



Use of multistate Bennett acceptance ratio method for free-energy calculations from enhanced sampling and free-energy perturbation

Yasuhiro Matsunaga¹ · Motoshi Kamiya² · Hiraku Oshima³ · Jaewoon Jung⁴ · Shingo Ito⁵ · Yuji Sugita^{3,4,5}

Received: 24 October 2022 / Accepted: 1 December 2022 / Published online: 14 December 2022
© The Author(s) 2022

Abstract

Multistate Bennett acceptance ratio (MBAR) works as a method to analyze molecular dynamics (MD) simulation data after the simulations have been finished. It is widely used to estimate free-energy changes between different states and averaged properties at the states of interest. MBAR allows us to treat a wide range of states from those at different temperature/pressure to those with different model parameters. Due to the broad applicability, the MBAR equations are rather difficult to apply for free-energy calculations using different types of MD simulations including enhanced conformational sampling methods and free-energy perturbation. In this review, we first summarize the basic theory of the MBAR equations and categorize the representative usages into the following four: (i) perturbation, (ii) scaling, (iii) accumulation, and (iv) full potential energy. For each, we explain how to prepare input data using MD simulation trajectories for solving the MBAR equations. MBAR is also useful to estimate reliable free-energy differences using MD trajectories based on a semi-empirical quantum mechanics/molecular mechanics (QM/MM) model and ab initio QM/MM energy calculations on the MD snapshots. We also explain how to use the MBAR software in the GENESIS package, which we call *mbar_analysis*, for the four representative cases. The proposed estimations of free-energy changes and thermodynamic averages are effective and useful for various biomolecular systems.

Keywords Multistate Bennett acceptance ratio · Umbrella sampling · Replica exchange molecular dynamics · Free-energy perturbation · Molecular dynamics · Enhanced conformational sampling · Free-energy calculation · QM/MM calculations

Introduction

Multistate Bennett acceptance method (MBAR) (Shirts and Chodera 2008) is widely used to estimate free energy changes between different states and thermodynamic averages at the states of interest using simulation trajectory data

after the simulations have been finished. MBAR is an extension of Bennett's acceptance ratio method (Bennett 1976) that was formulated for only two states to multiple states. Theoretically, the estimation using MBAR is superior to other estimators in that it has the lowest variance and is asymptotically unbiased (Shirts and Chodera 2008). MBAR can be interpreted as the limit of infinitesimal bin size in the weighted histogram analysis method (WHAM) (Kumar et al. 1992), (Souaille and Roux 2001), (Tan et al. 2012). It can avoid any biases introduced by the binning of trajectory data.

MBAR covers a broad range of applications where thermodynamic ensembles are required. They include the estimations of the free-energy differences including absolute/relative binding free-energy calculations (Chodera et al. 2011) and the ensemble averages including the potential of mean force (PMF) calculations from biased molecular dynamics (MD) trajectories, such as umbrella sampling (Torrie and Valleau 1977), or temperature replica-exchange MD (T-REMD) (Sugita and Okamoto 1999). Furthermore, MBAR has recently been applied to the rapid

✉ Yuji Sugita
sugita@riken.jp

¹ Graduate School of Science and Engineering, Saitama University, Saitama, Saitama 338-8570, Japan

² Institute for Molecular Science, Myodaiji, Okazaki, Aichi 444-8585, Japan

³ Laboratory for Biomolecular Function Simulation, RIKEN Center for Biosystems Dynamics Research, Kobe, Hyogo 650-0047, Japan

⁴ Computational Biophysics Research Team, RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan

⁵ Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan

parameterization of all-atom force-field parameters (Messerly et al. 2018) and coarse-grained model parameters (Shinobu et al. 2019).

One of the practical difficulties in MBAR is to solve the nonlinear simultaneous equations numerically. This requires a sufficient memory capacity on the order of the square of the number of states multiplied by the number of samples. The computational complexity in solving the equations drastically increases with the number of states, which is a bottleneck in large-scale applications with MBAR. To solve this problem, Zhang et al. (2015) and Tan et al. (2016) proposed stochastic solvers, where locally weighted histogram analysis is used to approximate the global solution of the MBAR equations. Recently, Ding et al. developed a rapid solver inspired by the fact that the MBAR equations can be derived as a Rao-Blackwell estimator (Ding et al. 2019).

Another difficulty in using the MBAR equations is how to prepare the input data using different types of MD simulation trajectories. The MBAR equations require the potential energies and thermodynamic parameters of the system as input data. However, in many practical applications, it is often not necessary to use all the potential energy terms in the MBAR equations. For example, in the case of umbrella sampling (US), the potential energies of the system excluding the restraint energies are canceled out in the MBAR equations. The restraint energies alone are sufficient as input data. Another example is the case of MD simulations at constant volume, where pV terms can be canceled out; thus, it is not necessary for solving the MBAR equations. Because it is difficult for non-experts to recognize which terms can be canceled out, users may attempt to prepare whole system's potential energies by themselves. This leads to tedious efforts in preparing unnecessary input data and limits the uses of the MBAR by non-expert users.

In this review, we summarize the basic theory of the MBAR equations and describe representative usages for different types of MD simulation trajectories, such as enhanced conformational sampling methods like umbrella sampling (US) (Torrie and Valleau 1977), temperature replica-exchange MD (T-REMD) (Sugita and Okamoto 1999), and replica exchange with solute tempering (REST/REST2/gREST) (Liu et al. 2005) (Wang et al. 2011) (Terakawa et al. 2011) (Kamiya and Sugita 2018), and free-energy perturbation (FEP) (Zwanzig 1954) (Tembre and Mc Cammon 1984) (Mey et al. 2020). MBAR is also useful for estimating more reliable free-energy changes using MD simulations based on semi-empirical quantum mechanics/molecular mechanics (QM/MM) and ab initio QM/MM calculations (Warshel and Levitt 1976) on the MD snapshots (Yagi et al. 2021). As examples, we present several computational results using the MBAR equations for free-energy estimator for simple biological systems. Finally, we will discuss our perspectives on free-energy estimation methods.

The classifications of MBAR applications

The MBAR equations

First, we define $\mathbf{x} \in \Gamma$ as the configurations of the target system to be simulated. Here, Γ is the configuration space. We also define that $\mathbf{n}(\mathbf{x})$ is the number of molecules of each of M components of the system, and $\boldsymbol{\mu}$ is the vector of chemical potentials of the corresponding components. In the MBAR formulation, state i is specified by a combination of potential energy function $U_i(\mathbf{x})$, inverse temperature β_i , pressure p_i , and/or chemical potential(s) $\boldsymbol{\mu}_i$, or their conjugate variables depending on the ensemble. Here, the inverse temperature β_i , the pressure p_i , and/or the chemical potential(s) $\boldsymbol{\mu}_i$ are given as external parameters in the input of MD simulation. For example, the inverse temperature β_i corresponds to that of the thermostat. Suppose that we obtain N_i uncorrelated samples in state i out of K states, and for each state i , $U_i(\mathbf{x})$, β_i , p_i , and $\boldsymbol{\mu}_i$ are defined. Then, the reduced potential energy $u_i(\mathbf{x})$ of state i is then defined by Shirts and Chodera (2008):

$$u_i(\mathbf{x}) = \beta_i [U_i(\mathbf{x}) + p_i V(\mathbf{x}) + \boldsymbol{\mu}_i^T \mathbf{n}(\mathbf{x})] \quad (1)$$

where $\mathbf{n}(\mathbf{x})$ are the numbers of molecules corresponding to the chemical potentials.

The dimensionless free energy of state i can be defined using the integral of the Boltzmann weight:

$$f_i = -\ln \int e^{-u_i(\mathbf{x})} d\mathbf{x} \quad (2)$$

Note that the above equations can use arbitrary weights other than the Boltzmann factor. The free energy difference between two states, $\Delta f_{ij} = f_i - f_j$, is estimated using the MBAR equations as discussed below.

The MBAR equations give us the best estimator for free energy differences between different states from MD simulation data. Let $\{\mathbf{x}_{in}\}_{n=1}^{N_i}$ be configurations sampled in state i , then the equations are given by:

$$\hat{f}_i = -\ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{\exp[-u_i(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_{jn})]} \quad (3)$$

Here, \hat{f}_i are the solutions to the MBAR equations. \hat{f}_i are determined up to an additive constant, so only their differences $\Delta \hat{f}_{ij} = \hat{f}_i - \hat{f}_j$ are meaningful. Since \hat{f}_i is found on the right-hand side of Eq. (3) as well as the left-hand side, the equation is solved by self-consistent iterations or the Newton–Raphson method until the convergence of \hat{f}_j . On the other hand, when the samples from different states do not have any overlaps in the configuration space, \hat{f}_i in Eq. 3 fail to converge, or \hat{f}_i could have large uncertainties even after convergence. The uncertainties can be estimated by

the analytical equation (given in (Shirts and Chodera 2008)) or by using the bootstrap method or the block averaging method. When the hidden energy barriers exist in a full dimensional space, the uncertainty in \hat{f}_i in the MBAR equation cannot decrease rapidly even with extended MD simulation trajectory data. To overcome this, better conformational sampling schemes applicable to higher dimensional spaces are necessary.

Once \hat{f}_i is obtained, thermodynamic averages in the unsampled state can be calculated by “extrapolating” or “interpolating” the MBAR equations from the samples obtained in simulations with K states. Let $u_{\text{target}}(\mathbf{x})$ be the reduced potential energy in the target state where the average $\langle A \rangle_{\text{target}}$ of the physical quantity $A(\mathbf{x})$ will be estimated. $\langle A \rangle_{\text{target}}$ can be derived as follows:

$$\langle A \rangle_{\text{target}} = \frac{\int A(\mathbf{x})e^{-u_{\text{target}}(\mathbf{x})} d\mathbf{x}}{\int e^{-u_{\text{target}}(\mathbf{x})} d\mathbf{x}} = \frac{\exp[-f_{\text{target},A}]}{\exp[-f_{\text{target}}]} \tag{4}$$

Here, $f_{\text{target},A}$ uses a non-Boltzmann weight because of the factor $A(\mathbf{x})$. The estimates of the free energies, $f_{\text{target},A}$, f_{target} , $\hat{f}_{\text{target},A}$, and \hat{f}_{target} , can be obtained by solving the above MBAR equations by regarding the ensembles as $(K + 1)$ -th and $(K + 2)$ -th states with $N_{K+1} = 0$ and $N_{K+2} = 0$, respectively:

$$\hat{f}_{\text{target}} = -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} \exp[-u_{\text{target}}(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_{jn})]} \tag{5}$$

$$\hat{f}_{\text{target},A} = -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} A(\mathbf{x}_{jn}) \exp[-u_{\text{target}}(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - u_k(\mathbf{x}_{jn})]} \tag{6}$$

When the delta function (or indicator function) on order parameter(s) (or collective variable(s)) is used as the physical quantity $A(\mathbf{x})$, then the PMF on those coordinate(s) can be obtained.

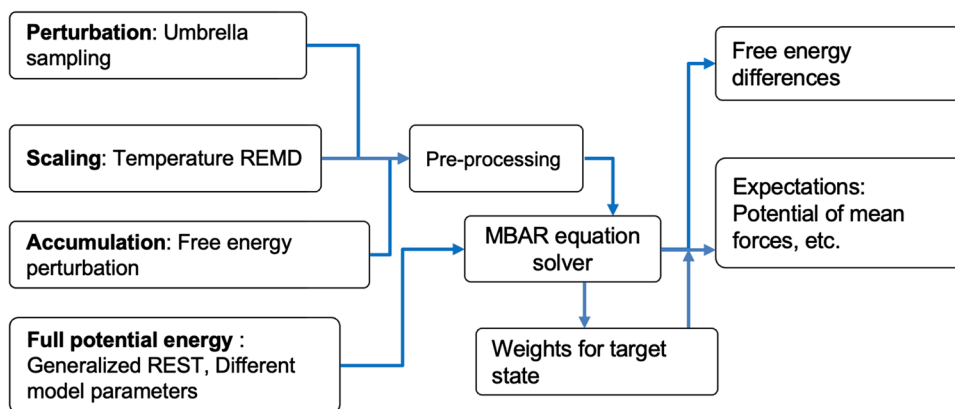
In principle, the MBAR equations can utilize the potential energy $U_i(\mathbf{x})$, volume $V(\mathbf{x})$, and other thermodynamic data as input data. However, evaluating the whole system’s potential energies for all combinations at different states is a time-consuming process. Obviously, some applications can simplify or skip this process. For this purpose, we categorize four major usages of the MBAR equations: (i) perturbation, (ii) scaling, (iii) accumulation, and (iv) full potential energy as shown in Fig. 1. In the following, we describe the reduced MBAR equations and input data for each case.

Perturbation: umbrella sampling

Here, we mean that “perturbation” is to add an additional term to the original potential energy that is used in MD simulations. Umbrella sampling (Torrie and Valleau 1977) imposes a restraint force(s) along collective variable(s) $\mathbf{z}(\mathbf{x})$ and is considered as a major example of “perturbation.” The potential energy used in umbrella sampling can be decomposed to $U_i(\mathbf{x}) = U_{\text{system}}(\mathbf{x}) + U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}))$. Here, $U_{\text{system}}(\mathbf{x})$ is the potential energy from the system’s force-field shared at all the umbrella windows, and $U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}))$ is a restraint energy for the i -th umbrella window. The quadratic form $k_i(\mathbf{z}(\mathbf{x}) - \mathbf{z}_i^{\text{ref}})^2$ is usually used for $U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}))$. In the context of umbrella sampling, states i in the MBAR analysis are specified by $U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}))$. When $U_i(\mathbf{x})$ is used in the MBAR equations, the terms of $U_{\text{system}}(\mathbf{x})$ are canceled out, so only $U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}))$ is needed as input. To simplify further, only k_i , $\mathbf{z}_i^{\text{ref}}$, and $\mathbf{z}(\mathbf{x}_{jn})$ are needed as the input data. The resulting MBAR equations for the (dimensionless) free energies of umbrella window systems are:

$$\begin{aligned} \hat{f}_i &= -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} \exp[-\beta U_{\text{restraint}}^{(i)}(\mathbf{z}(\mathbf{x}_{jn}))]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - \beta U_{\text{restraint}}^{(k)}(\mathbf{z}(\mathbf{x}_{jn}))]} \\ &= -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} \exp[-\beta k_i (\mathbf{z}(\mathbf{x}_{jn}) - \mathbf{z}_i^{\text{ref}})^2]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - \beta k_k (\mathbf{z}(\mathbf{x}_{jn}) - \mathbf{z}_k^{\text{ref}})^2]} \end{aligned} \tag{7}$$

Fig. 1 Classifications of four major applications for the MBAR equations and the schemes from input data to output results



Typically, umbrella sampling aims to obtain the PMF force along z under the ensemble of the target state as the restraint-free state. In this case, $u_{\text{target}}(\mathbf{x}) = 0$ is used as the reduced energy, which can be calculated without further input data.

Scaling: temperature replica exchange MD

In “scaling,” a scaling of the original potential energy by a factor is applied in MD simulations. We here consider T-REMD as an example of “scaling.” In the context of T-REMD, states i in the MBAR analysis are specified by β_i . The reduced potential energy used in T-REMD is $u_i(x) = \beta_i U_{\text{system}}(\mathbf{x})$, just a scaling of the original potential energy $U_{\text{system}}(\mathbf{x})$. In this case, the resulting MBAR equation is:

$$\hat{f}_i = -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} \exp[-\beta_i U_{\text{system}}(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp[\hat{f}_k - \beta_k U_{\text{system}}(\mathbf{x}_{jn})]} \quad (8)$$

Here, $\beta_i U_{\text{system}}(\mathbf{x}_{jn})$ can be calculated by multiplying $U_{\text{system}}(\mathbf{x}_{jn})$ by β_i . Therefore, the only inputs required for the MBAR equations are the set of $U_{\text{system}}(\mathbf{x}_{jn})$ and β_i . For the calculation of thermodynamic averages or PMF, the target ensemble, i.e., $u_{\text{target}}(\mathbf{x}) = \beta_1 U_{\text{system}}(\mathbf{x}_{jn})$, at the lowest temperature, β_1 . Note that making use of different temperatures like Eq. (8) using the WHAM instead of the MBAR requires an advanced reformulation of the WHAM theory, such as the parallel tempering WHAM (Chodera et al. 2007). Here, we assume the NVT ensemble, but in the NPT ensemble, we must also consider pV as the inputs (Paschek and García 2004) (Peter et al. 2016).

Accumulation: free-energy perturbation

“Accumulation” is defined as calculations that are accumulating or integrating the free energy difference between states. As a typical example of “accumulation,” we here consider free-energy perturbation (FEP) (Zwanzig 1954) (Tembre and McCammon 1984) (Mey et al. 2020), which is typically used to calculate the free energy difference in alchemical transformation. In FEP, multiple intermediate states are stratified with a one-dimensional control parameter λ ($0 \leq \lambda \leq 1$). By stratifying the start ($\lambda = 0$) and end ($\lambda = 1$) states, the distributions of reduced energies of neighboring states over λ can overlap with each other, resulting in more accurate estimates of the free-energy differences between neighboring states, $\Delta f_{i,i+1}$. Theoretically, the free-energy difference between the start and end states is obtained by accumulating the neighboring differences, $\Delta f_{\text{start,end}} = \sum \Delta f_{i,i+1}$. In FEP, the potential energy is described by $U(x; \lambda_i) = (1 - \lambda)U_{\text{start}}(x) + \lambda U_{\text{end}}(x)$. In the context of FEP, states i in the MBAR analysis are specified by $U(x; \lambda_i)$. In the MBAR analysis of FEP data, it is

ideal to prepare the reduced energies for all combinations of stratified states $u_i(\mathbf{x}_{jn})$, and to solve the full MBAR equations. However, for practical uses, it is often sufficient for keeping accuracy of the free-energy difference to care about the reduced energies only between neighboring states. It is thus not worth the cost to calculate the reduced energies for all combinations of the states (Paliwal and Shirts 2011). Also, the difference of the reduced energies between non-neighboring states often becomes too large, leading to the instability of energy calculations. In the case of scaling Lennard–Jones (LJ) interactions of a solute in water, the solute can overlap with water molecules in the end state ($\lambda = 0$, i.e., the fully decoupled state of the solute). The reduced energy of the start state ($\lambda = 1$; the fully coupled state of the solute) estimated using the configuration of the end state becomes infinite even though the soft-core modification (Beutler et al. 1994) (Zacharias et al. 1994, p.) is introduced into the LJ interactions.

In our framework, only the reduced energies of the neighboring states, i.e., $u_i(\mathbf{x}_{in})$, $u_{i+1}(\mathbf{x}_{in})$, are used as input data. The free-energy differences between the neighboring states are accumulated to obtain the difference between the start and end states. In this sense, MBAR for “perturbation” is essentially the same framework as the original BAR.

Full potential energy: REST, parameter tuning, and others

The last case, “full potential energy,” needs the whole system’s potential energies for solving the MBAR equations. In this context, state i in the MBAR analysis is specified by the whole systems’ potential energy. In replica exchange with solute tempering (REST) (Liu et al. 2005), REST2 (Wang et al. 2011) (Terakawa et al. 2011), and gREST (Kamiya and Sugita 2018), solute molecules, such as proteins or ligands (in REST/REST2), or part of solute molecules with full or partial potential energies (in gREST) are scaled as a solute region. In REST/REST2/gREST simulations, the solute region in each replica has a different temperature, while the temperatures in the solvent region in all replicas are the same, for instance, room temperature. Although only the scaled terms can be given to the MBAR equations ignoring the other canceling terms, most MD software does not write specific terms of the potential energy by default, but evaluates potential energies from updated parameters of the solute region. Thus, in terms of computational cost, there is no difference in preparing the potential energy for the entire system or the specific scaled terms. Another example is the tuning of force field parameters or model parameters (Messerly et al. 2018)(Shinobu et al. 2019). Since parameters affect the total energy in a complicated way, we need to prepare the potential energies of all the replicas in the whole system. In the case of parameter tuning, parameter sets, which have not yet been sampled, are used as the target state of the MBAR for predicting the behavior of new parameter sets.

Further complicated yet important case is a combination of umbrella sampling (perturbed case) and different potential energy functions (case of full potential energy) in hybrid quantum mechanics/molecular mechanic (QM/MM) calculations (Warshel and Levitt 1976) (Yagi et al. 2021). In the MBAR analysis using the QM/MM data, it is possible to reweight samples obtained by low-level theory (LL, e.g., classical force field or semi-empirical QM) with the energies of high-level theory (HL, e.g., ab initio QM) and obtain more accurate PMF. Because the computational cost of MD simulations with HL is far greater than with that with LL, various methods have been proposed to correct LL data with HL calculations as post-processing (Yagi et al. 2021). LL calculation is often conducted with umbrella sampling along a reaction coordinate $z(\mathbf{x})$. Then, the total potential energy of the LL with umbrella window becomes $U_i(x) = U_{\text{system}}^{\text{LL}}(\mathbf{x}) + U_{\text{restraint}}(z(\mathbf{x}))$. The target potential energy is the potential energy of the HL system without restraint potential, $u_{\text{target}}(\mathbf{x}) = U_{\text{system}}^{\text{HL}}(\mathbf{x})$. The required input data are the potential energies of the LL $U_{\text{system}}^{\text{LL}}(\mathbf{x}_n)$ and HL $U_{\text{system}}^{\text{HL}}(\mathbf{x}_n)$, and the restraint energies $U_{\text{restraint}}^{(i)}(z(\mathbf{x}_n))$. The resulting MBAR equations are:

$$\hat{f}_i = -\ln \frac{\sum_{j=1}^K \sum_{n=1}^{N_j} \exp[-\beta_i U_{\text{system}}^{\text{HL}}(\mathbf{x}_{jn})]}{\sum_{k=1}^K N_k \exp\left[\hat{f}_k - \beta_k \left\{ U_{\text{system}}^{\text{LL}}(\mathbf{x}_{jn}) + U_{\text{restraint}}^{(k)}(\mathbf{x}_{jn}) \right\}\right]} \quad (9)$$

Implementation

Based on the above classifications, we implemented a MBAR code, which we call `mbar_analysis` in the GENESIS software package (Jung et al. 2015) (Kobayashi et al. 2017). The implemented code preprocesses the input data for each one of the above classifications, and the input formats are thus simplified as much as possible. The code calls a common MBAR equation solver to estimate free-energy differences. The solver was implemented by combining a simple self-consistent iteration with the Newton–Raphson method. The calculation of the denominator of the MBAR equation was parallelized with multi-threads, resulting in a faster execution than the reference MBAR implementation, PyMBAR (Shirts and Chodera 2008). After obtaining the estimates, it gives us the free-energy difference, weights, and PMF according to the input parameters. The code was implemented in FORTRAN.

Demonstration of MBAR analysis using `mbar_analysis`

Test simulation systems

We first demonstrate the MBAR analysis in the case of “perturbation” using alanine-tripeptide in vacuum as a

target molecule. Using CHARMM36m force field (Huang et al. 2017), the umbrella sampling simulation was performed with GENESIS atdyn (Jung et al. 2015), (Kobayashi et al. 2017). The ω angle values were increased by 3 degrees between the centers of neighboring windows, resulting in 61 windows to investigate ω angle values from 0 to 180. A spring constant of 200 kcal/mol/rad² was applied. Temperature was controlled at 300 K by the stochastic velocity scaling method (Bussi et al. 2007). Long-range electrostatic interactions were treated without cutoff. Instantaneous ω angles corresponding to $z(\mathbf{x}_{jn})$ in the MBAR equations were extracted from the trajectory data with `trj_convert` tool in GENESIS. Extracted angle values, the spring constants, and the window centers were used as the input of the MBAR. For reference, the same calculation was performed with the WHAM implemented as `wham_analysis` tool in GENESIS.

To demonstrate the MBAR analysis of T-REMD data, we performed a T-REMD simulation of alanine-tripeptide in solution. The initial structures and parameters are the same as those used in GENESIS tutorials (<https://www.rccs.riken.jp/labs/cbrt/tutorials2022/>). We calculated the potential of mean force (PMF) in the ϕ and ψ dihedral angle space. Using CHARMM36 force field (Huang and MacKerell 2013) and TIP3P water molecules (Jorgensen et al. 1983), the simulation was performed with GENESIS spdyn. In MD, electrostatic interactions were treated by the smooth Particle Mesh Ewald (Essmann et al. 1995), and covalent bonds containing hydrogen atoms were constrained by the SHAKE (Ryckaert et al. 1977) or SETTLE (Miyamoto and Kollman 1992) algorithms. The temperature was controlled by the stochastic velocity scaling method (Bussi et al. 2007). These methods were kept in other applications. The system’s potential energy was extracted from the MD log file and sorted from replica-ID data to temperature-ID data by using `remd_convert` tool in GENESIS. Then, they were used as the input together with the temperatures of the replicas for the MBAR. Also, the dihedral angles ϕ and ψ were extracted (with `trj_convert` tool) from the trajectory data and used as the input to calculate PMF at 300 K. We compared PMF using the trajectory only at 300 K with that of all temperature trajectories reweighted using MBAR.

For a demonstration of the “accumulation” case, we calculated a mutation from the amino-acid side-chain analogue of Val to that of Trp in water. Twenty strata were used to divide λ into equal intervals from $\lambda = 0$ to $\lambda = 1$, and FEP was performed. Using CHARMM36 force field and TIP3P water model, the simulation was performed with the FEP implementation (Oshima et al. 2020) in GENESIS spdyn. To remove the instability of MD simulations near end states, the soft-core treatment is introduced to the LJ and electrostatic potentials (Zacharias et al.

a Umbrella sampling

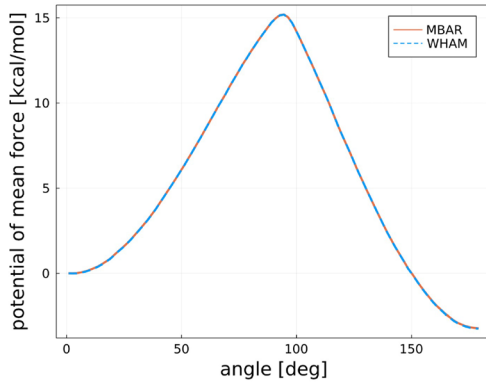
```
# STEP  CV
1      -3.889
2      -3.287
3      -0.976
4      -0.714
5      -1.768
6      -1.648
7      2.706
.....

[INPUT]
cvfile      = run_{}.tor
variable file

[OUTPUT]
pmffile     = pmf.dat

[MBAR]
nreplica    = 61
input_type  = US
temperature = 300.0
target_temperature = 300.0
rest_function1 = 1
grids1 = 0.0 180.0 101

[RESTRAINTS]
constant1   = 1.0 1.0 1.0 ...
reference1  = -180 -175 -160 ...
```



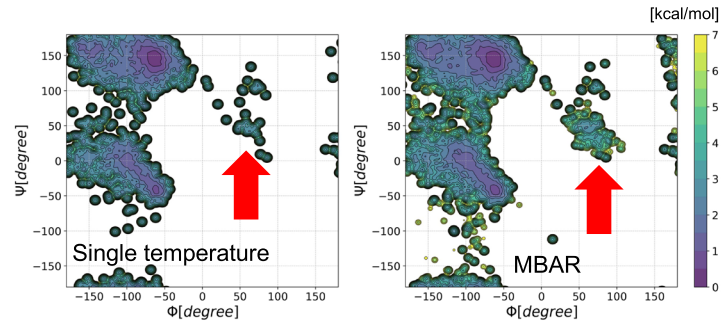
b Temperature REMD

```
# STEP  ene(temp1)
500    -38460.9662
1000   -38598.5814
1500   -38575.7761
2000   -38372.4267
2500   -38217.0582
3000   -38382.6585
.....

[INPUT]
cvfile = ../4_sort_dcd/remd_paramID{ }.pot

[OUTPUT]
fenefile = fene.dat
weightfile = weight{ }.dat

[MBAR]
nreplica    = 20
input_type  = REMD
dimension   = 1
temperature = 300.00 302.53 305.09 307.65
310.24 312.85 315.47 318.12 320.78 323.46
326.16 328.87 331.61 334.37 337.14 339.94
342.75 345.59 348.44 351.26
target_temperature = 300.0
```



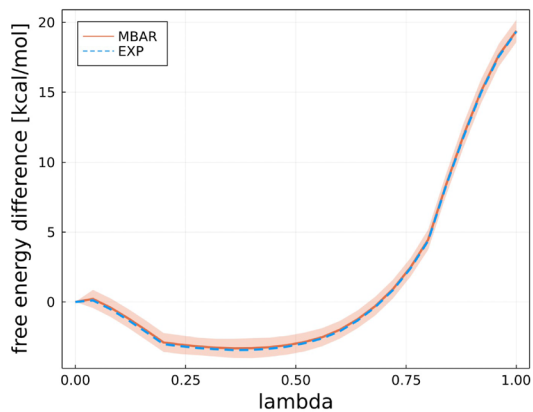
c Free energy perturbation

```
# STEP  Total_E_ref  Delta_E_rev  Delta_E_fwd
50000  -35094.1694   0.4146     0.0052
50050  -35078.3699   1.2473    -1.0525
50100  -35119.4119   1.1648    -0.9913
50150  -35143.7431   1.2738    -1.2400
50200  -35098.9353   0.9983    -0.8737
.....
```

```
[INPUT]
cvfile = ../4_prod/output/run.fepout{ }

[OUTPUT]
fenefile = fene.dat

[MBAR]
nreplica = 26
input_type = FEP
nblocks = 5
temperature = 300.00
```



d Generalized REST

```
# STEP  ene(temp1)  ene(temp2)  ene(temp3)  ene(temp4)
100    -38367.2280  -38367.1429  -38366.9888  -38366.8413
200    -38358.9920  -38358.8884  -38358.7094  -38358.5409
300    -38342.9412  -38342.7462  -38342.4928  -38342.2804
400    -38359.3222  -38358.7614  -38358.1599  -38357.7100
500    -38458.9490  -38459.2113  -38459.3891  -38459.4693
.....
```

```
[INPUT]
cvfile = ../remd_convert/eneout_param{ }

[OUTPUT]
fenefile = fene.dat
weightfile = weight{ }.dat

[MBAR]
nreplica = 4
input_type = REST
temperature = 300.0
target_temperature = 300.0
```

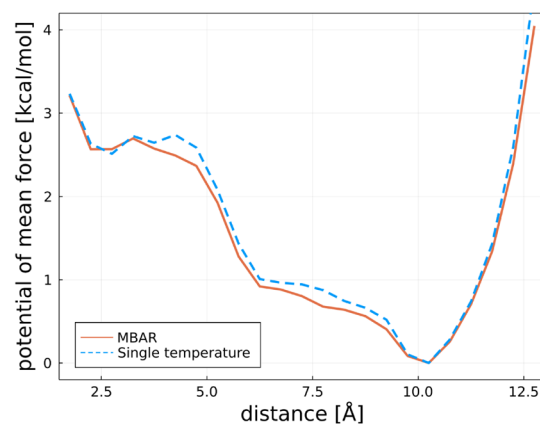


Fig. 2 Results of the MBAR analysis for four cases of our classification. **a** Potential of mean force (PMF) along ω dihedral angle of alanine-tripeptide in vacuum. The results of MBAR and WHAM are indicated by red solid line and blue dashed lines, respectively. **b** PMF in ϕ and ψ dihedral angles of alanine-tripeptide in water. The PMF calculated only from the trajectory of 300 K and all trajectories reweighted by the MBAR are shown. The region of the left-handed helix is indicated by red arrow. **c** Free-energy differences upon the mutation from the amino-acid side-chain analogue of Val to that of Trp as a function of λ , calculated by free-energy perturbation method. The shaded region indicates the uncertainties estimated by block averaging. EXP means the exponential averaging. **d** PMF along the distance between the terminal residues of alanine-tripeptide in water, calculated by gREST simulation. The PMF calculated only from the trajectory of 300 K and all trajectories reweighted by the MBAR are shown

1994) (Steinbrecher et al. 2011). The reduced energies of the neighboring states, i.e., $u_{i-1}(\mathbf{x}_{in})$, $u_i(\mathbf{x}_{in})$, $u_{i+1}(\mathbf{x}_{in})$ were generated with GENESIS spdyn, and they were used as the input for the MBAR analysis.

As a demonstration of a case of “full potential energy,” gREST of alanine-tripeptide in water (taken from GENESIS tutorials) was performed using four replicas of alanine-tripeptide as solute and temperature control by scaling all potential energy terms. Using CHARMM36 force field and TIP3P water model, the simulation was performed with GENESIS spdyn. The potential energies required in MBAR were generated using the optional function of GENESIS spdyn. The potential energy data of replicas were sorted from replica-ID data to temperature-ID data using remd_convert tool of GENESIS. We calculated PMF along the distance between the terminal residues (an oxygen of the ALA1 and a hydrogen of ALA3). We compared PMF using the trajectory at 300 K with that of all replicas’ trajectories reweighted by the MBAR.

Finally, as a demonstration of QM/MM with umbrella sampling MD simulation (QM/MM-US-MD), QM/MM-US-MD of malonaldehyde (MA) and p-Nitrophenyl phosphate (pNPP²⁻) were performed at the third-order extension Self-Consistent-Charge Density Functional Tight-Binding level (Gaus et al. 2011, p. 3). MA and pNPP²⁻ are encapsulated in a TIP3P water sphere with a radius of 20 Å. The reaction coordinate for MA is a linear combination of two distances between atoms involved in the proton transfer reaction. The reaction coordinate for pNPP²⁻ is the P-O distance involved in hydrolysis (Fig. 3a and b).

For MA, MD calculations were performed for 200 ps at 300 K using a spring constant of 40 kcal/mol/rad² in 21 umbrella windows centered along the reaction coordinate increased from -1.0 to 1.0 at 0.1 Å intervals. For pNPP²⁻, MD calculations were performed for 500 ps at 300 K using a spring constant of 300 kcal/mol/rad² in 18 umbrella windows centered along the reaction coordinate increased from -1.4 to 3.1 at 0.1 Å intervals. The 3ob (ophyd) parameter (Gaus

et al. 2013)(Gaus et al. 2014) was used as the Slater-Koster parameter, and CHARMM36m (Huang et al. 2017) and CGenFF (Vanommeslaeghe et al. 2009) (Yu et al. 2012) were used as the classical force fields. For reweighting with MBAR, potential energies were evaluated by single-point energy calculation for 2000 and 5000 samples for MA and pNPP²⁻ at the density functional theory (DFT) level using B3LYP/cc-pvdz (Lee et al. 1988) (Becke 1993) (Grimme et al. 2010) (Dunning 1989), respectively. DFTB and DFT calculations were performed using QSimulate (<https://qsimulate.com>), a fast quantum computation program package, in combination with GENESIS atdyn (Yagi et al. 2021).

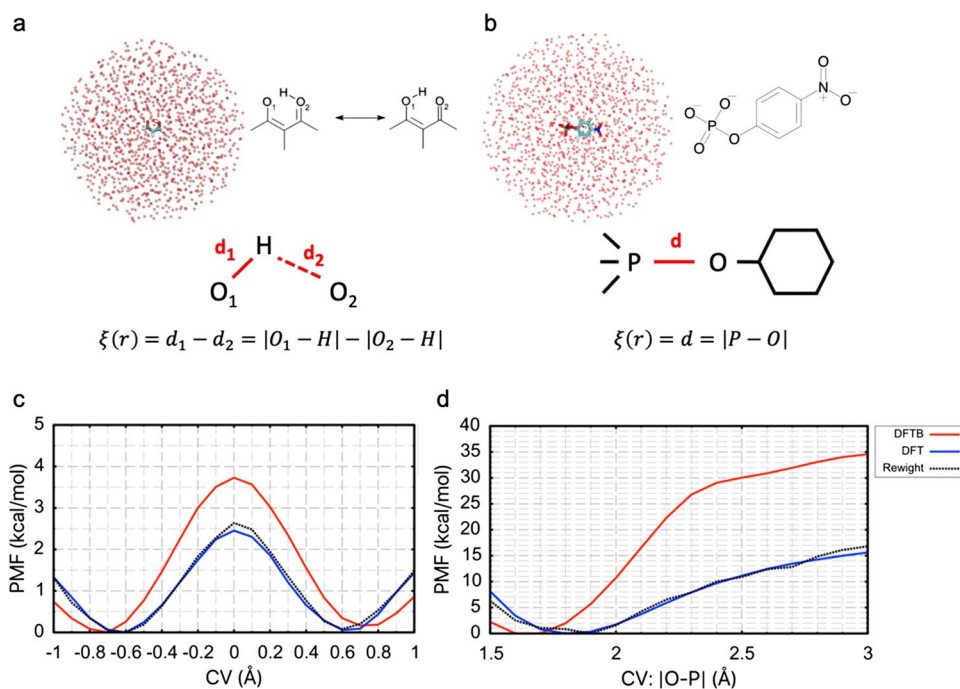
Demonstrations of mbar_analysis

Figure 2A shows the inputs required for the MBAR analysis and the results of the umbrella sampling of alanine-tripeptide in vacuum. As explained above, the only inputs required for umbrella sampling are the collective variable trajectories of umbrellas, $z(\mathbf{x}_{jn})$, the window centers, and the spring constants. Restraint energies, $U_{\text{restraint},i}(\mathbf{x}_{jn})$, required in the MBAR equations, are calculated internally and passed to the solver. In the PMF calculation with WHAM, the density of states is calculated after making the histogram of the trajectories, whereas in the MBAR, the density of states for each z is calculated after the weight of each sample point. Therefore, theoretically, MBAR can prevent any biases due to convolutions.

Figure 2b shows the inputs required for the subsequent MBAR analysis and the results of the T-REMD of alanine-tripeptide in water. As explained above, the only inputs required for T-REMD are the potential energies of sorted replicas and temperatures. The reduced energies $\beta_i U_{\text{system}}(\mathbf{x}_{jn})$, required in the MBAR equations, are calculated internally and passed to the solver. Comparing the PMF obtained from only the lowest temperature 300 K trajectory with that weighted all replica trajectories with MBAR, the left-handed helix state is well captured in MBAR. This is because the left-handed helix sampled at high temperature is not discarded but treated as the weighted samples by the MBAR.

Figure 2c shows the inputs required in the subsequent MBAR analysis and the results of the alchemical FEP from Val to Trp in water. As explained above, the only input required for FEP is the potential energies of the neighboring states. Our MBAR tool integrates the free-energy difference between the neighboring states to obtain the total difference between the start and end states. Inside the program, the MBAR equation for two states is called repeatedly. The free energy difference between Val and Trp shown in Fig. 2c is comparable with the exponential averaging, and matched with the result obtained with NAMD (Liu et al. 2012), indicating that the potential energy difference from

Fig. 3 Simulation system and result of QM/MM with umbrella sampling simulations. **a** The malonaldehyde is located at the center of the water droplet. **b** The pNPP²⁻ is located at the center of the water droplet. The collective variables used in the umbrella sampling are indicated in the figure. **c** The potential of mean force (PMF) of the malonaldehyde obtained by DFTB (red line), DFT (blue line), and re-weighted with the MBAR (black dotted line) of the malonaldehyde. **d** PMF of the pNPP²⁻



the distant state is not contributing. This is also consistent with the result of the exhaustive FEP benchmarks by Paliwal and Shirts (2011), where they showed that MBAR and BAR have comparative accuracies for many FEP data analyses.

Figure 2d shows the results of gREST for alanine-tripeptide in water and the inputs required in the MBAR analysis. The reduced energies required for MBAR were generated with GENESIS spdyn during the simulations. As shown in the figure, the potential energies with all scaling values are required as input for the MBAR analysis. Figure 2d compares the PMF at the lowest temperature (i.e., scaling factor) trajectory with that using MBAR. The figure shows that the PMF using MBAR well captures the stable conformations indicated at the lower PMF values compared to the PMF values using a single trajectory at 300 K.

Figure 3c and d shows the PMF obtained in DFTB, DFT, and reweighted PMF with MBAR. A total of 10 ps (10,000 MD steps) equilibration followed by 20 ps (20,000 MD steps) US-MD calculations were performed in MA and pNPP²⁻ to obtain PMF at the DFT level. In both systems, the PMF obtained from the DFT-level calculation and the PMF obtained from the reweighting are in good agreement. The computational cost for reweighting was 6.7% for MA (16.7% for pNPP²⁻) of the 30,000 MD steps of the brute-force DFT-US-MD calculation (only 2000 samples were reweighted for MA, and 5000 samples were reweighted for pNPP²⁻). Furthermore, as only the potential energy is required for reweighting, the gradient calculation can be omitted, further reducing the actual computational cost.

Discussion

MBAR is a very flexible method, which can handle various states, ranging from different temperature/pressure conditions to different force field parameters. On the other hand, this flexibility sometimes requires complicated inputs, even for straightforward analyses such as umbrella sampling. In this review, we classified typical applications of MBAR and simplifying input data for practical usages. Our classification would help non-expert users to apply the MBAR analysis to obtain unbiased thermodynamic data from biased MD simulation trajectories.

One important future application of MBAR would be feedback to simulation models, as shown by Messerly et al. (2018) and Shinobu et al. (2019). Since MBAR extrapolates and estimates statistics on parameter sets that have not yet been sampled, it helps us to optimize the parameters of the simulation model. When the number of parameters to be optimized is large, grid search becomes difficult even with the help of the MBAR equations. In this case, as recently shown by Wieder et al. (2021), by making the outputs of the MBAR equations differentiable with respect to potential energies and their force field parameters, efficient gradient search can be performed. It is actually similar to the training of neural network model parameters. On the other hand, the limitation of MBAR in terms of feedback to the simulation model is that the MBAR is just an estimator; thus, it can only estimate free-energy differences or the averaged data. Recent deep learning technologies are also flexible, and they can directly use modeling of the conformational density of molecules as well as statistical

averages. For example, Wang et al. recently succeeded in modeling conformational density from various temperature trajectories obtained by T-REMD using a diffusion model in which temperature is incorporated as one of the random variables (Wang et al. 2022). If it becomes possible to model the density and the uncertainty of parameters and conformations with these technologies, more efficient optimization will be possible in combination with MBAR.

Acknowledgements We thank Chigusa Kobayashi for her helpful comments on this work. We also thank the authors of GENESIS tutorials. Some simulation systems in this work were taken from the tutorials.

Author contribution Yasuhiro Matsunaga and Yuji Sugita contributed to the study conception and design. Software was developed by Yasuhiro Matsunaga, Motoshi Kamiya, and Jaewoon Jung. Data collection and analysis were performed by Yasuhiro Matsunaga, Motoshi Kamiya, Hiraku Oshima, and Shingo Ito. The first draft of the manuscript was written by Yasuhiro Matsunaga and Yuji Sugita, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by MEXT as “Program for Promoting Researches on the Supercomputer Fugaku” (biomolecular dynamics in a living cell, grant number: JPMXP1020200101 to Y.S. and Y.M.), JSPS KAKENHI (grant numbers: 20K21380 (to Y.M.), 19H05645, 21H05249 (to Y.S.), 21H05282 (to J.J.)), and the Cooperative Research Program of “Network Joint Research Center for Materials and Devices” (to Y.M.). We used the computational resources provided by the HPCI system research project (Project ID: hp200135, hp210177, and hp220170) and those in RIKEN Hokusai “BigWaterFall.”

Data availability The developed code has been released to the public as one of the analysis tools (called `mbar_analysis`) of MD simulation software GENESIS (Jung et al. 2015) (Kobayashi et al. 2017). This software is freely available from <https://www.r-ccs.riken.jp/labs/cbrt/>. The inputs used in the calculations in this review are available at https://github.com/matsunagalab/paper_mbar, except for the inputs for the QM/MM simulations, which are available from the authors upon reasonable request.

Declarations

Ethics approval This work does not contain any human or animal subjects.

Consent to participate This work does not contain any human or animal subjects.

Consent for publication This work does not contain any individual person’s data in any form.

Competing interests. The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98:5648–5652. <https://doi.org/10.1063/1.464913>
- Bennett CH (1976) Efficient estimation of free energy differences from Monte Carlo data. *J Comput Phys* 22:245–268. [https://doi.org/10.1016/0021-9991\(76\)90078-4](https://doi.org/10.1016/0021-9991(76)90078-4)
- Beutler TC, Mark AE, van Schaik RC et al (1994) Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* 222:529–539. [https://doi.org/10.1016/0009-2614\(94\)00397-1](https://doi.org/10.1016/0009-2614(94)00397-1)
- Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126:014101. <https://doi.org/10.1063/1.2408420>
- Chodera JD, Swope WC, Pitera JW et al (2007) Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J Chem Theory Comput* 3:26–41. <https://doi.org/10.1021/ct0502864>
- Chodera JD, Mobley DL, Shirts MR et al (2011) Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol* 21:150–160. <https://doi.org/10.1016/j.sbi.2011.01.011>
- Ding X, Vilseck JZ, Brooks CL (2019) Fast solver for large scale multi-state Bennett acceptance ratio equations. *J Chem Theory Comput* 15:799–802. <https://doi.org/10.1021/acs.jctc.8b01010>
- Dunning TH (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J Chem Phys* 90:1007–1023. <https://doi.org/10.1063/1.456153>
- Essmann U, Perera L, Berkowitz ML et al (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577–8593. <https://doi.org/10.1063/1.470117>
- Gaus M, Cui Q, Elstner M (2011) DFTB3: extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB). *J Chem Theory Comput* 7:931–948. <https://doi.org/10.1021/ct100684s>
- Gaus M, Goez A, Elstner M (2013) Parametrization and benchmark of DFTB3 for organic molecules. *J Chem Theory Comput* 9:338–354. <https://doi.org/10.1021/ct300849w>
- Gaus M, Lu X, Elstner M, Cui Q (2014) Parameterization of DFTB3/3OB for sulfur and phosphorus for chemical and biological applications. *J Chem Theory Comput* 10:1518–1537. <https://doi.org/10.1021/ct401002w>
- Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J Chem Phys* 132:154104. <https://doi.org/10.1063/1.3382344>
- Huang J, MacKerell AD (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* 34:2135–2145. <https://doi.org/10.1002/jcc.23354>
- Huang J, Rauscher S, Nawrocki G et al (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 14:71–73. <https://doi.org/10.1038/nmeth.4067>
- Jorgensen WL, Chandrasekhar J, Madura JD et al (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935. <https://doi.org/10.1063/1.445869>
- Jung J, Mori T, Kobayashi C et al (2015) GENESIS: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations: GENESIS. *Wiley Interdiscip Rev Comput Mol Sci* 5:310–323. <https://doi.org/10.1002/wcms.1220>

- Kamiya M, Sugita Y (2018) Flexible selection of the solute region in replica exchange with solute tempering: application to protein-folding simulations. *J Chem Phys* 149:072304. <https://doi.org/10.1063/1.5016222>
- Kobayashi C, Jung J, Matsunaga Y et al (2017) GENESIS 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms. *J Comput Chem* 38:2193–2206. <https://doi.org/10.1002/jcc.24874>
- Kumar S, Rosenberg JM, Bouzida D et al (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I the Method *J Comput Chem* 13:1011–1021. <https://doi.org/10.1002/jcc.540130812>
- Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B* 37:785–789. <https://doi.org/10.1103/PhysRevB.37.785>
- Liu P, Kim B, Friesner RA, Berne BJ (2005) Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc Natl Acad Sci* 102:13749–13754. <https://doi.org/10.1073/pnas.0506346102>
- Liu P, Dehez F, Cai W, Chipot C (2012) A toolkit for the analysis of free-energy perturbation calculations. *J Chem Theory Comput* 8:2606–2616. <https://doi.org/10.1021/ct300242f>
- Messlerly RA, Razavi SM, Shirts MR (2018) Configuration-sampling-based surrogate models for rapid parameterization of non-bonded interactions. *J Chem Theory Comput* 14:3144–3162. <https://doi.org/10.1021/acs.jctc.8b00223>
- Mey ASJS, Allen BK, Bruce Macdonald HE, et al (2020) Best practices for alchemical free energy calculations [Article v1.0]. *Living J Comput Mol Sci* 2: <https://doi.org/10.33011/livecoms.2.1.18378>
- Miyamoto S, Kollman PA (1992) Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* 13:952–962. <https://doi.org/10.1002/jcc.540130805>
- Oshima H, Re S, Sugita Y (2020) Prediction of protein–ligand binding pose and affinity using the gREST+FEP method. *J Chem Inf Model* 60:5382–5394. <https://doi.org/10.1021/acs.jcim.0c00338>
- Paliwal H, Shirts MR (2011) A benchmark test set for alchemical free energy transformations and its use to quantify error in common free energy methods. *J Chem Theory Comput* 7:4115–4134. <https://doi.org/10.1021/ct2003995>
- Paschek D, García AE (2004) Reversible temperature and pressure denaturation of a protein fragment: a replica exchange molecular dynamics simulation study. *Phys Rev Lett* 93:238105. <https://doi.org/10.1103/PhysRevLett.93.238105>
- Peter EK, Pivkin IV, Shea J-E (2016) A canonical replica exchange molecular dynamics implementation with normal pressure in each replica. *J Chem Phys* 145:044903. <https://doi.org/10.1063/1.4958325>
- Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5)
- Shinobu A, Kobayashi C, Matsunaga Y, Sugita Y (2019) Building a macro-mixing dual-basin Gō model using the Multistate Bennett Acceptance Ratio. *Biophys Physicobiology* 16:310–321. https://doi.org/10.2142/biophysico.16.0_310
- Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:124105. <https://doi.org/10.1063/1.2978177>
- Souaille M, Roux B (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun* 135:40–57. [https://doi.org/10.1016/S0010-4655\(00\)00215-0](https://doi.org/10.1016/S0010-4655(00)00215-0)
- Steinbrecher T, Joung I, Case DA (2011) Soft-core potentials in thermodynamic integration: comparing one- and two-step transformations. *J Comput Chem* 32:3253–3263. <https://doi.org/10.1002/jcc.21909>
- Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151. [https://doi.org/10.1016/S0009-2614\(99\)01123-9](https://doi.org/10.1016/S0009-2614(99)01123-9)
- Tan Z, Gallicchio E, Lapelosa M, Levy RM (2012) Theory of binless multi-state free energy estimation with applications to protein-ligand binding. *J Chem Phys* 136:144102. <https://doi.org/10.1063/1.3701175>
- Tan Z, Xia J, Zhang BW, Levy RM (2016) Locally weighted histogram analysis and stochastic solution for large-scale multi-state free energy estimation. *J Chem Phys* 144:034107. <https://doi.org/10.1063/1.4939768>
- Tembre BL, Mc Cammon JA (1984) Ligand-receptor interactions. *Comput Chem* 8:281–283. [https://doi.org/10.1016/0097-8485\(84\)85020-2](https://doi.org/10.1016/0097-8485(84)85020-2)
- Terakawa T, Kameda T, Takada S (2011) On easy implementation of a variant of the replica exchange with solute tempering in GROMACS. *J Comput Chem* 32:1228–1234. <https://doi.org/10.1002/jcc.21703>
- Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J Comput Phys* 23:187–199. [https://doi.org/10.1016/0021-9991\(77\)90121-8](https://doi.org/10.1016/0021-9991(77)90121-8)
- Vanommeslaeghe K, Hatcher E, Acharya C et al (2009) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* NA-NA. <https://doi.org/10.1002/jcc.21367>
- Wang L, Friesner RA, Berne BJ (2011) Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *J Phys Chem B* 115:9431–9438. <https://doi.org/10.1021/jp204407d>
- Wang Y, Herron L, Tiwary P (2022) From data to noise to data for mixing physics across temperatures with generative artificial intelligence. *Proc Natl Acad Sci* 119:e2203656119. <https://doi.org/10.1073/pnas.2203656119>
- Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249. [https://doi.org/10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9)
- Wieder M, Fass J, Chodera JD (2021) Teaching free energy calculations to learn from experimental data. *Biophysics*
- Yagi K, Ito S, Sugita Y (2021) Exploring the minimum-energy pathways and free-energy profiles of enzymatic reactions with QM/MM calculations. *J Phys Chem B* 125:4701–4713. <https://doi.org/10.1021/acs.jpcc.1c01862>
- Yu W, He X, Vanommeslaeghe K, MacKerell AD (2012) Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J Comput Chem* 33:2451–2468. <https://doi.org/10.1002/jcc.23067>
- Zacharias M, Straatsma TP, McCammon JA (1994) Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J Chem Phys* 100:9025–9031. <https://doi.org/10.1063/1.466707>
- Zhang BW, Xia J, Tan Z, Levy RM (2015) A stochastic solution to the unbinned WHAM equations. *J Phys Chem Lett* 7
- Zwanzig RW (1954) High-temperature equation of state by a perturbation method. I Nonpolar Gases *J Chem Phys* 22:1420–1426. <https://doi.org/10.1063/1.1740409>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.