JCTC Journal of Chemical Theory and Computation

Potential of Mean Force Calculations: A Multiple-Walker Adaptive Biasing Force Approach

K. Minoukadeh,*[,†,‡] C. Chipot,[§,||] and T. Lelièvre[†,‡]

CERMICS, École des Ponts ParisTech, 6−8 avenue Blaise-Pascal, 77455 Champs-sur-Marne, Marne-la-Vallée cedex 2, France, MICMAC Project-Team, INRIA Rocquencourt, 78153 Le Chesnay, France, Équipe de dynamique des assemblages membranaires, UMR 7565, Nancy Université, BP 239, 54506 Vandœuvre-lès-nancy Cedex, France, and Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801

Received October 5, 2009

Abstract: The adaptive biasing force (ABF) scheme is a powerful molecular-dynamics based method for overcoming barriers of the free-energy landscape. Integration of the mean force measured along a chosen reaction coordinate (RC) yields the so-called potential of mean force (PMF). The RC is a coarse-grained description of the transition mechanism. The mean force is estimated by accruing and averaging the instantaneous force exerted on the system. The PMF is then used to bias the standard dynamics of the system in order to improve sampling in the RC. We show that faster exploration of the reaction pathway can be achieved by running multiple walkers in parallel and exchanging information at fixed intervals in the course of the simulation. Numerical experiments performed on the prototypical deca-alanine peptide demonstrate that the convergence properties of the free-energy calculation are globally improved through a more efficient exploration of compact configurations reflected in parallel valleys of the free-energy landscape. Diffusion along the RC is further enhanced by a selection mechanism, whereby far-reaching walkers are cloned, replacing less effective ones.

1. Introduction

Central to the understanding of most processes of either physical, chemical, or biological interest, the determination of the underlying free-energy change occupies a prominent position in the arena of numerical simulations. Over the past decades, a variety of methods have been devised to compute free-energy differences efficiently (see, for example, refs 1 and 2). Roughly speaking, these methods can be classified into two main categories: (i) the free energy is computed directly, or (ii) its first derivative is determined and subsequently integrated. Perturbation techniques,[3] probability density function-based methods such as histogram methods,[4−6] nonequilibrium computations,[7] and adaptive biasing potential methods,[8,9] for example, fall into the first category. Thermodynamic integration[10] and adaptive biasing force methods,[11,12] which are the core of the present work, belong to the second category. Adaptive methods are designed to compute free-energy profiles and favor transitions between metastable states by using a current estimate of the free energy as a biasing potential.

In this contribution, we are interested in a particular class of adaptive methods, referred to as adaptive biasing force methods.[11−13] Specifically, we endeavor to investigate a novel implementation of this class of methods, using a number of walkers simulated in parallel, in the spirit of the ideas put forth by Lelièvre et al.[14] The advantage of the present, novel implementation is 3-fold. First, the parallelization is straightforward, and its theoretical parallel efficiency is very good since the only shared information is the biasing force, or the marginal law, namely, low-

* To whom correspondence should be addressed: E-mail: kimiya.minoukadeh@cermics.enpc.fr.
† École des Ponts ParisTech.
‡ INRIA Rocquencourt.
§ Nancy Université.
|| University of Illinois at Urbana−Champaign.

10.1021/ct900524t
Published on Web 03/04/2010

dimensional functions. In turn, this yields efficient, scalable algorithms to compute free-energy differences, well adapted to the massively parallel architecture of high-performance computers. Second, we will show that the implementation relying upon many walkers is particularly interesting when the reaction coordinate does not describe well all the metastabilities of the system, which, quite unfortunately, constitutes a generic situation for the vast majority of nontrivial molecular systems. This is typically the case, for instance, of the so-called bichannel scenario—namely, the free-energy landscape features two parallel valleys, which are orthogonal to the isocontours of the reaction coordinate—or, more generally, when several transition mechanisms are associated with a single reaction coordinate, which is, therefore, not sufficient to parametrize fully the transformation. The underlying idea is that, when many walkers are involved, they can visit more efficiently in parallel all the valleys in the direction of the reaction coordinate. A mathematical proof is currently underway to show that, in the limiting case of a very large number of walkers, and with suitable assumptions, the rate of convergence of the ABF method is in fact not limited by free-energy barriers orthogonal to the RC direction. Third, as will be detailed below, this new implementation allows selection mechanisms to be introduced, consisting of duplicating effective walkers, while deleting poor ones, according to a fitness function that ought to be chosen. An example of such a fitness function, which favors rapid exploration of the reaction coordinate, will be provided hereafter.

As a proof of concept, the present approach was probed on a realistic test case, using a high-level Tcl implementation of the algorithm in the scalable molecular dynamics program NAMD.[15−17] The efficiency of the overall procedure is, however, expected to be enhanced by embedding and optimizing the algorithm at a deeper level of the molecular-dynamics platform.

In the following section, the mathematical framework of the method is introduced and the adaptive biasing force method reviewed. Next, the discretization and implementation details are presented. The present contribution closes with a discussion of the numerical results obtained for the reversible folding of the paradigmatic deca-alanine peptide.

2. General Setting

In the canonical ensemble, a system of dimension d is equipped with the Boltzmann−Gibbs probability measure, i.e., the canonical measure:

$$\mu(dq) = \phi(q)dq = Z^{-1} \exp(-\beta V(q))\, \mathrm{d}q \qquad (1)$$

where ϕ is the density of the measure, $\beta = 1/(k_B T)$ is proportional to the inverse temperature, $q \in \mathbb{R}^d$ is the system configuration, $V : \mathbb{R}^d \to \mathbb{R}$ is the potential energy function, and $Z = \int_{\mathbb{R}^d} \exp(-\beta V(q))\, \mathrm{d}q$ is the normalization constant or the so-called partition function. To sample this measure, one can use the overdamped Langevin dynamics:[18]

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}\, \mathrm{d}W_t \qquad (2)$$

where $(X_t)_{t\geq 0}$ is the system trajectory and W_t is an $\mathbb{R}^d$-valued standard Brownian motion (or Wiener process). Under suitable regularity assumptions on the potential, the dynamics (eq 2) is ergodic and admits the canonical measure as its unique invariant measure. It must be emphasized that, for the sake of simplicity, the method is described in the framework of the overdamped dynamics. The method can, nevertheless, be generalized to the Langevin dynamics as is done in the numerical simulations at the end of the paper.

The canonical measure (eq 1) gives us microscopic information about the system, the probability that it is to be found at any particular point q in configuration space. A practitioner, however, is generally interested in some coarse-grained collective variable $\xi(q)$, where ξ is typically a smooth mapping from $\mathbb{R}^d$ to $\mathbb{R}$. In what follows, ξ will be referred to as the reaction coordinate (RC). The RC typically represents an end-to-end distance of a protein chain, a structural angle in a protein, or a measure of the evolution of a chemical system. If X is a random variable with probability μ, then $\xi(X)$ is a random variable with a law whose density ϕ^ξ is defined by

$$\phi^\xi(z) = \int_{\mathbb{R}^d} \phi(q)\, \delta(\xi(q) - z)\, \mathrm{d}q \qquad (3)$$

and the distribution $\phi^\xi(z)dz$ is called the marginal distribution of μ in ξ. The free energy, or so-called potential of mean force (PMF), A, is related to this marginal density in the following way

$$A(z) = -\beta^{-1} \ln \phi^\xi(z) \qquad (4)$$

Sampling the canonical measure using the standard overdamped Langevin dynamics (eq 2) can in fact be inefficient in practice. The convergence to equilibrium can be very slow due to metastable states where the dynamics remains trapped for long periods of time. To explore the whole configuration space, one often needs to overcome very large energy barriers. From Arrhenius's law, it follows that the typical time needed to overcome these barriers scales exponentially with the barrier height. Regular molecular-dynamics methods are, therefore, typically not used to calculate statistical averages for systems prone to metastabilities.

Several methods have been proposed to ameliorate sampling methods in these situations, such as the *blue moon* method[19] or importance sampling methods such as *umbrella sampling*.[20] More recently, adaptive importance sampling methods have been developed such as the *Wang−Landau* method[8] and the *adaptive biasing force* (ABF) method.[11] The latter and its variations will be the focus of this contribution.

Before detailing the ABF method, the reader is reminded that the main quantity of interest in the study of chemical reactions is a free-energy difference and not an absolute free energy. Free energies are, therefore, computed only up to an additive constant. The free-energy difference between two coarse-grained states, labeled by the RC values z_a and z_b, can be written as

$$\Delta A = A(z_b) - A(z_a) = \int_{z_a}^{z_b} A'(z)\,\mathrm{d}z \tag{5}$$

where $'$ is the derivative with respect to the collective variable value z and A' is called the mean force. The integrand can be shown to be the Boltzmann average of a real-valued function F^V, conditioned to being at a fixed point z in the reaction coordinate space

$$\begin{aligned} A'(z) &= \frac{\int_{\mathbb{R}^d} F^V(q)\exp(-\beta V(q))\,\delta(\xi(q) - z)\,\mathrm{d}q}{\int_{\mathbb{R}^d}\exp(-\beta V(q))\,\delta(\xi(q) - z)\,\mathrm{d}q} \\ &= \langle F^V(q)|\xi(q) = z\rangle_\mu \end{aligned} \tag{6}$$

where

$$F^V = \frac{\nabla V\cdot\nabla\xi}{|\nabla\xi|^2} - \beta^{-1}\nabla\cdot\left(\frac{\nabla\xi}{|\nabla\xi|^2}\right) \tag{7}$$

and $\langle\cdot\rangle_\mu$ represents the canonical average—i.e., the average with respect to the measure μ. Note that F^V is the negative projection of the force onto the RC plus some correction term. For the derivation of eq 7, the reader is referred to refs 21, 22, and 23. The aim of the ABF method, which will be detailed hereafter, is to compute A' as efficiently as possible.

3. Adaptive Biasing Force Methods

In this section, we will present the framework behind ABF methods for free energy computations.

3.1. Framework. The basic idea of ABF is to use the mean force estimate to bias the dynamics and help the system overcome free-energy barriers. An estimate of $A'(z)$ is obtained as the statistical average of the force field F^V at specified points z along the RC by accruing instantaneous forces $F^V(X_t)$ for a single system trajectory X_t when $\xi(X_t) = z$. In the long-time limit, one obtains a good approximation for A', and ΔA can be computed by numerical integration. The resulting biased dynamics is

$$\begin{cases} dX_t = -\nabla(V - A_t\circ\xi)(X_t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_t \\ A_t'(z) = \langle F^V(X_t)|\xi(X_t) = z\rangle \end{cases} \tag{8}$$

where $A_t\circ\xi$ denotes the composition of A_t with ξ, so that $A_t\circ\xi(x) = A_t(\xi(x))$, and A_t' is the estimated mean force. [Note that the gradient term in the biased dynamics can be rewritten as $-\nabla V + (A_t'\circ\xi)\nabla\xi$; thus, only estimated mean force information is needed and not the estimated free energy.] The estimated mean force is thus defined as a conditional average of $F^V(X_t)$ at a fixed value of $\xi(X_t)$. In practice, it can be approximated as an average over many walkers or as an *on-the-fly* average over the trajectory X_t (see the next section for more details). The above can be viewed as an overdamped Langevin dynamics, with the potential V replaced by the time varying potential $\mathscr{V}_t' = V - A_t\circ\xi$. In the following, ψ_t will denote the density of the law of X_t. The consistency of the method may be justified by noticing that, if a stationary state A_∞' for A_t' is obtained, then ψ_∞ is proportional to $\exp(-\beta\mathscr{V}_\infty')$, and thus $A_\infty' = A'$ since

$$\begin{aligned} A_\infty'(z) &= \langle F^V(q)|\xi(q) = z\rangle_{\exp(-\beta\mathscr{V}_\infty')} \\ &= \langle F^V(q)|\xi(q) = z\rangle_\mu = A'(z) \end{aligned}$$

As a result, A_t converges to A up to an additive constant, and the equilibrium marginal density in ξ is constant, since

$$\langle\delta(\xi(q) - z)\rangle_{\exp(-\beta\mathscr{V}_\infty')} = \frac{\langle\delta(\xi(q) - z)\rangle_\mu}{\exp(-\beta A(z))}$$

is constant by the definitions of μ and A in eqs 1 and 4, respectively. Precise convergence results can be found in ref 24. The aim of ABF is, therefore, to estimate the biasing force as efficiently as possible in order to bias the dynamics by reducing and eventually eliminating any force along ξ. It serves as an adaptive importance sampling method, driving the system out of its metastable states, using on-the-fly estimates of the mean force.

3.2. Calculating the Bias. Different approaches have been proposed in recent literature[11,24,25] for calculating the biasing force. There are two principal methods for computing A_t', which will serve as a basis of comparison in the present contribution.

Original ABF. The idea of the standard ABF method,[11] which involves one single walker, is to calculate averages using the whole trajectory of the system. The mean force is calculated by taking a trajectorial average of instantaneous forces at fixed z using one long system trajectory (see ref 1 for further details)

$$\langle F^V(X_t)|\xi(X_t) = z\rangle \simeq \frac{\int_0^t F^V(X_s)\,\delta(\xi(X_s) - z)\,\mathrm{d}s}{\int_0^t \delta(\xi(X_s) - z)\,\mathrm{d}s} \tag{9}$$

The mean force estimate is only computed once a trajectory reaches the value z in the RC; therefore, the denominator in the above equation is always nonzero for the RC values needed.

Multiple Walker (MW-)ABF. In a recent paper,[25] a new formulation of the ABF method has been proposed, consisting of running $R > 1$ trajectories of the ABF dynamics in parallel. The R walkers of the system follow a similar dynamics driven by independent Brownian motions. These multiple walkers then exchange information at fixed time intervals. The immediate gain of this new formulation is that one can take advantage of parallel computing to speed up convergence of the ABF method. Furthermore, with the use of a small number of walkers, we are able to overcome issues related to poorly chosen or oversimplified reaction coordinates, where other important slow degrees of freedom are overlooked. In such cases, metastabilities can be found at fixed ξ, as illustrated in Figure 1. With multiple walkers, it is likely that each walker will explore a different path or valley along ξ, whereas single-walker simulations could potentially take exponentially long times to fully explore the low energy states. This will be studied numerically in the final sections of the paper.

In the following, $(X_t^i)_{0\le i\le R-1}$ is the set of trajectories for the R walkers. Each trajectory X_t^i follows the dynamics (eq 8) with the Brownian motion W_t replaced with W_t^i. The

biasing force is then estimated using an average over all the trajectories and over all walkers:

$$\langle F^V(X_t)|\xi(X_t) = z\rangle \simeq \frac{\sum_{i=0}^{R-1}\int_0^t F^V(X_s^i)\,\delta(\xi(X_s^i) - z)\,\mathrm{d}s}{\sum_{i=0}^{R-1}\int_0^t \delta(\xi(X_s^i) - z)\,\mathrm{d}s} \tag{10}$$

Implementation details for the approaches discussed above will be discussed in the next section.

3.3. Enhancing Sampling through Selection. In addition to the exchange of information between walkers to compute the estimated mean force A_t', the MW-ABF method allows for resampling of the walkers according to their "importance". The success of the ABF method is strongly determined by the marginal distribution of walkers in the RC, given that the RC has been well chosen. One would, therefore, want to encourage walkers that are exploring undersampled regions of the RC and penalize those in oversampled regions. A selection mechanism[25−27] may be used to achieve this objective. It is implemented by a system of interacting walkers, where the walkers are cloned or killed at a rate defined by $S(t, z)$ over the values taken by the RC. The function $S(t, z)$ can be chosen as

$$S(t, z) = c\frac{\partial_{zz}\psi_t^\xi(z)}{\psi_t^\xi(z)} \tag{11}$$

where c is a positive constant and ψ_t^ξ, defined by

$$\psi_t^\xi(z) = \int_{\mathbb{R}^d}\psi_t(q)\,\delta(\xi(q) - z)\,\mathrm{d}q$$

represents the marginal distribution of walkers in the RC at time t. Here, $\psi_t(q)$ denotes the distribution of X_t. With this

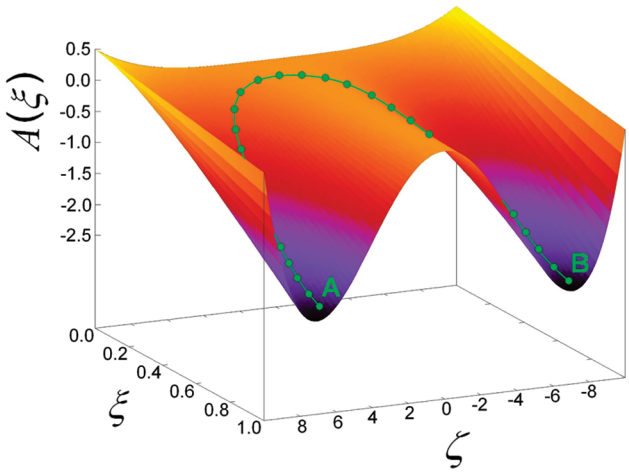

Figure 1. Example of a 2-dimensional free-energy surface exhibiting metastabilities at fixed ξ. The variable ζ represents another slow degree of freedom of the system, orthogonal to the RC. The standard ABF method relies on efficient sampling at fixed points in the RC, which is made difficult by the presence of such large energy barriers in the orthogonal directions. Using multiple walkers helps to overcome this issue as each one is very likely to explore a different pathway.

choice of the function S, it can be shown that the marginal density ψ^ξ satisfies the partial differential equation (in fact for a slightly modified version of the original adaptive dynamics (eq 8), see ref 25):

$$\partial_t\psi_t^\xi = (\beta^{-1} + c)\partial_{zz}\psi_t^\xi \tag{12}$$

The selection process thus accelerates the diffusion of the marginal distribution in the RC. The reason for this choice of S can also be understood when written in a finite difference form

$$S(t, z) \simeq \frac{3c}{\Delta z^2\psi_t^\xi(z)}\left[\frac{\psi_t^\xi(z - \Delta z) + \psi_t^\xi(z) + \psi_t^\xi(z + \Delta z)}{3} - \psi_t^\xi(z)\right] \tag{13}$$

where Δz is some small displacement in the RC. The quantity $S(t, z)$ at a given value z of the RC is, therefore, positive if the marginal density at this point is small compared to its local average, and negative otherwise. To implement the selection process, one can either continuously update birth and death times, initially drawn from an exponential distribution, as in ref 25, or resample the walkers according to their weights at fixed resampling times. The latter will be used for the simulations reported herein. At a resampling time t, each walker trajectory X_t^i is given a weight:[28]

$$w_t^i = K_t^{-1}\exp\left[\int_0^t S(s, \xi\{X_s^i\})\,\mathrm{d}s\right] \tag{14}$$

where $K_t = \sum_{i=0}^{R-1}\exp[\int_0^t S(s, \xi\{X_s^i\})\,\mathrm{d}s]$ is the normalization constant. Replicas are initially assigned a uniform weight, $w_0^i = 1/R$, which evolves in time. From eqs 13 and 14, it is now clear that a walker i that is often found in undersampled regions—in which case $S(t, \xi(X_t^i))$ is often positive—is given a stronger weight than walkers in oversampled regions—where $S(t, \xi(X_t^i))$ is often negative. The ith walker is then replicated on average Rw_t^i times. This procedure thus accelerates the convergence to a uniform distribution of the walkers in the RC, in accordance with eq 12.

Let us now give some details about the resampling procedure. To calculate the number of times a walker is to be copied, a systematic resampling method[29−31] is used, described briefly by the following algorithm. At a resampling time t:

Set $u \sim U(0, 1)$, $\bar{w}_0 = w_t^0$, $N_0 = \lfloor R \times \bar{w}_0 + u\rfloor$,
for $i = 1, ..., R - 1$
$\bar{w}_i = \bar{w}_{i-1} + w_t^i$,
$N_i = \lfloor R \times \bar{w}_i + u\rfloor - \lfloor R \times \bar{w}_{i-1} + u\rfloor$
end

where $U(0, 1)$ denotes a uniform distribution between 0 and 1, w_t^i is the normalized weight assigned to walker i as defined in eq 14, $\bar{w}_i$ is the cumulative sum of the weights, $\lfloor\cdot\rfloor$ is the integer part, and N_i is the number of copies of walker i to be generated. It is important to note that this algorithm guarantees that $\sum_{i=0}^{R-1}N_i = R$. After every resampling stage, the weights of all walkers are reset uniformly to the value $1/R$. The choice of the constant c in eq 11 is of paramount importance in the performance of the selection mechanism. This parameter should

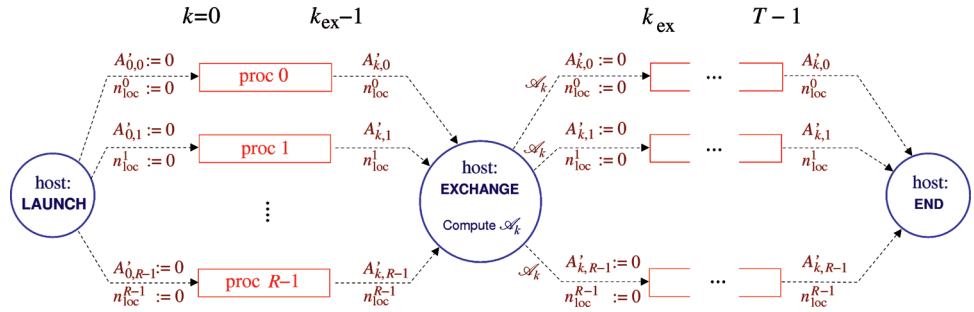

Figure 2. Schematic diagram of MW-ABF. The main script is executed on a host machine, which acts as the Tcl server. This machine launches the R walkers onto different processors via socket connections and, after every k_{ex} time steps, carries out exchange of information. This consists of reading in local variables from each processor; computing the total biasing force A_k' by means of eqs 21 and 22, sharing $\mathscr{A}_k'$ with all processors, and setting local variables to zero. This is carried out T/k_{ex} times until the program terminates.

be sufficiently large to accelerate the exploration along the RC, but not too large in case one walker is selected during the resampling stage (due to degeneracy of weights), which implies a very large variance of the estimator. This will be discussed further at the end of the next section.

4. Implementation Details

In this section, the implementation details of the adaptive biasing force methods are provided. The simulations reported in the present contribution have been carried out using the scalable molecular-dynamics code NAMD, but the algorithmic detail is by no means specific to this software package. The ABF methods have been implemented as Tcl scripts, for which the single-walker ABF method is already available. How the method is discretized will be spelled out hereafter, and the detail of the single-walker ABF method will be outlined before proceeding with the implementation of the MW-ABF method and selection.

We consider a reaction coordinate ξ taking values in the interval $[z_0, z_N]$, which is divided into N bins of size $\Delta z = (z_N - z_0)/N$. We denote by $\tilde{\xi}:\mathbb{R}^d \to \{0, ..., N-1\}$ a mapping from a configuration onto its associated bin in the RC

$$\tilde{\xi}(\cdot) = \left\lfloor \frac{\xi(\cdot) - z_0}{\Delta z} \right\rfloor$$

where $\lfloor \cdot \rfloor$ again denotes the integer part. In the following, functions and trajectories will be indexed by the number of time steps k, so that A_k' will be the mean force approximation and X_k will be the configuration of the system at time $k\Delta t$, for a time step Δt. Furthermore, with a slight abuse of notation, z will now denote the bin in the reaction coordinate, $z = \tilde{\xi}(X_t)$.

Original ABF Method. The reader is reminded that, in the standard ABF method, the biasing force is calculated for each bin using a trajectorial average, as in eq 9. The biasing force is in practice updated to include the current force observation. For $z \in \{0, ..., N-1\}$

$$A_k'(z) = \frac{n_{tot}(k-1, z)}{n_{tot}(k, z)} A_{k-1}'(z) + \frac{\mathbf{1}_{\tilde{\xi}(X_{k-1})=z}}{n_{tot}(k, z)} F^V(X_{k-1}) \tag{15}$$

where $\mathbf{1}_{\tilde{\xi}(X_k)=z}$ denotes the indicator function—taking value 1 if $\tilde{\xi}(X_k) = z$ and 0 otherwise—and

$$n_{tot}(k, z) = \sum_{s=0}^{k-1} \mathbf{1}_{\tilde{\xi}(X_s)=z} \tag{16}$$

is the total number of times the system trajectory has visited bin z. To justify eq 15, the expression in eq 9 is recast in its discretized form

$$A_k'(z) = \frac{\sum_{s=0}^{k-1} F^V(X_s) \mathbf{1}_{\tilde{\xi}(X_s)=z}}{\sum_{s=0}^{k-1} \mathbf{1}_{\tilde{\xi}(X_s)=z}} \tag{17}$$

Developing further, one subsequently obtains

$$\begin{aligned} A_k'(z) &= \frac{\sum_{s=0}^{k-2} F^V(X_s) \mathbf{1}_{\tilde{\xi}(X_s)=z} + F^V(X_{k-1}) \mathbf{1}_{\tilde{\xi}(X_{k-1})=z}}{n_{tot}(k, z)} \\ &= \frac{n_{tot}(k-1, z) A_{k-1}'(Z) + F^V(X_{k-1}) \mathbf{1}_{\tilde{\xi}(X_{k-1})=z}}{n_{tot}(k, z)} \end{aligned}$$

where the last line follows from eq 17 at time $k-1$.

MW-ABF. The basis for the multiple-walker implementation of ABF in NAMD can be found in a set of Tcl scripts written for parallel-tempering, replica-exchange simulations.[15,17] The scripts use Tcl server and socket connections to launch NAMD processes for each individual walker. Each walker is handled by a different computing unit. The NAMD processes run for a fixed number of time steps, then wait in order for the Tcl server to exchange information between walkers. Figure 2 is a synoptic diagram of the MW-ABF method. It is not necessary (and not desirable from a computational point of view) to exchange information at every time step. Exchange of information between walkers only occurs at every k_{ex} time steps (Figure 3). We therefore proceed as follows: the mean force approximation, denoted by $A_{k,i}'(\cdot)$, is evaluated locally on the computing unit, where the indices k and i represent respectively the number of time steps since the beginning of the simulation and the

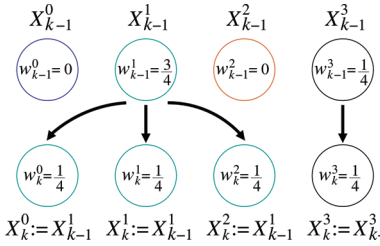

Figure 3. Selection mechanism for $R = 4$ walkers. If walker i has weight w^i_k at the time of selection, on average, Rw^i_k copies are made of this walker at the next step. In practice, this means that Rw^i_k walkers will be launched using the configuration and velocity files of walker i. Note that, in the above, $k = nk_{ex}$, the time at which selection is carried out.

computing unit running walker i. This quantity therefore depends solely on the trajectory of the walker of interest. Between exchange times, $k \in [nk_{ex},\ (n+1)k_{ex}]$, the mean force estimation evolves according to the update formula

$$A'_{k,i}(z) = \frac{n^i_{loc}(k-1,z)}{n^i_{loc}(k,z)} A'_{k-1,i}(z) + \frac{\mathbf{1}_{\tilde{\xi}(X^i_{k-1})=z}}{n^i_{loc}(k,z)} F^V(X^i_{k-1}) \tag{18}$$

Note that this is the same as eq 15, where A'_k is replaced by $A'_{k,i}$, X_k is replaced by X^i_k, and $n_{tot}(k,z)$ is replaced by $n^i_{loc}(k,z)$, the number of times walker i has entered bin z since the last exchange, defined in eq 19 below.

At every exchange time, the information gathered by each walker is collected and local variables on each processor are updated. This is formalized with the help of some further notation. We denote by

$$k_{last} = \lfloor k/k_{ex} \rfloor k_{ex}$$

the time of the last exchange. Next,

$$n^i_{loc}(k,z) = \sum_{s=k_{last}}^{k-1} \mathbf{1}_{\tilde{\xi}(X^i_s)=z} \tag{19}$$

denotes the number of times walker i has entered bin z since the last exchange, and

$$n^i_{tot}(k,z) = \sum_{s=0}^{k-1} \mathbf{1}_{\tilde{\xi}(X^i_s)=z} \tag{20}$$

is the total number of times walker i has entered bin z since the beginning of the simulation. Finally,

$$N_{loc}(k,z) = \sum_{i=0}^{R-1} n^i_{loc}(k,z) \text{ and } N_{tot}(k,z) = \sum_{i=0}^{R-1} n^i_{tot}(k,z)$$

denote respectively the total number of visits to bin z since the last exchange and the beginning of the simulation, over all the walkers.

At every exchange time, a local average A'_{loc} is calculated of the mean force estimated from the run of each individual walker:

$$A'_{loc}(k,z) = \frac{1}{N_{loc}(k,z)} \sum_{i=0}^{R-1} n^i_{loc}(k,z)\, A'_{k,i}(z) \tag{21}$$

The total biasing force, $\mathscr{A}'_k(z)$, to be shared between the walkers, is then also updated to include this new information

$$\mathscr{A}'_k(z) = \left[1 - \frac{N_{loc}(k,z)}{N_{tot}(k,z)}\right] \mathscr{A}'_{k-1}(z) + \frac{N_{loc}(k,z)}{N_{tot}(k,z)} A'_{loc}(k,z) \tag{22}$$

The latter quantity, $\mathscr{A}'_k$, is then communicated to each one of the walkers, and the variables n^i_{loc} and $A'_{k,i}$ are reset to zero.

The total biasing force in eq 22 is utilized by each walker in the steps following the exchange; however, new local information is also incorporated in order to speed up the diffusion in ξ. The biasing force applied to walker i in the simulations, therefore, writes

$$F^i_{bias,k}(z) = \left[1 - \frac{n^i_{loc}(k,z)}{N_{tot}(k,z)}\right] \mathscr{A}'_{k_{last}}(z) + \frac{n^i_{loc}(k,z)}{N_{tot}(k,z)} A'_{k,i}(z)$$

where $\mathscr{A}'_{k_{last}}$ is again the total mean force calculated at the preceding exchange interval, and $A'_{k,i}$ is the local mean force information, as defined in eq 18. [In practice, this force is actually only fully applied to the dynamics after a certain number of visits have been made to the bin, that is to say, after $N_{tot}(k,z) > N_{min}$, where in our simulations $N_{min} = 500$. If $N_{tot}(k,z) < N_{min}/2$, then no biasing force is added. Beyond that, the force is slowly introduced using a ramp function with scaling factor $\min(2N_{tot}/N_{min} - 1, 1)$.]

Selection. Resampling may be carried out at most every k_{ex} time steps, when the walkers exchange information. Selection is a technically costly process as NAMD must be exited and reloaded with new configuration and velocity files. For this reason, it is even advisible for it to be carried out less frequently. The computational complexity of the process is $O(R)$ using a systematic resampling method (see the previous section for the algorithm). For purposes of illustration, the resampling will be carried out as often as the interprocessor communication, namely, every k_{ex} time steps. The purpose of resampling is to improve the exploration in the RC. The weights of the walkers are adjusted according to the utility function $S(k,z)$, depending in practice upon the total distribution of the walkers:

$$S(k,z) = c\, \frac{N_{tot}(k_{last}, z+1) - 2N_{tot}(k_{last}, z) + N_{tot}(k_{last}, z-1)}{N_{tot}(k_{last}, z)} \tag{23}$$

The integral in eq 14 is calculated by summing the terms $S(k, \tilde{\xi}(X^i_k))$ over k for each walker i during each individual run. At the selection stage, when $k = k_{ex}$, the weights of the walkers are computed:

$$w^i_k = K_k^{-1} \exp\left[\sum_{s=k_{last}}^{k-1} S(s, \tilde{\xi}\{X^i_s\})\right]$$

where K_k is again the normalization constant. The walkers are then selected according to these weights using a

systematic resampling method, as described above. In practice, to generate N_i copies of walker i, the configuration and velocity files for walker i are passed to NAMD as the startup files for N_i walkers. Finally, after resampling, S is set to zero, so that all walker trajectories have equal weight.

For resampling to be effective, there are two main issues that need to be addressed. First, the constant c has to be chosen carefully: it must be large enough for the selection mechanism to be beneficial and small enough to avoid degeneration of weights, where all walkers are given zero weight except for one. Another issue to be addressed is when to stop resampling. Due to the technical costs of the selection mechanism, it is advised to impose a stopping criterion, so that selection is not applied throughout the whole simulation. The stopping criterion could depend on the sampling of the RC, or on the distribution of the weights. For the simulations herein, the latter criterion is used, for once the walkers begin to be equally weighted, the selection has no effect and ought to be stopped. In order to measure the distribution of the weights, we consider the relative entropy of the weights compared to a uniform distribution, defined by

$$E_w(t) = \sum_{i=0}^{R-1} w_i \log(Rw_i) \tag{24}$$

This can be understood as the difference between the entropy of weights and the entropy for the uniform weight distribution: $\sum_{i=0}^{R-1} w_i \log(w_i) - \log(1/R)$. This quantity is bounded above by $\log(R)$, in case of degeneracy, and is bounded below by 0, in case of uniform distribution of weights. A good stopping criterion for the selection algorithm is to end the process when the relative entropy is sufficiently small. In our simulations, the selection is stopped once

$$E_w(t) < \varepsilon \log(R) \tag{25}$$

where $0 < \varepsilon < 1$ is set closer to 0 for a stringent stopping criterion or closer to 1 for a weaker threshold.

5. Numerical Results

In this section, we present comparisons of the single-walker and multiple-walker ABF methods on the deca-alanine peptide in the gas phase, for which comprehensive studies have already been carried out.[12,32,33] All the simulations reported herein were performed with the molecular-dynamics code NAMD,[15−17] using the CHARMM27[34] force field. The 10-residue peptide chain has a total of 104 atoms and the RC has been chosen as the distance separating the center of mass of its first and last C−H pairs. To sample the full range of conformations from the α-helical conformation to the ensemble of extended structures, the range of values accessible to ξ varies from 12 to 32 Å. Additional tests are also carried out to study more compact conformations, where ξ varies between 4 and 16 Å. The system is kept within the assigned ranges by enforcing reflective boundary conditions.

The average forces were accumulated in bins of size $\Delta z = 0.1$ Å. The equations of motion were integrated employing Langevin dynamics with a time step $\Delta t = 0.5$ fs.

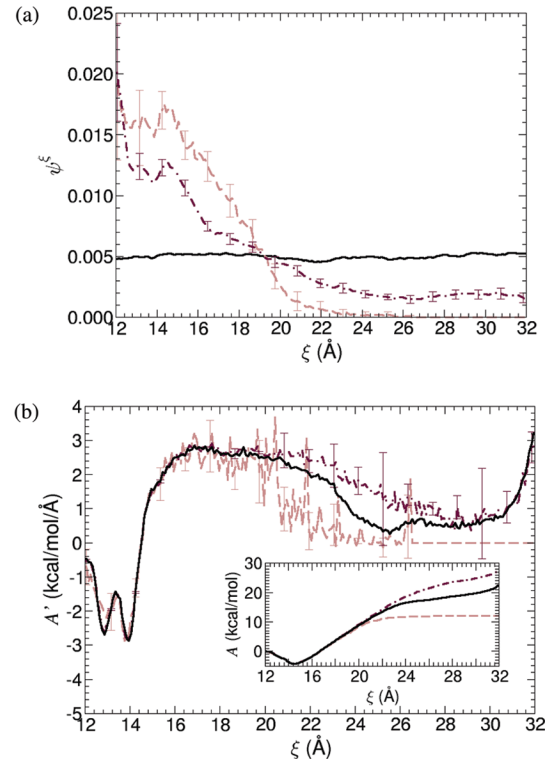

Figure 4. Results for ξ ranging from 12 to 32 Å (after 0.25 ns). The curves are averages of 20 independent single-walker (dashed lines) and 16-walker (dashed-dotted lines) simulations with error bars representing the 95% confidence intervals. Solid lines represent reference profiles, obtained from a single-walker run of 200 ns. (a) Density of marginal distribution in the RC. The multiple-walker simulation has explored the whole ξ-space whereas the single-walker simulations very rarely stretch beyond 22 Å. (b) Mean force and free-energy profiles (inset). For the multiple-walker simulations, we see the mean force profiles already nearly converged, whereas little information is gathered beyond 22 Å for the standard ABF simulation.

Electrostatic and van der Waals interactions were truncated smoothly beyond 11 Å.

We will first present the results for the conventional range of 12 to 32 Å, which spans conformations comprised between the α-helix to more extended structures. Next, results for more compact conformations—with ξ ranging from 4 to 16 Å—are presented, where stark differences can be observed between the single- and multiple-walker ABF simulations. Finally, we will study the impact of selection on walkers.

5.1. Reaction Coordinate Range: 12−32 Å. Starting from the α-helical conformation, R walkers of the system are launched with ABF dynamics, communicating at every $k_{ex} = 50\,000$ time steps (25 ps). Reference curves are obtained from a 200-ns simulation using the original ABF algorithm, featuring a single walker.

Figure 4 compares the sampling distribution, mean force, and free-energy profiles for single-walker and 16-walker simulations after 0.25 ns.

It may be observed from Figure 4a that the single-walker runs rarely manage to stretch beyond a distance of $\xi = 22$ Å, whereas the 16-walker simulations explore the whole reaction-coordinate space. Furthermore, in Figure 4b,

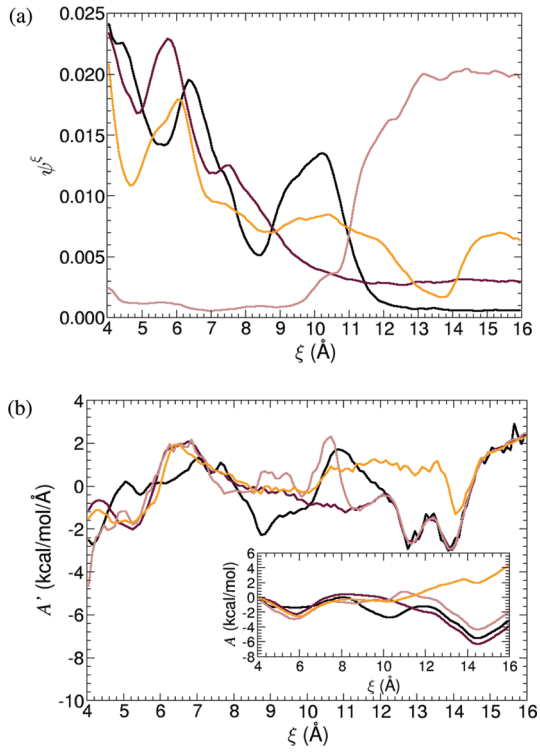

Figure 5. Results for ξ ranging from 4 to 16 Å using 1 walker (after 100 ns). Results are from four independent simulations. (a) Sampling along the RC. (b) Mean force approximations and free-energy profiles (inset): large discrepancies are observed, suggesting the presence of parallel valleys along ξ. Note that one of the free-energy profiles suggests a global minimum at $\xi = 6$ Å.

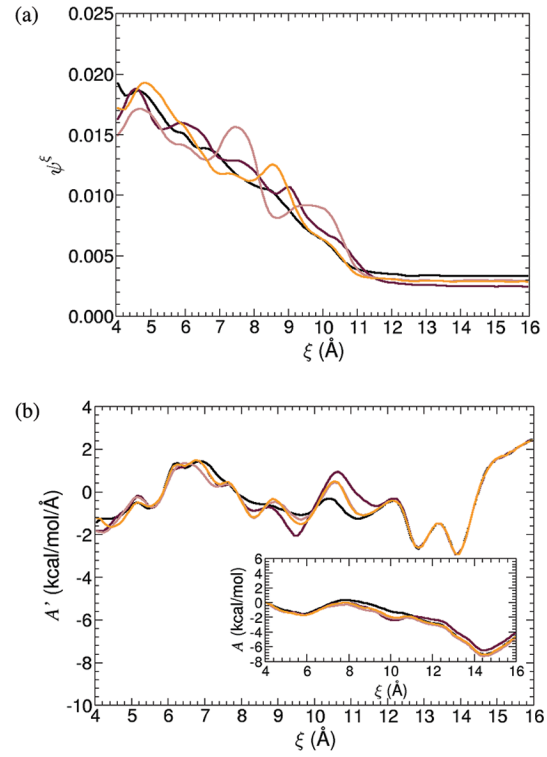

Figure 6. Results for ξ ranging from 4 to 16 Å using 32 walkers (results after 100 ns). Results are from four independent simulations. (a) Sampling along the RC. (b) Mean force approximations and free-energy profiles (inset). Sampling and mean force estimations are consistent with each other, and the α-helical conformation is recovered as the global free-energy minimum.

it is apparent that the mean force and free-energy profiles obtained by the 16-walker simulations are already qualitatively consistent with the reference curves.

5.2. Reaction Coordinate Range: 4−16 Å. As previously mentioned, convergence of the standard ABF method can be rather slow in the presence of metastabilities on the submanifold of conformations at a fixed value of ξ. This is generally the result of a poor choice of the RC, which does not capture all metastabilities of the system. In such a case—as depicted in Figure 1—several low energy conformations could be associated to a fixed value of ξ and separated by high-energy barriers. As highlighted in ref 32, this shortcoming arises when studying compact conformations of the deca-alanine peptide. In this article, an extension of the standard sampling window reveals a free-energy profile that exhibits a wide global minimum ranging from 4 to 12 Å. It is known, however, that the global minimum of the deca-alanine peptide is the α-helical conformation at about $\xi = 14$ Å (see refs 12 and 33). The present results can be explained by the fact that, in compact states, a great number of low-energy conformations are associated to a value of ξ of the RC, which are not fully explored by a standard, single-walker ABF simulation due to their separation by high free-energy barriers. These high free-energy barriers are generally insurmountable from the perspective of conventional MD simulations and can be viewed as kinetic traps that preclude the exploration of the full RC space over reasonable time scales. A recent study has helped to capture the various slow degrees of freedom for these compact structures by exploring multidimensional free-energy landscapes.[13]

The shortcomings discussed above can be advantageously circumvented using multiple walkers. The results obtained from 100-ns single- and multiple-walker simulations of the compact conformations are compared in Figures 5 and 6, respectively. Figure 5b depicts mean force estimations for four independent single-walker simulations. Even after a 100-ns simulation, large discrepancies are observed between the mean force profiles. As can be observed in the inset of Figure 5b, one free-energy profile has revealed a global minimum around $\xi = 6$ Å, which is, in most likelihood, artifactual. Figure 6 summarizes the results obtained from four independent 32-walker simulations. A marked improvement in the convergence of the mean force profiles is immediately apparent. This supports the speculation that there exist parallel valleys along the RC, each of a different nature, separated by high free-energy barriers. The present set of results is far more promising with a multiple-walker scheme.

Due to eventual traps in the parallel valleys, it is in fact likely that a T-nanosecond single-walker simulation will be less efficient than an R-walker simulation run for T/R nanoseconds. This argument is supported numerically by Figure 7, showing results for a 32-walker simulation after $100/32 \sim 3$ ns. The results are qualitatively consistent with Figure 6 and offer a far more reliable set of results than a 100 ns single-walker simulation. In this way, a

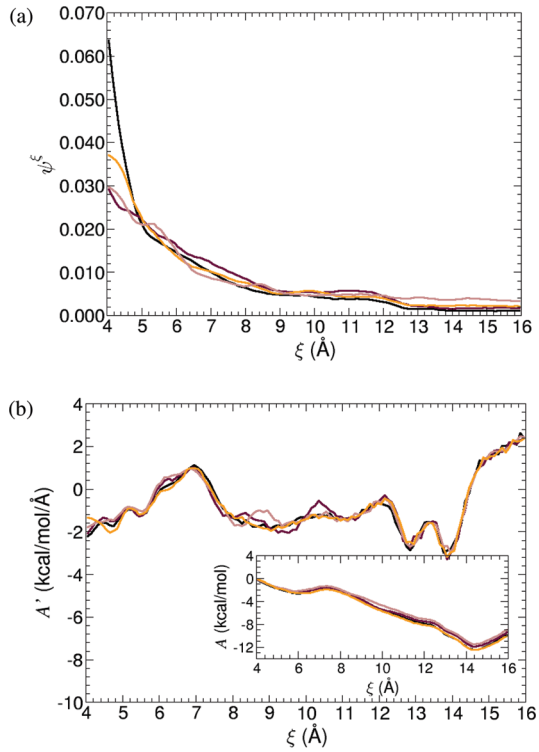

Figure 7. Results after 3 ns using 32 walkers. To compare results at constant total CPU time, we observe the results of a 32-walker simulation after 100/32 ∼ 3 ns. Results are from four independent simulations. (a) Sampling along the RC. (b) Mean force estimates and free-energy profiles (inset) are qualitatively very close to Figure 6b. The results show that a multiple-walker simulation can outperform a single-walker simulation at constant CPU time.

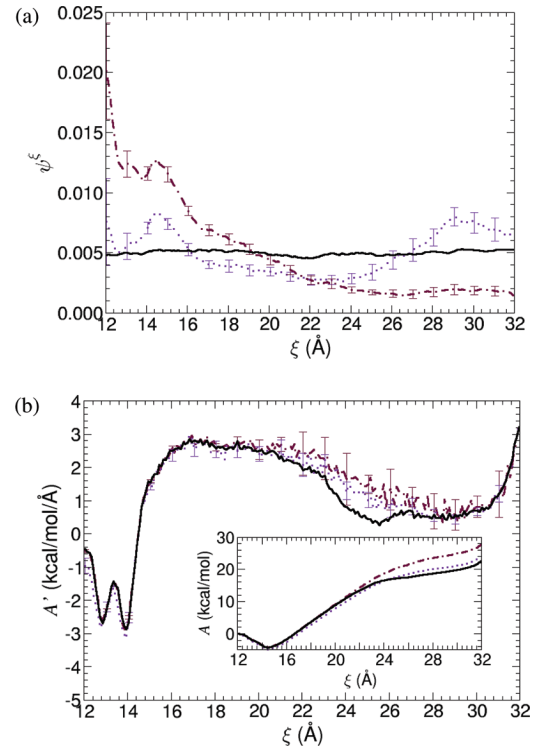

Figure 8. Results for ξ ranging from 12 to 32 Å using 16 walkers (results after 0.25 ns). Comparing results between a 16-walker run with (dotted lines) and without selection (dashed-dotted lines). The curves represent averages of 20 independent ABF simulations, and the error bars are 95% confidence intervals. Reference curves are shown as solid lines. (a) The sampling along ξ shows that simulations *with* selection provide a much more uniform distribution along the RC. (b) Mean force approximations and free energy difference profiles (inset): the free-energy profile for the simulation *with* selection is already very close to the reference curve.

multiple-walker implementation not only improves results at a constant wall time but at a *constant CPU time* as well.

5.3. Selection. In order to monitor the impact of the selection mechanism, we have chosen to study again the standard range of 12−32 Å. This choice is dictated by the topology of the free-energy landscape. It ought to be recalled here that the selection criterion is based solely on the position of the walkers along the RC. In the case of multiple, parallel valleys, this criterion can, therefore, impede convergence of the mean force, should many copies of one walker be generated in the same valley. It appears that selection is most effective in the presence of metastabilities in the RC only. Metastabilities in the orthogonal directions are well explored by multiple-walker simulations, albeit not always improved by selection.

Figures 8a and 8b compare the sampling and free-energy profiles determined by a 16-walker simulation after 0.25 ns. We observe a more uniform sampling and a better potential of mean force for the simulation *with* a selection mechanism. For these simulations, we used the selection constant $c = 0.0001$ in eq 23 and a stopping criterion as in eq 25 with $\varepsilon = 0.05$.

Figure 9 depicts $E_w(t)$, the relative entropy of the walker weights, during the first 2.5 ns of the simulation. It may be observed that $E_w(t)$ decreases during the ABF simulation,

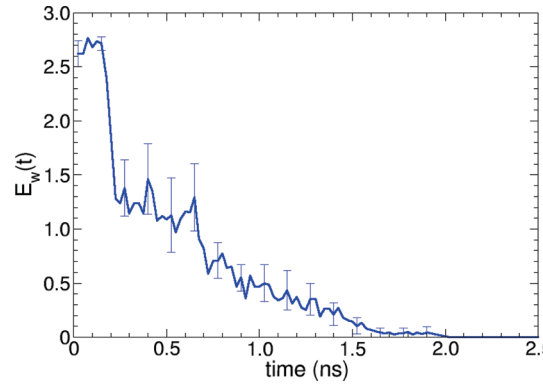

Figure 9. Relative entropy of weights. It can be seen that the $R = 16$ walkers are approximately of equal weight after about 1.5 ns of an ABF simulation. The selection is switched off after $E_w(t) < \varepsilon \log(16)$, where $\varepsilon = 0.05$. For these simulations, the selection constant in eq 23 is chosen as $c = 0.0001$.

as the biasing force converges. This result suggests that the walkers are then more free to move along the RC, and thus, each walker is of equal "importance". Once the walkers are more or less of equal weight, the selection becomes redundant and is, therefore, switched off, avoiding unnecessary computational effort.

6. Discussion

In the present contribution, we have demonstrated the applicability of the MW-ABF method to a prototypical biomolecular system. Importance sampling techniques are often held back by the difficulty of choosing good reaction coordinates. If the RC is chosen poorly, one is likely to encounter parallel valleys separated by large free-energy barriers, thereby making sampling at a fixed point along the RC very difficult. In such an event, a standard single-walker ABF simulation would lead to slow convergence, as was shown here. The system is biased only in the direction of the RC and, therefore, would be likely to linger in one valley for a long time before reaching another. We have shown that such shortcomings can be elegantly overcome using multiple walkers, through the proposed MW-ABF method. We emphasize that the use of multiple walkers is particularly beneficial when the choice of the model RC is suboptimal, where improvement has been demonstrated herein at constant CPU time. For a well chosen RC, the MW-ABF is not guaranteed to outperform single-walker ABF simulations at a fixed total CPU cost but still has the advantage of being easily parallelized. The selection process introduced herein can be employed profitably when encountering pronounced free-energy barriers along the RC. In the presence of parallel valleys, attention must be paid to avoid degeneration of weights, as this could lead to many walkers being kinetically trapped in the same valley, losing the main interest of the use of multiple walkers.

Acknowledgment. This work is supported in part by the Agence Nationale de la Recherche, under grants LN3M (ANR-05-CIGC-0003) and MEGAS (ANR-09-BLAN-0216-01). The authors are grateful to Gabriel Stoltz and Jérome Hénin for fruitful discussions and to the reviewers for many helpful comments.

References

(1) Chipot, C.; Pohorille, A. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Springer: New York, 2007.

(2) Lelièvre, T.; Rousset, M.; Stoltz, G. *Free Energy Computations: A Mathematical Perspective*; Imperial College Press, to appear.

(3) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(4) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

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

(6) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *124*, 124105.

(7) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.

(8) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.

(9) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302.

(10) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(11) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9196–9183.

(12) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904–2914.

(13) Hénin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. *J. Chem. Theory Comput.* **2010**, *6*, 35–47.

(14) Lelièvre, T.; Rousset, M.; Stoltz, G. *J. Comput. Phys.* **2007**, *222*, 624–643.

(15) Bhandarkar, M. *NAMD*; Theoretical Biophysics Group, University of Illinois and Becman Institute: Urbana, IL, 2003.

(16) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J. C.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comp. Phys.* **1999**, *151*, 283–312.

(17) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(18) Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II: Nonequilibrium Statistical Mechanics*, 2nd ed.; Springer: New York, 1991.

(19) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Lett.* **1989**, *156*, 472–477.

(20) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: New York, 1996.

(21) Ciccotti, G.; Lelièvre, T.; Vanden-Eijnden, E. *Comm. Pure Appl. Math.* **2008**, *61*, 3.

(22) Sprik, M.; Cicotti, G. *J. Chem. Phys.* **1998**, *109*, 7737–7744.

(23) den Otter, W. K.; Briels, W. J. *J. Chem. Phys.* **1998**, *109*, 4139.

(24) Lelièvre, T.; Rousset, M.; Stoltz, G. *Nonlinearity* **2008**, *21*, 1155–1181.

(25) Lelièvre, T.; Rousset, M.; Stoltz, G. *J. Chem. Phys.* **2007**, *126*, 134111.

(26) Assaraf, R.; Caffarel, M.; Khelif, A. *Phys. Rev. E* **2000**, *61*, 4566.

(27) Doucet, A.; de Freitas, N.; Gordon, N. J. *Sequential Monte Carlo Methods in Practice*; Springer: New York, 2001; Series Statistics for Engineering and Information Science.

(28) Del Moral, P. *Feynman-Kac Formulae Genealogical and Interacting Particle Systems with Applications*; Springer: New York, 2004.

(29) Douc, R.; Cappe, O.; Moulines, E. *Image and Signal Processing and Analysis, 2005. ISPA 2005* **2005**, 64–69.

(30) Kitagawa, G. *J. Comput. Graph. Stat.* **1996**, *5*, 1–25.

(31) Carpenter, J.; Clifford, P.; Fearnhead, P. *Technical Report, Department of Statistics*; University of Oxford: Oxford, U.K., 1999.

(32) Chipot, C.; Hénin, J. *J. Chem. Phys.* **2005**, *123*, 244906.

(33) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559.

(34) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

CT900524T