

# Fully Integrated Effective Fragment Molecular Orbital Method

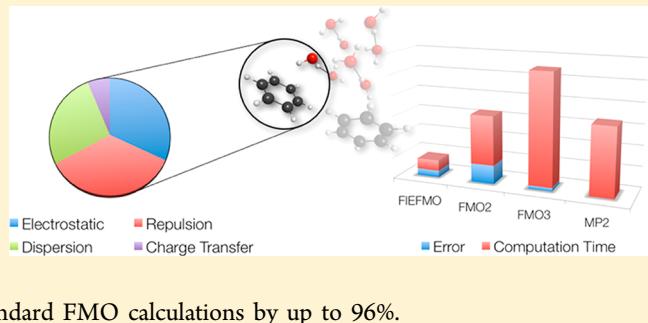
Spencer R. Pruitt,<sup>†</sup> Casper Steinmann,<sup>‡</sup> Jan H. Jensen,<sup>‡</sup> and Mark S. Gordon\*,<sup>†</sup>

<sup>†</sup>Department of Chemistry, Iowa State University, Ames, Iowa 50011, United States

<sup>‡</sup>Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

## Supporting Information

**ABSTRACT:** In this work, the effective fragment potential (EFP) method is fully integrated (FI) into the fragment molecular orbital (FMO) method to produce an effective fragment molecular orbital (EFMO) method that is able to account for all of the fundamental types of both bonded and intermolecular interactions, including many-body effects, in an accurate and efficient manner. The accuracy of the method is tested and compared to both the standard FMO method as well as to fully ab initio methods. It is shown that the FIEFMO method provides significant reductions in error while at the same time reducing the computational cost associated with standard FMO calculations by up to 96%.



## 1. INTRODUCTION

Modern computational chemistry methods strive to accurately model chemical systems using efficient computational algorithms. Unfortunately, it is difficult to reconcile both of these goals, since most methods that are widely viewed as the most accurate<sup>1</sup> also require the most computational effort. A very effective compromise is the application of fragmentation approaches to these computationally intensive methods. Many such fragmentation methods have been introduced in recent years,<sup>2–8</sup> with several showing the ability to accurately model large molecular systems. Methods such as the systematic molecular fragmentation (SMF) method,<sup>9–11</sup> molecular fractionation with conjugate caps (MFCC),<sup>12</sup> the molecular tailoring approach (MTA),<sup>13</sup> and the explicit polarization potential (X-Pol)<sup>14,15</sup> have all exhibited success in describing different chemical systems.

One such method, the fragment molecular orbital (FMO) method,<sup>16</sup> has been extensively developed<sup>17</sup> since the original implementation by Kitaura et al. Based upon a many-body expansion of the energy, the FMO method takes the effects of the entire system into account during each step of a given calculation through the use of an electrostatic potential (ESP). The FMO method, as well as other fragmentation methods,<sup>18</sup> also benefit from the relative ease with which calculations can be parallelized on modern computer architectures. This inherent parallelizability aids in lowering the computational cost of the most accurate ab initio methods.

While not a fragmentation method in the same vein as the FMO method, the effective fragment potential (EFP) method<sup>19–21</sup> was originally developed to accurately introduce solvent effects into chemical processes without the use of any fitted parameters. The importance of modeling chemistry in solution is apparent in many applications,<sup>22–32</sup> making the EFP method an attractive solution to the study of solvent effects and intermolecular interactions. Recently, the generalized EFP

method<sup>33,34</sup> (sometimes called EFP2) has been developed as an ab initio based method for capturing all intermolecular interactions including, but not limited to, solvent effects. As implemented in the General Atomic and Molecular Electronic Structure System (GAMESS) program package,<sup>35,36</sup> the EFP2 method requires a preliminary calculation to generate the potential for each fragment. The generated potentials are then incorporated into a production calculation on the chemical system of interest. Since the EFP method is essentially a classical model, it can only provide intermolecular interactions. Bond breaking/bond making processes must be described by quantum mechanics. EFP fragments have internally frozen geometries.

In an effort to combine the accuracy of the FMO method with the accuracy and speed of the EFP method for intermolecular interactions, Steinmann et al. recently developed a combined method called the effective FMO (EFMO) method.<sup>37,38</sup> The original EFMO method uses the fragmentation scheme from the FMO method and treats the separated fragment interactions using the Coulomb interaction of the EFP method. Additionally, the ESP used during standard FMO calculations is replaced with the many-body polarization of the EFP method. Other intermolecular interactions such as exchange repulsion, charge transfer, and dispersion that are present in the EFP method were not included in the EFMO method. The EFMO method also incorporates the intramolecular energy of each fragment, and it removes the restriction of frozen internal EFP geometries.

Although the original EFMO method has many benefits, the inclusion of only the EFP polarization and Coulomb interactions essentially limits the ability of the method to capture the dispersion energy to only interactions that are

Received: February 12, 2013

Published: March 27, 2013



directly calculated with quantum mechanics; that is, the nearest neighbor fragment-fragment interactions. The present work describes a fully integrated EFP/FMO method, called FIEFMO. The EFMO method described here includes all relevant intermolecular interactions (Coulomb, polarization, exchange repulsion, charge transfer, and dispersion). By including all fundamental types of intermolecular interactions, this FIEFMO method can also reduce the total number of explicit quantum mechanics (QM) calculations required, providing a significant reduction in computational cost. An additional benefit of this extension is the detailed analysis of the individual contributions to the total interaction energy, or energy decomposition analysis<sup>39,40</sup> (EDA), between two fragments provided by the EFP method. A dimer EDA can provide insights into the most important intermolecular interactions in a given chemical system.

The first section of this paper introduces the theoretical background of the FMO, EFP, and EFMO methods. Next, the FIEFMO method is tested on pure water clusters and mixtures of water and methanol molecules of varying size, comparing the energies to those from the FMO method and to fully ab initio energies. A dimer EDA using the FIEFMO method is also performed on a cluster of eight water molecules and two benzene molecules. Timings are then presented that compare the FIEFMO, EFMO, FMO2, FMO3, and fully ab initio methods, followed by conclusions and future directions.

## 2. METHODOLOGY

**2.1. The Fragment Molecular Orbital Method.** The FMO method is based on a many-body expansion of the total energy:

$$\begin{aligned} E = & \sum_I E_I + \sum_{I>J} (E_{IJ} - E_I - E_J) + \\ & \sum \{(E_{IJK} - E_I - E_J - E_K) - (E_{IJ} - E_I - E_J) \\ & - (E_{IK} - E_I - E_K) - (E_{JK} - E_J - E_K)\} \dots \end{aligned} \quad (1)$$

The expansion in eq 1 is usually truncated at three-body interactions. Each energy in eq 1 is calculated using quantum mechanical (QM) methods. Equation 1 includes energies for individual fragments  $I$  (monomers), fragment pairs  $IJ$  (dimers), and fragment triples  $IJK$  (trimers). The first two terms in eq 1 omit explicit trimer calculations and comprise the FMO2 method. The FMO3<sup>41</sup> method includes the third (trimer) term. The FMO2 and FMO3 energy expressions can be written as:

$$E^{\text{FMO2}} = \sum_I E_I + \sum_{I>J} (E_{IJ} - E_I - E_J) \quad (2)$$

$$\begin{aligned} E^{\text{FMO3}} = & E^{\text{FMO2}} + \sum_{I>J>K} \{(E_{IJK} - E_I - E_J - E_K) \\ & - (E_{IJ} - E_I - E_J) - (E_{IK} - E_I - E_K) \\ & - (E_{JK} - E_J - E_K)\} \end{aligned} \quad (3)$$

Each energy in eqs 2 and 3 can be calculated in one of two ways: as isolated  $n$ -mers with energy  $E_x^0$ , or in the presence of the rest of the fragments with energy  $E_x$  ( $x = I, J, K, IJ, IK, JK, IJK, \dots$ ). The standard FMO method calculates the energies  $E_x$  by incorporating an ESP derived from the densities of all other fragments.<sup>16</sup> The ESP has the form

$$V_{\mu\nu}^x = \sum_{K \neq x} (\mu_{\mu\nu}^K + \nu_{\mu\nu}^K) \quad (4)$$

$$\mu_{\mu\nu}^K = \sum_{A \in K} \langle \mu | (-Z_A / |\mathbf{r} - \mathbf{r}_A|) | \nu \rangle \quad (5)$$

$$\nu_{\mu\nu}^K = \sum_{\lambda\sigma \in K} D_{\lambda\sigma}^K (\mu\nu | \lambda\sigma) \quad (6)$$

where  $\mu_{\mu\nu}^K$  represents the nuclear attraction contribution to the energy and  $\nu_{\mu\nu}^K$  represents the two-electron contribution. Both of these terms are expressed in terms of one- (eq 5) and two- (eq 6) electron integrals over AOs  $\mu$  and  $\nu$  and are calculated for each of the surrounding monomers  $K$  with electron density  $D^K$ .

While the formalism up to this point provides a means to obtain accurate energies, efficient algorithms must be used to take advantage of modern computer hardware. The FMO implementation in GAMESS makes use of the generalized distributed data interface<sup>42</sup> (GDDI) to accomplish this goal. One performs each  $n$ -mer calculation on a separate computer node (coarse grained parallel); the availability of multiple cores within each node facilitates the fine-grained parallel calculation for each fragment. The two-level parallelism described here allows the FMO method to take advantage of massively parallel computers.<sup>43</sup>

Despite the computational savings that can be achieved through the use of the GDDI, the rapid increase in the number of dimer and trimer calculations with system size for large molecular systems still requires an increasingly significant computational effort. The number of  $n$ -mer calculations performed in an FMO calculation can be partially reduced by employing approximations for the interaction energies of  $n$ -mers that are farther apart than a unitless, predefined cutoff value  $R_{\text{cut}}$ . The unitless distance between two fragments,  $R_{IJ}$ , is given by

$$R_{IJ} = \min_{i \in I, j \in J} \left\{ \frac{|\vec{r}_i - \vec{r}_j|}{r_i^{\text{vdw}} + r_j^{\text{vdw}}} \right\} \quad (7)$$

where  $R_{IJ}$  is the relative minimum interatomic distance between fragment  $I$  and fragment  $J$  based on the van der Waals radii  $r^{\text{vdw}}$  of the atoms  $i$  and  $j$ . Through a reformulation of eq 2,  $E^{\text{FMO2}}$  can be expressed as

$$E^{\text{FMO2}} = \sum_I E_I + \sum_{I>J \leq R_{\text{cut}}} \Delta E_{IJ} + \sum_{I>J > R_{\text{cut}}} \Delta E_{IJ}^{\text{sep}} \quad (8)$$

where

$$\Delta E_{IJ} = (E_{IJ} - E_I - E_J) \quad (9)$$

In eq 8  $\Delta E_{IJ}^{\text{sep}}$  refers to the “separated dimers”; that is, those pairs of fragments that are separated by a distance greater than  $R_{\text{cut}}$ . There is a corresponding expression<sup>41</sup> to eq 8 for  $E^{\text{FMO3}}$ . The total number of fully QM dimers (term 2) in eq 8 can now be varied based on the user-defined value of  $R_{\text{cut}}$  (FMO2 default value = 2.0). The remaining (separated) dimers (last term in eq 8) are now approximated by

$$\begin{aligned} \Delta E_{IJ}^{\text{sep}} \cong & E_I + E_J + \text{Tr}(D^I u^{1,I(J)}) + \text{Tr}(D^J u^{1,J(I)}) \\ & + \sum_{\mu\nu \in I} \sum_{\lambda\sigma \in J} D_{\mu\nu} D_{\lambda\sigma} (\mu\nu | \lambda\sigma) + \Delta E_{IJ}^{\text{NR}} \end{aligned} \quad (10)$$

where  $u^{I,I(J)}$  and  $u^{J,I(I)}$  are the one-electron Coulomb potentials of the force exerted by fragment  $I$  ( $J$ ) on fragment  $J$  ( $I$ ). The electron–electron interaction (fifth term in eq 10) and the nuclear repulsion energy,  $\Delta E_{IJ}^{\text{NR}}$ , are also included. In contrast to the fully QM dimers (second term in eq 8), the standard SCF procedure is not employed during separated dimer calculations, resulting in significant time savings compared to QM dimer calculations. A recent development by Choi and Fedorov<sup>44</sup> reduces the computational cost of separated dimers even further by using eq 10 for dimers that are separated within a specified “intermediate” range and calculating the interactions for “far” separated dimers using a multipole expansion. The values used for the intermediate and far separated ranges may be defined by the user.

Two user-defined values that are analogous to  $R_{\text{cut}}$  also exist for the exclusion of electron correlation during MP2 or CC calculations ( $R_{\text{CORS}}_D$ ), as well as an array of four values ( $R_{\text{ITRIM}}$ ) used to completely neglect separated trimer interactions during FMO3 calculations. A third user-defined value ( $R_{\text{ESPPC}}$ ) uses the same distance definition to apply approximations to the ESP during FMO calculations.<sup>45</sup> For the present work,  $R_{\text{cut}}$  is always equal to  $R_{\text{CORS}}_D$  for FMO2 calculations and default values are used for all FMO3 calculations unless otherwise specified.

**2.2. Effective Fragment Potential Method.** An extension of the original EFP method<sup>19–21</sup> to the generalized (EFP2) method<sup>33,34</sup> facilitates the inclusion of solvent effects and the evaluation of intermolecular interactions for any molecular species. Only the EFP2 method (hereinafter referred to simply as EFP) is considered here. The EFP intermolecular energy is composed of five terms:

$$E^{\text{EFP}} = E^{\text{Coul}} + E^{\text{pol}} + E^{\text{disp}} + E^{\text{exrep}} + E^{\text{ct}} \quad (11)$$

$E^{\text{Coul}}$  is based on a distributed multipole analysis at atom centers and bond midpoints,<sup>46,47</sup> truncated at the octopole term. The polarization energy,  $E^{\text{pol}}$ , arises from the interaction of induced and permanent multipoles between fragments. The dipoles are iterated to self-consistency, enabling the EFP method to capture some of the many-body effects present in chemical systems. The third term in eq 11,  $E^{\text{disp}}$ , is expressed using an inverse  $R$  expansion:<sup>45</sup>

$$E^{\text{disp}} = \sum_n C_n R^{-n} \quad (12)$$

In the EFP method, this expansion is truncated at the leading induced dipole–induced dipole  $R^{-6}$  term. The contribution of the  $R^{-8}$  term is estimated as one-third of the  $R^{-6}$  term.<sup>48</sup> The coefficients  $C_n$  are derived from imaginary frequency-dependent polarizabilities integrated over the entire frequency range. In particular, the  $C_6$  coefficients are derived in terms of the interactions between pairs of localized molecular orbitals (LMOs), one on each molecular species.

The last two terms in eq 11,  $E^{\text{exrep}}$  and  $E^{\text{ct}}$ , are based on approximate energy expressions that depend on the intermolecular overlap of molecular orbitals. Since the EFP method uses frozen LMOs, the overlap expansion used for  $E^{\text{exrep}}$  can be reliably truncated at the quadratic term.<sup>49,50</sup> The calculation of  $E^{\text{exrep}}$  requires each generated potential to carry a basis set, making the EFP method basis set dependent. The charge transfer energy,<sup>51,52</sup>  $E^{\text{ct}}$ , between two fragments is calculated by considering the interaction between the occupied orbitals of one fragment and the virtual orbitals of a second fragment. The approximate formula used to calculate  $E^{\text{ct}}$  is based on a second-

order perturbative treatment of the intermolecular interactions. This formula is expressed in terms of canonical HF orbitals for a pair of fragments, using a truncated multipolar expansion (through quadrupoles) to represent the molecular ESP. An alternative method for the calculations of  $E^{\text{ct}}$  using quasiamionic minimal basis orbitals in place of the canonical HF orbitals is also available.<sup>53</sup> The addition of  $E^{\text{ct}}$  to the total energy results in a significant lowering of the energy for ionic and highly polar species. The original formulation<sup>51,52</sup> of  $E^{\text{ct}}$  is used throughout this work.

The first three terms in eq 11 can fail at short intermolecular distances. For example, the Coulomb interaction becomes repulsive at short distances and the polarization interaction becomes too attractive. The failures that occur at small distances can be mitigated through the addition of damping functions.<sup>54</sup> For the Coulomb term, an exponential damping function is used:<sup>32,56</sup>

$$f_{\text{damp}} = 1 - \exp(-\alpha R) \quad (13)$$

The parameter  $\alpha$  is obtained through the fitting of the damped multipole potential to the Hartree–Fock potential. An alternate approach to damping the short-range Coulomb interactions is through an approximation of the short-range charge penetration energy based on the intermolecular overlap.<sup>56,57</sup> The polarization term can be damped using either an exponential term as in eq 13 or Gaussian damping. The dispersion term can be damped using the Tang–Toennies<sup>48,58</sup> formula, or through the use of a damping formula based on the intermolecular overlap.<sup>56</sup> The latter is preferable since the intermolecular overlap is already calculated for the exchange repulsion and since no arbitrarily fitted parameter is required.

**2.3. The Effective Fragment Molecular Orbital Method.** The EFMO method was developed to integrate the FMO and EFP methods in an effort to provide a generally applicable, accurate, and efficient approach to large molecular systems. The original EFMO energy<sup>37</sup> is given by

$$E^{\text{EFMO}} = \sum_I E_I^0 + \sum_{I>J}^{R_{IJ} \leq R_{\text{cut}}} (\Delta E_{IJ}^0 - E_{IJ}^{\text{pol}}) + \sum_{I>J}^{R_{IJ} > R_{\text{cut}}} E_{IJ}^{\text{Coul}} + E_{\text{tot}}^{\text{pol}} \quad (14)$$

The standard fragment energies in the FMO method (which include the ESP of each of the other fragments) are replaced with the isolated energies described previously. The use of isolated fragment energies eliminates the need to calculate the ESP that is used in standard FMO method calculations. The many-body interaction energy formerly computed using the ESP is replaced by the total EFP polarization energy  $E_{\text{tot}}^{\text{pol}}$ . Each fragment pair polarization energy,  $E_{IJ}^{\text{pol}}$ , is subtracted from the corresponding dimer energy  $E_{IJ}^0$  to avoid double counting, since each  $E_{IJ}^{\text{pol}}$  is already contained in  $E_{\text{tot}}^{\text{pol}}$ . The separated dimer energies ( $\Delta E_{IJ}^{\text{sep}}$ ) in eq 8 are replaced by the EFP Coulomb energies  $E_{IJ}^{\text{Coul}}$ .

The original formulation of the total EFMO energy is replaced in this work to include all five components of the EFP energy as described in Section 2.2. The new expression for the FIEFMO energy is

**Table 1.** Average Total Number of Separated and QM Dimers for All Water Cluster Sizes and Values of  $R_{\text{cut}}^{\text{a}}$ 

$R_{\text{cut}}$	8 waters		16 waters		32 waters		64 waters	
	separated	QM	separated	QM	separated	QM	separated	QM
0.6	28	0	120	0	496	0	2016	0
0.8	16	12	93	27	444	52	1904	112
1.4	5	23	65	55	363	133	1730	286
2.0	0	28	30	90	237	259	1418	598

<sup>a</sup>"separated" indicates the number of dimers calculated with approximations. QM indicates the number of dimers calculated using quantum mechanics.

**Table 2.** Average Signed Errors (kcal/mol) Relative to Ab Initio Energies for the FIEFMO, EFMO, FMO2, and FMO3 Methods for Each Water Cluster Size and  $R_{\text{cut}}$  Value

$R_{\text{cut}}$	6-31++G(d,p)				6-311++G(3df,2p)			
	EFMO	FIEFMO	FMO2	FMO3	EFMO	FIEFMO	FMO2	FMO3
8 Water Molecules								
0.6	-74.9	10.4	-107.2		-72.9	4.6	-107.8	
0.8	-4.4	-5.3	-6.6		-4.0	-5.8	-8.1	
1.4	-7.7	-7.9	-13.0		-6.2	-6.5	-21.8	
2.0 <sup>a</sup>	-8.3	-8.3	-14.2	-0.8	-6.7	-6.7	-23.4	-1.2
16 Water Molecules								
0.6	-221.6	20.6	-310.7		-213.9	8.8	-312.7	
0.8	-13.8	-15.9	-23.2		-14.7	-19.4	-30.1	
1.4	-21.9	-23.0	-38.9		-19.8	-21.9	-70.7	
2.0 <sup>a</sup>	-25.6	-25.8	-48.2	-10.3	-22.5	-22.7	-85.5	-25.6
32 Water Molecules								
0.6	-456.3	63.1	-641.6		-446.8	19.6	-653.5	
0.8	-18.8	-23.1	-40.3		-29.8	-40.9	-61.0	
1.4	-38.7	-43.0	-79.8		-37.1	-46.9	-153.2	
2.0 <sup>a</sup>	-50.2	-51.1	-111.9	-52.1	-47.0	-48.4	-204.3	-115.6

<sup>a</sup> $R_{\text{cut}} = 2.0$  is the FMO default in GAMESS.

$$E^{\text{EFMO}} = \sum_I E_I^0 + \sum_{I>J}^{R_{IJ} \leq R_{\text{cut}}} (\Delta E_{IJ}^0 - E_{IJ}^{\text{pol}}) + \sum_{I>J}^{R_{IJ} > R_{\text{cut}}} (E_{IJ}^{\text{Coul}} + E_{IJ}^{\text{disp}} + E_{IJ}^{\text{exrep}} + E_{IJ}^{\text{ct}}) + E_{\text{tot}}^{\text{pol}} \quad (15)$$

Including the dispersion energy in the separated dimer energies allows the EFMO method to be used with correlated ab initio methods, such as second order Møller–Plesset perturbation theory (MP2) and coupled cluster theory (CC), to provide the dispersion energy for separated dimers, not just QM dimers.<sup>38</sup> Additionally, by including all intermolecular interactions, the user defined cutoff value  $R_{\text{cut}}$  can be reduced to neglect additional QM dimers. The reduction in QM dimers lowers the computational requirements of FIEFMO calculations relative to both the original EFMO and FMO calculations.

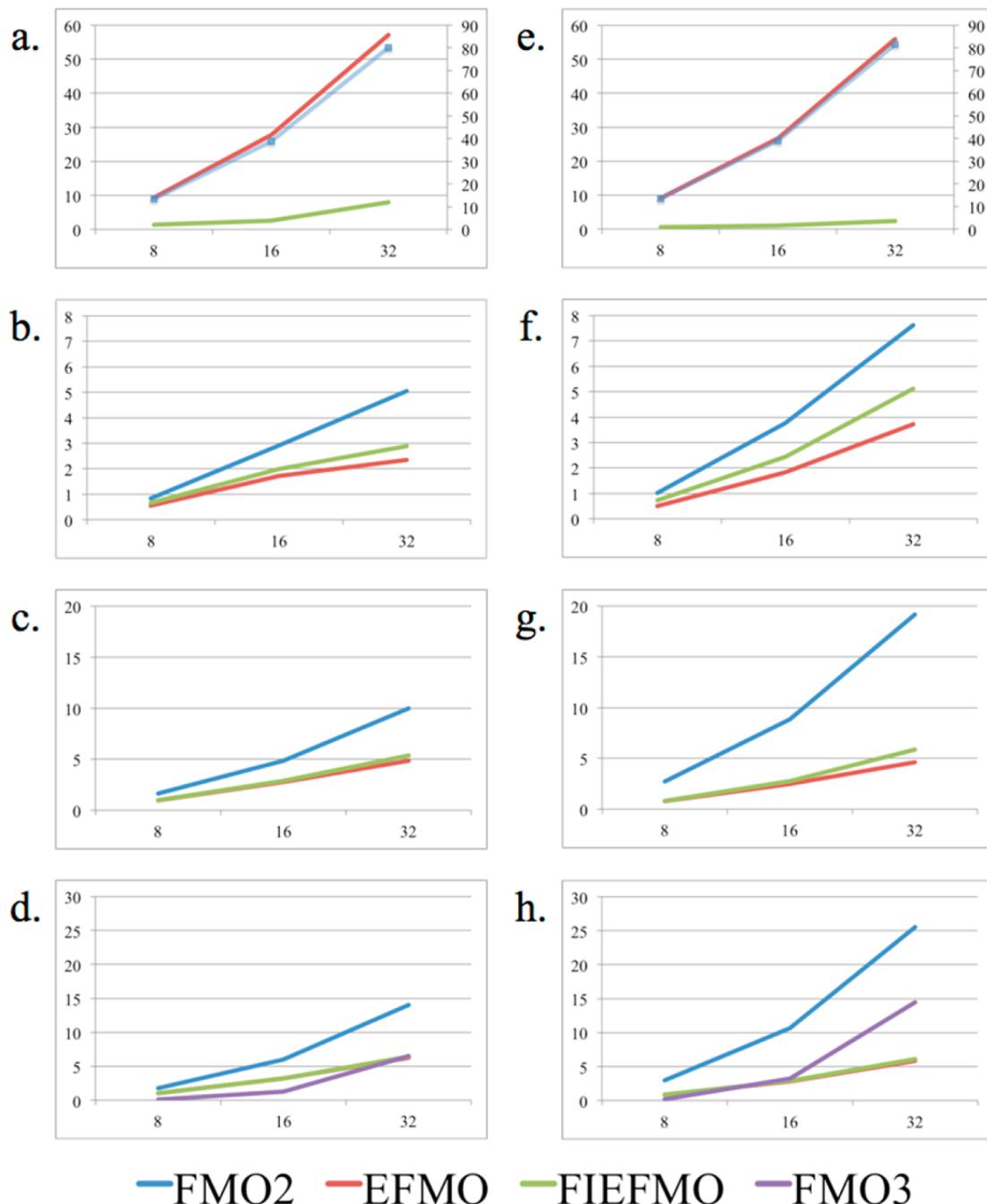
The FIEFMO method has been implemented in the GAMESS program package. All of the default screening parameters for the EFP method were used with the exception of the Coulomb energy, for which the overlap based damping<sup>56,57</sup> is used. Full alpha-polarizability tensors were used for all FIEFMO and EFMO calculations, in contrast to the approximated tensors<sup>59</sup> used in previous EFMO publications.<sup>37,38</sup>

### 3. RESULTS

The accuracy of the FIEFMO method is evaluated using a variety of molecular clusters. While the ability of the FIEFMO method to perform calculations on covalently bonded systems is not investigated in this manuscript, it is possible to perform such calculations. The topic of the FIEFMO method applied to covalently bonded systems will be the subject of a subsequent publication.

**3.1. Water Clusters.** Clusters of water molecules, shown in Supporting Information, Figures S1–S3, were used as test systems to assess the accuracy of the FIEFMO method versus the EFMO, FMO2, FMO3 and fully ab initio methods. Seven minimum energy structures for clusters of both 8 and 16 water molecules were chosen from previous work.<sup>60</sup> Larger clusters of 32 water molecules were obtained as part of a previous effort<sup>61,62</sup> using a Monte Carlo algorithm with simulated annealing (MC/SA).<sup>60,63</sup> Fully ab initio energies for each of the 8, 16, and 32 water clusters were calculated at the MP2 level of theory using two Pople-type basis sets<sup>64</sup> (6-31++G(d,p) and 6-311++G(3df,2p)).

Table 1 shows the various  $R_{\text{cut}}$  values tested and the average number of separated and QM dimer calculations performed for each cutoff value and cluster size. Recall that the smaller the value of  $R_{\text{cut}}$ , the fewer is the number of dimers that are calculated using a fully QM level of theory. For example, if  $R_{\text{cut}} = 0.6$ , no dimers are calculated with the QM level of theory for all of the water cluster sizes considered, so there are only separated dimer interactions for the FIEFMO, EFMO, and FMO2 methods. The largest value of  $R_{\text{cut}}$  in Table 1 (2.0, the



**Figure 1.** Comparison of error per water molecule ( $e/w$ ) for FIEFMO, EFMO, FMO2, and FMO3 for 8, 16, and 32 water molecules. Graphs a–d show  $e/w$  for the 6-31++G(d,p) basis set using  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0 respectively. Graphs e–h correspond to the 6-311++G(3df,2p) basis set results. In graphs a and e the FMO2 energy scale is shown on the right ordinate while the energy scale for all other methods is shown on the left ordinate. The number of water molecules is represented on the abscissa of each graph, and all energies represented on the ordinate are in kcal/mol.

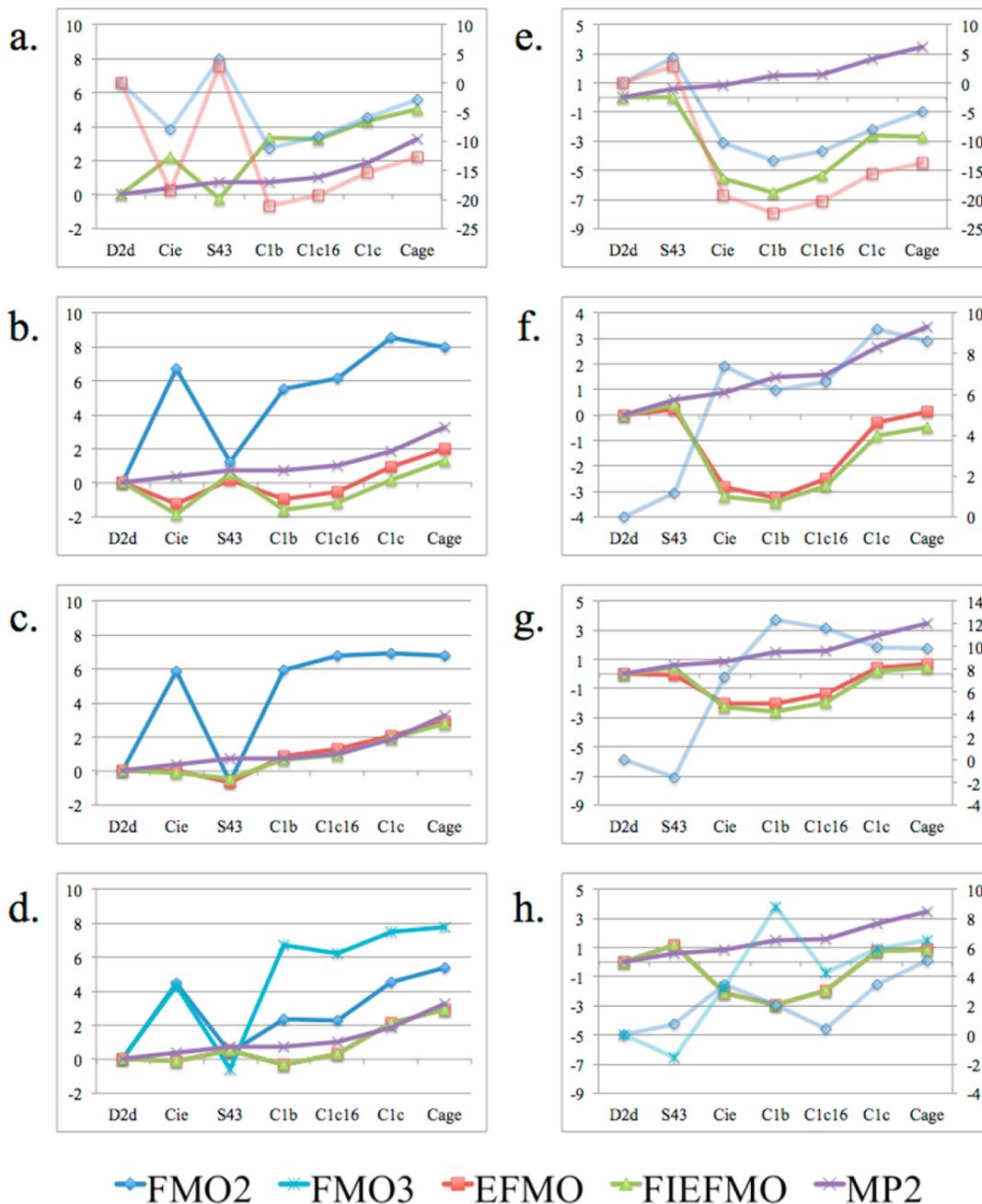
default value in GAMESS) results in the QM calculation of many more dimers, as many as 259 for the 32-water clusters.

**3.1.1. Average signed errors.** The total cluster energies produced by the FIEFMO method are compared with EFMO, FMO2, FMO3, and MP2 energies in Table 2. The average signed errors in Table 2 are calculated as

$$\text{Error} = \frac{\sum_{i=1}^n (E_i^X - E_i^{\text{MP2}})}{n} \quad (15a)$$

where  $n$  is the number of isomers and  $E_i^X$  is the FIEFMO, EFMO, FMO2, or FMO3 energy.

For the 6-31++G(d,p) basis set and  $R_{\text{cut}} = 0.6$ , the FIEFMO method produces consistently smaller errors than does FMO2. The FIEFMO errors average 10.4 kcal/mol for clusters of 8 water molecules and 63.1 kcal/mol for clusters of 32 water molecules. Average FMO2 errors are significantly larger, with an average error of -107.2 kcal/mol for clusters of 8 water molecules and -641.6 kcal/mol for clusters of 32 water molecules. The EFMO errors are closer to those produced by FMO2 than FIEFMO when  $R_{\text{cut}} = 0.6$ , with an average EFMO



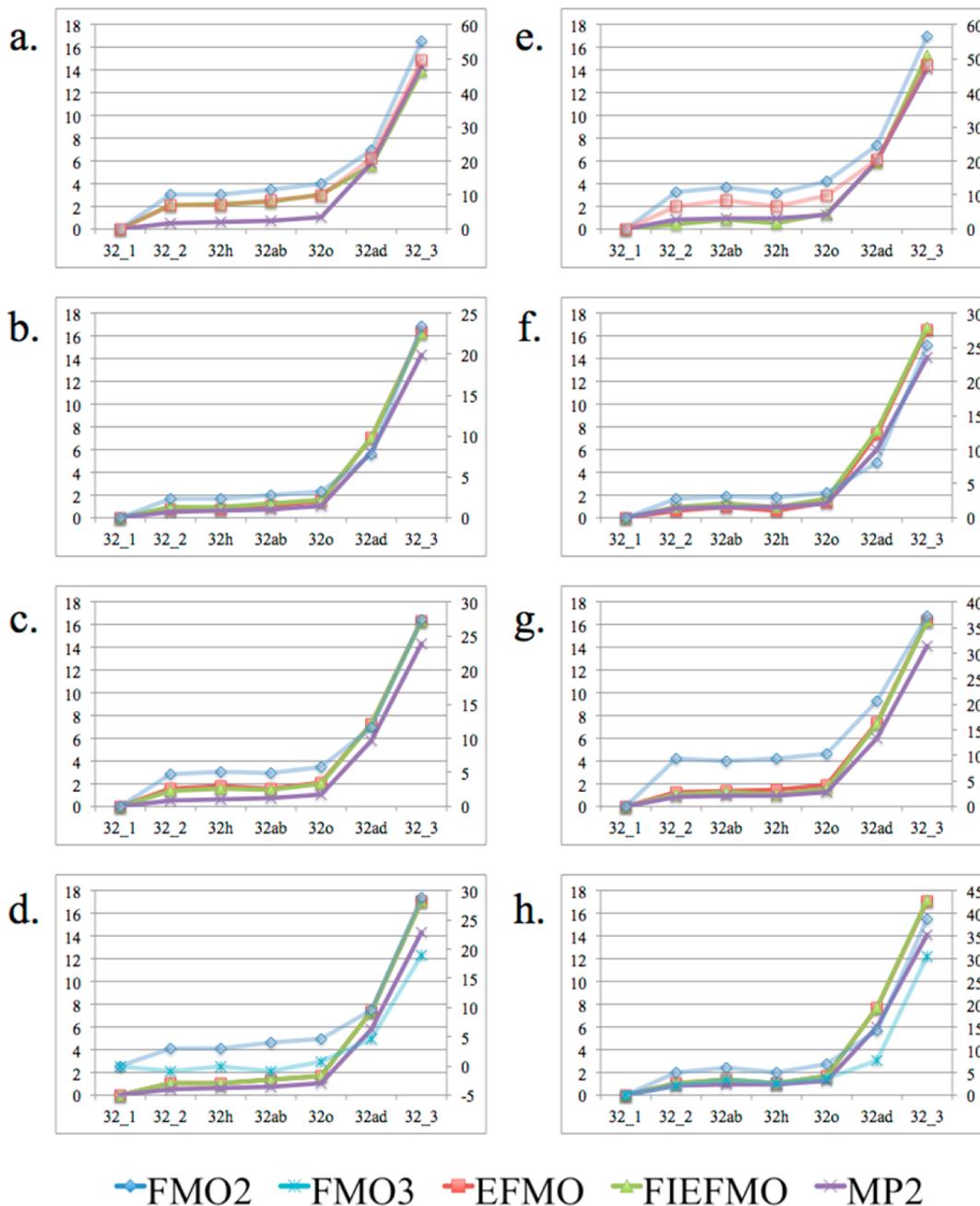
**Figure 2.** Comparison of relative energies for FIEFMO, EFMO, FMO2, FMO3, and fully MP2 for clusters of 16 water molecules. Graphs a–d show relative energies for the 6-31++G(d,p) basis set using  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0, respectively. Graphs e–h correspond to the 6-311++G(3df,2p) basis set results. In graphs a, e, f, g, and h the FMO2 energy scale is shown on the right ordinate while the energy scale for MP2 and FIEFMO results is shown on the left ordinate. The FMO3 energy scale is also on the right ordinate in graph h. Isomers are represented on the abscissa of each graph, and all energies represented on the ordinate are in kcal/mol.

error of  $-74.9$  kcal/mol for clusters of 8 water molecules compared to the average FMO2 error of  $-107.2$  kcal/mol. The large errors in the EFMO and FMO2 energies can be attributed to the lack of short-range intermolecular interactions that are included in the FIEFMO method.

Increasing  $R_{\text{cut}}$  to 0.8 provides a sharp decrease in the average EFMO and FMO2 errors for all cluster sizes, with errors between  $-4.4$  and  $-18.8$  kcal/mol for EFMO and errors between  $-6.6$  and  $-40.3$  kcal/mol for FMO2. The corresponding errors for the FIEFMO method decrease as well, falling in between  $-5.3$  and  $-23.1$  kcal/mol for all cluster sizes. The inclusion of additional QM dimers by increasing  $R_{\text{cut}}$

to 1.4 and then 2.0 shows a steady increase in average error for the FIEFMO, EFMO, and FMO2 methods, with the FIEFMO energies converging to the EFMO energies as the importance of short-range intermolecular interactions diminishes. Despite this increase in average error, the FIEFMO and EFMO methods produce smaller errors than the FMO2 method for all cluster sizes and values of  $R_{\text{cut}}$ .

The average FMO3 errors using the default values of  $R_{\text{cut}}$  and  $R_{\text{TRIM}}$  for clusters of 8 and 16 water molecules are also included in Table 2. For water clusters of 32 water molecules, the default value of  $R_{\text{ESPPC}}$  was increased from 2.5 to 3.5 to achieve SCF convergence for some of the QM trimer



**Figure 3.** Comparison of relative energies for FIEFMO, EFMO, FMO2, FMO3, and fully MP2 for clusters of 32 water molecules. Graphs a–d show relative energies for the 6-31++G(d,p) basis set using  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0, respectively. Graphs e–h correspond to the 6-311+ +G(3df,2p) basis set results. In all graphs the FMO2 and FMO3 energy scales are shown on the right ordinate while the energy scale for MP2 and FIEFMO results is shown on the left ordinate. Isomers are represented on the abscissa of each graph, and all energies are in kcal/mol.

calculations. It is important to note here that the use of a uniform field of point charges for the ESP<sup>45</sup> (i.e.,  $R_{\text{ESPPC}} = -1$ ) is capable of providing a significant reduction in the FMO2 and FMO3 error for water clusters.<sup>65</sup> However, this increase in accuracy is system dependent and therefore only the FMO3 default settings are investigated in this work. While the value of  $R_{\text{cut}}$  is varied for FMO2 in the present work,  $R_{\text{ESPPC}} = -1$  is not investigated, since this setting is not relevant for FIEFMO. For the smaller cluster sizes of 8 and 16 water molecules, FMO3 outperforms FIEFMO, EFMO, and FMO2 for all  $R_{\text{cut}}$  values, with average errors of -0.8 and -10.3 kcal/mol, respectively. However, for the largest cluster size of 32 water molecules, the FIEFMO method using  $R_{\text{cut}} = 0.6$  produces an average error

that is 11.0 kcal/mol smaller than that obtained using FMO3. For all values of  $R_{\text{cut}}$  greater than 0.6, the FIEFMO and EFMO methods outperform the FMO3 method by as much as 33.3 kcal/mol.

Errors obtained using the larger 6-311++G(3df,2p) basis set and  $R_{\text{cut}} = 0.6$  follow the same trend of increasing errors with cluster size reported above for the smaller basis set. In general, average FMO2 errors increase for all cluster sizes compared to those obtained using the smaller basis set. However, the average FIEFMO errors drop substantially for all cluster sizes when the larger basis set is used. For example, the average FIEFMO error for the 8-water clusters with  $R_{\text{cut}} = 0.6$  is 4.6 kcal/mol for the larger basis set vs 10.4 kcal/mol for the smaller basis set. The

Table 3. Binding Energy Per Water Molecule for FIEFMO, EFMO, FMO2, FMO3 and MP2 (kcal/mol)

$R_{\text{cut}}$	6-31++G(d,p)					6-311++G(3df,2p)				
	EFMO	FIEFMO	FMO2	FMO3	MP2	EFMO	FIEFMO	FMO2	FMO3	MP2
8 Water Molecules										
0.6	-20.1	-9.5	-24.2		-10.8	-18.4	-8.8	-22.8		-9.3
0.8	-11.3	-11.4	-11.6		-10.8	-9.8	-10.0	-10.3		-9.3
1.4	-11.7	-11.7	-12.4		-10.8	-10.1	-10.1	-12.1		-9.3
2.0	-11.8	-11.8	-12.5	-10.9	-10.8	-10.2	-10.2	-12.2	-9.5	-9.3
16 Water Molecules										
0.6	-24.2	-9.1	-29.8		-10.4	-22.4	-8.5	-28.6		-9.0
0.8	-11.2	-11.4	-11.8		-10.4	-9.9	-10.2	-10.9		-9.0
1.4	-11.7	-11.8	-12.8		-10.4	-10.3	-10.4	-13.4		-9.0
2.0	-12.0	-12.0	-13.4	-11.0	-10.4	-10.4	-10.4	-14.4	-19.8	-9.0
32 Water Molecules										
0.6	-25.1	-8.8	-30.8		-10.8	-23.2	-8.7	-29.7		-9.3
0.8	-11.4	-11.5	-12.1		-10.8	-10.2	-10.5	-11.2		-9.3
1.4	-12.0	-12.1	-13.3		-10.8	-10.4	-10.7	-14.1		-9.3
2.0	-12.4	-12.4	-14.3	-12.4	-10.8	-10.7	-10.8	-15.7	-12.9	-9.3

analogous comparison for the 32-water clusters is 19.6 kcal/mol vs 63.1 kcal/mol. This significant reduction in the FIEFMO method errors is due to the use of a better EFP potential generated with the larger basis set,<sup>49</sup> and can be observed in the EFMO errors as well. Increasing the number of QM dimers causes an increase in error for FMO2, FIEFMO, and EFMO. However, the errors produced by the FIEFMO method converge more quickly for each cluster size. For example, with  $R_{\text{cut}} = 0.8, 1.4$ , and  $2.0$  the FIEFMO method produces errors of  $-5.8, -6.5$ , and  $-6.7$  kcal/mol for the smallest clusters of 8 water molecules. The corresponding FMO2 errors are  $-8.1, -21.8$ , and  $-23.4$  kcal/mol. This behavior is consistent for all cluster sizes, with the FIEFMO method errors differing by less than 7.5 kcal/mol when  $R_{\text{cut}}$  is increased from 0.8 to 2.0.

In contrast to the relative decrease in errors of the FIEFMO and EFMO methods when using a larger basis set, the FMO3 average errors roughly double in magnitude when the basis set size is increased. This doubling of the error is evident for all cluster sizes, with the most significant increase from  $-52.1$  to  $-115.6$  kcal/mol for the 32 water clusters.

Now, consider the increase in absolute error per water molecule ( $e/w$ ) for each method relative to the fully ab initio results. Figure 1 shows the  $e/w$  for FIEFMO, EFMO, FMO2, and FMO3 for all cluster sizes and both basis sets. Graphs a through d present results for  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0, respectively, for the 6-31++G(d,p) basis set. Graphs e through f present the corresponding data for the 6-311++G(3df,2p) basis set. The absolute error per water molecule increases with increasing cluster size for all methods; however, the rate of increase is significantly larger for FMO2 compared to both EFMO and FIEFMO. Indeed, the  $e/w$  for the FIEFMO method and an  $R_{\text{cut}}$  value of 0.6 remains nearly flat, particularly when using the larger basis set. The rate of increase for FMO3 is comparable to EFMO and FIEFMO for the smaller basis set; however for the larger basis set the  $e/w$  for the FMO3 method rises sharply for the larger clusters of 32 water molecules.

While the FMO2 and FMO3 methods include some level of higher order many body interactions from the iterated ESP, the FIEFMO and EFMO methods clearly outperform the FMO2 method through the use of the EFP many body polarization while performing as many or fewer QM dimer calculations as the FMO2 method. This is particularly noteworthy, since the

calculation of the converged ESP during FMO2 and FMO3 calculations requires approximately 15 times the computational effort compared to the FIEFMO and EFMO methods many body polarization calculation. If a uniform field of point charges is used to represent the ESP ( $R_{\text{ESPPC}} = -1$ ) during FMO2 and FMO3 calculations, the monomer densities must still be iterated to self-consistency. While this alternative formulation of the ESP can lead to reductions in error, only modest computational savings are obtained, much less than is obtained with the EFMO and FIEFMO methods.

**3.1.2. Relative Energies.** Seven different isomers were chosen for each cluster size to test the ability of the FIEFMO, EFMO, FMO2, and FMO3 methods to reproduce relative energies compared to MP2 results. The same set of  $R_{\text{cut}}$  values (0.6, 0.8, 1.4, and 2.0) that were discussed in Section 3.1 are used. Clusters of 16 and 32 water molecules are discussed here. The behavior of 8-water clusters, shown in the Supporting Information, is similar to that of the larger clusters and is not discussed further.

Relative energies for the 16-water isomers are shown in Figure 2, following the same scheme as in Supporting Information, Figure S4. Results for FIEFMO, EFMO, and FMO2 using the smaller basis set are not as accurate as the results for the 8-water isomers. The most accurate results are now produced using  $R_{\text{cut}} = 1.4$  for both FIEFMO and EFMO, while the most accurate FMO2 results are produced when using  $R_{\text{cut}} = 2.0$ . The FMO3 method fails to reproduce the MP2 relative energies, overestimating the four higher energy isomers to a greater extent than even FMO2. The accuracy for all methods when using the larger basis set is relatively poor. The inability of any of the methods to accurately reproduce the fully MP2 relative energies of the 16 water molecule clusters may be due to the highly symmetric structural arrangements of the 16 water clusters, leading to unusually large electron delocalization.<sup>65</sup> These unique structures are not found in the more globular arrangements of the larger 32 water clusters.

When the size of the clusters is increased to 32 water molecules, there are no highly structured isomers; globular droplets are more prevalent. Figure 3 shows that for both the smaller and larger basis sets, the FIEFMO method consistently produces extremely accurate relative energies, in some cases (Figures 3b and 3e) nearly overlaying the MP2 curve exactly. The relative energies produced by the FMO2 method are not

dissimilar to the MP2 results; however, the two lowest energy structures are found to be farther apart in energy than the relative energies obtained by the FIEFMO method or by the MP2 method. Consequently, the FMO2 relative energy curve is shifted up, in all cases predicting the highest energy isomer to be greater than 10 kcal/mol higher than predicted by either FIEFMO or MP2; the largest difference in Figure 2e is ~40 kcal/mol. While the FMO3 method produces the correct isomer ordering compared to the fully MP2 results, the entire relative energy curve is shifted up in energy by as much as 15 kcal/mol.

**3.1.3. Average Binding Energy per Water Molecule.** As an additional test of the FIEFMO method, the average binding energy per molecule was calculated for all clusters and basis sets for the FIEFMO, EFMO, FMO2, and FMO3 methods and compared to the MP2 results in Table 3. Since each water molecule has the same internal geometry, the binding energy per molecule ( $E_{\text{binding/molecule}}$ ) can be easily evaluated for each cluster (eq 16) by subtracting the gas phase monomer energy ( $E^{\text{Monomer}}$ ), multiplied by the appropriate number of monomers, from the total energy of the cluster ( $E^{\text{Cluster}}$ ), and dividing by the number of monomers.

$$E_{\text{binding/molecule}} = \frac{E^{\text{Cluster}} - ((\# \text{of monomers}) \times E^{\text{Monomer}})}{n} \quad (16)$$

The average binding energy for each cluster size is obtained by summing  $E_{\text{binding/molecule}}$  and dividing by the number of isomers.

The FIEFMO method outperforms the FMO2 method for all clusters and basis sets investigated. The FMO2 errors compared to MP2 for the smallest basis set are in the range 2–19 kcal/mol, while the corresponding FIEFMO binding energies are in error by only 1–5 kcal/mol. For the larger basis set, the FIEFMO method produces even more accurate binding energies, with errors of 0.5 and 0.6 kcal/mol. The FMO2 errors do not change appreciably when the larger basis set is used, still falling in the range of 1–20 kcal/mol. The most accurate FMO2 binding energies are in error by 1.9–2.1 kcal/mol compared to MP2 for  $R_{\text{cut}} = 0.8$ . In comparison, the most accurate FIEFMO binding energies are in error by only 0.5 kcal/mol for  $R_{\text{cut}} = 0.6$ ; the largest FIEFMO error of 1.5 kcal/mol is comparable in magnitude to the most accurate FMO2 binding energy. In general, the EFMO binding energies mirror those of FMO2 for  $R_{\text{cut}} = 0.6$ , while the EFMO binding energies for all other values of  $R_{\text{cut}}$  are nearly identical to the FIEFMO binding energies. It is especially noteworthy that the largest difference in predicted binding energies between FMO2 and FIEFMO occurs consistently for  $R_{\text{cut}} = 0.6$ . This means that many fewer fully MP2 calculations are required if one uses the FIEFMO method that is presented here.

**3.2. Water–Methanol Mixtures.** Now, consider the second test system, clusters of water and methanol molecules. Geometries were obtained using the same MC/SA procedure outlined for the water clusters in Section 3.1. Isomers obtained from the MC/SA procedure were additionally optimized to their corresponding local minima using the EFP method. Clusters were a 50/50 mixture, consisting of three cluster sizes of 4 molecules, 8 molecules, and 16 molecules. Three isomers of each cluster size were chosen for a total of 9 clusters, shown in Supporting Information, Figure S5. Single point energies of each cluster were calculated using the FIEFMO, EFMO, FMO2, FMO3 methods and compared to fully MP2 energies using the same two Pople-type basis sets as discussed in Section

3.1. Table 4 shows the average number of separated and QM dimer calculations performed for each water/methanol cluster using the four  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0.

**Table 4. Average Total Number of Separated and QM Dimers for All Water/Methanol Cluster Sizes and Values of  $R_{\text{cut}}^a$**

$R_{\text{cut}}$	2 H <sub>2</sub> O/2 CH <sub>3</sub> OH		4 H <sub>2</sub> O/4 CH <sub>3</sub> OH		8 H <sub>2</sub> O/8 CH <sub>3</sub> OH	
	separated	QM	separated	QM	separated	QM
0.6	6	0	28	0	120	0
0.8	2	4	16	12	98	22
1.4	0	6	3	25	77	43
2.0	0	6	0	28	40	80

<sup>a</sup>“separated” indicates the number of dimers calculated with approximations. QM indicates the number of dimers calculated using quantum mechanics.

**3.2.1. Average Signed Errors.** Average signed errors are shown in Table 5, comparing the FIEFMO, EFMO, FMO2, and FMO3 energies to fully MP2 results. For the smaller basis set, the FMO2 errors follow the same general trend that is seen in the water cluster results; the largest error is produced for the  $R_{\text{cut}}$  value of 0.6, with a sharp drop in error when nearest neighbor QM dimer calculations are performed. As more QM dimer calculations are included, the FMO2 error rises. For the smallest cluster of 4 molecules the FMO2 error rises from −0.7 kcal/mol for  $R_{\text{cut}} = 0.8$  to −2.1 kcal/mol for  $R_{\text{cut}} = 2.0$ . The behavior of the FMO2 errors is comparable to the FIEFMO and EFMO results that fall in the ranges of −0.9 to −1.3 kcal/mol and −0.6 to −1.3 kcal/mol, respectively. However, as the cluster size is increased to 8 molecules the FMO2 error increases as well, to as much as 6.6 kcal/mol using the default  $R_{\text{cut}}$  value of 2.0. The FIEFMO and EFMO errors grow in a more modest fashion, increasing by only 2.9 kcal/mol for  $R_{\text{cut}} = 2.0$ . The FMO3 method produces the smallest average errors of 0.1 kcal/mol for the 4 molecule clusters and −1.4 kcal/mol for the cluster of 8 molecules. However, when the cluster size increases to 16 molecules the FMO3 error increases to −7.8 kcal/mol.

While the results produced for the range of  $R_{\text{cut}}$  values from 0.8 to 2.0 follow the same general trend as the pure water clusters, the results for the smallest  $R_{\text{cut}}$  value of 0.6 is dramatically different. For all of the pure water clusters, the most accurate FIEFMO results using the smaller basis set are produced when no QM dimer calculations were performed. The opposite behavior is observed for the water/methanol mixtures; when no QM dimer calculations are performed using the FIEFMO method the highest errors are reported, as much as 17.2 kcal/mol for the largest cluster of 16 molecules. When the  $R_{\text{cut}}$  value is set to 0.8 the EFMO results show some unusual behavior as the size of the system is increased to 16 molecules, producing an error of 3.5 kcal/mol. Similar behavior is seen for the larger basis set and, when compared to results from the FIEFMO method, a significant contribution from the dispersion energy (not included in the original EFMO method) is observed for these clusters. The net contribution from the short-range interactions is between 0.4 and 4.3 kcal/mol for the smaller basis set and consistently on the order of −6 kcal/mol for the larger basis set.

For the larger basis set the FIEFMO errors are reduced because of the improved potential generated using the larger basis set for all  $R_{\text{cut}}$  values except 0.8, where the errors increase

**Table 5. Average Signed Errors (kcal/mol) Relative to Ab Initio Energies for the FIEFMO, EFMO, FMO2 and FMO3 Methods for Each Water-Methanol Cluster Size and  $R_{\text{cut}}$  Value**

$R_{\text{cut}}$	6-31++G(d,p)				6-311++G(3df,2p)			
	EFMO	FIEFMO	FMO2	FMO3	EFMO	FIEFMO	FMO2	FMO3
2 Water Molecules +2 Methanol Molecules								
0.6	-15.8	2.2	-26.6		-15.6	-1.2	-27.6	
0.8	-0.6	-0.9	-0.7		-0.4	-0.8	-1.0	
1.4	-1.3	-1.3	-2.1		-0.8	-0.8	-3.8	
2.0	-1.3	-1.3	-2.1	0.1	-0.8	-0.8	-3.8	-0.3
4 Water Molecules +4 Methanol Molecules								
0.6	-34.3	6.0	-61.0		-34.0	-2.5	-63.0	
0.8	-0.3	-1.4	-2.2		0.0	-2.1	-4.0	
1.4	-3.2	-3.6	-6.4		-1.8	-2.4	-11.6	
2.0	-4.2	-4.2	-8.7	-1.4	-2.5	-2.6	-14.7	-4.8
8 Water Molecules +8 Methanol Molecules								
0.6	-74.9	17.2	-134.8		-76.0	-4.3	-140.8	
0.8	3.5	-0.5	-2.2		2.1	-4.9	-9.0	
1.4	-4.5	-6.9	-10.7		-3.4	-7.0	-23.7	
2.0	-10.1	-10.5	-24.5	-7.8	-7.0	-7.5	-42.6	-26.9

by up to  $-4.4$  kcal/mol. Errors for the FIEFMO method fall within the range of  $-0.8$  to  $-7.5$  kcal/mol, compared to a range of  $-0.5$  to  $-10.5$  kcal/mol for the smaller basis set. In contrast, the errors for the FMO2 method increase by an average of  $1.2$  kcal/mol for the 4-molecule clusters,  $3.8$  kcal/mol for the 8-molecule clusters, and  $11$  kcal/mol for the 16-molecule clusters. Errors for the FMO3 method also increase by  $0.4$  to  $19.1$  kcal/mol.

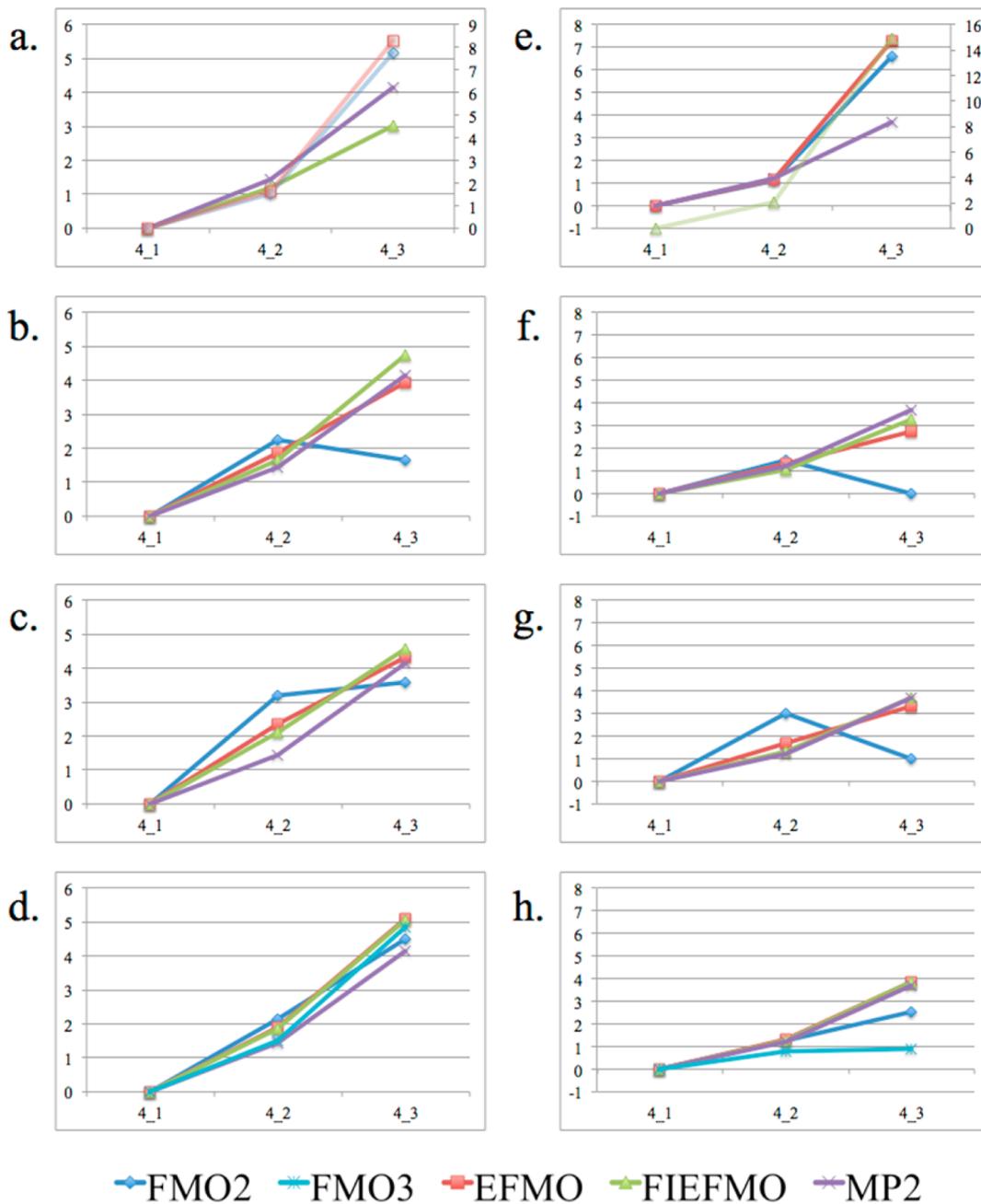
**3.2.2. Relative Energies.** The ability of the FIEFMO method to reproduce the relative energies of three isomers for each water–methanol cluster size is now compared to EFMO, FMO2, FMO3, and fully MP2 results. Results for the same four  $R_{\text{cut}}$  values and two basis sets previously described are shown in Figures 4 and 5. Only clusters of 8 and 16 molecules are discussed here. The behavior of clusters containing only 4 molecules, shown in the Supporting Information, is similar to that of the larger clusters.

Relative energies for larger cluster sizes of 8 molecules are shown in Figure 4 using the same layout as Figures 2 and 3. Energetics for the larger clusters follows the same general trend observed for the smaller basis set, with FIEFMO providing fairly good agreement with the smallest  $R_{\text{cut}}$  value. An increase in  $R_{\text{cut}}$  to  $0.8$  and  $1.4$  for this system produces poorer agreement between FMO2 and fully MP2 energies, while the FIEFMO and EFMO results improve. Only when  $R_{\text{cut}}$  is increased to the default setting of  $2.0$  do the FMO2 relative energies agree with the MP2 results. Results from the FMO3 method for the smaller basis set also provide relatively good agreement with fully MP2 results, slightly better than either FIEFMO or EFMO results in the case of isomer 4\_2. When the basis set is increased, all results remain fairly consistent when compared to the results from the smaller basis set. The agreement between FIEFMO and fully MP2 results for  $R_{\text{cut}} = 0.6$  is the only instance of a serious degradation in accuracy. The ordering of the isomers remains correct; however, the entire FIEFMO relative energy curve is shifted up in energy and is plotted on the right ordinate axis. Relative energies for FMO2 and FMO3 using the default value of  $R_{\text{cut}}$  show an underestimation of the energy for the highest energy isomer 4\_3. In all cases the FIEFMO method again outperforms the FMO2 and FMO3 methods in terms of accuracy.

Results obtained from the largest cluster of 16 molecules (Figure 5) shows the most variability among the accuracy of each method. Relative energies for the smaller basis set in Figures 5a through 5d show the FMO2 method both overestimating (Figure 5a) and underestimating (Figures 5b–d) the energies of isomers 8\_2 and 8\_3. All FIEFMO results, with the exception of Figure 4b, show very good agreement with fully MP2 relative energies. The EFMO results are quite poor for the smaller two  $R_{\text{cut}}$  values ( $0.6$  and  $0.8$ ), while the FMO3 method produces the poorest results of all methods (Figure 5d). An increase in basis set does little to improve the relative energies produced by the FMO2 and FMO3 methods. The FIEFMO results improve for  $R_{\text{cut}} = 0.8$  while the EFMO relative energies for  $R_{\text{cut}} = 0.6$  remain relatively unchanged in comparison to the smaller basis set. A small shift of  $\sim 3$  kcal/mol in the FIEFMO relative energy curve for  $R_{\text{cut}} = 0.6$  is also observed for the larger basis set; however, the ordering of the isomers remains correct.

**3.3. Energy Decomposition Analysis.** An additional benefit of the FIEFMO method lies in its ability to provide a detailed energy decomposition analysis (EDA) for each dimer interaction from a single calculation. A small test system of 8 water molecules and 2 benzene molecules (Figure 6) is used to demonstrate this capability. The accuracy of the EDA provided by the EFP method compared to fully ab initio methods has already been extensively studied elsewhere.<sup>56</sup> Therefore, this work addresses the efficacy of the newly developed FIEFMO method in providing such information.

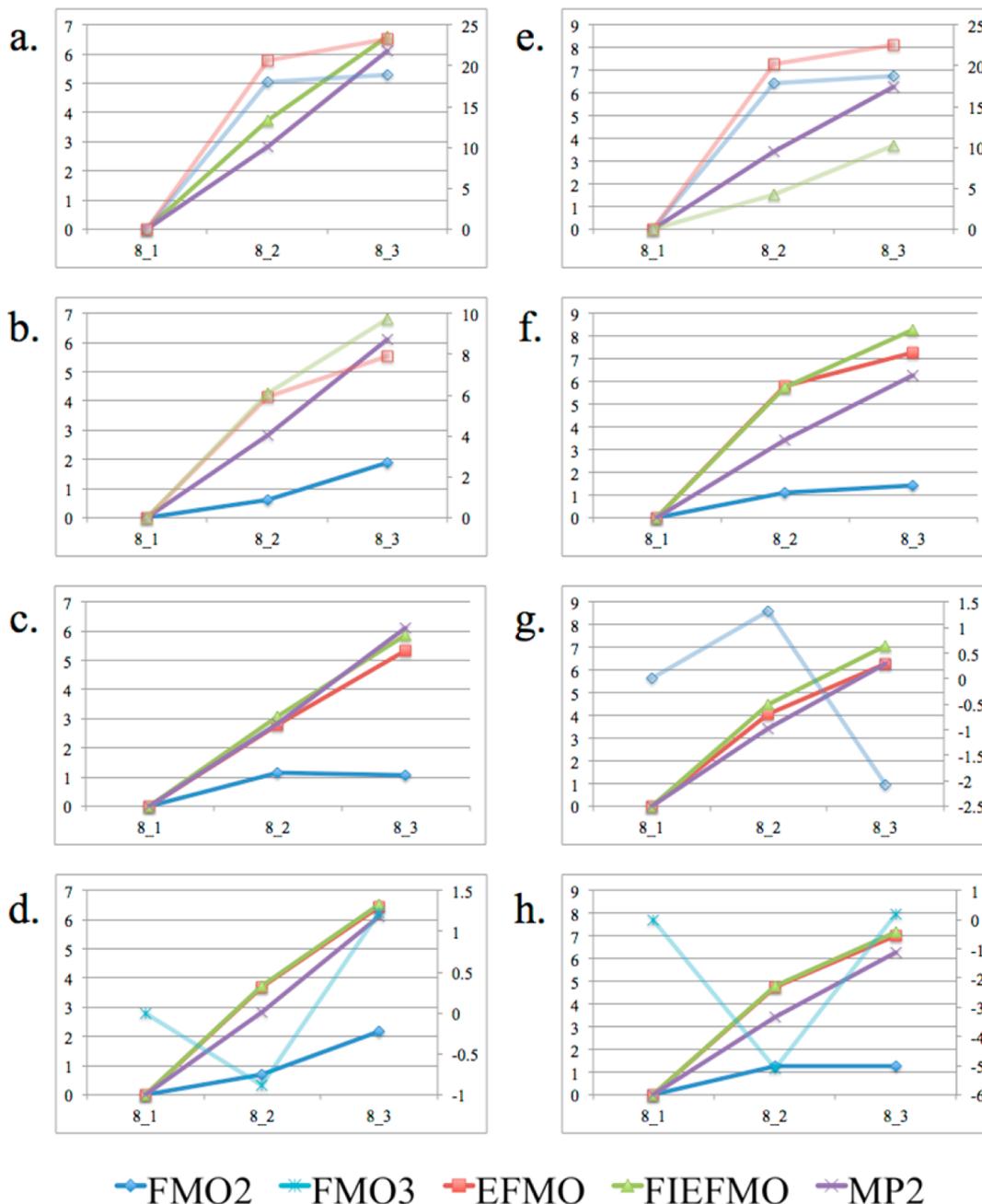
When an FIEFMO calculation on a molecular cluster is performed with a small enough  $R_{\text{cut}}$  value, in this work  $R_{\text{cut}} = 0.6$ , all QM dimer calculations are replaced with separated dimer interactions evaluated using the EFP2 method. This type of calculation is not only capable of a higher level of accuracy than the corresponding FMO2 and FMO3 calculations, but also provides insight into the individual intermolecular interactions through a detailed EDA. Four of the five types of intermolecular interactions described in the introduction are obtained for each dimer interaction, with the exception of the nonadditive polarization interaction. Figure 7 shows a decomposition of each dimer interaction into these four interaction types: electrostatic, dispersion, repulsion, and charge transfer. Dimer pairs (labeled in Figure 6) are listed along the



**Figure 4.** Comparison of relative energies for FIEFMO, EFMO, FMO2, FMO3, and fully MP2 for clusters of 4 water molecules and 4 methanol molecules. Graphs a–d show relative energies for the 6-31++G(d,p) basis set using  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0 respectively. Graphs e–h correspond to the 6-311++G(3df,2p) basis set results. In graph a the EFMO and FMO2 energy scale is shown on the right ordinate while the energy scale for MP2 and FIEFMO results is shown on the left ordinate. In graph e the FIEFMO energy scale is shown on the right ordinate while the energy scale for all other methods is shown on the left ordinate. Isomers are represented on the abscissa of each graph, and all energies are in kcal/mol.

abscissa while the ordinate shows the interaction energy in units of kcal/mol. Note the important contribution charge transfer makes to the total interaction energy of the cluster. Particularly in homogeneous water dimer interactions, charge transfer accounts for 0.02–0.7 kcal/mol of the total interaction energy. This contribution, when summed over all dimer interactions, plays an important role in the relative energies of different isomers, contributing over 5 kcal/mol to the total interaction energy of a single cluster. Dispersion interactions play an even more important role, not only in homogeneous water dimers but also in benzene-benzene and water-benzene

interactions.<sup>55</sup> Nearly 2 kcal/mol of interaction energy is accounted for by the dispersion term in the benzene-benzene interaction of this cluster, while the water-benzene dispersion energies add up to 7.9 kcal/mol of interaction energy for the entire cluster. At short intermolecular distances (less than 2.0 Å) the repulsion energy contributes nearly as much to the total interaction energy as the electrostatic energy. This information would not be as easily attainable were it not for the addition of the short-range intermolecular interactions in the FIEFMO method.



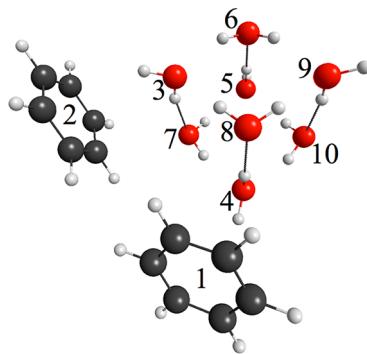
**Figure 5.** Comparison of relative energies for FIEFMO, EFMO, FMO2, FMO3, and fully MP2 for clusters of 8 water molecules and 8 methanol molecules. Graphs a–d show relative energies for the 6-31++G(d,p) basis set using  $R_{\text{cut}}$  values of 0.6, 0.8, 1.4, and 2.0 respectively. Graphs e–h correspond to the 6-311++G(3df,2p) basis set results. The following graphs show the specified results on the right ordinate axis while the energy scale for other methods is shown on the left ordinate: FMO2 energies in graphs a, e, and g; FIEFMO and EFMO energies in graphs a, b, and e; FMO3 energies in graphs d and h. Isomers are represented on the abscissa of each graph, and all energies are in kcal/mol.

The total interaction energy obtained from the FIEFMO method is compared to the total MP2 interaction energy in Table 6. The sum of two-body interactions from the FIEFMO method provides approximately 78% of the total interaction energy, with the addition of the many-body polarization energy bringing the total FIEFMO interaction energy to within 4.3 kcal/mol.

**3.4. Timings.** Timing comparisons are presented in Table 7 for the FIEFMO, EFMO, FMO2, FMO3, and MP2 methods, using the 6-311++G(3df,2p) basis set on 10 compute nodes, each containing a single 2.66 GHz hex core Intel Xeon X5650 CPU, connected with 4X Quad Data Rate InfiniBand. Timings

are compared for  $R_{\text{cut}}$  values of 0.6 and 2.0. The numbers of QM and separated dimers for each of these values of  $R_{\text{cut}}$  may be found in Tables 1 and 4.

For the smallest test system of 8 water molecules and 2 benzene molecules, the wall time for the fully MP2 calculation is 3955 s. The FMO2 method provides a reduction in wall clock time of ~1130 s with an error of -17.9 kcal/mol when  $R_{\text{cut}} = 2.0$ . The FIEFMO method produces an error of 6.6 kcal/mol while reducing the total wall time to ~602 s. The EFMO timings are nearly identical to the FIEFMO timings, taking 512 and 1994 s for  $R_{\text{cut}} = 0.6$  and 2.0 respectively. While the FMO3 method produces the smallest error of all four methods at -3.2



**Figure 6.** Cluster of 8 water molecules and 2 benzene molecules used for the energy decomposition analysis calculation. Molecules are labeled 1 through 10.

kcal/mol, the corresponding calculation takes over 2000 s longer than the fully MP2 calculation. This reflects the inefficiency of using the FMO3 method for such a small system.

For the system of 8 water molecules and 8 methanol molecules, the FIEFMO method again outperforms the FMO2 method in terms of accuracy and wall clock times. For  $R_{\text{cut}} = 2.0$ , the FMO2 method produces an error of -40.2 kcal/mol in 224 s, compared to an error of 8.1 kcal/mol produced in just 15 s using the FIEFMO method. The FMO3 error is nearly half that of the FMO2 error, but the corresponding calculation takes roughly an order of magnitude longer at 2268 s. For this slightly larger cluster, the FMO3 time is now about 1/3 of that of the full MP2 calculation. As was the case for the smaller test system, the FIEFMO timings are nearly identical to the EFMO timings.

The largest test system of 32 water molecules follows the same trend as the timings discussed for the smaller test systems. While the FMO2 method is capable of reducing wall clock times significantly compared to fully MP2 calculations, the errors are more than an order of magnitude larger than those

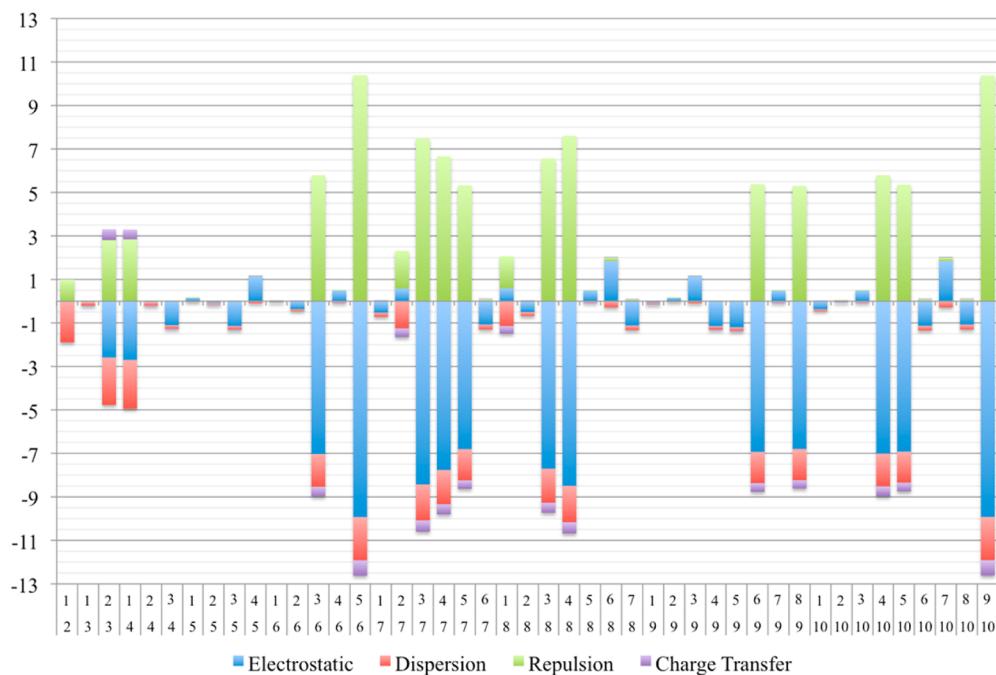
**Table 6. Contributions to the Total Intermolecular Interaction Energy (kcal/mol) for All Dimer Interactions As Well As the Total Polarization Energy in a Cluster of 8 Water Molecules and 2 Benzene Molecules Compared to the Total MP2 Interaction Energy**

	FIEFMO	MP2
Two-Body		
electrostatic	-93.5	
repulsion	86.9	
dispersion	-31.1	
charge transfer	-6.0	
polarization	-16.3	
Many-Body		
polarization	-12.8	
total	-72.9	-77.2

obtained using the FIEFMO method. Additionally, the difference between the smallest errors produced by the EFMO and FIEFMO methods begins to increase (-47.8 kcal/mol versus 19.5 kcal/mol) while the wall clock times are nearly identical. For all test system sizes, the FIEFMO method shows the ability to reduce wall clock times between 79 and 96% compared to the FMO2 method while producing superior results.

#### 4. CONCLUSIONS

The fully integrated EFMO method now includes all five intermolecular interactions that are present in the EFP method. The accuracy of the FIEFMO method was tested versus the standard EFMO, FMO2, FMO3, and MP2 methods. The FIEFMO method was shown to provide superior average errors, relative energies, and binding energies for all cluster sizes and basis sets compared to the FMO2 method. In many instances, the FIEFMO method also outperforms the FMO3 method in terms of accuracy. In addition to increased accuracy,



**Figure 7.** Energy contributions of each intermolecular interaction for all dimer interactions in a cluster of 8 water molecules and 2 benzene molecules. Molecule labeling on abscissa used from Figure 6 while the energy scale on the ordinate is in kcal/mol.

**Table 7.** Timing Comparison<sup>a,b</sup> for FIEFMO, EFMO, FMO2, FMO3, and Fully MP2 Energy Calculations Performed on Water Clusters

$R_{\text{cut}}$	6-311++(3df,2p)								
	EFMO		FIEFMO		FMO2		FMO3		MP2
	wall time	error	wall time	error	wall time	error	wall time	error	wall time
8 Water Molecules +2 Benzene Molecules									
0.6	512	-38.6	602	6.6	1414	-65.8			
2.0	1994	-4.8	2178	-4.9	2825	-17.9	6391	-3.2	3955
8 Water Molecules +8 Methanol Molecules									
0.6	10	-87.1	15	8.1	70	-149.8			
2.0	117	-7.7	151	-8.3	224	-40.2	2268	-22.1	6793
32 Water Molecules									
0.6	10	-457.7	8	19.5	53	-668.5			
2.0	105	-47.8	109	-49.2	184	-211.7	2806	-119.3	

<sup>a</sup>All timings performed on 10 nodes containing 6 CPU cores each. <sup>b</sup>Times in are in seconds, and errors are in kcal/mol.

the FIEFMO method is capable of providing a detailed EDA of short and mid range interactions in a single calculation, providing valuable information about the nature of intermolecular interactions in clusters. Simultaneously, one obtains significant reductions in wall clock times compared to the FMO2 and FMO3 methods. Through the reduction of the number of explicit QM dimers performed during FIEFMO method calculations, time savings of up to 96% compared to the FMO2 method were achieved while at the same time providing a more accurate estimate of the MP2 energies.

Future work on the FIEFMO method will include the implementation of fully analytic gradients for the newly added energy terms to enable geometry optimizations and molecular dynamics simulations.<sup>66,67</sup> With the addition of energy gradients, as well as the improved accuracy and reduction in computational requirements versus the FMO2 method, the FIEFMO method will be able to provide MP2-quality MD simulations at a fraction of the cost.

## ASSOCIATED CONTENT

### S Supporting Information

Further details are given in Figures S1–S6. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: mark@si.msg.chem.iastate.edu.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported in part by an SI2 grant (to M.S.G.) from the National Science Foundation and by the IRENE Collaborative Project (J.H.J.) funded by the EU framework 7 program.

## REFERENCES

- (1) Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer: Dordrecht, The Netherlands, 1995; pp 47–108.
- (2) Söderhjelm, P.; Aquilante, F.; Ryde, U. *J. Phys. Chem. B* **2009**, *113*, 11085.
- (3) Pomogaev, V.; Pomogaeva, A.; Aoki, Y. *J. Phys. Chem. A* **2009**, *113*, 1429.
- (4) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459.
- (5) Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1573.
- (6) Suárez, E.; Díaz, N.; Suárez, D. *J. Chem. Theory Comput.* **2009**, *5*, 1667.
- (7) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. *J. Phys. Chem. B* **2009**, *113*, 9646.
- (8) Deshmukh, M. M.; Gadre, S. R. *J. Phys. Chem. A* **2009**, *113*, 7927.
- (9) Deev, V. A.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102.
- (10) Collins, M. A.; Deev, V. A. *J. Chem. Phys.* **2006**, *125*, 104104.
- (11) Addicoat, M. A.; Collins, M. A. *J. Chem. Phys.* **2009**, *131*, 104103.
- (12) Zhang, D. W.; Xiang, Y.; Zhang, J. Z. H. *J. Phys. Chem. B* **2003**, *107*, 12039.
- (13) Babu, K.; Gadre, S. R. *J. Comput. Chem.* **2003**, *24*, 484.
- (14) Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657.
- (15) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346.
- (16) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701.
- (17) Fedorov, D. G., Kitaura, K., Eds.; *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*; CRC Press: Boca Raton, FL, 2009; pp 1–36.
- (18) Ganesh, V.; Dongare, R. K.; Balanarayanan, P.; Gadre, S. R. *J. Chem. Phys.* **2006**, *125*, 104109.
- (19) Jensen, J. H.; Day, P. N.; Gordon, M. S.; Basch, H.; Cohen, D.; Garmer, D. R.; Kraus, M.; Stevens, W. J. Modeling the Hydrogen Bond. *ACS Symp. Ser.* **1994**, *569*, 139.
- (20) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968.
- (21) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.
- (22) Petersen, C. P.; Gordon, M. S. *J. Phys. Chem. A* **1999**, *103*, 4162.
- (23) Balawender, R.; Safi, B.; Geerlings, P. *J. Phys. Chem. A* **2001**, *105*, 6703.
- (24) Merrill, G. N.; Webb, S. P. *J. Phys. Chem. A* **2003**, *107*, 7852.
- (25) Merrill, G. N.; Webb, S. P.; Bivin, D. B. *J. Phys. Chem. A* **2003**, *107*, 386.
- (26) Yoshikawa, A.; Morales, J. A. *J. Mol. Struct. (Theochem)* **2004**, *681*, 27.
- (27) Merrill, G. N.; Webb, S. P. *J. Phys. Chem. A* **2004**, *108*, 833.
- (28) Kemp, D. D.; Gordon, M. S. *J. Phys. Chem. A* **2005**, *109*, 7688.
- (29) Chandrakumar, K. R. S.; Ghanty, T. K.; Ghosh, S. K.; Mukherjee, T. *J. Mol. Struct. (Theochem)* **2007**, *807*, 93.
- (30) Kemp, D. A.; Gordon, M. S. *J. Phys. Chem. A* **2008**, *112*, 4885.
- (31) Ghosh, D.; Kosenkov, D.; Vanovschi, V.; Williams, C. F.; Herbert, J. M.; Gordon, M. S.; Schmidt, M. W.; Slipchenko, L. V.; Krylov, A. I. *J. Phys. Chem. A* **2010**, *114*, 12739.
- (32) Slipchenko, L. V.; Gordon, M. S. *J. Comput. Chem.* **2007**, *28*, 276.

- (33) Gordon, M. S.; Slipchenko, L. V.; Li, H.; Jensen, J. H. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177.
- (34) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. *J. Phys. Chem. B* **2009**, *113*, 9646.
- (35) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. J. *A. J. Comput. Chem.* **1993**, *14*, 1347.
- (36) Gordon, M. S.; Schmidt, M. W. Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry*; Dykstra, C. E., Frenking, G., Kim, K.S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, The Netherlands, 2005; pp 1167–1189.
- (37) Steinmann, C.; Fedorov, D. G.; Jensen, J. H. *J. Phys. Chem. A* **2010**, *114*, 8705.
- (38) Steinmann, C.; Fedorov, D. G.; Jensen, J. H. *PLoS One* **2012**, *7*, e41117.
- (39) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.
- (40) Chen, W.; Gordon, M. S. *J. Chem. Phys.* **1996**, *100*, 14316.
- (41) Fedorov, D. G.; Kitaura, K. *Chem. Phys. Lett.* **2006**, *433*, 182.
- (42) Fedorov, D. G.; Olson, R. M.; Kitaura, K.; Gordon, M. S.; Koseki, S. *J. Comput. Chem.* **2004**, *25*, 872.
- (43) Fletcher, G. D.; Fedorov, D. G.; Pruitt, S. R.; Windus, T. L.; Gordon, M. S. *J. Chem. Theory Comput.* **2012**, *8*, 75.
- (44) Choi, C. H.; Fedorov, D. G. *Chem. Phys. Lett.* **2012**, *543*, 159.
- (45) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. *Chem. Phys. Lett.* **2002**, *351*, 475.
- (46) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.
- (47) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U.K., 1996.
- (48) Adamovic, I.; Gordon, M. S. *Mol. Phys.* **2005**, *103*, 379.
- (49) Jensen, J. H.; Gordon, M. S. *Mol. Phys.* **1996**, *89*, 1313.
- (50) Jensen, J. H.; Gordon, M. S. *J. Chem. Phys.* **1998**, *108*, 4772.
- (51) Jensen, J. H. *J. Chem. Phys.* **2001**, *114*, 8775.
- (52) Li, H.; Gordon, M. S.; Jensen, J. H. *J. Chem. Phys.* **2006**, *124*, 214108.
- (53) Xu, P.; Gordon, M. S. in preparation
- (54) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300.
- (55) Slipchenko, L. V.; Gordon, M. S. *J. Phys. Chem. A* **2009**, *113*, 2092.
- (56) Slipchenko, L. V.; Gordon, M. S. *Mol. Phys.* **2009**, *107*, 999.
- (57) Kairys, V.; Jensen, J. H. *Chem. Phys. Lett.* **1999**, *315*, 140.
- (58) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726.
- (59) Minikis, R. M.; Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2001**, *105*, 3829.
- (60) Day, P. N.; Pacher, R.; Gordon, M. S.; Merrill, G. N. *J. Chem. Phys.* **2000**, *112*, 2063.
- (61) Mullin, J. M.; Gordon, M. S. *J. Phys. Chem. B* **2009**, *113*, 8657.
- (62) Mullin, J. M.; Gordon, M. S. *J. Phys. Chem. B* **2009**, *113*, 14413.
- (63) Metropolis, N.; Reosenbluth, A.; Teller, A. *J. Chem. Phys.* **1953**, *21*, 1087.
- (64) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- (65) Pruitt, S. R.; Addicoat, M. A.; Collins, M. A.; Gordon, M. S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7752.
- (66) Nagata, T.; Brorsen, K. R.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Phys.* **2011**, *134*, 124115.
- (67) Brorsen, K. R.; Minezawa, N.; Xu, F.; Windus, T. L.; Gordon, M. S. *J. Chem. Theor. Comput.* **2012**, *8*, 5008.