

REVIEW

Phase Space and Collective Variable Based Simulation Methods for Studies of Rare Events

Sanjib Paul^a, Nisanth N. Nair^b, and Harish Vashisth^a

^aDepartment of Chemical Engineering, University of New Hampshire, Durham, NH, USA;

^bDepartment of Chemistry, Indian Institute of Technology, Kanpur, Uttar Pradesh, India

ARTICLE HISTORY

Compiled May 27, 2019

ABSTRACT

In this work, we review molecular simulation methods for studies of rare events. Specifically highlighted are recent advances and biological applications of methods that utilize unbiased sampling in full phase space to discover reactive trajectories and their ensembles as well as methods that utilize biased sampling along collective variables to efficiently sample free energy landscapes. Among phase space methods, we discuss transition path sampling and its variants and highlight emerging themes in analyses of reactive trajectories via transition path theory, reactive islands, and Lagrangian descriptors. Among collective variable methods, we discuss techniques that allow efficient sampling in a high-dimensional CV space, either via replica-based approaches or by invoking an independent sampling protocol for each CV. Specifically, we highlight two variants of the metadynamics method, bias-exchange and parallel-bias metadynamics, and a hybrid method termed temperature-accelerated sliced sampling, that have been designed to overcome limitations of related approaches.

KEYWORDS

enhanced sampling; reaction coordinates; collective variables; free energy calculations; transition path sampling; metadynamics; temperature-accelerated sliced sampling

1. Introduction

Conformational dynamics and chemical reactions in complex molecular systems are often studied via molecular simulation methods with an aim to fully characterize thermodynamics and kinetics of underlying physicochemical processes. A key goal in these investigations is to resolve details of rarely occurring transitions between two metastable states of a system. Although the transition event itself is relatively rapid, it is rarely observed due to the existence of high free energy barriers that limit system fluctuations at equilibrium to long-lived metastable states. However, the observation and characterization of transition events reveal the relevant degrees of freedom driving processes and as a result provide mechanistic insights to enhance or diminish their rates of occurrence.

Specifically, major questions in atomic-scale simulation studies of rare events are centered around identification of key theoretical constructs (order parameters, collective variables, and reaction coordinates) to monitor the progress of the transition process, computation of free energy profiles along the postulated coordinates, and an estimation of rates. Transition state theory, which defines the transition state as a saddle point on the potential energy surface, was an early theoretical tool to study rare events [1–4]. However, it remains limited in studies of macromolecular systems [5], often due to the lack of knowledge of complex potential energy surfaces and/or their surface features, thereby making it difficult to unambiguously identify a single transition state or a unique transition state region; especially challenging is to identify the dividing hyperplane with the maximum transmission coefficient.

The viewpoint that one can study the ensemble of transition paths instead of focusing on the transition states is at the core of the transition path sampling (TPS) method that employs a Monte Carlo (MC) sampling strategy to construct the transition path ensemble (TPE) [6–8]. Importantly, TPS utilizes sampling in full phase space (configurational co-

ordinates and momenta) to harvest reactive trajectories and does not necessitate *a priori* knowledge of the reaction coordinate; generally it suffices to unambiguously demarcate the initial (reactant) and final (product) states.

In contrast, a suite of techniques that rely extensively on the definition of a reaction coordinate or a collective variable (CV) are rooted in the molecular mechanics based approach of molecular dynamics (MD) simulation [9]. Primarily due to challenges in directly observing rare events using conventional MD simulations, these approaches aim to enhance the sampling in configuration space via biased potentials or coupling to a high-temperature bath. Among many other methods in this class are umbrella sampling [10–13], metadynamics [14], driven-adiabatic free energy dynamics (d-AFED) [15,16], and temperature-accelerated molecular dynamics (TAMD) [17].

Given that several excellent reviews on theory and applications of these simulation methods already exist [5,7,18–33], we mainly focus in this short review on methodological advances that are recent and whose applications are just emerging or have not been highlighted in other reviews. Specifically in the context of applications to biological systems, we discuss the TPS method, two-variants of the metadynamics method (bias-exchange and parallel-bias), and a hybrid technique termed temperature-accelerated sliced sampling (TASS).

2. Transition Path Sampling

2.1. Overview

TPS is an algorithm to generate reactive trajectories connecting an initial (reactant) state and a final (product) state, which are generally considered separated by a high free-energy barrier [6,7,34] and between which no other long-lived metastable state exists. The method uses recursive MC sampling in trajectory space to create a new trajectory from

an old one, and the repeated sampling in this way generates an ensemble of transition paths in more probable regions. The transition path ensemble thus constructed is a set of fully dynamical trajectories and therefore facilitates computation of rare-event kinetics. In the following, we briefly describe key ingredients of a TPS simulation.

2.1.1. Definition of Stable States

Although the TPS method does not require any prior knowledge of a reaction coordinate or a set of CVs, it does require precise definitions of reactant and product states. To define these states, one should identify one or more low-dimensional order parameters along which stable basins can be clearly distinguished. In principle, the regions defining stable basins should be large enough to accommodate equilibrium fluctuations of the system and should not overlap to prevent sampling of non-reactive trajectories. Therefore, defining stable states is often a non-trivial task and it may require a substantial amount of trial to determine the degree to which a given criterion can be relaxed.

While there is no general rule to identify order parameters or CVs to distinguish between the reactant and product states, as it often depends on the specific process/problem under investigation, earlier studies on similar transition processes can provide useful inputs [33]. For example, in studies of chemical reactions, distances between the atoms involved in the bond formation/breaking are usually used as order parameters to identify the reactant and product states. In protein folding studies, number of native contacts, radius of gyration, number of backbone hydrogen bonds, and solvent accessible surface area are some common order parameters used to detect stable states. If unfolding of a protein is studied, then the distances between the hydrophobic residues or the number of water molecules within a cut-off distance of the hydrophobic residues can be considered to identify reactant and product states. For studying conformational fluctuations or

isomerization in a specific residue, torsional angles can be used to distinguish between reactant, product, and intermediate states.

2.1.2. Initial Reactive Trajectory

A key requirement in TPS is the need for an initial reactive trajectory connecting stable basins that will serve as a seed trajectory to generate subsequent transition paths. The initial transition path may be obtained by running a long conventional MD simulation, but often it turns out to be unsuccessful as the processes to be observed are statistically rare. There is no general recipe for obtaining an initial reactive trajectory, but biased or high-temperature MD simulations are often used to obtain the initial transition path. The trajectories generated in this way need not represent the TPE, because successive sampling during TPS moves these trajectories toward regions of high probability.

2.1.3. Transition Path Ensemble

Starting from an initial reactive trajectory, the TPS algorithm generates an ensemble of unbiased trajectories, connecting the predefined initial and final states, using a random walk MC procedure in trajectory space. First, a time-slice on the initial reactive trajectory is randomly selected and momenta of all atoms at this time-slice are slightly perturbed, keeping the positions of all atoms fixed. This step is followed by the conservation of total energy as well as total linear and angular momentum.

Through this procedure, a new point in the phase space called a shooting point is generated. Then, from the shooting point, a trial MD trajectory is propagated in both forward and backward directions in time. The new trial trajectory is accepted with a certain probability, only if it connects the initial and final states of the process, otherwise it is rejected. Once a new trajectory is generated, it becomes the old one and the same procedure is repeated. A set of properly weighted reactive trajectories form the TPE.

For a comprehensive statistical description of the TPE, we refer readers to the following article [34].

2.1.4. Identification of Transition States

The transition state of a process between the reactant state A and the product state B is identified from the TPE using the committor probability (p_B) calculation [35]. For a given configuration x , $p_B(x)$ is the fraction of trajectories initiated from x with random momenta reaching the state B. For transition state configurations, $p_B(x)$ is 1/2 because trajectories initiated from those points have equal probability to reach state A or state B. In TPS, configurations are extracted from the TPE and then a few hundreds of trajectories are generated from each of those configurations. Those configurations that show an equal probability to reach state A or B are identified as transition states.

A slightly different interpretation says that transition states are points in configuration space with the highest probability that trajectories passing through them are reactive [36,37]. The probability of a point (x) being on a transition path can be estimated through a Bayesian expression:

$$p(\text{TP}|x) = \frac{p(x|\text{TP})p(\text{TP})}{p_{eq}(x)}. \quad (1)$$

where, the conditional probability, $p(x|\text{TP})$, to find a point x on a given transition path can be calculated from the TPE, $p_{eq}(x)$ is the equilibrium distribution of x , and the normalizing factor $p(\text{TP})$ is the fraction of time spent in transition paths. Therefore, $p(\text{TP}|x)$ quantifies the “differences between the transition state average $[p(x|\text{TP})]$ and the stable average $[p_{eq}(x)]$ ” [34,36].

2.1.5. Calculation of Reaction Rates

As trajectories generated using TPS are truly dynamical, harvested without any bias, one can calculate the rate constant ($k_{A \rightarrow B}$) from the TPE. The $k_{A \rightarrow B}$ can be related to a time correlation function, $C(t)$ [38].

$$C(t) = \frac{\langle h_A(x_0)h_B(x_t) \rangle}{\langle h_A \rangle} \approx k_{A \rightarrow B}t \quad (2)$$

where $x_t = (q_t, p_t)$ is the state of the system at time t , and h_A/h_B are the characteristic functions of states A/B, where h_A or h_B would be 1 if the system is in state A or B, otherwise 0. For deterministic dynamics, the correlation function $C(t)$ is the probability of finding the system in state B at time t , provided that the system was in state A at $t = 0$. As the transition between states A and B is rare, the characteristic time of molecular motion (τ_{mol}) is much less than the typical reaction time (τ_{rxn}). Therefore, there exists a time region where $C(t)$ increases linearly; hence $C(t) \approx k_{A \rightarrow B}t$. The slope of this region gives the rate constant ($k_{A \rightarrow B}$) of the transition.

Transition interface sampling (TIS), a variant of TPS, is a more efficient way to obtain rates, especially when more than two stable or metastable states are involved in the transition [39,40]. For a given transition between states A and B, TIS divides the intermediate space into many interfaces using an order parameter λ and calculates the effective positive fluxes through the interfaces. The rate constant is then estimated using the following equation:

$$k_{A \rightarrow B} = \frac{\langle \Phi_{A, \lambda_1} \rangle}{h_A} \prod_{i=1}^{n-1} P(\lambda_{i+1} | \lambda_i) P(\lambda_B | \lambda_n). \quad (3)$$

The factor, $\frac{\langle \Phi_{A, \lambda_1} \rangle}{h_A}$, calculates the positive fluxes leaving the state A and reaching the first

interface λ_1 . $P(\lambda_i|\lambda_j)$ is the conditional probability of a trajectory reaching the interface λ_i , given that it comes from the state A and have crossed the interface λ_j . One advantage of TIS over TPS is that it allows path lengths to vary, thus becoming computationally less expensive and more applicable to diffusive dynamics. Unlike TPS, TIS ignores multiple recrossings at the transition state surface and concentrates only on the positive fluxes through the dividing interfaces for the rate calculation.

2.2. Applications

Since its development, TPS has been used to study a wide range of rare event problems. To highlight a few: ion pair dissociation [35], chemical dynamics of the protonated water [41], peptide isomerization [42], protein folding [43,44], DNA binding [45], catalysis [46–48], and conformational fluctuations of residues [49,50]. In the following discussion, we provide a brief description of biological applications of the TPS method in isomerization, folding, and catalysis, that share issues common to many (bio)physical problems.

(i) **Isomerization:** Alanine dipeptide has been used as a minimal model system to study conformational changes in biomolecules [51]. The conformational space of the dipeptide is often probed using the backbone dihedral angles (ϕ and ψ) that describe different states: C_{7eq} (in vacuum, $\phi \approx -86^\circ$ and $\psi \approx 68^\circ$; and in solution, $\phi \approx -80^\circ$ and $\psi \approx 160^\circ$), C_{ax} (in vacuum, $\phi \approx 50^\circ$ and $\psi \approx -50^\circ$), and α_R (in solution, $\phi \approx -80^\circ$ and $\psi \approx -30^\circ$). The transition between C_{7eq} and C_{ax} isomers of alanine dipeptide in vacuum and between C_{7eq} and α_R isomers in water has been studied using TPS [42].

In this study, all stable basins corresponding to configurations, C_{7eq} , C_{ax} and α_R were distinguished by ϕ and ψ angles of the dipeptide. For instance, in vacuum the C_{7eq} state was identified by $-150^\circ < \phi < -30^\circ$ and $0^\circ < \psi < 180^\circ$, and C_{ax}

state was identified by $30^\circ < \phi < 130^\circ$ and $-180^\circ < \psi < 0^\circ$. The initial trajectory was obtained by generating trajectories at high temperature from a configuration near the saddle point in forward and backward directions of time. Starting from the initial reactive trajectory, a collection of 1000 transition paths in vacuum and 256 transition paths in water were generated using the shooting and shifting algorithms [6]. Analyzing committor probability for configurations obtained from the transition state ensemble, it was shown that in both vacuum and water, only ϕ and ψ angles are insufficient to describe the progress of the transition, thereby revealing that another key variable, a torsional angle θ , should be incorporated to predict the correct dynamical pathway in vacuum. In solution, the solvent degrees of freedom should also be considered to understand the mechanism of transition. The rate constant $k_{C_{7eq} \rightarrow \alpha_R}$ was found to be 10 ns^{-1} .

- (ii) **Folding:** The folding pathways of a β -hairpin of the GB1 protein were investigated in explicit solvent using all-atom MD simulations and TPS [43]. This β -hairpin exhibits two metastable states ‘F’ and ‘H’ along with the completely folded (N) and unfolded states (U). These states were distinguished by several order parameters: number of native hydrogen bonds (N_{hb}), number of native contacts (N_{nc}), radius of gyration (R_g), number of broken native backbone hydrogen bonds (N_{nb}), the sum of the O-H distances of the backbone hydrogen bonds (R_{OH}), the minimum distance (d_{min}) between residues F52 and Y45 or W43, and the number of water molecules between these residues. These order parameters were first examined in a high temperature unfolding simulation to understand their contribution to different conformational states. More detailed definitions of all (meta)stable states (N, F, H and U) are provided in Ref. [43].

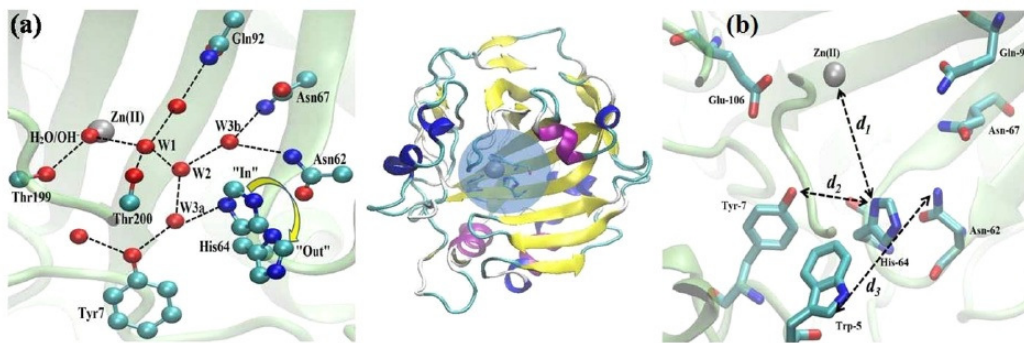


Figure 1. High-resolution crystal structure of HCA II (PDB: 2CBA [52], center) with (a) molecular model of the rate-determining proton transfer between the zinc-bound water and His-64 side chain mediated by hydrogen-bonded network of water molecules at the active site; (b) selected active site residues and key distances monitored in TPS. Adapted with permission from ref. [48]. Copyright 2018 American Chemical Society.

At 400K, an initial 2-ns long MD trajectory connecting states N and U and passing through metastable states F and H was generated. As TPS has been designed to study transitions between two states, the TPE for transitions between N and F, F and H, and H and U states were created separately. In the N to F transition, R_g slightly decreased and most of the hydrogen bonds remained intact. Two specific hydrogen bonds were observed to break simultaneously at state F, leading to state H. In the H to U transition, R_g increased and the distance between hydrophobic residues also increased, thereby creating space for water molecules.

The rate constant for the rate determining step, the F to H transition, was calculated using TIS. As F is a metastable state and can reach the N state within a few ns, the rate constant for N to H transition was calculated. The R_{OH} distance was used as an order parameter to make interfaces between states N and H and a few hundreds of paths were generated for each interface. The rate constant for unfolding at 300 K was found to be $0.20 \mu s^{-1}$, which was in good agreement with the experimental result ($0.17 \mu s^{-1}$).

(iii) **Catalysis:** A key system for applications in catalysis has been human carbonic anhydrase II (HCA II), a zinc-containing metalloenzyme, that catalyzes the reversible

hydration of carbon dioxide to bicarbonate. In the catalytic process, major steps include conduction of a proton from the zinc-bound water to a histidine residue (H64) and conformational changes in H64 in its protonated and unprotonated forms [53–57]. Specifically, H64 is known to exist in two (inward/outward) rotameric conformations, where the H64 side-chain faces toward or away from the active site. The important steps underlying the catalytic mechanism in HCA II have been extensively studied using a judicious combination of classical MD, the QM/MM approach, and TPS [48–50,58].

The first TPS study focused on the fluctuations of the unprotonated H64 [49], where two rotameric states of this residue were differentiated by a sidechain dihedral angle (χ_1). Starting from the initial reactive trajectory, created using the adaptive biasing force method [59], an ensemble of 150 transition paths was generated. A detailed inspection of the TPE revealed that the rotation of the H64 sidechain involves a narrow channel lined by W5 and N62 residues.

A subsequent TPS study [50], aimed to determine the reaction coordinate of the conformational transition in H64, used newer protocols including the aimless shooting version of TPS and the likelihood maximization technique (see Section 2.3). In this work, the initial trajectory was obtained from a 15-ns long conventional MD simulation. A total of 32 CVs (Fig. 1) comprising sidechain and backbone dihedral angles of seven active-site residues, three distance parameters, and the number of water molecules at the active site, were used. The reaction coordinate was found to be a linear combination of four CVs.

A more recent TPS study [48] on the fluctuations of protonated and unprotonated H64 used principle component analysis to reduce the dimensionality of the CV-set for free energy calculations along the optimum reaction coordinate and es-

259 timination of rates. The rotation of the protonated H64 was found to be 10 times
260 faster than that in the unprotonated state, largely due to electrostatic repulsion of
261 the protonated H64 resulting from the catalytic zinc-ion.

262 Using a combination of the QM/MM approach and TPS, this work also examined
263 the rate determining step of proton transfer from the zinc-bound water to the $N_{\delta 1}$
264 atom of H64. The stable states in this case were defined by an order parameter
265 called the mean path (MP) [48]. The initial trajectory was generated using steered
266 QM/MM MD simulations [60] and a total of 615 transition paths were generated
267 using the aimless shooting version of TPS. The initial set of 33 CVs were used to
268 optimize the reaction coordinate which was described by three principal modes. The
269 rate constant for the proton transfer step was found to be $1 \times 10^6 \text{ s}^{-1}$, in reasonable
270 agreement with the earlier experimental result [61].

271 The coupling of protein motions with the chemical events in catalysis have
272 also been examined using QM/MM and TPS for purine nucleoside phosphory-
273 lase (PNP) [62]. This enzyme catalyzes the reversible phosphorolysis of 6-oxypurine
274 deoxy-nucleoside to produce a purine base and the deoxy-ribose 1-phosphate [63,64].
275 This study probed the role of heavy isotopes in altering the transition state forma-
276 tion and the role of mutations in altering dynamics during catalysis [62]. Specifically,
277 it was found that the mutations in two neighboring residues (E258D and L261A)
278 removed the steric-hindrance to H257, thereby making it more prone to change con-
279 formation. The dynamics of the transition state formation were found to be restored
280 in the mutated PNP. Other applications of TPS for studying catalytic processes in-
281 clude studies on DNA polymerase β [65] and lactate dehydrogenase [46].

2.3. Reaction Coordinate and CV Determination

While the knowledge of reaction coordinates or CVs is not a requirement for applying the TPS method, it is feasible to develop approaches for determining these highly useful theoretical constructs; we refer readers to Ref. [19] for understanding subtle differences between order parameters, CVs, and the reaction coordinate. Briefly, transition path theory (TPT) [5] addresses these notions in full generality by defining the committor function that serves as an ideal reaction coordinate, i.e., it defines the probability that a trajectory initiated at a configuration x will first reach the product state B before reaching the reactant state A. The isosurfaces of the committor functions are known as isocommittor surfaces, where the isosurface with the committor value of 1/2 defines the surface with an equal probability of first reaching A or B. The committor function satisfies a backward Kolmogorov equation with boundary values between 0 and 1 to define reactant and product basins, respectively. While it is not possible to obtain committor by solving the backward Kolmogorov equation for high-dimensional complex molecular systems, several low-dimensional models of the committor can be constructed [19].

Based on the TPS method, a quantitative approach was developed by Peters *et al.* [66,67] to generate the reaction coordinate for a transition process, which differs from the original TPS approach in that, instead of perturbing momenta, they are freshly sampled from a Boltzmann distribution. This version of the TPS method is called aimless shooting. By design, aimless shooting creates most of the shooting points near the barrier region and helps in generating highly decorrelated trajectories.

To determine the reaction coordinate, the likelihood of a linear combination of a large number of CVs, computed at shooting points, are tested using the likelihood maximization method and finally the best combination is achieved by the Bayesian information criteria. To further improve the efficiency of the likelihood maximization method in obtaining

accurate transmission coefficient, inertial likelihood maximization (iLmax) method was developed [68,69]. iLmax utilizes the velocity information along with configurational coordinates to identify the optimum reaction coordinate. Recently, a two step screening method has been proposed to reduce the number of candidate CVs and thus making the task of likelihood maximization easier [48]. In this method, population distributions of CVs at the end points of shooting trajectories are calculated. Those CVs which show distributions at two different regions for two end points A and B are only selected for likelihood maximization. Both aimless shooting and the likelihood maximization methods have been used in several applications [50,58,70–73].

2.3.1. Relevance of Reactive Island Theory to Transition Path Sampling

A unique new perspective on the TPS method connects it to classical phase-space based reactive island (RI) theory in which reactive islands are manifolds emanating from a transition state and thereby mediate reaction pathways [74,75]. As a model system of reasonable complexity [75], the paradigmatic Müller-Brown Hamiltonian [76] was used to probe the nature of reactive trajectories via committor analysis and understanding its sensitivity to reactive islands. A new coordinate system based on normal modes at saddle points was adopted to unambiguously demarcate reactant and product states. It was argued that the RI hierarchy is intimately related to rare reactive trajectories because trajectories crossing the transition zone from the reactant to product states must sequentially pass through higher order to lower order RIs. Importantly, it was shown that committor functions are linked to the number and relative disposition of RIs relative to the shooting configurations. Although it is not feasible to visualize RIs in high-dimensional phase spaces of complex molecular systems, Lagrangian descriptors [77] (which are a measure of the arc length of the trajectory over the specified time-interval [75]) were found

to efficiently detect RI hierarchy. As a result, such descriptors are excellent candidates for designing new variants of TPS and could be used as CVs in the CV-based enhanced sampling methods described in the following sections.

3. CV-based Enhanced Sampling Methods

3.1. Overview

Free energy calculations are often carried out as a function of relevant coordinates. Given that the coordinates accurately quantifying dynamical evolution of the process (from the reactant to product states) are not known *a priori* in most cases, it is common practice to resolve the free energy (F) along a set of postulated collective coordinates, often termed as CVs [78], that are functions of the atomic Cartesian coordinates, *i.e.*, $\mathbf{s}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, where $\mathbf{r}_i = (x_i, y_i, z_i)$. Therefore, it is essential to sample all conformations accessible to the system for a given $\mathbf{s}$ to accurately compute $F(\mathbf{s})$, which is related to $P(\mathbf{s})$ [79], the probability distribution of $\mathbf{s}$ as: $F(\mathbf{s}) = -\beta^{-1} \ln P(\mathbf{s})$, where $\beta = (k_B T)^{-1}$ with k_B and T being the Boltzmann constant and temperature, respectively.

Given that the potential energy is a function of the full $3N$ -dimensional configuration space, $U(\mathbf{R})$, the configurations are visited with a probability, $P(\mathbf{R}) \propto \exp(-\beta U(\mathbf{R}))$. This means that the probability of visiting configurations higher in potential energy is significantly lower, and as a result, finite-length conventional MD trajectories remain non-ergodic in most cases. This problem can be alleviated by adding a bias potential ($V^b(\mathbf{R})$) that flattens large barriers separating metastable configurations, thereby making the biased probability distribution, $\tilde{P}(\mathbf{R}) \propto \exp(-\beta[U(\mathbf{R}) + V^b(\mathbf{R})])$; the biased distribution is related to the unbiased one as: $P(\mathbf{R}) = \tilde{P}(\mathbf{R}) \exp[\beta V^b(\mathbf{R})]$. Since obtaining $V^b(\mathbf{R})$ that flattens the potential energy landscape is highly challenging even for a moderate size system, one can incorporate the bias only along selected CVs and compute $F(\mathbf{s})$ via $P(\mathbf{s}) = \tilde{P}(\mathbf{s}) \exp[\beta V^b(\mathbf{s})]$.

3.1.1. Algorithms

The earlier technique of umbrella sampling [10] and the modern method of metadynamics [14] are methods of this flavor, where $V^b(s)$ has a harmonic potential form in the former but in latter it is constructed as a sum of Gaussian functions centered along the trajectory of $\mathbf{s}$. Other ways of obtaining $V^b(s)$ is in a variational manner, as in the variational enhanced sampling method [80], where it is constructed via a linear expansion using basis functions. Importantly, metadynamics is not only applicable for resolving free energy as a function of CVs but also for computing reaction rates [81–83].

Another way of enhancing sampling in MD is to utilize an extended Lagrangian approach in which CVs are coupled to a set of auxiliary/fictitious variables that dynamically evolve at a temperature higher than that of the physical system. Adiabatic separation between the physical and fictitious variables is achieved by increasing the mass of the fictitious variables and by introducing a higher fictitious friction coefficient as a thermostat parameter. This approach forms the basis for the TAMD method [17], which is closely related to the adiabatic/driven-adiabatic free energy dynamics (d-AFED) method [15,16,84]. A number of studies have employed the TAMD/d-AFED method to investigate several biophysical problems [85–96], and the method has been further improved by combining it with biased-sampling [97].

3.1.2. Practical Aspects

Some practical issues in applications of these CV-based methods need to be considered. For example, umbrella sampling is often used for a single CV, and rarely a full sampling of two-dimensional free energy landscapes is achieved [98], primarily because the efficiency of the method dramatically decreases with increasing dimensionality of the CV-set. In contrast, metadynamics is quite successful in sampling free energy landscapes in 2 or 3

CVs, but further increasing the dimensionality of CVs makes it challenging to obtain good sampling at a reasonable computational cost. Moreover, the self-guiding nature of the method pushes the system toward the direction of zero mean force, which may lead to sampling of uninteresting configurational states [99]. Therefore, a controlled sampling along a CV is difficult to achieve in metadynamics, as is also the case in the TAMD/d-AFED method. However, the TAMD/d-AFED method has the advantage that it can sample relatively high-dimensional free energy landscapes in an efficient manner [31]. For example, up to 700 CVs were simultaneously sampled using this method [100], thereby demonstrating the efficiency of this approach in exploring a high-dimensional CV space. This is primarily due to the fact that TAMD has been designed only for enhanced sampling of the underlying landscape without the need to reconstruct it. In the following, we discuss recent advances in these methods for computing free energies as a function of a high-dimensional CV-set.

3.2. Exploring High Dimensional Free Energy Landscapes

Although only few CVs may be sufficient to describe pertinent states (reactant, product, and transition) for a given process, hidden orthogonal coordinates may become crucial to efficiently sample conformational space for achieving convergence in free energy calculations [31]. In the following, we discuss techniques that aim to accomplish efficient conformational sampling in a high-dimensional CV-space. Specifically, we highlight bias-exchange and parallel-bias metadynamics, two variants of the metadynamics method, and a relatively recent method termed TASS.

3.2.1. *Bias-Exchange and Parallel-Bias Metadynamics Methods*

Bias-exchange metadynamics is a multiple-replica approach to sample a large number of CVs, where each replica is assigned to sample a low-dimensional CV-subset of the full CV-space using its own metadynamics bias potential [101]. Unlike replica-exchange MD [102], all replicas in bias-exchange metadynamics are maintained at the same temperature and exchanges between a pair of replicas follow a Metropolis-Hastings scheme although the exchange rates as well as convergence of the free energy can be enhanced via infinite-swapping or the Suwa-Todo algorithms [103]. We refer readers to Refs. [31,104] for additional methodological details. However, one drawback of this method is the requirement of exchanges between replicas, as this diminishes its performance. Some limitations of bias-exchange metadynamics are addressed in another variant of metadynamics, namely the parallel-bias metadynamics method [105]. Unlike bias-exchange metadynamics, this is a one replica method in which CVs are biased with one or two-dimensional biases. The heights of the Gaussian functions added along different CVs are scaled by a conditional factor in a manner that the bias potentials are balanced as they are built dynamically. A modification to this method has been recently proposed to increase its efficiency for sampling in higher dimensions [106].

3.2.2. *Applications of Bias-Exchange and Parallel-Bias Metadynamics*

Several applications of bias-exchange and parallel-bias metadynamics methods have been discussed in detail in Ref. [31] and therefore we only briefly highlight them here. The bias-exchange metadynamics method has been applied in studies of a number of biophysical problems including protein folding [107–115], ligand binding [116,117,117–123], conformational sampling [124–142], and similar studies in nucleic acid systems [143–146]. Similarly, the parallel-bias metadynamics method has been applied for studying protein

conformational sampling and ligand binding [105,147–149] as well as for studying complex chemical reaction pathways [150].

3.2.3. *Temperature Accelerated Sliced Sampling (TASS)*

A new approach, termed TASS [151], combines the advantages of umbrella sampling, metadynamics, and the TAMD/d-AFED method. In this method, the TAMD/d-AFED Lagrangian is modified by adding umbrella sampling and metadynamics biases on different CVs. Similar to the TAMD/d-AFED method, all CVs in a TASS simulation are coupled to fictitious variables dynamically evolving at a temperature higher than of the physical system.

The advantage of this method over conventional umbrella sampling, metadynamics, and the TAMD/d-AFED method is that a controlled exploration of a high-dimensional CV-space is possible. Particularly, free energy landscapes that are inherently flat and broad can be sampled efficiently with TASS [99]. Importantly, a high-dimensional free energy surface can be reconstructed by judiciously combining the weighted-histogram analysis method [11] and the Tiwary-Parrinello reweighting scheme [25] with the TAMD/d-AFED free energy estimator [79,152]. Therefore, TASS affords sampling of a large number of transverse coordinates and also allows usage of distinct orthogonal coordinates for different umbrella windows. It is also noted that the TASS method differs from the biased version of TAMD/d-AFED (UFED) [97], where all the CVs are biased with a high-dimensional biasing potential similar to metadynamics [31]. Owing to a controlled sampling as well as rapid convergence in free energies achievable by TASS, it is a potentially useful alternative to solely applying umbrella sampling, metadynamics, or TAMD/d-AFED for studying chemical reactions. In the following section, we highlight applications of the TASS method to biophysical systems.

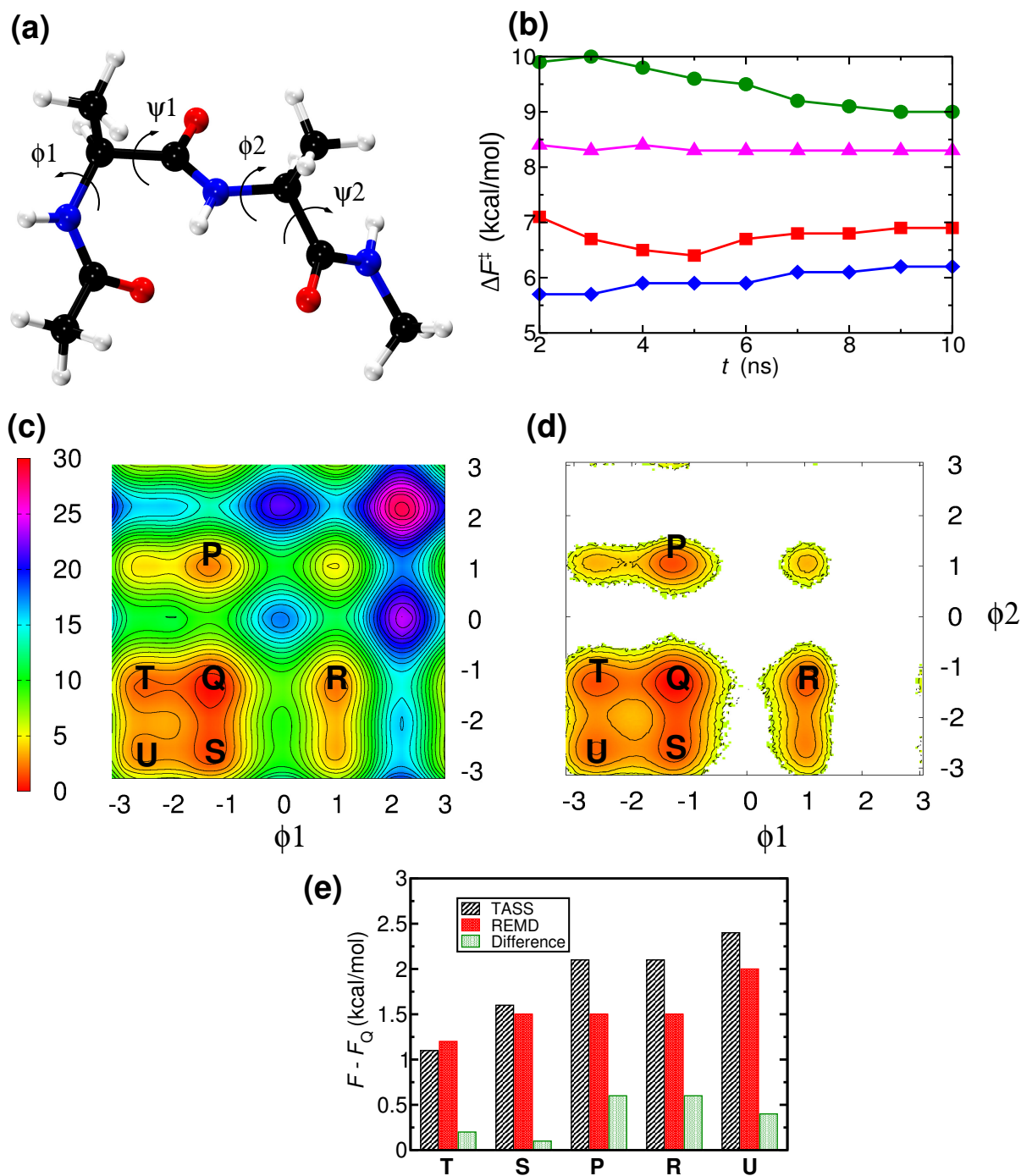


Figure 2. Free energy surface of alanine tripeptide (in vacuum) computed using the TASS method where ϕ_1 , ϕ_2 , ψ_1 and ψ_2 were chosen as the CVs. (a) Structure of alanine tripeptide with the definition of the CVs; (b) Convergence of free energy barriers for $P \rightarrow Q$, $Q \rightarrow P$, $R \rightarrow Q$, and $Q \rightarrow R$ are shown in $\blacksquare$, $\bullet$, $\blacklozenge$ and $\blacktriangle$, respectively; (c) Projection of $F(\phi_1, \psi_1, \phi_2, \psi_2)$ on the (ϕ_1, ϕ_2) space is shown; (d) $F(\phi_1, \phi_2)$ from parallel tempering simulation is shown for the reference. Contour lines are drawn for every 1 kcal mol⁻¹; and (e) Free energies of all the minima with respect to the free energy of Q from TASS and replica exchange simulations, and their differences are plotted. Reprinted from Awasthi S, Nair NN. Exploring high dimensional free energy landscapes: Temperature accelerated sliced sampling. J. Chem. Phys. 2017;146:094108, with the permission of AIP Publishing.

- (i) **Peptide conformational sampling:** While alanine dipeptide has been widely studied as a model biophysical system [51], alanine tripeptide (Figure 2a) has been studied to a limited extent. Therefore, the TASS method was applied to alanine tripeptide for resolving the free energy as a function of four Ramachandran torsional angles as CVs ($\phi_1, \psi_1, \phi_2, \psi_2$) (Figure 2a) [151]. The four fictitious/auxiliary variables corresponding to four torsional angles were kept at 900 K, while the physical system was maintained at 300 K. Here, ϕ_1 was biased using umbrella sampling (30 umbrella windows along ϕ_1 spanning $-\pi$ to $+\pi$) and ϕ_2 was biased using metadynamics. The reconstructed free energy surface $F(\phi_1, \psi_1, \phi_2, \psi_2)$ was then projected on the ϕ_1 - ϕ_2 sub-space for analyzing the convergence in free energies (Figure 2b,c). A reasonable convergence was achieved using an 8 ns long simulation per umbrella window. Importantly, even the regions with higher free energy values ($\phi_1 \in [1.5, 3.1]$) were efficiently sampled in TASS by virtue of the umbrella bias along ϕ_1 . In fact, μ s-scale parallel tempering simulations could not sample these regions of the free energy surface (Figure 2d). The free energy differences between various minima agreed very well with the results from the parallel tempering (Figure 2e) simulations, which in turn ascertained the accuracy of the method.
- (ii) **Chemical reactions in enzymes:** The TASS method has been combined with the density functional theory (DFT) based QM/MM MD methods to study chemical reactions in enzymatic systems [153]. At the DFT level, QM/MM based MD simulations are computationally intensive. As a result, enhanced sampling methods are desired to obtain quick convergence in free energies (typically within 10-20 ps). For the case of bond-formation reactions or the A+B type chemical reactions, a large number of conformational states can be representative of the reactant basin [99].

Therefore, an umbrella sampling bias along the bond-formation coordinate is preferred to efficiently sample the relevant conformational states in the reactant basin. However, in complex chemical reactions, several orthogonal coordinates also need to be sampled to get a converged free energy along the bond-formation CV. Thus, the TASS method is ideally suited for modeling such reactions [31].

In Ref. [153], the deacylation reaction of a covalent bond between the ring-opened aztreonam drug and a class C β -lactamase enzyme was modeled. Four CVs were chosen to study this reaction, among which the distance between the deacylating water oxygen and the carbonyl C of the substrate was biased using umbrella sampling to drive the reaction in a controlled manner. Here, the umbrella bias is computationally efficient as it avoids the sampling of those conformations in which water is far from the reaction site. In addition, metadynamics was used to bias the sampling of the proton transfer from the attacking water to the phenolic oxygen of Tyr150, while the relative orientations of Tyr150 and the two adjacent lysine residues (Lys67 and Lys315) were sampled only by TAMD/d-AFED. A satisfactory convergence in free energy was obtained within 8 ps of simulation per umbrella window and the computed barrier was in good agreement with the free energies computed from the experimental kinetic data. In addition, several other QM/MM based TASS simulations were reported by choosing up to 8 CVs [153].

(iii) **Product Release in Enzymes:** It was experimentally observed [154] that one of the active site Mg^{2+} ions is getting discharged with the pyrophosphate after the nucleotidyl transfer reaction in a sugar nucleotidyltransferases (GlmU). To validate this observation, Vithani et al. [154] first performed conventional umbrella sampling simulations by using as a CV the distance between the centre-of-mass of the releasing pyrophosphate and the active site (Figure 3a,b). Similar to experiments, the release

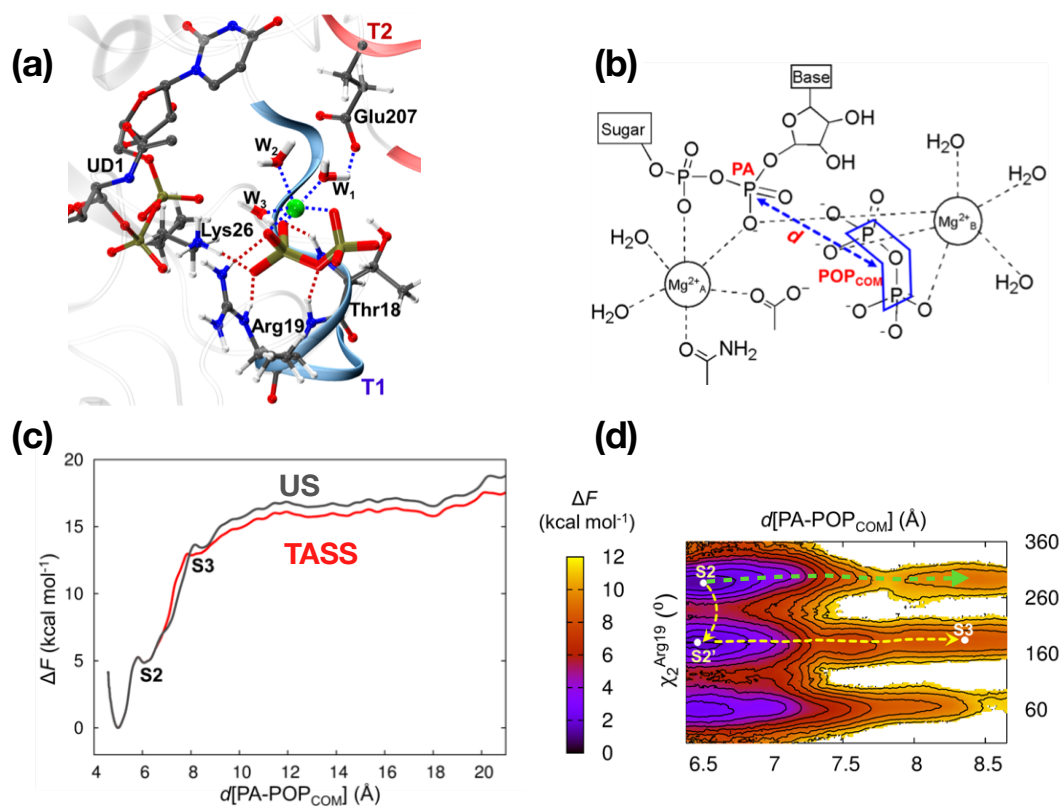


Figure 3. (a) Active site structure of pyrophosphate release in GlmU is shown in the intermediate state (S2). Interactions of pyrophosphate (red colored dotted line) and Mg²⁺ (blue colored dotted line) are depicted. Pyrophosphate is shown in stick representation, while Mg²⁺ is shown as a green colored sphere. Thr18, Arg19, Lys26, UD1 and water molecules (W₁, W₂ and W₃) are shown in ball-stick representation. Color code: gray (carbon), red (oxygen), blue (nitrogen), green (Mg²⁺) and white (hydrogen). T1 and T2 loop regions are shown by blue and red colored ribbons, respectively. (b) The distance CV ($d[\text{PA-POP}_{\text{COM}}]$) used for US is shown. Here POP_{COM} is the centre of mass of the P atoms and the bridging O atom in the pyrophosphate residue. (c) Free energy along $d[\text{PA-POP}_{\text{COM}}]$ from US and TASS simulations are compared. (d) Free energy surface resolved along $d[\text{PA-POP}_{\text{COM}}]$ and χ_2^{Arg19} (defining rotation around C_γ–C_δ bond of Arg19 side chain), as computed from the TASS simulation. The minimum energy taken identified from TASS is shown in yellow dotted lines while that explored in US is shown in green dotted lines. Reprinted from Structure, 26/3, Vithani et al., Mechanism of Mg²⁺-accompanied product release in sugar nucleotidyltransferases, 459-466, Copyright (2018), with permission from Elsevier [154].

of the Mg^{2+} ion with the pyrophosphate was also observed in these simulations, but the free-energy barrier was computed to be 18 kcal/mol. Such a high energy barrier is unusual for the product release step in enzymatic reactions. Therefore, it is likely an artefact of the umbrella sampling simulations resulting from a poor sampling of other relevant orthogonal coordinates.

To test this hypothesis, TASS simulations were performed by using 10 CVs, which included the torsional angles of an Arg residue along the pathway of the ligand dissociation (Figure 3a). TASS sampled several pathways for the product release, including the one observed in umbrella sampling simulations. Importantly, the lowest energy pathway found by TASS was 2 kcal/mol lower in free energy than the pathway in umbrella sampling simulations (Figure 3c,d). This difference was mainly attributed to a poor sampling of Arg19 conformational states in umbrella sampling simulations. TASS simulations also reaffirmed the view that the product release is a relatively slow process and the origin of this delayed release was ascribed to interactions with the Arg residue located at the exit along the product release pathway [154].

The TASS method has also been extended for studying chemical reactions in zeolites using the QM/MM approach [155]. While TASS has not been yet applied to study many other biophysical problems (e.g. ligand binding and protein folding), the method is potentially applicable and provides opportunities in future to study a broader class of biophysical problems.

4. Conclusions

In this review, we have highlighted methodological details and biological applications of phase space and CV-based methods for studying thermodynamics and kinetics of rare

biochemical and biophysical events. Specifically, we have discussed key ingredients and applications of the TPS method, a phase space technique for discovering reactive trajectories and computing reaction kinetics. Further highlighted are links of the committor functions to reactive island theory and the emerging concept of Lagrangian descriptors as a model for detecting island hierarchy. We then describe bias-exchange metadynamics and parallel-bias metadynamics as two CV-based methods utilizing a replica approach. We end the review by discussing details of a hybrid technique termed TASS, which combines the advantages of multiple and distinct CV-based methods including umbrella sampling, metadynamics, and the TAMD/d-AFED method.

Disclosure statement

Authors declare no conflicts of interest.

Funding

We are grateful for financial support provided by National Science Foundation (awards CBET-1554558 and OIA-1757371; HV).

References

- [1] Eyring H. The activated complex in chemical reactions. *J Chem Phys.* 1935;3(2):107–115.
- [2] Wigner E. The transition state method. *Trans Faraday Soc.* 1938;34:678.
- [3] Chandler D. Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J Chem Phys.* 1978;68(6):2959–2970.
- [4] Truhlar DG, Garrett BC. Variational transition state theory. *Annu Rev Phys Chem.* 1984;35(1):159–189.
- [5] E W, Vanden-Eijnden E. Transition-path theory and path-finding algorithms for the study of rare events. *Annu Rev Phys Chem.* 2010;61:391–420.
- [6] Dellago C, Bolhuis PG, Csajka FS, et al. Transition path sampling and the calculation of rate constants. *J Chem Phys.* 1998;108(5):1964–1977.
- [7] Bolhuis PG, Chandler D, Dellago C, et al. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu Rev Phy Chem.* 2002;53(1):291–318.
- [8] Dellago C, Bolhuis PG. Transition path sampling simulations of biological systems. In: *Atomistic approaches in modern biology.* Springer; 2006. p. 291–317.
- [9] Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Mol Biol.* 2002;9(9):646.
- [10] Torrie G, Valleau J. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J Comput Phys.* 1977;23(2):187–199.
- [11] Kumar S, Rosenberg JM, Bouzida D, et al. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem.* 1992;13(8):1011–1021.
- [12] Souaille M, Roux B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun.* 2001;135(1):40–57.
- [13] Rosta E, Hummer G. Free energies from dynamic weighted histogram analysis using unbiased markov state model. *J Chem Theory Comput.* 2015;11(1):276–285.
- [14] Laio A, Parrinello M. Escaping free-energy minima. *Proc Natl Acad Sci USA.* 2002;99(20):12562–12566.
- [15] Rosso L, Mináry P, Zhu Z, et al. On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. *J Chem Phys.* 2002;116:4389.

- [16] Rosso L, Abrams JB, Tuckerman ME. Mapping the backbone dihedral free-energy surfaces in small peptides in solution using adiabatic free-energy dynamics. *J Phys Chem B*. 2005;109:4162.
- [17] Maragliano L, Vanden-Eijnden E. A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. *Chem Phys Lett*. 2006;426(1):168–175.
- [18] Rohrdanz MA, Zheng W, Clementi C. Discovering mountain passes via torchlight: Methods for the definition of reaction coordinates and pathways in complex macromolecular reactions. *Annu Rev Phys Chem*. 2013;64(1):295–316.
- [19] Peters B. Recent advances in transition path sampling: accurate reaction coordinates, likelihood maximisation and diffusive barrier-crossing dynamics. *Mol Sim*. 2010;36(15):1265–1281.
- [20] Peters B. Reaction coordinates and mechanistic hypothesis tests. *Annual Rev Phys Chem*. 2016;67(1):669–690.
- [21] Vashisth H, Skiniotis G, Brooks III CL. Collective variable approaches for single molecule flexible fitting and enhanced sampling. *Chem Rev*. 2014;114(6):3353–3365.
- [22] Wales DJ. Exploring energy landscapes. *Annu Rev Phys Chem*. 2018;69(1):401–425.
- [23] Henkelman G. Atomistic simulations of activated processes in materials. *Annu Rev Mater Res*. 2017;47(1):199–216.
- [24] Zuckerman DM, Chong LT. Weighted ensemble simulation: Review of methodology, applications, and software. *Annu Rev Biophys*. 2017;46(1):43–57.
- [25] Valsson O, Tiwary P, Parrinello M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annual Rev Phys Chem*. 2016;67(1):159–184.
- [26] Allen RJ, Valeriani C, ten Wolde PR. Forward flux sampling for rare event simulations. *J Phys Condens Matter*. 2009;21(46):463102.
- [27] Noé F, Fischer S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr Opin Struct Biol*. 2008;18(2):154–162.
- [28] Meloni S, Ciccotti G. Free energies for rare events: Temperature accelerated MD and MC. *Eur Phys J Spec Top*. 2015 Sep;224(12):2389–2407.
- [29] Chong LT, Saglam AS, Zuckerman DM. Path-sampling strategies for simulating rare events in biomolecular systems. *Curr Opin Struct Biol*. 2017;43:88–94.
- [30] Zinovjev K, Tun I. Reaction coordinates and transition states in enzymatic catalysis. Wiley Inter-

discip Rev Comput Mol Sci. 2018;8(1):e1329.

[31] Awasthi S, Nair NN. Exploring high-dimensional free energy landscapes of chemical reactions. Wiley Interdiscip Rev Comput Mol Sci. 2018;0(0):e1398.

[32] West AMA, Elber R, Shalloway D. Extending molecular dynamics time scales with milestone: Example of complex kinetics in a solvated peptide. J Chem Phys. 2007;126(14):145104.

[33] Pietrucci F. Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead. Rev Phys. 2017;2:32–45.

[34] Dellago C, Bolhuis PG, Geissler PL. Transition path sampling. Adv Chem Phys. 2002;123:1–78.

[35] Geissler PL, Dellago C, Chandler D. Kinetic pathways of ion pair dissociation in water. J Phys Chem B. 1999;103(18):3706–3710.

[36] Hummer G. From transition paths to transition states and rate coefficients. J Chem Phys. 2004; 120(2):516–523.

[37] Best RB, Hummer G. Reaction coordinates and rates from transition paths. Proc Natl Acad Sci USA. 2005;102(19):6732–6737.

[38] Dellago C, Bolhuis PG, Chandler D. On the calculation of reaction rate constants in the transition path ensemble. J Chem Phys. 1999;110(14):6617–6625.

[39] van Erp TS, Moroni D, Bolhuis PG. A novel path sampling method for the calculation of rate constants. J Chem Phys. 2003;118(17):7762–7774.

[40] van Erp TS, Bolhuis PG. Elaborating transition interface sampling methods. J Comput Phys. 2005; 205(1):157–181.

[41] L Geissler P, Dellago C, Chandler D. Chemical dynamics of the protonated water trimer analyzed by transition path sampling. Phys Chem Chem Phys. 1999;1:1317–1322.

[42] Bolhuis PG, Dellago C, Chandler D. Reaction coordinates of biomolecular isomerization. Proc Natl Acad Sci USA. 2000;97(11):5877–5882.

[43] Bolhuis PG. Transition-path sampling of β -hairpin folding. Proc Natl Acad Sci USA. 2003; 100(21):12129–12134.

[44] Juraszek J, Bolhuis PG. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. Proc Natl Acad Sci USA. 2006;103(43):15859–15864.

[45] Hagan MF, Dinner AR, Chandler D, et al. Atomistic understanding of kinetic pathways for single

- base-pair binding and unbinding in dna. *Proc Natl Acad Sci USA*. 2003;100(24):13922–13927.
- [46] Basner JE, Schwartz SD. How enzyme dynamics helps catalyze a reaction in atomic detail: A transition path sampling study. *J Am Chem Soc*. 2005;127(40):13822–13831.
- [47] Quaytman SL, Schwartz SD. Reaction coordinate of an enzymatic reaction revealed by transition path sampling. *Proc Natl Acad Sci USA*. 2007;104(30):12253–12258.
- [48] Paul S, Paul TK, Taraphder S. Reaction coordinate, free energy, and rate of intramolecular proton transfer in human carbonic anhydrase II. *J Phys Chem B*. 2018;122(11):2851–2866.
- [49] Roy A, Taraphder S. Transition path sampling study of the conformational fluctuation of his-64 in human carbonic anhydrase II. *J Phys Chem B*. 2009;113(37):12555–12564.
- [50] Paul S, Taraphder S. Determination of the reaction coordinate for a key conformational fluctuation in human carbonic anhydrase II. *J Phys Chem B*. 2015;119(34):11403–11415.
- [51] Feig M. Is alanine dipeptide a good model for representing the torsional preferences of protein backbones? *J Chem Theory Comput*. 2008;4(9):1555–1564.
- [52] Hkansson K, Carlsson M, Svensson L, et al. Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. *J Mol Biol*. 1992;227(4):1192–1204.
- [53] Silverman DN, Lindskog S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc Chem Res*. 1988;21(1):30–36.
- [54] Maupin CM, Voth GA. Preferred orientations of his64 in human carbonic anhydrase II. *Biochemistry*. 2007;46(11):2938–2947.
- [55] Maupin CM, McKenna R, Silverman DN, et al. Elucidation of the proton transport mechanism in human carbonic anhydrase II. *J Am Chem Soc*. 2009;131(22):7598–7608.
- [56] Maupin CM, Castillo N, Taraphder S, et al. Chemical rescue of enzymes: Proton transfer in mutants of human carbonic anhydrase II. *J Am Chem Soc*. 2011;133(16):6223–6234.
- [57] Taraphder S, Maupin CM, Swanson JMJ, et al. Coupling protein dynamics with proton transport in human carbonic anhydrase II. *J Phys Chem B*. 2016;120(33):8389–8404.
- [58] Paul S, Paul TK, Taraphder S. Orthogonal order parameters to model the reaction coordinate of an enzyme catalyzed reaction. *J Mol Graph Model*. 2019;90:18–32.
- [59] Hnin J, Chipot C. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *The Journal of Chemical Physics*. 2004;121(7):2904–2914.

- [60] Bonomi M, Branduardi D, Bussi G, et al. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun.* 2009;180(10):1961 – 1972.
- [61] Duda D, Tu C, Qian M, et al. Structural and kinetic analysis of the chemical rescue of the proton transfer function of carbonic anhydrase II. *Biochemistry.* 2001;40(6):1741–1748.
- [62] Zoi I, Suarez J, Antoniou D, et al. Modulating enzyme catalysis through mutations designed to alter rapid protein dynamics. *J Am Chem Soc.* 2016;138(10):3403–3409.
- [63] Saen-oon S, Quaytman-Machleder S, Schramm VL, et al. Atomic detail of chemical transformation at the transition state of an enzymatic reaction. *Proc Natl Acad Sci USA.* 2008;105(43):16543–16548.
- [64] Antoniou D, Ge X, Schramm VL, et al. Mass modulation of protein dynamics associated with barrier crossing in purine nucleoside phosphorylase. *J Phys Chem Lett.* 2012;3(23):3538–3544.
- [65] Radhakrishnan R, Schlick T. Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β 's closing. *Proc Natl Acad Sci USA.* 2004;101(16):5970–5975.
- [66] Peters B, Trout BL. Obtaining reaction coordinates by likelihood maximization. *J Chem Phys.* 2006;125(5):054108.
- [67] Peters B, Beckham GT, Trout BL. Extensions to the likelihood maximization approach for finding reaction coordinates. *J Chem Phys.* 2007;127(3):034109.
- [68] Peters B. Inertial likelihood maximization for reaction coordinates with high transmission coefficients. *Chem Phys Lett.* 2012;554:248–253.
- [69] Joswiak MN, Doherty MF, Peters B. Ion dissolution mechanism and kinetics at kink sites on NaCl surfaces. *Proc Natl Acad Sci USA.* 2018;115(4):656–661.
- [70] Xi L, Shah M, Trout BL. Hopping of water in a glassy polymer studied via transition path sampling and likelihood maximization. *J Phys Chem B.* 2013;117(13):3634–3647.
- [71] Leitold C, Dellago C. Folding mechanism of a polymer chain with short-range attractions. *J Chem Phys.* 2014;141(13):134901.
- [72] Lechner W, Dellago C, Bolhuis PG. Reaction coordinates for the crystal nucleation of colloidal suspensions extracted from the reweighted path ensemble. *J Chem Phys.* 2011;135(15):154110.
- [73] Knott BC, Haddad Momeni M, Crowley MF, et al. The mechanism of cellulose hydrolysis by a

two-step, retaining cellobiohydrolase elucidated by structural and transition path sampling studies. J Am Chem Soc. 2014;136(1):321–329.

[74] Patra S, Keshavamurthy S. Classical-quantum correspondence in a model for conformational dynamics: Connecting phase space reactive islands with rare events sampling. Chem Phys Lett. 2015; 634:1–10.

[75] Patra S, Keshavamurthy S. Detecting reactive islands using lagrangian descriptors and the relevance to transition path sampling. Phys Chem Chem Phys. 2018;20:4970–4981.

[76] Müller K, Brown LD. Location of saddle points and minimum energy paths by a constrained simplex optimization procedure. Theor Chim Acta. ????.

[77] Mendoza C, Mancho AM. Hidden geometry of ocean flows. Phys Rev Lett. 2010;105(3):038501.

[78] Fiorin G, Klein ML, Hénin J. Using collective variables to drive molecular dynamics simulations. Mol Phys. 2013;111(22-23):3345–3362.

[79] Tuckerman ME. Statistical mechanics: Theory and molecular simulation. 1st ed. Oxford: Oxford University Press; 2010.

[80] Valsson O, Parrinello M. Variational approach to enhanced sampling and free energy calculations. Phys Rev Lett. 2014;113:090601.

[81] Tiwary P, Parrinello M. From metadynamics to dynamics. Phys Rev Lett. 2013;111(23):230602.

[82] Salvalaglio M, Tiwary P, Parrinello M. Assessing the reliability of the dynamics reconstructed from metadynamics. J Chem Theory Comput. 2014;10(4):1420–1425.

[83] Tiwary P, Limongelli V, Salvalaglio M, et al. Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. Proc Natl Acad Sci USA. 2015;112(5):E386–E391.

[84] Abrams JB, Tuckerman ME. Efficient and direct generation of multidimensional free energy surfaces via adiabatic dynamics without coordinate transformations. J Phys Chem B. 2008;112:15742.

[85] Abrams CF, Vanden-Eijnden E. Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proc Nat Acad Sci USA. 2010;107:4961–4966.

[86] Vashisth H, Maragliano L, Abrams CF. Dfg-flip in the insulin receptor kinase is facilitated by a helical intermediate state of the activation loop. Biophys J. 2012;102:1979 – 1987.

[87] Vashisth H, Brooks III CL. Conformational sampling of maltose-transporter components in cartesian collective variables is governed by the low-frequency normal modes. J Phys Chem Lett. 2012;

3(22):3379–3384.

- [88] Vashisth H, Skiniotis G, Brooks III CL. Using enhanced sampling and structural restraints to refine atomic structures into low-resolution electron microscopy maps. *Structure*. 2012;20(9):1453–1462.
- [89] Vashisth H, Storaska AJ, Neubig RR, et al. Conformational dynamics of a regulator of G-protein signaling protein reveals a mechanism of allosteric inhibition by a small molecule. *ACS Chem Biol*. 2013;8(12):2778–2784.
- [90] Maragliano L, Cottone G, Ciccotti G, et al. Mapping the network of pathways of CO diffusion in myoglobin. *J Am Chem Soc*. 2010;132:1010–1017.
- [91] Lapelosa M, Abrams CF. A computational study of water and co migration sites and channels inside myoglobin. *J Chem Theory Comput*. 2013;9:1265–1271.
- [92] Tzanov AT, Cuendet MA, Tuckerman ME. How accurately do current force fields predict experimental peptide conformations? an adiabatic free energy dynamics study. *J Phys Chem B*. 2014; 118:6539–6552.
- [93] Cortes-Ciriano I, Bouvier G, Nilges M, et al. Temperature accelerated molecular dynamics with soft-ratcheting criterion orients enhanced sampling by low-resolution information. *J Chem Theory Comput*. 2015;11:3446–3454.
- [94] Cuendet MA, Margul DT, Schneider E, et al. Endpoint-restricted adiabatic free energy dynamics approach for the exploration of biomolecular conformational equilibria. *J Chem Phys*. 2018; 149:072316.
- [95] Abrams JB, Rosso L, Tuckerman ME. Efficient and precise solvation free energies via alchemical adiabatic molecular dynamics. *J Chem Phys*. 2006;125:074115.
- [96] Cuendet MA, Tuckerman ME. Alchemical free energy differences in flexible molecules from thermodynamic integration or free energy perturbation combined with driven adiabatic dynamics. *J Chem Theory Comput*. 2012;8:3504–3512.
- [97] Chen M, Cuendet MA, Tuckerman ME. Heating and flooding: A unified approach for rapid generation of free energy surfaces. *J Comp Phys*. 2012;137:024102.
- [98] Kästner J. Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci*. 2011;1(6):932–942.
- [99] Awasthi S, Kapil V, Nair NN. Sampling free energy surfaces as slices by combining umbrella sampling and metadynamics. *J Comp Chem*. 2016;37:1413.

- [100] Lucid J, Meloni S, MacKernan D, et al. Probing the structures of hydrated nafion in different morphologies using temperature-accelerated molecular dynamics simulations. *J Phys Chem C*. 2013;117:774–782.
- [101] Piana S, Laio A. A bias-exchange approach to protein folding. *J Phys Chem B*. 2007;111(17):4553–4559.
- [102] Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett*. 1999;314(1-2):141–151.
- [103] Galvelis R, Sugita Y. Replica state exchange metadynamics for improving the convergence of free energy estimates. *J Comput Chem*. 2015;36(19):1446–1455.
- [104] Baftizadeh F, Cossio P, Pietrucci F, et al. Protein folding and ligand-enzyme binding from bias-exchange metadynamics simulations. *Curr Phys Chem*. 2012;2:79–91.
- [105] Pfaendtner J, Bonomi M. Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. *J Chem Theory Comput*. 2015;11:5062.
- [106] Prakash A, Fu CD, Bonomi M, et al. Biasing smarter, not harder, by partitioning collective variables into families in parallel bias metadynamics. *J Chem Theory Comput*. 2018;14(10):4985–4990.
- [107] Piana S, Laio A. Advillin folding takes place on a hypersurface of small dimensionality. *Phys Rev Lett*. 2008;101:208101.
- [108] Piana S, Laio A, Marinelli F, et al. Predicting the effect of a point mutation on a protein fold: The villin and advillin headpieces and their Pro62Ala mutants. *J Mol Biol*. 2008;375:460 – 470.
- [109] Todorova N, Marinelli F, Piana S, et al. Exploring the folding free energy landscape of insulin using bias exchange metadynamics. *J Phys Chem B*. 2009;113:3556–3564.
- [110] Pietrucci F, Laio A. A collective variable for the efficient exploration of protein beta-sheet structures: Application to SH3 and GB1. *J Chem Theory Comput*. 2009;5:2197–2201.
- [111] Juraszek J, Bolhuis PG. (un)folding mechanisms of the FBP28 WW domain in explicit solvent revealed by multiple rare event simulation methods. *Biophys J*. 2010;98:646–656.
- [112] Cossio P, Marinelli F, Laio A, et al. Optimizing the performance of bias-exchange metadynamics: Folding a 48-residue lysm domain using a coarse-grained model. *J Phys Chem B*. 2010;114:3259–3265.
- [113] Rossetti G, Cossio P, Laio A, et al. Conformations of the huntingtin n-term in aqueous solution

from atomistic simulations. FEBS Letters. 2011;585:3086–3089.

[114] Granata D, Camilloni C, Vendruscolo M, et al. Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. Proc Nat Acad Sci USA. 2013;110:6817–6822.

[115] Benetti F, Biarnes X, Attanasio F, et al. Structural determinants in prion protein folding and stability. J Mol Biol. 2014;426:3796–3810.

[116] Pietrucci F, Marinelli F, Carloni P, et al. Substrate binding mechanism of HIV-1 protease from explicit-solvent atomistic simulations. J Am Chem Soc. 2009;131:11811–11818.

[117] Herbert C, Schieberr U, Saxena K, et al. Molecular mechanism of SSR128129E, an extracellularly acting, small-molecule, allosteric inhibitor of FGF receptor signaling. Cancer Cell. 2013;23:489 – 501.

[118] Li Y, Lavey NP, Coker JA, et al. Consequences of depsipeptide substitution on the ClpP activation activity of antibacterial acyldepsipeptides. ACS Med Chem Lett. 2017;8:1171–1176.

[119] Duan M, Liu N, Zhou W, et al. Structural diversity of ligand-binding androgen receptors revealed by microsecond long molecular dynamics simulations and enhanced sampling. J Chem Theory Comput. 2016;12:4611–4619.

[120] Paloncýová M, Navrátilová V, Berka K, et al. Role of enzyme flexibility in ligand access and egress to active site: Bias-exchange metadynamics study of 1,3,7-trimethyluric acid in cytochrome P450 3A4. J Chem Theory Comput. 2016;12:2101–2109.

[121] Zhu L, Jiang H, Sheong FK, et al. A flexible domain-domain hinge promotes an induced-fit dominant mechanism for the loading of guide-dna into argonaute protein in thermus thermophilus. J Phys Chem B. 2016;120:2709–2720.

[122] Marinelli F, Kuhlmann SI, Grell E, et al. Evidence for an allosteric mechanism of substrate release from membrane-transporter accessory binding proteins. Proc Nat Acad Sci USA. 2011;108:E1285–E1292.

[123] Bisha I, Rodriguez A, Laio A, et al. Metadynamics simulations reveal a Na⁺ independent exiting path of galactose for the inward-facing conformation of vSGLT. PLoS Comput Biol. 2014 12;10:1–8.

[124] Michel J, Cuchillo R. The impact of small molecule binding on the energy landscape of the intrinsically disordered protein c-myc. PLoS One. 2012 07;7:1–13.

[125] Damas JM, Filipe LC, Campos SR, et al. Predicting the thermodynamics and kinetics of helix

formation in a cyclic peptide model. *J Chem Theory Comput.* 2013;9:5148–5157.

[126] Mahajan SP, Velez-Vega C, Escobedo FA. Tilting the balance between canonical and noncanonical conformations for the h1 hypervariable loop of a llama VHH through point mutations. *J Phys Chem B.* 2013;117:13–24.

[127] Do TN, Choy WY, Karttunen M. Accelerating the conformational sampling of intrinsically disordered proteins. *J Chem Theory Comput.* 2014;10:5081–5094.

[128] Zerze GH, Miller CM, Granata D, et al. Free energy surface of an intrinsically disordered protein: Comparison between temperature replica exchange molecular dynamics and bias-exchange metadynamics. *J Chem Theory Comput.* 2015;11:2776–2782.

[129] Camilloni C, Vendruscolo M. Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics. *J Am Chem Soc.* 2014;136:8982–8991.

[130] Camilloni C, Vendruscolo M. Using pseudocontact shifts and residual dipolar couplings as exact nmr restraints for the determination of protein structural ensembles. *Biochemistry.* 2015;54:7470–7476.

[131] cheng Chiu C, Singh S, de Pablo J. Effect of proline mutations on the monomer conformations of amylin. *Biophys J.* 2013;105:1227 – 1235.

[132] Pietrucci F, Mollica L, Blackledge M. Mapping the native conformational ensemble of proteins from a combination of simulations and experiments: New insight into the src-sh3 domain. *J Phys Chem Lett.* 2013;4:1943–1948.

[133] Corbi-Verge C, Marinelli F, Zafra-Ruano A, et al. Two-state dynamics of the SH3–SH2 tandem of abl kinase and the allosteric role of the n-cap. *Proc Natl Acad Sci USA.* 2013;110:E3372–E3380.

[134] Buchanan LE, Dunkelberger EB, Tran HQ, et al. Mechanism of iapp amyloid fibril formation involves an intermediate with a transient-sheet. *Proc Nat Acad Sci USA.* 2013;110:19285–19290.

[135] Scarabelli G, Provasi D, Negri A, et al. Bioactive conformations of two seminal delta opioid receptor penta-peptides inferred from free-energy profiles. *Biopolymers.* 2014;101(1):21–27.

[136] Ceballos JA, Giraldo MA, Cossio P. Effects of a disulfide bridge prior to amyloid formation of the ABRI peptide. *RSC Adv.* 2014;4:36923–36928.

[137] Li D, Liu M, Ji B. Mapping the dynamics landscape of conformational transitions in enzyme: The adenylate kinase case. *Biophys J.* 2015;109:647 – 660.

[138] Hoffmann KQ, McGovern M, Chiu Cc, et al. Secondary structure of rat and human amylin across force fields. *PLoS ONE*. 2015 07;10(7):1–24.

[139] Gómez-Sicilia À, Sikora M, Cieplak M, et al. An exploration of the universe of polyglutamine structures. *PLoS Comp Biol*. 2015;11(10):e1004541.

[140] Cristina P, Michela G, Laura B, et al. Metadynamics simulations rationalise the conformational effects induced by n-methylation of RGD cyclic hexapeptides. *Chem Eur J*. 2015;21:14165–14170.

[141] Camilloni C, Sala BM, Sormanni P, et al. Rational design of mutations that change the aggregation rate of a protein while maintaining its native structure and stability. *Sci Rep*. 2016;6:25559.

[142] Singh R, Bansal R, Rathore AS, et al. Equilibrium ensembles for insulin folding from bias-exchange metadynamics. *Biophys J*. 2017;112(8):1571–1585.

[143] Mlýnský V, Bussi G. Molecular dynamics simulations reveal an interplay between shape reagent binding and RNA flexibility. *J Phys Chem Lett*. 2018;9:313–318.

[144] Cunha RA, Bussi G. Unraveling Mg^{2+} -RNA binding with atomistic molecular dynamics. *RNA*. 2017;23:628–638.

[145] Bian Y, Tan C, Wang J, et al. Atomistic picture for the folding pathway of a hybrid-1 type human telomeric dna g-quadruplex. *PLoS Comput Biol*. 2014 04;10(4):1–11.

[146] Bian Y, Zhang J, Wang J, et al. Free energy landscape and multiple folding pathways of an h-type rna pseudoknot. *PLoS ONE*. 2015 06;10(6):1–16.

[147] Heller GT, Aprile FA, Bonomi M, et al. Sequence specificity in the entropy-driven binding of a small molecule and a disordered peptide. *J Mol Biol*. 2017;429:2772–2779.

[148] Löhr T, Jussupow A, Camilloni C. Metadynamic metainference: Convergence towards force field independent structural ensembles of a disordered peptide. *J Comp Phys*. 2017;146:165102.

[149] Prakash A, Baer MD, Mundy CJ, et al. Peptoid backbone flexibility dictates its interaction with water and surfaces: A molecular dynamics investigation. *Biomacromolecules*. 2018;19(3):1006–1015.

[150] Fu CD, Pfaendtner J. Lifting the curse of dimensionality on enhanced sampling of reaction networks with parallel bias metadynamics. *J Chem Theory Comput*. 2018;14:2516–2525.

[151] Awasthi S, Nair NN. Exploring high dimensional free energy landscapes: temperature accelerated sliced sampling. *J Chem Phys*. 2017;146:094108.

[152] Cuendet MA, Tuckerman ME. Free energy reconstruction from metadynamics or adiabatic free

energy dynamics simulations. J Chem Theory Comput. 2014;10:2975.

[153] Awasthi S, Gupta S, Tripathi R, et al. Mechanism and kinetics of aztreonam hydrolysis catalyzed by class-c β -lactamase: A temperature-accelerated sliced sampling study. J Phys Chem B. 2018; 122:4299–4308.

[154] Vithani N, Jagtap PKA, Verma SK, et al. Mechanism of Mg^{2+} -accompanied product release in sugar nucleotidyltransferases. Structure. 2018;26:459 – 466.

[155] Sahoo SK, Nair NN. Interfacing the core-shell or the drude polarizable force field with car-parrinello molecular dynamics for QM/MM simulations. Front Chem. 2018;6:275.