The Physical Basis of Nucleic Acid Base Stacking in Water

Ray Luo, Hillary S. R. Gilson, Michael J. Potter, and Michael K. Gilson
Center for Advanced Research in Biotechnology, Rockville, Maryland 20850 USA

ABSTRACT It has been argued that the stacking of adenyl groups in water must be driven primarily by electrostatic interactions, based upon NMR data showing stacking for two adenyl groups joined by a 3-atom linker but not for two naphthyl groups joined by the same linker. In contrast, theoretical work has suggested that adenine stacking is driven primarily by nonelectrostatic forces, and that electrostatic interactions actually produce a net repulsion between adenines stacking in water. The present study provides evidence that the experimental data for the 3-atom-linked bis-adenyl and bis-naphthyl compounds are consistent with the theory indicating that nonelectrostatic interactions drive adenine stacking. First, a theoretical conformational analysis is found to reproduce the observed ranking of the stacking tendencies of the compounds studied experimentally. A geometric analysis identifies two possible reasons, other than stronger electrostatic interactions, why the 3-atom-linked bis-adenyl compounds should stack more than the bis-naphthyl compounds. First, stacked naphthyl groups tend to lie further apart than stacked adenyl groups, based upon both quantum calculations and crystal structures. This may prevent the bis-naphthyl compound from stacking as extensively as the bis-adenyl compound. Second, geometric analysis shows that more stacked conformations are sterically accessible to the bis-adenyl compound than to the bis-naphthyl compound because the linker is attached to the sides of the adenyl groups, but to the ends of the naphthyl groups. Finally, ab initio quantum mechanics calculations and energy decompositions for relevant conformations of adenine and naphthalene dimers support the view that stacking in these compounds is driven primarily by nonelectrostatic interactions. The present analysis illustrates the importance of considering all aspects of a molecular system when interpreting experimental data, and the value of computer models as an adjunct to chemical intuition.

INTRODUCTION

Energetically favorable stacking interactions between nucleic acid bases in water are believed to play an important role in determining and stabilizing the secondary and tertiary structures of DNA and RNA, as reviewed elsewhere (Saenger, 1984; Kool, 1997). However, the physics of base-stacking is subject to debate. In a series of elegant experiments, Newcomb and Gellman (1994) synthesized compounds with two aromatic groups joined by a 3-atom linker that only allows association of the groups in near-parallel arrangements (Fig. 1,*a* and *b*). Significant stacking of adenyl groups was found by chemical shift analysis, but napthyl groups did not stack detectably. Newcomb and Gellman used these results to discuss the interactions between adenyl groups in terms of the hydrophobic effect, dispersion forces, and interactions between partial positive and negative charges on atoms. They reasoned that the hydrophobic and dispersion forces acting between naphthyl groups are similar to those acting between adenyl groups. Thus, the observation that only the adenyls associated suggested that the only remaining interaction in their analysis, interactions between partial atomic charges, was responsible for adenyl–adenyl stacking in water.

Friedman and Honig (1995) subsequently used an empirical force field and a continuum solvation model of water to examine energetic contributions to adenine stacking. They reported that the net effect of the partial atomic charges assigned to adenine in the AMBER (Weiner et al., 1984) force field was to oppose the stacking of adenines in water, due primarily to the energetic cost of desolvating these polar heterocycles. Instead, stacking was found to be driven primarily by the Lennard–Jones component of the force-field. In addition, gas-phase ab initio calculations indicated that the stability of stacked nucleic acids originates not in electrostatic interactions but in electron correlations (Sponer et al., 1996b; Alhambra et al., 1997; Hobza and Sponer, 1999).

These theoretical results appeared to conflict with the results of Newcomb and Gellman: if the partial atomic charges of adenyl groups opposed binding in water, then adenyl groups should stack less strongly than naphthyls, not more strongly as was observed experimentally. Newcomb and Gellman pointed out, however, that the theory used by Friedman and Honig had not been applied directly to the compounds studied experimentally, or indeed to any compounds containing naphthyl groups (Gellman et al., 1996).

The contrast between the conclusions reached by Newcomb and Gellman and the theoretical studies of Friedman and Honig is of concern, for it raises doubts about the validity of widely used methods for evaluating the physical basis of noncovalent association. It is therefore important to seek a resolution. The present paper addresses this issue with ab initio calculations, crystal structure analysis, geo-

---

Received for publication 16 June 2000 and in final form 25 September 2000.

Address reprint requests to Dr. Michael K. Gilson, Center for Advanced Research in Biotechnology, 9600 Gudelsky Dr., Rockville, MD 20850-3479. Tel.: 301-738-6217; Fax: 301-738-6255; E-mail: gilson@umbi.umd.edu.

Dr. Luo's present address is Department of Pharmaceutical Chemistry, University of California, San Francisco, 513 Parnassus Ave., San Francisco, CA 94143-0446.

© 2001 by the Biophysical Society

0006-3495/01/01/140/09 $2.00

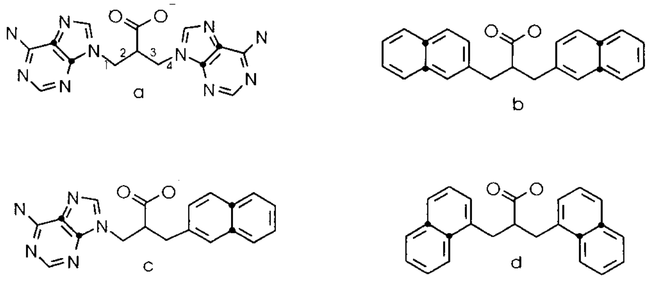

FIGURE 1 Diagrams of the (a) bis-adenyl (b) bis-naphthyl (c) mixed adenyl-naphthyl and (d) "sideways" bis-naphthyl compounds. Dots mark fusing atoms and numbers identify torsion angles ϕ_1–ϕ_4.

metric analysis, and molecular modeling of the specific compounds studied by Newcomb and Gellman with an energy model that is essentially the same as that used by Friedman and Honig.

METHODS

Ab initio calculations

Both Newcomb and Gellman (Gellman et al., 1996) and Friedman and Honig (Friedman and Honig, 1995) focus on the role of "interactions between partial charges" (Gellman et al., 1996), versus dispersion or Lennard–Jones interactions, in the stacking of adenyl and naphthyl groups. The partitioning of intermolecular forces into Coulombic and Lennard–Jones terms is typical of current empirical force fields. However, this partitioning is a simplification, and other contributions that could in principle play a role, such as electronic polarization and charge-transfer (Stone, 1996), can be accounted for only implicitly in the Coulombic and Lennard–Jones terms of force fields. Therefore, it is of interest to gain a more complete account of the energetics of aromatic stacking in the bis-adenyl and bis-naphthyl compounds via ab initio calculations.

Ab initio methods were applied as follows. Initial monomer geometries were generated with Quanta 98 (Molecular Simulations Inc., 1998) and subjected to Newton–Raphson (Press et al., 1989) energy-minimization in CHARMM (version 26) (Brooks et al., 1983) with the CHARMM 98 vacuum energy function. The minimizations were stopped when the energy gradient changed by $<1.0 \times 10^{-5}$ kcal/mol-Å per step. The geometries of the monomers were further optimized in Gaussian 98 (Frisch et al., 1998) at the MP2/6-31G** level and were not reoptimized for the complexes. Gaussian 98 (Frisch et al., 1998) was used to calculate gas-phase changes in potential energy for the association of adenines and naphthalenes, using the 6-31G*(0.25) basis set (Sponer et al., 1996b) at the MP2 level to improve the description of electron correlation interactions. The MP2 energies were corrected for basis set superposition error (Hobza and Sponer, 1999; Frisch et al., 1986). GAMESS (Schmidt et al., 1993) was then used to carry out reduced variation space (RVS) (Stevens and Fink, 1987) decompositions of the HF/6-31G*(0.25) stacking energies. Similar calculations were carried out for two adenines stacked in a B-DNA configuration and for adenine and naphthalene dimers in the most stable configurations found for the Newcomb and Gellman compounds via mining minima calculations (see below).

The electrostatic interaction energy is of particular interest in this study, because this term corresponds most closely to the interactions between partial charges discussed by Newcomb and Gellman. The RVS decomposition computes the electrostatic interaction between two molecules by obtaining their wavefunctions when they are widely separated, then moving them together without allowing their wavefunctions to change and computing their electrostatic interaction energy.

The accuracy of the present approach is supported by previous calculations showing that higher-level calculations of the electrostatic interactions between stacked nucleic acid bases agree with calculations at the MP2/6-31G*(0.25) level typically to within several tenths of a kcal/mol with occasional errors up to about 0.9 kcal/mol (Sponer et al., 1996a). A similar study using the same ab initio calculation and a continuum solvent model gave good agreement with thermodynamic measurements on base stacking (Florian et al., 1999). It has also been argued that the total stacking energies obtained from MP2/6-31G*(0.25) calculations are accurate since the modest underestimation of the dispersion attraction due to the limited size of the basis set tends to be compensated by the neglect of higher-order electron correlation terms (Hobza et al., 1997; Hobza and Sponer, 1999). Finally, comparisons between calculation and experiment for naphthalene trimers in the gas phase suggest that the MP2/6-31G level is adequate to deduce the equilibrium conformations of small aromatic clusters (Gonzalez and Lim, 1999). However, the MP2 method does not account fully for electron correlations. As a consequence, the decomposition of the non-electrostatic parts of the total interaction energies are best viewed as semiquantitative.

Crystal structure review

Examples of stacked adenyl and naphthyl groups were sought in the Cambridge Structure Database (Allen et al., 1979) to determine typical distances between the stacking planes. Systems in which bulky groups were interposed between the planes were rejected. Interplane distances were estimated as the shortest distance between any pair of atoms. The crystal structures of pure naphthalene show edge-face contacts, rather than planar stacking. However, some derivatives of naphthalene do provide examples of planar stacking of naphthyl groups, presumably because crystal packing is controlled by interactions among the substituents. Data were also collected for larger systems of 6-membered aromatic carbon rings, such as coronenes, which form parallel stacks.

Geometric analysis

Geometric analysis was used to determine whether the optimal separation of the aromatic rings could affect the stacking probabilities of the Newcomb and Gellman compounds. This approach isolates well-defined steric and geometric effects from the parameter- and method-dependencies of energy-based calculations. A conformation was considered geometrically accessible if it could be attained by rotations of the bonds of the 3-atom linker and did not have steric clashes. The four rotatable bonds of the linker, (ϕ_1, ϕ_2, ϕ_3, ϕ_4), were sampled systematically from 0 to 350° in 10° steps. (See Fig. 1 *a*.) Each of the resulting 36^4 different conformers was then checked for steric clashes between nonhydrogen atoms in the two aromatic groups. For simplicity, a single distance criterion of either 3.3 or 3.5 Å (see below) was used for all pairs of atoms. Sterically allowed conformations were then classified as stacked or unstacked, depending upon whether either of the two fusing atoms (see Fig. 1) of one aromatic ring was less than 5 Å away from either of the two fusing atoms of the other aromatic ring. Requiring that two pairs of fusing atoms, instead of one pair, be within 5 Å did not alter the conclusions of this qualitative analysis.

The overall stacking probability was computed as the ratio of the number of sterically accessible stacked conformations N_{stack} to the total number of sterically accessible conformations N_{acc}. The results were also visualized as contour plots (Fig. 2) showing the stacking probability $\mathrm{P}(\phi_2, \phi_3)$ as a function of the central two dihedral angles of the 3-atom linker:

$$\mathrm{P}(\phi_2, \phi_3) \equiv \frac{N_{\mathrm{stack}}(\phi_2, \phi_3)}{N_{\mathrm{acc}}}. \tag{1}$$

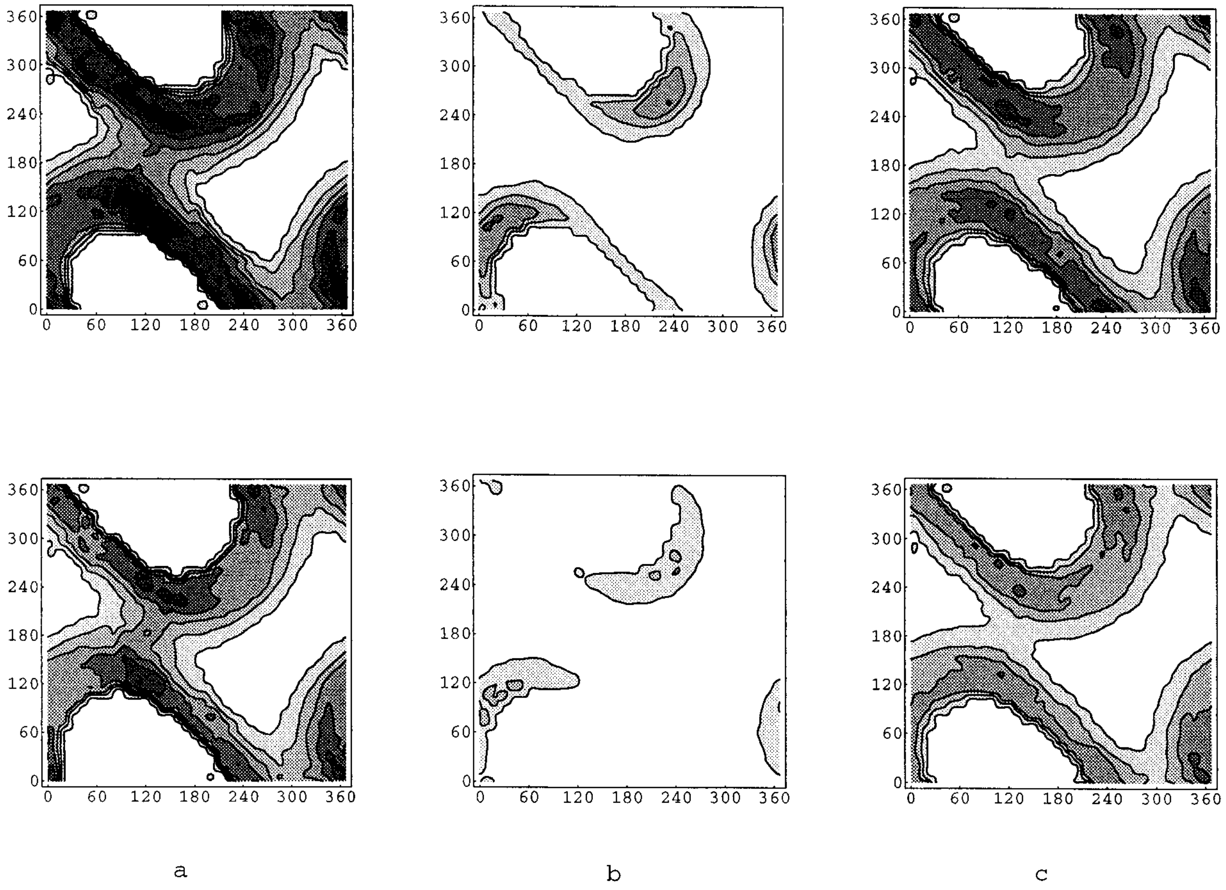

FIGURE 2 Contour plots of stacking probabilities as a function of the two central torsion angles ϕ_2 (abscissa) and ϕ_3 (ordinate) of the 3-atom linker (Fig. 1), for 3.3 Å (*top row*) or 3.5 Å (*bottom row*) clash criterion (see Methods). (a) Bis-adenyl; (b) bis-naphthyl; (c) "sideways" bis-naphthyl.

Here, $P(\phi_2, \phi_3)$ is the contribution of conformations with linker angles (ϕ_2, ϕ_3) to the overall stacking probability, and $N_{\text{stack}}(\phi_2, \phi_3)$ is the number of accessible, stacked conformations with linker angles (ϕ_2, ϕ_3). Dihedral angles ϕ_2, ϕ_3 are computed from the positions of the aliphatic carbons of the linker and the ring atoms to which they are bonded.

Force field calculations

The present study computes the stacking probabilities of the Newcomb and Gellman compounds with essentially the same energy model that Friedman and Honig used to analyze the energetics of adenine stacking (Friedman and Honig, 1995). Conformational distributions were computed with the recently developed "mining minima" algorithm (Head et al., 1997), which has been successfully applied to a range of problems (Luo et al., 1998; Luo et al., 1999a, b; David et al., 1999; Mardis et al., 1999, 2000; Luo and Gilson, 2000). This algorithm rapidly identifies each low-energy conformation i of the molecule and computes its stability via numerical evaluation of its configuration integral,

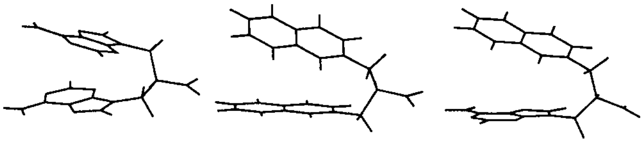

FIGURE 3 Global energy minima for bis-adenyl, bis-naphthyl, and mixed adenyl-naphthyl compounds, from mining minima calculations.

$$Z_i = \int_i \exp(-E(\mathbf{r})/kT)\, d\mathbf{r}. \tag{2}$$

Here $E(\mathbf{r})$ is the energy as a function of the conformation $\mathbf{r}$, kT is thermal energy, and the integral ranges only over conformations in well i. The probability of conformation i, P_i, is $P_i = Z_i/\Sigma_j Z_j$, where the sum ranges over all low-energy conformations. The probability of stacking is then the sum of P_i over all conformations meeting the criterion for stacking described in the Geometric Analysis section. The mining minima calculations sample over only torsion angles; bond lengths and bond angles are held fixed at initial geometries obtained by energy-minimization with CHARMM 26 (see above). Test calculations show that energy minimization from a different starting conformation changes the final results negligibly.

The energy function $E(\mathbf{r})$ used here is the sum of a potential energy and a solvation-free energy (Gilson et al., 1997). The potential energy is computed with the most recent CHARMM parameters available (Brooks et al., 1983; MacKerell et al., 1995, 1998; Foloppe and Mackerell, 2000; Duffy et al., 1993). The solvation-free energy is computed with a generalized Born (GB)/surface area model (Qiu et al., 1997; Still et al., 1990) where atomic self-energies are estimated with a charge-induced dipole term (Gilson and Honig, 1990). The surface tension is set to 6.4 cal/mol Å^2 (Friedman and Honig, 1995), and the molecular and solvent dielectric constants are set to 1 and 78, respectively. As described previously (Luo et al., 1998), the cavity radius of each atom is set to ½(σ + 1.4 Å), where σ is the CHARMM Lennard–Jones parameter appropriate to the atom and 1.4 is the radius of a water molecule.

Finite difference (FD) solutions of the Poisson equation (Warwicker and Watson, 1982; Klapper et al., 1986; Gilson et al., 1988) are believed to be more accurate than the faster generalized Born approximation used during the conformational sampling (Luo et al., 1999b). Therefore, the initial results of the mining minima calculation are corrected toward finite difference calculations as follows. For each energy minimum *i* identified with the mining minima algorithm, the electrostatic solvation free energy is calculated with both GB and FD, and the deviation of the GB from the FD result is subtracted from the conformational free energy of the minimum. This correction has been shown to improve the agreement of computation with experiment (Luo et al., 1999b). The surface area contribution to the free energy is included in a similar manner (Luo et al., 1999b).

RESULTS AND DISCUSSION

Ab initio and crystal structure analyses

Ab initio quantum mechanics was used to compare the contributions to the interactions between two adenines with those between two naphthalenes. Calculations were done for parallel-stacked, in-register, adenine and naphthalene dimers in vacuo, at ring separations of 3.4 to 4.0 Å. No attempt was made to find the global energy minima here because the calculations aim to compare the stacked dimers on an equal footing. Moreover, the global minima for these dimers are likely to be different from those for the Newcomb and Gellman compounds, where the relative disposition of the aromatic groups is constrained by the linker. However, additional ab initio calculations are reported below for the global energy minima found in the force field calculations.

Table 1 lists the total MP2 interaction energy (Total), the electrostatic energy (ES), and the nonelectrostatic energy (NonES) computed as Total minus ES. The nonelectrostatic energy consists of the electron correlation energy (Corr) computed as Total minus Hartree–Fock energy, along with the polarization (PL), exchange/repulsion (EX) and charge transfer (CT) energies from the RHF calculation.

As shown in Table 1, the total interaction energies of the adenines are less favorable than those for the naphthalenes, largely because the electrostatic repulsion between the adenines is greater than that between the naphthalenes. This is as expected, because the adenines are more polar and the dimers are stacked with like atoms on like. The nonelectrostatic parts of the interaction energies (NonES) shows energy minima of similar depth (−7.1 kcal/mol) for both dimers. This suggests that the nonelectrostatic attractions between stacked adenines and between stacked naphthalenes are of similar strength.

Table 1 also indicates that the optimum stacking distance for adenines is less than that for naphthalenes. This result is consistent with structural data in the Cambridge Structural Database (Allen et al., 1979). As summarized in Table 2, planar stacked adenyl groups tend to lie about 0.2 Å closer together than planar stacked naphthyls and other C_6 aromatic ring systems. That the distances in Table 2 are smaller than the optima inferred from the nonelectrostatic energies in Table 1 probably results primarily from our use of in-register dimers, which maximize the steric clashes between atomic orbitals; crystal structures typically show staggered conformations that permit shorter distances between the aromatic planes. Accordingly, previous ab initio structure optimizations have found inter-plane distances of 3.5 Å for staggered, stacked benzene (Hobza et al., 1994) and naphthalene (Gonzalez and Lim, 2000) dimers.

Geometric analysis

We conjectured that the greater stacking distance of naphthalene might reduce its ability to stack in the bis-naphthyl compound of Newcomb and Gellman, relative to the bis-adenyl compound. In particular, it seemed possible that the 3-atom linker might make it difficult for the naphthyls to form a parallel stack without some degree of steric clash. The geometric analysis described in Methods was used to compute the percentage of sterically accessible conformations that are stacked for the bis-adenyl and bis-naphthyl compounds, using steric distance criteria of 3.3 and 3.5 Å.

TABLE 1 Ab initio interactions energies and their components for parallel, in-register adenine and naphthalene dimers with no linker, as a function of separation distances (Dist).

		Energy (kcal/mol)						
Dimer	Dist (Å)	Total	ES	NonES	Corr	PL	EX	CT
Adenine	3.4	−3.23	3.15	−6.38	−12.89	−0.28	7.61	−0.82
	3.5	−3.71	3.35	−7.06	−9.80	−0.34	4.78	−1.70
	3.8	−3.59	2.85	−6.44	−6.66	−0.22	1.48	−1.04
	4.0	−3.01	2.46	−5.47	−5.30	−0.16	0.65	−0.66
Naphthalene	3.4	−3.76	0.45	−4.21	−16.35	−0.54	14.18	−1.50
	3.5	−4.80	1.10	−5.90	−14.02	−0.48	10.04	−1.44
	3.8	−5.38	1.71	−7.09	−9.17	−0.30	3.48	−1.10
	4.0	−4.81	1.67	−6.48	−7.07	−0.24	1.67	−0.84

Other symbols are defined in text.

TABLE 2 Inter-plane distances for stacked adenine and its derivatives and for naphthalene derivatives and other planar aromatic systems composed of 6-membered carbon rings

Adenyl		Naphthyl-related	
CSD Code	Distance (Å)	CSD Code	Distance (Å)
JOZZED (a)	3.283	COTZEQ (h)	3.680
LEZHIH (b)	3.200	DMBCRY (i)	3.459
MEADEN (c)	3.384	GICWUK (j)	3.399
MEADEN01 (d)	3.390	HADFEX (k)	3.524
SEFSOL (e)	3.414	HBZCOR01 (l)	3.451
VAVTOB (f)	3.256	KANMUH (m)	3.468
VIMREO (g)	3.389	PINTEL (n)	3.555
Mean	**3.33**	**Mean**	**3.51**

a) Adenine hydrogen peroxide
b) bis(Adenine) 7-hydroxy-8-(phenylazo)-1,3-naphthalene-disulfonate pentahydrate
c) 9-Methyladenine
d) 9-Methyladenine
e) 8-Ethyl-9-methyladenine dihydrate
f) 9-Vinyladenine
g) 8-Propyl-9-methyladenine hydrate
h) 1-Fluoro-9-(2-methyl-1-naphthyl)-fluorene
i) 6,12-Dimethyl-dibenzo(def,mno)chrysene,6,12-Dimethylanthanthrene
j) 7,14-Di-*n*-propylacenaphtho(1,2-k)fluoranthene
k) 2-(1-Naphthyl)-5-phenyl-1,3-oxazole
l) Hex-abenzo(bc,ef,hi,kl,no,qr)coronene
m) (Z)-9-(2-Nitro-2-phenylethenyl)anthracene
n) 1,1′-bis(Naphthyl)-di-selenium naphthalene.

Comparison of the two columns of results in Table 3 shows that increasing the clash radii of the aromatic groups decreases the stacking probabilities. This is purely a geometric effect, since these calculations do not use any energy functions other than those implicit in the steric interactions and the chemical bonds. Thus, the observation that the bis-adenyl compounds stack more effectively than the bis-naphthyls does not necessarily imply a greater intrinsic attraction between adenyls versus naphthyls. This observation offers at least a partial resolution of the apparent contradiction between experiment and theory that motivates the present study.

Furthermore, comparison of the first two rows of Table 3 shows that geometry dictates a much higher stacking probability for the bis-adenyl than the bis-naphthyl compound, even when the same steric criterion is used for both. This difference is studied via contour plots of stacking probabilities as a function of the two central dihedral angles of the 3-atom linker (Eq. 1). Figure 2 presents these graphs for both the 3.3 Å (*top row*) and 3.5 Å (*bottom row*) steric criteria. It is apparent that the range of dihedral angles for which the bis-adenyl compounds can stack is larger than that for the bis-naphthyl compounds. Moreover, the fraction of stacked conformations for a given set of linker dihedrals is consistently greater for the bis-adenyl compounds than for the bis-naphthyl compounds. These differences are traceable to the "sideways" orientation of the adenyls on the linker, which allows stacking to occur more readily and over a larger range of linker conformations. Thus, when the naphthyl groups are turned sideways by moving the linker from the 2- to the 1- carbon (Fig. 1 *d*), the percentage of stacked conformations rises (Table 3, *rows 2, 4*) and the contour plot now resembles that of the bis-adenyl compound (Fig. 2, *a* and *c*). The sideways orientation allows the rings to adopt stacked conformations even when the linker bonds numbered 1 and 4 in Fig. 1 *a* are splayed apart.

TABLE 3 Percentage of sterically accessible conformations that are stacked, for two different steric clash criteria

	Clash distance (Å)	
Compound	3.3	3.5
Bis-adenyl*	9.7	6.7
Bis-naphthyl*	1.0	0.3
Mixed*	3.2	1.6
Bis-naphthyl-1†	6.7	4.6

*Newcomb and Gellman, 1994
†A hypothetical compound linked at the 1- instead of the 2-carbon (Figure 1d).

The fact that the bis-adenyl compound has many more accessible stacked conformations than the bis-naphthyl compound that was studied experimentally raises serious concerns regarding the original interpretation of the Newcomb and Gellman study: even if the energies of interaction between adenyl and naphthyl groups were identical, the geometric differences between the bis-adenyl and bis-naphthyl compounds studied would cause the former to stack more than the latter. This analysis thus offers a second possible resolution of the apparent conflict between theory and experiment that was laid out in the Introduction.

The number of geometrically accessible conformations for the mixed adenyl–naphthyl compound is between that for the bis-adenyl and bis-naphthyl compounds. Experimentally, however, the mixed compound shows stronger association than the bis-adenyl compound. This indicates that steric interactions and geometry can provide only part of the explanation for the experimental results. The analysis given below indicates that specific patterns of electrostatic interactions between the aromatic groups also influence the stacking tendencies of these compounds.

Interestingly, prior experiments have already established that changing the points of attachment of a trimethylene linker to two adenyl groups changes the degree to which the adenyl groups stack (Leonard and Ito, 1973). The degree of stacking in such compounds is ultimately determined not only by the purely geometric considerations just discussed, but also by the detailed form of the interactions between the two aromatic groups; for example, by the interplay of their electrical multipoles. The message is nonetheless clear that the measured degree of stacking of aromatic groups joined by a restraining linker does not provide a direct readout of the intrinsic stacking tendency of the groups.

Conformational analysis

This section considers whether the energy model of Friedman and Honig (1995) can reproduce the data of Newcomb and Gellman (1994). The mining minima method was used with a continuum solvent model, as described in Methods, to compute the conformational distributions of the bis-adenyl, bis-naphthyl, and mixed adenyl-naphthyl compounds studied by Newcomb and Gellman (1994). The calculated stacking probabilities are 55, 17, and 85%, respectively. These results agree with the experimental data, which indicate that the bis-naphthyl compound stacks least and that the mixed compound stacks most (Newcomb and Gellman, 1994). This is despite the fact that the energy model is essentially that of Friedman and Honig, which was thought to be inconsistent with the experimental results. Surprisingly, then, the present calculations show no contradiction between the experimental and theoretical studies outlined in the Introduction.

This finding probably is explained in part by the geometric considerations presented above: the combination of the sideways orientation of the adenyl groups on their linker with the tendency of naphthyls to stack further apart leads to a higher fraction of sterically allowed stacked states for the bis-adenyl versus the bis-naphthyl compound.

Interestingly, mining minima calculations for a sideways bis-naphthyl compound in which the 3-atom linker attaches to carbon 1 of the naphthyl group (see above) predict 99% association between the naphthyl groups. This experimentally testable result is consistent with a recent experimental and computational study of the corresponding bis-indolyl compound, which is similarly nonpolar and is also linked in a sideways configuration (Pang et al., 1999). The increased association in the sideways bis-naphthyl compound appears to be attributable, at least in part, to the fact that the naphthyl groups in this compound are free to adopt edge–face conformations that are not sterically accessible for the original bis-naphthyl compound (Newcomb and Gellman, 1994). Edge–face naphthyl–naphthyl interactions are believed to be particularly stable, based upon crystallographic data and ab initio calculations (Abrahams et al., 1949; Brock and Dunitz, 1982; Gonzalez and Lim, 2000).

Further mining-minima calculations for a bis-naphthyl compound with a 4-atom linker that was studied experimentally (Compound **1** series a) (Newcomb et al., 1995) show only 20% association of the naphthyl groups, even though edge–face conformations are sterically accessible for this compound. The absence of clear-cut association in this case is consistent with the experimental interpretation (Newcomb et al., 1995), and presumably results at least partly from the increased entropy cost of association for this more flexible compound. That NMR shifts were detectable for this compound may result from the increased magnitude of ring-current shifts in edge–face versus face–face conformations (Ando and Webb, 1983).

As noted above, geometric considerations alone do not fully explain why the mixed adenyl-naphthyl compound appears to stack even more than the bis-adenyl compound. The following subsection provides further analysis of the forces driving stacking in this series of compounds.

Energy component analysis

The forces driving stacking in the mining minima calculations were analyzed by calculating changes in the Boltzmann-weighted averages of the nonelectrostatic and electrostatic energy terms for stacked versus extended conformations of the bis-adenyl and bis-naphthyl compounds, where the Boltzmann weights are obtained from the mining minima results. The electrostatic terms comprise Coulombic interactions and the electrostatic part of the solvation energy, whereas the nonelectrostatic terms comprise van der Waals interactions and the surface area part of the solvation-free energy.

The results, presented in Table 4, are consistent with those of Friedman and Honig, for they indicate that electrostatic interactions in net oppose the association of adenyl groups. On the other hand, the previous results showed this electrostatic cost to result primarily from desolvation, whereas here the cost is Coulombic in origin. This difference could result from the presence of the anionic linker in the compounds studied here and also from differences between the conformations examined.

Interestingly, the electrostatic cost of stacking the mixed compound is low relative to the bis-adenyl compound. This helps account for this compound's high degree of stacking. The small magnitude of the electrostatic cost of stacking for this compound is primarily attributable to the small magnitude of the Coulombic repulsion, which in turn is a function of the specific charge distributions of the stacked groups and of the constraints imposed by the 3-atom linker. It is worth noting that other hydrocarbon/heterocyclic aromatic stacks might not have such favorable electrostatic complementarity and might stack poorly as a consequence. Thus,

TABLE 4 Boltzmann-averaged energy contributions for stacked versus extended conformations of the three compounds studied experimentally

	Energy (kcal/mol)				
Compound	Total	NonES	ES+ESolv	ES	ESolv
Bis-adenyl	−2.94	−5.98	3.04	4.28	−1.24
Bis-naphthyl	0.19	−1.60	1.79	0.17	1.63
Mixed	−3.20	−4.78	1.58	1.05	0.53

Total, sum of all energy contributions; ES, Coulombic part of CHARMM force field, most closely comparable with ES term from RVS decomposition; ESolv, electrostatic part of solvation energy; ES+ESolv, net electrostatic energy, defined here as sum of Coulombic and electrostatic solvation terms, most closely comparable with electrostatic energy in Friedman and Honig (1995); NonES, nonelectrostatic energy computed as total energy minus the net electrostatic energy (Total − (ES + ESolv)).

although Guckian et al. (1996) report enhanced stacking of nonpolar base analogs with nucleic acid bases, this result may not be generalizable.

Table 4 shows that the average nonelectrostatic attraction between adenyl groups is stronger than that between naphthyl groups. This difference in the nonelectrostatic attraction of adenyls versus naphthyls may result in part from the fact that the bis-adenyls more readily adopt stacked conformations, due to the steric constraints discussed above. In addition, the CHARMM force-field ascribes a somewhat deeper Lennard–Jones energy well to stacked adenines than to stacked naphthalenes. The surface area solvation term contributes negligibly to the difference in stacking, because its coefficient is the same for all atom types and adenine and naphthalene are of nearly the same size.

That CHARMM attributes stronger van der Waals attractions to adenyls than to naphthyls appears inconsistent with the ab initio calculations above, which show little difference between the nonelectrostatic interaction energies of the two compounds. It is difficult to judge which result is most plausible. On one hand, the ab initio results are more directly linked to fundamental principles. On the other hand, CHARMM is parameterized to reproduce crystal sublimation energies, while ab initio calculations are not adapted to fit experimental data. In addition, there is no one-to-one correspondence between force field terms and quantum mechanics energy components: as discussed in Methods, force fields typically lump all intermolecular forces into the Coulombic and Lennard–Jones terms, whereas the RVS decomposition, for example, provides a more detailed breakdown. Given these uncertainties, it is of interest to gain an independent account of the energetics of aromatic stacking in the bis-adenyl and bis-naphthyl compounds by carrying out ab initio calculations for these compounds.

Accordingly, changes in energy were computed for the in vacuo assembly of adenines and naphthalenes from infinite separation into the global energy minima found in the mining minima calculations (Fig. 3), but with the linkers removed and replaced by hydrogens for computational simplicity. An additional calculations was done for the assembly of two adenines into a B-DNA conformation obtained from Quanta 98 (Molecular Simulations Inc., 1998).

TABLE 5 Ab initio energy analysis (MP2/6-31G*(0.25)) for the association of adenine and naphthalene dimers with no linker

		Energy (kcal/mol)	
Dimer	Conformation	ES	NonES
Adenine	Bis-adenyl min	2.37	−5.11
Naphthalene	Bis-naphthyl min	0.22	−2.94
Adenine	B-DNA	−0.12	−6.60

Bis-adenyl and bis-naphthyl min, global energy minima found with mining-minima calculations; B-DNA, adenine dimer in conformation obtained from QUANTA; ES, RVS electrostatic energy; NonES, total energy minus electrostatic component.

The results, presented in Table 5, confirm the dominance of nonelectrostatic interactions over electrostatic interactions for stacking of both adenine and naphthalene in these chemically and biologically relevant conformations. The qualitative agreement obtained between force field and ab initio calculations supports the validity of this conclusion. Although the ab initio results are for the gas phase, nonelectrostatic interactions are expected to dominate stacking in water also, because water imposes an electrostatic desolvation penalty and adds an attractive hydrophobic component to the nonelectrostatic term. Indeed, NMR analysis has revealed increasing association of indole groups in water with increasing temperature and decreased association on transfer from water to dimethylsulfoxide; this pattern suggests a solvent-mediated hydrophobic contribution to association in water (Pang et al., 1999). Thus, the present calculations support the view (Friedman and Honig, 1995; Sponer et al., 1996a, b; Alhambra et al., 1997; Hobza and Sponer, 1999) that the stacking of adenines, and of aromatic groups in general (Guckian et al., 2000) is driven primarily by nonelectrostatic interactions.

CONCLUSIONS

The present study shows that the experimental results of Newcomb and Gellman for linked adenyl and naphthyl groups are consistent with the theoretical evidence that nonelectrostatic forces primarily drive stacking. We have identified two possible reasons, other than stronger electrostatic interactions, why the bis-adenyl compounds should stack more than the bis-naphthyl compounds in the Newcomb and Gellman's experiment.

First, naphthyl groups tend to stack at a wider separation than adenyl groups, based upon both quantum calculations and a review of crystal structures; and a geometric analysis shows that the 3-atom linker in the compounds studied by Newcomb and Gellman allows more stacked conformations when the aromatic rings can approach each other more closely. On this basis alone, one would expect the bis-adenyl compounds to stack more than the bis-naphthyl compounds.

Second, geometric analysis shows that more stacked conformations are sterically accessible to the bis-adenyl compound than to the bis-naphthyl compound because the linker is attached to the sides of the adenyl groups, but is attached to the ends of the naphthyl groups. The significance of the attachment points of the linker in such molecules is confirmed by a prior experimental study (Leonard and Ito, 1973). As a consequence, the experimental data of Newcomb and Gellman cannot be used directly to compare the intrinsic forces acting between adenyl and naphthyl groups.

It is encouraging that the mining minima calculations give results consistent with experiment. This success is of particular interest because the present calculations rely upon an energy model that was thought to be inconsistent with the data (Friedman and Honig, 1995; Newcomb and Gellman, 1994). The correct ranking of stacking tendencies for the bis-adenyl, bis-naphthyl, and mixed compounds in these calculations presumably results in part from the geometric constraints imposed by the linker, as just discussed. However, the interactions between the specific charge distributions of the aromatic groups also plays a role, as illustrated by the results for the mixed compound. In addition, the CHARMM force field ascribes a greater Lennard–Jones well-depth to stacked adenines than to stacked naphthalenes. Higher level quantum studies and further tests of the CHARMM parameters against measured quantities, such as the heat of sublimation of solid naphthalene, could be useful in assessing the accuracy of the force field.

Finally, the present study highlights the fact that subtle differences between molecules can produce unexpected differences in their conformational preferences. This makes it difficult to use chemical intuition alone to interpret experimental studies, even for relatively small molecules. A careful computational analysis can provide further insight and can thus serve as a valuable adjunct to chemical intuition.

We thank Drs. C. Gonzalez, B. Honig, J. Jensen, V. Kairys, M. Krauss, T. S. Lee, and A. Mackerell for helpful discussions, and the referees for thoughtful comments and suggestions.

This work was supported by the National Institutes of Health (GM54043 to MKG) and by the National Institute of Standards and Technology.

REFERENCES

Abrahams, S. C., J. Monteath Robertson, and J. B. White. 1949. The crystal and molecular structure of naphthalene. II. Structure investigation by the triple Fourier series method. *Acta. Cryst.* 2: 238–244.

Alhambra, C., F. J. Luque, F. Gago, and M. Orozco. 1997. *Ab initio* study of stacking interactions in A- and B-DNA. *J. Phys. Chem. B.* 101: 3846–3853.

Allen, F. H., S. Bellard, M. D. Brice, B. A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink, B. G. Hummelink-Peters, O. Kennard, W. D. S. Motherwell, J. R. Rodgers, and D. G. Watson. 1979. The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis, and display of information. *Acta Crystallogr. B.* 35:2331–2339.

Ando, I., and G. A. Webb. 1983. Theory of NMR Parameters, Academic Press, NY.

Brock, C. P., and J. D. Dunitz. 1982. Temperature-dependence of thermal motion in crystalline naphthalene. *Acta. Cryst. B.* 38: 2218–2228.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* 4: 187–217.

David, L., R. Luo, M. S. Head, and M. K. Gilson. 1999. Computational study of kni-272, a potent inhibitor of HIV-1 protease: on the mechanism of preorganization. *J. Phys. Chem.* 103: 1031–1044.

Duffy, E. M., P. J. Kowalczyk, and W. L. Jorgensen. 1993. Do denaturants interact with aromatic hydrocarbons in water? *J. Am. Chem. Soc.* 115: 9271–9275.

Florian, J., J. Sponer, and A. Warshel. 1999. Thermodynamic parameters for stacking and hydrogen bonding of nucleic acid bases in aqueous solution: Ab Initio/Lanvgein Dipoles study. *J. Phys. Chem. B.* 103: 884–892.

Foloppe, N., A. D. Mackerell, Jr. 2000. All-atom empirical force field for nucleic acids. I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comp. Chem.* 21: 86–104.

Friedman, R. A., and B. Honig. 1995. A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.* 69: 1528–1535.

Frisch, M. J., J. E. Del Bene, J. S. Binkley, and H. F. Schaefer. 1986. Extensive theoretical-studies of the hydrogen-bonded complexes (H2O)2, (H2O)2H+, (HF)2, (HF)2H+, F2H-, and (NH3)2. *J. Chem. Phys.* 84: 2279–2289.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G., Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople. 1998. Gaussian 98, Revision A.6, Gaussian, Inc., Pittsburgh, PA.

Gellman, S. H., T. S. Haque, and L. F. Newcomb. 1996. New evidence that hydrophobic effect and dispersion are not major driving forces for nucleotide base stacking. *Biophys. J.* 71: 3523–3525.

Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon. 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72: 1047–1069.

Gilson, M. K., and B. Honig. 1990. The inclusion of electrostatic hydration energies in molecular mechanics calculations. *J. Comp.-Aided. Molec. Design.* 5: 5–20.

Gilson, M. K., K. A. Sharp, and B. H. Honig. 1988. Calculating the electrostatic potential of molecules in solution: method and error assessment. *J. Comp. Chem.* 9: 327–335.

Gonzalez, C., and E. C. Lim. 1999. *Ab initio* and MM3 studies of the conformational structures of naphthalene trimer: comparison with experiment. *J. Phys. Chem. A.* 103:1437–1441.

Gonzalez, C., and E. C. Lim. 2000. Quantum chemistry study of the van der Waals dimers of benzene, naphthalene, and anthracene: crossed (D2d) and parallel-displaced (C2h) dimers of very similar energies in the linear polyacenes. *J. Phys. Chem. A.* 104:2953–2957.

Guckian, K. M., B. A. Schweitzer, R. X. Ren, C. Sheils, P. L. Paris, D. C. Tahmassebi, and E. T. Kool. 1996. Experimental measurement of aromatic stacking affinities in the context of duplex DNA. *J. Am. Chem. Soc.* 118: 8182–8183.

Guckian, K. M., B. A. Schweitzer, R. X. Ren, C. J. Sheils, D. C. Tahmassebi, and E. T. Kool. 2000. Factors contributing to aromatic stacking in water: evaluation in the context of DNA. *J. Am. Chem. Soc.* 122: 2213–2222.

Head, M. S., J. A. Given, and M. K. Gilson. 1997. "Mining minima": direct computation of conformational free energy. *J. Phys. Chem.* 101: 1609–1618.

Hobza, P., M. Kabelac, J. Sponer, P. Mejzlik, and J. Vondrasek. 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 nonempirical beyond Hartree–Fock results. *J. Comp. Chem.* 18: 1136–1150.

Hobza, P., H. L. Selzle, and E. W. Schlag. 1994. Structure and properties of benzene-containing molecular clusters: nonempirical *ab initio* calculations and experiments. *Chem. Rev.* 94: 1767–1785.

Hobza, P., and J. Sponer. 1999. Structure, energetics and dynamics of the nucleic acid base pairs: Nonempirical *ab initio* calculations. *Chem. Rev.* 99: 3247–3276.

Klapper, I., R. Hagstrom, R. Fine, K. Sharp, and B. Honig. 1986. Focusing of electric fields in the active site of Cu,Zn Superoxide Dismutase. *Proteins: Struct. Func. Gen.* 1: 47–79.

Kool, E. T. 1997. Preorganization of DNA: design principles for improving nucleic acid recognition by synthetic oligonucleotides. *Chem. Rev.* 97: 1473–1487.

Leonard, N. J., and K. Ito. 1973. Controlled interaction between nucleic acid bases. Intramolecular stacking interactions between two adenine rings. *J. Am. Chem. Soc.* 95: 4010–4016.

Luo, R., L. David, H. Hung, J. Devaney, and M. K. Gilson. 1999a. Strength of solvent exposed salt-bridges. *J. Phys. Chem.* 103: 727–736.

Luo, R., and M. K. Gilson. 2000. Synthetic adenine receptors: Direct calculation of binding affinity. *J. Am. Chem. Soc.* 122: 2934–2937.

Luo, R., M. S. Head, J. A. Given, and M. K. Gilson. 1999b. Nucleic acid base-pairing and *N*-methylacetamide self-association in chloroform: affinity and conformation. *Biophys. Chem.* 78: 183–193.

Luo, R., M. S. Head, J. Moult, and M. K. Gilson. 1998. pKa shifts in small molecules and HIV protease: electrostatics and conformation. *J. Am. Chem. Soc.* 120: 6138–6146.

MacKerell, A., D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102: 3586–3616.

MacKerell, A. D. Jr., J. Wiókiewicz-Kuczera, and M. Karplus. 1995. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* 117: 11946–11975.

Mardis, K. L., B. J. Brune, P. Vishwanath, B. Giorgis, G. F. Payne, and M. K. Gilson. 2000. Intramolecular versus intermolecular hydrogen bonding in the adsorption of aromatic alcohols onto an acrylic ester sorbent. *J. Phys. Chem. B.* 104:4735–4744.

Mardis, K. L., A. J. Glemza, B. J. Brune, G. F. Payne, and M. K. Gilson. 1999. The differential adsorption of phenol derivatives onto a polymeric sorbent: a combined molecular modeling and experimental study. *J. Phys. Chem. B.* 103: 9879–9887.

Molecular Simulations Inc. 1998. Quanta 98, San Diego, CA.

Newcomb, L. F., and S. H. Gellman. 1994. Aromatic stacking interactions in aqueous solution: evidence that neither classical hydrophobic effects nor dispersion forces are important. *J. Am. Chem. Soc.* 116: 4993–4994.

Newcomb, L. F., T. S. Haque, and S. H. Gellman. 1995. Searching for minimum increments of hydrophobic collapse: flexible dinaphthyl carboxylates. *J. Am. Chem. Soc.* 117: 6509–6519.

Pang, Y., J. L. Miller, and P. A. Kollman. 1999. Computational and experimental studies of (2,2)-bis(indol1-yl-methyl)acetate suggest the importance of the hydrophobic effect in aromatic stacking interactions. *J. Am. Chem. Soc.* 121: 1717–1725.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1989. Numerical Recipes. The Art of Scientific Computing. Cambridge University Press, Cambridge, U.K. 254–259.

Qiu, D., P. S. Shenkin, F. P. Hollinger, and W. C. Still. 1997. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem.* 101: 3005–3014.

Saenger, W. 1984. Principles of Nucleic Acid Structure, Springer-Verlag, New York.

Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. 1993. General atomic and molecular electronic-structure system. *J. Comp. Chem.* 14: 1347–1363.

Sponer, J., J. Leszczynski, and P. Hobza. 1996a. Hydrogen bonding and stacking of DNA bases: a review of quantum-chemical *ab initio* studies. *J. Biomolec. Struct. Dynam.* 14: 117–135.

Sponer, J., J. Leszczynski, and P. Hobza. 1996b. 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. *J. Phys. Chem.* 100: 5590–5596.

Stevens, W. J., and W. H. Fink. 1987. Frozen fragment reduced variational space analysis of hydrogen-bonding interaction: application to the water dimer. *Chem. Phys. Lett.* 139: 15–22.

Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrickson. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112: 6127–6129.

Stone, A. J. 1996. The Theory of Intermolecular Forces, The International Series of Monographs on Chemistry. Clarendon Press, Oxford, UK. 5.

Warwicker, J., and H. C. Watson. 1982. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* 157: 671–679.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.