Supporting Information

Taming Rugged Free-Energy Landscapes using an Average Force

Haohao Fu[†], Xueguang Shao[†,‡], Wensheng Cai*[,†] and Christophe Chipot*[,#,§]

[†]Tianjin Key Laboratory of Biosensing and Molecular Recognition, Research Center for Analytical Sciences, College of Chemistry, Nankai University, Tianjin 300071, China

[‡]State Key Laboratory of Medicinal Chemical Biology, Tianjin 300071, China

[#]Laboratoire International Associé CNRS and University of Illinois at Urbana−Champaign, UMR n°7019, Université de Lorraine, BP 70239, F-54506 Vandœuvre-lès-Nancy, France

[§]Department of Physics, University of Illinois at Urbana−Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States

Algorithms not belonging to importance sampling

Milestoning. A number of nonintersecting hypersurfaces or slices of the free-energy surface are chosen to describe the reaction, referred to as "milestones".[1] A trajectory around a given milestone is collected until sampling reaches an adjacent one. The local sampling around each milestone is then gathered together to capture the global kinetics through post-treatment. It should be noted that although a transition coordinate is usually employed to help define the milestones, it is not a requirement of the method.

Generalized ensembles and weighted ensembles. In both generalized and weighted ensembles, the probability of each state is weighted by a factor, rendering sampling of configurational space uniform. High-energy states, which are notoriously difficult to be explore by means of Boltzmann sampling, are, therefore, more likely to be captured in generalized and weighted ensembles. In generalized-ensemble algorithms, such as replica-exchange molecular dynamics (REMD),[2] replica-exchange simulated tempering (REST)[3] and replica exchange with solute scaling (REST2),[4] different simulations with distinct temperatures are performed simultaneously. Low-temperature replicas can receive configurations from the high-temperature ones, hence improving the sampling ability of the former. In stark contrast, weighted-ensemble schemes like accelerated MD (aMD),[5] Gaussian accelerated MD (GaMD)[6] and integrated tempering enhanced sampling (ITS),[7] morph the potential energy landscape to make high-energy states more favorable with respect to the same ones on the original surface. Since the generalized- and weighted-ensemble schemes aim at achieving a random walk in the potential-energy space, the use of a transition coordinate is not needed for these algorithms.

Importance splitting is a class of algorithm that is mostly used with the Monte-Carlo approach,[8,9] and the discussion of which goes beyond the scope of this paper.

Importance-sampling algorithms other than ABF

Umbrella Sampling. In the original algorithm, an external potential is added to the potential energy function, resulting in an effectively flat free-energy landscape. Staging or stratifying is added by means of confinement potentials to improve further sampling.[10] In practice, since guessing the umbrella potential is difficult, the algorithm has evolved towards an over-stratification approach, whereby the reaction pathway is broken down in a large number of very narrow windows, or strata, in which sampling is confined by means of a harmonic potential,[11]

$$U = k\left(\xi - \xi_0\right)^2 \tag{S1}$$

where ξ is the transition coordinate. The restraint guarantees that sampling along ξ is close to ξ_0. The free-energy difference in each of the windows is less than $k_\mathrm{B}T$, where k_B is the Boltzmann constant and T, the temperature. The overall free-energy change along ξ can be obtained via the weighted histogram analysis method (WHAM),[12] or umbrella integration.[13]

Metadynamics. Metadynamics (MtD) adds Gaussian potentials to the oversampled region to help sampling escape free-energy minima, until all the valleys of the free-energy landscape are flooded,[14–16] as shown in Figure 1C,

$$U = \sum_t V(t) \tag{S2}$$

where $V(t)$ is the potential added to the potential energy function at time t. The true free-energy change along ξ can be recovered by a time-independent estimator.

The choice of an enhanced-sampling algorithm highly depends on the inherent properties of the molecular object of interest and the degrees of freedom at play. Since generalized- and weighted-ensemble schemes do not require a predefined transition coordinate, they may be used to explore a process that cannot be easily described by collective variables. However,

sampling in a generalized- or weighted-ensemble simulation is very random. It is, therefore, plausible that the process of interest will not be observed over the timescale of the simulation, especially in the case of complex molecular objects. On the contrary, sampling along the transition coordinate is guaranteed in importance-sampling simulations, in the limit of timescale separation. Hence, importance-sampling algorithms are often more cost-effective when a reasonable transition coordinate is available.

Ease to use of importance-sampling algorithms

In general, for simple molecular objects and structural transformations, ABF, eABF, meta-eABF/WTM-eABF and MtD/WT-MtD are equally easy to use, since no stratification strategy is needed and the complete free-energy profile can be determined from a single simulation, while US requires the most significant human intervention. For molecular objects and geometric transformations of moderate complexity, the use of meta-eABF/WTM-eABF may be the easiest, because the transition coordinate can span a single, large window, whereas a stratification strategy and multiple simulations are needed for most other ABF-based algorithms. MtD/WT-MtD may also be used without stratification strategy if the size of the Gaussians is carefully determined. For a very complex task, the ease or difficulty of using a given algorithm highly depends on the properties of the molecular object, and how the transition coordinate is coupled to the slowest degrees of freedom, i.e., timescale separation. (Table S1).

Table S1. Human intervention needed for running a simulation using different importance-sampling algorithms. Part of the human time can be spared depending on the molecular simulation and free-energy calculation engine.

Algorithm	Human intervention
ABF	• Merging windows if a stratification strategy is applied • Integration of free-energy gradients[a]
eABF	• Merging windows if a stratification strategy is applied • Integration of free-energy gradients[a]

meta-eABF/WTM-eABF	• Merging windows if a stratification strategy is applied[b] • Integration of free-energy gradients[a]
MtD/WT-MtD	• Optimize the size of the Gaussians for non-trivial molecular objects and structural transformations • Calculating free-energy profiles using a time-independent estimator[c]
US	• Extracting initial structures for each window from a steered MD trajectory, followed by appropriate equilibration simulations • Preparing and running multiple simulations[d] • Recovering the free-energy profiles by post-treatment[e]

[a]Automated in Colvars.
[b]Usually not needed.
[c]Implemented in Plumed.
[d]The number of simulations needed in US is usually much larger than that in ABF or MtD.
[e]By WHAM[12] or umbrella integration.[13]

Discussion of the choice of transition coordinates

In general, a good transition coordinate can describe all (or most) of the slow degrees of freedom of a given molecular object. Figure S1 provides two examples of a good and a bad transition coordinate.

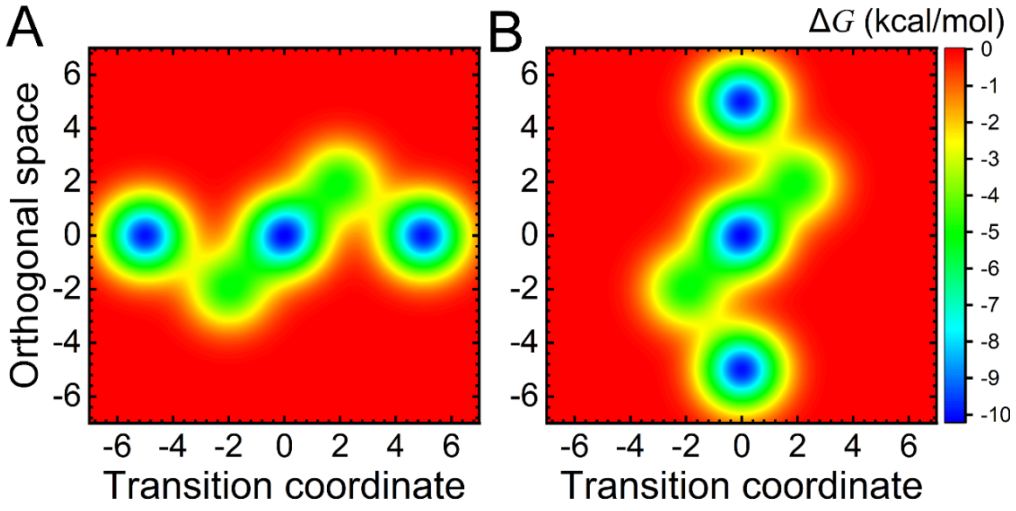

Figure S1. Graphical representation of a good (A) and a bad (B) transition coordinate. The former can characterize the slow degrees of freedom of the molecular object while the latter cannot.

A coarse method to examine the quality of a transition coordinate is to observe sampling along it during an ABF/eABF/meta-eABF/WTM-eABF simulation. A homogeneous, uniform sampling is an indicator of an appropriate transition coordinate. To rigorously check whether a chosen transition coordinate, ξ, is appropriate, a committor analysis ought to be performed. The committor function, p_{B}, at $\xi = \xi^{\ddagger}$, is defined as the probability that a trajectory will reach the product (B), before returning to the reactant (A), starting from the transition state. If the distributions of p_{B} are Gaussian-like with a peak at p_{B} = 0.5, the choice of transition coordinate is appropriate. A detailed introduction of committor analysis can be found elsewhere.[17,18]

According to eq. 1 in the main text, in the original ABF algorithm, the calculation of $\mathbf{F}_{\mathrm{bias}}$ requires the determination of $\left\langle \frac{\partial \ln|\mathbf{J}|}{\partial \xi} \right\rangle_{\xi}$, which involves the second derivative of the collective variable (CV). If one wants to implement a new CV in the context of an ABF calculation, the analytical expression of the CV, its first derivative and its second derivative must be implemented, which can be complicated in practice. eABF (and meta-eABF/WTM-eABF) obviates the need for an explicit analytical determination of $|\mathbf{J}|$, as mentioned in the main text, making knowledge of the second derivative of the CVs unnecessary. Moreover, in the current version of Colvars and Plumed, the first derivative of a CV can be automatically calculated by the built-in lepton library. Implementing a new CV for eABF/meta-eABF/WTM-eABF is, therefore, very straightforward in a modern free-energy calculation engine. It should be noted that eABF, meta-eABF and WTM-eABF are compatible with all the CVs already implemented in Colvars and Plumed, while ABF cannot be used with some of them due to the missing implementation of $\left\langle \frac{\partial \ln|\mathbf{J}|}{\partial \xi} \right\rangle_{\xi}$.

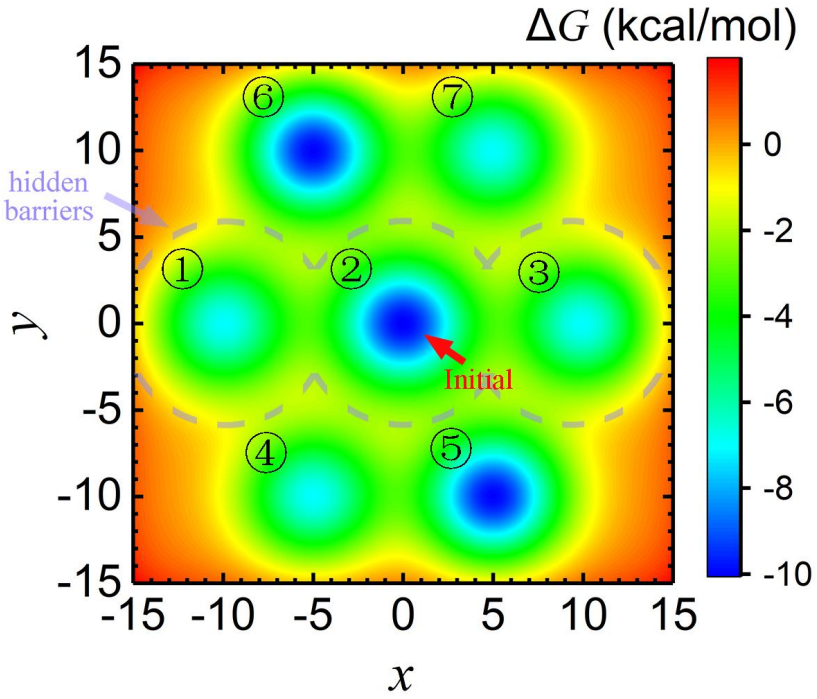

Figure S2. A *Gedanken* experiment showing the "hidden barriers". A hypothetical importance-sampling algorithm is used to explore the two-dimensional surface. *x* is selected as the one-dimensional transition coordinate, while *y* mirrors the orthogonal space. The initial state is located in basin 2. Since the importance-sampling algorithm accelerates the exploration along the transition coordinate, namely *x*, basin 1-3 can be easily investigated. Sampling of other local minima, however, is hampered due to the barriers along the *y*-direction. Two major basins, namely 5 and 6, therefore, may not be found during the simulation.

Detailed Introduction of Replica-Exchange WTM-eABF (REX-WTM-eABF) Algorithm

Approximately, any extended-Lagrangian-based calculation can be regarded as an umbrella-sampling-like simulation. The umbrella-integration unbiased estimator, can therefore, be used to calculate the free-energy landscape.[19,20] Here, the Metropolis criterion used in replica-exchange umbrella sampling (REUS)[21] is applied in an replica-exchange WTM-eABF (REX-WTM-eABF) algorithm. Similar criteria have also been adopted in other extended-Lagrangian-based methods, such as multi-scale sampling using temperature accelerated and replica exchange molecular dynamics（MuSTAR MD).[22]

In REX-WTM-eABF, the probability of exchanging two replicas is,

$$p(\text{exchange}) = \begin{cases} 1 \ (\Delta \leq 0) \\ e^{-\Delta} \ (\Delta > 0) \end{cases} \tag{S3}$$

and

$$\Delta = \beta_m(E_n - E_m) - \beta_n(E_n - E_m) \tag{S4}$$

where E is the sum of the force field energy and the elastic potential energy of the fictitious spring, and n and m, two replicas.

Comparison between Different Importance-Sampling Algorithms

In plain language, US is a histogram-based scheme, which improves the sampling relying on restraints and a stratification of the reaction pathway. ABF is a thermodynamic-integration-based approach, shaving the free energy barriers by applying biasing forces. MtD fills the valleys by gradually adding Gaussian potentials as the simulation proceeds. Due to these differences, each algorithm has its own advantages and limitations, as detailed in ref. [23].

To assess the computational efficiency of different importance-sampling schemes, we have calculated the free-energy landscapes characterizing the isomerization of deca-alanine and N-acetyl-N'-methylalanylamide (NANMA) in aqueous solution using WTM-eABF,

eABF and WT-MtD. As shown in Figure S3, for the deca-alanine case, the PMF obtained from 100 ns WTM-eABF simulation is almost identical to the reference, while those from eABF and WT-MtD simulations still differ significantly from the original one. For the NANMA case, the WTM-eABF simulation converged in 10 ns, while plain eABF and WT-MtD only sampled the low-free-energy region within the same time scale. The free-energy landscape obtained from the WT-MtD simulation, moreover, is slightly noisy. All in all, WTM-eABF outperforms the other two algorithms in sampling efficiency in both cases, suggestive of a reasonable choice for the standard free-energy calculations.

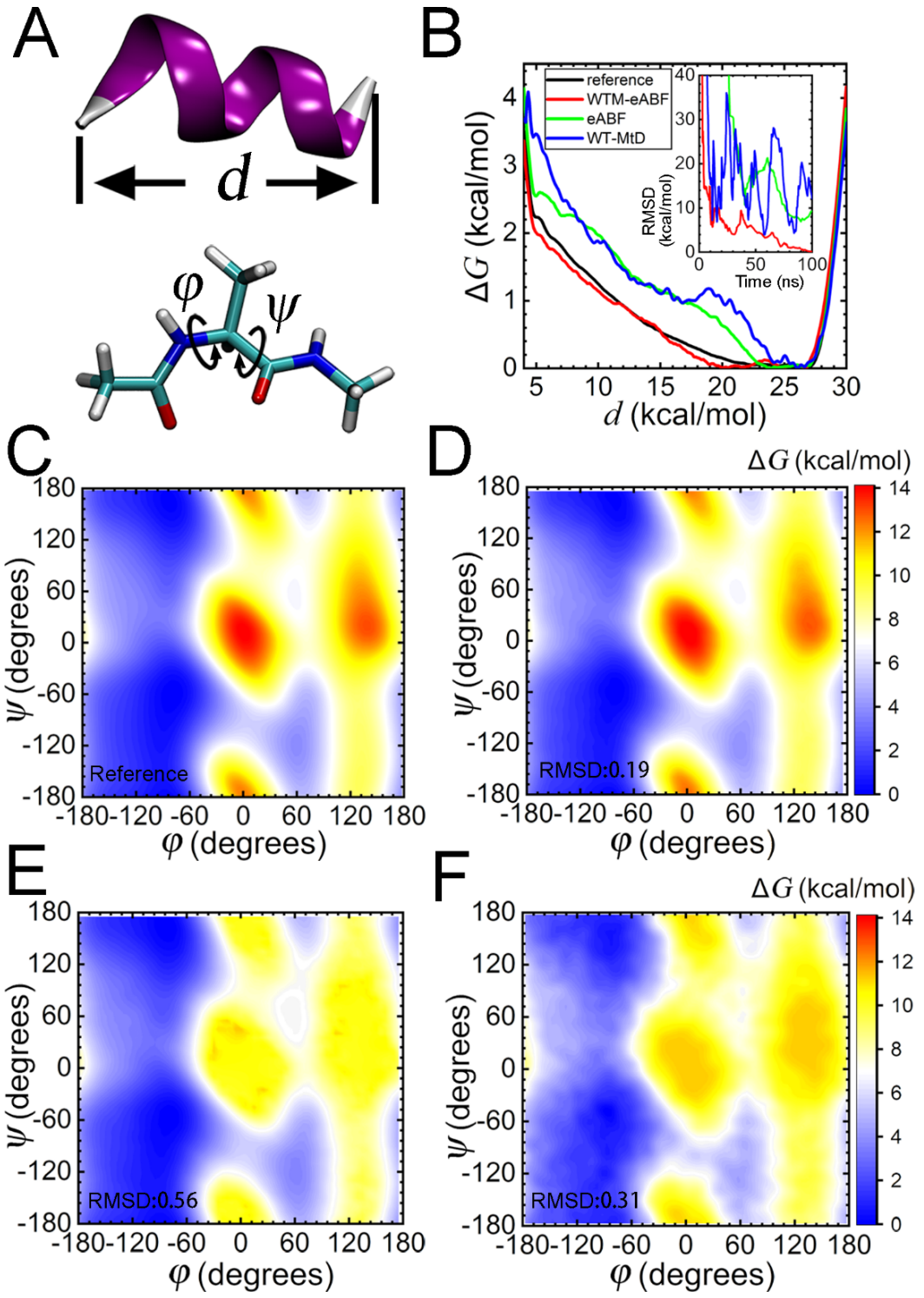

Figure S3. Performance comparison of WTM-eABF, eABF and WT-MtD. Molecular objects and transition coordinates (A). Both of the two molecules are solvated in aqueous solution. Free-energy changes along the end-to-end distance, d, of deca-alanine (B). Inset: Time evolution of the PMF root-mean-square deviation (RMSD) with respect to the reference.

Free-energy landscapes characterizing the isomerization of alanine dipeptide from the reference (C), WTM-eABF (D), eABF (E) and WT-MtD (F) simulations. The RMSD (in kcal/mol) of each two-dimensional free-energy landscape with respect to the reference is shown in the lower left corner of the corresponding panel. Both references are obtained from a more than 1-μs eABF simulation. Other one-dimensional PMFs and two-dimensional free-energy landscapes are from 100- and 10-ns importance-sampling calculations, respectively. Extra efforts have been made to optimize the MtD parameters.

```
colvar {
   name cv1

   ... // boundaries and width of cv1

   extendedlagrangian   on
   extendedFluctuation  0.1
   extendedTimeConstant   200
   expandBoundaries    on
   subtractappliedforce  on

   ... // definition of cv1
}

abf {
  colvars cv1
  fullSamples   500
  writeCZARwindowFile   on
}

metadynamics {
  colvars cv1
  hillWidth         5.0
  hillWeight        0.1
  wellTempered      on
  biasTemperature   4000
  keepFreeEnergyFiles on
}
```

Figure S4. Example configuration file for a WTM-eABF simulation using the Colvars module. Example files for running ABF, eABF, meta-eABF and WTM-eABF simulations are provided in GitHub (https://github.com/fhh2626/ABF-example-files-for-NAMD-and-OpenMM) and in the ZIP archive of the Supporting Information.

```
cv1: ... // definition of cv1

DRR ...
LABEL=eabf
ARG=cv1
FULLSAMPLES=500
KAPPA=2046.9939
FRICTION=1.0
TAU=0.2
TEMP=300
TEXTOUTPUT

... // setting of GRID_MIN, GRID_MAX, GRID_BIN, OUTPUTFREQ AND HISTORYFREQ

... DRR

METAD ...
LABEL=metad
ARG=eabf.cv1_fict
PACE=1000
SIGMA=0.35
HEIGHT=0.4
TEMP=300
BIASFACTOR=14

... // setting of GRID_MIN, GRID_MAX and GRID_BIN

... METAD
```

Figure S5. Example configuration file for a WTM-eABF simulation using the Plumed module. Example files for running eABF, meta-eABF and WTM-eABF simulations are provided in GitHub (https://github.com/fhh2626/ABF-example-files-for-NAMD-and-OpenMM) and in the ZIP archive of the Supporting Information.

Table S2. Standard Binding Free-Energy Calculations of p41:Abl-SH3 Complex using eABF and WTM-eABF

Contribution	eABF (ref .[24])		WTM-eABF (this work)[a]	
	Energy[b] (kcal/mol)	Simulation Time (ns)	Energy (kcal/mol)	Simulation Time (ns)
ΔG_c^{site}	-4.59	15	-4.59	15
$\Delta G_{\Theta}^{\text{site}}$	-0.16	8	-0.18	4
$\Delta G_{\Phi}^{\text{site}}$	-0.53	14	-0.57	7
$\Delta G_{\Psi}^{\text{site}}$	-0.21	12	-0.23	6
$\Delta G_{\theta}^{\text{site}}$	-0.15	4	-0.11	4
$\Delta G_{\varphi}^{\text{site}}$	-0.21	12	-0.21	6
$-\frac{\ln(S^*I^*C^\circ)}{\beta}$	-16.46	40	-16.22	20
ΔG_c^{bulk}	7.77	84	7.27	42
ΔG_o^{bulk}	6.61	-	6.61	-
$\Delta G_{\text{bind}}^{\circ}$	-7.95	189	-8.24	104
$\Delta G_{\text{bind}}^{\circ}$(exp)[c]	-7.99			
$\Delta G_{\text{bind}}^{\circ}$(MM-PBSA)[d]	-2.6			

[a]The result reported herein is from one of the four parallel simulations at T=300 K in Table 2, as a representative.

[b]Except for ΔG_o^{bulk} and ΔG_o^{bulk}, quantities are calculated by integration of the PMFs shown in Figure S6.

[c]From ref. [25].

[d]From ref. [26].

Table S3. Local minima and corresponding free energies in Figure 4A and 4B.

QM/MM (Figure 4A)		MM (Figure 4B)	
position of local mimima (ϕ_1, ϕ_2, ϕ_3) (°)	ΔG (kcal/mol)	position of local mimima (ϕ_1, ϕ_2, ϕ_3) (°)	ΔG(kcal/mol)
$(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3)$ = (-88, -78, -74, 146, 158, 149)[a]	0.0	$(\phi_1, \phi_2, \phi_3, \psi_1, \psi_2, \psi_3)$ = (-67, -66, -70, 150, 161, 156)[a]	0.0
(-80, 70, -70)	1.5	(-70, -150, -70)	1.0
(-80, -80, 70)	2.1	(-70, -70, 60)	1.1
(-80, 70, 70)	2.4	(-70, -70, -150)	1.2
(60, 70, -70)	2.5	(50, -70, -70)	1.2
		(-140, -70, -70)	1.3
		(-60, 50, -70)	1.6
		(-60, 50, 60)	1.9
		(-70, -150, 60)	2.0
		(50, 60, -70)	2.2
		(-150, -150, -70)	2.2
		(60, -70, -150)	2.3
		(50, -70, -140)	2.3
		(-140, -60, 60)	2.3
		(50, 50, 60)	2.4

[a]Additional simulations with a precision of 1° were added to identify the global minimum. It is noteworthy that (ψ_1, ψ_2, ψ_3) fluctuate a lot since they are not slow degrees of freedom.

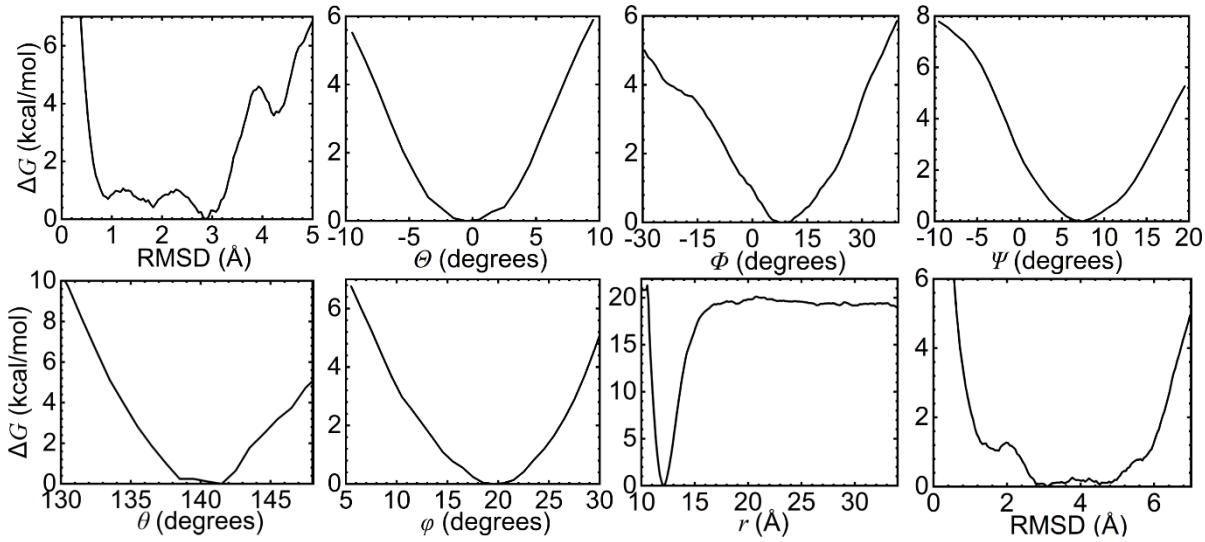

Figure S6. Free-energy profiles of an example of accurate binding free-energy calculation. The corresponding collective variables are shown in Figure 3. Reproduced with permission from ref. [24]. Copyright 2017 American Chemical Society.

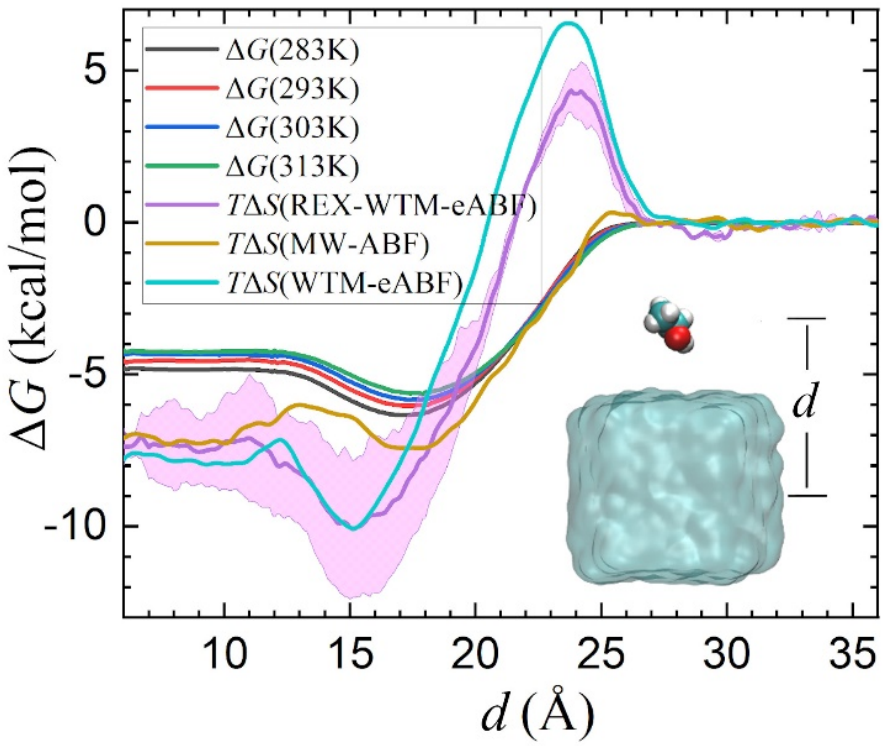

Figure S7. Calculation of $T\Delta S$ ($T = 298$ K) for ethanol hydration by 750-ns REX-WTM-eABF, 2.4-μs MW-ABF and 3-μs WTM-eABF simulations. The error bars for $T\Delta S$ (REX-WTM-eABF) correspond to the standard deviation of the quantity calculated using different values of ΔT. The error bars for $T\Delta S$ (MW-eABF) and $T\Delta S$ (WTM-eABF) are shown in Figure S8. Inset: transition coordinate.

The difficulty of exploring the region at the water-vacuum interface ($10\text{Å} \leq d \leq 24\text{Å}$) makes this model representative for testing the ability to estimate accurately the entropy change of importance-sampling algorithms. An alchemical route for entropy calculation circumvents the challenge of sampling at water-vacuum interface.[27] Therefore, the simulation time consumed by the geometrical and alchemical routes should not be compared directly.

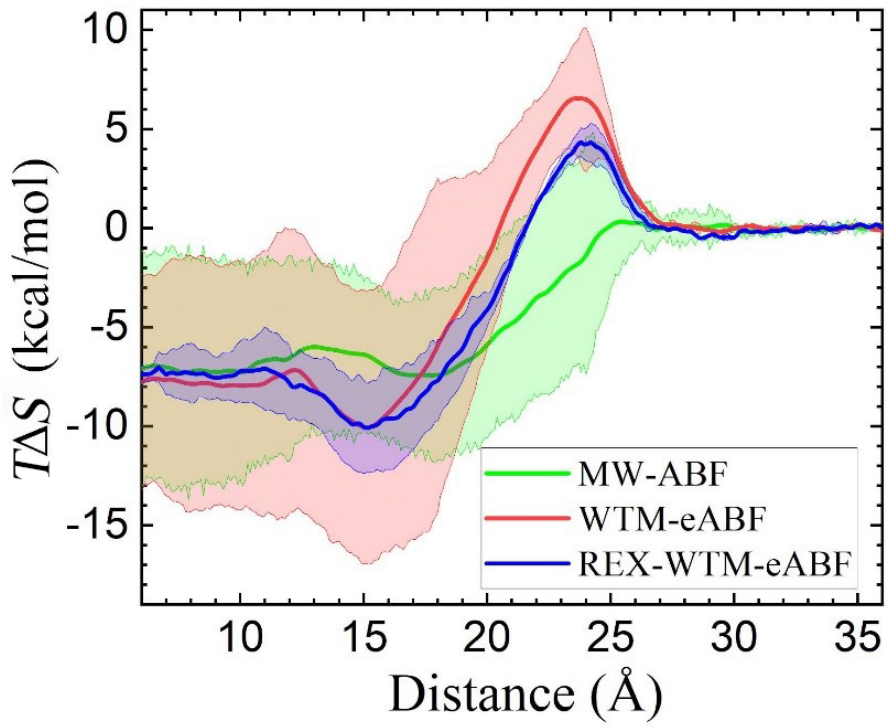

Figure S8. Calculation of $T\Delta S$ ($T = 298$ K) for ethanol hydration. All the curves are identical to those shown in Figure S7. The error bars for $T\Delta S$ (MW-eABF) and $T\Delta S$ (WTM-eABF) are shown in this figure for clarity.

Simulation Details

Table S4. Details of Simulations Performed in This Account

Processes	Models	Ensemble	Number of atoms	Box size (Å^3)
Permeation of 2',3'-dideoxyadenosine across a lipid bilayer	CHARMM[28] + TIP3P[29]	NPT	42093	69×69×86
p41:Abl-SH3 binding			11366	48×48×48
Ethanol hydration			4287	35×35×80
Conformational change of an alanine tripeptide	PM7[30] + TIP3P[29]		4515 (QM:42; MM:4473)	36×37×37

All the atomistic MD simulations reported herein were performed using the parallel, scalable program NAMD.[31] MOPAC[32] was used as the QM engine. Covalent bonds involving hydrogen atoms of (bio)organic molecules were constrained to their equilibrium length by means of the SHAKE/RATTLE[33,34] algorithms while the SETTLE[35] algorithm was used for water molecules. The r-RESPA multiple time-step algorithm[36] was employed to integrate the equations of motion with a time step of 2 and 4 fs for short- and long-range interactions, respectively. The particle mesh Ewald (PME)[37] scheme was utilized to estimate long-range electrostatic forces. A smoothed 12 Å cutoff was applied to truncate short-range electrostatic and van der Waals interactions. Periodic boundary conditions (PBCs) were applied in the three directions. NPT ensemble equilibrations were performed employing Langevin dynamics[38] and the Langevin piston pressure control.[39]

References

(1) West, A. M. A.; Elber, R.; Shalloway, D. Extending Molecular Dynamics Time Scales with Milestoning: Example of Complex Kinetics in a Solvated Peptide. *J. Chem. Phys.* **2007**, *126*, 145104.

(2) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(3) Mitsutake, A.; Okamoto, Y. Replica-Exchange Simulated Tempering Method for Simulations of Frustrated Systems. *Chem. Phys. Lett.* **2000**, *332*, 131–138.

(4) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **2011**, *115*, 9431–9438.

(5) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.

(6) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11*, 3584–3595.

(7) Yang, L.; Liu, C.-W.; Shao, Q.; Zhang, J.; Gao, Y. Q. From Thermodynamics to Kinetics: Enhanced Sampling of Rare Events. *Acc. Chem. Res.* **2015**, *48*, 947–955.

(8) Jegourel, C.; Legay, A.; Sedwards, S. Importance Splitting for Statistical Model Checking Rare Properties BT - Computer Aided Verification; Sharygina, N., Veith, H., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp 576–591.

(9) Cérou, F.; Guyader, A.; Rousset, M. Adaptive Multilevel Splitting: Historical Perspective and Recent Results. *Chaos An Interdiscip. J. Nonlinear Sci.* **2019**, *29*, 43108.

(10) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

(11) Kästner, J. Umbrella Sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932–942.

(12) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(13) Kästner, J.; Thiel, W. Bridging the Gap between Thermodynamic Integration and Umbrella Sampling Provides a Novel Analysis Method: "Umbrella Integration." *J. Chem. Phys.* **2005**, *123*, 144104.

(14) Huber, T.; Torda, A. E.; Van Gunsteren, W. F. Local Elevation: A Method for Improving the Searching Properties of Molecular Dynamics Simulation. *J. Comput. Aided. Mol. Des.* **1994**, *8*, 695–708.

(15) Grubmüller, H. Predicting Slow Structural Transitions in Macromolecular Systems: Conformational Flooding. *Phys. Rev. E* **1995**, *52*, 2893–2906.

(16) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566.

(17) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transion Path Sampling: Throwing Ropes Over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.

(18) Pan, A. C.; Sezer, D.; Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B* **2008**, *112*, 3432–3440.

(19) Zheng, L.; Yang, W. Practically Efficient and Robust Free Energy Calculations: Double-Integration Orthogonal Space Tempering. *J. Chem. Theory Comput.* **2012**, *8*, 810–823.

(20) Chen, H.; Fu, H.; Shao, X.; Chipot, C.; Cai, W. ELF: An Extended-Lagrangian Free Energy Calculation Module for Multiple Molecular Dynamics Engines. *J. Chem. Inf. Model.* **2018**, *58*, 1315–1318.

(21) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

(22) Yamamori, Y.; Kitao, A. MuSTAR MD: Multi-Scale Sampling Using Temperature Accelerated and Replica Exchange Molecular Dynamics. *J. Chem. Phys.* **2013**, *139*, 145105.

(23) Fu, H.; Shao, X.; Chipot, C.; Cai, W. Extended Adaptive Biasing Force Algorithm. An On-the-Fly Implementation for Accurate Free-Energy Calculations. *J. Chem. Theory Comput.* **2016**, *12*, 3506–3513.

(24) Fu, H.; Cai, W.; Hénin, J.; Roux, B.; Chipot, C. New Coarse Variables for the Accurate Determination of Standard Binding Free Energies. *J. Chem. Theory Comput.* **2017**, *13*, 5173–5178.

(25) Pisabarro, M. T.; Serrano, L. Rational Design of Specific High-Affinity Peptide Ligands for the Abl-SH3 Domain. *Biochemistry* **1996**, *35*, 10634–10640.

(26) Hou, T.; Chen, K.; McLaughlin, W. A.; Lu, B.; Wang, W. Computational Analysis and Prediction of the Binding Motif and Protein Interacting Partners of the Abl SH3 Domain. *PLOS Comput. Biol.* **2006**, *2*, 1–10.

(27) Comer, J.; Roux, B.; Chipot, C. Achieving Ergodic Sampling Using Replica-Exchange Free-Energy Calculations. *Mol. Simul.* **2014**, *40*, 218–228.

(28) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell Jr, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2016**, *14*, 71.

(29) 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.

(30) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters. *J. Mol. Model.* **2013**, *19*, 1–32.

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

(32) Stewart, J. J. P. MOPAC: A Semiempirical Molecular Orbital Program. *J. Comput. Aided. Mol. Des.* **1990**, *4*, 1–103.

(33) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(34) Andersen, H. C. Rattle: A "Velocity" Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.

(35) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(36) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992**, *97*, 1990–2001.

(37) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(38) Uhlenbeck, G. E.; Ornstein, L. S. On the Theory of the Brownian Motion. *Phys. Rev.* **1930**, *36*, 823–841.

(39) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.