JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

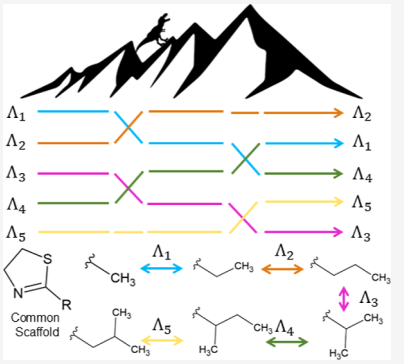
Multiple Topology Replica Exchange of Expanded Ensembles for Multidimensional Alchemical Calculations

Anika J. Friedman, Wei-Tse Hsu, and Michael R. Shirts*

Cite This: J. Chem. Theory Comput. 2025, 21, 230−240 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Relative free energy (RFE) calculations are now widely used in academia and the industry, but their accuracy is often limited by poor sampling of the complexes' conformational ensemble. To help address conformational sampling problems when simulating many relative binding free energies, we developed a novel method termed multiple topology replica exchange of expanded ensembles (MT-REXEE). This method enables parallel expanded ensemble calculations, facilitating iterative RFE computations while allowing conformational exchange between parallel transformations. These iterative transformations can be adaptable to any set of systems with a common backbone or central substructure. We demonstrate that the MT-REXEE method maintains thermodynamic cycle closure to the same extent as standard expanded ensemble calculations for both solvation free energy and relative binding free energy calculations. The transformations tested involve systems that incorporate diverse heavy atoms and multisite perturbations of a small molecule core resembling multisite λ dynamics, without necessitating modifications to the MD code. Our initial implementation is in GROMACS. We outline a systematic approach for the topology setup and provide instructions on how to perform inter-replica coordinate modifications. This work shows that MT-REXEE can be used to perform accurate and reproducible free energy estimates and prompts expansion to more complex test systems and other molecular dynamics simulation infrastructures.



1. INTRODUCTION

Alchemical free energy calculations have long been an important tool for accurately ranking the free energy difference between two physical states, but these methods have recently become much more common and widely used primarily due to improved accuracy and computational efficiency of these calculations.[1−4] The primary application of alchemical free energy calculations is in the lead optimization stage of drug development.[5−9] Many leading pharmaceutical companies use alchemical relative binding free energy (RBFE) calculations to discover compounds with higher binding affinity to their target protein.[10−12] The introduction of alchemical RBFE calculations to the workflow for these companies has been demonstrated to save time and money in the development of new drug compounds; however, there are still numerous limitations to the widespread application of these techniques.[8] Compared to other computational methods that have gained popularity in recent years, like machine learning predictions[13] or end-point methods (i.e., MM/GBSA),[14] alchemical free energy calculations are oftentimes significantly slower, which is balanced by their increased accuracy.[15−18] The accuracy of alchemical free energy calculations is somewhat limited by the accuracy of molecular mechanics force fields and the interplay between protein and small molecule force fields, as well as the handling of charge-changing mutations, which are being addressed well by other efforts,[19,20] but insufficient conformational sampling due to limitations in available computational resources is often a primary factor.[21,22] Numerous enhanced sampling methods have been developed to minimize simulation times necessary for the calculation of sufficiently accurate free energy estimates.[23−25] Despite the progress made in recent years, there are still many systems for which relative free energy (RFE) techniques are not normally applied due to the presence of flexible binding interfaces, significantly increasing the challenge of computing accurate free energy estimates.

Many researchers have worked to try to resolve or alleviate the issue of conformational sampling in alchemical RBFE calculations. Far too many methods have been proposed to review here,[23,26−35] but we highlight a few methods that have important similarities to what we propose in this study. Methods like Hamiltonian replica exchange (HREX) and expanded ensemble (EE) use the alchemical λ variable to their advantage by either allowing conformational swaps between λ intermediates, as in HREX, or by allowing the λ states to

Received: September 27, 2024
Revised: December 16, 2024
Accepted: December 17, 2024
Published: January 2, 2025

change while retaining the conformation from the previous state.[36−39]

Alchemical metadynamics allows simultaneous sampling along both alchemical and configurational collective variables to enhance configurational sampling in alchemical free energy calculations,[21] although this method requires prior knowledge of the slowest degrees of freedom to select useful collective variables. λ-Dynamics allows for the evolution of λ along the intrinsic free energy landscape, rather than using predetermined λ state spacing used by the other methods discussed previously.[40,41]

Our group previously developed the method of replica exchange of expanded ensembles (REXEE), which builds on the principles of HREX and EE to sample subsets of λ intermediate states, enhancing the flexibility and parallelizability of alchemical free energy calculations.[42] This paper introduces the multiple topology replica exchange of expanded ensembles (MT-REXEE) method, which further extends the REXEE method to couple several different alchemical transformations together to increase conformational swapping while computing the relative free energies of these transformations (Figure 1). This results in increased configurational sampling beyond what is achievable through the use of an alchemical variable alone by allowing alternate configurations that are generated with different ligands to be shared.

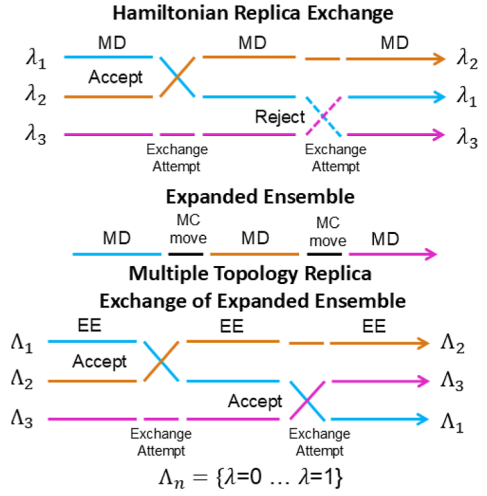


Figure 1. Traditional HREX involves running several parallel simulations, each of which has a constant λ value that varies from 0 to 1. EE calculations require only a single simulation that moves between λ states within a predetermined interval. MT-REXEE combines these methods to allow multiple parallel EE simulations to run and swap coordinates between replicas with one identical end state. Λ_i denotes the set of λ values varying from 0 to 1 for the *i*-th transformation.

The idea of increasing conformational sampling by doing many simultaneous transformations is not unique either, but our method approaches this implementation in a distinct manner. Multisite λ-dynamics (MSλD) allows sampling of multiple functional groups at multiple sites on a shared ligand core to be sampled within a single simulation.[43−46] MSλD has been demonstrated to produce accurate binding free energy estimates for relatively inflexible protein−ligand complexes significantly faster than traditional RBFE methods.[47−49] However, since this approach samples only a single protein ensemble while multiple ligands are sampled, the increased sampling is primarily on the part of the ligand. This becomes relevant for more flexible binding pockets, especially those binding pockets under ligand-induced conformational changes. These methods are also limited to the CHARMM[50] and OpenMM[51] simulation engines. In contrast, MT-REXEE samples separate protein and ligand conformations in each transformation and is implemented as a wrapper for the GROMACS[52] simulation engine, which, because the exchange process is independent of the MD engine, can easily be applied to any other engine with minor adjustments for handling engine-specific file extensions or data types. MSλD methods will generally be computationally less expensive if conformational sampling is sufficient in the protein pocket. However, MT-REXEE is likely to be more successful as a method specifically for instances in which MSλD or other less expensive methods fail to converge because of protein conformational sampling, such as the case of large flexible ligands in flexible binding pockets. MT-REXEE is also much more amenable to generalizations beyond the GROMACS implementation described in this paper as it requires no modification to the simulation code.

Although the primary application of RFE methods is small molecule binding free energy calculations for drug discovery, this is certainly not the only potential application. The increase in conformational sampling on behalf of both binding partners lends itself to more diverse problems such as binding free energy of protein−peptide complexes for the development of peptide drugs as well as protein−protein binding complexes.

In introducing the MT-REXEE method, we demonstrate first that the introduction of these conformational swaps between different alchemical transformations does not introduce additional error into either relative solvation free energy (RSFE) or RBFE calculations beyond existing methods. To demonstrate this, we utilize two relatively simple test systems. These simple test systems feature a single R group transformation on a common core. This is not a limitation to the method but rather a careful decision to select simple systems that can clearly identify fundamental errors in the method. First, we computed the RSFE of a set of aldehydes with varying numbers of carbons (4−16). We also examine RSFE and RBFE of major urinary protein 1 (MUP1) inhibitors, which are commonly used benchmarking targets for alchemical free energy methods.[44] We demonstrate that the results are statistically identical to those of well-sampled EE simulations, validating the approach and opening the door to more complex generalizations and applications that require the benefit of increased conformational sampling.

2. METHODS

2.1. Running Free Energy Simulations. *2.1.1. System Setup.* In this study, we use two test systems: a set of aldehydes and protein−thiazole binding complexes. The initial conformations for the solvation free energy calculations of the aldehyde systems (C_4OH_8−$C_{16}OH_{32}$) were constructed using Avogadro,[53] while those for calculations of both the solvation free energy and binding free energy of the protein−thiazole complexes were derived from the crystal structure (PDB 1IO6). All small molecules were parametrized using GAFF2,[54] and the protein was parametrized using the Amber *ff14sb* force field.[55] The relative topologies for all systems were created using PMX,[56] followed by postprocessing to remove all dihedral energies between dummy and real atoms within the hybrid topologies. The treatment of dummy atoms in this

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

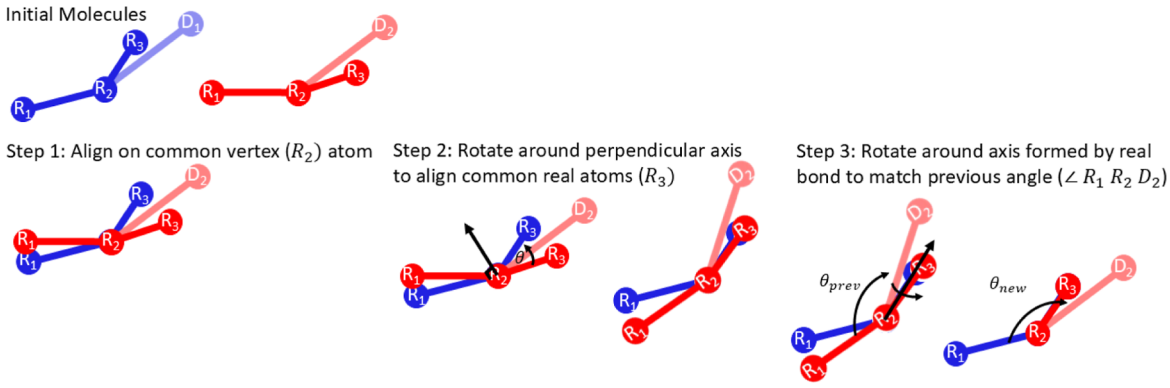


Figure 2. Diagram of the method in which missing atom coordinates are reconstructed during the conformational swaps. In this example, we construct atoms for the red dummy group ($D_{2,\mathrm{red}}$) for the blue molecule configuration. The blue molecule may sometimes contain alternate dummy groups ($D_{1,\mathrm{blue}}$), which will not show up in the new configuration, so these are ignored for the alignment protocol.

study follows the best practices presented in earlier studies.[57,58] For vacuum simulations, the simulation box was constructed to maintain a 3 nm distance between the molecule and the periodic boundary. The box for all simulations conducted in solvent was constructed to maintain a minimum distance of 1 nm between the molecule and the periodic boundary and then filled with TIP3P[59] solvent. All simulations were run in a modified version of GROMACS 2022.5,[60] which can be found on the official GitLab repository, and all code changes have been merged for inclusion in GROMACS 2025.

We used a standard procedure for equilibrating all of the systems. This procedure consisted of an initial energy minimization to a threshold of 50 kJ/mol·nm followed by a 100 ps *NVT* equilibration using the Bussi−Parrinello thermostat[61] maintaining a temperature of 300 K. A 100 ps *NPT* equilibration was also performed with the stochastic cell rescaling barostat[62] for all calculations performed in solvent to maintain a pressure of 1 atm. All simulations were performed with a 2 fs time step. Additional simulation setup information can be found for each simulation in the GitHub repository (https://github.com/shirtsgroup/MT-REXEE).

2.1.2. Validating Free Energy Simulations. Absolute solvation free energy (ASFE) calculations were performed to identify any error associated with the RFE calculation performed with both EE and MT-REXEE. The procedures for running both ASFE and RFE simulations were identical with the total aggregate simulation run time, which was 30 ns for relative vacuum simulations, 40 ns for relative solvent simulations, and 60 ns for ASFE and relative protein complex simulations. The λ-state spacing was determined for ASFE of both systems independently to maintain a sufficient off-diagonal state overlap of at least 0.1. Production simulations were run using the EE formalism[63,64] in GROMACS. A change in the Hamiltonian state was attempted every 100 time steps (or 200 fs), and state moves were made using the Metropolis−Gibbs algorithm.[65] The state weights were updated using the Wang−Landau algorithm[66] and were fixed when the Wang−Landau incrementor was less than the threshold of 0.001 k_BT. All settings for the expended ensemble simulations can be found for each simulation in the MT-REXEE GitHub repository (https://github.com/shirtsgroup/MT-REXEE).

2.1.3. Coordinate Modification. A generalized method for coordinate modification is implemented in the *ensemble_md* repository (https://github.com/shirtsgroup/ensemble_md). First, we determine which λ state is being sampled in each parallel simulation. If two simulations are at compatible end states in the final frame of the previous iteration, then a swap is proposed. In order for a configurational swap to occur, two simulations must share a common molecular end point in each of their transformations (i.e., A to B with $\lambda = 1$ and B to C with $\lambda = 0$, where B is the common end point being sampled). For example, simulation 1 samples the relative transformation from molecule A to B and simulation 2 samples the relative transformation from molecule B to C. If simulation 1 is sampling $\lambda = 1$ in the last frame of the trajectory and simulation 2 is sampling $\lambda = 0$, then a swap will be performed. Unlike traditional replica exchange methods, the two molecules in this transformation are in the same chemical state, and thus, all swaps are automatically accepted if they are in compatible states as explained above (Supporting Information Section 1).

The coordinate modification method described below was designed to depend on as small a region of both molecules as possible to remove dependence on the global conformation of the molecules. The initial coordinates for both molecules are read from the full-precision trajectory using MDTraj.[67] Prior to swapping coordinates, any discontinuities in the molecule due to periodic boundary conditions were resolved by shifting atoms across the periodic boundary to prevent alignment issues. This is a key difference between the REXEE and MT-REXEE methods. Additionally, in MT-REXEE, configurational swaps are always accepted because the two molecules are required to be sampling alchemically identical states.

The molecules are then aligned to enable the building in coordinates for dummy atoms that differ between the two adjacent replicas. The alignment occurs across the bond in which a different functional group must be substituted and, therefore, can occur multiple times depending on the differences in the structure of the dummy atoms. This would easily accommodate the introduction of multiple varying R group substitutions to the common scaffold, but ring-breaking and ring-forming modifications are not supported by the current modification scheme. To perform the conformational swap, we must define the atoms of interest shown in Figure 2, which include an anchor atom (R_2), an alignment atom (R_3), and an angle atom (R_1). The anchor atom is the first real atom with which the dummy group being reconstructed forms a covalent bond. The alignment atom is the atom that is also covalently bonded to the anchor atom and is real at the currently sampled λ but becomes a dummy atom at the other

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

end state. The angle atom is the real atom present in both molecules and is used to ensure that the new dummy group is added in a compatible configuration relative to the rest of the molecule. The angle atom must be covalently bonded to the anchor atom.

To perform the swap, we first perform a translation on all coordinates in the red molecule such that the anchor atom ($R_{1,red}$) is aligned with the anchor atom ($R_{1,blue}$) in the blue molecule. In step two, we define an axis perpendicular to the plane formed between the vertex ($R_{1,blue}$) and the common real atoms ($R_{2,red}$ and $R_{2,blue}$). We then perform a rotation about this axis to align R_2 in blue and red. In step 3, we define an axis between the vertex and the (now aligned) red real atom (R_2). The dummy group is now rotated around this axis such that the angle formed between the dummy and real sections of the molecule ($\angle R_1R_2D_2$) are the same as they were before the swap ($\theta_{prev} = \theta_{new}$). The final dummy coordinates are then reconstructed onto the original molecule, such that the new configuration will have atomic coordinates from the original blue molecule for all atoms other than the dummy atoms.

The coordinates for all atoms, including the constructed atoms, are then written to a GRO file with 7 decimal place precision. This method does not modify the simulation engine code, which necessitates human-readable structure input files to interface with the engine. We could additionally provide a full-precision binary trajectory file as an input, but the coordinate precision would be equivalent for single-precision GROMACS installations. The standard GRO file contains 3 decimal places, but we found that the additional precision was necessary to ensure that the systems energies match within machine precision before and after swaps occur. In order to achieve additional precision in the input coordinates, the coordinate modification file reads in a noncompressed trajectory (.trr format), as GROMACS can read additional precision in GRO files, but does not output higher precision in this format. There are no issues with clashing water molecules or protein residues as the atoms that have been constructed from this coordinate modification begin the simulation as fully noninteracting dummy atoms that are incapable of clashing with real atoms and the system is allowed to equilibrate before these interactions are restored. This configuration writing and swapping algorithm runs on a single core and takes 0.3−0.8 s of wall time, depending on the size of the system and whether periodic boundary breaks need to be corrected.

The introduction of the coordinate modification function thus makes a small contribution to the run time of the full algorithm. The wall times for a single 20 ps iteration for a vacuum, solvent, and complex simulation are 11.1 ± 0.2, 12.3 ± 0.2, and 28.5 ± 0.4 s, respectively, which are all significantly larger (14−57×) than the 0.3−0.8 s introduced per conformational swap including the file processing, meaning that introducing more complex program structure to perform conformational swap itself is not necessary. The variance in the time to perform the conformational swap is primarily due to whether 0, 1, or 2 of the molecules have breaks across the periodic boundary. The increase in wall time based on system size is minimal with a difference of ∼0.1 s between a system containing 20 atoms and one containing 35K atoms. For a total simulation length of 20 ns for vacuum simulations, the coordinate modification leads to a 1.5% increase in the total computational cost. The coordinate modification step scales with the degree of modification rather than the total number of atoms in the system, so the increase in computational cost decreases with system size for an increase in computational cost of 0.8% and 0.2% for the solvent and complex simulations, respectively.

The coordinate modification step must conserve the potential energy of the two segments of the molecule, which are run in separate simulations, i.e., the $\lambda = 0$ state of one simulation and the $\lambda = 1$ state of a companion simulation. If we are swapping coordinates between molecule A and molecule B as shown in Figure S11A, we need to reconstruct coordinates for the dummy hydrogen for the new configuration of molecule A and coordinates for the methyl group for the new configuration of molecule B. The new configuration for A would, thus, have a subset of atoms, which should have bonded parameters corresponding to the potential energy for the initial configuration in A and a subset corresponding to the initial configuration in B shown by molecule color in Figure S11B.

We computed the potential energy of bonds, angles, torsions, and van der Waals interactions of the MUP1 ligands (A, F shown in Figure 5A) used in the systems described in this paper before and after the conformational swap. The interactions were then separated based on the origin molecule for which the interactions should match (i.e., energies corresponding to blue atoms before the swap should match blue atoms after the swap based on Figure S11A,B). We computed these energies for three replicates of the acyl system in both vacuum and solvent as well as the MUP1 ligand system in both vacuum and solvent. On average, the total deviation in potential energy with this coordinate swapping method was $7.1 \times 10^{-5} \pm 8.6 \times 10^{-5}$ kcal/mol (Figure S11C). This error can be decreased to $2.4 \times 10^{-7} \pm 2.5 \times 10^{-7}$ kcal/mol by running GROMACS in double precision and increasing the number of decimal places in the GRO file from 7 to 10, which leads us to conclude that the dominant deviation in potential energy is due to the precision in the coordinates and not a systematic error being introduced. We note that the released code includes a "paranoid mode", which slows exchange but includes potential energy validation of each switch for validation.

2.1.4. MT-REXEE Protocol. Each MT-REXEE simulation was conducted using the same λ state spacing as that of the corresponding EE simulation and was run for the same total duration (following weight equilibration). We included redundant end states ($\lambda = 0$ and $\lambda = 1$) for these simulations in order to increase the frequency at which swaps were able to occur. We added these redundant end states such that 1/3 of the total number of states were at $\lambda = 0$, 1/3 were at $\lambda = 1$, and 1/3 were sampling intermediate λ states. The choice of 1/3 as transition states and 1/3 as end states was an empirical choice to ensure sufficient swaps between alchemical transformations and should be tuned for future applications. Redundant end states were utilized because there is currently no method within the GROMACS infrastructure to equilibrate weights for uneven state sampling. All λ states were initiated with 0 weights to match the EE procedure, and weights were equilibrated using MT-REXEE with weight equilibration determined by the Wang−Landau algorithm. The redundant end states are included during weight equilibration, and Monte Carlo swaps are allowed between nonadjacent λ states (i.e. any end states have an equal probability of swapping with any other state), producing even sampling across all λ states after equilibration (Figures S9 and S10). For both the solvation and binding free energies of both the aldehyde and thiazole systems, the conformational swaps were proposed every 20 ps

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

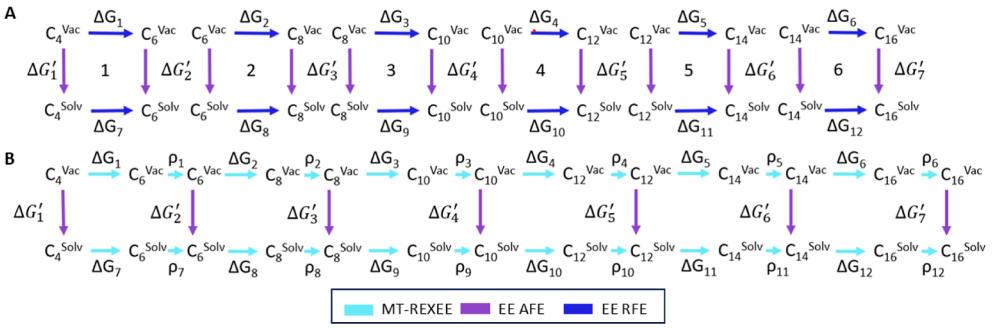


Figure 3. (A) Individual thermodynamic cycles for the calculation of $\Delta\Delta G$ of the iterative transformation from C4−C16 using EE simulations. The purple ASFE transformations and the blue RSFE transformations create independent thermodynamics cycles for each transformation. (B) Thermodynamic cycle for the calculation of $\Delta\Delta G$ for the iterative transformation from C4−C16 using a single MT-REXEE simulation with 6 replicas. Notably, the light blue RSFE transformations are connected through the conformational swaps between simulation instances.

unless otherwise stated. The 20 ps value was determined by running MT-REXEE simulations in a vacuum and solvent for the MUP1 ligand systems using different swapping rates (10, 20, 50, and 100 ps). We found that there was no significant difference in the free energy estimates while varying the swapping rate, but the round-trip time (time to sample all states) significantly decreased as the swapping frequency decreased faster than a frequency of 20 ps (Figure S8C). We also validated that altering the swapping frequency did not have any effect on the Wang−Landau weight equilibration time (Figure S8B). This inclusion of redundant states reduces the total fraction of intermediate time spent at intermediate states compared to that in EE simulations but, as demonstrated in the results section, significantly increases the exchange rate between states. The exact trade-off between these two effects will be system-dependent, and we do not attempt in this study to generalize the determination of this time scale to other systems.

2.2. Analyzing Free Energy Simulations. *2.2.1. Determining FE from Simulations.* For all simulations conducted in this study the free energy estimate was computed using the alchemlyb[68] wrapper around pymbar.[69] Only the trajectory information after weight equilibration according to the Wang−Landau incrementor criteria was used to compute the free energy. The number of uncorrelated samples was determined using the alchemlyb *statistical_inefficiency* function.[70] Three replicates were used for each simulation, and the free energy values reported are the mean of the three replicate simulations. The uncertainties reported are the standard error of the mean using the three independent replicas.

2.2.2. Evaluating Cycle Closure. RFE calculations rely on the closure of the thermodynamic cycle for each relative transformation. We evaluate whether our introduction of conformational swaps between different relative transformations affects the accuracy of our RFE estimates.

Figure 3A shows the thermodynamic cycles for each of the individual transformations for the solvation free energy for the aldehyde system. The path independence of free energies gives us eq 1, from which we can get eq 2.

$$\Delta G_1 + \Delta G_2' - \Delta G_7 - \Delta G_1' = 0 \tag{1}$$

$$\Delta\Delta G = \Delta G_1' - \Delta G_2' = \Delta G_7 - \Delta G_1 \tag{2}$$

Equation 2 assumes perfect matching between topologies, but there is potentially error associated with each individual simulation such that eq 1 becomes eq 3 where δ_1 is the sum of the error associated with all four calculations.

$$\text{Cycle Error} = |\Delta G_1 + \Delta G_1' - \Delta G_7 - \Delta G_2'| = |\delta_1| \tag{3}$$

In the MT-REXEE method, there is an additional source of error that does not apply to the individual EE calculations from the conformational swapping step. This additional error associated with the conformational swaps in a given MT-REXEE simulation is given by ρ_n in Figure 3B. This error likely comes from two sources: changes in the potential energy of the molecule before and after conformational swaps and changes in conformational sampling between MT-REXEE and EE. Dummy atoms have no nonbonded energies because the Lennard-Jones parameters (ϵ and σ) and charges (q) are all 0, but they still must contain some bonded interactions with the real atoms in order to prevent nonphysical atom movements and maintain the general molecule structure.[57] In this study, all torsions involving real and dummy atoms are set to 0 at the dummy end points, but there are bonds and angles that maintain their standard energy parameters in order to prevent nonphysical structures from being created. The presence of these dummies with some nonzero nonbonded terms could potentially produce changes in the configurational space sampled, whose effect must be evaluated.[57] We do not expect this contribution to the transformation free energy to be zero, but we do expect it to be nearly the same between transformations in the solution, in the vacuum, and in complex with proteins. The swapping function minimizes the potential energy difference to machine precision for fixed coordinates, and we assert that the error in free energy should be nearly equivalent in both the solvent and vacuum transformations such that $|\rho_n - \rho_{n+6}| \approx 0$. To evaluate this error, we assume that the deviation between the RFE obtained from EE and the value obtained from MT-REXEE can be attributed to this conformational swapping error given by $|\rho_n - \rho_{n+6}|$.

$$|\Delta\Delta G_{\text{EE}} - \Delta\Delta G_{\text{MT-REXEE}}| = |\rho_n - \rho_{n+6}| \tag{4}$$

Equation 4 assumes that the conformational sampling between the EE and MT-REXEE simulations is equivalent such that they would be expected to produce the same result if there was no additional error introduced by the conformational swapping step itself. This can reasonably be expected for the solvation free energy and binding free energy of the systems selected for this study as there are no significant free energy barriers that may lead to significant differences in the conformational sampling from the two methods. We validate the approach on these simpler systems as it will be much harder to ensure that there are no free energy errors associated with MT-REXEE for more complex systems, for which MT-

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

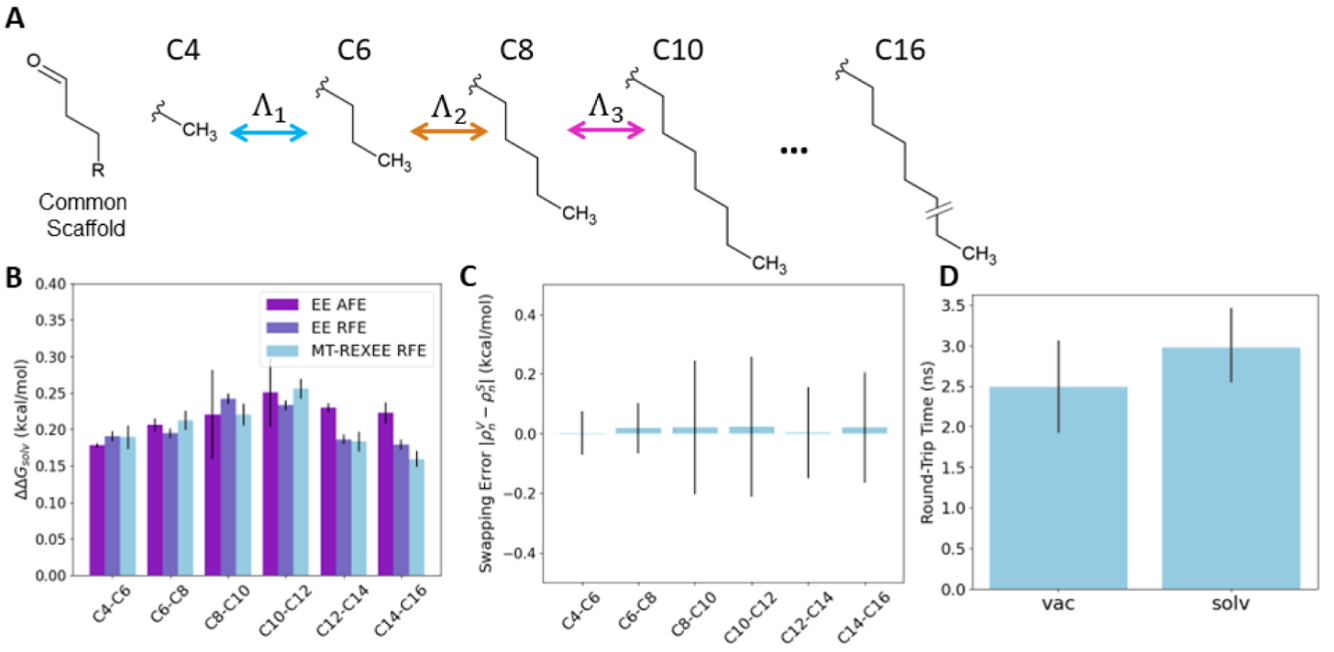


Figure 4. Validation of the aldehyde chain growth test case. (A) We compute the solvation free energy of a hydrocarbon-functionalized aldehyde as we vary the number of carbons in the chain, with error bars shown as the standard error of the measurements between three replicas. (B) We compute the RSFE using three different methods: EE AFE using the difference in EE ASFEs of each end state to compute a relative value, EE RFE using separate RSFE calculations for each transformation, and MT-REXEE runs parallel solvation free energy calculations with conformational swapping. (C) The error associated with the conformational swap process (see eq 4) is statistically insignificant. (D) The round-trip time shows the average duration for conformational swaps, which allow shared swapping between all simulations. The rate is on the order to 2−3 ns, and thus, a 20 ns simulation will have on average ∼8 full round trips.

REXEE will achieve increased sampling of the full conformational space compared to other methods, thus leading to differences in the RFE estimates despite minimal error introduced by the conformational swapping step.

2.2.3. Evaluating Replica Swapping. In order to evaluate the path of replica swaps throughout the simulation, we construct replica swap paths starting from the initial state of each replica. These replica swap paths follow the state being sampled for each conformation. All simulations are initiated from $\lambda = 0$ for their individual transformation. The transition path follows the conformation state, not the replica state, such that when a conformational swap occurs, the conformational state is now sampling the region for the adjacent replica. These transition paths can be used to determine the round-trip time. The round-trip time starts at $\lambda = 0$ for each transformation, and we follow the replica swap path until all end states are sampled in the simulation and return back to the original end state. This definition is important for nonlinear swapping paths for which swaps do not occur solely with adjacent replicas. The mean round-trip times were computed from all three replicate trajectories.

3. RESULTS

3.1. Validation Using the Aldehyde System. Our first step in validating this enhanced sampling method is to ensure that the introduction of these conformational swaps between adjacent relative transformations does not introduce additional errors into the calculation. We first use a simple test system in which we compute the solvation free energy for acyl groups functionalized with alkane chains of varying lengths from 4 to 16 carbons (Figure 4A). This system was selected because it is relatively computationally inexpensive, allowing us to ensure that we are exhaustively sampling the conformational space in both the EE and MT-REXEE calculations. This allows us to isolate the contribution of conformational swapping to the free energy estimate since there should be minimal variance between individual replicas over the time scale of 20−30 ns.

We first compute the ASFE of each aldehyde chain from 4 to 16 carbons and compare against available experimental data[71] in order to ensure that there are no significant issues affecting the overall molecular behavior with the parameterization of our molecules or other simulation settings common to all simulations. Where experimental data were available, our solvation free estimates properly followed the trend in solvation free energy, though they are not necessarily expected to be within experimental error (Figure S1). This allows us to move on to running the RSFE simulations using both the standard EE method and our MT-REXEE method.

We utilize the same λ state spacing for both the EE and MT-REXEE simulations with the exception of the inclusion of redundant end states (3 for vacuum simulations and 5 for solvent simulations) for the MT-REXEE simulations. The λ state spacing and simulation length were optimized to ensure sufficient state overlap and estimate convergence (Figures S2 and S3). These redundant end states do lead to a slight increase in the time needed to optimize the λ weights. This translates to mean vacuum weight equilibration times of 1.16 ± 0.34 and 2.41 ± 0.51 ns for EE and MT-REXEE, respectively, and solvent weight equilibration times of 2.56 ± 0.56 and 4.64 ± 0.65 for EE and MT-REXEE, respectively (Figure S8A). The MT-REXEE simulations have 3× as many λ states, so increasing the weight equilibration time by ∼2× is reasonable. In order to directly compare the two methods, the length of the simulation after weight equilibration was 20 ns for vacuum simulations and 30 ns for solvent simulations by using both EE and MT-REXEE.

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

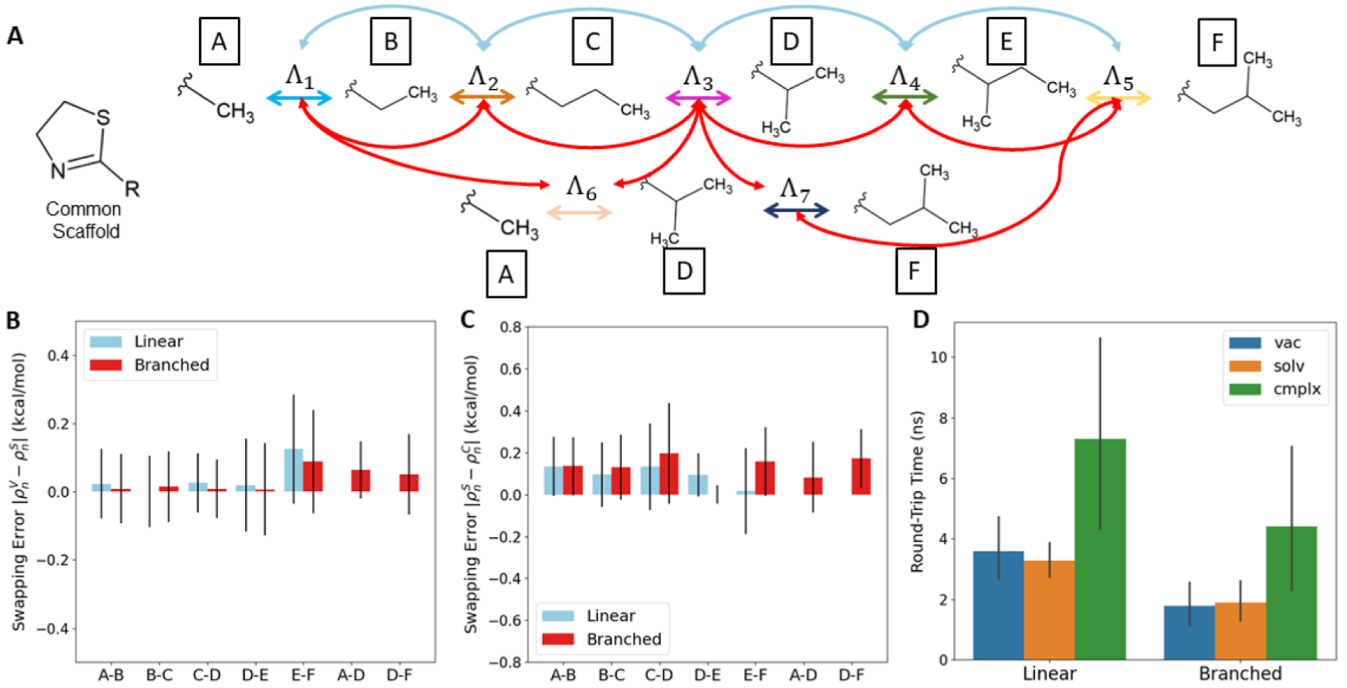


Figure 5. (A) The swapping network defines the relative transformations occurring in both the linear (blue) and branched (red) networks for both the RSFE and RBFE calculations. (B) The error in RSFE calculations introduced using the MT-REXEE method with both the linear and branched paths is statistically indistinguishable from zero, supporting the claim that any error introduced by changes in dummy atoms is equal in solvent and vacuum and cancels in the RFE. (C) This same cancellation of error also occurs for solvent and complex simulations used to compute the RBFE. (D) Introducing the branched swapping pattern reduces the round-trip time similarly in the vacuum, solvent, and complex simulations, though the round-trip time is increased substantially in simulations with more topologically complex connections.

As intended, there is not a statistically significant difference between the RFE estimates obtained from EE and MT-REXEE (Figures 4B and S7A). Using eq 4, we obtain the error likely introduced through the inclusion of configurational swapping steps in MT-REXEE, with values ranging from 0.0079 ± 0.134 to 0.025 ± 0.052 kcal/mol (Figure 4C). This error is within the statistical error for every one of these measurements, which validates that the MT-REXEE method does not introduce significant errors to the RFE estimate for this simple system.

We need to ensure that the limited error introduced is not due to the simulations becoming trapped in an individual replica as this would essentially become equivalent to a standalone EE simulation. We demonstrate this by comparing round-trip times for the replica swaps, that is, the time it takes for the configuration to sample all other replicas and return to its initial replica. This round-trip time was 2.49 ± 1.04 and 4.00 ± 1.54 for vacuum and solvent simulations, respectively (Figure 4D), averaged over 3 replicates. A minimum of two round trips occur per replica in the simulations conducted here with a 20 ps swapping frequency. We will explore later in the article how varying the rate of this swapping frequency also affects the round-trip time for the simulations.

3.2. MUP1 System Validation. We first computed the solvation free energies of the ligands to validate that the MT-REXEE method does not affect the cycle closure of the RFE measurements. As we did previously, the λ state spacing and simulation length were optimized to ensure sufficient state overlap and estimate convergence for all calculations (Figures S4−S6). For this system, in addition to computing a straight path of adding functional groups to the molecule, we also added alternate transformations to create a branched swapping network (Figures 5A and S7B,C). These calculations validated that increasing the complexity of the swapping network does not introduce significant additional error into the system. This demonstrates the generalization of this method to diverse systems for which complicated branched transformation maps may be necessary. The only limitation to swapping between different transformations is that one end state must be shared between the two replicas.

We observe no statistically significant differences in RSFE when comparing the standard EE simulations with the linear and branched paths in MT-REXEE simulations (Figure 5B). Using the same transformation network, we also computed RBFE values using both EE and MT-REXEE methods. The difference between the EE method and both the branched and linear transformation paths was found to be within simulation error (Figure 5C). These calculations demonstrate that the MT-REXEE method will produce the same RFE estimates as EE alone when both calculations are converged. We also demonstrate that introducing these additional swapping paths has a significant reduction in the round-trip time by increasing the frequency for which replicas were in compatible λ states to engage in confrontational swaps (Figure 5D). The significant increase in the round-trip time in complex simulations can mostly be attributed to individual replicas becoming temporarily stuck and swapping between a subset of allowed λ states, reducing the probability of two simulations visiting compatible λ states at the end of a given simulation. This effect can be reduced by lowering the Wang−Landau threshold for weight equilibration, though the uneven weight sampling was seen only during short time frames (5−10 ps), and over longer time frames (100 ps), the state sampling was even across λ states with the equilibration threshold used in this study.

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

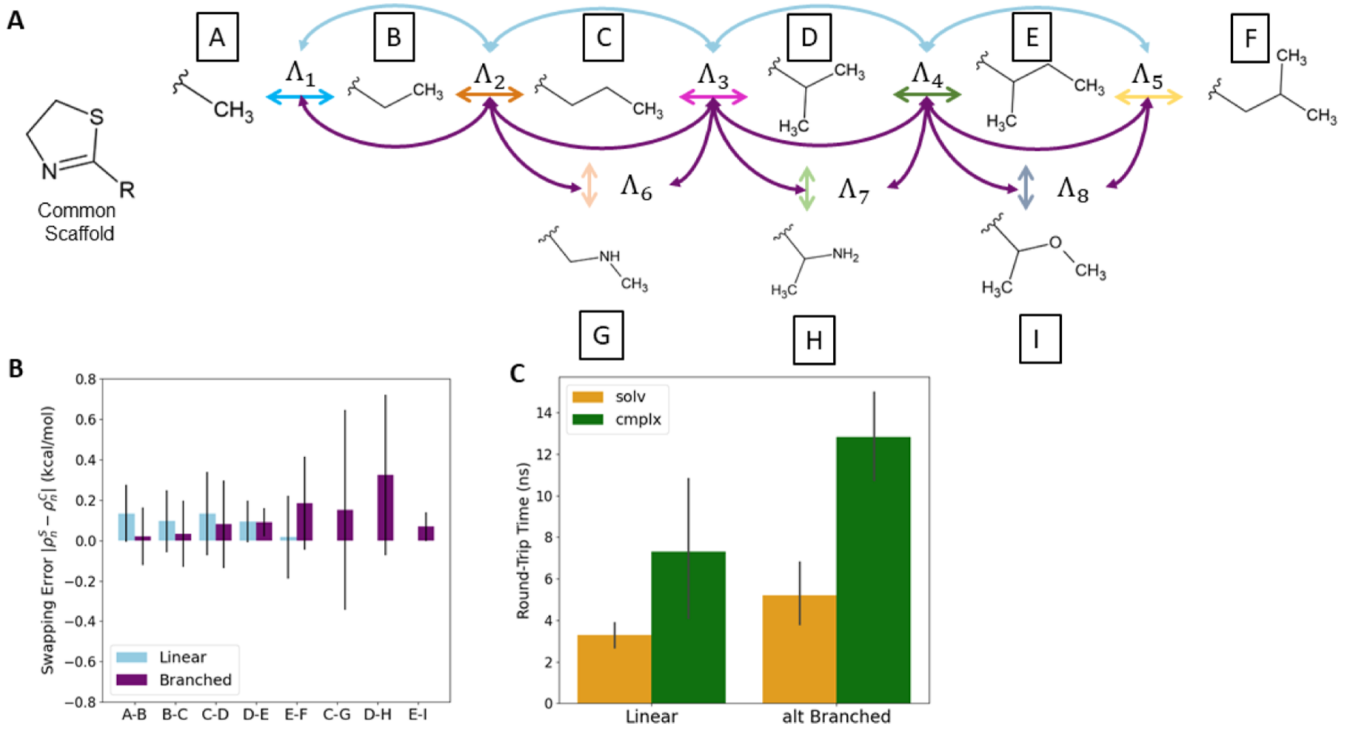


Figure 6. (A) The swapping network shown here introduces diversity in the heavy atoms involved in the relative transformations for both the complex and solvent simulations. (B) Despite this additional complexity, there is again no statistically significant error introduced to the computation of the RBFE for any of the transformations. (C) Introducing these additional relative transformations does increase the round-trip time by 60%, which is consistent with the 50% increase in the end states, which must be sampled to complete a round trip in this swapping network.

All of the relative solvation and binding free energy calculations presented thus far involve dummy hydrogen and carbon atoms only. To ensure that the method can be generalizable to a wider range of potential relative transformations, we included additional transformations that alter a heavy atom (in this case carbon) to another heavy atom (in this case oxygen or nitrogen) (Figures 6A and S7D). Introducing this diversity into the heavy atoms involved in the transformations resulted in no significant additional error in the free energy estimates (Figure 6B). The branched network was shown to significantly increase round-trip times as there are three additional end states that must be visited to complete a round trip (Figure 6C). As expected, introducing the branched network increases the round-trip time, which is approximately linearly related to the number of unique end states in the system as a given simulation can engage in only one conformational swap in each iteration. As we continue to increase the complexity of the transformation network by introducing additional unique end states, we also increase the importance of planning the swapping network to minimize the round-trip time.

4. DISCUSSION AND CONCLUSIONS

The MT-REXEE method allows the calculation of multiple relative free energies when there is a common scaffold for the transformations. In this study, we demonstrate that introducing conformational exchanges between replicas does not introduce error into the free energy calculation in the event that sufficient sampling is completed. We have shown that this is the case for both linear and branched swapping patterns, as well as relative transformations that involve a variety of heavy atoms.

The primary limitation of the current method is that this implementation requires redundant end states, which increases the Wang−Landau weight equilibration time, as well as increased the time it takes for free energy estimates to converge due to decreased sampling of intermediate states. The computational cost of running MT-REXEE scales linearly with the number of transformations being computed, the same as for any other large-scale scan of multiple alchemical mutations. The only difference with MT-REXEE is that communication between simulation tasks can be carried out relatively infrequently compared with replica exchange. For binding free energy calculations, the MT-REXEE method also samples separate protein configurations with each alchemical transformation, and this sharing between alchemical states may be very useful for obtaining accurate estimates for flexible protein systems. This is a clear advantage over λ-dynamics and related combined topology approaches, which compute multiple ligand alchemical changes significantly more efficiently but sample only one protein configurational ensemble.[44]

Further work will focus on demonstrating the benefit of this enhanced conformational sampling of MT-REXEE in larger systems with multiple variable R groups as well as integrating core hopping. The use of EE methods to run the individual transformations enables decoupling of computational resources from the number of alchemical intermediate states. The number of computational cores needed is instead dependent on the number of individual transformations being performed, which better lends itself to high-throughput calculations.

MT-REXEE is most likely to be well-suited for systems that feature large flexible proteins or ligands. This is particularly relevant for the rise of interest in the study of peptide drugs

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240

that are not well-suited for traditional free energy calculation methods due to their flexibility. Though drug discovery is the primary focus of most RFE studies, the application of these calculations can extend far beyond small molecule and peptide drugs. For example, it could be applied to enzyme engineering problems that involve protein−protein binding interfaces, which often involve significant rearrangement. As we expand to more diverse systems, it may also be desirable to combine MT-REXEE with other adaptive sampling techniques. Methods to adaptively improve sampling within the EE calculations themselves can easily be added,[72,73] and the logic of how much simulation time is spent simulating each edge in the alchemical transformation network can also be adjusted based any of the recent published methods.[74−77]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c01268.

Details of swap acceptance, comparison of aldehyde solvation free energy values to experiment, convergence and overlap matrices for aldehyde and MUP1 ligand systems in vacuum, solvent, and complex, replica values for all simulations using EE and MT-REXEE, mean equilibration times for vacuum and solvent MUP1MT-REXEE simulations, λ state sampling for vacuum simulations for aldehyde and MUP1 ligand systems, and potential energy differences before and after conformational swaps in the MUP1 ligand systems (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Michael R. Shirts − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0003-3249-1097; Email: michael.shirts@colorado.edu

Authors

Anika J. Friedman − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0002-5427-2779

Wei-Tse Hsu − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.4c01268

Author Contributions

A.J.F., W.-T.H., and M.R.S. conceptualized the project and designed the methodology. A.J.F. implemented the sampling method in Python with advice from W.-T.H. Experiments were performed and analyzed by A.J.F. A.J.F. wrote the original manuscript draft; W.-T.H. and M.R.S. edited and reviewed the manuscript. M.R.S. supervised the project and obtained resources.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was supported by the grant OAC-1835720 from the National Science Foundation and grant R01GM123296 from the National Institutes of General Medical Sciences. This work utilized the Alpine high-performance computing resource at the University of Colorado Boulder. Alpine is jointly funded by the University of Colorado Boulder, the University of Colorado Anschutz, and Colorado State University. M.R.S. has received an Open Science Fellowship from Psivant Therapeutics and consults for Relay therapeutics.

■ REFERENCES

(1) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.

(2) Song, L. F.; Merz, K. M. J. Evolution of Alchemical Free Energy Methods in Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60*, 5308−5318.

(3) York, D. M. Modern Alchemical Free Energy Methods for Drug Discovery Explained. *ACS Phys. Chem. Au* **2023**, *3*, 478−491.

(4) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. Chapter 4 Alchemical Free Energy Calculations: Ready for Prime Time? *Annu. Rep. Comput. Chem.* **2007**, *3*, 41−59.

(5) Abel, R.; Wang, L.; Mobley, D. L.; Friesner, R. A. A Critical Review of Validation, Blind Testing, and Real-World Use of Alchemical Protein-Ligand Binding Free Energy Calculations. *Curr. Top. Med. Chem.* **2017**, *17*, 2577−2585.

(6) Lee, T.-S.; Allen, B. K.; Giese, T. J.; Guo, Z.; Li, P.; Lin, C.; McGee, T. D. J.; Pearlman, D. A.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Xu, H.; Sherman, W.; York, D. M. Alchemical Binding Free Energy Calculations in AMBER20: Advances and Best Practices for Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60*, 5595−5623.

(7) Jorgensen, W. L. Efficient Drug Lead Discovery and Optimization. *Acc. Chem. Res.* **2009**, *42*, 724−733.

(8) Cournia, Z.; Chipot, C.; Roux, B.; York, D. M.; Sherman, W. Free Energy Methods in Drug Discovery: Current State and Future Directions. *ACS Symp. Ser. (Am. Chem. Soc.)* **2021**, *1397*, 1−38.

(9) Williams-Noonan, B. J.; Yuriev, E.; Chalmers, D. K. Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry. *J. Med. Chem.* **2018**, *61*, 638−649.

(10) Muegge, I.; Hu, Y. Recent Advances in Alchemical Binding Free Energy Calculations for Drug Discovery. *ACS Med. Chem. Lett.* **2023**, *14*, 244−250.

(11) Mobley, D. L.; Klimovich, P. V. Perspective: Alchemical Free Energy Calculations for Drug Discovery. *J. Chem. Phys.* **2012**, *137*, 230901.

(12) Homeyer, N.; Stoll, F.; Hillisch, A.; Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of their Accuracy in an Industrial Drug Design Context. *J. Chem. Theory Comput.* **2014**, *10*, 3331−3344.

(13) Cournia, Z.; Allen, B. K.; Beuming, T.; Pearlman, D. A.; Radak, B. K.; Sherman, W. Rigorous Free Energy Simulations in Virtual Screening. *J. Chem. Inf. Model.* **2020**, *60*, 4153−4169.

(14) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478−9508.

(15) Ngo, S. T.; Tam, N. M.; Pham, M. Q.; Nguyen, T. H. Benchmark of Popular Free Energy Approaches Revealing the Inhibitors Binding to SARS-CoV-2 Mpro. *J. Chem. Inf. Model.* **2021**, *61*, 2302−2312.

(16) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57*, 2911−2937.

(17) Luo, J.; Wei, W.; Waldispühl, J.; Moitessier, N. Challenges and Current Status of Computational Methods for Docking Small Molecules to Nucleic Acids. *Eur. J. Med. Chem.* **2019**, *168*, 414−425.
(18) Bitencourt-Ferreira, G.; de Azevedo, W. F. Development of a Machine-Learning Model to Predict Gibbs Free Energy of Binding for Protein-Ligand Complexes. *Biophys. Chem.* **2018**, *240*, 63−69.
(19) Lin, F.-Y.; MacKerell, A. D. Force Fields for Small Molecules. *Methods Mol. Biol.* **2019**, *2022*, 21−54.
(20) Lim, V. T.; Hahn, D. F.; Tresadern, G.; Bayly, C. I.; Mobley, D. L. Benchmark Assessment of Molecular Geometries and Energies from Small Molecule Force Fields. *Chem. Inf. Sci.* **2020**, *9*, 1390.
(21) Hsu, W.-T.; Piomponi, V.; Merz, P. T.; Bussi, G.; Shirts, M. R. Alchemical Metadynamics: Adding Alchemical Variables to Metadynamics to Enhance Sampling in Free Energy Calculations. *J. Chem. Theory Comput.* **2023**, *19*, 1805−1817.
(22) La Serra, M. A.; Vidossich, P.; Acquistapace, I.; Ganesan, A. K.; De Vivo, M. Alchemical Free Energy Calculations to Investigate Protein−Protein Interactions: The Case of the CDC42/PAK1 Complex. *J. Chem. Inf. Model.* **2022**, *62*, 3023−3033.
(23) Hénin, J.; Lelièvre, T.; Shirts, M. R.; Valsson, O.; Delemotte, L. Enhanced Sampling Methods for Molecular Dynamics Simulations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2022**, *4*, 1583.
(24) Jiang, W. Enhanced Configurational Sampling Approaches to Alchemical Ligand Binding Free Energy Simulations: Current Status and Challenges. *J. Phys. Chem. B* **2023**, *127*, 6835−6841.
(25) de Ruiter, A.; Oostenbrink, C. Advances in the Calculation of Binding Free Energies. *Curr. Opin. Struct. Biol.* **2020**, *61*, 207−212.
(26) Lee, T.-S.; Tsai, H.-C.; Ganguly, A.; York, D. M. ACES: Optimized Alchemically Enhanced Sampling. *J. Chem. Theory Comput.* **2023**, *19*, 472−487.
(27) Azimi, S.; Khuttan, S.; Wu, J. Z.; Pal, R. K.; Gallicchio, E. Relative Binding Free Energy Calculations for Ligands with Diverse Scaffolds with the Alchemical Transfer Method. *J. Chem. Inf. Model.* **2022**, *62*, 309−323.
(28) Zhang, S.; Giese, T. J.; Lee, T.-S.; York, D. M. Alchemical Enhanced Sampling with Optimized Phase Space Overlap. *J. Chem. Theory Comput.* **2024**, *20*, 3935−3953.
(29) Wu, J. Z.; Azimi, S.; Khuttan, S.; Deng, N.; Gallicchio, E. Alchemical Transfer Approach to Absolute Binding Free Energy Estimation. *J. Chem. Theory Comput.* **2021**, *17*, 3309−3319.
(30) Perthold, J. W.; Oostenbrink, C. Accelerated Enveloping Distribution Sampling: Enabling Sampling of Multiple End States While Preserving Local Energy Minima. *J. Phys. Chem. B* **2018**, *122*, 5030−5037.
(31) Xu, J.; Cao, X.-M.; Hu, P. Accelerating Metadynamics-Based Free-Energy Calculations with Adaptive Machine Learning Potentials. *J. Chem. Theory Comput.* **2021**, *17*, 4465−4476.
(32) König, G.; Glaser, N.; Schroeder, B.; Kubincová, A.; Hünenberger, P. H.; Riniker, S. An Alternative to Conventional λ-Intermediate States in Alchemical Free Energy Calculations: λ-Enveloping Distribution Sampling. *J. Chem. Inf. Model.* **2020**, *60*, 5407−5423.
(33) Hahn, D. F.; König, G.; Hünenberger, P. H. Overcoming Orthogonal Barriers in Alchemical Free Energy Calculations: On the Relative Merits of λ-Variations, λ-Extrapolations, and Biasing. *J. Chem. Theory Comput.* **2020**, *16*, 1630−1645.
(34) Bussi, G.; Laio, A. Using Metadynamics to Explore Complex Free-Energy Landscapes. *Nat. Rev. Phys.* **2020**, *2*, 200−212.
(35) Lagardère, L.; Maurin, L.; Adjoua, O.; El Hage, K.; Monmarché, P.; Piquemal, J.-P.; Hénin, J. Lambda-ABF: Simplified, Portable, Accurate, and Cost-Effective Alchemical Free-Energy Computation. *J. Chem. Theory Comput.* **2024**, *20*, 4481−4498.
(36) Procacci, P. Does Hamiltonian Replica Exchange via Lambda-Hopping Enhance the Sampling in Alchemical Free Energy Calculations? *Molecules* **2022**, *27*, 4426.
(37) Zhang, S.; Hahn, D. F.; Shirts, M. R.; Voelz, V. A. Expanded Ensemble Methods Can Be Used to Accurately Predict Protein-Ligand Relative Binding Free Energies. *J. Chem. Theory Comput.* **2021**, *17*, 6536−6547.
(38) Jiang, W.; Thirman, J.; Jo, S.; Roux, B. Reduced Free Energy Perturbation/Hamiltonian Replica Exchange Molecular Dynamics Method with Unbiased Alchemical Thermodynamic Axis. *J. Phys. Chem. B* **2018**, *122*, 9435−9442.
(39) Wan, S.; Tresadern, G.; Pérez-Benito, L.; van Vlijmen, H.; Coveney, P. V. Accuracy and Precision of Alchemical Relative Free Energy Predictions with and without Replica-Exchange. *Adv. Theory Simul.* **2020**, *3*, 1900195.
(40) Kong, X.; Brooks, C. L., III Λ-Dynamics: A New Approach to Free Energy Calculations. *J. Chem. Phys.* **1996**, *105*, 2414−2423.
(41) Hayes, R. L.; Cervantes, L. F.; Abad Santos, J. C.; Samadi, A.; Vilseck, J. Z.; Brooks, C. L. I. How to Sample Dozens of Substitutions per Site with λ Dynamics. *J. Chem. Theory Comput.* **2024**, *20*, 6098−6110.
(42) Hsu, W.-T.; Shirts, M. R. Replica Exchange of Expanded Ensembles: A Generalized Ensemble Approach with Enhanced Flexibility and Parallelizability. *J. Chem. Theory Comput.* **2024**, *20*, 6062−6081.
(43) Hayes, R. L.; Buckner, J.; Brooks, C. L. I. BLaDE: A Basic Lambda Dynamics Engine for GPU-Accelerated Molecular Dynamics Free Energy Calculations. *J. Chem. Theory Comput.* **2021**, *17*, 6799−6807.
(44) Robo, M. T.; Hayes, R. L.; Ding, X.; Pulawski, B.; Vilseck, J. Z. Fast Free Energy Estimates from λ-Dynamics with Bias-Updated Gibbs Sampling. *Nat. Commun.* **2023**, *14*, 8515.
(45) Raman, E. P.; Paul, T. J.; Hayes, R. L.; Brooks, C. L. I. Automated, Accurate, and Scalable Relative Protein−Ligand Binding Free-Energy Calculations Using Lambda Dynamics. *J. Chem. Theory Comput.* **2020**, *16*, 7895−7914.
(46) Hu, R.; Zhang, J.; Kang, Y.; Wang, Z.; Pan, P.; Deng, Y.; Hsieh, C.-Y.; Hou, T. Comprehensive, Open-Source, and Automated Workflow for Multisite λ-Dynamics in Lead Optimization. *J. Chem. Theory Comput.* **2024**, *20*, 1465−1478.
(47) Knight, J. L.; Brooks, C. L. I. Multisite λ Dynamics for Simulated Structure−Activity Relationship Studies. *J. Chem. Theory Comput.* **2011**, *7*, 2728−2739.
(48) Ding, X.; Hayes, R. L.; Vilseck, J. Z.; Charles, M. K.; Brooks, C. L. CDOCKER and λ-Dynamics for Prospective Prediction in D3R Grand Challenge 2. *J. Comput. Aided Mol. Des.* **2018**, *32*, 89−102.
(49) Champion, C.; Gall, R.; Ries, B.; Rieder, S. R.; Barros, E. P.; Riniker, S. Accelerating Alchemical Free Energy Prediction Using a Multistate Method: Application to Multiple Kinases. *J. Chem. Inf. Model.* **2023**, *63*, 7133−7147.
(50) Brooks, B.; Brooks, C.; MacKerell, A.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.; Post, C.; Pu, J.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(51) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13*, No. e1005659.
(52) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(53) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012**, *4*, 17.
(54) He, X.; Man, V. H.; Yang, W.; Lee, T.-S.; Wang, J. A Fast and High-Quality Charge Model for the Next Generation General AMBER Force Field. *J. Chem. Phys.* **2020**, *153*, 114502.
(55) 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.
(56) Gapsys, V.; de Groot, B. L. Pmx Webserver: A User Friendly Interface for Alchemistry. *J. Chem. Inf. Model.* **2017**, *57*, 109−114.
(57) Fleck, M.; Wieder, M.; Boresch, S. Dummy Atoms in Alchemical Free Energy Calculations. *J. Chem. Theory Comput.* **2021**, *17*, 4403−4419.
(58) Mey, A. S. J. S.; Allen, B. K.; Bruce Macdonald, H. E.; Chodera, J. D.; Hahn, D. F.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A.; Scheen, J.; Shirts, M. R.; Tresadern, G.; Xu, H. Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2020**, *2*, 18378.
(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(60) Bauer, P.; Hess, B.; Lindahl, E. *GROMACS 2022.5 Source Code*; Zenodo, 2023.
(61) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(62) Bernetti, M.; Bussi, G. Pressure Control Using Stochastic Cell Rescaling. *J. Chem. Phys.* **2020**, *153*, 114107.
(63) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. New Approach to Monte Carlo Calculation of the Free Energy: Method of Expanded Ensembles. *J. Chem. Phys.* **1992**, *96*, 1776−1783.
(64) Chodera, J. D.; Shirts, M. R. Replica Exchange and Expanded Ensemble Simulations as Gibbs Sampling: Simple Improvements for Enhanced Mixing. *J. Chem. Phys.* **2011**, *135*, 194110.
(65) Liu, J. S. Peskun's Theorem and a Modified Discrete-State Gibbs Sampler. *Biometrika* **1996**, *83*, 681−682.
(66) Belardinelli, R. E.; Pereyra, V. D. Wang-Landau Algorithm: A Theoretical Analysis of the Saturation of the Error. *J. Chem. Phys.* **2007**, *127*, 184105.
(67) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528−1532.
(68) Beckstein, O.; Dotson, D. L.; Wu, Z.; Wille, D.; Marson, D.; Kenney, I.; shuail Lee, H.; Lim, V.; Schlaich, A.; Alibay, I.; Hénin, J.; Barhaghi, M. S.; Merz, P.; Joseph, T.; Hsu, W.-T.; carter, h.; trje3733. *hl2500 Alchemistry/Alchemlyb: 2.3.1*; Zenodo, 2024.
(69) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.
(70) Chodera, J. D. A. Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 1799−1805.
(71) Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. Thermodynamics of Solute Transfer from Water to Hexadecane. *J. Chem. Soc., Perkin Trans.* **1990**, *2*, 291−300.
(72) Åberg, K. M.; Lyubartsev, A. P.; Jacobsson, S. P.; Laaksonen, A. Determination of Solvation Free Energies by Adaptive Expanded Ensemble Molecular Dynamics. *J. Chem. Phys.* **2004**, *120*, 3770−3776.
(73) Invernizzi, M.; Piaggi, P. M.; Parrinello, M. Unified Approach to Enhanced Sampling. *Phys. Rev. X* **2020**, *10*, 041034.
(74) Ries, B.; Alibay, I.; Swenson, D. W. H.; Baumann, H. M.; Henry, M. M.; Eastwood, J. R.; Gowers, R. J. *Kartograf: An Accurate Geometry-Based Atom Mapper for Hybrid Topology Relative Free Energy Calculations*, 2023.
(75) Liu, S.; Wu, Y.; Lin, T.; Abel, R.; Redmann, J. P.; Summa, C. M.; Jaber, V. R.; Lim, N. M.; Mobley, D. L. Lead Optimization Mapper: Automating Free Energy Calculations for Lead Optimization. *J. Comput. Aided Mol. Des.* **2013**, *27*, 755−770.
(76) Xu, H. Optimal Measurement Network of Pairwise Differences. *J. Chem. Inf. Model.* **2019**, *59*, 4720−4728.
(77) Ding, X.; Drohan, J. *Principled Approach for Computing Free Energy on Perturbation Graphs with Cycles*, 2024.

https://doi.org/10.1021/acs.jctc.4c01268
J. Chem. Theory Comput. 2025, 21, 230−240