THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Accelerating *Ab Initio* Quantum Mechanical and Molecular Mechanical (QM/MM) Molecular Dynamics Simulations with Multiple Time Step Integration and a Recalibrated Semiempirical QM/MM Hamiltonian

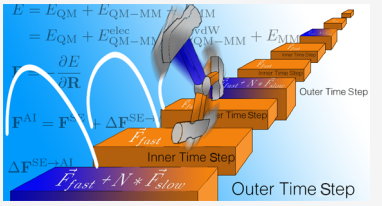
Xiaoliang Pan, Richard Van, Evgeny Epifanovsky, Jian Liu, Jingzhi Pu,* Kwangho Nam,* and Yihan Shao*

Cite This: https://doi.org/10.1021/acs.jpcb.2c02262

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Molecular dynamics (MD) simulations employing *ab initio* quantum mechanical and molecular mechanical (ai-QM/MM) potentials are considered to be the state of the art, but the high computational cost associated with the ai-QM calculations remains a theoretical challenge for their routine application. Here, we present a modified protocol of the multiple time step (MTS) method for accelerating ai-QM/MM MD simulations of condensed-phase reactions. Within a previous MTS protocol [Nam *J. Chem. Theory Comput.* **2014**, *10*, 4175], reference forces are evaluated using a low-level (semiempirical QM/MM) Hamiltonian and employed at inner time steps to propagate the nuclear motions. Correction forces, which arise from the force differences between high-level (ai-QM/MM) and low-level Hamiltonians, are applied at outer time steps, where the MTS algorithm allows the time-reversible integration of the correction forces. To increase the outer step size, which is bound by the highest-frequency component in the correction forces, the semiempirical QM Hamiltonian is recalibrated in this work to minimize the magnitude of the correction forces. The remaining high-frequency modes, which are mainly bond stretches involving hydrogen atoms, are then removed from the correction forces. When combined with a Langevin or SIN(R) thermostat, the modified MTS-QM/MM scheme remains robust with an up to 8 (with Langevin) or 10 fs (with SIN(R)) outer time step (with 1 fs inner time steps) for the chorismate mutase system. This leads to an over 5-fold speedup over standard ai-QM/MM simulations, without sacrificing the accuracy in the predicted free energy profile of the reaction.



INTRODUCTION

The combined quantum mechanical and molecular mechanical (QM/MM) method has been established as an essential tool in the computational modeling of solution-phase and enzymatic reactions.[1−11] In recent years, *ab initio* QM/MM (ai-QM/MM) calculations,[12−15] especially those based on a density functional theory (DFT)[16−20] description of the reactive region, have shed light on the mechanism of many enzymatic reactions. However, due to their high computational cost, the routine application of the ai-QM/MM potentials is often limited to minimum energy pathway (MEP) calculations, where only $O(10^2)$ QM/MM energy/force calculations are required to optimize each point along a discretized reaction pathway. Less common is the direct employment of ai-QM/MM models in the minimum free energy pathway (MFEP) simulations,[21−30] which typically involve $O(10^4)$ QM/MM energy/force calculations to sample each point/window along the free energy pathway and thus require hundreds of thousands (if not millions) of computer core-hours. Consequently, direct QM/MM MFEP simulations often resort to less accurate semiempirical QM/MM (se-QM/MM) Hamiltonians based on AM1,[31] PM3,[32] PM6,[33] or tight-binding DFT[34−37] methods. However, the accuracy of each se-QM model is not always guaranteed for a specific enzyme system of interest.

Going beyond direct QM/MM MFEP calculations, Gao[4] and Marí et al.[38] pioneered the dual-Hamiltonian strategy for simultaneously exploiting the accuracy of a high-level Hamiltonian and the efficiency of a low-level one. As illustrated in Figure 1, there are two general schemes to implement the dual-Hamiltonian strategy. In the reference potential method (i.e., indirect QM/MM simulations),[4,39−45] the configurational sampling is carried out using a reference low-level potential energy, and then, the low-level free energy results are corrected in a postprocessing process using thermodynamical perturbation.[46] This correction can also be achieved through non-equilibrium work simulations.[47−51]

Received: April 2, 2022
Revised: May 14, 2022

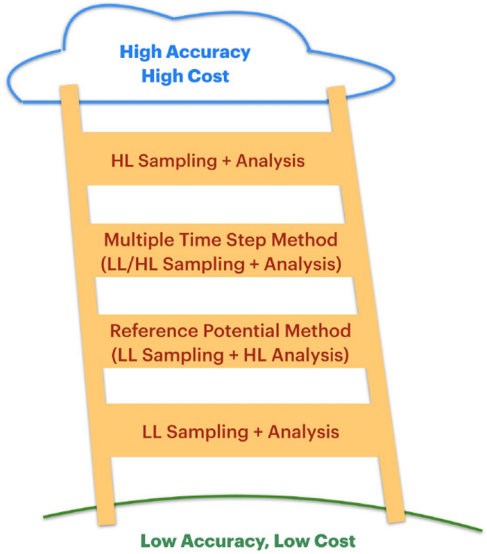


Figure 1. Hierarchy of free energy simulations using a lower-level (LL) Hamiltonian and/or a high-level (HL) Hamiltonian.

In the alternative multiple time step (MTS) approach,[52] the configurational sampling using the low-level QM/MM Hamiltonian is corrected on the fly via regular computation of high-level forces.[53] Dating back to the reversible reference system propagator algorithm (rRESPA) developed by Tuckerman and coworkers,[54] the MTS method assumes a separation of the total force of each configuration of a system into fast and slow components (see Figure 2). Within the MTS integration, the fast

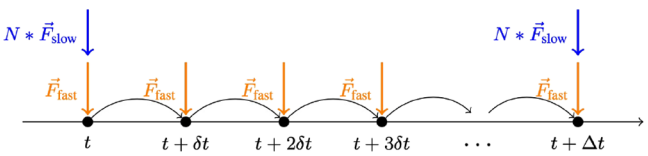


Figure 2. Illustration of MTS simulations with fast forces applied at each inner time step (with intervals δt) and slow forces at outer time steps (at intervals of Δt).

forces are applied at each step of a molecular dynamics (MD) simulation, which becomes inner time steps with intervals δt (typically ranging between 0.5 and 2.0 fs). The slow forces, on the other hand, are applied only at outer time steps with intervals $\Delta t = N \times \delta t$, i.e., N times longer. For the finite-temperature free energy simulations, the MTS algorithm can be combined with a Nosé−Hoover thermostat,[54] a Nosé−Hoover−Langevin thermostat with isokinetic constraints,[55−57] Langevin and Andersen thermostats.[57,58] Monte Carlo steps have also been introduced by Roux and coworkers to suppress the discretization error in MTS simulations.[58] For QM/MM simulations, MTS has been used by Nam to accelerate the modeling of solution-phase[53] and enzyme reaction pathways[59] and by Rothlisberger and coworkers to speed up the computation of solution-phase molecular vibrational frequencies.[60,61]

Following the MTS QM/MM enzyme simulations by Nam,[53] in this article, we focus on the application of the rRESPA method for combining se-QM/MM and ai-QM/MM Hamiltonians in MFEP simulations of condensed-phase reactions. In particular, the se-QM/MM forces are employed as the "fast" forces to propagate the trajectory at inner time steps, whereas the difference between ai-QM/MM and se-QM/MM forces is used as the "slow" correction forces at outer time steps. As we show in this work, in comparison to the "total" (ai-QM/MM) force, the difference forces (as shown in Figures S1 and S4 for the reactant and transition state regions of the chorismate mutase catalysis pathway) indeed oscillate at substantially reduced magnitudes. Indeed, such small force differences allowed Nam to apply Hartree−Fock/3-21G/MM correction forces at outer time steps of up to 3 fs to accurately simulate the S_N2 reaction between CH_3Cl and Cl^- in water, in which the AM1/MM or PM3/MM energy function was used as the lower-level Hamiltonian.[53]

In this work, we shall aim at a further increase in the outer time step of an MTS QM/MM free energy simulation of condensed-phase reactions and thus a more significant enhancement to the computational efficiency with these simulations. Clearly, the smaller the difference forces are (between se-QM/MM and ai-QM/MM models), the weaker the correction would be at the outer time steps, which can then make the MTS simulations more robust with larger outer time steps. Henceforth, we sought to reduce the difference forces via recalibrating the se-QM/MM Hamiltonian for the specific reaction of interest, following the idea of reaction path force matching parameterization of Zhou and Pu.[62] Such a recalibration of the se-QM/MM Hamiltonian and its application in MTS simulations are described in the Methods section, with our MTS simulation details provided in the Computational Details section. Computational results on the chorismate mutase system are presented and discussed in the Results and Discussion section. Finally, concluding remarks are made in the Conclusions section.

In comparison to the earlier MTS QM/MM simulations by Nam,[53] our implementation offers a couple of advantages. First, the recalibration of the se-QM model should allow us to use longer outer time steps, thus increasing the computational efficiency. Second, the use of our QM/MM-AC electrostatics model[63] enabled the use of the periodic boundary condition instead of a cluster model for force correction, thus ensuring a continuous potential energy surface. Finally, our MTS implementation utilizes the QMHub interface, which allows the use of QM methods in different QM software packages, such as those in the Q-Chem software package,[64] and a range of semiempirical methods, such as those implemented in the SQM module in AmberTools.[65,66]

METHODS

In QM/MM calculations, the system is typically divided into a QM subsystem and an MM subsystem. The potential energy of the whole system can be expressed as

$$\begin{aligned} E &= E_{QM} + E_{QM-MM} + E_{MM} \\ &= E_{QM} + E_{QM-MM}^{elec} + E_{QM-MM}^{vdW} + E_{MM} \end{aligned} \quad (1)$$

where the first two terms are evaluated together at a QM level of theory, while the last two terms are usually calculated at the MM level. In this work, we will focus on

$$\mathbf{F} = -\frac{\partial E}{\partial \mathbf{R}} \quad (2)$$

which is the corresponding force on both QM and MM atoms, with coordinates at $\mathbf{R}_{QM}$ and $\mathbf{R}_{MM}$. When the ai-QM method and se-QM are applied for the QM region, the force will be denoted as $\mathbf{F}^{AI}$ and $\mathbf{F}^{SE}$, respectively.

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Multiple Time Step (MTS) Integrator for ai-QM/MM MD Simulations. In our MTS QM/MM method, the ai-QM/MM forces for the entire system, $\mathbf{F}^{\mathrm{AI}}$, are decomposed into

$$\mathbf{F}^{\mathrm{AI}} = \mathbf{F}^{\mathrm{SE}} + \Delta\mathbf{F}^{\mathrm{SE}\to\mathrm{AI}} \tag{3}$$

where $\mathbf{F}^{\mathrm{SE}}$ refers to the se-QM/MM forces and the correction forces are defined as

$$\Delta\mathbf{F}^{\mathrm{SE}\to\mathrm{AI}} = \mathbf{F}^{\mathrm{AI}} - \mathbf{F}^{\mathrm{SE}} \tag{4}$$

each with $3N_{\mathrm{atom}}$ components [In contrast to the original MTS QM/MM method by Nam,[53] where the correction forces are calculated only for atoms within the real-space cutoff, in this work, both $\mathbf{F}^{\mathrm{AI}}$ and $\mathbf{F}^{\mathrm{SE}}$ are calculated using the QM/MM with augmentary charge (QM/MM-AC) electrostatic embedding scheme that we developed recently[63] to include the correction for the long-range QM/MM electrostatics.]. Accordingly, the Kolmogorov operator is

$$\mathcal{L}^{\mathrm{AI}} = \mathcal{L}_x + \mathcal{L}_p^{\mathrm{SE}} + \mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}} \tag{5}$$

with the individual operators being

$$\mathcal{L}_x = \sum_{i=1}^{3N_{\mathrm{atom}}} \dot{r}_i \frac{\partial}{\partial r_i} \tag{6a}$$

$$\mathcal{L}_p^{\mathrm{SE}} = \sum_{i=1}^{3N_{\mathrm{atom}}} F_i^{\mathrm{SE}} \frac{\partial}{\partial p_i} \tag{6b}$$

$$\mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}} = \sum_{i=1}^{3N_{\mathrm{atom}}} \Delta F_i^{\mathrm{SE}\to\mathrm{AI}} \frac{\partial}{\partial p_i} \tag{6c}$$

where r_i and p_i refer to the coordinates and momentum of each atom in the system.

A leapfrog step for molecular dynamics simulations

$$e^{\mathcal{L}\delta t} \approx e^{\mathcal{L}_x\delta t} e^{1/2\mathcal{L}_p\delta t} e^{1/2\mathcal{L}_p\delta t} = e^{\mathcal{L}_x\delta t} e^{\mathcal{L}_p\delta t} \tag{7}$$

involves one force evaluation, two half-time step momentum updates, i.e., $\mathbf{p}\left(t - \frac{1}{2}\delta t\right) \to \mathbf{p}(t)$ and $\mathbf{p}(t) \to \mathbf{p}\left(t + \frac{1}{2}\delta t\right)$, and one full-time step coordinate update, $\mathbf{r}(t) \to \mathbf{r}(t + \delta t)$. Note that the two momentum updates utilize the same force and can thus be combined. With the decomposition of the Liouville operator in eq 4, the propagation of an MTS trajectory (as illustrated in Figure 2) within the leapfrog algorithm follows

$$e^{\mathcal{L}^{\mathrm{AI}}\Delta t} \approx [e^{\mathcal{L}_x\delta t} e^{\mathcal{L}_p^{\mathrm{SE}}\delta t}]^{N-1} e^{\mathcal{L}_x\delta t} e^{(\mathcal{L}_p^{\mathrm{SE}} + N\mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}})\delta t} \tag{8}$$

This will generate a microcanonical (NVE) ensemble of our enzyme system.

MTS ai-QM/MM Canonical Ensemble Simulations with Langevin Dynamics. When a Langevin thermostat is adopted, one applies a Kolmogorov operator, $\mathcal{L}_T$, which corresponds to an additional (Ornstein−Uhlenbeck) update of the atomic momenta

$$\mathbf{p}(t) \leftarrow e^{-\gamma\delta t}\mathbf{p}(t) + \sqrt{\frac{1 - e^{-2\gamma\delta t}}{\beta}}\mathbf{M}^{1/2}\tilde{\boldsymbol{\eta}} \tag{9}$$

and depends on γ, the friction constant, $\tilde{\boldsymbol{\eta}}$, a random vector within a unit-variance Gaussian distribution, β, the inverse temperature $1/(k_{\mathrm{B}}T)$, and $\mathbf{M}$, a diagonal matrix with atomic masses.[67] With this thermostat added to our MTS leapfrog integration in eq 8, the evolution within each outer time step becomes

$$e^{\mathcal{L}^{\mathrm{AI}}\Delta t} \approx [e^{\mathcal{L}_T\delta t} e^{\mathcal{L}_x\delta t} e^{\mathcal{L}_p^{\mathrm{SE}}\delta t}]^{N-1} \times e^{\mathcal{L}_T\delta t} e^{\mathcal{L}_x\delta t} e^{(\mathcal{L}_p^{\mathrm{SE}} + N\mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}})\delta t} \tag{10}$$

within the "end" scheme for the thermostat[67] or

$$e^{\mathcal{L}^{\mathrm{AI}}\Delta t} \approx [e^{1/2\mathcal{L}_x\delta t} e^{\mathcal{L}_T\delta t} e^{1/2\mathcal{L}_x\delta t} e^{\mathcal{L}_p^{\mathrm{SE}}\delta t}]^{N-1} \times e^{1/2\mathcal{L}_x\delta t} e^{\mathcal{L}_T\delta t} e^{1/2\mathcal{L}_x\delta t} e^{(\mathcal{L}_p^{\mathrm{SE}} + N\mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}})\delta t} \tag{11}$$

within the "middle" scheme.[67]

MTS ai-QM/MM Simulations with SIN(R). The Stochastic-Iso-NH-RESPA [SIN(R)] algorithm employs a set of isokinetic constraints to control the resonance problem in MTS, leading to an increase in the outer time step.[55] When the "middle" thermostat scheme is applied to SIN(R), it yields more stable and accurate results, especially as the inner time step increases.[57] When the "middle" scheme SIN(R) with a leapfrog integrator is used, the propagator reads

$$e^{\mathcal{L}^{\mathrm{AI}}\Delta t} \approx [e^{1/2\mathcal{L}_x\delta t} e^{1/2\mathcal{L}_N\delta t} e^{\mathcal{L}_O\delta t} e^{1/2\mathcal{L}_N\delta t} e^{1/2\mathcal{L}_x\delta t} e^{\mathcal{L}_p^{\mathrm{SE}}\delta t}]^{N-1} \times e^{1/2\mathcal{L}_x\delta t} e^{1/2\mathcal{L}_N\delta t} e^{\mathcal{L}_O\delta t} e^{1/2\mathcal{L}_N\delta t} e^{1/2\mathcal{L}_x\delta t} e^{(\mathcal{L}_p^{\mathrm{SE}} + N\mathcal{L}_p^{\mathrm{SE}\to\mathrm{AI}})\delta t} \tag{12}$$

where the relevant Kolmogorov operators ($\mathcal{L}_x$, $\mathcal{L}_p$, $\mathcal{L}_N$, and $\mathcal{L}_O$) are defined in eqs 89−93 of ref 57. Specifically, the operator $\mathcal{L}_p$ updates the momentum, p_i, and the first set of thermodynamic variables, $v_{1,i}^{(j)}$ (j = 1,2, ..., L), for each nuclear degree of freedom (i = 1,2, ...,$3N_{\mathrm{atom}}$), i.e., its propagation is

$$p_i \leftarrow \frac{p_i + F_i s_i\left(\frac{\delta t}{2}\right)}{\dot{s}_i\left(\frac{\delta t}{2}\right)} \tag{13a}$$

$$v_{1,i}^{(j)} \leftarrow \frac{v_{1,i}^{(j)}}{\dot{s}_i\left(\frac{\delta t}{2}\right)} \tag{13b}$$

with

$$F_i s_i\left(\frac{\delta t}{2}\right) = \sqrt{m_i L k_{\mathrm{B}}T}\,\sinh\left(\frac{F_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}\right) + p_i\left[\cosh\left(\frac{F_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}\right) - 1\right] \tag{14a}$$

$$\dot{s}\left(\frac{\delta t}{2}\right) = \frac{p_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\sinh\left(\frac{F_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}\right) + \cosh\left(\frac{F_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}\right) \tag{14b}$$

that is slightly different from eq S77 in ref 57, namely, the argument of hyperbolic functions was changed from $\sqrt{\frac{F_i^2}{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}$ to $\frac{F_i}{\sqrt{m_i L k_{\mathrm{B}}T}}\frac{\delta t}{2}$ in order for $F_i s_i$ to retain the sign of F_i. Propagators of $\mathcal{L}_N$ and $\mathcal{L}_O$ follow eqs S78 and S79 of ref 57.

Recalibrating se-QM/MM Models with Force Matching. In order to improve the efficiency of MTS simulations, it is preferable to set the outer time step to be as large as possible while maintaining the stability and accuracy of the simulations. In MTS ai-QM/MM MD simulations, the length of the outer time step is bound by the magnitude and nature of the correction

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

forces $\Delta\mathbf{F}^{\mathrm{SE}\rightarrow\mathrm{AI}}$ in eq 4. To achieve longer outer time steps, the correction forces $\Delta\mathbf{F}^{\mathrm{SE}\rightarrow\mathrm{AI}}$ need to be sufficiently small for the relevant configurations. In the limit of a "perfect" se-QM model that can fully reproduce the high-level forces for any configuration, i.e., $\Delta\mathbf{F}^{\mathrm{SE}\rightarrow\mathrm{AI}} \rightarrow 0$, the outer step would never be needed, i.e., $\Delta t \rightarrow \infty$. In the other limit, where a null se-QM model is with substantial and random errors in the atomic forces, the outer step size would have to be equal to the inner step size to correct the forces at each time step, i.e., $\Delta t = \delta t$, achieving no speedup at all. In practice, the accuracy of an se-QM/MM model for a specific system will fall between the two limits. By minimizing $\Delta\mathbf{F}^{\mathrm{SE}\rightarrow\mathrm{AI}}$, more accurate and efficient samplings can thus be attained for MTS ai-QM/MM MD.

One way to improve the accuracy of an se-QM model is to recalibrate the se parameters through force matching. In this work, the force matching protocol was improved upon the one from our previous paper,[43] which was in turn based on the reaction path force matching (RP-FM) method.[62] Inspired by the force balance method,[68] where the objective function is defined as the weighted sum of a set of scaled properties of the system, such as energies, forces, and dipole moments, we defined the objective function in this work to include errors in energies, bond forces, Cartesian forces with bond forces removed, and a quadratic penalty term that restrains the parameters to their initial values. The final form of the objective function is

$$\chi^2 = W_E \frac{\sum_{i\in N_s}((E_i^{\mathrm{SE}} - \langle E^{\mathrm{SE}}\rangle) - (E_i^{\mathrm{AI}} - \langle E^{\mathrm{AI}}\rangle))^2}{\sum_{i\in N_s}(E_i^{\mathrm{AI}} - \langle E^{\mathrm{AI}}\rangle)^2} + W_F \frac{\sum_{i\in N_s}\sum_{a\in N_{\mathrm{bond}}} w_a |\mathbf{F}_{i,a}^{\mathrm{SE}} - \mathbf{F}_{i,a}^{\mathrm{AI}}|^2}{\sum_{i\in N_s}\sum_{a\in N_{\mathrm{bond}}} w_a |\mathbf{F}_{i,a}^{\mathrm{AI}}|^2} + W_F' \frac{\sum_{i\in N_s}\sum_{b\in N_{\mathrm{atom}}} |\mathbf{F}_{i,b}^{'\mathrm{SE}} - \mathbf{F}_{i,b}^{'\mathrm{AI}}|^2}{\sum_{i\in N_s}\sum_{b\in N_{\mathrm{atom}}} |\mathbf{F}_{i,b}^{'\mathrm{AI}}|^2} + \lambda \sum_c \left|\frac{\xi_c - \xi_{c,0}}{\xi_{c,0}}\right|^2 \quad (15)$$

where N_s is the number of sampled configurations, $\langle E^{\mathrm{AI}}\rangle$ and $\langle E^{\mathrm{SE}}\rangle$ are the average ai-QM/MM and se-QM/MM energies for the samples, $\mathbf{F}_{i,a}$ refer to the forces on the ath bond of the ith configuration, $\mathbf{F}'_{i,b}$ correspond to the Cartesian forces on the bth atom (with associated bond forces projected out), and ξ_c are the empirical parameters for the se-QM model. In this work, we set $W_E = 1$, $W_F = 1$, $W'_F = 1$, $\lambda = 10^{-5}$, and $w_a = 1/N_a$ and reparameterized the PM3 se-QM model following the procedure described below. Henceforth, the recalibrated PM3 model is hereafter denoted as PM3*, and the optimized parameters can be found in Table S1 of ref 69.

The Broyden−Fletcher−Goldfarb−Shanno (BFGS) algorithm as implemented in the SciPy package was used to perform the minimization, without any bounds on the parameters. The first-order derivatives with respect to the parameters were calculated using the finite difference method. It should be noted that, like in our previous paper,[43] the procedure will only find a local minimum that is close to the standard parameter set, which is not necessarily the global minimum in the parameter space. However, it is not an issue in practice for a couple of reasons. First, the standard parameter set has been calibrated for a wide range of systems, so moving too far from it (to reach the global minimum) could increase the risk of going into nonphysical regions of the parameter space. Second, the se-QM model in MTS ai-QM/MM MD simulations is only used as the low-level method for the configurational sampling, so the inaccuracy of the model (at a local minimum) will be corrected by the high-level corrections at outer time steps. Nevertheless, if necessary, for example, when the accuracy of the recalibrated se-QM model is not sufficient, more elaborate nonlinear parameter optimization algorithms, such as the genetic algorithm, can be employed.

COMPUTATIONAL DETAILS

In this study, chorismate mutase, which catalyzes the conversion from chorismate to prephenate,[70−72] was chosen as the model system to test the proposed MTS method for computing the free energy profile of enzyme reactions. Due to a relatively small substrate with 24 atoms, the enzyme has been one of the most widely used testing systems for QM/MM and method developments.[38,73−85]

Starting from the X-ray crystal structure (PDB ID 2CHT[70]) of *Bacillus subtilis* chorismate mutase complexed with a transition state analog, the setup and equilibration of the system followed the same protocol in our previous study[43] using AMBER software.[65,66] The last frame of a 2 ns classical trajectory was used as the starting structure for the QM/MM MD simulations in this study.

We first conducted short ai-QM/MM MD simulations under periodic boundary conditions, which were used as the reference calculations for se-QM/MM parameter recalibration. The QM subsystem consisted of the substrate chorismate (i.e., 24 QM atoms) and was described by the B3LYP/6-31G* level of theory[86−89] as implemented in the Q-Chem software package.[64] The MM subsystem included the enzyme, water solvent molecules, and sodium counterions in the center cell, as well as all atoms in the image cells (including the QM images). The MM subsystem was described by ff14SB/GAFF/TIP3P forces fields[90−92] as implemented in the AMBER software package.[65,66]

The QM/MM-AC electrostatic embedding scheme[63] was used to handle both short- and long-range QM-MM electrostatic interactions. The particle mesh Ewald (PME) method[93,94] was used to treat the MM-MM electrostatic interactions with a real-space cutoff of 10 Å. Both the QM-MM and MM-MM van der Waals interactions were treated classically using the Lennard-Jones potential as defined in the Amber force fields[90] and truncated at a cutoff of 10 Å. The SHAKE algorithm[95] was used to constrain all the bonds involving hydrogen atoms in the MM subsystem. Langevin dynamics with a friction coefficient of 5 ps^{-1} was performed at 300 K, and a time step of 1 fs was used for the MD integration. As noted in the Methods section, both "end" and "middle" schemes were used for the thermostat in the leapfrog integration within the AMBER software package. For the SIN(R) simulations, $L = 4$ sets of thermostat variables were coupled to each degree of freedom with $Q_1 = Q_2 = k_B T(0.05\ \mathrm{ps})^2$ for the thermostat coupling parameters and $\gamma = 1\ \mathrm{ps}^{-1}$ for the friction coefficient. For the MTS ai-QM/MM molecular dynamics simulations, the same setup was used as the regular ai-QM/MM MD, except that a 1 fs time step was used for the inner steps, whereas the time step for the outer steps varied from 2 to 10 fs.

The umbrella sampling technique[96] was employed to construct the direct ai-QM/MM free energy profile along the reaction coordinate. In the present study, the reaction coordinate was defined as the difference between the bond lengths of the breaking C−O and forming C−C bonds. Forty windows were evenly distributed along the reaction coordinates

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

ranging from −1.95 to 1.95 Å, and the force constant of the harmonic biasing potential was set to be 300 kcal mol^{-1} Å^{-2} for all the windows. For each window, a 60 ps ai-QM/MM MD simulation was performed. The snapshots were collected every 40 steps during the last 40 ps simulation, which resulted in 1000 snapshots for postprocessing analysis. The multistate Bennett acceptance ratio (MBAR)[97] as implemented in the *pymbar* package was utilized to estimate the free energy profile.

RESULTS AND DISCUSSION

Timescale Separation of Reference and Correction Forces. The fundamental assumption of an MTS QM/MM simulation is a timescale separation of the reference and correction forces. Especially, the correction forces oscillate with a small amplitude and/or more slowly than that of the reference forces. To test this assumption, we postprocessed a 200 fs NVT trajectory in both reactant and transition state regions that was generated using direct B3LYP/MM simulations (i.e., without MTS) and calculated the atomic forces on the QM atoms with both PM3/MM and PM3*/MM models.

Figures S1 and S4 show the x component of the B3LYP/MM forces on all QM atoms, as well as the correction forces to PM3/MM and PM3*/MM models (i.e., the difference between B3LYP/MM forces and those from these models). A clear separation of timescales between the reference and correction forces is observed for both PM3/MM and PM3*/MM models. This is even more evident after a Fourier transformation is performed, where Figures S2 and S5 show that the high-frequency (>2000 cm^{-1}) oscillations become quite small for the correction forces. More important are the correction forces of hydrogen atoms, as their oscillation determines the time step for MD simulation. Between PM3/MM and PM3*/MM, the recalibrated potential (green line) shows smaller and more smoothly varying correction forces than the uncalibrated one (orange line). Together, this allows for longer outer integration steps and the generation of more accurate trajectories.

Removing Fast Components from the Correction Force. In order to determine the maximum outer time steps allowed, we have investigated the stability of the MTS QM/MM trajectory with different outer time steps. As expected, the MTS QM/MM trajectory became unstable with the increase in the outer time step. When the outer time step increased to a certain point (4 fs for PM3 and 6 fs for PM3*), we observed hydrogen atoms from C−H and O−H bonds detached from the C or O atoms.

In order to mitigate this, we first tested using the hydrogen mass repartitioning method[98] with a scaling factor of 3 or 4 to reduce the stretching frequencies of C−H and O−H bonds, which has been used to increase the outer time step for MTS simulations with polarizable force fields.[99] However, unlike in classical force fields where bonds are not able to break because of the use of harmonic potentials, we observed that the bonds between heavy atoms were broken during the MTS QM/MM simulations with larger outer time steps because of the reduced masses of the heavy atoms that were bonded to hydrogen atoms. This suggested that the hydrogen mass repartitioning method is not a viable strategy to increase the time step for QM/MM simulations.

Subsequently, we adopted an alternative strategy, where the force corrections associated with C−H and O−H bonds in the system were projected out. As shown in the left column of Figures S3 and S6, the PM3→B3LYP force corrections (blue line) for most of these bonds were reduced after the PM3 model recalibration (orange line). However, the remaining forces still caused the breaking of C−H and O−H bonds with larger outer time steps. When these bond force components were removed from the force corrections in the outer time steps, it improved the stability of the simulation (i.e., keeping the C−H and O−H bonds intact) even with a larger outer time step. Hereafter, we denote it as the "remove force" (RF) version of our MTS QM/MM simulations.

Alternatively, we applied the SHAKE algorithm[95] to freeze the lengths of O−H and C−H bonds in the QM region (at values from the GAFF force field). While some encouraging results will be shown later, this strategy is not necessarily universally applicable in part because many reactions involve proton transfer and in part because C−H and O−H bond lengths might change substantially in some reactions (such as Menshutkin reactions).

Instead of removing the fast components of the correction force, one can in principle apply the MTS correction only on key slow degrees of freedom associated with the reaction coordinate, in the spirit of the reaction path force matching in the collective variable (RF-FM-CV) method.[100] The effectiveness of this strategy will be examined in future studies.

Microcanonical Ensemble MD Simulations. We tested the stability of the MTS ai-QM/MM MD simulations with different combinations of inner step models and outer step sizes using microcanonical (NVE) ensemble MD simulations. The setup was the same as the one described in the Computational Details section, and no thermostat was applied. All NVE simulations started from the same initial condition and ran for 10 ps. For the Δt = 3 fs case, the energy fluctuation is shown in Figure 3.

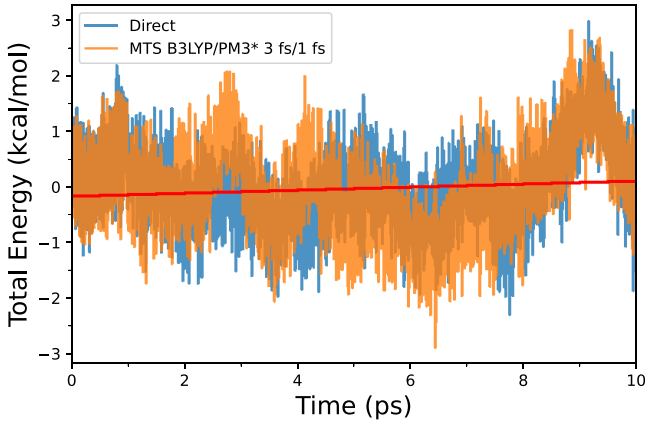


Figure 3. Fluctuation in the total energy (in kcal/mol) of the chorismate mutase system during a 10 ps MTS simulation with PM3*/MM at 1 fs inner time steps and B3LYP/6-31G*/MM force corrections at 3 fs outer time steps.

The energy fluctuation, which is measured by the standard deviation of the total energy along the trajectory, is listed in Table 1. Separately, the energy drifts, which correspond to the slope of the linear least-squares regression of the same data, are also listed. Only the trajectories that were stable during the 10 ps simulations were presented.

For the PM3 model, the MTS ai-QM/MM NVE MD was numerically stable up to Δt = 2 fs. Beyond that, the reactive region gets seriously distorted due to extra broken bonds. When the recalibrated PM3 model (PM3*) was used, the outer time step Δt could reach as far as 3 fs without artificial breakage of

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 1. Energy Fluctuations and Drifts in 10 ps NVE Simulations with B3LYP/MM as well as Several MTS Schemes

	fluctuation (kcal/mol)	drift (kcal/mol/ps)
B3LYP, 1 fs	0.74	0.042
B3LYP/PM3, 2 fs/1 fs	0.78	0.006
B3LYP/PM3*, 2 fs/1 fs	0.74	−0.043
B3LYP/PM3*, 3 fs/1 fs	0.82	0.028
B3LYP/PM3*, 4 fs/1 fs, SHAKE	0.63	−0.058
B3LYP/PM3*, 5 fs/1 fs, SHAKE	0.76	0.017
B3LYP/PM3*, 6 fs/1 fs, SHAKE	0.88	−0.067

chemical bonds involving hydrogen atoms. When the SHAKE algorithm was applied to constrain the C−H and O−H bonds in the QM region, the outer time steps could be increased up to 6 fs.

Overall, the energy fluctuation in all MTS simulations ranged from 0.63 to 0.88 kcal/mol, which was comparable to the 0.74 kcal/mol fluctuation for the single-Hamiltonian simulation using the B3LYP/MM energy function. The energy drifts were also comparable. This confirmed that our implementation of the MTS algorithm maintained the energy conservation, largely due to a rigorous treatment of QM/MM electrostatics in a periodic system using our QM/MM-AC model.

Free Energy Profile. Next, we calculated the free energy profile for each of the umbrella sampling simulations with our MTS QM/MM method, where the outer time steps were stretched to 8 fs with the use of Langevin dynamics without the issue of C−H and H−O bond breakage. With the "end" scheme for the thermostat, the free energy profiles (Figure 4A) are all in good agreement with the direct QM/MM simulations. In terms of the sampled pathways, overall, the MTS ones were in agreement with the direct QM/MM one (Figure 4B). However, noticeable deviations were observed in the transition state regions. Especially, the position of the highest-energy point on the free energy profiles was off by ∼0.1 Å in terms of the C−O bond length for the MTS simulations.

We also tested the "middle" scheme Langevin dynamics for the QM/MM MTS simulations using 6 and 8 fs outer time steps and obtained similar free energy profiles (Figure 4C). Very encouragingly, the sampled pathways for the "middle" scheme in Figure 4D were in better agreement with the direct QM/MM one than the ones from the "end" scheme simulations (Figure 4A). This supports that the "middle" scheme Langevin dynamics can give more accurate ensembles in the configurational space than its "end" scheme counterpart.[57,67]

So far, these free energy profiles were obtained from MTS simulations within our RF scheme, where the correction forces associated with O−H and C−H bonds in the QM region were removed at outer time steps. Also shown in Figure 4 are MTS simulation results with SHAKE constraints on O−H and C−H bonds in the QM region. The free energy profile in Figure 4E with an 6 fs outer time step also matched well with direct ai-QM/MM results. When the outer time step was further increased to 8 fs, the free energy profile yielded a barrier that was ∼1.0 kcal/mol higher. This suggested that an adoption of the SHAKE algorithm, when applicable, could produce less accurate results for MTS simulations with longer time steps than our RF scheme. We note, however, that this is not a completely consistent comparison because we have yet to perform the direct simulation (i.e., no-MTS ai-QM/MM simulation) with SHAKE.

Finally, as shown in Figure 4G,H, MTS QM/MM simulations with the SIN(R) thermostat remained stable with an outer time step of up to 10 fs. When a 6 fs outer time step was used, both the predicted free energy barrier and reaction free energy were well within 1 kcal/mol from the direct ai-QM/MM values. For 8 fs outer time step simulations, the predicted free energy barrier was ∼1.2 kcal/mol higher than the direct ai-QM/MM value. When the outer time step was increased to 10 fs, the differences in the free energy barrier and reaction free energy increased to *ca.* 1.8 and −2.5 kcal/mol, respectively.

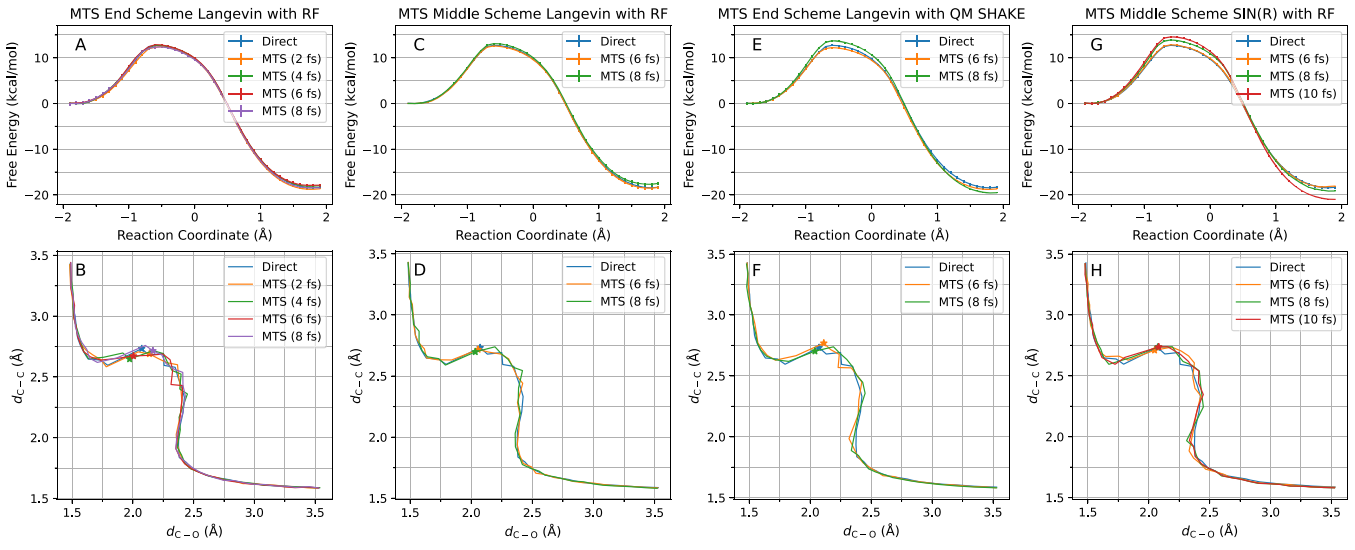


Figure 4. (A) Free energy profiles and (B) sampled pathways of the chorismate mutase reaction from direct and MTS simulations with "end" scheme Langevin dynamics and (C,D) results from "middle" scheme Langevin dynamics. (E,F) Results from MTS simulations with QM SHAKE. (G,H) Results from MTS simulations with the "middle" scheme SIN(R) thermostat. In MTS simulations, PM3*/MM forces are evaluated at each inner step (1 fs), while B3LYP/6-31G*/MM force evaluations are carried out at 2−10 fs intervals. The stars show the locations of the transition states on the pathways.

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

To further test the accuracy of the MTS simulations, we calculated the average potential energy per atom[57,67,101] for direct and MTS simulations in the transition state window of the umbrella sampling simulations (Figure S7). For all the models, the distributions of the average potential energy per atom were found to be consistent across the various outer time steps in use, which shows the robustness of our MTS model.

Speedup. We can use t_{QM}^{high} and t_{QM}^{low} to represent the time required for a single energy and force calculation for the QM region and the QM/MM interactions at high- and low-level methods, respectively. Meanwhile, t_{MM} corresponds to all other times required for a single MD step (including MM force calculation, MD integration, and communication overhead of calling external QM programs). The speedup of an MTS ai-QM/MM MD simulation relative to a standard ai-QM/MM one can then be estimated using

$$s = \frac{t_{QM}^{high} + t_{MM}}{\frac{1}{N}t_{QM}^{high} + t_{QM}^{low} + t_{MM}} \quad (16)$$

where the denominator—the average cost of each MTS time step—contains the cost of computing high-level QM/MM forces at the outer time steps (divided evenly over all the time steps).

In the limit that t_{QM}^{low} and t_{MM} are negligible compared to t_{QM}^{high}, the maximum speedup will tend to N. In practice, the speedup will be smaller than N, and the actual speedup depends on the ratio between t_{QM}^{high} and t_{QM}^{low} if t_{MM} is negligible. Figure 5 shows the

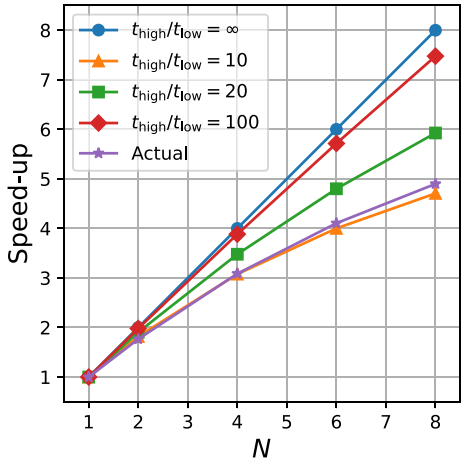


Figure 5. Theoretical and actual speedups for NVT calculations with the MTS integrator. The theoretical speedups are shown for various $t_{QM}^{high}/t_{QM}^{low}$ ratios. The actual speedup comes from MTS simulations of the chorismate mutase reaction with PM*/MM calculations at inner time steps and B3LYP/6-31G*/MM calculations at outer time steps.

maximum speedup of different $t_{QM}^{high}/t_{QM}^{low}$ ratios and actual speedup of the system under study. Our current implementation of the MTS ai-QM/MM method achieved a 5-fold speedup for the chorismate mutase system, when $N = 8$. For larger systems, more significant speedup is expected due to higher $t_{QM}^{high}/t_{QM}^{low}$ ratios.

CONCLUSIONS

In this work, we explored the use of a recalibrated PM3 model for the inner time steps of MTS QM/MM molecular dynamics simulations of chorismate mutase catalysis. The main findings were as follows:

- The recalibrated PM3 model led to smaller correction forces at the outer steps, which allowed the use of longer outer time steps.
- In order to prevent nonphysical breakage of the C−H and O−H bonds with the use of longer time steps (such as 8 fs), one can either project out the force corrections associated with these bonds, or when applicable, apply SHAKE constraints to these bonds in the QM region.
- There are a couple of alternative ways to increase the outer time steps in NVT MTS QM/MM simulations. For the simulation of the chorismate mutase reaction, the outer time step of Langevin dynamics simulations could be increased to 6 fs with SHAKE and 8 fs with the removal of C−H/O−H force corrections. The outer time step can reach 10 fs with SIN(R), if C−H/O−H force corrections are removed.
- With an 8 fs outer time step, it enabled a 5-fold speedup in the computational time for acquiring the chorismate mutase reaction free energy profile.

We are applying the QM/MM MTS simulation in the modeling of several other enzymes with larger QM regions, which will be reported in the future.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.2c02262.

The x components of atomic forces of the QM atoms along 200 fs DFT/MM MD trajectories in the reactant (Figure S1) and the transition state (Figure S4) regions of the chorismate mutase reaction, their corresponding Fourier transformations (Figures S2 and S5), the bond forces for the bonds that involve hydrogen atoms along the same trajectories (Figures S3 and S6), and the average potential energy per atom for direct and MTS simulations in the transition state window of the umbrella sampling simulations (Figure S7) (PDF)

AUTHOR INFORMATION

Corresponding Authors

Jingzhi Pu − *Department of Chemistry and Chemical Biology, Indiana University−Purdue University Indianapolis, Indianapolis, Indiana 46202, United States;* orcid.org/0000-0002-3042-335X; Email: jpu@iupui.edu

Kwangho Nam − *Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, Texas 76019, United States;* orcid.org/0000-0003-0723-7839; Email: kwangho.nam@uta.edu

Yihan Shao − *Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019-5251, United States;* orcid.org/0000-0001-9337-341X; Email: yihan.shao@ou.edu

Authors

Xiaoliang Pan − *Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019-5251, United States;* orcid.org/0000-0002-6399-4853

Richard Van − *Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019-5251, United States*

Evgeny Epifanovsky − *Q-Chem, Inc., Pleasanton, California 94588, United States*

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Jian Liu − *Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China;* orcid.org/0000-0002-2906-5858

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.2c02262

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health through grant R43GM133270 to E.E., K.N., and Y.S. J.P. and Y.S. are also supported by the National Institutes of Health (grant R01GM135392). K.N. is also supported by the National Institutes of Health (grants R01GM132481 and R01GM138472). J.L. acknowledges support from the National Natural Science Foundation of China (NSFC) Grant No. 21961142017 and from the Ministry of Science and Technology of China (MOST) Grant No. 2017YFA0204901. The authors thank the OU Supercomputing Center for Education & Research (OSCER) for the computational resources.

■ REFERENCES

(1) 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.

(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. A Priori Computation of a Solvent-Enhanced SN2 Reaction Profile in Water: the Menshutkin Reaction. *J. Am. Chem. Soc.* **1991**, *113*, 7796−7797.

(4) Gao, J. Absolute Free Energy of Solvation From Monte Carlo Simulations Using Combined Quantum and Molecular Mechanical Potentials. *J. Phys. Chem.* **1992**, *96*, 537−540.

(5) Gao, J.; Truhlar, D. G. Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.

(6) Senn, H. M.; Thiel, W. QM/MM Studies of Enzymes. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182−187.

(7) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go From Here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(8) Hu, H.; Yang, W. Free Energies of Chemical Reactions in Solution and in Enzymes with Ab Initio Quantum Mechanics/Molecular Mechanics Methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573−601.

(9) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(10) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708−2728.

(11) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM Simulations: What Are Some of the âĂIJBurning IssuesâĂİ? *J. Phys. Chem. B* **2021**, *125*, 689−702.

(12) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, N.Y, 1996.

(13) Helgaker, T.; JÃÿrgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory; Wiley: Chichester* ; New York, 2000.

(14) Sherrill, C. D. Frontiers in Electronic Structure Theory. *J. Chem. Phys.* **2010**, *132*, 110902.

(15) Cui, Q. Perspective: Quantum Mechanical Methods in Biochemistry and Biophysics. *J. Chem. Phys.* **2016**, *145*, 140901.

(16) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*, 1st ed.; International series of monographs on chemistry 16; Oxford Univ. Press [u.a.]: New York, NY, 1994; OCLC: 832732716.

(17) Neese, F. Prediction of Molecular Properties and Molecular Spectroscopy with Density Functional Theory: From Fundamental Theory to Exchange-cCupling. *Coord. Chem. Rev.* **2009**, *253*, 526−563.

(18) Burke, K. Perspective on Density Functional Theory. *J. Chem. Phys.* **2012**, *136*, 150901.

(19) Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Chem. Phys.* **2014**, *140*, 18A301.

(20) Yu, H. S.; Li, S. L.; Truhlar, D. G. Perspective: Kohn-Sham Density Functional Theory Descending a Staircase. *J. Chem. Phys.* **2016**, *145*, 130901.

(21) Hu, P.; Zhang, Y. Catalytic Mechanism and Product Specificity of the Histone Lysine Methyltransferase SET7/9: an ab initio QM/MM-FE Study with Multiple Initial Structures. *J. Am. Chem. Soc.* **2006**, *128*, 1272−1278.

(22) Wang, L.; Yu, X.; Hu, P.; Broyde, S.; Zhang, Y. A Water-Mediated and Substrate-Assisted Catalytic Mechanism for *Sulfolobus solfataricus* DNA Polymerase IV. *J. Am. Chem. Soc.* **2007**, *129*, 4731−4737.

(23) Wu, R.; Wang, S.; Zhou, N.; Cao, Z.; Zhang, Y. A Proton-Shuttle Reaction Mechanism for Histone Deacetylase 8 and the Catalytic Role of Metal Ions. *J. Am. Chem. Soc.* **2010**, *132*, 9471−9479.

(24) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **2011**, *133*, 8934−8941.

(25) Lior-Hoffmann, L.; Wang, L.; Wang, S.; Geacintov, N. E.; Broyde, S.; Zhang, Y. Preferred WMSA Catalytic Mechanism of the Nucleotidyl Transfer Reaction in Human DNA Polymerase kappa Elucidates Error-free Bypass of a Bulky DNA Lesion. *Nucleic Acids Res.* **2012**, *40*, 9193−9205.

(26) Wong, K.-Y.; Gu, H.; Zhang, S.; Piccirilli, J. A.; Harris, M. E.; York, D. M. Characterization of the Reaction Path and Transition States for RNA Transphosphorylation Models from Theory and Experiment. *Angew. Chem., Int. Ed.* **2012**, *51*, 647−651.

(27) Ganguly, A.; Thaplyal, P.; Rosta, E.; Bevilacqua, P. C.; Hammes-Schiffer, S. Quantum Mechanical/Molecular Mechanical Free Energy Simulations of the Self-Cleavage Reaction in the Hepatitis Delta Virus Ribozyme. *J. Am. Chem. Soc.* **2014**, *136*, 1483−1496.

(28) Stevens, D. R.; Hammes-Schiffer, S. Exploring the Role of the Third Active Site Metal Ion in DNA Polymerase Eta with QM/MM Free Energy Simulations. *J. Am. Chem. Soc.* **2018**, *140*, 8965−8969.

(29) Casalino, L.; Nierzwicki, L.; Jinek, M.; Palermo, G. Catalytic Mechanism of Non-Target DNA Cleavage in CRISPR-Cas9 Revealed by *Ab Initio* Molecular Dynamics. *ACS Catal.* **2020**, *10*, 13596−13605.

(30) Dürr, S. L.; Bohuszewicz, O.; Berta, D.; Suardiaz, R.; Jambrina, P. G.; Peter, C.; Shao, Y.; Rosta, E. The Role of Conserved Residues in the DEDDh Motif: the Proton-Transfer Mechanism of HIV-1 RNase H. *ACS Catal.* **2021**, *11*, 7915−7927.

(31) 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.

(32) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.

(33) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.

(34) 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* **1998**, *58*, 7260−7268.

(35) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M. Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes. *J. Phys. Chem. B* **2006**, *110*, 6458−6469.

(36) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method: Third-Order Expansion of the Density Functional Theory Total Energy and Introduction of a Modified Effective Coulomb Interaction. *J. Phys. Chem. A* **2007**, *111*, 10861−10873.
(37) Hou, G.; Zhu, X.; Elstner, M.; Cui, Q. A Modified QM/MM Hamiltonian with the Self-Consistent-Charge Density-Functional-Tight-Binding Theory for Highly Charged QM Regions. *J. Chem. Theory Comput.* **2012**, *8*, 4293−4304.
(38) Marí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Transition Structure Selectivity in Enzyme Catalysis: A QM/MM Study of Chorismate Mutase. *Theor. Chem. Acc.* **2001**, *105*, 207−212.
(39) Muller, R. P.; Warshel, A. Ab Initio Calculations of Free Energy Barriers for Chemical Reactions in Solution. *J. Phys. Chem.* **1995**, *99*, 17516−17524.
(40) Heimdal, J.; Ryde, U. Convergence of QM/MM free-energy perturbations based on molecular-mechanics or semiempirical simulations. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12592.
(41) Polyak, I.; Benighaus, T.; Boulanger, E.; Thiel, W. Quantum Mechanics/Molecular Mechanics Dual Hamiltonian Free Energy Perturbation. *J. Chem. Phys.* **2013**, *139*, No. 064105.
(42) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi-Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. *J. Chem. Theory Comput.* **2018**, *14*, 5583−5596.
(43) Pan, X.; Li, P.; Ho, J.; Pu, J.; Mei, Y.; Shao, Y. Accelerated Computation of Free Energy Profile at *ab initio* Quantum Mechanical/Molecular Mechanical Accuracy *via* a Semi-Empirical Reference Potential. II. Recalibrating Semi-Empirical Parameters with Force Matching. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20595−20605.
(44) Hu, W.; Li, P.; Wang, J.-N.; Xue, Y.; Mo, Y.; Zheng, J.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at *Ab Initio* Quantum Mechanical/Molecular Mechanics Accuracy via a Semiempirical Reference Potential. 3. Gaussian Smoothing on Density-of-States. *J. Chem. Theory Comput.* **2020**, *16*, 6814−6822.
(45) Wang, J.-N.; Liu, W.; Li, P.; Mo, Y.; Hu, W.; Zheng, J.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at *Ab Initio* Quantum Mechanical/Molecular Mechanics Accuracy via a Semiempirical Reference Potential. 4. Adaptive QM/MM. *J. Chem. Theory Comput.* **2021**, *17*, 1318−1325.
(46) Zwanzig, R. W. High−Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(47) Crooks, G. E. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.* **1998**, *90*, 1481−1487.
(48) Hummer, G.; Szabo, A. Free Energy Profiles from Single-Molecule Pulling Experiments. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 21441−21446.
(49) Ngo, V. A.; Kim, I.; Allen, T. W.; Noskov, S. Y. Estimation of Potentials of Mean Force from Nonequilibrium Pulling Simulations Using Both Minh-Adib Estimator and Weighted Histogram Analysis Method. *J. Chem. Theory Comput.* **2016**, *12*, 1000−1010.
(50) Hudson, P. S.; Woodcock, H. L.; Boresch, S. Use of Nonequilibrium Work Methods to Compute Free Energy Differences Between Molecular Mechanical and Quantum Mechanical Representations of Molecular Systems. *J. Phys. Chem. Lett.* **2015**, *6*, 4850−4856.
(51) Kearns, F. L.; Hudson, P. S.; Woodcock, H. L.; Boresch, S. Computing Converged Free Energy Differences Vetween Levels of Theory via Nonequilibrium Work Methods: Challenges and Opportunities. *J. Comput. Chem.* **2017**, *38*, 1376−1388.
(52) Abreu, C. R. A.; Tuckerman, M. E. Multiple Timescale Molecular Dynamics with Very Large Time Steps: Avoidance of Resonances. *Eur. Phys. J. B* **2021**, *94*, 231.
(53) 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.
(54) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(55) Leimkuhler, B.; Margul, D. T.; Tuckerman, M. E. Stochastic, Resonance-Free Multiple Time-Step Algorithm for Molecular Dynamics with Very Large Time Steps. *Mol. Phys.* **2013**, *111*, 3579−3594.
(56) Margul, D. T.; Tuckerman, M. E. A Stochastic, Resonance-Free Multiple Time-Step Algorithm for Polarizable Models That Permits Very Large Time Steps. *J. Chem. Theory Comput.* **2016**, *12*, 2170−2180.
(57) Zhang, Z.; Liu, X.; Yan, K.; Tuckerman, M. E.; Liu, J. Unified Efficient Thermostat Scheme for the Canonical Ensemble with Holonomic or Isokinetic Constraints via Molecular Dynamics. *J. Phys. Chem. A* **2019**, *123*, 6056−6079.
(58) Chen, Y.; Kale, S.; Weare, J.; Dinner, A. R.; Roux, B. Multiple Time-Step Dual-Hamiltonian Hybrid Molecular Dynamics − Monte Carlo Canonical Propagation Algorithm. *J. Chem. Theory Comput.* **2016**, *12*, 1449−1458.
(59) 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.
(60) Liberatore, E.; Meli, R.; Rothlisberger, U. A Versatile Multiple Time Step Scheme for Efficient *ab Initio* Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14*, 2834−2842.
(61) Kirsch, T.; Olsen, J. M. H.; Bolnykh, V.; Meloni, S.; Ippoliti, E.; Rothlisberger, U.; Cascella, M.; Gauss, J. Wavefunction-Based Electrostatic-Embedding QM/MM Using CFOUR through MiMiC. *J. Chem. Theory Comput.* **2021**, DOI: 10.1021/acs.jctc.1c00878.
(62) 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.
(63) Pan, X.; Nam, K.; Epifanovsky, E.; Simmonett, A. C.; Rosta, E.; Shao, Y. A Simplified Charge Projection Scheme for Long-Range Electrostatics in *ab initio* QM/MM Calculations. *J. Chem. Phys.* **2021**, *154*, No. 024115.
(64) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C. M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A., Jr.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T. C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S. P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y. C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z. Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J. D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C. P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W. Z.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113*, 184−215.

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(65) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(66) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198−210.
(67) Zhang, Z.; Liu, X.; Chen, Z.; Zheng, H.; Yan, K.; Liu, J. A Unified Thermostat Scheme for Efficient Configurational Sampling for Classical/Quantum Canonical Ensembles via Molecular Dynamics. *J. Chem. Phys.* **2017**, *147*, No. 034109.
(68) Wang, L.-P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9*, 452−460.
(69) Pan, X.; Yang, J.; Van, R.; Epifanovsky, E.; Ho, J.; Huang, J.; Pu, J.; Mei, Y.; Nam, K.; Shao, Y. Machine-Learning-Assisted Free Energy Simulation of Solution-Phase and Enzyme Reactions. *J. Chem. Theory Comput.* **2021**, *17*, 5745−5758.
(70) Chook, Y. M.; Ke, H.; Lipscomb, W. N. Crystal Structures of the Monofunctional Chorismate Mutase from Bacillus Subtilis and its Complex with a Transition State Analog. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 8600−8603.
(71) Kast, P.; Asif-Ullah, M.; Hilvert, D. Is Chorismate Mutase a Prototypic Entropy Trap? - Activation Parameters for the Bacillus Subtilis Enzyme. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(72) KienhÃűfer, A.; Kast, P.; Hilvert, D. Selective Stabilization of the Chorismate Mutase Transition State by a Positively Charged Hydrogen Bond Donor. *J. Am. Chem. Soc.* **2003**, *125*, 3206−3207.
(73) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. Insights into Chorismate Mutase Catalysis from a Combined QM/MM Simulation of the Enzyme Reaction. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.
(74) MartÃη, S.; AndrÃl's, J.; Moliner, V.; Silla, E.; TuÃśÃşn, I.; BertrÃąn, J. A QM/MM Study of the Conformational Equilibria in the Chorismate Mutase Active Site. The Role of the Enzymatic Deformation Energy Contribution. *J. Phys. Chem. B* **2000**, *104*, 11308−11315.
(75) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. Substrate conformational transitions in the active site of chorismate mutase: Their role in the catalytic mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 9032−9037.
(76) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. Understanding the Role of Active-Site Residues in Chorismate Mutase Catalysis from Molecular-Dynamics Simulations. *Angew. Chem., Int. Ed.* **2003**, *42*, 1508−1511.
(77) Marí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Preorganization and Reorganization as Related Factors in Enzyme Catalysis: The Chorismate Mutase Case. *Chem. − Eur. J.* **2003**, *9*, 984−991.
(78) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Insights Into Enzyme Catalysis from QM/MM Modelling: Transition State Stabilization in Chorismate Mutase. *Mol. Phys.* **2003**, *101*, 2695−2714.
(79) Lee Woodcock, H.; Hodošček, M.; Sherwood, P.; Lee, Y. S.; Schaefer, H. F., III; Brooks, B. R. Exploring the Quantum Mechanical/Molecular Mechanical Replica Path Method: a Pathway Optimization of the Chorismate to Prephenate Claisen Rearrangement Catalyzed by Chorismate Mutase. *Theor. Chem. Acc.* **2003**, *109*, 140−148.
(80) Szefczyk, B.; Mulholland, A. J.; Ranaghan, K. E.; Sokalski, W. A. Differential Transition-State Stabilization in Enzyme Catalysis: Quantum Chemical Analysis of Interactions in the Chorismate Mutase Reaction and Prediction of the Optimal Catalytic Field. *J. Am. Chem. Soc.* **2004**, *126*, 16148−16159.
(81) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of Chorismate Mutase Catalysis by QM/MM Modelling of Enzyme-Catalysed and Uncatalysed Reactions. *Org. Biomol. Chem.* **2011**, *9*, 1578.
(82) Burschowsky, D.; van Eerde, A.; Ökvist, M.; Kienhöfer, A.; Kast, P.; Hilvert, D.; Krengel, U. Electrostatic Transition State Stabilization Rather than Reactant Destabilization Provides the Chemical Basis for Efficient Chorismate Mutase Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 17516−17521.
(83) Freindorf, M.; Tao, Y.; Sethio, D.; Cremer, D.; Kraka, E. New Mechanistic Insights into the Claisen Rearrangement of Chorismate âĂŞ a Unified Reaction Valley Approach study. *Mol. Phys.* **2019**, *117*, 1172−1192.
(84) Ranaghan, K. E.; Shchepanovska, D.; Bennie, S. J.; Lawan, N.; Macrae, S. J.; Zurek, J.; Manby, F. R.; Mulholland, A. J. Projector-Based Embedding Eliminates Density Functional Dependence for QM/MM Calculations of Reactions in Enzymes and Solution. *J. Chem. Inf. Model.* **2019**, *59*, 2063−2078.
(85) Brickel, S.; Meuwly, M. Molecular Determinants for Rate Acceleration in the Claisen Rearrangement Reaction. *J. Phys. Chem. B* **2019**, *123*, 448−456.
(86) Becke, A. D. Density-functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(87) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(88) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(89) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(90) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(91) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(92) 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.
(93) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An *N*log(*N*) Method for Ewald sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(94) 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.
(95) 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.
(96) Torrie, G. M.; Valleau, J. P. Monte Carlo Free Energy Estimates Using Non-Boltzmann Sampling: Application to the Sub-Critical Lennard-Jones Fluid. *Chem. Phys. Lett.* **1974**, *28*, 578−581.
(97) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.
(98) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. Improving Efficiency of Large Time-Scale Molecular Dynamics Simulations of Hydrogen-Rich Systems. *J. Comput. Chem.* **1999**, *20*, 786−798.
(99) Lagardère, L.; Aviat, F.; Piquemal, J.-P. Pushing the Limits of Multiple-Time-Step Strategies for Polarizable Point Dipole Molecular Dynamics. *J. Phys. Chem. Lett.* **2019**, *10*, 2593−2599.
(100) Kim, B.; Snyder, R.; Nagaraju, M.; Zhou, Y.; Ojeda-May, P.; Keeton, S.; Hege, M.; Shao, Y.; Pu, J. Reaction Path-Force Matching in Collective Variables: Determining Ab Initio QM/MM Free Energy Profiles by Fitting Mean Force. *J. Chem. Theory Comput.* **2021**, *17*, 4961−4980.
(101) Sun, Z.; Kalhor, P.; Xu, Y.; Liu, J. Extensive Numerical Tests of Leapfrog Integrator in Middle Thermostat Scheme in Molecular Simulations. *Chin. J. Chem. Phys.* **2021**, *34*, 932−948.

https://doi.org/10.1021/acs.jpcb.2c02262
J. Phys. Chem. B XXXX, XXX, XXX−XXX