Electrostatic Energy in the Effective Fragment Potential Method: Theory and Application to Benzene Dimer

LYUDMILA V. SLIPCHENKO, MARK S. GORDON
Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50011

Received 11 May 2006; Revised 3 June 2006; Accepted 4 June 2006
DOI 10.1002/jcc.20520
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: Evaluation of the electrostatic energy within the effective fragment potential (EFP) method is presented. The performance of two variants of the distributed multipole analysis (DMA) together with two different models for estimating the charge penetration energies was studied using six homonuclear dimers. The importance of damping the higher order multipole terms, i.e. charge dipole, was also investigated. Damping corrections recover more than 70% of the charge penetration energy in all dimers, whereas higher order damping introduces only minor improvement. Electrostatic energies calculated by the numerical DMA are less accurate than those calculated by the analytic DMA. Analysis of bonding in the benzene dimer shows that EFP with inclusion of the electrostatic damping term performs very well compared to the high-level coupled cluster singles, doubles, and perturbative triples method. The largest error of 0.4 kcal/mol occurs for the sandwich dimer configuration. This error is about half the size of the corresponding error in second order perturbation theory. Thus, EFP in the current implementation is an accurate and computationally inexpensive method for calculating interaction energies in weakly bonded molecular complexes.

 J Comput Chem 28: 276–291, 2007

Key words: electrostatic screening; EFP; dispersion; benzene-benzene

Introduction

Modeling intermolecular interactions play an important role in studying liquids, molecular clusters, surface catalysis, and biosystems: it is of interest in many fields of chemistry, physics, biological sciences, and materials. However, accurate calculation of intermolecular interactions is a non-trivial task for quantum chemistry. High-level *ab initio* calculations can accurately describe weak intermolecular forces in systems of small size. However, the accuracy of *ab initio* calculations markedly depends on both the basis set size and dynamical correlation. As a result, such calculations are very computationally demanding and fast become impractical as the system size increases.

Alternatively, one can describe intermolecular interactions using perturbation theory, starting from the non-interacting (unperturbed) fragments. In the last decade, this idea has been developed within the context of the effective fragment potential (EFP) method.[1] In terms of perturbation theory, intermolecular interactions can be presented as a series of short- and long-range terms. Long-range interactions (interactions that are proportional to the distance according to $(1/R)^n$) include Coulomb, induction, and dispersion terms, whereas short-range interactions, which decay exponentially, consist of exchange-repulsion, charge-transfer, and screening terms. Screening terms are used to damp induction, dispersion, and Coulomb energies. All of the types of interactions described here, including screening, are present in the current EFP implementation.

The Coulomb (sometimes called electrostatic) interaction is a leading term in the bonding of many molecular complexes. To avoid very time-consuming evaluation of the integral over molecular charge distributions, in low-cost computational models it is useful to use a distributed multipole approximation, as described by Stone.[2,3] In this approach, the Coulomb potential is expanded in a series of terms in $(1/R)$:

$$V = V^{\text{charge}} + V^{\text{dipole}} + V^{\text{quadrupole}} + V^{\text{octopole}} + \cdots \qquad (1)$$

In the EFP method, the Coulomb potential is expanded up to octopoles, with expansion centers at each atom center and each bond midpoint.[4]

The Coulomb energy of two interacting molecules (A and B) can be evaluated according to the classical multipolar interaction:

$$E^{\text{ES}} = q_{\text{A}}V_{\text{B}}^{\text{ch}} + q_{\text{A}}V_{\text{B}}^{\text{dip}} + q_{\text{B}}V_{\text{A}}^{\text{dip}} + \cdots, \qquad (2)$$

where q_{A} is the charge on center A, V_{B}^{ch} is the potential due to the charge on B, etc. There are several possible procedures for distributing multipoles over molecular sites.[2,3,5–10] In Stone's distributed multipole analysis (DMA),[2,3] partitioning of the molecu-

Correspondence to: M. S. Gordon; e-mail: mark@si.fi.ameslab.gov

Contract grant sponsors: National Science Foundation NIRT; SciDAC grant, Department of Energy

lar density among atoms is carried out in basis-function space. However, it has been noted[10] that the DMA is unstable with respect to increase of the basis set. Several methods that partition the density in real space have also been introduced. Examples are the atoms-in-molecule procedure of Bader[5] and a similar method for distributed multipoles by Popelier.[9] Another interesting approach suggested by Harrision[6] uses the Hirschfeld partitioning method, in which the molecular electron density is assigned to atoms in proportion to the free atom densities at the same distances from the nuclei.[7,8] Recently, Stone reported a DMA approach, referred to here as the numerical Stone DMA, which combines partitioning in both basis-function and real space.[10]

In the current EFP method, the Coulomb energy is evaluated using Stone's DMA. However, the instability of this method with respect to increasing the basis set becomes especially problematic when diffuse functions are added to the basis set, or in systems with high orbital degeneracy, such as benzene. To potentially overcome these difficulties, the numerical Stone DMA procedure has been implemented into the EFP method. In the following sections the analytic and numerical DMA are discussed in detail, and their performance compared for a series of molecular dimers.

Representing the molecular electron density by a set of multipole expansions markedly decreases the computational cost of evaluating the Coulomb energy. However, this classical approximation is valid only when the electron densities of interacting molecules do not overlap, i.e. at large separations. As the intermolecular separation decreases and the molecular electron densities begin to overlap, the nuclei of one molecule become less shielded by their electron density and experience stronger attraction by the electron density of the other molecule. Thus, the real electrostatic interaction is somewhat stronger than that calculated by classical multipoles due to charge penetration. It appears that at equilibrium geometries of many molecular clusters, charge penetration plays a significant role. For example, as discussed in more detail later, the charge penetration is about 15% of the total electrostatic energy in the water dimer, and it is as large as 200% in the sandwich structure of benzene dimer. Of course, the charge penetration contribution to the total energy becomes even larger at shorter intermolecular separations.

It has been shown[4,11–13] that the charge penetration energy can be estimated by adding a damping (charge penetration) term to the classical multipolar potential. In the current implementation of the damping term, only the charge–charge part of the penetration energy is included.

In this work, expressions are presented for the higher order charge-penetration contributions; i.e., charge–dipole, dipole–dipole, and charge–quadrupole. This facilitates an analysis of the importance of the high-order damping in the framework of both analytic and numerical DMAs. This analysis provides insight regarding the benefits and drawbacks of analytic and numerical methods. The numerical results include analysis of the Coulomb (electrostatic) interaction in dimers of six solvent molecules and ethylene, as well as an in-depth investigation of the benzene dimer.

The benzene dimer, a prototype for π-π interactions, has attracted extensive theoretical and experimental attention.[14–31] π-π interactions are an important noncovalent interaction governing structures of proteins and DNA, self-assembly of aromatic macromolecules, and drug-intercalation into DNA. In spite of the ubiquity and importance of π-π interactions, many levels of theory cannot provide an adequate description of this type of interaction. In the present work, an in-depth study of π-π interactions predicted by the EFP method is compared with very accurate *ab initio* studies by Sherrill and coworkers.[18,19]

The structure of the paper is as follows: the next section presents theoretical aspects of the analytic and numerical Stone DMA techniques, and the evaluation of the formulas for the charge-penetration corrections. Computational details are given in the next section, followed by numerical results and discussion, and final remarks are presented in the last section.

Theory

Analytic DMA

The distributed multipolar analysis was proposed by Stone in 1981.[2,3] This DMA can be applied to any molecular density that is expanded in primitive Gaussian products. A Gaussian function centered at A is comprised of a Gaussian function, $\exp(-\xi_A(\mathbf{r}-A)^2)$, and a polynomial of a degree that corresponds to the angular momentum of the function, i.e., 0 for s, 1 for p, and so on. A product of two Gaussian functions centered at A and B is another Gaussian at a different center, whose polynomial degree is the sum of the degrees of the original polynomials, and the Gaussian part has the form $\exp(-(\xi_A+\xi_B)(\mathbf{r}-\mathbf{P})^2)$, where $\mathbf{P}$ is the overlap center $\mathbf{P}=(\xi_A\mathbf{A}+\xi_B\mathbf{B})/(\xi_A+\xi_B)$. Thus, the position of the product Gaussian is on the line connecting points A and B. The weighted sum of the products of the Gaussian functions is the total density, with individual coefficients being determined from the density matrix.

Any Gaussian product can be expressed as a sum of multipole moments of ranks up to the degree of the associated polynomial. Each product is shifted to the closest of the expansion centers. A set of DMA expansion centers is user-defined; usually this set consists of atoms and bond mid-points.

A main drawback of the analytic DMA is its instability with respect to increases in the size of the basis set. In a large basis set, and especially a basis set with diffuse functions, there are many ways to express a given charge distribution, and the most optimal one may be determined variationally. In another basis set, with a different set of exponents, the same charge distribution may be optimally represented by a different DMA expansion. This ambiguity increases as the size of the basis set increases, and it can result in very different and basis set-dependent distributed multipoles. However, as emphasized in Ref. 10, these different sets of multipoles still produce almost identical electrostatic potentials at Van der Waals distances. Therefore, the instability of multipole moments with increase of basis set does not seem to be critical. However, the divergence of the multipoles in a large basis set becomes much more troublesome when a molecule of interest contains degenerate orbitals. For example, for benzene the analytic DMA produces huge values of quadrupole and octopole moments even when a single diffuse function is added to the basis set. However, in the EFP method, the basis set must contain diffuse functions to correctly describe the exchange–repulsion interaction.[32] Therefore, in the EFP framework, the analytic DMA seems to be inappropriate for calculations on benzene and possibly some benzene derivatives.

Numerical DMA

As an alternative to the analytic DMA a numerical version of the method recently suggested by Stone[10] has been implemented. This numerical method is a mixture of the original DMA that partitions the density in basis function space and a numerical integration scheme that does partitioning in real space. The original method is used to partition the density of the most compact basis functions. This avoids any difficulties with numerical integration of strongly peaked functions. The numerical scheme is used for diffuse functions, thereby eliminating the strong basis set dependence of the original method.

The implementation of the numerical DMA described here closely follows the original implementation by Stone.[10] A grid of integration points is constructed around each atom. This consists of a 100-point radial grid and a 590-point angular Lebedev grid.[33] An Euler–Maclaurin quadrature is used for radial integration.[34] The radial grid is scaled by values of the Bragg–Slater covalent radii.[35] The softness of the boundary between atoms is controlled by Becke's cutoff parameter k.[36] Larger values of k give a sharper cutoff. Numerical results presented in this work were obtained with $k = 2$. The boundary between atoms can be set halfway between them or scaled by the ratio of the Bragg–Slater radii of each atom. The radii of the bond midpoints are determined as an average of the corresponding atomic radii. In this work, the Bragg–Slater radii were used to set the cutoffs between atoms. A switch between the analytic and numerical algorithms works as follows: if the sum of the exponents $\xi_A + \xi_B$ in a product of primitives is larger than the specified switch value, then the product function is considered to be compact and the original analytic integration procedure is performed. If the sum of exponents is smaller than the switch value, the numerical integration on the grid is employed. Following Stone's recommendation, a switch value of 4 was used. This switch produces moments very similar to those that are produced with the pure numerical algorithm.

Both the multipole moments and the resulting electrostatic energies depend on the parameters of the numerical integration scheme, that is, the position and sharpness of the cutoffs between the expansion centers. It is found that for some molecules, parameters used throughout this work (Bragg–Slater radii on each center and $k = 2$) perform best in terms of electrostatic energies, while for other molecules, other values of these parameters (e.g., equal radii on centers and $k = 4$) would give better results. However, there is no obvious reason for this trend, and so the same set of parameters has been used in all calculations presented here.

Electrostatic Screening

Charge penetration causes a decrease in the classical electrostatic energy of interacting multipoles due to the overlap of electronic densities. Stone illustrated the origin of charge penetration by considering the interaction of a hydrogen-like atom with nuclear charge Z with a proton.[2] The electrostatic potential of the electron of a hydrogen-like atom has the form

$$V(r) = -\frac{e}{R} + e\exp(-2ZR)\left(Z + \frac{1}{R}\right). \quad (3)$$

This expression can be rewritten as

$$V(r) = -\frac{e}{R}[1 - \exp(-2ZR)(1 + ZR)] = -\frac{e}{R}f^{\text{damp}}(2ZR). \quad (4)$$

Thus, the potential of an electron in a hydrogen-like atom differs from the classical charge potential by the damping function $f^{\text{damp}}(2ZR)$. This suggests that, in general, the multipole expansion of the electrostatic potential can be corrected for charge penetration by this damping function f^{damp}, which has exponential dependence on the separation R. Introducing a parameter $\alpha = 2Z$, one can rewrite the damping function for the charge potential as

$$f_1^{\text{damp}} = 1 - \exp(-\alpha R)\left(1 + \frac{\alpha R}{2}\right). \quad (5)$$

The parameter α can be determined by minimizing the difference between the quantum mechanical electrostatic potential and the damped multipolar expansion over a grid of points:

$$\Delta = \sum_{\text{grid}}[V_{\text{ab initio}} - V_{\text{damped multipole}}]^2. \quad (6)$$

More details on the fitting procedure will be given in Computational Details.

In the previous work on the charge-penetration correction to the electrostatic energy,[12] a simpler form of the damping function was used

$$f_2^{\text{damp}} = 1 - \exp(-\alpha R). \quad (7)$$

Comparison of two damping functions is given in Figure 1. Both types of damping functions go to unity for large R and fall off toward zero as R approaches zero. As seen from Figure 1a, f_1 and f_2 behave similarly at large R; however, they differ at short and medium distances (0–2 au).

Consider the hydrogen-like atom. For the spherically symmetric hydrogen-like atom, a one-dimensional grid can be employed to numerically determine parameter α for f_1 and f_2. Since the f_1 damping function gives the correct form for the potential, the parameter α does not depend on characteristics of the grid, such as the separation between grid points and the minimum and maximum distances from nucleus, r_{min} and r_{max}. Thus, in a hydrogen-like atom, α for f_1 will be determined uniquely. The damping function shown in Figure 1a has $\alpha = 2.5$. On the other hand, the optimal parameter for f_2 will depend on the properties of the grid. Figure 1a illustrates three different f_2 functions with $\alpha = 1.2$, 1.4, and 1.6. The corresponding errors relative to f_2 are given in Figure 1b. Note that a grid starting further away from the nucleus, i.e., the grid with larger r_{min}, will produce a larger value of α. For example, for a grid with $r_{\text{min}} = 0.5$ the optimal α would be 1.2, and for the grid with $r_{\text{min}} = 1.5$ the optimal $\alpha = 1.6$.

Of course, the form of the potential around each expansion center in a many-electron system differs from that in a hydrogen-like atom. However, as will be discussed in Results and Discussion, the general tendency shown in Figure 1 for damping functions f_1 and f_2 holds for many molecules.

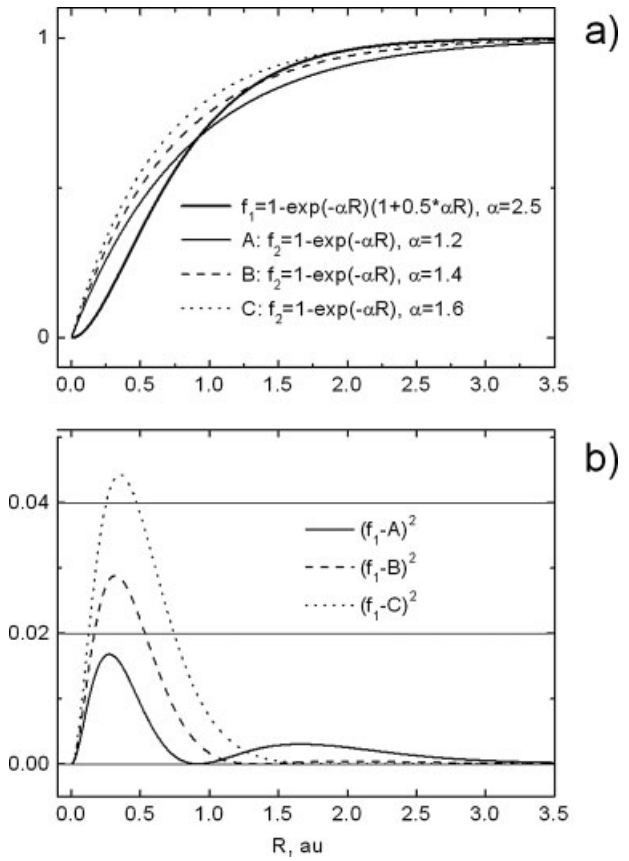

Figure 1. Electrostatic damping functions. (a) f_1 damping function with parameter $\alpha = 2.5$ and f_2 with three different $\alpha = 1.2$, 1.4, 1.6 are shown. f_1 and f_2 behave similarly at long distances but are very different at short distances. (b) Squared errors of f_2 with different values of α with respect to f_1 are shown.

The minimum distance from the nucleus is an important characteristic of the grid from a practical point of view because in any molecule, the potential in close proximity to a nucleus is not uniform. This makes the fitting process cumbersome. Note also that (see Fig. 1) f_2 always overestimates damping at small R and underestimates damping at large R. So, as $r_{\min}$ and, consequently α, increase, the short-range overestimation of damping will increase, and long-range underestimation will diminish. Therefore, if f_2 is used as a damping function, one must determine which region of molecular separation is more important for accurate damping. In general, it appears that molecules with strong electrostatic interactions (and shorter intermolecular distances) prefer damping parameters that average the error over all distances (i.e., smaller $r_{\min}$), whereas weakly interacting molecules prefer larger $r_{\min}$ in order to better describe damping at longer distances. However, in spite of the uncertainty in choosing the optimal $r_{\min}$, the simple exponential damping function has been shown to give reasonable charge penetration corrections for many molecular complexes.[12] Therefore, both $f_1$and f_2 are considered here as possible damping functions, referred to as model 1 and model 2. Thus, the following two damped charge potentials will be considered:

$$V_1^{\text{ch}} = \frac{q}{R}\left[1 - \exp(-\alpha R)\left(1 + \frac{\alpha R}{2}\right)\right], \tag{8a}$$

$$V_2^{\text{ch}} = \frac{q}{R}[1 - \exp(-\alpha R)]. \tag{8b}$$

The damped potentials of dipole and higher multipoles can be evaluated by expanding the damped charge potential as a Taylor series, similar to the procedure for obtaining classical multipole potentials. Consider the damped potential due to molecule A located at position **A** with particles of this molecule (a) at positions **a** relative to **A** (see Fig. 2). The potential is calculated at point **B**, separated from **A** by $\mathbf{R} = \mathbf{B} - \mathbf{A}$. The damped potential in model 2 can be expressed as

$$\begin{aligned} V_2^{\text{A}}(\mathbf{B}) &= \sum_a \frac{e_a}{4\pi\varepsilon_0|\mathbf{R}-\mathbf{a}|}(1 - \exp(-\alpha|\mathbf{R}-\mathbf{a}|)) = \sum_a \frac{e_a}{4\pi\varepsilon_0} \\ &\times \left[\begin{array}{l} \dfrac{1-\exp(-\alpha R)}{R} + a_\alpha\left(\dfrac{\partial}{\partial a_\alpha}\dfrac{(1-\exp(-\alpha|\mathbf{R}-\mathbf{a}|))}{|\mathbf{R}-\mathbf{a}|}\right)_{\mathbf{a}=0} \\ \quad + \dfrac{1}{2}a_\alpha a_\beta\left(\dfrac{\partial^2}{\partial a_\alpha \partial a_\beta}\dfrac{(1-\exp(-\alpha|\mathbf{R}-\mathbf{a}|))}{|\mathbf{R}-\mathbf{a}|}\right)_{\mathbf{a}=0} + \ldots \end{array}\right] \\ &= \sum_a \frac{e_a}{4\pi\varepsilon_0}\left[\frac{1-\exp(-\alpha R)}{R} - a_\alpha \nabla_\alpha \frac{1-\exp(-\alpha R)}{R}\right. \\ &\qquad \left. + \frac{1}{2}a_\alpha a_\beta \nabla_\alpha \nabla_\beta \frac{1-\exp(-\alpha R)}{R} - \ldots\right] \\ &= \frac{1}{4\pi\varepsilon_0}\left[q\frac{1-\exp(-\alpha R)}{R} - \hat{\mu}_\alpha \nabla_\alpha \frac{(1-\exp(-\alpha R))}{R}\right. \\ &\qquad \left. + \frac{1}{3}\hat{\Theta}_{\alpha\beta}\nabla_\alpha\nabla_\beta\frac{(1-\exp(-\alpha R))}{R} - \ldots\right]. \end{aligned} \tag{9}$$

Subscripts α and β in the last equation $= x$, y, or z, and ε_0 is the permittivity of free space. In atomic units $4\pi\varepsilon_0$ is unity. Therefore, within model 2, the damped dipole potential is

$$\begin{aligned} V_2^{\text{dip}} &= \hat{\mu}_\alpha \nabla_\alpha\left(\frac{1}{R}[1 - \exp(-\alpha R)]\right) \\ &= -\hat{\mu}_\alpha \frac{R_\alpha}{R^3}[1 - \exp(-\alpha R)(1 + \alpha R)], \end{aligned} \tag{10}$$

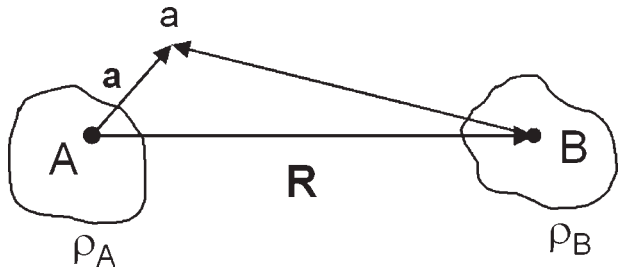

Figure 2. Definition of position vectors in two interacting molecules. (See text for explanations.)

and the damped quadrupole potential is

$$V_2^{\text{quad}} = \hat{\Theta}_{\alpha\beta}\nabla_\alpha\nabla_\beta\left(\frac{1}{R}[1-\exp(-\alpha R)]\right) = \hat{\Theta}_{\alpha\beta}\left\{\frac{3R_\alpha R_\beta - R^2\delta_{\alpha\beta}}{R^5}[1-\exp(-\alpha R)(1+\alpha R)] - \frac{R_\alpha R_\beta}{R^5}\left[\exp(-\alpha R)(\alpha R)^2\right]\right\}. \quad (11)$$

Note that the dipole potential has a single damping term, $\exp(-\alpha R)(1+\alpha R)$, which is applied to the whole classical dipole potential. That is, the damping function for the dipole potential is spherically symmetric, and the dipole potential is damped equally in every direction. This is not the case for the quadrupole potential, for which there is an additional term, $\exp(-\alpha R)(\alpha R)$,[2] that acts only on a part of the quadrupole potential. Thus, the damping function for the quadrupole potential is stronger in some directions than in others. This is also true for the higher-order multipoles. Note also that the angular dependence in quadrupole damping is produced by the largest magnitude (leading) term of the damping function. It is interesting that even though different order multipole potentials have different power dependences on $1/R$, the leading terms of the corresponding damping functions (within model 2) are always proportional to $1/R^*\exp(-\alpha R)$. Thus, the damping functions for different order multipoles fall off similarly with increase in R.

Following the same procedure, one can obtain damped dipole and quadrupole potentials within model 1:

$$V_1^{\text{dip}} = \hat{\mu}_\alpha\nabla_\alpha\left(\frac{1}{R}[1-\exp(-\alpha R)(1+1/2\alpha R)]\right) = -\hat{\mu}_\alpha\frac{R_\alpha}{R^3}\left[1-\exp(-\alpha R)(1+\alpha R+1/2(\alpha R)^2)\right], \quad (12)$$

$$V_1^{\text{quad}} = \hat{\Theta}_{\alpha\beta}\nabla_\alpha\nabla_\beta\left(\frac{1}{R}[1-\exp(-\alpha R)(1+1/2\alpha R)]\right) = \hat{\Theta}_{\alpha\beta}\left\{\frac{3R_\alpha R_\beta - R^2\delta_{\alpha\beta}}{R^5}[1-\exp(-\alpha R)(1+\alpha R+1/2(\alpha R)^2)] - \frac{R_\alpha R_\beta}{R^5}\left[\exp(-\alpha R)1/2(\alpha R)^3\right]\right\}. \quad (13)$$

As expected, the dipole and quadrupole damping functions are different for models 1 and 2. However, the general forms of the dipole and quadrupole damped potentials in the two models have some similarities. As in model 2, the dipole damping in model 1 is spherical, and the non-spherical term appears only in the quadrupole potential. Similar to model 2, the leading damping term in all multipoles is proportional to $\exp(-\alpha R)$.

Adding a damping term to the classical electrostatic potential assumes that the new potential is produced by a 3D electronic density *distribution*, rather than by a series of classical point multipoles. However, to maintain an analogy with the classical approach, one can expand the electron density in a series of densities that correspond to the multipoles:

$$\rho = \rho^{\text{ch}} + \rho^{\text{dip}} + \rho^{\text{quad}} + \rho^{\text{oct}} + \cdots. \quad (14)$$

Each term in this expansion can be determined from the corresponding multipole potential through the Poisson equation:

$$\nabla^2 V^{\text{multipole}} = -\frac{\rho^{\text{multipole}}}{\varepsilon_0}. \quad (15)$$

For example, the charge densities in model 1 and 2 are, respectively:

$$\rho_1^{\text{ch}} = \frac{q\alpha^3}{8\pi}\exp(-\alpha R), \quad (16a)$$

$$\rho_2^{\text{ch}} = \frac{q\alpha^2}{4\pi R}\exp(-\alpha R). \quad (16b)$$

Finding the density corresponding to a known potential is straightforward, because there is no need to solve the differential Poisson equation. One needs only to differentiate the known potential.

To calculate the interaction energy between two centers, A and B, one can use the Coulomb equation in the following form:

$$E_{\text{Cond}} = \int\frac{\rho_A\rho_B}{R_{AB}}\,dv_A dv_B = \int\rho_A V_B(\mathbf{A})\,dv = \frac{1}{2}\left(\int\rho_A V_B(\mathbf{A})\,dv_A + \int\rho_B V_A(\mathbf{B})\,dv_B\right)_A. \quad (17)$$

Substituting in the last equation the different terms of the multipolar potential and density, one can obtain different components of the electrostatic energy, e.g.

$$E^{\text{ch-ch}} = \frac{1}{2}\left(\int\rho_A^{\text{ch}}V_B^{\text{ch}}(\mathbf{A})\,dv_A + \int\rho_B^{\text{ch}}V_A^{\text{ch}}(\mathbf{B})\,dv_B\right), \quad (18a)$$

$$E^{\text{ch-dip}} = \frac{1}{2}\left(\int\rho_A^{\text{ch}}V_B^{\text{dip}}(\mathbf{A})\,dv_A + \int\rho_B^{\text{dip}}V_A^{\text{ch}}(\mathbf{B})\,dv_B\right). \quad (18b)$$

Integrals that appear in the expressions for the charge–charge electrostatic energy are analytic in both models; their solutions were given by Coulson in 1942.[37] Thus, the charge–charge interaction energies for models 1 and 2 are, respectively:

$$E_1^{\text{ch-ch}} = \frac{q_A q_B}{R}\left[\begin{array}{l}1-\exp(-\alpha R)\dfrac{\beta^4}{(\beta^2-\alpha^2)^2}\left(1-\dfrac{2\alpha^2}{\beta^2-\alpha^2}+\dfrac{\alpha R}{2}\right)\\ -\exp(-\beta R)\dfrac{\alpha^4}{(\alpha^2-\beta^2)^2}\left(1-\dfrac{2\beta^2}{\alpha^2-\beta^2}+\dfrac{\beta R}{2}\right)\end{array}\right], \quad (19a)$$

$$E_2^{\text{ch-ch}} = \frac{q_A q_B}{R}\left[1-\exp(-\alpha R)\frac{\beta^2}{\beta^2-\alpha^2}-\exp(-\beta R)\frac{\alpha^2}{\alpha^2-\beta^2}\right]. \quad (19b)$$

Here, α and β are damping parameters for centers A and B, respectively. To calculate the electron–nuclear attraction, one has to evaluate the interaction between the damped charge potential and the classical charge potential. Alternatively, the electron–nuclear attraction can be found by setting one of the screening parameters, α or β, to infinity (because the nucleus is not damped). The electron–nuclear attraction energies for models 1

and 2 are

$$E_1^{\text{ch-nucl}} = \frac{q_A Z_B}{R}\left(1-\exp(-\alpha R)\left(1+\frac{\alpha R}{2}\right)\right) + \frac{q_B Z_A}{R}\left(1-\exp(-\beta R)\left(1+\frac{\beta R}{2}\right)\right), \quad (20a)$$

$$E_2^{\text{ch-nucl}} = \frac{q_A Z_B}{R}(1-\exp(-\alpha R)) + \frac{q_B Z_A}{R}(1-\exp(-\beta R)). \quad (20b)$$

Therefore, the total charge–charge penetration energies in models 1 and 2 are:

$$E_{\text{pen1}}^{\text{ch-ch}} = -\frac{q_A q_B}{R}\left[\begin{array}{l}\exp(-\alpha R)\dfrac{\beta^4}{(\beta^2-\alpha^2)^2}\left(1-\dfrac{2\alpha^2}{\beta^2-\alpha^2}+\dfrac{\alpha R}{2}\right)\\ +\exp(-\beta R)\dfrac{\alpha^4}{(\alpha^2-\beta^2)^2}\left(1-\dfrac{2\beta^2}{\alpha^2-\beta^2}+\dfrac{\beta R}{2}\right)\end{array}\right], \\ -\frac{q_A Z_B}{R}\exp(-\alpha R)\left(1+\frac{\alpha R}{2}\right)-\frac{q_B Z_A}{R}\exp(-\beta R)\left(1+\frac{\beta R}{2}\right). \quad (21a)$$

$$E_{\text{pen2}}^{\text{ch-ch}} = -\frac{q_A q_B}{R}\left[\exp(-\alpha R)\frac{\beta^2}{\beta^2-\alpha^2}+\exp(-\beta R)\frac{\alpha^2}{\alpha^2-\beta^2}\right] - \frac{q_A Z_B}{R}\exp(-\alpha R)-\frac{q_B Z_A}{R}\exp(-\beta R). \quad (21b)$$

When the screening parameters α and β are equal, these equations have poles at $\alpha = \beta$. One can reevaluate the charge–charge integrals by setting $\alpha = \beta$, or calculate the limit of $\alpha \to \beta$. This leads to

$$E_1^{\text{ch-ch}} = \frac{q_A q_B}{R}\left[1-\exp(-\alpha R)\left(1+\frac{11}{16}\alpha R + \frac{3}{16}(\alpha R)^2+\frac{1}{48}(\alpha R)^3\right)\right] \quad (22a)$$

$$E_2^{\text{ch-ch}} = \frac{q_A q_B}{R}\left[1-\exp(-\alpha R)\left(1+\frac{1}{2}\alpha R\right)\right]. \quad (22b)$$

The total charge-penetration energies will change accordingly. Formulas for the charge–charge electrostatic and charge-penetration energies for model 2 were published previously in Ref. 12, where it was also shown that damping of only the charge–charge term gives reasonable results.

Damping of the higher-order terms was not included in the previous implementation, because the integrals required to evaluate the charge-penetration corrections to the higher terms of the electrostatic energy, i.e., charge–dipole, dipole–dipole, etc. [see eq. (18b)], do not have analytic solutions, and numerical analysis of these integrals was considered to be too CPU-intensive.

Here, an alternative approach is suggested for calculating the higher-order damping terms. Consider the interaction of a charge Q with a dipole μ (see Fig. 3). The charge corresponds to an electron density given by eqs. (16a) or (16b). If one assumes that the dipole is a system of two charges q^+ and q^- separated by a finite length d such that $q^+ d = \mu$, and that each of the two charges that produce the dipole corresponds to a density with the same form as that corresponding to charge Q, one can calculate the charge–dipole energy as the sum of two charge–charge energies in the limit of $d \to 0$. For the configuration shown in Figure 3 this will take the form:

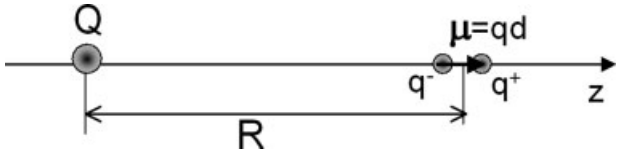

Figure 3. Interaction of charge Q and dipole μ.

$$\begin{aligned} E^{\text{ch-dip}} &= E^{Qq^-} + E^{Qq^+} = Qq^- f\left(R-\frac{d}{2}\right)+Qq^+ f\left(R+\frac{d}{2}\right) \\ &= \frac{Q\mu}{d}\left(-f\left(R-\frac{d}{2}\right)+f\left(R+\frac{d}{2}\right)\right) \\ &= Q\mu\frac{f\left(R+\frac{d}{2}\right)-f\left(R-\frac{d}{2}\right)}{d} = Q\mu\,\frac{\partial f(R)}{\partial R_z} \\ &= Q\mu\nabla_\alpha f(R), \end{aligned} \quad (23)$$

where $f(R)$ is the R-dependent part of the charge–charge electrostatic energy in models 1 or 2, see eqs. (19a), (19b), (22a), and (22b). This result shows that the charge–dipole energy can be obtained by differentiating the corresponding charge–charge energy. This is general for any orientation of charge and dipole. Using the same approach, one can show that the charge–quadrupole and dipole–dipole energies are second derivatives of the charge–charge energy, the charge–octopole energy is the third derivative, and so on. This can also be shown mathematically based on the fact that the dipole potential is a derivative of the charge-potential. The same is true for the densities, since they are connected to the potentials by the Poisson equation. The charge–dipole and dipole–dipole charge-penetration energies are

$$E_{\text{pen1}}^{\text{ch-dip}} = q_A\mu_B\frac{R_\alpha}{R^3}\left[\exp(-\alpha R)\frac{\beta^4}{(\beta^2-\alpha^2)^2} \times\left(\left(1-\frac{2\alpha^2}{\beta^2-\alpha^2}\right)(1+\alpha R)+\frac{(\alpha R)^2}{2}\right)+g(\beta R)\right] + Z_A\mu_B\frac{R_\alpha}{R^3}\exp(-\alpha R)\left(1+\alpha R+\frac{(\alpha R)^2}{2}\right) \quad (24)$$

$$E_{\text{pen1}}^{\text{dip-dip}} = -\mu_A\mu_B\frac{3_1 R_\alpha R_\beta - R^2\delta_{\alpha\beta}}{R^5} \times\left[\begin{array}{c}\exp(-\alpha R)\dfrac{\beta^4}{(\beta^2-\alpha^2)^2}\\ \left(\left(1-\frac{2\alpha^2}{\beta^2-\alpha^2}\right)(1+\alpha R)+\frac{(\alpha R)^2}{2}\right)+g(\beta R)\end{array}\right] - \mu_A\mu_B\frac{R_\alpha R_\beta}{R^5}\left[\exp(-\alpha R)\frac{\beta^4}{(\beta^2-\alpha^2)^2} \times\left(-\frac{2\alpha^2}{\beta^2-\alpha^2}(\alpha R)^2+\frac{(\alpha R)^3}{2}\right)+g(\beta R)\right] \quad (25)$$

$$E_{\mathrm{pen2}}^{\mathrm{ch-ch}} = q_A \mu_B \frac{R_\alpha}{R^3}\left[\exp(-\alpha R)\frac{\beta^2}{\beta^2-\alpha^2}(1+\alpha R) + g(\beta R)\right] + Z_A \mu_B \frac{R_\alpha}{R^3}\exp(-\beta R)(1+\beta R) \quad (26)$$

$$E_{\mathrm{pen2}}^{\mathrm{dip-dip}} = -\mu_A \mu_B \frac{3R_\alpha R_\beta - R^2\delta_{\alpha\beta}}{R^5} \times \left[\exp(-\alpha R)\frac{\beta^2}{\beta^2-\alpha^2}(1+\alpha R) + g(\beta R)\right] - \mu_A \mu_B \frac{R_\alpha R_\beta}{R^5}\left[\exp(-\alpha R)\frac{\beta^2}{\beta^2-\alpha^2}(\alpha R)^2 + g(\beta R)\right]. \quad (27)$$

The functions $g(\beta R)$ in eqs. (24)–(27) are damping functions with a parameter β; they can be obtained by interchanging α and β in the corresponding α-damping functions. The R-dependent part of the charge–quadrupole energy is the same as in the dipole–dipole energy. Expressions for the charge penetration energies with equal parameters α and β can be easily obtained by finding the $\alpha \to \beta$ limit in eqs. (24)–(27). It was noted earlier that the functional dependence on R in the charge–charge energies is similar to that of the corresponding potentials. This is also true for the higher-order energies.

Analytic gradients for all of the damping energy terms described earlier have been implemented. These formulas are easily obtained by differentiating the corresponding energy terms.

For consistency, if only the charge–charge energy term is screened, only the first (charge) part of the potential should be augmented by the damping function. If the charge–dipole damping term is added to the energy, the dipole potential also should be screened. Damping of the quadrupole potential allows introduction of the charge–quadrupole and dipole–dipole damping terms. In the following, two models are considered: (i) only the charge–charge part of the energy and the charge part of the potential are screened, and (ii) the charge–charge, charge–dipole, charge–quadrupole, and dipole–dipole terms of the energy, along with the charge, dipole, and quadrupole potentials are screened.

Computational Details

All calculations were performed with the quantum chemistry program GAMESS.[38]

Running EFP calculations is a two-step procedure. In the first MAKEFP step, the polarization and dispersion points are generated, as is the wave function required for further calculation of the exchange-repulsion term. The analytic or numerical Stone procedure is used to generate the electrostatic multipolar potential. The difference between the multipolar and *ab initio* potentials is minimized by introducing the damping parameter α. One can choose between damping models and levels of damping. The second step is to combine the potential(s) generated by the MAKEFP run(s) for each fragment and give a starting geometry of the molecular cluster of interest.

Optimization [see eq. (6)] of the damping parameters is performed on the grid of points using the Powell minimization algorithm. The performance of three different grids was examined: a cubic grid, geodesic grid,[39] and Connolly grid.[40,41] In the cubic grid, the fragment is placed in a three-dimensional Cartesian grid with user-defined spacing (usually, 0.5 bohr). Two concentric spheres are placed around each atom and the bond mid-point, and any grid points outside these spheres are discarded. The radii of the two spheres determine the minimum and maximum radii of grid. Previously,[12] the recommended values of these parameters for model 2 with charge–charge screening were 67 and 300% of the Van der Waals (vdW) radii of the corresponding atom. In geodesic and Connolly grids, points are placed on a number of spheres of increasing radii with centers at each atom and bond mid-point, starting from the sphere with minimum radius.

Geodesic and Connolly grids have different algorithms for distributing points on the surface of each sphere. In the Connolly grid, the points on each sphere are distributed to approximately preserve the point density, i.e., a sphere of larger radius has more points. The drawback of this approach is that the distribution of points on each sphere is not very uniform. In the geodesic grid, each sphere has the same number of points, distributed based on a tessellation scheme. As a result, the point distribution on each sphere is very uniform, but the angular point density is not preserved. The spacing between the spheres and the total number of spheres in both Connolly and geodesic grids can be varied. Of course, this determines the radial grid density and the maximum radius of the grid.

When the spacing between the points in each grid is small enough, the minimization procedure using any of the three grids results in similar parameters. However, the convergence of these parameters as the number of grid points is increased depends on the type of grid used. For example, water requires $\sim$16,000 points for the cubic grid, $\sim$8000 points for the Connolly grid, and $\sim$3500 points for the geodesic grid. At smaller numbers of points, i.e., at larger spacing, the Connolly and cubic grids produce large fluctuations in the optimized parameters. The instability of cubic and Connolly grids at small point densities occurs because the most important fitting region is near the nuclei where the potential is very strong; a large number of points is required in this region. Because the cubic and Connolly grids have isotropic point distributions over all space, a very large total number of points is required. On the other hand, the geodesic grid has a higher point density at smaller distances from the nuclei and, therefore, is more stable. The stability of the optimized damping parameters with respect to the grid size and density, combined with the smaller computational time, suggests that the geodesic grid is the most appropriate for the present purposes. The parameters of the geodesic grid used throughout this work are 25 spheres with spacing 0.1 vdW between them, where vdW refers to the Van der Waals radius; points on each sphere were distributed by an icosahedral pattern of third order (92 points on a sphere); the minimum radius of the grid (radius of the first sphere) is model-dependent, and is discussed in detail in the next Section.

The geometries of dimers considered in this work were obtained by the following procedure: first, second order perturbation theory (MP2)[42] and the 6-311G(d,p)[43,44] basis set were used to obtain equilibrium geometries of each monomer except benzene. These internal monomer geometries were kept frozen during MP2/6-311++G(d,p)[43–45] optimization of dimer geometries. These dimer geometries are referred to in the following as equilibrium geometries. Exceptions from the general procedure were made for ethylene and benzene dimers. For benzene, MP2/aug-cc-pVTZ[46,47] monomer geometries from Ref. 19 were used. Two

dimer configurations, the sandwich with $R = 3.7$ Å and the T-shaped form with $R = 5.0$ Å, were considered (see Fig. 8). For ethylene dimer, two possible structures were considered: D_{2d}, the global minimum, with $R = 3.8$ Å, and the sandwich structure with $R = 3.8$ Å (see Fig. 8).

EFPs for each monomer discussed in the Performance of Damping Models section were obtained at the monomer equilibrium geometry with the 6-31++G(d,p)[44,45,48] basis set (6-311++G(d,p) basis for water). Then, EFP single point energy calculations were performed at the dimer equilibrium geometries. The Morokuma–Kitaura analysis[49] in the EFP basis was performed for each dimer at its equilibrium geometry. In all cases, the results of Morokuma–Kitaura analysis were consistent with those of the reduced variational space analysis,[50] with differences 0.01 kcal/mol or less. The electrostatic energy from the Morokuma analysis was compared with the EFP electrostatic energy. Differences between the EFP and Morokuma electrostatic energies are analyzed in the next section.

For the benzene calculations in the Benzene Dimer section, the EFP was constructed using the 6-311++G(3df,2p) basis set, at the monomer geometry described earlier. This basis set was chosen, because the correlation consistent basis sets are not appropriate for the Hartree–Fock-based electrostatic and repulsion terms.

Results and Discussion

Performance of Damping Models

In this section the performance of damping models 1 and 2 with different levels of screening are analyzed. The key issues are (i) the relative performance and advantages and disadvantages of models 1 and 2; (ii) the relative importance of charge–charge damping versus high-order damping in models 1 and 2; and (iii) the performance of analytic and numerical DMAs in conjunction with different types of damping. Answering these questions will facilitate general recommendations regarding the evaluation of electrostatic interactions.

The performance of the damping models is analyzed by comparing the EFP electrostatic energy with the electrostatic energy from the Morokuma analysis in several different dimers. It is assumed that the difference between undamped EFP and Morokuma electrostatic energies is due to the charge-penetration:

$$E^{\text{pen}} = E_{\text{ES}}^{\text{multipole}} - E_{\text{ES}}^{\text{Morokuma}}. \tag{28}$$

Damping of the EFP (multipole) electrostatic energy should account for the charge-penetration energy and decrease the difference between EFP and Morokuma electrostatic energies. All figures in this section have the following format: differences between the EFP (damped and not damped) and Morokuma electrostatic energies are plotted versus the intermolecular separation (see Fig. 4) or versus grid parameter r_{min}, at the equilibrium geometries of the dimers (Figs. 5–7).

A comparison of different damping models for water dimer is presented in Figure 4. The abscissa is the relative O—O distance between two water molecules, such that 0 Å represents the equilibrium O—O distance, negative values bring the fragments closer together and positive values move them further apart along the line connecting the oxygen atoms. Figures 4a and 4b show the results based on the analytic and numerical DMAs, respectively.

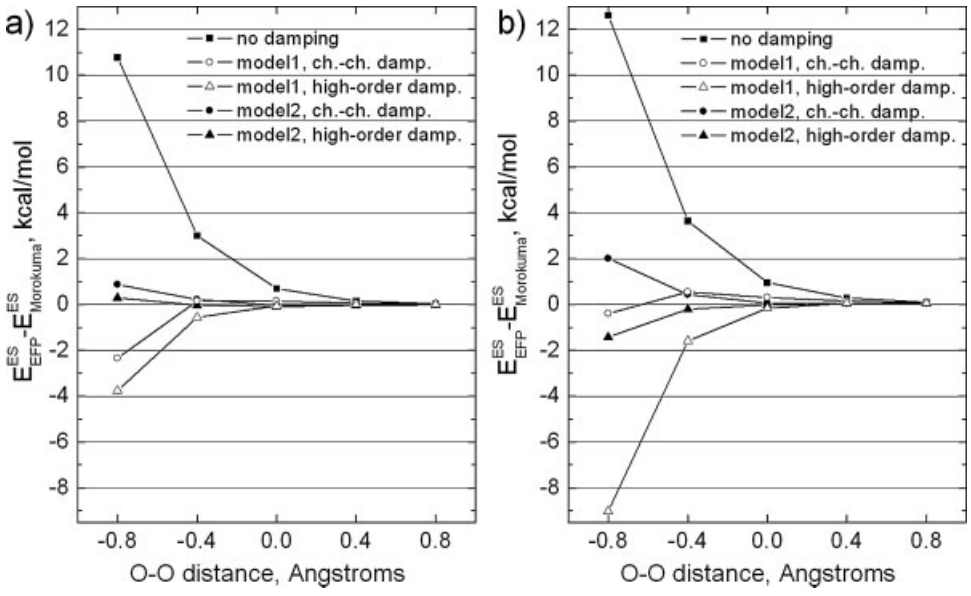

Figure 4. Errors in the electrostatic energy (kcal/mol) in the water dimer as a function of oxygen–oxygen distance. (a) Analytic DMA is used to assign multipole moments; (b) numerical DMA is used. Performance of damping models 1 and 2 with charge–charge only and high-order damping terms is shown. Optimal r_{min} (in parts of vdW) was chosen for each model: 0.5/0.6 for model 1 with charge–charge damping, 0.8/0.8 for model 1 with high-order damping, 0.5/0.5 for model 2 charge–charge damping, and 0.6/0.8 for model 2 high-order damping, where the first and second numbers correspond to analytic and numerical DMAs, respectively.

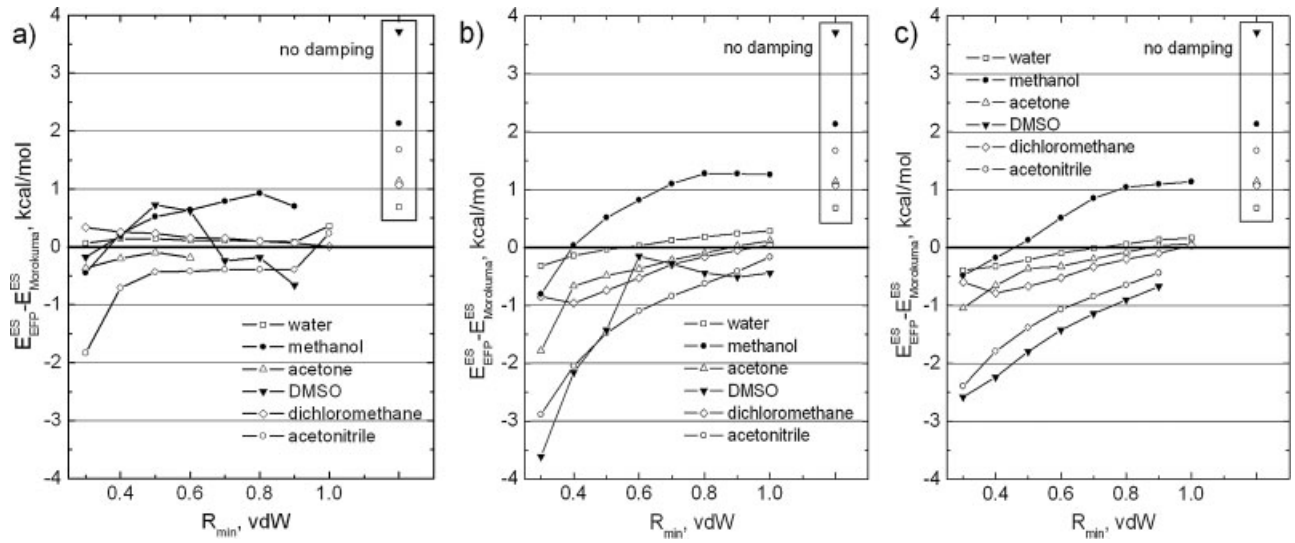

Figure 5. Errors in the electrostatic energies (kcal/mol) calculated using analytic DMA in 6 homomolecular dimers at their equilibrium geometries as a function of r_{min}. (a) Model 1 with charge–charge damping, (b) model 2 with charge–charge damping, and (c) model 2 with high-order damping is used for screening. Electrostatic errors without damping are given in the right hand side of each graph.

The optimal value of r_{min} (see below) was used for each curve. A horizontal straight line would correspond to exact agreement with the Morokuma electrostatic energy. For the analytic DMA, all damping models significantly improve the undamped multipolar energy. As expected, for both models 1 and 2, adding damping due to higher multipoles increases the magnitude of the multipolar electrostatic energy (the corresponding curve lies lower). Both charge–charge and higher-order damping curves of model 2 are nearly parallel to the abscissa, and so the charge-penetration correction is fully captured at all intermolecular distances in the range from -0.8 to $+0.8$ Å from equilibrium. On the other hand, for model 1, both charge–charge and high-order damping curves drop at shorter intermolecular distances, because the electrostatic energy in this region is overestimated; the high-order damping contributes to the overestimation. Similar behavior occurs in the model 1 curves for other dimers, and so the overestimation of the

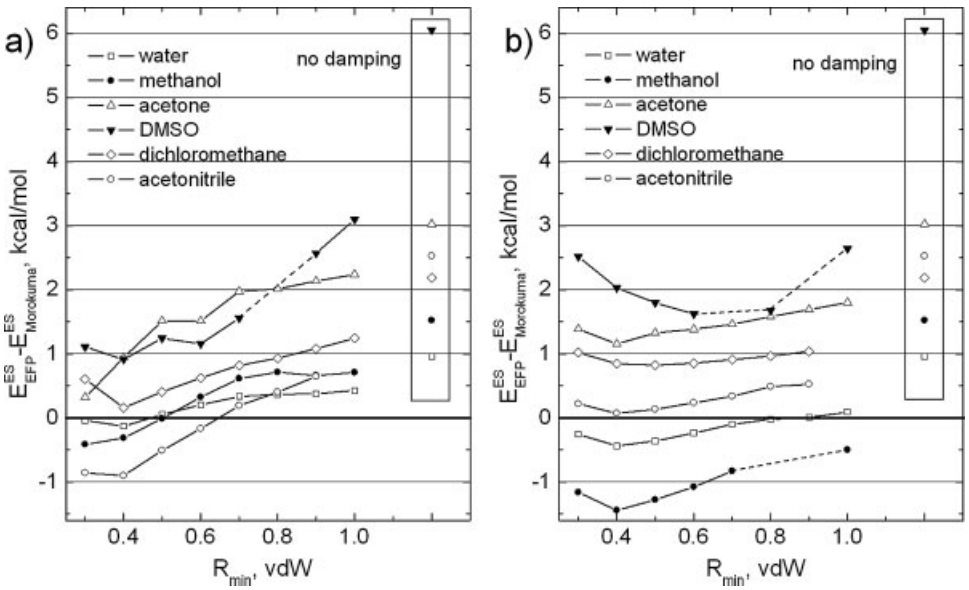

Figure 6. Errors in the electrostatic energies (kcal/mol) calculated using numerical DMA in 6 homomolecular dimers at their equilibrium geometries as a function of r_{min}. Model 2 with (a) charge–charge damping and (b) high-order damping was used for screening. Electrostatic errors without damping are given in the right part of each graph.

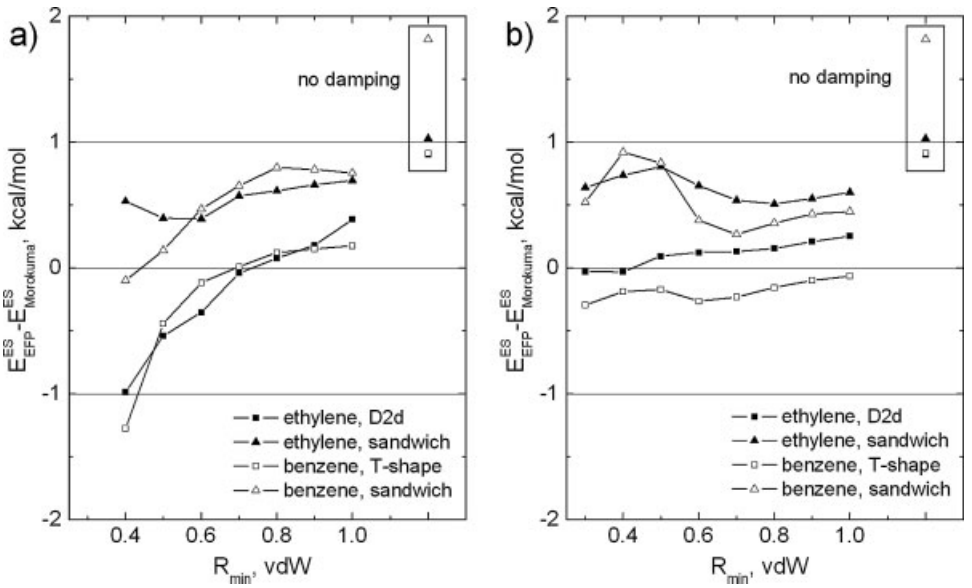

Figure 7. Errors in the electrostatic energies (kcal/mol) calculated using numerical DMA in 2 structures of the ethylene and benzene dimers as a function of r_{min}. Model 2 with (a) charge–charge damping and (b) high-order damping was used for screening. Electrostatic errors without damping are given in the right part of each graph.

electrostatic energy at short intermolecular distances is general for model 1. This probably means that the damping function $\int_1^{damp} = 1 - \exp(-\alpha R)\left(1 + \frac{\alpha R}{2}\right)$ overestimates damping near the nuclei, and that the correct curve for many molecules lies somewhere in between f_1 and f_2 (see Fig. 1).

Results for the numerical DMA, shown in Figure 4b, are less encouraging. The model 1 high-order damping curve strongly overestimates the electrostatic energy at short intermolecular distances. Model 2 curves deviate more from the correct (Morokuma) value than for the analytic DMA. The model 1 charge–charge damping curve is nearly parallel to the abscissa. The undamped curve rises more sharply and higher than for the analytic DMA. Thus, the numerical multipoles produce larger errors with respect to the reference Morokuma results. The worse performance of the damping models for the numerical DMA occurs because of intrinsic problems in the numerical analysis, as will be discussed in some detail later.

Figure 5 compares the performance of three damping models based on the analytic DMA for 6 typical solvent dimers: water, methanol, acetone, DMSO, dichloromethane, and acetonitrile. Each dimer is considered at its equilibrium geometry. The abscissa of each graph is the grid parameter r_{min} in fractions of the vdW radius of each atom; r_{min} determines the radius of the first sphere with grid points. As discussed in the Electrostatic Screening section, the results of model 2 are expected to be sensitive to this parameter. The results of model 1 are expected to be less sensitive to the choice of r_{min}. This is illustrated by the model 1 charge–charge damping results, shown in Figure 5a: there is only mild dependence on r_{min} for water, dichloromethane, and acetonitrile. The methanol and DMSO curves are still not parallel to the abscissa, with the DMSO curve being very irregular. Severe convergence problems occurred for acetone for r_{min} values $\gg$ 0.6 vdW. In general, the results shown in Figure 5a suggest that for model 1 with charge–charge damping, any value of r_{min} in the range 0.5 to 0.9 vdW can be used to obtain damping parameters. The damping procedure recovers 70–90% of the charge-penetration energy at equilibrium distances. As discussed for the water dimer, model 1 overestimates the electrostatic energy at shorter distances. Given these results, model 1 with high-order damping is not considered here.

Model 2 charge–charge and high-order damping curves are shown in Figures 5b and 5c. For all dimers, the errors in the electrostatic energy depend on r_{min}. In general, for dimers with stronger electrostatic interaction, and, consequently, shorter intermolecular separations, such as methanol and water, smaller values of r_{min} are preferable. For weakly interacting molecules, e.g., dichloromethane and acetone, larger r_{min} values are more appropriate. Higher order damping produces slightly more consistent results, with generally larger optimal r_{min} values. However, comparison of the two sets of curves does not exhibit significant improvement when higher order damping terms are included. This is most likely due to the fact that the analytic Stone DMA overestimates charges on atoms and underestimates higher order moments. Therefore, damping is dominated by the contribution of the charges.

Based on the foregoing, the optimal model 2 r_{min} are 0.4–0.5 vdW for methanol, 0.5–0.8 for water, and 0.7–0.9 for weakly bonded molecules. Note that r_{min} values of 1.0 vDW and higher are likely to result in poorly converged damping parameters and are not recommended. As discussed for model 1, model 2 damping recovers 70–95% of the charge-penetration energy, with absolute errors up to 0.7–0.8 kcal/mol.

Figure 6 shows curves obtained using model 2 and the numerical DMA. In general, the numerical DMA results are worse than those obtained with the analytic DMA. The higher order damping terms (Fig. 6b) seem to be especially troublesome. For some molecules, only 50% of the penetration energy is recovered, with maxi-

mum errors as large as 1.5 kcal/mol. It is interesting to examine the differences between the EFP electrostatic energies without damping and the Morokuma electrostatic energies, given in the vertical box on the right side of each graph. Recall that these differences are the charge penetration energies that must be corrected by the damping. The numerical and analytic charge penetration energies are rather different from each other. For example, these energies are, respectively, 6.0 and 3.4 kcal/mol for DMSO, 3.0 and 1.2 kcal/mol for acetone, and 1.5 and 2.2 kcal/mol for methanol. Thus, the damping correction clearly has some contribution from some intrinsic DMA inaccuracy. If one subtracts from all curves in Figures 6a and 6b, the difference between numerical and analytic undamped energies, the resulting damped curves look similar to those in Figures 5b and 5c. It is likely that the numerical DMA is not fully converged with respect to the multipole expansion (through octopoles) and the corresponding energy terms (through quadrupole–quadrupole). As discussed earlier, using different parameters in the numerical DMA results in different electrostatic energies, which can be favorable or unfavorable for a given molecule. Thus, the poor performance of the damping models with the numerical DMA is due to problems in the numerical DMA rather than the damping model. So, using the numerical DMA for calculating the electrostatic energies is possible but should be done with caution. If used, $r_{min} = 0.5$ is recommended for the charge–charge model 2 and $r_{min} = 0.7$–0.8 is recommended for model 2 with higher order damping.

Finally, consider two very weakly bound systems, ethylene dimer and benzene dimer. While the dimers discussed earlier are dipole dominant, ethylene and benzene are quadrupole molecules and their electrostatic interaction is very weak. So, electrostatic damping in these molecules may have different patterns than those observed for the dipole molecules.

Figure 7 presents the results for the T-shaped and sandwich structures of benzene dimer, and the D_{2d} and sandwich structures of ethylene dimer. The numerical DMA was employed with model 2 charge–charge (Fig. 7a) and higher-order (Fig. 7b). As explained in the Numerical DMA section, the numerical DMA was used for benzene because the analytic DMA diverges. The sandwich structures in ethylene and benzene dimers and the corresponding D_{2d} and T-shaped structures are expected to have similar bonding patterns. In comparison to benzene dimer, the ethylene sandwich structure used in these calculations has $R = 3.8$ A, similar to $R = 3.7$ A in the benzene sandwich. However, in contrast to the benzene dimer, the potential of the sandwich configuration of the ethylene dimer is unbound.[51] So, the ethylene and benzene electrostatic terms, with and without damping should also be similar. Indeed, the corresponding ethylene and benzene curves in Figure 7 are very similar. The D_{2d} and T-shaped structures have very similar charge-penetration energies, and they are improved by damping in a similar manner.

The penetration energy in the benzene sandwich structure is nearly twice as large as the penetration energy in sandwich ethylene, because the former has more interacting centers. Nonetheless, the damped curves for ethylene and benzene sandwich structures are similar. The screened electrostatic energy is underestimated in the sandwich structures. As a result, the electrostatic term will underestimate binding in the sandwich compounds. For ethylene and benzene, and probably, other quadrupole molecules, $r_{min} = 0.7$–0.9 vdW, with higher order damping is recommended.

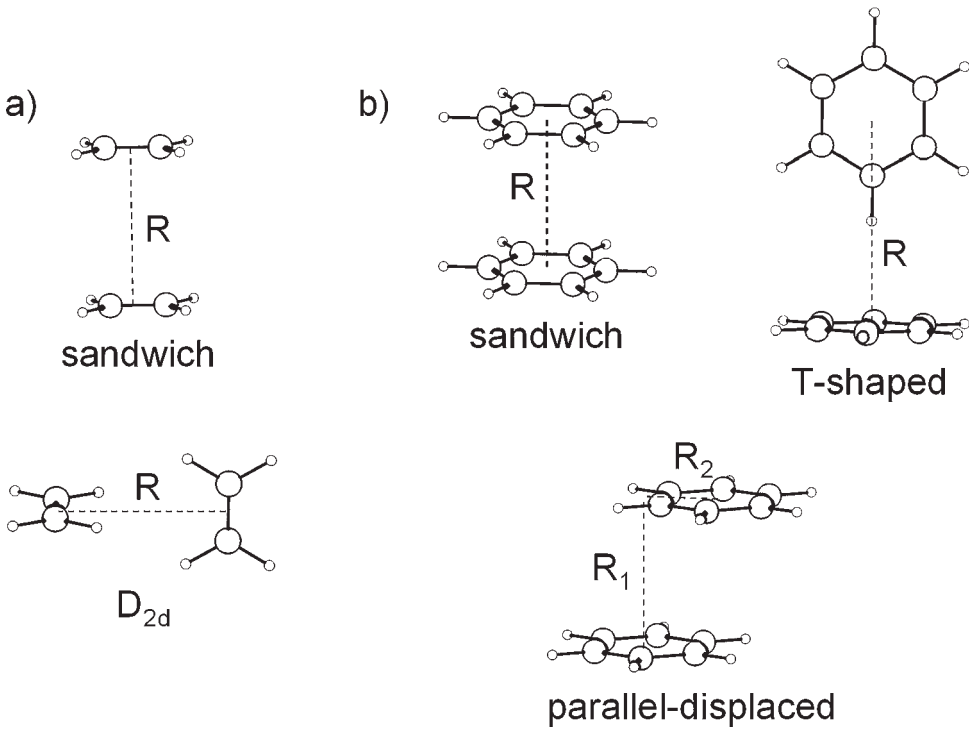

Figure 8. (a) Sandwich and D_{2d} configurations of the ethylene dimer. (b) Sandwich, T-shaped, and parallel-displaced configurations of the benzene dimer.

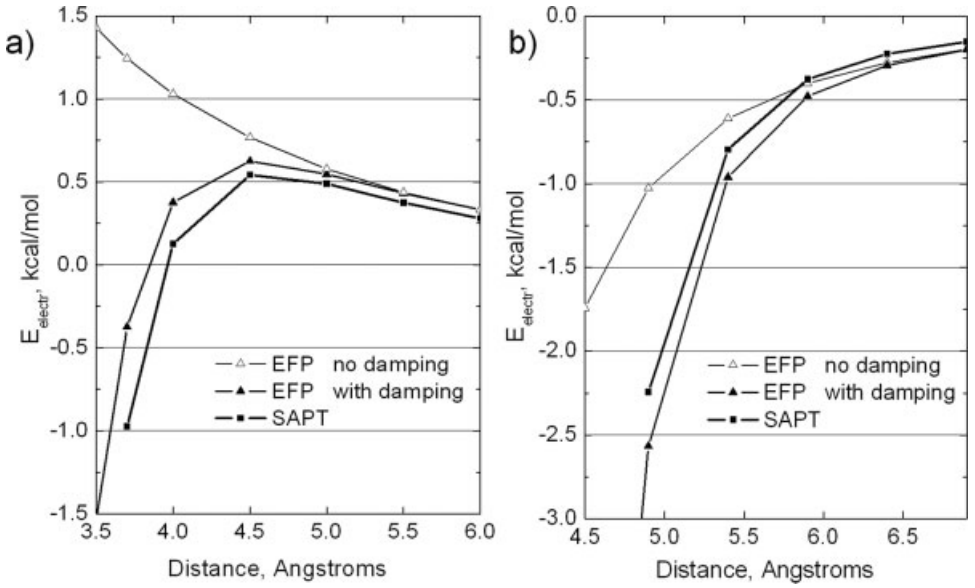

Figure 9. Comparison of the EFP and SAPT electrostatic energies (kcal/mol) in the benzene dimer for (a) sandwich configuration and (b) T-shaped configuration. Both damped and undamped EFP electrostatic curves are shown. SAPT energies from Ref. 18.

Benzene Dimer

The benzene dimer, the simplest among π-stacking systems, is both fundamental and very challenging to theory and experiment due to its small binding energy (2–3 kcal/mol). Combined theoretical and experimental studies suggest that there are two or three minima on the potential energy surface of the benzene dimer: The perpendicular T-shaped and parallel-slipped configurations, with the sandwich structure being highest in energy (see Fig. 8). A rotational experiment by Arunan and Gutowsky[23] determined a 4.96 Å separation between the benzene centers of mass in the T-shaped configuration. The binding energy of the dimer was determined to be $D_0 = 1.6 \pm 0.2$ kcal/mol by Krause et al.[28] and as 2.4 ± 0.4 kcal/mol by Grover et al.[26]

Accurate *ab initio* calculations for the benzene dimer require using both an extensive basis set with diffuse functions and a high level of dynamic correlation. The most comprehensive analysis of the benzene dimer has been presented by Sherrill and co-workers.[18,19] They combined MP2 energies using the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ bases with energies obtained from coupled-cluster singles and doubles with perturbative triple corrections [CCSD(T)][52] in the aug-cc-pVDZ basis to estimate CCSD(T)/aug-cc-pVQZ potential energy curves for the dimer. They also analyzed the nature of the π-π interactions in the T-shaped and sandwich configurations of the dimer using symmetry-adapted perturbation theory (SAPT).[53] SAPT partitions the total interaction energy of a complex into electrostatic, exchange-repulsion, polarization, and dispersion contributions. So, a qualitative analysis of the quality of the EFP results may be assessed by comparing EFP and SAPT/aug-cc-pVDZ potential curves for each energy term separately, and by comparing the total EFP potential energy curves for the benzene dimer with the CCSD(T)/aug-cc-pVQZ curves from Ref. 18. Since the SAPT/aug-cc-pVDZ relative energies are within 0.2–0.3 kcal/mol of those obtained from CCSD(T)/aug-cc-pVQZ, small EFP errors with respect to SAPT may be interpreted as small errors with respect to CCSD(T)/aug-cc-pVQZ.

Figures 9–12 present comparisons of the EFP and SAPT results for electrostatic, exchange-repulsion, polarization, and dispersion terms, respectively. The left hand side of each figure shows the potential energy curves for the sandwich configuration and the right hand sides give the T-shaped curves. The equilibrium inter-monomer distances (defined in Fig. 8) in the benzene dimer vary from $R = 3.7$ Å to 4.0 Å in the sandwich and from $R = 4.9$ Å to 5.1 Å in the T-shaped configurations depending on the level of theory and basis set used. The CCSD(T)/aug-cc-pVQZ values are $R = 3.9$ Å and $R = 5.0$ Å, respectively. These intermolecular separations are used as the reference values in the following discussion.

To emphasize the importance of electrostatic damping, the electrostatic curves are plotted both with and without damping in Figure 9. Model 2 with $r_{min} = 0.7$ vdW based on the numerical DMA was used for damping. For both sandwich and T-shaped structures, the EFP curves without damping underestimate the electrostatic interaction in the equilibrium region by more than 1 kcal/mol. The damping correction accounts for most of the charge-penetration energy, so that the damped EFP energies differ from the SAPT values by 0.2–0.3 kcal/mol at the equilibrium geometries. Damping in the sandwich structure underestimates the penetration energy whereas damping in the T-shaped configuration overestimates it. Thus, the equilibrium T-shaped configuration is over-stabilized by about 0.5 kcal/mol with respect to the sandwich geometry. Consequently, equilibrium intermonomer separations will tend to be larger in the sandwich and smaller in the T-shaped structures, because the difference between the EFP and SAPT energies becomes larger at shorter distances. It is necessary to stress the importance of electrostatic damping for a sys-

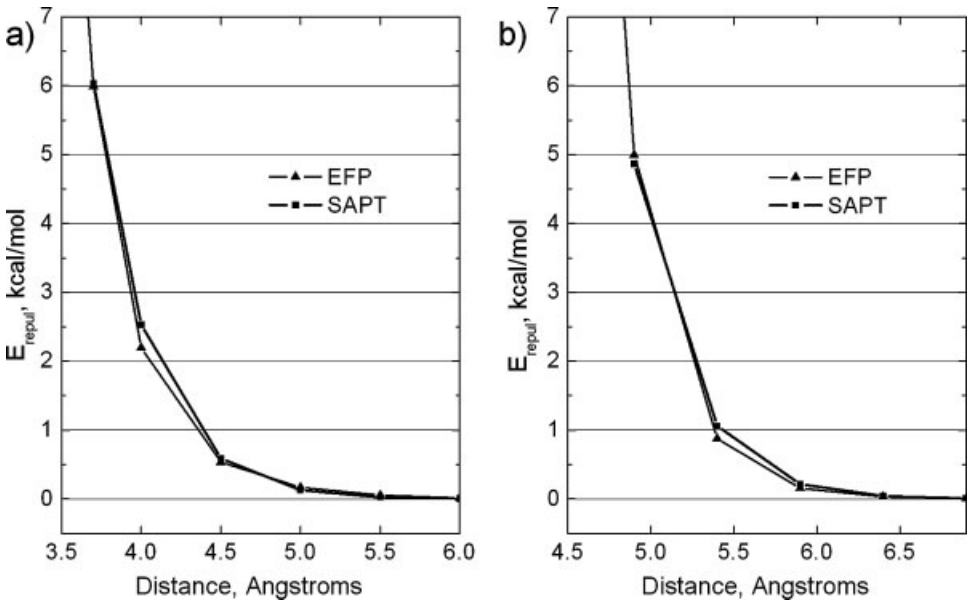

Figure 10. Comparison of the EFP and SAPT exchange-repulsion energies (kcal/mol) in the benzene dimer for (a) sandwich configuration and (b) T-shaped configuration. SAPT energies from Ref. 18.

tem like the benzene dimer, because both equilibrium geometries and binding energies would be significantly in error without the damping term.

The exchange-repulsion energy curves are shown in Figure 10. At the equilibrium sandwich geometry (Fig. 10a), the exchange-repulsion energy is underestimated by about 0.2–0.3 kcal/mol, whereas in the T-shaped structure it is in very good agreement with the SAPT exchange-repulsion energy. The net effect of this is an over-stabilization of the sandwich with respect to the T-shaped configuration by 0.2–0.3 kcal/mol.

The magnitudes of the polarization energies in the benzene dimer are small (see Fig. 11). The EFP polarization over-stabilizes the sandwich structure by 0.1 kcal/mol and under-stabilizes the T-shaped form by 0.3 kcal/mol. The net result is over-binding the sandwich by 0.4 kcal/mol at equilibrium.

The dispersion energy in benzene dimer (see Fig. 12) can be very important, depending on the isomer. The EFP dispersion curve for the sandwich configuration is in good agreement with the SAPT curve in the equilibrium region, and the curve for the T-shaped configuration underestimates the SAPT curve by only 0.3–

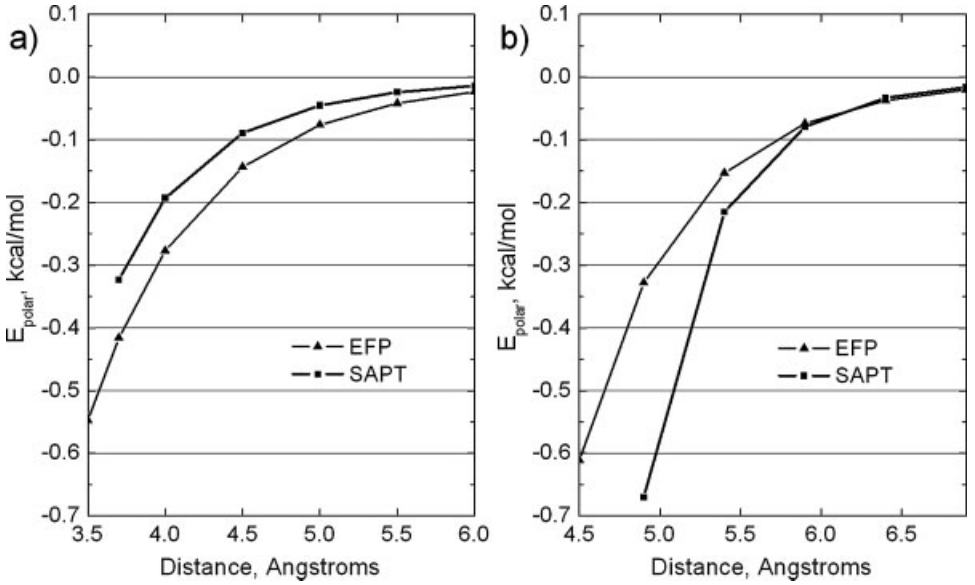

Figure 11. Comparison of the EFP and SAPT polarization energies (kcal/mol) in the benzene dimer for (a) sandwich configuration and (b) T-shaped configuration. SAPT energies from Ref. 18.

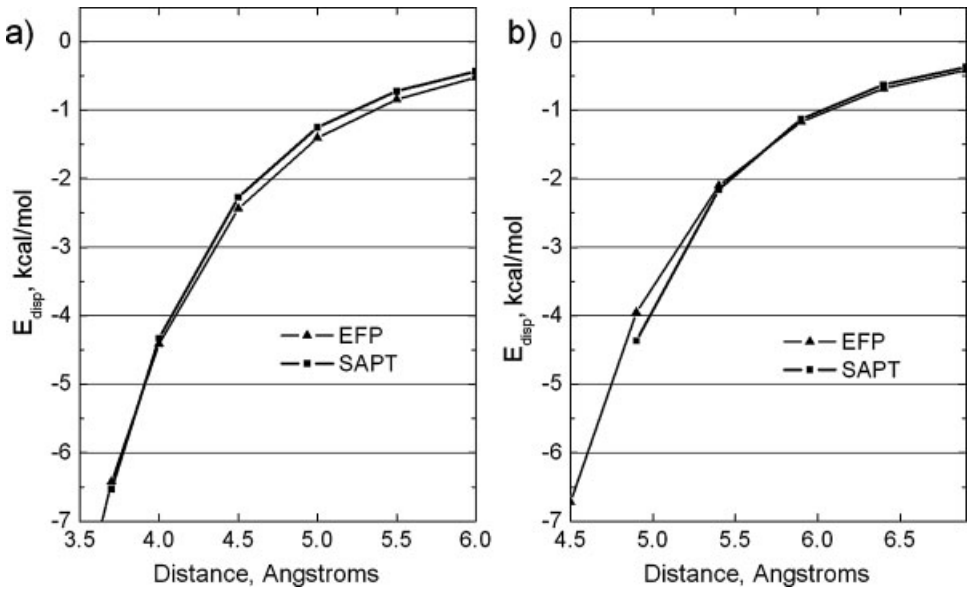

Figure 12. Comparison of the EFP and SAPT dispersion energies (kcal/mol) in the benzene dimer for (a) sandwich configuration and (b) T-shaped configuration. SAPT energies from Ref. 18.

0.4 kcal/mol. Similar to the polarization term, the EFP dispersion predicts the sandwich configuration to be too low by 0.3–0.4 kcal/mol with respect to the T-shaped dimer. In both dimer configurations the EFP dispersion curves are too shallow. This may be due to the omission of higher order terms in the dispersion energy.

Based on the preceding analysis of each energy term that contributes to the EFP interaction energy, one can conclude that the net effect of the EFP errors, relative to SAPT will be to over-bind the sandwich structure with respect to the T-shaped dimer by about 0.5 kcal/mol. This is indeed demonstrated by Figure 13, in which the total binding energies of both structures are plotted. Compared to CCSD(T), EFP over-binds the sandwich dimer by about 0.4 kcal/mol and under-binds the T-shaped structure by 0.1 kcal/mol. The equilibrium intermolecular separations calculated by EFP are 0.1–0.2 Å longer than those calculated by CCSD(T).

Figure 14 shows the potential energy curves for the parallel-displaced configuration with $R_1 = 3.4$, 3.6, and 3.8 Å, calculated by EFP and CCSD(T)/aug-cc-pVQZ. At the CCSD(T) equilibrium inter-monomer separation of 3.6/1.6 Å, the EFP and CCSD(T) binding energies differ by 0.8 kcal/mol. The EFP equi-

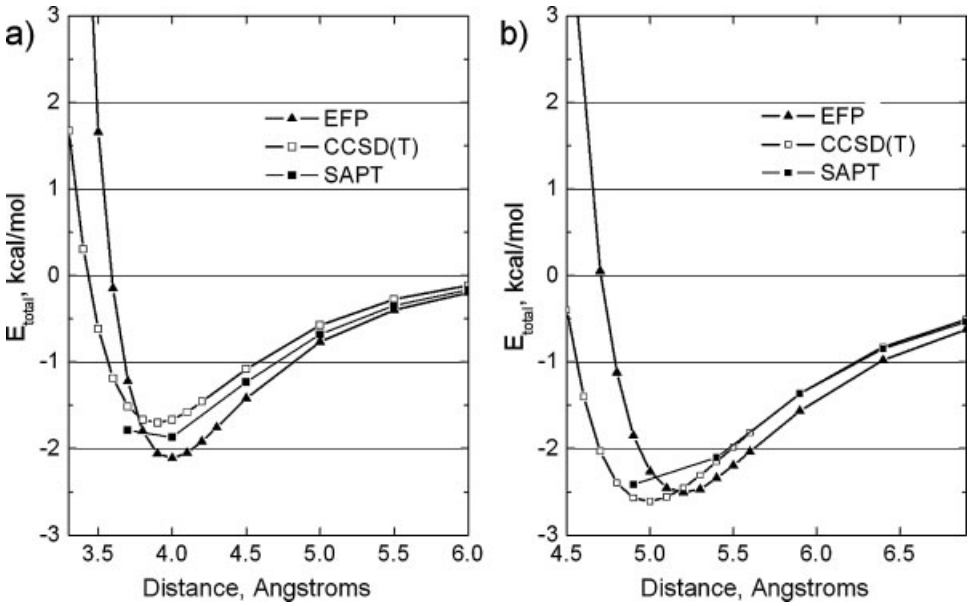

Figure 13. EFP, SAPT, and CCSD(T) potential energy curves (kcal/mol) for the (a) sandwich configuration and (b) T-shaped configuration of the benzene dimer. SAPT and CCSD(T) results from Ref. 18.

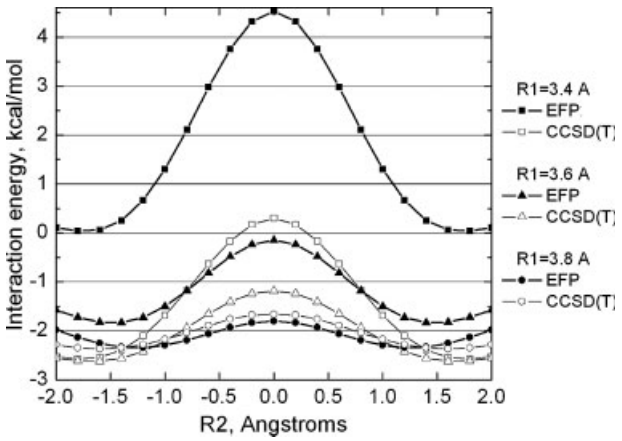

Figure 14. EFP and CCSD(T) potential energy curves (kcal/mol) for the parallel-displaced configuration of the benzene dimer. CCSD(T) results from Ref. 18.

librium geometry is $R_1 = 3.8$ Å, $R_2 = 1.2$ Å, versus $R_1 = 3.6$ Å, $R_2 = 1.6$ Å for CCSD(T). This is not in good agreement as was found for the other two isomers, although the surface here is very flat. In the parallel-displaced configuration, the separation between the monomers is the smallest among the three structures. Since the EFP electrostatic and presumably exchange-repulsion terms underestimate the binding energy at short distances in the sandwich configuration (see Figs. 9a and 10a), this is also probably true for the parallel-displaced structure. Consequently, the EFP potential energy curve in the parallel-displaced dimer with $R_1 = 3.4$ Å lies 2.5–4.0 kcal/mol higher than the corresponding CCSD(T) curve, and the EFP equilibrium geometry is displaced to larger R_1 and smaller R_2 values. However, the EFP curve with $R_1 = 3.8$ Å (equilibrium value) closely follows the corresponding CCSD(T) potential curve.

Table 1 summarizes the interaction energies of the three structures of the dimer calculated by the MP2 and CCSD(T) methods and by EFP. In all structures, relative to CCSD(T) with the same basis set, MP2 underestimates the equilibrium distances by 0.1–0.2 Å and overestimates the binding energies by 0.7–2.1 kcal/mol. The best agreement between MP2 and CCSD(T) is for the T-shaped structure, while the worst is for the parallel-displaced configuration. Comparing EFP with CCSD(T)/aug-cc-pVQZ, EFP overestimates the inter-monomer separations by 0.1–0.2 Å, and inaccuracies in the interaction energies are 0.1–0.4 kcal/mol. In general, the agreement between the EFP and CCSD(T) methods is very reasonable, and overall EFP is in better agreement with CCSD(T) than is MP2. This is striking in view of the orders of magnitude less computer time required by EFP. For example, a single-point energy calculation in the 6-311++G(3df,2p) basis set (660 basis functions) by MP2 requires 142 min of CPU time on one IBM Power5 processor, whereas the analogous EFP calculation requires only 0.4 s.

Conclusions

The charge-penetration energy, i.e., the decrease in the classical electrostatic interaction energy of two multipole expansions due to the overlap of the electron densities, is a significant part of the total electrostatic interaction. Introducing a charge-penetration correction into the EFP method significantly improves the accuracy of equilibrium geometries and binding energies of weakly interacting molecules.

This work investigated two different models for evaluating charge-penetration corrections and the importance of including damping corrections in multipole energy terms of higher order than charge–charge (charge–dipole, dipole–dipole, and charge–quadrupole). The accuracy of the damping models was studied in conjunction with alternative methods for distributing multipoles, i.e., analytic or numerical DMAs. Both damping models work reasonably well, accounting for more than 70% of the charge-penetration energy. Model 1 is more stable with respect to the grid parameter r_{min}, but sometimes experiences difficulties with finding optimal damping parameters. High-order damping within model 1 is not recommended. Electrostatic energies calculated by using model 2 are more dependent on r_{min}, but there are no difficulties with optimizing damping parameters. The analytic DMA produces more satisfactory electrostatic energies than the numerical DMA; the latter is probably not fully converged with respect to the highest multipole moment (octopole) used in this study. However, the numerical DMA can still be used when the analytic DMA breaks down. With

Table 1. Equilibrium Geometries (Å) and Interaction Energies (kcal/mol) for Different Configurations of the Benzene Dimer.

Method	Basis	Sandwich		T-shaped		Parallel-displaced		
		R	Energy	R	Energy	R_1	R_2	Energy
MP2[a]	aug-cc-pVDZ*[b]	3.8	−2.83	5.0	−3.00	3.4	1.6	−4.12
	aug-cc-pVTZ	3.7	−3.25	4.9	−3.44	3.4	1.6	−4.65
	aug-cc-pVQZ*[b]	3.7	−3.35	4.9	−3.48	3.4	1.6	−4.73
CCSD(T)[a]	aug-cc-pVDZ*[b]	4.0	−1.33	5.1	−2.24	3.6	1.8	−2.22
	aug-cc-pVQZ*[b]	3.9	−1.70	5.0	−2.61	3.6	1.6	−2.63
EFP	6-311++G (3df,2p)	4.0	−2.11	5.2	−2.50	3.8	1.2	−2.34

[a]Ref. 18.
[b]Basis sets as described in Ref. 18.

the analytic DMA, higher-order damping corrections in model 2 introduce a minor improvement in electrostatic energies of dipole dimers considered in this work. Using higher-order damping along with the numerical DMA produces less stable results. However, for quadrupole molecules, like ethylene and benzene, using high-order damping is more favorable.

Analysis of bonding in the benzene dimer, a prototype of aromatic π-π interactions, shows that the EFP method can accurately determine equilibrium geometries and interaction energies of this very challenging system. EFP errors with respect to the CCSD(T)/aug-cc-pVQZ method are 0.2 Å and 0.1 kcal/mol for the T-shaped configuration, 0.1 Å and 0.4 kcal/mol for the sandwich configuration, and 0.2/0.4 Å and 0.3 kcal/mol for the most challenging parallel-displaced structure. These errors are in general smaller than the corresponding MP2 errors at a small fraction of the computational cost. Therefore, it is concluded that EFP can be reliably used for studying π-π interactions in other and larger systems.

This study shows that in its present implementation, EFP is a computationally very inexpensive but accurate alternative to MP2. Moreover, the accuracy of EFP can be systematically improved by expanding perturbative series in the electrostatic, polarization, and dispersion terms. The low computational cost of EFP allows Monte-Carlo and molecular dynamic simulations for systems consisting of hundreds of molecules, which opens new horizons in studying intermolecular interactions.

Acknowledgment

The authors are very grateful to Professor David Sherrill for providing his SAPT and CCSD(T) data for benzene dimer.

References

1. Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. J Phys Chem A 2001, 105, 293.
2. Stone, A. J. The Theory of Intermolecular Forces; Oxford University Press: Oxford, 1996.
3. Stone, A. J Chem Phys Lett 1981, 83, 233.
4. Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. J Chem Phys 1996, 105, 1968.
5. Bader, R. Atoms in Molecules; Clarendon Press: Oxford, 1990.
6. Harrison, J. F. J Chem Phys 2003, 119, 8763.
7. Hirshfeld, F. L. Theor Chim Acta 1977, 44, 129.
8. Hirshfeld, F. L. J Mol Struct 1985, 130, 125.
9. Popelier, P. L. Atoms in Molecules: An Introduction; Prentice Hall: Harlow, 2000.
10. Stone, A. J. J Chem Theory Comput 2005, 1, 1128.
11. Kairys, V.; Jensen, J. H. Chem Phys Lett 1999, 315, 140.
12. Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. J Chem Phys 2000, 112, 7300.
13. Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. J Phys Chem A 2003, 107, 10353.
14. Hobza, P.; Selzle, H. L.; Schlag, E. W. J Am Chem Soc 1994, 116, 3500.
15. Hobza, P.; Selzle, H. L.; Schlag, E. W. J Phys Chem 1996, 100, 18790.
16. Jaffe, R. L.; Smith, G. D. J Chem Phys 1996, 105, 2780.
17. Park, Y. C.; Lee, J. S. J Phys Chem A 2006, 110, 5091.
18. Sinnokrot, M. O.; Sherrill, C. D. J Phys Chem A 2004, 108, 10200.
19. Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. J Am Chem Soc 2002, 124, 10887.
20. Spirko, V.; Engvist, O.; Soldan, P.; Selzle, H. L.; Schlag, E. W.; Hobza, P. J Chem Phys 1999, 111, 572.
21. Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. J Am Chem Soc 2002, 124, 104.
22. Tsuzuki, S.; Uchimaru, T.; Sugawara, K.; Mikami, M. J Chem Phys 2002, 117, 11216.
23. Arunan, E.; Gutowsky, H. S. J Chem Phys 1993, 98, 4294.
24. Bornsen, K. O.; Selzle, H. L.; Schlag, E. W. J Chem Phys 1986, 85, 1726.
25. Felker, P. M.; Maxton, P. M.; Schaeffer, M. W. Chem Rev 1994, 94, 1787.
26. Grover, J. R.; Walters, E. A.; Hui, E. T. J Phys Chem 1987, 91, 3233.
27. Janda, K. C.; Hemminger, J. C.; Winn, J. S.; Novick, S. E.; Harris, S. J.; Klemperer, W. J Chem Phys 1975, 63, 1419.
28. Krause, H.; Ernstberger, B.; Neusser, H. J. Chem Phys Lett 1991, 184, 411.
29. Law, K. S.; Schauer, M.; Bernstein, E. R. J Chem Phys 1984, 81, 4871.
30. Scherzer, W.; Kratzschmar, O.; Selzle, H. L.; Schlag, E. W. Z Naturforsch AA J Phys Sci 1992, 47, 1248.
31. Steed, J. M.; Dixon, T. A.; Klemperer, W. J Chem Phys 1979, 70, 4940.
32. Jenson, J. H.; Gordon, M. S. Mol Phys 1996, 89, 1313.
33. Lebedev, V. I.; Laikov, D. N. Dokl Akad Nauk 1999, 366, 741.
34. Murray, C. W.; Handy, N. C.; Laming, G. J. Mol Phys 1993, 78, 997.
35. Slater, J. C. J Chem Phys 1964, 41, 3199.
36. Becke, A. D. J Chem Phys 1988, 88, 2547.
37. Coulson, C. A. Proc Cambridge Philos Soc 1942, 38, 210.
38. 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, 1347.
39. Spackman, M. A. J Comput Chem 1996, 17, 1.
40. Connolly, M. L. J Appl Crystallogr 1983, 16, 548.
41. Besler, B. H.; Merz, K. M.; Kollman, P. A. J Comput Chem 1990, 11, 431.
42. Moller, C.; Plesset, S. Phys Rev 1934, 46, 618.
43. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. J Chem Phys 1980, 72, 650.
44. Hariharan, P. C.; Pople, J. A. Theor Chim Acta 1973, 28, 213.
45. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. J Comput Chem 1983, 4, 294.
46. Dunning, T. H. J Chem Phys 1989, 90, 1007.
47. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. J Chem Phys 1992, 96, 6796.
48. Hehre, W. J.; Ditchfield, R.; Pople, J. A. J Chem Phys 1972, 56, 2257.
49. Kitaura, K.; Morokuma, K. Int J Quantum Chem 1976, 10, 325.
50. Stevens, W. J.; Fink, W. H. Chem Phys Lett 1987, 139, 15.
51. Tsuzuki, S.; Tanabe, K. J Phys Chem 1992, 96, 10804.
52. Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. Chem Phys Lett 1989, 157, 479.
53. Jeziorski, B.; Moszynski, R.; Szalewicz, K. Chem Rev 1994, 94, 1887.