Nature of Nucleic Acid−Base Stacking: Nonempirical ab Initio and Empirical Potential Characterization of 10 Stacked Base Dimers. Comparison of Stacked and H-Bonded Base Pairs

Jiří Šponer,*,†,‡,§ Jerzy Leszczyński,‡ and Pavel Hobza†

J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague 8, Czech Republic; Department of Chemistry, Jackson State University, Jackson, Mississippi 39217; Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

Received: November 8, 1995[⊗]

Ab initio (MP2/6-31G*(0.25)) interaction energies were calculated for almost 240 geometries of 10 stacked nucleic acid−base pairs: A···A, C···C, G···G, U···U, A···C, G···A, A···U, G···C, C···U, and G···U; in some cases uracil was replaced by thymine. The most stable stacked pair is the G···G dimer (−11.3 kcal/mol), and the least stable is the uracil dimer (−6.5 kcal/mol). The interaction energies of H-bonded base pairs range from −25.8 kcal/mol (G···C) to −10.6 kcal/mol (T···T). The stability of stacked pairs originates in the electron correlation, while all the H-bonded pairs are dominated by the HF energy. The mutual orientation of the stacked bases is, however, primarily determined by the HF interaction energy. The ab initio base stacking energies are well reproduced by the empirical potential calculations, provided the atomic charges are derived by the same method as used in the ab initio calculations. Some contributions previously postulated to significantly influence base stacking (induction interactions, $\pi-\pi$ interactions) were not found. Base stacking was also investigated in six B-DNA and two Z-DNA base pair steps; their geometries were taken from the oligonucleotide crystal data. The many-body correction was estimated at the HF/MINI-1 level. The sequence-dependent variations of the total base pair step stacking energies range from −9.9 to −14.7 kcal/mol. The range of the calculated many-body corrections to the stacking energy is 2 kcal/mol. The ab initio calculations exclude the consideration that the unusual conformational properties of the CpA(TpG) steps might be associated with attractive induction interactions of the exocyclic groups of DNA bases and the aromatic rings of bases.

1. Introduction

Stacking and hydrogen bonding of DNA bases affects the three-dimensional structure of DNA.[1] The H-bonded complexes of DNA bases are dominated by the electrostatic contributions, well covered at the Hartree−Fock level of theory.[2] Stabilization of the stacked complexes of DNA bases is mostly assumed to be due to the dispersion energy which originates in the electron correlation.[2j] In addition, the dependence of stacking energy on intermolecular geometry of stacked DNA bases is complex,[2j,3] in contrast to the rather small number of well-defined minima of the H-bonded pairs. Therefore, the number of quantum-chemical studies on base stacking is smaller[2j,4] and their reliability is lower. Base stacking interactions influence the local conformational variability of DNA, which has been analyzed in a number of empirical potential studies.[1,3,5] However, differences among various empirical potentials are nonnegligible[5g] and it was also proposed that there are contributions not covered by the empirical potentials ($\pi-\pi$ interactions,[1b] induction attraction,[6] lone-pair interactions,[7] anisotropic short-range repulsion[2j,8]). Recently, we carried out the first consistent ab initio comparison, with inclusion of electron correlation, of a stacked and hydrogen-bonded DNA pair, cytosine dimer.[2j,1,8] Later, we characterized 30 hydrogen-bonded DNA base pairs at the HF/6-31G** level, including their vibrational analysis and intrinsic nonplanarity;[2k,l] the interaction energies were evaluated with inclusion of electron correlation. Here we extend the previous studies by performing correlated (MP2/6-31G*) calculations on 10 stacked base pairs (10−30 geometries per base pair), covering the stacking energy dependences on twist, displacement, and vertical separation of stacked bases, and including also geometries from high-resolved B-DNA and Z-DNA crystal structures. The ab initio calculations are combined with empirical potential calculations using a standard combination of a Lennard-Jones and atomic point charge terms, adjusted to reproduce the ab initio data. A comparison of the empirical potential and ab initio data should reveal whether there are energy contributions not covered by the standard empirical potentials.

2. Method

(a) Ab initio calculations were made with rigid DNA bases; planar MP2/6-31G* optimized geometries of guanine, adenine, uracil a thymine, and MP2/DZ(2d) optimized geometry of cytosine were used.[9] Ab initio supermolecule calculations were made at the second-order Møller−Plesset (MP2) level of theory. A standard split-valence 6-31G basis set was used, with a set of diffuse polarization functions having an exponent of 0.25 added to the second-row elements (designated 6-31G*(0.25)). The flat polarization functions improve the description of dispersion energy.[2j,10] Both HF and MP2 contributions to the interaction energy were corrected for basis set superposition error.[11]

The size of the systems under study, the complex nature of their conformational space, and the fact that available gradient optimization methods are not corrected for the basis set superposition error do not permit use of the gradient optimization

* Address correspondence to this author at J. Heyrovský Institute of Physical Chemistry.
† J. Heyrovský Institute of Physical Chemistry.
‡ Jackson State University.
§ Institute of Biophysics.
⊗ Abstract published in *Advance ACS Abstracts,* March 1, 1996.

0022-3654/96/20100-5590$12.00/0 © 1996 American Chemical Society

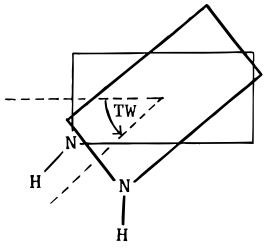

Figure 1. A sketch of how the twist angle TW is defined.

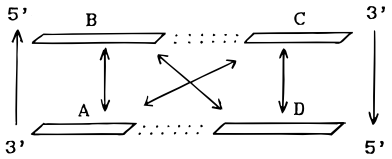

Figure 2. Base stacking of a base pair step AD···BC approximated as a sum of four base−base contributions: A···B, C···D, A···C, B···D.

for stacked DNA base pairs. Instead, we scanned the conformational space of stacked base pairs by performing a set of single-point calculations. First, the dependence of the interaction energy on twist was investigated for 10 undisplaced stacked dimers: C···C, G···G, A···A, U···U, G···C, G···A, A···U, G···U, A···C, and C···U. In undisplaced structures the centers of mass of the bases are stacked one above the other. Twist angle TW is a rotation of the upper base in the dimer around the axis passing through its center of mass and perpendicular to its plane. Uracil-containing pairs were studied due to steric difficulties caused by the thymine methyl group. For a twist angle of 0° the two N1−H1 bonds of pyrimidines and N9−H9 bonds of purines were colinear. The twist was introduced in the right-handed sense (Figure 1). A limited number of calculations were then made for displaced structures. All calculations were made with coplanar bases; vertical separation of bases was 3.3 or 3.4 Å,[5g,6,12] according to vertical base separations observed in crystals of DNA constituents[6] and all high-resolved oligonucleotide crystals (3.3−3.5 Å).[5g] In addition, we have found that the optimal vertical separation obtained at the MP2/6-31G*(0.25) level for antiparallel, undisplaced homodimers (A···A, C···C, G···G, and U···U) is 3.25−3.35 Å. It should be noted that for less favorable geometries the optimal vertical separation can be somewhat increased (parallel, undisplaced cytosine dimer is even unbound due to the dipole−dipole repulsion[2j]).

Finally, the MP2/6-31G*(0.25) stacking interaction energies were calculated for eight DNA base pair steps observed in oligonucleotide crystals. The following crystal data were used: The monoclinic B-DNA decamer $d(CCAACGTTGG)_2$[13] solved at a 1.3 Å resolution (database[14] designation 5DNB), the G5C6 step of the monoclinic decamer $d(CCAGGCCTGG)_2$[15] solved at a 1.6 resolution (1BD1), and the C2G3 and G3C4 steps of the $d(CGCGCG)_2$[16] Z-DNA hexamer, solved at a 0.9 Å resolution (2DCG). The 5DNB decamer has five crystallographically different base pairs due to the twofold symmetry. The crystal geometries of DNA bases were replaced by optimized DNA bases (here thymine was used),[9] which did not change the intermolecular geometry significantly. The sugar was replaced by a hydrogen. The MP2 interaction energies of stacked base pair dimers were obtained as a sum of four base−base contributions (see Figure 2)—two intrastrand and two interstrand stacks. The many-body effects were estimated at the HF/MINI-1 level using the same geometries as for the MP2 calculations. The supermolecule HF/MINI-1 interaction energies were calculated as a stacking of two H-bonded pairs, and as a sum of the four separate base−base contributions. The difference of these two energies gives the many-body correction E^{MB}. Many-body effects estimated in this way cover only the Heitler−London exchange and SCF deformation nonadditivities. The inclusion of dispersion and higher-order exchange nonadditivities would require at least the MP2 level of calculations, while the three-body nonadditivities are almost completely covered at the third-order Møller−Plesset level. Such calculations are beyond our present computer facilities. All calculations were made with the Gaussian92 set of programs.[17]

The empirical potential calculations were used for a qualitative analysis of the dependence of stacking energy on twist and displacement. The empirical potential interaction energy was calculated with the expression

$$E_{INT} = \sum -A_{ij}/r_{ij}^{\,6} + \sum B_{ij}/r_{ij}^{\,9} + 332\sum q_i q_j/r_{ij}$$

where A_{ij} and B_{ij} are empirical van der Waals constants, r_{ij} are distances between atoms i and j, and q_i are the atomic charges. The point charges were localized on all atomic centers and derived from the molecular electric potential (MEP)[18] at the MP2/6-31G*(0.25) and MP2/aug-cc-pVDZ[19] levels. The 6−9 Lifson−Hagler[20] (6-9LH) empirical potential with the interaction energies scaled by a factor of 0.7 was used as a van der Waals potential. The scaling factor approximately adjusted the absolute values of empirical potential energies to the ab initio values. This scaling factor was the only additional empirical parameter added to the original Lennard-Jones parameters. It should be noted that van der Waals energy evaluated with the empirical potential changes only insignificantly with twist and displacement angle, because it is proportional to the geometrical overlap of bases. The choice of the 6-9LH potential is due to our previous experience with this potential—the power of the repulsive term (9 or 12) is not important, as explained elsewhere.[5g] Other (scaled) Lennard-Jones potentials could be used without changing the results significantly because of the isotropic character of dispersion attraction and short-range repulsion. It should be emphasized that the purpose of this study, to reveal energy contributions not covered by the standard empirical potentials, strongly requires that the individual terms of the potential (mainly the structure-making coulombic term) are as close as possible to their respective contributions in the ab initio calculations. That is why the commonly used force fields (e.g., AMBER, GROMOS, CHARMM, OPLS, CVFF) were not used. An extensive comparison of the available high-level ab initio data for H-bonded[2k,l] and stacked base pairs with the currently used empirical force fields will be carried out separately.

3. Results and Discussion

3.1. Dependence of the Stacking Energy on Twist and Displacement. Figure 3 summarizes the dependence of base stacking energy on twist and displacement for all 10 stacked pairs. The solid lines present the dependence of the stacking energy on twist angle for the undisplaced dimers, calculated by the empirical potential with the MEP MP2/6-31G*(0.25) charges. The dotted lines are dependences of empirical potential energy on twist angle, but with the optimized (for every value of twist) displacement. The supermolecule MP2/6-31G*(0.25) ab initio data calculated for the undisplaced structures are marked by crosses, and the ab initio interaction energies obtained for optimum (empirical potential) displaced geometries are marked by circles. The crosses thus correspond to the solid lines and the open circles to the dotted curves.

The shape of the energy dependence for undisplaced and displaced pairs is similar, and, in the low-energy region of twist,

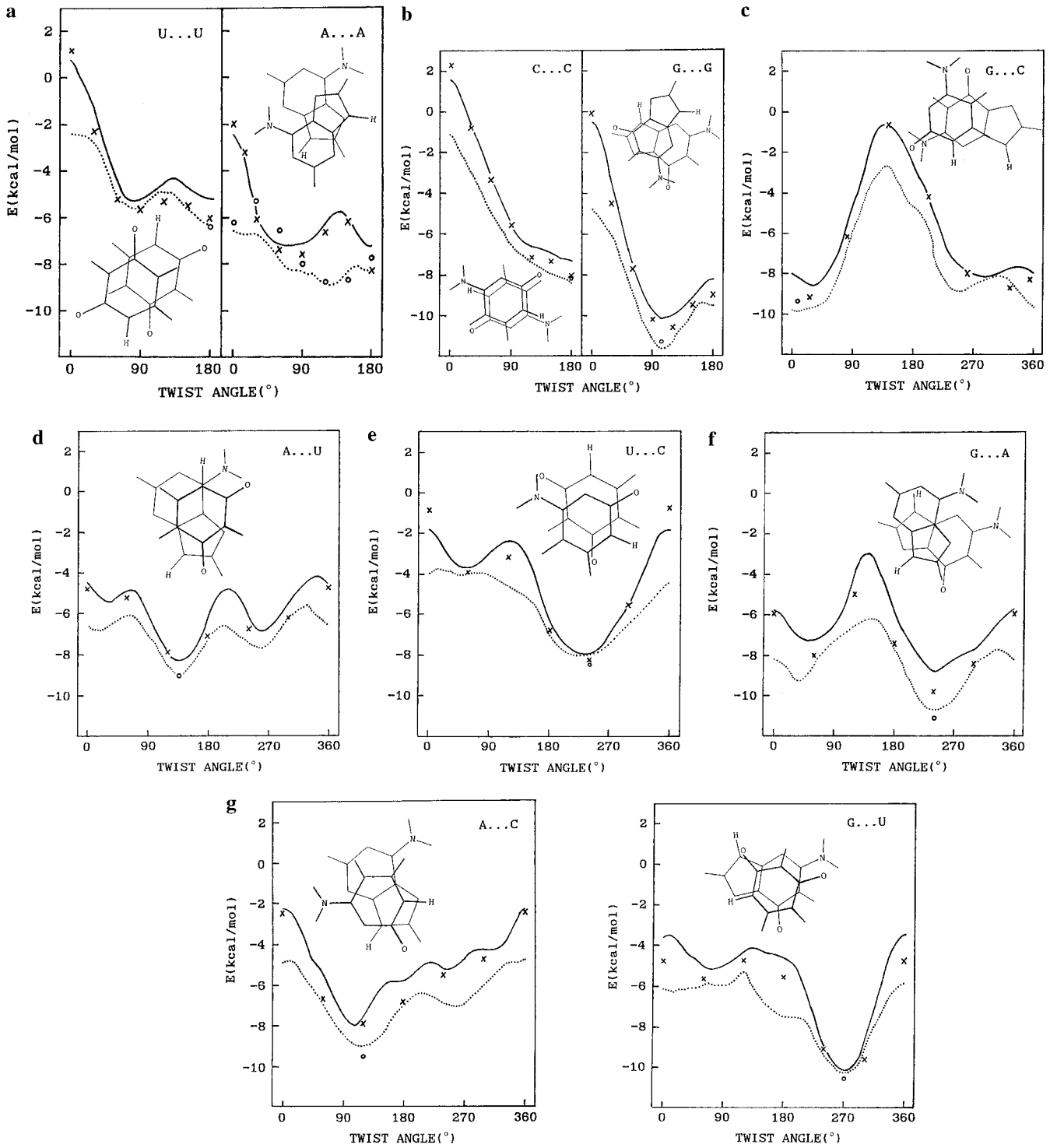

Figure 3. Dependence of the base stacking energy on twist: solid lines, the empirical potential energy, undisplaced dimer; dotted lines, empirical potential energy, optimized displacement. The MP2/6-31G*(0.25) supermolecule interaction energies are marked by crosses (undisplaced pairs) and circles (displaced pairs). The calculations were made with a constant vertical separation of bases (see Method) of 3.4 Å (C···C, G···G, G···C) and 3.3 Å (the remaining steps). Optimum stacking geometries of the 10 stacked pairs obtained by the empirical potential calculations are depicted. The exocyclic heteroatoms and H1 (H9) hydrogens of pyrimidines (purines) are indicated.

rather a small energy improvement is associated with the displacement optimization. The interaction energy is thus primarily determined by twist angle. The largest (empirical potential) differences between the undisplaced and displaced structures were found for the adenine dimer; for this pair a more detailed ab initio study of the displaced structures was made.

The agreement of the ab initio and empirical potential data is surprisingly good; their energy difference is never larger than 1.5 kcal/mol. The largest differences were found for the displaced structures of the adenine dimer (cf. the circles and the dotted line for A···A at the Figure 3)—see below. Table 1 summarizes the MP2/6-31G*(0.25) interaction energies calculated for the optimum geometry found by the empirical potential; these geometries are depicted in Figure 3. The largest stabilization energy was found for the guanine dimer, 11.3 kcal/mol, and the smallest for the uracil dimer, 6.5 kcal/mol. This is not surprising in light of the large polarizability and dipole moment of guanine (cf. Table 2).

Table 1 presents also interaction energies of the most stable hydrogen-bonded pairs, calculated at the MP2/6-31G*(0.25)//

TABLE 1: MP2/6-31G*(0.25) Interaction Energies (kcal/mol) of Stacked and H-Bonded DNA Base Pairs[a]

	stacked pairs				H-bonded pairs		
	E^{HF}	E^{COR}	E^{MP2}		E^{HF}	E^{COR}	E^{MP2}
G···G	−0.84	−10.47	−11.31	GGI	−25.08	+0.39	−24.69
G···A	+1.30	−12.47	−11.16	GAI	−12.22	−3.01	−15.23
G···U	−1.17	−9.45	−10.62	GTI	−14.23	−0.92	−15.15
C···A	+0.85	−10.35	−9.50	CAI	−10.83	−3.51	−14.34
G···C	−1.44	−7.87	−9.32	GCWC	−24.58	−1.23	−25.81
A···U	+1.25	−10.33	−9.08	TAH	−10.38	−2.94	−13.32
A···A	+4.01	−12.84	−8.83	AAI	−7.83	−3.72	−11.55
C···C	−2.09	−6.17	−8.26	CC	−16.15	−2.66	−18.81
C···U	−1.51	−7.01	−8.52	TCI	−8.68	−2.67	−11.44
U···U	+0.46	−6.98	−6.52	TTII	−9.29	−1.35	−10.64

[a] The optimized stacked pairs (cf. Figure 3) were found by the empirical potential calculations, and the H-bonded pairs were optimized at the HF/6-31G** level. The designation of the H-bonded pairs was taken from ref 2a.

HF/6-31G** level. All the hydrogen-bonded pairs are more stable than the optimum stacked ones.

Let us now discuss the several points of displaced adenine dimer, where the ab initio procedure gives smaller stabilization energies than the empirical potential. Figure 4a shows the undisplaced A···A dimer with a twist of 60°. For this structure, the MP2 calculation predicts (at vertical separation of 3.3 Å) the stabilization energy to be 7.4 kcal/mol, and the empirical potentials gives 7.9 kcal/mol. Figure 4b shows the dimer with displacement optimized by the empirical potential, which improves the empirical stacking stabilization to 8.2 kcal/mol. However, the ab initio method predicts a stacking destabilization (6.6 kcal/mol) due to this displacement. On the other hand, there is another displaced dimer with twist 60° having the empirical stabilization 8.2 kcal/mol (Figure 4c) which is also stabilized at the ab initio level (7.8 kcal/mol). We performed a set of calculations with increased vertical separation of adenines (up to 3.8 Å, not shown). This eliminated the differences between the ab initio and empirical potential data, which indicates that these differences are of a short-range origin. A similar positive ab initio energy peak was previously investigated for the antiparallel displaced cytosine dimer[2j,8] and could be attributed to the anisotropic character of the short-range repulsion,[8] not covered by potentials having spherically

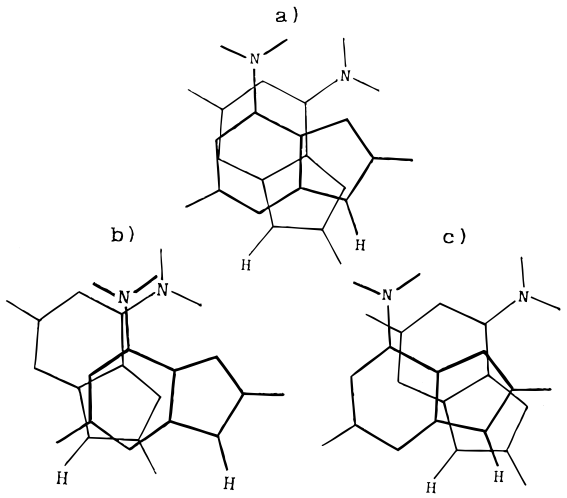

Figure 4. Three stacked adenine dimers with a twist of 60°. The MP2/6-31G*(0.25) interaction energies are −7.4 kcal/mol (a), −6.6 kcal/mol (b), −7.8 kcal/mol (c); the empirical potential stacking energies are −7.9 kcal/mol (a), −8.2 kcal/mol (b, c).

symmetrical atoms.[21] We also analyzed displaced structures of the U···U dimer, but no increased short-range repulsion was found.

For the four homopairs (C···C, U···U, G···G, and A···A), the empirical potential calculations underestimate the energy difference between the parallel and antiparallel undisplaced dimers. (The ab initio data are above the empirical potential curve for the parallel undisplaced structure and below it for the antiparallel undisplaced structure, see Figure 3.) The empirical potential calculations overestimate the short-range repulsion for the antiparallel structure but underestimate the short-range repulsion for the parallel structure.[8] This difference was not eliminated by varying the van der Waals parameters or using the 12th power of the repulsion term.[8] If the van der Waals term was adjusted to reproduce the short-range repulsion for the antiparallel structure, it failed for the parallel dimer, and vice versa.[8] A possible explanation is again the effect of anisotropic short-range repulsion—parallel structure has a number of overlapping atoms.

Attractive induction interactions between polar exocyclic groups and the delocalized electrons of aromatic rings were

TABLE 2: MEP Atomic Charges of DNA Bases (q, e) and Their Dipole Moments (μ, D) and Polarizabilities (α, au^3) Calculated at the MP2/aug-cc-pVDZ and MP2/6-31G*(0.25) (in parentheses) Levels[a]

	q							
	cytosine		adenine		guanine		thymine	
C2	0.967 (0.881)	N3	−0.715 (−0.659)	O6	−0.500 (−0.499)	O4	−0.531 (−0.486)
N1	−0.634 (−0.601)	C2	0.465 (0.452)	C6	0.524 (0.602)	C4	0.683 (0.586)
C6	0.238 (0.207)	N1	−0.694 (−0.662)	N1	−0.710 (−0.747)	N3	−0.618 (−0.573)
C5	−0.740 (−0.653)	C6	0.603 (0.638)	C2	0.856 (0.841)	C2	0.744 (0.666)
C4	1.055 (0.970)	C5	0.128 (0.013)	N3	−0.713 (−0.654)	N1	−0.520 (−0.475)
N3	−0.817 (−0.757)	C4	0.591 (0.538)	C4	0.480 (0.394)	C6	−0.043 (−0.059)
O2	−0.619 (−0.577)	N9	−0.568 (−0.503)	C5	0.194 (0.051)	C5	−0.041 (0.055)
N4	−1.100 (−1.067)	C8	0.243 (0.277)	N7	−0.593 (−0.498)	O2	−0.559 (−0.520)
H1	0.359 (0.354)	N7	−0.578 (−0.522)	C8	0.302 (0.236)	CM	−0.474 (−0.576)
H6	0.138 (0.129)	N6	−0.897 (−0.844)	N9	−0.602 (−0.434)	H3	0.361 (0.354)
H5	0.244 (0.224)	H2	0.071 (0.047)	N2	−1.053 (−0.977)	H1	0.364 (0.351)
H41	0.451 (0.442)	H8	0.129 (0.089)	H1	0.393 (0.340)	H6	0.186 (0.173)
H42	0.458 (0.449)	H9	0.402 (0.369)	H8	0.105 (0.091)	HM1	0.158 (0.177)
		H61	0.421 (0.392)	H9	0.415 (0.356)	HM2	0.133 (0.151)
		H62	0.399 (0.375)	H21	0.458 (0.420)	HM3	0.158 (0.177)
				H22	0.444 (0.419)		
μ	6.49 (6.27)		2.56 (2.55)		6.65 (6.45)		4.31 (4.01)
α[b]	63.6		76.7		84.5		67.9

[a] The corresponding planar optimized geometries of cytosine (MP2/DZ(2d)) and adenine, thymine and guanine (MP2/6-31G*) were taken from ref 9. [b] Polarizability was calculated at the HF/6-31G*(0.25) level.

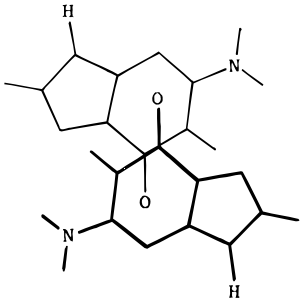

Figure 5. Guanine dimer with a stacking of carbonyl groups and aromatic rings. The MP2/6-31G*(0.25) interaction energy is −8.8 kcal/mol, and the empirical potential interaction energy is −8.9 kcal/mol. The geometry was taken from ref 3a where it corresponds to the local minimum on the potential energy surface.

TABLE 3: MP2/6-31G* Optimized Geometry (Cartesian Coordinates, Å), and the MEP Atomic Charges *q* (e) and Dipole Moment (μ, D) and Polarizability (α, au^3) of Uracil Calculated at the MP2/aug-cc-pVDZ Level[a]

	x	*y*	*z*	*q*
O4	0.000	0.000	0.000	−0.555 (−0.519)
C4	0.000	0.000	1.228	0.817 (0.744)
N3	1.211	0.000	1.944	−0.609 (−0.571)
C2	1.409	0.000	3.315	0.749 (0.681)
N1	0.214	0.000	4.023	−0.505 (−0.455)
C6	−1.029	0.000	3.431	0.133 (0.106)
C5	−1.176	0.000	2.088	−0.554 (−0.481)
O2	2.508	0.000	3.854	−0.554 (−0.519)
H5	−2.153	0.000	1.622	0.222 (0.198)
H3	2.059	0.000	1.383	0.349 (0.340)
H1	0.315	0.000	5.031	0.351 (0.334)
H6	−1.870	0.000	4.116	0.155 (0.144)
μ				4.37 (4.10)
α[b]				58.8

[a] The values in parentheses were obtained at the MP2/6-31G*(0.25) level. [b] Polarizability was calculated at the HF/6-31G*(0.25) level.

postulated to significantly stabilize the base stacking.[6,15] This case would be indicated by a disagreement between the ab initio and empirical potential data, because our empirical potential do not cover the induction contribution. We found no such stabilization for any dimer analyzed in this study. Figure 5 shows one example of a guanine dimer with an overlap of exocyclic carbonyl oxygens and the rings. A perfect agreement of the ab initio and empirical potential energy indicates that there is no significant stabilizing contribution originating in the induction interaction.

3.2. Basis Set Dependence of the Electrostatic Energy. In order to estimate the reliability of the electrostatic part of the MP2/6-31G*(0.25) ab initio interaction energies, the MEP charges were also derived at the MP2/aug-cc-pVDZ level (Tables 2 and 3). The electrostatic energies calculated by the MP2/6-31G*(0.25) and MP2/aug-cc-pVDZ charges were compared for all stacked dimers, covering twists 0−360° and displacements up to 3.0 A. The differences were generally within several tenths of kcal/mol, and the largest deviation found was smaller than 0.9 kcal/mol. The MP2/6-31G*(0.25) calculational level gives correct description of the electrostatic interactions.

3.3. Base Stacking in DNA Crystal Structures. Table 4 summarizes the MP2 interaction energy (E^{MP2}), its HF part (E^{HF}), and the empirical potential interaction energy (E^{EMP}) for the experimentally observed base pair steps. E^{MP2} is a sum of the HF interaction energy E^{HF} and the MP2 contribution to the correlation energy E^{COR}. For every step, the first two columns present the intrastrand contributions, the next two columns show the interstrand contributions (cf. Figure 2). E^{T} is the sum of the respective four base−base contributions. Its HF and MP2 values include also the many-body correction E^{MB}, calculated at the HF/MINI-1 level.

TABLE 4: Interaction Energies (kcal/mol) of Selected Stacked DNA Base Pair Steps[a]

step	C1.C2(G19.G20)−5DNB					
	C1C2	G19G20	C1G19	C2G20	E^{MB}	E^{T}
E^{HF}	+5.8	+10.0	−3.6	−2.8	+1.4	+10.7
E^{MP2}	−0.5	−3.4	−4.4	−3.0	+1.4	−9.9
E^{EMP}	−0.4	−2.9	−4.4	−3.3		−11.1
step	C5.G6(C15.G16)−5DNB					
	C5G6	C15G16	C5C15	G6G16	E^{MB}	E^{T}
E^{HF}	−1.1	−1.1	+2.3	+3.7	−0.5	+3.0
E^{MP2}	−4.7	−4.7	+0.8	−4.5	−0.5	−13.5
E^{EMP}	−5.3	−5.3	+0.8	−4.4		−14.3
step	C2.A3(T18.G19)−5DNB					
	C2A3	T18G19	C2T18	A3G19	E^{MB}	E^{T}
E^{HF}	+3.1	−2.1	+1.1	−0.6	−0.2	+1.4
E^{MP2}	−2.4	−5.3	−2.4	−2.4	−0.2	−12.7
E^{EMP}	−2.5	−5.6	−3.0	−3.0		−14.1
step	A3.A4(T17.T18)−5DNB					
	A3A4	T17T18	A3T17	A4T18	E^{MB}	E^{T}
E^{HF}	+5.2	+2.9	+0.1	−0.2	−0.3	+7.7
E^{MP2}	−6.3	−3.2	−1.5	−0.8	−0.3	−12.1
E^{EMP}	−6.7	−3.2	−2.1	−1.2		−13.2
step	A4.C5(G16T17)−5DNB					
	A4C5	G16T17	A4G16	C5T17	E^{MB}	E^{T}
E^{HF}	+5.2	+5.0	−0.7	+0.9	+0.2	+10.5
E^{MP2}	−4.7	−5.2	−3.4	+0.1	−0.2	−13.0
E^{EMP}	−4.5	−5.1	−3.9	−0.4		−13.9
step	G5.C6(G15.C16)−1DNB					
	G5C6	G15C16	G5G15	C6C16	E^{MB}	E^{T}
E^{HF}	−0.6	−0.6	+7.1	+4.8	−0.1	+10.4
E^{MP2}	−8.3	−8.3	+2.0	+3.0	−0.1	−11.8
E^{EMP}	−8.6	−8.6	+2.5	+3.1		−11.6
step	C2.G3(C10.G11)−2DCG					
	C2G3	C10G11	C2C10	G3G11	E^{MB}	E^{T}
E^{HF}	−1.7	−1.8	+2.8	+5.1	−0.7	+3.8
E^{MP2}	−8.7	−8.5	+1.5	+2.2	−0.7	−14.1
E^{EMP}	−9.2	−8.9	+1.4	+2.4		−14.3
step	G3.C4(G9.C10)−2DCG					
	G3C4	G9C10	G3G9	C4C10	E^{MB}	E^{T}
E^{HF}	−3.6	−3.8	+1.0	+5.7	−0.5	−1.1
E^{MP2}	−4.6	−5.0	+0.7	−2.2	−0.5	−11.6
E^{EMP}	−5.0	−5.3	+0.6	−2.0		−11.7

[a] E^{HF}, HF interaction energy; E^{MP2}, MP2 interaction energy (sum of E^{HF} and the electron correlation contribution); E^{EMP}, empirical potential energy calculated as a sum of the scaled 6-9LH van der Waals term and the Coulombic term with MP2/6-31G*(0.25) MEP charges; E^{MB}, many-body correction, obtained at the HF/MINI-1 level; E^{T}, total stacking energy of stacked base pair dimer; i.e., sum of the four base−base contributions, plus E^{MB} in case of the E^{HF} and E^{MP2} energies.

The calculations again show a very good agreement between the MP2 ab initio and empirical potential calculations; their difference is never larger than 0.6 kcal/mol for the individual base−base contributions. The many-body effects (covered by the HF calculations) are largest for the CC step (+1.4 kcal/mol), while they are smaller and mostly negative for the remaining steps.

The least stable base pair step is the C1C2 B-DNA step (−9.9

kcal/mol). Both cytosine and guanine have large dipole moments and the two intrastrand (almost parallel) CC and GG stacks are destabilized by the repulsive dipole−dipole interactions. This repulsion is partially eliminated by a significant overlap of the guanines,[13] giving a large dispersion attraction. The two cytosines are unstacked;[13] such mutual displacement decreases the dipole−dipole repulsion.[2j] The C1C2 step is mainly stabilized by the interstrand contributions (−7.4 kcal/mol). A quite different energy distribution was found for the AA B-DNA step (−13.0 kcal/mol). This step is mainly stabilized by the intrastrand stacking (−10.1 kcal/mol), while the interstrand contributions are rather small. Except the C1C2 B-DNA step, the other steps are stabilized by the intrastrand energy contribution. For the G5C6 B-DNA step the interstrand stacking is very repulsive (+5.0 kcal/mol). The most stable step among those analyzed is the C2G3 Z-DNA step (−14.1 kcal/mol), which is also characterized by a large intrastrand stabilization (−17.2 kcal/mol) and repulsive interstrand interactions (+3.7 kcal/mol). Because the GC Z-DNA step is rather stable (−11.7 kcal/mol), the base stacking in Z-DNA structure is not less stable than in the B-DNA. The C2A3 step of the 5DNB decamer exhibits a very unusual geometry[5c,13,15] with large slide (−2.5 Å), associated with large helical twist (51°) and negative base pair roll. Unusual conformational properties of the CpA steps are expected to play a unique role in biological processes.[22] It was proposed that attractive induction interactions of polar exocyclic groups with the delocalized electrons of the aromatic rings are responsible for the conformational properties of the CpA step.[15] However, Table 4 demonstrates that there is no significant difference between the ab initio and empirical potential data and there is thus no stabilization originating in the induction interactions. The base stacking in the CA step is rather stable (−12.7 kcal/mol) and can be fully covered by considering the electrostatic and van der Waals contributions.[5c]

The total HF interaction energy of the 8 steps analyzed varies from −1.1 kcal/mol to +10.7 kcal/mol, while its correlation contribution ranges from −10.5 kcal/mol to −23.5 kcal/mol. The stabilization of the stacked pairs thus originates in the electron correlation. The total MP2 ab initio stacking energy ranges from −9.9 to −14.1 kcal/mol. The MP2 interaction energy exhibits a much smaller sequence-dependent variability than its two separate contributions, E^{HF} and E^{COR}.

A similar rule seems to also hold for the interstrand and intrastrand stacking contributions. The intrastrand stacking energy varies from −3.9 to −17.2 kcal/mol, and the interstrand contribution varies from +5.0 to −7.4 kcal/mol. The sequence-dependent variations of the intrastrand and interstrand contributions compensate for each other, giving a rather low variability of the total stacking energy along the double helix.

It should be noted that all the above calculations were made in the gas phase. The stability of the particular steps within DNA can be influenced by solvent effects,[5f] entropy contributions, and other factors.

4. Conclusions

The most stable stacked pair is the G···G dimer (−11.3 kcal/mol), while the least stable is the uracil dimer (−6.5 kcal/mol). H-bonded pairs are more stable. The stability of stacked pairs originates in the electron correlation, while all the H-bonded pairs are dominated by the HF energy. The mutual orientation of the stacked bases is, however, primarily determined by the HF contribution to the interaction energy. The orientational dependence of stacking energy is dominated by changes of twist. The ab initio data are satisfactorily reproduced by the empirical potential calculations. Some contributions, previously postulated to significantly influence the base stacking (induction interactions, $\pi-\pi$ interactions), are negligible.

The HF and correlation contributions to the base stacking energy of DNA base pair steps show a rather large sequence-dependent variability. Also, the intrastrand and interstrand contributions exhibit a large variability. The sequence-dependent variations of the total base pair stacking energy are smaller, ranging from −9.9 to −14.7 kcal/mol. The range of the calculated many-body corrections to the stacking energy is 2 kcal/mol.

Acknowledgment. This study was supported in part by a ONR grant No. N00014-95-1-0049, by NIH grant No. 332090, and by grant 203/94/1490 from the grant agency of the Czech Republic. The Mississippi Center for Supercomputing Research is acknowledged for generous allotment of computer time. We thank M. J. Stewart for his helpful comments on the manuscript.

References and Notes

(1) (a) Calladine, C. R. *J. Mol. Biol.* **1982**, *161*, 343. (b) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025. (c) Drew, H. R.; Wing, R. M.; Takana, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 2179. (d) Kennard, O.; Hunter, W. N. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 1254.

(2) (a) Hobza, P.; Sandorfy, C.; *J. Am. Chem. Soc.* **1987**, *109*, 1302. (b) Colson, A.-O.; Besler, B.; Sevilla, M. D. *J. Phys. Chem.* **1993**, *97*, 13852. (c) Hrouda, V.; Florián, J.; Hobza, P. *J. Phys. Chem.* **1993**, *97*, 1542. (d) Aida, M. *J. Comput. Chem.* **1988**, *9*, 362. (e) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493. (f) Colson, A.-O.; Besler, B.; Close, D. M.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 661. (g) Florián, J.; Leszczynski, J. *J. Biomol. Struct. Dyn.* **1995**, *12*, 1055. (h) Sagarik, K. P.; Rode, B. M. *Inorg. Chim. Acta* **1983**, *78*, 177. (i) Anwander, E. H. S.; Probst, M. M.; Rode, B. M. *Biopolymers* **1990**, *29*, 757. (j) Hobza, P.; Šponer, J.; Polášek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792. (k) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965. (l) Šponer, J.; Leszczynski, J.; Hobza, P. *Computational Chemistry*: *Reviews of Current Trends*; World Scientific Publisher: Singapore, 1996; p 185.

(3) (a) Poltev, V. I.; Shulyupina, N. V. *J. Biomol. Struct. Dyn.* **1986**, *3*, 739. (b) Šponer, J.; Kypr, J. *J. Biomol. Struct. Dyn.* **1993**, *11*, 27.

(4) (a) Aida, M.; Nagata, C. *Chem. Phys. Lett.* **1982**, *86*, 44. (b) Aida, M., *J. Theor. Biol.* **1988**, *103*, 327. (c) Langlet, J.; Claverie, P.; Caron, F.; Boevue, J. C. *Int. J. Quantum Chem.* **1981**, *19*, 229. (d) Kudryatskaya, Z. G.; Danilov, V. I. *J. Theor. Biol.* **1976**, *59*, 303. (e) Ornstein, R.; Fresco, J. R. *Biopolymers* **1983**, *22*, 1979.

(5) (a) Šponer, J.; Kypr, J. *J. Biomol. Struct. Dyn.* **1990**, *7*, 1211. (b) Šponer, J.; Kypr, J. *J. Mol. Biol.* **1991**, *221*, 761. (c) Šponer, J.; Jursa, J.; Kypr, J. *Nucleosid. Nucleotid.* **1994**, *13*, 1669. (d) Srinivasan, A. R.; Torres, R.; Clark, W.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1987**, *5*, 459. (e) Ulyanov, N. B.; Zhurkin, V. B. *J. Biomol. Struct. Dyn.* **1984**, *2*, 361. (f) Friedman, R. A.; Honig, B. *Biopolymers* **1992**, *32*, 145. (g) Šponer, J.; Kypr, J. *J. Biomol. Struct. Dyn.* **1993**, *11*, 277.

(6) (a) Bugg, C. E.; Thomas, J. M.; Sundaralingam, M.; Rao, S. T. *Biopolymers* **1971**, *10*, 175.

(7) (a) Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709. (b) Šponer, J.; Burcl, R.; Hobza, P. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1357. (c) Šponer, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1994**, *12*, 671.

(8) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Comput. Chem.*, in press.

(9) Šponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161.

(10) (a) Kroon-Batenburg, L. M. J.; van Duijneveldt, F. B. *J. Mol. Struct.* **1985**, *121*, 185. (b) Hobza, P.; Zahradník, R. *Chem. Rev.* **1988**, *88*, 871. (c) Hobza, P.; Melhorn, A.; Čársky, P.; Zahradník, R. *J. Mol. Struct. (THEOCHEM)* **1986**, *138*, 387.

(11) (a) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553. (b) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. *Chem. Rev.* **1994**, *94*, 1873. (c) Chalasinsky, G.; Szczesniak, M. M. *Chem. Rev.* **1994**, *94*, 1723.

(12) Zhurkin, V. B.; Poltev, V. I.; Florentev, V. L. *Mol. Biol. (USSR)* **1980**, *14*, 1116 (in Russian).

(13) Privé, G. G.; Yanagi, K.; Dickerson, R. E. *J. Mol. Biol.* **1991**, *221*, 177.

(14) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Mayer, G. F.; Brice, M. D.; Rotgers, J. R.; Kennard, O.; Shimanouchi, T.; Tsaumi, M. *J. Mol. Biol.* **1977**, *112*, 535.

(15) Heinemann, U.; Alings, C. *J. Mol. Biol.* **1989**, *210*, 369.

(16) Wang, A. H.-J.; Quigley, G. J.; Kolpak, J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680.

(17) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. V.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Rob, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzales, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *GAUSSIAN 92*; Gaussian Inc.: Pittsburgh, PA.

(18) (a) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129. (b) Besler, B. H.; Merz, K. M. Jr.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(19) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(20) (a) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111. (b) Šponer, J.; Kypr J. In *Theoretical Biochemistry and Molecular Biophysics*; Beveridge, D. L., Lavery, R., Eds.; Adenine Press: New York, 1991; Vol. 1, p 271. (c) Šponer, J.; Kypr, J. *Gen. Physiol. Biophys.* **1989**, *8*, 257.

(21) (a) Stone, A. J.; Price, S. L. *J. Phys. Chem.* **1988**, *92*, 3325. (b) Stone, A. J.; Tong, C.-S. *J. Comput. Chem.* **1994**, *15*, 1377.

(22) (a) Nussinov, R. *Biochim. Biophys. Acta* **1986**, *866*, 93. (b) Schultz, S. C.; Shields, G. C.; Steitz, T. A. *Science* **1991**, *253*, 1001. (c) Pardue, M. L.; Lowenhaupt, K.; Rich, A.; Nordheim, A. *EMBO J.* **1987**, *6*, 1781. (d) Riazance-Lawrence, J. H.; Kang, H. S.; Chou, P. J.; Johnson, W. C.; Vorlíčková, M. *Biopolymers* **1994**, *34*, 1469. (e) Hamada, H.; Petrino, M. G.; Kakunaga, T. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6465.

JP953306E