Accelerated Computation of Free Energy Profile at ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi-Empirical Reference Potential. II. Recalibrating Semi-Empirical Parameters with Force Matching[†]

Xiaoliang Pan,[a] Pengfei Li,[b] Junming Ho[c], Jingzhi Pu[*d], Ye Mei,[*be] and Yihan Shao[*a]

An efficient and accurate reference potential simulation protocol is proposed for producing *ab initio* quantum mechanical molecular mechanical (AI-QM/MM) quality free energy profiles for chemical reactions in a solvent or macromolecular environment. This protocol involves three stages: (a) using force matching to recalibrate a semi-empirical quantum mechanical (SE-QM) Hamiltonian for the specific reaction under study; (b) employing the recalibrated SE-QM Hamiltonian (in combination with molecular mechanical force fields) as the reference potential to drive umbrella samplings along the reaction pathway; and (c) computing AI-QM/MM energy values for collected configurations from the sampling and performing weighted thermodynamic perturbation to acquire AI-QM/MM corrected reaction free energy profile. For three model reactions (identity S_N2 reaction, Menshutkin reaction, and glycine proton transfer reaction) in aqueous solution and one enzyme reaction (Claisen arrangement in chorismate mutase), our simulations using recalibrated PM3 SE-QM Hamiltonians well reproduced AI-QM/MM free energy profiles (at the B3LYP/6-31G* level of theory) all within 1 kcal/mol with a 20 to 45 fold reduction in the computer time.

1 Introduction

In the computational modeling of an enzyme or solute-phase reaction, the central task is to assess the free energy (FE) profile for the entire reaction.[1–4] Here, the most direct and accurate approach would be to explore the free energy landscape using a combined *ab initio* quantum mechanical molecular mechanical (AI-QM/MM) Hamiltonian,[5] where atoms in the reactive region are described by *ab initio* quantum mechanical (AI-QM) methods while all other enzyme and solvent atoms are described by molecular mechanics (MM) force fields. Indeed, direct AI-QM/MM free energy simulations have been performed on several key enzymatic reactions[6–12] and provided invaluable insights into the respective reaction mechanism.

Unfortunately, direct AI-QM/MM free energy simulations can be extremely costly. For a typical enzyme reaction, the construction of a reaction free energy profile usually requires 20–50 sampling windows along the reaction pathway, each of which with 100 ps or longer configuration sampling. Typically there are at least 50–100 QM atoms in the reactive region, a direct and well-converged AI-QM/MM free energy simulation could thus easily require 250,000 or more CPU hours. Such simulations are beyond the routine use of most computational and experimental chemists.

In this work, we focus on the *indirect* AI-QM/MM free energy simulation methodology, as pioneered by Gao, Warshel and others,[13–15] for a substantial reduction in the computational cost of AI-QM/MM free energy simulations. For a typical reaction, an indirect AI-QM/MM simulation would involve three stages: (a) configuration sampling and iterative pathway optimization with a lower-level Hamiltonian (i.e. reference potential); (b) post-processing analysis [using weighted histogram analysis method (WHAM),[16] multistate Bennett acceptance ratio (MBAR),[17] umbrella integration,[18] variational free energy profile[19,20] methods, etc] to acquire the lower-level reaction free energy profile; and (c) thermodynamical perturbation[21] to a target AI-QM/MM Hamiltonian to obtain a corrected free energy profile.

In recent years, semi-empirical QM/MM (SE-QM/MM) Hamil-

tonians have been proposed by Ryde, Thiel and others to be used as the reference potential in indirect AI-QM/MM reaction free energy simulations. [15,22] While very encouraging results have been obtained from such "SE-QM/MM → AI-QM/MM" simulations, further improvement to the methodology is needed before its routine use in reaction free energy simulations.

In a recent publication, [23] we considered the third and final stage of a "SE-QM/MM → AI-QM/MM" indirect free energy simulation. In particular, a theoretically rigorous weighted thermodynamical perturbation analysis was proposed for the free energy correction, which properly takes into consideration configuration weights [from the post-processing analysis of the SE-QM/MM trajectories using the MBAR method].

In this work, we will seek to improve another aspect of indirect reaction free energy simulations — SE-QM/MM sampling. Clearly, in these free energy calculations, we want all important configurations on the AI-QM/MM potential energy surface to be sampled during the SE-QM/MM simulation. But, most SE-QM/MM Hamiltonians display limited similarity ("overlap") to AI-QM/MM ones, which can lead to inadequate sampling of important AI-QM/MM configurations and in turn cause slow convergence and low accuracy in the AI-QM/MM-corrected free energy results. As a way to enhance the similarity, here we propose the use of reaction pathway force-matching[24] to calibrate semi-empirical QM models (against the target AI-QM model) for each specific reaction of interest. With a re-parameterized semi-empirical PM3 model (PM3*), for example, our indirect "PM3*/MM → B3LYP/MM" simulations will be shown to reproduce B3LYP/MM free energy profiles within 1 kcal/mol for four model condensed-phase reactions, while retaining the computational efficiency of indirect free energy simulations.

Overall, the force-matching indirect AI-QM/MM free energy simulation methodology in this work provides a simple and robust alternative to several other strategies (such as machine-learning, [25–27] non-equilibrium-work, [28] and multiple-time-step approaches [29,30]) for combining the use of SE-QM/MM and AI-QM/MM Hamiltonians in reaction free energy simulations, and to sequential sampling, [31,32] steered molecular dynamics [33], or metadynamics [34–36] approaches to directly accelerate AI-QM/MM free energy simulations.

2 Method

2.1 Weighted Thermodynamic Perturbation

An indirect reaction free energy simulation starts from umbrella sampling simulations using a low-level Hamiltonian. Due to the high efficiency of the low-level method comparing to the high-level method, this leads to more extensive samplings of the potential energy surface within a finite amount of computer time. Then the FE profile $F_L(\eta)$ and the corresponding uncertainty $\delta^2 F_L(\eta)$ at the low-level Hamiltonian, where η's are bin indices along the reaction pathway, can be estimated by one of umbrella sampling post-processing methods, such as WHAM, [16] MBAR, [17] umbrella integration, [18] and variational free energy profile [19,20] methods. MBAR is the preferred method in our indirect free energy simulations, because we are interested in not only the FE profile at the

low-level Hamiltonian, but also the MBAR weight factor for each frame in the umbrella sampling. In this work, the MBAR analysis was carried out by using the pymbar program.

The weighted thermodynamic perturbation (TP) method [23] has been proposed recently by some of us to estimate free energy (FE) profiles at the ab initio QM/MM level by correcting the FE profiles obtained from umbrella samplings with semi-empirical reference potentials. A detailed description of the weighted TP method can be found in the original paper, and only a brief introduction will be presented here.

When the free energy difference $\Delta F(\eta)$ and the corresponding uncertainty $\delta^2 \Delta F(\eta)$ between the high- and low-level Hamiltonians for each histogram bin is calculated by the weighted TP method, *frames in each bin are weighted by their weight factors from the MBAR analysis instead of being weighted equally* since the frames are sampled from biased simulations. Then, the FE profile at the high-level Hamiltonian can be obtained by adding the free energy difference to the FE profile at the low-level Hamiltonian for each bin of the histogram,

$$F_H(\eta) = F_L(\eta) + \Delta F(\eta). \tag{1}$$

To characterize the reliability of the TP calculations, the reweighting entropy was computed. For bin η, its reweighting entropy is defined as [37]

$$\mathscr{S}(\eta) = -\frac{1}{\ln N_\eta} \sum_{i_\eta=1}^{N_\eta} \mathscr{P}(\boldsymbol{x}_{i_\eta}) \ln \mathscr{P}(\boldsymbol{x}_{i_\eta}), \tag{2}$$

where

$$\mathscr{P}(\boldsymbol{x}_{i_\eta}) = \frac{w_L(\boldsymbol{x}_{i_\eta}) e^{-\beta[U_H(\boldsymbol{x}_{i_\eta}) - U_L(\boldsymbol{x}_{i_\eta})]}}{\sum_{i_\eta=1}^{N_\eta} w_L(\boldsymbol{x}_{i_\eta}) e^{-\beta[U_H(\boldsymbol{x}_{i_\eta}) - U_L(\boldsymbol{x}_{i_\eta})]}}. \tag{3}$$

Here $w_L(\boldsymbol{x}_{i_\eta})$ is the MBAR weight factor for the ith frame in bin η, and $U_H(\boldsymbol{x}_{i_\eta})$ and $U_L(\boldsymbol{x}_{i_\eta})$ are the potential energies for the ith frame in bin η at high- and low-level Hamiltonians, respectively.

Generally speaking, TP calculations with larger reweighting entropy lead to more reliable results. As all the one-sided methods to estimate free energy differences, the reliability of the weighted TP calculations heavily depends on the overlap in phase space between the sampled (low-level) and target (high-level) Hamiltonians. So it is generally advisable to choose the low-level Hamiltonian that best resembles the high-level one.

In this work, we will start by using PM3 method [38] and the B3LYP/6-31G* model [39–42] as the low- and high-level Hamiltonians, respectively, where limited overlap in phase space will be found for some cases. The semi-empirical and *ab intio* single point QM/MM calculations were carried out by using the sqm program from the AmberTools16 package [43] and Q-Chem 4.4, [44] respectively.

The FE profiles after the TP correction are usually prone to larger noises than the underlying low-level ones, simply because overlapping between the neighboring windows in umbrella sampling simulations can be more easily achieved than that between the two Hamiltonians in TP calculations. Following our previous work, Gaussian process regression (GPR) [45] was utilized to effectively reduce the noise in the free energy profiles after the weighted TP

correction. The Gaussian process regression was performed using the scikit-learn package.

2.2 Recalibrating SE-QM/MM Models with Force Matching

To improve the phase space overlap between the SE-QM/MM and AI-QM/MM Hamiltonians, the semi-empirical parameters were recalibrated through the force matching technique. Following the Reaction Path Force Matching (RP-FM)[24] procedure, 4 to 6 reference configurations were randomly picked from the trajectory of each window of the PM3/MM umbrella sampling simulations, yielding 144 – 192 configurations for each of the four reactions in this study. Then the atomic forces in Cartesian coordinates were calculated for the reference configurations using the AI-QM/MM Hamiltonian, which would serve as the target data. Then the semi-empirical parameters were optimized by minimizing the least-squares error between the atomic forces (nuclear gradients of energy) calculated from the SE-QM/MM and AI-QM/MM Hamiltonians, i.e.,

$$\mathscr{X}^2 = \sum_i \left| \vec{G}_i^H - \vec{G}_i^L \right|^2, \tag{4}$$

where the index i runs over all the atomic forces in the target data. In this work, only the atomic forces on the QM atoms were used for force matching. The Trust Region Reflective algorithm[46] as implemented in the SciPy package was used to perform the least-squares minimization, where the parameters were bounded between $\pm 5\%$ of their original values in the standard PM3 method.

In most AI-QM/MM calculations today, the QM-MM electrostatics is usually undamped (although it has been suggested[47]). In contrast, within the AmberTools17/SQM implementation of the SE-QM/MM models, which is employed in this work, the electrostatics interactions between SE-QM and MM regions were always damped.[48] Hereafter, we will use PM3-M to refer to a PM3/MM model where the damping QM-MM electrostatics was weakened (by setting qmi_oneBDD1, qmi_oneBDD2, and qmi_oneBDD3 to 0 in qm2_calc_rij_and_eqns.F90). After force matching, the recalibrated models will be denoted as PM3* and PM3*-M, with QM-MM electrostatics modified in the latter. The original PM3 and calibrated PM3*-M parameters for the four reactions under study are listed in Tables S1, S2, S3, S4 of the ESI. As shown in Table 1, the root-mean-square deviation in the PM3/MM forces on QM atoms with respect to B3LYP/MM forces are reduced by 2 – 3 fold after the recalibration.

3 Computational Details

3.1 Reaction Systems

Four well-studied reactions will be modeled in this work. Two chemical reactions in aqueous solution from our previous work,[23] namely identity S_N2 reaction ($CH_3Cl + Cl^- \longrightarrow Cl^- + CH_3Cl$, Fig. 1a) and intramolecular proton transfer in glycine from the neutral form to the zwitterion form ($NH_2CH_2COOH \longrightarrow NH_3^+CH_2COO^-$, Fig. 1c), were revisited in the current work. We also studied Menshutkin reaction ($NH_3 + CH_3Cl \longrightarrow CH_3NH_3^+ + Cl^-$, Fig. 1b) in aqueous solution, and the Claisen rearrangement of chorismate to prephenate catalyzed by chorismate mutase (Fig. 1d).
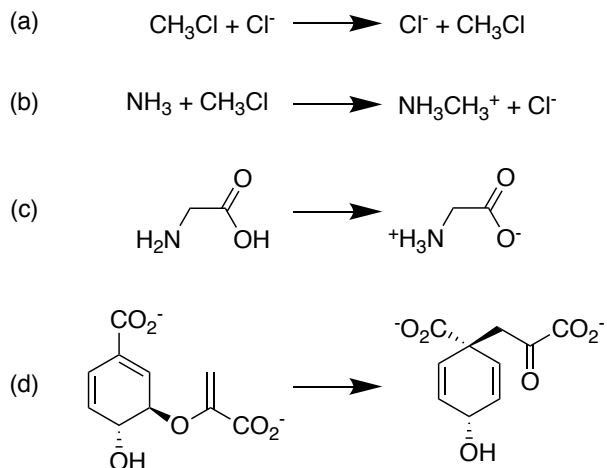
Fig. 1 Reaction schemes for (a) identity S_N2 reaction, (b) Menshutkin reaction, (c) glycine intramolecular proton transfer reaction, and (d) chorismate mutase reaction.

The mechanism of these reactions are already well understood, thanks to numerical experimental and computational investigations: identity S_N2,[25,27,31,49–55] Menshutkin,[26,27,56–60] glycine intramolecular proton transfer,[25,55,61–68] and chorismate mutase.[69–74] This work focuses on the development of efficient simulation protocols for generate AI-QM/MM-quality free energy profiles for these reactions, instead of the mechanisms.

3.2 Classical Simulations

For the aqueous systems, the solute was solvated by a TIP3P[75] water sphere of radius 25 Å, which is centered on the center atom (the heavy atom closest to the center-of-mass) of the solute. For the enzyme system, the initial structure was built on the X-ray crystal structure (PDB ID: 2CHT) of Bacillus subtilis chorismate mutase complexed with a transition state analog,[69] which was modified to the substrate chorismate manually, and solvated in a cubic box of TIP3P water with an initial size of $82\,\text{Å} \times 82\,\text{Å} \times 82\,\text{Å}$. Sodium counter ions were added to neutralize the system. The solutes and ligand were modeled with general Amber force field (GAFF),[76] whereas the enzyme was modeled using the AMBER ff14SB force field.[77]

Classical molecular dynamics simulations were first performed to relax the environments with the heavy atoms in the solutes and ligand restrained to their initial positions. For the aqueous systems, no cutoff was used to truncate the non-bonded interactions, and a harmonic wall potential with a force constant of $10\,\text{kcal}\,\text{mol}^{-1}\,\text{Å}^{-2}$ was applied to the water molecules to prevent them from moving too far from the center of the water sphere. For the enzyme system, periodic boundary conditions were employed, and the particle mesh Ewald (PME) method[78,79] was used to treat the electrostatic interactions, while the van der Waals interactions were truncated at a cutoff of 12 Å. For all the systems, the SHAKE algorithm[80] was used to constrain all bonds involving hydrogen atoms, and a time step of 2 fs was used for the MD integration. Langevin dynamics was performed at 300 K for 500 ps and 2 ns for the aqueous and enzyme systems, respectively.

Table 1 Root-Mean-Square Errors of Atomic Forces in Cartesian Coordinates $(\mathrm{kcal/mol/\mathring{A}})^{[a]}$

reaction	PM3	PM3-M	PM3*	PM3*-M
S_N2	10.53 ± 3.16	9.29 ± 2.27	3.66 ± 0.80	2.98 ± 0.65
MEN	15.17 ± 4.29	13.89 ± 3.27	4.99 ± 1.08	4.23 ± 1.05
PT	18.58 ± 2.14	19.34 ± 2.31	5.34 ± 1.00	5.23 ± 0.99
CM	16.11 ± 1.26	16.55 ± 1.25	7.04 ± 0.96	7.34 ± 1.07

[a] The SE-QM/MM forces were calculated from 144–192 configurations sampled along the SE-QM/MM reaction pathways.

3.3 QM/MM Simulations

The starting structures for the SE-QM/MM MD simulations were taken from the last frame of each of the classical simulation trajectories. During the SE-QM/MM MD simulations, the solutes and ligand were described by the PM3 semi-empirical method, while the rest of the system was still described by the TIP3P/ff14SB force fields used in the classical simulations. The SHAKE algorithm was only applied to the MM subsystem, and the integration time step was set to 1 fs.

For each of the aqueous systems, the center atom of the solute was restrained to its initial position to keep the solute staying at the center of the water sphere. For the enzyme system, the periodic system was truncated around the ligand where the residues and molecules whose atoms were all farther than 25 Å from the center atom of the ligand were removed from the system. The remaining residues whose atoms were all farther than 20 Å were restrained to their starting positions, and the same harmonic wall potential was applied to the remaining water molecules.

Umbrella samplings were used to estimate the free energy profile along a predefined reaction coordinate. For all the systems, the reaction coordinate was defined as $\eta = d_{\mathrm{break}} - d_{\mathrm{form}}$, where d_{break} and d_{form} were the bond lengths of the breaking and forming bonds, respectively. The centers of the harmonic biasing potentials were evenly distributed with an interval of 0.1 Å along the region of interest of the reaction coordinate, and a fixed force constant of the biasing potential was used for all the windows of each system. For the S_N2 reaction, 48 windows were used to cover η ranged from -2.35 to 2.35 Å, and the force constant was set to be $250 \, \mathrm{kcal \, mol^{-1} \mathring{A}^{-2}}$. For the Menshutkin reaction, 40 windows were used to cover η ranged from -1.95 to 1.95 Å, and the force constant was set to be $300 \, \mathrm{kcal \, mol^{-1} \mathring{A}^{-2}}$. For the glycine intramolecular proton transfer reaction, 24 windows were used to cover η ranged from -1.15 to 1.15 Å, and the force constant was set to be $250 \, \mathrm{kcal \, mol^{-1} \mathring{A}^{-2}}$. For the chorismate mutase system, 40 windows were used to cover η ranged from -1.95 to 1.95 Å, and the force constant was set to be $300 \, \mathrm{kcal \, mol^{-1} \mathring{A}^{-2}}$.

Hamiltonian replica exchange molecular dynamics (HREMD) was adopted to accelerate the convergence of the free energy simulation, and attempts of exchanging the biasing potentials between the neighboring windows were made every 100 fs. The systems were first equilibrated for 50 ps, and 100 ps production runs were conducted where the Cartesian coordinates were saved every 50 fs, which resulted in 2000 structures for subsequent MBAR and weighted TP analyses.

To get reference values for the free energy profiles at ab intio

level, we also performed AI-QM/MM umbrella sampling simulations using DFT (B3LYP/6-31G*) as the QM method, during which the systems were equilibrated for 50 ps for the aqueous system and 10 ps for the enzyme system, and the Cartesian coordinates were saved every 50 fs for the 100 ps and 20 ps production runs for the aqueous systems and the enzyme system, respectively. HREMD was not used for the AI-QM/MM umbrella sampling simulations.

Both of the classical and semi-empirical QM/MM MD simulations were performed by using the sander program [81] from the AmberTools17 package, whereas the *ab intio* QM/MM MD simulations were performed by using the sander/Q-Chem interface. [82]
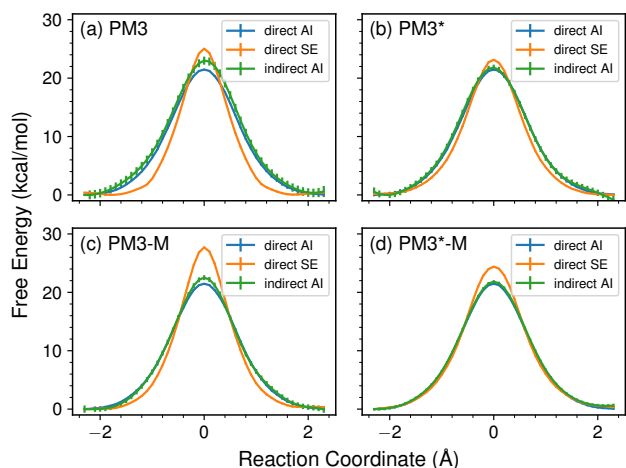
4 Results

4.1 Identity S_N2 Reaction: $CH_3Cl + Cl^- \longrightarrow Cl^- + CH_3Cl$

For the SE-QM/MM simulations with the standard PM3 parameters, the direct SE-QM/MM barriers are significantly higher than the direct AI-QM/MM barrier (Fig. 2a and c). With recalibrated PM3 parameters, the direct SE-QM/MM barriers are much closer to the direct AI-QM/MM barrier (Fig. 2b and d), but are still over 2 kcal/mol higher than the AI-QM/MM value. Comparing Fig. 2a, b and Fig. 2c, d, respectively, it can been seen that the direct SE-QM/MM barriers are increased when the damping of QM-MM electrostatics are reduced. This is likely due to the increased QM-MM electrostatic interactions stabilizing the reactant more than the transition state, which has a smaller dipole moment than the reactant. Overall, direct SE-QM/MM cannot fully recover the AI-QM/MM FE profile, even after a recalibration of the SE-QM parameters, so TP calculations are still needed to obtain AI-QM/MM-quality FE profiles.

After the SE-QM/MM FE profiles are corrected by using weighted TP, all of them are in better agreement with the direct AI-QM/MM FE profile. For the FE profile of the standard PM3 and the standard QM-MM electrostatics, the barrier is still over 1 kcal/mol higher than the direct AI-QM/MM one (Fig. 2a and Table 2), and the uncertainty is relatively large (0.7 kcal/mol, Table 2), which is in accordance with the relatively small reweighting entropies and the number of effective samples (Fig. S2a). When the recalibrated PM3 was used with the standard QM-MM electrostatics, the barrier difference is well within 1 kcal/mol. However, the uncertainty is still a bit high. Surprisingly, the indirect FE profile of the standard PM3 and the modified QM-MM electrostatics can already achieve good agreement with the direct AI FE profile, which suggests that part of the discrepancy between the AI-QM/MM and SE-QM/MM arises from the QM-MM electrostatics. The best result can be obtained by the recalibrated PM3 and the modified QM-MM

Table 2 Free Energy Barriers (in kcal/mol) for CH₃Cl + Cl⁻ $\longrightarrow$ Cl⁻ + CH₃Cl (S$_N$2), Menshutkin Reaction (MEN), Glycine Intramolecular Proton Transfer Reaction (PT), and Chorismate Mutase Reaction (CM)

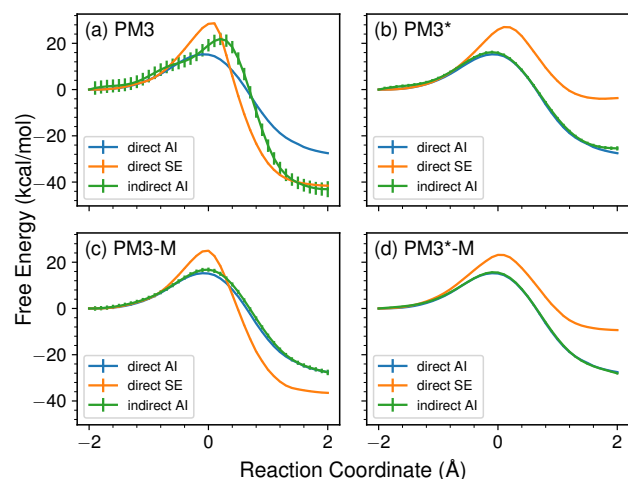| reaction	direct PM3	direct PM3*-M	B3LYP		
			indirect from PM3	indirect from PM3*-M	direct
S$_N$2	25.1 ± 0.1	24.4 ± 0.1	23.0 ± 0.7	21.8 ± 0.3	21.5 ± 0.1
MEN	28.7 ± 0.2	23.2 ± 0.1	21.9 ± 2.7	15.7 ± 0.3	15.3 ± 0.1
PT	23.1 ± 0.1	4.2 ± 0.1	3.3 ± 0.7	2.9 ± 0.3	2.9 ± 0.1
CM	33.7 ± 0.1	13.0 ± 0.1	18.2 ± 1.0	11.5 ± 0.7	12.1 ± 0.2

Fig. 2 Direct and indirect FE profiles for identity S$_N$2 reaction estimated by using SE-QM/MM Hamiltonians where (a) PM3, (b) PM3*, (c) PM3-M, and (d) PM3*-M are used as the QM method. The direct AI-QM/MM results are also shown for comparison. All the indirect FE profiles are smoothed by Gaussian process regression, and the unsmoothed FE profiles are presented in Fig. S1.
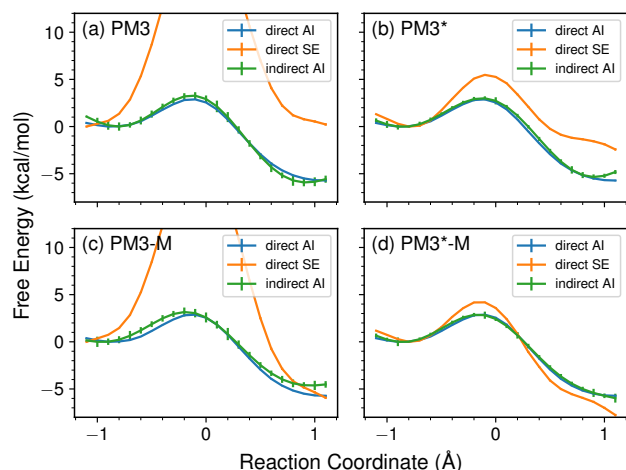

Fig. 3 Direct and indirect FE profiles for Menshutkin reaction estimated by using SE-QM/MM Hamiltonians where (a) PM3, (b) PM3*, (c) PM3-M, and (d) PM3*-M are used as the QM method. The direct AI-QM/MM results are also shown for comparison. All the indirect FE profiles are smoothed by Gaussian process regression, and the unsmoothed FE profiles are presented in Fig. S3.

electrostatics (Fig. 2d and Table 2), which reproduced the entire AI-QM/MM FE profile almost perfectly.

It should be noted that the estimated direct AI-QM/MM free energy barrier (21.5 kcal/mol) is 5 kcal/mol lower than the value of 26.5 kcal/mol determined from the experimental rate constant,[49] and the direct PM3 free energy barrier (25.1 kcal/mol) seems to have a better agreement with the experimental value. However, it does not mean PM3 is more reliable than B3LYP/6-31G* for this system, and the better agreement might come from a cancellation of errors. In general, to achieve good and reliable agreement with the experiments, more sophisticated functionals (such as M06-2X,[83] ωB97X-D,[84] and ωB97X-V[85]) and larger basis sets might be needed (see improved results in Table S5 with these functionals for both identity S$_N$2 and Menshutkin reactions). However, the objective of the current study is to reproduce target AI-QM/MM results with only SE-QM/MM sampling, and reproducing the experimental values is not a focus in this work.

4.2 Menshutkin Reaction: NH₃ + CH₃Cl $\longrightarrow$ CH₃NH₃⁺ + Cl⁻

Similar to the CH₃Cl + Cl⁻ $\longrightarrow$ Cl⁻ + CH₃Cl reaction, the direct SE-QM/MM simulations for the Menshutkin reaction with the standard PM3 parameters and the standard QM-MM electrostatics

(Fig. 3a and Table 2) also overestimated the free energy barrier comparing to the direct AI-QM/MM. On the other hand, the predicted reaction free energy was too low. By reducing the damping of the QM-MM electrostatics, both free energy barrier and reaction free energy were closer to the AI-QM/MM ones (Fig. 3c). Further, after recalibrating PM3 parameters, the barriers were slightly reduced for both cases (Fig. 3b and d). However, the reaction free energies became higher than the AI-QM/MM ones after switching to the recalibrated PM3 parameters.

After applying weighted TP corrections, the indirect FE profile of the standard PM3 and the standard QM-MM electrostatics failed to reproduce the AI-QM/MM one in this case, which is consistent with the low reweighting entropies and the small number of effective samples around the barrier region (Fig. S4a). Better agreement was observed for the indirect FE profile with the recalibrated PM3 model and the standard QM-MM electrostatics (Fig. 3b) and the indirect FE profile with the standard PM3 model and the modified QM-MM electrostatics (Fig. 3c), which is consistent with the reweighting entropies and the number of effective samples (Fig. S4b and c). Finally, the indirect FE profile of the recalibrated PM3 and the modified QM-MM electrostatics almost perfectly reproduced the direct AI-QM/MM one (Fig. 3d), even before GPR smoothing was applied (Fig. S3d).

4.3 Glycine Intramolecular Proton Transfer Reaction: NH$_2$CH$_2$COOH $\longrightarrow$ NH$_3^+$CH$_2$COO$^-$

Fig. 4 Direct and indirect FE profiles for glycine intramolecular estimated by using SE-QM/MM Hamiltonians where (a) PM3, (b) PM3*, (c) PM3-M, and (d) PM3*-M are used as the QM method. The direct AI-QM/MM results are also shown for comparison. All the indirect FE profiles are smoothed by Gaussian process regression, and the unsmoothed FE profiles are presented in Fig. S5.

The direct SE-QM/MM FE profile of the standard PM3 parameters and the standard QM-MM electrostatics (Fig. 4a) predicts a very high barrier and incorrect reaction free energy since the zwitterion form of glycine should be the dominant state in aqueous solution. Interestingly, by simply reducing the damping of the QM-MM electrostatics, the correct reaction free energy can be recovered (Fig. 4c) with direct SE-QM/MM, which suggests that the zwitterion and water interaction might be too weak in the standard QM-MM electrostatics of SE-QM/MM because of the damping. However, the free energy barrier is still too high and not fixed by reducing the damping. When the recalibrated PM3 was used with the standard QM-MM electrostatics (Fig. 4b), the barrier was largely reduced, though still higher than the direct AI-QM/MM barrier, and the relative free energy of the zwitterion form of glycine is still overestimated. When the modified QM-MM electrostatics was applied, the transition state and the product (zwitterion) were further stabilized in the direct recalibrated PM3 FE profile, which is already in a qualitative agreement with the direct AI-QM/MM one, even before the TP correction (Fig. 4d).

After applying TP corrections, the indirect FE profiles of the recalibrated PM3 are in good agreement with the direct AI-QM/MM one even before smoothing (Fig. S5b and d), whereas the ones of the standard PM3 show large fluctuations, which is consistent with the reweighting entropies and the number of effective samples (Fig. S6). After smoothing by Gaussian process regression, the predicted barriers of all the methods are well within 1 kcal/mol from the direct AI-QM/MM one, though the ones of the standard PM3 carry large uncertainties in the barriers. Similar to the S$_N$2 reaction, the best result came from the recalibrated PM3 and the modified QM-MM electrostatics, which predicted a nearly identical barrier (Table 2) as the direct AI-QM/MM and the whole FE profile can be overlay onto the direct AI-QM/MM one almost perfectly

(Fig. 4d).

4.4 Chorismate Mutase

Similar to the S$_N$2 reaction, the direct SE-QM/MM FE profile of the standard PM3 parameters and the standard QM-MM electrostatics (Fig. 5a) predicts a very high barrier but relatively accurate reaction free energy. In this case, reducing the damping of the QM-MM electrostatics does not change the overall direct SE-QM/MM FE profile much (Fig. 5c), which is expected since there is no large charge separation during the chorismate mutase reaction. After applying the TP correction, the barriers are much closer to the direct AI-QM/MM profile (Fig. 5a and c), however, the differences are still at 4–6 kcal/mol, and the fluctuations in the unsmoothed FE profile (Fig. S7a and c) are quite large, which is consistent with the relatively small reweighting entropies and the number of effective samples (Fig. S8a and c).
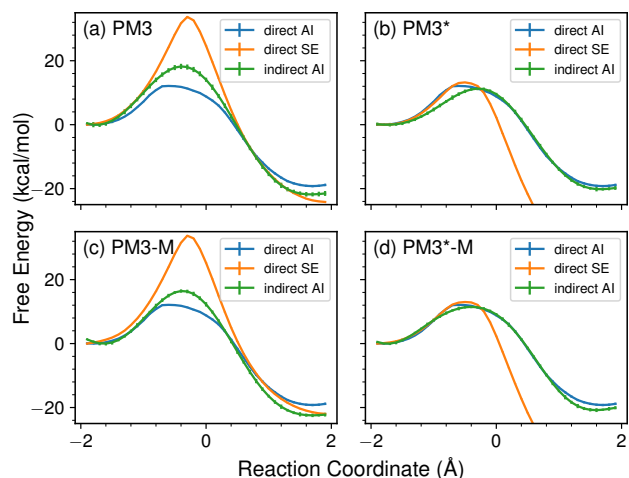
Fig. 5 Direct and indirect FE profiles for chorismate mutase reaction estimated by using SE-QM/MM Hamiltonians where (a) PM3, (b) PM3*, (c) PM3-M, and (d) PM3*-M are used as the QM method. The direct AI-QM/MM results are also shown for comparison. All the indirect FE profiles are smoothed by Gaussian process regression, and the unsmoothed FE profiles are presented in Fig. S7.

With the recalibrated PM3, the direct SE-QM/MM FE barriers are very close to the direct AI-QM/MM ones. Actually the part of the profiles on the reactant side are in very good agreement with each other, while the agreement of the profiles on the product side are much less satisfactory, which suggests that it is difficult for the direct SE-QM/MM calculations to reproduce the whole AI-QM/MM profile with one set of parameters within the search space of the current study. Fortunately, the AI-QM/MM profile can be recovered after the TP corrections were applied, despite of the large difference of free energy on the product side. It should be noted that the results of the chorismate mutase reaction have larger uncertainties compared with the other two reactions, so longer simulations might be desired to achieve more precise results. Also the deviation of the barrier seems larger than the other two reactions, even though still within 1 kcal/mol. However, it should be taken into consideration that the direct AI-QM/MM simulations for the chorismate mutase reaction have much shorter

Table 3 Estimated CPU Time (in 10^3 h) for the Computation of Free Energy Profiles for $CH_3Cl + Cl^- \longrightarrow Cl^- + CH_3Cl$ (S_N2), Menshutkin Reaction (MEN), Glycine Intramolecular Proton Transfer Reaction (PT), and Chorismate Mutase Reaction (CM)

			B3LYP		
reaction	direct PM3	direct PM3*	indirect from PM3	indirect from PM3*	direct[a]
S_N2	1.3	1.9	1.5	2.1	51.9
MEN	1.1	1.7	1.3	1.9	45.1
PT	0.6	1.0	0.8	1.2	26.0
CM	1.8	2.7	4.0	4.9	44.7

[a] For direct B3LYP simulations, the CPU time correspond to 100 ps simulation time per window for S_N2, MEN, and PT, and 20 ps simulation time per window for CM.

simulation time than the other reactions, so the precision of the barrier is not as high as the other reactions.

4.5 Computational Expense

The estimated CPU times for the computations of the direct and indirect QM/MM free energy profiles at the B3LYP/6-31G* level are listed in Table 3. For direct PM3*, the CPU time includes the time to generate the training structures using the standard PM3 model, compute AI-QM/MM forces for those structures, and calibrate the PM3 parameters with force matching, in addition to the cost to perform SE-QM/MM sampling and postprocessing. For indirect profiles, the CPU time includes the cost of the corresponding direct SE-QM/MM profiles and the cost to compute the single point energies at B3LYP/6-31G* for the saved configurations. Since the SE-QM/MM and AI-QM/MM calculations do not scale at the same rate with respect to the size of the QM subsystems, there is an optimal way to allocate the resources between the sampling and reweighting phases. In this study, we chose to sample the configurations at the same rate and use the same length of the simulation time for both AI-QM/MM and SE-QM/MM simulations, except for AI-QM/MM simulations for chorismate mutase reaction which has 1/5 of the simulation time as the other simulations. For the systems with smaller QM subsystems (6 atoms for S_N2, 9 atoms for MEN, and 10 atoms for PT), we can see roughly 25-fold accelerations. For CM, which has a larger QM subsystem (24 atoms), the efficiency enhancement would be about 45-fold, had the same length of AI-QM/MM simulations been performed.

5 Discussion

5.1 Recalibration of the SE-QM Method Improves the Configurational Sampling

To acquire accurate reaction free energy profiles with either single Hamiltonian or dual Hamiltonians, one needs (a) an adequate sampling of important configurations within each sampling window; and (b) accurate relative energies (and thus Boltzmann weights) among these configurations. These two factors are simultaneously ensured in direct AI-QM/MM simulations.

Through the use of TP, indirect "SE-QM/MM $\rightarrow$ AI-QM/MM" free energy simulations can only ensure accurate relative energy/weights among collected configurations. Therefore, the accuracy of such indirect simulations is entirely dependent on the sampling. As shown in Fig.6, PM3/MM simulations sample the

reactant/product regions for all four reactions reasonably well. Therefore, after weighted TP corrections, the indirect simulations well reproduced reaction free energies [Figs. 2a, 3a, 4a, 5a].

However, the PM3/MM simulation pathway deviated substantially from the target AI-QM/MM pathways around the transition state region. As a result, *indirect "PM3/MM $\rightarrow$ B3LYP/MM" simulations systematically overestimated the free energy barrier heights* by 1.5, 6,6, 0.4, and 6.1 kcal/mol, respectively, for the four reactions [Table 2 and Figs. 2a, 3a, 4a, 5a].
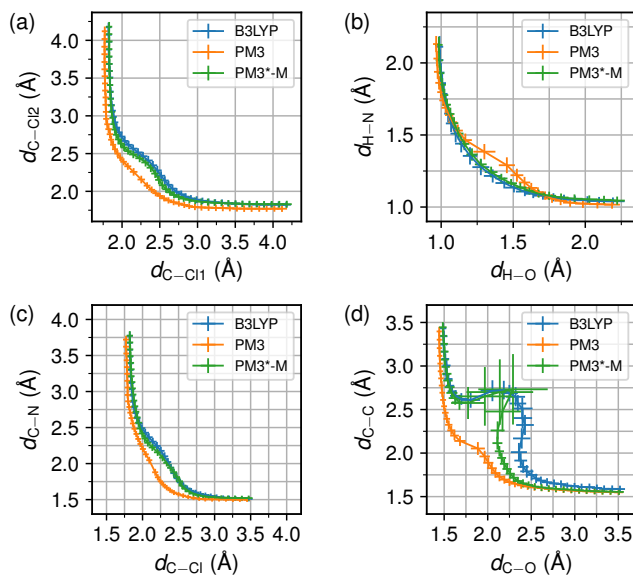
Fig. 6 Average lengths of the breaking and forming bonds in each window of the umbrella sampling simulations for (a) identity S_N2 reaction, (b) Menshutkin reaction, (c) glycine intramolecular proton transfer reaction, and (d) chorismate mutase reaction. Note that the pathways show only the sampling of the bond lengths, and do not correspond to the minimum free energy pathways.

Fig.6 clearly shows that a force-matching recalibration of the PM3 method substantially improved the sampling. As a matter of fact, it brought us nearly perfectly to AI-QM/MM sampling space for three aqueous reactions (S_N2, Menshutkin, and glycine proton transfer). For the chorismate mutase reaction, the PM3*-M/MM simulation also sampled configurations much closer to the AI-QM/MM pathway. However, for this reaction, it was much more challenging to reparameterize the PM3 method for the post-transition-state configurations.

5.2 Assessment of the Quality of Our Indirect Free Energy Simulations

Direct AI-QM/MM free energy barriers and pathways, which are used to assess the quality of our indirect "SE-QM/MM → AI-QM/MM" simulation of model reactions, will not be readily available in the study of other enzyme reactions. Therefore, here we will discuss a few other criteria to judge the quality of indirect free energy simulations.

First, a reliable TP-corrected free energy profile should be smooth enough without Gaussian process regression, which is the case for PM3*-M-based profiles in Figs. S1(d), S3(d), S5(d), and S7(d). Second, the reweighting entropy (defined in the method section) should be high enough. A threshold of 0.65 was recommended by some of us for solvation free energy calculations. Given the satisfactory PM3*-M-based indirect free energy profiles for three aqueous reactions, this threshold can probably be lowered to 0.2 for each bin along the reaction pathway (Figs. S2, S4, and S6). Caution should thus be taken for cases, such as the chorismate mutase reaction, where some reweighting entropy values are lower than 0.2.

Third, observing that Kish's effective sampling size [86,87] is not necessarily an integer, we have used an alternative definition, which corresponds to the smallest number of frames (within each bin) that brings us within 0.05 kcal/mol of the free energy correction (of that bin). As shown in Figs. S2, S4, S6 and S8, for some bins, a single frame can dominate the free energy correction to the standard PM3 model. For three aqueous reactions, the number of effective samples has increased to over 20 for all bins. On the other hand, even when a PM3*-M method is used, a single frame can still dominate the chorismate mutase reaction free energy correction in some bins.
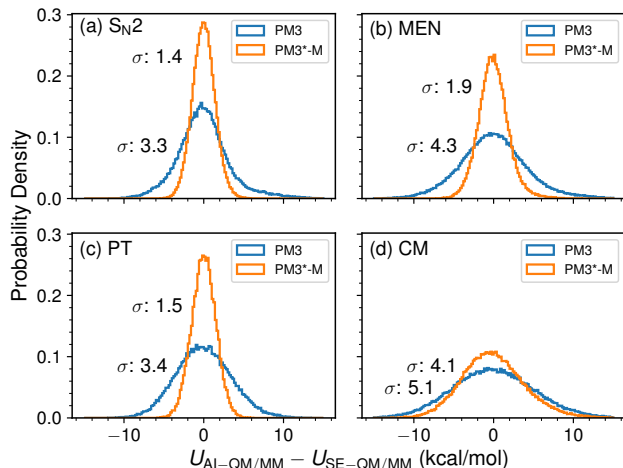
Fig. 7 Histograms of the relative energy differences between AI-QM/MM and SE-QM/MM Hamiltonians for (a) identity S_N2 reaction, (b) Menshutkin reaction, (c) glycine intramolecular proton transfer reaction, and (d) chorismate mutase reaction. The standard deviation (σ, in kcal/mol) is shown for each distribution of difference energies.

Lastly, several groups have suggested an analysis of the distribution of relative energy differences. [87–89] It was recommended that its standard deviation, σ, should be lower than 1.7 kcal/mol (~ 7 kJ/mol), [87] or 2.4 kcal/mol ($\sim 4\,k_BT$) [89] for reliable TP cal-

culations. The distribution of high-level and low-level energy differences for four systems in this study is shown in Fig. 7. The standard deviation is less than 2.4 kcal/mol for three aqueous reactions, whereas it is larger than the threshold for the chorismate mutase reaction. This is consistent with our observation based on all other criteria. Another criterion, Π, as developed by Wu and Kofke, [90] has yet to be analyzed for the reactions under study.

6 Conclusions

We proposed the recalibration of SE-QM models for a specific aqueous or enzyme reaction of interest, before performing an indirect AI-QM/MM free energy simulation of the reaction. For three aqueous reactions (S_N2, Menshutkin, and glycine proton transfer) and one enzyme reaction (Claisen rearrangement within chorismate mutase), we recalibrated the simple PM3 semi-empirical model for each reaction. The use of a recalibrated PM3 model in indirect simulations allowed us to

- shift SE-QM/MM simulation sampling closer to the AI-QM/MM reaction pathway,

- reproduce the free energy barriers of all four reactions within 1.0 kcal/mol,

- reduce the computer time by 20 – 45 fold as compared to direct AI-QM/MM simulations.

For future indirect AI-QM/MM reaction free energy simulations based on recalibrated SE-QM models, it was suggested to examine the smoothness of raw TP-corrected free energy profiles, compute reweighting entropies and the number of effective samples, and to analyze the distribution of relative energy differences to assess the quality of indirect free energy results.

The proposed simple strategy is expected to be valid for the modeling of many other enzyme systems, well beyond the four model reactions in this study. But, instead of using the difference between two bond distances as the reaction coordinate, finite-temperature string methods [9,91,92] will be needed to describe the reaction pathway.

There will certainly be difficult cases, where, even with a recalibration, a single SE-QM model is not sufficiently flexible to describe configurations along the entire react pathway. In those cases, more sophisticated strategies, such as accelerated dynamics, [33–35] sequential sampling, [31,32] and multiple-time-step approaches, [29,30] will be needed to accelerate direct AI-QM/MM simulations.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 A. Warshel, *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **32**, 425–43.

2 A. J. Mulholland, *Drug Discov. Today*, 2005, **10**, 1393–1402.

3 H. M. Senn and W. Thiel, *Curr. Op. Chem. Biol.*, 2007, **11**, 182–187.

4 H. Hu and W. Yang, *Ann. Rev. Phys. Chem.*, 2008, **59**, 573–601.

5 H. Lin and D. G. Truhlar, *Theor. Chem. Acc.*, 2007, **117**, 185–199.

6 P. Hu and Y. Zhang, *J. Am. Chem. Soc.*, 2006, **128**, 1272–1278.

7 L. Wang, X. Yu, P. Hu, S. Broyde and Y. Zhang, *J. Am. Chem. Soc.*, 2007, **129**, 4731–4737.

8 R. Wu, S. Wang, N. Zhou, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2010, **132**, 9471–9479.

9 E. Rosta, M. Nowotny, W. Yang and G. Hummer, *J. Am. Chem. Soc.*, 2011, **133**, 8934–41.

10 K.-Y. Wong, H. Gu, S. Zhang, J. A. Piccirilli, M. E. Harris and D. M. York, *Angew. Chem. Int. Ed.*, 2012, **51**, 647–51.

11 A. Ganguly, P. Thaplyal, E. Rosta, P. C. Bevilacqua and S. Hammes-Schiffer, *J. Am. Chem. Soc.*, 2014, **136**, 1483–96.

12 D. R. Stevens and S. Hammes-Schiffer, *J. Am. Chem. Soc.*, 2018, **140**, 8965–8969.

13 J. Gao, *J. Phys. Chem.*, 1992, **96**, 537–540.

14 R. P. Muller and A. Warshel, *J. Phys. Chem.*, 1995, **99**, 17516–17524.

15 J. Heimdal and U. Ryde, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12592–12604.

16 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1989, **63**, 1195–1198.

17 M. R. Shirts and J. D. Chodera, *J. Chem. Phys.*, 2008, **129**, 124105–10.

18 J. Kästner and W. Thiel, *J. Chem. Phys.*, 2005, **123**, 144104–5.

19 T.-S. Lee, B. K. Radak, A. Pabis and D. M. York, *J. Chem. Theory Comput.*, 2013, **9**, 153–164.

20 T.-S. Lee, B. K. Radak, M. Huang, K.-Y. Wong and D. M. York, *J. Chem. Theory Comput.*, 2014, **10**, 24–34.

21 R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420–1426.

22 I. Polyak, T. Benighaus, E. Boulanger and W. Thiel, *J. Chem. Phys.*, 2013, **139**, 064105.

23 P. Li, X. Jia, X. Pan, Y. Shao and Y. Mei, *J. Chem. Theory Comput.*, 2018, **14**, 5583–5596.

24 Y. Zhou and J. Pu, *J. Chem. Theory Comput.*, 2014, **10**, 3038–3054.

25 L. Shen, J. Wu and W. Yang, *J. Chem. Theory Comput.*, 2016, **12**, 4934–4946.

26 J. Wu, L. Shen, W. Yang, J. Wu, L. Shen and W. Yang, *J. Chem. Phys.*, 2017, **147**, 161732.

27 L. Shen and W. Yang, *J. Chem. Theory Comput.*, 2018, **14**, 1442–1455.

28 P. S. Hudson, H. L. Woodcock and S. Boresch, *J. Phys. Chem. Lett.*, 2015, **6**, 4850.

29 M. Tuckerman, B. J. Berne and G. J. Martyna, *J. Chem. Phys.*, 1992, **97**, 1990–2001.

30 K. Nam, *J. Chem. Theory Comput.*, 2014, **10**, 4175–4183.

31 H. Hu, Z. Lu, J. M. Parks, S. K. Burger and W. Yang, *J. Chem. Phys.*, 2008, **128**, 034105.

32 E. Rosta, M. Haranczyk, Z. T. Chu and A. Warshel, *J. Phys. Chem. B*, 2008, **112**, 5680–5692.

33 C. L. Ramirez, A. Zeida, G. E. Jara, A. E. Roitberg and M. A. Marti, *J. Chem. Theory Comput.*, 2014, **10**, 4609–4617.

34 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12562–12566.

35 A. Barducci, M. Bonomi and M. Parrinello, *WIRES Comput. Mol. Sci.*, 2011, **1**, 826–843.

36 G. Piccini and M. Parrinello, *arXiv:1904.02004 [physics]*, 2019.

37 M. Wang, P. Li, X. Jia, W. Liu, Y. Shao, W. Hu, J. Zheng, B. R. Brooks and Y. Mei, *J. Chem. Info. Model.*, 2017, **57**, 2476–2489.

38 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209–220.

39 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

40 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.

41 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

42 P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.

43 D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, I. Cheatham, T. E., T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, N. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York and P. A. Kollman, *AMBER 2016; University of California: San Francisco*, 2016.

44 Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits,

R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Mol. Phys.*, 2015, **113**, 184–215.

45 C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Mass., 3rd edn., 2008.

46 M. Branch, T. Coleman and Y. Li, *SIAM J. Sci. Comput.*, 1999, **21**, 1–23.

47 D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček and B. R. Brooks, *J. Chem. Phys.*, 2002, **117**, 10534–10547.

48 M. J. Field, P. A. Bash and M. Karplus, *J. Comput. Chem.*, 1990, **11**, 700–733.

49 W. J. Albery and M. M. Kreevoy, in *Adv. Phys. Org. Chem.*, Elsevier, 1978, vol. 16, pp. 87–157.

50 J. Chandrasekhar, S. F. Smith and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1984, **106**, 3049–3050.

51 J. Chandrasekhar, S. F. Smith and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1985, **107**, 154–163.

52 J. Chandrasekhar and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1985, **107**, 2974–2975.

53 G. Vayner, K. N. Houk, W. L. Jorgensen and J. I. Brauman, *J. Am. Chem. Soc.*, 2004, **126**, 9054–9058.

54 L. Song, W. Wu, P. C. Hiberty and S. Shaik, *Chem. Eur. J.*, 2006, **12**, 7458–7466.

55 Z. Lu and Y. Zhang, *J. Chem. Theory Comput.*, 2008, **4**, 1237–1248.

56 J. Gao, *J. Am. Chem. Soc.*, 1991, **113**, 7796–7797.

57 J. Gao and X. Xia, *J. Am. Chem. Soc.*, 1993, **115**, 9667–9675.

58 Y.-Y. Chuang, C. J. Cramer and D. G. Truhlar, *Int. J. Quantum Chem.*, 1998, **70**, 887–896.

59 H. Castejon and K. B. Wiberg, *J. Am. Chem. Soc.*, 1999, **121**, 2139–2146.

60 P. Su, F. Ying, W. Wu, P. C. Hiberty and S. Shaik, *ChemPhysChem*, 2007, **8**, 2603–2614.

61 G. Wada, E. Tamura, M. Okina and M. Nakamura, *Bull. Chem. Soc. Jpn.*, 1982, **55**, 3064–3067.

62 M. Slifkin and S. Ali, *J. Mol. Liq.*, 1984, **28**, 215–221.

63 N. Okuyama-Yoshida, M. Nagaoka and T. Yamabe, *J. Phys. Chem. A*, 1998, **102**, 285–292.

64 M. Nagaoka, N. Okuyama-Yoshida and T. Yamabe, *J. Phys. Chem. A*, 1998, **102**, 8202–8208.

65 I. Tuñón, E. Silla, C. Millot, M. T. C. Martins-Costa and M. F. Ruiz-López, *The Journal of Physical Chemistry A*, 1998, **102**, 8673–8678.

66 R. Karmacharya, D. Antoniou and S. D. Schwartz, *J. Phys. Chem. A*, 2001, **105**, 2563–2567.

67 K. Leung and S. B. Rempe, *J. Chem. Phys.*, 2005, **122**, 184506.

68 C. M. Aikens and M. S. Gordon, *J. Am. Chem. Soc.*, 2006, **128**, 12835–12850.

69 Y. M. Chook, H. Ke and W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 8600–8603.

70 Y. M. Chook, J. V. Gray, H. Ke and W. N. Lipscomb, *Journal of Molecular Biology*, 1994, **240**, 476–500.

71 P. Kast, M. Asif-Ullah and D. Hilvert, *Tetrahedron Lett.*, 1996, **37**, 2691–2694.

72 H. Lee Woodcock, M. Hodošček, P. Sherwood, Y. S. Lee, H. F. Schaefer III and B. R. Brooks, *Theor. Chem. Acc.*, 2003, **109**, 140–148.

73 F. Claeyssens, K. E. Ranaghan, N. Lawan, S. J. Macrae, F. R. Manby, J. N. Harvey and A. J. Mulholland, *Org. Biomol. Chem.*, 2011, **9**, 1578.

74 M. Sanchez-Martinez, M. Field and R. Crehuet, *J. Phys. Chem. B*, 2015, **119**, 1103–1113.

75 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

76 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

77 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

78 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

79 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.

80 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Chem. Phys.*, 1977, **23**, 327–341.

81 R. C. Walker, M. F. Crowley and D. A. Case, *J. Comput. Chem.*, 2008, **29**, 1019–1031.

82 A. W. Götz, M. A. Clark and R. C. Walker, *J. Comput. Chem.*, 2014, **35**, 95–108.

83 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.

84 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615.

85 N. Mardirossian and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2014, **16**, 9904.

86 L. Kish, *Survey Sampling*, Wiley, New York, 1995.

87 U. Ryde, *J. Chem. Theory Comput.*, 2017, **13**, 5745–5752.

88 E. R. Pinnick, C. E. Calderon, A. J. Rusnak and F. Wang, *Theor. Chem. Acc.*, 2012, **131**, 1146.

89 S. Boresch and H. L. Woodcock, *Mol. Phys.*, 2017, **115**, 1200–1213.

90 D. Wu and D. A. Kofke, *J. Chem. Phys.*, 2004, **121**, 8742–8747.

91 V. Ovchinnikov and M. Karplus, *J. Chem. Phys.*, 2014, **140**, 175103.

92 K. Zinovjev and I. Tuñón, *J. Phys. Chem. A*, 2017, **121**, 9764–9772.