JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Multi-state Approach to Chemical Reactivity in Fragment Based Quantum Chemistry Calculations

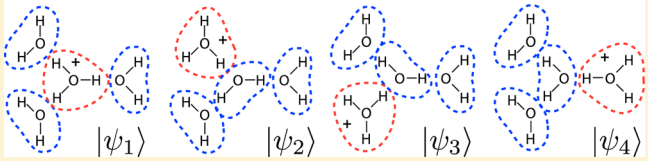
Adrian W. Lange[†] and Gregory A. Voth*[,‡]

[†]Leadership Computing Facility, Argonne National Laboratory, Argonne, Illinois 60439, United States

[‡]Department of Chemistry, James Franck Institute, Institute for Biophysical Dynamics, and Computation Institute, University of Chicago, Chicago, Illinois 60637, United States

ABSTRACT: We introduce a multistate framework for Fragment Molecular Orbital (FMO) quantum mechanical calculations and implement it in the context of protonated water clusters. The purpose of the framework is to address issues of nonuniqueness and dynamic fragmentation in FMO as well as other related fragment methods. We demonstrate that our new approach, Fragment Molecular Orbital Multistate Reactive Molecular Dynamics (FMO-MS-RMD), can improve energetic accuracy and yield stable molecular dynamics for small protonated water clusters undergoing proton transfer reactions.

$$|\Psi_{\mathrm{tot}}\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle + c_3|\psi_3\rangle + c_4|\psi_4\rangle$$



1. INTRODUCTION

The computational cost of *ab initio* quantum chemistry calculations is very high, usually scaling with respect to the number of electron basis functions as a function of some exponent [e.g., second order Møller−Plesset theory, MP2, generally scales as $O(N_{\mathrm{basis}}^5)$]. This cost is a serious hindrance to the modeling of increasingly larger and complex molecular systems with quantum chemistry. In recent years, though, several new approaches have been devised to approximate the full quantum calculation of a molecular system through some combination of relatively smaller, more feasible quantum calculations on subsets of the molecular system. In particular, we are referring to fragment methods in quantum chemistry.[1]

There are many fragment methods used in practice today, including the Fragment Molecular Orbital (FMO) method,[2−6] the Electrostatically Embedded Many-Body Expansion (EE-MBE) method,[7,8] the Explicit Polarization (X-Pol) method,[9] the X-Pol/Symmetry Adapted Perturbation Theory method,[10] the Molecules-in-Molecules method,[11] and the Binary Interaction Method,[12,13] just to name a few. It is not our intention to discuss all of them here, and we instead refer the reader to recent reviews on the topic.[1,14]

In this work, we will focus almost exclusively on the FMO method as our prototypical fragment method. Fragment methods all share the common notion that the entirety of a given molecular system can, in some prescribed way, be divided up into a set of fragments. However, the fragmentation of a system is not always straightforward, such as when fragmenting across covalent bonds or when chemical reactions take place. The important consequence is that the final energy of a fragment method can be highly dependent on the molecular fragmentation procedure. A given molecular system does not always have a unique fragmentation. Moreover, if molecular dynamics (MD) is propagated, a certain procedure might cause the fragmentation to vary between adjacent time steps, an issue known as *dynamic fragmentation*.[15]

The nonuniqueness and dynamic fragmentation issues of fragment methods are curiously similar to issues addressed by a certain class of classical molecular mechanics reactive force fields, specifically, the Multistate Empirical Valence Bond (MS-EVB) force fields.[16−18] In MS-EVB, a reactive molecular system is represented by some linear combination of multiple *states*, each state having a different bonding topology. These states form a basis set, and MD can be propagated with MS-EVB by determining the basis of states on-the-fly. However, some drawbacks of past MS-EVB models—and common to most popular classical force fields—are that they can require an elaborate fitting procedure for many parameters and that they might lack certain important physical properties, such as charge polarization.

In this paper, we suggest that FMO and MS-EVB are complementary. Nonuniqueness and dynamic fragmentation in FMO can be addressed by adopting the multistate approach of MS-EVB. Correspondingly, the accuracy of MS-EVB can be improved by basing it on FMO. We dub our new hybrid approach the Fragment Molecular Orbital Multistate Reactive Molecular Dynamics (FMO-MS-RMD) method. FMO-MS-RMD constitutes a general framework in the sense that the underlying multistate idea can be extended to any fragment method applying any level of electronic structure theory. It can also be adapted to many (though not necessarily all) chemically reactive systems, not just the example of proton transfer in this work. Compared to existing methods, FMO-MS-RMD is perhaps most closely related to the Effective Hamiltonian Molecular Orbital-Valence Bond (EH-MOVB) model of Gao

Received: June 17, 2013
Published: August 14, 2013

and co-workers,[19] but the two differ in a number of significant ways.

As an initial implementation of our concept, we introduce FMO-MS-RMD in the context of protonated water clusters treated with FMO at the MP2/cc-pVDZ level. Our formulation of FMO-MS-RMD embraces some degree of empiricism motivated by our experience with previous reactive MD multistate models,[16−18] yet it still retains much of the appealing *ab initio* flavor of FMO. After describing the framework, we present a series of results to exhibit the improvements made by FMO-MS-RMD. A discussion on how FMO-MS-RMD might be applied to other chemically reactive systems is presented in the Appendix.

2. THEORY

In FMO, the total system energy, E^{FMO}, is expressed via the many-body expansion,[1]

$$E^{\mathrm{FMO}} = \sum_{I}^{N_{\mathrm{frag}}} E_I + \sum_{I<J}^{N_{\mathrm{frag}}} (E_{IJ} - E_I - E_J) \tag{1}$$

For our purposes, we truncate the many-body expansion at two-body order (i.e., FMO2), although higher order terms could be incorporated to improve the energy accuracy.[6] The monomer energy, E_I, is the quantum mechanical (QM) energy of the I-th fragment embedded in the electrostatic field of all other fragments. Similarly, E_{IJ} is the QM energy of the dimer of fragments I and J embedded in the electrostatic field of all other fragments. These fragment QM energies could be computed at any level of electronic structure theory (e.g., Hartree−Fock, Density Functional Theory, Møller−Plesset theory, *etc.*). Traditionally, FMO computes the monomer energies and electrostatic field self-consistently until convergence.[3,5] However, for simplicity in the introduction of FMO-MS-RMD, the electrostatic field in our current formulation is represented by a set of fixed (i.e., static) point charges centered at each atom, which greatly simplifies our computational algorithm, our fitting procedure, and the computation of analytic gradients. That is, E^{FMO} in this work is a "frozen charge" FMO energy and is technically equivalent to an EE-MBE approach of Dahlke and Truhlar.[7] Nonetheless, we impose no particular restriction to how the electrostatic field is to be represented in the general framework of FMO-MS-RMD, and our choice to name our method with FMO is intended to reflect that point since FMO theory has been developed for a variety of embedding electrostatic fields. We acknowledge that the use of fixed point charges in our current formulation of FMO-MS-RMD as applied to protonated water is not the most sophisticated way to incorporate environmental electrostatics, perhaps not fully capturing certain electrostatic effects such as induction, but we anticipate that future renditions of FMO-MS-RMD will indeed include improved electrostatics, such as self-consistently determined Mulliken charges.

Note that E^{FMO} depends on the choice of fragmentation. Certain fragmentations are a better approximation to the true QM energy of the full molecular system, E^{QM}. However, arriving at the *best* fragmentation is not always immediately clear (see Figure 1).

FMO-MS-RMD aims to overcome fragmentation nonuniqueness by taking into consideration multiple fragmentations at once. The multistate *ansatz* of FMO-MS-RMD is that the total molecular system is a linear combination of different fragmentations (i.e., "states"), or different ways to arrange the chemical bonding topology, in other words. It is convenient to express this in the familiar bra−ket notation,

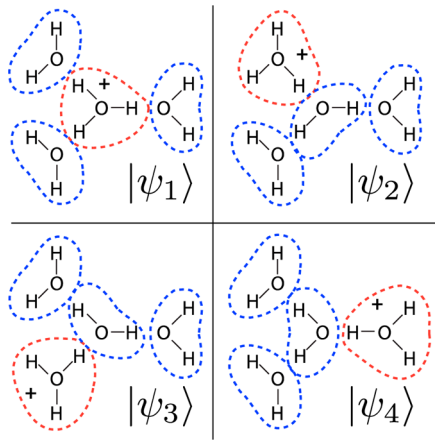


Figure 1. Schematic of four possible fragmentation states in a protonated tetramer of water. Dashed blue lines indicate neutral water fragments, and dashed red lines indicate positively charged hydronium fragments.

$$|\Psi_{\mathrm{tot}}\rangle = \sum_{A}^{N_{\mathrm{states}}} c_A |\Psi_A\rangle \tag{2}$$

where $|\Psi_A\rangle$ represents the A-th fragmentation state with normalized coefficient c_A. The FMO-MS-RMD total system state, $|\Psi_{\mathrm{tot}}\rangle$, is an approximation (in the basis of fragmentation states) to the full, nonfragmented molecular system corresponding to energy E^{QM}. The linear combination of states in eq 2 leads to a model Hamiltonian matrix with elements H_{AB}. For illustration, a schematic of possible fragmentation states for a protonated water tetramer is shown in Figure 1.

We then make the *ansatz* that each diagonal element is given by the FMO energy of the A-th fragmentation state, E_A^{FMO}, but with the addition of two empirical penalty terms,

$$H_{AA} = E_A^{\mathrm{FMO}} + E_A^{\mathrm{intra}} + E_A^{\mathrm{inter}} \tag{3}$$

Our motivation for the penalty terms is derived purely from experimentation. Based on our experience, relying on E_A^{FMO} alone is not sufficient to filter out "bad" fragmentation states with high energies that are very unlikely or are poor estimates of E^{QM}. We find that the penalty terms are necessary to disfavor such states. For example, a poor fragmentation might place an atom represented as an external charge too close to a QM fragment and strongly overpolarize the electron density of that fragment, rendering it much lower in energy than E^{QM} and causing a failure of the model. The empirical penalties are intended to compensate in this scenario by raising the energy of H_{AA}. For "good" fragmentations, though, the penalties are zero or relatively small compared to the FMO energy. In the implementation of FMO-MS-RMD presented here, we develop empirical formulas for these penalties specific to protonated water. Nonetheless, we believe that one can develop similar empirical formulas for other chemically reactive systems, yet it is quite possible that there may be some more *ab initio* approach available that we have not considered here.

The first term in eq 3, E_A^{intra}, is an intrafragment penalty based on *a priori* defined covalent bonds within each fragment. For water and hydronium fragments as in Figure 1, only O−H bonds are present, and we set E_A^{intra} to

dx.doi.org/10.1021/ct400516x | J. Chem. Theory Comput. 2013, 9, 4018−4025

$$E_A^{\text{intra}} = \sum_{I \in A}^{N_{\text{frag}}} \sum_{\text{bonds} \in I} k_{\text{intra}}[1 - f(r_{\text{OH}}^I, r_{\text{b}}, r_{\text{cut}})]^2 \quad (4)$$

where r_{OH}^I is an O−H bond distance for the I-th fragment (water or hydronium). We employ a well-known sigmoidal switching function,

$$f(r, r_{\text{in}}, r_{\text{out}}) = \begin{cases} 1 & \text{for } r \le r_{\text{in}} \\ (r_{\text{out}}^2 - r^2)^2(r_{\text{out}}^2 + 2r^2 - 3r_{\text{in}}^2)(r_{\text{out}}^2 - r_{\text{in}}^2)^{-3} & \text{for } r_{\text{in}} < r < r_{\text{out}} \\ 0 & \text{for } r \ge r_{\text{out}} \end{cases} \quad (5)$$

to establish a distance-based switching region for the intrafragment penalty. (We also use eq 5 in the interfragment penalty described below.) The switching region for eq 4 is between the bond distance parameter, r_{b}, and the cutoff distance, r_{cut}. By squaring the part of eq 6 in the square brackets, we guarantee a smooth and continuous derivative at the end points of the switching region. The parameters k_{intra}, r_{b}, and r_{cut} are left to be determined (see below). The penalty E_A^{intra} ensures that fragmentation states with unusually elongated bonds are disfavored and can be safely neglected in a state search (see Section 4).

The second penalty term in eq 3, E_A^{inter}, is an interfragment penalty based on the proximity of atoms in a donor reactive fragment (i.e., H_3O^+) to the atoms of a reactive acceptor fragment (i.e., H_2O). We set E_A^{inter} to

$$E_A^{\text{inter}} = \sum_{i \in H_2O}^{N_{\text{O}}} \sum_{j \in H_3O^+}^{N_{\text{H}}} k_{\text{inter}} f(r_{ij}, 0, r_{\text{cut}})^2 \quad (6)$$

where r_{ij} is the distance from O atom i in an acceptor water fragment to H atom j in the donor hydronium fragment. The switching function is the same as in eq 5, and it is also squared to guarantee a smooth and continuous derivative at r_{cut}. The parameters k_{inter} and r_{cut} are left to be determined (see below). The penalty E_A^{inter} disfavors fragmentations where a donor hydronium H atom is closer to an acceptor water fragment O atom than to the donor hydronium O atom, in which case a different fragmentation of the molecular system could be a better approximation to E^{QM}. Note that the interfragment penalty is zero beyond r_{cut}.

The off-diagonal coupling elements, H_{AB}, in the Hamiltonian matrix are representative of a chemical reaction occurring (a proton transfer in the current formulation) between states A and B. The formula for H_{AB} is empirical in this work, although a more *ab initio* based expression similar to that of the EH-MOVB model[19] could work as well. For protonated water, the H_{AB} elements are nonzero only for those pairs of states wherein the hydronium fragment in state A and the hydronium fragment in state B both share a single hydrogen atom. That is, the j-th hydrogen atom must belong to the hydronium fragment in state A as well as to the hydronium fragment in state B. This represents a proton transfer from state A to state B or vice versa. Also, the coupling is zero if the hydronium fragment in A and B share the same oxygen atom, which would be representative of a nonsensical self-reaction. We relate the empirical expression for coupling to the interfragment penalty in eq 6. It is simply the geometric mean of the contribution to E_A^{inter} from the H atom traveling between states A and B,

$$H_{AB} = \sqrt{h_{AB}h_{BA}} \quad (7)$$

with

$$h_{AB} = k_{\text{inter}} f(r_{ij}^{AB}, 0, r_{\text{cut}})^2 \quad (8)$$

where r_{ij}^{AB} is the distance between the shuttling H atom i in state A and the hydronium oxygen atom j in state B. The geometric mean in eq 7 ensures that the coupling smoothly goes to zero as the fragments A and B become distant (i.e., having O−H distances beyond r_{cut}), where a proton transfer is unlikely.

Once the model Hamiltonian matrix has been constructed, it is then diagonalized, yielding a set of eigenvalue energies, E_A, and a corresponding set of eigenvector coefficients c_A from eq 2. The total FMO-MS-RMD energy is assigned to be the minimum energy eigenvalue, $E^{\text{tot}} = \min(E_A)$. The force on each atom i can then be computed via the Hellmann−Feynman theorem,

$$F_i = -\langle \psi^{\text{tot}} | \partial \mathbf{H} / \partial \mathbf{x}_i | \psi^{\text{tot}} \rangle = -\sum_{A,B} c_A c_B \frac{\partial H_{AB}}{\partial \mathbf{x}_i} \quad (9)$$

For computing $\partial H_{AB}/\partial \mathbf{x}_i$, force contributions from the empirical penalty and couplings are straightforwardly derived. Then, with fixed embedding charges, as we use in our current model, the fully analytic FMO force for each state can be readily obtained as the linear combination of energy gradients analogous to eq 1.[8] The gradient of the FMO energy for each state is obtained as the analytic gradient for each QM fragment calculation (monomers and dimers) plus the force exerted on the surrounding embedding charges due to the QM fragment. We have verified that our implementation of the FMO-MS-RMD analytic gradient is accurate to at least 1.0×10^{-5} hartree/bohr (1.0×10^{-3} kcal/mol/Å) by comparison to numerical finite difference via the midpoint method.

3. PARAMETERIZATION

The above discussion defines the framework for FMO-MS-RMD in the context of protonated water. What remains to complete the implementation of the model is to decide upon a level of electronic structure and to obtain values for the model parameters. Currently, we choose MP2/cc-pVDZ for the QM fragment calculations in FMO, such that our model—compared with past MS-EVB models—naturally incorporates polarization, charge transfer, dispersion, and other many-body effects to the extent achieved by FMO. The full set of 6 adjustable parameters in our model is listed in Table 1. For contrast, MS-EVB3 requires 29 adjustable parameters, though many of them are fitted serially through a sequence of uncorrelated steps.[17]

Arriving at the values for the adjustable parameters in our current formulation is admittedly not completely straightforward. We have opted to take as simplistic of an approach to fitting as possible through a combination of experimentation and stochastic minimization. Once again, we use fixed point charges for the electrostatic embedding field in our FMO calculations. The charge values for water fragments are set to those of the SPC/E water model,[20] and the charge values for hydronium fragments are set to those of the MS-EVB3 model.[17] We assume, at least by precedent, that these are reasonable charges for modeling the condensed phase electrostatics of protonated water, and we anticipate that they might be transferrable to future condensed phase FMO-MS-RMD

Table 1. Parameters for the FMO-MS-RMD Protonated Water Model

parameter	value	units
$q_O^{H_2O}$	−0.8476	e
$q_H^{H_2O}$	0.4238	e
$q_O^{H_3O^+}$	−0.5	e
$q_H^{H_3O^+}$	0.5	e
k_{intra}	0.8	hartree
r_b	1.4	Å
k_{inter}	6.41×10^{-3}	hartree
r_{cut}	2.4	Å

calculations. The intra−fragment penalty parameter r_b [eq 4] was selected based on a scan of possible values in a subset of protonated water clusters from the Cambridge Cluster Database[21] (more details on this below) with the intent of penalizing fragmentations with unusually large bond lengths while not penalizing at all what were deemed the "best" fragmentations for each cluster. (Water or hydronium fragments with O−H bonds are less than r_b receive zero intrafragment penalty.) After settling on r_b, we experimented iteratively with a set of possible k_{intra} values in subsequent interfragment penalty fitting and MD stability tests. The interfragment penalty and coupling parameters [eq 6] were fit against MP2/cc-pVDZ absolute energies for two small protonated water clusters, $H_5O_2^+$ and $H_{13}O_6^+$, (dimer and hexamer, respectively) as a function of a proton transfer coordinate (Figures 2 and 3). The proton transfer data was

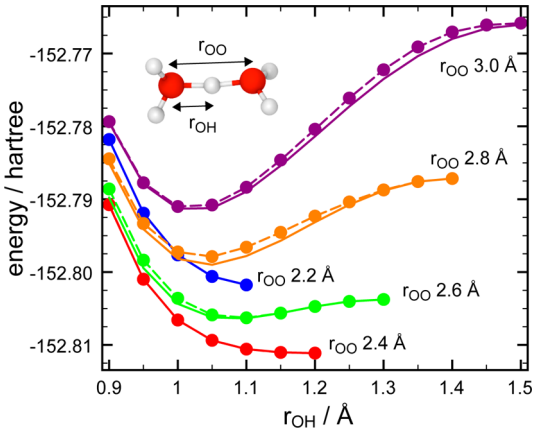


Figure 2. FMO-MS-RMD fit for $H_5O_2^+$. Solid lines correspond to the target MP2/cc-pVDZ absolute energy of the minimum energy pathway at the denoted r_{oo} constrained distance. Dashed lines with dots indicate the FMO-MS-RMD fit results. Fit data for r_{oo} = 2.2 Å and r_{oo} = 2.4 Å are essentially indistinguishable from the target energy. RMSE = 0.39 kcal/mol for this subset of data.

obtained as MP2/cc-pVDZ minimum energy pathways at constrained O−H and O−O distances of 2.2, 2.4, 2.6, 2.8, and 3.0 Å, amounting to 90 data points for fitting. Note that the proton transfer coordinates are symmetric about the maximum O−H distance for each O−O distance. To determine the best k_{inter} at a several values of r_{cut} and fixed experimental k_{intra} values, simulated annealing in the k_{inter} parameter space (at fixed r_{cut} values) was used to minimize the root-mean-square error (RMSE) between E^{QM} and E^{tot}. The combination of k_{inter} and r_{cut} for a given k_{intra} that produced the minimum error was selected. The results of this fitting are shown in Figures 2 and 3.

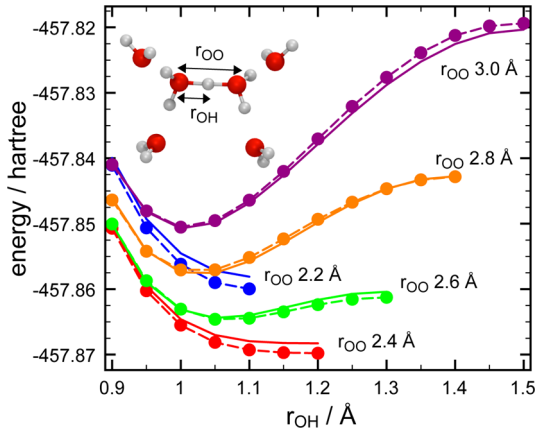


Figure 3. FMO-MS-RMD fit for $H_{13}O_6^+$. Solid lines correspond to the target MP2/cc-pVDZ absolute energy of the minimum energy pathway at the denoted r_{oo} constrained distance. Dashed lines with dots indicate the FMO-MS-RMD fit results. RMSE = 0.56 kcal/mol for this subset of data.

The final RMSE of the fit for all 90 data points is 0.48 kcal/mol. In addition to fitting, various short MD test runs were performed to verify that the set of parameters produced stable MD free of spuriously large forces as proton transfer reactions took place.

Because we have truncated our many-body expansion at two-body order, one might expect that FMO-MS-RMD is also exact for $H_5O_2^+$, having only two fragments. However, the empirical terms in FMO-MS-RMD spoil the exactness at two-body order. The RMSE of our fit for the subset of data points belonging to $H_5O_2^+$ is 0.39 kcal/mol. For comparison, the RMSE fit for the $H_{13}O_6^+$ subset of data points is 0.56 kcal/mol. Still, it is not our intention to perfectly replicate the full MP2 calculation but rather to find an acceptable approximation for propagating MD on a potential energy surface similar to MP2 achieved by FMO-MS-RMD.

4. IMPLEMENTATION

We have implemented our FMO-MS-RMD protonated water model as an external interface to a development version of the Q-Chem software package.[22] In our implementation, an FMO-MS-RMD calculation proceeds in four phases: (1) Given a molecular system geometry and an initial fragmentation guess, perform a state search (see below) to find states to form the basis, (2) compute the FMO energy (and gradient if needed) according to eq 1 for each state of the basis, performing a series of QM calculations for all fragment monomers and dimers, (3) compute the empirical terms to fill out the Hamiltonian matrix, (4) diagonalize the Hamiltonian and obtain the minimum energy eigenvalue as well as the analytic gradient if needed. MD can be propagated after phase 4 by using the gradient. In phase 2, the FMO calculations are noniterative. That is, all monomer and dimer calculations are subject only the fixed charge embedding field, which does not depend on the other FMO calculations. If FMO-MS-RMD were implemented with self-consistently determined charges, then phase 2 would involve a more complex procedure.

There is considerable parallelism available to an FMO-MS-RMD computation, and our code uses the Message Passing Interface (MPI) parallel paradigm to take advantage of it. Each FMO fragmentation state energy, E_A^{FMO}, is independent and, because we have opted to use fixed charges, so is every FMO

dx.doi.org/10.1021/ct400516x | J. Chem. Theory Comput. 2013, 9, 4018−4025

monomer and dimer fragment calculation contributing to E_A^{FMO}. Our current code runs all of these fragment calculations in parallel as separate calls to Q-Chem on each MPI rank. A comparison of timings is presented in the Section 5.

The practical determination of fragmentation states (phase 1 from above) for FMO-MS-RMD is a crucial aspect of implementation. It is can be unreasonable to include *every* possible fragmentation of a protonated water system, for which there is a roughly factorial number of combinations of oxygen and hydrogen atoms to form the waters and hydronium. (Other chemically reactive systems, though, might have few enough states to use a *complete* basis of fragmentations and not require such a state search.) Instead, we accomplish finding fragmentations using a state search algorithm similar to that of MS-EVB3.[17] The input for the search is a guess at a fragmentation of the molecular system, which we refer to as the *pivot state*. In practice, aside from an initializing guess, the pivot state we use comes from the immediately preceding step—perhaps a step from MD or geometry optimization. The pivot state is assigned as $|\Psi_A\rangle$ with the maximum square amplitude, c_A^2, from the previous step. Then, given the pivot state and system geometry, we perform a breadth-first search outwardly from the pivot state hydronium fragment. Each H atom on the hydronium donor fragment attempts to "hop" to a nearby O atom on an acceptor water fragment that is within the cutoff distance, r_{cut}. A successful hop forms a new possible fragmentation state from which another set of hops will be attempted in the next level of the search. The search terminates when all possible hops have been exhausted or when a certain threshold level of hops has been reached. In our calculations, we enforce a threshold of three hops out from the pivot hydronium. We note that our search implementation is capable of handling bifurcated hydrogen bonding water molecules, as has been shown previously to be important in MS-EVB3.[17]

In general, if the search algorithm is run to exhaustion, there will be about one fragmentation state per water in a protonated water system, wherein every water is a hydronium in some state. As the system grows, though, running the search to exhaustion is not feasible, and the hop threshold usually terminates the search. With a zero hop threshold, there is obviously only a single state, the pivot state. A one hop threshold will pick up one state for each water hydrogen bonded to the hydronium, typically three, such that the basis is formed by a total of four states. With two or more hops, it becomes less easy to estimate the number of states found in our search, as it is highly dependent on the geometry of the hydrogen bonding network. In the cluster calculations in this paper, we usually find somewhere in the range of 8−15 states with a three hop threshold.

5. RESULTS: PROTONATED WATER CLUSTERS

To gauge the accuracy of our FMO-MS-RMD model in systems different from our training set, we compare energies in a series of protonated water clusters obtained from the Cambridge Cluster Database.[21] The global minimum geometries according to the ASP potential[23] were used. The full MP2/cc-pVDZ absolute energy is compared to the FMO-MS-RMD absolute energy as well as to the single state whose FMO absolute energy was closest to the MP2 energy. Because FMO-MS-RMD is exact (i.e., the same as full MP2-ccpVDZ) for gas phase optimized geometry water and/or hydronium monomers, comparing the absolute energies of MP2, FMO, and FMO-MS-RMD is equivalent to comparing their respective binding energies, although neglecting basis set superposition error (BSSE). The energy error is computed as $\Delta E = (E - E^{QM})$, where E^{QM} is the MP2/cc-pVDZ energy of the cluster and E is either the FMO energy of the fragmentation state closest to E^{QM} or the FMO-MS-RMD energy. Results are presented in Figure 4.

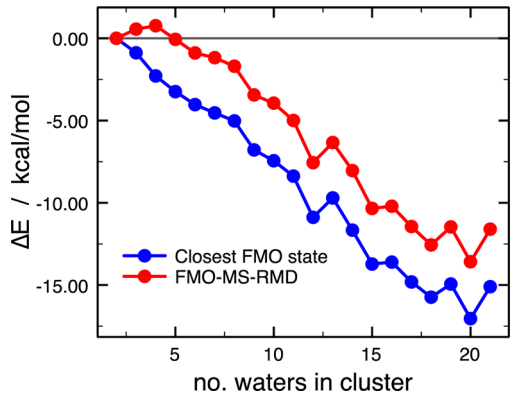


Figure 4. Absolute energy error in kcal/mol relative to MP2/cc-pVDZ of the closest FMO state and FMO-MS-RMD for the Cambridge cluster database global minimum geometries.

The FMO-MS-RMD error is consistently smaller in all cases except for the protonated water dimer. This shows that our FMO-MS-RMD model is at least more accurate (usually by about 3 kcal/mol) than the corresponding closest single state FMO energy. Nevertheless, the error grows to be quite large with increasing number of waters, becoming greater than 10 kcal/mol at 15 waters for FMO-MS-RMD. Although we do not make an in-depth investigation of it here, the increasing error in the FMO calculations is likely caused in part by BSSE, which will be quite significant with the relatively small cc-pVDZ atomic orbital basis set. Another important contribution to the error is likely from the lack of a self-consistent embedding field, which we have mentioned before. Presumably, one could then improve upon the error by employing a larger atomic orbital basis set and/or using self-consistently determined charges, which is something we expect can be explored in future renditions of FMO-MS-RMD.

In tests not shown here, we tracked the convergence of the energy error with respect to increasing the fragmentation basis as a function of changing the state search hop threshold. For these clusters, however, the error convergence is negligible beyond just one hop, where using two or three hops modifies the FMO-MS-RMD energies by less than 0.01 kcal/mol. We attribute the fast convergence to the clusters being at a minimum energy structure that seems to favor a predominantly "eigen-complex" motif (similar to what is depicted in Figure 1), which can be inferred from the FMO-MS-RMD eigenvector coefficients. Because of the eigen-complex character, hops beyond two from the pivot state are unfavorable, and the FMO-MS-RMD energy is quite well described at just one hop. However, we stress that it is crucial to MD stability to use more than one hop in the state search threshold to minimize or eliminate the impact of dynamic basis set discontinuities during proton transfers (see Section 6).

We also present timing data for the cluster series in Figure 5. FMO-MS-RMD clearly incurs more computational cost compared to regular FMO, since it involves multiple FMO calculations. The computational cost of FMO (truncated at

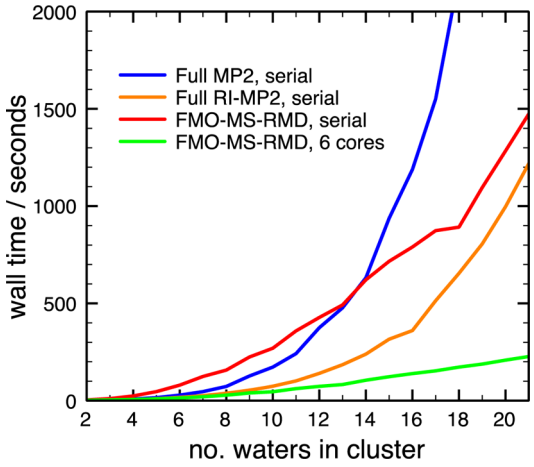


Figure 5. Wall times for MP2/cc-pVDZ, RI-MP2/cc-pVDZ, and FMO-MS-RMD on the Cambridge cluster database global minimum geometries. Calculations were performed on a 2.40 GHz 6-core Intel Xeon(R) CPU. MP2/cc-pVDZ data points for clusters with more than 17 waters are omitted for clarity.

dimer level) scales like $O(N_{\text{frag}}^2)$, and FMO-MS-RMD scales such as $O(N_{\text{states}}N_{\text{frag}}^2)$. This raises the concern that FMO-MS-RMD might be too costly, perhaps even worse than full MP2. On the contrary, Figure 5 shows that the total wall time for an FMO-MS-RMD energy calculation executed in serial scales better than the full MP2/cc-pVDZ energy calculation, crossing over to lesser wall time at a cluster size of about 14 waters. We also compare against the serial timing using RI-MP2, the Resolution of the Identity approximation for MP2, which tends to be much less costly than regular MP2. In Figure 5, we see that in serial RI-MP2 is actually faster than FMO-MS-RMD in these clusters. It does appear, though, judging by the slopes in Figure 5, that FMO-MS-RMD will eventually becoming faster than RI-MP2 in serial if the plot were extrapolated to larger water clusters. Computational speed aside, the memory requirements for FMO-MS-RMD scale linearly with system size, requiring no more core memory than that of a dimer FMO calculation. MP2 and RI-MP2 memory requirements are significantly larger and could limit their feasibility in larger molecular systems.

While the evidence that RI-MP2 can be faster than FMO-MS-RMD may seem discouraging, one of the biggest advantages of FMO-MS-RMD is that it is readily parallelized. Each fragmentation state constitutes a set of independent FMO calculations that can be distributed evenly across processors. To illustrate, we show that running FMO-MS-RMD in parallel with a mere six processors (parallelizing over fragmentation states) is dramatically faster than both the MP2 and the RI-MP2 serial calculations, and it exhibits a nearly perfect speedup. Our code implementation, though, has the ability to scale across many more processors by additionally distributing the individual monomer and dimer calculations, not just the states. For instance, in the MD calculations presented in Section 6, we use up to as many as 256 processors in parallel with appreciable parallel efficiency.

6. RESULTS: MOLECULAR DYNAMICS

Our goal in developing FMO-MS-RMD is not so much to produce more accurate absolute energies, as in Section 5, but more to be able to propagate stable MD based on FMO in a system undergoing chemical reactions. To put our model to the test, we ran a 30 ps MD trajectory with FMO-MS-RMD on a protonated heptamer water cluster ($H_{15}O_7^+$). MD was run with a 1.0 fs time step in the constant NVE ensemble to measure energy conservation, a metric we use here for MD stability.

Figure 6 shows the total (potential plus kinetic) energy across the whole trajectory. The mean energy is −335160.7

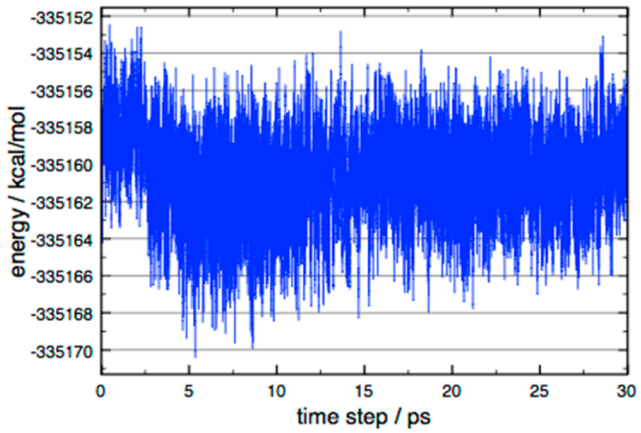


Figure 6. Total energy for the 30 ps MD trajectory of $H_{15}O_7^+$.

kcal/mol with a standard deviation of 2.3 kcal/mol. One can infer from Figure 6 that energy is not *perfectly* conserved, noting the slight drift that occurs from the beginning to about 10 ps. The energy drift, though, is comparable to the magnitude of energy fluctuations and is effectively continuous. Slow drift in total energy is known to be a feature of some multistate methods, such as MS-EVB3, and it is ultimately a result of small energy discontinuities brought about by the finite-sized dynamic basis of states used in the model Hamiltonian matrix, where weakly coupled states possibly appear/disappear as the molecular geometry and pivot state change throughout the trajectory. It is possible to improve energy conservation by including more states, but this comes at the cost of incurring more CPU time for a likely small correction. If one is able to include *all* possible fragmentation states under consideration (i.e., a complete basis of fragmentation states), then FMO-MS-RMD will not be subject to finite basis discontinuities whatsoever. Inclusion of all possible states is difficult for protonated water but might very well be feasible in other chemically reactive systems, for example, a two state model of some donor−acceptor reaction.

Nevertheless, we believe that the performance of our FMO-MS-RMD model with regard to energy conservation is acceptable and is demonstrably stable for the time scales in which we expect FMO-MS-RMD could be feasibly applied. Consider the stability of the MD as proton transfers takes place. Throughout the 30 ps trajectory, we observe qualitatively that nine distinct proton transfers occur, changing which fragment most resembles a hydronium. These proton transfers often happen with the shuttling proton shortly oscillating between two water fragments before finally being more completely accepted by one of the two fragments. This behavior can be observed more quantitatively by tracking the index of the fragment (or O atom) that belongs to the hydronium fragment in the pivot state (i.e., the state with the maximum c_A^2). This is shown in Figure 7. The pivot state hydronium "jumps" several times between fragment indexes, representative of a proton transfer. The pivot state hydronium is seen to traverse most of

dx.doi.org/10.1021/ct400516x | J. Chem. Theory Comput. 2013, 9, 4018−4025

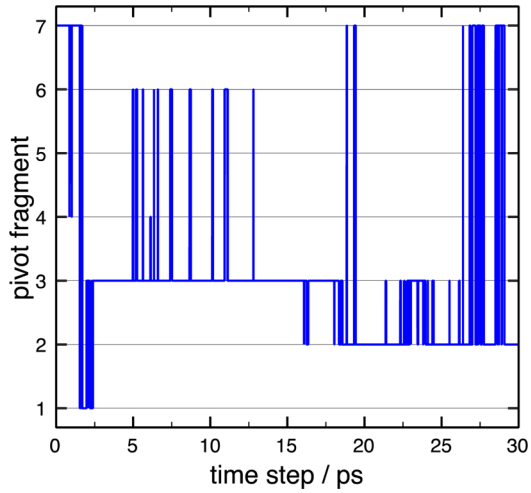


Figure 7. Pivot state hydronium fragment index throughout the 30 ps MD trajectory of $H_{15}O_7^+$. Jumps are indicative of proton transfers between fragments, which are often oscillatory before a complete transfer.

the fragments, indicating that the hydronium fragment migrates throughout the water cluster over the course of the MD. The observation that FMO-MS-RMD is free of spuriously large discontinuities in its total energy throughout these proton transfers supports our assertion that the model is indeed stable for propagating MD.

Furthermore, FMO-MS-RMD tends to follow a potential energy surface very similar to the full MP2/cc-pVDZ calculation. Figure 8 shows the evolution of the potential

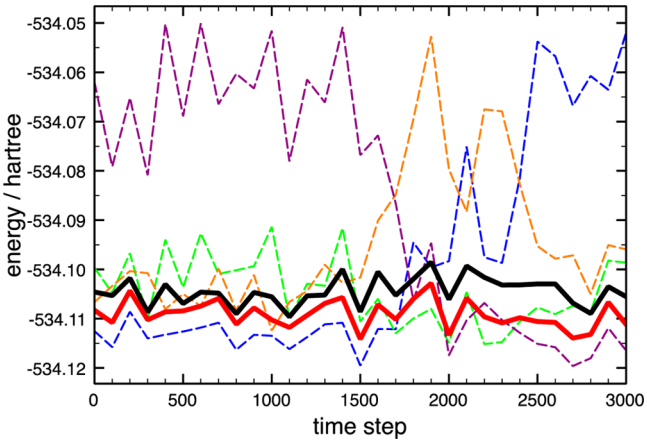


Figure 8. Comparison of four FMO state potential energies (blue, orange, green, and purple dashed lines) and FMO-MS-RMD potential energy (solid red line) to the full MP2/cc-pVDZ potential energy (solid black line) for first 3 ps of MD trajectory of $H_{15}O_7^+$. Data is shown at intervals of 100 time steps for visual clarity. Note that the FMO energies do not include the empirical terms E_A^{inter} and E_A^{intra}.

energy of FMO-MS-RMD, MP2/cc-pVDZ, and four selected fragmentation states extracted from the first 3 ps of the MD trajectory. The MP2/cc-pVDZ energy line is, of course, our target, and FMO-MS-RMD seems to mostly resemble it with a relatively small offset. Contrast this with the four FMO states. Initially, the FMO state represented with the blue line in Figure 8 is the "best" state, being the pivot state in FMO-MS-RMD and somewhat mimicking the MP2/cc-pVDZ line. But, at about 1600 fs, a proton transfer takes place, and the green line FMO state becomes the pivot state. The blue line FMO state then drifts out of favor to higher energies, and as another proton transfer takes place, the purple line FMO state comes into favor. Thus, we see that FMO-MS-RMD successfully and smoothly interpolates between these different fragmentation states in such a way that tends to maintain its similarity to the MP2/cc-pVDZ energy. This clearly displays the ability of FMO-MS-RMD to handle dynamic fragmentation. If instead we had chosen to use some sort of distance based criteria to determine which fragment should be a hydronium, then we would introduce significant energy discontinuities hopping from one fragmentation state to another during a proton transfer, akin to jumping between the FMO energy lines in Figure 8, which would very likely result in poor energy conservation.

8. CONCLUSIONS

We have introduced a new multistate framework for FMO calculations, which could potentially be extended to other fragment methods and any level of electronic structure theory. Our framework, FMO-MS-RMD, has been implemented as an initial example for protonated water clusters based upon the FMO/MP2/cc-pVDZ level of theory with noniterative FMO calculations. Our FMO-MS-RMD model includes a few empirically chosen functions to ensure proper interpolation between different fragmentation states, and we have shown that this model yields stable MD for a small water cluster where frequent proton transfer reactions take place. Our implementation of FMO-MS-RMD can be viewed as an improvement in terms of accuracy over our previous MS-EVB models. Moreover, FMO-MS-RMD tends to be somewhat more accurate for energy than FMO alone in our tests. FMO-MS-RMD is more computationally expensive than FMO, but it benefits greatly from parallel computing because each fragmentation state is independent. Future research in applying the FMO-MS-RMD model to the condensed phase with periodic boundary conditions and quantum mechanics/molecular mechanics (QM/MM) is underway in our group.

APPENDIX

Modeling Other Chemical Reactions

We have suggested throughout this work that FMO-MS-RMD represents a framework that could be expanded to various other chemical reactions. We discuss this further here with a few hypothetical examples. We do not examine the fine details or merits of modeling these reactions with FMO-MS-RMD but bring them up merely for illustration.

The key concept behind applying FMO-MS-RMD to modeling a chemical reaction step is to envision how a molecular system might be fragmented correspondingly into a reactant state and a product state. More generally, there may be several reactant and/or product states representative of a series of reactions, as is the case in protonated water. For example, FMO-MS-RMD is well suited to handle any chemical reaction that follows a donor−acceptor paradigm: D−X + A → D + X−A, where X is the atom or molecular group being transferred. D−X and A are separate fragments on the reactant side, and D and X−A are fragments on the product side. Proton transfer obviously follows this donor−acceptor paradigm. The well-known SN2 type reaction, which includes a wide variety of examples, also fits this paradigm; the donor is the leaving group, and the acceptor is the nucleophile. A concrete example of an

dx.doi.org/10.1021/ct400516x | J. Chem. Theory Comput. 2013, 9, 4018−4025

SN2 reaction that potentially could be studied with FMO-MS-RMD is the reaction of ethyl bromide with hydroxide (in aqueous solution): EtBr + OH^- → Br^- + EtOH, where the ethyl group is being transferred between the bromine donor and the hydroxide acceptor. Another possible reaction FMO-MS-RMD could be applied to is dissociation/association reactions (depending on direction): R−X → R + X. One simple concrete example might be aqueous heterolytic dissociation of KCl. Another example might be the individual dissociation/association steps of an SN1 type reaction. It is even conceivable that one could use FMO-MS-RMD to model more than one individual step of a multi-step chemical reaction all at once, where each intermediate reactant/product state is merely a different fragmentation of the molecular system.

Of course, the above just describes how to make a multi-state *ansatz* used for the FMO calculations. The remaining difficulty will be deciding on how to model off-diagonal coupling as well as any penalties. It is likely that one could develop distance based switching regions similar to what we have done for protonated water, but there are probably several other avenues one could take. This open topic is something we intend to explore in future work with FMO-MS-RMD.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gavoth@uchicago.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (NSF, Grant No. CHE-1214087), the Air Force Office of Scientific Research (AFOSR, Grant No. FA9550-13-1-0094), and by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract DE-AC02-06CH11357. This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357. This work was completed in part with resources provided by the University of Chicago Research Computing Center. AWL is an Argonne Leadership Computing Facility Early Science Project postdoctoral fellow.

REFERENCES

(1) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632−672.

(2) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular orbital method: An approximate computational method for large molecules. *Chem. Phys. Lett.* **1999**, *313*, 701−706.

(3) Fedorov, D. G.; Kitaura, K. Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *J. Phys. Chem. A* **2007**, *111*, 6904−6914.

(4) Fedorov, D.; Kitaura, K. Theoretical background of the Fragment Molecular Orbital (FMO) method and its implementation in GAMESS. In *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*; Fedorov, D., Kitaura, K., Eds.; Taylor & Francis: Boca Raton, FL, 2010; pp 1−268.

(5) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment molecular orbital method: Use of approximate electrostatic potential. *Chem. Phys. Lett.* **2002**, *351*, 475−480.

(6) Fedorov, D. G.; Kitaura, K. The importance of three-body terms in the fragment molecular orbital method. *J. Chem. Phys.* **2004**, *120*, 6832−6840.

(7) 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.

(8) Dahlke, E. E.; Truhlar, D. G. Electrostatically embedded many-body expansion for simulations. *J. Chem. Theory Comput.* **2008**, *4*, 1−6.

(9) Xie, W. S.; Song, L. C.; Truhlar, D. G.; Gao, J. L. The variational explicit polarization potential and analytical first derivative of energy: Towards a next generation force field. *J. Chem. Phys.* **2008**, *128*, 234108:1−9.

(10) Jacobson, L. D.; Herbert, J. M. An efficient, fragment-based electronic structure method for molecular systems: Self-consistent polarization with perturbative two-body exchange and dispersion. *J. Chem. Phys.* **2011**, *134*, 094118:1−17.

(11) 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.

(12) Hirata, S. Fast electron-correlation methods for molecular crystals: An application to the α, $\beta(1)$, and $\beta(2)$ modifications of solid formic acid. *J. Chem. Phys.* **2008**, *129*, 204104:1−11.

(13) Kamiya, M.; Hirata, S.; Valiev, M. Fast electron correlation methods for molecular clusters without basis set superposition errors. *J. Chem. Phys.* **2008**, *128*, 074103:1−11.

(14) Beran, G. J. O.; Hirata, S. Fragment and localized orbital methods in electronic structure theory. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7559−7561.

(15) Komeiji, Y.; Mochizuki, Y.; Nakano, T. Three-body expansion and generalized dynamic fragmentation improve the fragment molecular orbital-based molecular dynamics (FMO-MD). *Chem. Phys. Lett.* **2010**, *484*, 380−386.

(16) Schmitt, U. W.; Voth, G. A. The computer simulation of proton transport in water. *J. Chem. Phys.* **1999**, *111*, 9361−9381.

(17) Wu, Y. J.; Chen, H. N.; Wang, F.; Paesani, F.; Voth, G. A. An improved multistate empirical valence bond model for aqueous proton solvation and transport. *J. Phys. Chem. B* **2008**, *112*, 467−482.

(18) Knight, C.; Lindberg, G. E.; Voth, G. A. Multiscale reactive molecular dynamics. *J. Chem. Phys.* **2012**, *137*, 22A525:1−11.

(19) Song, L. C.; Mo, Y. R.; Gao, J. L. An effective hamiltonian molecular orbital-valence bond (MOVB) approach for chemical reactions as applied to the nucleophilic substitution reaction of hydrosulfide ion and chloromethane. *J. Chem. Theory Comput.* **2009**, *5*, 174−185.

(20) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(21) Hodges, M. P.; Wales, D. J. Global minima of protonated water clusters. *Chem. Phys. Lett.* **2000**, *324*, 279−288.

(22) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C. P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.

(23) Hodges, M. P.; Stone, A. J. Modeling small hydronium-water clusters. *J. Chem. Phys.* **1999**, *110*, 6766−6772.