

## Quantum Mechanical Fragment Methods Based on Partitioning Atoms or Partitioning Coordinates

Bo Wang, Ke R. Yang, Xuefei Xu, Miho Isegawa, Hannah R. Leverentz, and Donald G. Truhlar\*

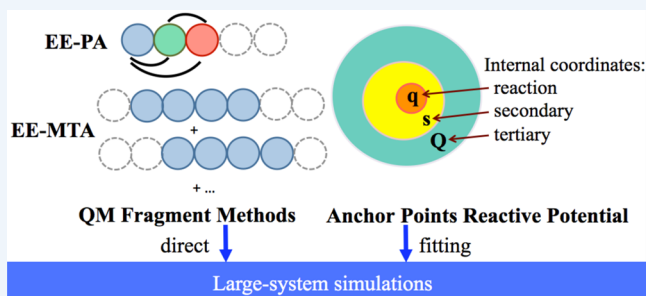
Department of Chemistry, Chemical Theory Center, and Supercomputing Institute, University of Minnesota, 207 Pleasant St. SE, Minneapolis, Minnesota 55455-0431, United States

**CONSPECTUS:** The development of more efficient and more accurate ways to represent reactive potential energy surfaces is a requirement for extending the simulation of large systems to more complex systems, longer-time dynamical processes, and more complete statistical mechanical sampling. One way to treat large systems is by direct dynamics fragment methods. Another way is by fitting system-specific analytic potential energy functions with methods adapted to large systems. Here we consider both approaches.

First we consider three fragment methods that allow a given monomer to appear in more than one fragment. The first two approaches are the electrostatically embedded many-body (EE-MB) expansion and the electrostatically embedded many-body expansion of the correlation energy (EE-MB-CE), which we have shown to yield quite accurate results even when one restricts the calculations to include only electrostatically embedded dimers. The third fragment method is the electrostatically embedded molecular tailoring approach (EE-MTA), which is more flexible than EE-MB and EE-MB-CE. We show that electrostatic embedding greatly improves the accuracy of these approaches compared with the original unembedded approaches.

Quantum mechanical fragment methods share with combined quantum mechanical/molecular mechanical (QM/MM) methods the need to treat a quantum mechanical fragment in the presence of the rest of the system, which is especially challenging for those parts of the rest of the system that are close to the boundary of the quantum mechanical fragment. This is a delicate matter even for fragments that are not covalently bonded to the rest of the system, but it becomes even more difficult when the boundary of the quantum mechanical fragment cuts a bond. We have developed a suite of methods for more realistically treating interactions across such boundaries. These methods include redistributing and balancing the external partial atomic charges and the use of tuned fluorine atoms for capping dangling bonds, and we have shown that they can greatly improve the accuracy. Finally we present a new approach that goes beyond QM/MM by combining the convenience of molecular mechanics with the accuracy of fitting a potential function to electronic structure calculations on a specific system. To make the latter practical for systems with a large number of degrees of freedom, we developed a method to interpolate between local internal-coordinate fits to the potential energy. A key issue for the application to large systems is that rather than assigning the atoms or monomers to fragments, we assign the internal coordinates to reaction, secondary, and tertiary sets. Thus, we make a partition in coordinate space rather than atom space. Fits to the local dependence of the potential energy on tertiary coordinates are arrayed along a preselected reaction coordinate at a sequence of geometries called anchor points; the potential energy function is called an anchor points reactive potential.

Electrostatically embedded fragment methods and the anchor points reactive potential, because they are based on treating an entire system by quantum mechanical electronic structure methods but are affordable for large and complex systems, have the potential to open new areas for accurate simulations where combined QM/MM methods are inadequate.



### 1. INTRODUCTION

A dynamics calculation requires a potential energy surface (PES) to govern the equations of motion. The goal of much current research is the calculation of the PES of a large system with the electronic structure of the entire system treated quantum mechanically at a reliable level. This usually involves fragmenting the system at one or more points in the calculation. There are many ways to do this, and there are even many ways of classifying the various approaches. We are especially interested in methods that yield a global or semiglobal PES, by which we mean that they can treat bond breaking or bond rearrangement rather than just the potential

for small-amplitude vibrations around an equilibrium structure or a single transition structure; it is especially important that they be able to treat chemical reactions. In this Account, we discuss three approaches to developing PESs for large systems that may be either reactive or unreactive. The first two, the electrostatically embedded many-body (EE-MB) approximation

**Special Issue:** Beyond QM/MM: Fragment Quantum Mechanical Methods

**Received:** February 17, 2014

**Published:** May 19, 2014

and the electrostatically embedded molecular tailoring approach (EE-MTA), yield implicit PESs suitable for direct dynamics, in which the potential energy, gradient, or Hessian is calculated by electronic structure methods each time the dynamics or sampling algorithm calls for it.<sup>1,2</sup> The third method, the anchor points reactive potential (APRP), is a systematic method for semiglobal fitting of an analytic function to the PES.

## 2. OVERLAPPING QUANTUM MECHANICAL FRAGMENT METHODS

The many-body (MB) expansion<sup>3</sup> provides the foundation for many fragment methods. It expresses the potential energy ( $V$ ) of a system as a finite sum of fragment potential energy terms, which can be calculated at any level of electronic structure theory or even at a combination of levels. The first term of the series involves only monomer energies ( $E_i$ ), the second term depends on both monomer and dimer ( $E_{ij}$ ) energies, the third term depends on monomer, dimer, and trimer ( $E_{ijk}$ ) energies, etc. For a system containing  $N$  monomers, the potential energy is

$$V = V_1 + V_2 + V_3 + \dots + V_N \quad (1)$$

where

$$V_1 = \sum_{i=1}^N E_i \quad (2)$$

$$V_2 = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \Delta E_{ij} \quad (3)$$

$$\Delta E_{ij} \equiv E_{ij} - E_i - E_j \quad (4)$$

$$V_3 = \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^N \Delta E_{ijk} \quad (5)$$

$$\Delta E_{ijk} \equiv E_{ijk} - \Delta E_{ij} - \Delta E_{jk} - \Delta E_{ik} - E_i - E_j - E_k \quad (6)$$

and so on. By truncating the MB expansion before the final term, one obtains an approximation of the total potential energy of the entire system. In terms of computational efficiency, it is undesirable to include  $V_4$  (or higher) because the number of tetramers in a system scales as  $N^4$ , which yields a scaling in computational cost that matches that of a conventional Hartree–Fock calculation or hybrid Kohn–Sham calculation on the entire system without fragmentation. Truncating the MB expansion after the second term yields the pairwise additive (PA) approximation, and truncating after the third term yields the three-body (3B) approximation.

The method described above is the straight MB method. Notice that the fragments (dimers, trimers, ...) overlap; that is, a given monomer belongs to more than one fragment. Because of this, all possible intermonomer interactions can be calculated at the desired level of electronic structure theory even at the pairwise additive level. In contrast, many fragment methods use disjointed fragments such that the interactions between monomers in different fragments are treated at a lower level than intrafragment interactions. Herbert et al.<sup>4</sup> have published an expression for a generalized many-body expansion that can be used to describe nearly all types of fragment methods.

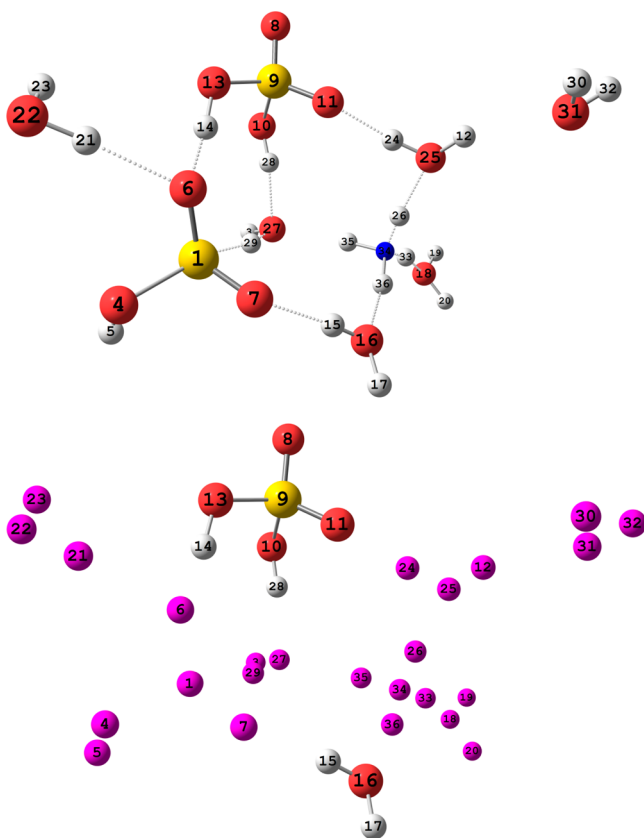
The MB approximation described in eqs 1–6 has been applied to a variety of water clusters.<sup>5–14</sup> The general consensus

emerging from these calculations is that in order to achieve quantitative accuracy, one must include cooperative effects: that is, one must use the 3B approximation or higher. It is desirable to find ways to include higher-order effects in the lower terms of the MB expansion of the PES that might allow greater accuracy at the PA level. Some approaches that our group has developed are the EE-MB method<sup>9</sup> and the many-body expansion of the correlation energy (MB-CE) method, which can also include electrostatic embedding, in which case it is called the EE-MB-CE method.<sup>10</sup>

The monomers not in the fragment currently under consideration (the primary fragment) are called the external monomers or the external subsystem. The EE-PA method approximates effects beyond  $V_2$  in a simple way: partial atomic charges obtained from isolated monomer electron densities are placed at the coordinates of the external nuclei and are included in the fragment energy calculations. In every case where we have compared PA and EE-PA calculations, we have found that including electrostatic embedding significantly improves the accuracy of the PA approximation. Over a series of water clusters ranging in size from 5 to 20 molecules, the mean unsigned deviation (MUD) from unfragmented calculations of binding energies at the same level of theory (MP2) is reduced from 1.31 to 0.07 kcal per mole of water molecules when electrostatic embedding is added to the PA approximation.<sup>10</sup> Similarly, over a series of five water 16-mers, the MUD from conventionally calculated binding energies at the same level of theory is reduced from 6.09 to 0.60 kcal per mole of water molecules.<sup>11</sup> Electrostatic embedding usually also improves the 3B approximation. One can also use “screened” or “smeared” charges rather than point charges to electrostatically embed the various fragments,<sup>12</sup> as discussed in section 3.

The EE-PA and EE-3B approximations have been applied successfully to clusters containing water, ammonia, and sulfuric acid and their proton transferred forms (as illustrated in Figure 1).<sup>15,16</sup> Most recently, we applied it to eight  $(\text{H}_2\text{O})_{26}$  clusters pulled out of a bulk water simulation; we presented CCSD(T)/CBS calculations as benchmarks.<sup>14</sup> By using the M11 exchange–correlation functional for the embedded dimers in the EE-PA approximation, we found mean unsigned errors of only 0.029 and 0.042 kcal per mole of water monomers for relative and absolute binding energies, respectively.<sup>14</sup> We have also applied the EE-3B method to a 64-mer.<sup>1</sup>

The MB-CE and EE-MB-CE approximations improve the accuracy of the PA approximation by including higher-order effects (such as cooperative induction) calculated at a less computationally costly level of theory. In these methods, an unfragmented PES calculation is performed on the entire system at the lower level of theory, and the many-body method is used to approximate the *difference* between the lower-level PES and what would be calculated at a higher level. In order to approximate this difference, the potential energy of each fragment must be calculated at both the lower and higher levels of theory. One can make the “EE-” version of the MB-CE approximation by including electrostatic embedding in the fragment calculations at both the lower and higher levels of theory. For the “-CE” approximations, the higher level of theory can be a post-Hartree–Fock correlated level of wave function theory, and the lower level of theory must be Hartree–Fock. We have tested the MB-CE and EE-MB-CE approximations on a wide variety of water clusters<sup>10,11,13</sup> and on metal–ligand complexes.<sup>17</sup> In all cases, both the MB-CE and EE-MB-CE



**Figure 1.** (top) Cluster containing one sulfuric acid molecule, one bisulfate ion (in which one oxygen atom is hidden), one ammonium ion, and six waters. (bottom) Seven of the nine fragments have been replaced by point charges, leaving one electrostatically embedded dimer.

approximations improve upon the already impressive accuracy of the corresponding EE-MB approximations.

Charge transfer between fragments is missing in methods with nonoverlapping fragments, but we have shown<sup>18</sup> that it can be treated reasonably accurately by EE-MB.

The molecular tailoring approach (MTA),<sup>19,20</sup> the general energy-based fragmentation approach,<sup>21</sup> and the molecules-in-molecules method<sup>22</sup> also involve overlapping fragments but have greater flexibility in the fragmentation scheme than the MB expansion. Rahalkar et al.<sup>23</sup> compared the MTA and fragment molecular orbital (FMO) method<sup>24</sup> for various sizes of water clusters, polypeptides, and a small protein. Their comparison shows that for most of the test systems MTA is more accurate than FMO2 (which is truncated at the two-body term) and comparable in both computational cost and accuracy to FMO3 (truncated at the three-body term). Rahalkar et al.<sup>23</sup> also summarized nine differences between the MTA and FMO

methods, including that MTA allows a more general basis for fragmentation and fragment overlap and allows a more physical treatment of broken bonds. The original MTA did not include electrostatic embedding.

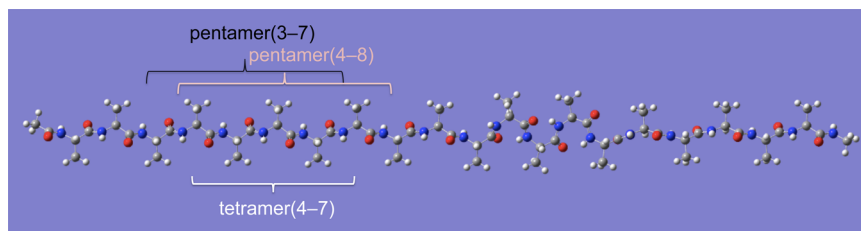
We suggested the EE-MTA, and we tested in the second order approximation.<sup>25</sup> At this order, the PES is given by

$$V = \sum_I E_I - \sum_{I,J} E_{I \cap J} - E_{OC} \quad (7)$$

where  $E_I$  is the electronic energy of fragment  $I$  that is embedded in the electrostatic field by the rest of system, and  $E_{I \cap J}$  is the energy (again electrostatically embedded) of the overlapping part of the two fragments, and  $E_{OC}$  corrects for overcounting Coulomb interactions involving the external subsystem; the correction greatly affects the accuracy, so a careful treatment<sup>21,25</sup> is required. In addition one must carefully choose the method for obtaining partial atomic charges, since the results can be sensitive to this; ideally the electrostatic embedding would be carried out without simplifying the embedding density to point charges.

For a chain polymer with 21 monomers, we may take each successive set of five monomers as a fragment. The fragments are then 17 pentamers (1–5 through 17–21), and the second term in eq 7 contains 16 tetramers (2–5 through 17–20) whose energies are subtracted to avoid overcounting monomers. See Figure 2. Or we may take the fragments as hexamers or other oligomers. Notice that the boundaries between fragments cut covalent bonds, so we must cap the dangling bonds; this is discussed in the next subsection. The primary role of the cap atom is to allow the fragment method to mimic the electronic structure of each fragment as closely as possible to its electronic structure in the whole peptide, but cap–cap interactions can be unphysical. Such interactions become significant as the fragment size is made small. To avoid this unphysical interaction, we recommend keeping the fragments large enough that the caps on the two edges of a given fragment are far from each other, where their interaction energy is small. If one considers higher-order terms involving monomers that appear in three or more fragments, computations of smaller fragments are required to keep the cost manageable, and cap–cap interactions generate some errors, so we only consider the second-order approximation.

We tested EE-MTA for the polypeptide Ace-(Ala)<sub>20</sub>-NMe.<sup>25</sup> Table 1 gives results for the relative conformational energy of an  $\alpha$ -helix and a parallel- $\beta$  sheet, and it shows that the deviation of EE-MTA from full QM calculation is smaller than that of conventional MTA by more than a factor of 10.



**Figure 2.**  $\alpha$ -Helix structure of the peptide, showing two of the pentamers and one of the tetramers.

**Table 1. Relative Energy (kcal/mol) of the  $\alpha$  Helix and Parallel- $\beta$  Sheet of Ace-(Ala)<sub>20</sub>-NMe<sup>a</sup>**

fragments	MTA	EE-MTA
trimers	49.1	-102.5
tetramers	-67.0	-114.6
pentamers	-75.9	-127.4
hexamers	-86.2	-128.0
heptamers	-102.8	-129.1
full calculation	-131.1	-131.1

<sup>a</sup>The cap atom is a hydrogen link atom. The embedding charges are AMBER RESP charges. The charge redistribution scheme is BRC2.<sup>35</sup> The partitioning is cap-C <sub>$\alpha$</sub> ...N-cap. The QM subsystems are treated by M06-2X/6-31G\*. See ref 25 for details.

### 3. NEW METHODS FOR ELECTROSTATIC EMBEDDING AND FRAGMENT BOUNDARIES

It is crucial to develop more robust schemes for electrostatic embedding in both combined quantum mechanical and molecular mechanical (QM/MM) methods and fragment QM methods. Two critical issues addressed in this section are the treatment of the fragment boundary across a covalent bond and the treatment of the electrostatic interactions between the primary and external subsystems.

The boundary between a QM fragment and the external subsystem may cut a covalent bond; the resulting dangling valence in the QM fragment must be handled properly. The conventional approach is to saturate a cut bond by a hydrogen link atom<sup>26,27</sup> or orbitals.<sup>28–30</sup> New approaches involve using an artificial atom parametrized (“tuned”) to better represent the external subsystem,<sup>31–36</sup> and we have developed a tuning scheme based on the CMS charge model that allows more consistency without constraints on the basis set.<sup>36</sup>

The partial atomic charges near the boundary need special treatment for a boundary through a covalent bond. When the bond is cut, the sum of the partial atomic charges of the surroundings and the charge of the QM region does not necessarily equal the charge of the original uncapped fragment. Balancing the external charges so that their sum plus the charge of the capped QM region equals the charge of the original system greatly improves the results.<sup>31,34,37</sup> For combined QM/MM calculations with a hydrogen link atom on a database of 15 protonation energies, balancing the charges reduced the mean unsigned deviation (MUD) from 19 to 9 kcal/mol. Capping with a fluorine link atom instead of a hydrogen link atom further lowered this to 5 kcal/mol. After balancing, the external charges near the boundary need to be redistributed and/or smeared to avoid overpolarization of the QM region.<sup>35,38–42</sup> Redistributing the charges with the BRC2 algorithm<sup>35</sup> (in which the adjusted charge on the MM atom directly bonded to the QM region is evenly redistributed to all MM atoms that are bonded to it) further lowered the MUD on the test set of protonation energies to 6 kcal/mol with a hydrogen link atom or 4 kcal/mol with a fluorine link atom.

Using a pseudopotential to tune the fluorine atom to correctly reproduce the sum of the partial charges of the uncapped portion of the quantum mechanical fragment<sup>35,36</sup> then lowered the MUD to 2 kcal/mol.

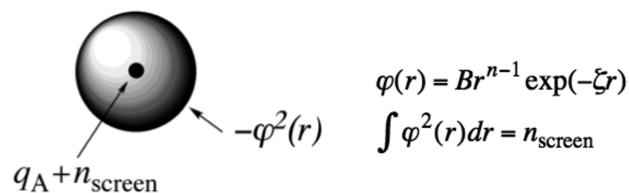
At this point, it is interesting to return to the discussion of the EE-MTA method. The error caused by cutting a covalent bond is less severe in EE-MTA since the boundary portions are canceled out in the MTA. Our comparison with a hydrogen link atom and a tuned fluorine link atom<sup>25</sup> do not yield

substantial differences in the energy deviation from the full QM calculation. This cancellation of errors from the boundary region is one of advantages of the EE-MTA formalism.

In electrostatically embedded calculations, the QM fragment is embedded in the electrostatic potential of the surroundings. In the results discussed so far, the electrostatic potential of the surroundings is represented by a set of atom-centered partial atomic charges. These point charges polarize the QM electron density. However, the electrostatic potential due to the external atoms is not accurately described by a collection of point charges.<sup>43–45</sup> One important correction is to include the charge penetration effect. As a point charge approaches a frozen atom, it sees more and more of the positive nuclear charge because the shielding of the nuclear charge by its electron density decreases. Various approaches have been proposed to include this charge penetration effect in the embedding potentials.<sup>39,42,46–48</sup> Our own solution to this problem is to treat the atoms of the secondary subsystem as charge distributions rather than point charges.<sup>48</sup> The conventional approximation for the electric potential of an atom with partial atomic charge  $q_A$  is  $q_A/r$  where  $r$  is the distance from the nucleus. However, for nonhydrogenic atoms, we put one electron in a spherical Slater-type orbital, with a charge of  $q_A + 1$  at the nucleus. Then the electric potential becomes  $q_A^*/r$  where

$$q_A^* = q_A + f(\zeta r) \exp(-2\zeta r) \quad (8)$$

where  $\zeta$  is a parameter depending on the atomic number of atom  $A$ , and  $f$  is a polynomial factor. For external hydrogen atoms, the treatment is the same except that the number  $n_{\text{screen}}$  of electrons in the Slater orbital is  $1 - q_A$  if  $q_A$  is positive. The procedure is illustrated in Figure 3.



**Figure 3.** A screened charge includes a point charge (for the nucleus and inner electron density) and a smeared charge (for the outer electron density). These replace a single partial atomic charge of  $q_A$  at the nucleus.

Since  $q_A^*/r$  is a simple radial function, this treatment of screening can be easily implemented in existing programs. The screened charge model has been tested in both the QM/MM method and the EE-MB method, and it showed good performance for both the electrostatic interactions and the polarization of the QM fragment.<sup>12,48</sup> For 414 test cases of noncovalent interactions, the MUD in the electrostatic energy (compared with symmetry-adapted perturbation theory<sup>49</sup> calculations) was reduced from 8 to 3 kcal/mol, and the MUD in the induction energy was reduced from 1.9 to 1.4 kcal/mol.<sup>48</sup> For five water hexamers, screened electrostatics reduced the MUD in binding energy from 0.22 kcal per mole of water molecules to 0.10 with EE-PA and from 0.09 to 0.04 with EE-3B.<sup>12</sup>

The above discussion is devoted to electrostatic embedding in the absence of solvent. A method to extend electrostatic embedding to reactions in the condensed phase is the use of charge response kernels.<sup>50–53</sup>

#### 4. A NEW WAY TO COMBINE QM AND MM: ANCHOR POINTS REACTIVE POTENTIAL

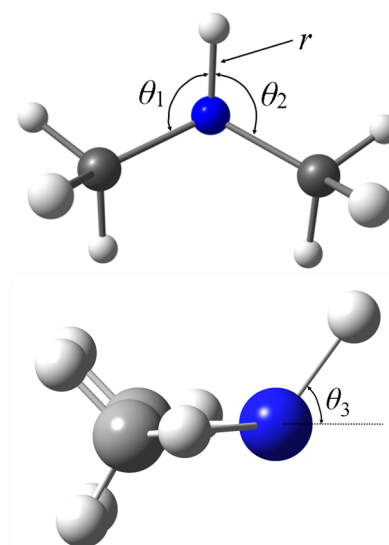
QM fragment methods are very efficient for electronic structure calculations of large systems. Much larger systems can be treated with QM/MM and QM fragment methods than with unfragmented methods. However, even the fragment methods discussed above are often too expensive for dynamics studies of moderate and large systems. For extensive sampling or long-time dynamics of large systems, we require an analytic representation of the PES. MM provides general functional forms and parameters for analytical potentials, and the same functional forms can be fit to electronic structure calculations to generate more accurate system-specific parameters,<sup>54,55</sup> but only for nonreactive systems; reactive analytic potentials require more general functional forms with parameters specific to a given system or delimited range of systems.<sup>56–60</sup> Recently developed methods based on permutation-invariant polynomials (PIP)<sup>61–63</sup> and interpolating moving least-squares (IMLS)<sup>64–66</sup> can describe bond dissociation processes and chemical reactions very well; but they are limited to very small systems, usually five or fewer atoms. Here we describe the new APRP method applicable for dynamics of larger systems.<sup>67</sup> In this method, degrees of freedom that execute large-amplitude motions, including bond breaking, are treated with accurate model potentials or general functional forms, and degrees of freedom that execute only small-amplitude vibrations are treated with molecular mechanics terms that are both system-specific and dependent on one or more reaction coordinates. This provides a general method of representing reactive PESs that combines the accuracy of QM with the ease of evaluation of MM as well as analytic gradients. Because only small-amplitude vibrations are treated for some degrees of freedom (which will be called the tertiary coordinates), this method yields only a semiglobal PES, but that is adequate for most chemical reactions.

The APRP method uses internal coordinates, and the internal coordinates are divided into three predefined groups: (1) reaction coordinates,  $\mathbf{q}$ ; (2) secondary coordinates,  $\mathbf{s}$ ; (3) tertiary coordinates,  $\mathbf{Q}$ . Then the PES is partitioned as

$$V = V^{[1]}(\mathbf{q}) + V^{[2]}(\mathbf{s}|\mathbf{q}) + V^{[3]}(\mathbf{Q}|\mathbf{q}) \quad (9)$$

where  $f(x|y)$  denotes a function with a dependence on  $x$  and parametric dependence on  $y$ . The three terms of eq 9 are called the primary, secondary, and tertiary potentials; the primary and secondary terms can be completely general. We emphasize the difference in philosophy from other fragment methods; here the primary and secondary subsystems are defined as subsets of internal coordinates rather than as subsets of atoms. As an example, Figure 4 shows coordinates for dimethyl amine.<sup>67</sup> The primary coordinate is the bond stretching coordinate  $r$ , and there are two secondary coordinates, the difference of two C–N–H bend angles,  $\theta_1 - \theta_2$ , and the out-of-plane H–N–C–C bend angle,  $\theta_3$ ; the other internal coordinates make up the tertiary coordinates.

Accurate quantum calculations are performed in the space of  $\mathbf{q}$  and  $\mathbf{s}$  for a set of fixed values of  $\mathbf{Q}$  corresponding to a reference geometry. General functional forms are used to fit  $V^{[1]}(\mathbf{q})$  and  $V^{[2]}(\mathbf{s}|\mathbf{q})$ . The tertiary potential  $V^{[3]}(\mathbf{Q}|\mathbf{q})$  is constructed by using tent functions to interpolate the anchor-point-specific, system-specific MM potentials between anchor points, which are partially optimized for selected  $\mathbf{q}$  values along reactive coordinate. The MM functional forms should be



**Figure 4.** Two views of dimethyl amine. (top) The primary coordinate  $r$  and the two bond angles ( $\theta_1$  and  $\theta_2$ ) used as secondary coordinates. (bottom) The out-of-plane bending angle  $\theta_3$  used as a secondary coordinate; this is the angle between the C–N–C plane and the vector from N to H.

globally well behaved even though they are parametrized only for small-amplitude vibrations.

The APRP method can also be applied to electronically nonadiabatic photodissociation reactions, for example, by fitting the diabatic potential energy surfaces<sup>68</sup> of phenol. The size of phenol prevents the use of conventional PES fitting methods to construct full-dimensional (33-D) PESs, but constructing APRPs remains practical.

#### 5. CONCLUDING REMARKS

We have reviewed two fragment methods with overlapping fragments (in overlapping fragments, a monomer appears in more than one fragment) and also a method where the system division is defined in terms of sets of coordinates rather than collections of monomers. We have made good progress on developing these methods for PESs of large systems, but there is considerable room for further development. For example, we could develop multilevel schemes where more important or more active fragments are treated at a higher level than others, and some fragments could even be replaced with an implicit solvent model.<sup>69</sup> There is also considerable room for improvement in the embedding potentials, where exchange–repulsion effects could be included by effective potentials.<sup>70–72</sup> We anticipate that fragment methods will continue to be improved.

#### ■ AUTHOR INFORMATION

##### Corresponding Author

\*E-mail: truhlar@umn.edu.

##### Notes

The authors declare no competing financial interest.

##### Biographies

**Bo Wang** was born in Beijing, China. He obtained his B.S. degree in chemistry from Peking University. He received his Ph.D. degree in chemistry at the University of Minnesota under the guidance of

Professor Don Truhlar, where he now continues by doing postdoctoral research on material simulations.

**Ke R. Yang** was born in Chenxi, China. He earned his B.S. degree in chemistry from Tsinghua University under the supervision of Professor Jun Li. He went to the United States to pursue his Ph.D. degree in Professor Don Truhlar's group in 2008. He is interested in excited states and dynamics.

**Xuefei Xu** was born in Heilongjiang, China, obtained her Ph.D. in 2006, and became a lecturer in 2006 at Xiamen University. She was a postdoctoral associate with Yehuda Haas at the Hebrew University of Jerusalem in 2008–2010. She is currently a research associate with Don Truhlar at the University of Minnesota.

**Miho Isegawa** received her Ph.D. with Prof. Shigeki Kato in 2009 from Kyoto University in Japan. After a postdoctoral stay at University of Minnesota (USA) supervised by Professors Don Truhlar and Jiali Gao, she is working with Professor Keiji Morokuma as a research fellow at Fukui Institute for Fundamental Chemistry in Kyoto University.

**Hannah R. Leverentz** received a Ph.D. in Chemical Physics at the University of Minnesota in Minneapolis, MN, under the direction of Don Truhlar and is now an Assistant Professor of Chemistry at Tusculum College in Greeneville, TN.

**Don Truhlar** was born in Chicago and received a Ph.D. in Chemistry from Caltech. Since 1969, he has been on the faculty of the University of Minnesota where he is now Regents Professor of Chemistry, Chemical Physics, Nanoparticle Science and Engineering, and Scientific Computation.

## ACKNOWLEDGMENTS

The authors are grateful to Hai Lin, Erin Dahlke, Helena Qi, Jeremy Tempkin, Alek Marenich, Jingjing Zheng, Elbek Kurbanov, and Elizabeth Amin for contributions to our research in this area. This work was supported in part by DOE (grant DE-FG02-86ER13579).

## REFERENCES

- (1) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Simulations. *J. Chem. Theory Comput.* **2008**, *4*, 1–6.
- (2) Speetzen, E. D.; Leverentz, H. R.; Lin, H.; Truhlar, D. G. Electrostatically embedded many-body expansion for large systems. In *Accurate Condensed-Phase Quantum Chemistry*; Manby, F. R., Ed.; CRC Press: Boca Raton, FL, 2011; pp 105–127.
- (3) Elrod, M. J.; Saykally, R. J. Many-Body Effects in Intermolecular Forces. *Chem. Rev.* **1994**, *94*, 1975–1997.
- (4) Richard, R. M.; Herbert, J. M. A Generalized Many-Body Expansion and a Unified View of Fragment-Based Methods in Electronic Structure Theory. *J. Chem. Phys.* **2012**, *137*, No. 064113.
- (5) Xantheas, S. S. Ab Initio Studies of Cyclic Water Clusters (H<sub>2</sub>O)<sub>n</sub>, n = 1–6. II. Analysis of Many-Body Interactions. *J. Chem. Phys.* **1994**, *100*, 7523–7534.
- (6) Kulkarni, A. D.; Ganesh, V.; Gadre, S. R. Many Body Interaction Analysis: Algorithm Development and Application to Large Molecular Clusters. *J. Chem. Phys.* **2004**, *121*, 5043–5050.
- (7) Anderson, J. A.; Crager, J.; Federoff, L.; Tschumper, G. S. Anchoring the Potential Energy Surface of the Cyclic Water Trimer. *J. Chem. Phys.* **2004**, *121*, 11023–11029.
- (8) Dahlke, E. E.; Truhlar, D. G. Assessment of the Pairwise Additive Approximation and Evaluation of Many-Body Terms for Water Clusters. *J. Phys. Chem. B* **2006**, *110*, 10595–10601.
- (9) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Large Systems, with Applications to Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 46–53.
- (10) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Correlation Energy, with Applications to the Calculation of Accurate Second-Order Møller-Plesset Theory Energies for Large Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 1342–1348.
- (11) Qi, H. W.; Leverentz, H. R.; Truhlar, D. G. Water 16-mers and Hexamers: Assessment of the Three-Body and Electrostatically Embedded Many-Body Approximations of the Correlation Energy or the Nonlocal Energy as Ways to Include Cooperative Effects. *J. Phys. Chem. A* **2013**, *117*, 4486–4499.
- (12) Tempkin, J. O. B.; Leverentz, H. R.; Wang, B.; Truhlar, D. G. Screened Electrostatically Embedded Many-Body Method. *J. Phys. Chem. Lett.* **2011**, *2*, 2141–2144.
- (13) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. Evaluation of the Electrostatically Embedded Many-Body Expansion and the Electrostatically Embedded Many-Body Expansion of the Correlation Energy by Application to Low-Lying Water Hexamers. *J. Chem. Theory Comput.* **2008**, *4*, 33–41.
- (14) Friedrich, J.; Yu, H.; Leverentz, H. R.; Bai, P.; Siepmann, J. I.; Truhlar, D. G. Water 26-mers Drawn from Bulk Simulations: Benchmark Binding Energies for Unprecedentedly Large Water Clusters and Assessment of the Electrostatically Embedded Three-Body and Pairwise Additive Approximations. *J. Phys. Chem. Lett.* **2014**, *5*, 666–670.
- (15) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. Application of the Electrostatically Embedded Many-Body Expansion to Microsolvation of Ammonia in Water Clusters. *J. Chem. Theory Comput.* **2008**, *4*, 683–688.
- (16) Leverentz, H. R.; Truhlar, D. G. The Electrostatically Embedded Many-Body Approximation for Systems of Water, Ammonia, and Sulfuric Acid and the Dependence of Its Performance on Embedding Charges. *J. Chem. Theory Comput.* **2009**, *5*, 1573–1584.
- (17) Kurbanov, E. K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. Analysis of the Errors in the Electrostatically Embedded Many-Body Expansion of the Energy and the Correlation Energy for Zn and Cd Coordination Complexes with Five and Six Ligands and Use of the Analysis to Develop a Generally Successful Fragmentation Strategy. *J. Chem. Theory Comput.* **2013**, *9*, 2617–2628.
- (18) Leverentz, H. R.; Maerzke, K. A.; Keasler, S. J.; Siepmann, J. I.; Truhlar, D. G. Electrostatically Embedded Many-Body Method for Dipole Moments, Partial Atomic Charges, and Charge Transfer. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7669–7678.
- (19) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. Molecular Tailoring Approach for Simulation of Electrostatic Properties. *J. Phys. Chem.* **1994**, *98*, 9165–9169.
- (20) Sahu, N.; Yeole, S. D.; Gadre, S. R. Appraisal of Molecular Tailoring Approach for Large Clusters. *J. Chem. Phys.* **2013**, *138*, No. 104101.
- (21) Li, W.; Li, S.; Jiang, Y. Generalized Energy-Based Fragmentation Approach for Computing the Ground-State Energies and Properties of Large Molecules. *J. Phys. Chem. A* **2007**, *111*, 2193–2199.
- (22) Mayhall, N. J.; Raghavachari, K. Molecules-in-Molecules: An Extrapolated Fragment-Based Approach for Accurate Calculations on Large Molecules and Materials. *J. Chem. Theory Comput.* **2011**, *7*, 1336–1343.
- (23) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. Molecular Tailoring Approach in Conjunction with MP2 and RI-MP2 Codes: A Comparison with Fragment Molecular Orbital Method. *J. Comput. Chem.* **2010**, *31*, 2405–2418.
- (24) Fedorov, D. G.; Kitaura, K. The Importance of Three-Body Terms in the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2004**, *120*, 6832–6840.
- (25) Isegawa, M.; Wang, B.; Truhlar, D. G. Electrostatically Embedded Molecular Tailoring Approach and Validation for Peptides. *J. Chem. Theory Comput.* **2013**, *9*, 1381–1393.
- (26) Singh, U. C.; Kollman, P. A. A Combined *ab Initio* Quantum-Mechanical and Molecular Mechanical Method for Carrying Out Simulations on Complex Molecular Systems: Applications to the CH<sub>3</sub>Cl + Cl<sup>-</sup> Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comput. Chem.* **1986**, *7*, 718–730.

- (27) Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- (28) Théry, V.; Rinaldi, D.; Rivail, J. L.; Maigret, B.; Ferenczy, G. G. Quantum Mechanical Computations on Very Large Molecular-Systems: The Local Self-Consistent Field Method. *J. Comput. Chem.* **1994**, *15*, 269–282.
- (29) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.
- (30) Pu, J.; Gao, J.; Truhlar, D. G. Generalized Hybrid Orbital Method for Combining Density Functional Theory with Molecular Mechanicals. *ChemPhysChem* **2005**, *6*, 1853–1865.
- (31) Zhang, Y.; Lee, T.-S.; Yang, W. A Pseudobond Approach to Combining Quantum Mechanical and Molecular Mechanical Methods. *J. Chem. Phys.* **1999**, *110*, 46–54.
- (32) Antes, I.; Thiel, W. Adjusted Connection Atoms for Combined Quantum Mechanical and Molecular Mechanical Methods. *J. Phys. Chem. A* **1999**, *103*, 9290–9295.
- (33) Dilabio, G. A.; Hurley, M. M.; Christiansen, P. A. Simple One-Electron Quantum Capping Potentials for Use in Hybrid QM/MM Studies of Biological Molecules. *J. Chem. Phys.* **2002**, *116*, 9578–9584.
- (34) Wang, B.; Truhlar, D. G. Combined Quantum Mechanical and Molecular Mechanical Methods for Calculating Potential Energy Surfaces: Tuned and Balanced Redistributed-Charge Algorithm. *J. Chem. Theory Comput.* **2010**, *6*, 359–369.
- (35) Wang, B.; Truhlar, D. G. Geometry Optimization Using Tuned and Balanced Redistributed Charge Schemes for Combined Quantum Mechanical and Molecular Mechanical Calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10556–10564.
- (36) Wang, B.; Truhlar, D. G. Tuned and Balanced Redistributed Charge Scheme for Combined Quantum Mechanical and Molecular Mechanical (QM/MM) Methods and Fragment Methods: Tuning Based on the CMS Charge Model. *J. Chem. Theory Comput.* **2013**, *9*, 1036–1042.
- (37) Walker, R. C.; Crowley, M. F.; Case, D. A. The Implementation of a Fast and Accurate QM/MM Potential Method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019–1031.
- (38) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. Computer Simulation of Zeolite Structure and Reactivity Using Embedded Cluster Methods. *Faraday Discuss.* **1997**, *106*, 79–92.
- (39) Das, D.; Eurenus, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodošček, M.; Brooks, B. R. Optimization of Quantum Mechanical Molecular Mechanical Partitioning Schemes: Gaussian Delocalization of Molecular Mechanical Charges and the Double Link Atom Method. *J. Chem. Phys.* **2002**, *117*, 10534–10547.
- (40) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. Zeolite Structure and Reactivity by Combined Quantum-Chemical–Classical Calculations. *J. Phys. Chem. B* **1999**, *103*, 6133–6141.
- (41) Lin, H.; Truhlar, D. G. Redistributed Charge and Dipole Schemes for Combined Quantum Mechanical and Molecular Mechanical Calculations. *J. Phys. Chem. A* **2005**, *109*, 3991–4004.
- (42) Amara, P.; Field, M. J. Evaluation of an Ab Initio Quantum Mechanical/Molecular Mechanical Hybrid-Potential Link-Atom Method. *Theor. Chem. Acc.* **2003**, *109*, 43–53.
- (43) Bentley, J. Atomic Multipole Expansions of Molecular Charge Densities. Electrostatic Potentials. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981; pp 63–84.
- (44) Stone, A. J. Distributed Multipole Analysis, or How to Describe a Molecular Charge Distribution. *Chem. Phys. Lett.* **1981**, *83*, 233–239.
- (45) Gadre, S. R.; Pundlik, S. S.; Shrivastava, I. H. A “Critical” Appraisal of Electrostatic Charge Models for Molecules. *Proc.—Indian Acad. Sci. (Chem. Sci.)* **1994**, *106*, 303–314.
- (46) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An Effective Fragment Method for Modeling Solvent Effects in Quantum Mechanical Calculations. *J. Chem. Phys.* **1996**, *105*, 1968–1986.
- (47) Cisneros, G. A.; Tholander, S. N.-I.; Parisel, O.; Darden, T. A.; Elking, D.; Perera, L.; Piquemal, J. P. Simple Formulas for Improved Point-Charge Electrostatics in Classical Force Fields and Hybrid Quantum Mechanical/Molecular Mechanical Embedding. *Int. J. Quantum Chem.* **2008**, *108*, 1905–1912.
- (48) Wang, B.; Truhlar, D. G. Including Charge Penetration Effects in Molecular Modeling. *J. Chem. Theory Comput.* **2010**, *6*, 3330–3342.
- (49) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (50) Morita, A.; Kato, S. Ab Initio Molecular Orbital Theory on Intramolecular Charge Polarization: Effect of Hydrogen Abstraction on the Charge Sensitivity of Aromatic and Nonaromatic Species. *J. Am. Chem. Soc.* **1997**, *119*, 4021–4032.
- (51) Higashi, M.; Truhlar, D. G. Combined Electrostatically Embedded Multiconfiguration Molecular Mechanics and Molecular Mechanical Method: Application to Molecular Dynamics Simulation of a Chemical Reaction in Aqueous Solution with Hybrid Density Functional Theory. *J. Chem. Theory Comput.* **2008**, *4*, 1032–1039.
- (52) Higashi, M.; Truhlar, D. G. Efficient Approach to Reactive Molecular Dynamics with Accurate Forces. *J. Chem. Theory Comput.* **2009**, *5*, 2925–2929.
- (53) Tsutsumi, J.; Yoshida, H.; Murdey, R.; Kato, S.; Sato, N. An Accurate Calculation of Electronic Contribution to Static Permittivity Tensor for Organic Molecular Crystals on the Basis of the Charge Response Kernel Theory. *J. Phys. Chem. A* **2009**, *113*, 9207–9212.
- (54) Cacelli, I.; Prampolini, G. Parametrization and validation of intramolecular force fields derived from DFT calculations. *J. Chem. Theory Comput.* **2007**, *3*, 1803–1817.
- (55) Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. JOYCE and ULYSSES: Integrated and User-Friendly Tools for the Parameterization of Intramolecular Force Fields from Quantum Mechanical Data. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3736–3751.
- (56) Schatz, G. C. The Analytic Representation of Electronic Potential-Energy Surfaces. *Rev. Mod. Phys.* **1989**, *61*, 669–688.
- (57) Hollebeek, T.; Ho, T.-S.; Rabitz, H. Constructing Multidimensional Molecular Potential Energy Surface from ab Initio Data. *Annu. Rev. Phys. Chem.* **1999**, *50*, 537–570.
- (58) Fernández-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. Modeling the Kinetics of Bimolecular Reactions. *Chem. Rev.* **2006**, *106*, 4518–4584.
- (59) Albu, T. V.; Espinosa-García, J.; Truhlar, D. G. Computational Chemistry of Polyatomic Reaction Kinetics and Dynamics: The Quest for an Accurate CH<sub>3</sub> Potential Energy Surface. *Chem. Rev.* **2007**, *107*, 5101–5132.
- (60) Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A., III ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A* **2008**, *112*, 1040–1053.
- (61) Braams, B. J.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces in High Dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577–606.
- (62) Bowman, J. M.; Braams, B. J.; Carter, S.; Chen, C.; Czako, G.; Fu, B.; Huang, X.; Kamarchik, E.; Sharma, A. R.; Shepler, B. C.; Wang, Y.; Xie, Z. Ab-Initio-Based Potential Energy Surfaces for Complex Molecules and Molecular Complexes. *J. Phys. Chem. Lett.* **2010**, *1*, 1866–1874.
- (63) Paukku, Y.; Yang, K. R.; Varga, Z.; Truhlar, D. G. Global ab Initio Ground-State Potential Energy Surface of N<sub>4</sub>. *J. Chem. Phys.* **2013**, *139*, No. 044309.
- (64) Dawes, R.; Thompson, D. L.; Guo, Y.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting Potential Energy Surfaces: Computing High-Density Potential Energy Surface Data from Low-Density ab Initio Data Points. *J. Chem. Phys.* **2007**, *126*, No. 184108.
- (65) Guo, Y.; Tokmakov, I.; Thompson, D. L.; Wagner, A. F.; Minkoff, M. Interpolating Moving Least-Squares Methods for Fitting

Potential Energy Surfaces: Improving Efficiency via Local Approximants. *J. Chem. Phys.* **2007**, *127*, No. 214106.

(66) Bender, J. D.; Doraiswamy, S.; Truhlar, D. G.; Candler, G. V. Potential Energy Surface Fitting by a Statistically Localized, Permutationally Invariant, Local Interpolating Moving Least Squares Method for the Many-Body Potential: Method and Application to  $N_4$ . *J. Chem. Phys.* **2014**, *140*, No. 054302.

(67) Yang, K. R.; Xu, X.; Truhlar, D. G. Anchor Points Reactive Potential for Bond-Breaking Reactions. *J. Chem. Theory Comput.* **2014**, *10*, 924–933.

(68) Xu, X.; Yang, K. R.; Truhlar, D. G. Diabatic Molecular Orbitals, Potential Energies, And Potential Energy Surface Couplings by the 4-Fold Way for Photodissociation of Phenol. *J. Chem. Theory Comput.* **2013**, *9*, 3612–3625.

(69) Marenich, A.; Cramer, C. J.; Truhlar, D. G. Generalized Born Solvation Model SM12. *J. Chem. Theory Comput.* **2013**, *9*, 609–620.

(70) Yudanov, I. V.; Nasluzov, V. A.; Neyman, K. M.; Rösch, N. Density Functional Cluster Description of Ionic Materials: Improved Boundary Conditions for MgO Clusters with the Help of Cation Model Potentials. *Int. J. Quantum Chem.* **1997**, *65*, 975–986.

(71) Nygren, M. A.; Pettersson, L. G. M.; Barandiarán, Z.; Seijo, L. Bonding between CO and the MgO(001) Surface: A Modified Picture. *J. Chem. Phys.* **1994**, *100*, 2010–2018.

(72) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.