Base Stacking in Cytosine Dimer. A Comparison of Correlated *Ab Initio* Calculations with Three Empirical Potential Models and Density Functional Theory Calculations

JIŘÍ ŠPONER,*[1,2] JERZY LESZCZYNSKI,[3] and PAVEL HOBZA[1]

[1]*J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague 8, Czech Republic;* [2]*Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic; and* [3]*Department of Chemistry, Jackson State University, Jackson, 39217 MS*

Received 3 March 1995; accepted 5 July 1995

ABSTRACT

Ab initio MP2/6-31G* interaction energies were calculated for more than 80 geometries of stacked cytosine dimer. Diffuse polarization functions were used to properly cover the dispersion energy. The results of *ab initio* calculations were compared with those obtained from three electrostatic empirical potential models, constructed as the sum of a Lennard-Jones potential (covering dispersion and repulsion contributions) and the electrostatic term. Point charges and point multipoles of the electrostatic term were also obtained at the MP2/6-31G* level of theory. The point charge MEP model (atomic charges derived from molecular electrostatic potential) satisfactorily reproduced the *ab initio* data. Addition of π-charges localized below and above the cytosine plane did not affect the calculated energies. The model employing the distributed multipole analysis gave worse agreement with the *ab initio* data than the MEP approach. The MP2 MEP charges were also derived using larger sets of atomic orbitals: cc-pVDZ, 6-311 + G(2d, p), and aug-cc-pVDZ. Differences between interaction energies calculated using these three sets of point charges and the MP2/6-31G* charges were smaller than 0.8 kcal/mol. The correlated *ab initio* calculations were also compared with the density functional theory (DFT) method. DFT calculations well reproduced the electrostatic part of interaction energy. They also covered some nonelectrostatic short-range effects which were not reproduced by the empirical potentials. The DFT method does not include the dispersion energy. This energy, approximated by an empirical term, was therefore added to the

*Author to whom all correspondence should be addressed.

Journal of Computational Chemistry, Vol. 17, No. 7, 841–850 (1996)

CCC 0192-8651 / 96 / 070841-10

DFT interaction energy. The resulting interaction energy exhibited an artifact secondary minimum for a 3.9–4.0 A vertical separation of bases. This defect is inherent in the DFT functionals, because it is not observed for the Hartree-Fock + dispersion interaction energy.

Introduction

Stacking and hydrogen bonding interactions of DNA bases represent an important source of conformational variability of DNA. Hydrogen bonding interactions are rather well understood and described. There is a number of quantum chemical studies convincingly demonstrating the dominating electrostatic contribution.[1-10] Hydrogen bonding seems to be satisfactorily described by the empircal potentials.[5,11] On the other hand, the nature of base stacking interactions is much less known. The differences among various empirical potentials are rather large,[12,13] and it cannot be ruled out that some important contributions are not covered by the simple empirical potentials at all.[14] Four rather contradictory opinions on the dominant contribution of the base stacking in DNA exist. Mostly, the electrostatic contribution[15-18] (described by a point charge model) or dispersion contribution[19-24] (included in the van der Waals term of empirical potential) are expected to be the leading terms. However, on the basis of X-ray geometries of base stacking, induction theory of stacking was proposed.[25] According to this theory, the dominant contributions are the interactions of polar exocyclic groups on the bases with delocalized electrons on the aromatic rings. These interactions were, for example, proposed to be responsible for the unusual high twist–high slide geometry of the CpA step, observed in B-DNA crystal structures.[26] However, recently a different explanation of conformational flexibility of the CpA steps was found.[27] It was also proposed that inclusion of so-called π-interactions[28] is quite important in calculations of base stacking.[29] An empirical potential "sandwich" model has been developed, which uses a set of π-point charges below and above the rings.[28,29] Also, the use of atomic multipoles instead of atomic charges is frequently recommended for the DNA base interactions,[30,31] though we are not aware of any calculations on base stacking using this method (with exception of our preceding article[32]).

The number of quantum chemical studies on base stacking[32-35] is much smaller than the number of studies on hydrogen-bonded complexes of DNA bases. We recently reported a first consistent comparison of a stacked and hydrogen-bonded DNA base pair (cytosine dimer)[32]; the calculations were made at the MP2/6-31G* level of theory. This analysis has shown that the geometry of the stacked cytosine dimer is primarily determined by the electrostatic interactions. (This does not mean, however, that the electrostatic contribution necessarily dominates the base stacking in a DNA double helix; see refs. 19 and 32.) Due to the large cytosine dipole moment, antiparallel orientation of the two stacked cytosines is strongly preferred. On the other hand, the stabilization of the stacked complex originates mainly in the electron correlation. The calculations did not give any support to the induction theory of stacking. It was also found that crystal structures of DNA bases and DNA constituents frequently exhibit unfavorable base stacking, which confirmed the previous empirical potential analysis by Poltev and Shulyupina.[36]

In this study, we extend our previous calculations by performing MP2/6-31G* calculations for 25 other geometries of cytosine dimer. The main aim of this article is to assess the ability of various less demanding methods to reproduce the benchmark correlated *ab initio* data. This comparison includes the DFT method (two nonlocal functionals are tested), the MEP point charge model (atomic charges derived from molecular electrostatic potential), the π-charges sandwich model, and the DMA (distributed multipole analysis) model.[31] A Lennard-Jones potential is used to include the dispersion–repulsion contribution. In addition, the MEP charges will be derived using larger sets of atomic orbitals: cc-pVDZ, 6-311 + G(2d, p), and aug-cc-pVDZ. This allows us to perform an indirect assessment of the MP2/6-31G* supermolecule *ab initio* interaction energies by comparing the electrostatic energies obtained using the MP2/6-31G*(0.25) MEP charges and charges obtained with more reliable basis sets.

Method

All calculations were made with coplanar DNA bases; the MP2/DZ(2d)[37] planar optimized geom-

etry of cytosine was used. The bases were perpendicular to the z-axis of coordinate system; the vertical separation (VS) of the bases is their distance along the z-axis.[22] In undisplaced structures, the centers of mass of the cytosines are stacked one above the other. Parallel structure means that, for a vertical separation of 0 A, the two cytosines coincide. Antiparallel geometry is achieved by rotating one of the cytosines by 180° (twist angle – TW) around the axis passing through the cytosine center of mass and perpendicular to the cytosine plane. In displaced structures, one of the cytosines is shifted in the x-y plane.[32] Displacement is determined by the displacement distance DI – the absolute value of displacement in A, and the displacement angle DA – the direction of displacement.[32] Figure 1 shows the geometry of cytosine dimer (antiparallel geometry, DI = 2.0 A) for displacement angles of 0°, 90°, 180°, and 270°.

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 with exponent of 0.25 added to the second-row elements[38,39] [designated 6-31G*(0.25)]. Both HF and MP2 contributions to the interaction energy were corrected for the basis set superposition error.[40-42]

Density functional theory (DFT) calculations were made using two nonlocal DFT functionals, BLYP[43-45] and Becke3LYP.[46] The following basis sets were used: 6-31G*, 6-31G*(0.25), 6-31 + G*, and DZP.[47] The present study is based on almost

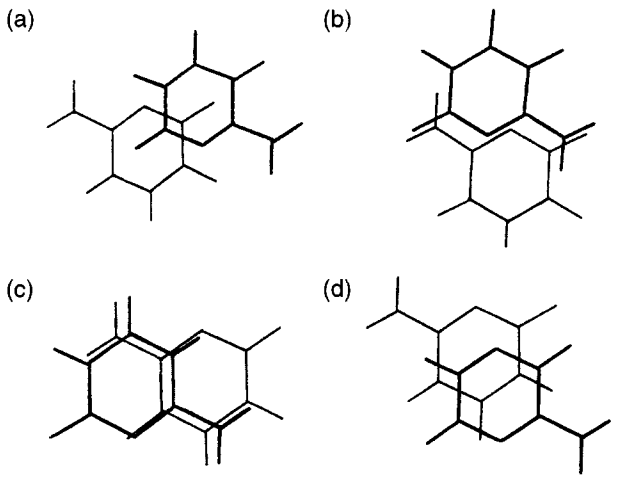

FIGURE 1. Four displaced antiparallel structures with the displacement distance DI = 2.0 A. (a) displacement angle (DA) = 0°, (b) DA = 90°, (c) DA = 180°, (d) DA = 270°.

100 DFT single-point calculations on the cytosine dimer. Also, the DFT interaction energies were corrected for the basis set superposition error. The correction is similar in magnitude to that at the Hartree-Fock (HF) level of theory. The DFT method does not include the intermolecular correlation energy.[48] This contribution was approximated using the empirical dispersion London-type expression,[39] with atomic polarizabilities and ionization potentials taken from the literature.[49] All *ab initio* and DFT calculations on the cytosine dimer were made with the Gaussian92/DFT sets of programs.[50]

All sets of point charges were derived from the molecular electric potential (MEP). Three calculational schemes[51-54] were tested. All these methods gave very similar results, and mostly the Merz-Kollman (MK) algorithm[51,52] was used. First, four sets (S1–S4) of MP2/6-31G*(0.25) atomic charges were derived to find out how different MEP models reproduce the respective supermolecule *ab initio* data:

- *S1 model*: 13 point charges localized on all atomic centers.
- *S2 model*: 13 point charges again localized on atomic centers, supplemented by the π-charges placed below and above all second-row atoms at a distance of 0.47 A from cytosine plane (16 point charges). Two other π-charges were added to the exocyclic carbonyl oxygen and amino group nitrogen atoms, lying in the cytosine plane, perpendicularly with respect to the O—C and N—C bonds (four charges). Their distance from the respective atomic centers was again 0.47 A. The distance of 0.47 A was recommended in refs. 28 and 29.
- *S3 model*: the same as S2, but another partial charge was placed at the center of mass of cytosine, and two others below and above it at a distance of 0.47 A from the molecular plane.
- *S4 model*: the same as S2, but a 0.9 A (instead of 0.47 A) distance was used for all the π-charges.

The S3 and S4 sets were used in order to find out whether the calculated interaction energies change upon a significant variation of the separation of the π-charges from the molecular plane, or upon placing additional charge centers on the aromatic ring. As is shown later, all four sets of charges give very similar electrostatic interaction energies, all

being fully in agreement with the *ab initio* supermolecule data. Therefore, no other set of charges was considered.

Instead, S1-type sets of point charges were derived using various basis sets, with and without inclusion of electron correlation. The following calculational levels were used: HF/STO-3G (this level of theory was originally used in parameterization of the AMBER[51,55,56] force field), HF/6-31G*(0.25), HF/6-31G*, BLYP/6-31G*(0.25), Becke3LYP/6-31G*(0.25), Becke3LYP/DZP, MP2/6-31G*, MP2/DZP, MP2/cc-pVDZ,[57] MP2/6-311 + G(2d, p), and MP2/aug-cc-pVDZ.[58] All sets of charges were derived using the Gaussian 92 set of programs.[50]

The empirical potential interaction energy was calculated using the expression

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

where A_{ij} and B_{ij} are empirical van der Waals constants, r_{ij} were distances between atoms or point charges i and j, and q_i were the MEP charges. The 6-9 Lifson-Hagler (6-9LH) force field[59] scaled by a factor of 0.7 was used as a van der Waals potential. The 6-9LH force field was used in a number of our previous studies,[12,13,20-23,27] and it provided a satisfactory description of base stacking in DNA crystal structures.[13] The scaling factor was applied in order to approximate the *ab initio* MP2/6-31G*(0.25) interaction energies. The choice of van der Waals potential has a negligible influence on the problems discussed in the present study. In fact, the change of van der Waals energy with twist and displacement angle is negligible, because the van der Waals energy is proportional to the geometrical overlap of bases.[32]

The atomic point multipoles (up to hexadecupoles) were derived at the MP2/6-31G*(0.25) level using the CADPAC[60] program package. The DMA electrostatic interaction energy was calculated using the ORIENT code[61] (all contributions up to the r^{-5} rank were included[61]) and combined with a scaled 6-9LH van der Waals potential.

Results and Discussion

AB INITIO SUPERMOLECULAR CALCULATIONS

Figure 2 shows the dependence of the MP2/6-31G*(0.25) interaction energy on various conformational parameters of the stacked cytosine dimer.

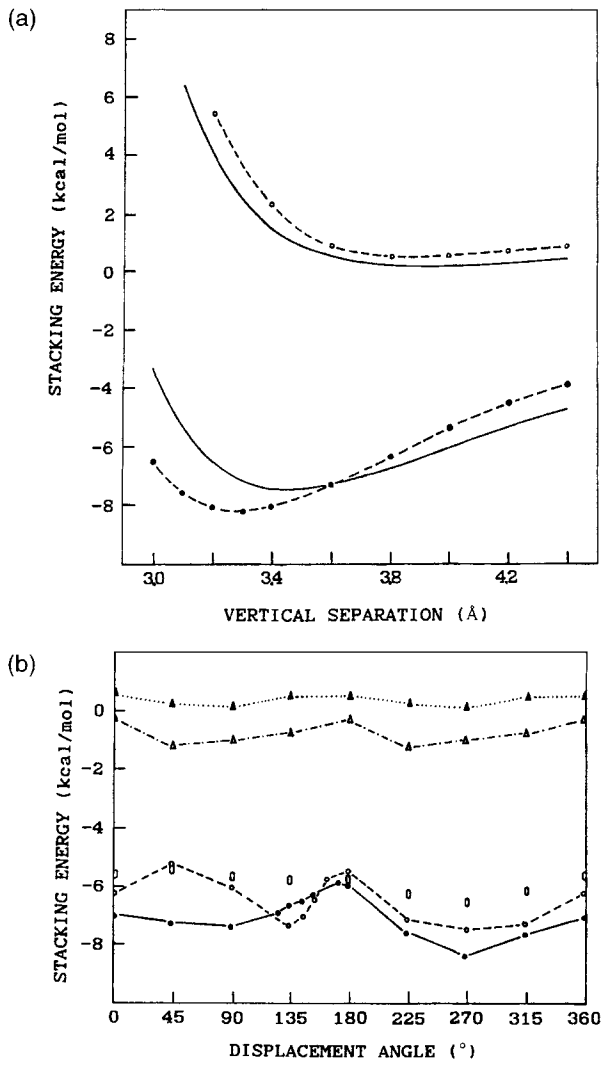

FIGURE 2. (a) Dependence of the MP2/6-31G*(0.25) interaction energy of the cytosine dimer on vertical separation (VS) of cytosines for parallel undisplaced (top circles) and antiparallel undisplaced (bottom circles) structures. The thin solid lines represent the respective empirical potential data. (b) Dependence of the MP2/6-31G*(0.25) interaction energy of the cytosine dimer on displacement angle: solid line — antiparallel structure, DI = 1.0 A, VS = 3.3 A; dashed line — antiparallel structure, DI = 2.0 A, VS = 3.3 A; dotted line — parallel structure, DI = 1.0 A, VS = 3.3 A; dot-dash line — parallel structure, DI = 2.0 A, VS = 3.3 A; 0 — antiparallel structure, DI = 1.0 A, VS = 3.8 A.

Figure 2a presents the dependence of interaction energy on the vertical separation of bases for the antiparallel (top circles) and parallel (bottom circles) undisplaced structures. The thin solid lines show the respective empirical potential data. The antiparallel geometry is very stable, −8.3

kcal/mol, with the optimum vertical separation of bases 3.25 A. This value seems to agree with the average vertical separation of base pairs (3.3–3.4 A) in DNA.[12,13] The parallel structure is purely repulsive. The energy difference between the parallel and antiparallel arrangements is due to the dipole–dipole interactions[32]; the antiparallel dimer is optimal. Figure 2b presents the dependence of interaction energy on the displacement angle (DA) for the displacement distances (DI) of 1.0 A (solid and dotted line for antiparallel and parallel structures, respectively) and 2.0 A (dash and dot-dash line for antiparallel and parallel structures, respectively). A zero (0) represents the data for the antiparallel dimer with DI = 1.0 A and vertical base separation of 3.8 A. The antiparallel geometry is always favored over the parallel arrangement. Two conformational regions for the antiparallel dimer deserve comment. First, the optimum geometry of the cytosine dimer was found for a displaced structure. (For DI = 1.0 A and DA = 270° A, the interaction energy is improved by 0.2 kcal/mol compared to the undisplaced structure.) This could be due to the interaction of cytosine quadrupoles; the quadrupole–quadrupole electrostatic term is repulsive for stacked structures and attractive for displaced structures.[62] Second, there is a sharp positive energy peak near DA = 180° mentioned already in the previous study.[32] In the present study, we calculated a number of other points in this region of conformational space. This peak is of a short-range origin, because it disappears if the vertical separation of bases is increased. It will be shown later that this peak is not reproduced by the empirical potentials, but it is reproduced by the DFT method.

The MP2/6-31G*(0.25) calculations represent the most sophisticated calculations yet published on a stacked DNA pair. Although both higher contributions to the correlation energy and extension of the basis set can still change the result,[63] the present calculations provide a unique set of data to assess the ability of other calculational methods to reproduce the benchmark *ab initio* calculations.

EMPIRICAL POTENTIAL CALCULATIONS

Let us first compare the *ab initio* data and the three electrostatic models: the standard MEP model (S1 set of point charges), the extended π-charges MEP model (S2–S4), and the DMA model. All methods satisfactorily reproduce the dependence of interaction energy on twist angle (not shown) and on displacement for the parallel structure (not shown). The empirical potential predicts somewhat larger optimum vertical base separation for the antiparallel structure (3.4 A compared to 3.25 A obtained by the MP2 procedure, cf. Fig. 2a), while it somewhat underestimates the short-range repulsion for the parallel structure. Note that the optimum vertical base separation calculated by the empirical potential can be easily varied by a small modification of the van der Waals part of the empirical potential.[12,13]

Figure 3 compares the MP2 data for the antiparallel dimer and DI = 1.0 A with the S1 MEP model and DMA calculations. For the sake of completeness, the DFT data are also presented. The extended MEP models give results very close to those obtained by the S1 model (not shown). The dot-dot-dash curve represents the net van der Waals contribution to the empirical potential energy. Figure 3 displays the large differences between the various methods. First, for the dependence shown, the DMA model significantly exaggerates the difference between the energy maximum and minimum compared to the MP2, DFT, and MEP data. The same result was obtained for displacement distance of 2.0 A (not shown). Second, the empirical potential models do not re-

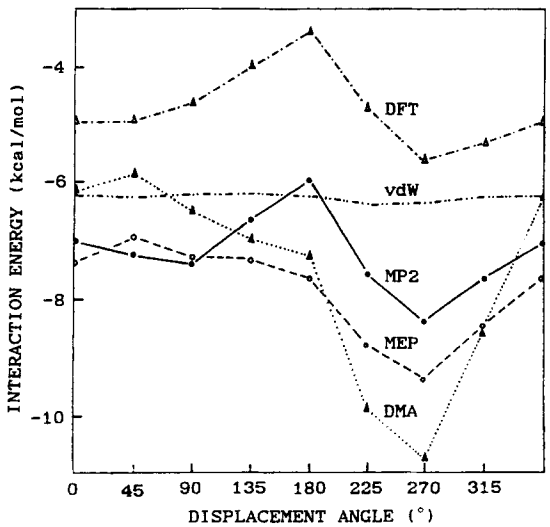

FIGURE 3. Dependence of interaction energy of the cytosine dimer on displacement angle DA, antiparallel structure, DI = 1.0 A, VS = 3.3 A. Solid line — MP2/6-31G*(0.25) data; dashed solid line — S1 MEP model + scaled 6-9LH potential; dotted line — DMA model + scaled 6-9LH potential; dot and dash line — BLYP/6-31G*(0.25) energy + empirical dispersion energy; dot-dot-dash line — scaled 6-9LH potential.

produce the positive peak near DA = 180° (for comparison with the *ab initio* data, the shape of the curve is the most important property, because the absolute value of energy can be adjusted by the van der Waals term). This peak is, however, reproduced by the DFT BLYP/6-31G*(0.25) data (Fig. 3), as well as by the BLYP/DZP and Becke3LYP/6-31G*(0.25) calculations.

Due to its short-range origin (see above), the peak should be reproduced by some short-range term rather than by the electrostatic term. However, the 6-9LH potential provides a very flat dependence of energy on DA. A number of other van der Waals potentials, including AMBER, provide a very similar energy dependence. It is possible to get an undulating van der Waals energy surface if substantial differences of atomic van der Waals radii are introduced, but this would lead to other defects.[13] The Lennard-Jones potentials assume spherical atoms, and this simplification could be responsible for their inability to reproduce the peak, because the short-range exchange repulsion is proportional to the overlap integral and is thus anisotropic.[64,65] Work is in progress to analyze this problem using the decomposition of interaction energy. Figure 4 demonstrates that the "repulsive" geometry is associated with several couples of overlapping atoms, including two N2...N1 contacts.

The difference between the MP2 supermolecule data and the MEP + 6-9LH potential data is never larger than 1.5 kcal/mol for more than 80 different structures of the cytosine dimer analyzed here, except the two antiparallel structures with vertical base separation of 3.1 and 3.0 A (Fig. 2a). The empirical potential calculations overestimate the short-range repulsion for the antiparallel structure and underestimate the short-range repulsion for the parallel structure. A variation of the van der Waals parameters as well as the use of a 12th power of the repulsion term did not eliminate this defect. 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. This finding could also be associated with the anisotropic nature of the short-range exchange repulsion contribution.

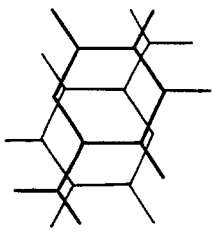

FIGURE 4. Stacked cytosine dimer, antiparallel structure, DA = 180°, DI = 1.0 A.

TABLE I.
The Dependence of Electrostatic Energy (kcal / mol) of Stacked Cytosine Dimer on Twist Angle (°) Calculated Using the S1 Set of Charges, and the Differences between Energies Obtained with the S1 Set and the Remaining Sets.

	Electrostatic Energy			
Twist	S1	S2 − S1	S3 − S1	S4 − S1
0	6.95	−0.10	−0.01	−0.09
30	4.89	−0.12	−0.11	−0.12
60	1.88	+0.19	+0.17	+0.18
90	−0.65	+0.22	+0.19	+0.20
120	−1.84	+0.14	+0.11	+0.10
150	−2.04	+0.10	+0.06	+0.05
180	−2.44	−0.02	−0.05	−0.03

A comparison of the standard S1 set of point charges with the sets S2–S4 (almost 1400 points were used, covering the full region of twist, and displacements up to 3 A) did not reveal differences larger than several tenths of kcal/mol. Thus there is no significant change in the calculated energies upon inclusion of the π-charges (see also Table I). It can be concluded that among the empirical potential models tested, the simplest standard S1 MEP model provides a very good approximation of the *ab initio* data for the stacked cytosine dimer. This does not mean that this approximation could not fail for other systems[66] or geometries not analyzed here; this uncertainty is a drawback of any empirical model. We are currently analyzing nine other stacked dimers of DNA bases by the MP2 *ab initio* method. Preliminary results (more than 150 points) again show a good agreement between *ab initio* data and the S1 MEP model, combined with the scaled 6-9LH potential.

The Hunter and Sanders π-charges model[28,29] differs from the π-charges model used here. Hunter calculated the σ and π electron distribution using the Del Re and HMO empirical methods, and the distance of the π-charges from the molecular plane was chosen rather arbitrarily. The old empirical procedures used are not reliable enough to calculate the σ and π electron distribution. Moreover, the calculated interaction energies could be sensitive to variation of the distance of π-charges from

the molecular plane. In this study, the out-of-plane π-charges were calculated using correlated *ab initio* calculations and the MEP procedure, which excluded any adjustable parameter in the calculations (except the atomic van der Waals radii used in the MEP fit; we used the standard values[50]). The MEP charges have no physical meaning, because they are fitted to reproduce the molecular electric potential. They effectively include the higher electric multipole moments.[51]

It was shown that the MEP point charge model reproduces the supermolecule *ab initio* data for the stacked cytosine dimer. However, a question arises: How reliable are the *ab initio* data? It is at present not possible to perform supermolecule calculations at a higher level of theory with large sets of atomic orbitals. However, because the MEP model reproduces the supermolecule data rather well, we can indirectly assess the reliability of the electrostatic part of the MP2/6-31G*(0.25) interaction energy. This is one of the leading contributions, and the twist-displacement energy dependence is primarily determined by the electrostatic term. Several other sets of MEP charges were derived, and the electrostatic energies calculated by them were compared. The full range of twist and displacements up to 3 A were considered. The results, which are encouraging, are summarized in Table II. The MP2/6-311 + G(2d, p) and MP2/aug-cc-pVDZ levels represent a reliable method to calculate the electrostatic potential of cytosine. The differences between the energies calculated using the MP2/6-31G*(0.25) charges and those obtained with the mentioned sophisticated basis sets are small, with the largest deviations smaller (in absolute value) than 0.8 kcal/mol. On the other hand, the HF/STO-3G MEP charges provide larger differences from the benchmark sets of charges. The charges obtained by standard 6-31G* and diffuse 6-31G*(0.25) basis sets are very similar. The DFT Becke3LYP/6-31G*(0.25) and BLYP/6-31G*(0.25) charges are very close to the MP2/6-31G*(0.25) charges. This confirms that the electrostatic interaction provided by the nonlocal DFT method is very close to the MP2 calculations. On the other hand, the results obtained by the HF/6-31G* MEP charges are far from those obtained by charges derived from correlated wave functions. This is due to the fact that the HF level of theory significantly overestimates the dipole moments of DNA bases.[32,67] However, if a recommended scaling factor of 0.85[67] is used for the HF/6-31G* MEP charges, the results become very close to the MP2/6-31G* MEP calculations. Table III summarizes several sets of the MEP charges.

DFT AND HF CALCULATIONS

It was shown earlier that the DFT calculations satisfactorily reproduce the twist-displacement de-

TABLE II.
Maximum (MAX) and Minimum (MIN) Electrostatic Energy Difference (kcal / mol) among Various Sets of MEP Charges.

Set A	Set B	MAX(B − A)	MIN(B − A)
MP2 / 6-31G*(0.25)	HF / STO-3G	2.36	−1.57
MP2 / 6-31G*(0.25)	HF / 6-31G*(0.25)	1.99	−2.28
MP2 / 6-31G*(0.25)	HF / 6-31G*(0.25)[a]	0.60	−0.50
MP2 / 6-31G*(0.25)	HF / 6-31G*	2.20	−2.62
MP2 / 6-31G*(0.25)	HF / 6-31G*[a]	0.67	−0.47
MP2 / 6-31G*(0.25)	BLYP / 6-31G*(0.25)	0.50	−0.27
MP2 / 6-31G*(0.25)	MP2 / DZP	1.20	−0.80
MP2 / 6-31G*(0.25)	MP2 / cc-pVDZ	0.30	−0.48
MP2 / 6-31G*(0.25)	MP2 / 6-311 + G(2d, p)	0.44	−0.76
MP2 / 6-31G*(0.25)	MP2 / aug-cc-pVDZ	0.36	−0.56
MP2 / aug-cc-pVDZ	HF / STO-3G	2.92	−1.79
MP2 / aug-cc-pVDZ	MP2 / 6-31G*(0.25)	0.56	−0.36
MP2 / aug-cc-pVDZ	MP2 / DZP	0.97	−0.43
MP2 / aug-cc-pVDZ	MP2 / cc-pVDZ	0.68	−0.79
MP2 / aug-cc-pVDZ	MP2 / 6-311 + G(2d, p)	0.19	−0.26

Twist angle was varied from 0° to 180°, displacement was scanned up to 3.0 A, vertical separation of cytosines was 3.3 A.
[a]The charges were scaled by a factor of 0.85.[67]

TABLE III.
The Cytosine Dipole Moment (D) and MEP Atomic Charges (au) Derived with Different Basis Sets and Methods.

MP2 / DZ(2d) Geometry of Cytosine[32]				MEP Charges				
	x(A)	y(A)	z(A)	6-31G*(0.25) MP2	6-31G*(0.25) BLYP	cc-pVDZ MP2	6-311 + G(2d, p) MP2	aug-cc-pVDZ MP2
C2	0.0000	0.0000	0.0	0.881	0.844	0.864	0.944	0.967
N1	0.0000	1.4174	0.0	−0.601	−0.557	−0.550	−0.623	−0.634
C6	1.1245	2.1722	0.0	0.207	0.224	0.187	0.206	0.238
C5	2.3480	1.5657	0.0	−0.653	−0.595	−0.660	−0.718	−0.740
C4	2.3336	0.1264	0.0	0.970	0.875	0.916	1.044	1.055
N3	1.2351	−0.6064	0.0	−0.757	−0.725	−0.763	−0.808	−0.817
O2	−1.0738	−0.5887	0.0	−0.577	−0.571	−0.560	−0.612	−0.619
N4	3.5186	−0.5430	0.0	−1.067	−0.926	−0.950	−1.123	−1.100
H1	−0.9236	1.8425	0.0	0.354	0.329	0.340	0.366	0.359
H6	0.9919	3.2549	0.0	0.129	0.106	0.133	0.148	0.138
H5	3.2687	2.1458	0.0	0.224	0.193	0.220	0.245	0.244
H41	4.4049	−0.0606	0.0	0.442	0.397	0.412	0.464	0.451
H42	3.4954	−1.5554	0.0	0.449	0.405	0.412	0.467	0.458
Dipole moment (D)				6.19	6.30	6.03	6.56	6.49

pendence of the MP2 interaction energy in the cytosine dimer. The DFT method even reproduces the positive peak of energy (Fig. 3), which is not reproduced by the empirical potentials. On the other hand, Figure 3 demonstrates that the DFT method combined with the empirical dispersion energy underestimates the total binding energy of the cytosine dimer, compared to the MP2 level. This could indicate that the empirical dispersion energy underestimates the intermolecular correlation contribution. (For the methylamine dimer, an opposite result was reported.[68]) However, the dependence of the interaction energy on vertical separation (VS) of the cytosines (Fig. 5) demonstrates

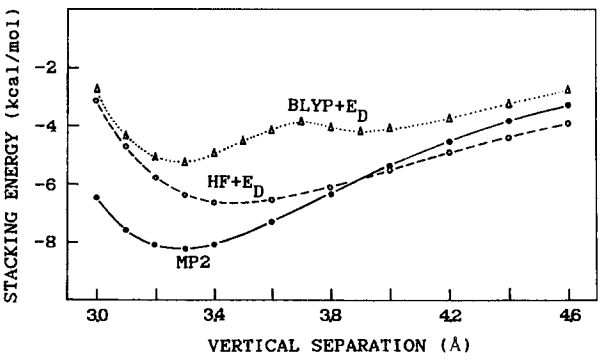

FIGURE 5. Dependence of the interaction energy of the cytosine dimer on vertical separation of bases for antiparallel undisplaced structure. Solid line — MP2 / 6-31G*(0.25) data; dashed line — HF / 6-31G*(0.25) data + empirical dispersion energy; dotted line — BLYP / 6-31G*(0.25) data + empirical dispersion energy.

that the case is more complicated. The BLYP/6-31G*(0.25) method gives an artificial flat secondary energy minimum at the VS 3.9–4.0 A (Fig. 5). BLYP/DZP and Becke3LYP/6-31G* calculations lead to the same result. This could be attributed to the fact that current nonlocal DFT functionals are inherently incorrect in the region of larger interatomic distances.[69] The DFT method needs some further improvements in order to describe the base stacking correctly.[69]

The HF energy combined with empirical dispersion energy gives a reasonable description of the dependence of interaction energy on VS without any false secondary minimum, though the slope of the dependence differs from the MP2 calculations. The HF method, however, provides incorrect twist-displacement dependence of stacking energy.[32]

Concluding Remarks

None of the simpler methods is able to reproduce fully the MP2 *ab initio* data on the stacked cytosine dimer. However, the MEP charges model represents a very good and fast computational tool to evaluate the base stacking energies, provided that the charges were derived with a sufficiently large basis set and with inclusion of electron correlation. The DFT method needs further improvement to calculate the base stacking satisfactorily. This method can also be used to derive the MEP charges.

Acknowledgments

This study was supported in part by NIH grant 332090 and by a contract (DAAL 03-89-0038) between the Army Research Office and the University of Minnesota for the Army High Performance Computing Research Center and JSU/Earth Observing System Program NCA & T Subcontract. The Mississippi Center for Supercomputing Research is acknowledged for generous allotment of computer time.

References

1. P. Hobza and C. Sandorfy, *J. Am. Chem. Soc.*, **109**, 1302 (1987).
2. A.-O. Colson, B. Besler, and M. D. Sevilla, *J. Phys. Chem.*, **97**, 13852 (1993).
3. V. Hrouda, J. Florián, and P. Hobza, *J. Phys. Chem.*, **97**, 1542 (1993).
4. M. Aida, *J. Comp. Chem.*, **9**, 362 (1988).
5. I. R. Gould and P. A. Kollman, *J. Am. Chem. Soc.*, **116**, 2493 (1994).
6. A.-O. Colson, B. Besler, D. M. Close, and M. D. Sevilla, *J. Phys. Chem.*, **96**, 661 (1992).
7. J. Florián and J. Leszczynski, *J. Biomol. Struct. Dyn.*, **12**, 1055 (1995).
8. K. P. Sagarik and B. M. Rode, *Inorg. Chim. Acta*, **78**, 177 (1983).
9. E. H. S. Anwander, M. M. Probst, and B. M. Rode, *Biopolymers*, **29**, 757 (1990).
10. J. Šponer and P. Hobza, *Chem. Phys.*, in press.
11. J. Pranata, S. G. Wierschke, and W. L. Jorgensen, *J. Am. Chem. Soc.*, **113**, 2810 (1991).
12. J. Šponer and J. Kypr, In *Theoretical Biochemistry and Molecular Biophysics*, Vol. I, D. L. Beveridge and R. Lavery, Eds., Adenine Press, New York, 1991, p. 271.
13. J. Šponer and J. Kypr, *J. Biomol. Struct. Dyn.*, **11**, 277 (1993).
14. J. Šponer and P. Hobza, *J. Am. Chem. Soc.*, **116**, 709 (1994).
15. A. Sarai, J. Mazur, R. Nussinov, and R. L. Jernigan, *Biochemistry*, **27**, 8498 (1988).
16. R. Rein, In *Intermolecular Interactions: From Diatomic to Bipolymers*: B. Pullman, Ed., Wiley-Interscience, New York, 1979, p. 307.
17. Z. G. Kudryatskaya and V. I. Danilov, *J. Theor. Biol.*, **59**, 303 (1976).
18. R. L. Ornstein, R. Rein, D. L. Breen, and R. D. MacElroy, *Biopolymers*, **17**, 2341 (1978).
19. R. A. Friedman and B. Honig, *Biopolymers*, **32**, 145 (1992).
20. J. Šponer and J. Kypr, *J. Biomol. Struct. Dyn.*, **7**, 1211 (1990).
21. J. Šponer and J. Kypr, *J. Mol. Biol.*, **221**, 761 (1991).
22. J. Šponer and J. Kypr, *J. Biomol. Struct. Dyn.*, **11**, 27 (1993).
23. R. E. Dickerson, *J. Mol. Biol.*, **166**, 419 (1983).
24. T. E. Haran, Z. Berkovich-Yellin, and Z. Shakked, *J. Biomol. Struct. Dyn.*, **2**, 397 (1984).
25. C. E. Bugg, J. M. Thomas, M. Sundaralingam, and S. T. Rao, *Biopolymers*, **10**, 175 (1971).
26. U. Heinemann and C. Alings, *J. Mol. Biol.*, **210**, 369 (1989).
27. J. Šponer, J. Jursa, and J. Kypr, *Nucleosid. Nucleotid.*, **13**, 1669 (1994).
28. C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, **112**, 5525 (1990).
29. C. A. Hunter, *J. Mol. Biol.*, **230**, 1025 (1993).
30. W. A. Sokalski, P. C. Hariharan, and J. J. Kaufman, *Int. J. Quantum Chem. Quantum Biol. Symp.*, **14**, 111 (1987).
31. S. L. Price and A. J. Stone, *J. Chem. Phys.*, **86**, 2859 (1987).
32. P. Hobza, J. Šponer, and M. Polášek, *J. Am. Chem. Soc.*, **117**, 792 (1995).
33. M. Aida and C. Nagata, *Chem. Phys. Lett.*, **86**, 44 (1982).
34. M. Aida, *J. Theor. Biol.*, **130**, 327 (1988).
35. J. Langlet, P. Claverie, F. Caron, and J. C. Boevue, *Int. J. Quantum Chem.*, **19**, 229 (1981).
36. V. I. Poltev and N. V. Shulyupina, *J. Biomol. Struct. Dyn.*, **3**, 739 (1986).
37. J. Šponer and P. Hobza, *J. Phys. Chem.*, **98**, 3161 (1994).
38. L. M. J. Kroon-Batenburg and F. B. van Duijneveldt, *J. Mol. Struct.*, **121**, 185 (1985).
39. P. Hobza and R. Zahradnik, *Chem. Rev.*, **88**, 871 (1988).
40. S. F. Boys and F. Bernardi, *Mol. Phys.*, **19**, 553 (1970).
41. F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. van Lenthe, *Chem. Rev.*, **94**, 1873 (1994).
42. G. Chalasinsky and M. M. Szczesniak, *Chem. Rev.*, **94**, 1723 (1994).
43. A. D. Becke, *Phys. Rev. A*, **38**, 3098, 1988.
44. C. Lee, W. Yang, and G. Paar, *Phys. Rev. B*, **37**, 785, 1987.
45. B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.*, **157**, 200 (1989).
46. A. D. Becke, *J. Chem. Phys.*, **98**, 5648 (1994).
47. T. H. Dunning, Jr., *J. Chem. Phys.*, **53**, 2823 (1970).
48. P. Hobza, J. Šponer, and T. Reschel, *J. Comput. Chem.*, **16**, 1315 (1995).
49. Y. K. Kang and M. S. Jhon, *Theoret. Chim. Acta*, **61**, 41 (1982).
50. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Rob, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzales, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, GAUSSIAN 92, Gaussian Inc., Pittsburgh, PA, 1995.
51. U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **5**, 129 (1984).
52. B. H. Besler, K. M. Merz, Jr., and P. A. Kollman, *J. Comput. Chem.*, **11**, 431 (1990).
53. L. E. Chirlian and M. M. Francl, *J. Comput. Chem.*, **8**, 894 (1987).
54. C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, **11**, 361 (1990).
55. S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, *J. Comput. Chem.*, **7**, 230 (1986).
56. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner, *J. Am. Chem. Soc.*, **106**, 765 (1984).

57. T. H. Dunning, Jr., *J. Chem. Phys.*, **90**, 1007 (1989).
58. R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.*, **96**, 6796 (1992).
59. S. Lifson, A. T. Hagler, and P. Dauber, *J. Am. Chem. Soc.*, **101**, 5111 (1979).
60. R. D. Amos, I. L. Alberts, J. S. Andrews, S. M. Colwell, N. C. Handy, D. Jayatilaka, P. J. Knowles, R. Kobayashi, N. Koga, K. E. Laidig, P. E. Maslen, C. W. Murray, J. E. Rick, J. Sanz, E. D. Simandiras, A. J. Stone, and M.-D. Su, CADPAC5, Cambridge 1992.
61. A. J. Stone, ORIENT 2.2, Cambridge, UK, 1992.
62. P. Hobza, H. L. Selzle, and E. W. Schlag, *J. Am. Chem. Soc.*, **116**, 3500 (1994).
63. P. Hobza, H. L. Selzle, and E. W. Schlag, *Chem. Rev.*, **94**, 1767 (1994).
64. A. J. Stone and S. L. Price, *J. Phys. Chem.*, **92**, 3325 (1988).
65. A. J. Stone and C.-S. Tong, *J. Comput. Chem.*, **15**, 1377 (1994).
66. D. E. Williams, *J. Comput. Chem.*, **15**, 719 (1994).
67. P. Cieplak, P. Bash, C. U. Singh, and P. A. Kollman, *J. Am. Chem. Soc.*, **109**, 6283 (1987).
68. J. Šponer, R. Burcl, and P. Hobza, *J. Biomol. Struct. Dyn.*, **11**, 1357 (1994).
69. N. C. Handy, "Advances in Computational Chemistry. Third Conference on Current Trends in Computational Chemistry," November 10–11, 1994, Vicksburg, MS.