	−9.8 (−8.6)	−2.2 (−2.7)	−0.1 (−0.1)	−12.0 (−11.4)
AA crystal#	−9.5	−2.3	0.0	−11.8
GG*	−4.6 (−4.9)	−6.9 (−5.6)	+2.0 (+2.1)	−9.5 (−8.3)
GG crystal#	−3.9	−7.4	+2.4	−8.9
TC*	−12.5 (−11.5)	+0.4 (−0.1)	+0.7 (+0.4)	−11.4 (−11.2)
CT*	−11.9 (−10.8)	−0.3 (−0.6)	+0.7 (+1.2)	−11.5 (−10.1)
AT*	−10.3 (−11.7)	−0.4 (−0.4)	+0.1 (+0.1)	−10.6 (−12.0)
AT Z-DNA	−10.6	−1.3	+0.6	−11.3
GC*	−18.1 (−18.5)	+4.0 (+3.9)	+0.9 (−0.9)	−13.2 (−15.6)
GC crystal#	−16.6	+5.0	+0.1	−11.5
GC Z-DNA	−18.2	+5.1	−0.5	−13.6
GC Z-DNA#§	−17.2	+3.7	−0.5	−14.0
GT*	−9.0 (−9.5)	−3.3 (−3.0)	+0.6 (0.0)	−11.8 (−12.5)
GT crystal#	−9.9	−3.3	+0.6	−12.6
TA*	−10.1 (−8.7)	−1.1 (−3.3)	0.0 (−0.2)	−11.2 (−12.2)
TA Z-DNA	−6.1	−5.3	+0.4	−11.0
CG*	−11.1 (−8.4)	−2.7 (−4.2)	+0.7 (+1.0)	−13.1 (−11.6)
CG crystal#	−9.4	−3.7	+0.8	−12.3
CG Z-DNA	−10.1	−1.6	−1.0	−12.7
CG Z-DNA#§	−10.3	−1.4	−0.9	−12.6
CA*	−8.5 (−6.9)	−4.4 (−5.3)	+0.6 (+0.9)	−12.3 (−11.3)
CA crystal#	−7.7	−4.8	+0.6	−11.9

E_{intra} is the sum of the two intrastrand stacking contributions, evaluated at the MP2/6-31G*(0.25) level; E_{inter} is the sum of the two interstrand stacking contributions, evaluated at the MP2/6-31G*(0.25) level; E_{MB} is the many-body correction evaluated at the HF/6-31G*(0.25) level. Total stacking energy $E_T = E_{intra} + E_{inter} + E_{MB}$. The glycosidic bonds were replaced by hydrogens.

*Standard B-DNA, geometry set I. Values in parentheses, geometry set II (see Method).

#Geometries from high-resolution crystal structures, interstrand and intrastrand contributions taken from Šponer et al. (1996a).

§Note that in our previous paper (Šponer et al., 1996a), the Z-DNA steps are mislabeled. They should be G2C3 and C3G4 instead of C2G3 and G3C4.

stacking in the GG steps originates in repulsive intrastrand dipole-dipole interactions. The interstrand contribution to the stacking energy ranges (for set I) from −6.9 kcal/mol (GG) to +4.0 kcal/mol (GC). The interstrand and intrastrand contributions tend to compensate each other. The large interstrand attraction in the CA step is due to the guanine-adenine interaction, which influences the unique stacking properties of CpA(TpG) steps (Šponer et al., 1994).

Nonadditivity of stacking

The many-body corrections are rather small and vary from −0.9 kcal/mol (GC, set II) to +2.4 kcal/mol (GG, crystal geometry). However, they influence the sequence dependence of stacking energy. Many-body correction changes the stability order for standard geometry I. Specifically, the GG step becomes the least stable upon inclusion of the many-body term. The many-body term also increases the energy difference between the most stable and least stable steps for geometry set II (from 4.4 to 7.4 kcal/mol) and for the crystal geometries (from 1.9 to 3.4 kcal/mol). The GG step with evidently repulsive intrastrand electrostatics has a repulsive many-body term. The GC step has a repulsive many-body term in geometry set I, but an attractive many-body term in geometry II with a difference between these two geometries of 1.8 kcal/mol. Thus the many-body term changes faster than all other contributions and is responsible for the energy difference of 2.3 kcal/mol in favor of geometry II. We investigated other geometries of the GC step, starting from the geometry in set I, with various values of Y-displacement, mimicking the I-II transition. The many-body term was initially +0.9 kcal/mol (Y-displacement of 0, twist 36°), and then was continuously reduced: +0.5 kcal/mol (−0.85 Å, 35°), +0.1 kcal/mol (−1.7 Å, 34°), −0.5 kcal/mol (−2.55 Å, 33°), and finally, −0.9 kcal/mol for geometry II. Table 1 clearly shows that the many-body term can be large in absolute value and highly variable in steps consisting of two GC base pairs. It is always negligible in steps consisting of two AT base pairs. The present results indicate that the bases are polarized by the electric field created by the other bases. To further illustrate the induction origin of nonadditivity, we considered a GG step with a twist of 0, which maximizes the intrastrand dipole-dipole repulsion (other parameters as above for geometry I). The very repulsive intrastrand electrostatic interaction is expected to be enhanced by the polarization. Indeed, the three-body term increased from +2.0 kcal/mol (twist of 36°) to +2.8 kcal/mol. Furthermore, we added a third GC base pair (i.e., two consecutive GG steps with a twist of 0), which increased the many-body correction to +6.9 kcal/mol, or +3.4 kcal/mol per base-pair step. The enhancement of polarization effects when increasing the size of the cluster (and intensity of the electric field) is well documented for smaller systems, such as formamide clusters and crystal (Suhai, 1994) and water clusters (Saykally, 1996, and references therein). Such effects are not included in the present empirical potentials for nucleic acids. Polarizable potentials are under development and are currently being testing for

simpler systems (Ding et al., 1995; Caldwell and Kollman, 1995; Bernardo et al., 1994; Sun et al., 1995; Alkorta and Perez, 1996; Engkvist et al., 1996; Aastrand et al., 1994).

Stacking in B-DNA crystals and Z-DNA

The crystal-stacking energies are similar to those for idealized structures. The central GpC step of the d(CCAGGCCTGG)$_2$ decamer is rather unstable. The reason is not known, although it has been suggested that central steps of crystals with twofold symmetry are more restricted by the crystal environment than the remaining steps (Šponer and Hobza, 1994b). Unfortunately, we cannot make a more exhaustive scan of the conformational space of the stacked dimers. In addition, the solvent (Friedman and Honig, 1992) and entropy contributions, both difficult to quantify, are likely to be important. A comparison with crystals could be hampered by artifacts in crystal structures, x-ray data, and refinement inaccuracies (Šponer and Kypr, 1993b). Our data confirm that stacking interactions are very flexible and allow a wide range of geometries with similar stability.

Base stacking in Z-DNA has a strength comparable to that of B-DNA for both CGC and TAT alternating sequences. Moreover, the individual contributions exhibit some similarity in B-DNA and Z-DNA. We did not find any striking explanation based on stacking of why alternating CG sequences readily form Z-DNA whereas alternating TA sequences do not. The present analysis does not cover the free energy of base stacking (Pearlman and Kollman, 1990; Dang et al., 1990). The sequence preference for the B-to-Z transition may be determined primarily by the solvent free energy and the solvent-accessible surface (Tagawa et al., 1989; Mooers et al., 1995).

Empirical potential calculations

An important question is, how accurately can empirical potentials approximate ab initio data? We recently compared MP2 interaction energies for almost 300 geometries of neutral stacked base dimers with data obtained by using a simple pairwise-additive empirical potential (Šponer et al., 1996a,b,d,e). The potential consisted of a common Lennard-Jones term and a standard Coulombic term with atom-centered point charges. Proper evaluation of the electrostatic term is crucial to achieving an agreement with the ab initio data. The charges were fitted to reproduce the molecular electric potential around isolated bases with the 6-31G*(0.25) basis set and with inclusion of the electron correlation using the MP2 method. More details concerning these empirical potential calculations, including a list of the atomic charges and specification of the van der Waals term, may be found in a previous paper (Šponer et al., 1996a). The agreement between the ab initio data and this empirical potential is remarkable, which strongly suggests that atom-centered point charges derived from molecular electric potentials with a sufficiently accurate method provide an excellent approximation of the electrostatic interaction energy. These charges effectively include higher multipoles. Thus there is no reason to use additional charges (for example, out-of-plane, "π") (Šponer et al., 1996a,b). Our findings support the use of the standard empirical potential form currently being used in many classical mechanical studies to treat electrostatics. The quality of such potentials depends mainly on the method by which the charges were derived (for a survey of recently introduced potentials, see Hobza et al., 1997) and the ability of the van der Waals term to reproduce the vertical separation of base pairs in DNA (Šponer and Kypr, 1993b).

The same comparison for B-DNA geometry set I is made in Table 2. There is again a good agreement between the empirical potential and ab initio data (the agreement for the other sets of geometries was similar (not shown); the data for crystal geometries were published previously (Šponer et al., 1996a)).

Table 2 shows all individual base-base contributions. For a given step, the first row contains the ab initio MP2 values. The next row shows the electrostatic energy evaluated with the potential-derived charges, and the numbers in parentheses show the total empirical potential interaction energy after the Lennard-Jones potential is added. The empirical potential and ab initio data correlate well. Because of the large dipole moments of guanine and cytosine, base-pair steps consisting of two GC pairs ensure larger variations of base-stacking energy along the helix. This stacking energy variability is much smaller for the remaining steps. The electrostatic interaction energies shown in Table 2 support this statement. The total values of electrostatic interaction energies do not vary significantly, from +0.2 kcal/mol (CG) to +3.4 kcal/mol (GG). However, the difference between the GC and AT pairs is revealed when the sum of absolute values of the four individual base-base electrostatic contributions is considered. This sum is 12–13 kcal/mol for the GG and GC steps, and smaller than 5 kcal/mol for the remaining eight steps. This quantity exhibits especially small variations for steps consisting of two AT pairs, 2.6–3.0 kcal/mol. The individual base-base electrostatic contributions in these three steps only vary from −0.3 kcal/mol to +1.1 kcal/mol. In other words, the GC step influences stacking more by electrostatic contributions, whereas AT stacking is more dispersion-controlled. It is interesting to note that this trend is promoted by the pyrimidine methylation. Reversion of methylation (U instead of T, and ^{5m}C instead of C) would reduce the above-mentioned difference between stacking properties of GC and AT base pairs (cf. also discussion in Wang and Kool, 1995).

Comparison with previous studies

Fig. 3 compares the present MP2 ab initio data (geometry set I without the many-body correction) with some older computational procedures. (Note that the geometries used here and those used in the previous studies differ, so Fig. 3 should be searched for qualitative trends only.) Among the previous quantum-chemical studies, the best agreement is

TABLE 2 Comparison of the ab initio stacking energies (without the many-body term), and the empirical potential data, standard B-DNA, geometry set I

		1···2	3···4	1···3	2···4	TOT
A1A2(T3T4)	MP2	−6.2	−3.6	−1.2	−1.0	−12.0
	POT	+0.7 (−6.8)	+0.8 (−3.5)	−0.2 (−1.4)	−0.3 (−1.9)	+1.0 (−13.7)
G1G2(C3C4)	MP2	−3.5	−1.1	−4.1	−2.8	−11.5
	POT	+4.3 (−4.3)	+3.5 (−0.7)	−3.1 (−4.4)	−1.3 (−3.4)	+3.4 (−12.8)
T1C2(G3A4)	MP2	−3.7	−8.8	+0.4	0.0	−12.1
	POT	+0.7 (−3.7)	−1.2 (−9.4)	+1.0 (−0.3)	+1.6 (0.0)	+2.1 (−13.4)
C1T2(A3G4)	MP2	−4.9	−7.0	0.0	−0.3	−12.2
	POT	−0.2 (−4.4)	+0.1 (−7.7)	+1.0 (−0.3)	+1.0 (−1.0)	+1.9 (−13.4)
A1T2(A3T4)	MP2	−5.1	−5.1	−0.8	+0.4	−10.6
	POT	+0.6 (−5.2)	+0.6 (−5.2)	+1.0 (−1.6)	+0.8 (+0.2)	+3.0 (−11.8)
G1C2(G3C4)	MP2	−9.0	−9.0	+1.4	+2.5	−14.2
	POT	−2.9 (−9.5)	−2.9 (−9.5)	+4.1 (+1.1)	+3.2 (+2.6)	+1.4 (−15.3)
G1T2(A3C4)	MP2	−4.5	−4.5	−3.2	−0.1	−12.3
	POT	+1.7 (−4.6)	+1.3 (−4.7)	−1.2 (−3.9)	+0.4 (−0.2)	+2.2 (−13.5)
T1A2(T3A4)	MP2	−5.0	−5.0	+0.3	−1.4	−11.2
	POT	+0.6 (−5.2)	+0.6 (−5.2)	+0.8 (+0.2)	+1.1 (−2.3)	+2.9 (−13.6)
C1G2(C3G4)	MP2	−5.6	−5.6	+1.1	−3.7	−13.4
	POT	−0.9 (−6.0)	−0.9 (−6.0)	+1.7 (+1.0)	+0.3 (−4.2)	+0.2 (−15.3)
C1A2(T3G4)	MP2	−3.8	−4.7	−0.9	−3.5	−12.9
	POT	0.0 (−5.3)	+1.0 (−4.0)	−0.4 (−1.0)	−0.4 (−4.4)	+0.1 (−14.7)

1...2, 3...4, 1...3, 2...4, The individual base-base contributions, as indicated in the first column; TOT, sum of the four individual base-base contributions; MP2, the ab ibitio values; POT, electrostatic energy approximated by the point charge Coulombic term with electrostatic potential-derived charges. Values in parentheses, electrostatic energy plus van der Waals Lennard-Jones empirical potential. The empirical potential parameters were taken from Šponer et al. (1996a). All data are in kcal/mol.

found with the pioneering ab initio study of Aida (1988), although her calculations underestimate the stacking energies because of the insufficient size of the basis set. The results of Ornstein et al. (1978) also tend to underestimate stacking energies and to significantly exaggerate the sequence dependence of stacking energy. Kudriatskaya and Danilov (not shown) predicted that the stacking energies vary from −6.6 kcal/mol (AT) to −23.6 kcal/mol (CG) (Kudriatskaya and Danilov, 1976). DeVoe and Tinoco (1962) calculated a range from −1.8 kcal/mol (TA, CG) to −15.9 kcal/mol (GC). The inability of previous attempts to properly predict the base stacking and their mutual inconsistency is due to the insufficient level of theoretical treatment imposed by the computer resources available. Let us note, however, that when going from the older studies to present high-level ab initio data, steady improvement can be seen. Furthermore, as the physical description of stacking becomes more complete, the predicted sequence dependence of stacking decreases sharply. Fig. 3 also shows the stacking energies obtained by Hunter (1993), using the empirical π-π interaction model (the data were extracted from the figures in Hunter (1993) and correspond to standard B-DNA structure with a propeller twist of 15°). His agreement with the present data is rather poor, and the absolute value of stacking is underestimated.

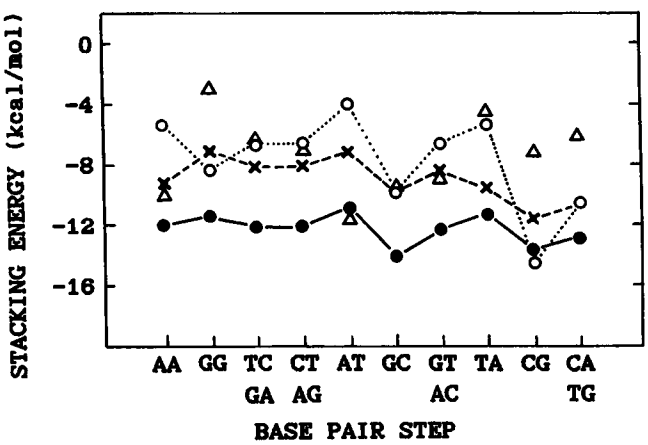

FIGURE 3 Comparison of the MP2/6–31G*(0.25) base-pair stacking energies (——; geometry I, without the many-body correction) for 10 B-DNA steps with previous studies: Aida (1988) (– – –), Ornstein et al. (1978) (·····), Hunter (1993) (△).

Base-(deoxy)ribose stacking in Z-DNA

Interaction energy

Crystal structures of Z-DNA, as well as some 2′,5′-linked dinucleotides, reveal a tight interaction of the sugar O4′ with the six-membered aromatic ring of adjacent purines (Wang et al., 1979) (Fig. 4). Fig. 5 shows the dependence of the base-sugar ab initio stacking energy and its components on the O4′-guanine plane distance. The starting geometry is based (see Methods) on a Z-DNA crystal structure (Gessner et al., 1989). The vertical distance has been varied by shifting the sugar along the guanine normal. The base-sugar interaction is attractive, and the stabilization originates in the electron correlation component of the interaction energy (basically dispersion attraction). The optimal energy, −2.7

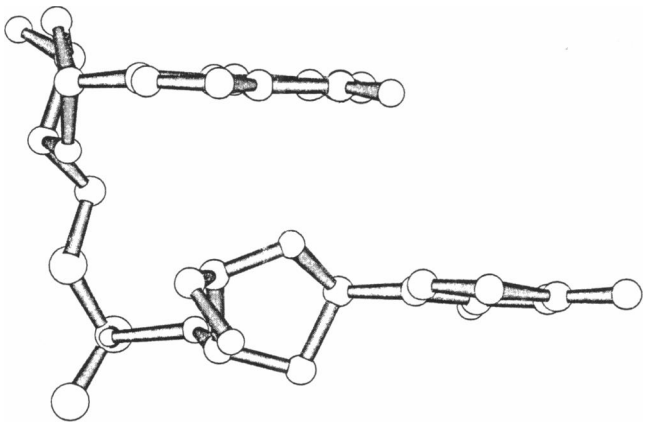

FIGURE 4 Graphical illustration of the base-deoxyribose interaction in a Z-form dinucleotide (dCpG). The guanidine is in the C3′-endo/syn configuration with the six-membered ring of the guanine base stacked over the neighboring cytidine sugar (C2′-endo). Pyrimidine-purine dinucleotides linked by a 2′,5′-phosphodiester bond have similar structures.

kcal/mol, is typical for such a contact and represents a nonnegligible stabilization.

We reevaluated the interaction energy by the empirical potential. The Lennard-Jones parameters for guanine and sugar were taken from the recent AMBER4.1 force field (Cornell et al., 1995). The atomic charges were derived from the molecular electric potential by the same MP2/6-31G*(0.25) method that was used in the ab initio interaction energy calculations. We obtained the value of −2.8 kcal/mol (electrostatic energy of +0.2 kcal/mol), i.e., the same as

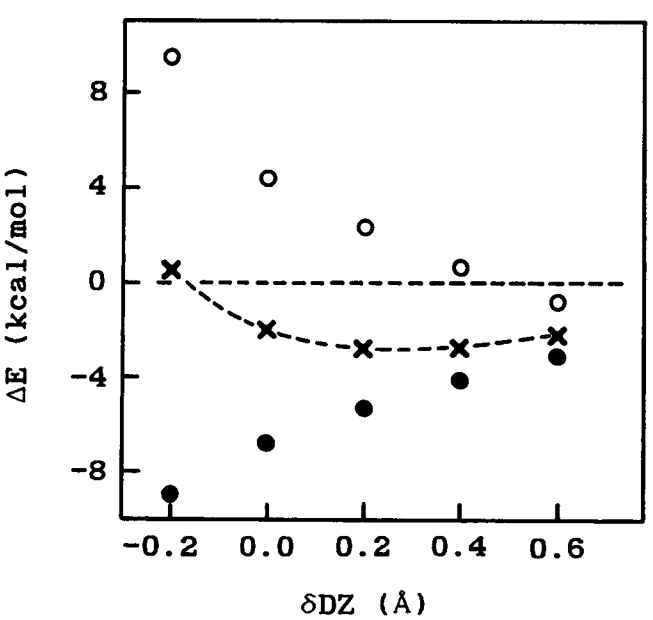

FIGURE 5 Dependence of the deoxyribose···guanine interaction energy on the variation of the separation of the monomers δZ ($\delta Z = 0$ corresponds to the crystal geometry from Gessner et al., 1989). ×, The MP2/6-31G*(0.25) interaction energy; ○, its HF component; ●, its electron correlation component.

using the quantum-mechanical procedure. It indicates that currently available empirical potentials are sufficient to evaluate the sugar-base stacking. Let us note that there is no reason to expect any stabilizing electrostatic interaction between the negatively charged oxygen of sugar and the delocalized π-electron clouds of guanine.

Molecular orbital analysis

We also analyzed the molecular orbitals (MOs) obtained from the HF procedure. The MO analysis does not rule out the attractive charge-transfer n-π^* interaction proposed by Egli and Gessner (1995). The lowest unoccupied molecular orbital (LUMO) is of the π^* type and is predominantly localized on the guanine six-membered ring. The HOMO-5 (fifth highest occupied molecular orbital), localized predominantly at the n-orbitals of the sugar oxygen, is energetically close, and their overlap is also geometrically favorable. However, all charge-transfer interactions are included in the HF interaction energy, which is, for the present system, slightly repulsive. There are thus two possibilities: 1) the charge-transfer n-π^* stabilization is very weak, or 2) charge-transfer n-π^* attraction is significant but is compensated by repulsive electrostatics between base and sugar. Unfortunately, the MO analysis does not allow any quantification of the magnitude of the charge-transfer attraction, and in the present calculations it is not possible to separate this contribution from the HF interaction energy. Nevertheless, the unambiguous conclusion of our study is that sugar-base stacking is stabilized by the dispersion attraction, as evidenced by the ratio between HF and correlation components of the interaction energy. Although we are not able to separate the n-π^* interaction from the other contributions, this contribution is, in contrast to force fields, properly included.

The influence of the sugar···base geometry

We carried out further ab initio calculations. The lateral position of the sugar with respect to guanine was varied by 0.5 Å in all directions at optimal vertical separation of monomers. For the most part, this resulted in worse interaction energies compared to the original geometry, but the energy did remain within −2 to −3 kcal/mol. This indicates that the observed geometry is close to optimum and that the potential energy surface is flat, which is also characteristic of van der Waals interactions. The only surprising result is the apparent compression of the crystal base-sugar dimer by 0.2–0.3 Å compared to the calculated optimum geometry, clearly shown in Fig. 5 (the empirical potential predicted a similar geometry). There is a rather significant repulsion in the original geometry with respect to the distance of the O4′ atom from the guanine plane. However, this compression appears to be an artifact of the particular structure used. We carried out another search using the structure derived from the average base-sugar geometries found in the data base. Here the O4′-base distance is close to optimum, whereas the

calculated optimal stabilization energy is the same. Geometric analysis of all Z-DNA dinucleotides in the Rutgers Nucleic Acid Database (Berman et al., 1992), using the method of Gabb et al. (1996), shows that the average distance from the sugar O4′ to the purine base plane is 2.935 ± 0.013 Å (range 2.58–3.35 Å) for guanine (N = 138) and 3.018 ± 0.019 Å (range 2.93–3.17 Å) for adenine (N = 10). Likewise, the average lateral distance of the sugar O4′ to the geometric center of the six-membered ring of the purine is 0.504 ± 0.005 Å (range 0.11–0.99 Å) for guanine and 0.361 ± 0.004 Å (range 0.0–1.14 Å) for adenine.

Interaction with a metal cation

It has been proposed that a coordination of divalent metal cations to the N7 and/or O6 atoms of guanine may polarize the base, leading to stabilization of the purine-ribose interaction (Egli and Gessner, 1995). We carried out MP2 calculations on a (guanine-Mg^{2+})···deoxyribose trimer to test this hypothesis. The Mg^{2+} cation was selected because of its strong interaction with guanine (Burda et al., 1996). The (G-Mg^{2+})···sugar interaction can be decomposed into three components: G···sugar and Mg^{2+}···sugar interaction energies, evaluated as pair-additive contributions, plus the three-body term. The effect of guanine polarization on sugar···base stacking is included in the three-body term. We have found that the (G-Mg^{2+})···sugar interaction is attractive, −11.2 kcal/mol. The attraction, however, originates primarily in long-range electrostatic interaction between the charged metal ion and the sugar (Mg^{2+}···sugar), which may be significantly reduced because of the solvent screening in DNA. The base-sugar interaction has not been changed substantially. Finally, the three-body correction (MP2/6-31G*(0.25)) is small and slightly repulsive (+0.3 kcal/mol). Hence no additional nonadditive stabilization of the guanine-sugar interaction through the metal-ion coordination has been found.

We also investigate the deoxyribose O4′ interaction with inosine and adenine. Both contacts are of a strength similar to that of the O4′-guanine interaction. Once again, the stabilization is due to dispersion attraction.

Role of the sugar C5′ group

The above calculations were carried out with the C5′ deoxyribose atom replaced by a hydrogen. However, the C5′ atom is not far from the base, so in some calculations the C5′ atom was replaced by a methyl group. This enhanced the base-sugar stacking attraction by about −1.1 to −3.7 kcal/mol at the optimum O4′-base distance. The respective change was +0.2 kcal/mol for the HF interaction energy and −1.3 kcal/mol for the correlation interaction energy, which undoubtedly shows that the increased stabilization is due to dispersion energy. Therefore, the C5′ group contributes significantly to the base-sugar stacking observed in Z-DNA. As shown above, base-pair stacking varies in B-DNA and Z-DNA within −8 to −16 kcal/mol per base-pair step; i.e., the average base-base contribution (there are four base···base contributions per step; Fig. 1) is −2 to −4 kcal/mol, not larger than the base-sugar stacking.

CONCLUDING REMARKS

Further analysis of the data base revealed that there is no water around the sugar···base contacts in Z-DNA. Water is excluded from the contact area by the phosphate group and neighboring bases. In some crystal structures it is possible for water to get close to the C2′/C3′ side of the sugar, but never to the base···sugar contact point. Thus the sugar···base stacking could be described as a dispersion-controlled hydrophobic interaction. A number of related interactions (C-H···O contacts and sugar···base stacking) were recently shown to exist in i-DNA (Kang et al., 1994; Berger et al., 1995, 1996). These interactions do not appear to be exposed to solvent, as in the sugar... base contact in Z-DNA (I. Berger, personal communication). The significance of hydrophobic interactions in DNA has recently been demonstrated by successful replacement of polar bases by non-hydrogen-bonding bases (Schweitzer and Kool, 1994, 1995).

The sugar-base stacking might be one of the sources of observed helical preferences of short $d(CG)_n$ oligonucleotides. For example, it is known that $d(CG)_3$ can form Z-DNA, whereas $d(GC)_3$ cannot (Quadrifoglio et al., 1984). The $d(CG)_3$ oligonucleotide is capable of forming three Z-form dinucleotide repeating units and hence six base-sugar interactions. $d(GC)_3$ can form two Z-form dinucleotides and hence only four base-sugar interactions. We suggest that this loss of stabilization contributes to the inability of $d(GC)_3$ to form Z-DNA. Another contribution could originate in the resistance of base-stacking forces in purine-pyrimidine steps (compared with the pyrimidine-purine steps) to adopt A-DNA conformation (Šponer and Kypr, 1991b). If we subtract the flanking GC steps (where the roll can easily be decreased), this contribution will destabilize the $d(CG)_3$ A-DNA oligonucleotide more than $d(GC)_3$.

CONCLUSIONS

1. The base-pair stacking energies of canonical B-DNA steps, crystal B-DNA steps, and Z-DNA steps were evaluated at the MP2/6-31G*(0.25) level with inclusion of Hartree-Fock nonadditivities. The stacking energies vary from −9.5 kcal/mol (GG) to −13.2 kcal/mol (GC) for standard B-DNA steps, whereas the values obtained for crystal and Z-DNA geometries are similar.

2. Many-body effects influence the sequence dependence and geometry dependence of stacking energy. The present values of many-body effects can still change, because the dispersion nonadditivity could not be assessed.

3. The present results substantially improve on older quantum-chemical studies.

4. Deoxyribose-purine stacking, observed in CG Z-DNA steps and some other structures, is stabilized by dispersion attraction, like base stacking. This interaction appears to gain little or no additional (three-body) stabilization through metal-ion coordination to guanine.

The Supercomputer Centre, Brno, and the Mississippi Center for Supercomputing research are acknowledged for generous allotment of computer time. We thank Imre Berger for stimulating discussion, and both referees for careful reading and very relevant comments.

This study was supported by grant 203/97/0029 from the GA ČR, by ONR grant N00014-95-1-0049, and National Institutes of Health grant 332090.

REFERENCES

Aastrand, P.-O., A. Wallqvist, and G. Karlström. 1994. Nonempirical intermolecular potential for urea-water system. *J. Chem. Phys.* 100: 1262–1269.

Aida, M. 1988. Ab initio molecular orbital study on the sequence-dependency of DNA conformation: an evaluation of intra-strand and inter-strand stacking interaction energy. *J. Theor. Biol.* 130:327–335.

Aida, M., and C. Nagata. 1982. Ab initio MO study on base stacking: adenine-adenine interaction in single-stranded polyadenylic acid (poly A). *Chem. Phys. Let.* 86:44–46.

Aida, M., and C. Nagata. 1986. An ab initio molecular orbital study on the stacking interaction between nucleic acid bases: dependence on the sequence and relation to conformation. *Int. J. Quantum Chem.* 29: 1253–1261.

Alkorta, I., and J. J. Perez. 1996. Molecular polarization potential maps of the nucleic acid bases. *Int. J. Quantum Chem.* 57:123–135.

Arnott, S., R. Chandrasekharan, D. L. Birdsall, A. G. W. Leslie, and R. L. Ratliff. 1980. Left handed DNA helices. *Nature.* 283:743–745.

Arunam, E., and H. S. Gutowsky. 1993. The rotational spectrum, structure and dynamics of benzene dimer. *J. Chem. Phys.* 98:4294–4296.

Berger, I., C.-H. Kang, A. Fredian, R. Ratliff, R. Moyzis, and A. Rich. 1995. Extension of the four-stranded intercalated cytosine motif by adenine·adenine base pairing in the crystal structure of d(CCCAAT). *Struct. Biol.* 2:416–425.

Berger, I., M. Egli, and A. Rich. 1996. Inter-strand C-H···O hydrogen bonds stabilizing four-stranded intercalated molecules: stereoelectronic effects of O4′ in cytosine-rich DNA. *Proc. Natl. Acad. Sci. USA.* 93:12116–12121.

Berman, H. M., W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.-H. Hsieh, A. R. Srinivasan, and B. Schneider. 1992. The Nucleic Acid Database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* 63:751–759.

Bernardo, D. N., Y. B. Ding, K. Krogh-Jespersen, and R. M. Levy. 1994. An anisotropic polarizable water model: incorporation of all-atom polarizabilities into molecular mechanics force fields. *J. Phys. Chem.* 98:4180–4187.

Brupbacher, Th., J. Makarewicz, and A. Bauder. 1994. Intermolecular dynamics of benzene-rare gas complexes as derived from microwave spectra. *J. Chem. Phys.* 101:9736–9746.

Burda, J. V., J. Šponer, and P. Hobza. 1996. An ab initio study of the interaction of guanine and adenine with various monovalent and bivalent metal cations (Li^+, Na^+, K^+, Rb^+, Cs^+, Cu^+, Ag^+, Au^+, Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}). *J. Phys. Chem.* 100:7250–7255.

Caldwell, J. W., and P. A. Kollman. 1995. Structure and properties of neat liquids using nonadditive molecular dynamics: water, methanol, and methylacetamide. *J. Phys. Chem.* 99:6208–6219.

Calladine, C. R., and H. R. Drew. 1986. Principles of sequence-dependent flexure of DNA. *J. Mol. Biol.* 192:907–918.

Carloni, P., and W. Andreoni. 1996. Platinum-modified nucleobase pairs in the solid state. A theoretical study. *J. Phys. Chem.* 100:17797–17800.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117: 5179–5197.

Dang, L. X., D. A. Pearlman, and P. A. Kollman. 1990. Why do A·T base pairs inhibit Z-DNA formation? *Proc. Natl. Acad. Sci. USA.* 87: 4630–4634.

DeVoe, H., and I. Tinoco, Jr. 1962. The stability of helical polynucleotides: base contributions. *J. Mol. Biol.* 4:500–517.

Ding, Y., D. N. Bernardo, K. Krogh-Jespersen, and R. M. Levy. 1995. Solvation free energies of small amides and amines from molecular dynamics/free energy perturbation simulations using pairwise additive and many-body polarizable potentials. *J. Phys. Chem.* 99:11575–11583.

Egli, M., and R. V. Gessner. 1995. Stereoelectronic effects of deoxyribose O4′ on DNA conformation. *Proc. Natl. Acad. Sci. USA.* 92:180–184.

Engkvist, O., P.-O. Aastrand, and G. Karlström. 1996. Intermolecular potential for the 1,2-dimethoxyethane-water complex. *J. Phys. Chem.* 100:6950–6957.

Florián, J., and J. Leszczynski. 1995. Theoretical investigation of the molecular structure of πk DNA base pair. *J. Biomol. Struct. Dyn.* 12:1055–1062.

Florián, J., and J. Leszczynski. 1996. Spontaneous DNA mutations induced by proton transfer in the guanine·cytosine base pairs: an energetic perspective. *J. Am. Chem. Soc.* 118:3010–3017.

Friedman, R. A., and B. Honig. 1992. The electrostatic contribution to DNA base stacking interactions. *Biopolymers.* 32:145–159.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc. 1995. Pittsburgh, PA.

Gabb, H. A., S. R. Sanghani, C. H. Robert, and C. Prevost. 1996. Finding and visualizing nucleic acid base stacking. *J. Mol. Graph.* 14:6–11.

Gessner, R. V., C. A. Frederick, G. J. Quigley, A. Rich, and A. H.-J. Wang. 1989. The molecular structure of the left-handed Z-DNA double helix at 1.0 Angstrom atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem.* 264:7921–7935.

Gorin, A. A., V. B. Zhurkin, W. K. Olson. 1995. B-DNA twisting correlates with base-pair morphology. *J. Mol. Biol.* 247:34–48.

Gould, I. R., and P. A. Kollman. 1994. Theoretical investigation of the hydrogen bond strengths in guanine-cytosine and adenine-thymine base pairs. *J. Am. Chem. Soc.* 116:2493–2499.

Halgren, T. A. 1995. Potential energy functions. *Curr. Opin. Struct. Biol.* 5:205–210.

Harvey, S. C. 1983. DNA structural dynamics: longitudinal breathing as a possible mechanism for the B-Z transition. *Nucleic Acids Res.* 11: 4867–4878.

Heinemann, U., and C. Alings. 1989. Crystallographic study of one turn of G/C rich B-DNA. *J. Mol. Biol.* 210:369–381.

Henson, B. F., G. V. Hartland, V. A. Venturo, and P. M. Felke. 1992. Raman-vibronic double-resonance spectroscopy of benzene dimer isotopomers. *J. Chem. Phys.* 97:2189–2208.

Herbert, A. G., J. R. Spitzner, K. Lowenhaupt, and A. Rich. 1993. Z-DNA binding protein from chicken blood nuclei. *Proc. Natl. Acad. Sci. USA.* 90:3339–3342.

Hobza, P., F. Hubálek, M. Kabeláč, P. Mejzlík, J. Šponer, and J. Vondráöek. 1996a. Ability of empirical potentials (AMBER, CHARMM, CVFF, OPLS, Poltev) and semi-empirical quantum chemical methods (AM1, MNDO/M, PM3) to describe H-bonding in DNA base pairs: comparison with ab initio results. *Chem. Phys. Lett.* 257:31–35.

Hobza, P., M. Kabeláč, J. Šponer, P. Mejzlík, and J. Vondráöek. 1997. Performance of empirical potentials (AMBER, CFF95, CVFF, CHARMM, OPLS, POLTEV), semiempirical quantum-chemical methods (AM1, MNDO/M, PM3) and ab initio Hartree-Fock method for interaction of DNA bases. Comparison with beyond-Hartree-Fock results. *J. Comput. Chem.* (in press).

Hobza, P., A. Melhorn, P. Carsky, and R. Zahradník. 1986. Stacking interactions: ab initio SCF and MP2 study on $(H_2O)_2$, $(H_2S)_2$, $(HCN)_2$, $(CH_2O)_2$, and (C_2H_4). *J. Mol. Struct.* 138:387–399.

Hobza, P., H. L. Selzle, and E. W. Schlag. 1994a. Potential energy surface of benzene dimer: ab initio theoretical study. *J. Am. Chem. Soc.* 116: 3500–3506.

Hobza, P., H. L. Selzle, and E. W. Schlag. 1994b. Structure and properties of benzene-containing molecular clusters: nonempirical ab initio calculations and experiments. *Chem. Rev.* 94:1767–1786.

Hobza, P., H. L. Selzle, and E. W. Schlag. 1996b. Potential energy surface for the benzene dimer: results of ab initio CCSD(T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced. *J. Phys. Chem.* 100:18970–18794.

Hobza, P., J. Šponer, and M. Poláöek. 1995. H-bonded and stacked DNA base pairs: cytosine dimer. Ab initio second order Møller-Plesset study. *J. Am. Chem. Soc.* 117:792–798.

Hunter, C. A. 1993. Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol.* 230:1025–1054.

Hutter, J., P. Carloni, and M. Parrinello. 1996. Nonempirical calculations of a hydrated RNA duplex. *J. Am. Chem. Soc.* 118:8710–8712.

Jaworski, A., W.-T. Hsieh, J. A. Blaho, J. E. Larson, and R. D. Wells. 1987. Left-handed DNA in vivo. *Science.* 238:773–777.

Kagawa, T. T., D. Stoddard, G. Zhou, and P. S. Ho. 1989. Quantitative analysis of DNA secondary structure from solvent-accesible surfaces: the B- to Z-DNA transition as a model. *Biochemistry.* 28:6642–6651.

Kang, C.-H., I. Berger, C. Lockshin, R. Ratliff, R. Moyzis, and A. Rich. 1994. Crystal structure of intercalated four-stranded $d(C_3T)$ at 1.4 Å resolution. *Proc. Natl. Acad. Sci. USA.* 91:11636–11640.

Kerr, I. M., and R. E. Brown. 1978. pppA2′p5′A2′p5′A: an inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells. *Proc. Natl. Acad. Sci. USA.* 75:256–260.

Krause, H., and H. J. Neusser. 1993. Dissociation energy of neutral and inonic benzene-noble gas dimers by pulsed field threshold ionization spectroscopy. *J. Chem. Phys.* 99:6278–6286.

Krishnan, R., and T. P. Seshadri. 1993. Stereochemistry of 2′-5′ linked nucleic acids: crystal and molecular structure of ammonium adenylyl-2′,5′-adenosine tetrahydrate: a core fragment of 2′-5′ oligo A's produced by interferon-induced adenylate synthetase. *J. Biomol. Struct. Dyn.* 10:727–745.

Krishnan, R., and T. P. Seshadri. 1994. Crystal structure of guanylyl-2′,5′-cytidine dihydrate: an analogue of msDNA-RNA junction in *Stigmatella aurantiaca. Biopolymers.* 34:1637–1646.

Kroon-Batenburg, L. M. J., and F. B. J. van Duijneveldt. 1985. The use of a moment-optimized DZP basis set for describing the interaction in the water dimer. *J. Mol. Struct.* 121:185–199.

Kudryatskaya, Z. G., and V. I. Danilov. 1976. Quantum mechanical study of bases interactions in various associates in atomic dipole approximation. *J. Theor. Biol.* 59:303–318.

Lavery, R. 1988. Junctions and bends in nucleic acids: a new theoretical and modelling approach. *In* Structure and Expression, Vol. 3, DNA Bending and Curvature. W. K. Olson, R. H. Sarma, M. H. Sarma, and M. Sundaralingam, editors. Adenine Press, Schenectady, NY. 191–211.

Lavery, R., and B. Hartmann. 1994. Modelling DNA conformational mechanics. *Biophys. Chem.* 50:33–45.

Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lavery, R., H. Sklenar, K. Zakrzewska, and B. Pullman. 1986. The flexibility of the nucleic acids: (II) the calculation of internal energy and applications to mononucleotide repeat DNA. *J. Biomol. Struct. Dyn.* 4:989–1014.

Lukomski, S., and R. D. Wells. 1994. Left-handed Z-DNA and in vivo supercoil density in the *Escherichia coli* chromosome. *Proc. Natl. Acad. Sci. USA.* 91:9980–9984.

Mooers, B. H. M., G. P. Schroth, W. W. Baxter, and P. S. Ho. 1995. alternating and non-alternating dG-dC hexanucleotides crystallize as canonical A-DNA. *J. Mol. Biol.* 249:772–784.

Müller, V., M. Takeya, S. Brendel, B. Wittig, and A. Rich. 1996. Z-DNA-forming sites within the human β-globin gene cluster. *Proc. Natl. Acad. Sci. USA.* 93:780–784.

Olson, W. K., A. R. Srinivasan, N. L. Marky, and V. N. Balaji. 1983. Theoretical probes of DNA conformation examining the B leads to Z conformation transition. *Cold Spring Harb. Symp. Quant. Biol.* 47: 229–241.

Olson, W. K., and Zhurkin, V. B. 1996. Twenty years of DNA bending. *In* Biological Structure and Dynamics, Proceedings of the Ninth Conversation. R. H. Sarma and M. H. Sarma, editors. Adenine Press, Schenectady, NY. 341–370.

Ornstein, R. L., R. Rein, D. L. Breer, and R. D. MacElroy. 1978. An optimized potential function for the calculation of nucleic acid interaction energies. I. Base stacking. *Biopolymers.* 17:2341–2360.

Pearlman, D. A., and P. A. Kollman. 1990. The calculated free energy effects of 5-methyl cytosine on the B to Z transition in DNA. *Biopolymers.* 29:1193–1209.

Prive, G. G., K. Yanagi, and R. E. Dickerson. 1991. Structure of B-DNA decamer CCAACGTTGG and isomorphous decamers CCAAGATTGG and CCAGCCTGG. *J. Mol. Biol.* 221:177–199.

Quadrifoglio, F., G. Manzini, and N. Yathindra. 1984. Short oligodeoxynucleotides with d(G-C)n sequence do not assume left-handed conformation in high salt conditions. *J. Mol. Biol.* 175:419–423.

Rahmouni, A. R., and R. D. Wells. 1989. Stabilization of Z-DNA in vivo by localized supercoiling. *Science.* 246:358–363.

Rich, A., 1994. Speculation on the biological roles of left-handed Z-DNA. *Ann. N.Y. Acad. Sci.* 726:1–17.

Rudnicki, W. R., and B. Lesyng, 1994. Applicability of commonly used atom-atom type potential energy functions in structural analysis of nucleic acids. The role of eletrostatic interactions. *Computers Chem.* 19:253–258.

Saenger, W., and U. Heinemann. 1989. Raison d'etre and structural model for the B-Z transition of poly d(G-C)poly d(G-C). *FEBS Lett.* 257: 223–227.

Saykally, R. 1996. Water clusters. *Science.* 271:929–933.

Schweitzer, B. A., and E. T. Kool. 1994a. Aromatic nonpolar nucleosides as hydrophobic isoesters of pyrimidine and purine nucleosides. *J. Org. Chem.* 59:7238–7242.

Schweitzer, B. A., and E. T. Kool. 1994b. Hydrophobic, non-hydrogen-bonding bases and base pairs in DNA. *J. Am. Chem. Soc.* 117: 1863–1872.

Smith, G. D., and R. L. Jaffe. 1996. Comparative study of force fields for benzene. *J. Phys. Chem.* 100:9624–9629.

Šponer, J., and P. Hobza. 1994a. Nonplanar geometries of DNA bases. Second order Møller-Plesset study. *J. Phys. Chem.* 98:3161–3164.

Šponer, J., and P. Hobza. 1994b. Bifurcated hydrogen bonds in DNA crystal structures. An ab initio quantum-chemical study. *J. Am. Chem. Soc.* 116:709–714.

Šponer, J., and P. Hobza. 1994c. G·C base pair in parallel-stranded DNA—a novel type of base pairing: an ab initio quantum chemical study. *J. Biomol. Struct. Dyn.* 12:671–680.

Šponer, J., and P. Hobza. 1996. Nonempirical ab initio calculations on DNA base pairs. *Chem. Phys.* 204:365–372.

Šponer, J., and P. Hobza. 1997. MP2 and CCSD(T) study on hydrogen bonding, aromatic stacking and nonaromatic stacking. *Chem. Phys. Lett.* 267:263–270.

Šponer, J., J. Jursa, and J. Kypr. 1994. Interactions between the guanine amino group and the adenine six membered ring stabilizes the unusual conformation of the CpA step in B-DNA. *Nucleosides Nucleotides.* 13:1669–1677.

Šponer, J., and J. Kypr. 1990. Base pair buckling can eliminate the interstrand purine clash in the C-G steps of B-DNA. *J. Biomol. Struct. Dyn.* 7:1211–1221

Šponer, J., and J. Kypr. 1991a. On the use of empirical potentials in studies of base stacking in DNA. *In* Theoretical Biochemistry and Molecular Biophysics, Vol. 1, DNA. D. L. Beveridge and R. Lavery, editors. Adenine Press, Schenectady, NY. 271–284.

Šponer, J., and J. Kypr. 1991b. Different intrastrand and interstrand contributions to stacking account for roll variations at the alternating purine-pyrimidine sequences in A-DNA and A-RNA. *J. Mol. Biol.* 221: 761–764.

Šponer, J., and J. Kypr. 1993a. Correlations among roll, rise, cup and stagger in DNA structures suggested by empirical potential calculations. *J. Biomol. Struct. Dyn.* 11:27–41.

Šponer, J., and J. Kypr. 1993b. Theoretical analysis of the base stacking in DNA. Choice of the force field and a comparison with the oligonucleotide crystal structures. *J. Biomol. Struct. Dyn.* 11:277–292.

Šponer, J., J. Leszczynski, and P. Hobza. 1996a. On the nature of nucleic acid base stacking. Nonempirical ab initio and empirical potential characterization of 10 stacked base pairs. Comparison of stacked and H-bonded base pairs. *J. Phys. Chem.* 100:5590–5596.

Šponer, J., J. Leszczynski, and P. Hobza. 1996b. Base stacking in cytosine dimer. A comparison of correlated ab initio calculations with three empirical potential models and density functional theory calculations. *J. Comput. Chem.* 17:841–850.

Šponer, J., J. Leszczynski, and P. Hobza. 1996c. Structures and energies of hydrogen-bonded DNA base pairs. A nonempirical study with inclusion of electron correlation. *J. Phys. Chem.* 100:1965–1974.

Šponer, J., J. Leszczynski, and P. Hobza. 1996d. Hydrogen bonding and stacking of DNA bases. A review of ab initio quantum-chemical studies. *J. Biomol. Struct. Dyn.* 14:117–135.

Šponer, J., J. Leszczynski, V. Vetterl, and P. Hobza. 1996e. Base stacking and hydrogen bonding in protonated cytosine dimer. The role of molecular ion-dipole and induction interactions. *J. Biomol. Struct. Dyn.* 13: 695–707.

Suhai, S. 1994. Structure and bonding in the formamide crystal: a complete fourth-order many-body perturbation theoretical study. *J. Chem. Phys.* 103:7030–7039.

Sun, Y., J. W. Caldwell, and P. A. Kollman. 1995. Molecular dynamics and free energy perturbation study of spherand complexation with metal ions employing additive and nonadditive force fields. *J. Phys. Chem.* 99: 10081–10085.

Wang, A. H-J., G. J. Quigley, F. J. Kolpack, J. L. Crawford, J. H. van Boom, G. A. van der Marel, and A. Rich. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature.* 282:680–686.

Wang, S., and E. T. Kool. 1995. Origins of the large differences in stability of DNA and RNA helices: C-5 methyl and 2′-hydroxyl effects. *Biochemistry.* 34:4125–4132.

Wölfl, S., C. Martinez, A. Rich, and J. A. Majzoub. 1996. Transcription of the human corticotropin-releasing hormone gene in NPLC cells is correlated with Z-DNA conformation. *Proc. Natl. Acad. Sci. USA.* 93: 3664–3668.