JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Acceleration of Semiempirical QM/MM Methods through Message Passage Interface (MPI), Hybrid MPI/Open Multiprocessing, and Self-Consistent Field Accelerator Implementations

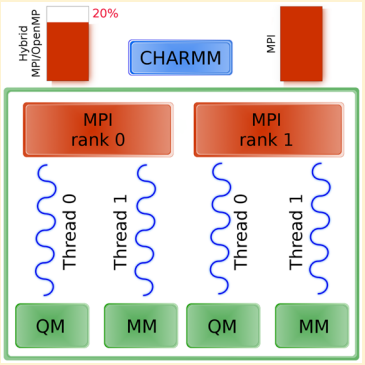
Pedro Ojeda-May*,†,‡ and Kwangho Nam*,†,§

†Department of Chemistry and ‡High Performance Computing Center North (HPC2N), Umeå University, Umeå SE-901 87, Sweden

§Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, Texas 76019-0065, United States

Supporting Information

ABSTRACT: The strategy and implementation of scalable and efficient semiempirical (SE) QM/MM methods in CHARMM are described. The serial version of the code was first profiled to identify routines that required parallelization. Afterward, the code was parallelized and accelerated with three approaches. The first approach was the parallelization of the entire QM/MM routines, including the Fock matrix diagonalization routines, using the CHARMM message passage interface (MPI) machinery. In the second approach, two different self-consistent field (SCF) energy convergence accelerators were implemented using density and Fock matrices as targets for their extrapolations in the SCF procedure. In the third approach, the entire QM/MM and MM energy routines were accelerated by implementing the hybrid MPI/open multiprocessing (OpenMP) model in which both the task- and loop-level parallelization strategies were adopted to balance loads between different OpenMP threads. The present implementation was tested on two solvated enzyme systems (including <100 QM atoms) and an S_N2 symmetric reaction in water. The MPI version exceeded existing SE QM methods in CHARMM, which include the SCC-DFTB and SQUANTUM methods, by at least 4-fold. The use of SCF convergence accelerators further accelerated the code by ∼12−35% depending on the size of the QM region and the number of CPU cores used. Although the MPI version displayed good scalability, the performance was diminished for large numbers of MPI processes due to the overhead associated with MPI communications between nodes. This issue was partially overcome by the hybrid MPI/OpenMP approach which displayed a better scalability for a larger number of CPU cores (up to 64 CPUs in the tested systems).



INTRODUCTION

The combined quantum mechanical and molecular mechanical (QM/MM) methods are one of the most efficient approaches to simulate enzymatic reactions in solvated environments.[1−4] In these methods, highly localized regions of the enzyme involved in the reaction, typically containing substrates and catalytic residues, are described by the ab initio (AI), density functional theory (DFT), or semiempirical (SE) QM methods, and the remainder of the enzyme and surrounding solvents are described by MM methods. Despite being efficient, however, the time scales sampled by these QM/MM methods are relatively too short compared with the time scales reached by the classical MM simulations: i.e., ≪1 ns/day for SE-QM methods and ≪10 ps/day for AI/DFT-QM methods[5] versus >50 ns/day for classical MM methods.[6] This makes it necessary to develop efficient and scalable QM/MM methods to enable long-time molecular dynamics (MD) simulations, which are essential, for instance, to explore the free energy landscape of enzymatic catalysis[7,8] while preserving the accuracy to reproduce the experimentally measured catalytic activities.

To understand what factors affect the performance of the QM/MM methods, both the MM and QM Hamiltonians are compared below to observe the difference in their methodologies and implementations using the serial and parallel programming models. In the MM calculation, the most time-consuming part is the evaluation of long-range interactions, for which efficient algorithms have been developed[9−12] and parallelized. The two most widely adopted parallel programming models are the message passage interface (MPI)[13,14] and the open multiprocessing (OpenMP)[15] protocols. These models are implemented in most modern molecular mechanics simulation packages, such as GROMACS,[16] LAMMPS,[17] AMBER,[18] CHARMM,[19] and NAMD.[20] Recently, an increasing number of MM packages also adopt the hybrid MPI/OpenMP programming model for better use of shared memory multicore and many-core CPU architectures.

In performing the periodic boundary MD simulations, the most widely adopted algorithms for MM electrostatic interactions are the Ewald summation method[9] and its variants.[10,21,22] In these methods, the MM electrostatic interactions are divided into two terms, namely, the real and

Received: March 29, 2017
Published: June 19, 2017

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

reciprocal space terms. The real space term is the sum of pairwise interactions that are evaluated explicitly up to a certain cutoff distance, and the reciprocal space term is computed in the k−space (i.e., the reciprocal space) with the classical Ewald[9] or particle mesh Ewald (PME) summation technique.[10] The PME method was also generalized to higher orders to include the long-range polarization and dispersion interactions.[21,23−25]

In performing the QM/MM calculations, although both the AI[26] and DFT[27] methods are the "gold" standard because of their high accuracy, their large computational expenses limit their use in long QM/MM MD simulations, making a typical length of the simulation to be less than 100 ps. For this reason, when studying catalytic mechanisms of enzymes, the AI/DFT QM/MM calculations are carried out mostly through optimizations of transition states[28] or reaction paths[29−31] with the assumption that entropy contributes little to the determined reaction mechanism. Although generally less accurate, the SE QM methods are one cheaper alternative to the AI/DFT methods, where computer-intensive electron integrals are replaced with relatively simpler functions with empirical parameters. The parameters are obtained by calibrating the methods against experiments or higher levels of QM theory. In this way, the SE QM methods outperform the AI/DFT methods in speed by 2−3 orders of magnitude. In addition, the methods can be reparameterized to achieve higher accuracy for specific reactions[29,31−38] These make the SE QM methods very attractive for performing long QM/MM MD simulations. Some successful examples include the study of catalytic mechanisms of enzymes[37,39,40] and ribozymes.[33,41]

Unlike the MM methods, however, the SE QM methods have been challenging to achieve a high degree of parallelization.[42−44] One major reason is the iterative nature of the self-consistent field (SCF) procedure of the QM methods, which is necessary to attain the converged electronic energy. Because each routine involved in the SCF iteration has to be executed sequentially, each of them has to be parallelized individually for the entire procedure to be scalable. This makes developing scalable SE QM/MM methods very challenging. Methods such as the extended Lagrangian methods[45−48] have been developed to avoid the need for SCF iterations by propagating density matrix elements as extra dynamical variables. These methods, however, require the simulations to be performed with reduced time steps to achieve a good energy conservation while increasing the overall computational cost.

In the present manuscript, we have developed efficient and scalable SE QM/MM methods based on the general-purpose modified neglect of diatomic overlap 97 (MNDO97) suite of programs.[49,50] The MNDO97 program was chosen because of its versatility supporting various SE QM methods, including the MNDO/d[51] and AM1/d methods.[33,41] These methods are important for studying enzymatic reactions involving transition metal cofactors and/or phosphoryl group transfers[8,33] in which the d orbitals play important roles. As a first step, a light version of the MNDO97 program was developed by rewriting the entire MNDO97 routines that are essential for the energy and gradient evaluations and integrating them with the main CHARMM code. Then, the QM/MM-PME method[50,52] was implemented for efficient evaluation of the long-range electrostatic interactions for both the QM and MM parts.

Finally, this light version of MNDO97 was accelerated with three approaches. In the first approach, the MPI parallelization was introduced to the SE QM/MM module by following the MPI implementation in CHARMM for consistency with other CHARMM parallel routines. In the second approach, in addition to the conventional direct inversion in the iterative subspace (DIIS)[53,54] extrapolation and pseudo-diagonalization algorithms,[55] two additional SCF convergence accelerators were implemented to reduce the total number of SCF iterations. They are the extended Lagrangian formulation with dissipation in the Born−Oppenheimer approximation (hereafter, DXL-BOMD)[56,57] and a variant of the Fock matrix dynamics (FD) method.[58,59] In DXL-BOMD, the auxiliary density, which approximates the fully converged electron density, is extrapolated and used as the guess density for each SCF calculation. By employing this method, the SCF iterations can be truncated to a predefined number of steps without sacrificing the total energy conservation. In FD, the Fock matrix is extrapolated directly at each MD step. The SCF iterations are then performed to the full SCF convergence as in the conventional SCF QM calculation.

The third approach, which was adopted to further improve the performance of the parallel SE QM/MM methods, was through balancing the workloads between processors and minimizing the network-intensive MPI communications in several routines involved in the QM/MM energy and gradient evaluations. Specifically, the OpenMP parallelization was introduced on top of the CHARMM MPI implementation, i.e., the hybrid MPI/OpenMP (MPI/OMP) parallelization, to the QM/MM energy routines. On the other hand, the remainder of the program, including the bonded and nonboned MM energies and QM/MM gradient routines, was parallelized by the task-level parallelizations. Because of this limitation, we allowed only two OpenMP threads per MPI in the present implementation.

The presently developed methods were tested using three different systems: (1) a fully solvated adenylate kinase (AdK), (2) insuline receptor kinase (IRK), and (3) the symmetric S_N2 reaction between Cl^- and CH_3Cl[36,52] in explicit waters. The first two systems were used to perform detailed performance and energy conservation analyses among the MPI, DXL-BOMD/FD accelerators, and hybrid MPI/OMP models. The third system was used to compare the reaction free energy profiles between the pure MPI and hybrid MPI/OMP models. Details of each system preparation and simulations are provided in the Computational Details section.

THEORY

Semiempirical Self-Consistent Field (SCF) Molecular Orbital (MO) Method. In the neglect of diatomic differential overlap (NDDO)-based SE methods, the total electronic energy of the QM system is determined by solving the Roothaan−Hall equation[60]

$$\mathbf{FC} = \mathbf{CE} \tag{1}$$

where $\mathbf{F}$ is the Fock matrix, and $\mathbf{C}$ and $\mathbf{E}$ are the MO coefficient and energy diagonal matrices, respectively. The Fock matrix can be written as[48,61]

$$\mathbf{F} = \mathbf{H}_{\mathrm{core}} + \mathbf{G}(\mathbf{P}) \tag{2}$$

where $\mathbf{H}_{\mathrm{core}}$ is the core Hamiltonian that includes one-electron contributions and $\mathbf{G}$ is the two-electron repulsion contributions. In the equation, $\mathbf{P}$ is the electron density matrix.

In the restricted closed-shell Hartree−Fock system, the electron density matrix is defined as

$$\mathbf{P} = 2\mathbf{C}_{\mathrm{occ}}\mathbf{C}_{\mathrm{occ}}^{\mathrm{T}} \tag{3}$$

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

where $\mathbf{C}_{occ}$ is the submatrix of the MO coefficient matrix, containing only the occupied MO contributions, and 2 implies that all occupied orbitals are doubly occupied. In the SCF MO procedure, because **G** depends explicitly on the density matrix **P**, eqs 1−3 are solved iteratively starting from an initial guess of **P** until the full convergence of the total electronic energy of the QM system. Then, the total electronic energy is determined by

$$E_{el} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}(H_{\mu\nu}^{core} + F_{\mu\nu}) \tag{4}$$

and the total energy of the QM system is obtained by adding the core−core repulsion energy as,

$$E_{QM}^{tot} = E_{el} + \sum_{A<B} E_{AB}^{core} \tag{5}$$

Periodic Boundary QM/MM Potential Energy Function. On the basis of the work by Nam et al.,[50] the total QM/MM energy under the periodic boundary conditions (PBC) can be written as

$$E_{QM/MM}^{tot} = E_{QM/MM}^{RS}[\rho; q] + E_{MM}[q; q] + \Delta E_{QM/MM}^{PB}[Q; q] \tag{6}$$

where ρ denotes the total charge density of the QM system including both the electron and nuclear charge distributions, q the MM point charges, and Q is the point charge mapped onto the QM atoms, such as the Mulliken charges.[62] The first term in eq 6, which is referred to as the real space term, is composed of two contributions

$$E_{QM/MM}^{RS}[\rho; q] = E_{QM-QM}^{RS}[\rho; \rho] + E_{QM-MM}^{RS}[\rho; q] \tag{7}$$

where the first term represents the SCF energy of the QM region (eq 5), and the second term is the (real-space) QM−MM interaction evaluated up to a certain cutoff distance. Conventionally, the QM−MM interactions have both the electrostatic and van der Waals interaction contributions as

$$E_{QM-MM}^{RS}[\rho; q] = E_{QM-MM}^{elec}[\rho; q] + E_{QM-MM}^{vdW} \tag{8}$$

where

$$E_{QM-MM}^{elec}[\rho; q] = \sum_{g}\sum_{G}\sum_{i\in g}\sum_{j\in G}\int \frac{q_i\rho(\mathbf{R}'_j)}{|\mathbf{r}_i - \mathbf{R}'_j|}\, d^3\mathbf{R}'_j \tag{9}$$

and

$$E_{QM-MM}^{vdW} = \sum_{g}\sum_{G}\sum_{i\in g}\sum_{j\in G}\left[\frac{A_{ij}}{|\mathbf{r}_i - \mathbf{R}_j|^{12}} - \frac{B_{ij}}{|\mathbf{r}_i - \mathbf{R}_j|^{6}}\right] \tag{10}$$

In eqs 9 and 10, g and G denote the MM and QM groups (see below), and **r** and **R** refer to the positions of the MM and QM atoms, respectively. In the present implementation, the so-called "group-based" cutoff scheme of the CHARMM energy function[63] was adopted for both the MM−MM and QM−MM nonbonded interactions. Namely, two groups, each enclosing a set of predefined QM or MM atoms constituting a functional unit in residue or molecule, are considered to interact if their centers of mass are within the real-space cutoff distance. However, because of the delocalized nature of the QM electron density, details regarding how the QM and MM groups interact are different from the way that MM groups interact with each other. In the MM−MM interactions, both the electrostatic and van der Waals interactions are evaluated only between the group pairs that are within the cutoff. In the QM−MM electrostatic interaction, any MM group that is within the cutoff distance from any QM group interact not only with that QM group but also with all other QM groups. On the other hand, the QM−MM van der Waals interactions are evaluated in the same way that the MM−MM interactions are evaluated.

The second term in eq 6 is the classical MM potential energy. The last term in eq 6 represents the PBC correction due to the periodic boundary Coulombic interactions that are not included in eq 9. This term is further decomposed into two terms

$$\Delta E_{QM/MM}^{PB}[Q; q] = \Delta E_{QM-QM}^{PB}[Q; Q] + \Delta E_{QM-MM}^{PB}[Q; q] \tag{11}$$

In this decomposition, the first term accounts for the interactions between the QM atoms and all QM images both described by the point charges. The second term includes all QM−MM electrostatic interactions between the QM and all MM atoms that are outside of the real space cutoff, including all MM images. In the method proposed by Nam et al.,[50] these two terms are evaluated as the sum of the reciprocal space contribution and the real-space correction, i.e.,

$$\Delta E_{QM-QM}^{PB}[Q; Q] = \Delta E_{QM-QM}^{real}[Q; Q] + E_{QM-QM}^{recip}[Q; Q] \tag{12}$$

and

$$\Delta E_{QM-MM}^{PB}[Q; q] = \Delta E_{QM-MM}^{real}[Q; q] + E_{QM-MM}^{recip}[Q; q] \tag{13}$$

The reciprocal space terms in both equations are solved either with the conventional Ewald summation method, resulting in the QM/MM-Ewald method,[50] or with the PME method, yielding the QM/MM-PME method.[52,64] When using the QM/MM-PME method, the only exception to note is the QM−QM pair interactions (eq 12), which are solved using the standard Ewald summation method,[9,50,52] because their evaluation via the PME method requires multiple FFTs during the SCF iteration. We refer to refs 50 and 52 for detailed descriptions of the two methods.

QM/MM Long-Range Correction Using a Switching Function. Because of the approximation introduced to the QM/MM-Ewald and QM/MM-PME methods to represent the QM charge density with atomic charges, the energy given in eq 6 experiences a slight discontinuity at the cutoff distance, which results in a slow energy drift during MD simulation. This occurs because the QM−MM interaction changes at the cutoff distance from the rigorous integral interactions between the QM charge density and MM point charges to the point charge−point charge interactions (i.e., changes from the multipole−monopole interactions to the monopole−monopole interactions). This discontinuity can be avoided by introducing a switching function to transform between the two interaction forms smoothly at the real space cutoff. Eq 9 is then modified as

$$E_{QM-MM}^{elec}[\rho; q] = \sum_{g}\sum_{G}\sum_{i\in g}\sum_{j\in G}\left[S_w(r_{gG})\int \frac{q_i\rho(\mathbf{R}'_j)}{|\mathbf{r}_i - \mathbf{R}'_j|} d^3\mathbf{R}'_j + (1 - S_w(r_{gG}))\frac{q_iQ_j}{|\mathbf{r}_i - \mathbf{R}_j|}\right] \tag{14}$$

where r_{gG} is the distance between the centers of mass of the QM and MM groups. The switching function S_w is defined as[63]

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

$$S_w(r) = \begin{cases} 1 & \text{if } r \le r_{on} \\ \dfrac{(r_{cut}^2 - r^2)^2(r_{cut}^2 + 2r^2 - 3r_{on}^2)}{(r_{cut}^2 - r_{on}^2)^3} & \text{if } r_{on} < r \le r_{cut} \\ 0 & \text{if } r > r_{cut} \end{cases} \quad (15)$$

where r_{on} and r_{cut} are the distances where the switching function is switched on and off, respectively. Because the long-range PB correction affects the total energy only from r_{cut}, eq 11 does not require any modification. For consistency, the switching function was included in the calculation of the gradients of eq 14. With the switching function, we obtained an energy conservation of less than 1 kcal/mol drift during the 100 ps microcanonical ensemble (NVE) MD simulation for the systems considered here. Notably, a similar switching function was introduced in the AMBER program[18,64] independently but at the level of the atom-based cutoff.

DXL-BOMD and FD SCF Accelerators. The standard SCF convergence accelerator in MNDO97 is the DIIS[53,54] method, which extrapolates the Fock matrix based on the previous iterations. Besides this, we have implemented the extended Lagrangian Born−Oppenheimer molecular dynamics with dissipation (DXL-BOMD)[56,57] and Fock matrix dynamics (FD) approaches.[58,59]

In the DXL-BOMD method, an auxiliary density matrix (D), which approximates the true electron density matrix (P), is used as an initial guess for each SCF calculation during MD. The accelerator works by extrapolating the auxiliary density for the next time step (D_{n+1}) based on the previous $K + 1$ MD steps, referred to as the extrapolation order, as

$$D_{n+1} = 2D_n - D_{n-1} + \kappa(P_n - D_n) + \alpha \sum_{k=0}^{K} c_k D_{n-k} \quad (16)$$

where P_n is the fully converged density matrix in the present MD step, α and κ are the adjustable parameters, and c_k is a set of coefficients. In the work by Niklasson et al.,[56] the parameters were determined by an optimization procedure and the coefficients by the overdetermined extrapolations. In the present implementation, those parameters and coefficients were adopted without additional optimization. Because the DXL-BOMD method uses the auxiliary density, which is restrained to the fully converged density, as a guess density for the new SCF iteration, the number of SCF iterations can be truncated to accelerate the entire SCF procedure without sacrificing accuracy. We refer to refs 56 and 57 for details of the DXL-BOMD method.

In the FD approach,[58,59] the Fock matrix elements are extrapolated based on the Fock matrix elements of the previous $N + 1$ MD steps, which determine the order of the Taylor expansion, i.e.,

$$F_{\mu\nu}(s) = \sum_{m=0}^{M} c_{\mu\nu,m} s^m \quad (17)$$

where s has the values $-N/2$, ..., $N/2$ (by symmetry arguments). In the original FD method,[58,59] the previous values of the $F_{\mu\nu}$ element are arranged as a vector array $\mathbf{b}$. The resultant linear equation is solved as

$$\mathbf{Ac} = \mathbf{b} \quad (18)$$

where $\mathbf{A}_{ki} = k^i$ with k ranges between $-N/2$ and $N/2$ and i between 0 to M. Because matrix $\mathbf{A}$ is independent of the Fock matrix elements, its inverse can be constructed at the beginning of the simulation. Then, when the first $N + 1$ values of the Fock matrix elements are obtained, the subsequent $F(s + 1)$ value is extrapolated using the information from the coefficients array.

In the present work, we have implemented two variants of the FD extrapolation method. The first method is a polynomial expansion as described previously.[58] In the second method, the second derivative of the Fock matrix is extrapolated similarly to the first method using the previous $N + 1$ steps and the Taylor expansion centered on the time step N. The extrapolated Fock matrix is obtained by the Verlet integration with the second derivative term as the acceleration

$$F(s + 1) = 2F(s) - F(s - 1) + dt^2 \frac{\partial^2 F(s)}{\partial t^2} \quad (19)$$

Subsequent to this extrapolation, the regular SCF iteration is carried out to the full convergence of the QM electronic energy.

CAPABILITIES OF THE DEVELOPED QM/MM MODULE IN CHARMM

In the present study, the entire SE QM/MM routines of the MNDO97 module required for the electronic energy/gradient evaluation were rewritten based on the CHARMM program (version c41a2) and earlier MNDO97 module implementation in CHARMM.[50] The main focus of the present implementation was to enhance the efficiency and parallel scalabilty of the SE QM/MM method in CHARMM to enable long QM/MM MD simulations. Therefore, we did not attempt in this work to support the entire functionality available in the original MNDO97 package,[49] and left out several important functions for future implementation. They include, for example, the open-shell unrestricted Hartree−Fock, NMR chemical shifts, conductor-like screening model (COSMO)[65,66] for implicit solvent treatment, and orthogonalization-corrected methods (OMx).[67−71] The presently supported functions include:

- The major NDDO-based semiempirical QM methods, including the MNDO,[72,73] AM1,[74] PM3,[75−77] MNDO/d,[51] and AM1/d methods,[33,41] and their energy and gradients to perform energy minimizations and MD simulations.
- The periodic and nonperiodic boundary conditions supported by CHARMM to perform the QM/MM-Ewald, QM/MM-PME, and cutoff-based QM/MM calculations.
- The hydrogen link-atom[2,78] and generalized hybrid orbital (GHO) QM/MM boundary methods[79] to allow truncation of the covalent bonds between the QM and MM regions.
- The DXL-BOMD and FD SCF accelerators, which can be used together with the serial, MPI, and hybrid MPI/OMP parallel modes.

Finally, we note that the functionalities described above can be combined with many features available in CHARMM, for example, the position, distance, and dihedral restraints, Umbrella sampling free energy simulations, and all energy minimization and MD algorithms.

COMPUTATIONAL DETAILS

Solvated Adenylate Kinase (AdK) and Insulin Receptor Kinase (IRK). To evaluate the performance of the implemented QM/MM methods, two protein systems, AdK and IRK, were prepared and tested for the various methods described in the Theory section.

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

For the AdK system, the *Aquifex aeolicus* X-ray structure (PDB ID: 4CF7[7]) was used. The enzyme was solvated with 14,605 TIP3P[80] water molecules and 43 Na^+ and 41 Cl^- neutralizing ions in an 89 Å rhombic dodecahedron (RHDO) box. The CHARMM force fields[81−83] and CMAP correction[84] were used to describe the protein, one Mg^{2+} ion, and two ADP molecules, in which the one at the ATP binding site was protonated at its terminal phosphate. The simulation box included 47,293 atoms in total. A cutoff of 9.5 Å was used for the nonbonded interactions with the 90 × 90 × 90 FFT grid for PME calculations. The system was first energy minimized for 100 steps using the ABNR method[63] with constraints applied to both the protein and ligands and then 200 steps with harmonic restraints to the same set of atoms. During this step, the force constant of the harmonic restraints was gradually reduced from 100 to 0 kcal/mol Å^2 in steps of 50. After energy minimizations, the system was heated from 0 to 298 K over 90 ps and equilibrated for 10 ps. In the test of the performance of the present QM/MM implementation, a total of 92 atoms were selected as the QM atoms, which included the two ADPs, one Mg^{2+} ion, and its four coordinated water molecules.

For IRK, the active conformation system was prepared based on the X-ray crystal structure of IRK (PDB ID: 1IR3[85]) with its activation loop phosphoryl groups removed. Details of the system preparation were presented elsewhere.[8] Briefly, 28,899 atoms were present in the simulation box in which 76 atoms were QM atoms. The QM region included a part of ATP, two Mg^{2+} ions, six water molecules, and the side chains of Asp148, Asn153, Asp166, and Tyr302. The GHO method[79] was used at the C4′ atom of ATP and at the Cα atoms of Asp148, Asn153, Asp166, and Tyr302. The electrostatic interactions were evaluated with the PME method with 9 Å real space cutoff and 80 × 80 × 80 FFT grid. The same cutoff distance was used for the van der Waals interactions.

In both systems, the MD simulations were carried out using the leapfrog Verlet with the 1 fs time step and SHAKE[86] on all hydrogens. The QM region was described by the AM1/d-PhoT SE QM model[33] for the MNDO97 method and by the AM1 model[74] for the SQUANTUM (SQM) . For the SCC-DFTB method, the parameters developed by Elstner and co-workers[38,87,88] were used.

S_N2 Reaction between Cl^- and CH_3Cl. The Cl^- and CH_3Cl molecules, which were placed in the center, were solvated with 2193 previously equilibrated waters in a 40 Å cubic box. The system was described by the AM1 model[74] for the QM region and the TIP3P[80] model for the waters. The QM/MM-PME algorithm was used for electrostatics with the FFT grid of 64 × 64 × 64 and cutoff distance of 10 Å. The difference of distances from the C atom of CH_3Cl to the two Cl atoms was used as the reaction coordinate to perform the Umbrella sampling free energy simulations[89]

$$X^R = r_{Cl-C} - r_{C-Cl} \quad (20)$$

The initial window was prepared at X^R = 3.0 Å and minimized for 500 steps with the ABNR method. Force constants were gradually changed from 10 kcal/mol Å^2 at 3.0 Å to 150 kcal/mol Å^2 at 0 Å. Because of the symmetry of the reaction, only one side of the reaction was simulated in the range between 3.0 and 0 Å. Each simulation window was equilibrated for 100 ps followed by 100 ps simulation for data collection. The time step was 1 fs, and SHAKE was applied to all hydrogen atoms. Potential of mean force (PMF) was determined using the weighted histogram analysis method[90,91] and symmetrized to cover the entire range of the reaction coordinate between −3.0 and 3.0 Å.

Hardware Specifications. Performance analyses for the systems studied in the present work were conducted on a Linux cluster. Each node contained two CPUs each equipped with an 8-core Intel Xeon E5-2660 "Sandy Bridge" processor at 2.2 GHz, resulting in a total of 16 CPU cores. In the case of the hybrid MPI/OMP implementation, an MPI pin domain scheme was used for thread distribution across the cores.

RESULTS AND DISCUSSION

Timing Analysis of the Serial MNDO97 Implementation in CHARMM. Four characteristic routines were identified to be the most time-consuming in the total energy evaluation. They are displayed in the column labeled as "Main flow" in Figure 1. The first routine is the PME routine,[10,21] which

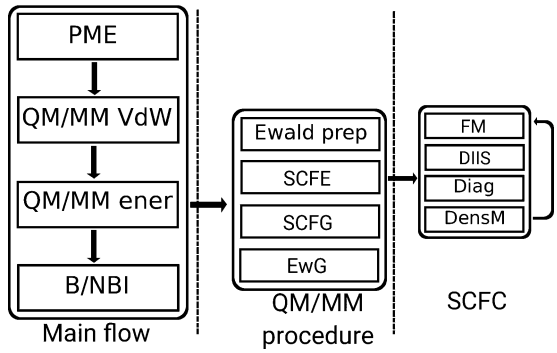


Figure 1. Flowchart of the present MNDO97/CHARMM implementation showing the most representative routines. The "Main flow" includes the PME, QM/MM van der Waals (VdW), SCF energy and gradients (QM/MM ener), and bonded/nonbonded interaction evaluation routines (B/NBI). The QM/MM energy and gradient routines ("QM/MM procedure") include all routines involved in the QM/MM nonbonded interaction, Ewald potential (Ewald prep), SCF energy (SCFE), and QM, QM-MM (SCFG), and Ewald gradient evaluations (EwG). The SCF iteration ("SCFC") is further decomposed into the Fock matrix construction (FM), DIIS acceleration (DIIS), Fock matrix diagonalization (Diag), and density matrix constrution (DensM), each of which is executed sequentially in each SCF iteration cycle.

computes the reciprocal space part of both the MM−MM and QM−MM electrostatic interactions. In the present implementation, we compute the PME interactions, especially at the beginning of the energy calculation to be used in the QM/MM SCF energy calculation. Second, the van der Waals (VdW) interactions between the QM and MM particles are computed with a switching function to smoothly turn off the interactions at the cutoff distance. Third, the SCF QM/MM energy and gradients are computed. Last, the bonded and nonbonded (electrostatic and van der Waals) interactions beween all MM pairs are computed up to a predefined cutoff distance. For each tested protein system, the timings of the four routines were monitored for the serial version of the code (BOMD with 1 MPI process) during the 1000 steps of MD simulation. In Figure 2a, their relative excution times are presented for the AdK system in which the most time-consuming routine is the QM/MM energy/gradient routine (52%) followed by the nonbonded (NB) interactions (28%) and PME routines (7%). The row data used in Figure 2a are also provided in Table 1 for both the AdK and IRK systems.

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

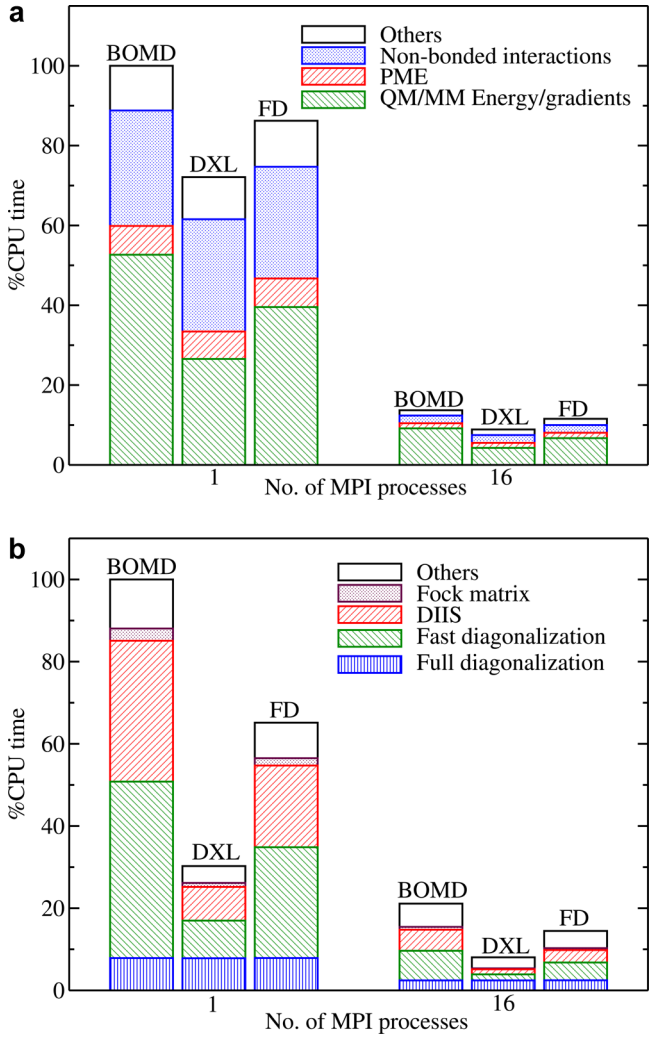


Figure 2. Timing histograms of (a) the main flow energy and (b) SCFC routines (Figure 1) of the implemented MNDO97 QM/MM methods for the AdK system. All timings were relative to the timings of the corresponding serial BOMD calculations.

Timing Analysis of the (Serial) SCF QM/MM Energy Routines. The QM/MM energy and gradients are determined through four sequential steps (see the QM/MM procedure in Figure 1): (1) the QM−QM Ewald sum and real space QM/MM correction evaluations (Ewald prep), (2) the SCF energy (SCFE) calculation, and the computations of the (3) SCF and (4) QM/MM Ewald gradients (SCFG and EwG, respectively). Timings of each of these steps are provided in Table 2. Among the four steps, the most time-consuming step is the SCFE step, which includes the computation of the core Hamiltonian (CH), electrostatic contribution to the core Hamiltonian (ECH), two-electron integrals (TEI) evaluation, and the CPU demanding SCF cycle (SCFC). The timings of each of these routines are provided in Table S1.

In SCFE, the SCF cycle (SCFC) routine is the most time-consuming. This happens because this routine was executed

Table 2. Execution Times (in Seconds) of Each of the QM/MM Energy and Gradient Routines (Figure 1) Including the QM/MM Nonbonded Interaction and Ewald Potential Evalutations (Ewald prep), SCF Energy (SCFE), QM/MM Real Space (SCFG), and PME Gradient Evaluations (EwG)[a]

method	MPI	Ewald prep	SCFE	SCFG	EwG
BOMD	1	13.71	878.63	215.85	155.71
		(10.10)	(343.06)	(154.25)	(96.15)
	16	2.66	185.61	14.05	16.77
		(2.06)	(94.32)	(9.81)	(10.79)
DXL-BOMD	1	13.89	265.87	217.91	155.09
		(10.38)	(120.46)	(156.13)	(103.09)
	16	2.63	70.73	14.08	17.04
		(2.09)	(41.37)	(9.89)	(10.93)
FD	1	13.76	572.45	216.81	156.71
		(10.36)	(208.36)	(156.79)	(102.73)
	16	2.75	127.18	13.99	18.25
		(2.08)	(60.86)	(9.89)	(10.73)

[a]All timings were based on the execution times of each routine accumulated over the 1000 MD steps for the AdK system and in parentheses for the IRK system.

Table 1. Timings of the Main Flow Energy Routines in CHARMM (Figure 1) Including the Particle Mesh Ewald (PME), QM/MM van der Waals (vdW) and SCF Energy/Gradient (Ener), and All MM Bonded (B) and Nonbonded Interaction (NBI) Routines[a]

			QM/MM		bond/nonbond		
method	MPI[b]	PME	vdW	Ener	NBI	B	total time[c]
BOMD	1	173.96	2.36	1269.25	697.03	3.37	2410.20
		(101.53)	(1.35)	(608.24)	(372.91)	(3.29)	(1221.00)
	16	30.67	0.99	220.92	46.83	1.04	330.00
		(19.78)	(0.85)	(116.73)	(25.79)	(0.46)	(183.60)
DXL-BOMD	1	165.69	2.36	639.99	677.92	3.37	1738.20
		(100.56)	(1.30)	(372.91)	(376.61)	(3.32)	(1006.20)
	16	30.69	1.01	102.56	47.07	1.13	214.20
		(19.74)	(0.84)	(64.73)	(25.8)	(0.45)	(130.80)
FD	1	172.74	2.37	953.17	674.49	3.39	2078.40
		(101.82)	(1.35)	(458.69)	(375.62)	(3.32)	(1090.20)
	16	32.67	1.73	161.60	46.77	0.75	278.40
		(19.71)	(0.86)	(84.74)	(25.81)	(0.49)	(151.20)

[a]All timings were based on the total elapsed times (in seconds) during the 1000 MD steps for the AdK system and in parentheses for the IRK system. [b]The total number of CPU cores used in the calculation: 1 for the serial and 16 for the 16 MPI processes. [c]The total exection time included the time for the nonbonded list update.

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

Table 3. Execution Times (in Seconds) of Different SCF Iteration Cycle Routines (SCFC in Table S1) Compared Between the Serial and Parallel and Different SCF Accelerator Implementations[a]

method	MPI	FM	DIIS	FastD	FullD	DensM
BOMD	1	26.12	301.24	377.31	69.06	80.87
		(11.73)	(113.73)	(140.88)	(30.34)	(18.02)
	16	6.12	45.03	63.31	21.35	29.54
		(6.35)	(22.68)	(22.87)	(12.64)	(17.65)
DXL-BOMD	1	8.52	72.02	80.47	68.77	23.36
		(3.79)	(29.27)	(33.46)	(27.71)	(5.59)
	16	2.04	10.90	12.94	21.32	8.01
		(2.15)	(5.92)	(5.68)	(12.35)	(5.36)
FD	1	15.68	174.69	236.94	69.20	52.61
		(5.38)	(54.33)	(75.51)	(30.10)	(10.32)
	16	4.08	26.59	38.10	21.57	19.66
		(3.78)	(11.24)	(12.55)	(12.47)	(10.76)

[a]Displayed data are for the computation of Fock matrix (FM), DIIS accelerator, fast (FastD) and full diagonalizations (FullD), and density matrix (DensM) for the AdK system. The timings for the IRK system are presented in parentheses.

~20 times to achieve full convergence of the electronic energy (10^{-6} eV threshold) at each MD step in both of the tested protein systems. Each SCFC iteration is comprised of four major substeps (see Figure 1): the Fock matrix construction (FM), DIIS extrapolation (DIIS), matrix diagonalization using both the fast[55] (FastD) and full[92] (FullD) diagonalization techniques (both labeled as Diag in Figure 1), and density matrix (DensM) construction, each of which is executed sequentially in each SCF iteration. Timings for these substeps are presented in Table 3, and a histogram of their timings is presented in Figure 2b, labeled as BOMD with the 1 MPI process. The most time-consuming step in SCFC is the matrix diagonalization using both the fast and full diagonalization techniques, which together account for 51% of the total execution time of the entire SCFC step. In the present MNDO97 implementation, the EISPACK routines based on the Housholder transformations were used in the full diagonalization,[92] and the so-called pseudodiagonalization in the FastD substep, respectively. The pseudodiagonalization technique works by a series of pseudoelimination of the occupied virtual pair Fock matrix elements,[55] which reduced the total execution time of the diagonalization step substantially compared to that of the FullD technique. Nevertheless, because FastD was executed ~18 times, whereas FullD was only ~2 times on average during each SCFC execution, FastD contributed more to the total execution time than FullD (Table 3 and Figure 2b). In Table S2, we also present the timings of each routine involved in the fullD step in which the backtransformation of the eigenvectors in the tridiagonal matrix form to the original symmetric matrix form is the slowest step.

Timing Analysis of the Parallel (MPI) MNDO97 QM/MM Implementation in CHARMM. The MPI parallelization was introduced to the serial MNDO97 QM/MM module by following the CHARMM MPI standard. During the implementation, we also optimized each routine in the SCF energy and gradient evaluations for improved performance of both the serial and parallel versions of the QM/MM methods based on the timings obtained in Tables 1–3 and Table S2. In particular, substantial efforts have been made to improve both the serial and parallel efficiencies of the two diagonalization routines (i.e., FastD and FullD). In some cases, several data array structures were modified to improve parallel scalability while retaining serial code efficiency. For example, routines in the full matrix diagonalization were parallelized only in the most expensive nested loops, such as the expensive legacy routines of the matrix tridiagonalization and backtransformation, to avoid high MPI communication overhead. This resulted in acceleration of the entire matrix diagonalization routines (Table 3).

As in the serial version, the execution times of the parallel version were analyzed. The results are presented in Figure 2, Tables 1–3, and Table S1 with the labels BOMD and 16 MPI processes. The results show that the total execution time of the QM/MM routines decreased subtantially. For example, the times spent on the QM/MM energy and gradient routines decreased by ~6-fold with 16 MPI processes compared to the timing of the single MPI BOMD calculation (Figure 2a). When the number of processors was increased to 64, the overall acceleration was ~8 fold. This acceleration was possible because of the reduction of the execution time of the diagonalization routines by a factor of 5 between 1 and 16 MPI processes. In contrast, because of the lack of proper parallelization in the SCF iteration routines, the SCC-DFTB and SQM methods showed no more than a factor of 2 acceleration in both the AdK and IRK systems.

In addition, the total time spent on each MPI communication routine was analyzed by applying the SCALASCA software.[93] As expected, the communication time increases with increasing number of MPI processes. It reached, for example, ~40% of the total execution time with 64 MPI processes. Results of the analysis are presented in Figure S1. This analysis indicates the ability to further improve our implementation.

Performance Enhancement by the DXL-BOMD and FD Methods. For the DXL-BOMD method, the single processor execution was ~17% faster than the BOMD method for the AdK system (Table S3). In the parallel version, we observed ~35% acceleration with 16 MPI processes. The performance enhancement was similar in the IRK system, i.e., ~21 and ~27% for the single and 16 processes, respectively (Table S4). For the FD approach, the single processor execution was ~13% faster than the BOMD counterpart for the AdK sytem. In the parallel version, the acceleration was ~19% with 16 processes (Table S3). In the IRK system, the acceleration was ~15 and ~12% for the 1 and 16 MPI processes, respectively (Table S4).

The detailed analyses of the DXL-BOMD and FD method timings are presented in Figure 2, and Tables 1–3, and Table S1 for both the AdK and IRK systems. These data were obtained with the ninth order expansion and five SCF steps for

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525–3536

the DXL-BOMD method and the fifth order expansion for the FD method (eq 19). In both tests, the order expansion and the number of SCF steps were chosen with the criterion that the total energy of the entire system drifted less than 1 kcal/mol during the 100 ps NVE MD simulation. The total energy conservation is compared in Figure 3a between the different SCF accelerators and also between the different parallelization models.

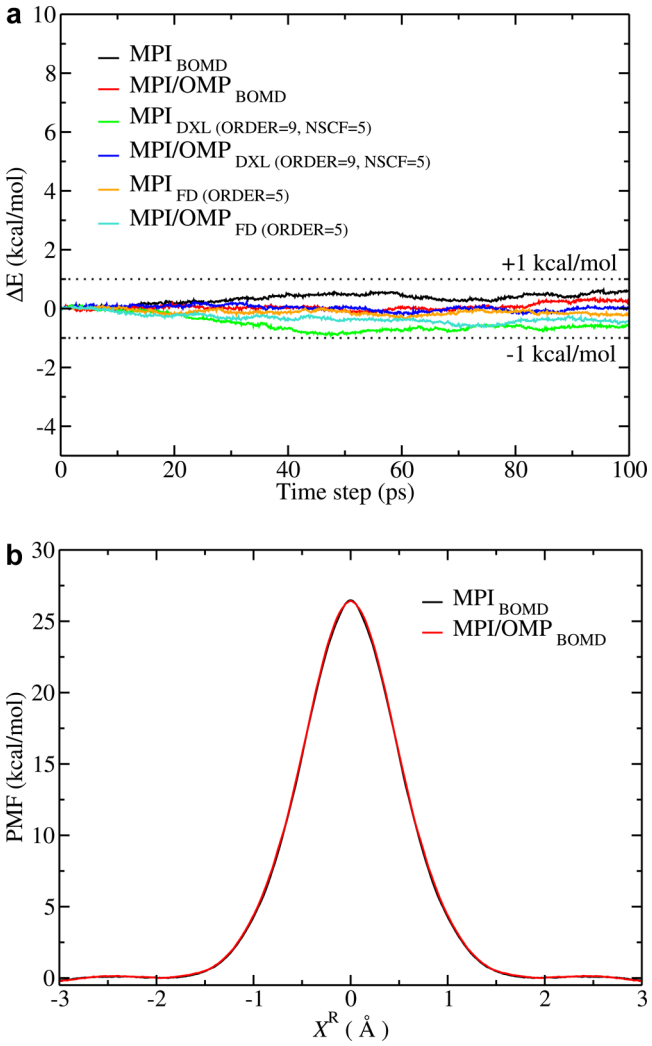


Figure 3. (a) Energy conservation of different parallel and SCF accelerator implementations for the AdK system. All implemented MNDO97 QM/MM methods showed less than 1 kcal/mol deviation of the total energy during 100 ps NVE MD simulations. (b) PMF profiles of the S_N2 reaction between CH_3Cl and Cl^- in water for different parallel implementations. In the test, the AM1 model[74] was used as the SE QM method.

Between the two SCF accelerators, the DXL-BOMD method was slightly faster than the FD method. This happens because the DXL-BOMD method does not require the full SCF convergence, and the total number of SCF steps are thus smaller with this method. In both methods, nevertheless, the performance reached a plateau as the number of CPU increased beyond 32 due to the increased communication overhead (Figure S1). On the other hand, as demonstrated by the presently tested systems, the performance gain by the DXL-BOMD and FD methods was expected to be higher in the parallel version with the increase in the size of the QM region.

Further Acceleration by the Hybrid MPI/OMP Method. To examine if the hybrid MPI/OMP approach can further accelerate the SE QM/MM calculations, it was implemented to the presently developed MNDO97 SE QM/MM methods. In the implementation, however, because implementing the hybrid approach to the entire CHARMM program was demanding given the size and complexity of the program, we applied the hybrid approach only to the SE QM/MM routines of the MNDO97 module and parallelized the MM and other energy evaluation routines with the "omp sections" work-sharing construct of OpenMP. The "omp sections" construct allowed groups of independent tasks to be simultaneously executed by OpenMP threads. With this approach, a reasonable performance gain was achieved for the tested systems with larger than 16 processes (Tables S3 and S4), which suggests that the fully functional hybrid MPI/OMP implementation would further improve the scalability of the SE QM/MM methods.

For the presented performance gain to be achieved, the overall execution times of different routines were analyzed first between 1 and 16 MPI processes (Tables 1−3). On the basis of the analysis, we have determined the routines to be executed in different OpenMP threads (via the "omp sections" construct). For instance, the PME and bonded/nonbonded routines can be executed in parallel without interfering with each other. Similarly, the SCFG and EwG routines can be run in different OpenMP threads once the SCF convergence is achieved. However, this task-level parallelization is not optimal. It is thus difficult to achieve a good load balance, especially when a large number of OpenMP threads are used. Because of this limitation, we allowed only two OpenMP threads per MPI process in the present implementation. In addition, two workflow schemes (Figure S2) were used to balance their loads with an increasing number of MPIs in which the switch between them is controlled by a user-defined variable. In each scheme, we defined two parallel regions (i.e., the "First" and "Second" parallel regions) each before and after the SCFE call and controlled by the "omp sections" constructs.

Two strategies were adopted in the hybrid MPI/OMP parallelization of the SCF energy (SCFE) routines. First, the SCF routines involved in the DIIS extrapolation and the full and fast matrix diagonalizations (Table 3 and Table S2) were parallelized with the loop-level parallelization. For example, either the inner or outer loop of each of those routines, depending on their overall performance, was split to be executed in different OpenMP threads. Then, the remainder of the SCFE calculation routines (Table S1) were parallelized with the task-level parallelization strategy to execute them as blocks of independent tasks similarly to the way presented in Figure S2. In Tables S3 and S4 and Figure 4, the performance of the hybrid MPI/OMP model was compared to the MPI-only model. In both systems (i.e., the AdK and IRK systems), the parallel performance was substantially better for the hybrid approach with larger than 16 CPUs compared with the MPI-only results. The performance gain was larger with the DXL-BOMD and FD methods in particular because they require a smaller number of MPI communications than that of the regular BOMD method. They resulted in as large as 14−15-fold acceleration relative to the serial BOMD results (Tables S3 and S4). However, for smaller than 8 CPUs, the hybrid model performed poorer because the current set of independent tasks

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

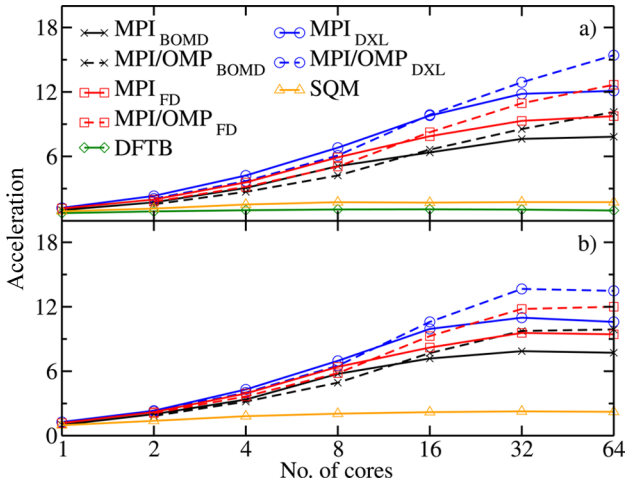


Figure 4. Comparison of the parallel scaling of the presently implemented MNDO97 QM/MM methods between the different parallelization models and SCF convergence acclerators for the (a) Adk and (b) IRK systems. For comparison, the scaling behaviors of the SCC-DFTB (DFTB) and SQUANTUM (SQM) methods are also presented.

causes unbalanced loads in the threads executing sections constructs.

Comparison of Serial and Parallel Timings of Different SE QM/MM Methods in CHARMM. The SE QM methods that were considered here for comparison included the presently developed MNDO97 MPI version, SCC-DFTB[94,95] (DFTB), and SQUANTUM (SQM) methods, which are all available in the CHARMM program. For the MNDO97 version, two parallelization options, pure MPI and hybrid MPI/OMP, were available and can be used alone or combined with the DXL-BOMD or FD convergence accelerator. The performance of the present MNDO97 implementations for the AdK and IRK systems is presented in Figure 4, and the raw data is presented in Tables S3 and S4, respectively. The results were based on the total elapsed time during the 10 ps QM/MM MD simulations (i.e., 10,000 MD steps) for each system. The resultant times were normalized against the times of the 1 MPI BOMD results and presented in Figure 4. In the case of DFTB, we used 10^{-9} eV as the SCF convergence criterion to achieve total energy conservation with less than 1 kcal/mol drift during the tested 10 ps MD simulation.

Overall, the MNDO97 MPI version showed a reasonable scaling up to 32 processes in both the AdK and IRK systems. In contrast, the SQM and DFTB methods showed very poor scaling behavior and no enhancement beyond 4 CPUs. This poor scaling is due to the lack of proper parallelization in the SQM and DFTB methods. For the hybrid MPI/OMP (MNDO97) model, we observed an additional acceleration compared to the pure MPI version with larger than 16 CPUs. For 16 CPUs, for example, a total of 7−8-fold acceleration was achieved with the MPI/OMP version relative to the single core (1 MPI) BOMD results. This value was increased by an additional factor of 2 when the number of CPUs was increased to 32 or 64. The largest acceleration was achieved when the hybrid MPI/OMP models were combined with the DXL-BOMD method. In this case, the method was scaled up to 64 CPUs (32 MPI processes and 2 threads per each MPI) with overall 15-fold acceleration for the AdK system and 14-fold for the IRK system. The results presented here for different sizes of the QM region suggest that the scaling behavior will improve further as the number of QM atoms increases because the overall contribution of the MPI communication to the total execution time would decrease with the increased Fock matrix size.

Finally, the accuracy of the present MPI and hybrid MPI/OMP models was examined by comparing the potentials of mean force (PMFs) for the S_N2 reaction between Cl^- and CH_3Cl in which the central methyl group is exchanged between the two Cl atoms. The results for both parallelization models are presented in Figure 3b. PMFs of both models agreed with each other with less than 0.1 kcal/mol difference along the entire reaction coordinate. The results are also in good agreement with the previous results for this chemical reaction.[52]

Potential Impact of the Present SE QM/MM Implementation. The present implementation of SE QM/MM methods offers the possibility to perform fast QM/MM MD simulations (in ns time scale) within reasonable computer time. For example, with the present implementation, we were able to run more than 1 ns MD simulation per day for the IRK system (Figure 5). As a comparison, pure MM simulations were only

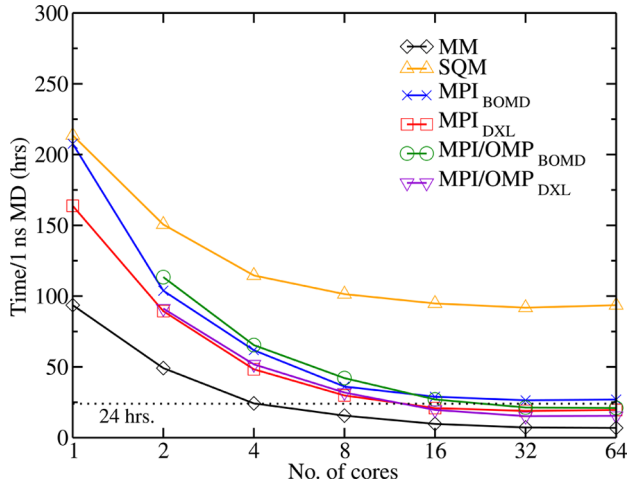


Figure 5. Projected time for a 1 ns MD simulation (in h) for the IRK system, as a function of the number of CPU cores, for the MM-only, SQM, MPI-only, MPI-DXL, and MPI/OMP models.

~2-times faster with the CHARMM program using the default PME and group-based cutoff schemes, whereas the other SE QM/MM methods (in CHARMM) took much longer. This is a substantial improvement over other SE QM/MM methods currently available not only in CHARMM but also in other MD packages, such as AMBER. The ability to perform long MD simulations has been an important issue in computational enzymology to obtain well-converged free energy values, especially if we want to accurately capture the impact of slow protein motions in catalysis.[8,96−102] To this end, the present implementation allows such a study in combination with other advanced simulation methods, such as the string[103−105] and metadynamics[106] methods.

CONCLUSIONS

In the present work, we have implemented efficient semi-empirical (SE) QM/MM methods in CHARMM based on the MNDO97 suite of programs and parallelized with the CHARMM MPI machinery. Scaling behavior of the present

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

implementation surpassed those of the previously implemented QM/MM methods in CHARMM, including the SCC-DFTB and SQUANTUM methods. The MPI implementation scaled well up to 32 processors for medium-sized systems, and beyond that number, the overheads associated with the MPI communications caused a decrease in performance. The performance and scaling behavior of the developed methods were further improved by implementing two different SCF convergence accelerators, DXL-BOMD and FD methods, and the hybrid MPI/OMP parallelization model. These methods showed excellent energy conservation (<1 kcal/mol) and parallel scaling that are better than those of the pure MPI model. At present, the best approach in our implementation is to combine the DXL-BOMD method with the hybrid MPI/OMP model. With this combination, the SE QM/MM simulation can be accelerated more than 14-fold relative to the single core BOMD simulation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00322.

Execution times of the SCFE and full matrix diagonalization routines, total MD timings of the AdK and IRK systems, and the OpenMP workload balancing schemes (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: pedro.ojeda-may@umu.se.

*E-mail: kwangho.nam@uta.edu. Phone: +1-817-272-1091.

ORCID

Pedro Ojeda-May: 0000-0001-9179-9441

Kwangho Nam: 0000-0003-0723-7839

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Authors acknowledge Dr. Guishan Zhang for providing the code and helpful discussion of the DXL-BOMD method. The present work was financially supported by grants from Kempestifelserna, Umeå University, Swedish Research Council (VR 2015-04114) and the University of Texas at Arlington. Computational resources were provided by the High Performance Computing Center North (HPC2N) and National Supercomputing Center (NSC) through the Swedish National Infrastructure for Computing (SNIC, Sweden) and the National Energy Research Scientific Computing Center (NERSC, USA).

■ 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) 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.

(3) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; John Wiley & Sons, Inc., 1996; pp 119−185.

(4) Cui, Q. Perspective: Quantum mechanical methods in biochemistry and biophysics. *J. Chem. Phys.* **2016**, *145*, 140901.

(5) Zhou, Y.; Ojeda-May, P.; Nagaraju, M.; Pu, J. In *Computational Approaches for Studying Enzyme Mechanism Part A*; Voth, G. A., Ed.; Methods in Enzymology; Academic Press, 2016; Vol. *577*, pp 185−212.

(6) Hynninen, A.-P.; Crowley, M. F. New faster CHARMM molecular dynamics engine. *J. Comput. Chem.* **2014**, *35*, 406−413.

(7) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; Alber, T.; Hagan, M. F.; Kern, D. The energy landscape of adenylate kinase during catalysis. *Nat. Struct. Mol. Biol.* **2015**, *22*, 124−131.

(8) Ojeda-May, P.; Li, Y.; Ovchinnikov, V.; Nam, K. Role of Protein Dynamics in Allosteric Control of the Catalytic Phosphoryl Transfer of Insulin Receptor Kinase. *J. Am. Chem. Soc.* **2015**, *137*, 12454−12457.

(9) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253−287.

(10) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An *Nlog*(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(11) Bowers, K. J.; Dror, R. O.; Shaw, D. E. Zonal methods for the parallel execution of range-limited N-body simulations. *J. Comput. Phys.* **2007**, *221*, 303−329.

(12) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. Classical Electrostatics for Biomolecular Simulations. *Chem. Rev.* **2014**, *114*, 779−814.

(13) M. P. Forum, *MPI: A Message-Passing Interface Standard*; University of Tennessee: Knoxville, TN, USA, 1994.

(14) Dongarra, J. Special issue on MPI: a message passage interface standard. *Int. J. Supercomput. Appl. High Perform. Comput.* **1994**, *8*, 167−168.

(15) Dagum, L.; Menon, R. OpenMP: an industry standard API for shared-memory programming. *IEEE Comput. Sci. Eng.* **1998**, *5*, 46−55.

(16) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(17) Plimpton, S. fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(18) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. *AMBER 2015*, University of California: San Francisco; **2015**.

(19) Brooks, B. R., III; Brooks, C. L.; Mackerell, A. D.; 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.; Hodošček, 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. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(20) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(21) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(22) Chang, C.-M.; Shao, Y.; Kong, J. Ewald mesh method for quantum mechanical calculations. *J. Chem. Phys.* **2012**, *136*, 114112.

(23) Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **2000**, *113*, 10913−10927.

(24) Simmonett, A. C.; Pickard, F. C., IV; Schaefer, H. F., III; Brooks, B. R. An efficient algorithm for multipole energies and derivatives

based on spherical harmonics and extensions to particle mesh Ewald. *J. Chem. Phys.* **2014**, *140*, 184101.
(25) Giese, T. J.; Panteva, M. T.; Chen, H.; York, D. M. Multipolar Ewald Methods, 1: Theory, Accuracy, and Performance. *J. Chem. Theory Comput.* **2015**, *11*, 436−450.
(26) Hehre, W. J., Radom, L., Schleyer, P. v. R., Pople, J. A., Eds. *Ab initio molecular orbital theory*; John Wiley: New York, 1986.
(27) Parr, R. G., Yang, W., Eds. *Density-functional theory of atoms and molecules*; Oxford University Press: New York, 1994.
(28) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Transition state stabilization and substrate strain in enzyme catalysis: ab initio QM/MM modelling of the chorismate mutase reaction. *Org. Biomol. Chem.* **2004**, *2*, 968−980.
(29) González-Lafont, A.; Truong, T. N.; Truhlar, D. G. Direct dynamics calculations with NDDO (neglect of diatomic differential overlap) molecular orbital theory with specific reaction parameters. *J. Phys. Chem.* **1991**, *95*, 4618−4627.
(30) Woodcock, H. L.; Hodošček, M.; Brooks, B. R. Exploring SCC-DFTB Paths for Mapping QM/MM Reaction Mechanisms. *J. Phys. Chem. A* **2007**, *111*, 5720−5728.
(31) Zhou, Y.; Pu, J. Reaction Path Force Matching: A New Strategy of Fitting Specific Reaction Parameters for Semiempirical Methods in Combined QM/MM Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 3038−3054.
(32) Rossi, I.; Truhlar, D. G. Parameterization of NDDO wavefunctions using genetic algorithms. An evolutionary approach to parameterizing potential energy surfaces and direct dynamics calculations for organic reactions. *Chem. Phys. Lett.* **1995**, *233*, 231−236.
(33) Nam, K.; Cui, Q.; Gao, J.; York, D. M. Specific Reaction Parametrization of the AM1/d Hamiltonian for Phosphoryl Transfer Reactions: H, O, and P Atoms. *J. Chem. Theory Comput.* **2007**, *3*, 486−504.
(34) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. Description of Phosphate Hydrolysis Reactions with the Self-Consistent-Charge Density-Functional-Tight-Binding (SCC-DFTB) Theory. 1. Parameterization. *J. Chem. Theory Comput.* **2008**, *4*, 2067−2084.
(35) Liang, S.; Roitberg, A. E. AM1 Specific Reaction Parameters for Reactions of Hydroxide Ion with Halomethanes in Complex Environments: Development and Testing. *J. Chem. Theory Comput.* **2013**, *9*, 4470−4480.
(36) Kuechler, E. R.; York, D. M. Quantum mechanical study of solvent effects in a prototype S_N2 reaction in solution: Cl^- attack on CH_3Cl. *J. Chem. Phys.* **2014**, *140*, 054109.
(37) Govender, K.; Gao, J.; Naidoo, K. J. AM1/d-CB1: A Semiempirical Model for QM/MM Simulations of Chemical Glycobiology Systems. *J. Chem. Theory Comput.* **2014**, *10*, 4694−4707.
(38) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, *10*, 1518−1537.
(39) Menegon Arantes, G.; Loos, M. Specific parametrisation of a hybrid potential to simulate reactions in phosphatases. *Phys. Chem. Chem. Phys.* **2006**, *8*, 347−353.
(40) Tejero, I.; González-Lafont, A.; Lluch, J. M. A PM3/d specific reaction parameterization for iron atom in the hydrogen abstraction catalyzed by soybean lipoxygenase-1. *J. Comput. Chem.* **2007**, *28*, 997−1005.
(41) Lopez, X.; York, D. M. Parameterization of semiempirical methods to treat nucleophilic attacks to biological phosphates: AM1/d parameters for phosphorus. *Theor. Chem. Acc.* **2003**, *109*, 149−159.
(42) Rendell, A. P.; Bliznyuk, A.; Huber, T.; Nobes, R. H.; Akhmatskaya, E. V.; Früchtl, H. A.; Kung, P. W.-C.; Milman, V.; Lung, H. Computational chemistry on Fujitsu vector-parallel processors: Development and performance of applications software. *Parallel Computing* **2000**, *26*, 887−911.
(43) Maia, J. D. C.; Urquiza Carvalho, G. A.; Mangueira, C. P.; Santana, S. R.; Cabral, L. A. F.; Rocha, G. B. GPU Linear Algebra Libraries and GPGPU Programming for Accelerating MOPAC Semiempirical Quantum Chemistry Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3072−3081.
(44) Wu, X.; Koslowski, A.; Thiel, W. Semiempirical Quantum Chemical Calculations Accelerated on a Hybrid Multicore CPU-GPU Computing Platform. *J. Chem. Theory Comput.* **2012**, *8*, 2272−2281.
(45) Car, R.; Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(46) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. *J. Chem. Phys.* **2001**, *114*, 9758.
(47) Herbert, J. M.; Head-Gordon, M. Curvy-steps approach to constraint-free extended-Lagrangian ab initio molecular dynamics, using atom-centered basis functions: Convergence toward Born-Oppenheimer trajectories. *J. Chem. Phys.* **2004**, *121*, 11542.
(48) Nam, K. Acceleration of Semiempirical Quantum Mechanical Calculations by Extended Lagrangian Molecular Dynamics Approach. *J. Chem. Theory Comput.* **2013**, *9*, 3393−3403.
(49) Thiel, W. *MNDO97*, version 5.0; University of Zurich: Zurich, Switzerland, 1998.
(50) Nam, K.; Gao, J.; York, D. M. An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations. *J. Chem. Theory Comput.* **2005**, *1*, 2−13.
(51) Thiel, W.; Voityuk, A. A. Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group. *J. Phys. Chem.* **1996**, *100*, 616−626.
(52) Nam, K. Acceleration of Ab Initio QM/MM Calculations under Periodic Boundary Conditions by Multiscale and Multiple Time Step Approaches. *J. Chem. Theory Comput.* **2014**, *10*, 4175−4183.
(53) Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chem. Phys. Lett.* **1980**, *73*, 393−398.
(54) Pulay, P. Improved SCF convergence acceleration. *J. Comput. Chem.* **1982**, *3*, 556−560.
(55) Stewart, J. J. P.; Császár, P.; Pulay, P. Fast semiempirical calculations. *J. Comput. Chem.* **1982**, *3*, 227−228.
(56) Niklasson, A. M. N.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C. J.; Holmström, E.; Zheng, G.; Weber, V. Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009**, *130*, 214109.
(57) Zheng, G.; Niklasson, A. M. N.; Karplus, M. Lagrangian formulation with dissipation of Born-Oppenheimer molecular dynamics using the density-functional tight-binding method. *J. Chem. Phys.* **2011**, *135*, 044122.
(58) Pulay, P.; Fogarasi, G. Fock matrix dynamics. *Chem. Phys. Lett.* **2004**, *386*, 272−278.
(59) Herbert, J. M.; Head-Gordon, M. Accelerated, energy-conserving Born-Oppenheimer molecular dynamics via Fock matrix extrapolation. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3269−3275.
(60) Roothaan, C. C. J. New developments in molecular orbital theory. *Rev. Mod. Phys.* **1951**, *23*, 69−89.
(61) Dral, P. O.; Wu, X.; Spoerkel, L.; Koslowski, A.; Weber, W.; Steiger, R.; Scholten, M.; Thiel, W. Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Theory, Implementation, and Parameters. *J. Chem. Theory Comput.* **2016**, *12*, 1082−1096.
(62) Mulliken, R. S. Electronic Population Analysis on LCAO[Single Bond]MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833.
(63) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(64) Walker, R. C.; Crowley, M. F.; Case, D. A. The implementation of a fast and accurate QM/MM potential method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019−1031.
(65) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536

(66) Khandogin, J.; Gregersen, B. A.; Thiel, W.; York, D. M. Smooth Solvation Method for d-Orbital Semiempirical Calculations of Biological Reactions. 1. Implementation. *J. Phys. Chem. B* **2005**, *109*, 9799−9809.
(67) Kolb, M. Ein neues semiempirisches Verfahren auf Grundlage der NDDO-Naeherung: Entwicklung der Methode, Parametrisierung und Anwendungen. Ph.D. Thesis, 1991.
(68) Kolb, M.; Thiel, W. Beyond the MNDO model: Methodical considerations and numerical results. *J. Comput. Chem.* **1993**, *14*, 775−789.
(69) Weber, W. Ein neues semiempirisches NDDO-Verfahren mit Orthogonalisierungskorrekturen: Entwicklung des Modells, Parametrisierung und Anwendungen. Ph.D. Thesis, 1996.
(70) Weber, W.; Thiel, W. Orthogonalization corrections for semiempirical methods. *Theor. Chem. Acc.* **2000**, *103*, 495−506.
(71) Scholten, M. Semiempirische Verfahren mit Orthogonalisierungskorrekturen: Die OM3 Methode. Ph.D. Thesis, 2003.
(72) Dewar, M. J. S.; Thiel, W. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.
(73) Dewar, M. J. S.; Thiel, W. Ground states of molecules. 39. MNDO results for molecules containing hydrogen, carbon, nitrogen, and oxygen. *J. Am. Chem. Soc.* **1977**, *99*, 4907−4917.
(74) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(75) Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.
(76) Stewart, J. J. P. Optimization of parameters for semiempirical methods II. Applications. *J. Comput. Chem.* **1989**, *10*, 221−264.
(77) Stewart, J. J. P. Optimization of parameters for semiempirical methods. III Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi. *J. Comput. Chem.* **1991**, *12*, 320−341.
(78) Singh, U. C.; Kollman, P. A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $CH_3Cl + Cl^-$ exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718−730.
(79) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.
(80) 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.
(81) MacKerell, A. D.; 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.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-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.
(82) Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(83) MacKerell, A. D.; Banavali, N.; Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **2000**, *56*, 257−265.
(84) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(85) Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO Journal* **1997**, *16*, 5572−5581.
(86) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(87) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338−354.
(88) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. *J. Phys. Chem. B* **2015**, *119*, 1062−1082.
(89) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(90) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(91) Rajamani, R.; Naidoo, K. J.; Gao, J. Implementation of an adaptive umbrella sampling method for the calculation of multidimensional potential of mean force of chemical reactions in solution. *J. Comput. Chem.* **2003**, *24*, 1775−1781.
(92) Garbow, B. S.; Boyle, J. M.; Dongarra, J. J.; Moler, C. B. *Matrix Eigensystem Routines - EISPACK Guide Extension*; Springer: Berlin, Heidelberg, 1977; pp 1−4.
(93) Geimer, M.; Wolf, F.; Wylie, B. J. N.; Ábrahám, E.; Becker, D.; Mohr, B. The Scalasca performance toolset architecture. *Concurrency Comput.: Pract. Exp.* **2010**, *22*, 702−719.
(94) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 7260−7268.
(95) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(96) Benkovic, S. J.; Hammes-Schiffer, S. A perspective on enzyme catalysis. *Science* **2003**, *301*, 1196−1202.
(97) Fan, Y.; Cembran, A.; Ma, S.; Gao, J. Connecting Protein Conformational Dynamics with Catalytic Function As Illustrated in Dihydrofolate Reductase. *Biochemistry* **2013**, *52*, 2036−2049.
(98) Schwartz, S. D.; Schramm, V. L. Enzymatic transition states and dynamic motion in barrier crossing. *Nat. Chem. Biol.* **2009**, *5*, 551−558.
(99) Kamerlin, S. C. L.; Warshel, A. At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1339−1375.
(100) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. Perspectives on Electrostatics and Conformational Motions in Enzyme Catalysis. *Acc. Chem. Res.* **2015**, *48*, 482−489.
(101) Kohen, A. Role of Dynamics in Enzyme Catalysis: Substantial versus Semantic Controversies. *Acc. Chem. Res.* **2015**, *48*, 466−473.
(102) Tuñón, I.; Laage, D.; Hynes, J. T. Are there dynamical effects in enzyme catalysis? Some thoughts concerning the enzymatic chemical step. *Arch. Biochem. Biophys.* **2015**, *582*, 42−55.
(103) Vanden-Eijnden, E.; Venturoli, M. Revisiting the finite temperature string method for the calculation of reaction tubes and free energies. *J. Chem. Phys.* **2009**, *130*, 194103.
(104) Ovchinnikov, V.; Karplus, M.; Vanden-Eijnden, E. Free energy of conformational transition paths in biomolecules: The string method and its application to myosin VI. *J. Chem. Phys.* **2011**, *134*, 085103.
(105) Ovchinnikov, V.; Nam, K.; Karplus, M. A simple and accurate method to calculate free energy profiles and reaction rates from restrained molecular simulations of diffusive processes. *J. Phys. Chem. B* **2016**, *120*, 8457−8472.
(106) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.

DOI: 10.1021/acs.jctc.7b00322
J. Chem. Theory Comput. 2017, 13, 3525−3536