JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Adaptive-Partitioning QM/MM Dynamics Simulations: 3. Solvent Molecules Entering and Leaving Protein Binding Sites

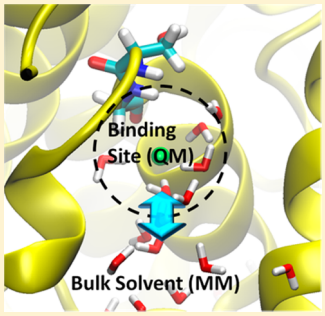
Soroosh Pezeshki,*[,†] Christal Davis,[†] Andreas Heyden,[‡] and Hai Lin*[,†]

[†]Chemistry Department, University of Colorado Denver, Denver, Colorado 80217-3364, United States

[‡]Department of Chemical Engineering, University of South Carolina, Columbia, South Carolina 29208, United States

Supporting Information

ABSTRACT: The adaptive-partitioning (AP) schemes for combined quantum-mechanical/molecular-mechanical (QM/MM) calculations allow on-the-fly reclassifications of atoms and molecules as QM or MM in dynamics simulations. The permuted-AP (PAP) scheme (*J. Phys. Chem. B* **2007**, *111*, 2231) introduces a thin layer of buffer zone between the QM subsystem (also called active zone) and the MM subsystem (also known as the environmental zone) to provide a continuous and smooth transition and expresses the potential energy in a many-body expansion manner. The PAP scheme has been successfully applied to study small molecules solvated in bulk solvent. Here, we propose two modifications to the original PAP scheme to treat solvent molecules entering and leaving protein binding sites. First, the center of the active zone is placed at a pseudoatom in the binding site, whose position is not affected by the movements of ligand or residues in the binding site. Second, the extra forces due to the smoothing functions are deleted. The modified PAP scheme no longer describes a Hamiltonian system, but it satisfies the conservation of momentum. As a proof-of-concept experiment, the modified PAP scheme is applied to the simulations under the canonical ensemble for two binding sites of the *Escherichia coli* CLC chloride ion transport protein, in particular, the intracellular binding site S_{int} discovered by crystallography and one putative additional binding site S_{add} suggested by molecular modeling. The exchange of water molecules between the binding sites and bulk solvent is monitored. For comparison, simulations are also carried out using the same model system and setup with only one exception: the extra forces due to the smoothing functions are retained. The simulations are benchmarked against conventional QM/MM simulations with large QM subsystems. The results demonstrate that the active zone centered at the pseudo atom is a reasonable and convenient representation of the binding site. Moreover, the transient extra forces are non-negligible and cause the QM water molecules to move out of the active zone. The modified PAP scheme, where the extra forces are excluded, avoids the artifact, providing a realistic description of the exchange of water molecules between the protein binding sites and bulk solvent.



I. INTRODUCTION

The adaptive algorithms for molecular dynamics (MD) simulation based on combined quantum-mechanical and molecular-mechanical (QM/MM)[1−16] calculations have received much attention recently. The adaptive algorithms permit on-the-fly reclassifications of atoms and/or groups as part of the QM or MM subsystems, which is not possible in conventional QM/MM setups. The adaptive algorithms thus automate the updates of the QM subsystem, which is of our primary interest, in complex and dynamical environments. This is important if the region of interest may exchange atoms or groups (e.g., solvent molecules) with the surroundings on the time scale of the simulation. With conventional QM/MM, one may have to include all solvent molecules as QM to be on the safe side. Because the QM subsystem can be kept as small as possible, adaptive algorithms hold the promise of using high-level QM methods and long simulation times, which can potentially lead to new insights.

The challenges of redefining an atom or molecule as QM or MM during an MD run are that the energy and forces will change abruptly, which can lead to not only numerical instabilities but also ambiguities in the result interpretations. To overcome those difficulties, adaptive algorithms introduce a buffer zone between the QM and MM subsystems to provide a smooth transition for the exchange of atoms or molecules between the two subsystems. The atoms and molecules in the buffer zone are often referred to as buffer groups, and they have dual QM and MM characteristics. As such, sometimes the QM subsystem is called the active zone and the MM subsystem the environmental zone to avoid possible confusions. Although not required, the active zone is often taken to be a sphere of radius R_{min}, and the buffer zone is configured to be a spherical shell of inner radius R_{min} and outer radius R_{max} (Figure 1). Depending on the applications of smoothing functions to energy or to forces, the adaptive algorithms can be grouped into two categories, that is, the energy-based and the force-based adaptive algorithms.

The energy-based algorithms apply smoothing functions to potential energy and compute forces as derivatives of the smoothed potential energy. Examples include the ONIOM-exchange of solvent (ONIOM-XS) algorithm by Kerdcharoen

Received: June 28, 2014

dx.doi.org/10.1021/ct500553x | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

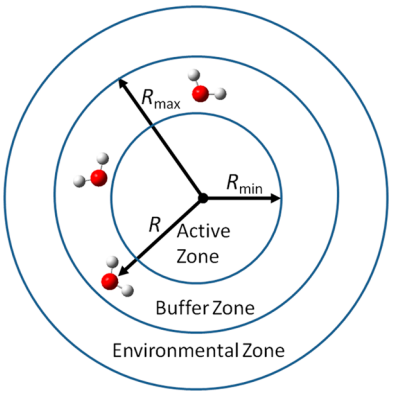


Figure 1. Buffer zone between the active zone treated by QM and the environmental zone treated by MM. A molecule will usually be considered in the buffer zone if its distance R to the center of the active zone satisfies $R_{min} < R < R_{max}$, where R_{min} and R_{max} are two predefined thresholds, and the molecule will be called a buffer group. Smoothing functions are applied to avoid the abrupt changes in the energy, or forces, or both when buffer groups are entering, leaving, and/or moving within the buffer zone.

and Morokuma,[17,18] and permuted adaptive-partitioning (PAP) and sorted adaptive-partitioning (SAP) schemes by Heyden et al.[19] with a recent extension by Pezeshki and Lin,[20] and the difference-based adaptive solvation (DAS) method by Bulo et al.[21,22] The force-based algorithms manipulate forces by applying smoothing functions to the forces or by simply assigning MM forces to the buffer groups. However, energy is not defined in the force-based algorithms. The methods in this category include the "hot-spot" method by Rode and co-workers,[23] the "bookkeeping" algorithm by Bulo et al.,[21] and the buffered-force scheme by Bernstein and co-workers.[24]

Whether the energy-based or the force-based algorithms are better is currently under debate. The energy-based schemes such as PAP and DAS are in principle more accurate, because they conserve both energy and momentum, and the conservations of energy and momentum are fundamental checks on the rigorousness of a given algorithm for MD simulations. However, the calculations using the energy-based schemes are often more expensive. Moreover, the extra forces due to the smoothing functions may not be negligible and can potentially cause artifacts. The extra forces can in principle be removed or minimized by aligning the QM and MM potentials but that is very difficult to achieve for polyatomic molecules because of the complicated dependence of the interaction energies on the relative orientations of one molecule with another. The force-based methods, on the other hand, are often simpler to implement and less expensive to apply. Moreover, they avoid the extra-force problem. However, the force-based algorithms lack a well-defined energy, making the interpretation of simulation data somewhat ambiguous, although it has been argued[25] that the force-based schemes can be used with free-energy techniques that require only forces or trajectories, for example, umbrella integration,[26] constrained dynamics,[27] meta-dynamics,[28] and adaptive biasing force.[29] It had been proposed that a bookkeeping term[21] obtained by path-integration over the artificial forces could be used as a correction to provide a conservation quantity; however, it was pointed out[30,31] that the use of the path-dependent bookkeeping term does not restore the conservation of energy. Furthermore, some of the force-based schemes, such as the hot-spot[23] and the buffered-force[24] schemes, do not conserve momentum. It is, therefore, not surprising that massive thermostats are required to stabilize the hot-spot and buffer-force simulations; otherwise, the system will heat up quickly. Comparisons of the energy-based and force-based adaptive schemes have been reported in the literature,[19,32] and it seems that the performance of many adaptive algorithms also depends to some extent on the specific QM level of theory and MM force fields that had been selected.[32] Clearly, there is an urgent need for further exploration and development of rigorous and efficient adaptive schemes.

The present contribution is the third step of our systematic development of the adaptive-partitioning schemes. In 2007, we proposed the PAP and SAP schemes and applied them to the solvation of ions and small molecules.[19] Recently, we extended the PAP scheme to treat large molecules and polymers, for which relocation of the QM/MM boundary passing through covalent bonds were automated on-the-fly.[20] However, our early implementations must place the active-zone center at a preselected atom or the center of mass of a set of preselected atoms; such a requirement may not be convenient when monitoring the exchange of atoms or groups of atoms between a protein binding site and bulk solvent, because the active-zone location will be affected by the diffusion of the ligand (if the active-zone center is associated with the ligand) or the movements of certain residues (if the residues are part of the active-zone center definition). It is desirable to find a new way to locate the active-zone center that is not affected by the movements of the ligand and residues. This issue will be addressed in the present contribution.

The second goal in this work is to modify the original PAP scheme to remove the extra forces due to the smoothing functions, which are proportional to the difference between the QM and MM potentials. In principle, keeping the energy interpolation is the best, because it permits having a Hamiltonian system. However, the key challenge of the original PAP scheme is to properly align the potentials of the groups in QM and MM system, such that the extra forces become negligible. For interactions between monatomic molecules and for systems that have very similar QM and MM potentials, the alignment can be done relatively easily.[19] However, when polyatomic species are involved, in general, the QM and MM levels of theory give very different multidimensional potentials for the same structures, complicating the alignment. Alignment of the potentials along one coordinate may not necessarily be meaningful for along another coordinate. Non-negligible extra forces may introduce artifacts in simulations. To deal with the extra forces in general, we propose here a modification to the original PAP algorithm. In contrast to the energy-based original PAP scheme, the modified PAP method is force-based. Although the modified PAP scheme does not describe Hamiltonian systems and cannot be applied in simulations under the microcanonical ensemble, it preserves the conservation of momentum, and it can be used for simulations under other ensembles.

The present contribution addresses the new way of defining active-zone centers and the modification for the extra forces. The treatments are introduced in section II. In section III, the new scheme was tested in the simulations of two binding sites of the *Escherichia coli* CLC chloride ion transport protein. The results and discussion are presented in section IV.

II. METHODOLOGY

II.A. Original PAP Scheme. In the present study, we have employed the PAP scheme for its superior numerical stability, but the new development can in principle be applied to SAP as well. The original PAP algorithm has been detailed in ref 19, and here, we will provide a brief outline only.

In a system with N buffer groups, the potential energy of the system is defined as

$$V = V^{A} + \sum_{i=1}^{N} P_i(V_i^{A} - V^{A}) + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} P_iP_j\left(V_{i,j}^{A} - \left[V^{A} + \sum_{r=i,j}(V_r^{A} - V^{A})\right]\right) + \sum_{i=1}^{N-2}\sum_{j=i+1}^{N-1}\sum_{k=j+1}^{N} P_iP_jP_k\left(V_{i,j,k}^{A} - \left(V^{A} + \sum_{r=i,j,k}(V_r^{A} - V^{A}) + \sum_{(p,q)=(i,j),(i,k),(j,k)}\left(V_{p,q}^{A} - \left(V^{A} + \sum_{r=(p,q)}(V_r^{A} - V^{A})\right)\right)\right)\right) + \ldots \quad (1)$$

where V^{A} is the energy determined with the groups in the active zone at the QM level; V_i^{A} with all active-zone groups and the i-th buffer-zone group at the QM level; $V_{i,j}^{A}$ with all active-zone groups, the i-th buffer-zone group, and the j-th buffer-zone group at the QM level;... $V_{1,2\cdots N}^{A}$ with all active-zone groups and all N buffer-zone groups at the QM level; and P_i the smoothing function of the i-th buffer-zone group in terms of the dimensionless reduced radial coordinate α_i:

$$P_i(\alpha_i) = \begin{cases} 0 & 1 < \alpha_i \\ -6\alpha_i^{5} + 15\alpha_i^{4} - 10\alpha_i^{3} + 1 & 0 < \alpha_i \le 1 \\ 1 & \alpha_i \le 0 \end{cases} \quad (2)$$

$$\alpha_i = \frac{R_i - R_{\min}}{R_{\max} - R_{\min}} \quad (3)$$

where R_i is the distance from the active-zone center to the i-th buffer-zone group, $R_{\min}$ and $R_{\max}$ are the inner and outer radii of the buffer zone, respectively. In total, 2^N QM calculations are to be performed. While this may seem quite expensive, the energy contributions of the terms in the series in eq 1 decrease rapidly, and truncation of the series is recommended to significantly reduce the number of embedded-QM calculations. The resultant small discontinuities in the energy and derivatives are controllable and have been shown to be insignificant if the series is truncated at the fifth order. That is, at no point in time is a QM/MM calculation performed with more than the active zone and 5 buffer zone groups treated at the QM level. Moreover, the calculations of $V_{i,j\ldots r}^{A}$ are parallel in nature, and thus, in principle, the wall clock is only limited by the calculation with the largest number of buffer groups.

II.B. Defining Active Zone in a Protein Binding Site. A protein binding site is usually a cavity surrounded by a few residues, where ligands can enter, stay, and leave. In general, the binding site is treated at the QM level and is configured to be the active zone. In practice, however, whether or not a particular ligand or residue in the active zone (and buffer zone) will be treated at the QM level depends on the specific question posed. If the ligand or residue does not need to be treated at the QM level of theory, it can be fixed as part of the environmental zone, regardless of how far it is from the active-zone center.

The original PAP scheme would have to place the active-zone center at an atom (or the center of mass of a number of atoms) of the ligand in the cavity and/or the residues that embrace the cavity. As mentioned earlier, this may present difficulty when monitoring solvent or ligand entering and leaving the binding site, because the location of the active zone will be affected by the movements of the atoms that are used to define the active-zone center. To fix the active zone in the space, we propose to use a pseudoatom X of sufficiently large mass as the center of the active zone. The pseudoatom will not interact with any QM, MM, or buffer groups. The forces acting on the pseudoatom are solely the extra forces due to the smoothing function. Because the pseudoatom has a large mass, it gains negligible acceleration and velocity, and its position is essentially unchanged during the simulation. As such, it will not be affected by the motions of the ligands (entering or leaving the binding site) or residues (flipping in or out of the binding site). We have tested different values for the pseudoatom mass parameter and found that 10^6 amu was a reasonable choice, which has led to negligible changes in the pseudoatom position during the simulations of the protein model that we employed without causing numerical difficulties in the calculations.

II.C. Removal of Extra Forces Due to Smoothing Functions. We propose to modify the original PAP algorithm by neglecting the extra forces due to the smoothing functions. Those extra forces act on the buffer groups and on the atoms that define the active-zone center. The use of a pseudoatom for the active center removes the extra forces on the atoms that are involved in the active-zone center definition but does not change the extra forces on the buffer groups. Discarding the extra forces does not interrupt the conservation of momentum, because Newton's Third Law of Motion is still obeyed. (The bookkeeping scheme by Bulo et al.[21] is another force-based method that conserves momentum.) However, the modified forces are not exact differentials and cannot be expressed as the derivatives of a potential. Deleting the extra forces is equivalent to adding *external* forces on the system to cancel out the extra forces due to the smoothing functions. Clearly, this leads to a non-Hamiltonian system. The treatment does provide a smooth interpolation of the forces acting on *all* involved QM, MM, and buffer group atoms. Note that this is different from other force-based adaptive algorithms where only the forces on certain atoms (usually the buffer group atoms) are smoothed; in those methods,[23,24] the momentum are not conserved because Newton's Third Law of Motion is violated. With the modified PAP scheme, the model system is coupled to a thermostat so that *NVT* simulations can be performed.

We note that our treatment differs from the "bookkeeping" scheme[21] in that we do not include an extra "pseudo-energy" term obtained by path integral along the propagated trajectory over the extra forces due to smoothing functions. In the bookkeeping scheme, the extra pseudoenergy term was used as correction to the potential energy. However, it has been found that the bookkeeping term is path-dependent,[21] and its inclusion does not restore the conservation of energy.[30,31]

III. COMPUTATION

The modified PAP method was tested by MD simulations on two binding sites of the *Escherichia coli* CLC Cl^- ion transport protein, EcCLC. The CLC family of transport proteins[33]

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

translocate Cl^- ions across membranes, and they are associated with a variety of critical physiological and cellular processes such as neuroexcitation, cell-volume regulation, organic solute transport, and muscle contraction.[34−38] The crystal structures[39,40] of the transmembrane domain of EcCLC and its E148Q mutant revealed three Cl^- binding sites: the extracellular binding site S_{ext}, the central binding site S_{cen}, and the intracellular binding site S_{int}. Computational studies also suggested an additional binding site S_{add}.[41−47] Of particular interest to us are the two binding sites S_{int} and S_{add}, which are located near the two ends of the pore. They are connected to bulk solvent through the vestibules. The S_{add} site is more open to bulk solvent than S_{int}. The possible exchanges of water molecules and ions between those binding sites and bulk solvent provide an excellent test for the modified PAP scheme.

A partially solvated droplet model of the transmembrane domain of the EcCLC protein (Figure 2) was employed in the test calculations. The crystal structure of the transmembrane domain (PDB code 1OTS[40]) contains two identical subunits, each of which forms a pore that functions independently.[39,40] As such, only one of the subunits (chain A) was used to construct the model system. The protonation states of the histidine residues were assigned according to the electrostatic calculations in ref 42. The gating residue E148 was protonated, and its side chain was manually rotated to align with the side chain of the gating residue in the E148Q mutant, whose crystal structure (PDB code 1OTU[40]) was deemed representing the open state of the channel.[39,40] Two water droplets (density = 1.0 g/mL) were added at the intracellular and extracellular vestibules of the protein. The water molecules were separated into an inner layer and an outer layer. Those in the outer layer, that is, with a distance to the center of the droplet greater than 10 Å, were restrained by applying harmonic potentials (force constant of 10 kcal/mol·Å^2) to the oxygen atoms. Those in the inner layer, which have distances to the center of the droplet equal to or smaller than 10 Å, were free to move in the simulations. The restraints on the outer layer water molecules prevented water molecules from evaporating into vacuum. Six Cl^- ions were added to the system. In particular, three ions were placed in the experimentally determined binding sites S_{int}, S_{cen}, and S_{ext}, one ion at the putative additional binding site S_{add}[41−46,48] near residue R147 suggested by computations, and one ion at a random location within each water droplet (with the distance <10 Å from the droplet center). In total, there were 9606 atoms, including 926 water molecules. The total charge of the model system was +3 e. The Cl^- ions at S_{int}, S_{cen}, S_{ext}, and S_{add} are denoted Cl_{int}, Cl_{cen}, Cl_{ext}, and Cl_{add}, respectively.

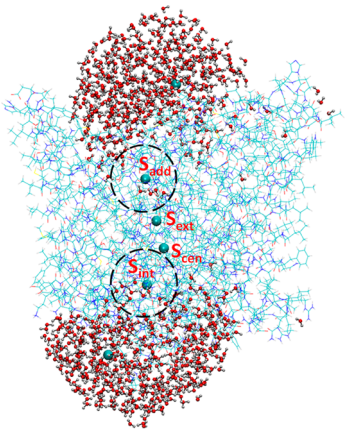


Figure 2. Partial solvation model for EcCLC chloride transport protein, where the protein is shown as lines, water as balls and sticks, and Cl^- ions as spheres in cyan. S_{int}, S_{cen}, and S_{ext} indicate the experimentally revealed intracellular, central, and extracellular binding sites, respectively.[39,40] The additional binding site suggested by computations is S_{add}.[41−46,48] The dashed circles indicate the active zones centered at S_{int} and S_{add} in the adaptive-partitioning QM/MM simulations.

The system was first equilibrated for 100 ps at the MM level using the CHARMM22[49] force fields for the protein and the TIP3P[50] model for water. Next, two sets of MD simulations at the QM/MM level employing the modified PAP scheme were performed. In one set of simulations, the pseudoatom designated as the center of the active zone was set to the position of the Cl^- ion after the MM equilibration in S_{int}, and in the other set of simulations, it was set to the position of the Cl^- ion after the MM equilibration in S_{add}. Accordingly, the pseudoatoms were labeled as X_{int} and X_{add} in the S_{int} and S_{add} simulations, respectively. The buffer zones were 0.5 Å thick with R_{min} = 6.0 Å and R_{max} = 6.5 Å. Each water molecule (or ion) was considered as one group in the adaptive partitioning. Because our goal was to monitor the exchange of molecules between the binding sites and bulk solvent, for simplicity, we chose the mechanical embedding scheme and described all residues at the MM level. The semiempirical PM3[51,52] model was chosen for the QM calculations due to its high computational efficiency. The CHARMM22/TIP3P combination was again used for the involved MM calculations. Each QM/MM simulation ran for 50 ps. Both the MM and QM/MM simulations were done under an *NVT* ensemble at 300 K. The time step was set to 1.0 fs. A cut off of 14 Å was used for the nonbonded interactions, with a linear tapering switched on at 13 Å. Periodic boundaries were not used for the droplet model. A Nosé−Hoover[53,54] thermostat was employed with a coupling constant of 4.0 fs. All calculations were carried out using a new version of the QMMM[55] program, which called TINKER[56] for MM gradient calculations and MNDO[57] for QM gradient calculations. The QMMM program then synthesized the MM and QM gradients and propagated the QM/MM trajectory. The results obtained by the modified PAP scheme are labeled Cal1.

To explore what differences the extra forces due to smoothing functions will bring up, we also carried out simulations under the same conditions but with the extra forces retained on the pseudoatoms and buffer groups. Each pseudoatom was assigned a large mass (10^6 amu), and its position was propagated similar to an ordinary atom (according to the extra forces acting on it). The large mass ensured that the pseudoatom stayed essentially at its original location and did not move out of the active zone during the simulations. The simulation results with the extra forces applied are labeled Cal2.

In addition, we carried out conventional QM/MM simulations with *large* QM subsystems using the same protein model. Two simulations were performed, one for S_{int} and the other for S_{add}, employing the same setup as that in adaptive QM/MM except for the selections and treatments of the QM atoms. For the conventional QM/MM simulation of S_{int}, the QM subsystem contained Cl_{int} as well as the Cl^- ion and all water molecules in the inner layer of the water droplet at the intracellular vestibule. For the simulation of S_{add}, included in the QM subsystem are Cl_{add} as well as the Cl^- ion and all water

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

molecules in the inner layer of the water droplet at the extracellular vestibule. To facilitate comparisons with the adaptive QM/MM, we defined Zone 1 in conventional QM/MM to be a sphere of R = 6.0 Å centered at the initial position of the Cl^- ion (Cl_{int} in the S_{int} simulation and Cl_{add} in the S_{add} simulation). Zone 1 is the counterpart in conventional QM/MM to the active zone in adaptive QM/MM. Similarly, we defined Zone 2 to be a spherical shell centered at the initial position of the Cl^- ion with the inner and outer radii of 6.0 and 6.5 Å. Zone 2 is the counterpart to the buffer zone in adaptive QM/MM. The results in the conventional QM/MM simulations with large QM subsystem are labeled QM and serve as the reference to benchmark the new PAP method.

IV. RESULTS AND DISCUSSION

IV.A. Extra Forces Due to Smoothing Functions. As mentioned above, the extra forces due to smoothing functions are one challenge in the energy-based adaptive QM/MM algorithms. The extra forces are basically determined by two factors: the gradient of the smoothing function and the difference between the energies $V^A_{i,j...r}$ and V^A. This point is the clearest when there is only one buffer group present, for which eq 1 reduces to

$$V = V^A + P_1(V_1^A - V^A) \quad (4)$$

The buffer group is treated by MM in the calculations of V^A and by QM in the calculations of $V_1{}^A$, respectively. The gradient of the smoothing function P_1 depends on the actual form of the smoothing function, but it should be in the order of 1/Å, because the smoothing function varies from 0 to 1 from R_{min} to R_{max}, and $R_{max} - R_{min} \leq 1$ Å is often the choice to limit the number of buffer groups. The potential energies V_1^A and V^A have been aligned such that they agree with each other at $R_1 = \infty$. A few kcal/mol difference between V_1^A and V^A are not surprising at $R_{min} < R_1 < R_{max}$, leading to extra forces of 10−20 kcal/mol·Å, which are too large to be neglected. The extra forces will act on the buffer group as well as the atoms that are used to define the active-zone center (the pseudoatom X in this work). When more than one buffer group is present, the net extra forces acting on the active-zone center will be the vector sum of the individual extra forces.

As a numerical check on the extra force, we plot the norm of the net extra forces acting on the pseudoatoms in Figure 3a for the S_{int} simulation and in Figure 3b for the S_{add} simulation. Only the Cal2 data are shown, since those forces were removed in Cal1. The extra forces were sizable in magnitude, usually around 10 kcal/mol/Å and occasionally spiking up to 60 kcal/mol·Å. Although the extra forces were only transient, their effects on a particular dynamical property of interest were not negligible, as revealed in further analysis below. Alignment of V_1^A and V^A near $R_1 = R_{min}$ or R_{max} should be able to significantly reduce the extra forces, but that is very challenging to achieve due to the complication brought by the dipole orientations of water molecule, and we are not aware of any simple method for doing the potential energy alignment in such a case. Another way of reducing the extra forces is to increase the thickness of the buffer zone ΔR ($\Delta R = R_{max} - R_{min}$). However, this method is of limited use in practice. Note that the extra forces decrease as $(\Delta R)^{-1}$, whereas the number of buffer groups increases as ΔR^3. The fast-growing computational costs can quickly make the calculations too expensive to be feasible, limiting the applicability of this approach.

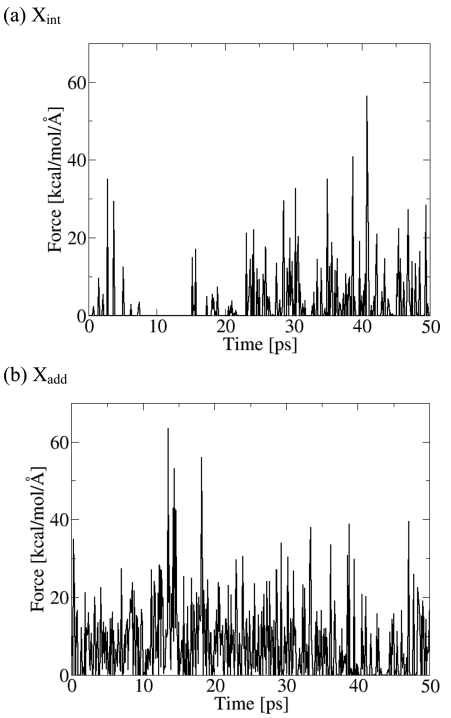


Figure 3. Norm of net extra forces due to smoothing functions on the pseudoatoms X for the (a) S_{int} and (b) S_{add} simulations. Only the Cal2 (calculations with those extra forces) results are shown.

IV.B. Active Zone Center. The Cl^- at the binding site S_{int} is coordinated by two backbone N−H groups of residues S107 and G108. Figure 4(a and b) shows the distances between X_{int} and the two N atoms for the simulation of S_{int} in Cal1 and Cal2, respectively. Overall, the locations of X_{int} were very stable in both calculations, about 3.8 Å away from the N atoms. The stability is expected due to the backbone rigidity of S107 and G108. Although the extra forces due to smoothing functions caused X_{int} to accelerate in Cal2, the velocity gained by X_{int} was negligible during the simulation owing to its assigned large mass (see Figure S3a in the Supporting Information). For comparison, we have plotted the distances between Cl_{int} and the two N atoms of S107 and G108 in Figure 4c and d for Cal1 and Cal2, respectively. Although Cl_{int} was kept near the initial position during the simulations, larger fluctuations can be seen in the Cl_{int} plots than in the X_{int} plots, with the distances spiking up to 5 Å. The larger variations in the distances are partly due to perturbations by the motions of water molecules nearby, which are very flexible as compared with the S107 and G108 backbones. The observations above suggested that X_{int} is a faithful representation of the binding site S_{int}.

The Cl^- ion at the binding site S_{add} is substantially stabilized by the attraction due to the positively charged residue R147. The distances between X_{add} and the central C atom and the closest N atom of the guanidinium cation for the simulations of S_{add} are given in Figure 4e and f for Cal1 and Cal2, respectively, and the corresponding distances for Cl_{add} are provided in Figure 4g and h, respectively. In contrast to the rigid backbones of S107 and G108, the R147 side chain was more flexible, with which Cl_{add} trended to move along. The fact that S_{add} is more open to bulk solvent than S_{int} is also made the position of Cl_{add} less stable because it was more prone to the perturbations by water molecules. Because X_{add} always stayed at its initial position, larger fluctuations were revealed in the X_{add} plots than

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

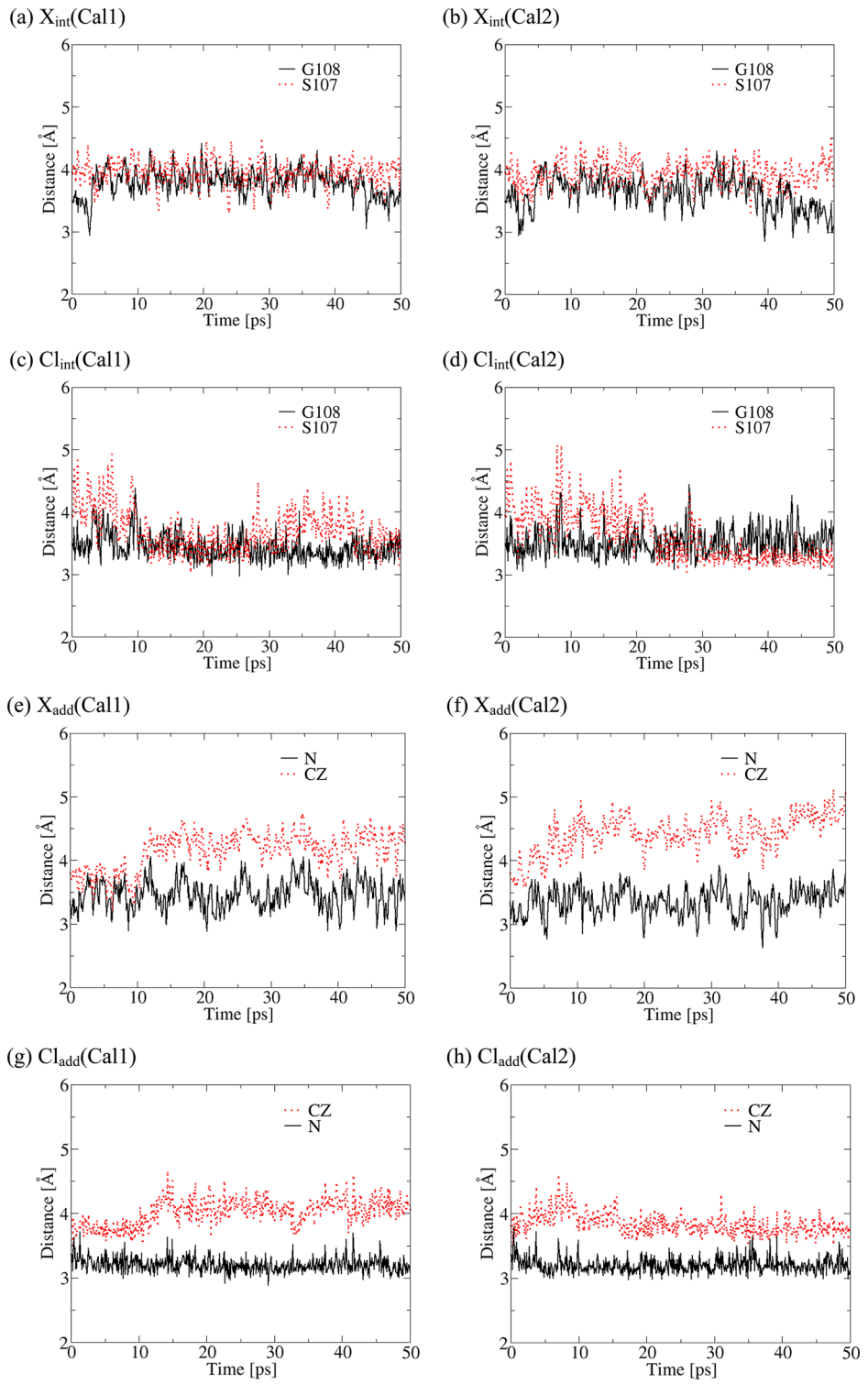


Figure 4. Distances from the pseudoatoms X or the Cl^- ions to selected heavy atoms of the residues that coordinate the Cl^- ions in the binding sites (the backbone N atoms of S107 and G108 in the S_{int} simulations and the central CZ and the closest N atom of the guanidinium cation of R147 in the S_{add} simulations, respectively). Cal1 denotes the calculations without the extra forces due to smoothing functions, and Cal2 denotes the calculations with those extra forces.

in the Cl_{add} plots. Interestingly, the X_{add} plot for Cal1 showed the same trend as the Cl_{add} plot, including the increase of the distance to the CZ atom from 3.8 to 4.3 Å due to the R147 side chain rotation near 12 ps. As opposed to S_{int}, the S_{add} binding site was not revealed in crystallographic experiments but was suggested by computations; the theoretical studies all predicted that S_{add} was near R147, but they differed from each other in the exact location of S_{add}.[41−48] In such a case, X_{add} serves as a convenient approximation for the binding site, because its position will not be affected by the movements of flexible side chains.

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

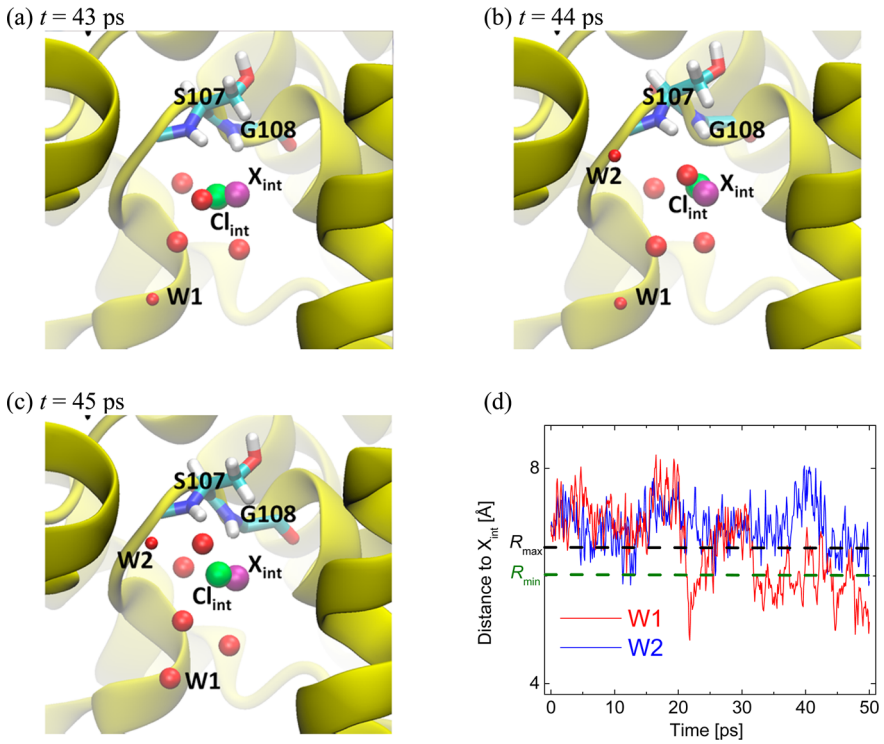


Figure 5. (a–c) Snapshots from the S_{int}(Cal1) simulations. The protein is shown as cartoon in yellow, with S107 and G108 in licorice (N, blue; C, cyan; O, red; and H, white). Cl_{int} and X_{int} are indicated as green and purple spheres, respectively, and water in the active and buffer zones as large and small red spheres, respectively. Water molecule W1 moved from the buffer zone into the active zone after t = 45 ps, whereas W2 was in the environmental zone at t = 43 ps (and not shown in a) but diffused into the buffer zone after t = 44 ps. (d) Distances from X_{int} to W1 and from X_{int} to W2 over the simulation. The black dashed line indicates the boundary between the environmental and buffer zones, and the green dashed line between the buffer and active zones.

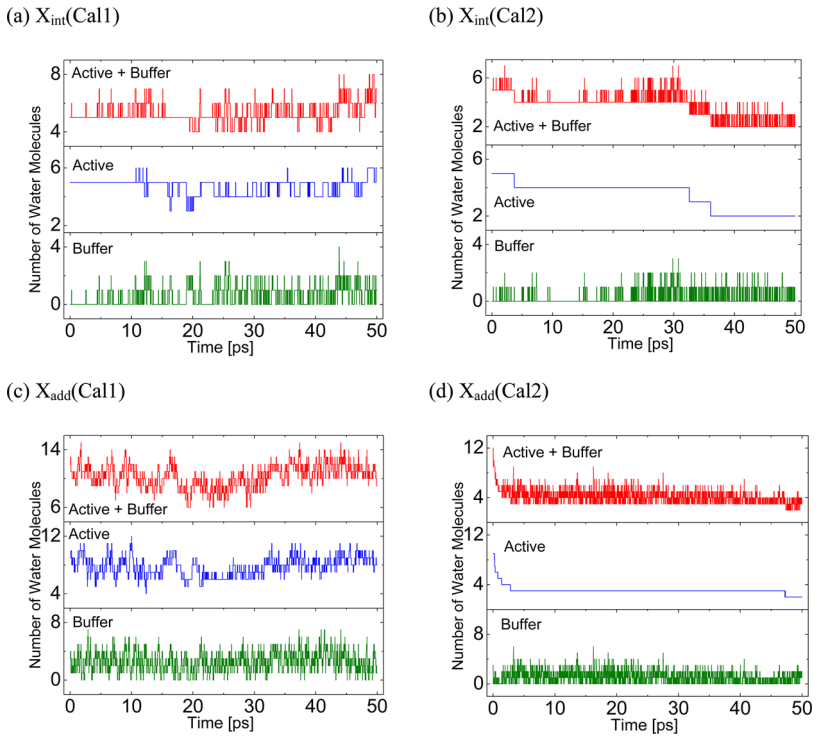


Figure 6. Numbers of water molecules in the active zone (blue) and buffer zone (green) during the MD simulations. X_{int} indicates the S_{int} simulations, and X_{add} the S_{add} simulations. Cal1 and Cal2 denote, respectively, the calculations without and with the extra forces due to smoothing functions.

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX–XXX

The differences between X_{int} and X_{add} can also be seen in the radial distribution functions (RDF) for the X_{int}−Cl_{int} and X_{add}−Cl_{add} pairs, which are shown in Figure S4 in the Supporting Information. Apparently, the Cl^- ions wandered around the corresponding pseudoatoms during the 50 ps simulations. Whereas Cl_{int} stayed rather close to X_{int}, as the density dropped to 0 at r = 1.5 Å, Cl_{add} could travel much farther away from X_{add}, as the distribution extended to r > 2.5 Å.

IV.C. Exchanges of QM and MM Solvent Molecules. Although both Cal1 and Cal2 stabilized the location of the active zone through the use of a pseudoatom as the active-zone center, they led to different pictures for solvent molecules entering and leaving the binding sites. First, we look at the Cal1 results. Visual inspection of the saved trajectories confirmed the exchange of water molecules between the protein binding site and bulk solvent. As an example, three snapshots from the S_{int} simulations are shown in Figure 5a−c. The time intervals between the snapshots are 1 ps. It can be seen that the Cl^- ion was wandering around X_{int}, while the two residues, S107 and G108, underwent only minor movements during the 2 ps time interval. Moreover, one water molecule (W1) entered the active zone from the buffer zone, and another water molecule (W2) drifted into the buffer zone from the environmental zone. Consequently, the number of QM water molecules increased from 4 at t = 43 ps to 5 at t = 45 ps. Further analysis of the locations of W1 and W2 is given in Figure 5d, which displays the distances from X_{int} to W1 and from X_{int} to W2 over the simulation. Both W1 and W2 were initially located in the environmental zone near the boundary between the buffer and environmental zones, and they visited all three (environmental, buffer, and active) zones many times in the following 50 ps. Whereas W2 remained basically near the environmental-buffer boundary, W1 gradually diffused into the active zone at the end. The numbers of water molecules present in the active and the buffer zones during the simulations are provided in Figure 6a and c for the S_{int} and S_{add} simulations, respectively. The statistics of those numbers are tabulated in Table 1. The fluctuations in the plots imply exchanges of water molecules between the active, buffer, and environmental zones. Because the S_{add} site is more open than S_{int} to bulk solvent, the X_{add} plot features more of buffer groups (up to 7) than the X_{int} plot (up to 4) and also fluctuates more frequently. For both S_{int} and S_{add}, the numbers of water molecules in the active zones were largely maintained at 5 and 8, respectively, despite fluctuations, implying equal probabilities of solvent molecules entering and leaving the active zone.

The scenario suggested by Cal1 agreed with the QM results obtained by conventional QM/MM with large QM subsystems. Figure S7 in the Supporting Information plots the numbers of water molecules over time in Zone 1 and Zone 2 obtained at the QM level by conventional QM/MM (Zones 1 and 2 are counterparts to the active and buffer zones in adaptive QM/MM, respectively). The plots are qualitatively similar to those in Cal1 in that the numbers of water were largely maintained despite fluctuations. The average numbers of water molecules were 5 in Zone 1 and 10 in Zone 2 (Table 1), close to those in the active zone (5) and buffer zone (8) in Cal1. Although one would not expect exact agreements because the QM interactions between the water molecules in Zone 1 and Zone 2 differed from the MM interactions between the water molecules in the active and buffer zones, the similarity between the Cal1 and the QM results did suggest that the modified PAP scheme provided a realistic description for the exchange of solvent molecules between the EcClC binding sites and bulk solvent.

Table 1. Statistics of the Numbers of Water Molecules in the Active and Buffer Zones in the Adaptive QM/MM Simulations in Comparisons with Those in Zone 1 and Zone 2 in Conventional QM/MM Simulations[a]

	zone	avg.	standard deviation	min.	max.
			S_{int}		
Cal1	active	4.7	0.6	3	6
	buffer	0.6	0.7	0	4
Cal2	active	3.4	1.0	2	5
	buffer	0.3	0.5	0	3
QM[b]	1	5.4	0.8	4	7
	2	0.7	0.7	0	3
			S_{add}		
Cal1	active	7.8	1.3	4	12
	buffer	2.6	1.3	0	7
Cal2	active	3.1	0.6	2	9
	buffer	1.1	0.9	0	6
QM[b]	1	9.8	1.0	7	13
	2	1.8	1.2	0	7

[a] S_{int} and S_{add} denote simulations where the QM subsystem located at the S_{int} and S_{add} binding sites, respectively. The extra forces due to smoothing functions were deleted in Cal1 but retained in Cal2. Conventional QM/MM used large QM subsystems that included the two Cl^- ions and all movable water molecules in the intracellular (or extracellular) vestibule for S_{int} (or S_{add}) simulations. Zones 1 and 2 are the counterparts in the conventional QM/MM to the active and buffer zones in the adaptive QM/MM, respectively. [b] By conventional QM/MM with large QM subsystems.

Turned to Cal2, we discover a quite different picture. Both visual inspections of the saved trajectories and the plots of the numbers of water molecules in the active and buffer zones over time (Figure 6b and d) revealed the moving of water molecules out of the active zone. Initially, there were 5 and 9 QM water molecules in the S_{int} and S_{add} simulations, respectively. By the end of the 50 ps simulations, both numbers had decreased monotonously to 2. We found that in each simulation, the two QM water molecules in the first solvation shell of the Cl^- ion stayed, but the other QM water molecules that were more distant from the Cl^- ion gradually diffused into the buffer zone. When those QM water molecules moved into the buffer zone, they did not re-enter the active zone; instead, they wandered in the buffer zone and eventually continued the journey into the environmental zone. When MM water molecules traveled to the buffer zone, they were not able to enter the active zone; they could only stay in the buffer zone or go back to the environmental zone. Apparently, the moving of the solvent molecules out of the active zone in Cal2 (as compared with the "free" exchange of water molecules between the active and environmental zones in Cal1) was caused by the extra forces due to smoothing functions that were acting on the buffer groups, because those forces were the only meaningful difference between Cal1 and Cal2. The extra forces prevented the buffer groups from getting into the active zone but not from getting into the environmental zone in the simulations. The unbalance in the numbers of water molecules entering and leaving the active zone caused the reduction in the number of QM water molecules in Cal2. The results clearly indicate that the extra forces, though transient, if not negligibly small, will lead to artifacts in simulations. The treatment of neglecting those extra forces is necessary in such situations.

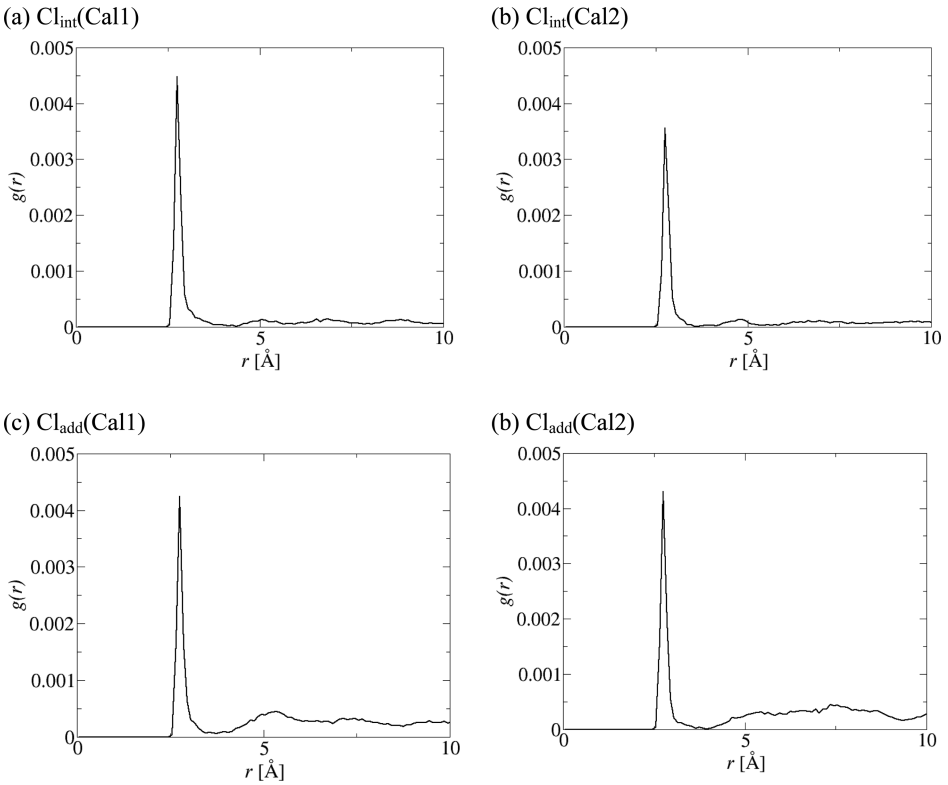


Figure 7. Radial distribution functions $g_{Cl\text{-}OW}$, where Cl═Cl$_{int}$ in the S$_{int}$ and Cl═Cl$_{add}$ in the S$_{add}$ simulations, respectively, and OW is the water oxygen. Cal1 and Cal2 denote, respectively, the calculations without and with the extra forces due to smoothing functions.

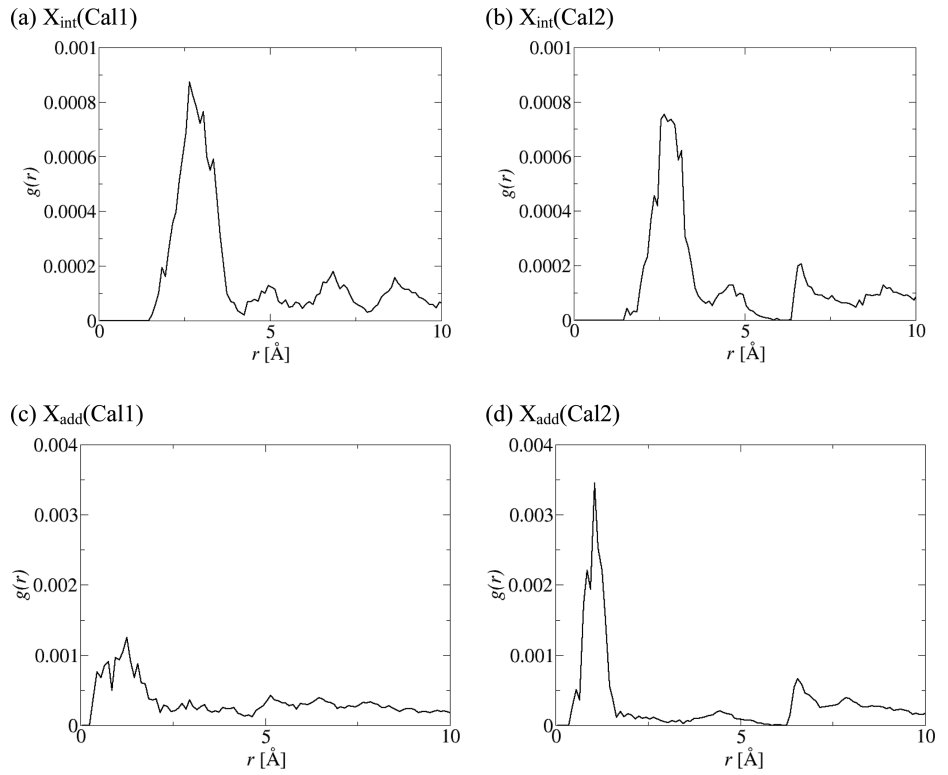


Figure 8. Radial distribution functions $g_{X\text{-}OW}$, where X = X$_{int}$ in the S$_{int}$ and X = X$_{add}$ in the S$_{add}$ simulations, respectively, and OW is the water oxygen. Cal1 and Cal2 denote, respectively, the calculations without and with the extra forces due to smoothing functions.

IV.D. Solvation Structures of Bound Cl$^-$ Ions. Water molecules assist the solvation of the bound Cl$^-$ ions that are coordinated by the backbone amine groups (for Cl$_{int}$) or by the side chain (for Cl$_{add}$). Figure 7 provides the Cl−OW RDF for the adaptive QM/MM, where OW denotes the water oxygen atom. Figure S8 in the Supporting Information shows the QM

dx.doi.org/10.1021/ct500553x | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

RDF results obtained by conventional QM/MM with large QM subsystems. Each RDF displays a sharp first peak around r = 2.6 Å, which corresponds to the water molecules that complete the first solvation shells of the Cl^- ion. It can easily be seen that the second peak and beyond are higher in the Cl_{add}−OW curves than in the Cl_{int}−OW curves, in line with that S_{add} is more open to bulk solvent than S_{int} is. Overall, the RDFs by Cal1 agree reasonably well for r < 6 Å with the QM RDFs. The agreements between the Cal2 and QM are less satisfactory, especially in the case of Cl_{add}−OW. It is noticeable that the water density between r = 5 Å and r = 6 Å in Cl_{add}−OW(Cal2) had been shifted to r > 6 Å. This tendency is echoed in the plots of the integrated coordination numbers of water oxygen around the Cl^- ions (Cal1 and Cal2 in Figure S5 and QM in Figure S8 in the Supporting Information). For example, at r = 6 Å, the integrated coordination number reaches 5 in Cl_{int}−OW(Cal1) and Cl_{int}−OW(QM) and 10 in Cl_{add}−OW(Cal1) and Cl_{add}−OW(QM), respectively. To reach the same values, one would have to go to r = 7 Å in the corresponding Cal2 plots.

It is interesting to take a look at the pseudoatom-OW RDF in Figure 8. Both the X_{int}−OW(Cal1) and X_{int}−OW(Cal2) curves resemble the Cl_{int}−OW(Cal1) and Cl_{int}−OW(Cal2) curves, respectively. The locations of the first and second peaks were found near r = 3 Å and r = 5 Å, respectively. The reason for the similarities between the Cl_{int}−OW and X_{int}−OW plots is obvious: Cl_{int} stayed rather close to X_{int} during the simulations. In contrast, the X_{add}−OW plots are rather different from the Cl_{add}−OW plots. This is understandable because Cl_{add} traveled rather far away from X_{add} during the simulations. Interestingly, the X_{add}−OW curves display a peak near r = 1.2 Å, indicating that a water molecule was moving closer toward the initial binding position of Cl_{add} once Cl_{add} left that position. Just as the Cl−OW RDF, the X−OW RDFs also exhibit the shifting of the density in the range of r = 5 to 6 Å to r > 6 Å, as confirmed in the plots of the integrated coordination number in Figure S6 in the Supporting Information.

V. SUMMARY

In this work, we present a modified PAP scheme of QM/MM for the exchange of solvent molecules between a protein binding site and bulk solvent. Two treatments are introduced. The first is to place the active-zone center at a pseudoatom in the binding site, whose location is not affected by the movements of ligands or residues. The second treatment is to delete the extra forces due to smoothing functions that act on the buffer groups and the atoms involved in the active-zone center definition. The new scheme, although no longer describing a Hamiltonian system, obeys Newton's Third Law of Motion and conserves momentum. The modified PAP algorithm was applied in the simulations of two binding sites S_{int} and S_{add} of the *Escherichia coli* CLC chloride ion transport protein. The results revealed frequent exchange of water molecules between the binding sites and the bulk solvent, consistent with benchmark QM results obtained by conventional QM/MM with large QM subsystems. Furthermore, as demonstrated by the differences between the Cal1 and Cal2 results, the extra forces due to smoothing functions on the buffer groups, when retained, caused artifacts that QM water molecules moved away from the active-zone, though those forces are transient. The overall agreements between the Cal1 and QM results in the water molecules entering and leaving the binding site as well as in the solvation structures of the bound Cl^- ions suggest that the modified PAP scheme was able to provide a realistic description for the exchange of solvent molecules between the protein binding site and bulk solvent.

While deleting the extra forces due to smoothing functions provides a convenient shortcut to avoid the artifacts caused by those forces, a theoretically more rigorous treatment will be to align the QM and MM potentials to minimize or even eliminate the extra forces, which are proportional to the differences between the QM and MM potentials. However, such parametrizations are very difficult to achieve in general because of the multidimensional nature of the potentials. The problem is perhaps more challenging for polar molecules such as water than for nonpolar molecules due to the various orientations of the molecular dipole. The treatment in this work (the QM and MM potentials agree with each other at $R = \infty$) is probably not the optimal alignment for this specific problem, but it is generally applicable and is not affected by the relative orientations of the involved species. It is desirable to explore general and efficient methods in the future to carry out potential alignments, especially for anisotropic polyatomic species.

■ ASSOCIATED CONTENT

Supporting Information

Total energy (Figure S1), temperature (Figure S2), and velocities of the pseudoatoms X (Figure S3) over time in MD simulations, radial distribution functions $g_{X\text{-}Cl}$ (Figure S4), and integrated coordination numbers of water oxygen around the Cl^- ions (Figure S5) and around the pseudoatoms X (Figure S6), numbers of water molecules in Zone 1 and Zone 2 (Figure S7), and radial distribution functions and integrated coordination numbers of water oxygen around Cl (Figure S8) at the QM level by conventional QM/MM with large QM subsystems. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: 303-352-3889. Fax: 303-556-4776. Email: soroosh.pezeshki@ucdenver.edu.

*Email: hai.lin@ucdenver.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by National Science Foundation (CHE-0952337). This work used the Extreme Science and Engineering Discovery Environment (XSEDE) under grant CHE-140070, which is supported by National Science Foundation grant number ACI-1053575. H.L. thanks the Camille & Henry Dreyfus Foundation for support (TH-14-028). C.D. thanks the University of Colorado Denver for an Undergraduate Research Opportunity Program award.

■ REFERENCES

(1) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(2) Singh, U. C.; Kollmann, P. A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH_3Cl + Cl^- exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718−730.

(3) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.
(4) Gao, J. Methods and applications of combined quantum mechanical and molecular mechanical potentials. *Rev. Comput. Chem.* **1996**, *7*, 119−185.
(5) Friesner, R. A.; Beachy, M. D. Quantum mechanical calculations on biological systems. *Curr. Opin. Struct. Biol.* **1998**, *8*, 257−262.
(6) Gao, J.; Thompson, M. A., Eds. *Combined Quantum Mechanical and Molecular Mechanical Methods*, ACS Symp. Ser. 712; American Chemical Society: Washington, DC, 1998; p 310.
(7) Ruiz-López, M. F.; Rivail, J. L. Combined quantum mechanics and molecular mechanics approaches to chemical and biochemical reactivity In *Encyclopedia of Computational Chemistry*, von Ragué Schleyer, P., Ed.; Wiley: Chichester, 1998; Vol. *1*, pp 437−448.
(8) Monard, G.; Merz, K. M., Jr. Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Acc. Chem. Res.* **1999**, *32*, 904−911.
(9) Hillier, I. H. Chemical reactivity studied by hybrid QM/MM methods. *THEOCHEM* **1999**, *463*, 45−52.
(10) Hammes-Schiffer, S. Theoretical perspectives on proton-coupled electron transfer reactions. *Acc. Chem. Res.* **2000**, *34*, 273−281.
(11) Sherwood, P. Hybrid quantum mechanics/molecular mechanics approaches. In *Modern Methods and Algorithms of Quantum Chemistry*, Grotendorst, J., Ed.; John von Neumann-Instituts: Jülich, 2000; Vol. *3*, pp 285−305.
(12) Gao, J.; Truhlar, D. G. Quantum mechanical methods for enzyme kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(13) Morokuma, K. New challenges in quantum chemistry: Quests for accurate calculations for large molecular systems. *Philos. Trans. R. Soc. A* **2002**, *360*, 1149−1164.
(14) Lin, H.; Truhlar, D. G. QM/MM: What have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(15) Senn, H. M.; Thiel, W. QM/MM methods for biological systems. *Top. Curr. Chem.* **2007**, *268*, 173−290.
(16) Pezeshki, S.; Lin, H. Recent developments in QM/MM methods towards open-boundary multi-scale simulations. *Mol. Simul.* **2014**, DOI: 10.1080/08927022.2014.911870.
(17) Kerdcharoen, T.; Morokuma, K. ONIOM-XS: An extension of the ONIOM method for molecular simulation in condensed phase. *Chem. Phys. Lett.* **2002**, *355*, 257−262.
(18) Kerdcharoen, T.; Morokuma, K. Combined quantum mechanics and molecular mechanics simulation of Ca^{2+}/ammonia solution based on the ONIOM-XS method: Octahedral coordination and implication to biology. *J. Chem. Phys.* **2003**, *118*, 8856−8862.
(19) Heyden, A.; Lin, H.; Truhlar, D. G. Adaptive partitioning in combined quantum mechanical and molecular mechanical calculations of potential energy functions for multiscale simulations. *J. Phys. Chem. B* **2007**, *111*, 2231−2241.
(20) Pezeshki, S.; Lin, H. Adaptive-partitioning redistributed charge and dipole schemes for QM/MM dynamics simulations: On-the-fly relocation of boundaries that pass through covalent bonds. *J. Chem. Theory Comput.* **2011**, *7*, 3625−3634.
(21) Bulo, R. E.; Ensing, B.; Sikkema, J.; Visscher, L. Toward a practical method for adaptive QM/MM simulations. *J. Chem. Theory Comput.* **2009**, *5*, 2212−2221.
(22) Park, K.; Gotz, A. W.; Walker, R. C.; Paesani, F. Application of adaptive QM/MM methods to molecular dynamics simulations of aqueous systems. *J. Chem. Theory Comput.* **2012**, *8*, 2868−2877.
(23) Kerdcharoen, T.; Liedl, K. R.; Rode, B. M. A QM/MM simulation method applied to the solution of Li^+ in liquid ammonia. *Chem. Phys.* **1996**, *211*, 313−323.
(24) Bernstein, N.; Varnai, C.; Solt, I.; Winfield, S. A.; Payne, M. C.; Simon, I.; Fuxreiter, M.; Csanyi, G. QM/MM simulation of liquid water with an adaptive quantum region. *Phys. Chem. Chem. Phys.* **2012**, *14*, 646−656.
(25) Várnai, C.; Bernstein, N.; Mones, L.; Csányi, G. Tests of an adaptive QM/MM calculation on free energy profiles of chemical reactions in solution. *J. Phys. Chem. B* **2013**, *117*, 12202−12211.
(26) Kästner, J.; Thiel, W. Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: "Umbrella integration". *J. Chem. Phys.* **2005**, *123*, 144104/1−5.
(27) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.* **1989**, *156*, 472−477.
(28) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(29) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(30) Heyden, A.; Truhlar, D. G. Conservative algorithm for an adaptive change of resolution in mixed atomistic/coarse-grained multiscale simulations. *J. Chem. Theory Comput.* **2008**, *4*, 217−221.
(31) Praprotnik, M.; Poblete, S.; Delle Site, L.; Kremer, K. Comment on "Adaptive multiscale molecular dynamics of macromolecular fluids". *Phys. Rev. Lett.* **2011**, *107*, 099801/1−2.
(32) Bulo, R. E.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. Multiscale modeling of chemistry in water: Are we there yet? *J. Chem. Theory Comput.* **2013**, *9*, 5567−5577.
(33) White, M. M.; Miller, C. A voltage-gated anion channel from the electric organ of Torpedo-Californica. *J. Biol. Chem.* **1979**, *254*, 161−166.
(34) Maduke, M.; Miller, C.; Mindell, J. A. A decade of ClC chloride channels: Structure, mechanism, and many unsettled questions. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 411−438.
(35) Fahlke, C. Ion permeation and selectivity in ClC-type chloride channels. *Am. J. Physiol. Renal. Physiol.* **2001**, *280*, F748−757.
(36) Dutzler, R. The ClC family of chloride channels and transporters. *Curr. Opin. Struct. Biol.* **2006**, *16*, 439−446.
(37) Jentsch, T. J. CLC chloride channels and transporters: From genes to protein structure, pathology, and physiology. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 3−36.
(38) Verkman, A. S.; Galietta, L. J. V. Chloride channels as drug targets. *Nat. Rev. Drug. Discovery* **2009**, *8*, 153−171.
(39) Dutzler, R.; Campbell, E. B.; Cadene, M.; Chait, B. T.; MacKinnon, R. X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature* **2002**, *415*, 287−294.
(40) Dutzler, R.; Campbell, E. B.; MacKinnon, R. Gating the selectivity filter in ClC chloride channels. *Science* **2003**, *300*, 108−112.
(41) Smith, M.; Lin, H. Charge delocalization upon chloride ion binding in ClC chloride ion channels/transporters. *Chem. Phys. Lett.* **2011**, *502*, 112−117.
(42) Faraldo-Gómez, J. D.; Roux, B. Electrostatics of ion stabilization in a ClC chloride channel homologue from *Escherichia coli*. *J. Mol. Biol.* **2004**, *339*, 981−1000.
(43) Yin, J.; Kuang, Z.; Mahankali, U.; Beck, T. L. Ion transit pathways and gating in ClC chloride channels. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 414−421.
(44) Bostick, D. L.; Berkowitz, M. L. Exterior site occupancy infers chloride-induced proton gating in a prokaryotic homolog of the ClC chloride channel. *Biophys. J.* **2004**, *87*, 1686−1696.
(45) Cohen, J.; Schulten, K. Mechanism of anionic conduction across ClC. *Biophys. J.* **2004**, *86*, 836−845.
(46) Gervasio, F. L.; Parrinello, M.; Ceccarelli, M.; Klein, M. L. Exploring the gating mechanism in the ClC chloride channel via metadynamics. *J. Mol. Biol.* **2006**, *361*, 390−398.
(47) Church, J.; Pezeshki, S.; Davis, C.; Lin, H. Charge transfer and polarization for chloride ions bound in ClC transport proteins: Natural bond orbital and energy decomposition analyses. *J. Phys. Chem. B* **2013**, *117*, 16029−16043.
(48) Chen, T. Y.; Miller, C. Nonequilibrium gating and voltage dependence of the ClC-0 Cl^- channel. *J. Gen. Physiol.* **1996**, *108*, 237−250.
(49) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.;

Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(50) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(51) Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.

(52) Stewart, J. J. P. Optimization of parameters for semiempirical methods II. Applications. *J. Comput. Chem.* **1989**, *10*, 221−264.

(53) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1982**, *52*, 255−268.

(54) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.

(55) Lin, H.; Zhang, Y.; Pezeshki, S.; Truhlar, D. G. *QMMM*, Version 1.4.0.CO; University of Minnesota: Minneapolis, 2014.

(56) Ponder, J. W. *TINKER*, Version 5.1; Washington University: St. Louis, MO, 2010.

(57) Thiel, W. *MNDO2005*, Version 7.0; Max-Planck-Institut für Kohlenforschung: Müheim an der Ruhr, Germany, 2005.