Physical Nature of Ethidium and Proflavine Interactions with Nucleic Acid Bases in the Intercalation Plane

Karol M. Langner,[†] Pawel Kedzierski,[†] W. Andrzej Sokalski,*[,†] and Jerzy Leszczynski[‡]

Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland, and Jackson State University, Jackson, Mississippi 39217

Received: November 24, 2005; In Final Form: March 22, 2006

On the basis of the crystallographic structures of three nucleic acid intercalation complexes involving ethidium and proflavine, we have analyzed the interaction energies between intercalator chromophores and their four nearest bases, using a hybrid variation−perturbation method at the second-order Møller−Plesset theory level (MP2) with a 6-31G(d,p) basis set. A total MP2 interaction energy minimum precisely reproduces the crystallographic position of the ethidium chromophore in the intercalation plane between UA/AU bases. The electrostatic component constitutes the same fraction of the total energy for all three studied structures. The multipole electrostatic interaction energy, calculated from cumulative atomic multipole moments (CAMMs), was found to converge only after including components above the fifth order. CAMM interaction surfaces, calculated on grids in the intercalation planes of these structures, reasonably reproduce the alignment of intercalators in crystal structures; they exhibit additional minima in the direction of the DNA grooves, however, which also need to be examined at higher theory levels if no crystallographic data are given.

Introduction

Interactions between small ligands and DNA or RNA are relevant for many biological mechanisms involving nucleic acids, such as translation and transcription—their interaction may lead to minor and major groove binding, intercalation, and other binding modes. Intercalative binding is achieved by the insertion of an aromatic ligand or its fragment between adjacent base pairs along the DNA/RNA strand,[1] inducing significant structural perturbations to its helical structure and disturbing its biological function. Intercalation was first predicted and characterized by Lerman[2] and others.[3,4] Chemically, intercalation in nucleic acids can be considered to be a specific case of aromatic stacking interaction.[5]

Among the first known intercalators were ethidium and proflavine (see Figure 1 for their structures). Many intercalating molecules, such as anthracycline and acridine derivatives, are used in chemotherapy.[6] Understanding their interactions with nucleic acids is valuable for many fields—most notably for medicinal chemistry, where it may aid the rational design of novel drugs[7] and make it possible to govern their behavior, for instance, by triggering intercalative capabilities.[8] The multitude of possible derivatives and their specificity for certain nucleic acid sequences makes intercalating agents especially useful as nucleic acid dyes.[9]

Alongside noncovalent interactions within the intercalation site, there are a number of significant energetic contributions to the intercalation process,[1,10] which makes assessing the consequences of structural modifications to intercalators a difficult task. While it has been known that stacking interactions are a driving force in intercalation, only lately Kubar et al. have performed free energy calculations that compare well with experimental values and confirm the crucial role of noncovalent interactions.[11] Also, most intercalating molecules used in practice have side chains, and their mechanism of action additionally involves binding to one or both of the DNA grooves and to molecules other than nucleic acids.

There have been relatively few, compared to experimental research, exhaustive theoretical studies on the nature of intercalation into DNA, its mechanism, and dependence on ligand structure.[11−13] The high level of computational techniques necessary to gain reliable information on electron correlation effects in stacked nucleic acid bases and intercalation complexes[13−15] makes the rigorous description of stacking and intercalation mechanisms one of the most demanding problems concerning nucleic acids. This situation is analogous to the challenges encountered when numerically studying various stacked aromatic systems, namely, intercalators stacked with single bases or base pairs,[15−17] stacked nucleic acids,[18−20] and other systems.[21]

Hill et al.[19] have pointed out that the electrostatic term may be indicative of relative stacking interaction energies. Medhi et al.[12] and Bondarev et al.[16] have shown that electrostatic interactions in the binding of positively charged chromophores with nucleic bases are significant, other researchers have reported similar conclusions for various stacked aromatic systems,[22−24] and some also have emphasized the role of nonelectrostatic interactions.[25] Furthermore, recent experimental investigations into quantitative structure−free energy relationships indicate the electrostatic control of aromatic stacking interactions.[26] If electrostatic interaction energy has a predictive value for intercalator−intercalant geometry of at least qualitative accuracy, then it can indeed be used to reasonably model the intercalation mechanism.

The aim of this study is to analyze in detail the interaction energy between intercalator and intercalant, based on crystallographic structure fragments of complexes involving cationic intercalators: ethidium$^{(+1)}$−UA/AU,[27] ethidium$^{(+1)}$−CG/GC,[28]

* To whom correspondence should be addressed. E-mail: sokalski@mml.ch.pwr.wroc.pl.
[†] Wroclaw University of Technology.
[‡] Jackson State University.

10.1021/jp056836b CCC: $33.50
Published on Web 04/22/2006

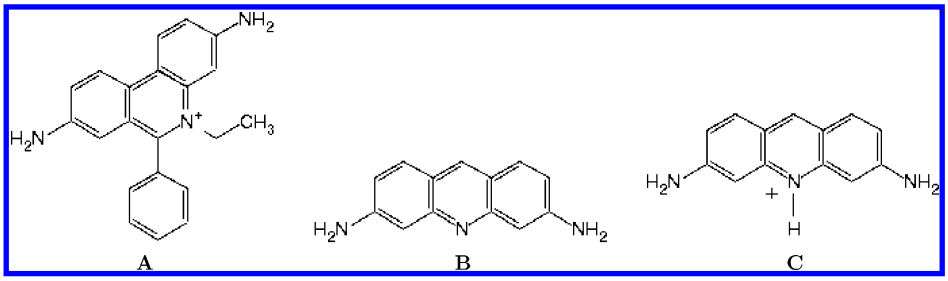

Figure 1. Structures of the intercalators studied: ethidium bromide (A), neutral proflavine (B), and cationic proflavine (C).

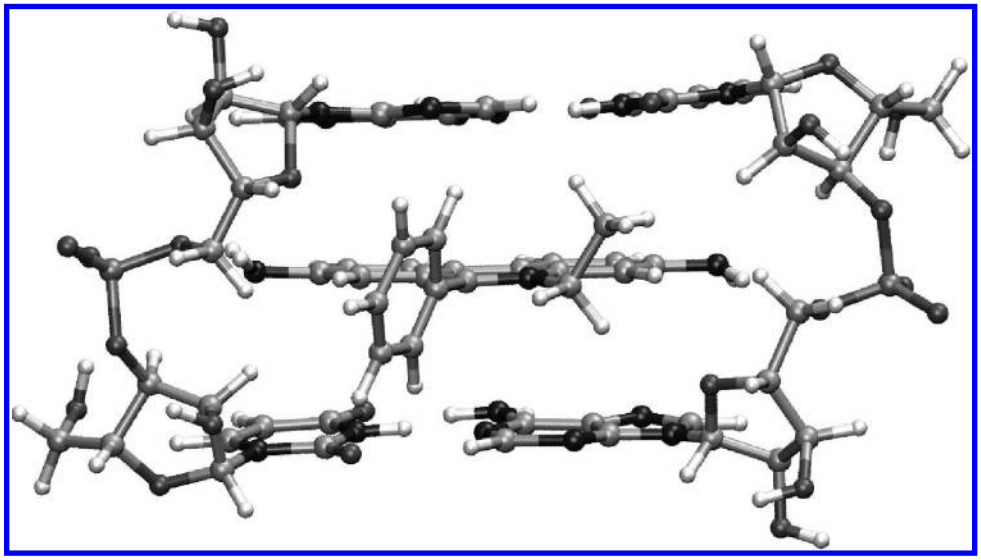

Figure 2. Intercalation site of the ethidium$^{(+1)}$−UA/AU complex as taken from the crystal structure of Jain and Sobell.[27]

and proflavine$^{(+1)}$−CG/AU.[29] In particular, we address the nature of the interactions in these systems by comparing the different energy components. The structural formulas of ethidium and proflavine are provided in Figure 1. Discussion is focused on the Eth$^{(+1)}$−UA/AU complex, for which more extensive calculations were performed; a molecular representation of this intercalation site is shown in Figure 2.

Due to the large sizes of the systems involved, they were partitioned into smaller segments. The nucleic acid backbone was separated from the nearest two base pairs. Ethidium was also divided, into its chromophore and side chain parts (see Figure 3); this study deals mostly with the ethidium chromophore. The proflavine molecule, having no side chain, was left unchanged. Such an approach, which recognizes a certain independence of the roles of the chromophore and side chain, is akin to that used by Medhi et al. in a related study.[12] A practical question can be addressed by this approach: whether the alignment of the chromophore between base pairs is determined by local interactions, especially electrostatic, or rather by other factors, such as side chain interactions.

Interaction energies were evaluated in this study pairwise, as the separate two-body interactions between the chromophores and each of their four nearest bases. These energies were summed to yield a "total" interaction between a chromophore and its four nearest bases. In doing so, we neglect many-body interactions, especially the polarization of hydrogen bonded base pairs. Reha et al.[13] emphasize that the electrostatic and one-electron properties of base pairs differ substantially from those of single bases and that the smallest complete model for the intercalation process is represented by the intercalator−base pair cluster. They also present evidence, based on an ethidium−AT/TA complex, that extending it to a model comprising the intercalator and two base pairs does not bring significant

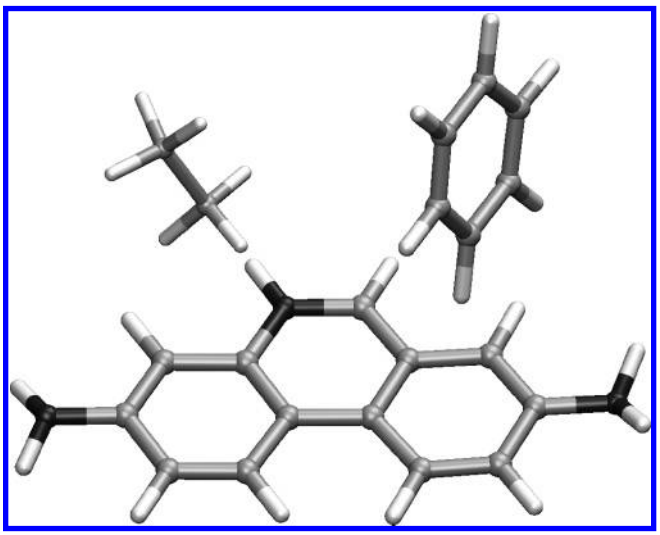

Figure 3. Fragmentation of the ethidium intercalator shown in Figure 2 into its chromophore, side chain, and side ring.

improvement in terms of the interaction energy. While we recognize the importance of treating base pairs as entities for calculating their stabilization energies in stacked complexes, in this study, we explore the suitability of a noninteracting base model for evaluating stabilization energies in intercalation complexes and in reproducing chromophore alignment in the intercalation plane. As reference and support for this approach, we also present in one case the results of calculations for a complete base pair model.

The interaction energies between chromophores and their nearest bases were subject to decomposition analysis, providing insight into the physical nature of the chosen experimental intercalation structures at and in the immediate vicinity of the

intercalation site. Also, cumulative atomic multipole moments (CAMMs) up to rank nine (512-pole moments) were determined for all molecular fragments and multipolar interaction energies were estimated on their basis. The predictive value of multipole electrostatics for chromophore alignment was demonstrated by calculating the interactions between CAMM expansions, representing frozen monomers in the studied complexes, on a grid in the plane of intercalation. This was further compared with other interaction components along a chosen path toward the major groove in the $Eth^{(+1)}$−UA/AU complex.

Methods

Interaction Energy Components. The interaction energy of each base−chromophore cluster was analyzed using a hybrid variation−perturbation scheme.[30,31] In this procedure, the MP2 level interaction energy, $\Delta E^{(2)}_{MP2}$, is decomposed into first-order components, namely, electrostatic $\Delta E^{(1)}_{el}$ and Heitler−London exchange $\Delta E^{(1)}_{ex}$, and higher-order delocalization $\Delta E^{(R)}_{del}$ and correlation $\Delta E^{(2)}_{corr}$ terms. Basis set dependence is considerably reduced by calculating these energies consistently in a dimer basis set (D).[31] The first-order electrostatic interaction energy is further decomposed into components attributed to a multipole expansion $\Delta E^{(1)}_{el,mtp}$ and penetration effects $\Delta E^{(1)}_{el,pen}$ (the difference between the total electrostatic interaction and the multipole term). The multipole term here was estimated from a distributed multipole analysis (DMA) through rank 3 (up to octupoles) in spherical harmonics representation.[32] In addition, numerically equivalent[33] Cartesian CAMMs (cumulative atomic multipole moments)[34] up to 512-poles have been used in this study. The total MP2 interaction energy can therefore be presented as the sum

$$\Delta E^{(2)}_{MP2}(D) = \Delta E^{(1)}_{el,mtp} + \Delta E^{(1)}_{el,pen} + \Delta E^{(1)}_{ex}(D) + \Delta E^{(R)}_{del}(D) + \Delta E^{(2)}_{corr}(D) \quad (1)$$

Such a series of interaction energy components gives rise to a hierarchy of gradually simplified theoretical models, in which each next approximation is defined by neglecting an additional term from the right-hand side of eq 1. This hierarchy of theory levels is also ordered by decreasing computational effort, as described in the parentheses after each term in the sequence

$$\Delta E^{(1)}_{el,mtp}(A^2) < \Delta E^{(1)}_{el}(N^4) < \Delta E^{(1)}(N^4) < \Delta E_{SCF}(N^4) < \Delta E_{MP2}(N^5) \quad (2)$$

where A is the number of atoms and N the number of basis set functions. The term $\Delta E^{(1)}$ is the first-order interaction energy, which includes electrostatic and exchange components, and ΔE_{SCF} originates from conventional Hartree−Fock (HF) self-consistent field (SCF) energies. This decomposition scheme has been successfully used to analyze interactions between biomolecules and in crystals.[35−39]

Multipole Electrostatic Interaction Energy in the Intercalation Plane. To test the usefulness of the multipole part of the electrostatic interaction energy as a predictor of chromophore alignment during intercalation, a two-dimensional conformation space of possible chromophore positions in the intercalation plane was examined with respect to translation. The *intercalation plane* was defined by fitting a plane to the atom coordinates of the chromophore, which is close to planar. Such a map hints at the viable positions of the chromophore between the studied base pairs. A similar approach has been chosen previously by Medhi et al.,[12] who compared the electrostatic potential and binding energies of ethidium and 9-aminoacridine in idealized models of various intercalation sites.

The multipole electrostatic interaction energy was estimated at each point on a grid in the intercalation plane using cumulative atomic multipole moments (CAMMs).[34] The evaluation of electrostatic interactions in molecular systems based on multipole expansions of electron density is generally accepted to be reliable for a wide range of molecular systems, and various methods for partitioning the electron density and calculating multipoles have been studied.[40,41] There are some reservations, however, especially regarding their convergence at small intermolecular distances, since in practice the length of a multipole expansion is finite and usually limited to the first few terms. In particular, molecular multipole expansions have been shown to fail in describing interactions already for simple systems.[42] Another emerging problem concerns electrostatic penetration effects ($\Delta E^{(1)}_{el,pen}$ in eq 1), namely, the accurate evaluation of their magnitude and importance for the stabilization of molecular complexes.[20]

A CAMM of the order klm for the ith atomic center, an element of a multipole tensor of rank $l = k + l + m$, is denoted as $M^{CAMM}_{klm,i}$. It can be obtained by transforming the corresponding atomic multipole moment in Cartesian representation, $\langle x^k y^l z^m \rangle_i$, which contributes to the expectation value of the molecular moment operator through the sum $\langle x^k y^l z^m \rangle = \sum_i \langle x^k y^l z^m \rangle_i$, to its local coordinate system ($x_i y_i z_i$) through the relation[34]

$$M^{CAMM}_{klm,i} = \langle x^k y^l z^m \rangle_i - \sum_{\substack{k'\geq 0 \\ k'l'm' \neq klm}}^{k} \sum_{l'\geq 0}^{l} \sum_{m'\geq 0}^{m} \binom{k}{k'} \binom{l}{l'} \binom{m}{m'} \times x_i^{k-k'} y_i^{l-l'} z_i^{m-m'} M_i^{k'l'm'} \quad (3)$$

Therefore, $M_{000,i}$ denotes the atomic monopole moment (net charge), ($M_{100,i}$, $M_{010,i}$, $M_{001,i}$) is the atomic dipole vector, and so forth. Atomic multipole moments can be calculated from electron density distributions as

$$\langle x^k y^l z^m \rangle_i = Z_i x_i^k y_i^l z_i^m - \sum_{I\in i}^{AO} \sum_{J}^{AO} P_{IJ} \langle I | x^k y^l z^m | J \rangle \quad (4)$$

where Z_i is the nuclear charge of the ith atom, $\langle I | x^k y^l z^m | J \rangle$ is a one-electron multipole integral for the atomic orbitals I and J, and P_{IJ} is the corresponding one-electron density matrix element. The transformation described by eq 3 makes CAMMs invariant of the coordinate system used with respect to a unitary transformation. The CAMM expansion has recently been applied in various studies,[43−45] including short-range intramolecular interactions,[30,46] and CAMMs seem to provide better conformational transferability than other atomic multipole moments.[47] This property is especially valuable when many conformations of the same molecule are considered.

Calculation Details. Coordinates for the three studied intercalation complexes were obtained from X-ray crystal structure data: $Eth^{(+1)}$−UA/AU,[27] $Eth^{(+1)}$−CG/GC,[28] and $PF^{(+1)}$−CG/AU.[29] Hydrogen atoms were added to the systems using the Reduce program,[48] and their coordinates were optimized using GAMESS[49] with a PM3 model Hamiltonian (all nonhydrogen atoms were kept frozen). In the two optimized structures containing ethidium (the $Eth^{(+1)}$−UA/AU intercalation site is shown in Figure 2), the chromophore of the intercalator was separated from its side chains, as illustrated in Figure 3.

TABLE 1: Energy Components (upper part), Theory Levels (middle part), and Their Selected Ratios (lower part) to the MP2 Interaction Energies between Intercalator chromophores and their four nearest bases in the complexes Eth$^{(+1)}$−UA/AU, Eth$^{(+1)}$−GC/CG, and PF$^{(+1)}$−AU/CG, in kcal/mol

	Eth$^{(+1)}$−AU/UA[a]		Eth$^{(+1)}$−GC/CG	PF$^{(+1)}$−AU/CG	PF$^{(0)}$−AU/CG[b]
	single bases	base pairs	single bases	single bases	single bases
$\Delta\epsilon^{IX}_{CAMM}$ [c]	−10.5	−10.3	−13.0	−14.4	1.9
$\Delta E^{(1)}_{el,mtp}$ (DMA)[c]	−13.5	−10.9	−16.8	−17.6	3.1
$\Delta E^{(1)}_{el,pen}$	−12.5	−14.4	−12.3	−13.2	−14.9
$\Delta E^{(1)}_{ex}$	32.3	32.0	33.2	35.0	37.7
$\Delta E^{(R)}_{del}$	−4.7	−4.8	−6.5	−2.4	−5.3
$\Delta E^{(2)}_{corr}$	−33.3	−33.5	−32.8	−34.5	−36.8
$\Delta E^{(1)}_{el}$	−26.0	−25.3	−29.1	−30.8	−11.8
$\Delta E^{(1)}$	6.3	6.6	4.2	4.2	25.9
ΔE_{SCF}	1.6	2.0	−2.4	−2.5	20.6
$\Delta E^{(2)}_{MP2}$	−31.7	−31.5	−35.1	−37.0	−16.2
$E^{(1)}_{el,mtp}/\Delta E_{MP2}$	0.43	0.35	0.48	0.48	−0.19
$\Delta\epsilon^{IX}_{CAMM}/\Delta E_{MP2}$	0.33	0.33	0.37	0.39	−0.12
$\Delta E^{(1)}_{el}/\Delta E_{MP2}$	0.82	0.80	0.83	0.83	0.73

[a] In the case of Eth$^{(+1)}$−UA/AU, results for the single base and base pair models are compared. [b] The last column shows values for a hypothetical variant of the proflavine complex, in which proflavine is neutral. [c] The quantities ΔE are energies obtained from decomposition analysis, and $\Delta\epsilon^{IX}_{CAMM}$ denotes the rank nine multipole interaction energy calculated based on CAMM multipoles.

For all structures, the four nearest bases were separated from the DNA backbone and the cut bonds were terminated by hydrogen atoms.

Interaction energies between the separated subsystems (e.g., the chromophore and four base pairs) were calculated at the MP2 level using the 6-31G(d,p) basis set and decomposed into the contributions described by eq 1 using a direct implementation of the variation−perturbation scheme in a modified version[38] of the GAMESS software package.[49]

CAMM multipoles[34] up to rank nine (512-pole moments) were generated by custom code from SCF density matrixes with corrections for electron correlation at the MP2 level, as obtained from GAMESS[49] using the same 6-31G(d,p) basis set for consistency. Multipoles were subsequently transformed, and their interaction energies were calculated (using a formula derived by Cipriani and Silvi[50] for the elements of the Cartesian multipole interaction tensor) in an exponent-truncated scheme[51] with Python scripts[52] utilizing the SciPy library.[53]

Results and Discussion

Interaction Energy Components. Among the components of the interaction energies between the intercalator chromophores and their four nearest bases in the crystallographic structures (Table 1), the largest terms, namely, correlation and exchange, have opposite signs and comparable values for the basis set used and therefore largely cancel out in the final sum. It should be noted, however, that electron correlation effects are underestimated here; the same calculations for larger basis sets yield a larger dispersion component, whereas the exchange term remains the same. For instance, changing the basis set to 6-31G(2d,2p) increases the total interaction energy, $\Delta E^{(2)}_{MP2}$, by about 40% (see Supporting Information, section 1).

It is clear that ab initio methods that neglect even one of these components (neither HF nor density functional theory (DFT) methods cover electron correlation effects correctly) are unsuitable for studying intercalation processes, an observation that is consistent with previous reports.[11,13] It is worth noting, on the other hand, that some force fields have proved to satisfactorily reproduce binding energies of intercalators to nucleic acid bases.[13,17]

The electrostatic term ($\Delta E^{(1)}_{el}$ in Table 1) is a significant part of the total interaction energy $\Delta E^{(2)}_{MP2}$, in line with results already reported for intercalation complexes[11] and other stacked aromatic systems.[19,21,22,24] Additionally, the delocalization component is the smallest and thus relatively insignificant, in accordance with the results of Luo et al.[25]

A high correlation between the electrostatic terms and the total interaction energies is eminent for the three studied charged complexes, therefore the ratio $\Delta E^{(1)}_{el}/\Delta E^{(2)}_{MP2}$ is almost constant—from 0.80 to 0.83. The same is true for the CAMM interaction energy ($\Delta\epsilon^{IX}_{CAMM}/\Delta E^{(2)}_{MP2}$ ranges from 0.33 to 0.39) but not for the multipole term attained from DMA analysis, $\Delta E^{(1)}_{el,mtp}$. The same trend has been noticed recently for stacked benzene dimers,[21] DNA bases,[19] and other aromatic complexes.[24]

The above observation pertains to the interaction energy components of the three charged systems studied here. For the hypothetical variant PF$^{(0)}$−AU/CG, in which proflavine is deprived of a proton and therefore neutral (see Figure 1), the multipole component is significantly smaller. This makes the overall electrostatic and $\Delta E^{(2)}_{MP2}$ interaction energy smaller as well, an observation already made previously.[11] In this case, correlation and electrostatic penetration seem to be the stabilizing components.

The change in the interaction energy components due to isolating the chromophore from its side chains and terminating the cut bonds with hydrogen atoms in the Eth$^{(+1)}$−AU/UA complex is illustrated in Table 2. The sum of $\Delta E^{(2)}_{MP2}$ between the four bases and the chromophore with its side chain intact amounts to −32.74 kcal/mol, as compared with −31.69 in the case of the chromophore alone, which gives a difference underestimated by about 2.5% relative to the interaction of the side chain alone.

Table 1 also shows the interaction energy for the CAMM expansions representing the chromophores and base pairs, up to rank nine. The order n of the multipole interaction (written as $\Delta\epsilon^{n}_{CAMM}$) denotes the maximum $1/R^{n+1}$ term in an exponent truncation of the interaction between all pairs of multipoles; alternatively, this is the maximum sum of the ranks of interacting multipoles. The CAMM interaction energies given in Table 2, $\Delta\epsilon^{IX}_{CAMM}$, are converged, which is illustrated in Figure 4: the

TABLE 2: Energy Components (upper part) and Theory Levels (lower part) of the Interaction Energies in the Eth$^{(+1)}$−UA/AU Intercalation Complex[a]

	ΔE_{chain}	$\Delta E_{\text{chrom+chain}}$	$\Delta E_{\text{chrom+chain}} - \Delta E_{\text{chrom}} - \Delta E_{\text{chain}}$	ΔE_{ring}
$\Delta E^{(1)}_{\text{el,mtp}}$ (DMA)	0.5	−13.5	0.5	−1.2
$\Delta E^{(1)}_{\text{el,pen}}$	−1.1	−13.7	0.1	−2.3
$\Delta E^{(1)}_{\text{ex}}$	3.4	35.3	0.4	5.9
$\Delta E^{(R)}_{\text{del}}$	−0.5	−5.1	−0.1	−1.2
$\Delta E^{(2)}_{\text{corr}}$	−2.5	−35.7	−0.1	−3.8
$\Delta E^{(1)}_{\text{el}}$	−0.6	−27.2	0.6	−3.5
$\Delta E^{(1)}$	2.8	8.0	1.0	2.4
ΔE_{SCF}	2.3	3.0	1.0	1.2
$\Delta E^{(2)}_{\text{MP2}}$	−0.2	−32.7	0.8	−2.6

[a] The columns, from left to right, present interactions between the bases and (1) the ethidium side chain (ΔE_{chain}), (2) the ethidium chromophore with the side chain intact ($\Delta E_{\text{chrom+chain}}$), (3) the difference between the chromophore and side chain calculated together and separately ($\Delta E_{\text{chrom+chain}} - \Delta E_{\text{chrom}} - \Delta E_{\text{chain}}$), and (4) the second side chain of ethidium consisting of an aromatic ring (ΔE_{ring}). All values are sums of the contributions from the four bases, given in kcal/mol.

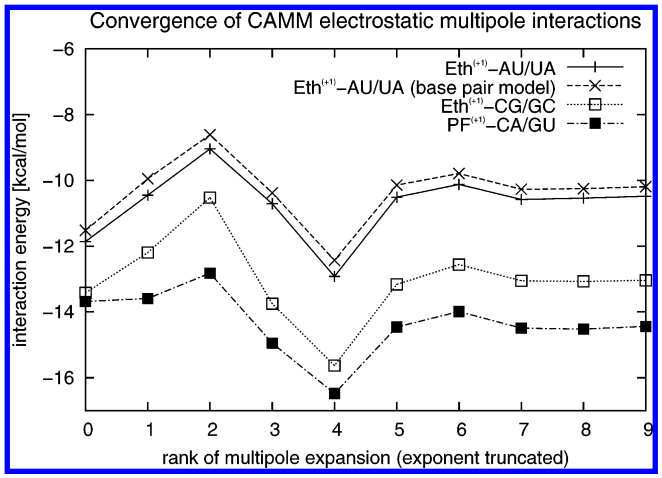

Figure 4. Multipole interaction energies for the three studied intercalation complexes at consecutive ranks, $\Delta\epsilon^{n}_{\text{CAMM}}$. Results for the single base and base pair model are compared in the case of the Eth$^{(+1)}$−UA/AU complex. For numerical values, see section 2 of the Supporting Information.

$\Delta\epsilon^{n}_{\text{CAMM}}$ for all three studied complexes are similar and converge only for $n > 5$. The lack of convergence for lower ranks is due to small intermolecular distances (the normal distance between intercalator and bases is 3.4 Å) and multiple contacts between atoms, characteristic of stacking complexes. It is worthwhile to note that the interactions of the corresponding molecular multipole expansions are divergent (see Supporting Information, section 2), which is expected considering the previous failures of molecular expansions for relatively simpler systems.[42] The weak convergence found in this case may be possibly improved by adding additional, off-atom centers to the multipole expansion. These energies exhibit a similar trend already for ranks above two, which make them useful in comparative studies, such as dealing with the effect of base pair sequence.

The slow convergence of the electrostatic multipole term is relevant in the context of penetration effects. Since the magnitude of these effects is calculated as the difference between the total electrostatic interaction and its multipole component, it will be inaccurate if the mulitpole expansion used does not provide a converged value. Previous studies using atomic expansions have rarely proceeded beyond octupole moments. For instance, Toczylowski et al.[20] have recently published an exhaustive study on the electrostatics of hydrogen-bonded and stacked DNA bases, in which a distributed multipole analysis (DMA) was expanded through hexadecapoles. In light of our results, this does not suffice, since for such systems multipole interactions start to converge only above rank five (see Figure 4) and using DMA moments up to rank four results in an overestimate. In turn, this may lead to an incorrect estimate of the penetration term. The electrostatic penetration term could be represented in an atomic multipole expansion as advocated by Jug et al.[54]

Another interesting observation is that treating base pairs as a whole, as suggested by Reha et al.,[13] does not significantly change this trend in the Eth$^{(+1)}$−AU/UA complex. While there is a constant difference of ∼0.4 kcal/mol in this case between the base pair and single base models used, the convergence trends behave similarly. The fact that these values do not largely differ hints that base pair polarization, while influencing the properties of the bases themselves, does not qualitatively change their interactions with the intercalator.

Finally, the CAMM interaction energy for the complex Eth$^{(+1)}$−GC/CG can be compared with the multipole electrostatic binding energy obtained by Medhi et al.,[12] which was based on the same structure and obtained from a DMA. The discrepancy between the value in that study (−7.9 kcal/mol) and our values (−16.8 kcal/mol from DMA analysis and −13.0 kcal/mol for the converged CAMM interaction) probably originates from the fact that Medhi et al. used an idealized model and from the differences in the electron density source (our multipole moments were calculated from density matrixes, including corrections for electron correlation at the MP2 level, the former were from SCF wave functions).

Intercalator Alignment in the Intercalation Plane. The interaction energy components were also examined for other positions of the ethidium chromophore in the intercalation plane of the Eth$^{(+1)}$−UA/AU complex, in the immediate vicinity of the intercalation site—including distances of ±0.5 Å and ±2.0 Å toward the major groove and at the same distances in the perpendicular direction. The interaction energy at various levels of theory and its components at these points are presented in Figure 5 and Figure 6, respectively. The ΔE_{MP2} energy has a minimum very close to the crystallographic position of the chromophore (zero on the plots)—a parabola fit gives a diagonal offset (originating from two perpendicular directions in the plane of intercalation) of 0.09 Å and an energy difference of about 0.05 kcal/mol. This supports the notion that factors other than local interactions—that is, those between the intercalator's chromophore and its nearest bases—are of minor importance for chromophore alignment and justifies the fragmentation method adopted for ethidium in this study.

The CAMM multipole interaction energy between the chromophores and the four nearest bases was calculated for each system on a grid in the intercalation plane as described in the Methods section. This energy may be presented in the form of a surface plot as a function of the displacement of the chromophore's geometric center from the crystal position—as in Figure 7 for the Eth$^{(+1)}$−UA/AU system. Steric constraints caused by the DNA side chain are illustrated here by the shaded region, where the distance between any pair of atoms is smaller than the sum of their van der Waals radii, scaled by a factor of 0.5. Similar plots were also evaluated in the intercalation planes of the other two studied cationic systems, Eth$^{(+1)}$−CG/GC and PF$^{(+1)}$−AU/CG (see Supporting Information, section 3); all three surface plots have a significant central minimum. On the other hand, the corresponding multipole interaction surface for

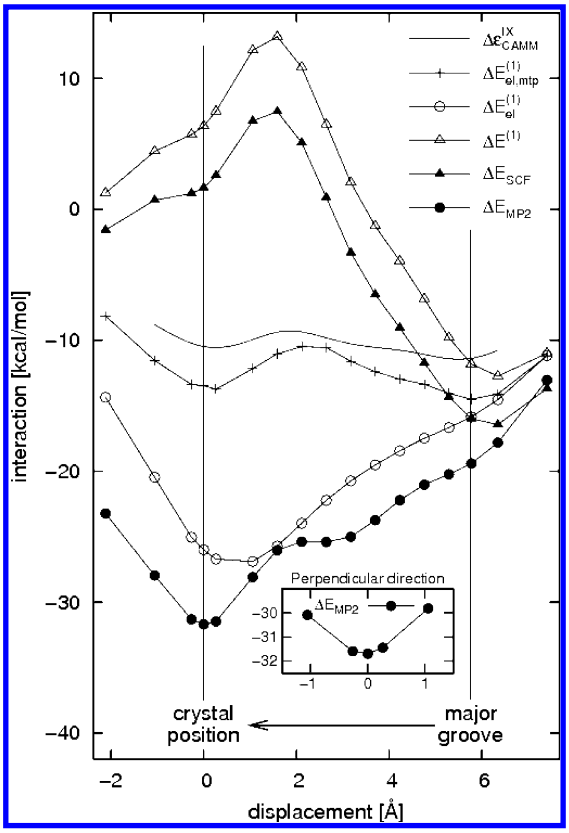

Figure 5. Interaction energy between the ethidium chromophore and its four nearest bases (Eth$^{(+1)}$−AU/UA complex) at various levels of theory on a path in the intercalation plane, which connects the crystallographic position of the chromophore and the minimum found in the major groove of the converged multipole interaction surface (bold segment in Figure 7).

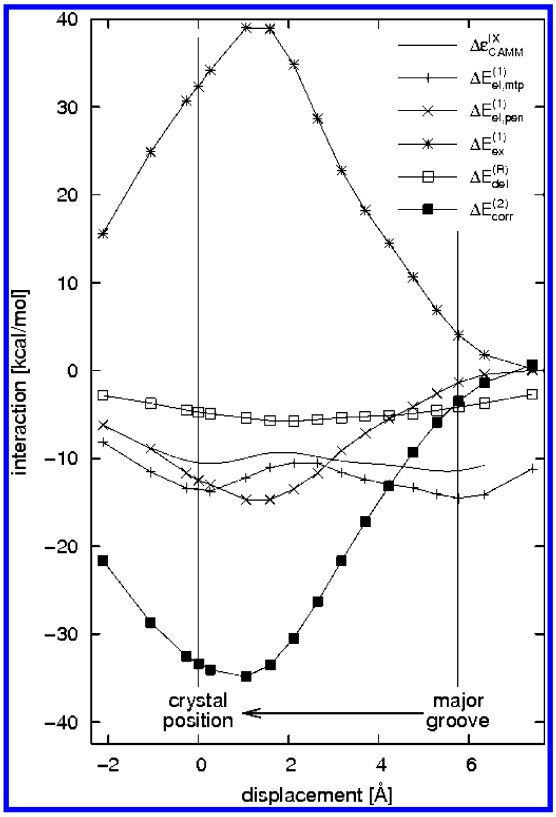

Figure 6. Energy components of the interaction energy between the ethidium chromophore and its four nearest bases (Eth$^{(+1)}$−AU/UA complex) on a path in the intercalation plane, which connects the crystallographic position of the chromophore and the minimum found in the major groove of the converged multipole interaction surface (bold segment in Figure 7).

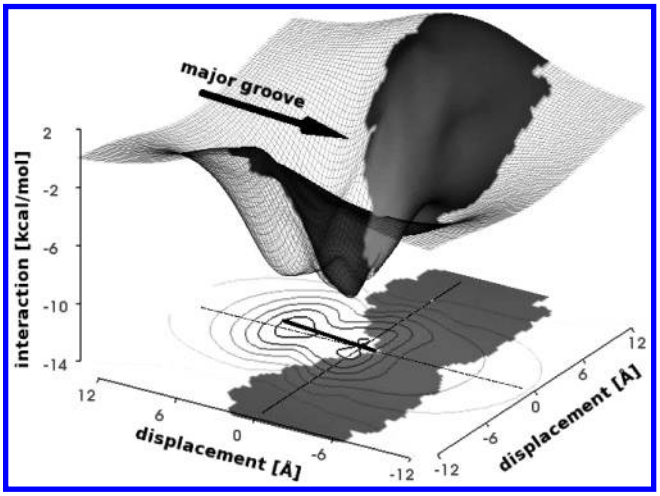

Figure 7. CAMM electrostatic interaction energy surface, $\Delta\epsilon^{IX}_{CAMM}$, between the ethidium chromophore and UA/AU bases, produced by translating the chromophore in the plane of intercalation with a step of approximately 0.26 Å. The bottom axes denote the displacement of the chromophore's geometric center from the crystal position. Areas inaccessible due to steric constraints are shaded. The direction from the major groove (positive displacements) is indicated by the arrow, and the path used in Figures 5 and 6 is represented by the bold segment on the contour plot below the surface.

the neutral proflavine complex is highly irregular and does not reproduce the crystal binding site in any reasonable way. It should be noted that, similar to the CAMM interaction at the crystallographic site (Figure 4), these maps as a whole also converge only after including fifth-order contributions (see Supporting Information, section 4).

The deep central minimum in Figure 7 has a value of −11.0 kcal/mol at the coordinates (−1.3 Å, 0.0 Å), 0.3 kcal/mol below that of the crystal position. This minimum is accompanied by a second one, at the coordinates (1.5 Å, 0.5 Å) with a value of −10.7 kcal/mol. Similar distances from the crystal position were obtained for the $\Delta\epsilon^{IX}_{CAMM}$ surfaces of the other two cationic intercalation complexes: 1.8 Å for Eth$^{(+1)}$−GC/CG and 1.7 Å for PF$^{(+1)}$−AU/GC, which demonstrates the accuracy of this approach in reproducing the alignment of chromophores between base pairs in these types of systems. In all cases, two central minima were located on opposite sides of the crystallographic positions.

An additional minimum is also present in Figure 7 in the direction of the major groove, at a distance of 5.7 Å from the crystal position at −11.5 kcal/mol. This minimum is observed for only the two complexes involving ethidium and at all interaction ranks. If similar studies were to be performed for model systems without crystallographic data, then it would be necessary to study all the obtained minima with a more exact method.

When an intercalator approaches the intercalation site from the major groove, it should experience an increase of penetration and dispersion interactions. To further investigate the relevance of the off-center minimum in $\Delta\epsilon_{CAMM}$ for the total interaction energy in the intercalation plane, the components of ΔE_{MP2} were calculated for the ethidium chromophore located along a path connecting the crystallographic binding site and this minimum in the Eth$^{(+1)}$−UA/AU structure, up to 8 Å away. This path simulates, in a very simplified way, the movement of the chromophore when entering the intercalation site (see section

5 in Supporting Information for an animation). Figures 5 and 6 present the components calculated at different levels of theory along this path. The most evident conclusion is that the profiles of the exchange ($\Delta E_{ex}^{(1)}$) and correlation ($\Delta E_{corr}^{(2)}$) terms have similar shapes and opposite signs. These two components, along with the electrostatic penetration term, cancel each other out at distances above 5 Å. It is interesting to note that the first-order interaction ($\Delta E^{(1)}$) and SCF interaction energy (ΔE_{SCF}) completely fail to reproduce the crystallographic binding site even along this one-dimensional path.

Furthermore, besides the full MP2 interaction energy, the multipole components (both the one calculated from CAMM multipoles, $\Delta\epsilon_{CAMM}$, and from the interaction energy decomposition, $\Delta E_{el,mtp}^{(1)}$) best reproduce the crystallographic binding site along this path, more precisely than the entire electrostatic interaction energy $\Delta E_{el}^{(1)}$ (which includes penetration effects) or the correlation component $\Delta E_{corr}^{(2)}$. On the other hand, the magnitude of the multipole component is significantly smaller than that of $\Delta E_{MP2}^{(2)}$ but becomes more dominant as the chromophore is withdrawn from the intercalation site. Although the MP2 interaction energy does not have a second minimum where $\Delta\epsilon_{CAMM}^{IX}$ does, it is not monotonic along the studied path and exhibits a plateau around 2.6 Å from the binding site.

Conclusions

(1) The MP2 interaction energy—between the ethidium chromophore and its nearest bases in the $Eth^{(+1)}$—UA/AU complex—reproduces within 0.1 Å the crystallographic binding site in the intercalation plane (Figure 5). This demonstrates that local interactions alone may decide about the alignment of the chromophore between base pairs, whereas the ethidium side chain and even steric constraints with the nucleic acid backbone are of minor importance (Table 2).

(2) An analysis of the MP2 interaction energy indicates that the electrostatic term comprises the same percentage of the total interaction energy for all three studied intercalation complexes ($\Delta E_{el}^{(1)}/\Delta E_{MP2}^{(2)}$ is around 0.82 for the three studied complexes). Less than half of this originates from the multipole component ($\Delta\epsilon_{CAMM}/\Delta E_{MP2}^{(2)}$ was around 0.35 in all three cases), illustrating the magnitude of electrostatic penetration effects. It should be kept in mind that these values will be smaller for larger basis sets but can be expected to correlate with the total energy.

(3) Multipole interaction energies for the three studied systems do not converge fast with rising rank (Figure 4), due to small intermolecular distances (3.4 Å) and multiple atom—atom contacts. In stacked aromatic systems, it is necessary to consider atomic multipoles above rank five to achieve convergence. A converged multipole interaction term is in turn required to obtain an accurate estimate of electrostatic penetration effects.

(4) The binding site in the intercalation plane can be reasonably reproduced at a lower computational cost than the MP2 level, by examining the interaction energy surface of the chromophore between nucleotides with CAMM multipoles (Figure 7). The accuracy is worse and in the studied systems was no less than about 1.3 Å. Such an approach may be used to qualitatively evaluate the alignment of charged intercalators between base pairs. In contrast, when the system involving proflavine is modified so that the intercalator is neutral, the multipole component has no predictive value and its corresponding interaction surface is highly irregular. Besides a minimum corresponding to the crystallographic binding site, additional minima in the direction of the DNA grooves are found and all should be supplemented by more exact calculations if the actual alignment is unknown.

Acknowledgment. The calculation results presented in this work were obtained at the Wroclaw Center for Networking and Supercomputing (WCSS), Poznan Supercomputing and Networking Center (PCSS), and Interdisciplinary Center for Mathematical and Computational Modeling at Warsaw University. This study was facilitated by NSF-CREST Grant No. HRD-0318519 and support from Wroclaw University of Technology. K.M.L. would like to thank Robert Gora for providing the working version of his software that implements the interaction energy decomposition scheme and for valuable comments. K.M.L. would also like to thank the National Science Foundation for financial support during the 2003 Summer Institute at CCMSI.

Supporting Information Available: Section 1: Dependence of interaction energy components on the basis set. Section 2: Convergence of $\Delta\epsilon_{CAMM}$ multipole interactions for the crystallographic structures. Section 3: Converged multipole interaction surfaces in the intercalation plane for the other studied intercalation complexes. Section 4: Convergence of $\Delta\epsilon_{CAMM}$ multipole interaction surfaces. Section 5: Animation of the studied path on the intercalation plane. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Graves, D. E.; Velea, L. M. *Curr. Org. Chem.* **2000**, *4* (9), 915—929.
(2) Lerman, L. S. *J. Mol. Biol.* **1961**, *3* (1), 18.
(3) Waring, M. J. *J. Mol. Biol.* **1965**, *13* (1), 269.
(4) Lepecq, J. B.; Paoletti, C. *J. Mol. Biol.* **1967**, *27* (1), 87.
(5) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2* **2001**, (5), 651—669.
(6) Waring, M. J. *Annu. Rev. Biochem.* **1981**, *50*, 159—192.
(7) Cashman, D. J.; Kellogg, G. E. *J. Med. Chem.* **2004**, *47* (6), 1360—1374.
(8) Starcevic, K.; Karminski-Zamola, G.; Piantanida, I.; Zinic, M.; Suman, L.; Kralji, M. *J. Am. Chem. Soc.* **2005**, *127* (4), 1074—1075.
(9) Fantacci, S.; de Angelis, F.; Sgamellotti, A.; Marrone, A.; Re, N. *J. Am. Chem. Soc.* **2005**, *127* (41), 14144—14145.
(10) Chaires, J. B. *Biopolymers* **1997**, *44* (3), 201—215.
(11) Kubar, T.; Hanus, M.; Ryjacek, F.; Hobza, P. *Chem.—Eur. J.* **2005**, *12* (1), 280—290.
(12) Medhi, C.; Mitchell, J. B. O.; Price, S. L.; Tabor, A. B. *Biopolymers* **1999**, *52* (2), 84—93.
(13) Reha, D.; Kabelac, M.; Ryjacek, F.; Sponer, J.; Sponer, J. E.; Elstner, M.; Suhai, S.; Hobza, P. *J. Am. Chem. Soc.* **2002**, *124* (13), 3366—3376.
(14) Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2002**, *124* (39), 11802—11808.
(15) Seio, K.; Ukawa, H.; Shohda, K.; Sekine, M. *J. Biomol. Struct. Dyn.* **2005**, *22* (6), 735—746.
(16) Bondarev, D. A.; Skawinski, W. J.; Venanzi, C. A. *J. Phys. Chem. B* **2000**, *104* (4), 815—822.
(17) Dracinsky, M.; Castano, O. *Phys. Chem. Chem. Phys.* **2004**, *8* (6), 1799—1805.
(18) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61* (1), 3—31.
(19) Hill, G.; Forde, G.; Hill, N.; Lester, W. A.; Sokalski, W. A.; Leszczynski, J. *Chem. Phys. Lett.* **2003**, *381* (5—6), 729—732.
(20) Toczylowski, R. R.; Cybulski, S. M. *J. Chem. Phys.* **2005**, *123* (15).
(21) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124* (1), 104—112.
(22) Price, S. L.; Stone, A. J. *J. Chem. Phys.* **1987**, *86* (5), 2859—2868.
(23) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116* (11), 4993—4994.
(24) Perez-Casas, S.; Hernandez-Trujillo, J.; Costas, M. *J. Phys. Chem. B* **2003**, *107* (17), 4167—4174.
(25) Luo, R.; Gilson, H. S. R.; Potter, M. J.; Gilson, M. K. *Biophys. J.* **2001**, *80* (1), 140—148.
(26) Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Am. Chem. Soc.* **2005**, *127* (4), 8594—8595.
(27) Jain, S. C.; Sobell, H. M. *J. Biomol. Struct. Dyn.* **1984**, *1* (5), 1161—1177.
(28) Jain, S. C.; Sobell, H. M. *J. Biomol. Struct. Dyn.* **1984**, *1* (5), 1179—1194.

(29) Aggarwal, A.; Islam, S. A.; Kuroda, R.; Neidle, S. *Biopolymers* **1984**, *23* (6), 1025−1041.

(30) Sokalski, W.; Kedzierski, P.; Grembecka, J.; Dziekonski, P.; Strasburger, K. *Computational Molecular Biology*; Elsevier-Science: Amsterdam, The Netherlands, 1999; Chapter 10, pp 369−396.

(31) Sokalski, W. A.; Roszak, S.; Pecul, K. *Chem. Phys. Lett.* **1988**, *153* (2−3), 153−159.

(32) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233−239.

(33) Spackman, M. A. *J. Chem. Phys.* **1986**, *85*, 6587−6601.

(34) Sokalski, W. A.; Poirier, R. A. *Chem. Phys. Lett.* **1983**, *98* (1), 86−92.

(35) Sokalski, W. A.; Kedzierski, P.; Grembecka, J. *Phys. Chem. Chem. Phys.* **2001**, *3* (5), 657−663.

(36) Gora, R. W.; Bartkowiak, W.; Roszak, S.; Leszczynski, J. *J. Chem. Phys.* **2002**, *117* (3), 1031−1039.

(37) Dyguda, E.; Grembecka, J.; Sokalski, W. A.; Leszczynski, J. *J. Am. Chem. Soc.* **2005**, *127* (6), 1658−1659.

(38) Gora, R. W.; Sokalski, W. A.; Leszczynski, J.; Pett, V. B. *J. Phys. Chem. B* **2005**, *109* (5), 2027−2033.

(39) Szefczyk, B.; Mulholland, A. J.; Ranaghan, K. E.; Sokalski, W. A. *J. Am. Chem. Soc. 126*, 16148−16159.

(40) Popelier, P. L. A.; Joubert, L.; Kosov, D. S. *J. Phys. Chem. A* **2001**, *105* (35), 8254−8261.

(41) Volkov, A.; Coppens, P. *J. Comput. Chem.* **2004**, *25* (7), 921−934.

(42) Qian, W. L.; Krimm, S. *J. Phys. Chem. A* **2005**, *109* (25), 5608−5618.

(43) Kedzierski, P.; Wielgus, P.; Sikora, A.; Sokalski, W. A.; Leszczynski, J. *Int. J. Mol. Sci.* **2004**, *5*, 186−195.

(44) Gresh, N.; Kafafi, S. A.; Truchon, J. F.; Salahub, D. R. *J. Comput. Chem.* **2004**, *25* (6), 823−834.

(45) Rinaldi, D.; Bouchy, A.; Rivail, J. L.; Dillet, V. *J. Chem. Phys.* **2004**, *120* (5), 2343−2350.

(46) Strasburger, K.; Sokalski, W. A. *Chem. Phys. Lett.* **1994**, *221* (1−2), 129−135.

(47) Kedzierski, P.; Sokalski, W. A. *J. Comput. Chem.* **2001**, *22* (10), 1082−1097.

(48) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285* (4), 1735−1747.

(49) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14* (11), 1347−1363.

(50) Cipriani, J.; Silvi, B. *Mol. Phys.* **1982**, *45* (2), 259−272.

(51) Stolarczyk, L. Z.; Piela, L. *Int. J. Quantum Chem.* **1979**, *15* (6), 701−711.

(52) van Rossum, G. Python Reference Manual CWI Report CS-R9524, 1995.

(53) Jones, E.; Oliphant, T.; Peterson, P.; et al. *SciPy: Open source scientific tools for Python*, 2001.

(54) Koster, A. M.; Kolle, C.; Jug, K. *J. Chem. Phys.* **1993**, *99* (2), 1224−1229.