Effective Fragment Method for Modeling Intermolecular Hydrogen-Bonding Effects on Quantum Mechanical Calculations

Jan H. Jenson  
*Iowa State University*

Paul N. Day  
*Iowa State University*

Mark S. Gordon  
*Iowa State University*, mgordon@iastate.edu

Harold Basch  
*Bar Ilan University*

Drora Cohen  
*Bar Ilan University*

Follow this and additional works at: [http://lib.dr.iastate.edu/chem_pubs](http://lib.dr.iastate.edu/chem_pubs)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/chem_pubs/268](http://lib.dr.iastate.edu/chem_pubs/268). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).
Effective Fragment Method for Modeling Intermolecular Hydrogen-Bonding Effects on Quantum Mechanical Calculations

Abstract
The effective fragment potential (EFP) method is introduced as a way to model the effect of intermolecular hydrogen bonds on molecules described by standard quantum mechanical (QM) methods. The chemical system of interest is divided into two regions: an "active region" (AR) described by QM, and a "spectator region" (SR) that influences the AR via hydrogen bonding. The SR is replaced by an EFP which describes the interaction by three terms: electrostatics, polarization, and exchange repulsion. The potentials are derived from separate ab initio calculations on the prototypical interactions represented by the spectator region. The method is currently being implemented in the quantum chemistry code GAMESS. Some applications involving water in the SR are presented.

Disciplines
Chemistry

Comments

Rights
Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

Authors
Chapter 9

Effective Fragment Method for Modeling Intermolecular Hydrogen-Bonding Effects on Quantum Mechanical Calculations

Jan H. Jensen¹, Paul N. Day¹,⁴, Mark S. Gordon¹, Harold Basch², Drora Cohen², David R. Garmer³,⁵, Morris Kraus³, and Walter J. Stevens³

¹Department of Chemistry, Iowa State University, Ames, IA 50011
²Department of Chemistry, Bar Ilan University, Ramat Gan 52100, Israel
³Center for Advanced Research in Biotechnology, National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD 20850

The effective fragment potential (EFP) method is introduced as a way to model the effect of intermolecular hydrogen bonds on molecules described by standard quantum mechanical (QM) methods. The chemical system of interest is divided into two regions: an “active region” (AR) described by QM, and a “spectator region” (SR) that influences the AR via hydrogen bonding. The SR is replaced by an EFP which describes the interaction by three terms: electrostatics, polarization, and exchange repulsion. The potentials are derived from separate ab initio calculations on the prototypical interactions represented by the spectator region. The method is currently being implemented in the quantum chemistry code GAMESS. Some applications involving water in the SR are presented.

Hydrogen bonding is one of the most important forms of intermolecular interaction. It is a critical component of biomolecular structure, molecular recognition, and protic solvent effects to name a few. Efficient computational models that describe hydrogen bonding accurately are thus essential for studies of such topics. One such model, the effective fragment potential (EFP) method, is introduced here.

Philosophy Behind The EFP Method

Initial Assumptions. The wavefunction of a chemical system of interest is divided into an “active region” (AR) and a “spectator region” (SR). The AR is the region in which chemical changes (e.g. bond breaking/making) occur. The chemistry in the AR is influenced by intermolecular hydrogen bonds to the SR. Thus, no covalent bonds

⁴Current address: WL/MLPJ Building 651, 3005 P Street, Suite 1, Wright-Patterson Air Force Base, OH 45433
⁵Current address: Department of Physiology and Biophysics, Mount Sinai Medical Center, New York, NY 10029

0097-6156/94/0569-0139$08.00/0 © 1994 American Chemical Society

connect the AR and SR. If one initially neglects the overlap of the two regions (errors introduced by this neglect will be discussed later), it is possible to relate the inter-region interaction to the properties of the isolated regions, in a general way.

Following Buckingham (1), the total Hamiltonian is defined as the sum of the AR and SR Hamiltonians plus an interaction term, $V$:

$$H' = H_{AR} + H_{SR} + V.$$  

(1)

When overlap, and hence electron exchange, is neglected one can treat the electrons as belonging to one or the other of the two regions. The wavefunction of the unperturbed system, in which neither region is perturbed by the presence of the other, can then be written as the product of the isolated AR- and SR-wavefunctions. This wavefunction is an eigenfunction of $H_{AR} + H_{SR}$ whose eigenvalue is the sum of the energies of the isolated AR and isolated SR. The energy due to the interaction, and resulting perturbation, of the two wavefunctions can be obtained through perturbation theory with $V$ as the perturbation, and is then

$$E' = E_{AR}^{(0)} + E_{SR}^{(0)} + E_{AS}^{(1)} + E_{AS}^{(2)} + \ldots$$  

(2)

In this case $V$ describes purely Coulombic interactions and classical interpretations can be given to each energy term. The first order energy corresponds to the electrostatic interactions of the static AR and SR charge distributions. The second order energy is comprised of two polarization energies (AR polarizing SR and SR polarizing AR) and a dispersion energy. The total interaction energy correct to second order is therefore

$$E' = E_{AR}^{(0)} + E_{SR}^{(0)} + E_{AS}^{(1)} + E_{AR}^{Pol} + E_{SR}^{Pol} + E_{AS}^{Disp}.$$  

(3)

At smaller inter-region distances, where electron exchange becomes important, the total unperturbed wavefunction must be antisymmetrized and is no longer an eigenfunction of $H_{AR} + H_{SR}$. Hence, it is not obvious how to relate the exchange repulsion energy, that must be added to the total energy, to the properties of the individual regions.

**Further Assumptions.** The following points are particular to the EFP implementation. (1) The wavefunction of the SR is replaced by an EFP comprised of effective potentials that simulate SR influence on the AR wavefunction. The AR wavefunction is described with standard *ab initio* MO theory. (2) The internal structure of the SR does not change, and the SR Hamiltonian (and resulting energy, $E_{SR}^{(0)}$) can thus be ignored. (3) The internal energy of the AR includes $E_{AR}^{Pol}$, since the AR wavefunction automatically responds to the presence of the EFP in the course of the energy evaluation,

$$E_{AR} = E_{AR}^{(0)} + E_{AR}^{Pol}.$$  

(4)

(4) The dispersion term ($E_{AS}^{Disp}$ in equation 3) is presumed to have negligible effects on the AR-electronic structure, based on the $R^{-6}$-distance dependence (1). When exchange repulsion becomes important (at small inter-region separations such that the charge distributions overlap), an additional term, $E_{Rep}$, must be added. The total energy of the system is then
The effective fragment potentials are added to the one-electron part of the AR hamiltonian, so the total energy in the AO basis may be rewritten as

\[ E' = E + \sum_{\mu, \nu} P_{\mu \nu} V_{\mu \nu} + V_N^{(\phi \phi)} . \]  

The second and third terms describe the interactions of the EFP with the electrons and nuclei of the AR, respectively:

\[ V_{\mu \nu} = \int d \tau \phi^*_\mu (1) \phi_\nu (1), \]  

\[ V_N^{(\phi \phi)} = \sum_A Z_A V^{(\phi \phi)} (A). \]  

The EFP can further be divided into electrostatic, polarization, and exchange repulsion contributions, cf. equation 5.

\[ V^{(\phi \phi)} (1) = \sum_k V^{\text{Electrostatic}}_k (1) + \sum_l V^{\text{Polarization}}_l (1) + \sum_m V^{\text{Repulsion}}_m (1), \]

where \( K, L, \) and \( M \) are the total number of reference points associated with the respective potentials. The first term is the molecular electrostatic potential (MEP) of the isolated SR. The second term represents the change in this MEP induced by the AR wavefunction. The third term is a repulsive potential that describes the exchange repulsion between the AR and SR. The nuclear part of the EFP consists only of the first two terms since the exchange repulsion is a purely electronic effect.

A key feature of the EFP approach is that these potentials are derived from separate \textit{ab initio} calculations. The previous discussion stated that the electrostatic and polarization terms can be rigorously derived from separate calculation of SR properties. This is not rigorously possible for the exchange repulsion term. The next section describes how each component of the potential is obtained, and how the potential is used during the derivation of the AR wavefunction.

**Constructing An Effective Fragment Potential**

**Electrostatic Interactions.** The electrostatic interaction dominates the hydrogen bond energy. Buckingham (1) has shown that this interaction potential can be related to the properties of the free molecules by expanding the molecular electrostatic potential (MEP) of one charge distribution in a multipolar expansion about an expansion point, \( k \). Thus the electrostatic interaction potential of the AR and SR can be expressed as

\[ V^{\text{Electrostatic}}_k (1) = \frac{q_k}{r_{1k}} - \sum_a \mu_a^k F_a^k (r_{1k}) - \frac{x_{12k}^{x_{12k}}}{} \sum_a \sum_b \Theta_{ab}^k U_{ab}^k (r_{1k}) - \frac{x_{12k}^{x_{12k}}}{} \sum_a \sum_b \sum_c \Omega_{abc}^k U_{abc}^k (r_{1k}) + \ldots \]
Here, $q_k$ is the net charge of the SR charge distribution, $\mu$, $\Theta$, and $\Omega$ are the dipole, quadrupole, and octupole, respectively, of the SR, and $F$, $F'$, and $F''$ are the electric field, field gradient, and field second derivative operators, due to the AR, at point $k$. As with the perturbative analysis described above, this expansion is only rigorous if the molecules have non-overlapping charge distributions.

In general, an infinite number of terms is required to get an exact expansion of the MEP. However, by choosing several expansion points ($K$ in equation 8) for a given molecule, the expansion's convergence can be greatly accelerated. Numerous schemes (2-4) have been developed to efficiently describe the MEP. The efficiency is usually determined by comparing the accuracy of the fitted MEP, relative to the quantum mechanical MEP, to the number of terms in the expansion. The electrostatic part of the EFP can be any expansion, but a compact expansion obviously reduces computational expense.

The distributed multipolar analysis (DMA) of Stone (5-6) has been found to give well-converged multipolar expansions for several small test molecules (7). This permits (but does not require) truncation at the quadrupole term at expansion points at the atom centers and bond midpoints, the expansion centers recommended by Stone et al. (6). Multipolar expansions of each gaussian product density element are evaluated at the expansion centers closest to the density element. Thus, the best expansion points coincide with large concentrations of gaussian product centers, e.g. atoms and bond mid-points. This is an approximation to the method of Rabinowitz, et al. (8) in which each of the $N(N+1)/2$ gaussian product centers in the basis set is used as an expansion point. While this yields finite expansions at each point, it results in an unwieldy number of points that are basis set dependent.

**Charge Penetration.** Typical hydrogen bonded distances between two atoms are generally shorter than the sum of their van der Waals radii, indicating that the atomic charge distributions are overlapping to a non-negligible extent. As mentioned previously, the form of the interaction potential in equation 9 is rigorous only for non-overlapping charge distributions. The multipolar expansion is not an accurate representation of the exact quantum mechanical MEP inside the region of significant charge density [one definition of this region is the 0.001 au charge density envelope (9)]. As the charge distributions interpenetrate, the MEP seen by one molecule due to the charge density on another molecule is significantly altered, due to the overlap of the two charge densities. Since nuclei generally are outside the overlap region they are effectively deshielded, leading to an effective increase in nuclear charge and thus an effective increase in electron-nuclear attraction. Charge penetration effects are therefore always attractive. Neglecting this charge penetration effect can result in serious errors.

The penetration effects are included in the EFP model by fitting the multipolar expansion of the MEP to the exact quantum mechanical MEP of the isolated spectator molecule. This is done by adding a penetration potential to each multipolar expansion, and optimizing penetration parameters to obtain the best fit to the accurate quantum mechanical MEP of the isolated SR. Preliminary test calculations on neutral atoms (7) indicate that the penetration effects decay rapidly with distance, and can be modeled with a single gaussian. The gaussian form facilitates easy implementation in integral evaluation and derivative schemes. Thus, by introducing adjustable parameters $\alpha_k$ and $\beta_k$ and making the substitution

$$V_k^{\text{Elec}}(1) \rightarrow (1 - \beta_k e^{-\alpha_k r_{12}^2}) V_k^{\text{Elec}}(1),$$

(10)
in the electrostatic part of the EFP, intermolecular electrostatic interactions were consistently reproduced to within 5% or less of \textit{ab initio} values at van der Waals distances (7).

**Polarization.** As indicated by equation 3, a part of the intermolecular interaction energy arises from the change in electronic structure in one molecule due to the presence of another, i.e. polarization. This interaction can be expressed in terms of properties of the isolated molecules, i.e. molecular multipolar polarizabilities, \(I\) in an expression similar to that for the electrostatic interaction,

\[
E_{\text{Pol}}^{\text{Pol}} = -\frac{1}{2} \sum_{\alpha \beta} \sum_{\gamma} \alpha_{\alpha \beta \gamma} F_{\alpha} F_{\beta} - \sum_{\alpha \beta} \sum_{\gamma} \Delta \mu_{\alpha \beta} \Delta \mu_{\beta} - \frac{1}{2} \sum_{\alpha \beta} \sum_{\gamma} \sum_{\delta} \Delta \mu_{\alpha \beta} \Delta \mu_{\beta \delta} - \cdots
\]

Here \(\alpha\) is the dipole polarizability tensor, and \(A\) and \(C\) are dipole-quadrupole and quadrupole-quadrupole polarizability tensors, respectively. The field and field gradient terms (\(F\) and \(F'\)) are similar to those in equation 9. In the EFP methodology, several expansion points (\(L\) in equation 8) are used to describe the polarizability of the SR. This leads to accelerated convergence and allows the individual expansions to be truncated after the first term while still maintaining some description of the higher order effects. Thus, the polarization term in the EFP is given by

\[
V_{l}^{\text{Pol}} = -\sum_{\alpha \beta} \sum_{\gamma} \Delta \mu_{\alpha \beta} \omega_{\alpha \beta \gamma} F_{\alpha}(\mathbf{r}_{l}) F_{\beta}(\mathbf{r}_{l})
\]

where \(\Delta \mu\) is referred to as the induced dipole moment at point \(l\). Since \(\langle F_{\beta}(\mathbf{r}_{l}) \rangle\), the expectation value of the field due to the AR at point \(l\), depends on the final wavefunction, the polarization term is non-linear with respect to the wavefunction. This is solved by obtaining an initial guess at the induced dipole, e.g. calculated by using the current electric field, and iterating to self-consistency. Since the distributed polarizabilities within an EFP are derived from fully-coupled SCF calculations (see below), the induced dipoles within an EFP molecule are not required to interact directly. This is an approximation if they arise from a non-uniform field.

A few methods have been developed to obtain distributed polarizabilities (10-13). The method most compatible with the EFP methodology is the one due to Garmer and Stevens (13) in which the molecular polarizability is decomposed in terms of localized molecular orbital (LMO) contributions. Each LMO polarizability is given by the (numerical) first derivative of the LMO dipole \(\mu_{l}'\) with respect to a uniform field:

\[
\alpha_{\alpha \beta}^{l} = \lim_{F_{\gamma} \rightarrow 0} \frac{\mu_{l}'(F_{\gamma}) - \mu_{l}'(0)}{F_{\gamma}}
\]

Thus, \(L\) in equation 8 is the number of LMOs in the SR, and each point, \(l\), is at the position of the LMO centroid of charge. Summing \(\alpha_{xy}^{l}\) over all LMOs gives the \(xy\) component of the total molecular polarizability. It is important to note that the
molecular polarizability tensor is symmetric, i.e. the sum of, say, the xy- and the yx- components of the LMO-α’s are equal, but that this is not necessarily true for each individual LMO-α.

Several tests have been performed (7) to compare the distributed polarizability model to the standard molecular polarizability and to Hartree-Fock SCF results. It is found that the distributed model generally reproduces the SCF results better than the single polarizable point model. The average error in energy for the former is around 10-20%, thus the majority of the polarization energy in equation 11 can be modeled through the first term.

**Exchange Repulsion.** Exchange repulsion can dominate the intermolecular interaction energy at distances where the charge density of two molecules overlap significantly. It arises in part from the fact that charge density in the overlap region is depleted, leading to a decrease in electron-nuclear attraction and thus a net repulsive interaction (14). As pointed out previously, no rigorous way to describe this repulsion in terms of properties of the isolated molecules exists. Thus, approximate methods for implementing this effect in EFP calculations must be tested on molecular systems for which this exchange repulsion energy has been calculated explicitly. In the EFP method reported here, the exchange repulsion energy representation is chosen to be as simple as possible, requiring only one-electron integrals and depending only on the density of the AR. The general approach used here is to calculate the exchange repulsion energy for a variety of intermolecular geometric arrangements for a complex (A...B) of interest, and fit the resulting energy surface to some functional form. The exchange repulsion energy, EXO, is calculated by using the energy decomposition scheme of Morokuma and Kitaura (14-15). Alternatively, it can be redefined as the difference between the total ab initio energy and the electrostatic plus polarization energy, which implicitly includes any charge-transfer and dispersion effects.

Repulsive effective potentials (REPs) are chosen here as a way to implement the exchange repulsion interaction in the EFP methodology. Here, a REP consists of several \( M \) in equation 8) linear combinations of gaussians,

\[
V_{\text{rep}}^m(l) = \sum_j \beta_{m,j} r_m^l e^{-\alpha_{m,j} r_m^l}.
\]  

where the coefficients \( \beta \) and \( \alpha \) have been optimized to reproduce the exchange repulsion energy surface (ERES) of a pair of molecules (A and B), for a given \( M \), \( J \), and \( n \). This fitting of the ERES is accomplished by minimizing the following error function,

\[
\Delta = \sum_p \left( \frac{\langle \Psi_A | \sum_m V_{\text{rep}}^m | \Psi_A \rangle_p - EXO_p}{EXO_p^2} \right)^2.
\]  

Here, \( P \) is the number of points on the ERES, corresponding to various orientations of A and B; \( EXO_p \) is the exchange repulsion energy at point \( p \), and \( \langle \Psi_A | \sum_m V_{\text{rep}}^m | \Psi_A \rangle_p \) is the exchange repulsion energy due to the REP of molecule B interacting with the unperturbed wavefunction of A at point \( p \). Molecule B is then the molecule to be
replaced by an EFP, and molecule A is the molecule in the AR to be described by quantum mechanics. It will be shown below that the repulsive part of the EFP can be used in calculations where the AR is a molecule other than A, and still give reasonable results. This transferability is not necessarily a given, so it has been established through testing. Alternatively, new repulsive potentials must be obtained for each molecular species used in EFP calculations.

Locating Stationary Points

While the internal structure of the of the EFP is fixed, its position relative to the AR is not. So, to obtain the optimum interaction energy, the overall structure must be optimized. This is achieved through a standard Newton-Raphson procedure. Since the internal EFP-geometry is fixed, each EFP adds six degrees of freedom to the system. The six degrees of freedom chosen are the three Cartesian components of the overall translation of the EFP relative to an arbitrary origin and the three Cartesian components of the rotation vector around the EFP center of mass (COM). The corresponding energy derivatives, depicted schematically in Figure 1, are the Cartesian components of the net force and total torque around the COM.

![Figure 1. Schematic representation of the transformation of the Cartesian gradient components on a fragment to internal coordinate components defined relative to the center of mass (COM) of the EFP.](image)

These new derivatives are obtained through the following transformation.

\[
F^A_x = \sum_a \left( \frac{\partial E'}{\partial x_a} \right), \quad (16a)
\]

\[
F^A_{\theta_x} = \sum_a \left( (R_a - R_A) \times F^A \right)_{x_a} + \sum_a \sigma^{(e)} \cdot (x_a). \quad (16b)
\]

Here, \( F^A_x \) and \( F^A_{\theta_x} \) are the \( x \) components of the total translational force and torque due to all terms on fragment A, respectively and \( [v]_x \) is the \( x \) component of vector \( v \). Similar equations apply for the \( x \) and \( y \) components. The last term in equation 16b describes the torque induced on the fragment multipoles by the electric field of the \( ab \) \( initio \) system. The detailed expressions for the energy derivatives and torques are given elsewhere (17).
Applications

In the applications of the EFP method described in this section the SR region is taken to be a water molecule. The EFP used for the water molecule is described by five multipolar expansion points (corrected for charge penetration), four polarizable points, and three exchange repulsion points. All terms in the EFP are evaluated at the RHF/CEP-31G* (18) level of theory (d orbital exponent=0.85) using a fixed water geometry with bond lengths of 0.957Å and a bond angle of 104.52°. The multipole expansion points are located at the three nuclei and at the two bond midpoints. Multipoles through octupoles are included in the evaluation of the energy and its derivatives. Multipoles have been evaluated from ab initio calculations on the water molecule by the method described by Stone (5-6). Four effective polarizability points are used, located at the centroids of the four valence localized molecular orbitals. These polarizabilities are obtained from ab initio calculations carried out under the influence of an electric field. Three exchange repulsion points are included in the effective fragment potential, one located at each nucleus. The contribution to the interaction potential from each of these points is given by a sum of two spherical gaussians. The potential was fit to the ERES of the water dimer. All calculations were performed with a local version of the GAMESS (19) program.

Water Dimer. The water dimer system is chosen as an initial test case for the EFP method. The effect of replacing either the hydrogen bond donor or acceptor water molecule with an EFP-water is compared to all-ab initio calculations. Properties of interest include the dimer structure, interaction energy, and vibrational frequencies, evaluated at the RHF/CEP-31G* level of theory. In addition, the effect of polarization functions on the interaction energy is studied.

Table I lists the optimized structure of the water dimer. The most important geometric parameters in the effective fragment calculations are the internal coordinates of the ab initio molecule. These values are underlined in Table I. The values marked with an asterisk are fixed in the effective fragment method. Table I indicates that the internal structure of the ab initio water molecule is predicted quite accurately by the effective fragment method. In both effective fragment calculations, the bond lengths in the solute molecule agree with those from the full ab initio calculation to within 0.001 Å, and the bond angle agrees to better than 0.1°. As for the relative positions of the two molecules, the hydrogen bond length is off by at most 1% (0.022Å for EFP=acceptor). The orientational angle θ is off by as much as 11° when the EFP acts as the acceptor. This also represents the largest difference in structure between the two EFP calculations.

Table II gives the interaction energies for the two water molecules at the equilibrium structures obtained from each of the three types of calculations with three basis sets. In the calculations with the CEP-31G* basis set the interaction energies predicted by the H-donor EFP calculation and by the H-acceptor EFP calculation are less than the 5.0 kcal/mol predicted in the all-ab initio calculation by 0.2 kcal/mol and 0.3 kcal/mol, respectively. This is virtually unchanged when the basis set quality is increased by adding p polarization functions (p orbital exponent=1.0) on the hydrogens in the ab initio water molecule. Larger discrepancies, 1.2 and 2.6 kcal/mol, arise when the oxygen polarization functions are removed. The source of these discrepancies is the fact that the EFP models a CEP-31G* water molecule and the EF calculations therefore resemble calculations with one CEP-31G water and one CEP-31G* water. Such all-ab initio calculations result in optimized interaction energies of 6.9 kcal/mol and 5.6 kcal/mol for d functions on only the donor or acceptor water, respectively. These more sophisticated full ab initio calculations are in better agreement with the EFP calculations.

Table III gives the harmonic vibrational frequencies and vibrational zero-point energy (ZPE) changes obtained from hessian calculations on the dimer geometries in
Table I. RHF/CEP-31G(d) Optimized Geometries. The * Marks Frozen EFP Coordinates, While The Underlined Numbers Refer To The Internal Water Structure

<table>
<thead>
<tr>
<th></th>
<th>All-ab initio</th>
<th>EFP=Donor</th>
<th>EFP=Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(O1-H2)</td>
<td>0.952</td>
<td>0.951</td>
<td>0.957*</td>
</tr>
<tr>
<td>a(H2-O1-H3)</td>
<td>106.4</td>
<td>106.4</td>
<td>104.5*</td>
</tr>
<tr>
<td>R</td>
<td>2.04</td>
<td>2.05</td>
<td>2.07</td>
</tr>
<tr>
<td>θ</td>
<td>41</td>
<td>40</td>
<td>52</td>
</tr>
<tr>
<td>r(O4-H5)</td>
<td>0.955</td>
<td>0.957*</td>
<td>0.955</td>
</tr>
<tr>
<td>r(O4-H6)</td>
<td>0.950</td>
<td>0.957*</td>
<td>0.950</td>
</tr>
<tr>
<td>a(H5-O4-H6)</td>
<td>105.9</td>
<td>104.5*</td>
<td>105.8</td>
</tr>
</tbody>
</table>

Table II. Interaction Energies For The Water Dimer In Kcal/mol

<table>
<thead>
<tr>
<th>Basis Set</th>
<th>All-ab initio</th>
<th>EFP=Donor</th>
<th>EFP=Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEP-31G</td>
<td>7.6</td>
<td>6.4</td>
<td>5.0</td>
</tr>
<tr>
<td>CEP-31G(d)</td>
<td>5.0</td>
<td>4.8</td>
<td>4.7</td>
</tr>
<tr>
<td>CEP-31G(d,p)</td>
<td>4.9</td>
<td>4.7</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Table III. Harmonic Frequencies (cm⁻¹) Of The RHF/CEP-31G(d) Water Dimer

<table>
<thead>
<tr>
<th>Frequency</th>
<th>All-ab initio</th>
<th>EFP=Donor</th>
<th>EFP=Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A''</td>
<td>138</td>
<td>134</td>
<td>130</td>
</tr>
<tr>
<td>2. A'</td>
<td>145</td>
<td>163</td>
<td>144</td>
</tr>
<tr>
<td>3. A''</td>
<td>151</td>
<td>187</td>
<td>204</td>
</tr>
<tr>
<td>4. A'</td>
<td>173</td>
<td>233</td>
<td>210</td>
</tr>
<tr>
<td>5. A'</td>
<td>342</td>
<td>346</td>
<td>414</td>
</tr>
<tr>
<td>6. A''</td>
<td>605</td>
<td>559</td>
<td>642</td>
</tr>
<tr>
<td>7. A'</td>
<td>1808</td>
<td>1809</td>
<td>ff</td>
</tr>
<tr>
<td>8. A'</td>
<td>1831</td>
<td>ff</td>
<td>1836</td>
</tr>
<tr>
<td>9. A'</td>
<td>4074</td>
<td>ff</td>
<td>4093</td>
</tr>
<tr>
<td>10. A'</td>
<td>4113</td>
<td>4113</td>
<td>ff</td>
</tr>
<tr>
<td>11. A'</td>
<td>4214</td>
<td>ff</td>
<td>4221</td>
</tr>
<tr>
<td>12. A''</td>
<td>4234</td>
<td>4236</td>
<td>ff</td>
</tr>
</tbody>
</table>

ΔZPE (kcal/mol) 2.2 2.3 2.5
Table I. Only numerical hessians are available in effective fragment calculations. To ensure accuracy, the maximum component of the gradient of each geometry was reduced to less than $10^{-6}$ Hartree/Bohr, and the symmetrical displacements around the minimum were reduced to 0.001 Bohr. For the water dimer this generally leads to frequencies that are within 4.1% of analytical results. The harmonic analysis in the effective fragment calculations shows an overestimation of the frequencies associated with the internal coordinates of the \textit{ab initio} molecule by 0.06%, 0.0%, and 0.05% when the H-donor is replaced with a fragment and by 0.3%, 0.5%, and 0.2% when the H-acceptor is replaced with a fragment. For the frequencies associated with the relative motion of the two waters, the H-donor effective fragment calculation agrees quite well with the full \textit{ab initio} calculation, except for frequencies #3 and #4, for which the fragment results (187 and 233 cm$^{-1}$) are 19% and 26% greater than the \textit{ab initio} frequencies. In the H-acceptor EF calculation, frequencies #3-6 are up to 26% (for #3) higher than in the all-\textit{ab initio} calculation. However, these deviations translate to only minor ($\geq$0.3 kcal/mol) errors in the zero point energy (ZPE) correction to the interaction energy.

Water-Formamide. In order to evaluate the more general usefulness of the effective fragment potential for the water molecule, we need to study its interaction with AR molecules other than water. The interaction between the formamide molecule and the water molecule is of interest in biochemistry because formamide is the simplest prototype for a peptide linkage. In an \textit{ab initio} study by Jasien and Stevens (20) four stationary points were found on the RHF/DZP (21) formamide-water potential energy surface, within the constraint of C$_{s}$ symmetry. We have carried out \textit{ab initio} geometry optimizations in C$_{i}$ symmetry on this system with the water molecule replaced by an effective fragment. In addition we have done full \textit{ab initio} geometry optimizations in C$_{1}$ symmetry. Both lead to three C$_{1}$ minima similar to three of the C$_{s}$ structures. The fourth structure was a C$_{s}$ transition state at the all-\textit{ab initio} level of theory.

Figure 2 depicts the three C$_{1}$ minima, labeled I-III, that were located by both all-\textit{ab initio} and EF optimizations. Selected structural parameters are listed for the full \textit{ab initio} (bold) and EF calculations. In the effective fragment calculations on the three minima, the length of the hydrogen bonds between the two molecules is longer than predicted by the \textit{ab initio} calculations by just 0.04 (structure III; 2%) to 0.17 (structure I; 8%) Å. The orientational angles obtained in the effective fragment calculations do not agree exactly with the \textit{ab initio} calculations either, but are qualitatively correct.

The geometric parameters of greatest interest, the internal coordinates of the formamide molecule, are compared in Table IV. The first column in Table IV lists the internal coordinates of a lone formamide molecule in its equilibrium configuration, as obtained in an \textit{ab initio} calculation with the DZP basis, and the other six columns list the change in these coordinates caused by the presence of a water molecule. For each of the three minimum energy configurations, results are listed both from full \textit{ab initio} calculations and from effective fragment calculations. While the changes in these coordinates are small, the effective fragment method consistently predicts nearly the same perturbation in these internal coordinates as was obtained in the full \textit{ab initio} calculations.

Table V lists the interaction energies for structures I-III. For the three minima on the potential energy surface, the interaction energies obtained in the effective fragment calculations differ from those obtained from the full \textit{ab initio} calculations by 1.6 (19%), 0.9 (15%), and 0.5 (9%) kcal/mol. Although the effective fragment
Figure 2. RHF/DZP optimized structures of three formamide-water complexes, with and without the water replaced by an EFP. The bold numbers are structural parameters for all-\textit{ab initio} calculation, and may be compared to the numbers obtained in the EFP calculations. Bond lengths are given in ångstroms and bond angles in degrees.

Table IV. RHF/DZP Internal Coordinates Of The Isolated Formamide Molecule, And The Change In These Coordinates Caused By An \textit{Ab Initio} Or EFP Water In Structures I-III. Bond Lengths In Angstroms And Angles In Degrees

<table>
<thead>
<tr>
<th>Formamide</th>
<th>Structure I \textit{ab initio}</th>
<th>Structure I EFP</th>
<th>Structure II \textit{ab initio}</th>
<th>Structure II EFP</th>
<th>Structure III \textit{ab initio}</th>
<th>Structure III EFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(C-N)</td>
<td>1.353</td>
<td>-0.009</td>
<td>-0.007</td>
<td>-0.006</td>
<td>-0.005</td>
<td>-0.004</td>
</tr>
<tr>
<td>r(C-O)</td>
<td>1.196</td>
<td>0.009</td>
<td>0.007</td>
<td>0.007</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>r(C-H)</td>
<td>1.092</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.003</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>r(N-H)</td>
<td>0.995</td>
<td>0.005</td>
<td>0.005</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>r(N-H')</td>
<td>0.992</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.005</td>
<td>0.004</td>
</tr>
<tr>
<td>a(OCN)</td>
<td>124.9</td>
<td>0.2</td>
<td>0.2</td>
<td>-0.5</td>
<td>-0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>a(HCN)</td>
<td>122.3</td>
<td>-1.0</td>
<td>-0.8</td>
<td>-0.4</td>
<td>-0.2</td>
<td>-0.3</td>
</tr>
<tr>
<td>a(HNC)</td>
<td>121.3</td>
<td>-0.7</td>
<td>-0.7</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>a(H'NC)</td>
<td>119.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>-0.5</td>
</tr>
<tr>
<td>d(OCNH)</td>
<td>0.0</td>
<td>1.6</td>
<td>2.6</td>
<td>0.4</td>
<td>-0.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table V. RHF/DZP Interaction Energies For Three Water-Formamide Complexes In Kcal/mol

<table>
<thead>
<tr>
<th>Structure</th>
<th>All-\textit{ab initio}</th>
<th>Water-EFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-8.3</td>
<td>-6.7</td>
</tr>
<tr>
<td>II</td>
<td>-6.0</td>
<td>-5.1</td>
</tr>
<tr>
<td>III</td>
<td>-5.3</td>
<td>-4.8</td>
</tr>
</tbody>
</table>
method underestimates the interaction energies, it does correctly predict that structure I is considerably more stable (by 1.6 kcal/mol, compared to 2.3 kcal/mol in the \textit{ab initio} case) than structures II or III, and that structures II and III are comparatively close in energy. Clearly, one arrives at the same qualitative picture of the water-formamide interaction based on both methods.

The fact that the EF method does not do quite as well at predicting the formamide-water interaction energy as it did at predicting the water dimer interaction energy is probably due to the exchange repulsion part of the effective fragment potential, which was fit to the water dimer interaction. More sophisticated exchange-repulsion potentials which explicitly take into account the overlap between the AR and the SR may provide improved transferability. Considering the simple form of the potential used here, the effective fragment method does quite well. Since the formamide-water system is small enough to be treated in a full \textit{ab initio} calculation, this system could in principle be used to construct an exchange repulsion potential that might be more accurate in modeling the hydration of larger peptide systems.

Conclusion and Future Directions

The effective fragment potential (EFP) method is introduced as a way to model intermolecular hydrogen bonds and their effect on quantum mechanical wavefunctions. It is shown that the effect of a water molecule on the \textit{ab initio} wavefunctions of water and formamide can be modeled relatively accurately by introducing an EFP in the \textit{ab initio} one-electron Hamiltonian. The potentials are obtained from other \textit{ab initio} calculations on the isolated water molecule and water dimer. Thus for the formamide-water complexes no parameters in the EFP model have been adjusted to reproduce the interaction. It is therefore encouraging to find only relatively modest deviations from calculations in which both the formamide and the water are treated quantum mechanically. The method is in principle extendible to model any intermolecular hydrogen bond.

Current research is focused on a parallelizing the EFP code in GAMESS, as well as including energy and gradient terms that describe EFP-EFP interactions. The latter would allow, for example, to surround an \textit{ab initio} wavefunction with more than one EFP water to approximately model the first solvation shell in aqueous solvation.

Acknowledgments

This research was supported, in part, by grant No. 88-00406 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. Two of the authors (MSG, JHJ) acknowledge the support of the National Science Foundation (CHE89-11911).

Literature Cited

9. JENSEN ET AL.  
Modeling Intermolecular Hydrogen-Bonding Effects  


RECEIVED June 2, 1994