Accurate First Principles Model Potentials for Intermolecular Interactions

Mark S. Gordon,[1] Quentin A. Smith,[1] Peng Xu,[1] and Lyudmila V. Slipchenko[2]

[1]Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50011; email: mark@si.msg.chem.iastate.edu, schrodinator@gmail.com, pxu@iastate.edu

[2]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907; email: lslipchenko@purdue.edu

Annu. Rev. Phys. Chem. 2013. 64:553–78

The *Annual Review of Physical Chemistry* is online at physchem.annualreviews.org

This article's doi: 10.1146/annurev-physchem-040412-110031

Keywords

effective fragment potential, EFP, charge transfer, dispersion, exchange repulsion, QM/EFP

Abstract

The general effective fragment potential (EFP) method provides model potentials for any molecule that is derived from first principles, with no empirically fitted parameters. The EFP method has been interfaced with most currently used ab initio single-reference and multireference quantum mechanics (QM) methods, ranging from Hartree-Fock and coupled cluster theory to multireference perturbation theory. The most recent innovations in the EFP model have been to make the computationally expensive charge transfer term much more efficient and to interface the general EFP dispersion and exchange repulsion interactions with QM methods. Following a summary of the method and its implementation in generally available computer programs, these most recent new developments are discussed.

1. INTRODUCTION

CT: charge transfer

QM: quantum mechanics

MP2: second-order perturbation theory

CC: coupled cluster

MM: molecular mechanics

EFP: effective fragment potential

LMO: localized molecular orbital

Intermolecular interactions play a central role in many areas of chemical, biological, and materials sciences. Perhaps the most obvious example is the impact of solvents on the properties, reactivity, spectroscopy, and dynamics of a solute. Other examples include the following: the theory of liquids and liquid properties, including the mixing (or nonmixing) of different liquids; interfacial phenomena that occur between different phases, such as in membrane processes, electrochemistry, and the lipid bilayer; agostic interactions in organometallic chemistry; π-stacking interactions in the DNA double helix; and polymer aggregation to form clusters.

Depending on the nature of the species involved, there can be several different fundamental origins of intermolecular interactions. These can include Coulombic effects (sometimes called electrostatics), polarization (induction), dispersion (sometimes referred to as van der Waals interactions), exchange repulsion that arises due to the Pauli principle, charge transfer (CT) interactions, and cross terms among these types of phenomena. Commonly, several of these interaction energy components make significant contributions to the net intermolecular interaction energy. Therefore, it is important to have a theoretical method that can capture all of the interaction energy components with an acceptable accuracy. Otherwise, it is difficult to compare the effects of different types of species, for example, polar versus nonpolar solvents.

Of course, quantum mechanics (QM) methods that include electron correlation, such as second-order perturbation theory (MP2) (1) and coupled cluster (CC) theory (2, 3), naturally include all of the aforementioned interaction energy components, and methods like symmetry adapted perturbation theory (4, 5) facilitate the interpretation of fully QM interaction energies in terms of physically meaningful components. However, correlated QM methods are very computationally demanding. Even the most efficient method, MP2, scales $\sim N^5$, where N measures the size of the system (e.g., number of basis functions). So such methods rapidly become intractable for large clusters and especially if one wishes to perform molecular dynamics simulations.

Molecular mechanics (MM) force fields are of course orders of magnitude less computationally demanding than QM methods. However, most MM force fields rely on many empirically fitted parameters that must be obtained for each different type of system. The parameter sets are often not transferable from system type to system type, thereby making it difficult to draw comparative conclusions. In addition, the most commonly used MM force fields do not contain all of the key components discussed above.

There are a small number of model potentials that are largely classical in nature but are derived from rigorous QM. Two notable examples are the SIBFA (sum of interactions among fragments ab initio computed) (6) force field and the effective fragment potential (EFP) (7–10) method. The present work focuses on the EFP method.

There are currently two versions of the EFP method, called EFP1 and EFP2. The original EFP1 interaction energy (7, 8) may be written as

$$E(\mathrm{EFP1}) = E^{Coul} + E^{pol} + E^{rem}. \quad (1)$$

In Equation 1 E^{Coul} represents the Coulomb interaction, computed according to the distributed multipole analysis suggested by Stone (11, 12), with the expansion points located at the atom centers and the bond midpoints. The polarization term E^{pol} is determined using a tensor sum of localized molecular orbital (LMO) polarizability tensors that are centered at the LMO centroids. The third term in Equation 1, E^{rem}, is a remainder term that is obtained by subtracting the first two terms from the total QM interaction energy of the water dimer and fitting the remainder to a functional form that depends on whether one is considering a QM-EFP interaction or an EFP-EFP interaction. The QM interaction energy has been obtained using Hartree-Fock (EFP1/HF)

and density functional theory (DFT) with the B3LYP functional (13, 14) [EFP1/DFT (15)]. In the EFP1/HF method, E^{rem} contains the exchange repulsion and CT terms. In the EFP1/DFT version, E^{rem} also includes short-range correlation via the correlation functional. The EFP1 method has been interfaced with most QM methods. In addition to HF and DFT, these methods include time-dependent DFT (16), singly excited configuration interaction (CIS) (17), multiconfiguration self-consistent field (MCSCF), MP2 and multireference MP2 (MRPT2) (18), and CC theory (19). Most applications of the EFP1 method have focused on aqueous solvent effects on ground and electronically excited state properties and processes (16–18, 20–49).

MCSCF: multiconfiguration self-consistent field

2EI: two-electron integral

The EFP2 interaction energy may be written as

$$E(\text{EFP2}) = E^{Coul} + E^{pol} + E^{exrep} + E^{disp} + E^{ct}. \quad (2)$$

The first two terms in Equation 2 are the same as those in Equation 1. The remaining terms in Equation 2 (the exchange repulsion, dispersion, and CT, respectively) are derived from first principles, so there are no empirically fitted parameters in EFP2, and the EFP2 method is not limited to water. Three of the terms in Equation 2 (E^{Coul}, E^{pol}, and E^{disp}) must be damped at either short or long intermolecular distances. Both Tang-Toennies (50) and overlap damping (51) schemes have been implemented. Generally, the overlap damping approach is recommended. Analytic energy gradients have been derived and implemented for all of the EFP2 terms.

In several applications (10, 52–60), it has been demonstrated that the EFP2 method can accurately predict the broad range of intermolecular interactions, ranging from those dominated by Coulomb interactions (e.g., water-water hydrogen bonds) to those dominated by dispersion (e.g., π-stacking interactions in DNA base pairs). The level of accuracy is generally equivalent to that of MP2, at orders of magnitude lower computational cost (61). EFP2 interaction energies are often closer to those obtained with CC methods than are the corresponding MP2 values (54, 57) and are considerably better than most functionals, including many that rely on the somewhat ad hoc addition of empirical dispersion terms (62). However, note that recently Szalewicz and coworkers (63) have developed a more systematic and rigorous dispersion method for DFT that appears to be promising.

While the EFP2 method is orders of magnitude faster than correlated QM methods, it is still slower than most simple MM methods. The main reason for this is that the CT term (64), evaluated as the interaction between occupied orbitals in one fragment with unoccupied orbitals in another fragment, requires the calculation of many approximate two-electron integrals (2EIs) over the entire (occupied plus virtual) orbital space. Very recently, a modified approach to the CT term has been developed (65) that significantly reduces the computer time requirements. Until recently, only the first two components in Equation 2 had been implemented for QM-EFP interactions, consequently limiting the study of multiple solvents. The QM-EFP terms for exchange repulsion (66) and dispersion (67) have now been derived and implemented, and the corresponding analytic gradients are in progress. These new developments, which are the primary focus of the present work, greatly expand the applicability of the EFP2 method. The EFP2 method is fully implemented in the GAMESS (General Atomic and Molecular Electronic Structure System) (68, 69) program and partially implemented in Q-Chem (70, 71).

2. SUMMARY OF THE EFP METHOD

2.1. EFP as a Force Field

The EFP method originated as a water potential to describe hydration effects on molecules of biological relevance (72). This first water potential (later referred to as EFP1), as illustrated in

EOM-CCSD: equation-of-motion coupled cluster with single and double excitations

Equation 1, contains the most important intermolecular interaction terms (Coulomb, polarization, and exchange repulsion) for the study of water structure and water behavior (7, 8, 15). The Coulomb and polarization terms are represented by distributed multipoles expanded through octopoles and by distributed LMO polarizabilities, respectively. The repulsion term of the exponential form is fitted to an exponential function, to reproduce either HF (EFP1/HF) or B3LYP DFT (EFP1/DFT) (15) energies of a set of ~200 water dimer geometries. The dispersion term (which is relatively small but generally not negligible for water) was subsequently introduced by fitting EFP1 to MP2 to produce an EFP1/MP2 method. The explicitly fitted dispersion term in EFP1/MP2 has the form $C_6/R^6 + C_8/R^8$. Starting in 1994, the EFP1 method has been interfaced with a variety of QM methods, including HF, DFT, MP2, CIS (17), MCSCF, MRPT2 (18), time-dependent DFT (16), and EOM-CCSD (equation-of-motion coupled cluster with single and double excitations) (19), in a QM/MM fashion in which all EFP1 terms are embedded as one-electron integrals in the QM Hamiltonian. The QM/EFP1 approach has been used in multiple studies of structures, chemical equilibria, reaction mechanisms, and electronic excitations and dynamics in aqueous systems (16–18, 29, 31, 32, 49, 73).

The popularity of EFP1 (mainly EFP1/HF and EFP1/DFT) is due to (*a*) the robustness of the water potential; (*b*) the simplicity of the functional form, which is important for interfacing it with QM methods; and (*c*) the simplicity of the EFP1 user interface in GAMESS. The successful implementation and applications of the EFP1 method prompted the effort to generalize this polarizable water potential to any solvent without the use of empirical parameterization. This generalization of the EFP approach requires a physically meaningful but computationally affordable formulation of all interaction energy components, including the repulsion and dispersion terms. Therefore, the recent development efforts of the general EFP method (also referred to as EFP2) have focused on the exchange repulsion, charge transfer (CT) (which was implicitly included in the EFP1 repulsive term), and dispersion terms and interfacing these components of the interaction energy with a QM region for QM/EFP schemes.

The five EFP2 interaction terms are summarized in Equation 2 (9, 10, 74, 75). These five terms may be grouped into long-range interactions that are $(1/R)^n$ distance dependent and short-range interactions that decay exponentially. The Coulomb, polarization, and dispersion interactions are long-range interactions that can be derived using the first (Coulomb) and second (polarization and dispersion) orders of Rayleigh-Schrödinger perturbation theory. The exchange repulsion, CT, and damping terms are short-range terms that can be represented as functions of the intermolecular overlap.

The Coulomb portion of the electrostatic interaction, E^{Coul}, is obtained using the Stone distributed multipolar analysis (11, 12). E^{pol} arises from the interaction of distributed induced dipoles on one fragment with a field due to multipoles and induced dipoles on the other fragments. The induced dipoles are created by anisotropic LMO polarizabilities. The number of polarizability points is equal to the number of bonds and lone pairs in the system; the core orbitals are typically excluded. The induced dipoles are iterated to self-consistency, so some many-body effects are captured (76).

The dispersion interaction energy can be expressed as the London expansion in inverse R, $E = \sum_n C_n/R^n$, with $n \geq 6$. The leading term, with $n = 6$, corresponds to the instantaneous dipole-induced dipole interactions. Distributed C_6 coefficients are derived from the (imaginary) frequency dependent polarizabilities integrated over the entire frequency range (77, 78). Centered at LMOs, dynamic polarizability tensors are obtained using the time-dependent HF method. In addition, the contribution of the $n = 8$ term is estimated as one-third of the $n = 6$ term. In the remainder of this work, the term EFP is used, rather than EFP2.

At short interfragment separations, the classical multipolar expansion diverges, leading to an incorrect asymptotic ($R \to 0$) behavior of the Coulomb, polarization, and dispersion terms. For

CHARGE TRANSFER

Although Mulliken (83) long ago used theory to anticipate the importance of ground state CT in intermolecular interactions, it now appears that electron delocalization may play an even more widespread role in aqueous chemistry than previously suspected (84). Ion-water CT (85–92) and the affinity of ions for aqueous interfaces (130–133), as well as the influence of ion-water and electron-water interactions on spectroscopy, biocatalysis, and nanoengineering (134–138), are subjects of intense current interest. However, despite multiple studies, the magnitude and importance of ion-water CT remain subjects of significant controversy. The intrinsic connection between CT and polarization makes the analysis of the CT interactions ambiguous. Consequently, formulations of CT range from those in which CT is considered to be an artificial term arising from incompleteness of the basis set to those, like natural bond analysis (139, 140), in which CT plays a predominant role in intermolecular binding. It is also unclear at present whether CT may be included as a stabilizing energy term or whether the actual transfer of charge is required for quantitative prediction of structure and dynamics at interfaces (86, 88, 92, 141). New fundamental studies of the origins of bonding in terms of the components discussed here will hopefully shed some light on this important problem (142).

the polarization interaction, the divergence of the multipolar expansion may result in so-called polarization collapse: It becomes impossible mathematically to obtain finite values of the self-consistent induced dipoles. To avoid the artifacts of the short-range behavior, one can modify the EFP Coulomb, polarization, and dispersion terms by damping functions (51, 54, 79). Different forms of the damping functions have been developed over time; the current recommendations are to use parameter-free overlap-based damping functions for Coulomb and dispersion terms and Gaussian-like damping for polarization (51).

The exchange repulsion term E^{exrep} is unique to the EFP method. The interaction between fragments is derived as an expansion in the intermolecular overlap, truncated at the quadratic term (80–82). The necessary overlap and kinetic energy integrals for each pair of fragments that carry a basis set and localized wave function are calculated on the fly.

From a quantum mechanical point of view, the CT energy lowering occurs due to interactions of the occupied orbitals on one fragment with the virtual orbitals on another fragment. The CT term is especially significant in polar or ionic species. The EFP CT term is derived based on a second-order perturbative treatment of the intermolecular interactions. The original implementation of the CT term (64) uses canonical HF orbitals of individual fragments and a multipolar expansion of the electrostatic potential of the fragment. A new implementation that incorporates localized rather than canonical virtual orbitals is discussed in Section 3.1. While the CT term is the most computationally expensive EFP component and is often omitted, its importance for predicting structures, dynamics, and spectroscopic signatures remains to be fully determined (6, 83–92).

The availability of analytic gradients for all of the EFP terms facilitates straightforward geometry optimizations of small clusters, as well as Monte Carlo or molecular dynamics simulations of larger clusters or bulk systems. Thus, the EFP method can be viewed as a first principles–based alternative to classical force fields for modeling properties of molecular clusters and bulk systems (see the sidebar, Charge Transfer).

2.2. QM/EFP Interface

Combined QM/MM methods, in which a QM approach is used for the region in which chemistry (e.g., bond making/bond breaking) occurs and a model potential is used for the observer region,

were introduced by Warshel (93–95) and others (96–105). QM/MM methods are widely used for large systems (e.g., a solute in a solvent, a reacting system on a surface, organometallic complexes with large, bulky ligands) that are difficult to treat by QM methods alone. The EFP provides a sophisticated QM/EFP interface in the spirit of the QM/MM paradigm.

The Coulomb and polarization terms in the QM/EFP interface are treated as one-electron contributions, V^{Coul} and V^{pol}, to the quantum mechanical Hamiltonian H^{QM}:

$$E^{QM-EFP} = \langle \Psi | H^{QM} + V^{Coul} + V^{pol} | \Psi \rangle. \quad (3)$$

Polarization in a QM/EFP system is treated self-consistently via a two-level iterative procedure. The lower level treats the convergence of the induced dipoles in the presence of the frozen ab initio wave function. The higher level is a standard HF iterative cycle in which the wave function is updated based on the converged values of the induced dipoles from the lower level. Convergence of the two-level procedure yields self-consistent induced dipoles and the ab initio wave function.

For the EFP1 water potential, the only remaining term is the exchange repulsion interaction that is obtained using fitted parameters as described above. Consequently, the QM/EFP1 interface for the energy and analytic gradient is completely developed and available in GAMESS. The development and implementation of the QM/EFP2 method have been more challenging due to the presence of the complex exchange repulsion, dispersion, and CT terms, each of which is derived from first principles. Recent progress in this direction is described in Section 3. This more general QM/EFP2 method is important, since it extends the study of solvent effects on the broad range of chemical and biological problems to all solvents, from polar protic solvents to nonpolar aprotic ones.

The implementation of the QM/EFP2 interface has been partially implemented in both GAMESS and Q-Chem (10, 66, 67, 106, 107). These implementations allow one to use the QM methods CCSD(T), EOM-CCSD, time-dependent DFT, CIS, CIS(D), MCSCF, and multireference perturbation theory for ground and excited electronic states with EFP (106–108). It has been shown, for example, for excited states that inclusion of the Coulomb and polarization terms into the QM Hamiltonian provides the majority of solvation effects such as solvatochromic shifts in polar or polarizable solvents (106, 107). The polarizable environment interacts differently with each electronic state of a solute so that the effective Hamiltonians of the different states differ by state-specific V^{pol} terms. Currently, one can either treat the polarizable environment fully self-consistently for each electronic state of interest or decouple the solute and solvent and solve the eigenvalue problem for the electronic excited states with a constant (frozen) response of the EFP environment corresponding to its electronic ground state value (17, 29). The latter approach, which preserves the orthogonality of the electronic states, has been shown to account for the overwhelming majority of the impact of the solvent on excited states. If necessary, the interaction of the excited state wave function with the polarizable environment can be included perturbatively (17, 29, 106, 107).

2.3. Using the EFP Method

The original implementation of the EFP method is in GAMESS. It contains complete energy and gradient codes for the EFP1 water potential both for fragment-only and QM/EFP1 jobs, as well as the energy and gradient for the EFP2 potential in fragment-only calculations. EFP1 is interfaced with most QM methods available in GAMESS, as well as the fragment molecular orbital method (109, 110). The polarizable continuum model can be used for representing solvation at the boundaries of EFP1 or QM/EFP1 systems (111).

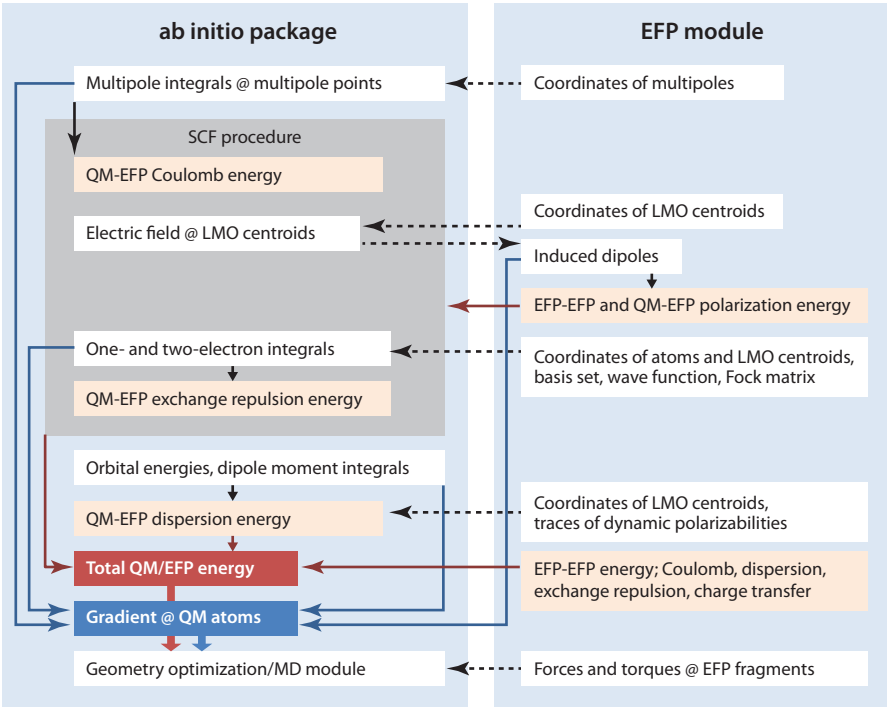

Figure 1

Interface between an electronic structure package and an effective fragment potential (EFP) module. Abbreviations: LMO, localized molecular orbital; MD, molecular dynamics; QM, quantum mechanics.

Recently, the EFP method was also implemented in Q-Chem (10). The Q-Chem implementation contains energy and gradient terms of the general EFP potential for fragment-fragment computations. QM/EFP electrostatic, polarization, and dispersion energy terms have been also implemented, allowing EFP to interface with many excited state methods available in Q-Chem (106–108).

Since the EFP method may be viewed as an accurate substitute for a classical force field in QM/MM simulations, it can be implemented as a module with a straightforward interface to various computational packages (see **Figure 1**). Interfacing EFP to different packages (and electronic structure methods and algorithms uniquely available in these packages) provides multiple opportunities to extend methodologies originally designed for gas phase chemistry to condensed phase and extended systems, without a significant increase in the computational cost. Such an EFP module has been developed using components (112, 113).

An effective potential for a fragment contains the following information. First, an electrostatic term includes coordinates of atoms and bond midpoints, and distributed multipoles (charges, dipoles, quadrupoles, octopoles) at these points; in the case of a short-range exponential screening, each point should contain one screening parameter as well. Second, a polarization term comprises coordinates of LMO centroids and distributed polarizability tensors (with nine components each) at these points; in the case of a short-range Gaussian screening, one generic screening parameter is added. Third, the dispersion term consists of the coordinates of the LMO centroids and the traces of the distributed dynamic polarizability tensors calculated at 12 predetermined imaginary frequencies at the LMO centroids. Fourth, an exchange repulsion term includes the coordinates

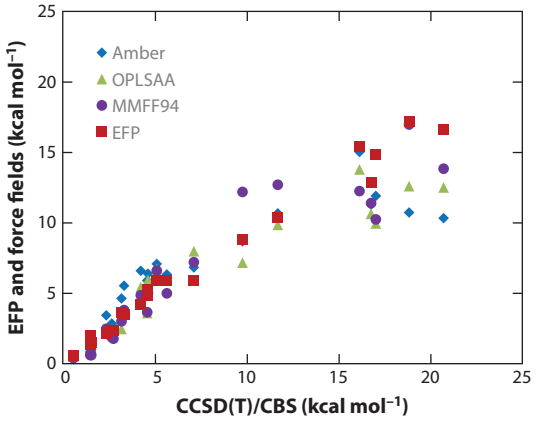

Figure 2

Total interaction energies for S22 data-set dimers calculated with the effective fragment potential (EFP) and molecular mechanics force fields (126) (AMBER, OPLSAA, MMFF94) compared with CCSD(T)/CBS (equation-of-motion coupled cluster with single, double, and triple excitations/complete basis set) (127).

of the atoms and the LMO centroids, a basis set, a localized wave function (occupied orbitals only), and a Fock matrix in a localized basis. Finally, a CT term consists of the coordinates of the atoms, a basis set, a canonical or localized wave function (all active orbitals), and orbital energies.

In order to simplify usage of the EFP method by nonexperts, both GAMESS and Q-Chem distributions provide a standardized set of effective potentials for common solvents, organic molecules, and DNA bases (10, 61). These standardized effective potentials were used to evaluate the accuracy of the EFP method on the S22 (114) and S66 (115) data sets of intermolecular noncovalent interactions (61). Mean absolute deviations (MADs) of EFP interaction energies with respect to CCSD(T) in complete basis set data are 0.91 kcal mol^{-1} and 0.61 kcal mol^{-1} for the S22 and S66 sets, respectively. The relative errors of the EFP interaction energies are 11–12%. Thus the accuracy of the EFP method for the description of intermolecular interactions is comparable to that of MP2, while the computational cost of EFP is several orders of magnitude less. The EFP method generally outperforms classical force fields (61) (see **Figure 2**).

The formal computational scaling of the EFP method for fragment-fragment interactions is $O(N^2)$, where N is the number of fragments. If Coulomb and dispersion interactions cost one unit, polarization would cost two units, and exchange repulsion would cost five units. CT in the original implementation with canonical orbitals would cost 50 units, but this component becomes more computationally affordable (20–30 units) in the new localized-orbital implementation. For systems with large N, scaling of the short-range exchange repulsion and CT terms decreases to $O(N)$, employing distance-based screening of the overlap integrals. Then, the total effective cost of the EFP calculations is determined by the long-range electrostatic and polarization terms.

The computational cost of QM/EFP calculations is typically determined by the cost of the corresponding QM calculation in the gas phase. However, in systems with a very large number of effective fragments, the cost of fragment-fragment calculations could become dominant.

3. NEW EFP DEVELOPMENTS

As noted above, the EFP-EFP and EFP-QM Coulomb and polarization interaction terms, and the corresponding damping terms, are well established. This section focuses on recent developments

regarding the other terms in the interaction energy: the EFP-EFP charge transfer (CT) and the EFP-QM exchange repulsion and dispersion interactions.

CMO: canonical molecular orbital

QUAMBO: quasiatomic minimal-basis-set orbital

3.1. The EFP-EFP Charge Transfer Interaction

The energy lowering due to the interaction between the occupied orbitals on one EFP2 fragment and the virtual orbitals of another is defined as the EFP-EFP CT energy. The approximate formula for the EFP-EFP CT interaction is derived from a second-order perturbative approach in which the SCF canonical molecular orbitals (CMOs) and orbital energies are generated in a preparative SCF calculation (64).

As a pair-wise interaction, the EFP CT energy of A induced by B is approximated as

$$CT^{A(B)} = 2\sum_{i}^{occA}\sum_{n}^{virB}\frac{1}{1-\sum_{m}^{allA}(S_{mn})^2}\frac{V_{in}^{EFB}-\sum_{m}^{allA}S_{mn}V_{im}^{EFB}}{(F_{ii}^{A}-T_{nn})} \times\left[V_{in}^{EFB}-\sum_{m}^{allA}S_{nm}V_{im}^{EFB}+\sum_{j}^{occB}S_{ij}\left(T_{nj}-\sum_{m}^{allA}S_{nm}T_{mj}\right)\right], \quad (4)$$

and the CT energy of B induced by A is

$$CT^{B(A)} = 2\sum_{j}^{occB}\sum_{m}^{virA}\frac{1}{1-\sum_{n}^{allB}(S_{mn})^2}\frac{V_{jm}^{EFA}-\sum_{n}^{allB}S_{mn}V_{jn}^{EFA}}{(F_{jj}^{B}-T_{mm})} \times\left[V_{jm}^{EFA}-\sum_{n}^{allB}S_{nm}V_{jn}^{EFA}+\sum_{i}^{occA}S_{ij}\left(T_{mi}-\sum_{n}^{allB}S_{nm}T_{ni}\right)\right]. \quad (5)$$

In Equations 4 and 5, S and T are the intermolecular overlap and kinetic energy integrals; the abbreviations *occ*, *vir*, and *all* refer to sums over occupied, virtual, and all orbitals; and

$$V_{in}^{nuc\,B}+\sum_{j}^{occB}(2\langle in\,|jj\rangle-\langle ij\,|nj\rangle)\approx V_{in}^{nuc\,B}+\sum_{j}^{occB}2\langle in\,|jj\rangle\approx V_{in}^{EFB}. \quad (6)$$

In Equation 6, the electrostatic potential is represented with multipole expansion points. The one-electron integrals $V^{EFA/B}$, which account for the electrostatic potential of the other fragment, are the most time-consuming terms in the CT equations. Furthermore, these one-electron terms in Equations 4 and 5 loop over not only the occupied molecular orbitals, but also the virtual molecular orbitals of each fragment. For example, for a typical EFP calculation, the recommended basis set is 6-311++G(3df,2p) (54, 61, 64, 80, 81, 116–118). For a water molecule using this basis set, there are 60 virtual CMOs and only five occupied MOs. In general, this large number of virtual MOs is the bottleneck for calculating the EFP-EFP CT interaction. Other EFP-EFP interactions (e.g., exchange repulsion) only require loops over occupied MOs on the fly. The time-consuming steps (e.g., coupled-perturbed HF calculations for the polarization and dispersion) are completed in preparatory calculations. Consequently, CT calculations are typically 20–30 times slower than those for the other terms (64).

To reduce the computational cost of the CT calculations, quasiatomic minimal-basis-set orbitals (QUAMBOs) (119) are employed as an alternative to the CMOs. The number of QUAMBOs is the number of minimal-basis-set orbitals of the molecule (119), so there are much fewer virtual orbitals than in the previous CT implementation. More importantly, these QUAMBO virtual orbitals serve as a basis to expand the valence virtual orbitals (VVOs), which comprise the most important part of the virtual space. The deviation of the QUAMBOs (119) amounts to minimizing

the mean square deviation

$$\langle A_j - A_j^* | A_j - A_j^* \rangle = 2[1 - \langle A_j | A_j^* \rangle] = 2[1 - (D_j)^{1/2}], \quad (7)$$

RVS: reduced variational space

where A_j are the QUAMBOs, A_j^* are the free-atom minimal-basis valence orbitals, and

$$D_j = \sum_n \langle \phi_n | A_j^* \rangle^2 + \sum_p \langle \psi_p | A_j^* \rangle^2, \quad (8)$$

where ϕ_n are the occupied SCF molecular orbitals, and ψ_p are selected orbitals from the space spanned by the virtual SCF molecular orbitals. The subscript j runs over the number of minimal-basis-set valence atomic orbitals for the particular atom in the molecule, and n and p go up to the number of valence occupied orbitals and VVOs, respectively.

To simultaneously minimize Equation 7 for all QUAMBOs, it is algorithmically equivalent to maximize the sum

$$\sum_j D_j = \sum_j \sum_n \langle \phi_n | A_j^* \rangle^2 + \sum_j \sum_p \langle \psi_p | A_j^* \rangle^2. \quad (9)$$

This maximization is ultimately achieved by maximizing the sum over the virtual orbitals ψ_p, the second term of Equation 8:

$$\psi - sum = \sum_j \sum_p \langle \psi_p | A_j^* \rangle^2 = \sum_p \sum_v \sum_w T_{vp} T_{wp} B_{vw}, \quad (10)$$

where $\varphi_p = \sum_v \phi_v T_{vp}$, with T being the expansion matrix and $B_{vw} = \sum_j \langle \phi_v | A_j^* \rangle \langle \phi_w | A_j^* \rangle = \sum_j a_{vj}^* a_{wj}^*$.

Note that v and w run over all the SCF virtual orbitals. By choosing the columns of the matrix T as the eigenvectors of the matrix B with the largest p eigenvalues β_p (i.e., $\sum_w B_{vw} T_{wp} = \beta_p T_{vp}$), one finds that $\psi - \text{sum} = \sum_p \sum_v \beta_p T_{vp} T_{vp} = \sum_p \beta_p$ is then at its maximum. Once the matrix T is determined, the set of P VVOs, ψ_p, and subsequently the expansion coefficients of QUAMBOs in terms of SCF MOs, a_{nj} and a_{vj}, can be obtained. Since the determination of QUAMBOs is basis set independent, the resulting resolution of MOs in terms of QUAMBOs is intrinsic to the exact wave function (119). In fact, QUAMBOs can be viewed as slightly deformed atomic orbitals in a molecular environment.

The use of QUAMBOs in the EFP-EFP CT energy and gradient calculation dramatically reduces the computational cost of EFP calculations. Typically, a reduction of 50% or more of the total CPU time reduction is observed (**Table 1**). **Figure 3** compares the CT energies calculated by QUAMBOs versus CMOs to reduced variational space (RVS) analysis (120, 121) for several dimer systems at five different basis sets. At equilibrium distances, QUAMBO CT energies are in better agreement with the benchmarks in the majority of cases. At longer separations, both QUAMBO and CMO CT energies agree well with RVS results. At distances shorter than the equilibrium separation, QUAMBOs tend to underestimate the CT energy, whereas CMOs tend to overestimate the energy. However, when the largest basis set is used, 6-311++G(3df,2p), both types of orbitals lead to the underestimation of the CT interaction (**Figure 3**). In general, QUAMBO-predicted CT interactions agree better with RVS values in the region around equilibrium and at longer distances. Since QUAMBOs do not span the whole SCF virtual space, it may seem counterintuitive that QUAMBOs achieve better results than using the full canonical space. This is partly due to the approximate nature of the CT formulation (e.g., neglect of exchange integrals, approximate Coulomb integrals by multipole interactions), which leads to a cancellation of errors.

Table 1 The total CPU time (in seconds) for an EFP-EFP energy or gradient calculation of $(NH_4^+\text{-}NO_3^-)_4$ at various basis sets using either canonical molecular orbitals (CMOs) or quasiatomic minimal-basis-set orbitals/valence virtual orbitals (VVOs)

		Energy		Gradient	
Basis set	Number of basis functions	CMO	VVO	CMO	VVO
6–31+G(d,p)	460	2.12	1.04	8.63	3.48
6-31++G(d,p)	476	2.24	0.94	9.22	3.73
6-31++G(df,p)	676	3.71	1.67	16.36	6.00
6-311+G(d,p)	572	2.86	1.29	11.90	4.56
6-311++G(3df,2p)	1,060	7.54	3.25	36.70	11.45

3.2. QM-EFP Exchange Repulsion

Exchange repulsion, the only repulsive component of the EFP method, is a purely quantum-mechanical effect due to the Pauli exclusion principle. The exchange repulsion between two EFP fragments is expressed as an expansion in the intermolecular overlap, truncated at the second order (80). The approximations used have been shown to work much better for LMOs (80) than for CMOs. For the QM-EFP exchange repulsion interaction, the QM molecule is in the canonical basis during the SCF iterations, so many of the approximations cannot be applied.

Equation 11 describes the QM-EFP2 exchange repulsion (82). The exchange repulsion Fock operator (82) (Equation 12) is added to the QM Fock matrix as a one-electron contribution to account for the presence of EFPs:

$$E^{XR} = -2\sum_{i\in A}\sum_{j\in B}(ij\,|\,ij) - 2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[2(V_{ij}^{A}+G_{ij}^{A})+\sum_{l\in B}F_{jl}^{B}S_{li}\right]$$
$$+2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[\sum_{k\in A}S_{kj}(F_{ik}^{A}+V_{ik}^{B}+J_{ik}^{B})+S_{ij}(V_{jj}^{A}+J_{jj}^{A})-\sum_{k\in A}S_{kj}(ik|jj)\right], \quad (11)$$

$$V_{mi}^{XR} = -\sum_{j\in B}(mj\,|\,ij) - \frac{1}{2}\sum_{j\in B}S_{mj}\left[2(V_{ij}^{A}+G_{ij}^{A})+\sum_{l\in B}F_{jl}^{B}S_{li}\right]$$
$$-\frac{1}{2}\sum_{j\in B}\left[2(V_{mj}^{A}+G_{mj}^{A})+\sum_{l\in B}F_{jl}^{B}S_{lm}\right]$$
$$-\sum_{k\in A}\sum_{j\in B}S_{kj}\,[4(kj\,|\,mi)-(km\,|\,ji)-(ki\,|\,jm)]$$
$$+\sum_{j\in B}S_{mj}\left[\sum_{k\in A}S_{kj}(F_{ik}^{A}+V_{ik}^{B}+J_{ik}^{B})-(ik\,|\,jj)+S_{ij}(V_{jj}^{A}+J_{jj}^{A})\right]$$
$$+\sum_{j\in B}S_{ij}\left[\sum_{k\in A}S_{kj}(F_{mk}^{A}+V_{mk}^{B}+J_{mk}^{B})-(mk\,|\,jj)\right]$$
$$+\frac{1}{2}\sum_{n\in A}\sum_{k\in A}\sum_{j\in B}S_{kj}S_{nj}\,[4(nk\,|\,im)-(nm\,|\,ik)-(ni\,|\,mk)]+2\sum_{k\in A}\sum_{j\in B}S_{kj}^{2}(jj\,|\,mi), \quad (12)$$

where

$$(ij|ij) = \int\int \phi_i^*(r_1)\phi_j(r_1)\frac{1}{r_{12}}\phi_i^*(r_2)\phi_j(r_2)dr_1dr_2, \quad (13)$$

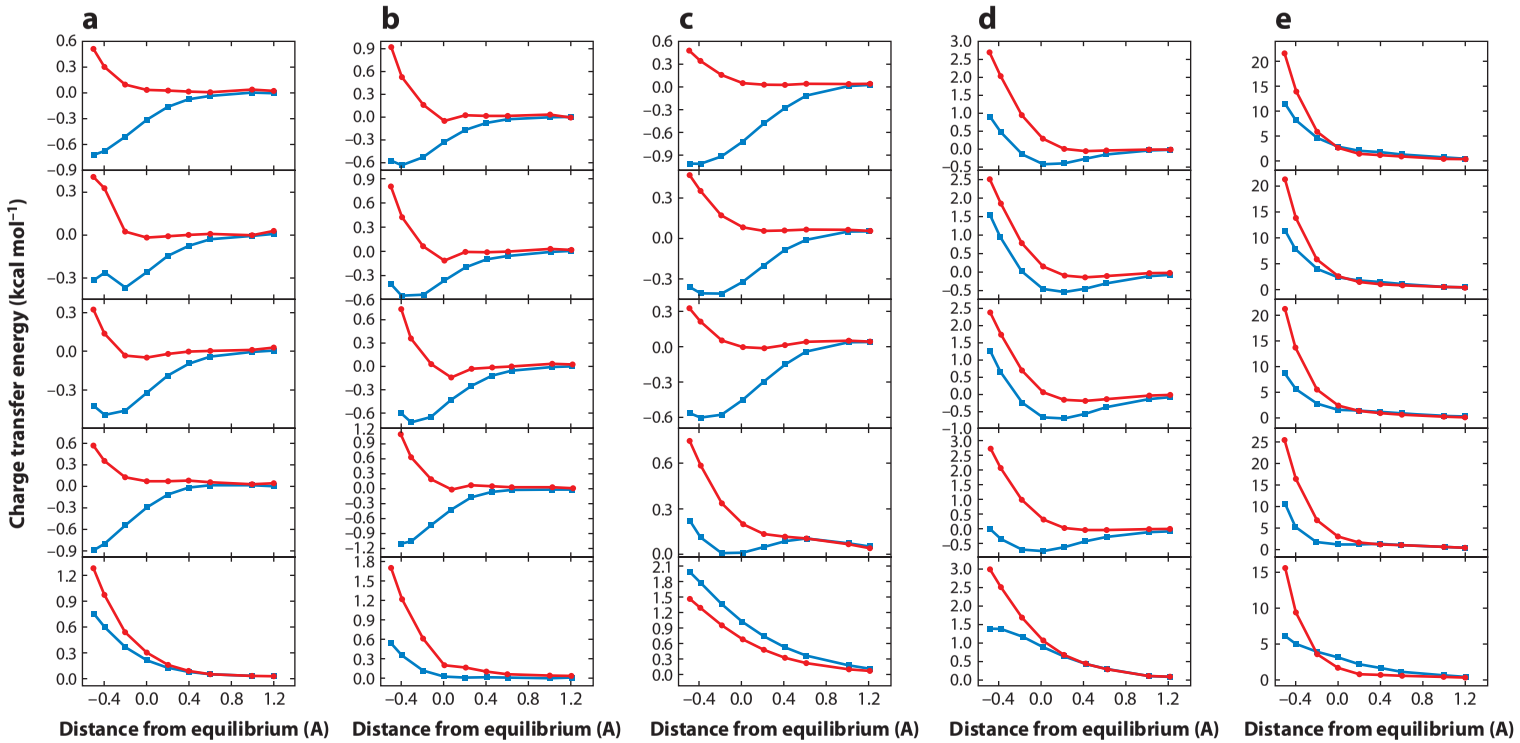

Figure 3

Differences in the charge transfer (CT) energies (in kcal mol^{-1}) calculated using canonical molecular orbital (CMO) (*blue squares*) or quasiatomic minimal-basis-set orbitals/valence virtual orbitals (QUAMBOs/VVOs) (*red circles*) compared to reduced variational space (RVS) values for (*a*) water dimer, (*b*) water-methanol, (*c*) ammonia-water, (*d*) ammonium-water, and (*e*) ammonium-nitrate with five basis sets: 6-31+G(d,p), 6-31++G(d,p), 6-31++G(df,p), 6-311++G(d,p), and 6-311++G(3df,2p). The *x* axis is the distance from the equilibrium separation (0.0); positive distance displacements represent increases relative to the equilibrium. The energy difference is defined as the CT energy from the CMO or VVO minus the CT energy from the RVS calculation. Since the CT energy is always negative, a positive number in the figure means an underestimated CT energy and vice versa.

$$S_{ij} = \int \phi_i(r_1)\phi_j(r_1)dr_1, \tag{14}$$

$$V_{ij}^{A} = \sum_{I\in A}\int \phi_i^*(r_1)\frac{Z_I}{R_{1I}}\phi_j(r_1)dr_1, \tag{15}$$

$$G_{ij}^{A} = 2J_{ij}^{A} - K_{ij}^{A}, \tag{16}$$

$$J_{ij}^{A} = \sum_{k\in A}(ij|kk) = \sum_{k\in A}\iint \phi_i^*(r_1)\phi_j(r_1)\frac{1}{r_{12}}\phi_k^*(r_2)\phi_k(r_2)dr_1dr_2, \tag{17}$$

$$K_{ij}^{A} = \sum_{k\in A}(ik|jk) = \sum_{k\in A}\iint \phi_i^*(r_1)\phi_k(r_1)\frac{1}{r_{12}}\phi_j^*(r_2)\phi_k(r_2)dr_1dr_2, \tag{18}$$

$$F_{ik}^{A} = T_{ik}^{A} + V_{ik}^{A} + G_{ik}^{A}, \tag{19}$$

$$T_{ij}^{A} = \int \phi_i(r_1)\left(-\frac{1}{2}\nabla_1^2\right)\phi_j(r_1)dr_1. \tag{20}$$

All orbitals above refer to molecular orbitals, and the orbitals i and j are always on QM molecule A and EFP molecule B, respectively.

Many of the approximations used for the exchange repulsion require LMOs, in particular, the spherical Gaussian overlap (SGO) approximation (122). For the QM-EFP interaction, the SGO approximation is applied to the atomic basis functions of the QM molecule and the EFP LMOs (Equation 23). In addition, the 2EIs are approximated as an electrostatic potential represented by the multipole expansion (Equations 24 and 25). The one-electron nuclear attraction term, V_{jj}^{A}, is replaced by a classical point-charge approximation, $-Z_I/R_{jI}$ (Equation 26). These approximations lead to the following expressions for the exchange repulsion energy and Fock operator (66, 82):

SGO: spherical Gaussian overlap approximation

$$E^{XR} \approx -2\sum_{i\in A}\sum_{j\in B}(ij\,|\,ij)^{SGO} - 2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[2(V_{ij}^{A} + G_{ij}^{A}) + \sum_{l\in B}F_{jl}^{B}S_{li}\right]$$
$$+2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[\sum_{k\in A}S_{kj}(F_{ik}^{A} + V_{ik}^{EFP,B}) + S_{ij}\left(\sum_{I\in A}-Z_I R_{Ij}^{-1} + \sum_{k\in A}V_{kk}^{j}\right) - \sum_{k\in A}S_{kj}V_{ik}^{j}\right], \tag{21}$$

$$V_{mi}^{XR} \approx -\sum_{j\in B}(mj\,|\,ij)^{SGO} - \frac{1}{2}\sum_{j\in B}S_{mj}\left[2(V_{ij}^{A} + G_{ij}^{A}) + \sum_{l\in B}F_{jl}^{B}S_{li}\right]$$
$$-\frac{1}{2}\sum_{j\in B}\left[2(V_{mj}^{A} + G_{mj}^{A}) + \sum_{l\in B}F_{jl}^{B}S_{lm}\right]$$
$$-\sum_{k\in A}\sum_{j\in B}S_{kj}\,[4(kj\,|\,mi) - t(km\,|\,ji) - (ki\,|\,jm)]$$
$$+\sum_{j\in B}S_{mj}\left[\sum_{k\in A}S_{kj}(F_{ik}^{A} + V_{ik}^{EFP,B}) - V_{ik}^{j} + S_{ij}\left(\sum_{I\in A}Z_I R_{Ij}^{-1} + V_{kk}^{j}\right)\right]$$
$$+\sum_{j\in B}S_{ij}\left[\sum_{k\in A}S_{kj}(F_{mk}^{A} + V_{mk}^{EFP,B}) - V_{mk}^{j}\right]$$
$$+\frac{1}{2}\sum_{n\in A}\sum_{k\in A}\sum_{j\in B}S_{kj}S_{nj}[4(nk\,|\,im) - (nm\,|\,ik) - (ni\,|\,mk)] + 2\sum_{k\in A}\sum_{j\in B}S_{kj}^{2}V_{mi}^{j}, \tag{22}$$

where

$$(ij\,|\,ij)^{SGO} = \sum_{\mu}^{AO}\sum_{\nu}^{AO} C_{\mu i}C_{\nu i}(\mu j\,|\,\nu j)^{SGO}$$

$$\approx \sum_{\mu}^{AO}\sum_{\nu}^{AO} C_{\mu i}C_{\nu i}\frac{2}{\sqrt{\pi}}\sqrt{\frac{2\alpha_{\mu j}\alpha_{\nu j}}{\alpha_{\mu j}+\alpha_{\nu j}}}S_{\mu j}S_{\nu j}F_0\left[\frac{1}{4}\left(\frac{2\alpha_{\mu j}\alpha_{\nu j}}{\alpha_{\mu j}+\alpha_{\nu j}}R^2_{\mu\nu}\right)\right], \quad (23)$$

F_0 is the incomplete gamma function $F_0[t] = \frac{1}{2}\left(\frac{\pi}{t}\right)^{1/2} erf\,(t^{1/2})$, and

$$F^A_{ik} + V^A_{ik} + J^B_{ik} \approx F^A_{ik} + V^{ES,B}_{ik} \approx F^A_{ik} + V^{EFP,B}_{ik}, \quad (24)$$

$$(ik\,|\,jj) \approx (i\,|\,(r_1 - R_J)^{-1}\,|\,k) = V^j_{ik}, \quad (25)$$

$$V^A_{jj} = \sum_{j\in B}\sum_{I\in A}\frac{-Z_I}{R_{jI}}. \quad (26)$$

It has been shown in a previous paper (66) that the 2EIs in which one center is on an EFP and the other three centers are on the QM molecule cannot be approximated by the SGO approximation. To compensate for the consequent loss of efficiency, the exchange repulsion Fock matrix is added to the updated QM Fock matrix only every fourth iteration (66) since the EFPs are considered to be small perturbations. Furthermore, Schwartz inequality screening is applied to avoid calculating many small 2EIs.

Since exchange repulsion is an HF phenomenon, the RVS analysis is employed as the benchmarking method. For small homogeneous clusters (e.g., trimer, tetramer, pentamer), the errors are well within 4 kcal mol^{-1} (**Table 2**). For homogeneous clusters, the predicted QM-EFP exchange repulsion depends on the position and orientation of the QM molecule relative to the EFP fragments. This distance dependence is evident for a medium-sized test system, $(H_2O)_{16}$, for which the RVS analysis is too costly and the all-EFP exchange repulsion energy is considered as the benchmark. For the $(H_2O)_{16}$ configuration in **Figure 4**, better agreement with the benchmark is achieved when an outer edge water molecule is QM, compared to that when a more embedded water is QM. Another source of error is the approximations used in the EFP-EFP exchange repulsion, since the total exchange repulsion energy of a system is the sum of all the pair-wise QM-EFP and EFP-EFP exchange repulsion energies. The different approximations used for EFP-EFP and QM-EFP exchange repulsion contribute to the variation in the total exchange repulsion energy.

For heterogeneous clusters, some QM species, such as acetone and acetonitrile, appear to be particularly sensitive to the approximations. This was previously noticed for EFP-EFP calculations (81). For acetone, it has been suggested that one needs to go beyond the SGO approximation (81). An expression for the exchange repulsion gradient with respect to the QM nuclei (Equation 14) between the QM molecule and an EFP fragment has been derived from the exact energy expression (Equation 11) and is being implemented:

$$\frac{\partial E^{XR}}{\partial x_a} = -2\sum_{i\in A}\sum_{j\in B}(ij\,|\,ij)^a - 2\sum_{i\in A}\sum_{j\in B}S^a_{ij}\left[2(V^A_{ij} + G^A_{ij}) + \sum_{l\in B}F^B_{jl}S_{li}\right]$$

$$-2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[2(V^{A^a}_{ij} + G^{A^a}_{ij}) + \sum_{l\in B}F^B_{jl}S^a_{li}\right]$$

$$+2\sum_{i\in A}\sum_{j\in B}S^a_{ij}\left[\sum_{k\in A}(S_{kj}(F^A_{ik} + V^B_{ik} + 2J^B_{ik}) + S_{ij}(V^A_{jj} + 2J^A_{jj}) - \sum_{k\in A}S_{kj}(ik\,|\,jj))\right]$$

$$
\begin{aligned}
&+2\sum_{i\in A}\sum_{j\in B}S_{ij}\left[\begin{array}{l}\sum_{k\in A}(S^a_{kj}(F^A_{ik}+V^B_{ik}+2J^B_{ik})+S_{kj}(F^{A^a}_{ik}+V^{B^a}_{ik}+2J^{B^a}_{ik}))\\ \quad +S^a_{ij}(V^A_{jj}+2J^A_{jj})+S_{ij}(V^{A^a}_{jj}+2J^{A^a}_{jj})\\ -\sum_{k\in A}[S^a_{kj}(ik\,|\,jj)+S_{kj}(ik\,|\,jj)^a]\end{array}\right]\\
&+\sum_{m\in A}\sum_{i\in A}\sum_{j\in B}S^a_{mi}\left[\begin{array}{l}2(mj\,|\,ij)+S_{mj}\left(2(V^A_{ij}+G^A_{ij})+\sum_{l\in B}F^B_{jl}S_{li}\right)\\ \quad +S_{ij}\left(2(V^A_{mj}+G^A_{mj})+\sum_{l\in B}F^B_{jl}S_{lm}\right)\\ -S_{mj}\left(\sum_{k\in A}S_{kj}(F^A_{ik}+V^B_{ik}+2J^B_{ik})+S_{ij}(V^A_{jj}+2J^A_{jj})-\sum_{k\in A}S_{kj}(ik\,|\,jj)\right)\\ -S_{ij}\left(\sum_{k\in A}S_{kj}(F^A_{mk}+V^B_{mk}+2J^B_{mk})+S_{mj}(V^A_{jj}+2J^A_{jj})-\sum_{k\in A}S_{kj}(mk\,|\,jj)\right)\end{array}\right]\\
&+\sum_{m\in A}\sum_{i\in A}\sum_{n\in A}\sum_{j\in B}S^a_{mn}S_{ij}\\
&\quad\times\left[2[4(ij\,|\,mn)-(im\,|\,jn)-(in\,|\,jn)]-\sum_{k\in A}S_{kj}[4(ik\,|\,mn)-(im\,|\,kn)-(in\,|\,km)]\right]\\
&-\sum_{m\in A}\sum_{i\in A}\sum_{k\in A}\sum_{j\in B}S^a_{mk}S_{ij}[S_{mj}(F^A_{ik}+V^B_{ik}+2J^B_{ik})+S_{kj}(F^A_{im}+V^B_{im}+2J^B_{im})\\
&\quad+4S_{ij}(jj\,|\,mk)-S_{mj}(ik\,|\,jj)-S_{kj}(im\,|\,jj)].
\end{aligned}
\tag{27}
$$

Table 2 Exchange repulsion energies (in kcal mol^{-1}) obtained from benchmark calculations and quantum mechanics (QM)-EFP calculations

Exchange repulsion	Benchmark	QM-EFP best agreement	QM-EFP worst agreement
$(H_2O)_3$	15.0	16.6	17.0
$(MeOH)_3$	13.5	13.5	16.2
$((CH_3)_2CO)_3$	5.6	4.4	4.2
$(CH_3CN)_3$	5.1	4.3	3.9
$(CH_2Cl_2)_3$	1.1	1.0	2.4
3DMSO	10.1	8.1	Not converged
$(H_2O)_4$	29.3	28.4	27.4
$(H_2O)_5$	39.1	36.7	35.3
$(H_2O)_6$-bag	42.5	43.6	35.3
$(H_2O)_6$-boat	43.3	40.9	36.5
$(H_2O)_6$-book	43.8	43.8	39.8
$(H_2O)_6$-cage	40.9	41.7	38.6
$(H_2O)_6$-cyclic	45.0	41.8	36.2
$(H_2O)_6$-prism	39.8	40.1	41.3
$(H_2O)_{16}$	118.3	118.6	123.5

The third and fourth columns show the smallest and the largest deviations from the reduced variational space (RVS) interaction energies when different molecules are treated ab initio. For all systems except $(H_2O)_{16}$, the benchmark results were obtained from the RVS analysis. The $(H_2O)_{16}$ benchmark value was obtained from an all-EFP calculation. All of the cluster structures were optimized with RHF/6-31+G(d,p), and the EFPs were generated with the 6-311++G(3df,2p) basis set.

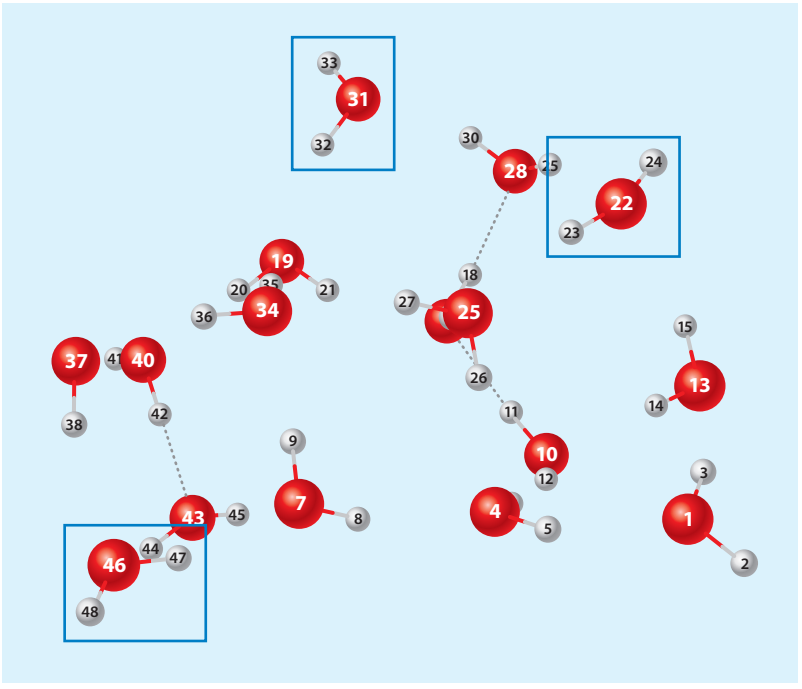

Figure 4

An $(H_2O)_{16}$ configuration. The boxed water molecules give better exchange repulsion energy when treated with quantum mechanics.

The derivatives of the molecular orbital coefficients are treated by expanding with orbital response terms (66, 123):

$$\frac{\partial C_{\mu i}}{\partial x_a} = \sum_m^{MO} U^a_{mi} C_{\mu m}. \tag{28}$$

These response terms are eliminated in the usual manner (66, 123):

$$\begin{aligned} U^a_{mi} + U^a_{im} &= -S^a_{im}, \\ S^a_{im} &= \sum_{\mu\nu}^{AO} C_{\mu i} C_{\nu m} \frac{\partial S_{\mu\nu}}{\partial x_a}. \end{aligned} \tag{29}$$

A 2EI (the four molecular orbitals can be either on the QM molecules or on EFP fragments) derivative $(ij|kl)^a$ is defined as

$$\begin{aligned} (ij\,|\,kl)^a &= C^*_{\mu i} C_{\nu j} C^*_{\lambda k} C_{\sigma l} (\mu\nu\,|\,\lambda\sigma)^a \\ &= C^*_{\mu i} C_{\nu j} C^*_{\lambda k} C_{\sigma l} \left[\left(\frac{\partial \phi_\mu}{\partial a} \phi_\nu \,|\, \phi_\lambda \phi_\sigma \right) + \left(\phi_\mu \frac{\partial \phi_\nu}{\partial a} \,|\, \phi_\lambda \phi_\sigma \right) + \left(\phi_\mu \phi_\nu | \frac{\partial \phi_\lambda}{\partial a} \phi_\sigma \right) + \left(\phi_\mu \phi_\nu | \phi_\lambda \frac{\partial \phi_\sigma}{\partial a} \right) \right]. \end{aligned} \tag{30}$$

Therefore,

$$J^{Ba}_{ik} = \sum_{l \in B} (ik\,|\,ll)^a, \tag{31a}$$

$$J_{jj}^{A^a} = \sum_{m\in A} (jj \mid mm)^a, \tag{31b}$$

$$G_{ij}^{A^a} = 2J_{ij}^{A^a} - K_{ij}^{A^a} = 2\sum_{k\in A} (ij \mid kk)^a - \sum_{k\in A} (ik \mid jk)^a, \tag{31c}$$

and the one-electron nuclear attraction integral derivative is defined as

$$\begin{aligned} V_{ij}^{A^a} &= \sum_{\mu}\sum_{\nu} C_{\mu i} C_{\nu j} \sum_{I\in A} \left(\phi_\mu \left| \frac{Z_I}{R_{1I}} \right| \phi_\nu\right)^a \\ &= \sum_{\mu}\sum_{\nu} C_{\mu i} C_{\nu j} \left[\sum_{I\in A} \left(\frac{\partial \phi_\mu}{\partial a} \left| \frac{Z_I}{R_{1I}} \right| \nu\right) + \sum_{I\in A} \left(\mu \left| \frac{Z_I}{R_{1I}} \right| \frac{\partial \phi_\nu}{\partial a}\right)\right] \end{aligned} \tag{32}$$

and similarly for $V_{ik}^{B^a}$, etc. If the atomic orbital is on the EFP fragment, the particular derivative with respect to the QM nuclear displacement will be zero.

A much simpler expression for the gradient with respect to the EFP nuclei (Equation 17) is obtained because the EFP molecular orbitals do not change during the optimization:

$$\begin{aligned} \frac{\partial E^{XR}}{\partial x_b} &= -4\sum_{i\in A}\sum_{j\in B} (ij^b \mid ij) - 2\sum_{i\in A}\sum_{j\in B} S_{ij}^b \left[2(V_{ij}^A + G_{ij}^A) + \sum_{l\in B} F_{jl}^B S_{li}\right] \\ &\quad - 2\sum_{i\in A}\sum_{j\in B} S_{ij} \left[2((i \mid V^A \mid j^b) + G_{ij}^{A^b}) + \sum_{l\in B} (F_{jl}^{B^b} S_{li} + F_{jl}^B S_{li}^b)\right] \\ &\quad + 2\sum_{i\in A}\sum_{j\in B} S_{ij} \\ &\quad \times \left[\begin{array}{l} \displaystyle\sum_{k\in A} S_{kj}^b (F_{ik}^A + V_{ik}^A + 2J_{ik}^A) + S_{kj}\left(\left(i \left| \sum_{J\in B} \frac{Z_J(x_1 - x_J)}{r_{1J}^3} \right| k\right) + 2\sum_{l\in B} (ik \mid l^b l)\right) \\ \displaystyle + S_{ij}^b (V_{jj}^A + 2J_{jj}^A) + S_{ij} V_{jj}^{A^b} + 2S_{ij} J_{jj}^{A^b} - \sum_{k\in A} S_{kj}^b (ik \mid jj) - \sum_{k\in A} S_{kj} (ik \mid jj)^b \end{array}\right]. \end{aligned} \tag{33}$$

3.3. QM-EFP Dispersion

The dispersion energy arises from the interaction between an instantaneous multipole on molecule A and an induced multipole on molecule B. The derivation of the QM-EFP dispersion energy (67) expression begins from the Rayleigh-Schrödinger perturbation theory expression:

$$E^{disp} = -\sum_{\substack{m\neq 0\\ n\neq 0}} \frac{\langle 0_A 0_B \mid \hat{V} \mid mn\rangle \langle mn \mid \hat{V} \mid 0_A 0_B\rangle}{E_m^A + E_n^B - E_0^A - E_0^B}, \tag{34}$$

where m and n are excited states of molecules A and B, respectively; 0 is the ground state; E_m^A is the energy of molecule A in state m; and $\hat{V}$ is the perturbation operator, which encompasses all electrostatic interactions between the molecules and here is expressed as the multipole expansion. The first term in the London expansion of the dispersion energy, $E^{disp} = C_6/R^6 + C_7/R^7 + \ldots$, corresponds to the dipole-dipole term of $\hat{V}$, i.e.,

$$\hat{V} \approx T_{ab}^{AB} \hat{\mu}^A \hat{\mu}^B. \tag{35}$$

Here T_{ab}^{AB} is the electric field gradient, scaling as $1/R^3$ in the intermolecular distance, and $\hat{\mu}^A$ is the dipole operator on molecule A. Substituting Equation 35 into Equation 34 and separating the

resulting expression into integrals on A and integrals on B, one finds that the C_6/R^6 term becomes

$$E_6^{disp} = -\frac{2\hbar}{\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB} T_{cd}^{AB} \int_0^\infty d\omega \sum_{\substack{m\neq 0\\ n\neq 0}} \frac{\omega_{m0}^A \langle 0_A|\hat{\mu}_a^A|m\rangle\langle m|\hat{\mu}_c^A|0_A\rangle}{\hbar((\omega_{m0}^A)^2+\omega^2)} \frac{\omega_{n0}^B \langle 0_B|\hat{\mu}_b^B|n\rangle\langle n|\hat{\mu}_d^B|0_B\rangle}{\hbar((\omega_{n0}^B)^2+\omega^2)}. \quad (36)$$

Equation 36 can be recast in terms of dynamic polarizability tensors α over the imaginary frequency range $i\omega$:

$$\alpha_{ab}(i\omega) = 2\sum_n \frac{\omega_{n0}\langle 0|\hat{\mu}_a|n\rangle\langle n|\hat{\mu}_b|0\rangle}{\hbar(\omega_{n0}^2+\omega^2)}, \quad \text{where} \quad \omega_{n0} = \omega_n - \omega_0. \quad (37)$$

In EFP-EFP dispersion, both the portion of Equation 36 relating to A and the portion relating to B are recast in terms of α; α is calculated during the process of generating the fragment potential (prior to using the potentials in an EFP calculation). However, because the calculation of the dynamic polarizability tensors is computationally expensive, it is not practical to calculate α on the fly for the QM molecule (A). Therefore, only that portion of Equation 36 that corresponds to EFP molecule B is expressed in terms of α, giving

$$E_6^{EFP-AI} = -\frac{1}{\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB} T_{cd}^{AB} \sum_{m\neq 0} \langle 0_A|\hat{\mu}_a^A|m\rangle\langle m|\hat{\mu}_c^A|0_A\rangle \int_0^\infty d\omega \frac{\omega_{m0}^A}{(\omega_{m0}^A)^2+\omega^2}\alpha_{bd}^B(i\omega). \quad (38)$$

An approximation to convert from sum over states to CMOs is invoked. Additionally, a distributed polarizability tensor is used on EFP molecule B. An isotropic approximation is made to eliminate off-diagonal terms, which do not contribute significantly to the total dispersion energy and are time-consuming to calculate. The product of electrostatic tensors T^{AB} is expressed as $6/R_{kj}^6$, where R_{kj} is the distance between QM occupied orbitals k and EFP LMOs j. This yields

$$E_6^{EFP-AI} = -\frac{6}{\pi}\sum_{j\in B}\sum_k^{occ}\sum_r^{vir} \frac{1}{R_{kj}^6}\langle k|\hat{\mu}|r\rangle\langle r|\hat{\mu}|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{(\omega_{rk}^A)^2+\omega^2}\alpha^j(i\omega). \quad (39)$$

Because previous EFP studies have suggested that an LMO-based formalism gives faster convergence and superior results, Equation 39 is recast in terms of LMOs. Boys localization, performed on the valence orbitals, produces orthogonal transformation matrices $L_{k\ell}$, which express LMOs ℓ in terms of canonical MOs k. The EFP-QM C_6 expression becomes

$$C_6^{\ell v} = \sum_a^{x,y,z}\sum_k^{\text{valence}}\sum_{k'}^{\text{valence}} L_{k\ell}\left[\sum_r^{vir}\langle k|\mu_a|r\rangle\langle r|\mu_a|k'\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{(\omega_{rk}^A)^2+\omega^2}\alpha^v(i\omega)\right]L_{k'\ell}. \quad (40)$$

The integral over the imaginary frequency range is calculated using a 12-point Gauss-Legendre numerical quadrature. The differences between the virtual and occupied orbital energies $\varepsilon_{rk} = \varepsilon_r - \varepsilon_k$ are used in place of ω_{rk}.

The final form of the EFP-QM dispersion energy is

$$E_{disp}^{EFP-AI} = \frac{4}{3}\sum_{\ell\in A}\sum_{v\in B}\frac{F_6^{\ell v} C_6^{\ell v}}{R_{\ell v}^6}. \quad (41)$$

Multiplication by 4/3 serves as an estimation of higher-order multipole contributions. $F_6^{\ell v}$ is a damping function, which accounts for short-range exchange dispersion and charge penetration effects. Two damping functions are available: a Tang-Toennies and a parameter-free overlap-based damping function (50, 51).

For a set of dimers examined (benzene, CH_2Cl_2, H_2O, NH_3, CH_3OH, HF, CH_4, Ar, and H_2 dimers), the EFP-QM dispersion agrees closely with the EFP-EFP dispersion for a given damping

Table 3 Effective fragment potential (EFP)-EFP and EFP–quantum mechanics (QM) C_6 coefficients

	QM (H donor)-EFP (H acceptor)[a]	EFP (H donor)-QM (H acceptor)[a]	EFP-EFP[a]	Expt.
HF	17.0 (−10.5%)	15.4 (−18.9%)	15.3 (−19.5%)	19.0[b]
H_2O	43.0 (−5.3%)	40.6 (−10.6%)	39.3 (−13.4%)	45.4[b]
NH_3	81.2 (−7.0%)	82.1 (−6.0%)	78.1 (−10.5%)	87.3[b]
CH_3OH	197.7 (−11.1%)	197.2 (−11.3%)	195.8 (−11.9%)	222.2[b]
	EFP-QM		**EFP-EFP**	**Expt.**
Ar	67.4 (+4.8%)		60.6 (−5.8%)	64.3[c]
H_2	10.2 (−15.7%)		11.4 (−5.8%)	12.1[b]
CH_4	120.3 (−7.2%)		120.4 (−7.1%)	129.6[b]
CH_2Cl_2	843.0		755.8	—
C_6H_6	2,087 (+21.1%)		1,805 (+4.8%)	1,723[c]

C_6 coefficients were calculated for dimers at equilibrium geometries, except for benzene (C_6H_6), which was calculated for a sandwich structure. For nonsymmetrical dimers, the C_6 coefficient may vary depending on which monomer is modeled with EFP and which with the QM method (Hartree-Fock); where applicable, both values are shown.

[a]Calculated using the 6-311++G(3df,3p) basis set for all dimers other than benzene, for which the 6-311++G(3df,2p) basis set was used.

[b]From References 12 and 128.

[c]From Reference 129.

function. Both methods agree well with symmetry adapted perturbation theory (4, 5) dispersion plus exchange dispersion values for a given basis set. Example C_6 coefficient values appear in **Table 3**, with the percent error relative to experimental values.

4. SUMMARY AND PROGNOSIS

The EFP is a very accurate method for treating intermolecular interactions, including solvent effects. A generalized EFP (sometimes called EFP2) contains all of the essential physics, while avoiding the need for empirically fitted parameters. Recently, the most computationally demanding EFP term, the charge transfer (CT) interaction, has been made much more efficient by including only the most important part of the virtual space, the part that is defined by the valence analog of the valence occupied space: the QUAMBOs. The use of QUAMBOs decreases the computational cost by about a factor of two, while maintaining the accuracy of the original method.

For EFP to be a truly general and useful method, it is important to have a complete interface with QM methods. This has now been accomplished for the exchange repulsion and dispersion EFP energy components. Expressions for the analytic gradients for each of these terms have been derived, and the implementations are in progress. The CT term is the one remaining component to be interfaced with QM methods. Even though this term is the most time-consuming EFP component, its cost will be very small relative to QM methods, especially since the most sensible QM methods to use with EFP are those that include electron correlation [e.g., DFT, MP2, CCSD(T)].

Now, consider the remaining limitations of the EFP method. A time-consuming aspect of the EFP process is the preliminary setup, called MAKEFP, which requires preparatory HF and time-dependent HF calculations. If one uses the recommended large basis set for a large molecule, this

can require considerable time and resources. For this reason, the establishment of the EFP library has been important, and this library needs to be expanded to include many more molecules.

Currently, EFPs are rigid species that are not able to change their geometries in the process of simulating chemical processes. One way to accomplish flexibility is to employ the new effective fragment molecular orbital (EFMO) method that incorporates most features of the EFP method into the fragment molecular orbital method (124). However, it is desirable to have at least torsional flexibility even when no QM component is present. Torsional flexibility can be achieved by splitting a molecule along a chosen bond into separate effective fragments and then substituting the broken covalent bonds by harmonic potentials. All torsional and noncovalent interactions between the fragments are treated in the standard EFP manner. An effort in this direction is in progress (125).

All of the EFP components are expressed in terms of a series of some kind. Some of these components, notably the dispersion, are terminated after the leading term. There will be instances (e.g., for charged species) when higher-order terms will be important.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by a grant (to M.S.G.) from the Air Force Office of Scientific Research. L.V.S. acknowledges support from the National Science Foundation (grant CHE-0955419) and Purdue University.

LITERATURE CITED

1. Moller C, Plesset MS. 1934. Note on an approximation treatment for many-electron systems. *Phys. Rev.* 46:618–22
2. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. 1989. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* 157:479–83
3. Bartlett RJ, Musiał M. 2007. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* 79:291–352
4. Jeziorski B, Moszynski R, Szalewicz K. 1994. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* 94:1887–930
5. Moszynski R. 1996. Symmetry-adapted perturbation theory for the calculation of Hartree-Fock interaction energies. *Mol. Phys.* 88:741–58
6. Gresh N, Cisneros GA, Darden TA, Piquemal JP. 2007. Anisotropic, polarizable molecular mechanics studies of inter- and intramolecular interactions and ligand-macromolecule complexes: a bottom-up strategy. *J. Chem. Theory Comput.* 3:1960–86
7. **Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, et al. 1996. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.* 105:1968–86**
8. Gordon MS, Freitag MA, Bandyopadhyay P, Jensen JH, Kairys V, Stevens WJ. 2001. The effective fragment potential method: a QM-based MM approach to modeling environmental effects in chemistry. *J. Phys. Chem. A* 105:293–307
9. **Gordon MS, Slipchenko LV, Li H, Jensen JH. 2007. The effective fragment potential: a general method for predicting intermolecular forces. *Annu. Rep. Comput. Chem.* 3:177–93**
10. **Ghosh D, Kosenkov D, Vanovschi V, Williams CF, Herbert JM, et al. 2010. Noncovalent interactions in extended systems described by the effective fragment potential method: theory and application to nucleobase oligomers. *J. Phys. Chem. A* 114:12739–54**

7. Describes original implementation of the EFP1 water model.

9. Provides an overview of the EFP2 method.

10. Describes the implementation of the EFP method in Q-Chem and the EFP fragment library.

11. Stone AJ. 1981. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* 83:233–39
12. Stone AJ. 1996. *The Theory of Intermolecular Forces*. New York: Oxford Univ. Press
13. Becke AD. 1993. Density-functional thermochemistry. 3. The role of exact exchange. *J. Chem. Phys.* 98:5648–52
14. Lee CT, Yang WT, Parr RG. 1988. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. *Phys. Rev. B* 37:785–89
15. Adamovic I, Freitag MA, Gordon MS. 2003. Density functional theory based effective fragment potential method. *J. Chem. Phys.* 118:6725–32
16. Yoo S, Zahariev F, Sok S, Gordon MS. 2008. Solvent effects on optical properties of molecules: a combined time-dependent density functional theory/effective fragment potential approach. *J. Chem. Phys.* 129:144112
17. Arora P, Slipchenko LV, Webb SP, Defusco A, Gordon MS. 2010. Solvent-induced frequency shifts: configuration interaction singles combined with the effective fragment potential method. *J. Phys. Chem. A* 114:6742–50
18. DeFusco A, Ivanic J, Schmidt MW, Gordon MS. 2011. Solvent-induced shifts in electronic spectra of uracil. *J. Phys. Chem. A* 115:4574–82
19. Leang SS, Slipchenko LV, Gordon MS. 2012. Manuscript in preparation
20. Chen W, Gordon MS. 1996. The effective fragment model for solvation: internal rotation in formamide. *J. Chem. Phys.* 105:11081–90
21. Merrill GN, Gordon MS. 1998. Study of small water clusters using the effective fragment potential model. *J. Phys. Chem. A* 102:2650–57
22. Webb SP, Gordon MS. 1999. Solvation of the Menshutkin reaction: a rigorous test of the effective fragment method. *J. Phys. Chem. A* 103:1265–73
23. Bandyopadhyay P, Gordon MS. 2000. A combined discrete/continuum solvation model: application to glycine. *J. Chem. Phys.* 113:1104–9
24. Day PN, Pachter R, Gordon MS, Merrill GN. 2000. A study of water clusters using the effective fragment potential and Monte Carlo simulated annealing. *J. Chem. Phys.* 112:2063–73
25. Bandyopadhyay P, Gordon MS, Mennucci B, Tomasi J. 2002. An integrated effective fragment–polarizable continuum approach to solvation: theory and application to glycine. *J. Chem. Phys.* 116:5023–32
26. Adamovic I, Gordon MS. 2005. Solvent effects on the S_N2 reaction: application of the density functional theory-based effective fragment potential method. *J. Phys. Chem. A* 109:1629–36
27. Mullin JM, Gordon MS. 2009. Alanine: Then there was water. *J. Phys. Chem. B* 113:8657–69
28. Mullin JM, Gordon MS. 2009. Water and alanine: from puddles(32) to ponds(49). *J. Phys. Chem. B* 113:14413–20
29. DeFusco A, Minezawa N, Slipchenko LV, Zahariev F, Gordon MS. 2011. Modeling solvent effects on electronic excited states. *J. Phys. Chem. Lett.* 2:2184–92
30. Netzloff HM, Gordon MS. 2004. The effective fragment potential: small clusters and radial distribution functions. *J. Chem. Phys.* 121:2711–14
31. Kina D, Nakayama A, Noro T, Taketsugu T, Gordon MS. 2008. Ab initio QM/MM molecular dynamics study on the excited-state hydrogen transfer of 7-azaindole in water solution. *J. Phys. Chem. A* 112:9675–83
32. Kina D, Arora P, Nakayama A, Noro T, Gordon MS, Taketsugu T. 2009. Ab initio QM/MM excited-state molecular dynamics study of coumarin 151 in water solution. *Int. J. Quantum Chem.* 109:2308–18
33. Atadinc F, Günaydin H, Özen AS, Aviyente V. 2005. A quantum mechanical approach to the kinetics of the hydrogen abstraction reaction $H_2O_2 + \bullet OH \rightarrow HO_2 + H_2O$. *Int. J. Chem. Kinetics* 37:502–14
34. Ferreira DEC, Florentino BPD, Rocha WR, Nome F. 2009. Quantum mechanical/effective fragment potential (QM/EFP) study of phosphate monoester aminolysis in aqueous solution. *J. Phys. Chem. B* 113:14831–36
35. Bandyopadhyay P. 2008. Assessment of two surface Monte Carlo (TSMC) method to find stationary points of $(H_2O)_{15}$ and $(H_2O)_{20}$ clusters. *Theor. Chem. Acc.* 120:307–12

29. Provides an overview of the QM/EFP methodology and applications for the electronic excited states.

36. Kemp DA, Gordon MS. 2005. Theoretical study of the solvation of fluorine and chlorine anions by water. *J. Phys. Chem. A* 109:7688–99
37. Kemp DA, Gordon MS. 2008. An interpretation of the enhancement of the water dipole moment due to the presence of other water molecules. *J. Phys. Chem. A* 112:4885–94
38. Merrill GN, Webb SP. 2003. Anion-water clusters $A^-(H_2O)_{1-6}$, A = OH, F, SH, Cl, and Br: an effective fragment potential test case. *J. Phys. Chem. A* 107:7852–60
39. Merrill GN, Webb SP, Bivin DB. 2003. Formation of alkali metal/alkaline earth cation water clusters, $M(H_2O)_{1-6}$, M = Li^+, Na^+, K^+, Mg_2^+, and Ca_2^+: an effective fragment potential (EFP) case study. *J. Phys. Chem. A* 107:386–96
40. Merrill GN, Webb SP. 2004. The application of the effective fragment potential method to molecular anion solvation: a study of ten oxyanion-water clusters, $A^-(H_2O)_{1-4}$. *J. Phys. Chem. A* 108:833–39
41. Merrill GN, Fletcher GD. 2008. A microsolvation approach to the prediction of the relative enthalpies and free energies of hydration for ammonium ions. *Theor. Chem. Acc.* 120:5–22
42. Chandrakumar KRS, Ghanty TK, Ghosh SK, Mukherjee T. 2007. Hydration of uranyl cations: effective fragment potential approach. *J. Mol. Struct.* 807:93–99
43. Petersen CP, Gordon MS. 1999. Solvation of sodium chloride: an effective fragment study of $NaCl(H_2O)_n$. *J. Phys. Chem. A* 103:4162–66
44. Yoshikawa A, Morales JA. 2004. The onset of dissociation in the aqueous LiOH clusters: a solvation study with the effective fragment potential model and quantum mechanics methods. *J. Mol. Struct.* 681:27–40
45. Balawender R, Safi B, Geerlings P. 2001. Solvent effect on the global and atomic DFT-based reactivity descriptors using the effective fragment potential model: solvation of ammonia. *J. Phys. Chem. A* 105:6703–10
46. Safi B, Balawender R, Geerlings P. 2001. Solvent effect on electronegativity, hardness, condensed Fukui functions, and softness, in a large series of diatomic and small polyatomic molecules: use of the EFP model. *J. Phys. Chem. A* 105:11102–9
47. Day PN, Pachter R. 1997. A study of aqueous glutamic acid using the effective fragment potential method. *J. Chem. Phys.* 107:2990–99
48. Song J, Gordon MS, Deakyne CA, Zheng WC. 2004. Theoretical investigations of acetylcholine (ACh) and acetylthiocholine (ATCh) using ab initio and effective fragment potential methods. *J. Phys. Chem. A* 108:11419–32
49. Sok S, Willow SY, Zahariev F, Gordon MS. 2011. Solvent-induced shift of the lowest singlet $\pi \rightarrow \pi^*$ charge-transfer excited state of *p*-nitroaniline in water: an application of the TDDFT/EFP1 method. *J. Phys. Chem. A* 115:9801–9
50. Tang KT, Toennies JP. 1984. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* 80:3726–41
51. Slipchenko LV, Gordon MS. 2009. Damping functions in the effective fragment potential method. *Mol. Phys.* 107:999–1016
52. Adamovic I, Gordon MS. 2006. Methanol-water mixtures: a microsolvation study using the effective fragment potential method. *J. Phys. Chem. A* 110:10267–73
53. Adamovic I, Li H, Lamm MH, Gordon MS. 2006. Modeling styrene-styrene interactions. *J. Phys. Chem. A* 110:519–25
54. Slipchenko LV, Gordon MS. 2007. Electrostatic energy in the effective fragment potential method: theory and application to benzene dimer. *J. Comput. Chem.* 28:276–91
55. Smith T, Slipchenko LV, Gordon MS. 2008. Modeling π-π interactions with the effective fragment potential method: the benzene dimer and substituents. *J. Phys. Chem. A* 112:5286–94
56. Smith QA, Gordon MS, Slipchenko LV. 2011. Benzene-pyridine interactions predicted by the effective fragment potential method. *J. Phys. Chem. A* 115:4598–609
57. Smith QA, Gordon MS, Slipchenko LV. 2011. Effective fragment potential study of the interaction of DNA bases. *J. Phys. Chem. A* 115:11269–76
58. Hands MD, Slipchenko LV. 2012. Intermolecular interactions in complex liquids: effective fragment potential investigation of water–tert-butanol mixtures. *J. Phys. Chem. B* 116:2775–86
59. Slipchenko LV, Gordon MS. 2009. Water-benzene interactions: an effective fragment potential and correlated quantum chemistry study. *J. Phys. Chem. A* 113:2092–102

60. Pranami G, Slipchenko L, Lamm MH, Gordon MS. 2009. Coarse-grained intermolecular potentials derived from the effective fragment potential: application to water, benzene, and carbon tetrachloride. In *Multi-Scale Quantum Models for Biocatalysis*, ed. DM York, T-S Lee, pp. 197–218. New York: Springer
61. Flick JC, Kosenkov D, Hohenstein EG, Sherrill CD, Slipchenko LV. 2012. Accurate prediction of non-covalent interaction energies with the effective fragment potential method: comparison of energy components to symmetry-adapted perturbation theory for the S22 test set. *J Chem. Theory Comput.* 8:2835–43
62. Leang SS, Pruitt SR, Xu P, Gordon MS. 2012. Manuscript in preparation
63. Podeszwa R, Cencek W, Szalewicz K. 2012. Efficient calculations of dispersion energies for nanoscale systems from coupled density response functions. *J. Chem. Theory Comput.* 8:1963–69
64. Li H, Gordon MS, Jensen JH. 2006. Charge transfer interaction in the effective fragment potential method. *J. Chem. Phys.* 124:214108
65. Xu P, Gordon MS. 2012. Manuscript in preparation.
66. Kemp D, Rintelman J, Gordon M, Jensen J. 2010. Exchange repulsion between effective fragment potentials and ab initio molecules. *Theor. Chem. Acc.* 125:481–91
67. Smith QA, Ruedenberg K, Gordon MS, Slipchenko LV. 2012. The dispersion interaction between quantum mechanics and effective fragment potential molecules. *J. Chem. Phys.* 136:244107
68. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, et al. 1993. General atomic and molecular electronic-structure system. *J. Comput. Chem.* 14:1347–63
69. Gordon MS, Schmidt MW. 2005. Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry*, ed. CE Dykstra, G Frenking, KS Kim, GE Scuseria, pp. 1167–89. Amsterdam: Elsevier
70. Kong J, White CA, Krylov AI, Sherrill D, Adamson RD, et al. 2000. Q-Chem 2.0: a high-performance ab initio electronic structure program package. *J. Comput. Chem.* 21:1532–48
71. Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, et al. 2006. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* 8:3172–91
72. Jensen JH, Day PN, Gordon MS, Basch H, Cohen D, et al. 1994. Effective fragment method for modeling intermolecular hydrogen-bonding effects on quantum-mechanical calculations. *ACS Symp. Ser.* 569:139–51
73. Minezawa N, Silva ND, Zahariev F, Gordon MS. 2011. Implementation of the analytic energy gradient for the combined time-dependent density functional theory/effective fragment potential method: application to excited-state molecular dynamics simulations. *J. Chem. Phys.* 134:054111
74. Gordon MS, Mullin JM, Pruitt SR, Roskop LB, Slipchenko LV, Boatz JA. 2009. Accurate methods for large molecular systems. *J. Phys. Chem. B* 113:9646–63
75. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV. 2011. Fragmentation methods: a route to accurate calculations on large systems. *Chem. Rev.* 112:632–72
76. Li H, Netzloff HM, Gordon MS. 2006. Gradients of the polarization energy in the effective fragment potential method. *J. Chem. Phys.* 125:194103
77. Amos RD, Handy NC, Knowles PJ, Rice JE, Stone AJ. 1985. Ab initio prediction of properties of CO_2, NH_3, and CO_2-NH_3. *J. Phys. Chem.* 89:2186–92
78. Adamovic I, Gordon MS. 2005. Dynamic polarizability, dispersion coefficient C_6 and dispersion energy in the effective fragment potential method. *Mol. Phys.* 103:379–87
79. Freitag MA, Gordon MS, Jensen JH, Stevens WJ. 2000. Evaluation of charge penetration between distributed multipolar expansions. *J. Chem. Phys.* 112:7300–6
80. Jensen JH, Gordon MS. 1996. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. *Mol. Phys.* 89:1313–25
81. Jensen JH, Gordon MS. 1998. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application to the effective fragment potential method. *J. Chem. Phys.* 108:4772–82
82. Jensen JH. 2001. Intermolecular exchange-induction and charge transfer: derivation of approximate formulas using nonorthogonal localized molecular orbitals. *J. Chem. Phys.* 114:8775–83
83. Mulliken RS. 1952. Molecular compunds and their spectra. 2. *J. Am. Chem. Soc.* 74:811–24

75. Presents a thorough review of fragmentation techniques.

84. Ben-Amotz D. 2011. Unveiling electron promiscuity. *J. Phys. Chem. Lett.* 2:1216–22
85. Zhao Z, Rogers DM, Beck TL. 2010. Polarization and charge transfer in the hydration of chloride ions. *J. Chem. Phys.* 132:014502
86. Ramesh SG, Re SY, Hynes JT. 2008. Charge transfer and OH vibrational frequency red shifts in nitrate-water clusters. *J. Phys. Chem. A* 112:3391–98
87. Cappa CD, Smith JD, Messer BM, Cohen RC, Saykally RJ. 2006. Effects of cations on the hydrogen bond network of liquid water: new results from X-ray absorption spectroscopy of liquid microjets. *J. Phys. Chem. B* 110:5301–9
88. Robertson WH, Johnson MA, Myshakin EM, Jordan KD. 2002. Isolating the charge-transfer component of the anionic H bond via spin suppression of the intracluster proton transfer reaction in the $NO^- \cdot H_2O$ entrance channel complex. *J. Phys. Chem. A* 106:10010–14
89. Thompson WH, Hynes JT. 2000. Frequency shifts in the hydrogen-bonded OH stretch in halide-water clusters: the importance of charge transfer. *J. Am. Chem. Soc.* 122:6278–86
90. Lee AJ, Rick SW. 2011. The effects of charge transfer on the properties of liquid water. *J. Chem. Phys.* 134:184507
91. Vacha R, Rick SW, Jungwirth P, de Beer AGF, de Aguiar HB, et al. 2011. The orientation and charge of water at the hydrophobic oil droplet-water interface. *J. Am. Chem. Soc.* 133:10204–10
92. Piquemal JP, Cisneros GA, Reinhardt P, Gresh N, Darden TA. 2006. Towards a force field based on density fitting. *J. Chem. Phys.* 124:104101

93. Introduces the QM/MM idea.

93. Warshel A, Levitt M. 1976. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103:227–49

94. Warshel A. 1979. Calculations of chemical processes in solutions. *J. Phys. Chem.* 83:1640–52
95. Luzhkov V, Warshel A. 1991. Microscopic calculations of solvent effects on absorption spectra of conjugated molecules. *J. Am. Chem. Soc.* 113:4491–99
96. Gao J. 1996. Hybrid quantum and molecular mechanical simulations: an alternative avenue to solvent effects in organic chemistry. *Acc. Chem. Res.* 29:298–305
97. Lin YL, Gao JL. 2007. Solvatochromic shifts of the $n \rightarrow \pi^*$ transition of acetone from steam vapor to ambient aqueous solution: a combined configuration interaction QM/MM simulation study incorporating solvent polarization. *J. Chem. Theory Comput.* 3:1484–93
98. Gao JL, Byun K. 1997. Solvent effects on the $n \rightarrow \pi$ transition of pyrimidine in aqueous solution. *Theor. Chem. Acc.* 96:151–56
99. Thompson MA, Schenter GK. 1995. Excited states of the bacteriochlorophyll-b dimer of *Rhodopseudomonas viridis*: a QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* 99:6374–86
100. Poulsen TD, Kongsted J, Osted A, Ogilby PR, Mikkelsen KV. 2001. The combined multiconfigurational self-consistent-field/molecular mechanics wave function approach. *J. Chem. Phys.* 115:2393–400
101. Kongsted J, Osted A, Mikkelsen KV, Christiansen O. 2002. Dipole and quadrupole moments of liquid water calculated within the coupled cluster/molecular mechanics method. *Chem. Phys. Lett.* 364:379–86
102. Kongsted J, Osted A, Mikkelsen KV, Christiansen O. 2002. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Mol. Phys.* 100:1813–28
103. Kongsted J, Osted A, Mikkelsen KV, Christiansen O. 2003. Linear response functions for coupled cluster/molecular mechanics including polarization interactions. *J. Chem. Phys.* 118:1620–33
104. Aidas K, Kongsted J, Osted A, Mikkelsen KV, Christiansen O. 2005. Coupled cluster calculation of the $n \rightarrow \pi^*$ electronic transition of acetone in aqueous solution. *J. Phys. Chem. A* 109:8001–10
105. Thiel W. 2009. QM/MM methodology: fundamentals, scope, and limitations. In *Multiscale Simulation Methods in Molecular Sciences*, ed. H Grotendorst, pp. 203–14. Jülich, Ger.: NIC
106. Kosenkov D, Slipchenko LV. 2010. Solvent effects on the electronic transitions of *p*-nitroaniline: a QM/EFP study. *J. Phys. Chem. A* 115:392–401
107. Slipchenko LV. 2010. Solvation of the excited states of chromophores in polarizable environment: orbital relaxation versus polarization. *J. Phys. Chem. A* 114:8824–30
108. Ghosh D, Isayev O, Slipchenko LV, Krylov AI. 2011. The effect of solvation on vertical ionization energy of thymine: from microhydration to bulk. *J. Phys. Chem. A* 115:6028–38

109. Nagata T, Fedorov DG, Kitaura K, Gordon MS. 2009. A combined effective fragment potential–fragment molecular orbital method. I. The energy expression and initial applications. *J. Chem. Phys.* 131:024101
110. Nagata T, Fedorov DG, Sawada T, Kitaura K, Gordon MS. 2011. A combined effective fragment potential–fragment molecular orbital method. II. Analytic gradient and application to the geometry optimization of solvated tetraglycine and chignolin. *J. Chem. Phys.* 134:034110
111. Li H, Gordon MS. 2007. Polarization energy gradients in combined quantum mechanics, effective fragment potential, and polarizable continuum model calculations. *J. Chem. Phys.* 126:124112
112. Gaenko A, Windues TL, Sosonkina M, Gordon MS. 2012. *J. Chem. Theory Comput.* Manuscript submitted
113. Kaliman I, Slipchenko LV. 2012. Manuscript in preparation
114. Jurečka P, Šponer J, Černý J, Hobza P. 2006. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* 8:1985–93
115. Řezáč J, Riley KE, Hobza P. 2011. S66: a well-balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* 7:2427–38
116. Krishnan R, Binkley JS, Seeger R, Pople JA. 1980. Self-consistent molecular-orbital methods 0.20. Basis set for correlated wave functions. *J. Chem. Phys.* 72:650–54
117. Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PV. 1983. Efficient diffuse function augmented basis sets for anion calculations 0.3. The 3-21+G basis set for 1st-row elements, Li-F. *J. Comput. Chem.* 4:294–301
118. Frisch MJ, Pople JA, Binkley JS. 1984. Self-consistent molecular-orbital methods 0.25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* 80:3265–69
119. Lu WC, Wang CZ, Schmidt MW, Bytautas L, Ho KM, Ruedenberg K. 2004. Molecule intrinsic minimal basis sets. I. Exact resolution of ab initio optimized molecular orbitals in terms of deformed atomic minimal-basis orbitals. *J. Chem. Phys.* 120:2629–37
120. Stevens WJ, Fink WH. 1987. Frozen fragment reduced variational space analysis of hydrogen bonding interactions: application to the water dimer. *Chem. Phys. Lett.* 139:15–22
121. Chen W, Gordon MS. 1996. Energy recomposition analyses for many-body interaction and applications to water complexes. *J. Phys. Chem.* 100:14316–28
122. Jensen JH. 1996. Modeling intermolecular exchange integrals between nonorthogonal molecular orbitals. *J. Chem. Phys.* 104:7795–96
123. Yamaguchi Y, Masamura Y, Goddard JD, Schaefer HF III. 1994. *A New Dimension to Quantum Chemistry*. New York: Oxford Univ. Press
124. Steinmann C, Fedorov DG, Jensen JH. 2010. Effective fragment molecular orbital method: a merger of the effective fragment potential and fragment molecular orbital methods. *J. Phys. Chem. A* 114:8705–12
125. Haupert L, Slipchenko LV. 2012. Manuscript in preparation
126. Paton RS, Goodman JM. 2009. Hydrogen bonding and π-stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions. *J. Chem. Inf. Model.* 49:944–55
127. Takatani T, Hohenstein EG, Malagoli M, Marshall MS, Sherrill CD. 2010. Basis set consistent revision of the S22 test set of noncovalent interaction energies. *J. Chem. Phys.* 132:144104
128. Meath WJ, Kumar A. 1990. Reliable isotropic and anisotropic dipolar dispersion energies, evaluated using constrained dipole oscillator strength techniques, with application to interactions involving H_2, N_2, and the rare gases. *Int. J. Quantum Chem.* 38:501–20
129. Gross EK, Ullrich CA, Gossmann UJ. 1995. Density functional theory of time-dependent systems. In *Density Functional Theory*, ed. EKU Gross, RM Dreizler, pp. 149–72. New York: Plenum
130. Jungwirth P, Tobias DJ. 2006. Specific ion effects at the air/water interface. *Chem. Rev.* 106:1259–81
131. Ghosal S, Hemminger JC, Bluhm H, Mun BS, Hebenstreit ELD, et al. 2005. Electron spectroscopy of aqueous solution interfaces reveals surface enhancement of halides. *Science* 307:563–66
132. Petersen PB, Saykally RJ. 2008. Is the liquid water surface basic or acidic? Macroscopic versus molecular-scale investigations. *Chem. Phys. Lett.* 458:255–61
133. Xu M, Spinney R, Allen HC. 2009. Water structure at the air-aqueous interface of divalent cation and nitrate solutions. *J. Phys. Chem. B* 113:4102–10

134. Fan YB, Chen X, Yang LJ, Cremer PS, Gao YQ. 2009. On the structure of water at the aqueous/air interface. *J. Phys. Chem. B* 113:11672–79
135. Berne BJ, Weeks JD, Zhou RH. 2009. Dewetting and hydrophobic interaction in physical and biological systems. *Annu. Rev. Phys. Chem.* 60:85–103
136. Larsen RE, Glover WJ, Schwartz BJ. 2010. Does the hydrated electron occupy a cavity? *Science* 329:65–69
137. Bakker HJ, Skinner JL. 2010. Vibrational spectroscopy as a probe of structure and dynamics in liquid water. *Chem. Rev.* 110:1498–517
138. Herbert JM, Jacobson LD. 2011. Nature's most squishy ion: the important role of solvent polarization in the description of the hydrated electron. *Int. Rev. Phys. Chem.* 30:1–48
139. Schenter GK, Glendening ED. 1996. Natural energy decomposition analysis: the linear response electrical self energy. *J. Phys. Chem.* 100:17152–56
140. Glendening ED. 2005. Natural energy decomposition analysis: extension to density functional methods and analysis of cooperative effects in water clusters. *J. Phys. Chem. A* 109:11936–40
141. Kumar R, Wang F-F, Jenness GR, Jordan KD. 2010. A second generation distributed point polarizable water model. *J. Chem. Phys.* 132:014309
142. Schmidt MW, Ruedenberg K, Gordon MS. 2012. Manuscript in preparation

Annual Review of
Physical Chemistry
Volume 64, 2013

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Physical Chemistry* articles may be found at http://physchem.annualreviews.org/errata.shtml