

Seeding the Multi-dimensional Nonequilibrium Pulling for Hamiltonian Variation: Indirect Nonequilibrium Free Energy Simulations at QM levels

Zhaoxi Sun^{1*} and Qiaole He²

¹*Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Institute of Theoretical and Computational Chemistry, Peking University, Beijing 100871, China*

²*AI Department of Enzymaster (Ningbo) Bio-Engineering Co., Ltd., North Century Avenue 333, 315100 Ningbo, China*

*To whom correspondence should be addressed: z.sun@pku.edu.cn

Abstract

The combination of free energy simulations in the alchemical and configurational spaces provides a feasible route to access the thermodynamic profiles under a computationally demanding target Hamiltonian. Normally, due to the significant differences between the computational cost of ab initio quantum mechanics (QM) calculations and those of semi-empirical quantum mechanics (SQM) and molecular mechanics (MM), this indirect method could be used to obtain the QM thermodynamics by combining the SQM or MM results and the SQM-to-QM or MM-to-QM corrections. In our previous works, a multi-dimensional nonequilibrium pulling framework for Hamiltonian variations has been introduced based on bidirectional pulling and bidirectional reweighting. The method performs nonequilibrium free energy simulations in the configurational space to obtain the thermodynamic profile along the conformational change pathway under a selected computationally efficient Hamiltonian, and uses the nonequilibrium alchemical method to correct or perturb the thermodynamic profile to that under the target Hamiltonian. The BAR-based method is designed to achieve the best generality and transferability and thus leads to modest (~20 folds) speedup. In this work, we explore the possibility of further accelerating the nonequilibrium free energy simulation by employing unidirectional pulling and using the selection criterion to obtain the initial configurations used to initiate nonequilibrium trajectories following the idea of adaptive steered molecular dynamics (ASMD). A single initial condition is used to seed the whole multi-dimensional nonequilibrium free energy simulation and the sampling is performed fully in the nonequilibrium ensemble. Introducing very short ps-length equilibrium

sampling to grab more initial seeds could also be helpful. The ASMD scheme estimates the free energy difference with the unidirectional exponential average (EXP), but it does not follow exactly the requirements of the EXP estimator. Another deficiency of the seeding simulation is the inherently sequential or serial pulling due to the inter-segment dependency, which triggers some problems in the parallelizability of the simulation. Numerical tests are performed to grasp some insights and guidelines for using this selection-criterion-based ASMD scheme. The presented selection-criterion-based multi-dimensional ASMD scheme follows the same perturbation network of the BAR-based method, and thus could be used in various Hamiltonian-variation cases.

Keywords: Steered Molecular Dynamics, Multi-dimensional Nonequilibrium Pulling, Seeding Simulations, Exponential Average, Gaussian Approximation

1. Introduction

All-atom molecular dynamics (MD) simulation is now a feasible tool to access the atomistic motions in complex systems.¹⁻³ The statistically meaningful estimates of observables require extensive sampling of the phase space, and the sampling time depends on the intrinsic properties of the system, the description or Hamiltonian, and the sampling strategy. Probability distributions are often satisfactorily transformed to the free energy profiles or landscapes. The differences between the free energies of different states depict their relative stabilities.

Enhanced sampling simulations provide a computationally feasible route to obtain converged thermodynamics in complex systems.⁴⁻⁹ They sample the system on modified energy landscapes to enhance the sampling efficiency and explore the phase space effectively.¹⁰⁻¹³ There are various types of enhanced sampling techniques. For instance, the traditional umbrella sampling¹⁴⁻¹⁶ adds a series of harmonic biasing potentials along the configurational collective variable (CV) to enhance the sampling efficiency in specific regions of phase space. The replica exchange method¹⁷⁻²² designs a series of systems similar to the interested one but with higher flexibility in some degrees of freedom and attempts to exchange configurations in different equilibrium ensembles to enhance the barrier crossing and conformational search. The nonequilibrium steered MD (SMD) method shares similar features of the equilibrium umbrella sampling but adds time-independent biasing potential to drive the system from one state to another. The SMD approach is less frequently used, but is observed to be promising in various case studies.²³⁻³⁰ As the energy landscape is modified, proper post-processing methods are required to recover the statistics in the original unperturbed ensemble. Theoretically rigorous reweighting estimators in free energy simulations could be obtained based on free energy perturbation (FEP).³¹ The estimator itself has many problems. For instance, the sample-size hysteresis problem^{10, 32, 33} introduces significant biases into the finite-sample estimates, and the statistical error is heavily underestimated.^{34, 35} These statistical problems could be avoided to some extent by combining the perturbations from multiple directions. In the two-state case, the statistically optimal estimator named Bennett Acceptance Ratio (BAR)^{36, 37} uses the Fermi weighting function to process the data points, while in the multi-state situation the generalization of BAR named MBAR^{38, 39} is of the highest statistical efficiency. The results obtained from perturbation-based reweighting estimators rely on the magnitude of phase space overlap between different states, and those of the neighboring states are often higher than the non-neighboring cases. As a result, the BAR estimates are often identical to the MBAR ones. The nonequilibrium generalization of the FEP derivatives replaces the energy difference with the nonequilibrium work (NEW) accumulated during nonequilibrium SMD simulations, and shares the same

statistical problems as the equilibrium perturbation schemes. The Jarzynski's Identity (JI)⁴⁰ corresponds to the exponential average (EXP) or FEP estimator, and the Crooks' Equation (CE)⁴¹ is the nonequilibrium scenario of BAR. The equilibrium and nonequilibrium perturbation-based schemes are observed to achieve similar efficiency and accuracy in the construction of the potential of mean force (PMF) in various cases.⁴²⁻⁴⁶

Complex processes often involve the rearrangements of multiple regions of the system, and it is difficult to find one or several proper CVs to describe these motions. Defining an optimal set of CVs requires a deep understanding of the dynamics of the system under investigation, which is very hard even for experienced researchers. For instance, the binding/unbinding event of protein-ligand complexes may involve significant conformational changes of the protein, which is hard to be captured with several CVs. The binding pathway may involve multiple ligand-residue interactions, some of which may be necessary to be included in the definition of the CV set. Considering the complexity of these problems, some alternative ways to obtain the variation of thermodynamics of the process could be preferred. The alchemical method only considers the differences between thermodynamics at physical end states. It avoids the direct simulation along the physical transformation pathway by defining an artificial alchemical order parameter.⁴⁷⁻⁵³ The free energy profile along this non-physical pathway is constructed and the overall free energy difference is obtained. As the alchemical method relies on the construction of a thermodynamic cycle to determine the relative free energy of different states,^{54,55} it could be viewed as an indirect regime to obtain the free energy difference. The alchemical method could also be used to perturb the description of the system.⁵⁶⁻⁵⁸ For instance, the ab initio quantum mechanics (QM) results could be obtained by employing the alchemical method to perturb the semi-empirical QM (SQM) results.³⁴ The SQM results could also be obtained in a similar way by perturbing the thermodynamics obtained under some molecular mechanics (MM) force fields.⁵⁹ Due to the significant differences between the computational costs of different Hamiltonians, performing such perturbation could render some speedups compared with direct simulations under the computationally demanding Hamiltonians. In the case that the energetics under different Hamiltonians show significant differences, the staging technique should be used to improve the convergence behavior. However, equilibrium sampling in the intermediate states is computationally demanding, which actually degrades the efficiency of the indirect method. The situation is further aggravated when the configurational sampling in the intermediate state is difficult to converge.^{55, 60-62} In this case, the nonequilibrium approach could be useful, as the intermediate-state sampling is avoided to some extent.

The combination of the enhanced sampling simulations in the configurational and alchemical spaces provides the multi-dimensional picture of the thermodynamic landscapes.⁶³⁻⁷¹ The PMF along the

configurational CV at one alchemical state depicts the variations of thermodynamics in the process under that alchemical description/Hamiltonian, while the PMF along the alchemical CV at one configurational state shows the variations of some observables under different alchemical Hamiltonians or descriptions. Our previous works have provided a multi-dimensional nonequilibrium pulling framework for Hamiltonian variations, which features the stratification strategy, bidirectional pulling and bidirectional reweighting.^{34, 59, 72} The BAR-based scheme obtains the thermodynamics along the configurational CV under some computationally efficient Hamiltonians such as SQM Hamiltonians, and perturbs the results to some ab initio levels with the alchemical method. Bidirectional pulling is used along both of these two perturbation pathways. However, as has been pointed out in the discussions of the previous works, the convergence of the simulation could be easily achieved in some cases,^{34, 59, 72} and unidirectional pulling could be sufficient. Therefore, we explore the possibility of unidirectional pulling in the current work. A computationally efficient unidirectional pulling scheme in the SMD regime is the adaptive steered MD (ASMD) method, which relies on the stratification strategy and uses some selection criteria to obtain the initial configuration for the next nonequilibrium pulling segment from the configurations of the previous pulling segment.⁷³⁻⁷⁸ In this way, the initial configurational sampling in the SMD simulations is avoided and the efficiency is consequently improved. As only one or several configurations are used to spawn the whole multi-dimensional nonequilibrium pulling simulations, the method could be considered as a seeding SMD approach. In the following parts, we would construct the multi-dimensional pulling framework in the ASMD regime and provide extensive tests on the perturbation parameters, aiming at providing some guidelines for using the method.

2. Methodology

The theoretical framework in the current work focuses on the variation of Hamiltonians or descriptions of the system. Each Hamiltonian defines a unique microscopic state of the simulated system, and we use the Hamiltonian H to describe the status of the system. The differences between the Hamiltonians could be, for instance, the details of the multi-scale treatment (e.g., the QM theory, the basis set, and the QM region). The microstate (k_1, k_2) is described with the state-specified Hamiltonian H_{k_1, k_2} , the two dimensions of which describe the conformational rearrangement and the description/model change, respectively. Consider the case that the thermodynamics under a target Hamiltonian (\dots, K_2) is pursued. We explore the configurational space with $k_2 = 1$, and then perturb the results to another Hamiltonian $k_2 = K_2$. An illustration of the

thermodynamic cycle is depicted in Fig. 1a. Unidirectional pulling and unidirectional reweighting are performed in all perturbation parts, and only transformations drawn with solid arrows are performed. The reason that the thermodynamics along the configurational CV under the target Hamiltonian H_{\dots,k_2} could be efficiently obtained by the indirect method is that the conformational sampling could be much more time-consuming than the alchemical transformation. Rearrangements of many atoms or groups are involved in the process of interest.

During each nonequilibrium transformation in the configurational space, the time-dependent harmonic potential V given below is used to drive the system from one state to another,⁷⁹

$$V(\mathbf{q}) = \frac{k}{2} (\xi(\mathbf{q}) - \xi_0(t))^2 \quad (1).$$

Here, k is the force constant, \mathbf{q} is the coordinate vector, $\xi_0(t)$ denotes the time-dependent protocol for CV variation defining the configurational transformation, and ξ refers to the current value of the CV. A large force constant could be used to suppress the fluctuations of the CV and keep the value of the CV very close to the predefined driving protocol, namely achieving the stiff spring limit.^{59, 78, 80-85} A small time step should be used to avoid unstable dynamics and the resulting perturbations of the distributions.^{34, 86-89} The change of the alchemical order parameter follows exactly the predefined protocol.

Although the whole process could be simulated in one long pulling simulation, the dissipation is large and the waiting time before useful feedback is long. Therefore, to achieve a better numerical behavior and obtain faster user feedback, the whole pulling excursion could be divided into a series of shorter segments. As each of them could be finished in a much shorter simulation time and could be simulated independently, the parallelizability is improved and the output could be accessed much faster. Such pleasingly parallel computation also avoids the slowdown triggered by communications overhead, thus maximizing the performance. As the sampling in the configurational space is harder than that in the alchemical space, we only stratify the configurational sampling into K conformational states and $K-1$ segments. Note that for periodic configurational CV, the number of the conformational states is the same as that of the segments. As the convergence of the nonequilibrium pulling simulations depends heavily on the phase space overlap between different states and the magnitude of the perturbations during the nonequilibrium pulling process, the pulling simulations are only performed between neighboring states.

In our previous BAR-based pulling framework, we aim at achieving the best generality and transferability. Thus, the bidirectional statistically optimal estimator of BAR is employed. It achieves faster convergence compared with unidirectional EXP, and the upper bound of the statistical error is larger for

BAR than EXP.^{34, 35} However, the BAR-based scheme involves the pulling process from the target Hamiltonian $H_{\dots K_2}$ to the selected Hamiltonian for configurational sampling $H_{\dots 1}$, which might slightly increase the computational costs. In cases that the $H_{\dots 1}$ -to- $H_{\dots K_2}$ perturbation is not difficult to converge, some accelerations could be achieved by only performing unidirectional pulling from $H_{\dots 1}$ to $H_{\dots K_2}$, as depicted in Fig. 1a. In this case, the EXP or JI estimator estimates the free energy difference by exponentially averaging the microscopic nonequilibrium works accumulated during nonequilibrium pulling simulations initiated from the i th state to the j th state,^{40, 90} namely

$$\Delta A_{ij} = -\ln \left\langle e^{-W_{ij}} \right\rangle_i \quad (2).$$

Here, ΔA denotes the dimensionless free energy difference, W_{ij} represents the dimensionless work accumulated during the nonequilibrium pulling initiated from state i and ended in state j , and $\langle \dots \rangle_i$ represents the canonical average over nonequilibrium realizations initiated from state i . As the nonequilibrium transformations are performed between neighboring states, we have $j=i+1$ here. The estimator is valid for any pulling speeds, and such calculation is historically called the fast growth simulation. An alternative estimator applicable for extremely slow pulling speeds (i.e., reversible pulling) is the slow growth method.⁹¹⁻⁹³ In this case, the ordinary average of the microscopic nonequilibrium works is used to estimate the free energy difference, namely

$$\Delta A_{ij} = W_a = \left\langle W_{ij} \right\rangle_i \quad (3).$$

The slow growth method is a variant of the equilibrium integration method thermodynamic integration.⁹¹ The equation is valid for reversible pulling. However, practical simulations are all of finite lengths. Thus, the estimates are intrinsically biased.

Although the fast growth method is theoretically rigorous, as it is based on exponential averaging, it suffers from the same numerical problems of EXP, e.g., the sample size hysteresis and the underestimation of the statistical error. Cumulant expansion provides a way to better the numerical estimates from finite-length simulations. The expanded components are on the exponent and the free energy difference could be expressed as a linear combination of cumulants.⁹⁴⁻⁹⁶ The distribution of the nonequilibrium works in near-equilibrium pulling is close to Gaussian according to the central limit theory. For normally distributed data, only the first two terms in cumulant expansion survive, as higher-order cumulants are all zeros. Therefore, the Gaussian approximated EXP estimate GEXP could be expressed as

$$\Delta A_{ij} = \mu_{W_{ij}} - \frac{\sigma_{W_{ij}}^2}{2} \quad (4).$$

Here, the mean of the works is μ , and σ is the standard deviation of the nonequilibrium works. The 1st term is exactly the same as the slow growth estimate, and the 2nd term is an estimate of the dissipation of the pulling process, which corrects the bias of the slow growth result.

The statistical errors of the above three estimators could be obtained with the normal error propagation procedure. Note that for the EXP estimator, there is an upper bound for the analytical statistical error, which is found to be the thermal energy $k_B T$ in our previous work.³⁵ The other two estimators of W_a and GEXP do not have this behavior.

The inputs of the above estimators and the corresponding statistical errors should be statistically independent. Thus, to initiate nonequilibrium pulling trajectories, equilibrium sampling in the initial state is required. The fluctuation of some observables could be used to define uncorrelated configurations. For instance, the autocorrelation of the reaction coordinate used to describe the conformational change was used to define uncorrelated initial configurations in dihedral flipping in peptides and nucleotide systems.^{11, 78, 85, 97} Another example is the derivative of the alchemical Hamiltonian, which has been widely used in many alchemical free energy calculations.⁹⁸⁻¹⁰³ To obtain uncorrelated configurations, the autocorrelation time of the selected observable in each state τ_i is calculated and the whole dataset is subsampled by the statistical inefficiency $\phi_i = 1 + 2\tau_i$. The statistical inefficiency $\phi_{\text{eq},i}$ is an estimate of the computational cost of each independent sample from equilibrium simulations, and the length of pulling simulations $\phi_{\text{NEW},i}$ should be added to define the sampling time required for an independent sample in nonequilibrium pulling simulations, namely

$$\phi_i = \phi_{\text{NEW},i} + \phi_{\text{eq},i} \quad (5).$$

The pulling time $\phi_{\text{NEW},i}$ for unidirectional pulling is the same as the pulling time in each segment, while that for bidirectional pulling should be multiplied by a factor of 2.

By using the above estimators to determine the free energy difference between all neighboring states, the variation of the free energy in the whole process could be obtained, namely

$$\Delta A_{1k} = \sum_{i=1}^{k-1} \Delta A_{i,i+1} \quad (6).$$

A reference state is often selected in the representation of the relative free energies. Here, the 1st state is

selected and its free energy is set to zero. Then, the relative free energy of each state could be written as,

$$A_k = \Delta A_{1k} = \sum_{i=1}^{k-1} \Delta A_{i,i+1} \quad (7).$$

The procedure could be applied to both the alchemical and the configurational CVs. Combining the free energy profiles along these two CVs, the two-dimensional free energy surface depicting the variation of the thermodynamic profile in the configurational and alchemical spaces is obtained. The subscripts are altered in the multi-dimensional case. There are K_1 states and the state is numbered by k_1 for the configurational CV, while there are K_2 states and the state is numbered by k_2 for the alchemical CV. The reference state A_{11} has a free energy of zero, and the relative free energy of the state (k_1, k_2) is expressed as,

$$A_{k_1 k_2} = \Delta A_{k_1 1, 11} + \Delta A_{k_1 k_2, k_1 1} = \sum_{i=1}^{k_1-1} \Delta A_{i,1,i+1,1} + \sum_{j=1}^{k_2-1} \Delta A_{k_1, j, k_1, j+1} \quad (8).$$

Here, $\Delta A_{k_1 1, 11}$ is the free energy difference between the reference configurational state 1 and the configurational state k_1 at the 1st alchemical state, and $\Delta A_{k_1 k_2, k_1 1}$ represents the perturbation term to change the alchemical state from 1 to k_2 at the k_1 th configurational state. The statistical fluctuation of the free energy estimates often makes the PMF noisy. Thus, the free energy profiles are smoothed with some curve fitting methods. Here, we use the Savitzky-Golay filter to increase the signal-to-noise ratio.

The unidirectional perturbation framework presented above is a simple alternative to our previous BAR-based scheme. The free energy estimates could be obtained by performing initial configurational sampling in each Hamiltonian state, pulling the system between neighboring states, and using the EXP, GEXP or W_a estimator to get the free energy difference. However, we expect to further accelerate the simulations by introducing the selection criterion for initial configurations. The idea of the selection criterion arises from the fact that the ensemble of nonequilibrium configurations should include equilibrium snapshots.^{73, 74, 77} Therefore, the last configurations of the i -to- $(i+1)$ pulling simulations could possibly include the equilibrium structure(s) in the $(i+1)$ th state. For small perturbations (i.e., slow pulling speeds), the existence of such equilibrium configuration(s) is expected to be extremely possible. Therefore, the initial configuration for the next stage (i.e., the $(i+1)$ -to- $(i+2)$ segment) could be obtained by selecting some of the last configurations of the i -to- $(i+1)$ segment. Here, we employ the naïve Jarzynski's scheme, where the configuration with the work closest to the EXP estimate is selected.⁷³⁻⁷⁸ An illustration of the ASMD selection criterion is presented at the center of Fig. 1b. The configuration with the microscopic work closest to the EXP estimate is

considered as an independent sample extracted from the equilibrium ensemble. The ASMD simulation with the JI criterion is used to construct the free energy profile along the configurational CV, and the obtained initial configuration in each stage is also used to initiate the alchemical transformation. In this way, the need for equilibrium sampling in each intermediate is totally eliminated, thus accelerating the staged SMD simulations.

Although the selection criterion is useful to speed up the simulation, it has some drawbacks. First, the whole nonequilibrium pulling simulations are initiated from a single configuration, which leads to some correlations of the trajectories. As a result, the nonequilibrium works do not follow the requirement of the free energy estimators. There could be some biases in the free energy estimates, and the estimate of the corresponding statistical uncertainty is not theoretically rigorous. Whether this initial-configuration-induced bias would be significant requires some numerical tests. Second, although the whole pulling process is divided into a series of shorter segments, as the simulation starts from one segment (e.g., the reference state), all later segments need to wait for the output of the previous segments. Such inter-segment dependency leads to the poor parallelizability of ASMD simulations. Parallelization is possible only for the trajectories in the current segment and the computation of intra- and inter-molecular interactions in each simulation box. The former leads to N_{traj} independent jobs for each segment, which could be run in parallel delightfully. However, the later parallelization suffers from parallel slowdown, and its efficiency depends on the size of the system, the details of the distributed computation and so on. Although the two parallelization schemes could be used simultaneously, the efficiency of the parallel computing is not as good as the previous BAR-based scheme, where all segments are independent and there are $N_{\text{segments}} * N_{\text{traj}}$ pulling simulations that could be run in parallel pleasingly. The limited parallelizability of ASMD simulations slows down the results output, which is further aggravated by the poor statistical behavior of the EXP estimator. The unidirectional EXP estimator often requires longer pulling times and larger sample sizes than the bidirectional BAR. Thus, a much longer waiting time is needed before user feedback, which is not a satisfactory behavior for protocol tests. However, the situation could be bettered to some extent if there are a series of similar systems (e.g., mutants of the same protein-ligand complex) under investigation. These similar systems could be simulated with the same protocol independently.

Convergence check is indispensable in free energy simulations. It ensures the reproducibility and the reliability of the simulation outputs, and thus should be tested in the first place. In the normal cases, we often monitor the sample-size and pulling-speed dependence of various ensemble averages.^{34, 59, 78, 100, 104, 105} As

the free energy is the generating function for other thermodynamic properties, we often focus on the free energy difference and the corresponding statistical uncertainty.^{35, 100, 104, 105} Convergence is achieved when the free energy profile does not change with larger sample sizes or slower pulling speeds. The convergence diagnostics in the current ASMD case is a bit different. As the ASMD simulations along the configurational CV uses a predefined number of trajectories N_{traj} , we need to rerun the simulations with different N_{traj} and check whether the results are invariant of this value. As the alchemical transformation is initiated from the configurations obtained during the ASMD simulations along the configurational CV and is finished in a single stage, we only need to test whether the number of trajectories is sufficient to converge the alchemical perturbation. The test of the pulling speed in ASMD simulations follows the same procedure as the normal staged SMD simulations. The same simulation procedure is repeated with different pulling speeds, and the convergence is reached when the free energy profile does not change with slower pulling speeds.

A final note we would add on the perturbation network is that the multi-dimensional nonequilibrium framework for Hamiltonian variation is not necessarily limited to the configurational-alchemical CV case. The configurational CV could be altered to some alchemical CVs. For instance, in relative or absolute binding free energy calculation, the mutation or annihilation/creation of the ligand(s) is described with an alchemical CV. The configurational CV in the above perturbation network could be replaced by this alchemical CV and the indirect free energy calculation could be used to obtain the higher-level free energy variation along that alchemical CV. Namely, the multi-dimensional nonequilibrium pulling is performed in the alchemical-alchemical case, with the first alchemical CV describing the changes of the ligand and the second alchemical CV specifying the level of theory. As the free energy profile along the alchemical CV is generally not useful, performing only end-state corrections could be more efficient. Then, the nonequilibrium transformation network is composed of two parts. The first one is the ASMD simulation along the alchemical CV describing the ligand change, and the second part is the ASMD simulation to perturb the level of theory.

3. Computational Details

The Hamiltonian perturbation framework could be used to change various details of the system. In the current numerical test, we focus on the change of the QM theory. We still use the dihedral flipping system formed by caps of biomolecules N-methylacetamide (NMA) formed by ACE-NME (caps of protein) in vacuo as an example. The reaction coordinate used to describe the dihedral flipping is the backbone C-C-N-C dihedral. We explore the configurational space with some computationally feasible Model A (e.g., SQM

levels), and then perform alchemical transformations to obtain the results under another Hamiltonian Model B, e.g., ab initio QM or still some SQM Hamiltonians. An illustration of the thermodynamic cycle for the indirect free energy simulation is presented in Fig. 1b. In our previous works, we have considered SQM Hamiltonians of Austin Model 1 (AM1), the Parametrized Model number 6 (PM6) and the Modified Neglect of Diatomic Overlap (MNDO), and ab initio QM ones such as the Hartree-Fock method (HF), the second-order perturbation theory (MP2), the Becke 3-parameter Lee–Yang–Parr (B3LYP) functional, and the ω B97X-D functional. This time we consider another SQM Hamiltonian with good accuracy but is less frequently used. Recife Model 1 (RM1)¹⁰⁶ is an improved version of AM1. It provides better descriptions of various observables and thus is considered as a target SQM Hamiltonian in the current work. The HF method is still used as an ab initio target in this work. Therefore, in this work, we aim at obtaining the RM1 and HF results indirectly with the selection-criterion-based multi-dimensional nonequilibrium pulling simulations. The above indirect free energy calculations at (S)QM levels are relatively easy to converge.³⁴ However, when MM Hamiltonians are included in the indirect simulations, e.g., indirect SQM results from MM simulations and MM-to-SQM perturbations, the convergence behavior is not good. Thus, we also consider the challenging case that the indirect RM1 free energy simulation of the same NMA system is performed by combining direct free energy simulations with AMBER14SB¹⁰⁷ and the AMBER14SB-to-RM1 nonequilibrium perturbations.

The nonequilibrium trajectories are simulated independently and a single core is distributed to each trajectory. This avoids the parallel slowdown that degrades the performance of the simulation. In Table S1, we present the single-core timing data. The costs of the SQM simulations are extremely similar, and thus only one value is reported for all SQM Hamiltonians. The HF simulation is much more costly than the SQM ones, and the MM simulation is much faster than the SQM ones.

As a single structure is used to spawn the whole ASMD pulling process, we test the influence of using different configurations as the initial seed. The ASMD pulling starts at the 0° state, which is used as the reference state with a free energy of zero. The system is constructed in vacuo and equilibrated for 50 ps under the AMBER14SB¹⁰⁷ force field. Then, we shift to the computationally efficient Hamiltonian (e.g., AMBER14SB or SQM levels) and equilibrate the system for another 50 ps. Finally, we perform the production run to extract independent configurations to investigate the influence of the initial seed, and extract 8 configurations every 90 ps to avoid any possible correlations between the initial seeds (the autocorrelation time of this system is sub-ps).

The whole dihedral flipping process from 0° to 360 ° is divided into 120 segments with 3 ° increments.

The force constant of 2000 kcal/mol·rad² is used to achieve the stiff spring limit. As for the pulling speed and the sample size for each segment, some related statistics of ASMD simulations could be obtained from our previous works.⁷⁸ Our previous experience on the dihedral flipping case indicates that 20-sample and 40-sample estimates are identical in the configurational pulling. Thus, the sample size in each segment along the configurational CV is 20. The simulation outcome shows a higher sensitivity on the pulling speed. The pulling speed of 8 ps per 5° was used in our previous work, which corresponds to a pulling time of approximately 5 ps for the 3° segment used in this work. The previous work is performed under the MM Hamiltonian while in the current work we use some SQM Hamiltonians. As the fluctuations under different Hamiltonians are different, some tests should be performed on the pulling speed to ensure that our results are converged on this degree of freedom. 5 ps, 10 ps, and 15 ps are tested for the current 3° segment in the configurational space. The alchemical transformation uses the initial configurations obtained with the selection criterion in the ASMD simulations along the configurational CV. In the previous BAR-based case, bidirectional alchemical transformations finished in several time steps are sufficient to converge the results.³⁴ In the current EXP-based case, the statistical efficiency is a bit worse and thus we employ longer pulling times for the alchemical transformation. The relaxation time between successive perturbations is set to 1 time step and the transformation is finished in 10 or 100 time steps, corresponding to a change of 0.1 or 0.01 per time step for the alchemical order parameter.

In all MD simulations, the time step is set to 0.5 fs to ensure stable and accurate dynamics,⁸⁹ and Langevin dynamics¹⁰⁸ with the collision frequency of 5 ps⁻¹ are implemented for temperature regulation at 300 K. As the simulation is performed in vacuo, there is no cutoff applied in our simulations. Note that although the current numerical experiment is performed in vacuo, the Hamiltonian-variation framework could be straightforwardly applied in condensed-phase simulations, which has been performed in our previous works.^{34, 59} We use the AMBER¹⁰⁹ suite for MD simulations and Gaussian 09¹¹⁰ for ab initio QM calculations. All statistical analyses are obtained with homemade codes.

4. Result and discussion

4.1. (S)QM-from-SQM Simulations.

The pulling-speed and sample-size dependence of the estimates.

As discussed in the last part of the methodology section, the convergence diagnostics should be performed in the first place before analyzing detailed results. We choose the target RM1 Hamiltonian as an example. The pulling-speed dependence of the free energy profile in direct free energy simulations at the

RM1 level is presented in Fig. 2a. The results from nonequilibrium pulling initiated from 3 different initial seeds are presented. The 10 ps/segment (i.e., 20000 steps) PMF is virtually identical to the 15 ps/segment results, and minor differences are observed for the 5 ps/segment result. The systematic bias in the 5 ps/segment estimate (i.e., 10000 steps) shows obvious dependence on the initial seed. Therefore, in the nonequilibrium pulling along the configurational CV, we use the pulling speed of 10 ps/segment. Note that there is no overestimation of the computational cost in the direct free energy simulation, which is crucial to obtain a fair estimate of the acceleration ratio of the indirect scheme. We then check the sample-size dependence of the indirect estimates obtained with the pulling time of 10 steps in the alchemical space. In Fig. 2b, the comparison between the indirect estimates obtained from the AM1 simulations with different sample sizes and different estimators for the AM1-to-RM1 perturbation from one initial seed and the direct RM1 PMF is presented. We can see that the 5-sample, 10-sample and 50-sample estimates obtained from the EXP, GEXP and W_a estimators are all identical, which indicates that the sample size is not a bottleneck for the convergence of the nonequilibrium free energy simulation. Namely, the estimates are numerically stable and could be converged easily. The difference between the indirect and direct estimates are small, but we expect to further minimize this deviation.

The agreement between the EXP, GEXP and W_a estimates does not necessarily indicate the convergence of the nonequilibrium pulling simulations. The configurational sampling in the equilibrium perturbation scheme is performed in the equilibrium ensemble, while that in the nonequilibrium pulling case required by JI or CE is performed in both the equilibrium ensemble and the nonequilibrium pulling. The ASMD scheme uses a single structure to seed a series of nonequilibrium simulations, from which the microscopic nonequilibrium works are extracted. As a result, these microscopic nonequilibrium works are not independent and thus do not fully satisfy the assumption of the JI estimator. The inter-trajectory correlation is significant when the nonequilibrium pulling simulation is very short. For instance, when the nonequilibrium pulling is finished in just one step, the nonequilibrium work reduces to the energy difference in the equilibrium perturbation scheme, and the nonequilibrium works from different pulling simulations are exactly the same. Significant bias is introduced in this case. Further, the EXP, GEXP and W_a estimates are exactly the same, which provides no hints on the dissipation of the nonequilibrium trajectories. Therefore, to achieve a sufficient level of configurational sampling, the nonequilibrium trajectories should not be too short. The calculation of the statistical error also has this problem. As the analytical formula of the statistical uncertainty requires independent works as input, the statistical error could be underestimated in ASMD

simulations. The initial-configuration-induced systematic bias is significant for the alchemical transformation due to the short length of such pulling simulation, but is very small for the configurational part, as the long pulling time in the latter simulations enables sufficient sampling in the nonequilibrium ensemble.

Although the EXP, GEXP and W_a estimates are virtually identical, obvious differences could be observed for their statistical errors. We focus on the alchemical perturbation to the RM1 level and the results obtained with the pulling speed of 10 ps/segment in the configurational space and the pulling time of 10 steps in the alchemical space are presented in Fig. 3a. We can see that the statistical uncertainty of the EXP estimator is much smaller than the others for all of the three alchemical transformations from SQM Hamiltonians to the RM1 level considered in the current work. The GEXP and W_a estimators have extremely similar statistical uncertainties, which is expected considering the agreement between the free energy estimates obtained from these two estimators. The first term of GEXP is the ordinary average of the microscopic works W_a , and the second term relates to the variance of the work distribution. When the GEXP and W_a estimates are virtually identical, the width of the work distribution is negligible, and the resulting statistical errors of these two estimators are extremely similar. Fig. 3b shows the dependence of the statistical uncertainty on the pulling speed in the configurational space for the AM1-to-RM1 transformation. Still, the EXP uncertainty is much smaller than the others. The statistical uncertainties obtained with the pulling speeds of 10 ps/segment and 15 ps/segment are extremely similar, which indicates that the pulling time in the configurational space is sufficiently long and the simulation outcome does not depend on this pulling speed.

Indirect vs direct.

After testing the pulling speed and the sample size in nonequilibrium free energy simulations, we then test the probably most influencing parameter, the initial seed. As discussed in the previous part, the initial configuration has a larger impact on the alchemical part due to the short transformation length of the nonequilibrium pulling, while in the configurational space the lack of sampling in the equilibrium ensemble (i.e., the initial configuration) is compensated by the sampling in the nonequilibrium ensemble during the long pulling time. To provide a clearer presentation of the initial-configuration-induced bias, we compare the indirect estimates obtained from different initial seeds with the direct RM1 result. The indirect estimates obtained from PM6, AM1 and MNDO are presented in Fig. S1 and S2. If not explicitly mentioned, the alchemical perturbation is finished in 10 steps in these cases.

For the PM6 Hamiltonian, longer pulling time in the configurational space does not necessarily result in better indirect estimates, as shown in Fig. S1a-b. The indirect estimates with the pulling speed of 10 ps/segment in the configurational space fluctuate around the direct RM1 result, as shown in Fig. S1a, c, d and e. This is expected as the sampling in the nonequilibrium ensemble is insufficient for the alchemical transformation. The situation could be bettered by performing longer pulling simulations in the alchemical space. As shown in Fig. S1f, the indirect result becomes closer to the direct RM1 estimate when the alchemical perturbation is finished in 100 steps.

Systematic investigations at the AM1 and MNDO levels could be more informative. The indirect estimates obtained from ASMD simulations initiated from one initial configuration with the pulling speed of 10 ps/segment in the configurational space and the pulling time of 10 steps in the alchemical space are presented in Fig. 4a, and the indirect estimates obtained from 8 initial seeds with the same pulling scheme are presented in Fig. S2a-h. We can see that the indirect estimates seem to fluctuate around the RM1 result. For some initial seeds, the accord between the indirect and direct estimates could be good, while for the others the agreement is not so satisfactory. The results obtained with slower pulling speeds in the configurational space are presented in Fig. S2i-k, from which we observe that longer pulling time in the configurational space does not help to improve the indirect estimates.

As has been discussed above, to better the indirect estimate, we need to perform more sampling either in the equilibrium ensemble or in the nonequilibrium one. If we want to distribute the sampling in the equilibrium ensemble, we need more uncorrelated configurations to initiate the nonequilibrium alchemical transformation and average over these simulations to obtain a more accurate and stable estimate, which could be achieved by either equilibrating the initial configurations obtained with the selection criterion from one ASMD trail or performing multiple ASMD simulations initiated from different seeds along the configurational CV. The former scheme could be easily performed and the results are expected to be good, while the performance of the latter regime is not straightforward and thus our numerical test focuses on this scheme. The results averaged over 5 initial seeds are presented in Fig. 4b, from which we can see that the indirect estimate becomes more stable and smoother and the agreement between the indirect and direct estimates is improved. The mean absolute error (MAE) is calculated to provide some numerical metrics to assess the deviation of the indirect estimate from the direct result, which is presented in Table 1. We can see that the MAE of the initial-configuration-averaged estimate is much smaller than the individual estimates. Another way to improve the result is distributing the sampling in the nonequilibrium ensemble, namely lengthening the pulling time in the alchemical space. We employ a longer pulling time of 100 steps in the

alchemical transformation with a change of 0.01 per time step for 5 of the initial seeds, and the results are presented in Fig. 4c and Fig. S3a-d. The indirect estimates are effectively improved by using a slower pulling speed along the alchemical CV. Again, we use a longer pulling time in the configurational space in Fig. S3e-f, where we observe that the pulling speed in the configurational space has little impact on the simulation outcome. The indirect estimate could be improved by using both the slower-pulling-speed and initial-seed-averaging schemes, the result of which is shown in Fig. 4d. The MAEs of the 0.01 indirect estimates shown in Table 1 are smaller than the corresponding 0.1 components, and the seed-averaged result achieves an MAE of about 0.09 kcal/mol, which is technically speaking negligible in free energy calculations. The state-specified deviations of the indirect estimates from the direct reference PMF are presented in Fig. S4, from which we can see that the error of the indirect scheme is small in all configurational states. Although both the slower-pulling-speed and the initial-seed-averaging regimes could improve the indirect estimate, their relative efficiency could differ in practical cases. Practical considerations based on own experience are converging the sample size (i.e., initial seeds) first and then lengthening the pulling time for convergence check on this degree of freedom.

Hitherto, we have gained some insights into using the ASMD-based indirect nonequilibrium free energy simulation scheme. We then employ the scheme to obtain the HF result indirectly. We still use the same pulling speed in nonequilibrium pulling, i.e., 10 ps/segment in the configurational space and 10 or 100 steps in the alchemical transformation. The MAEs of 7 trails with a change of 0.1 per step in the alchemical transformation and that of one trail with a change of 0.01 per step are presented in Table S2. We can see that the HF result could be obtained with the ASMD-based indirect scheme with an MAE \sim 0.25 kcal/mol. The previous RM1-from-SQM indirect simulations do not result in a speedup over the direct pulling in the configurational space, but the current HF-from-SQM one does accelerate the simulation. Thus, we then calculate the speedup of the indirect scheme. When performing free energy simulations in the configurational space, it requires 24 ns to converge the free energy profile. The pulling time of 10 steps is used for the alchemical transformation. 5 samples are sufficient to converge the alchemical term for each initial seed and 5 repeats/seeds are averaged to obtain the final indirect result with very good convergence behavior. When only a single initial seed is used and the initial conditions in ASMD segments are determined by the selection criterion, the need for equilibrium configurational sampling to obtain the initial configurations is eliminated and thus ϕ_{eq} is set to zero. However, to grab different initial seeds to perform a seed-averaging procedure, short equilibrium sampling in each intermediate state is required. However, as the autocorrelation time is much shorter than the pulling time in each segment, this computational cost is still

negligible. The simulation time at the HF level is scaled by the simulation speeds in Table S1 to obtain the effective sampling time at the SQM level. The resulting speedup of the indirect scheme is about 1000-fold, as shown in Table 2. Here, the computational cost of the ps-length further sampling for more initial seeds is not included in the calculation. However, this contribution is negligible and has little effect on the acceleration ratio. If we use the pulling time of 100 steps for the alchemical transformation, the resulting indirect estimate would be extremely accurate and the speedup would be about 200 folds. The computational speedup could also be compared with the statistics of the BAR-based scheme in our previous work,³⁴ which is presented in Table S3. As the BAR-based scheme achieves faster convergence than the ASMD regime in the configurational sampling,⁷⁸ the speedup of the ASMD-based indirect scheme is relatively modest compared with the ~1000-fold speedup shown in the previous table. However, the speedup of the indirect scheme is still two orders of magnitude.

Some more insights about speedups.

In the BAR-based scheme, the speedup of the indirect scheme could be approximated as⁷²

$$\text{speedup} = \frac{\phi_{\text{NEW,conf}} * t_{\text{AB}} + \phi_{\text{eq}} * t_{\text{AB}}}{\phi_{\text{NEW,conf}} + \phi_{\text{NEW,alchem}} * t_{\text{AB}} + (2 + t_{\text{AB}}) \phi_{\text{eq}}} \quad (9).$$

Here, t_{AB} represents the relative computational cost of a single evaluation of energetics under the two Hamiltonians, which is used to scale the sampling times under different Hamiltonians to make them directly comparable. The exact value of this term depends on the selection of the two models/Hamiltonians and the practical implementation (e.g., software) when performing the calculation. $\phi_{\text{NEW,conf}}$ is the pulling time in each segment along the configurational CV, which is determined by the magnitude of perturbation along this CV. $\phi_{\text{NEW,alchem}}$ denotes the pulling time in alchemical transformations. ϕ_{eq} is the statistical inefficiency or the sampling cost required to get an independent initial condition. The equation is valid under the following four assumptions.⁷² First, the equal-sample-size rule is applied to all segments in the multi-dimensional nonequilibrium pulling perturbation network. For more detailed discussions about the differences between the equal-time and equal-sample-size protocols, please refer to our previous works.^{59, 72, 78, 104, 105} Second, the pulling times in the forward (from Hamiltonian A to Hamiltonian B) and backward (from B to A) alchemical transformations are the same $\phi_{\text{NEW,alchem}}$. Third, the pulling times in all segments along the configurational CV are the same $\phi_{\text{NEW,conf}}$. Fourth, the autocorrelation times in all microstates are the same ϕ_{eq} . In the SQM-to-QM case, as the computational cost of the ab initio QM level is much higher than the

SQM one ($t_{AB} \gg 1$) and the pulling time along the configurational CV $\phi_{\text{NEW,conf}}$ is short and similar to the statistical inefficiency ϕ_{eq} due to the high statistical efficiency of BAR, the computational cost of the direct SQM free energy simulation under the model A in the denominator is negligible. As a result, the speedup of the indirect scheme could be further simplified to the ratio of the sampling time in the direct QM free energy simulation to that in the SQM-to-QM perturbation, namely

$$\text{speedup} = \frac{\phi_{\text{NEW,conf}} + \phi_{\text{eq}}}{\phi_{\text{NEW,alchem}} + \phi_{\text{eq}}} \quad (10).$$

The above equation tells us that the acceleration ratio of the bidirectional perturbation network depends on the length of the pulling time along the alchemical pathway, that along the configurational CV, and the equilibrium autocorrelation time due to the need of generating the initial conditions. Due to the stratified protocol and the high statistical efficiency of bidirectional perturbations, the pulling time along the physical CV is similar to the autocorrelation time, and the pulling time along the alchemical CV depends on the similarities of the two Hamiltonians (i.e., models A and B). Whether $\phi_{\text{NEW,conf}}$ is longer than $\phi_{\text{NEW,alchem}}$ determines whether the indirect scheme is faster or more computationally efficient than direct simulations. Another point in the above equation is that the equilibrium sampling ϕ_{eq} is non-negligible in practical situations. When the autocorrelation time is much longer than the pulling time along both CVs, the direct and indirect schemes are of similar efficiencies and thus the merit of indirect simulations vanishes.

However, the situation is different in the current ASMD case. All transformations are performed unidirectionally and no equilibrium sampling is required (or very short ps-length in the initial-seeds-averaging procedure). To get a more detailed insight into the acceleration ratio in this situation, we then derive the formula of the speedup ratio under the current unidirectional nonequilibrium perturbation network with selection-criterion-based ASMD pulling. For the direct free energy simulations along the configurational CV, the computational costs under the two models A and B are

$$\text{cost}_{\text{direct,B}} = N_{\text{segments}} * N_{\text{traj}} * \phi_{\text{NEW,conf}} * t_{AB} \quad (11),$$

$$\text{cost}_{\text{direct,A}} = N_{\text{segments}} * N_{\text{traj}} * \phi_{\text{NEW,conf}} \quad (12).$$

As the systematic bias in direct free energy simulations could be effectively eliminated by lengthening the pulling time along the configurational pathway, there is no need to generate equilibrium samples and thus no ϕ_{eq} exists in the above equations. For the alchemical transformation, the initial-seeds-averaging procedure

seems necessary to remove the systematic bias, and we consider the number of seeds to be N_{seeds} . As has been discussed previously, there are basically two ways to obtain the initial seeds. The first one is generating independent configurations in the initial state and repeating the whole ASMD pulling along the configurational pathway, which requires short ps-length sampling in the initial reference state and N_{seeds} full ASMD pulling along the configurational CV. The second one is equilibrating the close-to-equilibrium configuration obtained with the selection criterion from one ASMD trail, which requires short ps-length sampling in each segment. As only a single ASMD trail is defined in the computational cost of Eq. (12), the second scheme is considered in the following derivation. Note that the computational cost of the first scheme could be similarly derived. In the second scheme, for each initial seed, the number of pulling trajectories along the alchemical pathway is $N_{\text{traj,alchem}}$. Then, the A-to-B unidirectional alchemical transformation costs

$$\text{cost}_{\text{A-to-B}} = N_{\text{segments}} * N_{\text{seeds}} * \left(N_{\text{traj,alchem}} * \phi_{\text{NEW,alchem}} * t_{\text{AB}} + \phi_{\text{eq}} \right) \quad (13).$$

The speedup of the indirect scheme can be expressed as

$$\begin{aligned} \text{speedup} &= \frac{\text{cost}_{\text{direct,B}}}{\text{cost}_{\text{direct,A}} + \text{cost}_{\text{A-to-B}}} \\ &= \frac{N_{\text{traj}} * \phi_{\text{NEW,conf}} * t_{\text{AB}}}{N_{\text{traj}} * \phi_{\text{NEW,conf}} + N_{\text{seeds}} * \left(N_{\text{traj,alchem}} * \phi_{\text{NEW,alchem}} * t_{\text{AB}} + \phi_{\text{eq}} \right)} \end{aligned} \quad (14).$$

In the current illustrative calculations (dihedral flipping in ACE-NME), the number of trajectories along the configurational CV and the total number of trajectories along the alchemical CV are similar, i.e., $N_{\text{traj}} \approx N_{\text{seeds}} * N_{\text{traj,alchem}}$. Thus, the above speedup could be further simplified as

$$\text{speedup} = \frac{\phi_{\text{NEW,conf}} * t_{\text{AB}}}{\phi_{\text{NEW,conf}} + \phi_{\text{NEW,alchem}} * t_{\text{AB}} + N_{\text{seeds}} * \phi_{\text{eq}}} \quad (15).$$

The numerator is solely the pulling time along the configurational CV under the target Hamiltonian B, while the denominator is composed of the pulling time along the configurational CV under the Hamiltonian A, the A-to-B alchemical transformation term, and the short equilibration procedure to grab several initial seeds under the Hamiltonian A. Due to the low statistical efficiency of unidirectional estimators, the pulling time along the configurational space $\phi_{\text{NEW,conf}}$ is often much longer than the autocorrelation time ϕ_{eq} . Further, the number of seeds is often small (~5 in the current illustrative calculations). As a result, the last term in the denominator is much smaller and even negligible compared to $\phi_{\text{NEW,conf}}$, and the equation could be further

simplified as

$$\text{speedup} = \frac{\phi_{\text{NEW,conf}} * t_{\text{AB}}}{\phi_{\text{NEW,conf}} + \phi_{\text{NEW,alchem}} * t_{\text{AB}}} \quad (16).$$

The first term in the denominator is the contribution of the direct free energy simulation under the model A, while the second term arises from the A-to-B unidirectional pulling. Some basic insights about their sizes could be seen in Table 2 and S1, where the two terms in the denominator are of similar magnitudes and thus the equation cannot be further simplified. We then can compare the dependence of the speedups in the current ASMD Eq. (16) with the previous BAR-based case Eq. (9). Compared with the BAR-based case, the acceleration ratio of the ASMD case does not have the contribution from the equilibrium sampling at the computationally costly Hamiltonian B ($\phi_{\text{eq}} * t_{\text{AB}}$), and the equilibrium sampling at the computationally feasible Hamiltonian A (ϕ_{eq}) is also negligible due to the selection criterion and the significant length of $\phi_{\text{NEW,conf}}$. As the computational costs between the two models A and B (t_{AB}) must be larger than 1 to achieve computational gains for the indirect scheme, we consider a special case that t_{AB} is a large number ~ 3000 , namely the current HF-from-SQM case. For the BAR-based case, the speedup ratio in Eq. (9) reduces to Eq. (10), where in the denominator the contribution from the direct free energy simulation under the model A becomes negligible but the cost of equilibrium sampling still exists. By contrast, for the ASMD-based case, the contribution of the direct free energy simulation under the model A in Eq. (16) is still non-negligible, as shown in Table 2. Therefore, the basic differences between the acceleration ratios of the current ASMD and the previous BAR-based schemes lie in the contributions from equilibrium sampling for initial conditions and the direct free energy simulation under the computationally feasible model A. However, a key feature that remains unchanged is that the main contribution of the computational costs of the indirect scheme still comes from the alchemical transformation.

4.2. Indirect SQM-from-MM Simulations.

The above examples of indirect free energy simulations at SQM and QM levels provide evidence of the applicability of the multi-dimensional ASMD framework for Hamiltonian variations. The perturbation between SQM and QM Hamiltonians in the ACE-NME dihedral flipping case is relatively easy to converge due to the similarities of these models.³⁴ We then test the protocol in a challenging case, where the same ACE-NME system is selected but the two end states (i.e., Model A and B in Fig. 1) in the alchemical perturbation are MM and SQM Hamiltonians. Such MM-to-SQM perturbation is very difficult to converge

or get rid of the systematic bias even in this simple system, as shown in our previous work.⁵⁹ The computationally efficient Model A is AMBER14SB,¹⁰⁷ while the target SQM level is still RM1. The same computational protocol as the previous SQM-to-QM simulations is used in this case. Namely, we conduct direct MM simulations with the pulling speed of 10 ps per 3 ° segment and the sample size of 20 in each segment, after which nonequilibrium MM-to-SQM transformations are performed in 100 time steps with a perturbation speed of 0.01 per time step for the alchemical order parameter. The initial-seed-averaged results are presented in Fig. 5. We can see that even with long pulling time (100 time steps), the alchemical perturbation is still biased and the indirect PMF is obviously different from the direct RM1 result. Therefore, longer pulling time is required for fully converged results. We thus repeat the alchemical transformation with a slower pulling speed of 1000 time steps (0.001 per time step) and the results are also presented in Fig. 5. Still, obvious differences between direct and indirect RM1 results could be observed. The computational cost of the indirect scheme with 100 time steps for the alchemical transformation is shown in Table 3, where the computational cost of the indirect scheme is still much lower than the direct scheme. The relative efficiency of the 1000-time-steps protocol is about 1.5, which suggests that the indirect scheme is still computationally cheaper than direct free energy simulations even with 1000 time steps (i.e. 5 ps) for alchemical transformation. However, for both the 0.01 and 0.001 protocol, the agreement between the direct and indirect results are not very good, which indicates that the indirect scheme could be unreliable and inefficient in indirect SQM-from-MM simulations. The relative efficiency becomes 1 when the alchemical transformation is performed in 7.8 ps (~1560 time steps). Although we could further lengthen the alchemical transformation, the indirect scheme is already not a computationally efficient option. Thus, here we try another solution of changing/refining the low-level description to improve the phase space overlap between different Hamiltonians.^{111, 112}

The non-satisfactory convergence behavior in indirect SQM free energy simulations arises from the dissimilarities of MM and SQM Hamiltonians.^{113, 114} We thus seek for a better MM model that is more similar to the target SQM model to improve the convergence. The generalized force-matching scheme¹¹⁵ is employed to refit the AMBER14SB force field. The basic idea of the method is to optimize a force field that reproduces the energy, the atomic force,¹¹⁶ and other selected properties of a target Hamiltonian simultaneously.¹¹⁵ In our force-matching fitting, the energy and every component of the atomic forces¹¹⁷ are included in the objective function, and the initial/reference model used to initiate the refitting is AMBER14SB. The configurational ensemble used is generated in a 20 ns gas-phase unbiased simulation with the original AMBER14SB parameters. As the backbone dihedral under investigation has a partial

double-bond character and the free energy barrier along its flipping pathway is rather high (~ 10 kcal/mol), some enhanced sampling techniques should be employed to sample this degree of freedom thoroughly. As the configurational sampling in this phase aims at extracting representative structures rather than obtaining correct thermodynamic distributions in a specific state, a high temperature of 1300 K is used in the unbiased simulation to improve the sampling efficiency. The weight of each sample in parameter optimization could influence the refitting outcome, but for simplicity we apply the uniform weighting scheme, which would not lead to significant problems. The sampling interval is set to 2 ps according to autocorrelation analysis. Note that other schemes such as the dihedral scan or biasing specific CV (e.g., the backbone dihedral) could also be used to generate the configurational ensemble. In order to avoid overfitting and the existence of unphysical results (e.g., unreasonable force constants), the parameter space to explore is restrained in the neighborhood of the AMBER14SB parameter set with the L2 (harmonic) regularization term. The atom-type symmetry of force field parameters is also included in the optimization. Only the bonded interactions including the bond stretching, angle (bending), and dihedral terms are refitted, and the non-bonded terms including the atomic charges and vdW terms remain fixed. As our target level is RM1, this SQM model is used to calculate the reference data. The resulting refitted force field (force-matching RM1 or FM-RM1) produces a description that is more similar to the target RM1 Hamiltonian. To assess the quality of the refitted force field in reproducing the RM1 energetics, we re-generate 10 ns trajectories at 1300 K and 300 K, respectively, and calculate the root-mean-squared error (RMSE) and MAE of the energy with the original parameter set (i.e., AMBER14SB) and the newly fitted model. The higher temperature (i.e., 1300 K) enables the evaluation of the accuracy of the energetics in the ensemble that the configurations used in refitting are generated. As both high-energy and low-energy regions are explored, this high-temperature evaluation reflects the quality of the refitted model in the whole relevant configurational space. By contrast, the low-temperature (300 K) set only samples local fluctuations and thus is used to evaluate the accuracy of the refitted model in the low-energy regions/minima. The correlations between the MM and target RM1 energetics are presented in Fig. S5, where we can see that in the refitted model yields smaller deviations from the reference RM1 result in both the high- and low-temperature situations. Namely, both the high- and low-energy regions are described in a more accurate way by the new model. We also calculate the RMSE of atomic forces from these structures. Upon refitting the AMBER14SB force field, the RMSE of atomic forces is improved from 23.9 kcal/(mol $\cdot\text{\AA}\cdot\text{atom}$) to 13.1 kcal/(mol $\cdot\text{\AA}\cdot\text{atom}$) for configurations sampled at 300 K and from 37.6 kcal/(mol $\cdot\text{\AA}\cdot\text{atom}$) to 25.0 kcal/(mol $\cdot\text{\AA}\cdot\text{atom}$) for configurations generated at 1300 K. The accuracy improvement is more pronounced for the ensemble of configurations generated at the high

temperature, which indicates that the original force field (i.e., AMBER14SB) does provide a good description of the simulated system in low-energy regions.

We then repeat the multi-dimensional ASMD simulations to obtain the RM1 PMF indirectly from direct MM (refitted AMBER14SB) simulations and nonequilibrium MM-to-RM1 perturbations, the results of which are presented in Fig. 5. Compared with the indirect RM1 estimates obtained from AMBER14SB sampling, the refitted force field provides estimates closer to the direct RM1 result. Thus, the refitted model (i.e., refitted AMBER14SB) serves as a better model in indirect free energy simulations. The convergence of the indirect scheme could be achieved with 100 time steps for the alchemical order parameter, leading to an acceleration ratio close to the statistics shown in Table 3. We should note that although the indirect estimate is improved upon MM refitting, the accord between the direct and indirect estimates is still not as good as the previous examples where transformations between SQM or ab initio QM Hamiltonians are performed. To further improve the indirect estimate, longer nonequilibrium pulling or more elaborate refitting of the MM force field could be considered.

5. Conclusion

The combination of the alchemical and configurational free energy simulations provides a computationally feasible alternative to obtain the thermodynamics at ab initio QM levels. The free energy landscape obtained from enhanced sampling simulations under a computationally feasible Hamiltonian could be perturbed to that under another Hamiltonian with the alchemical method. In our previous works, we constructed a multi-dimensional nonequilibrium pulling framework for Hamiltonian variation based on bidirectional pulling and bidirectional reweighting. The BAR-based method is generally applicable and the acceleration of the indirect method is about an order of magnitude compared with the direct free energy simulations under some ab initio QM Hamiltonians. As bidirectional pulling could be costly in some cases, in the current work, we alter the bidirectional perturbation framework to a unidirectional one and use the exponential average, its Gaussian approximated form, and the ordinary average to estimate the free energy difference. The staged SMD simulation is further accelerated with the selection criterion for the initial configuration used to spawn nonequilibrium trajectories. This ASMD-based Hamiltonian variation framework is more than 100-fold faster than the direct free energy simulations under some ab initio QM Hamiltonian.

A worth noting behavior of the ASMD-based scheme is the distribution of the sampling. In the normal staged SMD pulling with JI or CE for free energy estimates, equilibrium sampling in each intermediate state

is performed to obtain the initial configuration for kicking off the pulling simulations. The sampling is distributed in both the equilibrium ensemble and the nonequilibrium counterpart. However, as the nonequilibrium trajectories are initiated from one initial configuration in ASMD simulations, some alternations of the simulation scheme are required. We can distribute some sampling times in the equilibrium ensemble by equilibrating the initial configuration obtained from the selection criterion, or performing ASMD simulations from different initial seeds to obtain more initial configurations. In this way, more independent initial configurations are obtained and nonequilibrium alchemical pulling simulations from these uncorrelated configurations could improve the convergence behavior of the indirect scheme. Alternatively, we can sample only in the nonequilibrium ensemble by pulling it slower in the alchemical space, which also improves the convergence behavior of the indirect scheme. The initial-seed-averaging and slower-pulling-speed schemes could be combined to achieve better numerical behaviors.

Although the JI-based ASMD scheme could be efficient for multi-dimensional nonequilibrium pulling, there are still some weaknesses. As the whole pulling simulation is initiated from one or several configurations, the nonequilibrium pulling in the configurational space is performed in a serial way. Specifically in the current dihedral flipping case, the simulation is initiated from the 0° state and the pulling is performed from this reference state to the other in a serial manner. The 3° -to- 6° pulling simulations need to wait for the outcome of the 0° -to- 3° simulations, and the 6° -to- 9° pulling simulations need to wait for the outcome of the 3° -to- 6° simulations. This inter-segment dependence limits the efficiency of parallelism and we can only parallelize the nonequilibrium trajectories of the current segment. Although the calculation of the energetics of each simulation could also be parallelized, the efficiency of this parallel computation is degraded compared with the delightfully parallel BAR-based scheme even for parallel-optimized dynamic engines (i.e., the so-called parallel slowdown due to communications overhead). A factor that further aggravates this parallel issue is the need for slower pulling speeds when using the exponential average. Compared with bidirectional reweighting, the lower statistical efficiency of EXP requires a longer pulling time to obtain converged estimates and thus the waiting time for each segment is lengthened. The overall outcome of these factors is a longer waiting time before user feedback, which is not a satisfactory property for protocol diagnosing. However, if the user already has some experiences on the convergence behavior of the system (e.g., after the protocol testing procedure), or if there are a series of systems with similar properties (e.g., a series of mutants of some protein-ligand complexes), the EXP-based ASMD scheme could save the computational costs significantly. Note that the tests of a series of different protocols could also be performed independently, and the parallelization in this case could also be useful. However, the user still

needs to wait for the output of the longer/longest protocol with the higher/highest reliability before he/she is able to confirm whether the faster protocols are reliable.

Supporting Information Description

The single-core timing information of SQM and ab initio QM simulations, MAEs of the indirect estimates at ab initio QM levels, comparison between the direct and indirect estimates at the RM1 level initiated from different seeds with different pulling speeds, the quality of energetics produced by the AMBER14SB and the force-matching refitted force fields are provided in the supporting information.

Conflict of Interest Statement

There are no conflicts of interest to declare.

Acknowledgement

Part of the simulation was performed on the high-performance computing platform of the Center for Life Science (Peking University). Dr. Zhaoxi Sun is supported by the PKU-Boya Postdoctoral Fellowship. We thank anonymous reviewers for valuable comments and critical reading.

Preprint Acknowledgement

Research presented in this article has been posted on a preprint server prior to publication. The corresponding preprint article can be found here: <http://dx.doi.org/10.26434/chemrxiv.13634987>.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Sharma, K. D.; Kathuria, P.; Wetmore, S. D.; Sharma, P., Can modified DNA base pairs with chalcogen bonding expand the genetic alphabet? A combined quantum chemical and molecular dynamics simulation study. *Phys. Chem. Chem. Phys.* **2020**.
2. de Ruiter, A.; Oostenbrink, C., Advances in the calculation of binding free energies. *Curr. Opin. Struct. Biol.* **2020**, 61, 207-212.
3. Pecina, A.; Eyrilmez, S. M.; Köprülüoğlu, C.; Miriyala, V. M.; Lepšík, M.; Fanfrlík, J.; Řezáč, J.; Hobza, P., SQM/COSMO Scoring Function: Reliable Quantum - Mechanical Tool for Sampling and Ranking in Structure - Based Drug Design. *ChemPlusChem* **2020**, 85, 2362-2371.
4. Echeverria, I.; Amzel, L. M., Helix propensities calculations for amino acids in alanine based peptides using Jarzynski's equality. *Proteins: Structure, Function, and Bioinformatics* **2010**, 78, 1302-1310.
5. Lee, T. S.; Radak, B. K.; Huang, M.; Wong, K. Y.; York, D. M., Roadmaps through free energy landscapes calculated using the multi-dimensional vFEP approach. *J. Chem. Theory Comput.* **2014**, 10, 24-34.
6. Sun, Z.; Wang, X.; Zhang, J. Z. H., Protonation-dependent Base Flipping in The Catalytic Triad of A Small RNA. *Chem. Phys. Lett.* **2017**, 684, 239-244.
7. Moraca, F.; Amato, J.; Ortuso, F.; Artese, A.; Pagano, B.; Novellino, E.; Alcaro, S.; Parrinello, M.; Limongelli, V., Ligand binding to telomeric G-quadruplex DNA investigated by funnel-metadynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, 114, E2136-E2145.
8. Wang, X.; Deng, B.; Sun, Z., Thermodynamics of helix formation in small peptides of varying length in vacuo, in implicit solvent, and in explicit solvent. *J. Mol. Model.* **2019**, 25, 3.
9. Sun, Z.; Wang, X., Thermodynamics of Helix formation in small peptides of varying length in vacuo, implicit solvent and explicit solvent: Comparison between AMBER force fields. *Journal of Theoretical and Computational Chemistry* **2019**, 1950015.
10. Gore, J.; Ritort, F.; Bustamante, C., Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12564-12569.
11. Sun, Z.; Zhang, J. Z. H., Thermodynamic Insights of Base Flipping in TNA Duplex: Force Fields, Salt Concentrations, and Free-Energy Simulation Methods. *CCS Chemistry* **2021**, 3, 1026-1039.
12. Lemkul, J. A.; Savelyev, A.; MacKerell Jr, A. D., Induced Polarization Influences The Fundamental Forces in DNA Base Flipping. *J. Phys. Chem. Lett.* **2014**, 5, 2077-2083.
13. Sun, Z.; He, Q.; Li, X.; Zhu, Z., SAMPL6 host-guest binding affinities and binding poses from spherical-coordinates-biased simulations. *J. Comput.-Aided Mol. Des.* **2020**, 34, 589-600.
14. Mezei, M., Adaptive Umbrella Sampling: Self-consistent Determination of the Non-Boltzmann Bias. *J. Comput. Phys.* **1987**, 68, 237-248.
15. Hooft, R. W.; van Eijck, B. P.; Kroon, J., An Adaptive Umbrella Sampling Procedure in Conformational Analysis using Molecular Dynamics and Its Application to Glycol. *J. Chem. Phys.* **1992**, 97, 6690-6694.
16. Kästner, J., Umbrella sampling. *Wiley Interdisip. Rev. Comput. Mol. Sci.* **2011**, 1, 932-942.
17. Fukunishi, H.; Watanabe, O.; Takada, S., On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.* **2002**, 116, 9058-9067.
18. Itoh, S. G.; Damjanovic, A.; Brooks, B. R., pH replica-exchange method based on discrete protonation states. *Proteins* **2011**, 79, 3420-36.
19. Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C., Improved Efficiency of Replica Exchange Simulations through Use of a Hybrid Explicit/Implicit Solvation Model. *J. Chem. Theory Comput.* **2006**, 2, 420.
20. Sabri, D. D.; Roitberg, A. E., Optimization of Umbrella Sampling Replica Exchange Molecular Dynamics by Replica Positioning. *J. Chem. Theory Comput.* **2013**, 9, 4692-4699.
21. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, 314, 141-151.
22. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, 113, 6042-6051.
23. Procacci, P., Accuracy, precision, and efficiency of nonequilibrium alchemical methods for computing free energies of

- solvation. I. Bidirectional approaches. *J. Chem. Phys.* **2019**, 151, 144113.
24. Chelli, R.; Marsili, S.; Barducci, A.; Procacci, P., Recovering the Crooks equation for dynamical systems in the isothermal-isobaric ensemble: a strategy based on the equations of motion. *J. Chem. Phys.* **2007**, 126, 044502.
25. Procacci, P., Unbiased free energy estimates in fast nonequilibrium transformations using Gaussian mixtures. *J. Chem. Phys.* **2015**, 142, 2690-2693.
26. Nicolini, P.; Frezzato, D.; Chelli, R., Exploiting configurational freezing in nonequilibrium Monte Carlo simulations. *J. Chem. Theory Comput.* **2011**, 7, 582-593.
27. Chelli, R.; Marsili, S.; Barducci, A.; Procacci, P., Generalization of the Jarzynski and Crooks nonequilibrium work theorems in molecular dynamics simulations. *Phys. Rev. E* **2007**, 75, 050101.
28. Nerattini, F.; Chelli, R.; Procacci, P., II. Dissociation free energies in drug-receptor systems via nonequilibrium alchemical simulations: application to the FK506-related immunophilin ligands. *Phys. Chem. Chem. Phys.* **2016**, 18, 15005-15018.
29. Chelli, R.; Gellini, C.; Pietrapperia, G.; Giovannelli, E.; Cardini, G., Path-breaking schemes for nonequilibrium free energy calculations. *J. Chem. Phys.* **2013**, 138, 214109.
30. Chelli, R.; Procacci, P., A potential of mean force estimator based on nonequilibrium work exponential averages. *Phys. Chem. Chem. Phys.* **2009**, 11, 1152-1158.
31. Zwanzig, R. W., High Temperature Equation of State by A Perturbation Method. *J. Chem. Phys.* **1954**, 22, 1420-1426.
32. Wood, R. H.; Muhlbauer, W. C. F.; Thompson, P. T., Systematic errors in free energy perturbation calculations due to a finite sample of configuration space: sample-size hysteresis. *J. Phys. Chem.* **1991**, 95, 6670-6675.
33. Zuckerman, D. M.; Woolf, T. B., Theory of a systematic computational error in free energy differences. *Physical Review Letters* **2002**, 89, 180602.
34. Sun, Z., BAR-based multi-dimensional nonequilibrium pulling for indirect construction of QM/MM free energy landscapes: from semi-empirical to ab initio. *Phys. Chem. Chem. Phys.* **2019**, 21, 21942-21959
35. Wang, X.; Sun, Z., A Theoretical Interpretation of Variance-based Convergence Criteria in Perturbation-based Theories. *arXiv preprint arXiv:1803.03123* **2018**.
36. Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S., Equilibrium Free Energies from Nonequilibrium Measurements using Maximum-likelihood Methods. *Physical review letters* **2003**, 91, 140601.
37. Bennett, C. H., Efficient Estimation of Free Energy Differences from Monte Carlo data. *J. Comput. Phys.* **1976**, 22, 245-268.
38. Tan, Z., On a Likelihood Approach for Monte Carlo Integration. *J. Am. Stat. Assoc.* **2004**, 99, 1027-1036.
39. Shirts, M. R.; Chodera, J. D., Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, 129, 124105.
40. Jarzynski, C., A Nonequilibrium Equality for Free Energy Differences. *Physical Review Letters* **1997**, 78, 2690-2693.
41. Mallick, K.; Moshe, M.; Orland, H., Supersymmetry and Nonequilibrium Work Relations. *arXiv preprint arXiv:0711.2059* **2008**.
42. Ballard, A. J.; Jarzynski, C., Replica exchange with nonequilibrium switches: enhancing equilibrium sampling by increasing replica overlap. *J. Chem. Phys.* **2012**, 136, 194101.
43. Vaikuntanathan, S.; Jarzynski, C., Escorted free energy simulations: improving convergence by reducing dissipation. *Physical Review Letters* **2008**, 100, 190601.
44. Dickson, A.; Dinner, A. R., Enhanced Sampling of Nonequilibrium Steady States. *Annual Review of Physical Chemistry* **2010**, 61, 441-459.
45. Procacci, P.; Marsili, S., Energy dissipation asymmetry in the non equilibrium folding/unfolding of the single molecule alanine decapeptide. *Chem. Phys.* **2010**, 375, 8-15.
46. Crespo, A.; Martí, M. A.; Estrin, D. A.; Roitberg, A. E., Multiple-Steering QM-MM Calculation of the Free Energy Profile in Chorismate Mutase. *Journal of the American Chemical Society* **2005**, 127, 6940-6941.
47. Swope, W. C., A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, 76, 637.
48. Procacci, P.; Chelli, R., Statistical Mechanics of Ligand-Receptor Noncovalent Association, Revisited: Binding Site and Standard State Volumes in Modern Alchemical Theories. *J. Chem. Theory Comput.* **2017**, 13, 1924-1933.

49. Shirts, M. R.; Pande, V. S., Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* **2005**, *122*, 134508.
50. Hummer, G.; Pratt, L. R.; Garcia, A. E., Hydration free energy of water. *J. Phys. Chem.* **1995**, *99*, 14188-14194.
51. Pham, T. T.; Shirts, M. R., Identifying low variance pathways for free energy calculations of molecular transformations in solution phase. *J. Chem. Phys.* **2011**, *135*, 034114.
52. Huai, Z.; Sun, Z., Titration of Adenine in a GA mismatch with Grand Canonical Simulations. *Journal of Computational Biophysics and Chemistry* **2020**, *20*, 165-173.
53. Huai, Z.; Yang, H.; Sun, Z., Binding thermodynamics and interaction patterns of human purine nucleoside phosphorylase-inhibitor complexes from extensive free energy calculations. *J. Comput.-Aided Mol. Des.* **2021**.
54. Sun, Z.; Wang, X.; Zhang, J. Z., Theoretical understanding of the thermodynamics and interactions in transcriptional regulator TtgR-ligand binding. *Phys. Chem. Chem. Phys.* **2020**, *22*, 1511-1524.
55. Procacci, P., Methodological uncertainties in drug-receptor binding free energy predictions based on classical molecular dynamics. *Curr. Opin. Struct. Biol.* **2021**, *67*, 127-134.
56. Mikulskis, P.; Cioloboc, D.; Andrejić, M.; Khare, S.; Brorsson, J.; Genheden, S.; Mata, R. A.; Söderhjelm, P.; Ryde, U., Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host-guest binding energies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 375-400.
57. Genheden, S.; Martinez, A. I. C.; Criddle, M. P.; Essex, J. W., Extensive all-atom Monte Carlo sampling and QM/MM corrections in the SAMPL4 hydration free energy challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 187-200.
58. Woods, C. J.; Manby, F. R.; Mulholland, A. J., An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J. Chem. Phys.* **2008**, *128*, 152-159.
59. Wang, X.; He, Q.; Sun, Z., BAR-Based Multi-Dimensional Nonequilibrium Pulling for Indirect Construction of a QM/MM Free Energy Landscape. *Phys. Chem. Chem. Phys.* **2019**, *21*, 6672-6688
60. Huai, Z.; Yang, H.; Li, X.; Sun, Z., SAMPL7 TrimerTrip host-guest binding affinities from extensive alchemical and end-point free energy calculations. *J. Comput.-Aided Mol. Des.* **2021**, *35*, 117-129.
61. Procacci, P.; Guarnieri, G., SAMPL7 blind predictions using nonequilibrium alchemical approaches. *J. Comput.-Aided Mol. Des.* **2021**, *35*, 37-47.
62. Sun, Z., SAMPL7 TrimerTrip Host-Guest Binding Poses and Binding Affinities from Spherical-Coordinates-Biased Simulations. *J. Comput.-Aided Mol. Des.* **2021**, *35*, 105-115.
63. Bentzien, J.; Muller, R. P.; Florián, J.; Warshel, A., Hybrid ab initio quantum mechanics/molecular mechanics calculations of free energy surfaces for enzymatic reactions: the nucleophilic attack in subtilisin. *J. Phys. Chem. B* **1998**, *102*, 2293-2301.
64. Polyak, I.; Benighaus, T.; Boulanger, E.; Thiel, W., Quantum mechanics/molecular mechanics dual Hamiltonian free energy perturbation. *J. Chem. Phys.* **2013**, *139*, 578.
65. Sun, Z.; Zhu, T.; Wang, X.; Mei, Y.; Zhang, J. Z., Optimization of convergence criteria for fragmentation methods. *Chem. Phys. Lett.* **2017**, *687*, 163-170.
66. Liu, W.; Sakane, S.; And, R. H. W.; Doren, D. J., The Hydration Free Energy of Aqueous Na⁺ and Cl⁻ at High Temperatures Predicted by ab Initio/Classical Free Energy Perturbation: 973 K with 0.535 g/cm³ and 573 K with 0.725 g/cm³. *J. Phys. Chem. A* **2002**, *106*, 1409-1418.
67. Olsson, M. A.; Söderhjelm, P.; Ryde, U., Converging ligand-binding free energies obtained with free-energy perturbations at the quantum mechanical level. *J. Comput. Chem.* **2016**, *37*, 1589-1600.
68. Luzhkov, V.; Warshel, A., Microscopic models for quantum mechanical calculations of chemical processes in solutions: LD/AMPAC and SCAAS/AMPAC calculations of solvation energies. *J. Comput. Chem.* **1992**, *13*, 199-213.
69. Wesolowski, T.; Warshel, A., Ab Initio Free Energy Perturbation Calculations of Solvation Free Energy Using the Frozen Density Functional Approach. *J. Phys. Chem.* **1994**, *98*, 5183-5187.
70. Zheng, Y. J.; Merz, K. M., Mechanism of the human carbonic anhydrase II-catalyzed hydration of carbon dioxide. *Journal of the American Chemical Society* **1992**, *114*, 10498-10507.
71. Plotnikov, N.; Kamerlin, S. C. L.; Warshel, A., ParaDynamics: An Effective and Reliable Model for Ab Initio QM/MM Free Energy Calculations and Related Tasks. *J. Phys. Chem. B* **2011**, *115*, 7950-62.

72. Sun, Z.; Liu, Z., BAR-based Multi-dimensional Nonequilibrium Pulling for Indirect Construction of QM/MM Free Energy Landscapes: Varying the QM Region. *Adv. Theory Simul.* **2021**.
73. Ozer, G.; Valeev, E. F.; Quirk, S.; Hernandez, R., Adaptive Steered Molecular Dynamics of the Long-Distance Unfolding of Neuropeptide Y. *J. Chem. Theory Comput.* **2010**, 6, 3026-3038.
74. Ozer, G.; Quirk, S.; Hernandez, R., Adaptive steered molecular dynamics: Validation of the selection criterion and benchmarking energetics in vacuum. *J. Chem. Phys.* **2012**, 136, 215104.
75. Soares, C. M.; Bureau, H. R.; Merz, D. R.; Hershkovits, E.; Quirk, S.; Hernandez, R., Constrained Unfolding of a Helical Peptide: Implicit versus Explicit Solvents. *Plos One* **2015**, 10, e0127034.
76. Ozer, G.; Keyes, T.; Quirk, S.; Hernandez, R., Multiple branched adaptive steered molecular dynamics. *J. Chem. Phys.* **2014**, 141, 064101.
77. Ozer, G.; Quirk, S.; Hernandez, R., Thermodynamics of Decaalanine Stretching in Water Obtained by Adaptive Steered Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2012**, 8, 4837-4844.
78. Wang, X.; Xingzhao, T.; Boming, D.; John Z. H., Z.; Sun, Z., BAR-based Optimum Adaptive Steered MD for Configurational Sampling. *J. Comput. Chem.* **2019**, 40, 1270-1289.
79. Hummer, G.; Szabo, A., From the Cover: Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proceedings of the National Academy of Science* **2001**, 98, 3658-3661.
80. Hummer, G.; Szabo, A., Free Energy Reconstruction from Nonequilibrium Single-molecule Pulling Experiments. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 3658-3661.
81. Hummer, G.; Szabo, A., Free Energy Surfaces from Single-molecule Force Spectroscopy. *Accounts of chemical research* **2005**, 36, 504-513.
82. Paramore, S.; Ayton, G. S.; Voth, G. A., Extending the Fluctuation Theorem to Describe Reaction Coordinates. *J. Chem. Phys.* **2007**, 126, 051102.
83. Balsara; Stepaniants; Izrailev; Oono; Schulten, Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophysical Journal* **1997**, 73, 1281.
84. Marsili, S.; Procacci, P., Free energy reconstruction in bidirectional force spectroscopy experiments: The effect of the device stiffness. *J. Phys. Chem. B* **2010**, 114, 2509-2516.
85. Wang, X.; Sun, Z., Determination of Base Flipping Free Energy Landscapes from Nonequilibrium Stratification. *J. Chem. Inf. Model.* **2019**, 59, 2980-2994.
86. Fass, J.; Sivak, D. A.; Crooks, G. E.; Beauchamp, K. A.; Leimkuhler, B.; Chodera, J. D., Quantifying configuration-sampling error in Langevin simulations of complex molecular systems. *Entropy* **2018**, 20, 318.
87. Sivak, D. A.; Chodera, J. D.; Crooks, G. E., Time Step Rescaling Recovers Continuous-Time Dynamical Properties for Discrete-Time Langevin Integration of Nonequilibrium Systems. *J. Phys. Chem. B* **2014**, 118, 6466-6474.
88. Sivak, D. A.; Chodera, J. D.; Crooks, G. E., Using Nonequilibrium Fluctuation Theorems to Understand and Correct Errors in Equilibrium and Nonequilibrium Simulations of Discrete Langevin Dynamics. *Phys. Rev. X* **2013**, 3, 011007.
89. Sun, Z., A Benchmark Test on the Leapfrog Integrator and its Middle Alternative. [10.26434/chemrxiv.13727839.v1](https://arxiv.org/abs/10.26434/chemrxiv.13727839.v1) **2021**.
90. Jarzynski, C., Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **1997**, 56, 5018-5035.
91. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, 3, 300.
92. Pearlman, D. A.; Kollman, P. A., The lag between the Hamiltonian and the system configuration in free energy perturbation calculations. *Journal of Chemical Physics* **1989**, 91, 7831-7839.
93. Straatsma, T. P.; Mccammon, J. A., Treatment of rotational isomers in free energy calculations. II. Molecular dynamics simulation study of 18-crown-6 in aqueous solution as an example of systems with large numbers of rotational isomeric states. *J. Chem. Phys.* **1989**, 91, 3631-3637.
94. Rodriguez, A.; Tsallis, C., A generalization of the cumulant expansion. Application to a scale-invariant probabilistic model. *Journal of mathematical physics* **2010**, 51, 073301.
95. Hummer, G., Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* **2001**, 114, 7330-7337.
96. Kubo, R., Generalized cumulant expansion method. *Journal of the Physical Society of Japan* **1962**, 17, 1100-1120.

97. Sun, Z.; Wang, X.; Zhang, J. Z. H.; He, Q., Sulfur-substitution-induced base flipping in the DNA duplex. *Phys. Chem. Chem. Phys.* **2019**, *21*, 14923-14940.
98. Huai, Z.; Shen, Z.; Sun, Z., Binding Thermodynamics and Interaction Patterns of Inhibitor-Major Urinary Protein-I Binding from Extensive Free-Energy Calculations: Benchmarking AMBER Force Fields. *J. Chem. Inf. Model.* **2021**, *61*, 284-297.
99. Sun, Z.; Wang, X.; Zhang, J. Z., Determination of Binding Affinities of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Inhibitors from Free Energy calculation. *Chem. Phys. Lett.* **2019**, *723*, 1-10.
100. Sun, Z.; Wang, X.; Song, J., Extensive Assessment of Various Computational Methods for Aspartate's pKa Shift. *J. Chem. Inf. Model.* **2017**, *57*, 1621-1639.
101. Sun, Z.; Wang, X.; Zhao, Q.; Zhu, T., Understanding Aldose Reductase-Inhibitors interactions with free energy simulation. *Journal of Molecular Graphics and Modelling* **2019**, *91*, 10-21.
102. Wang, X.; Sun, Z., Understanding PIM-1 kinase inhibitor interactions with free energy simulation. *Phys. Chem. Chem. Phys.* **2019**, *21*, 7544-7558.
103. Paliwal, H.; Shirts, M. R., A Benchmark Test Set for Alchemical Free Energy Transformations and Its Use to Quantify Error in Common Free Energy Methods. *J. Chem. Theory Comput.* **2011**, *7*, 4115-34.
104. Wang, X.; Tu, X.; Zhang, J. Z. H.; Sun, Z., BAR-based Optimum Adaptive Sampling Regime for Variance Minimization in Alchemical Transformation: The Nonequilibrium Stratification. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2009-2021.
105. Sun, Z. X.; Wang, X. H.; Zhang, J. Z. H., BAR-based Optimum Adaptive Sampling Regime for Variance Minimization in Alchemical Transformation. *Phys. Chem. Chem. Phys.* **2017**, *19*, 15005-15020.
106. Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P., RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J. Comput. Chem.* **2006**, *27*, 1101-1111.
107. 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.
108. Pastor, R. W.; Brooks, B. R.; Szabo, A., An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **1988**, *65*, 1409-1419.
109. Case, D. A.; Cheatham, T. E.; Tom, D.; Holger, G.; Luo, R.; Merz, K. M.; Alexey, O.; Carlos, S.; Bing, W.; Woods, R. J., The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668-1688.
110. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G., GAUSSIAN09, Gaussian, Inc., Wallingford, CT, USA,(2009). *Google Scholar* **2016**.
111. Lameira, J. S.; Kupchenko, I.; Warshel, A., Enhancing Paradynamics for QM/MM Sampling of Enzymatic Reactions. *J. Phys. Chem. B* **2016**, *120*, 2155.
112. Plotnikov, N. V.; Warshel, A., Exploring, refining, and validating the paradynamics QM/MM sampling. *J. Phys. Chem. B* **2012**, *116*, 10342-10356.
113. Hudson, P. S.; Boresch, S.; Rogers, D. M.; Woodcock, H. L., Accelerating QM/MM free energy computations via intramolecular force matching. *J. Chem. Theory Comput.* **2018**, *14*, 6327-6335.
114. Pinnick, E. R.; Calderon, C. E.; Rusnak, A. J.; Wang, F., Achieving fast convergence of ab initio free energy perturbation calculations with the adaptive force-matching method. *Theor. Chem. Acc.* **2012**, *131*, 1-11.
115. Morado, J.; Mortenson, P. N.; Verdonk, M. L.; Ward, R. A.; Essex, J. W.; Skylaris, C.-K., ParaMol: A Package for Automatic Parameterization of Molecular Mechanics Force Fields. *J. Chem. Inf. Model.* **2021**, *61*, 2026-2047.
116. Ercolessi, F.; Adams, J. B., Interatomic potentials from first-principles calculations: the force-matching method. *EPL (Europhysics Letters)* **1994**, *26*, 583.
117. 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.

Table 1. Mean absolute error of the indirect RM1 estimates in kcal/mol. 8 different initial seeds are used to initiate the multi-dimensional ASMD pulling. The pulling speed in the configurational space is 10 ps per 3° segment. Two pulling speeds including a change of 0.1 per time step and 0.01 per time step are tested in the alchemical space. The errors for the averaged indirect estimates over the last 5 initial configurations are also shown. We can see that the systematic errors of the indirect estimates are reduced upon increased pulling time in the alchemical space. Also, the seeds-averaged results are significantly improved over independent trails.

Hamiltonians Trail	0.1 to RM1		0.01 to RM1	
	AM1	MNDO	AM1	MNDO
1	0.44	0.28	0.23	0.07
2	0.37	0.43	0.13	0.15
3	0.43	0.41	0.13	0.13
4	0.37	0.70	0.11	0.14
5	0.27	0.22	0.23	0.12
6	0.79	1.13	-	-
7	0.35	0.37	-	-
8	0.21	0.34	-	-
average over 5 trails	0.11	0.30	0.08	0.10

Table 2. Efficiency comparison of direct and indirect free energy simulations under the HF Hamiltonian. The calculation of the computational cost is similar to the case of our previous BAR-based method. There are N_{segments} segments in stratified pulling and the number of realizations in each segment is N_{traj} . The total simulation time in the direct scheme is given by $N_{\text{segments}} * N_{\text{traj}} * (\phi_{\text{NEW}} + \phi_{\text{eq}})$, while the total simulation time in the indirect scheme is the sum of $N_{\text{segments,SQM}} * N_{\text{traj,SQM}} * (\phi_{\text{NEW,SQM}} + \phi_{\text{eq,SQM}})$ at SQM level and $N_{\text{traj,SQM}\rightarrow\text{QM}} * (\phi_{\text{NEW,SQM}\rightarrow\text{QM}} + \phi_{\text{eq,SQM}})$ in the SQM-to-QM correction. As the selection criterion eliminates the equilibrium sampling for initial configurations, ϕ_{eq} becomes 0. The simulation time at QM level is scaled by the ratio of computational cost under QM Hamiltonian and that under SQM Hamiltonian in Table S1 to provide the effective simulation time at SQM level, enabling the direct comparison between computational costs of different levels of theory. Note that the sample size for the correction term is the sum of 5 repeats of the 10-step SQM-to-QM transformation, corresponding to the sample size of 25 for the SQM->HF perturbation. If the pulling time of 100 steps is used for the alchemical transformation, the speedup of the indirect method would be about 200 folds.

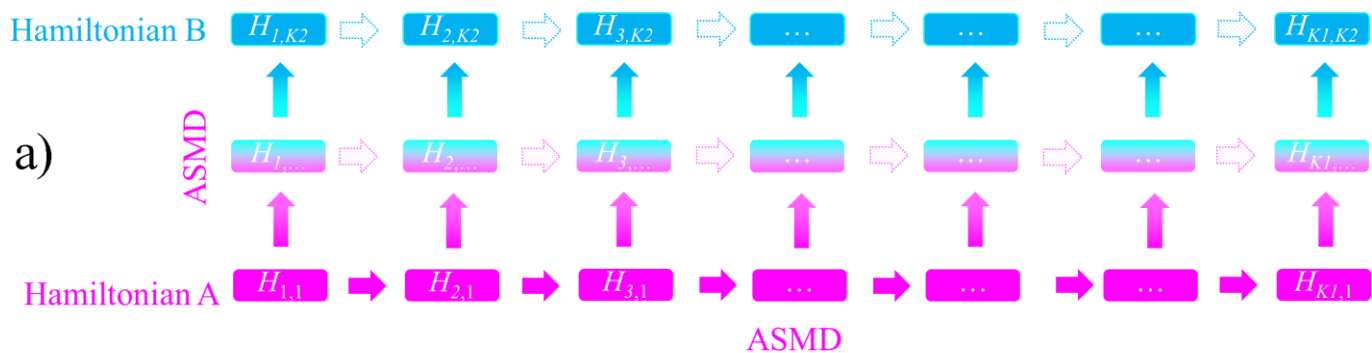
Terms Simulation	ϕ_{eq} for each initial configuration (ps)	ϕ_{NEW} in each segment (ps)	Number of segments	Number of realizations per segment	Total simulation time (ps) scaled to SQM Hamiltonian	Relative efficiency
direct SQM	0	10	120	20	24000.00	3343.64
SQM->HF	0	0.005	120	25	50154.60	-
HF->SQM	-	-	-	-	0.00	-
indirect HF	-	-	-	-	74154.60	1082.16
direct HF	0	10	120	20	80247357.45	1.00

Table 3. Efficiency comparison of direct and indirect free energy simulations under the SQM Hamiltonian.

There are N_{segments} segments in stratified pulling and the number of realizations in each segment is N_{traj} . The total simulation time in the direct scheme is given by $N_{\text{segments}} * N_{\text{traj}} * (\phi_{\text{NEW}} + \phi_{\text{eq}})$, while the total simulation time in the indirect scheme is the sum of $N_{\text{segments,SQM}} * N_{\text{traj,SQM}} * (\phi_{\text{NEW,SQM}} + \phi_{\text{eq,SQM}})$ at MM level and $N_{\text{traj,SQM} \rightarrow \text{QM}} * (\phi_{\text{NEW,SQM} \rightarrow \text{QM}} + \phi_{\text{eq,SQM}})$ in the MM-to-SQM correction. As the selection criterion eliminates the equilibrium sampling for initial configurations, ϕ_{eq} becomes 0. The simulation time at SQM level is scaled by the ratio of computational cost under SQM Hamiltonian and that under MM Hamiltonian in Table S1 to provide the effective simulation time at MM level, enabling the direct comparison between computational costs of different levels of theory. Note that the sample size for the correction term is the sum of 5 repeats of the 100-step MM-to-SQM transformation, corresponding to the sample size of 25 for the MM->SQM perturbation. If the pulling time of 1000 steps (i.e. $\phi_{\text{NEW}} = 5$ ps) is used for the alchemical transformation, the relative efficiency of the indirect method would be about 1.5.

Simulation	ϕ_{eq} for each initial configuration (ps)	ϕ_{NEW} in each segment (ps)	Number of segments	Number of realizations per segment	Total simulation time (ps) scaled to MM Hamiltonian	Relative efficiency
direct MM	0	10	120	20	24000.00	40.39
MM->SQM	0	0.5	120	25	60578.06	-
MM->SQM	-	-	-	-	0.00	-
indirect SQM	-	-	-	-	84578.06	11.46
direct SQM	0	10	120	20	969249.03	1.00

Fig. 1. a) An illustration of the Hamiltonian perturbation framework via EXP-based multi-dimensional nonequilibrium free energy calculations. The Hamiltonian perturbation is performed between neighboring states with nonequilibrium transformations. The Hamiltonian of the system at the k_1 th configurational state and k_2 th alchemical state is represented as H_{k_1,k_2} . The target free energy landscape is at the Hamiltonian state $K_2 = 2$. The indirect scheme performs direct free energy simulations at the Hamiltonian state $k_2 = 1$ to explore the configurational space, and adds the $H_{k_1,1}$ -to- $H_{k_1,2}$ correction term to perturb the thermodynamic profile to the result at the target Hamiltonian state. The unidirectional arrows represent unidirectional pulling and unidirectional reweighting of the EXP or JI estimator. Only the transformations described with solid arrows are performed due to efficiency considerations in nonequilibrium free energy simulations. b) The thermodynamic cycle describing the dihedral flipping process with different descriptions of the ACE-NME system. The reaction coordinate is the backbone C-C-N-C dihedral. The free energy simulation exploring the configurational space is performed at some computationally efficient levels (i.e., model A), and the A-to-B perturbation term is used to obtain the thermodynamic profiles at the target levels (i.e., model B). An illustration of the ASMD selection criterion is shown at the center of the thermodynamic cycle. The configuration with the microscopic work closest to the EXP estimate is considered as an equilibrium configuration and thus is used as the initial configuration for the next pulling segment along the configurational CV and also to initiate the alchemical pulling transformation.



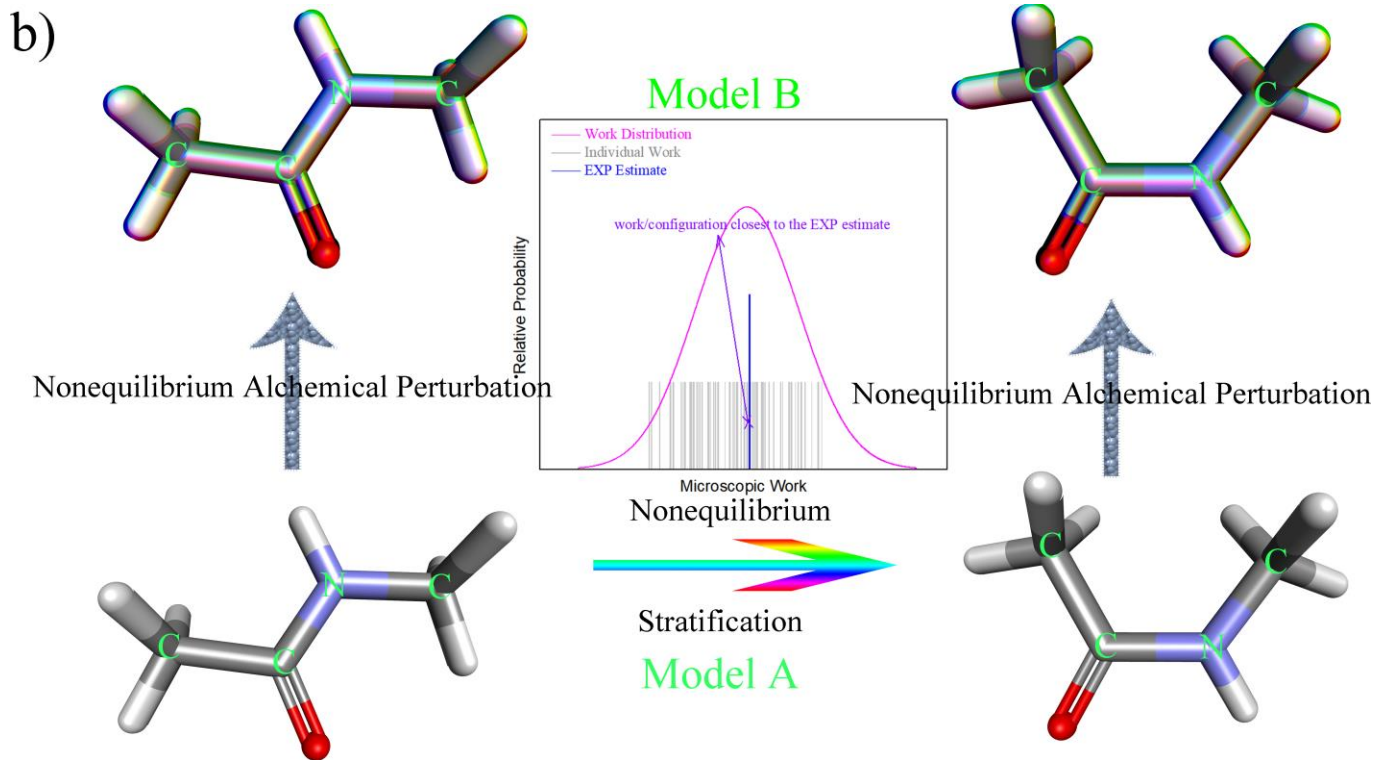


Fig. 2. a) The free energy profiles in the direct free energy simulation under the RM1 Hamiltonian with different pulling speeds. The number in the legend represents the number of MD steps for each nonequilibrium trajectory in each segment. As we are using 0.5 fs time steps, the pulling times of 10000 steps, 20000 steps and 30000 steps are equivalently 5 ps, 10 ps and 15 ps. The bias in the 5 ps/segment estimate shows seed-dependence, and the pulling speed of 10 ps per segment is sufficiently slow for converged estimates of the free energy profiles. b) The indirect estimates at the RM1 level with different sample sizes for the correction term in one seeding SMD trail. The alchemical order parameter is varied in 10 MD steps with a change of 0.1 per step. We can see that the 5-sample, 10-sample and 50-sample estimates with different estimators are virtually identical.

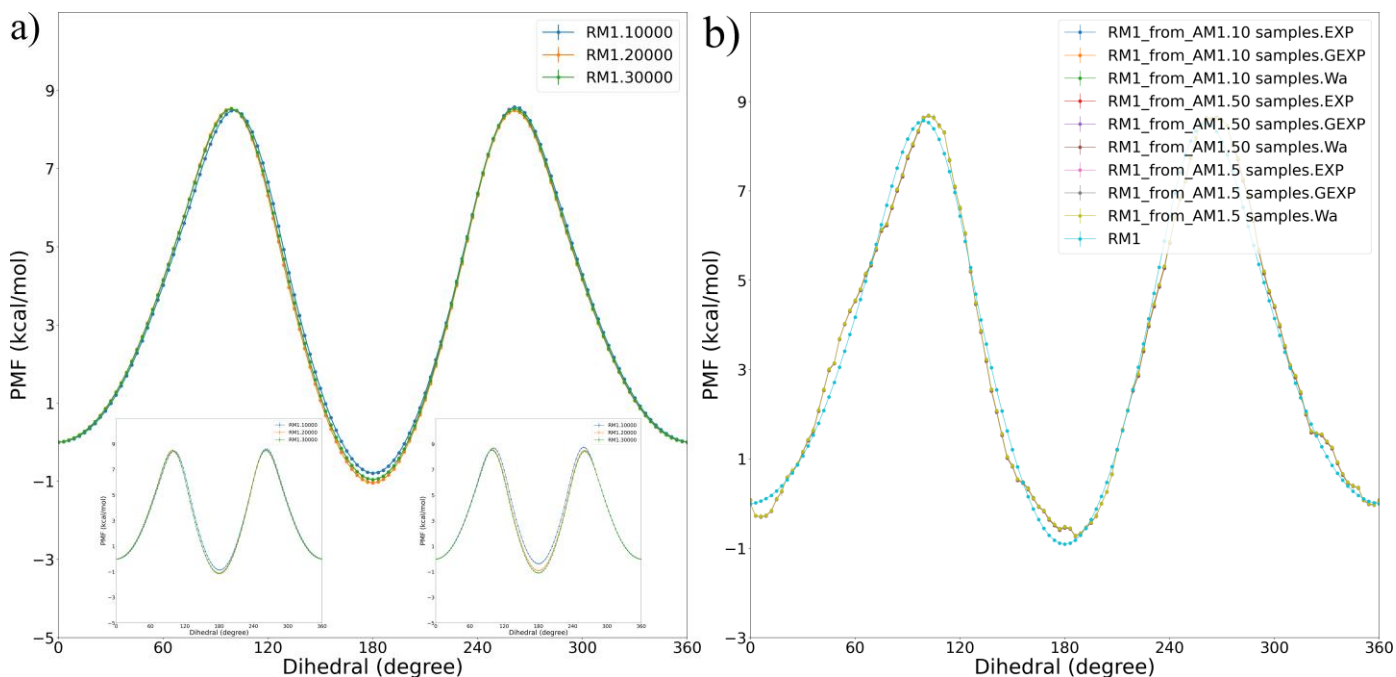


Fig. 3. The statistical uncertainty for the correction terms. a) The multi-dimensional SMD simulation is initiated from the 1st configuration with the pulling speed of 10 ps/segment. The AM1, MNDO and PM6 Hamiltonians are used to explore the configurational space, and the correction term is used to perturb the results to the RM1 level. As the differences between different SQM Hamiltonians and the target RM1 are different, the statistical errors for different SQM-to-RM1 corrections obtained from the same estimator are different. The statistical error of the GEXP estimator is extremely similar to that of the ordinary average Wa, due to the negligible dissipation or the width of the work distribution. As a result, the GEXP and the Wa estimates are virtually identical. b) The AM1-to-RM1 correction averaged over 5 initial configurations. The pulling speed of 0.1/step is used along the alchemical CV, while two pulling speeds of 10 ps/segment and 15 ps/segment are used along the configurational CV. We can see that the statistical error shows little dependence on the pulling speed in the configurational space.

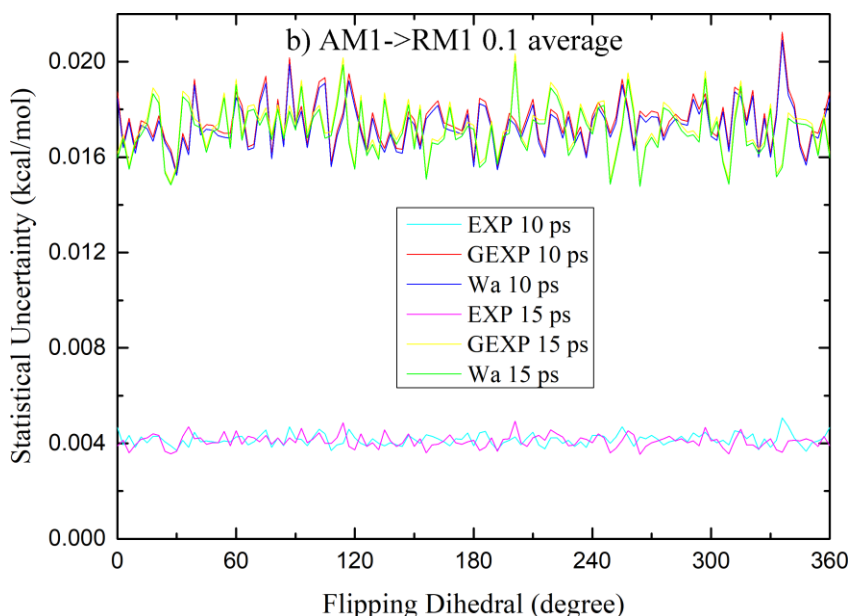
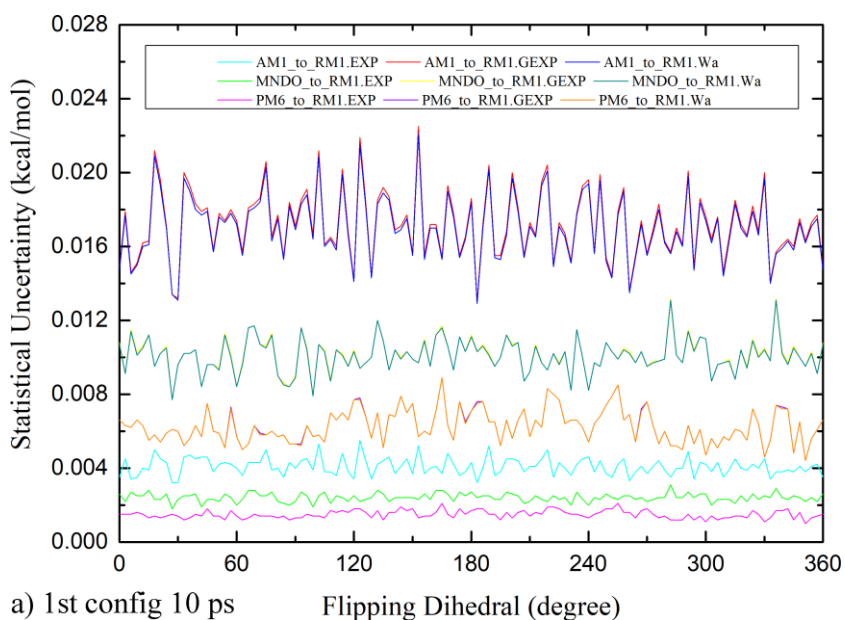


Fig. 4. Comparison between the direct and indirect results at the RM1 level with the pulling speed of 10 ps per 3° segment for nonequilibrium pulling in the configurational space. a) and c) are obtained from the same initial configuration to initiate the multi-dimensional ASMD simulation, while b) and d) are the averaged results over 5 initial configurations. For a) and b), the alchemical order parameter is varied in 10 MD steps with a change of 0.1 per step, while for c) and d), the alchemical CV is changed in 100 MD steps with a perturbation of 0.01 per time step. More detailed results for different initial configurations and pulling speeds are given in Fig. S2 and S3.

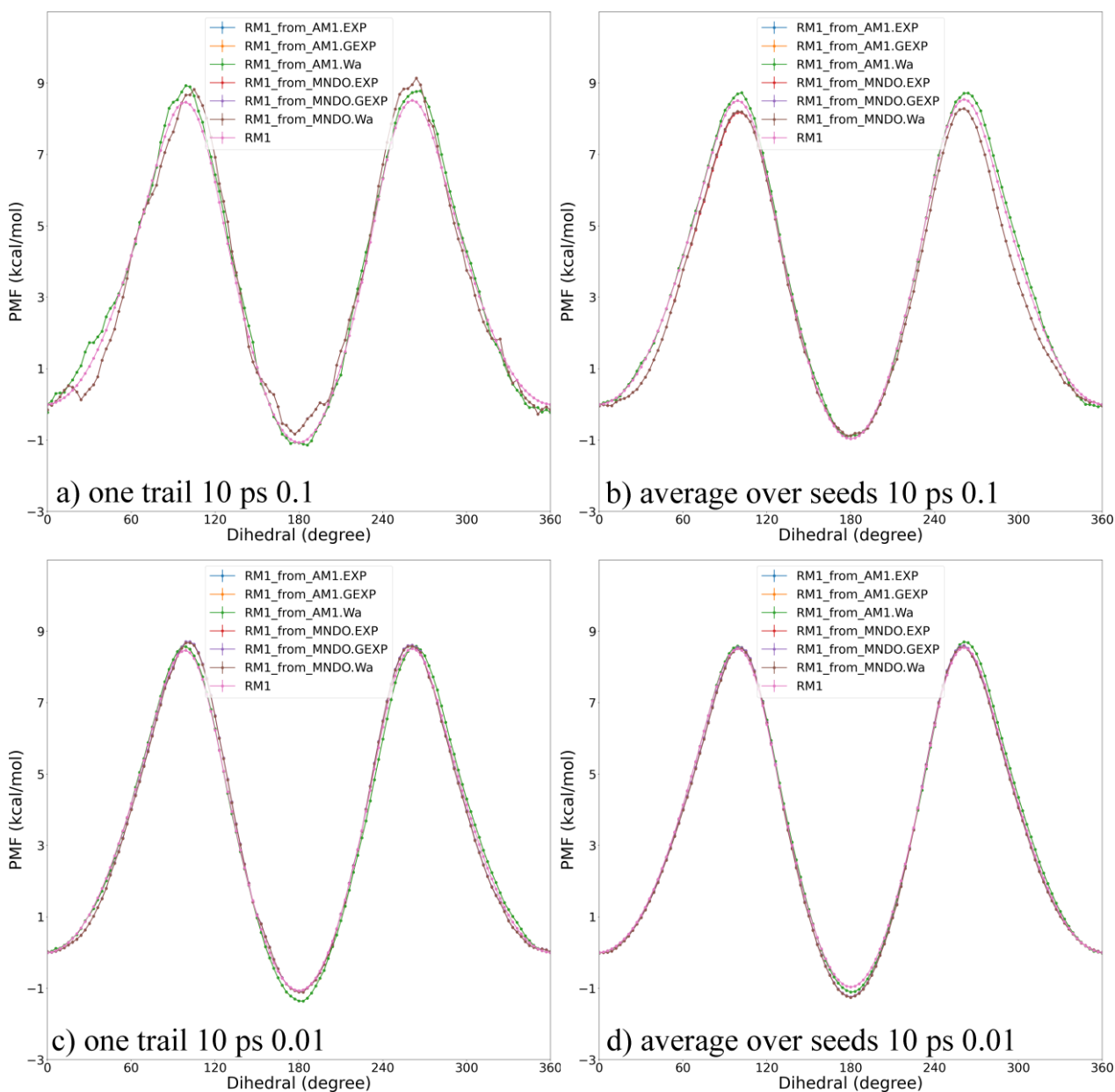
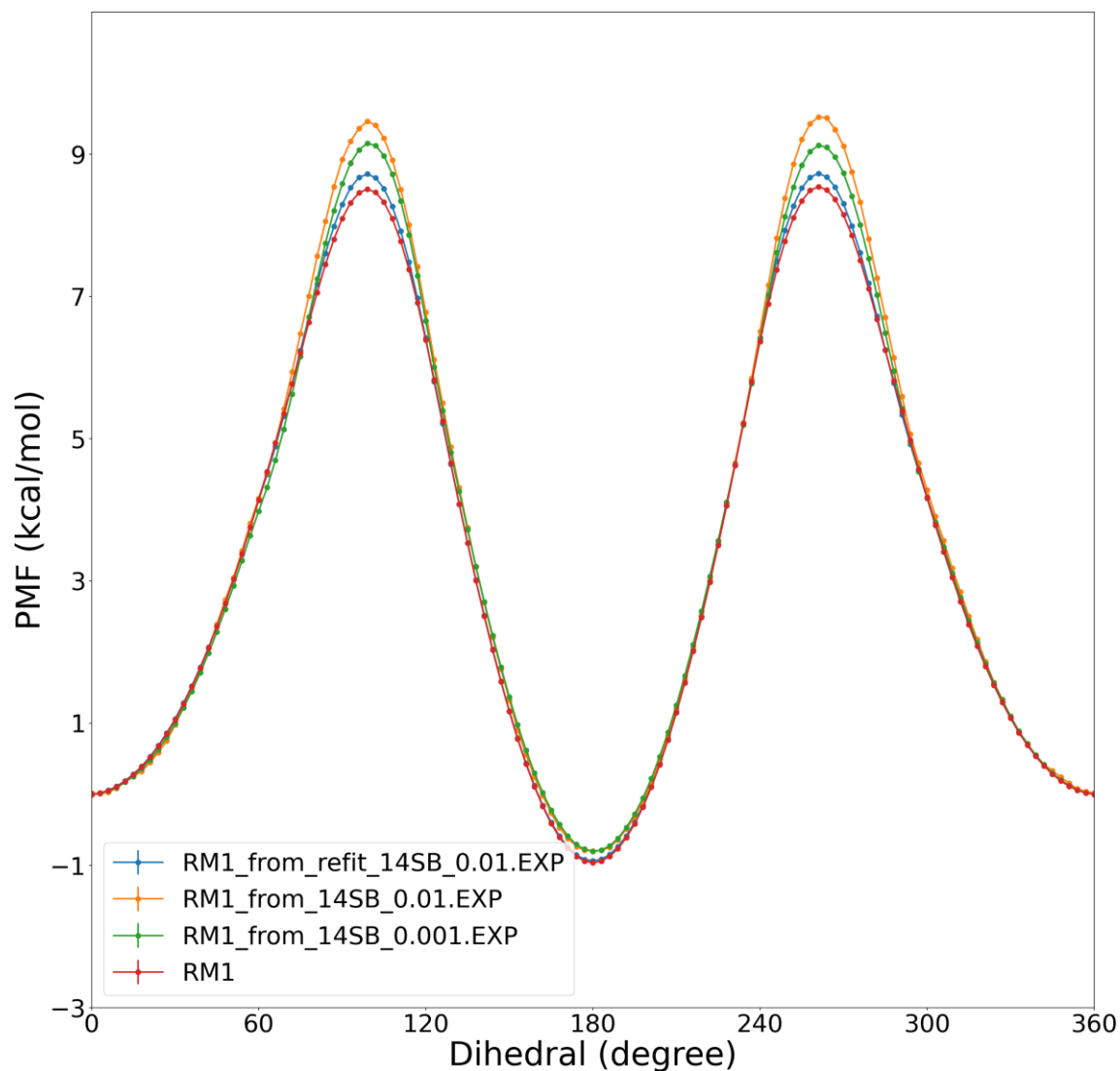


Fig. 5. Comparison between the direct and seeds-averaged indirect results at the RM1 level obtained from direct MM simulations and MM-to-RM1 transformations. The force fields used are AMBER14SB and its refitted version. The pulling speed along the configurational CV is set to 10 ps/segment, and the alchemical order parameter is varied in 100 or 1000 time steps.



**Supporting Information: Seeding the Multi-dimensional Nonequilibrium Pulling
for Hamiltonian Variation: Indirect Nonequilibrium Free Energy Simulations at
QM levels**

Zhaoxi Sun^{1*} and Qiaole He²

*¹Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Institute of
Theoretical and Computational Chemistry, Peking University, Beijing 100871, China*

*²AI Department of Enzymaster (Ningbo) Bio-Engineering Co., Ltd., North Century Avenue 333, 315100 Ningbo,
China*

*To whom correspondence should be addressed: z.sun@pku.edu.cn

Table S1. The timing information (in ns/day) of MM, SQM and QM calculations on a computing node with 384 GB memory and the CPU used is Intel(R) Xeon(R) Platinum 8160. A single core is used for this benchmark, which avoids the influence of parallelization-related issues. As the nonequilibrium trajectories are independent, the practical simulation speed is the same as these statistics. As the memory on each computing node is very large, no memory-related issue in QM calculations would influence the speed of calculation. The basis set of 6-31G* is employed in ab initio QM calculations. Different SQM simulations are of very similar computational costs, while the speed of ab initio QM calculations depends on the level of theory. Here, we only tested the HF Hamiltonian.

Hamiltonian Terms	SQM	HF	MM
speed(ns/day)	46.5	0.013907	1877.92
speedup(SQM/QM)	1.00	3343.64	0.02

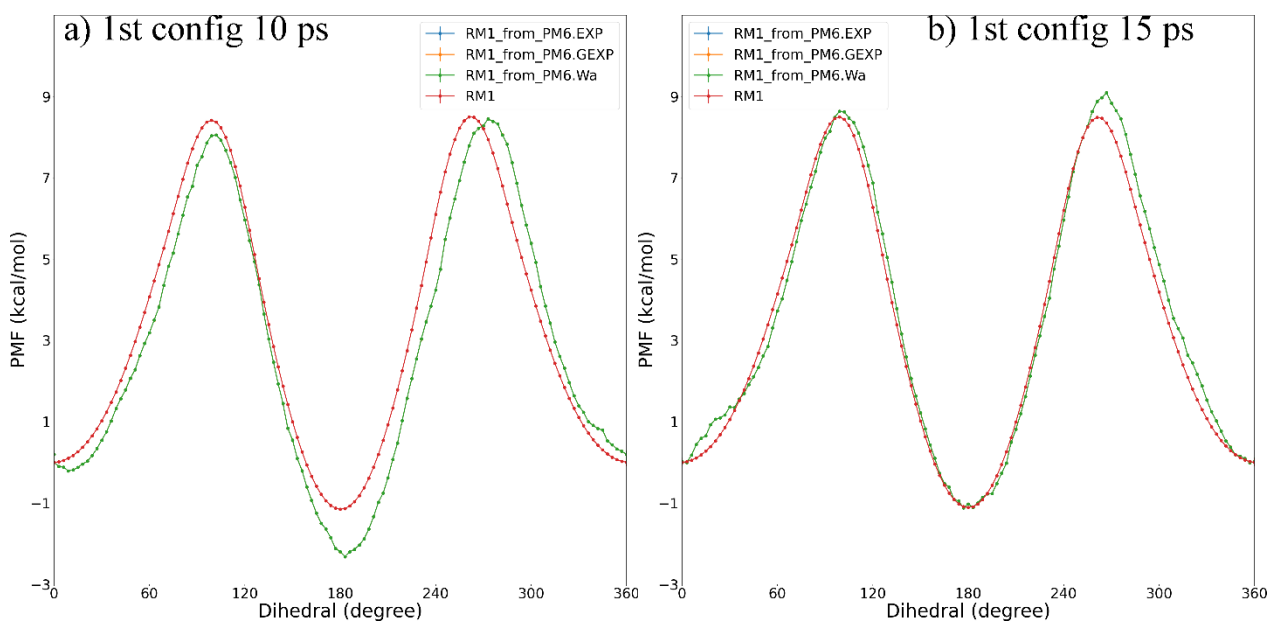
Table S2. Mean absolute error of the indirect estimates at ab initio QM levels in kcal/mol. 7 different initial configurations are tested to initiate the multi-dimensional ASMD pulling. The pulling speed in the configurational space is 10 ps per 3° segment. The alchemical perturbation is finished in 10 time steps with a change of 0.1 per step in the first 7 trails, while the last trail 0.01 uses 100 steps.

Hamiltonians Trail	HF		
	AM1	MNDO	RM1
1	0.31	0.35	0.60
2	0.83	0.79	1.33
3	0.88	0.54	0.93
4	0.56	0.91	0.75
5	0.50	0.44	0.67
6	0.81	1.21	0.47
7	0.52	1.23	0.67
0.01	0.20	0.30	0.21

Table S3. Efficiency comparison of direct and indirect nonequilibrium free energy simulations with the current JI-based ASMD method and the previously proposed BAR-based method. The total simulation time in direct scheme is given by $N_{\text{segments}} * N_{\text{traj}} * (\phi_{\text{NEW}} + \phi_{\text{eq}})$, while the total simulation time in the indirect scheme is the sum of $N_{\text{segments,small}} * N_{\text{traj,small}} * (\phi_{\text{NEW,small}} + \phi_{\text{eq,small}})$ under the SQM Hamiltonian and $N_{\text{traj,small}\rightarrow\text{large}} * (\phi_{\text{NEW,small}\rightarrow\text{large}} + \phi_{\text{eq,small}}) + N_{\text{traj,large}\rightarrow\text{small}} * (\phi_{\text{NEW,large}\rightarrow\text{small}} + \phi_{\text{eq,large}})$ in the SQM-to-QM correction. N_{segments} is the number of segments and N_{traj} is the number of realizations per segment. The simulation time under the QM Hamiltonian is scaled by the ratio of computational cost QM/SQM in Table S1 to provide the effective simulation time at the SQM level, enabling the direct comparison between computational costs from methods. The statistics for the BAR-based method are borrowed from our previous work, Phys. Chem. Chem. Phys. 2019, 21, 21942-21959. As the BAR-based method is faster than ASMD in the construction of the free energy profiles in the configurational space, its computational cost of the direct free energy simulation (3.78 ns in the reference) is smaller than the ASMD one (24 ns). Therefore, the speedup of the multi-dimensional ASMD method is relatively modest compared with the ~1000-fold speedup shown in the previous table.

Terms Simulation	ϕ_{eq} for each initial configuration (ps)	ϕ_{NEW} in each segment (ps)	Number of segments	Number of realizations per segment	Total simulation time (ps) scaled to SQM Hamiltonian	Relative efficiency
direct SQM	0	10	120	20	24000.00	526.62
SQM->HF	0	0.005	120	25	50154.60	-
HF->SQM	-	-	-	-	0.00	-
indirect HF	-	-	-	-	74154.60	170.44
direct HF	0.05	0.5x2=1	180	20	12638958.80	1.00

Fig. S1. Comparison between the direct and indirect estimates of the free energy profiles at the RM1 level initiated from different seeds (i.e., configurations) with different pulling speeds along the configurational CV or the alchemical CV. The configurational space is explored with the ASMD scheme with 20 samples in each stage at the PM6 level, and the RM1 result is obtained by the combination of the PM6 result and the PM6-to-RM1 unidirectional pulling. The pulling time ‘x ps’ denotes the pulling time for each segment along the configurational CV. The nonequilibrium transformation in the alchemical space is performed in 5 fs (i.e., 10 time steps with a change of 0.1 per step) for the first 5 subplots, while a smaller perturbation (0.01 per time step) and thus a longer pulling time is used for the last subplot. The exponential average EXP, the Gaussian approximation GEXP, and the ordinary average Wa are extremely similar, leading to overlaps of these curves in the plot. We can see that different initial configurations could lead to different systematic errors with a faster pulling speed along the alchemical CV, which could be eliminated when a slower pulling speed is employed.



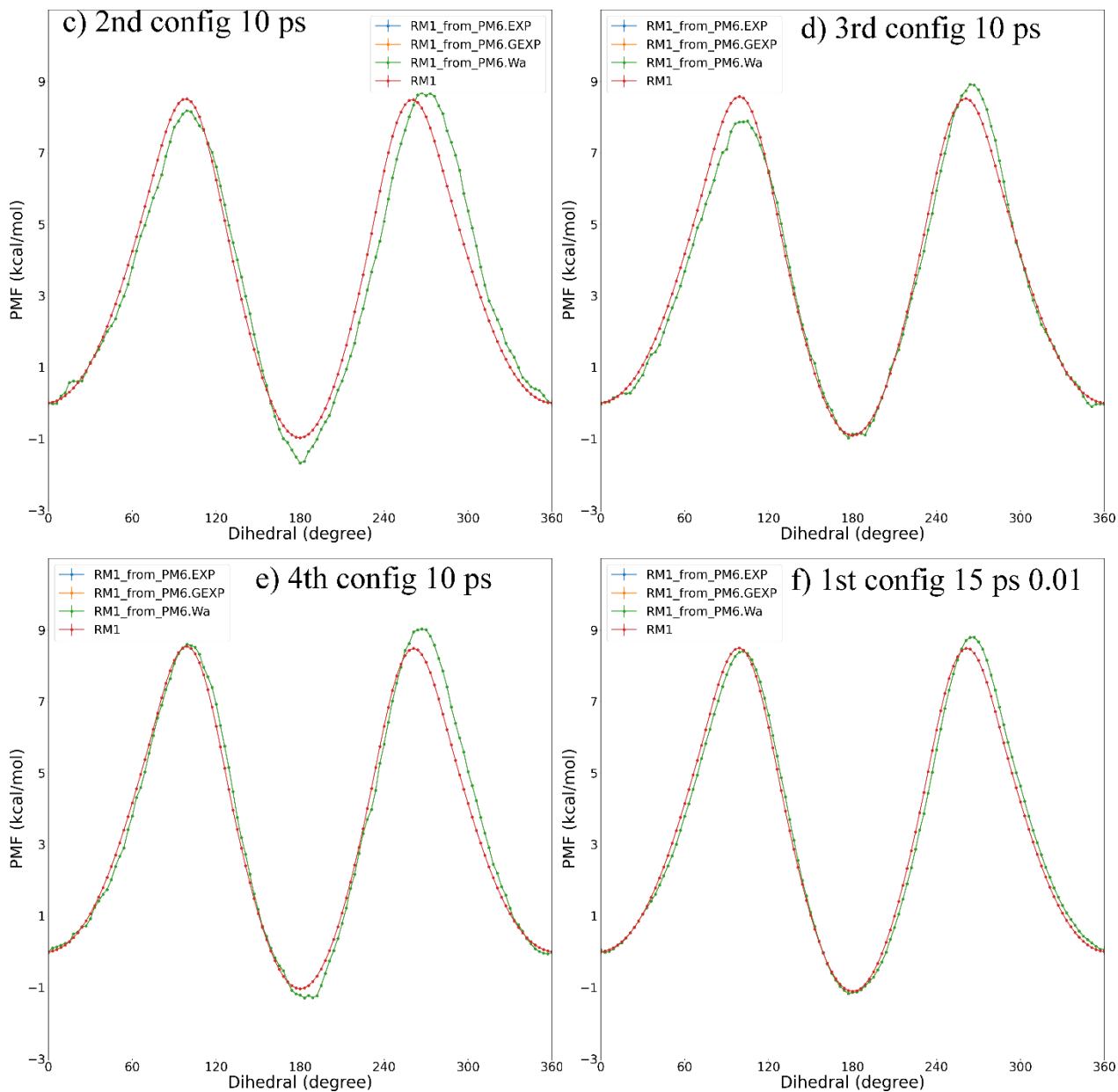
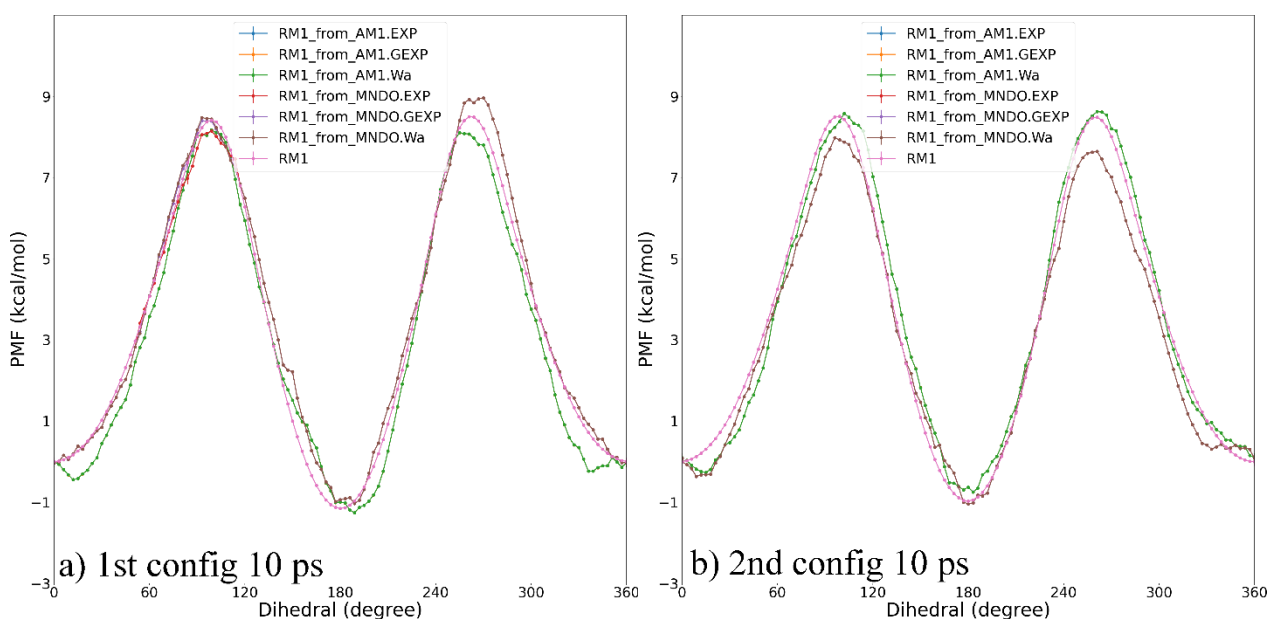
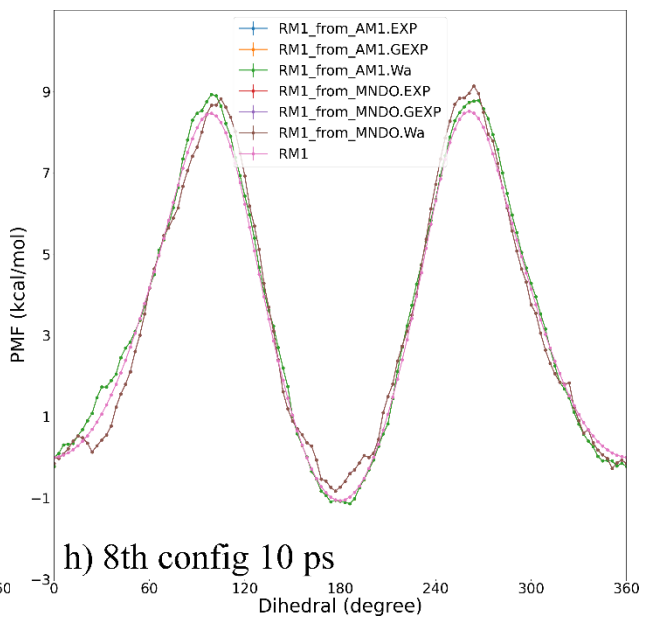
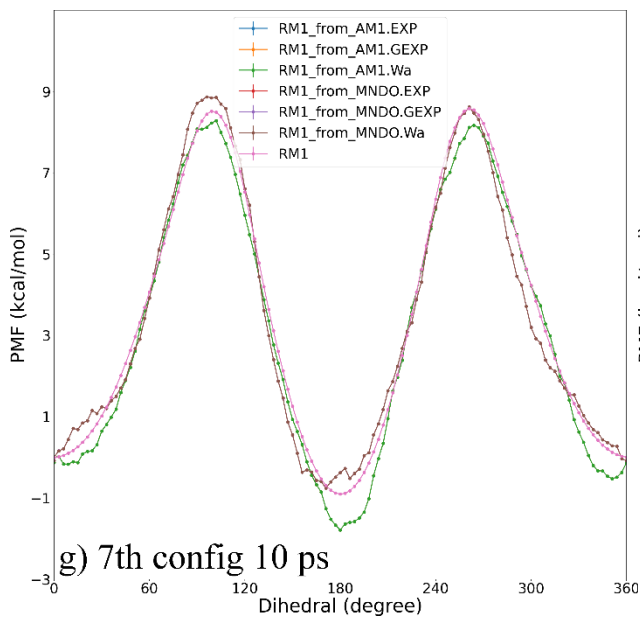
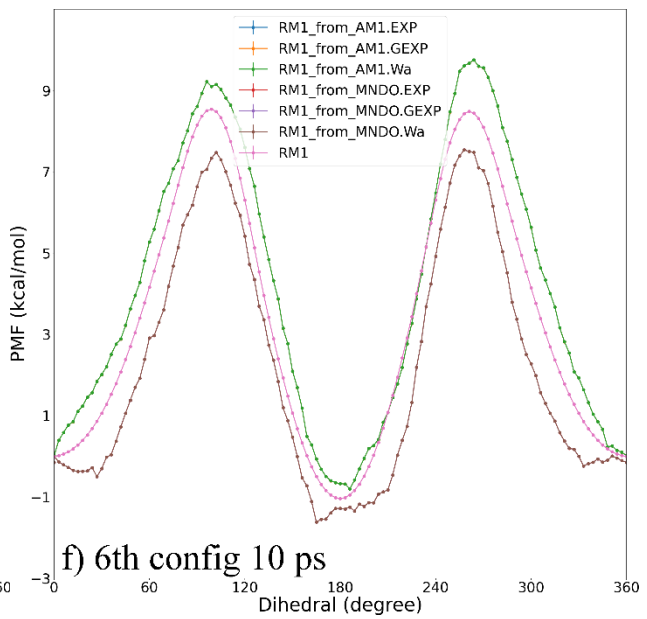
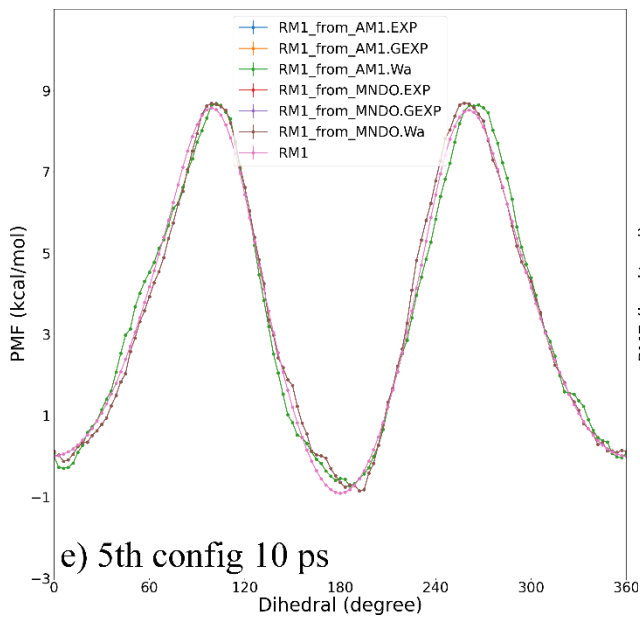
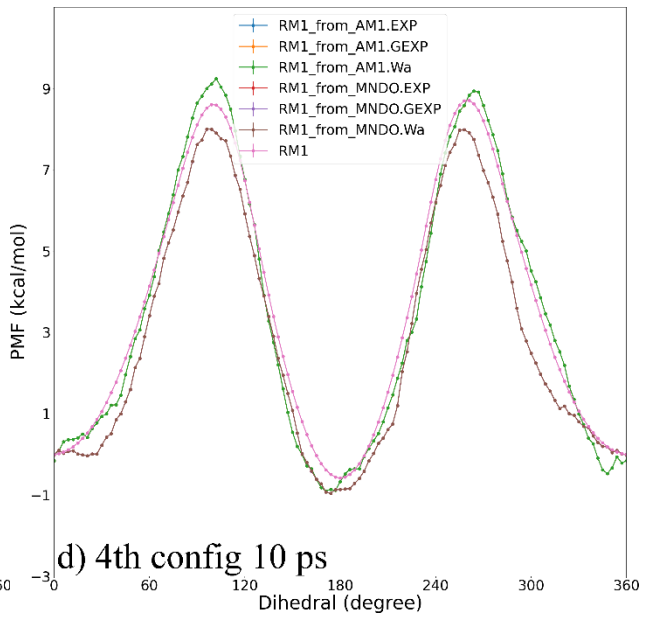
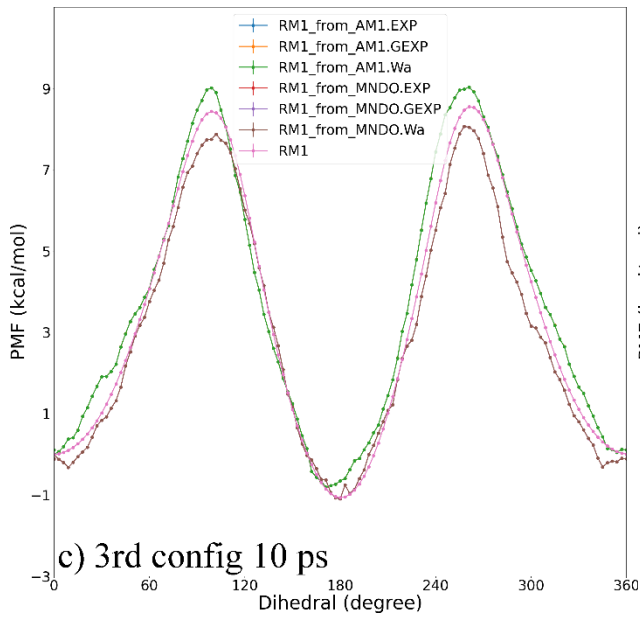
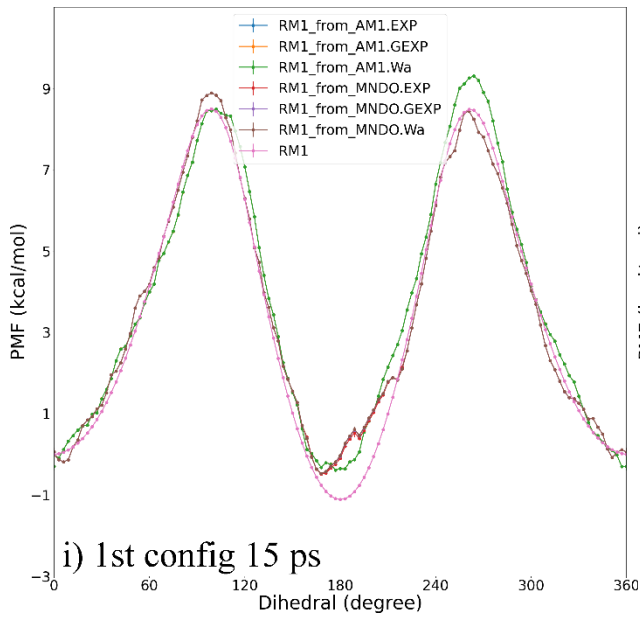


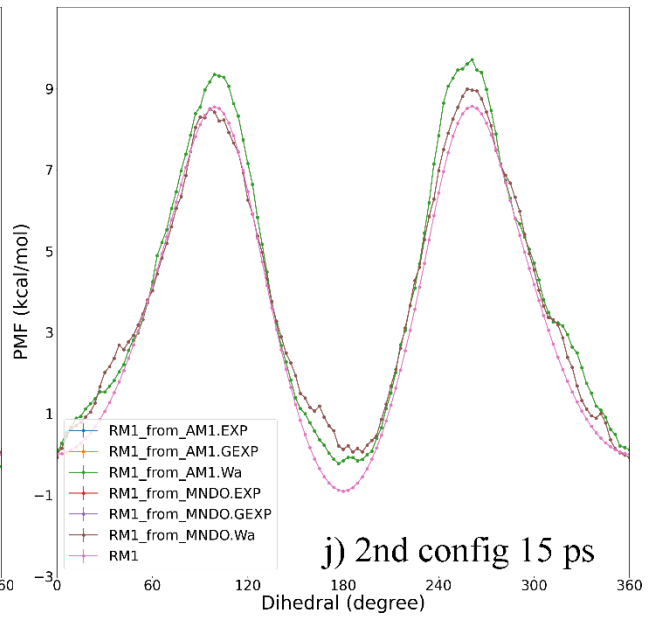
Fig. S2. Comparison between the direct and indirect estimates of the free energy profiles at the RM1 level initiated from different seeds (i.e., configurations) with different pulling speeds along the configurational CV. The nonequilibrium transformation in the alchemical space is performed in 5 fs (i.e., 10 time steps with a change of 0.1 per step). The configurational space is explored with the ASMD scheme with 20 samples in each stage under the AM1 or MNDO Hamiltonians, and the RM1 result is obtained by the combination of the AM1 or MNDO result and the AM1-to-RM1 or MNDO-to-RM1 unidirectional pulling. The pulling time ‘x ps’ denotes the pulling time for each segment along the configurational CV. 8 different initial configurations are used for the seeding SMD simulations with the pulling speed of 10 ps per segment in the first 8 subplots a-h), while the first 3 configurations are used to initiate the seeding SMD simulations with a slower pulling speed of 15 ps per segment for the last 3 subplots i-k). The exponential average EXP, the Gaussian approximation GEXP, and the ordinary average Wa are extremely similar, leading to overlaps of these curves in the plot. Different initial configurations introduce different systematic errors, and averaging over these configurations could eliminate this systematic error, as shown in the main article. The pulling speeds of 10 ps/segment and 15 ps/segment along the configurational CV have little influence on the systematic error of the indirect results, as the latter is mainly introduced in the alchemical perturbation term.



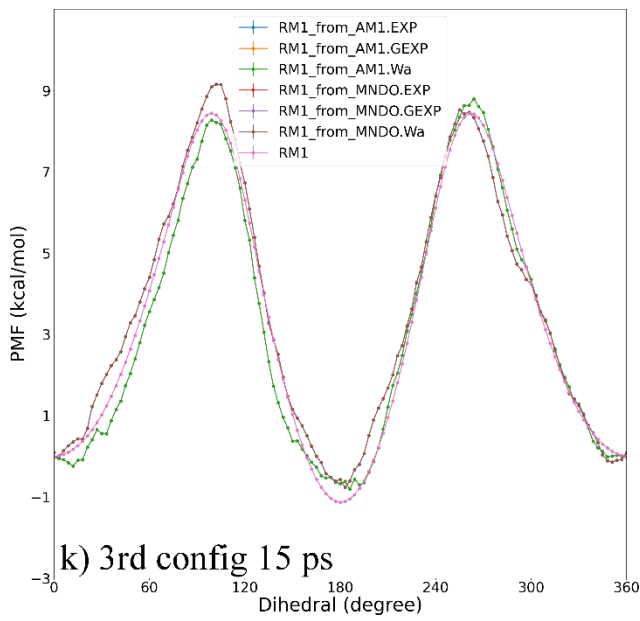




i) 1st config 15 ps

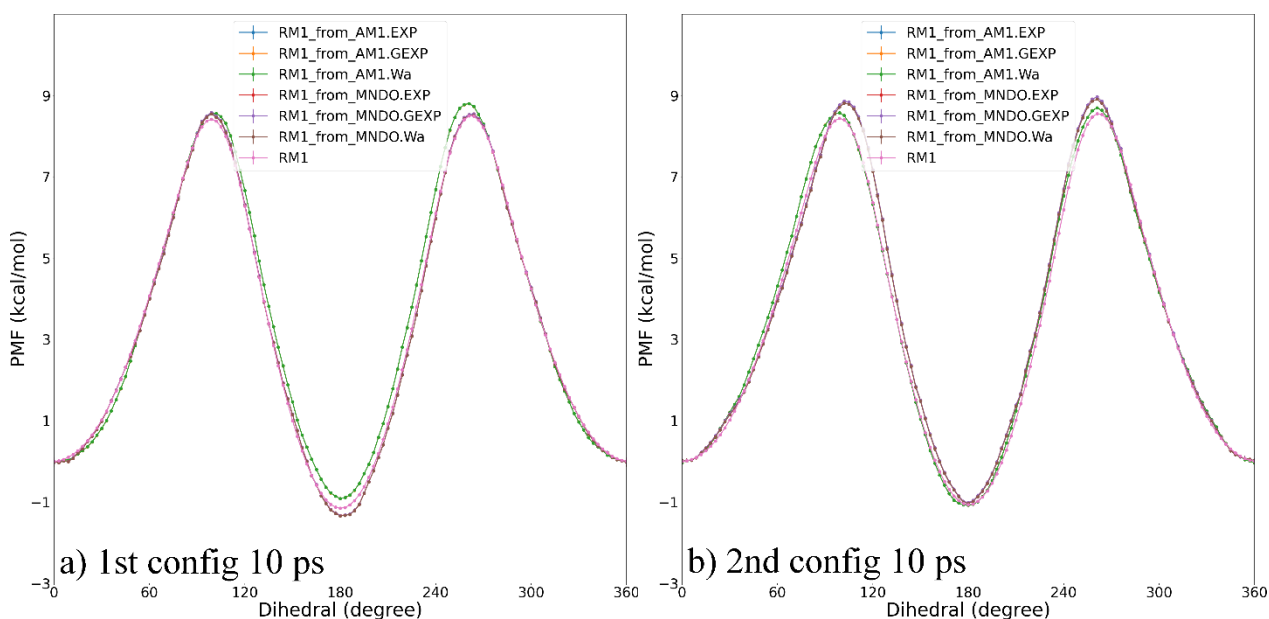


j) 2nd config 15 ps



k) 3rd config 15 ps

Fig. S3. Comparison between the direct and indirect estimates of the free energy profiles at the RM1 level initiated from different seeds (i.e., initial configurations) with different pulling speeds along the configurational CV. The nonequilibrium transformation in the alchemical space is performed in 50 fs (i.e., 100 time steps with a change of 0.01 per step). The configurational space is explored with the ASMD scheme with 20 samples in each stage under the AM1 or MNDO Hamiltonians, and the RM1 result is obtained by combining the AM1 or MNDO result and the AM1-to-RM1 or MNDO-to-RM1 unidirectional pulling. The pulling time ‘x ps’ denotes the pulling time for each segment along the configurational CV. Different initial configurations are used for the seeding SMD simulations with the pulling speed of 10 ps per segment for the first 4 subplots a-d), while the first 2 configurations are used to initiate the seeding SMD simulations with a slower pulling speed of 15 ps per segment for the last 2 subplots e-f). The exponential average EXP, the Gaussian approximation GEXP, and the ordinary average Wa are extremely similar, leading to overlaps of these curves in the plot. Compared with the indirect estimates obtained with a faster pulling speed along the alchemical CV shown in the previous figure, the current results are much closer to the direct free energy estimates, which indicates that the sampling in the nonequilibrium ensemble successfully eliminates the systematic bias introduced in the alchemical perturbation term. The pulling speed in the configurational space have little impact on the outcome.



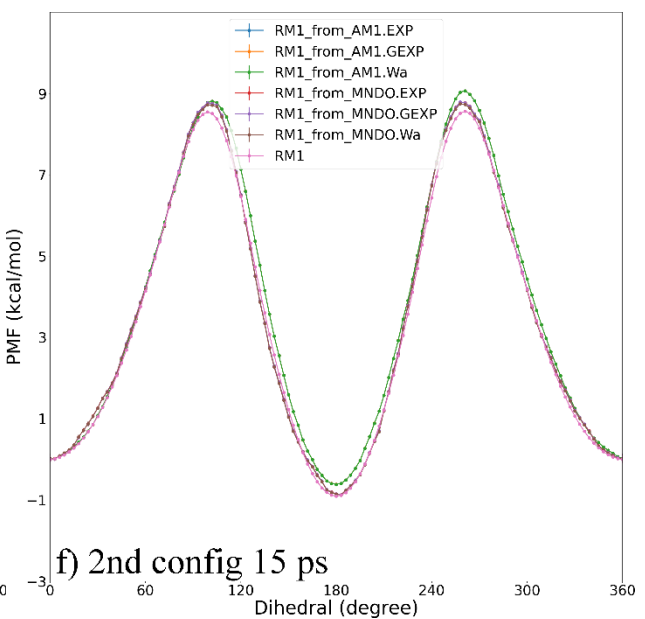
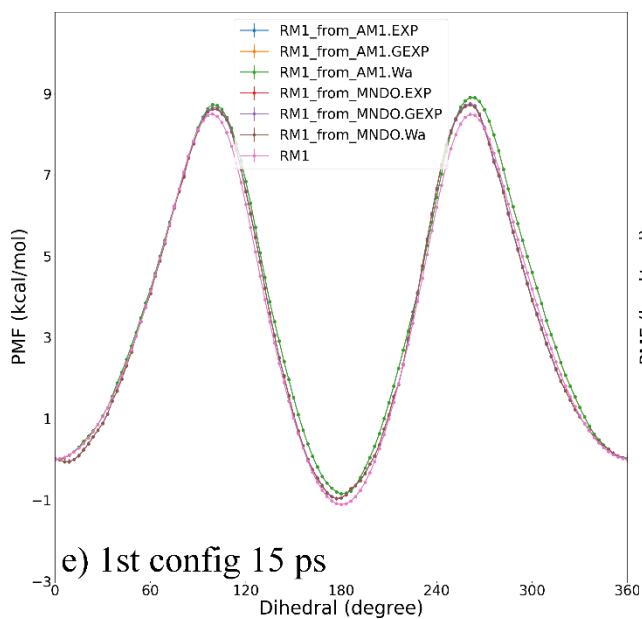
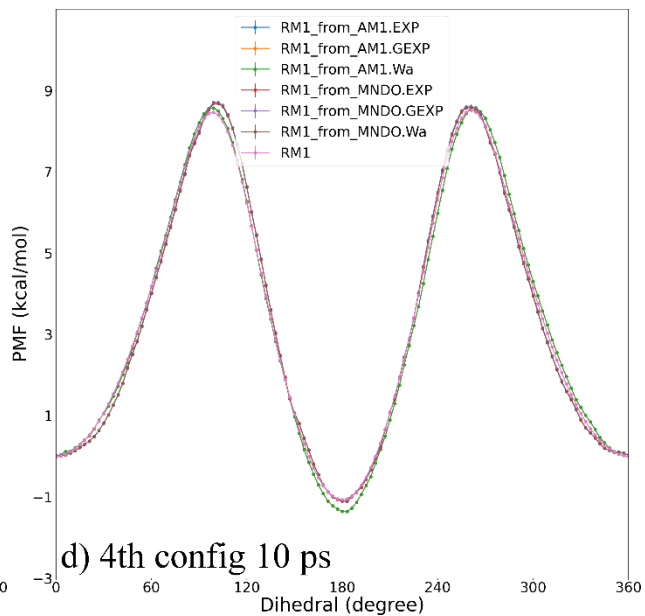
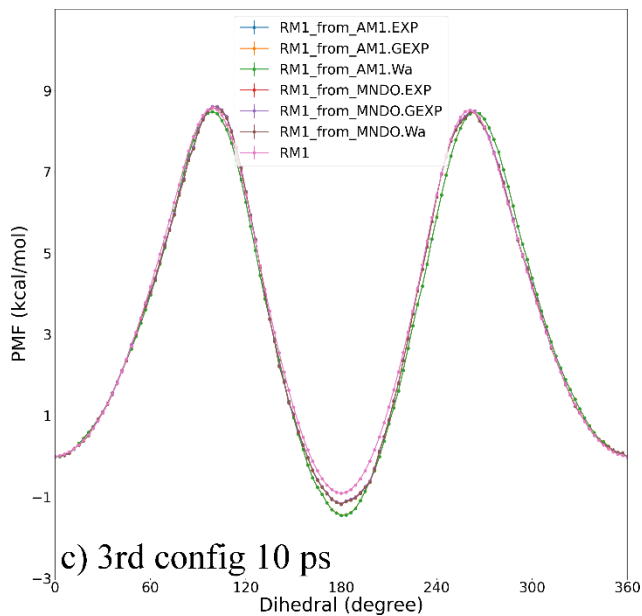


Fig. S4. The deviation of the indirect EXP, GEXP and Wa estimates from the direct result at the RM1 level for the initial-seed-averaged simulations with 100 steps in the alchemical transformation.

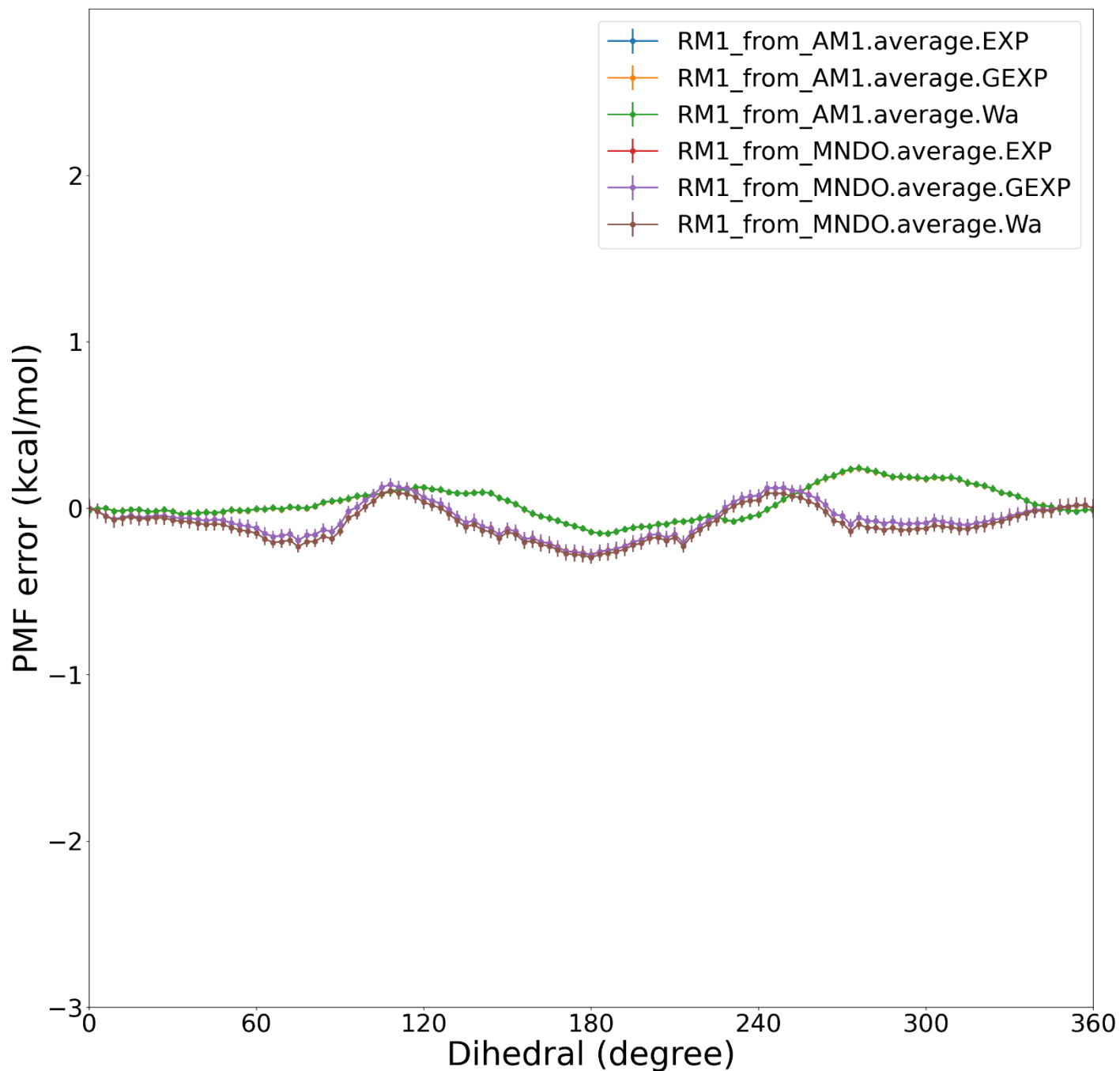


Fig. S5. The correlations between the MM and SQM (RM1) energetics calculated from 10 ns trajectories generated at 1300 K and 300 K. The sampling interval is 2 ps and there are 5000 independent configurations in total. The RMSE and MAE of the original parameter set and the newly fitted force-matching set are also presented. The RMSE of atomic forces is also improved from 23.9 kcal/(mol·Å·atom) to 13.1 kcal/(mol·Å·atom) at 300 K and from 37.6 kcal/(mol·Å·atom) to 25.0 kcal/(mol·Å·atom) at 1300 K.

