JCTC Journal of Chemical Theory and Computation
Letter
pubs.acs.org/JCTC

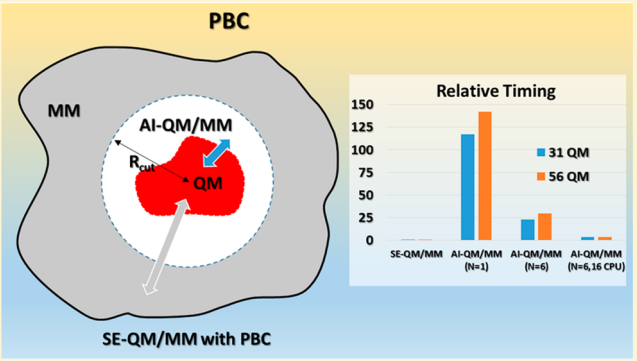
Acceleration of Ab Initio QM/MM Calculations under Periodic Boundary Conditions by Multiscale and Multiple Time Step Approaches

Kwangho Nam*,†

†Department of Chemistry and Computational Life Science Cluster (CLiC), Umeå University, 901 87, Umeå, Sweden

Supporting Information

ABSTRACT: Development of multiscale ab initio quantum mechanical and molecular mechanical (AI-QM/MM) method for periodic boundary molecular dynamics (MD) simulations and their acceleration by multiple time step approach are described. The developed method achieves accuracy and efficiency by integrating the AI-QM/MM level of theory and the previously developed semiempirical (SE) QM/MM-Ewald sum method [*J. Chem. Theory Comput.* **2005**, *1*, 2] extended to the smooth particle-mesh Ewald (PME) summation method. In the developed methods, the total energy of the simulated system is evaluated at the SE-QM/MM-PME level of theory to include long-range QM/MM electrostatic interactions, which is then corrected on the fly using the AI-QM/MM level of theory within the real space cutoff. The resulting energy expression enables decomposition of total forces applied to each atom into forces determined at the low-level SE-QM/MM method and correction forces at the AI-QM/MM level, to integrate the system using the reversible reference system propagator algorithm. The resulting method achieves a substantial speed-up of the entire calculation by minimizing the number of time-consuming energy and gradient evaluations at the AI-QM/MM level. Test calculations show that the developed multiple time step AI-QM/MM method yields MD trajectories and potential of mean force profiles comparable to single time step QM/MM results. The developed method, together with message passing interface (MPI) parallelization, accelerates the present AI-QM/MM MD simulations about 30-fold relative to the speed of single-core AI-QM/MM simulations for the molecular systems tested in the present work, making the method less than one order slower than the SE-QM/MM methods under periodic boundary conditions.



1. INTRODUCTION

The combined quantum mechanical and molecular mechanical (QM/MM) method is an effective and widely used approach for theoretical study of chemical reactions in biological molecules and in water.[1−3] In the QM/MM method, the simulated molecular system is partitioned into a subsystem containing a relatively small number of atoms, such as substrates and catalytic residues of enzymes, represented by quantum mechanics and a subsystem composed of the rest of the system represented by classical mechanics. Since the QM subsystem contains in most cases less than a few hundred atoms, the method is more cost-effective than treating the entire system quantum mechanically. However, while it is desirable to use ab initio (AI) quantum mechanical or density functional theory (DFT) methods,[4−6] the large computational expense of AI-QM/DFT methods, even for small QM regions, precludes their use in long QM/MM molecular dynamics (MD) simulations. Instead, less accurate and approximate quantum mechanical methods, such as the AM1,[7] PM3,[8] and DFT-based tight-binding methods,[9−11] have been used in many applications.[12,13]

Long-range electrostatics has a profound impact on the stability of periodic boundary MD simulations[14,15] and thus on chemical reactions that take place in enzymes and in water.[12,16,17] However, it has not been rigorously taken into account in most AI-QM/MM MD simulations under periodic boundary conditions (PBC) due mainly to lack of algorithms evaluating them efficiently. Instead, most conventional AI-QM/MM simulations have been carried out either by ignoring all MM charges outside of the so-called cutoff distance[18,19] or by including all MM charges of the system (i.e., no-cutoff).[20,21] Although simulations without cutoff yield in principle a more reliable result than simulations with cutoff,[22] they are computationally more expensive than simulations with cutoff.

One of the most rigorous and widely used methods to evaluate long-range electrostatic interactions under PBC is the Ewald summation method.[23,24] For the semiempirical QM (SE-QM) methods, a linear-scaling algorithm for performing the QM/MM Ewald summation calculations has been developed (hereafter the QM/MM-Ewald method).[22] This method is

Received: July 2, 2014
Published: September 15, 2014

dx.doi.org/10.1021/ct5005643 | *J. Chem. Theory Comput.* 2014, 10, 4175−4183

implemented in the CHARMM[22,25] and the AMBER[26] programs. There are also other SE-QM/MM methods developed for efficient evaluations of the long-range QM/MM interactions under PBC.[27−29] Although the QM/MM-Ewald method is general and can be extended to the conventional AI-QM/MM method, extension of the method to the AI-QM/MM method is not straightforward. For example, in the original work by Nam et al.,[22] Mulliken charges[30] were used to represent the QM image charges. Although they are reasonable for the SE-QM methods with minimal basis sets, they are basis set dependent and potentially cause instability of the resulting method, in particular, if large basis sets are used.[31,32] Recently, Holden et al.[32] developed an approach based on the ChElPG charges[33] to overcome the limitation of the Mulliken charge-based QM/MM-Ewald method. Their method, although promising, is computationally intensive and thus less practical for long AI-QM/MM MD simulations.

In this letter, development of an efficient multiscale (MS) QM/MM simulation method is presented. The developed method includes the long-range QM/MM electrostatic interaction energy evaluated at the SE-QM/MM level, utilizing the efficiency of the smooth particle-mesh Ewald (PME) sum algorithm (denoted as the SE-QM/MM-PME method),[34] and the energy correction at the AI-QM/MM level. Here, we use the term "multiscale" to point out that energy functions with different accuracy and length scales are integrated to the method. Specifically, an interpolation is made between the (high-accuracy) real space AI-QM/MM energy and the (low-accuracy) periodic boundary SE-QM/MM-PME energy on the fly to determine total energy of the simulated system. We also present the method to be further accelerated by applying the multiple time step reference system propagator algorithm[35] without sacrificing much accuracy. Therein, acceleration is achieved by minimizing the number of computer-intensive AI-QM/MM calculations. The developed methods are highly accurate and stable, yielding results that are comparable to single time step AI-QM/MM MD simulation results. In the following section, we briefly describe the derivation of the developed MS AI-QM/MM method. Computational details are presented in section 3. The results and discussion are presented in section 4, followed by conclusion in section 5.

2. THEORY

2.1. Ab Initio QM/DFT-based QM/MM Method under Periodic Boundary Conditions (PBC). On the basis of the QM/MM-Ewald sum method developed by Nam et al.,[22] the total energy of the AI-QM/MM method under PBC is expressed as

$$E_{\mathrm{tot}}^{\mathrm{AI-QM/MM}} = E_{\mathrm{AI-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] + \Delta E_{\mathrm{AI-QM/MM}}^{\mathrm{PB}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] + E_{\mathrm{MM}}[q^{\mathrm{MM}}; q^{\mathrm{MM}}] \quad (1)$$

where the charge distribution of the entire system is partitioned into a charge density ρ^{AI} (i.e., electron density and nuclear charges) of the QM subsystem and an MM charge distribution q^{MM}, which is represented by partial atomic charges of force fields. In the present work, the notation used by Nam et al.[22] is closely followed. In eq 1, $E_{\mathrm{AI-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}]$ is the interaction energy within the "real space" cutoff modeled by the conventional QM/MM potential:[3]

$$E_{\mathrm{AI-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] = E_{\mathrm{AI-QM/MM}}^{0}[\rho^{\mathrm{AI}}; \rho^{\mathrm{AI}}] + E_{\mathrm{AI-QM/MM}}^{\mathrm{elec}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] + E_{\mathrm{AI-QM/MM}}^{\mathrm{vdW}} \quad (2)$$

where $E_{\mathrm{AI-QM/MM}}^{0}$ is the energy for the QM charge distribution, $E_{\mathrm{AI-QM/MM}}^{\mathrm{elec}}$ the interaction energy between the QM charge distribution and the MM charges within the real space cutoff distance, and $E_{\mathrm{AI-QM/MM}}^{\mathrm{vdW}}$ the van der Waals interaction energy between the QM atoms and the MM atoms modeled by a classical Lennard-Jones potential, respectively. The last term in eq 1 is the classical force field energy of the MM subsystem, in which the CHARMM potential[36] is used in the present work.

The second term, $\Delta E_{\mathrm{AI-QM/MM}}^{\mathrm{PB}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}]$, in eq 1 is the periodic boundary QM/MM correction energy formally defined as a difference between the total energy of the QM subsystem interacting with the MM subsystem under PBC and the real space QM/MM energy, i.e.,

$$\Delta E_{\mathrm{AI-QM/MM}}^{\mathrm{PB}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] = (E_{\mathrm{tot}}^{\mathrm{AI-QM/MM}} - E_{\mathrm{MM}}[q^{\mathrm{MM}}; q^{\mathrm{MM}}]) - E_{\mathrm{AI-QM/MM}}^{RS}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] \quad (3)$$

To solve eq 3 approximately, we first represent the charge density of the QM subsystem by a set of point charges, denoted as q^{AI}, derived from ρ^{AI}, and apply the regular Ewald or the particle-mesh Ewald sum (PME) methods[34] (see Appendices A and B). Since eq 3 is applied only to atom pairs that are separated longer than the real space cutoff distance, q^{AI} is a reasonable approximation of the real QM charge distribution for those atom pairs. The resulting energy is then

$$E_{\mathrm{tot}}^{\mathrm{AI-QM/MM}} \approx E_{\mathrm{AI-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] + \Delta E_{\mathrm{AI-QM/MM}}^{\mathrm{PB}}[q^{\mathrm{AI}}; q^{\mathrm{MM}}] + E_{\mathrm{MM}}[q^{\mathrm{MM}}; q^{\mathrm{MM}}] \quad (4)$$

The first two terms of eq 4 are solved together by the self-consistent field (SCF) procedure[3,22] or by the extended Lagrangian MD approach,[37−39] in which the $E_{\mathrm{QM/MM}}^{\mathrm{vdw}}$ term (see eq 2) is determined separately. If the PME method is used, the resulting method is referred to as the QM/MM-PME method in this work, to distinguish it from the QM/MM-Ewald method (see Appendix A for the difference between the two methods).

2.2. Long-Range Interaction Energy Approximation of the Multiscale (MS) AI-QM/MM Method. To derive the multiscale AI-QM/MM method and to overcome the challenge of extending the Mulliken charge based QM/MM-Ewald (and QM/MM-PME) method to the AI-QM/MM level of theory, a reference low-level theory is introduced to eq 4. In the present work, we use the semiempirical QM (SE-QM) methods as the low-level theory and modify eq 4 as

$$E_{\mathrm{tot}}^{\mathrm{AI-QM/MM}} = \{E_{\mathrm{SE-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{SE}}; q^{\mathrm{MM}}] + (E_{\mathrm{AI-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{AI}}; q^{\mathrm{MM}}] - E_{\mathrm{SE-QM/MM}}^{\mathrm{RS}}[\rho^{\mathrm{SE}}; q^{\mathrm{MM}}])\} + \{\Delta E_{\mathrm{SE-QM/MM}}^{\mathrm{PB}}[q^{\mathrm{SE}}; q^{\mathrm{MM}}] + (\Delta E_{\mathrm{AI-QM/MM}}^{\mathrm{PB}}[q^{\mathrm{AI}}; q^{\mathrm{MM}}] - \Delta E_{\mathrm{SE-QM/MM}}^{\mathrm{PB}}[q^{\mathrm{SE}}; q^{\mathrm{MM}}])\} + E_{\mathrm{MM}}[q^{\mathrm{MM}}; q^{\mathrm{MM}}] \quad (5)$$

dx.doi.org/10.1021/ct5005643 | J. Chem. Theory Comput. 2014, 10, 4175−4183

where ρ^X and q^X represent the charge density and the QM charge distribution determined from the "X" level of theory (i.e., X = AI or SE), respectively. The equation is exact given the assumption introduced in eq 4 (i.e., $\Delta E^{PB}_{X-QM/MM}[\rho^X;q^{MM}] \approx \Delta E^{PB}_{X-QM/MM}[q^X;q^{MM}]$) and is rearranged to

$$E^{AI-QM/MM}_{tot} = E^{RS}_{SE-QM/MM}[\rho^{SE};\, q^{MM}] + \Delta E^{RS}_{cor} + \Delta E^{PB}_{SE-QM/MM}[q^{SE};\, q^{MM}] + \Delta\Delta E^{PB}_{cor} + E_{MM}[q^{MM};\, q^{MM}] \quad (6)$$

where

$$\Delta E^{RS}_{cor} = E^{RS}_{AI-QM/MM}[\rho^{AI};\, q^{MM}] - E^{RS}_{SE-QM/MM}[\rho^{SE};\, q^{MM}] \quad (7)$$

and

$$\Delta\Delta E^{PB}_{cor} = \Delta E^{PB}_{AI-QM/MM}[q^{AI};\, q^{MM}] - \Delta E^{PB}_{SE-QM/MM}[q^{SE};\, q^{MM}] \quad (8)$$

Here, we introduce an approximation to eq 8 that $\Delta\Delta E^{PB}_{cor} = 0$. Equation 6 is then simplified to

$$E^{AI-QM/MM}_{tot} = E^{SE-QM/MM}_{tot} + \Delta E^{RS}_{cor} \quad (9)$$

where

$$E^{SE-QM/MM}_{tot} = E^{RS}_{SE-QM/MM}[\rho^{SE};\, q^{MM}] + \Delta E^{PB}_{SE-QM/MM}[q^{SE};\, q^{MM}] + E_{MM}[q^{MM};\, q^{MM}] \quad (10)$$

Equation 10 is identical to the result of Nam et al.[22] In the present work, we also use the Mulliken charges for q^{SE}.

Given the assumption introduced to eq 8 (i.e., $\Delta\Delta E^{PB}_{cor} = 0$), eq 9 implies that the correction of the total energy at the AI-QM/MM level of theory occurs at the distance range within the real space cutoff. Here, we refer the interaction within the real space cutoff, including the interaction between QM atoms, as the short-range interaction and the interaction between atom pairs that are separated more than the cutoff distance as the long-range interaction, which is represented collectively by $\Delta E^{PB}_{SE-QM/MM}$ in eq 10. Equation 9 involves three energy evaluations: one at the SE-QM/MM-Ewald or the SE-QM/MM-PME level and the other two at the real space AI-QM/MM and SE-QM/MM levels to produce the short-range correction term (eq 7). Since the AI-QM/MM calculation is carried out with MM atoms that are within the real space cutoff distance (i.e., short-range), the present method is more efficient than the conventional no-cutoff AI-QM/MM method, the computational cost of which scales as $O(N^{QM} \times N^{MM})$, where N^{QM} and N^{MM} are the number of QM and MM atoms in the system, respectively. The resulting energy expression is similar to the energy interpolation scheme used in the ONIOM method.[40] The way that the short- and the long-range energies are integrated to the total QM/MM energy in the present methods, however, differs from the ONIOM method[40] as well as from the other dual-level QM/MM method.[41] In particular, the present method uses a multiscale representation of the electron density of the QM subsystem; i.e. the QM charge distribution is represented by the AI-QM/DFT level of theory for short-range interactions, by the SE-QM level for long-range interactions, and by the Mulliken point charges for the periodic images of the QM subsystem, respectively. In this work, we refer to the present method simply as the AI-QM/MM-PME method unless otherwise specified. By contrast, the commonly used cutoff/no-cutoff AI-QM/MM method is referred to as the conventional AI-QM/MM method.

2.3. Multiple Time Step (MTS) Integrator for the Multiscale AI-QM/MM-PME Method. Equation 9 suggests that the total atomic forces can be decomposed into the forces determined at the low-level theory and the correction forces at the high-level theory,

$$\mathbf{F}^{AI-QM/MM}_{tot} = \mathbf{F}^{SE-QM/MM}_{tot} + \Delta\mathbf{F}^{RS}_{cor} \quad (11)$$

If the low-level theory describes reasonably well chemical systems to be simulated, the correction forces will be small relative to the total forces of the low-level theory and fluctuate slowly over time. This time scale separation between the two force components makes possible the application of the multiple time step (MTS) algorithms, in particular, the reversible reference system propagator algorithm (rRESPA),[35] to accelerate the entire calculation.

In the rRESPA method, the time evolution of a phase space point $\Gamma(t) = \{\mathbf{r}(t), \mathbf{p}(t)\}$ of positions and conjugate momenta of an N_T−particle system ($N_T = N^{QM} + N^{MM}$) at time t to a new point $\Gamma(t + \Delta t)$ at $t + \Delta t$ is achieved by

$$\Gamma(t + \Delta t) = e^{iL\Delta t}\Gamma(t) \quad (12)$$

In eq 12, L is the classical Liouville operator defined as

$$iL = iL_x + iL_F = \sum_{k=1}^{3N_T} \dot{r}_k \frac{\partial}{\partial r_k} + \sum_{k=1}^{3N_T} F_k \frac{\partial}{\partial p_k} \quad (13)$$

where the first term is applied to the position (r_k) and the second term to the force acting on the kth degree of freedom (F_k), respectively. By applying eq 11, the force component of eq 13 becomes

$$iL_F = iL_{F_{SE}} + iL_{F_{cor}} \quad (14)$$

where $iL_{F_{SE}}$ and $iL_{F_{cor}}$ correspond to the terms contributed by $\mathbf{F}^{SE-QM/MM}_{tot}$ and $\Delta\mathbf{F}^{RS}_{cor}$, respectively. Finally, the Trotter factorization of the Liouville time propagator of eq 12 yields the rRESPA time integrator

$$\Gamma(t + \Delta t) = e^{iL_{F_{cor}}\Delta t/2}\left(e^{iL_{F_{SE}}\delta t/2}e^{iL_x\delta t}e^{iL_{F_{SE}}\delta t/2}\right)^N \times e^{iL_{F_{cor}}\Delta t/2}\Gamma(t) \quad (15)$$

where δt and $\Delta t = N\delta t$ are the time step for the inner and the outer loops, respectively. In the present work, the rRESPA integrator is realized using a modified form of the velocity Verlet integrator.

Equation 15 indicates that the inner loop evolves with the update of positions and forces at the SE-QM/MM level (using either the SE-QM/MM-Ewald or SE-QM/MM-PME method) with the δt time step, while the correction forces at the AI-QM/MM level are applied at the outer loop, at every Nth δt time step. Since the main computational bottleneck of the present AI-QM/MM-PME method is the AI-QM/MM calculation, the application of the rRESPA technique will result in a speed-up of the method, roughly proportional to N relative to the time of the single time step (STS) AI-QM/MM-PME calculation (i.e., $N = 1$) (see Results and Discussion). Recently, Steel has applied the rRESPA method to accelerate the ab initio molecular dynamics (AIMD), utilizing the fact that the correlation energy varies slowly over time.[42] The present multiple time step method and the method by Steel[42] differ from other MTS methods using the rRESPA algorithm.[19,35,43]

dx.doi.org/10.1021/ct5005643 | J. Chem. Theory Comput. 2014, 10, 4175−4183

In those methods, total atomic forces are decomposed based on distance separation between atoms, in which slowly varying long-range interactions are subjected to the large time step integration. By contrast, the present method and the method by Steel perform the force decomposition based on the levels of theory that are used to determine the total energy of the simulated system, in which the large time step is applied to the slowly varying correction energy determined at the high-level theory (i.e., the AI-QM/MM level in the present work and the correlation energy in the work by Steel). The present method also differs from the method by Steel, which is for the full AIMD,[42] while the present method is for the QM/MM MD simulations under periodic boundary conditions, in which multiscale decomposition of total energy is performed (eq 9). Additionally, since the SE-QM method is used as the low-level QM theory, any AI-QM/DFT method can be used as the high-level QM theory in our method. This makes the present method far more efficient than that of Steel,[42] in which the Hartree−Fock (HF) theory is employed as the low-level theory, thus computationally more demanding than our method.

2.4. Nose−Hoover Temperature Control Algorithm for the MTS AI-QM/MM Method. To perform constant temperature MD simulation, the Nose−Hoover (NH) algorithm[44] is applied to the rRESPA multiscale AI-QM/MM method. In the Nose−Hoover dynamics, the conserved quantity is

$$H_{\mathrm{NH}} = H_{\mathrm{system}}(\mathbf{r}, \mathbf{p}) + \frac{p_\eta^2}{2m_\eta} + (3N_T + 1)k_{\mathrm{B}}T\eta \tag{16}$$

where $H_{\mathrm{system}}(\mathbf{r},\mathbf{p})$ is the N_T-particle system Hamiltonian, η and p_η an extended system variable and conjugate momentum, m_η a thermal inertia parameter controlling the rate of temperature fluctuation of the simulated molecular system, T a reference temperature, and k_{B} the Boltzmann constant, respectively. Then, the equations of motion for the kth particle of the system are

$$\dot{\mathbf{r}}_k = \frac{\mathbf{p}_k}{m_k} \text{ and } \dot{\mathbf{p}}_k = -\frac{\partial}{\partial \mathbf{r}_k} E_{\mathrm{tot}}^{\mathrm{AI-QM/MM}} - \frac{p_\eta}{m_\eta}\mathbf{p}_k \tag{17}$$

and for the thermostat variable,

$$\dot{\eta} = \frac{p_\eta}{m_\eta} \text{ and } \dot{p}_\eta = \sum_{k=1}^{N_T} \frac{\mathbf{p}_k^2}{m_k} - 3N_T k_{\mathrm{B}}T \tag{18}$$

respectively.[45] We note that eq 16 is not Hamiltonian but is conserved throughout the simulation.

The corresponding Liouville operator, based on the derivation by Tuckerman et al.,[35] is

$$\begin{aligned} iL_{\mathrm{NH}} &= iL_x + iL_{F_{\mathrm{SE}}} + iL_{F_{\mathrm{cor}}} + iL_{\eta p} + iL_\eta + iL_{F_\eta} \\ &= iL'_{in} + iL_{F_{\mathrm{cor}}} \end{aligned} \tag{19}$$

where

$$iL_{\eta p} = -\sum_{k=1}^{3N_T} \dot{\eta} p_k \frac{\partial}{\partial p_k},\ iL_\eta = \dot{\eta}\frac{\partial}{\partial \eta},\ iL_{F_\eta} = \dot{p}_\eta \frac{\partial}{\partial p_\eta} \tag{20}$$

and

$$iL'_{in} = iL_x + iL_{F_{\mathrm{SE}}} + iL_{\eta p} + iL_\eta + iL_{F_\eta} \tag{21}$$

In the present work, we modify the Liouville time integrator (eq 15) such that the thermostat variable is propagated with δt in the inner loop, yielding the Nose-Hoover MTS integrator

$$e^{iL_{\mathrm{NH}}\Delta t} = e^{iL_{F_{\mathrm{cor}}}\Delta t/2}\left(e^{iL'_{in}\delta t}\right)^N e^{iL_{F_{\mathrm{cor}}}\Delta t/2} \tag{22}$$

where

$$\begin{aligned} e^{iL'_{in}\delta t} = {} & e^{iL_{F_\eta}\delta t/2} e^{iL_\eta \delta t/2} e^{iL_{\eta p}\delta t/2} e^{iL_{F_{\mathrm{SE}}}\delta t/2} e^{iL_x \delta t} e^{iL_{F_{\mathrm{SE}}}\delta t/2} e^{iL_{\eta p}\delta t/2} \\ & \times e^{iL_\eta \delta t/2} e^{iL_{F_\eta}\delta t/2} \end{aligned} \tag{23}$$

Since the inner loop integration is much faster than the outer loop integration, the overhead caused by the thermostat is negligible, while a better temperature control can be achieved.[35] In addition, since eq 23 is the standard Liouville integrator for the single time step Nose−Hoover MD, the implementation of the NH thermostat to the present MTS AI-QM/MM-PME method is straightforward; i.e. the only modification needed for the implementation of the rRESPA algorithm is the outer loop integration presented in eq 22.

3. IMPLEMENTATION AND COMPUTATIONAL DETAILS

The multiple time step AI-QM/MM-PME method has been implemented into a modified version of the CHARMM program[25] (version c35a2) and interfaced with the GAMESS-UK program (version 8.0).[46] Two systems are prepared to test the methods. The first system contains an acetone molecule in a cubic box of 560 TIP3P waters[47] (lattice length of 25.4 Å). For PME, we use a 30 × 30 × 30 Fast Fourier Transform (FFT) grid (approximate grid size of 0.85 Å), 10.0 Å cutoff for the real space electrostatic interactions, the sixth order B-spline interpolation, and Ewald κ = 0.34 Å^{-1}. The van der Waals interactions are evaluated with the switch function to turn the interaction to zero smoothly between 9.0 and 10.0 Å. All nonbonded interaction lists are generated using the heuristic update option available in CHARMM. The acetone molecule is represented by the semiempirical AM1 method[7] as the low-level QM theory and the HF/3-21G method as the high-level QM theory, respectively. All water molecules are represented by the fixed-geometry TIP3P model. Long-range electrostatic interactions between the QM subsystem and the periodic MM and QM images are described by the SE-QM/MM-PME method (see the Theory section). All MD simulations are carried out with δt = 0.5 fs for the inner loop time integration, and N is varied between 1 and 8.

The second system is prepared to evaluate accuracy of the developed method in describing chemical reactions in water. The reaction simulated is the S_N2 nucleophilic substitution reaction between CH_3Cl and Cl^- in water. The solutes (CH_3Cl and Cl^-) are solvated with a rhombic dodecahedron box of 990 TIP3P waters (lattice length 35.0 Å). PME is evaluated with 40 × 40 × 40 FFT grid, the sixth order B-spline interpolation, and a 9.5 Å real space cutoff. The Ewald κ value was 0.34 Å^{-1}. The van der Waals interactions are evaluated with the switch function between 8.0 and 9.5 Å, and all nonbonded pair lists are generated heuristically. The solute molecules are represented by the AM1 method as the low-level theory and the HF/3-21G method as the high-level theory, respectively. The reaction coordinate (ξ) is defined as the distance difference between the cleaving bond (between C and Cl_1) and the forming bond (between C and Cl_2), in which Cl_1 and Cl_2 refer the leaving and nucleophilic Cl atoms, respectively. Since the reaction is symmetric, the umbrella sampling free energy simulations[48] are

dx.doi.org/10.1021/ct5005643 | J. Chem. Theory Comput. 2014, 10, 4175−4183

carried out in the range between −2.5 and 0.0 Å of the reaction coordinate. A total of 29 separate umbrella sampling simulations are carried out with 0.1 Å spacing between neighboring windows for -2.5 Å $\leq \xi <$ -0.3 and 0.05 Å spacing for -0.3 Å $\leq \xi \leq$ 0.0 Å, respectively. The harmonic force constant for the umbrella potential ranges between 60.0 kcal/mol Å^2 and 90.0 kcal/mol Å^2. Each simulation is run for 10 ps with δt = 0.5 fs, and the last 9.5 ps simulation is used to determine the potentials of mean force (PMF) by the weighted histogram analysis method (WHAM).[49]

4. RESULTS AND DISCUSSION

4.1. Energy Conservation. Stability of the multiple time step (MTS) AI-QM/MM-PME method is tested by microcanonical (NVE) ensemble MD simulations for an acetone molecule in water. All simulations are initiated at 300 K and run for 1.0 ps with δt = 0.5 fs. In Figure 1A, the deviation of the MTS energy from the STS energy at every time point is presented to compare the conservation of the total energy (kinetic energy plus potential energy). We also present the conservation of the total energy for different N values in Figure S1A, in which the STS AI-QM/MM-PME simulation is extended to 5 ps, and the conservation of the total energy using the B3LYP[50,51]/6-31G* level of theory as the high-level QM method in Figure S1B. The results show that the total energy of the MTS MD remains close to the energy of the single time step (STS; i.e., N = 1) MD during the simulated time (1.0 ps). The inset of Figure 1A presents the fluctuation of the total energy for each simulation, in which the fluctuation of the total energy is determined as

$$\langle \delta \rangle = \frac{1}{N_{\text{steps}}} \sum_{i=1}^{N_{\text{steps}}} |E_{\text{tot}}(i) - \langle E_{\text{tot}} \rangle| \tag{24}$$

where E_{tot} is the total energy, $\langle X \rangle$ the average of X over the length of the MD simulation, and N_{steps} the total number of MD steps, respectively. Although the total energy fluctuation increases slowly with the increase of N, the total energy does not drift systematically from the STS value as shown in Figures 1A and S1B, suggesting that the MTS AI-QM/MM-PME method produces stable MD trajectories and does not cause any systematic drift of the total energy.

In Figure 1B, we compare the total energy conservation for the Nose−Hoover canonical (NVT) ensemble MD simulations. The length and the initial conditions of the simulations are identical to the NVE simulation presented in Figure 1A. Since the total energy is not a conserved quantity in the Nose−Hoover MD, the fluctuation of the total energy is slightly larger for the NVT simulations than the results from the NVE MD simulations (see Figure 1B inset). Nevertheless, the MTS approach produces very similar energy to the corresponding STS results throughout the entire 1.0 ps MD simulation.

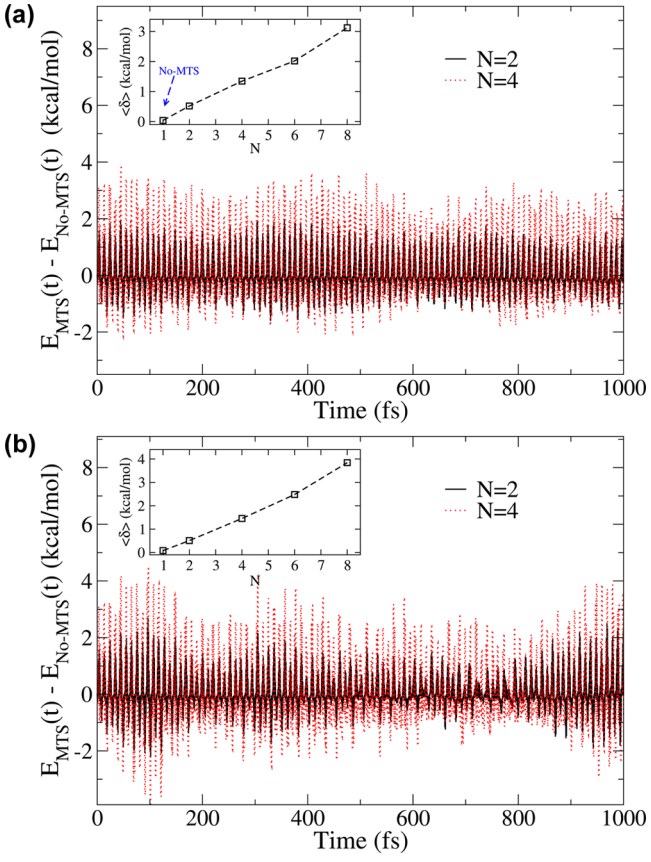


Figure 1. Conservation of the total energy (kinetic energy plus potential energy) of an acetone molecule in water (A) in the microcanonical (NVE) ensemble and (B) in the canonical (NVT) ensemble molecular dynamics (MD) simulations. In each figure, the deviation of the total energy of the multiple time step (MTS) AI-QM/MM-PME method from that of the single time step AI-QM/MM-PME method at each time point is plotted as a function of the simulation time. For clear presentation of the results, only the results with N = 2 and 4 are presented in both cases. The inset gives the total energy fluctuation (eq 24) as a function of N between 1 and 8 for both MD simulations. In all simulations, the acetone molecule is represented by the AM1 method as the low-level theory and the HF/3-21G method at the high-level theory, and water molecules are represented by the TIP3P water model, respectively.

4.2. Speed-up of the Multiple Time Step AI-QM/MM-PME Method. Two systems are prepared to test the relative speed-up of the multiple time step AI-QM/MM-PME method. The first system consists of 31 QM atoms (10 H atoms and 21 heavy atoms) and 2926 MM atoms. The QM subsystem here is represented by the B3LYP/6-31G level of theory as the high-level theory and the AM1 method as the low-level theory, respectively. The second system consists of 56 QM atoms (22 H atoms and 34 heavy atoms) and 28 843 MM atoms. The QM region in this case is represented by the B3LYP/3-21G level of theory as the high-level theory and the AM1 method as the low-level of theory, respectively. The speed-up is calculated based on 100 steps of MD simulation with δt = 0.5 fs. The results are presented in Figure 2. Although both systems scale almost linearly with the increase of N, there is also a slight slow-down at large N. In the tests, since the electron density determined from previous AI-QM/MM calculation (i.e., Δt time earlier) is used as the initial guess at each AI-QM/MM calculation, more SCF iterations are needed to reach full convergence for the simulations with a large N value. This results in a slow-down of the overall calculation.

To test the timing of the method, we have carried out the SE-QM/MM-PME calculations for the same systems to produce reference data. The (single time step) AI-QM/MM-PME calculations are more than 100-fold slower than the SE-QM/MM results for both systems (117-fold for the first system and 142-fold for the second system, respectively), whereas the N = 6 MTS AI-QM/MM-PME calculations are only 23-times slower for the 31 QM atom system and 30-times slower for

dx.doi.org/10.1021/ct5005643 | J. Chem. Theory Comput. 2014, 10, 4175−4183

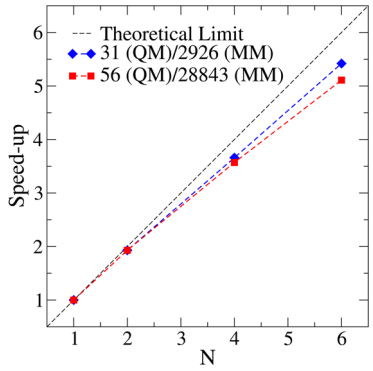


Figure 2. Relative speed-up of the multiple time step AI-QM/MM-PME methods. For each system, 100 steps of MD simulations are carried out for each N with δt = 0.5 fs as the time step for the inner loop integration. The N = 1 result is for the single time step calculation. The theoretical limit line represents the maximum speed-up that can be achieved by the MTS approach.

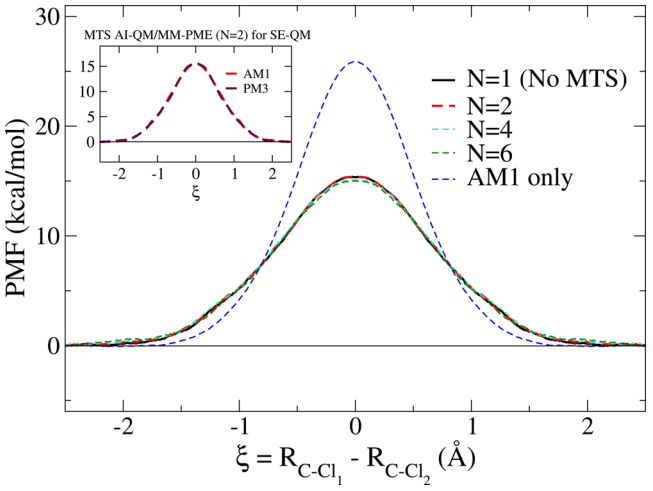


Figure 3. Potential of mean force (PMF) profiles of the S_N2 reaction between CH_3Cl and Cl^- in water. The results from the MTS simulations are indicated with their corresponding N values. The MTS simulations are carried out using the AM1 method as the low-level QM theory and the HF/3-21G method as the high-level QM theory. The "AM1 only" results are from the AM1 QM/MM-PME simulations. The inset compares effects of the low-level QM theory on the PMF profiles, in which the AM1 and PM3 methods are used as the low-level QM theory, respectively, and N = 2 for the MTS AI-QM/MM-PME calculations.

the 56 QM atom system than their corresponding SE-QM/MM-PME results, respectively. All these calculations are carried out using a single CPU core. Since AI-QM/DFT methods have, in general, better parallel scalability than the SE-QM methods,[39,52] additional speed-up is possible by carrying out the calculations using multiple CPUs. For example, the overall computational cost of the MTS AI-QM/MM-PME simulation (N = 6) using 16 CPU cores is only 4-times higher for both systems than the single core SE-QM/MM-PME results, which is a substantial improvement of speed compared to the single core STS AI-QM/MM results.

4.3. Reaction: Potential of Mean Force (PMF). We test the accuracy of the AI-QM/MM-PME method in describing chemical reactions by the umbrella sampling free energy (FE) simulations[48] of the S_N2 reaction between Cl^- and CH_3Cl in the water. The results are presented in Figure 3. The results show that the MTS approach produces PMF profiles that are very close to the STS AI-QM/MM-PME result. In Figure 3, we also show the result from the SE-QM/MM-PME calculations using the AM1 model as the SE-QM method, which yield a PMF profile that is very different from the AI-QM/MM-PME results. The difference between the SE-QM/MM-PME and the AI-QM/MM-PME results suggests that the inaccuracy of the underlying SE-QM method is well corrected by the present multiscale approach. As a further test of this, we carried out the umbrella sampling MTS AI-QM/MM-PME simulations (with N = 2) using the PM3 method as the low-level SE-QM method. The results are presented in the inset of Figure 3. The results show that both MTS simulations yield essentially identical results, confirming that the results are independent of the underlying low-level SE-QM method. This occurs partly because the long-range interactions are less sensitive to the detailed changes of the electron density of the QM subsystem as far as the low-level theory produces reasonable results at long-range. This appears to be the case for the AM1 and PM3 methods.[53,54] In Figure S2, we also present the PMF results using different AI-QM/DFT methods and different real space cutoff distance values, showing the results are highly dependent on the level of AI-QM/DFT method but less dependent on cutoff distance, which is desired. Together, these results support the reliability of the developed multiple time step AI-QM/MM-PME method and the assumption in eq 9. This is important because a direct comparison of the present method to the "real" AI-QM/MM-PME method is not possible due to the absence of such a method. Although there is the method developed by Holden et al.,[32] several differences between the two methods, such as the use of the regular Ewald versus the PME methods and the ChElPG versus the Mulliken charges, make a direct comparison between them less meaningful.

Interestingly, the MTS simulations produce essentially identical PMF profiles, and only small differences are observed between simulations with different N values. For example, the N = 6 simulations yield the barrier height only lower by 0.37 kcal/mol than the results of the STS simulations, the largest difference among the tested MTS simulations (0.00 kcal/mol for N = 2 and 0.26 kcal/mol for N = 4, respectively). Although the errors appear to be systematic (i.e., the error increases as N increases), they are likely within the range of the sampling errors of the present FE simulations (10 ps for each umbrella sampling window). The present results show that the developed MTS AI-QM/MM-PME methods are promising for simulation studies of chemical reactions that occur in enzymes and in water, in which given the large speed-up of the MTS approach, the length of the MD simulations can be extended substantially longer than that of the simulations that are commonly carried out with the conventional AI-QM/MM approach. In addition, since the present method includes the periodic boundary long-range interactions explicitly in the total energy, the resulting MD simulations are expected to be more reliable than the simulations with cutoffs.

5. CONCLUSIONS

The present letter describes the implementation and performance of the multiscale ab initio (AI) QM/MM method combined with the reversible reference system propagator algorithm (rRESPA) for efficient periodic boundary molecular dynamics (MD) simulations. In the developed method, the long-range QM/MM electrostatic interactions are evaluated at

the semiempirical (SE) QM/MM-PME level of theory, and the resulting energy is corrected by carrying out the AI-QM/MM calculation at short-range. This multiscale separation of the total energy into the low-level energy and the high-level correction energy enables the decomposition of the total atomic forces into two corresponding force terms and the application of the rRESPA multiple time step technique. The rRESPA technique, in particular, allows us to minimize the number of the time-consuming AI-QM/MM calculations during MD simulations, resulting in a substantial speed-up of the entire calculation. Test calculations show that the resulting MD simulations are highly stable and accurate compared to the single time step MD simulation results. Taken together, the present multiple time step multiscale AI-QM/MM method enables long AI-QM/MM molecular dynamics simulations of large biological systems under periodic boundary conditions, with a fraction of time that is needed for conventional AI-QM/MM simulations. Finally, since the large acceleration of the present method is achieved mainly by minimizing the number of time-consuming AI-QM/MM calculations, we expect to achieve further speed-up of the method by applying algorithms enabling faster SCF calculations, such as the Fock-matrix extrapolation[55,56] or the Lagrangian Born−Oppenheimer MD algorithm.[57]

APPENDIX A: DERIVATION OF THE QM/MM-PME METHOD

The periodic boundary correction term, $\Delta^{\mathrm{PB}}_{\mathrm{QM/MM}}[q^{\mathrm{QM}};q^{\mathrm{MM}}]$, given in eq 10 can be decomposed into two terms

$$\Delta E^{\mathrm{PB}}_{\mathrm{QM/MM}}[q^{\mathrm{QM}};\, q^{\mathrm{MM}}] = \Delta E^{\mathrm{PB}}_{\mathrm{QM-QM}}[q^{\mathrm{QM}};\, q^{\mathrm{QM}}] + \Delta E^{\mathrm{PB}}_{\mathrm{QM-MM}}[q^{\mathrm{QM}};\, q^{\mathrm{MM}}] \tag{A1}$$

where q^{QM} and q^{MM} denote the Mulliken charges[30] of the QM subsystem and the classical MM point charges, respectively. The first term corresponds to the long-range interactions between the QM atoms in the primary cell and all periodic QM image charges, and the second term the interactions between the QM atoms in the primary cell and all MM charges outside of the real space cutoff distance and all periodic MM images, respectively. Each term in eq A1 can be further decomposed into a real-space correction term and a reciprocal-space term. They are

$$\Delta E^{\mathrm{PB}}_{\mathrm{QM-QM}} = \Delta E^{\mathrm{recip}}_{\mathrm{QM-QM}} + \Delta E^{\mathrm{real}}_{\mathrm{QM-QM}} \tag{A2}$$

and

$$\Delta E^{\mathrm{PB}}_{\mathrm{QM-MM}} = \Delta E^{\mathrm{recip}}_{\mathrm{QM-MM}} + \Delta E^{\mathrm{real}}_{\mathrm{QM-MM}} \tag{A3}$$

respectively. The definition of each term in eqs A2 and A3 is

$$\Delta E^{\mathrm{recip}}_{\mathrm{QM-QM}} = \frac{1}{2}\sum_{\alpha=1}^{N^{\mathrm{QM}}}\sum_{\beta=1}^{N^{\mathrm{QM}}} q_\alpha^{\mathrm{QM}} q_\beta^{\mathrm{QM}} \times \left[\frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}\exp(2\pi i\mathbf{m}\cdot\mathbf{R}_{\alpha\beta})\right] \tag{A4}$$

$$\Delta E^{\mathrm{real}}_{\mathrm{QM-QM}} = -\frac{1}{2}\sum_{\alpha=1}^{N^{\mathrm{QM}}}\sum_{\beta=1}^{N^{\mathrm{QM}}} q_\alpha^{\mathrm{QM}} q_\beta^{\mathrm{QM}}\left[\frac{\mathrm{erf}(\kappa|\mathbf{R}_{\alpha\beta}|)}{|\mathbf{R}_{\alpha\beta}|}\right] \tag{A5}$$

$$\Delta E^{\mathrm{recip}}_{\mathrm{QM-MM}} = \sum_{\alpha=1}^{N^{\mathrm{QM}}}\sum_{j=1}^{N^{\mathrm{MM}}} q_\alpha^{\mathrm{QM}} q_j^{\mathrm{MM}} \times \left[\frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}\exp(2\pi i\mathbf{m}\cdot\mathbf{r}_{\alpha j})\right] \tag{A6}$$

and

$$\Delta E^{\mathrm{real}}_{\mathrm{QM-MM}} = -\sum_{\alpha=1}^{N^{\mathrm{QM}}}\sum_{j=1}^{N^{MM}} \mathrm{q}_\alpha^{\mathrm{QM}} q_j^{\mathrm{MM}}\left[\frac{\mathrm{erf}(\kappa|\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|}\right] \tag{A7}$$

where V is a total volume of the system, $\mathbf{R}_{\alpha\beta} = \mathbf{R}_\beta - \mathbf{R}_\alpha$ and $\mathbf{r}_{\alpha j} = \mathbf{r}_j - \mathbf{R}_\alpha$, in which $\mathbf{R}$ and $\mathbf{r}$ denote the position of the QM and MM atoms, respectively. The Ewald parameter κ controls the convergence of the real space and the reciprocal space terms in eqs 1 and 10 and eqs A4 to A7, and $\mathbf{m} = m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^* + m_3\mathbf{a}_3^*$ for integers m_k (k = 1,2,3), where $\mathbf{a}_1^*$, $\mathbf{a}_2^*$, and $\mathbf{a}_3^*$ are the reciprocal-space lattice vectors.

If eqs A2 to A7 are used to determine the Ewald periodic boundary correction term (i.e., eq A1), the resulting method is the QM/MM-Ewald method.[22] To derive the QM/MM-PME method, eq A6 is rearranged to

$$\Delta E^{\mathrm{recip}}_{\mathrm{QM-MM}} = \frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2} S_{\mathrm{QM}}(\mathbf{m})S_{\mathrm{MM}}(-\mathbf{m}) \tag{A8}$$

where the structure factors of the QM and MM subsystems are

$$S_{\mathrm{QM}}(\mathbf{m}) = \sum_{\alpha=1}^{N^{\mathrm{QM}}} q_\alpha^{\mathrm{QM}}\exp(2\pi i\mathbf{m}\cdot\mathbf{R}_\alpha) \tag{A9}$$

and

$$S_{\mathrm{MM}}(-\mathbf{m}) = \sum_{j=1}^{N^{\mathrm{MM}}} q_j^{\mathrm{MM}}\exp(-2\pi i\mathbf{m}\cdot\mathbf{r}_j) \tag{A10}$$

respectively.

To apply the Cardinal B-spline interpolation of order n with FFT grid $K_1 \times K_2 \times K_3$, the exponential function of eqs A9 and A10 is expanded to

$$\exp(2\pi i\mathbf{m}\cdot\mathbf{r}) = \exp\left(2\pi i\frac{m_1 f_1}{K_1}\right)\exp\left(2\pi i\frac{m_2 f_2}{K_2}\right)\exp\left(2\pi i\frac{m_3 f_3}{K_3}\right) \tag{A11}$$

where $f_\alpha = K_\alpha\mathbf{a}_\alpha^*\cdot\mathbf{r}$ is a scaled fractional coordinate for α = 1, 2, 3. In the derivation, we follow closely the notation introduced by Essmann et al.[34] For even n, each component in eq A11 can be determined approximately by

$$\exp\left(2\pi i\frac{m_\alpha f_\alpha}{K_\alpha}\right) \approx b_\alpha(m_\alpha)\sum_{k=-\infty}^{\infty} M_n(f_\alpha - k)\exp\left(2\pi i\frac{m_\alpha k}{K_\alpha}\right) \tag{A12}$$

In the equation,

dx.doi.org/10.1021/ct5005643 | J. Chem. Theory Comput. 2014, 10, 4175−4183

$$b_\alpha(m_\alpha) = \exp\left(2\pi i(n-1)\frac{m_\alpha}{K_\alpha}\right) \times \left[\sum_{k=0}^{n-2} M_n(k+1)\exp\left(2\pi i\frac{m_\alpha k}{K_\alpha}\right)\right]^{-1} \quad \text{(A13)}$$

where $M_n(x)$ is the nth order Cardinal B-spline function of a variable x and is determined by a recursion relation

$$M_n(x) = \frac{x}{n-1}M_{n-1}(x) + \frac{n-x}{n-1}M_{n-1}(x-1) \quad \text{(A14)}$$

and

$$M_2(x) = \begin{cases} 1-|x-1|, & \text{for } 0 \le x \le 2 \\ 0, & \text{for } x < 0 \text{ or } x > 2 \end{cases} \quad \text{(A15)}$$

Given the B-spline interpolation of eq A11, the structure factor of eqs A9 and A10 is written as

$$S_X(\mathbf{m}) = b_1(m_1)b_2(m_2)b_3(m_3)F(Q_X)(\mathbf{m}) \quad \text{(A16)}$$

where X = QM or MM, and $F(A)$ denotes the discrete FFT of an array A. The charge array Q_X is given by,

$$Q_X(\mathbf{m}) = \sum_j^{N^X} q_j^X\Big[\sum_{n_1,n_2,n_3} M_n(s_{j1}-n_1K_1)M_n(s_{j2}-n_2K_2)M_n(s_{j3}-n_3K_3)\Big] \quad \text{(A18)}$$

where $s_{j\alpha} = f_{j\alpha} - m_\alpha$ for $\alpha = 1, 2, 3$, and X is again QM or MM, respectively. With the definition of Q and $S(\mathbf{m})$, we rewrite $E^{\text{recip}}_{\text{QM-MM}}$ in eq A8 as

$$\Delta E^{\text{recip}}_{\text{QM-MM}} = \frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}B(\mathbf{m})F(Q_{QM})(\mathbf{m}) \times F(Q_{MM})(-\mathbf{m}) = \sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1} Q_{QM}(\mathbf{m})[\theta_{PB}{}^*Q_{MM}](\mathbf{m}) \quad \text{(A19)}$$

where

$$B(\mathbf{m}) = |b_1(m_1)|^2|b_2(m_2)|^2|b_3(m_3)|^2 \quad \text{(A20)}$$

$$\theta_{PB} = F\left(\frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}B(\mathbf{m})\right) \quad \text{(A21)}$$

and A^*B denotes a convolution A of B.

Equation A19 is the PME expression of eq A6. Although it is possible to apply a similar approach to eq A4, it would be less efficient than evaluating eq A4 explicitly, because the PME version of eq A4 requires discrete FFT at each SCF cycle, which is time-consuming. Therefore, we use the regular Ewald sum method for the QM-QM interactions given in eq A4. A similar approach was taken by Walker et al.[26] Since the number of QM atoms is much smaller than N^{MM} in most QM/MM calculations, the computational overhead for evaluating eq A4 is negligible compared to the computational cost of solving eq A19.

APPENDIX B: DISCUSSION OF THE QM/MM-PME METHOD

To facilitate the discussion of the QM/MM-PME method presented in Appendix A, we rewrite ΔE^{PB}_{QM-QM} and ΔE^{PB}_{QM-MM} using a pair potential $\Delta\omega$ and a potential $\Delta\psi$, i.e.,

$$\Delta E^{\text{PB}}_{\text{QM-QM}} = \frac{1}{2}\sum_{\alpha=1}^{N^{\text{QM}}}\sum_{\beta=1}^{N^{\text{QM}}} q_\alpha^{\text{QM}}q_\beta^{\text{QM}}\Delta\omega(\mathbf{R}_\alpha, \mathbf{R}_\beta) \quad \text{(B1)}$$

and

$$\Delta E^{\text{PB}}_{\text{QM-MM}} = \sum_{\alpha=1}^{N^{\text{QM}}} q_\alpha^{\text{QM}}\Delta\psi(\mathbf{R}_\alpha) \quad \text{(B2)}$$

In these equations, eqs A4,A5,A7, and A19 are used to define $\Delta\omega$ and $\Delta\psi$:

$$\Delta\omega(\mathbf{R}_\alpha, \mathbf{R}_\beta) = \frac{1}{\pi V}\sum_{\mathbf{m}\neq 0}\frac{\exp(-\pi^2\mathbf{m}^2/\kappa^2)}{\mathbf{m}^2}\exp(2\pi i\mathbf{m}\cdot\mathbf{R}_{\alpha\beta}) - \frac{\text{erf}(\kappa|\mathbf{R}_{\alpha\beta}|)}{|\mathbf{R}_{\alpha\beta}|} \quad \text{(B3)}$$

and

$$\Delta\psi(\mathbf{R}_\alpha) = \sum_{m_1=0}^{K_1-1}\sum_{m_2=0}^{K_2-1}\sum_{m_3=0}^{K_3-1} Q^\alpha_{\text{QM}}(\mathbf{m})[\theta_{\text{PB}}{}^*Q_{MM}](\mathbf{m}) - \sum_{j=1}^{N^{\text{MM}}} q_j^{MM}\left[\frac{\text{erf}(\kappa|\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|}\right] \quad \text{(B4)}$$

where

$$Q^\alpha_{\text{QM}}(\mathbf{m}) = \sum_{n_1,n_2,n_3} M_n(s_{\alpha 1}-n_1K_1)M_n(s_{\alpha 2}-n_2K_2)M_n(s_{\alpha 3}-n_3K_3) \quad \text{(B5)}$$

Here, the pair potential $\Delta\omega$ is used for the QM−QM interaction, because the QM charges have to be recalculated at each SCF step. The pair potential is computed and stored as a $N^{\text{QM}} \times N^{\text{QM}}$ matrix at the beginning of each energy calculation. The energy term, $\Delta E^{\text{PB}}_{\text{QM-QM}}$, is then updated by a simple matrix multiplication with the Mulliken charge vector at each SCF cycle. Since $N^{\text{QM}} \ll N^{\text{MM}}$ in most QM/MM calculations, the computational cost of eq B1 is relatively small. For the QM-MM interaction given in eq B2, the MM charge distribution does not change during the SCF cycle. Therefore, $\Delta\psi$ at each QM site is precomputed and stored at the beginning of each energy calculation. The update of the $\Delta E^{\text{PB}}_{\text{QM-MM}}$ energy is then carried out at each SCF cycle by a simple dot-product operation between the QM Mulliken charge vector and the $\Delta\psi$ vector. Last, since the Mulliken charges are used in the present work to represent the QM charge distribution, the Fock matrix correction term corresponding to each term in eqs B1 and B2 can be determined straightforwardly by a chain rule, as was suggested by Nam et al.[22]

ASSOCIATED CONTENT

Supporting Information

Figures presenting the performance of the presented method with the B3LYP level of theory and different real-space cutoff distances. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: kwangho.nam@chem.umu.se.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the grant from the Umeå University, Umeå, Sweden. We acknowledge Dr. Gordon Driver for carefully reading the manuscript. Computational resources were provided by the National Energy Research Scientific Computing Center (NERSC) and the Swedish National Infrastructure for Computing (SNIC) at High Performance Computing Center North (HPC2N) and National Supercomputer Centre (NSC) at Sweden.

REFERENCES

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(2) Gao, J. *Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials*; VCH: New York, 1996; Vol. 7.
(3) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(4) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573.
(5) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. *Angew. Chem.* **2006**, *118*, 7010.
(6) McGeagh, J. D.; Ranaghan, K. E.; Mulholland, A. J. *Biochim. Biophys. Acta* **2011**, *1814*, 1077.
(7) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(8) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
(9) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
(10) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.
(11) Frauenheim, Th.; Seifert, G.; Elsterner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. *Phys. Status Solidi B* **2000**, *217*, 41.
(12) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186.
(13) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188.
(14) York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345.
(15) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155.
(16) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210.
(17) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*, 4th ed.; Wiley-VCH: Weinheim, Germany, 2010.
(18) Hu, H.; Lu, Z.; Parks, J. M.; Burger, S. K.; Yang, W. *J. Chem. Phys.* **2008**, *128*, 034105.
(19) Hu, H.; Lu, Z.; Yang, W. *J. Chem. Theory Comput.* **2007**, *3*, 390.
(20) Zhang, Y.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459.
(21) Yang, J.; Lior-Hoffmann, L.; Wang, S.; Zhang, Y.; Broyde, S. *Biochemistry* **2013**, *52*, 2828.
(22) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 2.
(23) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. Lond. A* **1980**, *373*, 27.
(24) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(25) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(26) Walker, R. C.; Crowley, M. F.; Case, D. A. *J. Comput. Chem.* **2008**, *29*, 1019.
(27) Gao, J.; Alhambra, C. *J. Chem. Phys.* **1997**, *107*, 1212.
(28) Dehez, F.; Martins-Costa, M. T. C.; Rinaldi, D.; Millot, C. *J. Chem. Phys.* **2005**, *122*, 234503.
(29) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715.
(30) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.
(31) Jacobson, L. D.; Herbert, J. M. *J. Chem. Phys.* **2011**, *134*, 094118.
(32) Holden, Z. C.; Richard, R. M.; Herbert, J. M. *J. Chem. Phys.* **2013**, *139*, 244108.
(33) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.
(34) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(35) Tuckerman, M.; Bern, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.
(36) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(37) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2001**, *114*, 9758.
(38) Herbert, J. M.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 11542.
(39) Nam, K. *J. Chem. Theory Comput.* **2013**, *9*, 3393.
(40) Vreven, T.; Morokuma, K. *J. Chem. Phys.* **2000**, *113*, 2969.
(41) Marti, S.; Moliner, V.; Tunon, I.; Williams, I. H. *J. Phys. Chem. B* **2005**, *109*, 3707.
(42) Steele, R. P. *J. Chem. Phys.* **2013**, *139*, 011102.
(43) Luehr, N.; Markland, T. E.; Martinez, T. J. *J. Chem. Phys.* **2014**, *140*, 084116.
(44) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
(45) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.
(46) Guest, M. F.; Bush, I. J.; Van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. *Mol. Phys.* **2005**, *103*, 719.
(47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(48) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187.
(49) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011.
(50) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(51) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(52) Berzigiyarov, P. K.; Zayets, V. A.; Ginzburg, I. Y.; Razumov, V. F.; Sheka, E. F. *Int. J. Quantum Chem.* **2002**, *88*, 449.
(53) Ferenczy, G. G.; Reynolds, C. A.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 159.
(54) Bakowies, D.; Thiel, W. *J. Comput. Chem.* **1996**, *17*, 87.
(55) Pulay, P.; Fogarasi, G. *Chem. Phys. Lett.* **2004**, *386*, 272.
(56) Herbert, J. M.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3269.
(57) Zheng, G.; Niklasson, A. M. N.; Karplus, M. *J. Chem. Phys.* **2011**, *135*, 044122.