JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Quantum Fragment Based *ab Initio* Molecular Dynamics for Proteins

Jinfeng Liu,[†] Tong Zhu,[*,†,‡] Xianwei Wang,[§] Xiao He,[*,†,‡] and John Z. H. Zhang[*,†,‡,∥]

[†]State Key Laboratory of Precision Spectroscopy, Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062, China

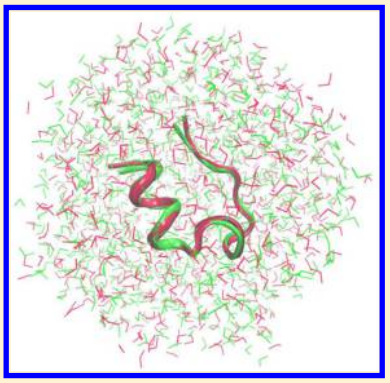
[‡]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

[§]Center for Optics & Optoelectronics Research, College of Science, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China

[∥]Department of Chemistry, New York University, New York, New York 10003, United States

Supporting Information

ABSTRACT: Developing *ab initio* molecular dynamics (AIMD) methods for practical application in protein dynamics is of significant interest. Due to the large size of biomolecules, applying standard quantum chemical methods to compute energies for dynamic simulation is computationally prohibitive. In this work, a fragment based *ab initio* molecular dynamics approach is presented for practical application in protein dynamics study. In this approach, the energy and forces of the protein are calculated by a recently developed electrostatically embedded generalized molecular fractionation with conjugate caps (EE-GMFCC) method. For simulation in explicit solvent, mechanical embedding is introduced to treat protein interaction with explicit water molecules. This AIMD approach has been applied to MD simulations of a small benchmark protein Trpcage (with 20 residues and 304 atoms) in both the gas phase and in solution. Comparison to the simulation result using the AMBER force field shows that the AIMD gives a more stable protein structure in the simulation, indicating that quantum chemical energy is more reliable. Importantly, the present fragment-based AIMD simulation captures quantum effects including electrostatic polarization and charge transfer that are missing in standard classical MD simulations. The current approach is linear-scaling, trivially parallel, and applicable to performing the AIMD simulation of proteins with a large size.



1. INTRODUCTION

Molecular dynamic (MD) simulation is currently the most important tool for computational study of structural and dynamical properties of biomolecules.[1−3] At the heart of the MD approach is the force field that describes intermolecular interactions of the system. Thus, the accuracy and reliability of simulation results depend heavily on the accuracy of the force field[4−6] employed in the simulation. The traditional route followed in MD is to determine the force field in advance. Typically, the full interactions in the system are broken up into pairwise analytical functions that describe bond stretching, bond angle bending, torsions, and nonbond terms including van der Waals and Coulomb interactions.

Despite widely successful applications of the current force fields in biomolecular simulations, the simplified, predefined pairwise force field has serious drawbacks. One of the major simplifications in widely used force fields is that the atomic charges are prefixed, and there is no explicit treatment of electrostatic polarization and charge transfer.[7] This approximation is known to be problematic when performing the MD simulation for the study of protein−ligand interaction and protein folding.[8,9] Although there has been great interest in the development of a polarizable force field for biomolecular systems, generally accepted models have not emerged, and several alternatives for the treatment of polarizability, including use of induced dipoles,[10−12] fluctuating charges,[13,14] and Drude oscillators,[15,16] remain under active development. In addition, point charge representation itself is problematic when the interacting pairs are very close to each other.

Quantum mechanics, in principle, can provide accurate potential energy function for biomolecules, including important quantum effects. Following the pioneering work of Car and Parrinello,[17−19] the so-called *ab initio* molecular dynamics (AIMD) method was developed to address these problems. In the AIMD approach, atomic forces are calculated by QM methods such as HF and DFT, whereas the nuclear dynamics of the system is described by classical mechanics. AIMD is currently a popular and expanding computational tool employed to study physical, chemical, and biological phenomena. In the previous work of Wei et al.,[20] the AIMD simulation was utilized in studying gas-phase conformational dynamics of an alanine dipeptide analogue. It was found that transformation between two stable conformations of alanine dipeptide named C5 and $C7_{eq}$ occurs on the picosecond time scale, whereas classical molecular dynamics using popular force fields does not yield a transition even after nanoseconds. Klein and coworkers[21] studied the solute−solvent charge transfer in a

Received: June 15, 2015
Published: October 23, 2015

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

solvated glycine dipeptide by performing the density functional theory (DFT) based AIMD simulations and confirmed that solute−solvent interactions involve a significant amount of charge transfer. Martinez and co-workers also found that AIMD are consistently better than classical force fields at reproducing experimental structures for proteins with disordered regions.[22−24]

However, most of the previous AIMD studies focused on relatively small systems such as polypeptides. This is mainly because the application of QM methods has been limited by its computational cost which increases rapidly with the size of the system being studied. To extend the applicability of rigorous QM methods to large systems, considerable effort has been made to developing linear-scaling and/or fragmentation methods. Among the existing approaches, the fragmentation approach has attracted much attention.[25−32] The fragmentation approach takes advantage of the chemical locality of molecular systems and assumes that local regions of the system are only weakly influenced by atoms that are far away from the region of interest.[30,33−36] In these methods, the system is divided into subsystems whose energies are calculated separately by the QM method with iterations, and finally the property of the whole system can be obtained by taking a proper combination of the properties of individual fragments. The fragmentation method is attractive in several aspects including easy implementation of parallelization without extensively modifying the existing QM programs and straightforward application at all levels of *ab initio* electronic structure theories. A range of fragmentation approaches for the QM calculation of large molecular systems has been proposed in recent years,[37] and they are extensively reviewed in refs 25 and 26.

In this paper, we employ an electrostatically embedded generalized molecular fractionation with conjugate caps (EE-GMFCC) method for quantum calculation of protein potential energy.[38] In the EE-GMFCC scheme, fragment-based energies of neighboring residues and interaction energies of non-neighboring residues that are spatially in close contact are computed by quantum mechanics, whereas the interaction energies between distant non-neighboring residues are treated by charge−charge Coulomb interactions. The EE-GMFCC method is computationally efficient and linear-scaling with a low prefactor. An individual fragment normally contains fewer than 60 atoms and thus can be carried out using standard quantum chemistry methods. Numerical studies showed that the EE-GMFCC approach can accurately reproduce the full system potential energy of proteins at the HF, DFT, and MP2 levels.[38] Based on the potential energy of the protein computed with the EE-GMFCC method, the energy gradients (atomic forces) are derived and used in the MD simulation of protein dynamics. The paper is organized as follows. Section 2 provides a description of the EE-GMFCC approach and the calculation of atomic forces. In section 3, we performed AIMD simulations on a small protein Trpcage to validate the new method. Finally, a brief summary is given in section 4.

2. THEORY AND METHODOLOGY

A. EE-GMFCC Approach for Protein Energy and Atomic Forces. The EE-GMFCC method was developed based on the original molecular fractionation with conjugate caps (MFCC) approach.[39] In this approach, a given protein with N amino acids is decomposed into N individual amino acid units by cutting through the peptide bond as illustrated in Figure 1.

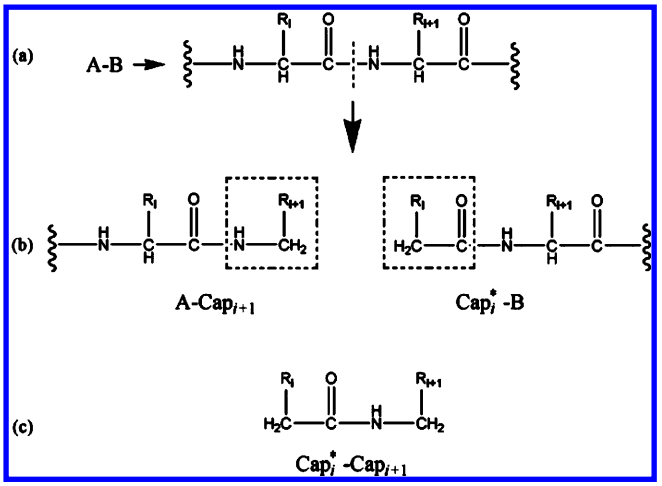


Figure 1. MFCC scheme in which the peptide bond is cut in panel (a) and the fragments are capped with Cap$_{i+1}$ and its conjugate Cap$_i^*$ in panel (b), where i represents the index of the i^{th} amino acid in the given protein. The atomic structure of the concap is shown in panel (c) defined as the fused molecular species Cap$_i^*$−Cap$_{i+1}$.

To preserve the local chemical environment of the remaining fragment for each separated amino acid, a pair of conjugate caps is designed to saturate each fragment. Hydrogen atoms are added to terminate the molecular caps to avoid dangling bonds (see Figure 1).

However, the MFCC scheme only includes the self-energy of individual residue and interaction energy between sequentially connected tripeptides. To obtain the total energy expression for proteins, we also proposed a generalized molecular fractionation with conjugate caps/molecular mechanics (GMFCC/MM) scheme,[40] as shown in Figure 2. In this scheme, classical force field interactions are introduced to represent the long-range interaction between distant non-neighboring fragments, whereas a generalized concap (Gconcap, as shown in Figure 2) was

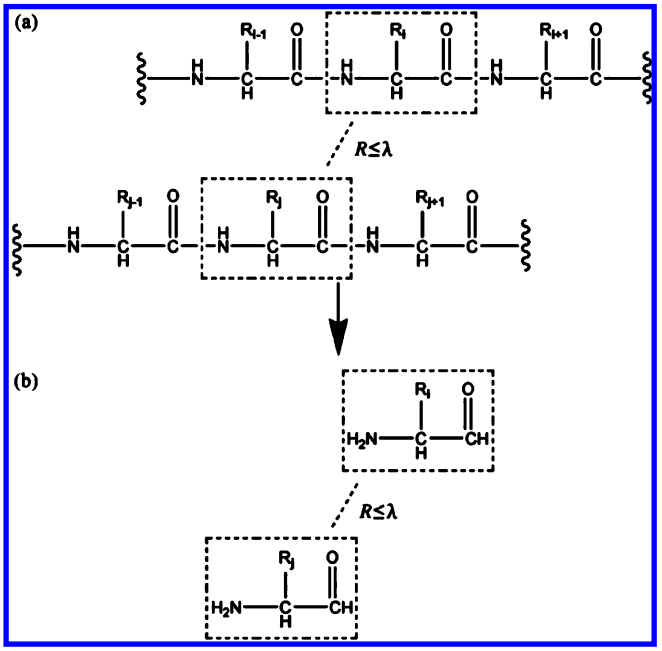


Figure 2. (a) Generalized concap (Gconcap) scheme, when the distance (R) between two non-neighboring residues i and j is within a distance threshold λ ($R \leq \lambda$). (b) Atomic structure of Gconcap.

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

introduced to account for the two-body QM interaction energies between short-range non-neighboring fragment interactions.

Several previous studies have shown that the electrostatic polarization arising from the environment is important for including the many-body effect in fragmentation methods.[41,42] Therefore, in the EE-GMFCC approach,[38] the QM calculation of each fragment is embedded in the electrostatic field of the point charges representing the remaining fragments of the protein to account for the environmental effect. In the gas phase, the total energy of a protein can then be expressed using the following formula

$$\begin{aligned} E_{\text{protein}}^{\text{EE\_GMFCC}} &= E_{\text{fragment}} - E_{\text{concap}} + E_{\text{two-body}} - E_{\text{double-counting}} \\ &= \sum_{i=2}^{N-1} \tilde{E}(\text{Cap}_{i-1}^{*}\text{A}_i\text{Cap}_{i+1}) - \sum_{i=2}^{N-2} \tilde{E}(\text{Cap}_i^{*}\text{Cap}_{i+1}) \\ &\quad + \sum_{\substack{i,j>i+2,\\ |R_i-R_j|\le\lambda}} [\tilde{E}(\text{A}_i\text{A}_j) - \tilde{E}(\text{A}_i) - \tilde{E}(\text{A}_j)] \\ &\quad - \left\{ \sum_{k,l}\sum_{m,n} \frac{q_{m(k)}q_{n(l)}}{R_{m(k)n(l)}} - \sum_{\substack{i,j>i+2,\\ |R_i-R_j|\le\lambda}} \sum_{m',n'} \frac{q_{m'(i)}q_{n'(j)}}{R_{m'(i)n'(j)}} \right\} \end{aligned} \tag{1}$$

where i and j denote the index for the i^{th} and j^{th} residues, respectively. The $\tilde{E}$ represents the summation of the self-energy of the QM zone (atoms treated by QM) shown in the parentheses and the interaction energy between the QM zone and background charges which represents the rest of the system (MM zone)

$$\begin{aligned} \tilde{E}(\text{QM\_zone}) &= E^{\text{QM\_zone}} \\ &\quad + \sum_{k}^{N_{\text{MM}}} q_k \left[\sum_{\alpha} \frac{Z_\alpha}{|\mathbf{R}_k - \mathbf{R}_\alpha|} - \int \mathrm{d}\mathbf{r} \frac{\rho_{\text{QM\_zone}}(\mathbf{r})}{|\mathbf{R}_k - \mathbf{r}|} \right] \end{aligned} \tag{2}$$

where N_{MM} is the number of background charges, q_k and $\mathbf{R}_k$ are the charge and coordinates of the k^{th} background charge, and Z_α and $\mathbf{R}_\alpha$ are the nuclear charge and coordinates of the α^{th} nuclei. Thus, the first term $(\tilde{E}(\text{Cap}_{i-1}^{*}\text{A}_i\text{Cap}_{i+1}))$ in eq 1 represents the summation of the self-energy of fragment i (the i^{th} residue A_i capped with a left cap Cap_{i-1}^{*} and a right cap Cap_{i+1}) and the interaction energy between it and its background charges. Then the self-energy of concap along with the interaction energy between the concap and its background charges $(\tilde{E}(\text{Cap}_i^{*}\text{Cap}_{i+1}))$ should be deducted.

Furthermore, if the distance of any two atoms between residue A_i and A_j is less than or equal to the distance threshold λ, the interaction energy between these two residues are described by quantum mechanics. The distance threshold used in this work is 4.0 Å, since the EE-GMFCC calculated protein potential energies have been shown to be converged at this value.[38]

The last term in eq 1 approximately cancels the double counting of the interaction energy by charge−charge interactions approximately. The $q_{m(k)}$ represents the point charge of the m^{th} atom in the fragment k. It has been proved in our previous study that k = (H)Cap$^{*}_{i-1}$-NH and l = CO-$\text{A}_{i+2}\text{A}_{i+3}$...$\text{A}_N$ (i = 2, 3, ..., N−2).[38] Through the monomer and dimer calculations, one- and two-body Coulomb, exchange, and correlation interactions are included nearly exactly in the QM calculation, while the three- and all higher-order Coulomb effects are approximated through the embedding field. As shown in the previous study, the potential energy of protein calculated by the EE-GMFCC approach is very close to the full-system result. Tables S1−S4 show the coordinates of the initial structure including two fragments and a concap created by the EE-GMFCC method for AIMD in the gas phase. For more details about energy calculation with the EE-GMFCC approach, please refer to ref 38.

To obtain atomic forces, we need to compute the differentiation of $E_{\text{protein}}^{\text{EE\_GMFCC}}$ with respect to nuclear coordinates. For a given atom m

$$\mathbf{F}_m = -\nabla_m E_{\text{protein}}^{\text{EE\_GMFCC}} \tag{3}$$

In the calculation of every single fragment, there are both atoms in the QM zone and the MM zone. For atom m, if it is in the QM zone, its force can be calculated by

$$\begin{aligned} \mathbf{F}_m^{\text{QM}} &= -\nabla_m \tilde{E}(\text{QM\_zone}) \\ &= -\nabla_m E^{\text{QM\_zone}} - \sum_{k}^{N_{\text{MM}}} q_k \nabla_m \\ &\quad \left[\sum_{\alpha} \frac{Z_\alpha}{|\mathbf{R}_k - \mathbf{R}_\alpha|} - \int \mathrm{d}\mathbf{r} \frac{\rho_{\text{QM\_zone}}(\mathbf{r})}{|\mathbf{R}_k - \mathbf{r}|} \right] \end{aligned} \tag{4}$$

which can be obtained directly from the QM calculation, with the wave function polarized by background charges. On the other hand, if the atom m is represented by the background charge, its force can be calculated by

$$\mathbf{F}_m^{\text{BC}} = q_m \mathbf{E}_{\text{QM\_zone}}(\mathbf{R}_m) \tag{5}$$

where $\mathbf{E}_{\text{QM\_zone}}(\mathbf{R}_m)$ is the electric field at the position of atom m due to the QM zone:

$$\begin{aligned} \mathbf{E}_{\text{QM\_zone}}(\mathbf{R}_m) &= \sum_{\alpha} \frac{Z_\alpha(\mathbf{R}_m - \mathbf{R}_\alpha)}{|\mathbf{R}_m - \mathbf{R}_\alpha|^3} \\ &\quad - \int \mathrm{d}\mathbf{r} \frac{\rho_{\text{QM\_zone}}(\mathbf{r}) \times (\mathbf{R}_m - \mathbf{r})}{|\mathbf{R}_m - \mathbf{r}|^3} \end{aligned} \tag{6}$$

Then, the final force of the atom m in the protein can be obtained by combining its forces calculated in every fragment and deducting the double counting term by direct derivative of the charge−charge interactions in the last term of eq 1.

B. *Ab Initio* Molecular Dynamics for the Protein System. In this study, AIMD simulations were performed by combining the EE-GMFCC approach with a modified version of the Amber12 package.[43] For every step of the simulation, atomic forces of the protein were calculated using the EE-GMFCC approach and then passed to the MD engine (the Sander module) of Amber. We selected a small protein Trpcage (PDBID: 1L2Y, 20 residues with 304 atoms, the structure of the first isomer in the PDB file was selected) to assess the performance of this method. All the titratable residues were assigned to the protonation states corresponding to a pH of 7. When the AIMD simulation was performed in the explicit solvent, the protein was solvated in a water ball with 2547 TIP3P water molecules. A confining potential with a force constant of 10 kcal/mol·Å was added to avoid boiling off water molecules in the simulation. Prior to the AIMD simulation, the whole system was relaxed by 5000 steps with constraints on the protein, and then the system was heated to 298 K by 25 ps heating simulation. The resulting geometry was used as the

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

initial structure for the AIMD simulation. All of the QM calculations were carried out at the M062X/6-31G* level[47,48] using the Gaussian09 package.[49] For comparison, we also performed classical MD simulations for this protein in solution with the same time scale using the Amber ff99SB force field. The time step used in this work is 1 fs, which is typically adopted by most *ab initio* QM/MM MD simulations. A detailed analysis on the time step−RMS(Energy) relationship in AIMD with the FMO approach can be found in previous studies of Y. Komeiji et al.[44] and K. R. Brorsen et al.[45,46]

3. RESULTS AND DISCUSSION

A. AIMD in the Gas Phase. The first test is for the gas-phase MD simulation of a benchmark small protein (Trpcage). The snapshot of the Trpcage structure at the end of an 8.8 ps AIMD simulation is plotted in Figure 3, along with the initial structure in the AIMD simulation. As shown in Figure 3, the structure of this protein is largely retained in the AIMD simulation. The fluctuation of the total energy, potential energy, and the structural changes of the protein during the simulation (in terms of the RMSD of backbone heavy atoms with respect to the initial structure) are shown in Figure 4. It can be seen from the figure that the energies are stable but with a small down drift. The standard errors of the potential energy and the total energy are, respectively, 0.042 and 0.048 hartree.

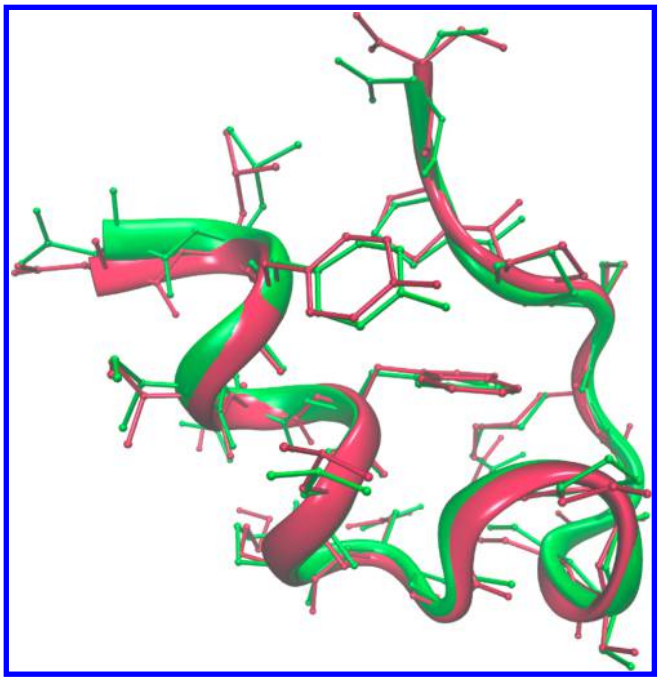


Figure 3. Comparison of the initio structure (red) and the last snapshot (green) of Trpcage in the gas-phase AIMD simulation.

For the AIMD simulation in the gas phase, the backbone RMSD shows a steady increase (Figure 4), indicating that the initial structure of the protein is not stable in the gas phase. This correlates with a small down drift of energy as the protein tries to relax the structure by lowing the energy as shown in Figure 4. This phenomenon is expected in the gas-phase simulation of protein. Overall, the smoothness in energy and RMSD curves from the calculation indicates that the AIMD result is reliable.

The EE-GMFCC method is trivially parallelizable since each fragment can be calculated independently on separate processors. For this system, it took about 7 min to perform the one-step MD simulation using 96 2.66 GHz processers (12 processers for the QM calculation of each fragment), and this time can be reduced to approximately 2.8 min if 240 processers were used. In fact, if the number of processors increases linearly with the size of the system, the time to perform every MD step will approximately keep constant. With further optimization and improvement in computer capacity, this approach will become more practical for large protein systems.

B. AIMD in an Explicit Water Environment. It has been long recognized that water plays an important role in protein structure and dynamics.[50,51] The EE-GMFCC approach can also be applied in explicit solvent environment, and the method is similar to that in the gas phase. The total energy of the protein−solvent system with EE-GMFCC can be expressed as the summation of protein energy, protein−solvent interaction energy, and solvent energy as follows

$$E_{\text{total}} = E_{\text{protein}}^{\text{EE-GMFCC}} + E_{\text{water}}^{\text{MM}} + E_{\text{protein-water}}^{\text{QM/MM}} \tag{7}$$

To reduce the computational cost, water molecules and their interactions are described by the classical force field method

$$E_{\text{water}}^{\text{MM}} = \sum_{i}^{N_{\text{water}}} E_{\text{water}}^{\text{intra}} + \sum_{\substack{\text{nonbonded} \\ \text{atom pairs } m,n}} \left\{ E_{m,n}^{\text{VDW}} + \frac{1}{4\pi\varepsilon_0} \frac{q_m q_n}{r_{m,n}} \right\} \tag{8}$$

$$E_{\text{water}}^{\text{intra}} = k_{\text{OH}}(r\text{-}r_0)^2 + k_{\text{HOH}}(\theta\text{-}\theta_0)^2 \tag{9}$$

$$E_{m,n}^{\text{VDW}} = \frac{A_{m,n}}{r_{m,n}^{12}} - \frac{B_{m,n}}{r_{m,n}^{6}} \tag{10}$$

where q_m, q_n, r_0, θ_0, k_{OH}, k_{HOH}, $A_{m,n}$, and $B_{m,n}$ are force field parameters taken from the Amber ff99SB force field.

The interaction between protein and solvent is described under the QM/MM scheme. The coupling between QM and MM parts was treated by mechanical embedding, which means that the EE-GMFCC calculation of protein is performed in the gas phase (i.e., without background charges of water molecules in the QM calculation of every fragment), and the interaction between protein and water is described by classical force fields

$$E_{\text{protein-water}}^{\text{QM/MM}} = \sum_{m}^{N_{\text{protein}}} \sum_{n}^{3*N_{\text{water}}} \left(E_{m,n}^{\text{VDW}} + \frac{1}{4\pi\varepsilon_0} \frac{q_m q_n}{r_{m,n}} \right) \tag{11}$$

where N_{protein} is the number of atoms in protein, and N_{water} is the number of water molecules. The atomic charges of protein and water are taken from the Amber ff99SB force field. Then, the atomic force of a given atom m in the system can be expressed as

$$\begin{aligned} \mathbf{F}_m &= -\nabla_m E_{\text{total}} \\ &= -\nabla_m E_{\text{protein}}^{\text{EE-GMFCC}} - \nabla_m E_{\text{water}}^{\text{MM}} - \nabla_m E_{\text{protein-water}}^{\text{QM/MM}} \end{aligned} \tag{12}$$

where $\nabla_m E_{\text{protein}}^{\text{EE-GMFCC}}$ is the same as that in eq 3. $\nabla_m E_{\text{water}}^{\text{MM}}$ and $\nabla_m E_{\text{protein-water}}^{\text{QM/MM}}$ are the standard gradient expressions of the classical MM force field.

Electrostatic embedding was not utilized in this study mainly because the TIP3P water model is not necessarily well-suited for interacting with the QM electron density.[52] A previous study by Laaksonen and co-workers[53] has demonstrated that

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

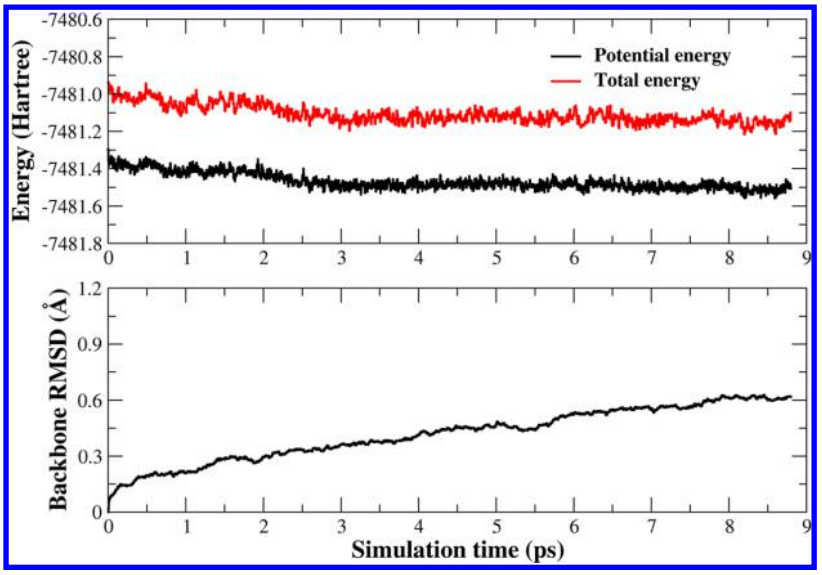


Figure 4. Time evolution of the total energy, potential energy (upper panel), and backbone RMSD (lower panel) with respect to the initial structure in the AIMD simulation of Trpcage in the gas phase.

when the TIP3P water model was utilized in QM/MM calculations with electrostatic embedding, the coupling between QM and MM regions was too strong. To avoid such a problem, some techniques have been developed such as by modifications of the QM-MM Columbic and VDW interactions. However, there does not exist a widely accepted algorithm yet. It is also worth noting that electronic embedding is not always superior to mechanical embedding, and an extensive study of comparing different QM/MM approaches can be found in the work by Hu et al.[54]

We performed the AIMD simulation of Trpcage solvated in explicit water molecules, and the results are shown in Figures 5 and 6. The standard errors of the potential energy and total energy of the whole system are, respectively, 0.102 and 0.107 hartree. In comparison to the simulation in the gas phase, the energy curves are more stable, although the fluctuations of the energies of the whole system (Trpcage plus waters) are slightly larger than those in the gas phase due to the inclusion of solvent molecules in the system. If the energies of water molecules are excluded from the system energy, the standard error of the potential energy of the protein is only 0.024 hartree, which is much smaller than that in the gas phase. Also, the RMSD of the protein backbone is more stable, indicating that the protein structure is more stable in solution than in the gas phase. Furthermore, it can be seen that the protein structure under the AIMD simulation is more stable than that under the Amber ff99SB force field. The overall backbone RMSD in the AIMD simulation is around 0.5 Å, as compared to a value of 0.8 Å under the Amber ff99SB force field. Since the current empirical force fields are built up from individual amino acids or small sample systems such as dipeptides, etc. without calibration for the overall protein structures, it is not hard to understand that the accuracy of their description of the protein structure is uncertain at best.[55−57]

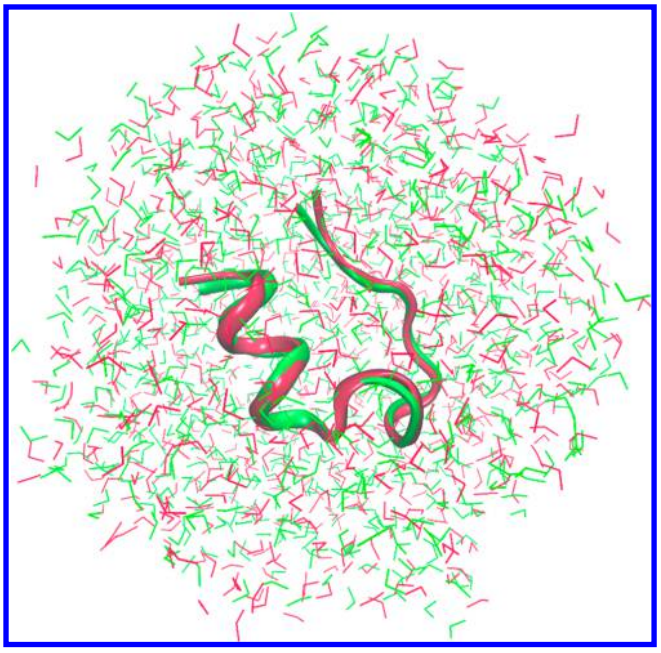


Figure 5. Comparison of the initial structure (red) and that from the last snapshot (green) of Trpcage from the AIMD simulation with explicit solvent.

In contrast, the AIMD approach is based on the *ab initio* calculation of interaction energy for given protein configurations without empirical parameters and is thus superior in terms of accuracy. For example, the AIMD simulation includes QM effects such as electrostatic polarization and charge transfer, both are lacking in standard force fields.[14,58,59] In the EE-GMFCC approach, interactions of sequentially connected tripeptides and residues are treated by QM. Therefore, the local charge transfer effect is included in addition to polarization. To illustrate the electrostatic polarization and charge transfer effects in the protein, we calculated the atomic charges by Mulliken population analysis when the QM calculation was performed for each fragment. In this scheme, the atomic charge of atom k in the protein can be calculated as follows

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

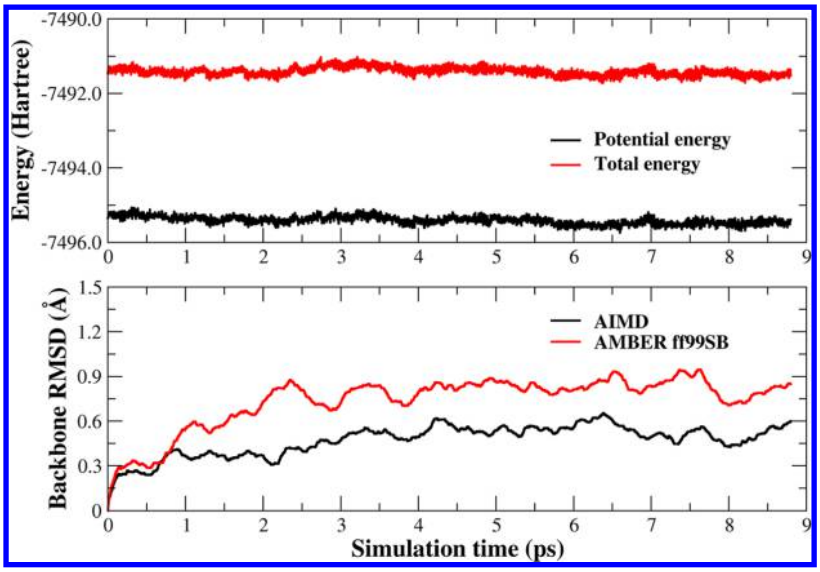


Figure 6. Time evolution of the total energy and the potential energy of the system (Trpcage+waters, upper panel) and protein backbone RMSD (lower panel) with respect to the initial structure in the AIMD simulation with explicit solvent. Protein backbone RMSD computed by the empirical Amber ff99SB force field was also shown for comparison.

$$q_k = \sum_{i=2}^{N-1} q_k(\mathrm{Cap}^*_{i-1}\mathrm{A}_i\mathrm{Cap}_{i+1}) - \sum_{i=2}^{N-2} q_k(\mathrm{Cap}^*_i\mathrm{Cap}_{i+1}) + \sum_{\substack{i,j>i+2,\\ |R_i-R_j|\le\lambda}} (q_k(\mathrm{A}_i\mathrm{A}_j) - q_k(\mathrm{A}_i) - q_k(\mathrm{A}_j)) \quad (13)$$

where $q_k(\mathrm{Cap}^*_{i-1}\mathrm{A}_i\mathrm{Cap}_{i+1})$ denotes the Mulliken charge of atom k that belongs to the fragment $\mathrm{Cap}^*_{i-1}\mathrm{A}_i\mathrm{Cap}_{i+1}$. To avoid double counting, the atom k's charge in the concap $\mathrm{Cap}^*_i\mathrm{Cap}_{i+1}$ needs to be deducted. The third term in eq 13 includes the two-body effect between nonsequentially connected residues that are spatially close. Figure S1 plots the atomic charges, obtained from both EE-GMFCC and full system QM calculations, from the last snapshot of the AIMD simulation in solution. As shown, the calculated results from the EE-GMFCC calculation are very close to that from the full system QM calculation. This is in line with expectations and demonstrates that the polarization and charge transfer effect are preserved in the EE-GMFCC method.

The partial charges of backbone nitrogen and oxygen atoms calculated using the EE-GMFCC scheme are shown in Figure 7a. For comparison, Amber ff99SB charges of those atoms are also presented. As shown in Figure 7a, the fluctuation of EE-GMFCC calculated charges is larger than that of the Amber ff99SB force field. In addition, atoms in the same type of residue but in the different environment have different charges due to electrostatic polarization and charge transfer effects. For instance, the atomic charges of backbone nitrogen in residues GLY10 and GLY11 are −0.6566*e* and −0.6071*e*, respectively. In contrast, in the Amber ff99SB force field, these two nitrogen atoms have the same charge of −0.4157*e*. Based on atomic charges calculated by EE-GMFCC, we also calculated the excess charge (calculated by subtracting the formal charge of the isolated residue at pH of 7) of each residue, which is shown in Figure 7b. For some charged residues, such as LYS8, the excess charge is as large as −0.30*e*, indicating that the charge transfer effect is significant and cannot be neglected. The excess charges of the same type of amino acid at different locations in the protein are different as well, depending on their local

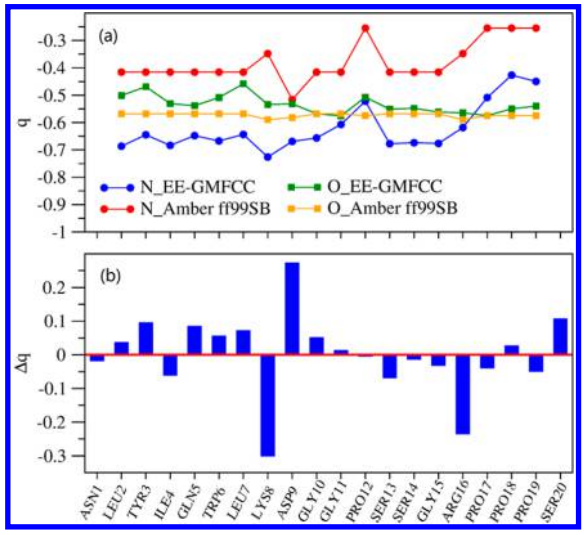


Figure 7. (a) Partial charges (in atomic units) of the backbone N and O atoms in each residue of Trpcage. (b) Excess charge (in atomic units) for each residue in Trpcage.

chemical environment. For example, the excess charge of SER13 is −0.068*e*, much larger than that of SER14 (−0.013*e*), which may originate from the fact that SER13 forms a hydrogen bond with GLY10. The Mulliken charges calculated in this study are utilized to demonstrate that the electrostatic polarization and charge transfer effects are included in the AIMD simulation. Other atomic charge models such as Natural Bond Orbital (NBO) analysis and restrained electrostatic potential charge fitting (RESP) can also be combined with EE-GMFCC for other biological applications.

4. CONCLUSION

In this study, we demonstrated the applicability of a recently developed electrostatically embedded generalized molecular fractionation with conjugate caps method (EE-GMFCC) for performing *ab initio* molecular dynamic (AIMD) simulations

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

on a small protein, both in the gas phase and in the explicit water environment.

In the EE-GMFCC approach, the entire protein is divided into small fragments according to certain criteria, and the total energy of protein is calculated by taking a linear combination of the QM energy of the neighboring residues and two-body QM interaction energy between non-neighboring residues that are spatially in close contact. All the fragment calculations are embedded in the electrostatic field of the point charges representing the remaining amino acids in the protein, which accounts for the electrostatic polarization effect of the protein environment. The interaction energies between distant non-neighboring residues are treated by charge−charge interactions. Atomic forces of protein atoms are calculated from direct differentiation of the EE-GMFCC energies. For the AIMD simulation in explicit water, the mechanically embedded QM/MM scheme is used to treat protein−water interactions.

The above AIMD simulations show that this AIMD approach is potentially powerful and attractive for studying protein dynamics. As compared to classical force fields, the AIMD method is more balanced without any priori structural biases that are inherent in some predefined force fields. Furthermore, important QM effects such as electrostatic polarization and charge transfer are included intrinsically in the AIMD simulation. The calculation of Mulliken charges of protein atoms shows that the charge transfer between residues is significant and cannot be simply ignored.

Further development of this approach will focus on three aspects. First is the accurate treatment of protein−solvent interaction. A problem with the mechanical embedding approach is that it completely ignores the electrostatic polarization effect of the QM system by the MM region. To fix this problem, one can apply more rigorous methods to calculate protein−water interactions, such as using full QM methods to describe water molecules in the framework of the EE-GMFCC approach, or one can modify the classical force field parameters of the water model to make them more suitable for electrostatic embedding. Research along this line is underway in our laboratory. The second is to further improve the efficiency of the AIMD method. There is still more room to lower its computational cost. For example, a precompiled Gaussian package was used to perform all QM calculations in this work for convenience. This can be changed to other faster software packages or more efficient algorithms such as TERACHEM which can run on massively parallel graphics processing unit (GPU).[60,61] The optimization in combining QM methods with the MM code as well as the choice of the *ab initio* method and basis sets can also improve the efficiency of the AIMD method. In addition, the EE-GMFCC method itself can be optimized to increase its computational efficiency. The last one is how to deal with the periodic boundary condition in the AIMD simulation. This is still a developing field and has recently attracted increased attention.[62−64]

As a fragment based approach, the EE-GMFCC method is linear-scaling with a low prefactor and trivially parallelizable. The execution wall time required for this AIMD simulation is independent of the size of the protein if the number of processors increases linearly as the number of protein residues. With continued advance in computer technology, this method will become more and more practical for the AIMD simulation of larger proteins. It is also attractive to apply this method to perform the calculation for structure optimization, vibrational spectrum, or other properties of proteins using high-level *ab initio* methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00558.

Tables S1−S4, coordinates of the initial structure including two fragments and a concap created by the EE-GMFCC method for AIMD in the gas phase; Figure S1, comparison of atomic charges, obtained from both EE-GMFCC and full system QM calculations, from the last snapshot of the AIMD simulation in solution (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: tzhu@lps.ecnu.edu.cn.
*E-mail: xiaohe@phy.ecnu.edu.cn.
*E-mail: zhzhang@phy.ecnu.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grants No. 21433004, 21403068, and 21303057) and Shanghai Putuo District (Grant 2014-A-02). X.H. is also supported by the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20130076120019) and the Fundamental Research Funds for the Central Universities. We thank the Supercomputer Center of East China Normal University for providing us computational time.

■ REFERENCES

(1) Cheatham, T. E., III; Kollman, P. A. Molecular dynamics simulation of nucleic acids. *Annu. Rev. Phys. Chem.* **2000**, *51*, 435−471.
(2) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(3) Karplus, M.; Petsko, G. A. Molecular dynamics simulations in biology. *Nature* **1990**, *347*, 631−639.
(4) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Chandra Singh, U.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(5) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **1986**, *7*, 230−252.
(6) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.
(7) Ji, C.; Mei, Y. Some Practical Approaches to Treating Electrostatic Polarization of Proteins. *Acc. Chem. Res.* **2014**, *47*, 2795−2803.
(8) Duan, L. L.; Mei, Y.; Zhang, D.; Zhang, Q. G.; Zhang, J. Z. H. Folding of a Helix at Room Temperature Is Critically Aided by Electrostatic Polarization of Intraprotein Hydrogen Bonds. *J. Am. Chem. Soc.* **2010**, *132*, 11159−11164.
(9) Tong, Y.; Mei, Y.; Li, Y. L.; Ji, C. G.; Zhang, J. Z. H. Electrostatic Polarization Makes a Substantial Contribution to the Free Energy of Avidin-Biotin Binding. *J. Am. Chem. Soc.* **2010**, *132*, 5137−5142.
(10) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.
(11) Wang, J.; Cieplak, P.; Li, J.; Hou, T.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

Calculations I: Parameterization of Atomic Polarizability. *J. Phys. Chem. B* **2011**, *115*, 3091−3099.
(12) Wang, J.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M.; Lei, H.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* **2011**, *115*, 3100−3111.
(13) Xiao, X.; Zhu, T.; Ji, C. G.; Zhang, J. Z. H. Development of an Effective Polarizable Bond Method for Biomolecular Simulation. *J. Phys. Chem. B* **2013**, *117*, 14885−14893.
(14) Zhu, T.; Xiao, X.; Ji, C.; Zhang, J. Z. H. A New Quantum Calibrated Force Field for Zinc-Protein Complex. *J. Chem. Theory Comput.* **2013**, *9*, 1788−1798.
(15) Lamoureux, G.; MacKerell, A. D.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **2003**, *119*, 5185−5197.
(16) Lamoureux, G.; Roux, B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.
(17) Laasonen, K.; Pasquarello, A.; Car, R.; Lee, C.; Vanderbilt, D. Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 10142−10153.
(18) Remler, D. K.; Madden, P. A. Molecular dynamics without effective potentials via the Car-Parrinello approach. *Mol. Phys.* **1990**, *70*, 921−966.
(19) White, J. A.; Bird, D. M. Implementation of gradient-corrected exchange-correlation potentials in Car-Parrinello total-energy calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *50*, 4954−4957.
(20) Wei, D. Q.; Guo, H.; Salahub, D. R. Conformational dynamics of an alanine dipeptide analog: An ab initio molecular dynamics study. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2001**, *64*, 011907.
(21) Dal Peraro, M.; Raugei, S.; Carloni, P.; Klein, M. L. Solute-solvent charge transfer in aqueous solution. *ChemPhysChem* **2005**, *6*, 1715−1718.
(22) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619−2628.
(23) Isborn, C. M.; Mar, B. D.; Curchod, B. F. E.; Tavernelli, I.; Martinez, T. J. The Charge Transfer Problem in Density Functional Theory Calculations of Aqueously Solvated Molecules. *J. Phys. Chem. B* **2013**, *117*, 12189−12201.
(24) Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Charge Transfer and Polarization in Solvated Proteins from Ab Initio Molecular Dynamics. *J. Phys. Chem. Lett.* **2011**, *2*, 1789−1793.
(25) Collins, M. A.; Bettens, R. P. A. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.* **2015**, *115*, 5607.
(26) 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.
(27) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678−5796.
(28) Collins, M. A.; Cvitkovic, M. W.; Bettens, R. P. A. The Combined Fragmentation and Systematic Molecular Fragmentation Methods. *Acc. Chem. Res.* **2014**, *47*, 2776−2785.
(29) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643−5677.
(30) Pruitt, S. R.; Bertoni, C.; Brorsen, K. R.; Gordon, M. S. Efficient and Accurate Fragmentation Methods. *Acc. Chem. Res.* **2014**, *47*, 2786−2794.
(31) Ramabhadran, R. O.; Raghavachari, K. The Successful Merger of Theoretical Thermochemistry with Fragment-Based Methods in Quantum Chemistry. *Acc. Chem. Res.* **2014**, *47*, 3596−3604.
(32) Li, S.; Li, W.; Ma, J. Generalized Energy-Based Fragmentation Approach and Its Applications to Macromolecules and Molecular Aggregates. *Acc. Chem. Res.* **2014**, *47*, 2712−2720.
(33) Fedorov, D. G.; Asada, N.; Nakanishi, I.; Kitaura, K. The Use of Many-Body Expansions and Geometry Optimizations in Fragment-Based Methods. *Acc. Chem. Res.* **2014**, *47*, 2846−2856.
(34) Gao, J.; Truhlar, D. G.; Wang, Y.; Mazack, M. J. M.; Loeffler, P.; Provorse, M. R.; Rehak, P. Explicit Polarization: A Quantum Mechanical Framework for Developing Next Generation Force Fields. *Acc. Chem. Res.* **2014**, *47*, 2837−2845.
(35) He, X.; Zhu, T.; Wang, X.; Liu, J.; Zhang, J. Z. H. Fragment Quantum Mechanical Calculation of Proteins and Its Applications. *Acc. Chem. Res.* **2014**, *47*, 2748−2757.
(36) Xu, X.; Nakatsuji, H.; Ehara, M.; Lü, X.; Wang, N. Q.; Zhang, Q. E. Cluster modeling of metal oxides: the influence of the surrounding point charges on the embedded cluster. *Chem. Phys. Lett.* **1998**, *292*, 282−288.
(37) 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.
(38) Wang, X.; Liu, J.; Zhang, J. Z. H.; He, X. Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps Method for Full Quantum Mechanical Calculation of Protein Energy. *J. Phys. Chem. A* **2013**, *117*, 7149−7161.
(39) Zhang, D. W.; Zhang, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J. Chem. Phys.* **2003**, *119*, 3599−3605.
(40) He, X.; Zhang, J. Z. H. The generalized molecular fractionation with conjugate caps/molecular mechanics method for direct calculation of protein energy. *J. Chem. Phys.* **2006**, *124*, 184703.
(41) 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.
(42) 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.
(43) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(44) 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.
(45) Brorsen, K. R.; Minezawa, N.; Xu, F.; Windus, T. L.; Gordon, M. S. Fragment Molecular Orbital Molecular Dynamics with the Fully Analytic Energy Gradient. *J. Chem. Theory Comput.* **2012**, *8*, 5008−5012.
(46) Brorsen, K. R.; Zahariev, F.; Nakata, H.; Fedorov, D. G.; Gordon, M. S. Analytic Gradient for Density Functional Theory Based on the Fragment Molecular Orbital Method. *J. Chem. Theory Comput.* **2014**, *10*, 5297−5307.
(47) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(48) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157−167.
(49) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.;

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905

Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*; Gaussian, Inc.: Wallingford, CT, USA, 2009.

(50) Pal, S. K.; Peon, J.; Zewail, A. H. Biological water at the protein surface: Dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1763−1768.

(51) Levy, Y.; Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.

(52) Takenaka, N.; Kitamura, Y.; Koyano, Y.; Nagaoka, M. An improvement in quantum mechanical description of solute-solvent interactions in condensed systems via the number-adaptive multiscale quantum mechanical/molecular mechanical-molecular dynamics method: Application to zwitterionic glycine in aqueous solution. *J. Chem. Phys.* **2012**, *137*, 024501.

(53) Tu, Y. Q.; Laaksonen, A. On the effect of Lennard-Jones parameters on the quantum mechanical and molecular mechanical coupling in a hybrid molecular dynamics simulation of liquid water. *J. Chem. Phys.* **1999**, *111*, 7519−7525.

(54) Hu, L.; Soederhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761−777.

(55) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.

(56) Li, Y.; Gao, Y.; Zhang, X.; Wang, X.; Mou, L.; Duan, L.; He, X.; Mei, Y.; Zhang, J. Z. H. A coupled two-dimensional main chain torsional potential for protein dynamics: generation and implementation. *J. Mol. Model.* **2013**, *19*, 3647−3657.

(57) Gao, Y.; Li, Y.; Mou, L.; Hu, W.; Zheng, J.; Zhang, J. Z. H.; Mei, Y. Coupled Two-Dimensional Main-Chain Torsional Potential for Protein Dynamics II: Performance and Validation. *J. Phys. Chem. B* **2015**, *119*, 4188−4193.

(58) Dudev, T.; Devereux, M.; Meuwly, M.; Lim, C.; Piquemal, J.-P.; Gresh, N. Quantum-Chemistry Based Calibration of the Alkali Metal Cation Series (Li+-Cs+) for Large-Scale Polarizable Molecular Mechanics/Dynamics Simulations. *J. Comput. Chem.* **2015**, *36*, 285−302.

(59) Sakharov, D. V.; Lim, C. Zn protein simulation including charge transfer and local polarization effects. *J. Am. Chem. Soc.* **2005**, *127*, 4921−4929.

(60) Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222−231.

(61) Akimov, A. V.; Prezhdo, O. V. Large-Scale Computations in Chemistry: A Bird's Eye View of a Vibrant Field. *Chem. Rev.* **2015**, *115*, 5797−5890.

(62) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. Multipolar Ewald Methods, 1: Theory, Accuracy, and Performance. *J. Chem. Theory Comput.* **2015**, *11*, 436−450.

(63) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. Multipolar Ewald Methods, 2: Applications Using a Quantum Mechanical Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 451−461.

(64) Lu, X.; Cui, Q. Charging Free Energy Calculations Using the Generalized Solvent Boundary Potential (GSBP) and Periodic Boundary Condition: A Comparative Analysis Using Ion Solvation and Oxidation Free Energy in Proteins. *J. Phys. Chem. B* **2013**, *117*, 2005−2018.

DOI: 10.1021/acs.jctc.5b00558
J. Chem. Theory Comput. 2015, 11, 5897−5905