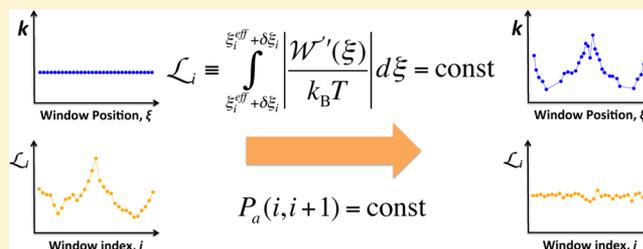# Theory of Adaptive Optimization for Umbrella Sampling

Soohyung Park* and Wonpil Im*

Department of Molecular Biosciences and Center for Bioinformatics, The University of Kansas, 2030 Becker Drive, Lawrence, Kansas 66047, United States

S *Supporting Information*

ABSTRACT: We present a theory of adaptive optimization for umbrella sampling. With the analytical bias force constant obtained from the constrained thermodynamic length along the reaction coordinate, the windows are distributed to optimize the overlap between neighbors. Combining with the replica exchange method, we propose a method of adaptive window exchange umbrella sampling. The efficiency gain in sampling by the present method originates from the optimal window distribution, in which windows are concentrated to the region where the free energy is steep, as well as consequently improved random walk.

$$\mathcal{L}_i \equiv \int_{\xi_i^{eff}+\delta\xi_i}^{\xi_i^{eff}+\delta\xi_i} \left| \frac{\mathcal{W}'(\xi)}{k_B T} \right| d\xi = \text{const}$$

$$P_a(i, i+1) = \text{const}$$

INTRODUCTION

The potential of mean force (PMF), $\mathcal{W}(\mathbf{X})$, as a function of a set of collective variables (CVs), $\mathbf{X}$, can be formally written as

$$\mathcal{W}(\mathbf{X}) = -k_B T \ln\left\{ \int d\mathbf{R}\, \delta[\mathbf{X}'(\mathbf{R}) - \mathbf{X}] e^{-\beta U(\mathbf{R})} \right\} \tag{1}$$

where k_B is the Boltzmann constant, T is the temperature, $\beta = (k_B T)^{-1}$, $\delta(\mathbf{z})$ is a delta function of $\mathbf{z}$, U is the potential energy, and $\mathbf{R}$ is a particular configuration of the system. Clearly, the quality of the calculated PMF relies on sampling in the configuration space, $\{\mathbf{R}\}$. Extensive sampling to obtain the PMF of a complex system is generally difficult even by long molecular dynamics (MD) or Monte Calro (MC) simulations. To better sample along the CVs, especially in steep-PMF regions, the umbrella sampling (US) method[1] has been extensively employed.

There have been many excellent US methods such as the adaptive US,[2] self-healing US,[3] Gaussian mixture adaptive US,[4] and self-learning adaptive US[5] to name a few. Except the last one, all the mentioned methods aim to achieve uniform sampling along the CVs by applying an adaptively estimated umbrella potential, which would eventually allow diffusive sampling along the CVs. The last method aims to efficiently determine (multidimensional) free energy landscape around the transition (minimum free energy) pathway by gradually adding windows over low free energy regions.

A typical system for (regular) US simulations consists of a set of windows with umbrella potentials that restrain the CVs near the center of each window, and the simulation data are postanalyzed to calculate the PMF by the weighted histogram analysis method[6] or the mean force integration.[7,8] For meaningful and efficient US simulations, the overlap between neighboring windows should be sufficient and well controlled. However, there has been no general theory to control the fine overlap because the window centers are system-dependently

shifted to the nearest PMF minimum. In order to control the overlap systematically, there should be a quantitative measure of the overlap between a pair of windows. The average acceptance probability P_a is such a measure, which has been originally derived in the theory of replica exchange (REX) methods.[9−13] P_a describes the overlap between two replicas quantitatively and thus can be directly applied to control the overlap between neighboring windows in US as well.

Recently, there have been efforts in the systematic control of the overlap between windows for the US REX method.[9,14] The first approach in this direction is our previous work,[9] where we derived a simple relationship between the window force constant k and the window spacing d based on an analytical expression for P_a,

$$z_{opt}^2 = \beta k d^2 / 2 \tag{2}$$

where $z_{opt} = 0.8643$, which provides the optimal window overlap (i.e., $P_a \approx 0.39$) based on the first passage time optimization of REX methods.[9−11] Equation 2 assumes that the PMF has minimal effects on shifts of the window centers. However, when the effect of the PMF becomes significant, the window parameters based on eq 2 are no longer optimal, although it still provides a good initial guess. In addition, the optimal number of windows for efficient US simulations still needs to be determined in an empirical (and system-dependent) manner.

More recently, Dashti and Roitberg[14] proposed a protocol for the optimization of US REX by realizing uniform P_a between replicas, which provides a way to determine where

to put windows and how many windows are needed. However, the bias force constant were kept constant and there were only the examples with the fixed number of windows. Without optimization of the bias force constant, the statistical uncertainty in the calculated PMF from US would not be minimized[15,16] and thus its convergence would be slow.

The concept of the thermodynamic length[15−18] provides a rigorous way to minimize the statistical uncertainty by finding a pathway that minimizes the dissipation,[15,16,18] which have been applied in optimization of thermodynamic processes.[19−21] Using the information theory, the connection between the thermodynamic length and the statistical uncertainty in the free energy calculation has been studied.[15,22−24] In these studies, it has been shown that the statistical uncertainty is minimized when intermediate states are spaced along a minimum distance pathway connecting these states,[15,16,18] which lies on the geodesic connecting the end states.[15,16]

In this work, to resolve these general issues in US simulations (efficient sampling with minimal statistical uncertainty), we present a theory of adaptive optimization for umbrella sampling, which is based on equidistant window distribution along the thermodynamic length. Here, we consider one-dimensional US ($\mathbf{X} = \xi$), which can be readily combined with REX methods.[25,26] In the following, the theoretical background for the present method is described. Then, we derive an analytical expression for the optimal (PMF-dependent) bias force constant for US and describe an adaptive optimization that distributes the windows with the optimal bias force constant by constraining P_a. We validate the method and explore its sampling efficiency along the hidden (or orthogonal) degrees of freedom (DOF) by using a toy-model. We then examine the efficacy and the limitation of the present method by applying it to a more realistic (and complicated) example, the assembly of the transmembrane domain of glycophorin A (GpA-TM).

THEORETICAL BACKGROUND

Let us consider a system whose canonical distribution of ξ is given by $p(\xi) = \exp[-\beta \mathcal{W}(\xi)]/Z$, where $Z = \int d\xi \exp[-\beta \mathcal{W}(\xi)]$. The corresponding (regular) US system consists of N windows and each window i is under the effective potential, $U_{\text{eff}}(\mathbf{R}) = U(\mathbf{R}) + w_i(\xi|\xi_i)$, where $w_i(\xi|\xi_i)$ is the window potential to restrain sampling of ξ near ξ_i ($i = 0, 1,...,N - 1$). We assume that the sampling range Λ is continuous (i.e., $\Lambda = [\xi_{\min}, \xi_{\max}]$). The sampling range at window i, Λ_i, can be conveniently described by the variance of sampled ξ, $\sigma_{\xi_i}^2$, and then, their average for the entire windows becomes $\sigma_{\bar{\xi}}^2 = \sum_i \sigma_{\xi_i}^2/N$. For the uniform sampling along ξ, $\sigma_{\bar{\xi}}^2 = \sigma_\xi^2$ should hold, which is equivalent to the minimization of the total variance, $\sum_i \sigma_{\xi_i}^2$ (see Figure 1A for schematic description). Note that the adaptive bias force and potential methods[27−30] aim to achieve this criterion by flattening the effective potential along ξ.

However, in the framework of US, this criterion does not imply the optimal sampling for efficient PMF calculations because the variance of the estimated PMF at each window, $\sigma_{\mathcal{W}_i'}^2$, is larger where the PMF changes faster and is smaller near the PMF minima, which increases the statistical errors and thus results in slow convergence of the calculated PMF. It has been known that the minimization of the statistical errors in the calculated free energy is equivalent to that of the total variance
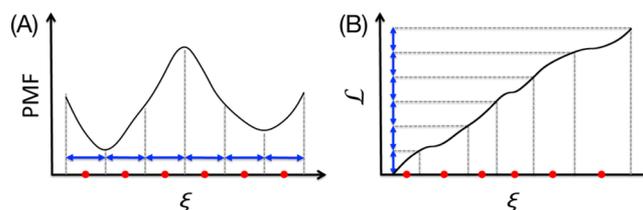

Figure 1. Schematic description of the equidistant window distribution with respect to (A) the sampling range Λ and (B) the thermodynamic length $\mathcal{L}$ calculated from the PMF in (A) by eq 3. By equipartitioning Λ or $\mathcal{L}$ (blue arrows), the boundaries for each window can be easily determined (vertical gray dashed lines) and the midpoint can be assigned to the center of each window (red circles). Note that the windows are not equally distributed along ξ in (B).

of the PMF estimated from intermediate states (i.e., windows in US) connecting two end states.[15,16,23,24] In the context of US, the equidistant window distribution along the thermodynamic length corresponds to the window distribution that minimizes $\sum_i \sigma_{\mathcal{W}_i'}^2$, which would result in the maximized PMF convergence (with the least statistical errors). Therefore, we consider the thermodynamic length as an alternative metric rather than ξ.

The thermodynamic length $\mathcal{L}$ along a given path connecting two states can be defined as the length traveled along the free energy coordinate.[15−18] Assuming that the PMF is known (or can be estimated from the data of the previous US simulations), $\mathcal{L}$ can be written as

$$\mathcal{L} = \int_{\xi \in \Lambda} d\xi |\overline{W'}(\xi)| = \int_{\xi_{\min}}^{\xi_{\max}} d\xi |\overline{W'}(\xi)| \tag{3}$$

where $\bar{z} \equiv \beta z$. Similarly, we define the thermodynamic length for a given window i, $\mathcal{L}_i$, as the length between two states, $\xi_i^{\text{eff}} - \delta \xi_i$ and $\xi_i^{\text{eff}} + \delta \xi_i$, where ξ_i^{eff} is the center of the window and $\delta \xi_i$ is its half-range, which is defined by $\delta \xi_i \equiv \alpha \sigma_{\xi_i}$. Then, we define the range $\Lambda_i = [\xi_i^{\text{eff}} - \delta \xi, \xi_i^{\text{eff}} + \xi_i]$ as the thermally accessible range of ξ at window i, that is, $2k_B T$ range of ξ with respect to the effective PMF,[9] which is the sum of the PMF and the window potential. The above definition is reasonable because the majority of sampled ξ is within this range. With the Gaussian approximation of the sampled ξ-distribution[31] (i.e., the harmonic approximation of the effective PMF), one can determine $\alpha = 2$ (see below). Note that Λ_i is wider than that from the equipartition of $\mathcal{L}$ for optimal overlap between neighboring windows. The sampling along the free energy coordinate (PMF) would be optimal when $\mathcal{L}_i = \lambda$ (constant) for all the windows[16] (schematically shown in Figure 1B), which results in the following constraint:

$$\mathcal{L}_i = \int_{\xi_i^{\text{eff}} - \delta \xi_i}^{\xi_i^{\text{eff}} + \delta \xi_i} d\xi |\overline{W'}(\xi)| = \lambda \tag{4}$$

A naturally arising question is how to determine the optimal bias force constant (eq 12) for each window that satisfies eq 4.

Optimal Bias Force Constant. To address this question, let us consider a typical harmonic window potential, $w_i(\xi|\xi_i) = k_i(\xi - \xi_i)^2/2$. We start with the assumption that the distribution of sampled ξ, $p_b(\xi|k_i,\xi_i)$, can be well approximated by a Gaussian distribution $p_G(\xi|k_i^{\text{eff}}, \xi_i^{\text{eff}})$, that is, $p_b(\xi|k_i, \xi_i) \approx p_G(\xi|k_i^{\text{eff}}, \xi_i^{\text{eff}})$, where k_i^{eff} is the effective bias force constant and ξ_i^{eff} is a shifted center from the (original or input) window center ξ_i due to the PMF effect:[31]

$$p_b(\xi|k_i, x_i) \propto \exp[-\overline{W}(\xi) - \overline{k_i}(\xi - \xi_i)^2/2]$$

$$p_G(\xi|k_i^{eff}, \xi_i^{eff}) \propto \exp[-\overline{k_i^{eff}}(\xi - \xi_i^{eff})^2/2] \tag{5}$$

where $\overline{k_i^{eff}} = 1/\sigma_{\xi,i}^2$. The effective PMF, $\mathcal{W}_i^{eff}(\xi) \equiv \mathcal{W}(\xi) + k_i(\xi - \xi_i)^2/2$, is then approximated to $\mathcal{W}_i^{eff}(\xi) \approx k_i^{eff}(\xi - \xi_i^{eff})^2/2$. Near ξ_i^{eff}, $\mathcal{W}(\xi)$ can be approximated to

$$\mathcal{W}(\xi) \approx \mathcal{W}(\xi_i^{eff}) + \mathcal{W}'(\xi_i^{eff})\Delta\xi_i + \frac{\mathcal{W}''(\xi_i^{eff})}{2}(\Delta\xi_i)^2 \tag{6}$$

where $\Delta\xi_i = \xi - \xi_i^{eff}$. By substituting eq 6 into eq 5 and comparing the exponents between $p_b(\xi|k_i,\xi_i)$ and $p_G(\xi|k_i^{eff}, \xi_i^{eff})$, one can obtain the relations between the window parameters and the effective ones:

$$k_i^{eff} - k_i = \mathcal{W}''(\xi_i^{eff})$$

$$\xi_i^{eff} - \xi_i = \mathcal{W}'(\xi_i^{eff})/[\mathcal{W}''(\xi_i^{eff}) - k_i^{eff}] \tag{7}$$

It should be noted that, without properly chosen k_i^{eff}, eq 4 is not satisfied in general. Therefore, for an optimal US, one needs to find a set of $\{k_i^{eff}, \xi_i^{eff}\}$ that satisfies eq 4 for all the windows. The corresponding set of window parameters ($\{k_i, \xi_i\}$) for the subsequent US simulations are then determined by eq 7.

An analytic solution of eq 4 can be obtained as follows. We start from rewriting eq 4 as

$$\mathcal{L}_i = \int_{\xi_i^{eff}-\delta\xi_i}^{\xi_i^{eff}} d\xi|\overline{W}'(\xi)| + \int_{\xi_i^{eff}}^{\xi_i^{eff}+\delta\xi_i} d\xi|\overline{W}'(\xi)| = \lambda \tag{8}$$

By noting that $|\mathcal{W}'(\xi)| \approx |\mathcal{W}_i' + \mathcal{W}_i''\Delta\xi_i| \leq |\mathcal{W}_i'| + |\mathcal{W}_i''\Delta\xi_i|$ for ξ near ξ_i^{eff}, we obtain

$$\lambda \leq 2|\overline{W}_i'|\delta\xi_i + |\overline{W}_i''|\delta\xi_i^2 = 2|\overline{W}_i'|(\alpha\sigma_{\xi,i}) + |\overline{W}_i''|(\alpha\sigma_{\xi,i})^2 \tag{9}$$

where $\mathcal{W}_i' = \mathcal{W}'(\xi_i^{eff})$ and $\mathcal{W}_i'' = \mathcal{W}''(\xi_i^{eff})$. Then, by assuming the equality in eq 9 and solving it, we obtain

$$\overline{k_i^{eff}} = \frac{2\alpha^2}{\lambda^2}\left[\overline{W_i'}^2 + \frac{\lambda}{2}|\overline{W_i''}|\right.$$
$$\left. + \sqrt{\left(\overline{W_i'}^2 + \frac{\lambda}{2}|\overline{W_i''}|\right)^2 - \frac{\lambda^2}{4}\overline{W_i''}^2}\right] \tag{10}$$

The (unknown) parameters α and λ in eq 10 can be determined as follows. By noting that the effective PMF can be approximated to a harmonic function, the thermodynamic length with respect to $\mathcal{W}_i^{eff}(\xi)$ for Λ_i can be written as

$$\mathcal{L}_i^{eff} = \int_{\xi_i^{eff}-\delta\xi_i}^{\xi_i^{eff}+\delta\xi_i} d\xi\left|\frac{d\mathcal{W}_i^{eff}}{d\xi}\right| = \overline{k_i^{eff}}\sigma_{\xi_i}^2\alpha^2 = \alpha^2 \tag{11}$$

Because we defined Λ_i as the thermally accessible range, $\mathcal{L}_i^{eff} = 4$, we obtain $\alpha = 2$. Hereafter, we omit the window index unless it is necessary, and consider k and k^{eff} as a function of ξ. One can show that $\lambda \leq \alpha^2$ for US, which implies that the maximum $\mathcal{L}_i$ cannot exceed $\mathcal{L}_i^{eff}(= 4)$. By matching $\mathcal{L}_i$ and

$\mathcal{L}_i^{eff}$ ($\lambda = \alpha^2 = 4$), we obtain an analytical expression for the optimal effective bias force constant $k^{eff}(\xi)$

$$\overline{k^{eff}}(\xi) = \frac{\overline{W'}^2}{2} + |\overline{W''}| + \sqrt{\left(\frac{\overline{W'}^2}{2} + |\overline{W''}|\right)^2 - \overline{W''}^2} \tag{12}$$

Adapative Optimization for Umbrella Sampling. With the effective bias force constant (eq 12) that satisfies eq 4 (i.e., the equidistant constraint along the thermodynamic length), the next question for optimal US is how to distribute windows with well controlled overlap between neighbors. To achieve this, we make use of P_a (the average acceptance probability) as a quantitative measure of the overlap between a pair of windows, which can be formally written for US as

$$P_a \approx \int\int d\xi_1 d\xi_2 \, P_{ex} \, p_G(\xi_1|k_i^{eff}, \xi_i^{eff}) p_G(\xi_2|k_j^{eff}, \xi_j^{eff}) \tag{13}$$

where subscripts 1 and 2 are indices for a pair of windows i and j, $P_{ex} = \min\{1,\exp(-\beta\Delta)\}$ is the metropolis criterion for the exchange between the pair with Δ being the difference between the effective potential before and after the exchange of the windows. In the theories of the optimization of REX methods[10−13] and in our previous work,[9] the optimal performance of window exchange umbrella sampling (WEUS) is achieved when P_a between neighboring windows is a constant whose value is in 0.28−0.4, depending on the exchange schemes. In this work, we set P_a to 0.4, which is optimal for the so-called "even-odd" exchange scheme.[9,10] For efficient evaluation of eq 13, we obtained an accurate analytic expression for P_a (see Appendix for details).

With eqs 7, 12, and 13, for a given $\mathcal{W}(\xi)$ calculated after a (short) US simulation, the windows can be optimally redistributed over Λ as follows. Let us consider a window at one fixed end-point, $\xi_0 = \xi_{min}$. The effective window parameters, ξ_0^{eff} and $k_0^{eff} = k^{eff}(\xi_0^{eff})$, and the (updated) bias force constant k_0 are determined from eqs 7 and 12. The ξ_1^{eff} and k_1^{eff} of the next window 1 are then determined by satisfying the constraint, $P_a = 0.4$, and the corresponding ξ_1 and k_1 are obtained by eq 7. This procedure is repeated until Λ is covered by the windows. It should be noted that the extensive evaluation of P_a can be efficiently performed by using the analytic expression for eq 13. After the redistribution, an estimate of the optimal number of windows, N_{opt}, is obtained as an outcome, which may be different from N for the previous simulation.

In a typical US simulation, however, N is not changed. By a simple modification of eq 12 and the window distribution procedure, one can still optimize US with a fixed N. First, let us consider $k^{eff}(\xi)$ when $N > N_{opt}$ ($\rho < 1$ where $\rho = N_{opt}/N$). In this case, one can naturally scale λ by ρ ($\lambda \to \rho\lambda$) in eq 4. Then, $k^{eff}(\xi)$ for optimal US with N windows can be obtained by substituting $\mathcal{W}'$ and $\mathcal{W}''$ with $\mathcal{W}'/\rho$ and $\mathcal{W}''/\rho$ in eq 12 as

$$\overline{k^{eff}}(\xi) = \frac{\overline{W'}^2}{2\rho^2} + \frac{|\overline{W''}|}{\rho} + \sqrt{\left(\frac{\overline{W'}^2}{2\rho^2} + \frac{|\overline{W''}|}{\rho}\right)^2 - \frac{\overline{W''}^2}{\rho^2}} \tag{14}$$

When $N < N_{opt}$ ($\rho > 1$), the above scaling cannot be applied because $k(\xi)$ could be negative around the PMF minima. By decreasing the target value of P_a for the distribution of windows, $k^{eff}(\xi)$ for $\rho = 1$ can still be used in the optimization

procedure. However, lowering P_a implies localized sampling around the window centers. Therefore, to maintain P_a, we sacrifice the optimal sampling by adding a bias force constant

$$k_a^{\text{eff}}(\xi) = (1 - \rho^{-1})|\mathcal{W}''(\xi)| \tag{15}$$

so that the same scaling can be applied for $\rho > 1$ as well. Therefore, the optimization of the distribution of N windows can be performed with the adjustable parameter ρ. This modified optimization procedure can be readily combined with the WEUSMD method, which is termed adaptive WEUSMD (aWEUSMD).

■ SIMULATION DETAILS

To validate and explore the sampling efficiency of aWEUSMD along the hidden (or orthogonal) DOF, we considered two systems: (1) a toy model and (2) the assembly of the GpA-TM.

The toy model is designed to examine not only the sampling along an explicit CV but also an efficiency gain in sampling along the orthogonal DOF if any. In the model, a particle (whose mass is the same as phosphorus) is moving under the potential $U(x,y) = U_0(x,y) + U_R(x,y)$ (in kcal/mol). As shown in Figure 2A, it has four Gaussian wells and barriers along x- and y-directions, which are described by $U_0(x,y)$:

$$U_0(x, y) = 4 - 2e^{-\|\mathbf{r}-\mathbf{r_1}\|^2/2} - 4e^{-\|\mathbf{r}-\mathbf{r_2}\|^2/2} - 2e^{-\|\mathbf{r}-\mathbf{r_3}\|^2/2}$$
$$- 4e^{-\|\mathbf{r}-\mathbf{r_4}\|^2/2} + H_y e^{-y^2/2} + H_x e^{-x^2/2} \tag{16}$$

where $\mathbf{r} = (x,y)$, $\mathbf{r_1} = (4,4)$, $\mathbf{r_2} = (-4,4)$, $\mathbf{r_3} = (-4,-4)$, $\mathbf{r_4} = (4,-4)$, $H_x = 5$, and H_y is the adjustable barrier along y-direction. The particle is restrained in a bounded area by $U_R(x,y)$:

$$U_R(x, y) = 1.5[R(|x| - 5)^2 + R(|y| - 5)^2]$$
$$R(z) = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

By varying H_y from 5 to 6.5 with the increment of 0.5, we performed 120-ns Langevin dynamics MD simulations of USMD, WEUSMD, and aWEUSMD for a set of the number of windows, $N \in \{21,31,41,51,61\}$, with two fixed end points at $x_{\min} = -6$ Å and $x_{\max} = 6$ Å. Initially, the windows were equally distributed along x with $d = (x_{\max} - x_{\min})/(N - 1)$ and k was obtained from eq 2. We have performed nine simulations for each set of H_y and N, where the initial x-coordinate was set to

each window center. The initial y-coordinate y_0 was set to $y_0 = 3.8 + 0.05r$, where $r \in \{0,1,\cdots,8\}$.

For a more realistic (and complicated) example, we considered the assembly of the right-handed GpA-TM dimer because it has been extensively studied for transmembrane (TM) helix assembly[32−35] and represents one of the tightest interfacial packing (with the well-known GxxxG motif) among all the known TM helix interfaces. Such tight packing makes it challenging to sample the conformational space (e.g., right-handed to left-handed dimer at short helix−helix distance r_{HH}) and thus makes the GpA-TM to be a valuable target to assess the sampling power.[35] We defined the GpA-TM as [72]EITLI IFGVM AGVIG TILLI SYGIR[96], in which we have chosen C_α atoms of residues 76−92 to define the helix principal axis for r_{HH} and the crossing angle Ω.[36,37] For the NMR structure (PDB: 1AFO),[38] the calculated r_{HH} and Ω are 6.59 Å and −39.2°, respectively.[39] To test and compare the sampling power of WEUSMD and aWEUSMD, we prepared three sets of initial configurations of GpA-TM: left-handed (IS1, $\Omega < 0$), right-handed (IS2, $\Omega > 0$), and parallel (IS3, $\Omega \approx 0$) helix−dimer interfaces.

The initial configurations were generated from the NMR structure by translating each helix along r_{HH} and then by rotating the helices along the helix−helix distance vector for a desired Ω. A total of 80 windows were generated for each initial configuration set in a r_{HH} range 5.2−21 Å with $d = \Delta r_{HH} = 0.2$ Å and the restraint force constant k obtained from eq 2 for WEUSMD (initial k for aWEUSMD). Then, we have performed 200-ns WEUSMD and aWEUSMD for IS1, IS2, and IS3. For computational efficiency, the IMM1 implicit membrane model[40] was used with the hydrophobic thickness of 23 Å to mimic a dimyristoylphosphatidylcholine (DMPC) lipid bilayer.

All the simulations were performed using CHARMM.[41] For GpA-TM simulations, we used the SHAKE algorithm and default IMM1 options. We used a time step of 2 fs and the collision frequency $\gamma = 5$ ps^{-1}, which is a typical value for Langevin dynamics with proteins. Window exchanges were attempted every 1 ps and controlled by the REPDSTR module[42] in CHARMM. In aWEUSMD, the window parameters were updated every 0.2 ns. The adaptive update were controlled by in-house Python scripts, which take the window parameters and the time series of ξ (x or r_{HH}) for all the windows from the previous simulation as input. The derivatives of the PMF ($\mathcal{W}'$ and $\mathcal{W}''$) were calculated, for the numerical stability, using the cubic spline interpolation of $\mathcal{W}'$ obtained by the umbrella integration method.[31] The resulting derivatives were used to obtain the updated window parameters for the next simulation by eqs 7 and 12−15. We limited the maximum shift of ξ_i in one update cycle by $d/4$ for its smooth evolution.

■ RESULTS AND DISCUSSION

Toy Model. In this section, we show the results from a representative parameter set, $H_y = 6$ and $N = 31$ (see Supporting Information section S1 for the results from all the tested parameter sets). The one-dimensional (1D) PMF, $\mathcal{W}(x)$, calculated from the results of USMD, WEUSMD, and aWEUSMD for $y_0 = 4.1$ are shown in Figure 2B. The accuracy of the aWEUSMD-PMF is the best among the three methods, while the WEUSMD-PMF is marginally better than the USMD-PMF. The most obvious reason for the improved
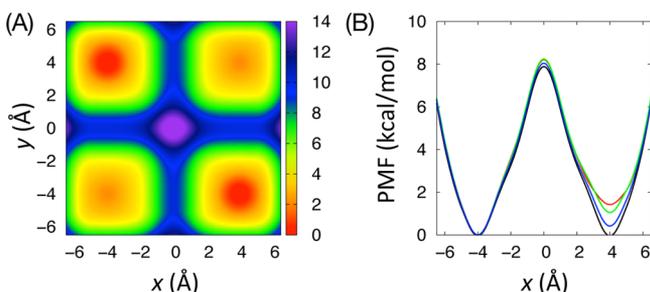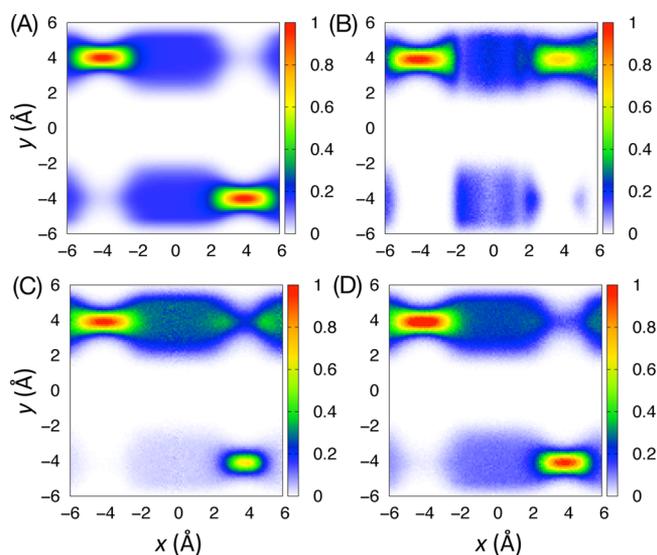

Figure 2. (A) Potential energy, $U(x,y)$, of the model system with $H_y = 6$ in eq 16. (B) The one-dimensional PMF, $\mathcal{W}(x)$, calculated from $U(x,y)$ (black) and the results of 120-ns simulation data from USMD (red), WEUSMD (green), and aWEUSMD (blue) with $N = 31$ windows. $\mathcal{W}(x)$ from $U(x,y)$ is calculated by eq 1 and those from simulations are calculated by WHAM.

Figure 3. Conditional probability, $P(y|x)$, calculated from (A) $U(x,y)$, (B) USMD, (C) WEUSMD, and (D) aWEUSMD for $H_y = 6$ and $N = 31$. $P(y|x)$ is normalized by $\max\{P(y|x)\| |x| \leq 6$ and $|y| \leq 6\}$ and the binwidth is 0.1 Å for both x- and y-directions.
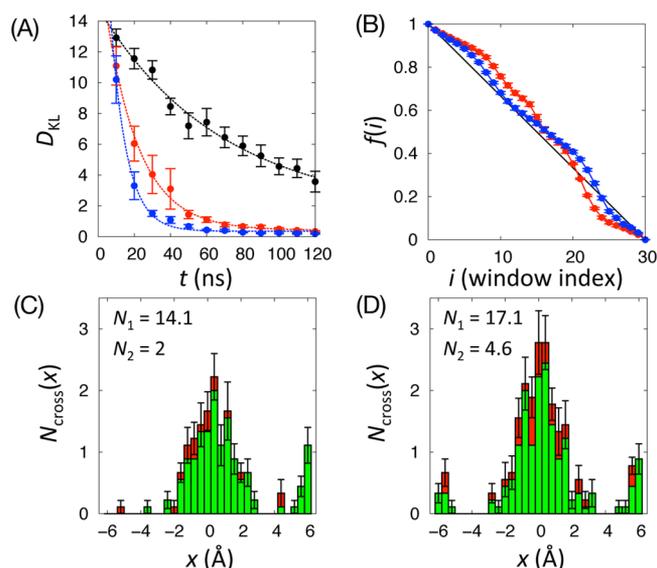

Figure 4. (A) Kullback–Leibler divergence D_{KL} calculated from 10-ns blocks of $P(y|x)$ from USMD (black), WEUSMD (red), and aWEUSMD (blue) by eq 18. Shown together are the exponential fits of D_{KL} (dashed lines). (B) The fraction of replicas that have visited the lowest-index window, $f(i)$, in eq 19, from WEUSMD (red) and aWEUSMD (blue) simulations. Shown together is $f(i)$ for the ideal random walk of replicas (black line). (C–D) The histogram of the number of y-barrier crossing, $N_{cross}(x)$, from (C) WEUSMD and (D) aWEUSMD simulations. $N_{cross}(x)$ is the sum of the number of crossing from $y > 0$ to $y < 0$, $N_1(x)$ (green), and that in the opposite direction, $N_2(x)$ (red). The total number of y-barrier crossing is denoted by N_{cross} and similarly by N_1 and N_2. In all the figures, the data are averaged over nine independent simulations and the error bars represent the standard errors.

accuracy of the PMFs from WEUSMD and aWEUSMD over USMD is the regular window exchange attempts, which facilitates the sampling of configurations. This improvement is clearly shown in Figure 3 by comparing the conditional probability, $P(y|x) \equiv P(x,y)/P(x)$, where $P(x)$ and $P(x,y)$ are the populations sampled by simulations.

The sampling power (i.e., the accuracy of the sampling compared to the reference) can be quantitively estimated by the Kullback–Leibler divergence[43]

$$D_{KL}[P_r(y|x)\|P(y|x)] = \sum_{i,j} \frac{P_r(y_j|x_i)}{N_{bin,x}} \ln\left[\frac{P_r(y_j|x_i)}{P(y_j|x_i)}\right] \quad (18)$$

where $P_r(y|x)$ is calculated from $U(x,y)$, $N_{bin,x}$ is the number of bin size along x, and x_i and y_j are the i- and j-th bins along x and y, respectively. Smaller values of D_{KL} indicate better agreement of $P(y|x)$ with $P_r(y|x)$, and D_{KL} vanishes when they are identical. In addition, the relaxation time of D_{KL}, τ, can be a measure of the efficiency, estimated by fitting $D_{KL}(t)$ to an exponential function: $D_{KL}(t) = a+b \exp(-t/\tau)$. As shown in Figure 4A, the WEUSMD-D_{KL} ($\tau = 18.6 \pm 1.1$ ns) decays much faster than USMD-D_{KL} ($\tau = 68.5 \pm 14.4$ ns), and aWEUSMD-D_{KL} ($\tau = 8.91 \pm 0.47$ ns) relaxes even faster; similar results are obtained for most parameter sets examined (Supporting Information Figure S1 and Table S1).

The efficiency gain in aWEUSMD over WEUSMD can be attributed to the optimal window distribution ($P_a = 0.4$) with $k^{eff}(x)$, where the windows are concentrated in steep PMF (high $|\mathcal{W}'|$) regions. In these regions, the y-barrier is lower than those around 1D-PMF minima ($x = 4$ or -4 in Figure 2B), which results in better sampling along y-direction in aWE-USMD, that is, the increased number of y-barrier crossing N_{cross} (Figures 4C and D). As expected, there is no meaningful increase in N_{cross} from WEUSMD compared to that from USMD (Supporting Information Figure S2 and Table S2). Considering the fact that aWEUSMD optimizes WEUSMD only along the explicit CV (i.e., the x axis), the increase in N_{cross} (along the y axis) is noteworthy.

In addition, there is an efficiency gain from the improved random walk of replicas, which can be quantified as the fraction of replicas that have visited the lowest-index window most recently,[11,44]

$$f(i) = n_0(i)/[n_0(i) + n_{N-1}(i)] \quad (19)$$

where $n_0(i)$ is the accumulated number of replicas with traveling state "0" monitored at window i during the course of simulation and $n_{N-1}(i)$ is similarly defined. The traveling state of a given replica is assigned (or updated) to "0" or "$N - 1$" only when it reaches at the lowest- or the highest-index window, 0 or $N - 1$, respectively. For the ideal random walk of replicas, $f(i)$ should be linear.[11,44] By comparing eq 3 with eq 25 in ref 45 (a work for the optimization of replica's random walk), it is clear that aWEUSMD indeed improves the random walk as well, as shown in Figure 4B; similar results are obtained for most parameter sets examined (Supporting Information Figure S3).

GpA-TM Assembly. In this section, we show the results from the WEUSMD and aWEUSMD for the assembly of the GpA-TM starting with the initial configurations with left-handed (IS1) and right helix–dimer interfaces (IS2) (see Supporting Information section S2.2 for the results from the parallel TM dimer initial configurations, IS3). Figure 5 shows the 1D-PMF as a function of r_{HH}, $\mathcal{W}(r_{HH})$, for IS1 and IS2. There is significant difference between the IS1-PMF and the IS2-PMF from WEUSMD, which is originated from incomplete sampling (see Figures 6A and C). The agreement between the IS1-PMF and the IS2-PMF from aWEUSMD is significantly
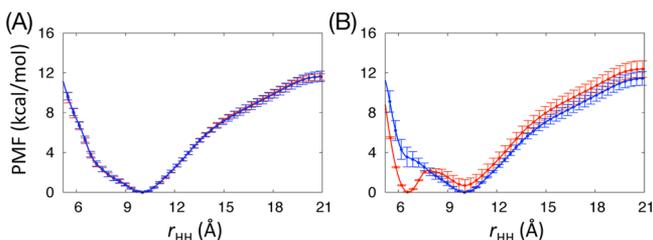
Figure 5. PMF as a function of r_{HH}, $\mathcal{W}(r_{HH})$, calculated from WEUSMD (red) and aWEUSMD (blue) starting from the initial configurations of (A) left-handed (IS1) and (B) right-hand (IS2) helix-dimer interfaces. For the PMF calculation, we used trajectories from 40 ns to 200 ns. The error bars represent the standard deviations calculated from 16 10-ns block average PMFs.

improved, which we attribute to the improved sampling power of aWEUSMD over WEUSMD.

The improved sampling efficiency of aWEUSMD over WEUSMD is supported by the conditional probability $P(\Omega|r_{HH})$, as shown in Figure 6. The agreement between $P(\Omega|r_{HH})$ calculated from aWEUSMD for IS1 and IS2 is improved up to shorter $r_{HH} \approx 7$ Å (Figures 6B and 6D) compared to the results from WEUSMD, which show good agreement for $r_{HH} > 9$ Å. This improvement allows the sampling of the right-handed NMR structure-like conformations in aWEUSMD with IS1 (Figure 6B and Supporting Information Figure S4B), indicating that aWEUSMD indeed promotes the sampling along Ω, where no restraint potential was applied.

The sampling power of aWEUSMD and WEUSMD for the GpA-TM assembly are quantitatively compared by D_{KL} of $P_t(\Omega|r_{HH})$ with respect to $P_r(\Omega|r_{HH})$ from the results of 2D-WEUSMD in our previous work,[35] where $P_t(\Omega|r_{HH})$ is calculated from the trajectories up to time t. As shown in Figure 7, the aWEUSMD-D_{KL} relaxes faster to its estimated steady-state value, D_{KL}^{ss}, compared to the WEUSMD-D_{KL}. For IS1, the estimates of (D_{KL}^{ss}, τ) from aWEUSMD and WEUSMD are $(0.515 \pm 0.004, 10.27 \pm 0.53 \text{ ns})$ and $(1.538 \pm 0.002, 14.20 \pm 0.76 \text{ ns})$, respectively. For IS2, these estimates are $(0.185 \pm$
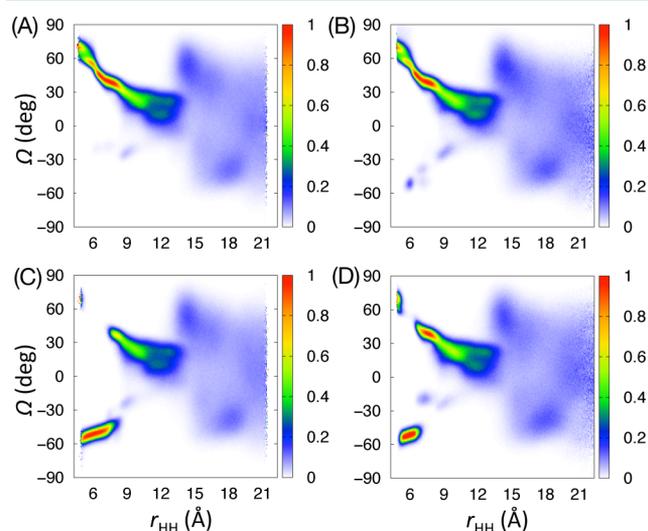

Figure 6. Conditional probability, $P(\Omega|r_{HH})$, calculated from the results of (A) WEUSMD and (B) aWEUSMD for IS1. $P(\Omega|r_{HH})$ calculated from the results of (C) WEUSMD and (D) aWEUSMD for IS2. $P(\Omega|r_{HH})$ is normalized by $\max\{P(\Omega|r_{HH})\}$ and the binwidth is $\Delta r_{HH} = 0.05$ Å and $\Delta \Omega = 6°$.
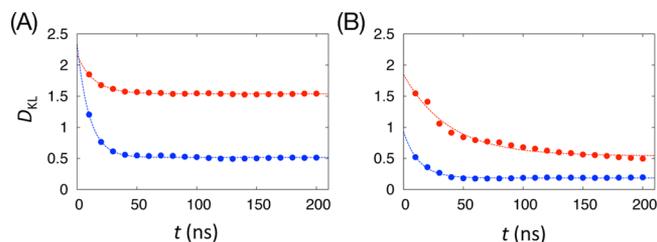

Figure 7. Kullback−Leibler divergence D_{KL} calculated using $P_t(\Omega|r_{HH})$ from WEUSMD (red) and aWEUSMD (blue) for (A) IS1 and (B) IS2 by eq 18 with respect to $P_r(\Omega|r_{HH})$ calculated from 2D-PMF in ref 35. $P_t(\Omega|r_{HH})$ is calculated from trajectories up to time t. Shown together are the exponential fits of D_{KL} (dashed lines).

$0.003, 12.84 \pm 0.89 \text{ ns})$ and $(0.541 \pm 0.021, 37.89 \pm 3.59 \text{ ns})$ for aWEUSMD and WEUSMD, respectively. Although D_{KL} is a quantity that dictates both the accuracy and the efficiency of the sampling, it should be noted that this quantity is not always available in general because the probability distribution of the reference state needs to be known *a prior*.

In aWEUSMD, as shown in Figure 8, the window parameters relax to their limiting values once the PMF is converged (albeit with fluctuations). The relaxation time of the window centers (including the effective ones) was estimated to be about 4 ns (20 update cycles). The estimated τ of the bias force constants (including the effective ones) are about 20 ns and 26.5 ns for IS1 and IS2, respectively. Interestingly, τ for the bias force constants are well correlated to that of D_{KL}. This good correlation enables us to estimate the efficiency by τ of window parameters without the need of D_{KL}. However, it should be noted that the converged window parameters does not guarantee the convergence of the PMF to the true one because aWEUSMD does not actively sample along the hidden DOF.

■ DISCUSSION

A fundamental requirement for proper US is that the sampling along ξ should cover entire Λ with sufficient overlap between neighboring windows. With poorly chosen window parameters, the sampling along ξ could be fragmentized into disconnected partitions in Λ. Such poor sampling would be observed in the cases, for example: (1) equidistant window distribution with too high bias force constant and (2) window distribution with too low bias force constants around steep barriers along ξ. In these cases, the quality of the calculated PMF would be poor because of the lack of the overlap between neighboring windows in certain regions (especially around the barriers). This is an intrinsic limitation (overlap problem) of any US-type methods.
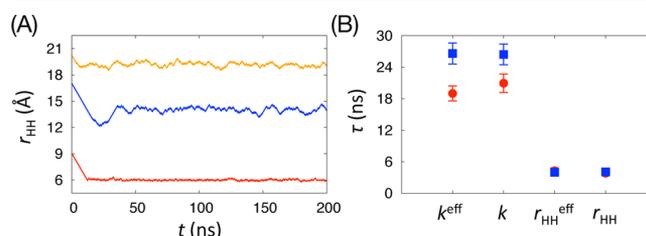

Figure 8. (A) Time-series of windows center ($r_{HH,i}$) for three windows ($i = 19, 59,$ and 75). (B) The average relaxation time τ of window parameters for IS1 (red) and IS2 (blue). The error bars are the standard errors of the average τ over the entire windows.

Such an overlap problem can be alleviated in the present optimization method because of (1) the minimization of the shift in the windows center by the adaptively optimized (PMF-dependent) bias force constants and (2) the concentrated window distribution in the steep PMF regions with well-controlled overlap between neighboring windows by constraining P_a. However, it should be noted that, when $N \ll N_{opt}$ ($\rho \gg 1$), aWEUSMD with fixed N cannot overcome this problem although it performs better than WEUSMD and USMD (data not shown). By relaxing the fixed N condition, aWEUSMD would be free of the overlap problem along ξ.

In this work, the efficiency of all the tested methods (USMD, WEUSMD, and aWEUSMD) was estimated by the relaxation time (τ) of D_{KL}. We observed that τ of D_{KL} is well correlated with that of window parameters which would be stable once the PMF is converged. Therefore, even in a situation in which one cannot calculate D_{KL} because of unknown reference state, the sampling efficiency (especially, convergence rate) can be estimated by τ of the window parameters (see Figure 8).

As shown in Figure 4A, the efficiency and sampling power of aWEUSMD and WEUSMD are significantly better than those of USMD due to the REX. The quality of REX is improved by aWEUSMD compared to that in WEUSMD, which is more clearly shown for smaller N but becomes less obvious as N increases (Figure S3). In Figure 9, we plotted $f(i)$ obtained
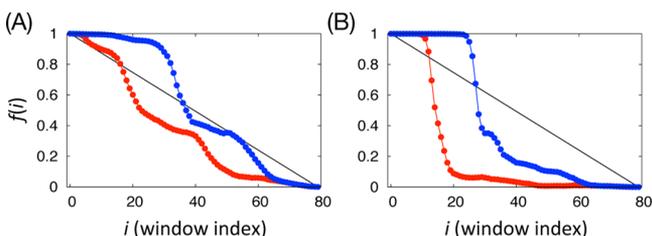

Figure 9. Fraction of replicas that have visited the lowest-index window, $f(i)$, in eq 19, from WEUSMD (red) and aWEUSMD (blue) simulations for (A) IS1 and (B) IS2. Shown together is $f(i)$ for the ideal random walk of replicas (black line).

from the results of WEUSMD and aWEUSMD for the GpA-TM assembly. Although the quality of REX was improved by aWEUSMD, the significant nonlinearity in $f(i)$ indicates that aWEUSMD has limited power in optimization of the flux of replicas, which we attribute to the incomplete sampling along the hidden DOF due to high barriers (>6 kcal/mol).

The REX itself does not improve the barrier crossing along the hidden DOF for a given replica (see Supporting Information Figure S2). However, in aWEUSMD, where the windows are more densely distributed in the steep PMF regions (with optimal bias force constants), there are more barrier crossing events than WEUSMD and USMD. This implies that the present adaptive optimization method indeed improves the sampling efficiency of US along the hidden DOF. However, it should be noted that aWEUSMD does not actively sample the hidden DOF, as attempted in some recent efforts.[28,29] Therefore, the efficiency gain in aWEUSMD over WEUSMD is limited by the barrier height along the hidden DOF. The results from the toy-model suggests that aWEUSMD can overcome the barriers along the hidden DOF, whose height is up to about 6 kcal/mol. It is consistent with the results from the GpA-TM assembly in that the sampling along Ω is improved up to $r_{HH} \approx 7$ Å, at which the barrier height along Ω is about 6 kcal/mol.[35]

CONCLUDING REMARKS

In this work, we present a theory of adaptive optimization for umbrella sampling for optimal PMF calculations based on two constraints: (1) the uniform thermodynamic length for all the windows ($\mathcal{L}_i = 4$) and (2) the uniform average acceptance probability between neighboring windows ($P_a = 0.4$). From the first constraint, we obtained an analytic expression for the effective bias force constant $k^{eff}(\xi)$. Then, the windows are distributed by constraining $P_a = 0.4$, which is known to be optimal for the exchange scheme chosen for the present work.

As an immediate consequence of the present theory, an estimate of the optimal number of windows N_{opt} and a parameter ρ ($=N_{opt}/N$) are obtained. The latter provides a criterion from which one can judge whether the given number of windows N is sufficient or not: when $\rho \approx 1$, N is close to N_{opt}. Sampling along ξ by aWEUSMD would be proper as long as $\rho \leq 1$, whose efficiency can be conveniently estimated by the relaxation time (τ) of the window parameters. However, when $N \ll N_{opt}$, aWEUSMD with fixed N cannot overcome the overlap problem, which is an intrinsic limitation of any US-type methods. By relaxing the fixed N condition, aWEUSMD would be free of the overlap problem along ξ.

The efficiency gain in aWEUSMD over WEUSMD originates from the optimal window distribution concentrated in the steep PMF regions, where more barrier crossing along the hidden (or orthogonal) DOF are feasible, and also from the consequently improved random walk of replicas. As a result, the synergetic improvements in the sampling along both explicit and hidden DOF by aWEUSMD make the PMF converge faster than WEUSMD. However, it should be noted that the advantage of aWEUSMD over WEUSMD is limited by the barrier height along the hidden DOF. The results from the two test cases suggest that aWEUSMD can help sampling along the hidden DOF, whose barrier height is up to about 6 kcal/mol.

In this work, we fixed N in aWEUSMD for comparison with USMD and WEUSMD with the same N. If we relax this condition, a better performance of aWEUSMD would be achieved by employing self-learning scheme, analogous to that in the self-learning adaptive US.[5] In this scheme, windows are added gradually until N reaches N_{opt} starting from aWEUSMD with small N. Then, aWEUSMD would be free of the overlap problem and would exploit its other advantages over WEUSMD (improved random walk of windows and sampling along the hidden DOF) simultaneously. This version of aWEUSMD (self-learning aWEUSMD) would be more useful in one-dimensional US simulations for efficient sampling and PMF calculations.

APPENDIX: AN ACCURATE ANALYTIC EXPRESSION FOR P_A

Here, we derive an accurate analytical approximate expression for P_a between a pair of windows i and j for arbitrary window parameters (ξ_i, k_i) and (ξ_j, k_j), whose corresponding effective window parameters, (ξ_i^{eff}, k_i^{eff}) and (ξ_j^{eff}, k_j^{eff}). Formally, P_a between the pair of windows is given by the integral,

$$P_a = \iint d\xi_1 d\xi_2 P_{ex} p_b(\xi_1|k_i, \xi_i) p_b(\xi_2|k_j, \xi_j)$$
$$\approx \iint d\xi_1 d\xi_2 P_{ex} p_G(\xi_1|k_i^{eff}, \xi_i^{eff}) p_G(\xi_2|k_j^{eff}, \xi_j^{eff}) \quad (20)$$
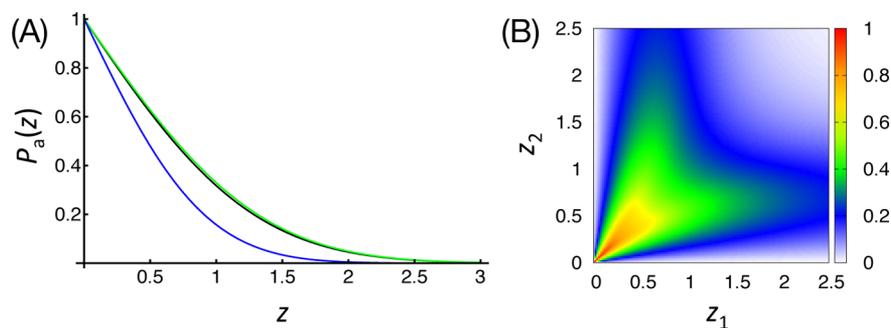
Figure 10. (A) The average acceptance probability, $P_a(z)$, calculated from the exact solution, eq 4 in ref 9. (black), and the equal z limit of eq 27, eq 29 (green). Shown together is $\mathrm{erfc}(z)$ (blue), which is the first term of the approximate expression of P_a given by eq 29. (B) The average acceptance probability as a function of z_1 and z_2, $P_a(z_1,z_2)$, given by eq 27.

where subscripts 1 and 2 are indices for a pair of windows i and j, $P_{ex} = \min\{1, \exp(-\beta\Delta)\}$ is the metropolis criterion for the exchange between the pair, and Δ is

$$
\begin{aligned}
\Delta &= \frac{1}{2}[k_i^{eff}(\xi_2 - \xi_i^{eff})^2 + k_j^{eff}(\xi_1 - \xi_2^{eff})^2 \\
&\quad - k_i^{eff}(\xi_1 - \xi_i^{eff})^2 + k_j^{eff}(\xi_2 - \xi_2^{eff})^2] \\
&= \frac{1}{2}(\xi_2 - \xi_1)[(k_j^{eff} - k_i^{eff})(\xi_2 + \xi_1) \\
&\quad - 2(k_j^{eff}\xi_j^{eff} - k_i^{eff}\xi_i^{eff})]
\end{aligned}
\tag{21}
$$

It can be shown easily that

$$
P_a = 2 \iint_{\Delta>0} d\xi_1 d\xi_2 p_G(\xi_1|k_i^{eff}, \xi_i^{eff}) p_G(\xi_2|k_j^{eff}, \xi_j^{eff})
\tag{22}
$$

Simply, one can evaluate P_a by the numerical integration of eq 13 (or eq 22). However, its analytical expression can provide deeper understanding and insights, which is helpful in further applications. For example, when $k_i^{eff} = k_j^{eff}$, the exact analytical expression of eq 13 is available,[9] which is utilized to obtain eq 2 for the relationship between the window force constant and the window spacing for WEUSMD. In general cases (when $k_i^{eff} \neq k_j^{eff}$), there has been no such analytical expression for P_a, which we are in need of, for fast and accurate evaluation of eq 13 in order to redistribute windows in our adaptive optimization for US method.

The starting point is rewriting eq 13 in a symmetric form as

$$
P_a = 1 - \int_0^\infty dx K_j(x) K_i'(x) + \int_0^\infty dx K_j'(x) K_i(x)
\tag{23}
$$

where, for $m \in \{i,j\}$, the kernels in eq 23 are defined as

$$
K_m(x) \equiv \frac{1}{2}\left[\mathrm{erf}\left(\frac{x - f_m |\Delta\xi_{j,i}^{eff}|}{\sqrt{2}\,\sigma_m} \right) + \mathrm{erf}\left(\frac{x + f_m |\Delta\xi_{j,i}^{eff}|}{\sqrt{2}\,\sigma_m} \right) \right]
\tag{24}
$$

and $\mathrm{erf}(x)$ is the error function of x, $f_m = z_m^2/(z_i^2 - z_j^2)$, $z_m \equiv |\Delta\xi_{j,i}^{eff}|/(\sqrt{2}\sigma_m)$, $\Delta\xi_{j,i}^{eff} \equiv |\xi_j^{eff} - \xi_i^{eff}|$, and $\sigma_m = (\beta k_m^{eff})^{-1/2}$. Hereafter, we assume that $k_j^{eff} > k_i^{eff}$ ($z_i > z_j$). When $k_j^{eff} < k_i^{eff}$, one can use the result after exchanging indices i and j.

To make eq 23 be integrable, we introduced a Gaussian approximation of the error function,

$$
\mathrm{erf}(x) \approx \mathrm{sign}(x)[1 - \exp(2c_1|x| + c_2 x^2)]
\tag{25}
$$

where $\mathrm{sign}(x)$ is the sign function of x. With 100 000 equally distributed supporting points along x in a range of $[1.0 \times$

$10^{-20}, 5.0]$, the coefficients c_1 and c_2 are determined from the numerical fit to $\log[1 - \mathrm{erf}(x)] = 2c_1 x + c_2 x^2$ as $c_1 = -0.425672$ and $c_2 = -0.913347$, respectively. After substitution of eqs 24 and 25 into eq 23, the integrals become those of Gaussian, whose analytic solution is available[46]

$$
\begin{aligned}
&\int dx\, e^{-(ax^2 + 2bx + c)} \\
&= \frac{1}{2}\sqrt{\frac{\pi}{a}} \exp\left[\frac{b^2 - ac}{a} \right] \mathrm{erf}\left(\sqrt{a}\,x + \frac{b}{\sqrt{a}} \right)
\end{aligned}
\tag{26}
$$

Using eq 26, we obtained an analytic approximate expression for P_a as a function of z_1 and z_2,

$$
P_a(z_i, z_j) = P_0(z_i, z_j) + P_1(z_i, z_j) + P_f(z_i, z_j) + P_R(z_i, z_j)
\tag{27}
$$

where

$$
P_0(z_i, z_j) = \frac{\mathrm{erfc}(z_i) + \mathrm{erfc}(z_j)}{2} + \frac{\mathrm{erfc}(fz_i) - \mathrm{erfc}(fz_j)}{2}
\tag{28a}
$$

$$
\begin{aligned}
P_1(z_i, z_j) &= \frac{e^{-z_i^2}}{4\sqrt{a_{i,j}}}\left[\Omega\left(\frac{b_{i,j,-1}}{\sqrt{a_{i,j}}} \right) - \Omega\left(\frac{b_{i,j,1}}{\sqrt{a_{i,j}}} \right) \right] \\
&+ \frac{e^{-z_j^2}}{4\sqrt{a_{j,i}}}\left[\Omega\left(\frac{b_{j,i,-1}}{\sqrt{a_{j,i}}} \right) - \Omega\left(\frac{b_{j,i,1}}{\sqrt{a_{j,i}}} \right) \right]
\end{aligned}
\tag{28b}
$$

$$
\begin{aligned}
P_f(z_i, z_j) &= \frac{e^{-f^2 z_i^2}}{4\sqrt{a_{i,j}}}\left[\Omega\left(\frac{b_{i,j,-f}}{\sqrt{a_{i,j}}} \right) - \Omega\left(\frac{b_{i,j,f}}{\sqrt{a_{i,j}}} \right) \right] \\
&- \frac{e^{-f^2 z_j^2}}{4\sqrt{a_{j,i}}}\left[\Omega\left(\frac{b_{j,i,-f}}{\sqrt{a_{j,i}}} \right) - \Omega\left(\frac{b_{j,i,f}}{\sqrt{a_{j,i}}} \right) \right]
\end{aligned}
\tag{28c}
$$

2726

dx.doi.org/10.1021/ct500504g | J. Chem. Theory Comput. 2014, 10, 2719−2728

$$P_R(z_i, z_j) = \frac{1}{2\sqrt{a_{i,j}}}\left[e^{-z_i^2} W\left(\frac{b_{i,j,-1}}{\sqrt{a_{i,j}}}, \frac{\sqrt{a_{i,j}} z_i^3}{z_i^2 - z_j^2} \right) \right.$$

$$\left. + e^{-f^2 z_i^2} W\left(\frac{b_{i,j,-f}}{\sqrt{a_{i,j}}}, \frac{\sqrt{a_{i,j}} z_i^3}{z_i^2 - z_j^2} \right) \right]$$

$$+ \frac{1}{2\sqrt{a_{j,i}}}\left[e^{-z_j^2} W\left(\frac{b_{j,i,1}}{\sqrt{a_{j,i}}}, \frac{\sqrt{a_{j,i}} z_j^3}{z_i^2 - z_j^2} \right) \right.$$

$$\left. + e^{-f^2 z_j^2} W\left(\frac{b_{j,i,-f}}{\sqrt{a_{j,i}}}, \frac{\sqrt{a_{j,i}} z_j^3}{z_i^2 - z_j^2} \right) \right]$$

(28d)

Here, $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$ is the complementary error function, $\Omega(x) = \exp(x^2)\mathrm{erfc}(x)$, $W(x,y) = \exp(x^2)\mathrm{erfc}(x + y)$, $f = f_i + f_j$, and $a_{m,n} = 1 - c_2(z_m/z_n)^2$ and $b_{m,n,r} = rz_m - c_1(z_n/z_m)$ for $m \in \{i, j\}$ and $n \in \{i,j\}$ ($m \neq n$).

We compared the exact expression for P_a for $k_i^{eff} = k_j^{eff}$ ($z_i = z_j \equiv z$), eq 4 in ref 9, with the equal z limit of eq 27 obtained by taking limit $z_i \to z$ and $z_j \to z$,

$$P_a(z) = \lim_{z_i, z_j \to z} P_a(z_i, z_j) = \mathrm{erfc}(z)$$

$$+ \frac{\exp(-z^2)}{2\sqrt{1 - c_2}}\left[\Omega\left(\frac{-z - c_1}{\sqrt{1 - c_2}} \right) - \Omega\left(\frac{z - c_1}{\sqrt{1 - c_2}} \right) \right]$$

(29)

As shown in Figure 10A, eq 29 agrees excellently with the exact solution. The absolute error is less than 1.48% (at $z = 0.856$) for any $z \geq 0$ and the relative error is less than 8.94% (at $z = 2.459$) for $z \in [0,6]$, which verified the accuracy of eq 27. In Figure 10B, eq 27 as a function of z_1 and z_2, $P_a(z_1,z_2)$, is shown, in which the green area is the region of the interest ($P_a \sim 0.4$) for the optimal overlap between neighboring windows.

ASSOCIATED CONTENT

Ⓢ Supporting Information

S1: Data for the toy model from all the tested parameter sets. S2: Additional data for the GpA-TM assembly. This material is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Authors
*E-mail:wonpil@ku.edu.
*E-mail:spark94@ku.edu.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

ABBREVIATIONS

PMF, potential of mean force; CV, collective variable; MD, molecular dynamics; MC, Monte Carlo; US, umbrella sampling; REX, replica exchange; DOF, degrees of freedom; WEUS, window exchange US; aWEUSMD, adaptive WEUSMD

REFERENCES

(1) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

(2) Mezei, M. *J. Comput. Phys.* **1987**, *68*, 237−248.

(3) Marsili, S.; Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V. *J. Phys. Chem. B* **2006**, *110*, 14011−14013.

(4) Maragakis, P.; van der Vaart, A.; Karplus, M. *J. Phys. Chem. B* **2009**, *113*, 4664−4673.

(5) Wojtas-Niziurski, W.; Meng, Y.; Roux, B.; Berneche, S. *J. Chem. Theory Comput.* **2013**, *9*, 1885−1895.

(6) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(7) Vaneerden, J.; Briels, W. J.; Harkema, S.; Feil, D. *Chem. Phys. Lett.* **1989**, *164*, 370−376.

(8) Lee, J.; Im, W. *Phys. Rev. Lett.* **2008**, *100*, 018103.

(9) Park, S.; Kim, T.; Im, W. *Phys. Rev. Lett.* **2012**, *108*, 108102.

(10) Lingenheil, M.; Denschlag, R.; Mathias, G.; Tavan, P. *Chem. Phys. Lett.* **2009**, *478*, 80−84.

(11) Nadler, W.; Hansmann, U. H. E. *Phys. Rev. E* **2007**, *75*, 026109.

(12) Kofke, D. *J. Chem. Phys.* **2002**, *117*, 6911−6914.

(13) Predescu, C.; Predescu, M.; Ciobanu, C. *J. Phys. Chem. B* **2005**, *109*, 4189−4196.

(14) Dashti, D. S.; Roitberg, A. E. *J. Chem. Theory Comput.* **2013**, *9*, 4692−4699.

(15) Crooks, G. E. *Phys. Rev. Lett.* **2007**, *99*, 100602.

(16) Shenfeld, D. K.; Xu, H.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *Phys. Rev. E* **2009**, *80*, 046705.

(17) Weinhold, F. *J. Chem. Phys.* **1975**, *63*, 2479−2483.

(18) Salamon, P.; Berry, R. S. *Phys. Rev. Lett.* **1983**, *51*, 1127−1130.

(19) Salamon, P.; Nulton, J. *Europhys. Lett.* **1998**, *42*, 571−576.

(20) Schaller, M.; Hoffmann, K.; Siragusa, G.; Salamon, P.; Andresen, B. *Comput. Chem. Eng.* **2001**, *25*, 1537−1548.

(21) Nulton, J.; Salamon, P. *J. Non-Equilib. Thermodyn.* **2002**, *27*, 271−288.

(22) Bennett, C. *J. Comput. Phys.* **1976**, *22*, 245−268.

(23) Shirts, M.; Bair, E.; Hooker, G.; Pande, V. *Phys. Rev. Lett.* **2003**, *91*, 140601.

(24) Maragakis, P.; Spichty, M.; Karplus, M. *Phys. Rev. Lett.* **2006**, *96*, 100602.

(25) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(26) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.

(27) Darve, E.; Rodriguez-Gomez, D.; Pohorille, A. *J. Chem. Phys.* **2008**, *128*, 144120.

(28) Zheng, L.; Chen, M.; Yang, W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20227−20232.

(29) Zheng, L.; Chen, M.; Yang, W. *J. Chem. Phys.* **2009**, *130*, 234105.

(30) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.

(31) Kästner, J.; Thiel, W. *J. Chem. Phys.* **2005**, *123*, 144104.

(32) Hènin, J.; Pohorille, A.; Chipot, C. *J. Am. Chem. Soc.* **2005**, *127*, 8478−8484.

(33) Zhang, J.; Lazaridis, T. *Biophys. J.* **2006**, *91*, 1710−1723.

(34) Li, P.-C.; Miyashita, N.; Im, W.; Ishido, S.; Sugita, Y. *J. Comput. Chem.* **2014**, *35*, 300−308.

(35) Park, S.; Im, W. *J. Chem. Theory Comput.* **2013**, *9*, 13−17.

(36) Im, W.; Lee, J.; Kim, T.; Rui, H. *J. Comput. Chem.* **2009**, *30*, 1622−1633.

(37) Lee, J.; Im, W. *J. Comput. Chem.* **2007**, *28*, 669−680.

(38) MacKenzie, K.; Prestegard, J.; Engelman, D. *Science* **1997**, *276*, 131−133.

(39) Chothia, C.; Levitt, M.; Richardson, D. *J. Mol. Biol.* **1981**, *145*, 215−250.

(40) Lazaridis, T. *Proteins* **2003**, *52*, 176−192.

(41) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.;

2727

dx.doi.org/10.1021/ct500504g | *J. Chem. Theory Comput.* 2014, 10, 2719−2728

Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(42) Woodcock, H. L., III; Hodoscek, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F., III; Brooks, B. R. *J. Comput. Chem.* **2007**, *28*, 1485−1502.

(43) Kullback, S.; Leibler, R. A. *Ann. Math. Statist.* **1951**, *22*, 79−86.

(44) Trebst, S.; Huse, D. A.; Troyer, M. *Phys. Rev. E* **2004**, *70*, 046701.

(45) Escobedo, F. A.; Martinez-Veracoechea, F. J. *J. Chem. Phys.* **2007**, *127*, 174103.

(46) *Handbook of Mathematical Functioins With Formulas, Graphs, and Mathematical Tables*; Abromowitz, M., Stegun, I. A., Eds.; Dover:: New York, 1972.