CHAPTER 1

Free-Energy Calculations with Metadynamics: Theory and Practice

Giovanni Bussi,[a] and Davide Branduardi[b]

[a]*Statistical and Molecular Biophysics Group, International School for Advanced Studies (SISSA), Trieste, IT 34136, Italy*
[b]*Theoretical Molecular Biophysics Group, Max Planck Institute of Biophysics, Frankfurt am Main D-60438, Germany*

INTRODUCTION

Molecular dynamics (MD) is a powerful tool in modern chemistry that allows one to describe the time evolution of a computational model for a complex molecular system.[1–3] Typical models range from being highly accurate where energy and forces are computed with advanced and expensive quantum chemistry methods to faster but less accurate empirically parameterized force fields at atomistic or coarser resolution. The power of these techniques lies in their ability to reproduce experimental observable quantities accurately while, at the same time, giving access to the mechanistic details of chemical reactions or conformational changes at very high spatial resolution – typically at atomistic scale. For this reason, MD is often used to complement experimental investigations and to help in interpreting experiments and in designing new ones. Moreover, thanks to new parallelization algorithms and to the continuous improvements in computer hardware driven by Moore's law, the range of application of these techniques has grown exponentially in the past decades and can be expected to continue growing.

Reviews in Computational Chemistry, Volume 28, First Edition.
Edited by Abby L. Parrill and Kenny B. Lipkowitz.

In spite of its success, however, MD is still limited to the study of events on a very short timescale. Indeed, depending on the required accuracy and on the available computational resources, MD can provide trajectories for events happening on the timescale of picoseconds (quantum chemistry) to microseconds (empirical force fields). Thus, many interesting phenomena, namely, chemical reactions, protein folding and aggregation, and macromolecular rearrangement are still out of reach of direct investigation using straightforward MD trajectories. Besides the optimization of computer software (e.g., Ref. 4) and/or hardware (e.g. Refs. 5, 6), it is a possible complementary strategy to alleviate this issue by using algorithms where the time evolution is modified to sample more frequently the event under investigation. Then, appropriate postprocessing techniques are necessary to recover unbiased properties from the accelerated trajectories.

Many algorithms to accelerate MD simulations have been designed in the past decades, and a discussion of all of them is out of the scope of this chapter. Some of these algorithms are based on increasing the temperature of the simulated system (e.g., parallel tempering[7] and solute tempering[8]), while others are based on exploiting an *a priori* knowledge of the investigated transition to design a proper order parameter to both describe and accelerate it. This last class includes umbrella sampling,[9] adaptive biasing force,[10] metadynamics,[11] self-healing umbrella sampling,[12] and other methods that keep the selected order parameters at an artificially high temperature.[13–15] This chapter focuses on metadynamics, which was first introduced in 2002[11] and then improved with several variants in the past decade. Metadynamics has been employed successfully in several fields, ranging from chemical reactions[16] to protein folding[17] and aggregation,[18] molecular docking,[19] crystal structure prediction,[20] and nucleation.[21] A further push in the diffusion of metadynamics application has been its availability in a few widespread molecular dynamics codes[22–24] and in open-source plugins.[25–27]

The main goal of this chapter is to provide an entry-level tutorial for metadynamics. In Section "Molecular Dynamics and Free-Energy Estimation" we provide an introduction to the basic concepts of molecular dynamics and of free-energy calculations. In Section "A Toy Model: Alanine Dipeptide" we introduce a toy model that will then be used for subsequent examples. Section "Biased Sampling" is devoted to the introduction of biased sampling. In Sections "Adaptive Biasing with Metadynamics" and "Well-Tempered Metadynamics" metadynamics is introduced, and Section "Metadynamics How-To" provides a practical how-to for performing a free-energy calculation with metadynamics. For all the simulations described in that section a sample input file for the open-source package PLUMED 2[26] is given in the Appendix. In the remaining sections, a quick overview of some of the latest improvements in the field is given, followed by a concluding section.

MOLECULAR DYNAMICS AND FREE-ENERGY ESTIMATION

Molecular Dynamics

In classical MD,[1−3] the Hamilton equations of motion are solved numerically to follow in real time the propagation of a collection of atoms. For a system of N_{at} atoms with coordinates q_i, momenta p_i, and masses m_i, a potential energy $U(q)$ should be defined. Notice that we use here q, without subscript, meaning the full $3N_{\text{at}}$-dimensional vector containing all the atomic positions. The Hamilton equations of motion will then read

$$\dot{q}_i = \frac{p_i}{m_i} \tag{1a}$$

$$\dot{p}_i = -\frac{\partial U(q)}{\partial q_i} \tag{1b}$$

Here with $\dot{x}$ we mean the derivative with respect to the time of the variable x. The potential energy function $U(q)$ describes the interatomic interactions. These interactions are sometimes defined in terms of empirically parameterized force fields, which provide a cheap and reasonably accurate approximation for $U(q)$, and sometimes the interactions are obtained by solving the Schrödinger equation for the electrons (*ab initio* calculations), to allow studying phenomena such as electron transfer and chemical reactions. In our examples we will only use empirical potentials. However, the specific definition of U is totally irrelevant for what concerns the discussed methodologies, which often rely on *ab initio* calculations.

For a system evolving according to the Hamilton equations [1a] and [1b], the total energy $H(p,q) = \sum_i \frac{p_i^2}{2m_i} + U(q)$ is conserved, so that only configurations that have a total energy exactly equal to the initial one are explored. This will provide the correct properties for an isolated system. On the contrary, whenever a system is coupled to an external bath, the transfer of energy between the system and the bath implies that different values of the total energy are accessible. More precisely, the phase space point (p,q) will be explored with a probability $P(p,q)$, which, in the case of a thermal bath, corresponds to the canonical ensemble:

$$P(p,q)dpdq \propto \exp\left(-\frac{\sum_i \frac{p_i^2}{2m_i} + U(q)}{k_B T}\right) dqdp \tag{2}$$

where k_B is the Boltzmann constant and T is the temperature of the thermal bath. Within MD, this is typically done by adding a so-called "thermostat" to the Hamilton equations.[2] Strictly speaking, a thermostat alters the dynamical properties, so that the latter could lose their physical meaning. This in turn depends a lot on the details of the adopted thermostat. Nonetheless, irrespective of the thermostat, MD can always be used to generate configurations according to the canonical distribution.

A crucial aspect of using MD for sampling the canonical distribution is the so-called *ergodic hypothesis*: if a system is simulated long enough all the states pertaining to the canonical ensemble will be explored, each with its own correct statistical weight. Unfortunately, this hypothesis cannot be proven for most of the systems and, even when verified, the length of a simulation necessary for this hypothesis to be exploited in calculating ensemble averages is often far out of reach for numerical simulations. This has profound consequences and led in the past decades to the development of several enhanced sampling algorithms that were designed to alleviate this difficulty.

Free-Energy Landscapes

The canonical distribution of a system sampled via MD carries full information about its thermodynamic properties. However, this probability distribution is of very little use. Indeed, the space on which it is defined (i.e., the set of all possible positions and velocities of all the atoms in a system) is huge – it is a $6N_{\text{at}}$ dimensional space – so that this function is completely unintelligible. For this reason, molecular systems are often analyzed in terms of collective variables (CVs) rather than atomic coordinates. A CV is a function of the atomic coordinates that is capable of describing the physics behind the process under investigation. As an example, for an isomerization process a reasonable CV could be a torsional angle, whereas for a proton transfer two reasonable CVs could be the distances between the hydrogen and each of the two involved heavy atoms. Because the CVs are functions of the atomic coordinates, we shall indicate them as $s_\alpha(q)$, with $\alpha = 1, \ldots, N_{\text{CV}}$ and let N_{CV} be equal to the number of CVs used. In short, the CVs represent a sort of coarse description of the system, which can be used to analyze a given process in low dimensionality. A basic requirement for a CV is being able to distinguish all the peculiar states of interest without lumping together states that are very different physicochemically.

When analyzing a molecular system using CVs $s(q)$, a role equivalent to that of the potential energy is played by the Helmholtz free energy $F(s)$. $F(s)$ is defined in such a manner that the probability of observing a given value of s is

$$P(s)ds \propto \exp\left(-\frac{F(s)}{k_B T}\right) ds \qquad [3]$$

It can be shown that the relationship between $U(q)$ and $F(s)$ can be written explicitly as

$$F(s_0) = -k_B T \log \int dq \delta(s(q) - s_0) \exp\left(-\frac{U(q)}{k_B T}\right) + C \quad [4]$$

where the Dirac δ selects all the microscopic configurations corresponding to a specific value s_0 of the CV and C is an arbitrary constant.

A typical free-energy landscape for an activated event is shown in Figure 1. Here one can appreciate a stable state denoted by "A" (low free energy, thus high probability), a metastable one denoted by "B" (slightly higher free energy, thus lower probability), and an intermediate region (very high free energy, thus very small probability). A proper definition of metastability is out of the scope of this chapter. The height of the free-energy barrier compared with the value of the thermal energy $k_B T$ affects the probability of observing the system in the intermediate region and thus the typical time required to go from "A" to "B" or vice versa. When CVs are appropriately chosen, the transition rate between minima "A" and "B" can be estimated according to transition state theory by an Arrhenius-like formula:

$$\nu_{A\to B} = \nu_0 \exp^{-\frac{\Delta F^{\ddagger}}{k_B T}} \quad [5]$$

where $\Delta F^{\ddagger}$ is the free-energy difference between the starting minimum and the transition state and ν_0 is a prefactor. Notably, the ratio between forward ($\nu_{A\to B}$) and backward ($\nu_{B\to A}$) rates is equal to $e^{\frac{F_A - F_B}{k_B T}}$, according to detailed balance. Thus, when appropriate CVs are used, the free-energy landscape provides a quantitative picture of the transition in terms of reactants and products stability and transition rates.

Evaluating the free-energy landscape defined by Eq. [4] is usually a daunting task, as it would require the calculation of a multidimensional integral in $3N_{at}$ dimensions. For this reason, the typical approach employed to compute

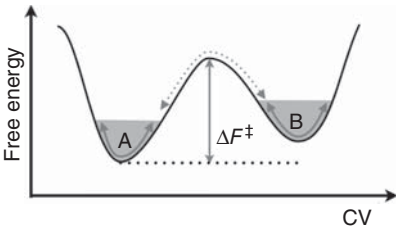

Figure 1 A model double-well potential. The stable state is denoted by "A" and the metastable one is denoted by "B." The gray shaded regions are those where the system fluctuates because of its thermal energy. The region in between is very unlikely to be explored, therefore making the transition from "A" to "B" less probable to occur.

free-energy landscapes is based on conformational sampling by means of MD or Monte Carlo (MC) methods. Indeed, if one is capable of producing a sequence of conformations (i.e., a trajectory) that are distributed according to the canonical ensemble, the free energy can be computed from the histogram of the visited conformations as

$$F(s) = -k_B T \log N(s) \qquad [6]$$

where $N(s)$ counts how many times the value s of the CVs has been explored. Because typical CVs are noninteger numbers, the histogram $N(s)$ is typically accumulated using a binning procedure. As discussed earlier, MD simulations allow one to produce such distribution via the ergodic hypothesis although, as we will show in the following section, this might be problematic even for very simple cases.

A TOY MODEL: ALANINE DIPEPTIDE

In this section we introduce a simple toy model that presents all the features of a typical real-life molecular system but still is small enough to allow the reader to readily test the concepts by means of inexpensive molecular dynamics simulations. This system is alanine dipeptide (ACE-ALA-NME), which is a small peptide in nonzwitterionic form (see Figure 2).

We describe interatomic interactions using the functional form $U(q)$ defined by the CHARMM 27 force field,[29] a standard force field available in many biomolecular-oriented molecular dynamics codes. For the sake of simplicity, all the simulations are carried out in vacuum as the free-energy landscape in such conditions presents several interesting features and has been studied previously using many different free-energy methods (see, e.g., Refs. 30–35) thus being a perfect test bed for the discussion.

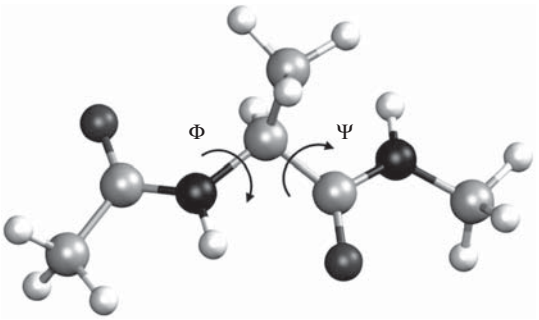

Figure 2 Molecular representation of alanine dipeptide. The two Ramachandran dihedral angles are denoted with Φ and Ψ. All the molecular representations are produced with VMD.[28]

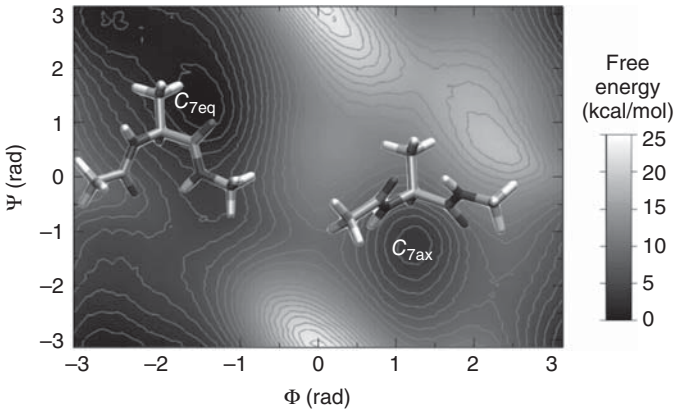

Figure 3 Representation of the free-energy landscape for alanine dipeptide as a function of the two Ramachandran dihedral angles Φ and Ψ. Each isoline accounts for 1 kcal/mol difference in free energy. The two main minima, namely, C_{7eq} and C_{7ax}, are labeled.

Alanine dipeptide in vacuum presents a peculiar free-energy landscape that can be rationalized in terms of Ramachandran dihedral angles. It displays two main minima: C_{7eq} and C_{7ax} (see Figure 3), placed around $\Phi \simeq -1.41$ rad, $\Psi \simeq 1.25$ rad and $\Phi \simeq 1.26$ rad, $\Psi \simeq -1.27$ rad, respectively. These two basins are separated by barriers around 8 kcal/mol that are remarkably higher than thermal energy at 300 K ($k_B T = 0.597$ kcal/mol), therefore presenting a typical case of metastability. In such a small system the differences in potential energy landscape are comparable with free-energy differences. Therefore, the Arrhenius equation can be used as a model to roughly estimate the rate of transition from one state to the other

$$\nu = \nu_0 \exp^{-\frac{\Delta U^{\ddagger}}{k_B T}} \qquad [7]$$

where $\Delta U^{\ddagger}$ is the potential energy difference between one metastable state, say C_{7eq}, and one of the transition states toward C_{7ax}. By assuming that the prefactor ν_0, as an upper bound, corresponds to the carbon–carbon vibration frequency in the force field we get $\nu_0 = 5 \times 10^9$/s. By using a barrier of 8 kcal/mol and $k_B T = 0.597$ kcal/mol we obtain a rate of 7×10^3 events per second, thus each barrier crossing could take about a millisecond. Unfortunately, such timescales are inaccessible to MD simulations that nowadays can reach several microseconds at most. Therefore, it is easy to understand why the problem of acquiring statistics of events that allow us to estimate the relative probability of metastable states is one of the grand challenges in computer simulations. For this reason some different techniques, also called *enhanced sampling* methods, have been devised through the years. Some of them will be the subject of the next sections.

BIASED SAMPLING

We have seen in Section "A Toy Model: Alanine Dipeptide" that even for a very small and simple molecule some transitions could be hindered by free-energy barriers. If the system is going to sample only minimum "A" or minimum "B" in Figure 1 and no transitions are observed, then Eq. [6] is completely useless to evaluate the free-energy difference between "A" and "B." Because the low transition rate is due to the fact that a nonlikely region (the barrier) is in between the two minima, it is intuitive that increasing the probability of sampling the barrier could alleviate this problem. Furthermore, because the barrier height affects the transition rate in an exponential manner (Eq. [5]), it is clear that changing the barrier could easily lead to a dramatic change in the observed rates.

As first step, it is important to analyze what happens if an additional potential $V(s)$, which acts only on the CVs s, is added to the physical one $U(q)$. The resulting potential will be $U(q) + V(s(q))$, so that the explored conformations will be distributed according to a biased canonical distribution

$$P'(p,q) \propto \exp\left(-\frac{\sum_i \frac{p_i^2}{2m_i} + U(q) + V(s(q))}{k_B T}\right) \qquad [8]$$

If one tries to evaluate the free-energy landscape from such a biased distribution of conformations, one will end up in a different free energy F', which is related to the original one by

$$\begin{aligned} F'(s_0) &= -k_B T \log \int dq \delta(s(q) - s_0) \exp\left(-\frac{U(q) + V(s(q))}{k_B T}\right) + C' \\ &= V(s_0) + F(s_0) + C' - C \end{aligned} \qquad [9]$$

where the part relative to the momenta has been integrated out and C' is another arbitrary constant. The correct free-energy landscape can then be recovered from the biased one up to an arbitrary constant by simply subtracting the bias potential $V(s)$.

Now, by supposing that the free-energy landscape is *a priori* known, at least approximately, it is possible to imagine performing a biased simulation, using $V(s) = -\widetilde{F}(s)$ as a bias where $\widetilde{F}(s)$ is our estimate for the free-energy landscape. The resulting simulation will explore the distribution,

$$P'(s) \propto \exp\left(-\frac{F(s) + V(s)}{k_B T}\right) = \exp\left(-\frac{F(s) - \widetilde{F}(s)}{k_B T}\right) \qquad [10]$$

which is an almost flat distribution. More precisely, this distribution is more and more flat if the estimate of the free-energy landscape is more and more accurate. In the ideal case, the free-energy barrier disappears because all the values of s are equally likely. This method is known as the *umbrella sampling* method.[9]

The efficiency of umbrella sampling depends on the accuracy of the initial guess of the free-energy landscape. An error of even a few k_BT could provide incorrect statistics because it would be unlikely to overcome the residual free-energy barriers. As we will see in the next section, this problem can be solved using an iterative procedure to build the bias potential.

ADAPTIVE BIASING WITH METADYNAMICS

Several algorithms have been proposed in the literature to progressively build a bias potential suitable for an umbrella sampling calculation.[10,36,37] We focus here on metadynamics.[11,38]

Metadynamics was originally introduced as a coarse dynamics in collective variable space, much in the spirit of Ref. 39. This dynamics was then accelerated by adding a penalty to the already visited states, similarly to the taboo search method[40] and the local elevation method.[41] Later, a version based on an extended Lagrangian formalism was introduced.[16] Presently, the most adopted variant is the continuous one.[38] In this chapter, we will only focus on this latter version.

In metadynamics (MetaD), a history-dependent bias potential is built during the simulation as a sum of repulsive Gaussians in the CV space. This is illustrated in Figure 4. These Gaussians are centered on the explored points in the CV space, have a preassigned width (σ) and height (w_G), and are deposited every τ_G time units as the simulation proceeds. The bias potential at time t thus reads

$$V(s,t) = \sum_{i=1}^{t/\tau_G} w_G \ \exp\left[-\sum_{\alpha=1}^{N_{CV}} \frac{\left(s_\alpha - s_\alpha\left(i\tau_G\right)\right)^2}{2\sigma_\alpha^2}\right] \qquad [11]$$

Notice that the width needs to be fixed for each of the CVs (σ_α). These widths determine the binning on each CV, that is, they can be used to state which distance in CV space should be considered as negligible. Intuitively, Gaussians are repeatedly added to the potential according to the explored states, such that they discourage the system from visiting again already visited configurations in the CV space. Notice that this procedure is done in the (low dimensional) CV space. Trying to do so in the full configurational space could not work in practice, because in that representation the system will practically never explore twice the same point. Contrarily, in the CV space, the system has a reasonable chance to explore several times the same value of s, albeit at different microscopic configurations q.

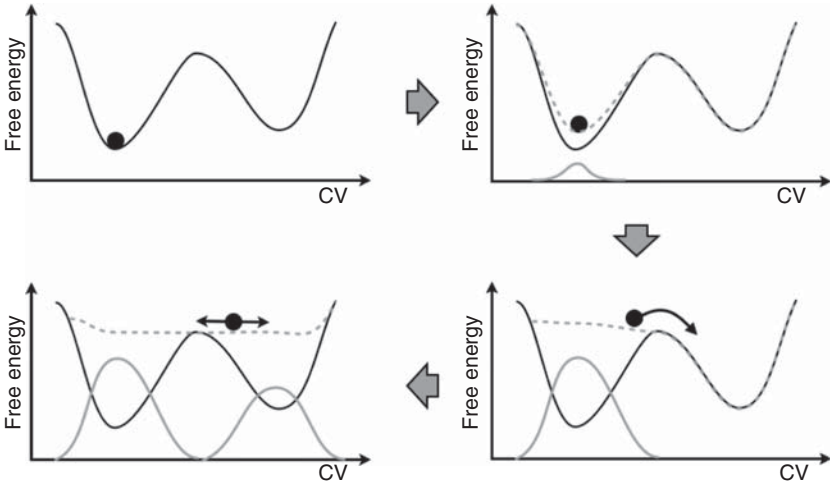

Figure 4 A sketch of the process of metadynamics. First the system evolves according to a normal dynamics, then a Gaussian potential is deposited (solid gray line). This lifts the system and modifies the free-energy landscape (dashed gray line) in which the dynamics evolves. After a while the sum of Gaussian potentials fills up the first metastable state and the system moves into the second metastable basin. After this the second metastable basin is filled, at this point, the system evolves in a flat landscape. The summation of the deposited bias (solid gray profile) provides a first rough negative estimate of the free-energy profile.

Several parameters should be chosen in a MetaD simulation and we will discuss how this is done in Section "Metadynamics How-To". Here we just observe that the w_G and τ_G parameter in Eq. [11] are not independent, and that, if they are chosen within meaningful ranges, what really matters is their rate $\omega = w_G/\tau_G$, also known as the *deposition rate*. In the limit of small τ_G, the expression for the history-dependent bias becomes Eq. [12].

$$V(s,t) = \omega \int_0^t dt' \exp\left[-\sum_{\alpha=1}^{N_{CV}} \frac{\left(s_\alpha - s_\alpha\left(t'\right)\right)^2}{2\sigma_\alpha^2}\right] \qquad [12]$$

Equivalently, it is possible to state that the bias potential $V(s)$ evolves with time according to the following equation of motion.

$$\dot{V}(s,t) = \omega \exp\left[-\sum_{\alpha=1}^{N_{CV}} \frac{\left(s_\alpha - s_\alpha\left(t\right)\right)^2}{2\sigma_\alpha^2}\right] \qquad [13]$$

The combination of this equation of motion with the Hamilton equations [1] describes completely the time evolution of a MetaD simulation. Another important observation is that the Hamiltonian dynamics is perturbed only through

the force term produced by $V(s,t)$. Thus, thinking that very wide Gaussians could compensate for the barrier in a faster way is incorrect as these may produce smaller forces.

It can be understood intuitively from Figure 4 that such a procedure not only discourages the exploration of the already visited states in the CV space, but also provides an immediate estimate for the underlying free-energy surface. As more and more Gaussians are added to the bias, the system will explore a larger and larger fraction of the CV space. Because Gaussians are most likely added at points where the effective total free energy $F(s) + V(s)$ is lower, their effect will tend to flatten the $F(s) + V(s)$ function. After a suitable "filling" time, the bias will start growing parallel to itself, and one can expect to directly estimate $F(s)$ as $-V(s)$, but for an additional arbitrary constant. This hand-waving conclusion has been tested accurately on realistic systems.[38] It can be supported by a rigorous analytical derivation that is based on assuming the CVs are evolving by means of a stochastic Langevin dynamics:[42]

$$\dot{s} = -\frac{D}{k_B T}\frac{\partial F}{\partial s} + \sqrt{2D}\xi \qquad [14]$$

where D is the diffusion coefficient of the CV and ξ a random noise. Within this framework, the error made in using $-V(s)$ as an estimator for $F(s)$ can be quantified from the function

$$\varepsilon(s) = F(s) + V(s) - \frac{1}{\Omega}\int ds[F(s) + V(s)] \qquad [15]$$

where the integral is taken over the CV space and Ω is a normalization factor equal to the total volume of the integration region. Recalling that $F(s)$ is defined up to an arbitrary constant, the integral in Eq. [15] allows this arbitrary constant to be removed. For Langevin dynamics, it can be shown that $\varepsilon(s)$ is not exactly zero, but its difference from zero fluctuates in time. The fluctuations have squared amplitude $\langle \varepsilon^2 \rangle \propto \omega/D$. This leads to two important observations:

- The statistical error grows with the deposition rate. Thus, a compromise should be found: ω should be large enough to quickly fill the basins in the free-energy landscape, but should also be small enough to limit the error in the free-energy estimation.
- The statistical error is larger for slower CVs (i.e., smaller D) because for such variables Gaussians are deposited on top of one another thus acting like an effective higher ω. This influences the compromise discussed above and implies that for slower descriptors one must choose a smaller deposition rate.

Because the free-energy estimate can be shown to be free from systematic errors, at least for model systems, it is always possible to increase its accuracy

by taking a time average of the result, as first proposed by Micheletti et al.[43] It is thus in principle sufficient to prolong a simulation enough to decrease the error above any desired threshold. This, however, can lead to another problem of MetaD: because MetaD is a flat histogram method it tries to sample the whole CV space. Consequently, the simulated system can be pushed toward states with nonphysically high free energy and might drift the simulation toward thermodynamically nonrelevant configurations. Additional restraining potentials on the CVs may alleviate this effect as discussed, for example, in Ref. 44. Nevertheless, because restraining potentials can add small artifacts in the free-energy estimation at the border, a better strategy was introduced in Ref. 45 and is based on the idea of neglecting forces due to the Gaussian potential when the CV is outside the desired range.

Reweighting

One of the drawbacks of metadynamics, which is shared with all the methods based on biasing CVs, is that those CVs should be chosen before performing the simulation, and their choice typically affects the accuracy of the final result. However, it is sometimes very useful to compute free energies as functions of CVs that differs from the biased CV. This can be done by an *a posteriori* analysis. This kind of analysis on MetaD simulations has been introduced in Ref. 46. It is based on the weighted histogram analysis method.[47] Although typically used in the framework of bias exchange metadynamics (see Section "Bias Exchange Metadynamics") where a large number of CVs is used, this technique can be applied straightforwardly to normal MetaD.

WELL-TEMPERED METADYNAMICS

Standard MetaD, introduced in the previous section, has two well-known problems:

- Its estimate for the free-energy landscape does not converge but fluctuates around an estimate that, at least for simplified systems, can be demonstrated to be unbiased.
- Because it is a flat histogram method, it tries to sample the whole CV space. This can push the simulated system toward states with nonphysically high free energy and might drift the simulation toward thermodynamically nonrelevant configurations.

As discussed, both these problems have been recognized and tackled respectively by taking time averages[43] and by using restraining potentials.[44,45] An alternative method that addresses both the problems in an elegant fashion,

called well-tempered metadynamics (WTMetaD), has been introduced in Ref. 48.

In WTMetaD, the rule for the bias update (Eq. [13]) is modified slightly to:

$$\dot{V}(s,t) = \omega \exp\left[-\frac{V(s(t),t)}{k_B \Delta T}\right] \exp\left[-\sum_{\alpha=1}^{N_{CV}} \frac{\left(s_\alpha - s_\alpha(t)\right)^2}{2\sigma_\alpha^2}\right] \qquad [16]$$

Equivalently, each Gaussian, when deposited, is scaled down by a factor $e^{-\frac{V(s(t),t)}{k_B \Delta T}}$, where the bias potential has been evaluated at the same point where the Gaussian is centered and ΔT is an input parameter measured in temperature units. Equation [16] implies that, after the initial filling, Gaussians of different height are added in different regions of the CV space. In particular, on top of deep wells, where a sizable bias has been already accumulated, the additional Gaussians have small height. In contrast, at the border of the explored region, where the bias is still small, the additional Gaussians have large height. In the long time limit, when the bias potential tends to grow parallel to itself, the simulated system should systematically spend more time on the regions where smaller Gaussians are used, that is, on top of deeper wells. This disrupts the flat histogram properties of the method and in turn implies that the sum $V(s) + F(s)$ is no longer encouraged to become flat.

To estimate this deviation from the flat histogram behavior quantitatively, one should assume that the added Gaussians are narrower than the features of the underlying free-energy landscape and thus can be considered equivalent to δ functions with the proper normalization:

$$\exp\left[-\sum_{\alpha=1}^{N_{CV}} \frac{\left(s_\alpha - s_\alpha(t)\right)^2}{2\sigma_\alpha^2}\right] \approx \delta(s - s(t)) \prod_{\alpha=1}^{N_{CV}} (\sqrt{2\pi}\sigma_\alpha) \qquad [17]$$

This allows us to approximate the bias update rule (Eq. [16]) as

$$\dot{V}(s,t) = \omega \exp\left(-\frac{V(s(t),t)}{k_B \Delta T}\right) \prod_{i=\alpha}^{N_{CV}} (\sqrt{2\pi}\sigma_\alpha) \delta(s - s(t)) \qquad [18]$$

In the long time limit, the CV s will be distributed according to a biased canonical distribution $P(s) \propto e^{-\frac{F(s)+V(s,t)}{k_B T}}$, and the bias will grow according to

$$\dot{V}(s,t) \propto \exp\left(-\frac{V(s(t),t)}{k_B \Delta T} - \frac{V(s,t) + F(s)}{k_B T}\right) \qquad [19]$$

By direct substitution, one finds the condition for the bias to grow uniformly, that is, for $\dot{V}(s,t)$ to be independent of s:

$$V(s) = -\frac{\Delta T}{T + \Delta T}(F(s) - C(t)) \qquad [20]$$

This last equation indicates that in WTMetaD the bias does not tend to become the negative of the free energy but is instead a fraction $\frac{\Delta T}{T+\Delta T}$ of it. Thus, it only partially compensates existing free-energy barriers by an *a priori* known scaling factor. This factor can be tuned adjusting the ΔT input parameter. Moreover, in the long time limit, the system will explore the biased canonical distribution:

$$P(s) \propto e^{-\frac{F(s)+V(s,t)}{k_B T}} \propto e^{-\frac{F(s)}{k_B (T+\Delta T)}} \quad [21]$$

As a consequence of the bias potential, the collective variable is exploring the canonical ensemble at an effective temperature $T + \Delta T$. Notice that the other microscopic variables are still sampled using a thermostatted MD at temperature T. In this sense, WTMetaD is related to other methods where an equivalent effect is obtained by extending the configurational space and by exploiting adiabatic separation of the CVs from the microscopic fluctuations.[13–15]

In short, WTMetaD allows performing a simulation where, in the long time limit, the effective temperature of one or more selected CVs is kept at an arbitrarily high value $T + \Delta T$. For $\Delta T \to \infty$, standard metadynamics is recovered (see Section "Adaptive Biasing with Metadynamics"). For $\Delta T = 0$, unbiased sampling is recovered.

This last feature of WTMetaD is of great advantage as it allows limiting the exploration of the CV space only to regions of reasonable free energy. Indeed, by fixing ΔT according to the height of the typical free-energy barrier for the problem under consideration, one will avoid overcoming barriers that are much higher than that.

Reweighting

As discussed in Section "Reweighting," it is sometimes useful to compute free energy as a function of *a posteriori* chosen CVs different from the biased ones. This is typically done in WTMetaD using a reweighting scheme introduced in Ref. 49. Here the time derivative of the partition function is estimated on the fly and allows consistent combination of data obtained at different stages of the simulation. An alternative has also been recently proposed[35] on the basis of the classical umbrella sampling formula.[9] Both these techniques are specific for WTMetaD.

METADYNAMICS HOW-TO

When adopting a new technique, the most frequent questions novice users ask are about how to choose the right input parameters. In MetaD in its various flavors there are three aspects that should be taken into account, namely:

- The choice of CV(s).

- The width of the deposited Gaussian potential.
- The energy rate at which the Gaussian potential is grown and, for WTMetaD, the parameter ΔT that determines the schedule for decreasing it along the simulation.

All of these decisions are deeply intertwined and we briefly outline here the main criteria for their choice, pushing on the practical implications and error diagnostics that allow one to decide whether the simulation parameters are well chosen or not.

The choice of the various parameters in MetaD generally relies on the following formula[38,42] that determines the error of a MetaD simulation under the assumption that the dynamics can be approximated by the Langevin equation:

$$\bar{\varepsilon} = C(d)\sqrt{\frac{S\delta s}{D}\omega \; k_B T} \qquad [22]$$

Here $C(d)$ is a prefactor that depends on the dimensionality of the problem (i.e., the number of CVs included), S is the dimension of the domain to be explored, δs is the width of the Gaussian potentials, D is the diffusion coefficient of the variable in the chosen space, and ω is the energy deposition rate for the Gaussian potential. This equation has several nuances that need to be explained, and here we will do it by considering alanine dipeptide as a "real-life" example.

The Choice of the CV(s)

Understanding the system one wants to study is pivotal in devising the correct CV(s) one needs. So, at any level, an accurate bibliographic search is very important in understanding the key factors in performing any kind of MetaD calculation. As an example, in computational biology problems, a mutation experiment that inhibits folding might be a signature of an important side-chain interaction that one needs to take into account.

More importantly it is worth understanding a key point of all the enhanced sampling techniques involving biasing one or more collective coordinate: a good descriptor might not be a good biasing coordinate. As a matter of fact, it may happen that different conformations might be distinguishable by using a specific dihedral as CV but if this is the outcome of a more complex chain of events and not the cause, one might end up by being disappointed by the result obtained by MetaD or other biased-sampling methods. Therefore, it is very important to understand that the lack of convergence of a free-energy calculation is often due to a suboptimal choice of CVs rather than the specific choice of the enhanced sampling technique.

In Section "Advanced Collective Variables" we will discuss *ad hoc* solutions for finding CVs that are optimal for specific problems. For the time being, we refer to the example of alanine dipeptide and try to identify the relevant

order parameters. Assume that we have no prior knowledge of the importance of the Ramachandran dihedral angles, and we intend to describe the transition between C_{7eq} and C_{7ax} with one single order parameter. Intuitively, the C_{7eq} conformer is more extended than is the C_{7ax} conformation (see Figure 3). Therefore, "gyration radius," defined as

$$r_{\mathrm{gyr}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} q_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} q_i\right)^2} \qquad [23]$$

where the summation runs on N heavy atoms, could be considered as a reasonable guess.

A simple molecular dynamics run of 200 ps in both basins can show that this descriptor is able to distinguish one state from the other (see Figure 5) because most of the sampling for C_{7eq} is concentrated at high values, whereas C_{7ax} exhibits low values of radius of gyration.

It is worth mentioning here that a better order parameter should be able to map the two minima onto two completely different regions of the CV space. The implication of having a partial overlap of the populations for the two basins, when they are structurally dissimilar as in this case, is that the kinetics cannot be simply retrieved through the free-energy landscape but requires the calculation of the transmission coefficient.[2] Nevertheless, for the sake of argument, we keep the choice as simple as possible to reproduce the possible pathological behaviors that one can encounter when experimenting with metadynamics.

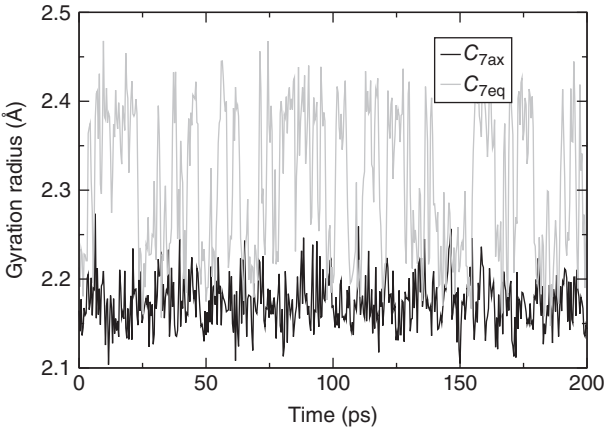

Figure 5 The value of gyration radius calculated on the heavy atoms for a trajectory of 200 ps of MD for alanine dipeptide in the two minima. Note that the average values of the gyration radius in the two cases are rather different, so this can be considered a possible choice for a CV.

The Width of the Deposited Gaussian Potential

The width(s) of the deposited Gaussian potentials is/are usually adjusted to adapt to the underlying free-energy landscape. Therefore, a preliminary unbiased MD run is generally performed, and the shape of the initial free-energy well is estimated. In this way one might figure out how wide the metastable basin is in the space of CVs. In a large system, to avoid performing extra heavy computations, the final part of the equilibration run can be postprocessed for this purpose. More specifically, for each CV, the probability density along the sampled CV range can be computed and, by assuming that the underlying free-energy landscape is a quadratic minimum, a Gaussian can be fitted on it. Under such assumptions, the addition of a Gaussian potential having height $k_B T$ and a width equal to the standard deviation of the fitted Gaussian can perfectly compensate the free-energy landscape so that the resulting landscape in the neighborhood should be flat. It should be taken into account that metadynamics does not aim at compensating perfectly the free-energy landscape at each step, though it aims at doing it on average in the long run (either completely, in normal MetaD of Section "Adaptive Biasing with Metadynamics," or partially, in WTMetaD of Section "Well-Tempered Metadynamics"); therefore, the Gaussian width σ of MetaD is generally set as a fraction of the standard deviation of such fitted Gaussian. One-half or one-third is generally suitable. Choosing a smaller width would capture better the free-energy features but would also give slower filling time and produce a steeper force that could interfere with integrator stability and therefore produce irrecoverable errors. Using a Gaussian fitting is done here with didactic purpose only: very often one just calculates the standard deviation out of the trajectory. This is illustrated in Figure 6.

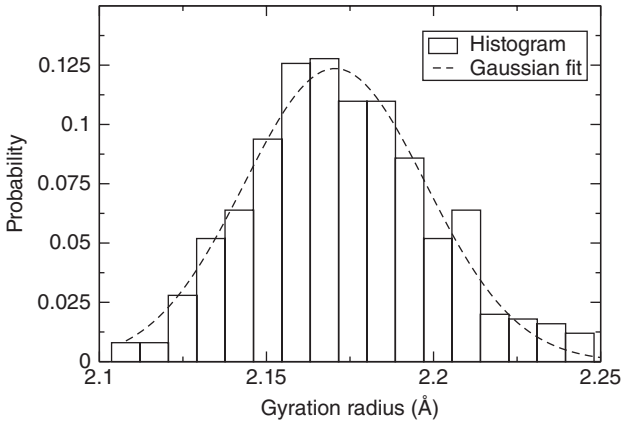

Figure 6 The probability of gyration radius calculated on the heavy atoms for a trajectory of 200 ps of MD for alanine dipeptide for the C_{7eq} minimum. The fitted normal distribution is centered in 2.17 Å and has standard deviation of 0.15 Å.

Moreover, it is worth noting that the chosen width for one region may not be optimal for other regions. This is particularly important with those CVs that present an anisotropic compression of phase space (see below in Section "Adaptive Gaussians" for a discussion of this issue and a possible solution).

The Deposition Rate of the Gaussian Potential

The deposition rate for the Gaussian potential can be expressed as the rate between the Gaussian height and the time interval between subsequent Gaussian depositions, that is, $\omega = w_G/\tau_G$. This rate can thus be tuned by adjusting both these parameters.

The choice of ω is crucial in metadynamics simulations because this parameter affects both the error and the speed at which the free-energy basins will be filled. It is wise to make it small enough to reduce the error but on the other side it must be sufficiently large to allow a reasonable filling rate of the free-energy landscape and to allow resampling the same point in CV many times. This balance will eventually produce accurate statistics that correspond to a good estimate of the free energy.

To avoid abrupt changes in the total potential when adding new Gaussians, one typically chooses the Gaussian height w_G as a fraction of the thermal energy k_BT. A typical choice is on the order of 0.1 kcal/mol for systems simulated at room temperature. Once this is fixed, the deposition time τ_G is typically adjusted to tune the deposition rate $\omega = w_G/\tau_G$.

As a rule of thumb, this can be done through a short free dynamics as shown in Figure 7. It is evident that the typical autocorrelation time of the CV is of the order of a picosecond or less. This sets a rough estimate for τ_G, which,

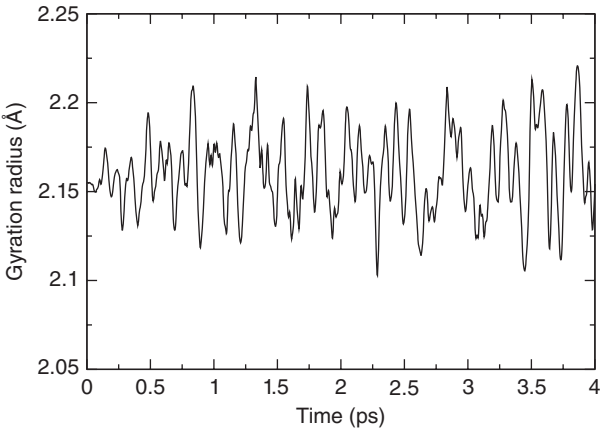

Figure 7 The value of gyration radius calculated on the heavy atoms for a trajectory of 4 ps of MD for alanine dipeptide in the C_{7ax} minimum. The values are acquired at each timestep of 2 fs.

for the following examples, will be chosen as 1.2 ps. By choosing the τ_G in this manner, the system is able to relax between one deposition and the other, to avoid placing the next Gaussian potential on top of the previously deposited one. This latter situation would produce an effective Gaussian potential with twice the height. In other words, this information encodes to some extent an estimate of the diffusion coefficient in the collective variable space that appears in Eq. [22].

A First Test Run Using Gyration Radius

With the parameters obtained from the analysis carried out above we can start a MetaD run. The parameters are $\sigma = 0.07$ Å, $w_G = 0.1$ kcal/mol, $\tau_G = 1.2$ ps. Using these parameters, we performed a 4-ns-long simulation. The time evolution of the CV is shown in Figure 8a. In a typical MetaD evolution the system spends a bit of time in a single basin around 2.2 Å of gyration radius, which is progressively filled, and this allows jumping after a few ps to the next minimum at 2.4 Å. The system is progressively biased and can reach very elongated as well as compressed states. While moving between the two extrema it should be noted that, as a result of the already filled free-energy landscape, the system is not stuck anymore in the starting basin: this is exactly the intended purpose of metadynamics.

It is crucial here to observe the so-called "recrossing" events, which means that the putative transition state is being explored by metadynamics many times. This ensures that the interesting states are sampled exhaustively, which is

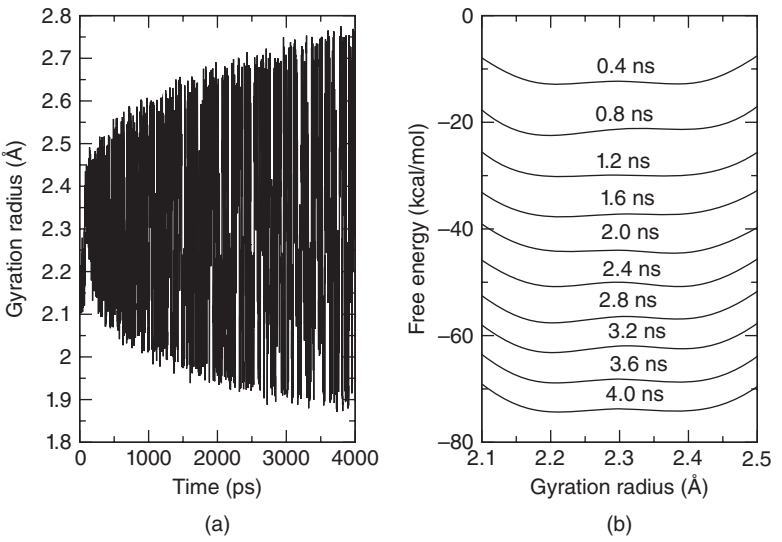

Figure 8 The time evolution of the chosen CV along the metadynamics (a) and the evolution of the bias (b).

important (but not sufficient) to obtain a correct result. In this case, as the two minima identified via straight MD are located at 2.1 Å and 2.4 Å one should look carefully for a number of transition in these regions, which occurs several dozens of times in the presented example. In realistic cases one may expect recrossing events of the order of tens.

In Figure 8b the evolution of the negative values of the bias is shown, which is the best approximation to the free-energy surface at a given time. It is worth noting that one should not worry about the absolute value of the bias to move progressively toward negative values, as the interesting features reside in the free-energy differences within one single profile and not in their absolute change from one time to the other. Interestingly, there is no sizable barrier between the two metastable states. This already casts some doubts about the quality of the CV employed. Indeed, if there is no barrier, one could have expected the transition to happen spontaneously in a few ps of simulation. This is a crucial point because critical judgment of consistency of both biased and unbiased simulations are very important to determine the quality of the results. Very often one can observe cases where an unexpected minimum occurs, which has characteristics not justified by any structural stabilizing feature. In such cases an unbiased MD run starting from a conformation taken from that minimum could allow one to clarify if the metastability is real or if it is an artifact of the CV configuration space coarsening.

An important point is shown in Figure 8a: in standard MetaD the extension of the explored portion of the CV domain increases with time. This generally gives access to higher energies and for many applications this has a negative effect. As an example, in proteins, one can reach energies that allow the unfolding of the secondary structure which, unless the algorithm is able to refold it, completely invalidates the simulation. Very often one has to artificially limit the exploration to prevent such effects by adopting confining potentials as discussed before. In our example the increased range explored implies that the landscape presents small changes from one time to the other in the region of interest, as the Gaussians are now deposited on a much broader range.

In Figure 9 (black lines) we report a set of free energies that are collected every 400 ps. Within 400 ps the system is in fact able to sample the gyration radius domain (2.1 Å < RGyr < 2.5 Å) several times; therefore, each realization of the free energy can be considered as a new estimate of the landscape. Under such conditions we can align all the landscapes by subtracting the average value and it becomes immediately evident that the landscapes may be very different from one another. This nonhomogeneity in the free-energy estimates is a source of error that is referred as "hysteresis" and is generally considered to be the most evident hint of a bad choice of CVs.

Frequently, there exist other significant diagnostics that can highlight problems in the CV choice and may help in devising a more suitable CV. For example, using an independent structural comparison consisting of a different descriptor, such as root mean square deviation (RMSD) after optimal

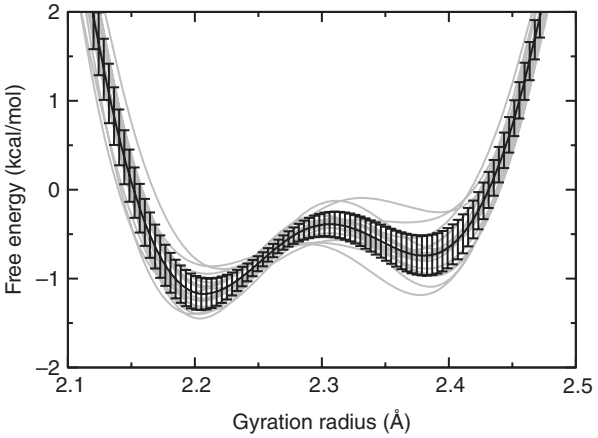

Figure 9 Average free-energy profile (cross-hatched thick profile) and associated error bars for different free-energy estimates aligned on their average value (gray profiles).

fit,[50] can reveal if the used CVs are resampling the expected states or if they are moving the simulation in unwanted phase space regions. Consider as an example alanine dipeptide in Figure 10, where the structural alignment with respect to C_{7ax} and C_{7eq} reference structure is reported. It is evident that C_{7ax} appears only within the first 500 ps (RMSD < 1 Å) of the simulation and never reappears even if the corresponding value of the gyration radius appears frequently in the simulation. In this case MetaD is not accelerating the sampling between the target states although many recrossing events are observed in the CV space.

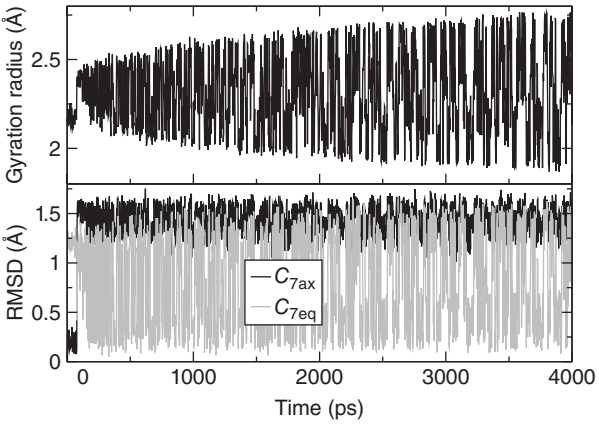

Figure 10 Time evolution of the radius of gyration (upper panel) along with Root Mean Square Deviation (RMSD) after optimal fit onto C_{7ax} and C_{7eq} reference structures.

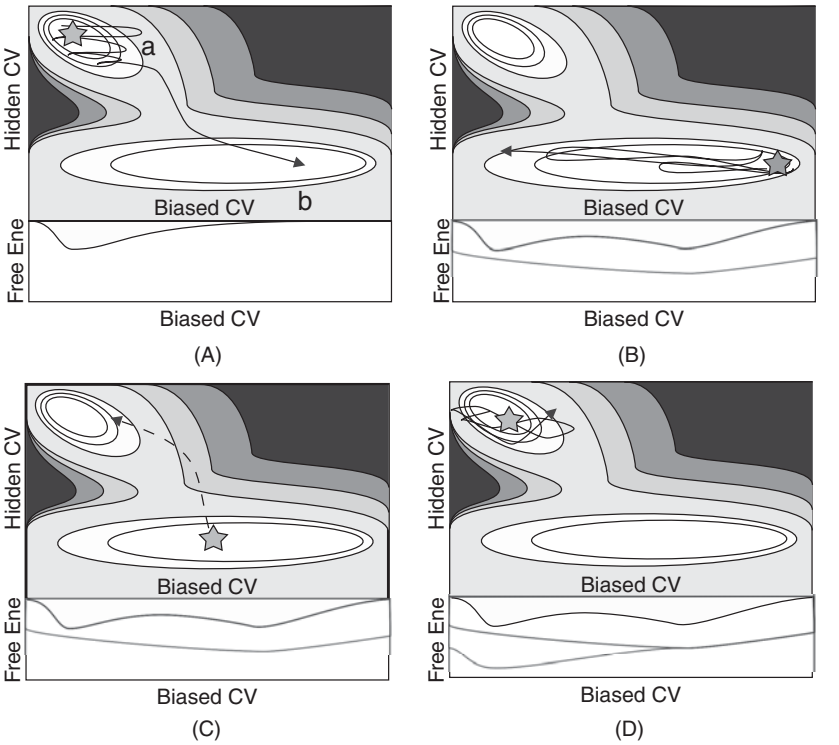

Figure 11 A representative chart of what may happen when making a wrong choice of CVs in MetaD. In the bottom of each panel is shown a sum of the potential to illustrate the effects on the estimated free-energy landscapes.

Figure 11 clarifies what happens in this system. Suppose we are interested in the difference between state *a* and *b* in Figure 11, which correspond to the C_{7ax} and C_{7eq} conformers, respectively. When performing MetaD along the variable on the horizontal axis the forces are added only in this direction, and the system is pushed away from minimum *a* thus reaching minimum *b*. This jump may have produced a change in a "hidden" CV that could be due to a peculiar free-energy landscape (as shown in Figure 11) or by a fortuitous jump in a parallel valley that is kinetically accessible. Once reaching *b*, MetaD will push the system in the valley located on the bottom (see Figure 11A) because the MetaD force is acting only in the horizontal direction and cannot control any further jump into the *a* basin, which is separated by a barrier. As a result, one gets a free-energy profile that resembles the bottom basin almost exclusively, leaving only a minimal trace of the initial basin (see Figure 11B). The system will only jump back to basin *a* because of an overaccumulation of potential or by a fortuitous thermal fluctuation (see Figure 11C). Whenever this occurs, the free energy changes and becomes similar to that of the projection of the *a* basin

onto the biased CV again (see Figure 11D). In this case the free-energy profile may change considerably at distant times and, consequently, show hysteresis.

This scenario points to the fact that the selected set of CVs is not suitable and a different choice should be made.

A Better Collective Variable: Φ Dihedral Angle

To illustrate the significance of CVs selection in MetaD and in other enhanced sampling schemes, the same calculation is now repeated but using the Ramachandran angle Φ as the CV. The σ was chosen according to the same procedure outlined above to be 0.1 rad while keeping all the other parameters identical.

The outcome using this CV choice is presented in Figure 12 and it is evident, especially in the progress of the deposited bias, that the shape of the deposited potential is conserved and grows approximately parallel to each realization. This results in a more uniform error as shown in Figure 13 where all the free energies are more consistent with one another.

Furthermore, the two metastable states are repeatedly visited as shown in Figure 14 where in both cases there exist structures with an RMSD < 1 Å along the whole trajectory. In the last case, the metadynamics run can be considered to provide a reliable free-energy profile.

Conceptually, the change in CV can be regarded as a sort of warping of the free-energy landscape. This is represented pictorially in Figure 15. In this

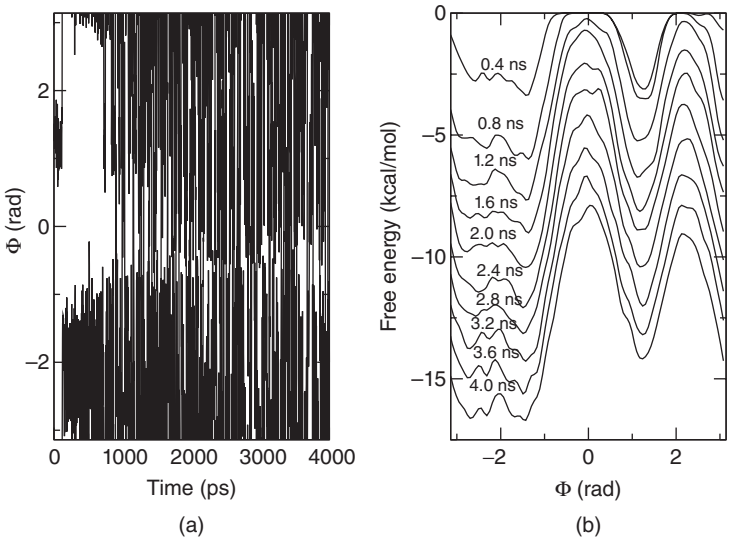

Figure 12 Time evolution of the Ramachandran dihedral angle Φ during metadynamics (a) and evolution of the deposited bias over time (b).

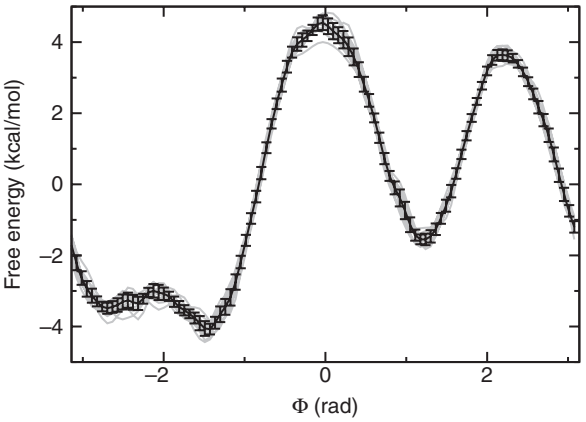

Figure 13 Average free-energy profile when using Φ as the CV in metadynamics. The error is reduced and the free-energy profiles (gray lines) are very coherent with one another.

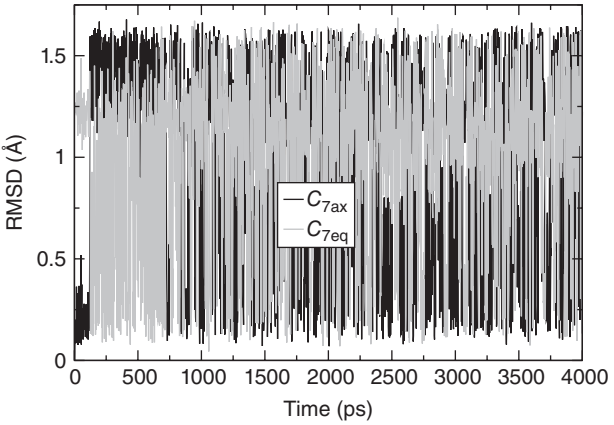

Figure 14 RMSD from the two C_{7ax} and C_{7eq} reference structures during the metadynamics run that uses the Ramachandran dihedral angle Φ as CV in metadynamics.

case no barrier appears on the hidden CV and consequently forces applied by MetaD are now in the direction connecting the two states.

Well-Tempered Metadynamics Using Gyration Radius

We turn here to describe the case of WTMetaD because recently it is becoming widely applied and its error control and associated diagnostics differ slightly from the case of the normal MetaD. In WTMetaD there is an additional parameter to be set, namely, ΔT, which technically regulates the decay factor of the

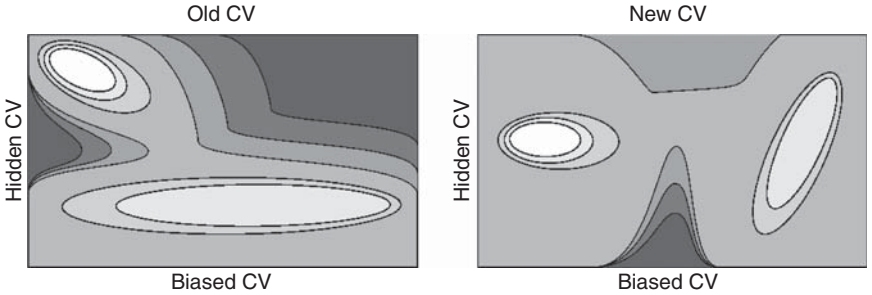

Figure 15 A graphical interpretation of the deformation of the landscape induced by the change of CV. Now no barrier appears on the hidden CV, and the two states do not project on top of each other but instead are well separated.

potential deposition rate. As discussed in Section "Well-Tempered Metadynamics," this can be regarded as a factor that determines the effective temperature of the enhanced CV. Thus, a wise way to choose ΔT is by having a rough estimate of the barrier. In our case, as the barrier is expected to be 8 kcal/mol, which is around 15 $k_B T$ units, we set $(T + \Delta T) = 15T$. This ensures that in the long run the bias will allow passing the barrier or, in other words, that before decreasing the Gaussian height dramatically a potential will be built that is able to overcome the barrier.

Unlike MetaD, WTMetaD provides only an exploration limited to the low free-energy region of the landscape. This is clearly visible from Figure 16 where

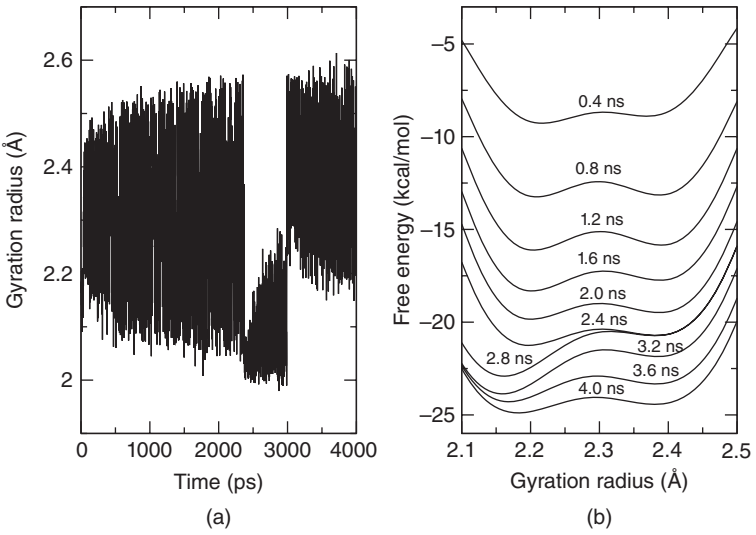

Figure 16 The time evolution of the chosen CV along the metadynamics (a) and the evolution of the bias (b).

the space visited is limited to 2 Å < RGyr < 2.6 Å, which is remarkably less than the non-well-tempered case. This is of course a big advantage over standard MetaD because the exploration can be restricted to only the interesting region where activated processes take place and prevents the sampling of high free-energy regions where irreversible transitions might occur. Another visible difference is that the profiles (Figure 16b) are getting closer to one another as time passes. This is the effect of the Gaussian height diminishing and, therefore, the deposited potential converging. Moreover, because a correct WTMetaD is expected to converge, one typically estimates the free energy from the final bias potential without averaging.

Nevertheless, pathological CVs still show their "hysteresis" issues (see the sudden change in profile at 2.8 ns, Figure 16b), which has a remarkable counterpart in the CV evolution in the Figure 16a. This aberrant behavior generally arises from the system beginning to explore a different free-energy surface after having performed a jump in a so-called "hidden" CV (see Figure 11). At that stage the underlying free-energy landscape changes, and the system stays trapped in a region, therefore adding more and more bias.

From these considerations, it becomes evident that the Gaussian height evolution gives a hint about when this is happening as shown in Figure 17a. Here, while the system seems to decrease the Gaussian height, as it is continuously resampling the same CV space, at 2.4 ns a sudden jump appears. This is usually associated with a jump into a nonsampled region where Gaussian height increases. What happens from the structural point of view is that the state C_{7ax} is occasionally resampled. This is clearly visible from the RMSD, which is reported in Figure 17b. In the schematic of Figure 11 this change corresponds to a jump on the hidden CV from basin *b* to the target basin of *a*. Once in C_{7ax}, the system finds a very narrow basin in free energy, which was

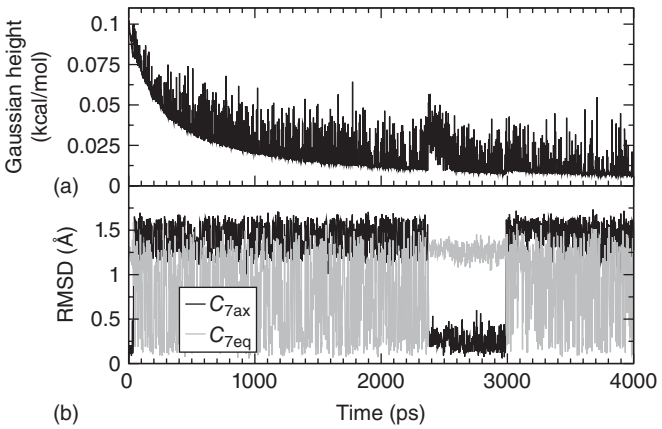

Figure 17 The time evolution of the height of the potential deposited (a) and the RMSD with respect to the two C_{7ax} and C_{7eq} states (b).

not accounted for by the bias potential accumulated by WTMetaD (almost flat in this region), so it starts accumulating potential to escape and eventually it does at 3 ns, where it reaches once again C_{7eq}. At this point the potential starts washing away the changes because of the C_{7ax} basin, and the old landscape is retrieved, leaving the user in doubt about which is the correct landscape.

Once again, a bad choice of CVs makes the transition from C_{7ax} to C_{7eq} rather fortuitous. This indicates that the selected CV is not a good variable for biasing, with either WTMetaD or MetaD.

It may also happen in WTMetaD that the growth of the bias potential is a poor diagnostic tool for detecting "hysteresis," especially when the system shows transitions at very long time and at low Gaussian height. In this respect the probability histogram of the CVs calculated over different trajectory segments can be more indicative of underlying problems. Indeed, in WTMetaD the system should converge to a probability $P(s) \propto e^{-\frac{F(s)}{k_B(T+\Delta T)}}$. Any divergence from this behavior should be looked at with caution as it might be a signature of poor convergence. This is evident in Figure 18 where the trajectory in the CV space is divided in four blocks of 1 ns each, and the probability distribution is calculated for each block. Note that the third and the fourth blocks are remarkably different from the first two and from one another.

Well-Tempered Metadynamics Using Dihedral Angle Φ

We revise here the last example to show how the diagnostics change when the CV is well chosen. As before, we adopt the Φ Ramachandran dihedral angle as a biased CV.

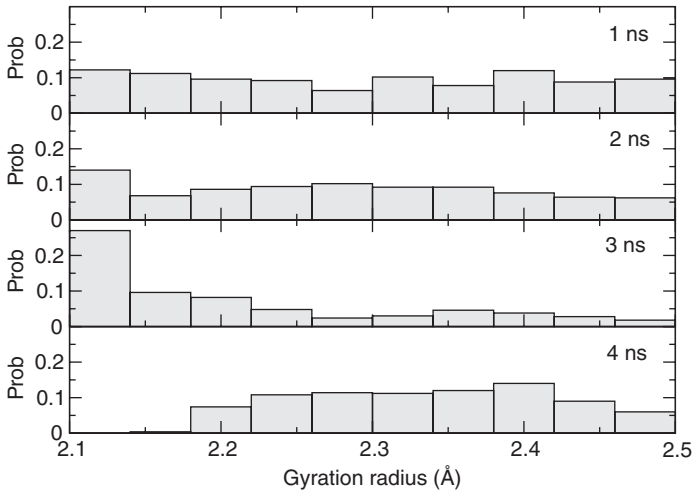

Figure 18 The probability distribution of finding the system in the CV range in four blocks of the simulation. It is evident that this distribution does not converge.

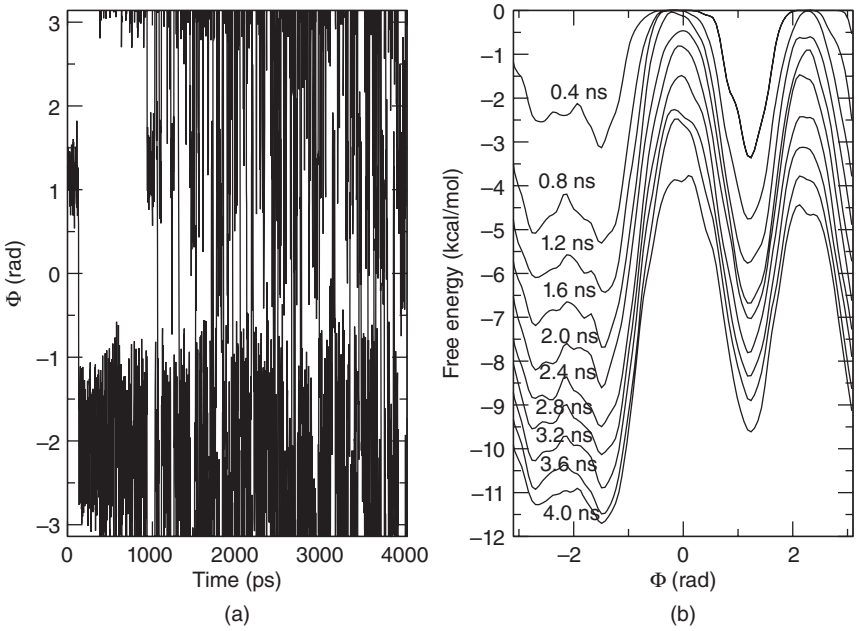

Figure 19 (a) CV evolution with time. (b) The evolution of the best estimate of the free-energy landscape from WTMetaD when using Φ as CV.

While the time evolution of the free-energy landscape (see Figure 19) is similar to the conventional MetaD case (Figure 12) we point out that WTMetaD is depositing less potential than the standard case, because of the diminishing Gaussian height.

The evolution of Gaussian height is also more regular as it can be seen in Figure 20a. Discontinuities in the decay of the Gaussian height are frequent and coincide with the jumps between C_{7ax} and C_{7eq} states, which are repeatedly visited as seen in the RMSD plot (Figure 20b).

Finally, the probability evolution is also more homogeneous with time as one might expect in the case of no hysteresis (see Figure 21).

ADVANCED COLLECTIVE VARIABLES

As discussed in the previous section, one of the most difficult issues in running metadynamics simulations is that of choosing a proper set of CVs. The limited reliability of metadynamics calculations is often mistakenly perceived as a limitation of the method although the real reason is generally due to the descriptors used.

A good CV that is able to distinguish reactants and products may not be sufficient in driving the reaction reversibly. This is due to the fact that a

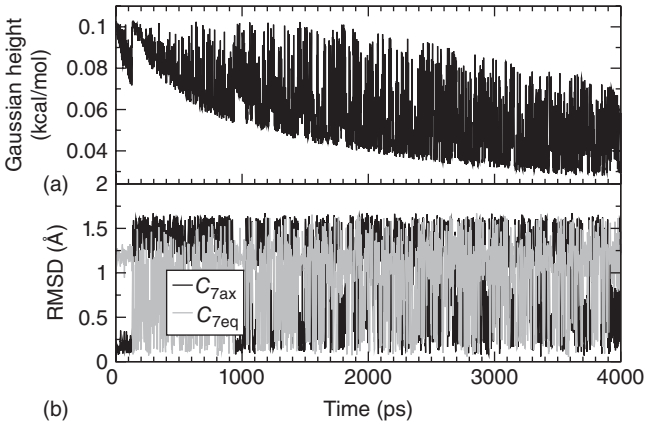

Figure 20 (a) The time evolution of the Gaussian height for WTMetaD with Φ CV. (b) RMSD with respect to the two states C_{7ax} and C_{7eq}.

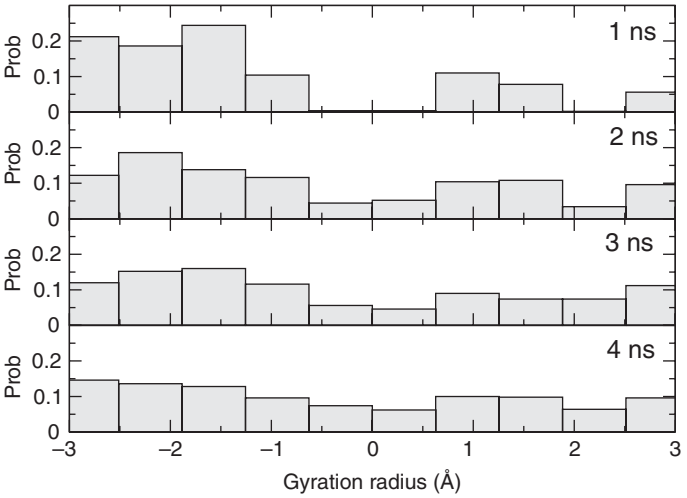

Figure 21 The probability distribution of finding the system in the CV range in four blocks of the simulation. It is evident that this distribution converges smoothly and consistently presents a higher probability for angles corresponding to the two metastable states.

change that is observable during a process might be triggered by a different (and hidden) physical cause. Thus, driving the observable is not guaranteed to facilitate the transition.

We also discussed how important it is to devise a CV that can also produce a force in a direction that permits the system to travel between all relevant configurations.

For these reasons it is typically more useful to explore multiple sets of CVs rather than to tune the other metadynamics parameters. Moreover, moving toward a better sampling by attempting various CV sets always implies a better understanding of the physicochemical process being studied.

Given the inherent difficulty of making such a choice, much effort in the past decade has been focused on developing CVs suitable to describe correctly a variety of conformational changes. We review here a few of the most popular choices used with metadynamics but which can also be used in other schemes like umbrella sampling, thermodynamic integration, and steered molecular dynamics.

Path-Based Collective Variables

The computational cost of metadynamics can scale as $S^{N_{CV}}$ where S is the extent of the CV space and N_{CV} is the number of CVs used. Often, it is inherently difficult to reduce N_{CV} because of its nature of being intrinsically multidimensional. For example, protein folding is due to an interplay of backbone flexibility, hydrogen bonding, hydrophobic collapse, and dehydration intertwined with one another. Another example is chemical reactions in enzymes involving several hydrogen bonds and backbone distortions. Sampling extensively all those degrees of freedom independently is of little help because most of them move in a concerted fashion, along a hypothetical one-dimensional reaction tube that changes character as one progresses from reactant to products. The need to find a free energy along an adapted "reaction tube" was first perceived by Ensing and Klein[51] who performed an explorative metadynamics run, followed by a refinement through an umbrella sampling calculation along the obtained minimum free-energy path. Later on a general functional form was found for this "reaction tube" to be sampled directly through MetaD, which is the purpose of the so-called path collective variables.

By assuming that the user already has M available snapshots $\{\widetilde{q}(i)\}$ (also called "nodes") that are representative of the transition and that can be obtained via high temperature MD or steered MD, a progress variable can then be built:[34]

$$s(q) = \frac{1}{M-1} \frac{\sum_{i=1}^{M} (i-1) \exp(-\lambda \|q - \widetilde{q}(i)\|)}{\sum_{i=1}^{M} \exp(-\lambda \|q - \widetilde{q}(i)\|)} \qquad [24]$$

where λ is a parameter that should be chosen according to the average interframe distance, and $\|q - \widetilde{q}(i)\|$ is the distance between the current molecular dynamics snapshot and the $\widetilde{q}(i)$ node. Intuitively, this variable provides a continuous fractional index ranging from 0 (the reactant) to 1 (the product) that allows one to map the conformations along an ideal "progress along the path" variable.

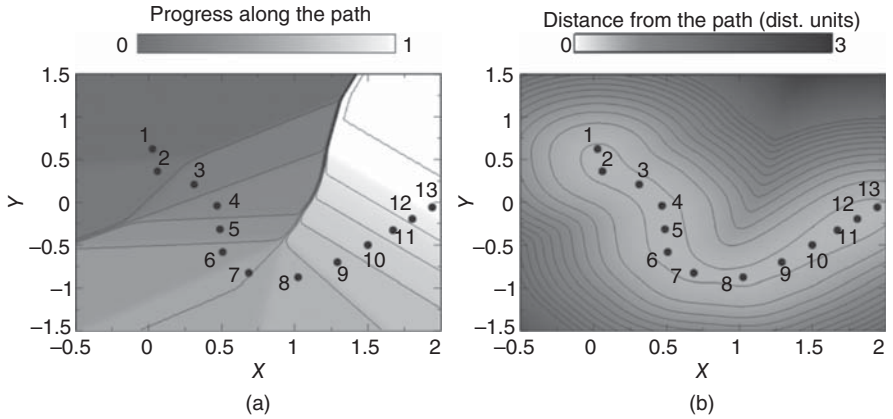

Figure 22 A two-dimensional example of the path variables and the foliation induced by the progress along the path (a) and the distance from the path (b).

Similarly, a complementary variable can be defined that provides the distance from the closest node:[34]

$$z(q) = -\frac{1}{\lambda} \ln \sum_{i=1}^{M} \exp(-\lambda \|q - \widetilde{q}(i)\|) \qquad [25]$$

A two-dimensional example is shown for a number of nodes in Figure 22. The foliation induced by Eq. [24] (Figure 22a) is similar to that produced by Voronoi tessellation.[52] The gradient of the variable changes direction along the path to adapt to the reaction character. Nevertheless, it shows some pathological behaviors at specific points where the indexing can change abruptly; this typically happens at large distances from the path. That is why defining the distance from the path by Eq. [25] (Figure 22b) is particularly helpful: the exploration of this variable can be artificially limited to prevent high-distance regions from being sampled. Alternatively, this variable can be used as an extra CV for metadynamics, in an attempt to explore paths that are far from those defined by the nodes.

In this framework the definition of the metrics $\|q - \widetilde{q}(i)\|$ plays a key role. An arbitrary reduced representation could be used in the spirit of the string method in CV space.[33] In general, the Cartesian coordinates of a subset of atoms are widely used and compared via mean square deviation after optimal alignment:[50]

$$\|q - \widetilde{q}(i)\| = \sum_{k,\alpha} \left[q_{k,\alpha} - cm(q)_\alpha - \sum_{\beta} \mathbf{R}(q, \widetilde{q}(i))_{\alpha,\beta} [\widetilde{q}(i)_{k,\beta} - cm(\widetilde{q}(i))_\beta] \right]^2 \qquad [26]$$

where $\mathbf{R}(q, \widetilde{q}(i))_{\alpha,\beta}$ is the optimal alignment matrix that superimposes q to $\widetilde{q}(i)$ and $cm(q)_\alpha$ is the α component of the center of mass of q. This metric was used in a variety of contexts including conformational transitions,[34,53] enzymatic catalysis,[54] chemical reactions,[55,56] and folding.[57] Notwithstanding the simplicity in its setup, the choice of the initial nodes must be done carefully. This set of "omnicomprehensive" Cartesian coordinates, although able to include most of the effects, is likely to increase artificially the distance between the molecular dynamics run and the nodes because multiple and irrelevant sources of noise contribute to it. Therefore, other metrics were proved to be useful like the comparison of contact maps,[57] which correctly detected the unfolded versus molten globule transition in folding. Also, the chirality indexes[58] that rely on the internal degrees of freedom associated with backbone conformations were found to be reliable.

An additional benefit from using this set of CVs is that one is not limited by the rough initial input. Instead, an iterative procedure can be set up to refine the frames and end up with an optimal description for the metrics employed.[34] This makes it possible to provide a better description of the saddle point, which can be quantified by computing the committor probability.[59] This procedure can be conveniently simplified by using an iterative steered-MD procedure as shown in Refs. 54, 55. More recently, a new integrated procedure was also introduced[60] on the basis of performing MetaD along the path while simultaneously recording the position orthogonal to the path. At regular intervals along the simulation, the path is evolved. At the end of the procedure it is possible to obtain both the free-energy landscape along the path and the optimized path. This is possible because MetaD allows the free energy along the path to adapt while the definition of the path itself is changed.

Collective Variables Based on Dimensional Reduction Methods

Another way to tackle the problem of defining the CVs is to renounce "general-purpose" CVs (e.g., number of hydrogen bonds, coordination of water, gyration radius in a typical folding problem) and to deduce them directly from the behavior of the system during the simulations.

The objective of these methods is to obtain a representation of the system in terms of a very small number of parameters. Typically, those parameters will be a complex but automatically determined function of the components of an intermediate representation of the system, in which the dimensionality of the problem can involve as many as several hundred degrees of freedom. Examples of such an intermediate representation might be the position of the heavy atoms in alanine dipeptide, the C_α carbons, or all the residue-to-residue distances in a protein.

The most straightforward use of this intermediate coarse-graining representation is through principal component analysis (PCA)[61] of a trajectory.

This produces a set of eigenvectors and eigenvalues that decompose the system motion into different amplitude motions acting on an average conformation. It is then possible to use the projections of the position of the system onto a subset of the eigenvectors as CVs in a MetaD simulation.[62] Generally, projections are performed onto the eigenvectors connected to large amplitude motions to accelerate slow conformational transitions.[63] It is worth noting the strong similarity with conformational flooding[37] in which the potential from the PCA is calculated only once and is not built adaptively as in MetaD.

The system under study can display many basins, each of them with a peculiar average structure and a specific set of eigenvectors, which might be very different from one another. For this reason Tribello and coworkers developed "reconnaissance metadynamics"[64] in which the dynamics is analyzed on the fly and onto which a Gaussian mixture model is fitted. The system is then assigned to one of the basins defined by the Gaussian, and the potential is added in one dimension (the distance from the center of the fitted Gaussian). This creates an "onion-like" potential that builds up to a point where it is detected by its fall into a new basin where the Gaussian mixture fitting is repeated. This method is very effective in sampling the configurational space, but no straightforward procedure exists yet for extracting free-energy profiles from the obtained trajectories.

Other methods exist besides PCA-based approaches that are able to extract a pattern from computed statistics. One of them is the so-called "classical multidimensional scaling method"[65] in which a fictitious, low-dimensional distribution of points is calculated to reproduce the distance relations among the points in the intermediate dimensionality. An example is the map of the distances in RMSD for configurations of dialanine. There, the dihedral angle Φ rotation is very similar to the distances that one would obtain from a suitable distribution of points along a ring. This implies that not all the coordinates of the intermediate representation are useful, because the landscape in this case is effectively two-dimensional.

This idea of using a high-dimensional data to fit a low-dimensional representation was exploited by Spiwok[66] who used this scaled representation to define the following modified path collective variable

$$s_P(q) = \frac{\sum_{i=1}^{M} P_i \exp(-\lambda \| q - \tilde{q}(i) \|)}{\sum_{i=1}^{M} \exp(-\lambda \| q - \tilde{q}(i) \|)} \qquad [27]$$

where the value of P is a general (vectorial) property of the node $\tilde{x}(i)$ that can be chosen arbitrarily. In particular, the authors chose the coordinates P_x, P_y, and P_z from an isomap multidimensional scaling derived from extensive sampling of a small molecule, much in the spirit of the work from Das et al.[67] In this way MetaD on that scaled representation can be performed effectively.

More recently, Ceriotti et al.[68] further developed the multidimensional scaling concepts for molecular representations and proposed a scheme called

"sketch-map." Their approach is a filtered multidimensional scaling algorithm that can reproduce the distances of the intermediate representation in a specific window of lengths, while coarsening the distances that are both higher and lower than this window. They proved this makes the problem more tractable and allows for a much broader range of scaling possibilities compared to standard multidimensional scaling. Eventually, they also figured out a novel way to perform MetaD that can be used specifically with sketch-map.[69]

Template-Based Collective Variables

The predictive simulation of protein folding at atomistic resolution is one of the grand challenges in the biophysical community and in statistical mechanics. The problem consists in predicting, with only the knowledge of the residues' sequence, the most stable state, which, in general, is known to be associated with its function. General purposes CVs, as distance or coordination numbers, are ineffective in reversibly folding, even for a small sequence of aminoacids.

Therefore, because proteins generally exist as one of several recurrent structural motifs, it is plausible to use a CV that is a similarity measure with minimal building blocks of one of those motifs. One then evaluates the number of times this building block is present in the actual MD snapshot. This is the central idea behind the template-based collective variables of Pietrucci and Laio.[70]

In detail they first identified an ideal building block for every structural pattern that could be either an α-helix, a parallel β-sheet, or an antiparallel β-sheet. They did that by isolating from the CATH[71] database all segments containing that specific motif using the STRIDE[72] definition and then by clustering them with distance-RMSD (dRMSD) comparison. The choice of dRMSD makes the computational comparison particularly efficient because it avoids calculating the rotation matrices and their derivatives. Once the center of the cluster is identified for each motif, the ideal structural motifs $\mathbf{d}_\alpha$, $\mathbf{d}_{\text{para}-\beta}$, $\mathbf{d}_{\text{anti}-\beta}$ can be retrieved as a list of pairwise distances. Each template is used to compute three distinct CVs: α-helix content, parallel β-sheet content, and antiparallel β-sheet content. For each MD snapshot the structural motif content of a given consecutive protein segment can be estimated simply by (e.g., for the α-helix case):

$$S_\alpha = \sum_i n[\text{dRMSD}(\mathbf{d}(q_i), \mathbf{d}_\alpha)] \qquad [28]$$

where $\mathbf{d}(q_i)$ includes the atoms of the block of three residues starting with atom q_i. The function n assigns more weight to those structures having a low dRMSD with respect to the template structure and therefore can be seen as a counter. Its functional form is:

$$n[\text{dRMSD}] = \frac{1 - (\text{dRMSD}/r_0)^8}{1 - (\text{dRMSD}/r_0)^{12}} \qquad [29]$$

where r_0 is a radius that discriminates if the segment is similar to the template or not and is generally set to 1.0 Å. It is important to note that in the dRMSD calculation, only C_α, N, C_β, and carbonyl oxygen atoms are included, and that the entire sequence is scrolled by steps of a single residue. The scrolling pattern depends on the type of secondary structure.

In this way the authors successfully produced a varied ensemble of secondary structures. More recently, these variables were used in conjunction with an extensive bias exchange MetaD (see Section "Bias Exchange Metadynamics") on VAL60 to sample, with atomistic resolution, plausible folds containing secondary structures[73] well beyond the variety contained in the PDB database.

Importantly, such strategy could be adapted for use well beyond the actual scope of protein folding and might be useful where other kinds of template are available.

Potential Energy as a Collective Variable

In the last three subsections we saw how carefully designed CV can be used as suitable reaction coordinates. We consider here the possibility of using the potential energy as a CV.

This idea was first proposed by Bartels and Karplus[74] in the context of adaptive umbrella sampling, and later used for plain metadynamics by Micheletti et al.[43] When a flat histogram distribution of potential energy is enforced, the system under investigation is able to freely explore values of the potential energy in a large window. When the potential energy grows, the explored conformations are equivalent to those explored at a higher temperature, so that, roughly speaking, the simulation performs a sort of annealing where the effective temperature is allowed to increase and decrease. At high temperature, the effective barriers are largely decreased, and the system is able to cross free-energy barriers easily. In this sense, achieving a flat histogram distribution of potential energy can be considered as comparable to performing a multicanonical simulation[75] and is also strictly related to Wang-Landau sampling.[76]

In a later work, Bonomi et al.[77] showed that it is possible to perform a WTMetaD simulation on the potential energy by sampling the so-called "well-tempered ensemble." This ensemble is characterized by a parameter γ, which is related to the usual parameter ΔT by $\gamma = \frac{T+\Delta T}{T}$. In the well-tempered ensemble the distribution of the potential energy is not flat but instead is related to the canonical distribution $P(U)$ by $P_{\mathrm{WTE}}(U) \propto P(U)^{1/\gamma}$. The effect of the γ parameter is to increase the fluctuation of the potential energy by a factor γ. This should be compared with a multicanonical ensemble,[75] where the fluctuations of the potential energy are infinite. In Ref. 77 the authors also showed how this amplification of the potential energy fluctuations can be used in practice to enhance the efficiency of parallel tempering simulations and in particular how this method can be used to decrease the number of replicas

required for parallel tempering simulations by orders of magnitude. Moreover, thanks to a suitable reweighting scheme (see Section "Reweighting"), it is possible to extract the free-energy landscape as a function of any *a posteriori* chosen CV. The combination of parallel tempering and well-tempered ensemble can also be used in conjunction with plain metadynamics, where standard collective variables such as number of hydrogen bonds or hydrophobic-core size are also biased.[78]

Interestingly, in a recent paper it has been shown that it is possible to dissect the potential energy and use as a CV only a relevant portion of it, namely, the Debye–Hückel interaction between two molecules.[79] Although in that paper this CV was biased using steered molecular dynamics, the extension to umbrella sampling or metadynamics is straightforward.

IMPROVED VARIANTS

Besides optimizing the choice of the CVs, several variants of the basic schemes outlined in Sections "Adaptive Biasing with Metadynamics" and "Well-Tempered Metadynamics" have also been introduced. We will discuss here three of them, which are based on the idea of running multiple metadynamics in parallel with different degrees of coupling. We will also describe a recently introduced scheme aimed at simplifying the choice of Gaussian width and at adapting it on the fly.

Multiple Walkers Metadynamics

As pointed out in Section "Adaptive Biasing with Metadynamics," a critical choice in metadynamics involves the deposition rate. In particular, a higher deposition rate is required for a faster initial filling of the free-energy landscape, but that would affect the final error. The multiple walkers algorithm is designed to exploit concurrent metadynamics simulations to allow a very fast filling albeit using a slow deposition rate.

In multiple walkers metadynamics[80] N_w metadynamics simulations (also referred to as *walkers*) are run concurrently, possibly on different machines. All these simulations contribute to the growth of a unique bias potential, which thus grows at a speed that is N_w times larger than for a single simulation. It has been shown heuristically (see Ref. 80) and analytically for a Langevin system (see Ref. 42) that the resulting error is the same as that expected from a single walker using the same Gaussian height and deposition time. This means that when using N_w walkers a filling time acceleration by a factor N_w can be obtained without increasing the error or, alternatively, the squared error can be decreased by a factor N_w without increasing the filling time.

In multiple walkers metadynamics there is no efficiency gain in computing time, as the same accuracy could be obtained by performing just a single simulation N_w-times longer. However, because the walkers are weakly coupled, this algorithm can be run easily on a parallel machine or even on a weakly interconnected cluster. Only a shared file system is required to allow interwalker communication in typical implementations. Moreover, it is not even necessary to start all the walkers at the same time, and there is no practical problem if one or more of the simulations is interrupted at some point.

Multiple walkers metadynamics can also be used in the well-tempered algorithm.[81] In the long time limit the error prefactor will be decreased when increasing the number of walkers because the walkers will provide independent statistics biased by the same potential.

As a final remark, we note that the theoretically perfect scaling of multiple walkers metadynamics can be reached only if the walkers are started from independent configurations taken from the proper equilibrium distribution. This can be a difficult task, as the free-energy landscape is not known *a priori*.

Replica Exchange Metadynamics

Replica exchange molecular dynamics is a technique where many simulations are performed at the same time using different control parameters. A typical example is parallel tempering,[7] where different temperatures are used. Hamiltonian replica exchange,[82] however, provides a more general formulation where different force field parameterizations are used in different replicas. Replicas are ordered to form a sort of ladder. The first step typically represents experimental conditions whereas the last step represents an artificially modified system (e.g., with a fictitiously high temperature in parallel tempering) where transitions are more likely to be observed. Coordinate swaps are attempted with a chosen time interval. This brings information from highly ergodic modified simulations to the typically frustrated room temperature simulation. To achieve sampling of the canonical ensemble, coordinate exchanges must be accepted according to a Metropolis rule where the acceptance is given by

$$\alpha = \min\left(1, \frac{e^{-\beta_i U_i\left(q_j\right)-\beta_j U_j(q_i)}}{e^{-\beta_i U_i(q_i)-\beta_j U_j(q_j)}}\right) \qquad [30]$$

Here U_i is the potential energy used for replica i, β_i its inverse thermal energy, and q_i represents the coordinates of all the system for the ith replica.

Replica exchange methods are typically expensive as they require many (sometimes hundreds) simulations to be run synchronously, often on a large supercomputer facility. Their advantage is that they allow us to accelerate sampling using a minimal amount of *a priori* knowledge about the specific transition under study. This contrasts with metadynamics, which is very cheap from a computational standpoint but requires physical insight to select proper CVs.

Thanks to their complementarity, metadynamics and replica exchange can nonetheless be optimally combined.[17] To do so one has to perform N_r simultaneous metadynamics simulations and adjust the Metropolis criterion to include the bias potential

$$\alpha = \min\left(1, \frac{e^{-\beta_i\left(U_i\left(q_j\right)+V_i(s(q_j),t)\right)-\beta_j(U_j(q_i)+V_j(s(q_i),t))}}{e^{-\beta_i(U_i(q_i)+V_i(s(q_i),t))-\beta_j(U_j(q_j)+V_j(s(q_j),t))}}\right) \quad [31]$$

The combined algorithm has the advantage with respect to metadynamics of accelerating all the degrees of freedom of a system without any *a priori* knowledge, albeit being more expensive. Consequently, if the high temperature replica is ergodic enough, hysteresis effects (see Section "A First Test Run Using Gyration Radius") because of a suboptimal choice of the CVs are strongly moderated or even disappear. When compared with replica exchange alone, replica exchange metadynamics improve the sampling of energetically unfavorable points such as metastable minima or transition states, provided they are defined on a CV space that is *a priori* identifiable. In this sense, replica exchange metadynamics takes the best from both worlds.

Notably, metadynamics can also be combined with replica exchange techniques other than parallel tempering, for example, solute tempering.[83]

Bias Exchange Metadynamics

One of the problems in metadynamics is that the maximum number of CVs that can be used in practice is limited. In particular, it is very difficult to converge metadynamics simulations performed using more than three CVs. This limits the applicability of the method to cases where a few appropriate CVs are known *a priori*. However, if one is able to define a super-set of CVs, which is very likely to include the appropriate CVs, multiple metadynamics simulations can be conveniently combined by means of bias exchange metadynamics (BEMD).[84]

In BEMD, several metadynamics simulations acting on different CVs are performed at the same time. Each of the simulations employs one or more CVs, and there might be overlap between the sets of CVs chosen for each simulation. For instance, the first replica could use two-angle CVs ψ and ϕ, the second replica could use only ϕ, and the third only ψ. Then, from time to time, coordinate exchanges are proposed and accepted with the usual Metropolis criterion. Like parallel tempering simulations, if the acceptance rate for the exchange moves is sufficiently high and one replica samples ergodically the configurational space, ergodicity is ensured in all of them. Whereas in parallel tempering the ergodic simulation is typically the one at the highest temperature, in BEMD any of the replicas could be ergodic depending on how virtuous the CV(s) used for that replica is(are) to describe the relevant conformational

transitions. In addition, far fewer replicas are needed than for a conventional parallel tempering-metadynamics calculation.

Using BEMD it is thus possible to replace an (unfeasible) metadynamics that would use ten or twenty CVs with ten or twenty weakly coupled one-dimensional metadynamics simulations and retain an accuracy that is determined by the replica containing the optimal CV. Moreover, the statistics of each of the replicas benefits from the exchange. Accordingly, the exploration of phase space by each replica is much higher compared to that of a simple MetaD. We finally point out that in BEMD the statistics produced by all the replicas should be used to calculate the final free-energy landscape via a reweighting procedure.[46]

Adaptive Gaussians

Up to this section, the shape of the repulsive potential adopted by MetaD was assumed to be of the form of Eq. [11] with σ_α fixed at the beginning of the simulation. As presented in Section "Metadynamics How-To," the variance of the CV from a simple MD run is used in devising the values of each σ_α, so, this value may change dramatically depending on the starting configuration. This is what might happen in the case of a protein folding problem, assuming the use of a set of distances as CVs. While the folded state is well defined and the variance measured for many CV might be small, this is not the case for an unfolded state that is slowly diffusing and is moving in a broad and featureless free-energy landscape requiring a larger σ_α.

In addition, the problem of selecting an adequate σ is tightly coupled with the definition of the CVs. For small molecules binding to a protein, coordination numbers are often used. These have a sigmoidal functional form that is tailored to change from one to zero when the small molecule is unbinding, but remains constant at values close to zero when the substrate is unbound. In this last case the free dynamics will provide very small fluctuations in CV toward zero, compared to that calculated for the bound state where the value is changing.

An additional intricacy to be considered is that two variables are often coupled as in the case of the two Ramachandran dihedral angles Φ and Ψ. These angles contain three atoms in common. Therefore, some sort of coupling should be introduced into the scheme to account for their interdependence.

We recently proposed[35] a more general multivariate Gaussian instead of Eq. [11]:

$$V(s,t) = \sum_{i=1}^{t/\tau_G} w_G \ \exp\left[-\frac{1}{2}\sum_{\alpha,\beta}\left(s_\alpha - s_\alpha\left(i\tau_G\right)\right)\sigma^{-2}_{\alpha,\beta}(s_\beta - s_\beta(i\tau_G))\right] \qquad [32]$$

Here α and β runs over the biased CVs and σ is a matrix that, unlike the standard MetaD case, might contain nonzero diagonal elements. Furthermore, in

this scheme the σ matrix is not fixed at the beginning of the simulation but adapts on the fly. To find this σ we proposed two approaches: one termed "Dynamically Adapted" (DA) MetaD and the other "Geometry Adapted" (GA) MetaD.

In the DA scheme the σ matrix is determined with an estimate over the space traveled in CV space:

$$\sigma^2_{\alpha,\beta}(t) = \frac{1}{\tau_D}\int_0^t dt'[s_\alpha(t') - \overline{s}_\alpha(t')][s_\beta(t') - \overline{s}_\beta(t')]\exp[-(t-t')/\tau_D] \qquad [33]$$

where the parameter τ_D is a typical time over which the drift is calculated. Intuitively, this is the time span that the Gaussian potential has to keep track of and is the only parameter to be set. $\overline{s}_\alpha(t)$ is the average position over the last portion of the trajectory defined as:

$$\overline{s}_\alpha(t) = \frac{1}{\tau_D}\int_0^t dt' s_\alpha(t') \quad \exp[-(t-t')/\tau_D] \qquad [34]$$

The advantage of this scheme is that the system detects the underlying landscape and adapts to it on the fly. A different approach is the GA scheme that takes into account only the compression of the phase space induced by the definition of the CV. In GA the σ matrix depends uniquely on the instantaneous conformation through

$$\sigma^2_{\alpha,\beta}(q) = \sigma^2_G \sum_i \frac{\partial s_\alpha(q)}{\partial q_i}\frac{\partial s_\beta(q)}{\partial q_i} \qquad [35]$$

In this formulation the only parameter to be set is σ_G, which is the effective extension of the Gaussian measured in the Cartesian space of atomic coordinates.

It is worth highlighting that in both formulations only one parameter has to be set and it replaces all the individual σ_α required by standard MetaD, to reduce substantially the number of parameters.

In addition, such schemes change both the width and orientation of the Gaussian potential on the fly thus taking into account the local coupling within variables (for the GA) or the local correlations in the diffusion matrix (for the DA).

In both cases, an estimate of the free energy cannot be done with the usual protocol because changing Gaussian width corresponds to using variable-size bins in the estimation of a probability. The correct free energy can be retrieved through a Torrie–Valleau-like formula[9] applied to the well-tempered case, when, for long simulation times, the bias increases constantly over the whole domain:

$$\widetilde{F}(s,t) = -k_B T \log N(s,t) - V(s,t). \qquad [36]$$

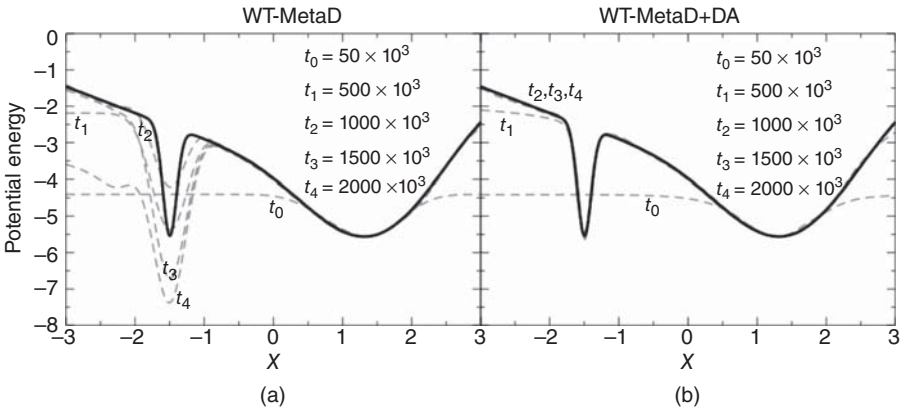

Figure 23 A one-dimensional Langevin model potential (black thick line) on which WTMetaD is applied (a) and on which WTMetaD coupled with the DA scheme is used. In dashed gray line the estimate of the free-energy landscape at different times along the simulation. The DA scheme always delivers a better and faster converging free-energy landscape. (Reprinted (adapted) with permission from Ref. 35. Copyright (2012) American Chemical Society.)

The improvement over standard metadynamics is remarkable as shown in Figure 23 for the case of a one-dimensional Langevin model. While standard WTMetaD may fail whenever σ is chosen too large, the DA scheme automatically tunes to the feature of the energy landscape.

CONCLUSION

Metadynamics is an established method aimed at accelerating molecular dynamics simulations and subsequently recovering free-energy landscapes. Its power and flexibility arise from the fact that it allows us to exploit the chemical and physical insight of the process under investigation to optimally spend the computational effort. However, in spite of its apparent simplicity, using metadynamics is nontrivial due to the input parameters that should be tuned for the specific application and also to the difficulty in choosing the proper collective variables for the simulation. Nevertheless, when done properly it also allows a deep understanding of the process under investigation to be reached, which is the ultimate goal of performing molecular simulations.

This chapter provides an introductory guide that should allow a practitioner to choose the input parameters of a metadynamics simulation and to optimize them for his/her own specific application. A particular focus was placed on avoiding the typical mistakes that can be done when preparing such a simulation and on how they can be solved practically. An assessment of the most widespread variants of the method was also outlined. This chapter is not

intended to be complete, as some of the most recent improvements have not been discussed (see, e.g., Refs. 85–90). Nevertheless, this introduction will be of value as a starting point for further explorations in the literature.

ACKNOWLEDGMENTS

The authors are grateful to Alessandro Barducci, Massimiliano Bonomi, Michele Ceriotti, Alessandro Laio, and Fabrizio Marinelli for carefully reading the manuscript and providing several useful suggestions.

APPENDIX A: METADYNAMICS INPUT FILES WITH PLUMED

In this section we want to introduce some typical input files to perform MetaD and WTMetaD by using the PLUMED plugin in its most recent version 2.0.[26] PLUMED is an open-source software that consists of a library enabling users to perform various enhanced sampling methods (among which is MetaD). PLUMED is interfaced and can be combined with a variety of MD codes, including GROMACS,[4] NAMD,[91] LAMMPS,[92] and Quantum-ESPRESSO.[93] These interfaces allow a modeler to use several free-energy methods in a variety of fields thus fostering the applicability, dissemination, and cross-validation of those techniques, at various levels of theory. Moreover, PLUMED provides a single executable that allows PLUMED to be used as a stand-alone tool.

For the examples carried out in Section "Metadynamics How-To" we used the GROMACS 4.5.5[4] classical MD engine. In all the following examples energies are reported in kcal/mol and distances in Å.

PLUMED is not provided with GROMACS code but is separately downloadable from http://www.plumed-code.org. It requires the user to apply a patching procedure and to recompile the MD software according to the procedure reported in its manual. After doing that, GROMACS is simply instructed to use PLUMED on the command line (e.g., `mdrun -plumed input.dat`).

The first example refers to the case where MetaD is performed on gyration radius reported in Section "A First Test Run Using Gyration Radius." Note that in PLUMED one can also monitor other variables that might be helpful in checking the results of the simulation (in this case Φ and Ψ dihedral angles):

Example

```
# choose units - by default PLUMED uses kj/mol, nm, and ps
UNITS ENERGY=kcal/mol LENGTH=A
# these three lines define the atoms involved in the group
# named "all"
# the name of the group is arbitrary. If not defined PLUMED
```

```
# gives a default
all: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
# this defines the gyration radius using a group
# defined by atoms in "all" and this value will have name "rg"
rg: GYRATION ATOMS=all
# this line sets the metadynamics hills height (HEIGHT) in
# kcal/mol
# and deposition time (PACE) in timestep. Here each
# timestep is 2 fs.
# this produces also a HILLS file containing the
# center and widths of Gaussians
meta: METAD ARG=rg SIGMA=0.07 HEIGHT=0.1 PACE=600
# additional variables can be set to monitor the simulation
# Phi:
t1: TORSION ATOMS=5,7,9,15
# and Psi:
t2: TORSION ATOMS=7,9,15,17
# This produces a COLVAR file that contains the
# values of the variables t1 and t2
# calculated every STRIDE steps, which can be more often than
# HILLS
PRINT ARG=t1,t2 STRIDE=100 FILE=COLVAR
```

At the end of the run one obtains three output files from PLUMED: a logfile, a COLVAR file that is produced by the `PRINT` command, and a HILLS file that is produced by the `METAD`. The PLUMED logfile in GROMACS is embedded in the `md.log` file and the user is strongly encouraged to explore it to understand if PLUMED has interpreted the commands correctly. The COLVAR file reports the time series of the CVs that are included in the input, and the HILLS file reports the centers, the widths, and the heights of the repulsive Gaussian potentials added by MetaD. At the end of the simulation one can calculate the sum of the deposited potential by using the standalone executable of PLUMED with the command `plumed sum_hills --hills HILLS`.

In the second example we report the input from Section "A Better Collective Variable: Φ Dihedral Angle" where the Φ dihedral angle was used. Please note that adding more arguments in the `ARG` field and corresponding values in `SIGMA` allows PLUMED to perform multidimensional metadynamics.

Example

```
UNITS ENERGY=kcal/mol LENGTH=A
all: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
rg: GYRATION ATOMS=all
t1: TORSION ATOMS=5,7,9,15
# now METAD takes t1 as an argument and the sigma is changed
# according to its fluctuation
```

```
# note that you can give as ARG only a CV which is already
# defined
meta: METAD ARG=t1 SIGMA=0.1 HEIGHT=0.1 PACE=600
# this variable again is only monitored in COLVAR
t2: TORSION ATOMS=7,9,15,17
PRINT ARG=t1,t2 STRIDE=100 FILE=COLVAR
```

Note that the `TORSION` collective variable is now periodic. The `sum_hills` tool of PLUMED knows automatically that the domain is periodic.

As a last example we report the use of WTMetaD as discussed in Section "Well-Tempered Metadynamics Using Dihedral Angle Φ."

Example

```
UNITS ENERGY=kcal/mol LENGTH=A
all: GROUP ATOMS=1,5,6,7,9,11,15,16,17,19
rg: GYRATION ATOMS=all
t1: TORSION ATOMS=5,7,9,15
# this is a multiple line command
# BIASFACTOR and TEMP are required to enable
# well-tempered metad
METAD ...
    LABEL=meta
    ARG=t1 SIGMA=0.1 HEIGHT=0.1 PACE=600
    BIASFACTOR=15 TEMP=300
... METAD
t2: TORSION ATOMS=7,9,15,17
PRINT ARG=t1,t2 STRIDE=100 FILE=COLVAR
```

More information regarding PLUMED and the various tools included in it along with some more tutorial examples can be retrieved from http://www.plumed-code.org.

REFERENCES

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, 1987.
2. D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Academic Press, London, second ed., 2002.
3. M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, Oxford, 2012.
4. B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J. Chem. Theory Comput.*, **4**(3), 435–447 (2008). GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation.

5. D. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. Dror, M. Eastwood, J. Bank, J. Jumper, J. Salmon, Y. Shan, and W. Wriggers, *Science*, **330**, 341–346 (2010). Atomic-Level Characterization of the Structural Dynamics of Proteins.
6. K. Lindorff-Larsen, S. Piana, R. Dror, and D. Shaw, *Science*, **334**, 517–520 (2011). How Fast-Folding Proteins Fold.
7. Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, **314**(1–2), 141–151 (1999). Replica-Exchange Molecular Dynamics Method for Protein Folding.
8. P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, *Proc. Natl. Acad. Sci. U. S. A.*, **102**(39), 13749–13754 (2005). Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water.
9. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, **23**(2), 187–199 (1977). Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling.
10. E. Darve and A. Pohorille, *J. Chem. Phys.*, **115**(20), 9169–9183 (2001). Calculating Free Energies Using Average Force.
11. A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **99**(20), 12562–12566 (2002). Escaping Free-Energy Minima.
12. S. Marsili, A. Barducci, R. Chelli, P. Procacci, and V. Schettino, *J. Phys. Chem. B*, **110**(29), 14011–14013 (2006). Self-Healing Umbrella Sampling: A Non-Equilibrium Approach for Quantitative Free Energy Calculations.
13. J. VandeVondele and U. Rothlisberger, *J. Phys. Chem. B*, **106**(1), 203–208 (2002). Canonical Adiabatic Free Energy Sampling (CAFES): A Novel Method for the Exploration of Free Energy Surfaces.
14. L. Rosso, P. Minary, Z. Zhu, and M. E. Tuckerman, *J. Chem. Phys.*, **116**(11), 4389–4402 (2002). On the Use of the Adiabatic Molecular Dynamics Technique in the Calculation of Free Energy Profiles.
15. L. Maragliano and E. Vanden-Eijnden, *Chem. Phys. Lett.*, **426**, 168–175 (2006). A Temperature Accelerated Method for Sampling Free Energy and Determining Reaction Pathways in Rare Events Simulations.
16. M. Iannuzzi, A. Laio, and M. Parrinello, *Phys. Rev. Lett.*, **90**(23), 238302 (2003). Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics.
17. G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello, *J. Am. Chem. Soc.*, **128**(41), 13435–13441 (2006). Free-Energy Landscape for β-Hairpin Folding from Combined Parallel Tempering and Metadynamics.
18. F. Baftizadeh, X. Biarnes, F. Pietrucci, F. Affinito, and A. Laio, *J. Am. Chem. Soc.*, **134**(8), 3886–3894 (2012). Multidimensional View of Amyloid Fibril Nucleation in Atomistic Detail.
19. F. L. Gervasio, A. Laio, and M. Parrinello, *J. Am. Chem. Soc.*, **127**(8), 2600–2607 (2005). Flexible Docking in Solution Using Metadynamics.
20. R. Martoňák, D. Donadio, A. Oganov, and M. Parrinello, *Nat. Mater.*, **5**, 623–626 (2006). Crystal Structure Transformations in SiO_2 from Classical and Ab Initio Metadynamics.
21. F. Trudu, D. Donadio, and M. Parrinello, *Phys. Rev. Lett.*, **97**(10), 105701 (2006). Freezing of a Lennard-Jones Fluid: From Nucleation to Spinodal Regime.
22. CPMD, http://www.cpmd.org/, Copyright IBM Corp 1990–2008, Copyright MPI für Festkörperforschung Stuttgart 1997–2001, 1990–2008.
23. CP2K, http://www.cp2k.org/, 2012.
24. S. Marsili, G. Signorini, R. Chelli, M. Marchi, and P. Procacci, *J. Comput. Chem.*, **31**(5), 1106–1116 (2009). ORAC: A Molecular Dynamics Simulation Program to Explore Free Energy Surfaces in Biomolecular Systems at the Atomistic Level.
25. M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, *Comput. Phys. Comm*, **180**(10), 1961–1972 (2009). PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics.

26. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, *Comput. Phys. Commun.*, **185**(2), 604–613 (2014). PLUMED 2: New Feathers For An Old Bird.
27. G. Fiorin, M. L. Klein, and J. Hénin, *Mol. Phys.*, **111**(22–23), 3345–3362 (2013). Using Collective Variables To Drive Molecular Dynamics Simulations.
28. W. Humphrey, A. Dalke, and K. Schulten, *J. Molec. Graphics*, **14**, 33–38 (1996). VMD – Visual Molecular Dynamics.
29. A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, W. Watanabe, J. Wiorkiewicz-Kunczera, D. Yin, and M. Karplus, *J. Phys. Chem. B*, **102**(18), 3586–3616 (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins.
30. T. Lazaridis, D. J. Tobias, C. Brooks, and M. E. Paulaitis, *J. Chem. Phys.*, **95**(10), 7612–7625 (1991). Reaction Paths and Free Energy Profiles for Conformational Transitions: An Internal Coordinate Approach.
31. D. J. Tobias and C. L. Brooks, *J. Phys. Chem.*, **96**(9), 3864–3870 (1992). Conformational Equilibrium in the Alanine Dipeptide in the Gas Phase and Aqueous Solution: A Comparison of Theoretical Results.
32. C. Bartels and M. Karplus, *J. Comput. Chem.*, **18**(12), 1450–1462 (1997). Multidimensional Adaptive Umbrella Sampling: Application to Main Chain and Side Chain Peptide Conformations.
33. L. Maragliano, A. Fischer, and E. Vanden-Eijnden, *J. Chem. Phys.*, **125**(2), 024106 (2006). String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces.
34. D. Branduardi, F. L. Gervasio, and M. Parrinello, *J. Chem. Phys.*, **126**(5), 054103 (2007). From A to B in Free Energy Space.
35. D. Branduardi, G. Bussi, and M. Parrinello, *J. Chem. Theory Comput.*, **8**(7), 2247–2254 (2012). Metadynamics with Adaptive Gaussians.
36. M. Mezei, *J. Comput. Phys.*, **68**(1), 237–248 (1987). Adaptive Umbrella Sampling: Self-Consistent Determination of the Non-Boltzmann Bias.
37. H. Grubmüller, *Phys. Rev. E*, **52**(3), 2893 (1995). Predicting Slow Structural Transitions in Macromolecular Systems: Conformational Flooding.
38. A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, *J. Phys. Chem. B*, **109**(14), 6714–6721 (2005). Assessing the Accuracy of Metadynamics.
39. C. W. Gear, I. G. Kevrekidis, and C. Theodoropoulos, *Comput. Chem. Eng.*, **26**, 941–963 (2002). "Coarse" Integration/Bifurcation Analysis via Microscopic Simulators: Micro-Galerkin Methods.
40. D. Cvijovic and J. Klinowski, *Science*, **267**, 664–666 (1995). Taboo Search - An Approach To The Multiple Minima Problem.
41. T. Huber, A. E. Torda, and W. F. van Gunsteren, *J. Comput.-Aided Mol. Des.*, **8**, 695–708 (1994). Local Elevation: A Method for Improving the Searching Properties of Molecular Dynamics Simulation.
42. G. Bussi, A. Laio, and M. Parrinello, *Phys. Rev. Lett.*, **96**(9), 090601 (2006). Equilibrium Free Energies from Nonequilibrium Metadynamics.
43. C. Micheletti, A. Laio, and M. Parrinello, *Phys. Rev. Lett.*, **92**(17), 170601 (2004). Reconstructing the Density of States by History-Dependent Metadynamics.
44. Y. Crespo, F. Marinelli, F. Pietrucci, and A. Laio, *Phys. Rev. E*, **81**(5), 055701 (2010). Metadynamics Convergence Law in a Multidimensional System.
45. F. Baftizadeh, P. Cossio, F. Pietrucci, and A. Laio, *Curr. Phys. Chem.*, **2**, 79–91 (2012). Protein Folding and Ligand-Enzyme Binding from Bias-Exchange Metadynamics Simulations.
46. F. Marinelli, F. Pietrucci, A. Laio, and S. Piana, *PLoS Comput. Biol.*, **5**(8), e1000452 (2009). A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations.

47. S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen, and P. Kollman, *J. Comput. Chem.*, **13**(8), 1011–1021 (1992). The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method.
48. A. Barducci, G. Bussi, and M. Parrinello, *Phys. Rev. Lett.*, **100**(2), 020603 (2008). Well-Tempered Metadynamics: A Smoothly-Converging and Tunable Free-Energy Method.
49. M. Bonomi, A. Barducci, and M. Parrinello, *J. Comput. Chem.*, **30**(11), 1615–1621 (2009). Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics.
50. S. K. Kearsley, *Acta Cryst. A*, **45**, 208–210 (1989). On the Orthogonal Transformation Used for Structural Comparison.
51. B. Ensing and M. L. Klein, *Proc. Natl. Acad. Sci. U. S. A.*, **102**(19), 6755–6759 (2005). Perspective on the Reactions Between F^- and CH_3CH_2F: The Free Energy Landscape of the E2 and S_N2 Reaction Channels.
52. E. Vanden-Eijnden and M. Venturoli, *J. Chem. Phys.*, **130**(19), 194103 (2009). Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies.
53. Y. A. Mantz, D. Branduardi, G. Bussi, and M. Parrinello, *J. Phys. Chem. B*, **113**(37), 12521–12529 (2009). Ensemble of Transition State Structures for the Cis-Trans Isomerization of N-Methylacetamide.
54. A. Lodola, D. Branduardi, M. De Vivo, L. Capoferri, M. Mor, D. Piomelli, and A. Cavalli, *PLoS One*, **7**(2), e32397 (2012). A Catalytic Mechanism for Cysteine N-Terminal Nucleophile Hydrolases, as Revealed by Free Energy Simulations.
55. D. Branduardi, M. De Vivo, N. Rega, V. Barone, and A. Cavalli, *J. Chem. Theory Comput.*, **7**(3), 539–543 (2011). Methyl Phosphate Dianion Hydrolysis in Solution Characterized by Path Collective Variables Coupled with DFT-Based Enhanced Sampling Simulations.
56. G. A. Gallet, F. Pietrucci, and W. Andreoni, *J. Chem. Theory Comput.*, **8**(11), 4029–4039 (2012). Bridging Static and Dynamical Descriptions of Chemical Reactions: An Ab Initio Study of CO2 Interacting with Water Molecules.
57. M. Bonomi, D. Branduardi, F. L. Gervasio, and M. Parrinello, *J. Am. Chem. Soc.*, **130**(42), 13938–13944 (2008). The Unfolded Ensemble and Folding Mechanism of the C-Terminal GB1 Beta-Hairpin.
58. A. Pietropaolo, L. Muccioli, R. Berardi, and C. Zannoni, *Proteins*, **70**(3), 667–677 (2008). A Chirality Index for Investigating Protein Secondary Structures and Their Time Evolution.
59. W. Ren, E. Vanden-Eijnden, and P. Maragakis, *J. Chem. Phys.*, **123**(13), 134109 (2005). Transition Pathways in Complex Systems: Application of the Finite-Temperature String Method to the Alanine Dipeptide.
60. G. Daz Leines and B. Ensing, *Phys. Rev. Lett.*, **109**(2), 020601 (2012). Path Finding on High-Dimensional Free Energy Landscapes.
61. A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, *Proteins*, **17**(4), 412–425 (1993). Essential Dynamics of Proteins.
62. V. Spiwok, P. Lipovová, and B. Králová, *J. Phys. Chem. B*, **111**(12), 3073–3076 (2007). Metadynamics in Essential Coordinates: Free Energy Simulation of Conformational Changes.
63. L. Sutto, M. D'Abramo, and F. L. Gervasio, *J. Chem. Theory Comput.*, **6**(12), 3640–3646 (2010). Comparing the Efficiency of Biased and Unbiased Molecular Dynamics in Reconstructing the Free Energy Landscape of Met-Enkephalin.
64. G. A. Tribello, M. Ceriotti, and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **107**(41), 17509–17414 (2010). A Self-Learning Algorithm for Biased Molecular Dynamics.
65. I. Borg and P. Groenen, *Modern Multidimensional Scaling*, Springer, New York, 2005.
66. V. Spiwok and B. Králová, *J. Chem. Phys.*, **135**(22), 224504 (2011). Metadynamics in the Conformational Space Nonlinearly Dimensionally Reduced by Isomap.
67. P. Das, M. Moll, H. Stamati, L. E. Kavraki, and C. Clementi, *Proc. Natl. Acad. Sci. U. S. A.*, **103**(26), 9885–9890 (2006). Low-Dimensional, Free-Energy Landscapes of Protein-Folding Reactions by Nonlinear Dimensionality Reduction.

68. M. Ceriotti, G. A. Tribello, and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **108**(32), 13023–13028 (2011). Simplifying the Representation of Complex Free-Energy Landscapes Using Sketch-Map.
69. G. A. Tribello, M. Ceriotti, and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **109**(14), 5196–5201 (2012). Using Sketch-Map Coordinates to Analyze and Bias Molecular Dynamics Simulations.
70. F. Pietrucci and A. Laio, *J. Chem. Theory Comput.*, **5**(9), 2197–2201 (2009). A Collective Variable for the Efficient Exploration of Protein Beta-Sheet Structures: Application to SH3 and GB1.
71. A. Cuff, O. C. Redfern, L. Greene, I. Sillitoe, T. Lewis, M. Dibley, A. Reid, F. Pearl, T. Dallman, A. Todd, R. Garratt, J. Thornton, and C. Orengo, *Structure*, **17**(8), 1051–1062 (2009). The CATH Hierarchy Revisited – Structural Divergence in Domain Superfamilies and the Continuity of Fold Space.
72. D. Frishman and P. Argos, *Proteins: Structure, Function, and Bioinformatics*, **23**(4), 566–579 (1995). Knowledge-Based Protein Secondary Structure Assignment.
73. P. Cossio, A. Trovato, F. Pietrucci, F. Seno, A. Maritan, and A. Laio, *PLoS Comput. Biol.*, **6**(11), e1000957 (2010). Exploring the Universe of Protein Structures Beyond the Protein Data Bank.
74. C. Bartels and M. Karplus, *J. Phys. Chem. B*, **102**(5), 865–880 (1998). Probability Distributions for Complex Systems: Adaptive Umbrella Sampling of the Potential Energy.
75. B. Berg and T. Neuhaus, *Phys. Rev. Lett.*, **68**(1), 9–12 (1992). Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions.
76. F. Wang and D. P. Landau, *Phys. Rev. Lett.*, **86**(10), 2050–2053 (2001). Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States.
77. M. Bonomi and M. Parrinello, *Phys. Rev. Lett.*, **104**(19), 190601 (2010). Enhanced Sampling in the Well-Tempered Ensemble.
78. M. Deighan, M. Bonomi, and J. Pfaendtner, *J. Chem. Theory Comput.*, **8**(7), 2189–2192 (2012). Efficient Simulation of Explicitly Solvated Proteins in the Well-Tempered Ensemble.
79. T. N. Do, P. Carloni, G. Varani, and G. Bussi, *J. Chem. Theory Comput.*, **9**(3), 1720–1730 (2013). RNA/Peptide Binding Driven by Electrostatics – Insight From Bidirectional Pulling Simulations.
80. P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, *J. Phys. Chem. B*, **110**(8), 3533–3539 (2006). Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics.
81. C. Melis, P. Raiteri, L. Colombo, and A. Mattoni, *ACS Nano*, **5**(12), 9639–9647 (2011). Self-Assembling of Zinc Phthalocyanines on ZnO (1010) Surface through Multiple Time Scales.
82. H. Fukunishi, O. Watanabe, and S. Takada, *J. Chem. Phys.*, **116**(20), 9058–9067 (2002). On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction.
83. C. Camilloni, D. Provasi, G. Tiana, and R. Broglia, *Proteins*, **71**(4), 1647–1654 (2007). Exploring the Protein G Helix Free-Energy Surface by Solute Tempering Metadynamics.
84. S. Piana and A. Laio, *J. Phys. Chem. B*, **111**(17), 4553–4559 (2007). A Bias-Exchange Approach to Protein Folding.
85. L. Zheng, M. Chen, and W. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, **105**(51), 20227–20232 (2008). Random Walk in Orthogonal Space to Achieve Efficient Free-Energy Simulation of Complex Systems.
86. S. Singh, C. Chiu, and J. de Pablo, *J. Stat. Phys.*, **145**(4), 932–945 (2011). Flux Tempered Metadynamics.
87. M. Chen, M. Cuendet, and M. Tuckerman, *J. Chem. Phys.*, **137**(2), 024102 (2012). Heating and Flooding: A Unified Approach for Rapid Generation of Free Energy Surfaces.

88. Q. Zhu, A. Oganov, and A. Lyakhov, *Cryst. Eng. Commun.*, **14**(10), 3596–3601 (2012). Evolutionary Metadynamics: A Novel Method to Predict Crystal Structures.

89. P. Tiwary and M. Parrinello, *Phys. Rev. Lett.*, **111**(23), 230602 (2013). From Metadynamics to Dynamics.

90. M. McGovern and J. de Pablo, *J. Chem. Phys.*, **139**, 084102 (2013). A Boundary Correction Algorithm for Metadynamics in Multiple Dimensions.

91. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *J. Comput. Chem.*, **26**(16), 1781–802 (2005). Scalable Molecular Dynamics with NAMD.

92. S. Plimpton, *J. Comput. Phys.*, **117**(1), 1–19 (1995). Fast Parallel Algorithms for Short-Range Molecular Dynamics.

93. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, *J. Phys.: Cond. Matter*, **21**(39), 395502 (2009). Quantum Espresso: A Modular And Open-Source Software Project For Quantum Simulations Of Materials.