Optimization of replica exchange temperature ladder under the well-tempered ensemble

Liu, Yang; Li, Weifeng; Mu, Yuguang

2018

Liu, Y., Li, W., & Mu, Y. (2018). Optimization of replica exchange temperature ladder under the well-tempered ensemble. Chemical Physics Letters, 711, 66-72. doi:10.1016/j.cplett.2018.09.036

https://hdl.handle.net/10356/136890

https://doi.org/10.1016/j.cplett.2018.09.036

Optimization of replica exchange temperature ladder under the well-tempered ensemble

Yang Liu [a,b], Weifeng Li [a], and Yuguang Mu [b,*]

[a] School of Physics, Shandong University, 27 Shandanan Road, Jinan City, Shandong Province, China 250100

[b] School of Biological Sciences, Nanyang Technological University (NTU), 60 Nanyang Drive, Singapore 637551

[*]Corresponding author. E-mail address: ygmu@ntu.edu.sg

ABSTRACT

Selection of temperature sequence and bias factor is a critical step for the preparation of Parallel Tempering simulations in a Well-Tempered Ensemble (PT-WTE). We provide a scheme in this study to generate the temperature sequence and the corresponding bias factor optimally. The number of replicas for a complete coverage of a specific temperature range is adjustable in this scheme, while keeping the average acceptance probability (AAP) between neighboring replica-pairs unchanged. Two series of PT-WTE simulations were tested with the number of replicas as 16 and 8 respectively. This work leads to a better understanding and application of PT-WTE simulations.

Keywords: PT-WTE; temperature sequence optimization

1. Introduction

Molecular dynamics simulations at room-temperature are usually of much interest but easily trapped in the metastable states. Simulations at higher-temperatures with correspondingly higher kinetic energies are able to overcome the energy barrier between different metastable states, however, the free energy surfaces (FES) obtained are different from the one at room temperature. Parallel tempering (PT) has now become a commonly used advanced sampling simulation technique in constructing FES when the simulated systems have rugged energy landscapes [1]. In PT, a series of replicas, whose temperatures are arranged in an incremental ladder, are simulating simultaneously. The enhanced sampling is realized by allowing the exchange of configurations of lower- and higher-temperature replicas. In practice, the exchange is attempted every specified number of simulation time-steps, with an acceptance probability described by [2]:

$$P(T_1 \leftrightarrow T_2) = \min\left(1, e^{(\beta_1-\beta_2)\cdot(U_1-U_2)}\right) \qquad (1)$$

where $\beta_1 = 1/(k_B T_1)$, $\beta_2 = 1/(k_B T_2)$, k_B is the Boltzmann constant, and U_1, U_2 are the instantaneous potential energies of the two exchanging configurations.

PT achieves the highest sampling efficiency when the average acceptance probability (AAP) between neighboring replica pairs is uniform [3]. In such a case, a well-designed temperature sequence would be required for the optimized performance of PT. Researchers have developed several methods to generate the temperature sequence. For instance, Kofke showed that for systems with a constant heat capacity, a geometric progression of temperature sequence assured the same AAP among all the neighboring replica pairs [4-6]. Rathore et al. found that AAP was relevant to the value of $\frac{U(T_i)-U(T_{i-1})}{\frac{1}{2}(\sigma(T_i)+\sigma(T_{i-1}))}$, where $U(T_i)$ is the average potential of replica at temperature T_i, and $\sigma(T_i)$ is the distribution width of the Gaussian-like potential at that temperature [3]. Before using this algorithm, short simulations for the system at several different temperatures are necessary to estimate the average potential and the potential distribution width. Then the temperatures for PT could be obtained by iteratively solving the equation $\frac{U(T_i)-U(T_{i-1})}{\frac{1}{2}(\sigma(T_i)+\sigma(T_{i-1}))} = \frac{U(T_{i+1})-U(T_i)}{\frac{1}{2}(\sigma(T_{i+1})+\sigma(T_i))}$. Patriksson et al. came up with a much friendlier using method [7], on the basis of the assumption of constant heat capacity: description of $U(T_i)$ and $\sigma(T_i)$ in the iterative equation were detailed down to the level of degree of freedom, thus users only needed to provide the information of system size and number of constraints to get the optimal allocation of temperatures with a specified AAP. In another case, Prakash et al. discussed the generation of such a temperature sequence for the system with a variable heat capacity [8].

However, even with the maximum efficiency of PT, there is still a big obstacle that limiting its wider applications. The requirements that the highest temperature must be high enough to keep all the replicas diffusive in phase space and the temperature gap between neighboring replicas must be small enough to keep sufficient exchange, determine the lower bound of the number of replicas, which is usually from dozens to more than one hundred. That is a disaster for researchers with limited computational resources, especially when simulating systems being large. The Well-Tempered Ensemble (WTE) framework can alleviate this issue [9]. Intrinsically, WTE is well-tempered meta-dynamics [10], but using the system potential as a collective variable (CV). When PT is combined with WTE (PT-WTE), the added biases enhance the fluctuations of potential energies of each replica, making it possible to have sufficient exchanges between replicas with larger temperature gaps. This approach has been proved its efficiency in the free energy surface study of peptide-surface interaction [11], peptide oligomerization [12], peptide conformational transition [13], flexible molecules binding [14], and so on. Deighan et al. reported that with a comparable sampling efficiency, the number of replicas could dramatically decrease from 100 in a standard PT to 10 in PT-WTE, and the efficiency in PT-WTE could be further enhanced by adopting a reasonably larger bias factor [15]. However, the criterion of choosing a proper bias factor for WTE has not been established so far.

Inspired by the works of Rathore et al. [3], Patriksson et al. [7] and Prakash et al. [8], we provided an operable scheme to generate reasonable temperature intermediates and a corresponding bias factor for PT-WTE within a given temperature range. By taking different bias factors, the number of replicas could be adjusted in this scheme to keep the desired AAP unchanged. Our method was demonstrated on an explicit solvated mini-protein system. Two series of PT-WTE simulations with different bias factors and number of replicas were carried out. Discussions on the choice of proper bias factors were given.

2. Computational details

2.1 Potential energy probability distribution function in WTE

In WTE, potential energy $U = U(R)$ is used as CV, where R is the full set of atomic coordinates. The bias potential $V(U,t)$ converges to [9,10]:

$$V(U, t \to \infty) = -(1-\gamma^{-1})F(U) \qquad (2)$$

where $F(U)$ is the underlying free energy of the system and γ is the bias factor which, by definition, could be expressed as $\gamma = (T+\Delta T)/T$, with T is the temperature of the system and ΔT has the dimension of a temperature. Within an irrelevant constant, $F(U)$ is defined as

$$F(U) = -\beta^{-1}\ln P(U) = -\beta^{-1}\ln\int dR\delta\big(U-U(R)\big)e^{-\beta U(R)} = U - \beta^{-1}lnN(U) \qquad (3)$$

where $P(U) = \int dR\delta\big(U-U(R)\big)e^{-\beta U(R)} = lnN(U)e^{-\beta U}$ is the energy probability distribution function in the canonical ensemble, and $N(U) = \int dR\delta(U-U(R))$ is the number of states of energy U. By combining Eq. (2) and Eq. (3), the biased potential of the system $U_\gamma(R)$ could be written as,

$$U_\gamma(R) = U(R) - (1-\gamma^{-1})\Big(U(R) - \beta^{-1}lnN\big(U(R)\big)\Big) \qquad (4)$$

thus the partition function of the WTE Z_γ is

$$Z_\gamma = \int dRe^{-\beta U_\gamma(R)} = \iint dRdU\delta\big(U-U(R)\big)e^{-\beta\left[U(R)-\left(1-\gamma^{-1}\right)(U(R)-\beta^{-1}lnN(U(R)))\right]} \qquad (5)$$

and could be easily simplified as:

$$Z_\gamma = \int dUP(U)^{\gamma^{-1}} \qquad (6)$$

If the energy probability distribution function in the canonical ensemble is considered as a Gaussian distribution, which is usually true for the water-dominated systems [16], we have $P(U) \propto e^{-\frac{(U-\mu)^2}{2\sigma^2}}$, where μ is the average potential and σ is the standard deviation of the distribution. In WTE,

energy probability distribution function $P_\gamma(U)$ is tuned by the bias factor but still have a Gaussian shape: $P_\gamma(U) = P(U)^{\gamma^{-1}} \propto e^{-\frac{(U-\mu)^2}{2\gamma\sigma^2}}$. This change leads to the same average potential, μ, but $\sqrt{\gamma}$ times larger distributional width in the WTE when compared with the canonical ensemble. These larger potential energy fluctuations guarantee the intact acceptance probability of configurational exchange between replicas with larger temperature gaps in PT-WTE. After normalization, $P_\gamma(U)$ could be written as:

$$P_\gamma(U) = \frac{1}{\sigma\sqrt{2\pi\gamma}} e^{-\frac{(U-\mu)^2}{2\gamma\sigma^2}} \qquad (7)$$

2.2 Prediction of temperature sequences and bias factors for PT-WTE

In PT, the AAP between two neighboring replicas $\langle P_{acc}(T_1 \leftrightarrow T_2)\rangle$ could be estimated as:

$$\langle P_{acc}(T_1 \leftrightarrow T_2)\rangle = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{U_1} P_{(2)}(U_2) \cdot P_{(1)}(U_1) dU_2 + \int_{U_1}^{+\infty} e^{(\beta_1-\beta_2)\cdot(U_1-U_2)} \cdot P_{(2)}(U_2) \cdot P_{(1)}(U_1) dU_2 \right) dU_1 \qquad (8)$$

where $P_{(1)}(U_1)$ and $P_{(2)}(U_2)$ are the probability of having instantaneous potential energy, U_1 and U_2, for replicas at temperature T_1 and T_2, respectively. $P_{(1)}(U_1)$ and $P_{(2)}(U_2)$ should have Gaussian shapes with their normalized expression as: $P_{(1)}(U_1) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(U_1-\mu_1)^2}{2{\sigma_1}^2}}$ and $P_{(2)}(U_2) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(U_2-\mu_2)^2}{2{\sigma_2}^2}}$ respectively, where μ_1, σ_1, μ_2 and σ_2 are the average value and standard deviation of each individual replica. Equivalently, Eq. (8) could be described as [7]:

$$\langle P_{acc}(T_1 \leftrightarrow T_2)\rangle = \int_{-\infty}^{0} P_{(1,2)}(U_{1,2}) dU_{1,2} + \int_{0}^{+\infty} e^{(\beta_1-\beta_2)\cdot U_{1,2}} \cdot P_{(1,2)}(U_{1,2}) dU_{1,2} \qquad (9)$$

where $P_{(1,2)}(U_{1,2})$ is the probability of having energy different $U_{1,2} = U_1 - U_2$, with its normalized form: $P_{(1,2)}(U_{1,2}) = \frac{1}{\sigma_{1,2}\sqrt{2\pi}} e^{-\frac{(U_{1,2}-\mu_{1,2})^2}{2{\sigma_{1,2}}^2}}$,where $\mu_{1,2} = \mu_1 - \mu_2$ and $\sigma_{1,2} = \sqrt{{\sigma_1}^2 + {\sigma_2}^2}$ are the new average and standard deviation respectively.

Based on the same principle, AAP between replica pairs in PT-WTE is also predictable:

$$\langle P_{\gamma,acc}(T_1 \leftrightarrow T_2)\rangle = \int_{-\infty}^{0} P_{\gamma,(1,2)}(U_{1,2}) dU_{1,2} + \int_{0}^{+\infty} e^{\frac{1}{\gamma}(\beta_1-\beta_2)\cdot U_{1,2}} \cdot P_{\gamma,(1,2)}(U_{1,2}) dU_{1,2} \qquad (10)$$

where

$$P_{\gamma,(1,2)}(U_{1,2}) = \frac{1}{\sigma_{1,2}\sqrt{2\pi\gamma}} e^{-\frac{(U_{1,2}-\mu_{1,2})^2}{2\gamma\sigma_{1,2}{}^2}} \qquad (11)$$

It should be noticed that the instantaneous acceptance probability is tuned by γ when compared to the canonical ensemble, since the Boltzmann factor is tuned by γ, based on Eq. (6). In Eq. (10), the AAP is a function of the potential average and potential distributional width of the two replicas, as well as the bias factor. This relationship can be further simplified by replacing the potential average and distributional width with the expression of their individual temperatures, such as the linear relationships $\mu(T) = A_1 + B_1 T$ and $\sigma(T) = A_2 + B_2 T$ (A_1, B_1, A_2, and B_2 are four constants) for systems with constant heat capacity [7]. Consequently, the AAP is only related to the bias factor and the two temperatures of the involved replicas. To cover a temperature gap from T_1 to T_n with n replicas in PT-WTE, and have a desired AAP $\langle P_{\gamma,acc} \rangle$, we will obtain a solvable equation set composed of $n-1$Eq. (10) which correspondingly contains $n-1$ unknown variables including $n-2$ temperatures from T_2 to T_{n-1} and the bias factor γ. The temperature ladder and the bias factor are generated by solving the equation set. WTE framework demonstrates its advantages when combined with PT: the number of replicas is adjustable by changing the value of bias factor, in which way considerable amounts of computational resources are saved.

2.3 Simulation parameters

All simulations were performed with the help of Gromacs package [17] (version 4.6.7) and plumed plug-in [18] (version 2.1.2). An explicitly solvated tryptophan-cage protein model (PDB code 1L2Y [19]) was simulated to compare the performance of PT-WTE under different bias factors and number of replicas. All interactions were described by the AMBER99SB force field [20]. Two series of PT-WTE simulations, including S_{16}, a 16-replica PT-WTE simulation with bias factor $\gamma = 37.9$, and S_8, an 8-replica PT-WTE simulation with $\gamma = 174.2$ were simulated for 200 ns per replica. To further confirm the improved sampling efficiency by the optimized temperature sequence as compared with the geometric one, another 16-replica PT-WTE simulation ($S_{16\text{-geo}}$) with bias factor $\gamma = 37.9$ but a geometric ladder was performed for 200 ns per replica as well. Totally we have 8 μs length of trajectories in this study.

The desired average acceptance probabilities and the temperature boundary for S_{16}, $S_{16\text{-geo}}$, and S_8 were set to 0.5 and 300K ~ 500K, respectively. Their detailed temperature sequences are listed in the Results and Discussion section. The simulations were prepared as following before the production sampling: different unfolded structures were generated from an unbiased MD simulation at 500K and were used as the initial configurations for different replicas. Each structure was then solved in a 6*6*6 nm^3 water box which contained 6769 TIP3P water molecules [21] and one chlorine ion as the counter ion.

After energy minimization of the simulation box, 500 ps of equilibrium simulation were carried out in NVT ensemble with position restraints of the protein. We then removed the position restraints and constructed the WTE framework by a 30 ns simulation. During this process, the potential energy was used as CV with biasing potentials of Gaussian width of 500 kJ/mol and Gaussian height of 2 kJ/mol, and the Gaussian potential was deposited every 500 calculation steps. It was then followed by a 200 ns production run keeping the added biasing potential with no more new biases added to the system. The NVT ensemble was adopted during the whole simulation with the V-rescale temperature coupling method to maintain the desired temperature [22]. A leap-frog algorithm was applied for the integration of Newton equation of motion, and the LINCS algorithm [23] and the SETTLE algorithm [24] were used to constrain all the covalent bonds in protein and water with a time step of 2 fs. The cutoff for both van der Walls (VDW) interactions and short-range electrostatic interactions were set to 1.0 nm. The long-range electrostatic interactions were handled by the Particle mesh Ewald (PME) [25] algorithm. An exchange was attempted between neighboring replicas every 500 calculation steps.

3. Results and discussions

3.1 Generation of temperature sequences and bias factors

Short unbiased MD simulations (10 ns each) were performed at different temperatures for the fitting of the potential energy average $\mu(T)$ and the potential energy distributional width $\sigma(T)$. Potential energies at different temperatures follow Gaussian distributions (Fig. 1a). Raw data of the potential averages are shown in Fig. 1b as black dots. A linear fitting (Fig. 1b blue line) was first performed as suggested by Patriksson et al. [7], however, instead of a proportional increase of the potential average to the temperature, we did observe a decay of the slope. This slope has a dimension of heat capacity, indeed it can be considered, by definition, as a reduced heat capacity of the system with the exclusion of the kinetic energy part. In our simulation, this reduced heat capacity decreased by $135.2\ kJ \cdot mol^{-1} \cdot K^{-1}$ as the temperature increases from 300K to 500K (Fig. 1c).

It is well known that the heat capacity for proteins is usually not a unique value at different temperatures [26], especially at the phase transition temperature, as shown by Rathore et al. [3]. In our simulation systems, we have almost 7000 water molecules. As a result, the water molecules contribute much more to the heat capacity than the protein. Experiments proved the decrease of the heat capacity of water at a constant volume at high temperature [27,28]. According to the calculation of the online water heat capacity calculator [29], the heat capacity decrease of a water system with comparable size at the same temperature range was equal to $109\ kJ \cdot mol^{-1} \cdot K^{-1}$. This decrease, which should come from the potential energy part since the kinetic energy part is supposed to be a constant, was in the same order of magnitude as the decrease of our system, indicating that the major change of heat capacity comes from water molecules.

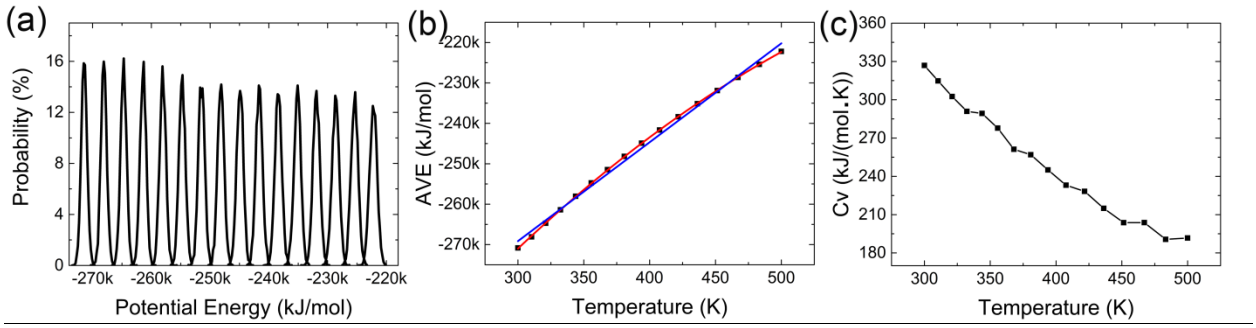

Fig. 1. Thermodynamic properties of the system. Potential energy distributions in canonical ensemble are Gaussian-shaped (figure a). Quadratic fitting (figure b, red curve) is more precise than linear fitting (figure b, blue line) for the potential average (AVE) and temperature relationship. The heat capacity decays almost linearly with the temperature (figure c). The heat capacity is calculated according to $C_V(T) = \sigma(T)^2/(k_B T^2)$.

The expectation of this study is to provide the PT-WTE in a specific temperature range with an optimal temperature ladder, which is supposed to be sparser than that in normal PT. However, along with the heat capacity decay, a linear fitting of the potential average would bring errors into the temperature ladder, and unfortunately, the errors eventually overestimate the AAPs in higher temperature region and underestimate the AAPs in lower temperature. Thus the linear fitting doesn't apply to our case and a more precise fitting method is required. Since the isochoric heat capacity of water decreases with temperature in a linear manner [30], so does the heat capacity in our system (Fig. 1c). This means the potential average is quadratically related to the temperature. We assume the expression of the potential average as $\mu(T) = A_1 + B_1 T + C_1 T^2$ and the heat capacity as $C_V(T) = B_1 + 2C_1 T$, correspondingly, the potential distributional width is expressed as $\sigma(T) = \sqrt{k_B B_1 T^2 + 2k_B C_1 T^3}$, based on $C_V(T) = \sigma(T)^2/(k_B T^2)$. The $\mu_{1,2}$ and $\sigma_{1,2}$ in Eq. (10) are then written as:

$$\mu_{12} = B_1(T_1 - T_2) + C_1\left({T_1}^2 - {T_2}^2\right) \qquad (12)$$

Table 1. Calculated bias factors and temperature sequences for simulation series S_{16} and S_8. A geometric sequence is listed as a reference.

Series	Bias factors	Temperature sequences (K)							
$S_{16\text{-geo}}$	37.9	300.00	310.39	321.14	332.27	343.78	355.69	368.01	380.76
		393.95	407.60	421.72	436.33	451.44	467.08	483.26	500.00
S_{16}	37.9	300.00	309.35	319.10	329.26	339.87	350.96	362.57	374.73
		387.50	400.93	415.09	430.05	445.90	462.74	480.72	500.00
S_8	174.2	300.00	320.52	342.99	367.71	395.09	425.69	460.27	500.00

and

$$\sigma_{12} = \sqrt{k_B B_1({T_1}^2 + {T_2}^2) + 2k_B C_1({T_1}^3 + {T_2}^3)} \qquad (13)$$

respectively. The quadratic function provided a good prediction of potential average (Fig. 1b, red curve) with an R-squared value of 0.99994, and the three constants A_1, B_1, and C_1, were fitted with values of $-391990.46\ kJ/mol$, $499.38\ kJ/(mol \cdot K)$, and $-0.32\ kJ/(mol \cdot K^2)$ respectively. The resultant temperature sequences (S_{16} and S_8) and the corresponding bias factors (Table 1) could be obtained by solving the equation set composed of Eq. (10), with the substitution of Eq. (12) and Eq. (13), as well as the specified temperature range (300 K ~ 500 K), the number of replicas (16 or 8) and the desired AAPs (0.5). Theoretically, this algorithm is capable of obtaining temperature intermediates and bias factors for simulations with any number of replicas. In this work, we only use a 16-replica PT-WTE and an 8-replica PT-WTE as demonstrations and compare their performances. A large amount of computational resources is saved as compared to the PT simulation in canonical ensemble, where more than 120 replicas are required for the same temperature range and the same AAP, based on the Temperature Generator [7].

3.2 Sampling efficiency of PT-WTE simulations with different bias factors

One of the commonly used indicators to evaluate the efficiency of PT is the round-trip time (RTT), which refers to the time that a replica spends to traverse all the temperatures and return to the start. RTTs in this study are shown in Table 2 and their average values are 18.7 ± 21.2 ns and 5.1 ± 1.3 ns for S_{16} and S_8 respectively. Interestingly, the RTT values directly correlate with the number of replicas and decline with the increase of bias factors, indicating a higher PT efficiency happens with a larger bias factor (or fewer replicas). In a recent study, the RTT was found to be around 5 ns in a 10-replica simulation [15]. Our simulation is comparable to their work at the account of our frequency of exchange attempt (1 ps^{-1}) is lower than theirs (2.5 ps^{-1}).

The RMSD of C_α atoms with respect to those of the NMR structure is calculated, and one conformation is considered fully folded when its RMSD value is smaller than 0.2 nm [15]. The biases were removed based on the work of Tiwary et al. [31] during data processing. FES of the 300K ensemble in S_{16} (Fig. 2a) is able to sample the folded and unfolded states of trp-cage with a free

Table 2. The RTTs for replicas in different simulation series (continuous trajectories).

Series	RTTs (ns)															
$S_{16\text{-}geo}$	15.4	18.2	22.2	40	8	11.1	28.6	12.5	9.1	10	20	9.1	15.4	25	33.3	14.3
S_{16}	11.8	15.4	12.5	11.1	10.5	8.7	15.4	16.7	20	15.4	9.5	15.4	11.8	13.3	100	11.8
S_8	7.7	4.4	5.7	4.8	3.2	4.2	5.3	5.9								

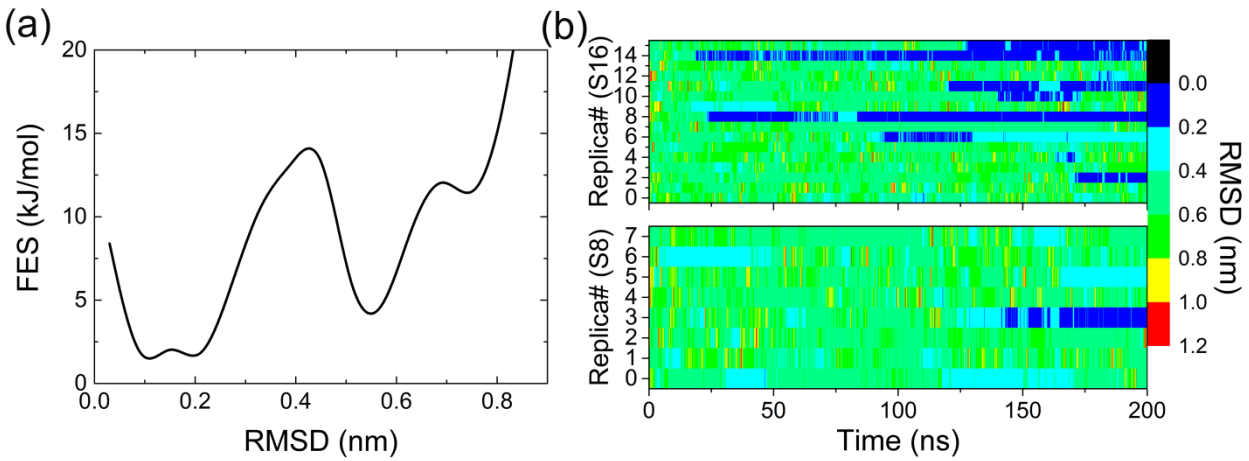

Fig. 2. Conformational sampling of the two simulation series. FESs of 300K ensemble in S_{16} (a) is able to reproduce the folded local minimum precisely. The time evolution of the C_α RMSD of each continuous trajectory is shown in figure (b), and the RMSD values are averaged over 100 ps time block with reweighting.

energy difference of -1.4 kJ/mol, which is comparable to the experimental value of -2.1 kJ/mol [32]. However, FES in S_8 (Fig. S1 in supplementary materials, with a folded-unfolded free energy difference equal to 1.3 kJ/mol) shows quite poor sampling of the folded states. The reason is shown by the RMSD change of each continuous trajectory in Fig. 2b, where we only monitored one of the replicas spent considerable time in the folded state in the S_8 simulation while there were more in the S_{16} simulation. Sampling of the $S_{16\text{-geo}}$ is also worse than that of the S_{16} (Fig. S1, with a folded-unfolded free energy difference of 0.5 kJ/mol), although the number of folded replicas is very close in the S_{16} and $S_{16\text{-geo}}$ simulations (Fig. S2). Compared to the S_{16} simulation, the $S_{16\text{-geo}}$ and S_8 simulation require much longer simulation time to converge.

If a folding event is defined by a transition of C_α RMSD from a value larger than 0.8 nm to that smaller than 0.2 nm, we only confirmed 0.94 folding events per replica in the S_{16} simulation, and even less (0.88 per replica) in the S_8 simulation. To enhance the sampling from this perspective, a recommended solution is to apply bias based on other structure-related CVs, as proved by the above mentioned work [15], where the backbone hydrogen bonds and hydrophobic core contacts were used as CVs in a PT-WTE simulation of the trp-cage system with 10 replicas, the folding events per replica increased to almost 30 times in 250 ns simulation. We also performed short (50 ns) simulation of S_{16} with these two CVs (more details are available in the supplementary material). The RMSD change of each continuous trajectory is shown in Fig. S3, and the folding events is already increased to 1.2 times per replica in this 50 ns simulation.

3.3 A smaller bias factor works better with our algorithm

The AAPs in the S_{16} and S_8 simulation series are nearly uniform (Table 3), and their averages are 0.439±0.031 and 0.424±0.072 respectively, which are both close to the predicted value of 0.5. These results indicate that our algorithm generates satisfactory temperature ladders for the two cases. By

contrast, the AAPs in the $S_{16\text{-geo}}$ simulation shift to a higher value in the higher temperature region, with an average of 0.442±0.051. From the temperature sequences in Table 1, our algorithm gets rid of the AAP shift by narrowing the temperature gaps in the lower temperature region while expanding them in the higher temperature region. From Table 3, we speculate that the poorer conformational sampling in the $S_{16\text{-geo}}$ than in the S_{16} is relevant to the uneven distribution of AAPs of the former, since some relatively lower AAPs in the lower temperature region would slow the construction of the lower temperature ensemble.

It is still notable that the AAP sequence in the S_8 simulation (bias factor = 174.2) is more dispersive than that in the S_{16} simulation (bias factor = 37.9). Such a behavior for the series under a larger bias factor is a consequence of the abnormal potential distributions. The biased potential distributions of the two simulation series (Fig. 3a and 3b) exhibit the Gaussian-shaped curves when using the bias factor of 37.9 and the misshaped distribution when the bias factor equals to 174.2.

In WTE, a history dependent bias [10]

$$V(U,t) = k_B \Delta T ln\left(1 + \frac{\omega N(U,t)}{k_B \Delta T}\right) \qquad (14)$$

is added to the Hamiltonian with a rate of

$$\dot{V}(U,t) = \omega e^{-\left(\frac{V(U,t)}{k_B \Delta T}\right)} \delta_{U,U(t)} \qquad (15)$$

where ω is the initial bias deposition rate, $N(U,t) = \int_0^t \delta_{U,U(t')}\, dt'$ is the histogram of potential U in the biased simulation, which rises linearly with simulation time. The Gaussian deposition rate at a given point U, $\omega e^{-\left(\frac{V(U,t)}{k_B \Delta T}\right)} \delta_{U,U(t)}$ decreases to zero if the time spent there is long enough, and that's the prerequisite to hold up Eq. (2) and the sign to reach a converged WTE. This is the foundation of the whole WTE framework and ensures the effectiveness of Eq. (6) and Eq. (10). To verify if the WTE is well established, we checked the time evolution of Gaussian height. Practically we used a finite width Gaussian to replace the $\delta_{U,U(t)}$. The Gaussian height decay in a CV range ($\mu - \sigma_G \sim \mu + \sigma_G$, where μ is the potential average and σ_G is the Gaussian width in the 300 K ensemble) is plotted as a function of visiting times (Fig. 3c). In 30 ns WTE-establishing time, Gaussian height

Table 3. The AAPs for neighboring replica-pairs. Lower temperature replica-pairs are on the left.

Series	AAPs of neighboring replica-pairs														
$S_{16\text{-geo}}$	0.39	0.42	0.36	0.46	0.42	0.36	0.46	0.42	0.48	0.52	0.52	0.44	0.47	0.52	0.43
S_{16}	0.43	0.42	0.48	0.40	0.48	0.42	0.46	0.36	0.46	0.45	0.44	0.43	0.44	0.44	0.48

S_8	0.48	0.36	0.56	0.39	0.41	0.33	0.44								

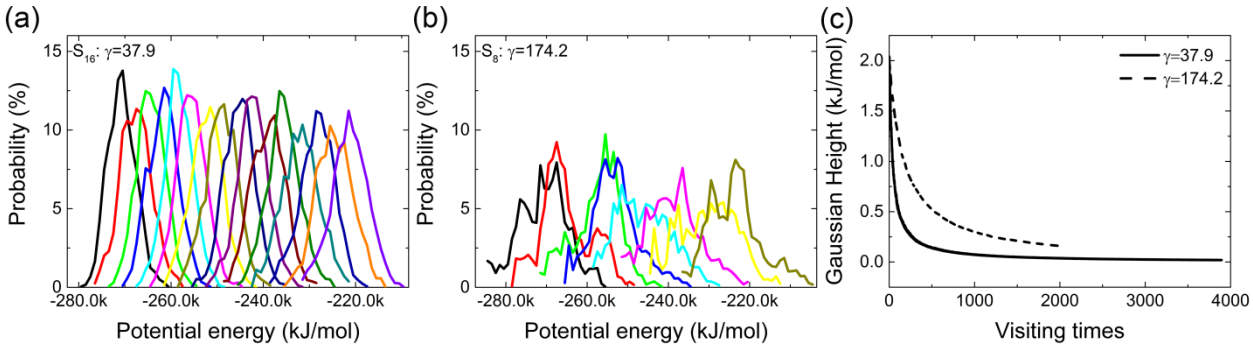

Fig. 3. Distribution of the biased potential energy in S_{16} (a) and S_8 (b). The Gaussian shapes become poorer in S_8. The Gaussian height decay of 300 K ensembles during the 30 ns WTE-establishing period (c). Only the regions around the potential average are counted. The x-axis is the times that the samples visit the considered region.

with bias factor of 37.9 decreased to a negligible value (0.019 kJ/mol), while that with bias factor of 174.2 is still sizeable (0.159 kJ/mol). As a result, the distributions deviate from the ideal Gaussian shapes in S_8, thus misleading the solution of Eq. (10) and leading to a more disperse AAP sequence. Based on these results, we speculate that the larger bias factor retards the convergence and offers a worse performance of PT-WTE.

3.4 Improvement of the performance of large bias factor

The performance of a large bias factor is improvable by extending the WTE-establishing time to achieve the convergence. To obtain a small value of $\omega e^{-(\frac{V(U,t)}{k_B\Delta T})}\delta_{U,U(t)}$, the $\frac{V(U,t)}{k_B\Delta T}$ should be large enough, which, based on Eq. (14), means the value of $\frac{\omega N(U,t)}{k_B\Delta T}$ needs to be large enough. Noting that $\Delta T = T(\gamma - 1)$ is increasing with bias factor, however, $N(U,t)$ grows linearly with the simulation time, indicating that for a specific bias factor, the convergence is always achievable by elongating the sampling time. For example, we extended the WTE-establishing time of S_8 simulation from 30 ns to 100 ns and obtained a further reduced Gaussian height (Fig. 4a, 0.054 kJ/mol), and correspondingly, the biased potential distribution in the following 20 ns PT-WTE simulation is more Gaussian-shaped (Fig. 4b). The AAPs were also improved to a value of 0.454±0.058 with a smaller standard deviation.

The performance of the improved S_8 simulation was still not as good as S_{16}. Here we make a rough evaluation based on the $\frac{\omega N(U,t)}{k_B\Delta T}$ values of two simulation cases. For a fully converged simulation, the biased potential energy is supposed to distribute in a Gaussian shape. We represent the

histogram of $N(U,t)$ from Eq. (14) in the region of $\mu-\sigma_G{\sim}\mu+\sigma_G$ in a period of time as $N(\mu-\sigma_G{\sim}\mu+\sigma_G)$, then $N(\mu-\sigma_G{\sim}\mu+\sigma_G)$ is countable based on:

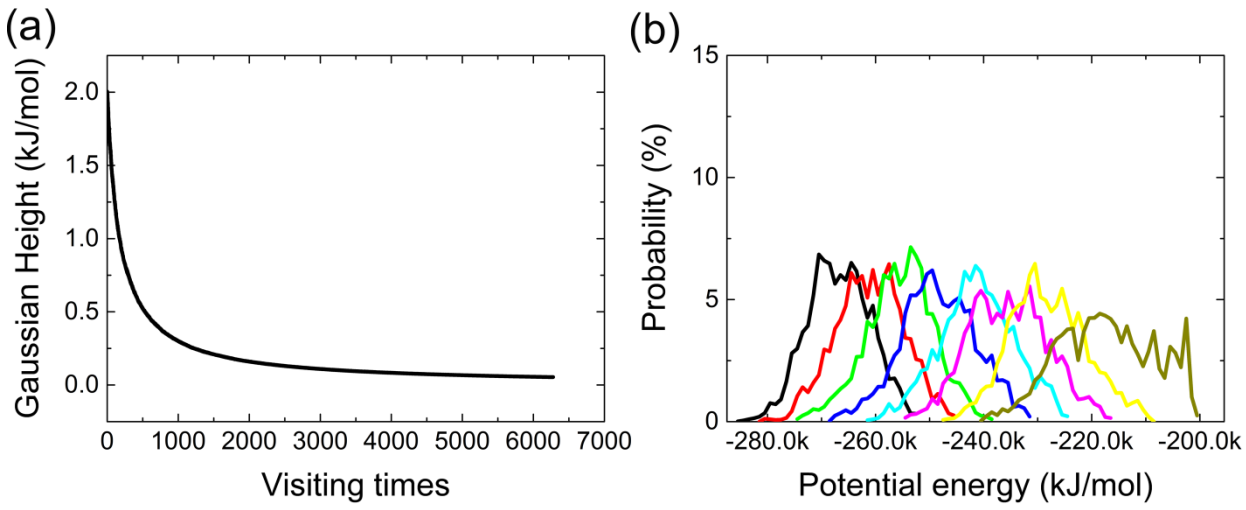

Fig. 4. Improvement of S_8 simulation by extending the WTE-establishing time to 100 ns. The Gaussian height decay of 300K ensemble is further reduced (a), and correspondingly the potential distribution is more Gaussian-shaped (b).

$$N(\mu-\sigma_G{\sim}\mu+\sigma_G)=N(-\infty{\sim}+\infty)\cdot\int_{\mu-\sigma_G}^{\mu+\sigma_G}P(U)dU \qquad (16)$$

where $N(-\infty{\sim}+\infty)$ is equal to the total number of Gaussian deposition steps, and $P(U)$ is expressed as Eq. (7). The ratio of $\frac{\omega N_{16}(\mu-\sigma_G{\sim}\mu+\sigma_G)}{k_B\Delta T_{16}}:\frac{\omega N_8(\mu-\sigma_G{\sim}\mu+\sigma_G)}{k_B\Delta T_8}$ (subscript 16 for S_{16} and 8 for S_8) is simplified to:

$$\frac{N_{16}(-\infty{\sim}+\infty)\cdot\int_{\mu-\sigma_G}^{\mu+\sigma_G}P_{16}(U)dU}{\gamma_{16}-1}:\frac{N_8(-\infty{\sim}+\infty)\cdot\int_{\mu-\sigma_G}^{\mu+\sigma_G}P_8(U)dU}{\gamma_8-1} \qquad (17)$$

In Eq. (17), $P_8(U)$ has a larger width than $P_{16}(U)$, thus fewer samples visit a specific CV region in equal simulation time in S_8 than in S_{16}. This is also illustrated in Fig. 3c, where the visiting times for the selected region in S_8 is only half as many as in S_{16}. The expansion of simulation time only enlarges the $N(-\infty{\sim}+\infty)$ term. If Eq. (17) has a value of 1, we will have the ratio of $N_{16}(-\infty{\sim}+\infty):N_8(-\infty{\sim}+\infty)\approx 1:10$, which means that in our simulation, the S_8 reaches the same convergence level as S_{16} in 300 ns of WTE-establishing time. Clearly the cost is huge to build up the WTE framework with a large bias factor. Thus, an oversized bias factor in our algorithm is not recommended, in spite of the reduction of the total number of replicas.

4. Conclusion

In this study, we come up with an algorithm to generate the temperature sequence and the corresponding bias factor for PT-WTE simulations to achieve desired exchange acceptance rates. The number of replicas that are required to cover a specific temperature range is adjustable by taking bias factor into consideration keeping all AAPs between neighboring replica-pairs to a fixed value.

We have tested two simulation series with the number of replicas as 16 and 8 and bias factors as 37.9 and 174.2, respectively. Bias factor of 174.2 is proved to be oversized for the application of the algorithm. This algorithm is beneficial for the preparation of PT-WTE or PTMetaD-WTE simulations with greatly reduced computational resource requirement.

Acknowledgment

We thank the funding support from a Singapore Ministry of Education Academic Research Fund Tier 3 (MOE2013-T3-1-002), a Singapore Ministry of Health Industry Alignment Fund (NMRC/MOHIAFCAT2/003/2014) and MOE Tier 1 Grant RG 146/17. We also thank the National Supercomputing Centre (NSCC), Singapore for providing the computational resources. Yang Liu acknowledges the support of PhD Research Scholarships from MOE.

References

[1] D.J. Earl, M.W. Deem, Phys. Chem. Chem. Phys. 7 (2005) 3910.
[2] Y. Sugita, Y. Okamoto, Chem. Phys. Lett. 314 (1999) 141.
[3] N. Rathore, M. Chopra, J.J. de Pablo, J. Chem. Phys. 122 (2005) 024111.
[4] D.A. Kofke, J. Chem. Phys. 117 (2002) 6911.
[5] D.A. Kofke, J. Chem. Phys. 121 (2004) 1167.
[6] D.A. Kofke, J. Chem. Phys. 120 (2004) 10852.
[7] A. Patriksson, D. van der Spoel, Phys. Chem. Chem. Phys. 10 (2008) 2073.
[8] M.K. Prakash, A. Barducci, M. Parrinello, J. Chem. Theory. Comput. 7 (2011) 2025.
[9] M. Bonomi, M. Parrinello, Phys. Rev. Lett. 104 (2010) 190601.
[10] A. Barducci, G. Bussi, M. Parrinello, Phys. Rev. Lett. 100 (2008) 020603.
[11] M. Deighan, J. Pfaendtner, Langmuir 29 (2013) 7999.
[12] A. Barducci, M. Bonomi, M.K. Prakash, M. Parrinello, Proc. Natl. Acad. Sci. U S A 110 (2013) E4708.
[13] F. Doro, G. Saladino, L. Belvisi, M. Civera, F.L. Gervasio, J. Chem. Theory. Comput. 11 (2015) 1354.
[14] Y. Liu, Y. Liu, C.-P.M. B., Y. Mu, Sci. Rep. 7 (2017).
[15] M. Deighan, M. Bonomi, J. Pfaendtner, J. Chem. Theory. Comput. 8 (2012) 2189.
[16] A. Amadei, M.E.F. Apol, A. Di Nola, H.J.C. Berendsen, J. Chem. Phys. 104 (1996) 1560.
[17] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, Comp. Phys. Comm. 95 (1995) 43.
[18] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R.A. Broglia, Comp. Phys. Comm. 180 (2009) 1961.
[19] J.W. Neidigh, R.M. Fesinmeyer, N.H. Andersen, Nat. Struct. Biol. 9 (2002) 425.
[20] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Proteins 65 (2006) 712.
[21] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.
[22] G. Bussi, D. Donadio, M. Parrinello, J. Chem. Phys. 126 (2007) 014101.
[23] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, J. Comput. Chem. 18 (1997) 1463.
[24] S. Miyamoto, P.A. Kollman, J. Comput. Chem. 13 (1992) 952.
[25] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, J. Chem. Phys. 103 (1995) 8577.

[26] M. Hackel, H.J. Hinz, G.R. Hedwig, J. Mol. Biol. 291 (1999) 197.
[27] T. Kuroki, N. Kagawa, H. Endo, S. Tsuruno, J.W. Magee, J. Chem. Eng. Data 46 (2001) 1101.
[28] I.M. Abdulagatov, V.I. Dvoryanchikov, A.N. Kamalov, J. Chem. Eng. Data 43 (1998) 830.
[29] E. ToolBox, Water-Heat Capacity (Specific Heat), https://www.engineeringtoolbox.com/specific-heat-capacity-water-d_660.html, 2004.
[30] J. Pátek, J. Hrubý, J. Klomfar, M. Součková, A.H. Harvey, J. Phys. Chem. Ref. Data 38 (2009) 21.
[31] P. Tiwary, M. Parrinello, J. Phys. Chem. B 119 (2015) 736.
[32] L. Qin, S.A. Pabit, A.E. Roitberg, S.J. Hagen, J. Am. Chem. Soc. 124 (2002) 12952.