Published in final edited form as:

Mol Simul. 2016 ; 42(13): 1046–1055. doi:10.1080/08927022.2015.1121541.

Unconstrained Enhanced Sampling for Free Energy Calculations of Biomolecules: A Review

Yinglong Miao[1,2] and **J. Andrew McCammon**[1,2,3]

Yinglong Miao: yimiao@ucsd.edu

[1]Howard Hughes Medical Institute, University of California San Diego, La Jolla, CA 92093

[2]Department of Pharmacology, University of California San Diego, La Jolla, CA 92093

[3]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093

Abstract

Free energy calculations are central to understanding the structure, dynamics and function of biomolecules. Yet insufficient sampling of biomolecular configurations is often regarded as one of the main sources of error. Many enhanced sampling techniques have been developed to address this issue. Notably, enhanced sampling methods based on biasing collective variables (CVs), including the widely used umbrella sampling, adaptive biasing force and metadynamics, have been discussed in a recent excellent review (Abrams and Bussi, Entropy, 2014). Here, we aim to review enhanced sampling methods that do not require predefined system-dependent CVs for biomolecular simulations and as such do not suffer from the hidden energy barrier problem as encountered in the CV-biasing methods. These methods include, but are not limited to, replica exchange/parallel tempering, self-guided molecular/Langevin dynamics, essential energy space random walk and accelerated molecular dynamics. While it is overwhelming to describe all details of each method, we provide a summary of the methods along with the applications and offer our perspectives. We conclude with challenges and prospects of the unconstrained enhanced sampling methods for accurate biomolecular free energy calculations.

Keywords

Biomolecules; Enhanced Sampling; Unconstrained; Free Energy

1. Introduction

Ever since the first simulation of bovine pancreatic trypsin inhibitor (BPTI) in 1977[1], molecular dynamics (MD) has embraced a new era for its wide applications in computational modeling of biomolecules[2]. The simulated systems range from proteins to membrane lipids and nucleic acids. Based on a molecular force field model, MD solves the Newton's equations of motion for all atoms in the system along time. From the atomic trajectory, one is able to extract structural and dynamical information on the biomolecules that can be compared and used to interpret experimental data obtained from nuclear magnetic resonance (NMR)[3–5] and neutron scattering[6, 7]. MD simulations have been

also found use in refinement of structures determined by cryo-electron microscopy[8, 9], X-ray crystallography[10–12] and NMR[13, 14].

In addition, MD simulations have been used to calculate free energies that determine the behavior of biomolecules at or near equilibrium. In the simplest form, the free energy of biomolecules in a canonical ensemble is calculated as $F = -k_B T \ln Q$, where k_B is the Boltzmann's constant, T is the system temperature and Q is the system partition function. Free energies describe the spontaneity of biomolecular transitions from one configuration state to another. They are central to understanding the structure, dynamics and function of the biomolecules. Examples include ligand binding to proteins, conformational change of large proteins during enzyme catalysis or cell signaling, and permeation of small-molecule drugs through cell membrane as discussed in a long list of previous review articles[15–18].

Whereas MD simulations provide a detailed picture of the biomolecular dynamics in atomistic detail, current MD simulations are often limited to tens of microseconds for average-sized proteins or even shorter timescales for larger systems[19]. Different functional states of biomolecules are often separated by high energy barriers (e.g., 8–12 kcal/mol) and structural transitions across the energy barriers take place over milliseconds or even longer timescales[20, 21]. Therefore, there remains a large gap between the MD simulation timescales and those of biological processes. This often leads to insufficient sampling or unconverged simulations of the biomolecules. It has been described as one of the main sources of error in biomolecular free energy calculations[15, 17, 18].

2. Why unconstrained enhanced sampling simulations are needed

To overcome the insufficient sampling problem of free energy calculations, many enhanced sampling techniques have been developed during the last several decades as reviewed earlier[22–27]. The first class of these methods is based on biasing collective variables (CVs) of biomolecules. They include the widely used umbrella sampling[28, 29], blue moon sampling[30], adaptive biasing force[31, 32], metadynamics[33, 34], conformational flooding[35], string method[36, 37], λ-dynamics[38], adiabatic free energy dynamics[39], temperature accelerated MD[40] and targeted MD[41–43].

The above enhanced sampling methods greatly improve the convergence of free energy calculations, but they require predefined CVs and often suffer from hidden energy barriers. In Figure 1, a hidden energy barrier is illustrated for enhanced sampling simulations that are based on biasing CVs. Without the loss of generality, there are two local energy minimum states for a given system along the y direction. If only x is defined as the CV, apparently converged sampling of the free energy surface along the x direction will be obtained from the corresponding enhanced sampling simulations, but not for the hidden energy barrier along the y direction. Therefore, when important CVs are missed during the simulation setup, the CV-biasing enhanced sampling methods often suffer from insufficient sampling of the hidden energy barrier(s), which leads to poor convergence of the simulations. Abrams and Bussi noted in their recent review[22] that: "*the drawback of all CV-biasing approaches is the risk that the chosen CV space does not provide the most faithful representation of the true spectrum of metastable subensembles and the barriers that separate them. Guaranteeing*

that sampling of CV space is not stymied by hidden barriers must be of paramount concern in the continued evolution of such methods".

Improved enhanced sampling methodology has been developed to address the "hidden barrier" issue, notably the orthogonal space random walk (OSRW)[44, 45], the essence of which is to perform a simultaneous random walk in the orthogonal space along the directions of a CV and its generalized force. Moreover, there exists a second class of enhanced sampling methods that do not require predefined system-dependent CVs and as such do not suffer from the hidden barrier problem as encountered in the CV-biasing methods, such as replica exchange (RE)[46, 47] or parallel tempering (PT)[48], self-guided molecular/Langevin dynamics (SGMD and SGLD)[49–52], essential energy space random walk (EESRW)[53–55] and accelerated molecular dynamics (aMD)[56, 57]. Compared with the CV-biasing enhanced sampling methods, these methods are advantageous to explore biomolecular configurations without the need for *a priori* knowledge. They provide unconstrained enhanced sampling to explore biomolecular structural transition pathways (e.g., protein folding and ligand binding) and identify possible unknown intermediate configuration states of the biomolecules.

3. Unconstrained enhanced sampling methods

Here, we consider a list of unconstrained enhanced sampling methods, including the RE/PT, SGMD/SGLD, EESRW and aMD. It is not possible to describe all details of each method in this review. Instead, we provide a summary of the methods along with their applications and offer our perspectives.

3.1 Replica Exchange/Parallel Tempering

As illustrated in Figure 2, Replica Exchange/Parallel Tempering (RE/PT) methods are based on running multiple copies (replicas) of parallel MD (or Monte Carlo) simulations at different temperatures and exchanging system configurations (or equivalently exchanging temperatures) between pairs of replicas at a certain frequency to achieve enhanced sampling. The replicas are exchanged between each pair of neighboring temperatures at a probability that meets the Metropolis criterion. It ensures that the generalized ensemble of the system satisfies a detailed balance and converges towards an equilibrium distribution. Meanwhile, the replica exchanges avoid the trapping of biomolecules in local energy-minimum configurations as simulated at low, room temperature and facilitate barrier crossing events in the biomolecules.

Consider a system of N atoms with positions $r \equiv \{\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N\}$ and momenta $p \equiv \{\boldsymbol{p}_1, \cdots, \boldsymbol{p}_N\}$. In the canonical ensemble at temperature T, each state $x \equiv (r,p)$ is weighted by the Boltzmann factor:

$$W_B(x;T) = \exp\{-\beta H(r,p)\}, \quad (1)$$

where $\beta = 1/k_B T$ with k_B as the Boltzmann constant and $H(r,p)$ is the system Hamiltonian that is given by the sum of the potential energy $V(r)$ and kinetic energy $K(p)$.

In RE/PT, the generalized ensemble consists of M non-interacting replicas of the original system in the canonical ensemble, each at a different temperature T_m ($m = 1, \cdots, M$). There is a one-to-one correspondence between the replicas and temperatures. The label i ($i = 1, \cdots, M$) for replicas is a permutation of the label m ($m = 1, \cdots, M$) for temperatures, *i.e.*, $i = i(m) \equiv f(m)$, and vice versa $m = m(i) \equiv f^{-1}(i)$. $f(m)$ is a permutation function of m and $f^{-1}(i)$ is its inverse. Let $X = \left(x_1^{[i(1)]}, \cdots, x_M^{[i(M)]}\right)$ stand for a "state" in the generalized ensemble. Each "substate" $x_m^{[i]}$ is specified by the coordinates $r^{[i]}$ and momenta $p^{[i]}$ of N atoms in replica i at temperature T_m: $x_m^{[i]} = \left(r^{[i]}, p^{[i]}\right)_m$. Because the replicas are non-interacting, the weight factor for the state X in this generalized ensemble is given by the product of the Boltzmann factors of each replica:

$$W_{RE}(X) = \prod_{i=1}^{M} \exp\left\{-\beta_{m(i)} H\left(x_{m(i)}^{[i]}\right)\right\} = \exp\left\{-\sum_{i=1}^{M} \beta_{m(i)} H\left(x_{m(i)}^{[i]}\right)\right\}. \quad (2)$$

To help the biomolecules escape from local energy minimum states, we exchange a pair of replicas with neighboring temperatures at probability $w(X \to X')$. Such replica exchange needs to satisfy a detailed balance in order to converge towards an equilibrium distribution:

$$W_{RE}(X) w(X \to X') = W_{RE}(X') w(X' \to X). \quad (3)$$

This can be met when one uses the Metropolis criterion to calculate the exchange probability:

$$w(X \to X') \equiv w\left(x_m^{[i]} | x_n^{[j]}\right) = \min(1, \exp(-\Delta)), \quad (4)$$

where $\Delta = (\beta_n - \beta_m)[V(r^{[i]}) - V(r^{[j]})]$, and i, j, m and n are related by the permutation functions before the exchange: $i = f(m)$, $j = f(n)$.

Compared with the simulated tempering (ST)[58, 59] and mulicanonical algorithms (MUCA)[60] that perform random walk in the temperature and potential energy spaces, respectively, the RE/PT methods are advantageous with *a priori* known weight factors. Canonical ensemble properties are readily obtained from the RE/PT simulations. Many variants of the RE/PT methods have been also developed for further enhanced sampling[47], such as the replica-exchange multicanonical algorithm[61, 62], replica-exchange simulated tempering[63], multicanonical replica-exchange[61], simulated tempering replica-exchange[63] and *ab initio* replica-exchange[64]. Moreover, multidimensional extension of REMD, multidimensional replica-exchange[65], has led to the development of the replica exchange umbrella sampling[66] and Hamiltonian REMD[67, 68] methods.

Instead of using spatial CVs, the RE/PT methods are based on exchanging temperature between parallel MD or Monte Carlo simulations. This allows for unconstrained

conformational sampling of the biomolecules, which has enabled a wide range of applications, for example, in conformational transitions of Argon fluid[69] and peptides[70, 71], protein folding[46, 72–74], protein structural refinement[75], ligand binding free energy calculations[65], DNA base stacking[76], nucleic acid conformational change[67], etc.

However, it has been noted that the RE/PT methods do not work for processes with the first-order phase transitions. Moreover, the number of replicas required for converged RE/PT simulations is proportional with the square root of the system number of degrees of freedom. This suggests that for systems of increasing size, a significantly larger number of replicas are needed are needed to run the simulations. This often requires expensive computational resources. One solution for this is to first run a short RE/PT simulation to determine the mulitcanonical weight factor and then perform with this weight factor a regular multicanonical simulation with high statistics. This is reflected in the replica-exchange multicanonical algorithm[61, 62] and replica-exchange simulated tempering method[63, 77].

3.2 Self-guided molecular/Langevin Dynamics

Self-guided molecular dynamics (SGMD)[49, 50] was developed to improve the conformational searching efficiency of MD simulations. It is based on the rationale that the conformational searching efficiency is often limited by slow systematic motions. To accelerate these systematic motions, a self-guiding force was introduced into the Newton's equation of motion:

$$m_i\ddot{\boldsymbol{r}}_i=\boldsymbol{f}_i+\boldsymbol{g}_i, \quad (5)$$

where m_i is the mass of atom i, $\ddot{\boldsymbol{r}}_i$ is the acceleration of atom i with altered systematic motion, $\boldsymbol{f}_i$ is the atom instantaneous force derived from force field, $\boldsymbol{g}_i$ is the guiding force calculated as:

$$\boldsymbol{g}_i=\lambda\frac{1}{t_1}\int_{t-t_1}^{t}(\boldsymbol{f}_i+\boldsymbol{g}_i)d\tau, \quad (6)$$

where λ is a guiding factor and t_L is the local averaging time. Moreover, bonded substructures were defined in SGMD to describe the local structural rigidity of a molecule[50]. It was suggested that most of the atomic motions due to bonded interactions within each substructure are fast. In contrast, systematic motions that result from non-bonded interactions between atoms of different substructures are generally slow. Thus the self-guiding force was applied to enhance these systematic motions only. With this, SGMD was shown to greatly improve the sampling of conformational transitions in the alanine dipeptide and a 16-residue synthetic polypeptide that undergoes conformational change between a fully extended conformation and a folded helix[50].

However, several drawbacks were found during applications of SGMD. Among those noted by Wu and Brooks[51], the self-guiding force results in an unwanted alteration of the conformational distribution in SGMD simulations. Moreover, the guiding force derived from

the local average of actual forces may not be sufficient to enhance conformational searching in stochastic dynamics simulations. To address these issues, the local average of friction forces that appear in a Langevin dynamics (LD) simulation was proposed to calculate the guiding force instead, which led to the development of the self-guided Langevin dynamics (SGLD) method. In SGLD, the self-guiding force was introduced into the Langevin equation:

$$\dot{\boldsymbol{p}}_i = \boldsymbol{f}_i + \boldsymbol{g}_i - \gamma_i \boldsymbol{p}_i + \boldsymbol{R}_i, \quad (7)$$

where $\boldsymbol{p}_i$ is the momentum of atom i, $\boldsymbol{f}_i$ is the atom instantaneous force derived from the force field, γ_i is the collision frequency and $\boldsymbol{R}_i$ is the random force that is related to γ_i and the system temperature T by $\langle \boldsymbol{R}_i(0)\, \boldsymbol{R}_i(t) \rangle = 2 m_i k_B T \gamma_i \delta(t)$, where k_B is the Boltzmann constant and $\delta(t)$ is the Dirac delta function. In the case that the time derivative of momentum can be dropped from the Langevin equation (*i.e.*, high friction regime), the guiding force $\boldsymbol{g}_i$ is calculated as:

$$\boldsymbol{g}_i = \lambda \gamma_i \langle \boldsymbol{p}_i \rangle_L - \xi \gamma_i \boldsymbol{p}_i, \quad (8)$$

where λ is a guiding factor that controls the strength of the guiding force. $\langle \boldsymbol{p}_i \rangle_L$ denotes local average of $\boldsymbol{p}_i$ over time t_L and can be calculated efficiently during simulation via:

$$\langle \boldsymbol{p}_i \rangle_L(t) \approx \left(1 - \frac{\partial t}{t_L}\right) \langle \boldsymbol{p}_i \rangle_L(t - \partial t) + \frac{\partial t}{t_L} \boldsymbol{p}_i(t). \quad (9)$$

The local averaging acts like a low frequency filter that reduces the high frequency components of the motion while keeping its low frequency contributions. The parameter, ξ, is an energy conservation factor used to cancel net energy input from the guiding force:

$$\xi = \frac{\lambda \sum_i \langle \boldsymbol{p}_i \rangle_L \cdot \dot{\boldsymbol{r}}_i}{\sum_i \boldsymbol{p}_i \cdot \dot{\boldsymbol{r}}_i}.$$

According to Eqn. (7), when the collision frequency and random force are equal to zero, SGLD is then reduced to SGMD. In this context, MD and LD are special cases of SGMD and SGLD, respectively, when the guiding force is zero. Up to date a list of SGLD variants has been developed using different formula of the guiding force (see Ref. [78] for details). Particularly, SGMD calculates the guiding force with nonbonded interactions as described above [49, 50], SGLD uses momenta[51] and is also referred to as SGLDp, and SGLDfp uses both forces and momenta[79]. In the most recent development of SGLD, the self-guiding force is modified according to a generalized Langevin equation using $\boldsymbol{g}_i = \lambda \gamma_i \langle \boldsymbol{p}_i \rangle_L - \left(1 - \sqrt{1 - \lambda}\right) \langle \boldsymbol{R}_i \rangle_L$ to sample exactly the *NVT* and *NPT* ensembles, i.e., SGLD-GLE[52].

Along with the theoretical development of the SGMD and SGLD methods, they have been applied for enhanced sampling of a wide range of systems, such as peptide conformational change[50], surface adsorption of complementary peptides[80], flexible fitting of biomolecular structures into electron microscopic density maps[81], protein folding and conformational rearrangements[49, 51], molecular docking[82], ligand binding[83, 84], crystallization and phase transitions[85], and so on.

3.3 Essential Energy Space Random Walk

Essential Energy Space Random Walk (EESRW)[53, 54] is an enhanced sampling method based on scaling perturbations of the biomolecular essential potential energy. Here, the essential potential energy U_s is defined as the sum of the energy terms that determine the local conformations of a region of interest. In case that the solute in a solution system is selected as the target region, U_s can be decomposed to two parts: the self-energy component U_{ss} that is calculated as the sum of self-interaction terms within the solute except for the bond, angle and improper energy terms and the interaction energy component U_{se} that is calculated as the sum of all the interaction terms between the solute and solvent. The sum of the remaining energy terms gives the environmental energy component U_e with the system total energy $U_0 = U_s + U_e$. Provided the U_s essential energy, a biasing potential is added to the biomolecular system to flatten its energy surface for enhanced sampling[53, 54]:

$$U = U_0 + f_m(U_s), \quad (10)$$

where U_0 is the original system potential energy, U_s is the essential potential energy and $f_m(U_s)$ is a relatively small Gaussian-shaped repulsive potential calculated similar to the metadynamics[33, 34] approach:

$$f_m(U_s) = \sum_{\substack{t_i=\tau,2\tau,\cdots \\ t_i<t}} h \exp\left(-\frac{|U_s - U_s(t_i)|^2}{2w^2}\right), \quad (11)$$

where τ is the frequency at which the Gaussian repulsive potential is added, h is the Gaussian height and w is the Gaussian width. EESRW has been implemented in CHARMM and successfully applied to accelerate conformational transitions in the alanine dipeptide, leucine sidechain and a “pentanelike” biomolecular model system[53, 54].

However, the U_{ss} and U_{se} essential energy components were found to fluctuate in various magnitudes with even opposite directions during many conformational transitions of biomolecules (e.g., alanine dipeptide). A one-dimensional potential biasing scheme in the original EESRW method is not able to capture such energy fluctuations effectively. To resolve this issue, EESRW was generalized to a two-dimension-EESRW (2D-EESRW) scheme[55], in which the modified system potential is given by:

$$U = U_0 + f_m(U_{ss}, U_{se}), \quad (12)$$

where U_{ss} and U_{se} are the essential self-energy and interaction energy components, respectively, and $f_m(U_{ss}, U_{se})$ is the negative of the free energy $G(U_{ss}, U_{se})$ that is obtained from the free energy profile along (U_{ss}, U_{se}). Similar to Eqn. (11), the $f_m(U_{ss}, U_{se})$ term is calculated by repetitively adding a relatively small Gaussian-shaped repulsive potential:

$$f_m(U_{ss}, U_{se}) = \sum_{\substack{t_i=\tau,2\tau,\cdots \\ t_i<t}} h\exp\left(-\frac{|U_{ss}-U_{ss}(t_i)|^2}{2w_{ss}^2}\right)\exp\left(-\frac{|U_{se}-U_{se}(t_i)|^2}{2w_{se}^2}\right), \quad (13)$$

where τ is the frequency at which the Gaussian repulsive potential is added, h is the Gaussian height, and w_{ss} and w_{se} are the widths of Gaussians corresponding to the essential potential energies U_{ss} and U_{se}, respectively.

The 2D-EESRW was shown to accelerate conformational sampling of the alanine dipeptide and aspartate-arginine peptide model systems[55]. Compared with the original EESRW, the 2D-EESRW simulations exhibit significantly higher rates of conformational transitions in the model systems. It was also suggested that the EESRW methods could be extended to higher dimension schemes when three or more essential energetic terms are needed to account for their distinct interplays during biomolecular conformational transitions.

3.4 Accelerated Molecular Dynamics (aMD)

Figure 3 depicts the central idea of the original Accelerated Molecular Dynamics (aMD) that was developed as a potential-biasing method for enhanced sampling[56]. It accelerates the conformational sampling of biomolecules often by adding a non-negative boost potential to the potential energy surface when the system potential is lower than a reference energy[56, 86, 87]:

$$V^*(r)=V(r), \quad V(r) \geq E,$$

$$V^*(r)=V(r)+\Delta V(r), V(r)<E. \quad (14)$$

where $V(r)$ is the original potential, E is the reference energy, and $V^*(r)$ is the modified potential. The boost potential, $\Delta V(r)$is given by:

$$\Delta V(r)=\frac{(E-V(r))^2}{\alpha+E-V(r)}, \quad (15)$$

where α is the acceleration factor. As the acceleration factor α decreases, the potential energy surface is flattened and biomolecular transitions between the low-energy states are increased (Figure 3).

In aMD, the boost potential was initially applied to only the dihedral energy term based on the finding that the dihedral torsions are determining factors of the protein conformations[56]. This "dihedral-boost" was found, however, to be still insufficient for conformational sampling of many biomolecules. Thus another boost was added to the system total potential energy, which accelerates diffusion of the solvent molecules and provides further enhanced sampling of the biomolecules. This led to the development of "dual-boost" aMD[86].

In addition, a number of different forms of aMD have been developed to improve the sampling statistics of bimolecular configurations, including the windowed aMD designed to protect the high-energy barriers[88], adaptive aMD[89], rotatable aMD or RaMD[90] and selective aMD[91]. With the enhanced sampling power, aMD has also been combined with many other techniques, including the thermodynamic integration[92], replica exchange[93–95], constant pH MD[96] and *ab initio* MD [97, 98], for various classical and quantum computational simulations of biomolecules.

A wide range of applications has shown that aMD works well for enhanced sampling of biomolecules[87]. For example, aMD simulations provide significant speed-up of the peptide conformational transitions[56, 99, 100], lipid diffusion and mixing[101], conformational transitions in proteins[89, 102–105], protein folding[106, 107] and protein-ligand binding[108]. Notably, hundreds-of-nanosecond aMD simulations are to capture millisecond-timescale events in both globular and membrane proteins[109–112].

While aMD is powerful for enhanced conformation sampling, its accuracy (particularly reweighting) for free energy calculations has attracted lots of attention[113]. In theory, frames of aMD simulations can be reweighted by the Boltzmann factors of the corresponding boost potential (i.e., $e^{\Delta V/k_BT}$) and averaged over each bin of selected reaction coordinate(s) to obtain the canonical ensemble, an algorithm termed "exponential average". However, exponential reweighting is known to suffer from large statistical noise in practical calculations[87, 111, 113, 114] because the Boltzmann reweighting factors are often dominated by a very few frames with high boost potential. To reduce the energetic noise, the RaMD[90] and selective aMD[91] methods were developed to apply boost potential to only rotatable dihedrals or certain regions of the simulated system. Alternatively, scaled MD flattens the biomolecular potential energy surface with a scaling factor[114], a similar notion used in EESRW[53, 54]. It enables population-based reweighting and significantly improves the free energy calculations. However, scaled MD does not allow selective acceleration of biomolecular conformational changes like aMD.

Overall, the boost potential in aMD simulations of proteins is on the order of tens to hundreds of kcal/mol, which is much greater in magnitude and wider in distribution than that of CV-biasing simulations (e.g., several kcal/mol in metadynamics). It has been a long-standing problem to accurately reweight aMD simulations and recover the original free energy landscapes, especially for large proteins[108, 111]. A recent study showed that when the boost potential follows near-Gaussian distribution, cumulant expansion to the second order provides improved reweighting of aMD simulations compared with the previously used exponential average and Maclaurin series expansion reweighting methods[115]. The

reweighted free energy profiles are in good agreement with the long-timescale cMD simulations as demonstrated on alanine dipeptide and fast-folding proteins[107]. However, such improvement is limited to rather small systems (e.g., proteins with less than ~35 amino acid residues)[107]. In simulations of larger systems, the boost potential exhibits significantly wider distribution and does not allow for accurate reweighting.

In order to achieve both unconstrained enhanced sampling and accurate energetic reweighting for protein simulations, Gaussian Accelerated Molecular Dynamics (GaMD) was developed by applying a harmonic boost potential instead to smoothen the system potential energy surface[57]. As shown in Figure 4, when the system potential $V(\vec{r})$ drops below a threshold energy E, the energy surface is modified by adding a boost potential as:

$$\Delta V(\vec{r}) = \frac{1}{2}k(E - V(\vec{r}))^2, V(\vec{r}) < E. \quad (16)$$

where k is the harmonic force constant. The two adjustable parameters E and k are automatically determined by applying the following three criteria. First, for any two arbitrary potential values $V_1(\vec{r})$ and $V_2(\vec{r})$ found on the original energy surface, if $V_1(\vec{r}) < V_2(\vec{r})$, ΔV should be a monotonic function that does not change the relative order of the biased potential values, i.e., $V_1^*(\vec{r}) < V_2^*(\vec{r})$. Secondly, if $V_1(\vec{r}) < V_2(\vec{r})$, the potential difference observed on the smoothened energy surface should be smaller than that of the original, i.e., $V_2^*(\vec{r}) - V_1^*(\vec{r}) < V_2(\vec{r}) - V_1(\vec{r})$. By combining the first two criteria and plugging in the formula of $V^*(\vec{r})$ and ΔV, we obtain:

$$V_{\max} \leq E \leq V_{\min} + \frac{1}{k}. \quad (17)$$

where $V_{\min}$ and $V_{\max}$ are the system minimum and maximum potential energies. To ensure that Eqn. (17) is valid, k has to satisfy: $k \leq \frac{1}{V_{\max} - V_{\min}}$. Let us define $k \equiv k_0 \bullet \frac{1}{V_{\max} - V_{\min}}$, then $0 < k_0 \leq 1$. Thirdly, the standard deviation of ΔV needs to be small enough (i.e., narrow distribution) to ensure accurate reweighting using cumulant expansion to the second order[115]: $\sigma_{\Delta V} = k(E - V_{avg})\sigma_V \leq \sigma_0$ where V_{avg} and σ_V are the average and standard deviation of the system potential energies, $\sigma_{\Delta V}$ is the standard deviation of ΔV with σ_0 as a user-specified upper limit (e.g., $10k_BT$) for accurate reweighting. When E is set to the lower bound $E = V_{\max}$ according to Eqn. (17), k_0 can be calculated as:

$$k_0 = \min\left(1.0, k_0'\right) = \min\left(1.0, \frac{\sigma_0}{\sigma_V} \cdot \frac{V_{\max} - V_{\min}}{V_{\max} - V_{\text{avg}}}\right). \quad (18)$$

Alternatively, when the threshold energy E is set to its upper bound $E=V_{\min}+\frac{1}{k}$, k_0 is set to:

$$k_0=k_0'' \equiv \left(1-\frac{\sigma_0}{\sigma_V}\right)\cdot\frac{V_{\max}-V_{\min}}{V_{\text{avg}}-V_{\min}}; \quad (19)$$

if k_0'' is calculated between *0* and *1*. Otherwise, k_0 is calculated using Eqn. (18), instead of being set to 1 directly as described in Ref. [57].

By adaptively adding a harmonic boost potential to smoothen the system energy surface, GaMD provides both unconstrained enhanced sampling and free energy calculation of biomolecules. Important statistical properties of the system potential, such as the average, maximum, minimum and standard deviation values, are used to calculate the simulation acceleration parameters. After the threshold energy E is calculated using Eqn. (17), the effective harmonic constant k_0 in the range of 0 to 1 determines the magnitude of the applied boost potential. With greater k_0, higher boost potential is added to the original energy surface, which provides enhanced sampling of biomolecules across the decreased energy barriers (Figure 4). The boost potential follows Gaussian distribution and accurate reweighting is obtained through cumulant expansion to the second order. Without the need to set predefined CVs, GaMD enables unconstrained enhanced sampling of the biomolecules. As demonstrated earlier, free energy profiles obtained from GaMD simulations allow us to identify distinct low energy states and characterize the protein folding and ligand binding pathways quantitatively[57].

4. Future directions

It remains a grand challenge to bridge the gap between the timescales of MD simulations and those of biological processes in the near future. Enhanced sampling, especially the unconstrained methods as discussed in this review, will continue to help computational modeling of biomolecules. Along with the theoretical development, there is a need to extend their applications to systems of increasing size and complexity that are more relevant to biological studies and computer-aided drug design. Most currently simulated systems are often limited to short peptides (e.g., alanine dipeptide) and small proteins. We need to reach for larger systems of biological and pharmaceutical interest, e.g., functional enzymes, membrane proteins, cellular protein complexes, virus particles, etc.

While the enhanced sampling methods discussed above have been found successful in various applications, they often suffer from one or more limitations. In most cases, it will be beneficial to combine two or more of the above techniques for further enhanced sampling. For example, it has been shown that the combination of RE with ST and MUCA in the REXST and REMUCA algorithms is able to reduce the computational cost without requiring too many MD replicas in the pure RE simulations. Based on a potential-biasing approach, GaMD mainly accelerates transitions across enthalpic energy barriers. Improvement for its application to systems with high entropic barriers is still needed. On this regard, GaMD can be potentially combined with the RE/PT algorithms like in

REXAMD[93, 94] for further enhanced sampling. Notably, the combination of parallel tempering and metadynamics (PT-MetaD)[116] has been shown to facilitate enhanced sampling of biomolecules over entropic barriers. Moreover, the REXAMD that combines RE and aMD methods have been found helpful in free energy calculations. Similarly, one can combine GaMD that provides improved reweighting and RE for more accurate free energy calculations. In addition, we can use the essential potential energy as described in the EESRW method to improve the potential-biasing sampling methods such as GaMD.

Finally, in order to obtain accurate free energy calculations, rigorous error analysis (particularly the reweighting) is still needed for the enhanced sampling methods. Take the reweighting of aMD/GaMD simulations as an example. Because the boost potential ΔV is physically equivalent to nonequilibrium work W, reweighting of aMD simulations for calculating free energies can be expressed by the Jarzynski equality[117]: $e^{-\beta\Delta F} = \langle e^{-\beta\Delta V} \rangle$. When the three ΔF estimators and different aMD reweighting techniques examined in Refs. [118] and [115] are compared, the mean work estimator is found equivalent to the approximation using Maclaurin series expansion, the fluctuation-dissipation (FD) theorem estimator corresponds to the cumulant expansion to the 2nd order (also referred to as the "Gaussian approximation"[18]) and the Jarzynski estimator corresponds to direct "exponential average" calculation. Detailed analysis showed that in the near-equilibrium regime, the mean work estimator gives comparatively larger errors and the Jarzynski estimator is more accurate than the FD estimator when the number of W (or ΔV) data points $N \leq 16$. However, the FD estimator gives the smallest error with increasing N (see Figure 4 in Ref. [118]). This is consistent with our previous aMD reweighting study that shows the cumulant expansion to the 2nd order is the most accurate compared with the Maclaurin series expansion and exponential average, for which the number of ΔV values N used for aMD reweighting is on the order of 10^2–10^6[115]. In summary, the bias and error of free energy calculations were examined rigorously in Ref. [118]. Similar error estimates need to be performed for GaMD simulations, as well as the other enhanced sampling methods, when they are applied for free energy calculations.

5. Conclusions

Enhanced sampling has been found highly useful in the last several decades for exploring the configuration space of biomolecules and improving the convergence of free energy calculations. In this review, we focused on the enhanced sampling methods that do not require predefined system-dependent CVs for biomolecular simulations and as such do not suffer from the hidden energy barrier problem as encountered in the CV-biasing methods. These methods are based on generalized ensemble simulations of temperature (for RE/PT), local average force (SGMD and SGLD) and potential energy (EESRW, aMD and GaMD). This allows for generally unconstrained enhanced sampling of biomolecules, particularly the pathways during their biological function such as protein folding and ligand binding. While theoretical and computational work is still needed to improve the sampling statistics and accuracy, wider applications are expected for the unconstrained enhanced sampling methods in free energy calculations of biomolecules and computer-aided drug design.

Acknowledgments

We thank Donald Hamelberg, Yuji Sugita, Xiongwu Wu and Wei Yang for valuable discussions and suggestions. This work was supported by NSF (grant MCB1020765), NIH (grant GM31749), Howard Hughes Medical Institute, National Biomedical Computation Resource (NBCR), and the national supercomputer centers (the XSEDE awards TG-MCA93S013 and TG-MCB140011 and the NERSC award project M1925).

References

1. McCammon JA, et al. Dynamics of Folded Proteins. Nature. 1977:585–590. [PubMed: 301613]
2. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nature Structural Biology. 2002:646–652. [PubMed: 12198485]
3. Miao Y, et al. Coupled Flexibility Change in Cytochrome P450cam Substrate Binding Determined by Neutron Scattering, NMR, and Molecular Dynamics Simulation. Biophysical Journal. 2012:2167–2176. [PubMed: 23200050]
4. Maragakis P, et al. Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. J Phys Chem B. 2008:6155–6158. [PubMed: 18311962]
5. Yang S, et al. Measuring similarity between dynamic ensembles of biomolecules. Nature Methods. 2014:552–554. [PubMed: 24705474]
6. Smith J, et al. Dynamics of Myoglobin - Comparison of Simulation Results with Neutron-Scattering Spectra. Proceedings of the National Academy of Sciences of the United States of America. 1990:1601–1605. [PubMed: 2304919]
7. Miao YL, et al. Temperature-Dependent Dynamical Transitions of Different Classes of Amino Acid Residue in a Globular Protein. Journal of the American Chemical Society. 2012:19576–19579. [PubMed: 23140218]
8. Chan KY, et al. Cryo-electron microscopy modeling by the molecular dynamics flexible fitting method. Biopolymers. 2012:678–686. [PubMed: 22696404]
9. Trabuco LG, et al. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure. 2008:673–683. [PubMed: 18462672]
10. Lindorff-Larsen K, et al. How fast-folding proteins fold. Science. 2011:517–520. [PubMed: 22034434]
11. Brunger AT, et al. Crystallographic R factor refinement by molecular dynamics. Science. 1987:458–460. [PubMed: 17810339]
12. McGreevy R, et al. xMDFF: molecular dynamics flexible fitting of low-resolution X-ray structures. Acta Crystallogr D Biol Crystallogr. 2014:2344–2355. [PubMed: 25195748]
13. Torda AE, van Gunsteren WF. The Refinement of Nmr Structures by Molecular-Dynamics Simulation. Computer Physics Communications. 1991:289–296.
14. Lehtivarjo J, et al. Combining NMR ensembles and molecular dynamics simulations provides more realistic models of protein structures in solution and leads to better chemical shift prediction. Journal of Biomolecular Nmr. 2012:257–267. [PubMed: 22314705]
15. Kollman P. Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. Chemical Reviews. 1993:2395–2417.
16. Vanden-Eijnden E. Some Recent Techniques for Free Energy Calculations. Journal of Computational Chemistry. 2009:1737–1747. [PubMed: 19504587]
17. Christ CD, et al. Basic Ingredients of Free Energy Calculations: A Review. Journal of Computational Chemistry. 2010:1569–1582. [PubMed: 20033914]
18. Pohorille A, et al. Good practices in free-energy calculations. J Phys Chem B. 2010:10235–10253. [PubMed: 20701361]
19. Johnston JM, Filizola M. Showcasing modern molecular dynamics simulations of membrane proteins through G protein-coupled receptors. Current Opinion in Structural Biology. 2011:552–558. [PubMed: 21764295]
20. Henzler-Wildman K, Kern D. Dynamic personalities of proteins. Nature. 2007:964–972. [PubMed: 18075575]

21. Lewandowski JR, et al. Direct observation of hierarchical protein dynamics. Science. 2015:578–581. [PubMed: 25931561]
22. Abrams C, Bussi G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. Entropy. 2014:163–199.
23. Spiwok V, et al. Enhanced sampling techniques in biomolecular simulations. Biotechnol Adv. 2014 In press, doi: j.biotechadv.2014.11.011.
24. Dellago C, Bolhuis PG. Transition Path Sampling and Other Advanced Simulation Techniques for Rare Events. Advanced Computer Simulation Approaches for Soft Matter Sciences Iii. 2009:167–233.
25. Gao YQ, et al. Thermodynamics and kinetics simulations of multi-time-scale processes for complex systems. International Reviews in Physical Chemistry. 2008:201–227.
26. Liwo A, et al. Computational techniques for efficient conformational sampling of proteins. Curr Opin Struct Biol. 2008:134–139. [PubMed: 18215513]
27. Christen M, van Gunsteren WF. On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: A review. J Comput Chem. 2008:157–166. [PubMed: 17570138]
28. Kastner J. Umbrella sampling. Wiley Interdisciplinary Reviews-Computational Molecular Science. 2011:932–942.
29. Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics. 1977:187–199.
30. Ciccotti G, et al. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. Chemphyschem. 2005:1809–1814. [PubMed: 16144000]
31. Darve E, et al. Adaptive biasing force method for scalar and vector free energy calculations. Journal of Chemical Physics. 2008:144120. [PubMed: 18412436]
32. Darve E, Pohorille A. Calculating free energies using average force. Journal of Chemical Physics. 2001:9169–9183.
33. Laio A, Parrinello M. Escaping free-energy minima. Proceedings of the National Academy of Sciences of the United States of America. 2002:12562–12566. [PubMed: 12271136]
34. Laio A, Gervasio FL. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. Reports on Progress in Physics. 2008:126601.
35. Grubmüller H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. Physical Review E, American Physical Society. 1995:2893–2906.
36. Maragliano L, et al. String method in collective variables: Minimum free energy paths and isocommittor surfaces. Journal of Chemical Physics. 2006:024106. [PubMed: 16422570]
37. Pan AC, et al. Finding transition pathways using the string method with swarms of trajectories. Journal of Physical Chemistry B. 2008:3432–3440.
38. Knight JL, Brooks CL. lambda-Dynamics Free Energy Simulation Methods. Journal of Computational Chemistry. 2009:1692–1700. [PubMed: 19421993]
39. Rosso L, Tuckerman ME. An adiabatic molecular dynamics method for the calculation of free energy profiles. Molecular Simulation. 2002:91–112.
40. Abrams CF, Vanden-Eijnden E. Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proceedings of the National Academy of Sciences of the United States of America. 2010:4961–4966. [PubMed: 20194785]
41. Schlitter J, et al. Targeted Molecular-Dynamics Simulation of Conformational Change - Application to the T[--]R Transition in Insulin. Molecular Simulation. 1993:291–308.
42. Ferrara P, et al. Targeted molecular dynamics simulations of protein unfolding. Journal of Physical Chemistry B. 2000:4511–4518.
43. Schlitter J, et al. Targeted Molecular-Dynamics - a New Approach for Searching Pathways of Conformational Transitions. Journal of Molecular Graphics. 1994:84–89. [PubMed: 7918256]
44. Zheng LQ, et al. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. Proceedings of the National Academy of Sciences of the United States of America. 2008:20227–20232. [PubMed: 19075242]

45. Zheng LQ, et al. Simultaneous escaping of explicit and hidden free energy barriers: Application of the orthogonal space random walk strategy in generalized ensemble based conformational sampling. Journal of Chemical Physics. 2009:234105. [PubMed: 19548709]
46. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chemical Physics Letters. 1999:141–151.
47. Okamoto Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. Journal of Molecular Graphics & Modelling. 2004:425–439. [PubMed: 15099838]
48. Hansmann UHE. Parallel tempering algorithm for conformational studies of biological molecules. Chemical Physics Letters. 1997:140–150.
49. Wu X, Wang S. Crossing the time scale of protein folding through self-guided molecular dynamics simulation. Journal of Molecular Graphics & Modelling. 1998:290–290.
50. Wu XW, Wang SM. Self-guided molecular dynamics simulation for efficient conformational search. Journal of Physical Chemistry B. 1998:7238–7250.
51. Wu XW, Brooks BR. Self-guided Langevin dynamics simulation method. Chemical Physics Letters. 2003:512–518.
52. Wu X. Self-guided Langevin dynamics via generalized Langevin equation. Journal of Computational Chemistry. 2015
53. Li H. Essential energy space random walk via energy space metadynamics method to accelerate molecular dynamics simulations. Journal of Chemical Physics. 2007:094101. [PubMed: 17824726]
54. Zheng LQ, Yang W. Essential energy space random walks to accelerate molecular dynamics simulations: Convergence improvements via an adaptive-length self-healing strategy. Journal of Chemical Physics. 2008:014105. [PubMed: 18624468]
55. Lv C. Generalized essential energy space random walks to more effectively accelerate solute sampling in aqueous environment. Journal of Chemical Physics. 2012:044103. [PubMed: 22299857]
56. Hamelberg D, et al. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. Journal of Chemical Physics. 2004:11919–11929. [PubMed: 15268227]
57. Miao Y, et al. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. Journal of Chemical Theory and Computation, American Chemical Society. 2015:3584–3595.
58. Marinari E, Parisi G. Simulated Tempering - a New Monte-Carlo Scheme. Europhysics Letters. 1992:451–458.
59. Lyubartsev AP, et al. New Approach to Monte-Carlo Calculation of the Free-Energy - Method of Expanded Ensembles. Journal of Chemical Physics. 1992:1776–1783.
60. Berg BA, Neuhaus T. Multicanonical Algorithms for 1st Order Phase-Transitions. Physics Letters B. 1991:249–253.
61. Sugita Y, Okamoto Y. Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. Chemical Physics Letters. 2000:261–270.
62. Mitsutake A, et al. Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. I. Formulation and benchmark test. Journal of Chemical Physics. 2003:6664–6675.
63. Mitsutake A, Okamoto Y. Replica-exchange simulated tempering method for simulations of frustrated systems. Chemical Physics Letters. 2000:131–138.
64. Ishikawa Y, et al. Ab initio replica-exchange Monte Carlo method for cluster studies. Chemical Physics Letters. 2001:199–206.
65. Sugita Y, et al. Multidimensional replica-exchange method for free-energy calculations. Journal of Chemical Physics. 2000:6042–6051.
66. Murata K, et al. Molecular dynamics simulations of DNA dimers based on replica-exchange umbrella sampling. I. Test of sampling efficiency. Journal of Theoretical & Computational Chemistry. 2005:411–432.

67. Curuksu J, Zacharias M. Enhanced conformational sampling of nucleic acids by a new Hamiltonian replica exchange molecular dynamics approach. Journal of Chemical Physics. 2009:104110. [PubMed: 19292526]
68. Meli M, Colombo G. A Hamiltonian Replica Exchange Molecular Dynamics (MD) Method for the Study of Folding, Based on the Analysis of the Stabilization Determinants of Proteins. International Journal of Molecular Sciences. 2013:12157–12169. [PubMed: 23743827]
69. Nishikawa T, et al. Replica-exchange Monte Carlo method for Ar fluid. Progress of Theoretical Physics Supplement. 2000:270–271.
70. Mitsutake A, et al. Replica-exchange multicanonical and multicanonical replica-exchange Monte Carlo simulations of peptides. II. Application to a more complex system. Journal of Chemical Physics. 2003:6676–6688.
71. La Penna G. Molecular dynamics of C-peptide of ribonuclease A studied by replica-exchange Monte Carlo method and diffusion theory. Chemical Physics Letters. 2003:609–619.
72. Pitera JW, Swope W. Understanding folding and design: Replica-exchange simulations of "Trp-cage" fly miniproteins. Proceedings of the National Academy of Sciences of the United States of America. 2003:7587–7592. [PubMed: 12808142]
73. Sugita Y, Okamoto Y. An analysis on protein folding problem by replica-exchange method. Progress of Theoretical Physics Supplement. 2000:402–403.
74. Nguyen PH, et al. Free energy landscape and folding mechanism of a beta-hairpin in explicit water: A replica exchange molecular dynamics study. Proteins-Structure Function and Bioinformatics. 2005:795–808.
75. Kannan, S.; Zacharias, M. Proteins: Structure, Function, and Bioinformatics. Wiley Subscription Services, Inc., A Wiley Company; 2010. Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent; p. 2809-2819.
76. Murata K, et al. Free energy calculations for DNA base stacking by replica-exchange umbrella sampling. Chemical Physics Letters. 2004:1–7.
77. Mitsutake A, Okamoto Y. Replica-exchange extensions of simulated tempering method. Journal of Chemical Physics. 2004:2491–2504. [PubMed: 15281846]
78. Wu XW. Efficient and Unbiased Sampling of Biomolecular Systems in the Canonical Ensemble: A Review of Self-Guided Langevin Dynamics. Advances in Chemical Physics. 2012; 150:255–326. [PubMed: 23913991]
79. Wu XW, Brooks BR. Force-momentum-based self-guided Langevin dynamics: A rapid sampling method that approaches the canonical ensemble. Journal of Chemical Physics. 2011:204101. [PubMed: 22128922]
80. Sheng YB, et al. Adsorption of an Ionic Complementary Peptide on the Hydrophobic Graphite Surface. Journal of Physical Chemistry C. 2010:454–459.
81. Wu XW, et al. Targeted conformational search with map-restrained self-guided Langevin dynamics: Application to flexible fitting into electron microscopic density maps. Journal of Structural Biology. 2013:429–440. [PubMed: 23876978]
82. Chandrasekaran V, et al. A computational modeling and molecular dynamics study of the Michaelis complex of human protein Z-dependent protease inhibitor (ZPI) and factor Xa (FXa). Journal of Molecular Modeling. 2009:897–911. [PubMed: 19172319]
83. Yang CY, et al. Solution conformations of wild-type and mutated Bak BH3 peptides via dynamical conformational sampling and implication to their binding to antiapoptotic Bcl-2 proteins. Journal of Physical Chemistry B. 2004:1467–1477.
84. Varady J, et al. Competitive and reversible binding of a guest molecule to its host in aqueous solution through molecular dynamics simulation: Benzyl alcohol/beta-cyclodextrin system. Journal of Physical Chemistry B. 2002:4863–4872.
85. Choudhary D, Clancy P. Application of accelerated molecular dynamics schemes to the production of amorphous silicon. Journal of Chemical Physics. 2005
86. Hamelberg D, et al. Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. Journal of Chemical Physics. 2007:155102. [PubMed: 17949218]

87. Markwick PRL, McCammon JA. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. Physical Chemistry Chemical Physics. 2011:20053–20065. [PubMed: 22015376]

88. Sinko W, et al. Protecting High Energy Barriers: A New Equation to Regulate Boost Energy in Accelerated Molecular Dynamics Simulations. Journal of Chemical Theory and Computation. 2012:17–23. [PubMed: 22241967]

89. Markwick PRL, et al. Adaptive Accelerated Molecular Dynamics (Ad-AMD) Revealing the Molecular Plasticity of P450cam. Journal of Physical Chemistry Letters. 2011:158–164. [PubMed: 21307966]

90. Doshi U, Hamelberg D. Improved Statistical Sampling and Accuracy with Accelerated Molecular Dynamics on Rotatable Torsions. Journal of Chemical Theory and Computation. 2012:4004–4012. [PubMed: 26605567]

91. Wereszczynski J, McCammon JA. Using Selectively Applied Accelerated Molecular Dynamics to Enhance Free Energy Calculations. Journal of Chemical Theory and Computation. 2010:3285–3292. [PubMed: 21072329]

92. de Oliveira CAF, et al. Coupling accelerated molecular dynamics methods with thermodynamic integration simulations. Journal of Chemical Theory and Computation. 2008:1516–1525. [PubMed: 19461868]

93. Fajer M, et al. Replica-Exchange Accelerated Molecular Dynamics (REXAMD) Applied to Thermodynamic Integration. Journal of Chemical Theory and Computation. 2008:1565–1569. [PubMed: 19461870]

94. Fajer M, et al. Using Multistate Free Energy Techniques to Improve the Efficiency of Replica Exchange Accelerated Molecular Dynamics. Journal of Computational Chemistry. 2009:1719–1725. [PubMed: 19421994]

95. Arrar M, et al. w-REXAMD: A Hamiltonian Replica Exchange Approach to Improve Free Energy Calculations for Systems with Kinetically Trapped Conformations. Journal of Chemical Theory and Computation. 2013:18–23. [PubMed: 23316122]

96. Williams SL, et al. Coupling Constant pH Molecular Dynamics with Accelerated Molecular Dynamics. Journal of Chemical Theory and Computation. 2010:560–568. [PubMed: 20148176]

97. Bucher D, et al. On the Use of Accelerated Molecular Dynamics to Enhance Configurational Sampling in Ab Initio Simulations. Journal of Chemical Theory and Computation. 2011:890–897. [PubMed: 21494425]

98. Pierce LCT, et al. Accelerating chemical reactions: Exploring reactive free-energy surfaces using accelerated ab initio molecular dynamics. Journal of Chemical Physics. 2011:174107. [PubMed: 21548673]

99. Hamelberg D, et al. Phosphorylation effects on cis/trans isomerization and the backbone conformation of serine-proline motifs: Accelerated molecular dynamics analysis. Journal of the American Chemical Society. 2005:1969–1974. [PubMed: 15701032]

100. De Oliveira CAF, et al. Estimating kinetic rates from accelerated molecular dynamics simulations: Alanine dipeptide in explicit solvent as a case study. Journal of Chemical Physics. 2007

101. Wang Y, et al. Enhanced Lipid Diffusion and Mixing in Accelerated Molecular Dynamics. Journal of Chemical Theory and Computation. 2011:3199–3207. [PubMed: 22003320]

102. Baron R, et al. E9-Im9 Colicin DNase-Immunity Protein Biomolecular Association in Water: A Multiple-Copy and Accelerated Molecular Dynamics Simulation Study. Journal of Physical Chemistry B. 2008:16802–16814.

103. Grant BJ, et al. Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics. Plos Computational Biology. 2009:e1000325. [PubMed: 19300489]

104. Bucher D, et al. Accessing a Hidden Conformation of the Maltose Binding Protein Using Accelerated Molecular Dynamics. Plos Computational Biology. 2011:e1002034. [PubMed: 21533070]

105. de Oliveira CAF, et al. Large-Scale Conformational Changes of Trypanosoma cruzi Proline Racemase Predicted by Accelerated Molecular Dynamics Simulation. Plos Computational Biology. 2011:e1002178. [PubMed: 22022240]

106. Doshi U, Hamelberg D. Achieving Rigorous Accelerated Conformational Sampling in Explicit Solvent. Journal of Physical Chemistry Letters. 2014:1217–1224. [PubMed: 26274474]

107. Miao Y, et al. Accelerated molecular dynamics simulations of protein folding. J Comput Chem. 2015:1536–1549. [PubMed: 26096263]

108. Kappel K, et al. Accelerated Molecular Dynamics Simulations of Ligand Binding to a Muscarinic G-protein Coupled Receptor. Quarterly Reviews of Biophysics. 2015:479–487. [PubMed: 26537408]

109. Pierce LCT, et al. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. Journal of Chemical Theory and Computation. 2012:2997–3002. [PubMed: 22984356]

110. Miao Y, et al. Activation and dynamic network of the M2 muscarinic receptor. Proc Natl Acad Sci U S A. 2013:10982–10987. [PubMed: 23781107]

111. Miao Y, et al. Free Energy Landscape of G-Protein Coupled Receptors, Explored by Accelerated Molecular Dynamics. Physical Chemistry Chemical Physics. 2014:6398–6406. [PubMed: 24445284]

112. Miao Y, et al. Allosteric Effects of Sodium Ion Binding on Activation of the M3 Muscarinic G-Protein Coupled Receptor. Biophysical Journal. 2015:1796–1806. [PubMed: 25863070]

113. Shen TY, Hamelberg D. A statistical analysis of the precision of reweighting-based simulations. Journal of Chemical Physics. 2008:034103. [PubMed: 18647012]

114. Sinko W, et al. Population Based Reweighting of Scaled Molecular Dynamics. Journal of Physical Chemistry B, American Chemical Society. 2013:12759–12768.

115. Miao Y, et al. Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. Journal of Chemical Theory and Computation. 2014:2677–2689. [PubMed: 25061441]

116. Bussi G, et al. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. Journal of the American Chemical Society. 2006:13435–13441. [PubMed: 17031956]

117. Jarzynski C. Nonequilibrium equality for free energy differences. Physical Review Letters. 1997:2690–2693.

118. Gore J, et al. Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements. Proceedings of the National Academy of Sciences of the United States of America. 2003:12564–12569. [PubMed: 14528008]

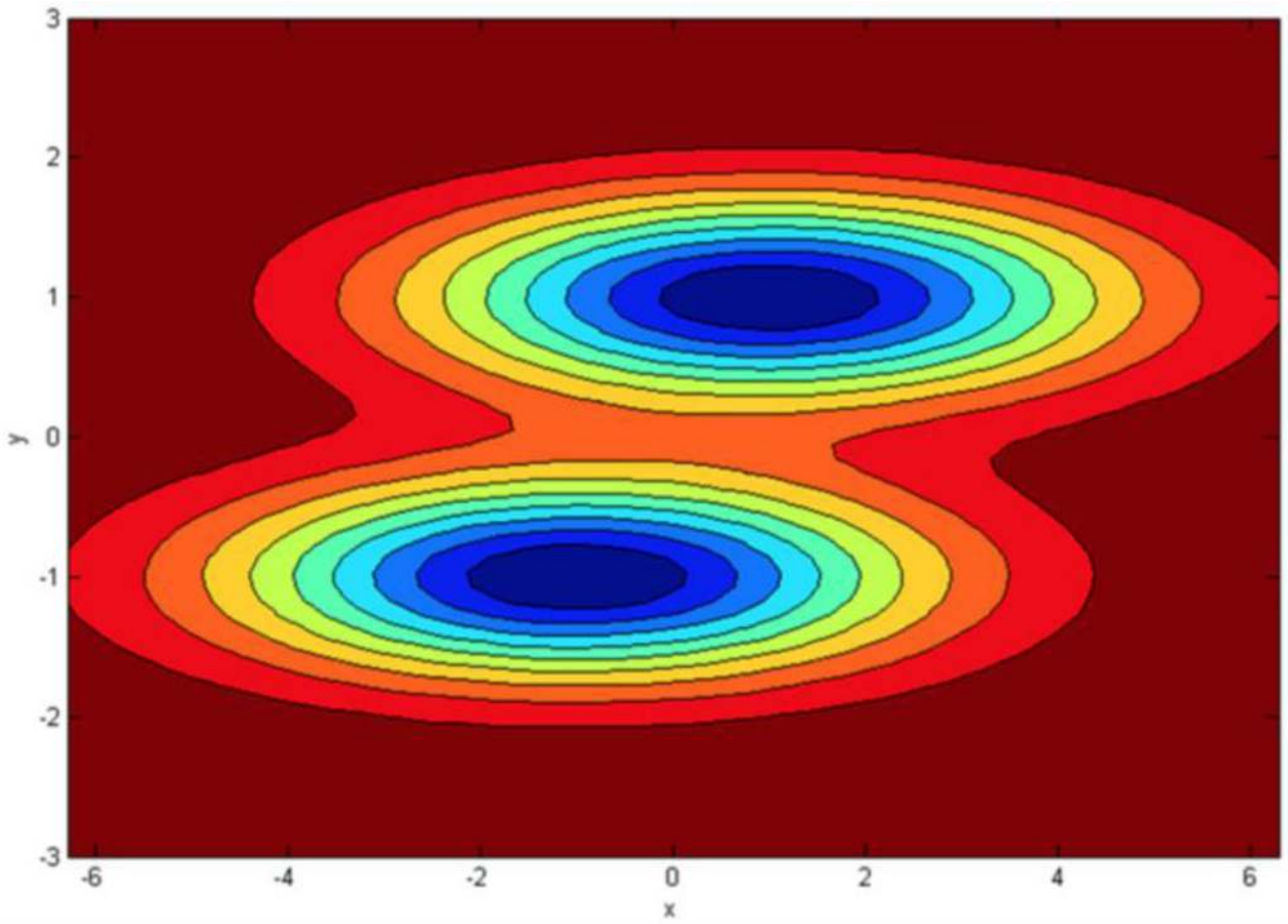

Fig. 1.

Illustration of an example hidden energy barrier along the *y* direction between two local energy minima.

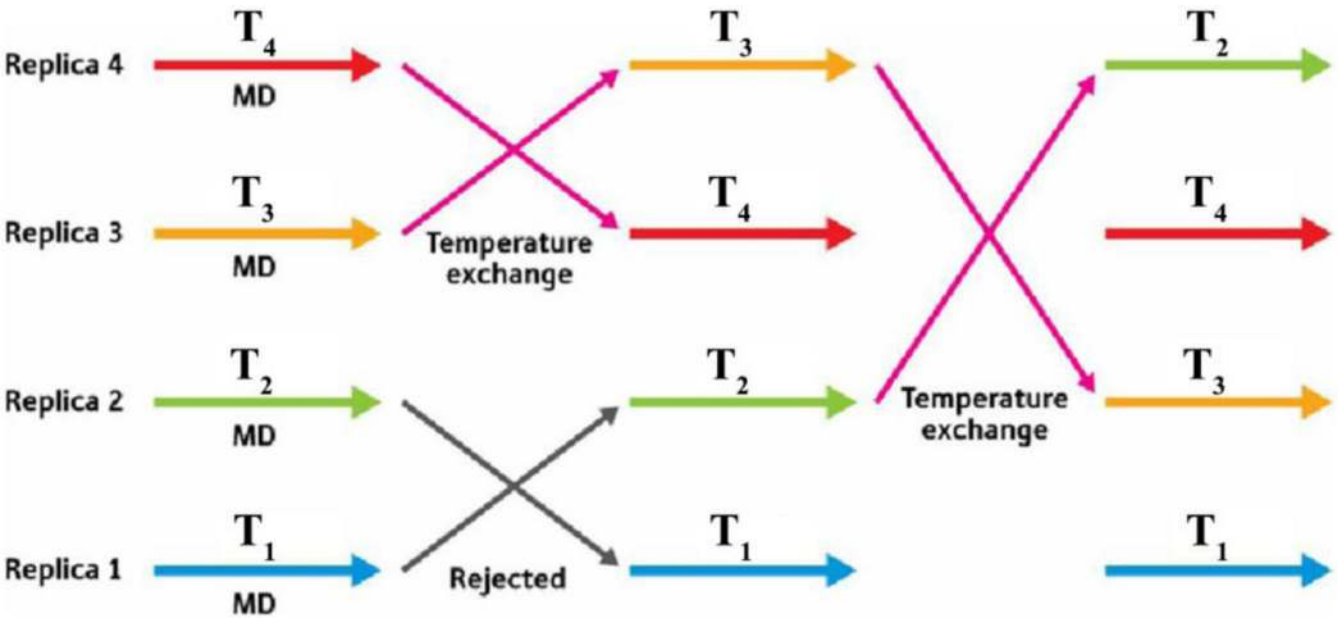

Fig. 2.

Schematic illustration of Replica Exchange Molecular Dynamics (REMD): a number of M (*e.g.,* 4 as shown here) non-interacting replicas of MD runs are performed in the canonical ensemble, each at a different temperature T_m ($m = 1, \cdots, M$). At a certain simulation time interval, we exchange each pair of replicas with neighboring temperatures at a probability that meets the Metropolis criterion. The generalized ensemble of the system then satisfies a detailed balance and converges towards an equilibrium distribution during the simulation.

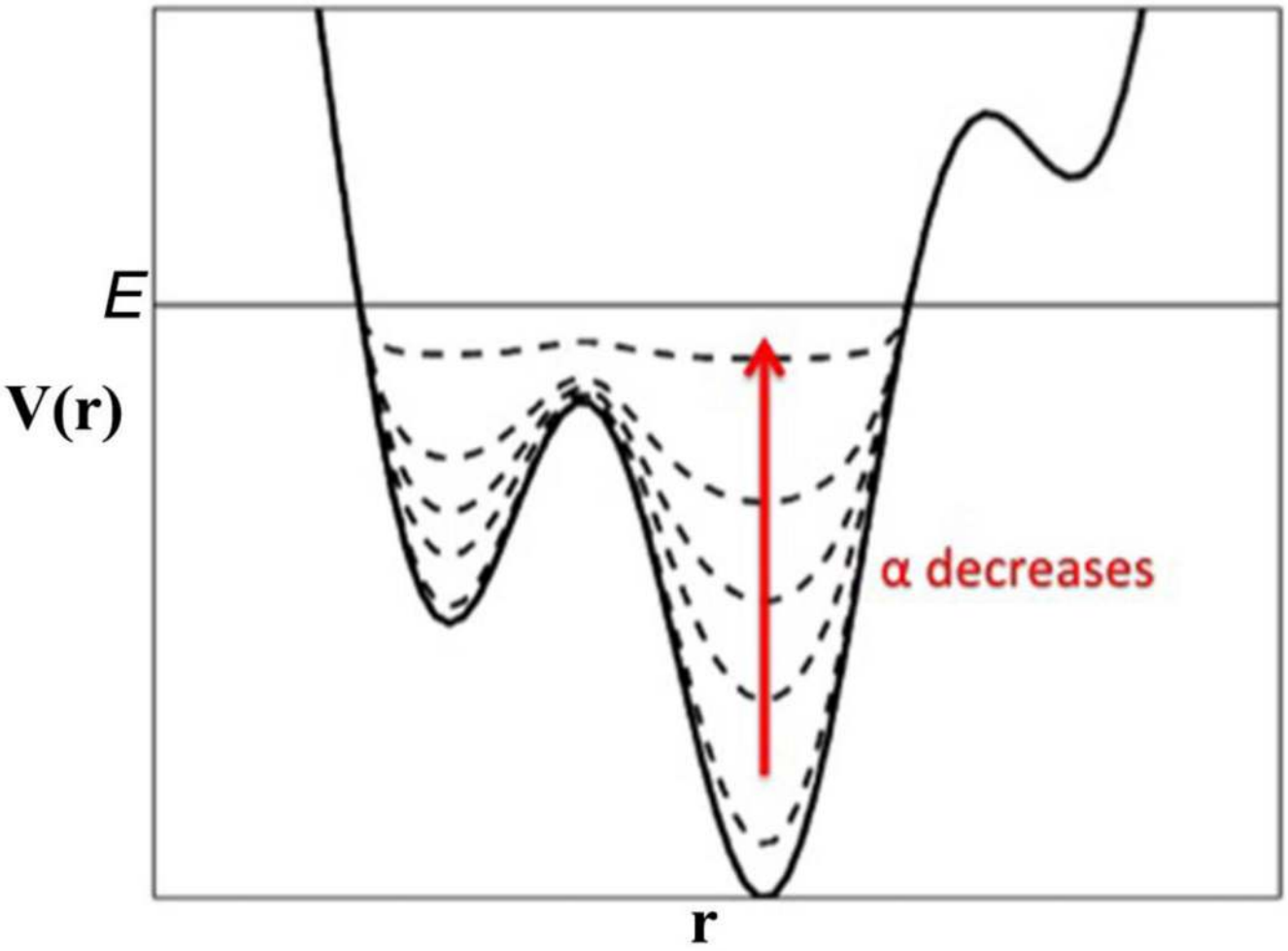

Fig. 3.

Schematic illustration of Accelerated Molecular Dynamics (aMD): a non-negative boost potential is added to biomolecular potential surface when the system potential is lower than a threshold energy *E*. As the acceleration factor α decreases, the potential energy surface is flattened more and transitions between different low-energy states become increased.

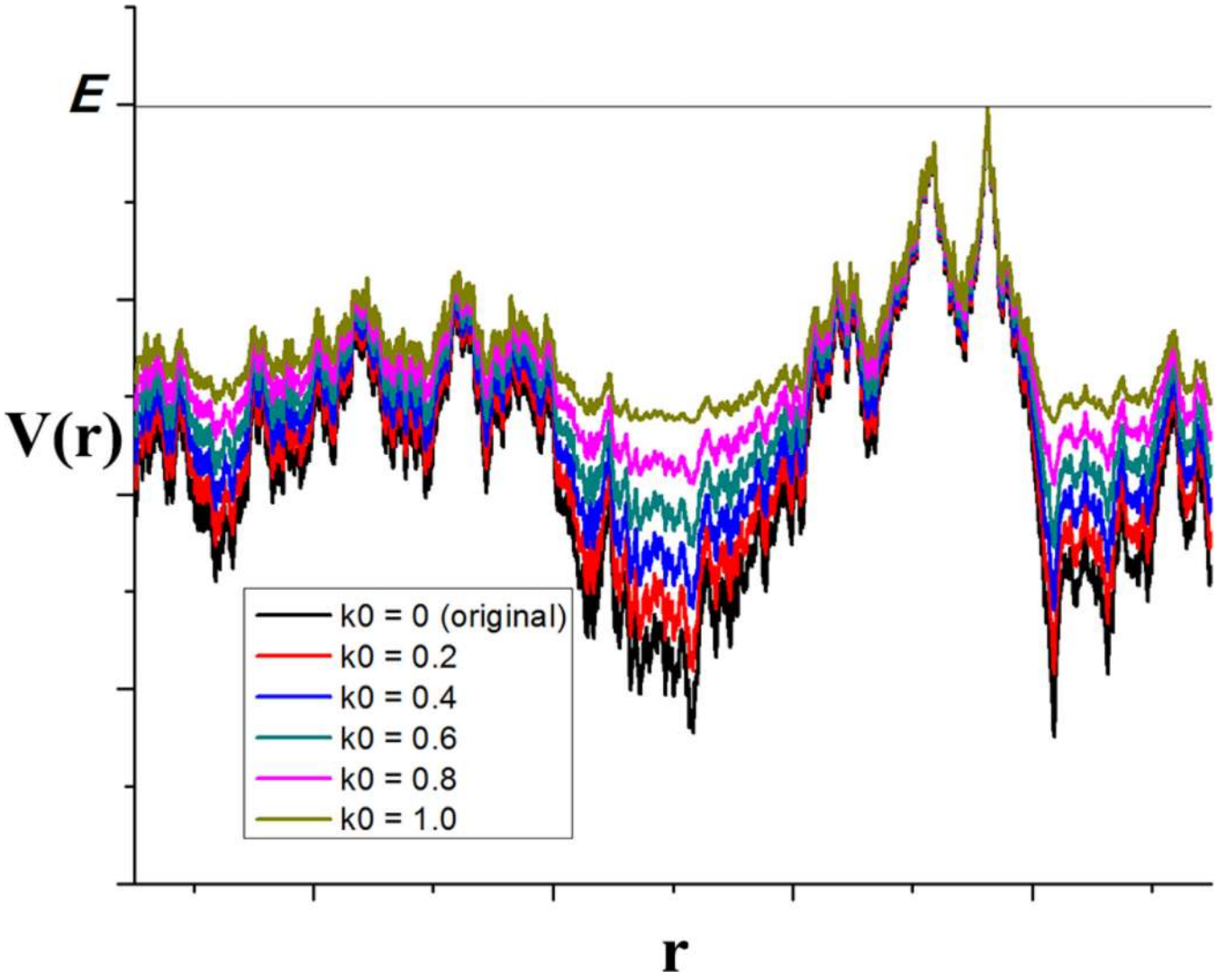

Fig. 4.

Schematic illustration of Gaussian Accelerated Molecular Dynamics (GaMD): when the threshold energy is set to the maximum potential ($E = V_{max}$), the system potential energy surface is smoothened by adding a harmonic boost potential that follows Gaussian distribution. The effective harmonic constant k_0 in the range of 0 to 1 determines the magnitude of the applied boost potential. With greater k_0, higher boost potential is added to the original energy surface in conventional molecular dynamics (cMD), which provides enhanced sampling of biomolecules across decreased energy barriers.