Title

Advances and Challenges in Milestoning Simulations for Drug-Target Kinetics.

Journal

Journal of Chemical Theory and Computation, 20(22)

Authors

Ojha, Anupam

Votapka, Lane

Amaro, Rommie

Publication Date

2024-11-26

DOI

10.1021/acs.jctc.4c01108

Peer reviewed

JCTC Journal of Chemical Theory and Computation

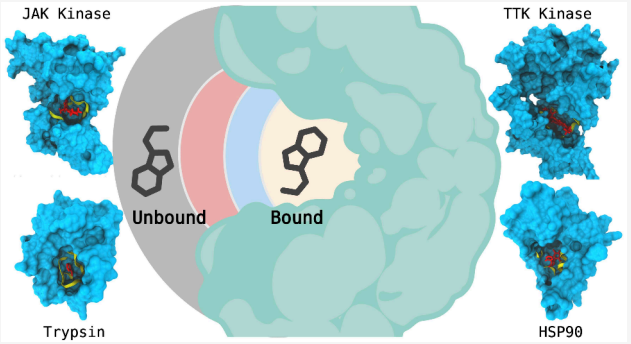
pubs.acs.org/JCTC
Perspective

Advances and Challenges in Milestoning Simulations for Drug−Target Kinetics

Anupam Anand Ojha,* Lane W. Votapka, and Rommie E. Amaro*

Cite This: J. Chem. Theory Comput. 2024, 20, 9759−9769
Read Online

ABSTRACT: Molecular dynamics simulations have become indispensable for exploring complex biological processes, yet their limitations in capturing rare events hinder our understanding of drug−target kinetics. In this Perspective, we investigate the domain of milestoning simulations to understand this challenge. The milestoning approach divides the phase space of the drug−target complex into discrete cells, offering extended time scale insights. This Perspective traces the history, applications, and future potential of milestoning simulations in the context of drug−target kinetics. It explores the fundamental principles of milestoning, highlighting the importance of probabilistic transitions and transition time independence. Markovian milestoning with Voronoi tessellations is revisited to address the traditional milestoning challenges. While observing the advancements in this field, this Perspective also addresses impending challenges in estimating drug−target unbinding rate constants through milestoning simulations, paving the way for more effective drug design strategies.



1. INTRODUCTION

Atomistic molecular dynamics (MD) simulations have emerged as a powerful tool in studying complex biological processes such as protein folding, protein−membrane interactions, and drug−target interactions. Substantial growth in algorithms, software, and hardware capabilities has extended the time scale of simulations to observe such processes, however, rigorously characterizing the kinetics and thermodynamics of such processes is still challenging. MD simulations are constrained to femtosecond-order timesteps, limiting their ability to capture complex conformational transitions or rare events (milliseconds or longer) in biological systems. Many different types of enhanced sampling methods enable researchers to bridge this gap in the time scale.[1,2] Enhanced sampling methods can be broadly grouped into two categories. The first category introduces a bias potential to accelerate transitions between conformational states. In some instances, including replica exchange MD,[3] selective integrated tempering,[4] and Gaussian accelerated MD,[5] the bias is applied to the whole system. In other instances, including metadynamics[6] and variationally enhanced sampling,[7] the bias is selectively applied to only certain forces within the system (so-called "collective variables"). The second category of enhanced sampling methods, termed path sampling methods, focuses on sampling transition regions. These methods include transition path sampling,[8] weighted ensemble simulations,[9] transition interface sampling,[10] and milestoning simulations.[11]

Drug−target binding and unbinding kinetics are critical determinants in diverse biological phenomena, including enzymatic catalysis, cellular signal transduction, and immune system activation. These molecular interactions equip cells to transduce external stimuli into biochemical signals, facilitating essential physiological functions. Drugs characterized by extended residence times demonstrate prolonged occupancy within the active site of the target, thereby extending their pharmacological effects. In contrast, fast-dissociating drugs tend to be pharmacologically less efficacious. Therefore, a comprehensive description of the kinetic profile is indispensable for the rational design and optimization of targeted therapeutic agents.[12−18] While the advancements in MD simulations have been pivotal for various biological processes, their application to drug−target kinetics presents unique challenges and opportunities.[19−27] Drug−target interactions often occur over varying time scales, on the order of seconds, hours, or even longer. Time scale limitations constrain conventional MD simulations to observe rare events. Moreover, large-scale conformational changes at the binding site, the

Received: September 6, 2024
Revised: October 30, 2024
Accepted: October 31, 2024
Published: November 7, 2024

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

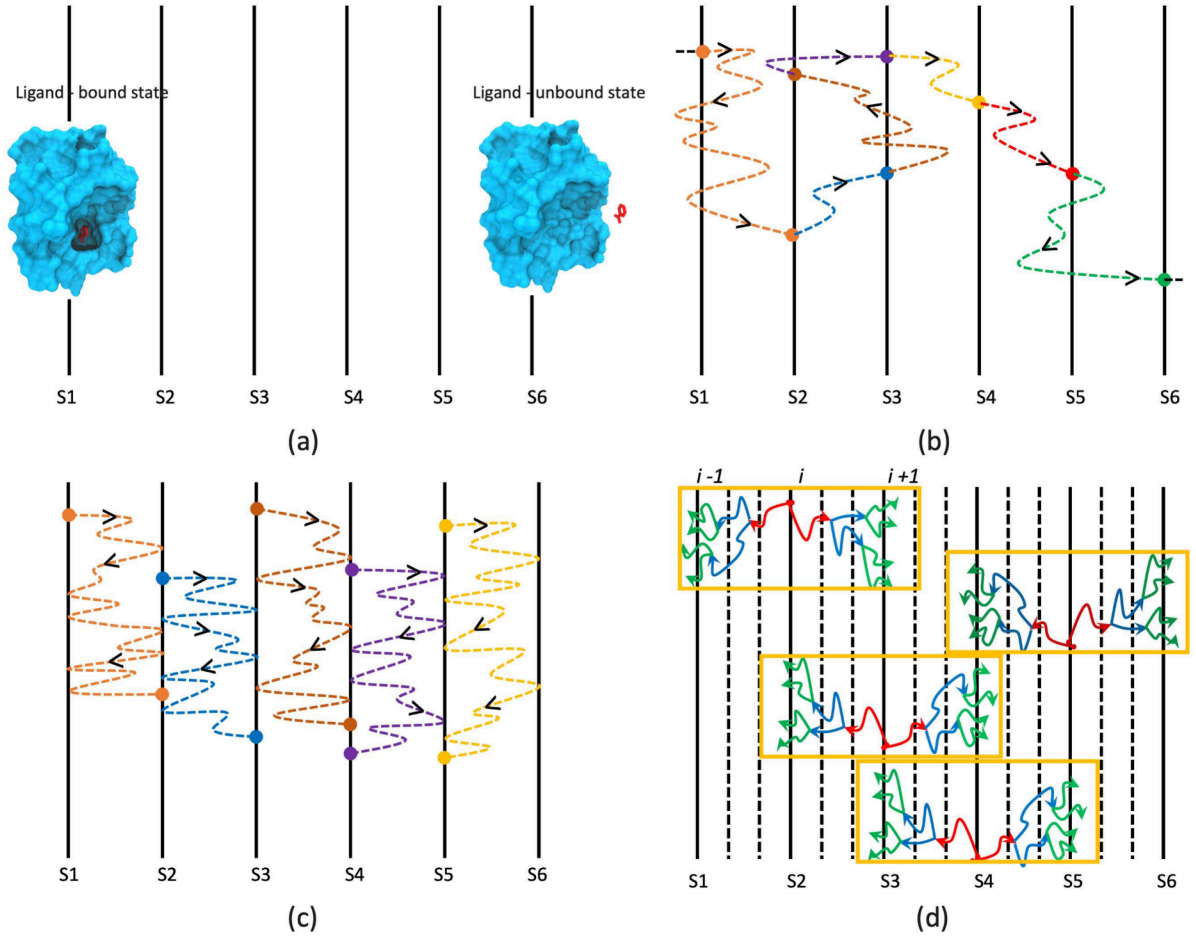


Figure 1. (a) Phase space of the receptor−ligand complex divided by milestones into discrete cells. (b) Schematic representation of a long ergodic trajectory crossing a set of six milestones. (c) The Markovian milestoning scheme, where replicas of the system evolve independently within the milestones with the collision rule or reflective boundary conditions such that the overall flux is maintained while ensuring equilibrium distribution within each milestone aligns with the Boltzmann−Gibbs distribution. (d) The weighted ensemble milestoning (WEM) method, where the milestone (solid lines) is further partitioned into bins (dotted) with the WE method employed within these bins. Each milestone is a starting point, while its neighbors serve as targets. Trajectories are halted upon reaching any milestone, unlike traditional steady-state simulations.

accuracy of force fields accounting for bound-state polarization, and experimental validation pose additional challenges for investigating drug−target kinetics.

Recently, milestoning theory has emerged as a robust and efficient approach for investigating drug−target kinetics. Milestoning simplifies the analysis of complex processes by dividing them into a series of transitional states, referred to as milestones. These milestones, dispersed across the configurational space of the system, surpass the limitations of conventional MD simulations and simulate the kinetics of biomolecular processes over extended timeframes. This perspective covers advancements and challenges within milestoning simulations for drug−target kinetics, providing an overview of the history, applications, challenges, and future prospects of milestoning simulations in drug−target kinetics.

2. MILESTONING THEORY: HISTORY AND CURRENT DEVELOPMENTS

2.1. Traditional Milestoning Approach. Milestoning is a computational method that simplifies the analysis of complex processes by breaking them down into a series of transitions between intermediate states, known as milestones, within the configurational space of the system (Figure 1a).[11,28−30] This method enables the simulation of the overall kinetics of the system over extended time scales, surpassing the limitations of traditional MD simulations. There are two underlying assumptions behind the milestoning theory.[31] First, successive transitions between milestones $i(t')$ at specific simulation times t' follow a probabilistic sequence $i_1(t_1)$, $i_2(t_2)$, $i_3(t_3)$, ..., which is determined by the likelihood $\rho_{i,j}(t)$ of successfully transitioning from one milestone i to the next milestone j within time t, ensuring that the overall statistical behavior of this progression mimics that of a discrete-time Markov process (eq 1):

$$\rho_{i_0,i_1}(t_1)\rho_{i_1,i_2}(t_2 - t_1)\rho_{i_2,i_3}(t_3 - t_2)... \quad (1)$$

In this process, the order of steps is essential, and the probability $\rho_{i,j}$ of each transition depends on the outcome of the previous one, creating a coherent and predictable sequence. Second, the transition times τ between successive milestone crossings are statistically independent. Based on these assumptions, the probability p of observing a given trajectory (given transition rates) depends on the number of observed transitions N_{ij} in a simulation, the total time R_i spent after crossing milestone i and before crossing any other milestone, and a rate matrix Q with components $q_{ij} = \frac{N_{ij}}{R_i}$ and is given by eq 2:

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

$$p(\text{traj}|Q) = \prod_{i=1}^{N} \prod_{j \neq i} q_{ij}^{N_{ij}} e^{-q_{ij} R_i} \tag{2}$$

Equation 2 forms the basis of milestoning theory. While this formulation assumes a single long trajectory, in practice, the milestoning method involves reinitializing short MD simulations on each milestone. Specific sets of milestones, defined as optimal milestones, satisfy the first assumption precisely, simplifying the analysis of kinetic properties like the mean first passage times (MFPT). Optimal milestones are difficult to provide or approximate. Hence, an alternative approach involves iteratively propagating trajectories between milestones to achieve the exact milestoning methodology, such that optimal milestones are not needed. While this approach does not rely on any assumptions, it is important to note that the precision of these calculations comes at a significantly higher computational cost, often by an order of magnitude.

In overdamped systems, milestoning has proven to be well-suited for accurately capturing the probabilistic transitions between states. This approach accounts for physical details of system dynamics, ensuring that the milestones selected lead to precise kinetic analyses.[31] If one approximates the placement of optimal milestones, or utilizes an approach such as exact milestoning to compensate for nonoptimal milestones, one ensures a specific probabilistic transition pattern between milestones. Optimal milestones are identified as isosurfaces of the committor function, an important concept in transition path theory. Utilizing isocommittor surfaces as milestones guarantees the desired probabilistic behavior during transitions. While the second assumption may not hold universally, exact computations of MFPTs are still achievable with these optimal milestones, making them useful for efficient kinetic analysis in complex systems.

Reinitialization of MD trajectories from the milestones may be used to compensate for the suboptimal placement of milestones. While having optimal milestones ensures that transition probabilities between milestones are truly constant, regardless of where the system has crossed the milestone surface, optimal milestones can only be approximated and are expensive to obtain, and it becomes necessary for one to effectively reinitialize MD simulations from the milestones in a manner that produces the correct transition probabilities and transition times. One may account for the absence of optimal milestones by estimating the first hitting point distribution (FHPD), which relies on a density $\rho_{ij}(x)$ at position, x, which refers to the probability of the first point a trajectory hitting milestone j given that the trajectory initiated from milestone i. A set of optimal milestones would have a density $\rho_j(x)$ that does not depend on the originating milestone i. To obtain an FHPD, transition path theory (TPT) provides the framework, where the FHPD density is defined by eq 3:

$$\rho_j(x) = \frac{e^{-\beta V(x)} |\nabla q(x)|}{\int_{S_j} e^{-\beta V(x)} |\nabla q(x)| \, dx} \tag{3}$$

where $V(x)$ is the potential energy of the system at position x, $\beta = 1/k_B T$ with temperature T and Boltzmann constant, k_B, and $\nabla q(x)$ is the gradient of the committor function q at position x.

For the exact calculation of MFPTs, it is important to understand that the second assumption about the statistical independence of transition times between milestones only sometimes holds true. However, we can still compute the MFPT accurately with optimal milestones.[30,32] When a trajectory visits milestones successively, the total duration of such a sequence is the sum of transition times between milestone crossings (Figure 1b). These transition times can be correlated, leading to the need to describe their statistical properties through a joint probability density, which is a complex task. However, regardless of their correlation, one can always calculate the mean duration of such sequences accurately. To compute the mean first passage time from one milestone to another, a modified transition probability matrix turns the last milestone into a sink state, preventing the process from leaving it once reached and enabling the calculation of MFPTs to be straightforward. This is a significant advantage of the milestoning approach, as it simplifies the calculation of kinetic properties.

Historically, MD simulations were limited to short time scales and small spatial resolutions due to the computational resources required for atomically detailed systems. Early methods often relied on long MD trajectories that, while conceptually simple, were inefficient for studying kinetic processes over biologically relevant time scales. The introduction of path-splitting methods such as milestoning simulations enabled efficient calculation of kinetic and thermodynamic properties by using many short trajectories mapped onto coarse space variables, reducing computational costs while maintaining atomistic details. In its early stages, milestoning simulations provided approximate predictions for processes such as ion permeation through membranes[33] or small molecule permeation through lipid bilayers,[34] with moderate accuracy. During the initial phases of SEEKR development, milestoning was employed to compute the kinetics of simple interactions, such as sodium-chloride encounters, and to estimate the k_{on} rate constants for simple biological systems, including the binding of superoxide dismutase with its natural substrate, the superoxide anion, and the N-terminal domain of troponin C with calcium ion.[35] A significant advancement for SEEKR was the accurate estimation of both k_{on} and k_{off} rates for simple host−guest complexes, along with precise ΔG predictions,[36] demonstrating its ability to capture the kinetics and thermodynamics of these interactions. With the development of state-of-the-art algorithms and computational resources, milestoning can now predict kinetics with high precision for complex biological systems and large, highly branched, flexible ligands, with applications extending to complex drug−target unbinding events with residence time predictions ranging from seconds to hours.[37−39]

2.1.1. Milestoning versus Markov State Models. Milestoning and Markov state models (MSMs) focus on transitions between distinct states in the configurational space of the system. While the milestoning approach assumes statistically independent transition events between surfaces in phase space (milestones), MSMs focus on the transitions between states represented by regions within the phase space of the system.[40−42] MSMs discretize the configuration space into distinct states and model transitions between these states using a Markov process. The milestoning approach, on the other hand, defines milestones as interfaces between cells in the phase space. In other words, instead of dividing the entire configuration space of a system into finely segmented states (as is the case of MSMs), the milestoning approach identifies key points or conditions within the phase space that are

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

strategically placed to capture significant events or transitions. This method simplifies the analysis of complex processes by focusing on these critical transitions. This approach allows for a more manageable and computationally efficient exploration by reducing the need for a detailed computation and analysis of every possible state, thus making it particularly useful for systems where identifying critical transitions is more relevant than detailed state-by-state dynamics. Both methods have their advantages and limitations. Milestoning is advantageous for its simplicity, computational efficiency, and accuracy by focusing on key transitions, making it suitable for specific analyses, such as calculating the unbinding rate constant for drug−target interactions. MSMs, while providing a broader view of the system dynamics, are suited to certain types of detailed state-by-state analyses, such as studying the folding pathways of proteins where multiple alternative intermediate states exist. Comparative studies with methods such as MSMs, transition interface sampling[10] (TIS), and forward flux sampling[43] (FFS) provide valuable insights into the efficacy of milestoning for studying system kinetics. This method offers a comprehensive approach to understanding complex processes, making it a valuable tool for researchers in various fields.

2.1.2. Milestoning versus Weighted Ensemble Simulations. The weighted ensemble[9] (WE) and milestoning approaches are path-splitting strategies to accelerate rare events in MD simulations. Instead of relying on a single long simulation, both methods partition the phase space of the system by discrete cells (for milestoning simulations), or bins (for WE simulations) based on a predefined CV and aim to provide unbiased estimates of key observables, such as rate constants and equilibrium state populations. In milestoning, "cells" are typically 1D/2D Voronoi cells determined by geometric partitioning through the Voronoi tessellation approach. Similar in concept, "bins" are discrete regions defined along the CVs that segment the phase space for redistributing and reweighting trajectory walkers. Although neither "bins" in WE, nor "cells" in milestoning are, necessarily, Voronoi tessellations. The key differences lie in the sampling strategies of these two methods. The WE method employs intermittent communication between parallel simulations at fixed intervals, with the frequency determined by the resampling time. The trajectories are periodically reweighted and resampled to accelerate the conformational search of rare events while maintaining a constant flux of trajectories. The transition rate between initial and final states is estimated using the Hill relation,[44] which relates the flux of transitions to the MFPT. This method explores the entire CV space comprehensively, such as the large-scale conformational transitions. However, the application of the WE method in calculating kinetics is computationally expensive, and it is frequently combined with other enhanced sampling methods to improve computational efficiency.[45,46]

2.2. Markovian Milestoning with Voronoi Tessellations. Conventional milestoning is a powerful technique for accelerated MD simulations but often presents several challenges. Milestones must be strategically placed to capture transitions in a complex potential energy landscape, but determining their optimal locations can be complicated and difficult to predict a priori. In conventional milestoning, when estimating transition probabilities and transition times, short trajectories are initiated on each milestone and run until they reach another milestone.[30] This process necessitates reinitializing trajectories from the milestones. The challenge here lies in determining the correct probability distribution for reinitialization, which in conventional milestoning is done using the equilibrium distribution of the system restricted to the milestones (typically obtained via umbrella sampling). However, this equilibrium distribution is usually not accurate for reinitialization, and it can be challenging to compute the FHPD explicitly. The exact milestoning procedure effectively continues trajectories that have arrived directly from a previous milestone, allowing one to obtain a closer estimate of the FHPD, and thus, more correct results.[47]

To tackle the prior-mentioned challenges in traditional milestoning, a procedure was introduced in the milestoning scheme, i.e., Markovian milestoning with Voronoi tessellations (MMVT).[48] There are three key ideas in MMVT. First, this procedure defines milestones as the edges of a Voronoi tessellation generated based on a set of points in Cartesian space. Each milestone corresponds to a portion of a hyperplane, and their boundaries are naturally determined by the regions where two milestones intersect. Second, MMVT assumes that the evolution of the system follows a continuous-time Markov jump process, and successive transitions between the milestones are statistically independent. Third, independent replicas of the system are sampled by running unbiased MD simulations in individual Voronoi cells, with a collision rule that constrains the dynamics of the system within these cells by implementing reflective boundary conditions (Figure 1c). With the transition counts between edges (milestones) within the Voronoi cells, a transition matrix is constructed by a maximum likelihood analysis to calculate the MFPT for the transition from the initial to the target state.

2.3. Weighted Ensemble Milestoning and Markovian Weighted Ensemble Milestoning Approaches. Weighted ensemble milestoning (WEM) combines the efficiency of WE simulations with the theoretical framework of the milestoning approach (Figure 1d).[49] The milestoning approach involves dividing the phase space of the complex into noninteracting hypersurfaces and estimating the probability flux between them. At the same time, the WE simulation method employs parallel trajectories with predefined weights.[9] Evolving trajectories are pruned or split based on progress toward a target state, focusing the computing effort on functional transitions. The configurational space spanned by the progress coordinates is divided into bins, and each bin keeps instantaneous trajectories or walkers constant by pruning or splitting them based on their weights. The WEM approach involves dividing the phase space between milestones into bins and employing the WE approach to sample trajectories within these bins, aiming to improve or accelerate the convergence of transition probabilities and time scales while reducing computational resources. The WEM scheme has been implemented on model systems, including a 1D double-well potential, a high-dimensional potential, and alanine dipeptide while reproducing kinetics from MD and other simulation methods with a reduced computational time. The applications of the WEM method are limited to model systems, and expanding its validation to complex receptor−ligand systems would be useful for future research.

The Markovian weighted ensemble milestoning (M-WEM) approach is another variant of the WEM approach, where the Markovian milestoning scheme (as discussed previously) is combined with the WE simulations to study the kinetics and free energy of rare events in systems of interest.[50] This method employed the dynamic binning scheme in the WE simulations.

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

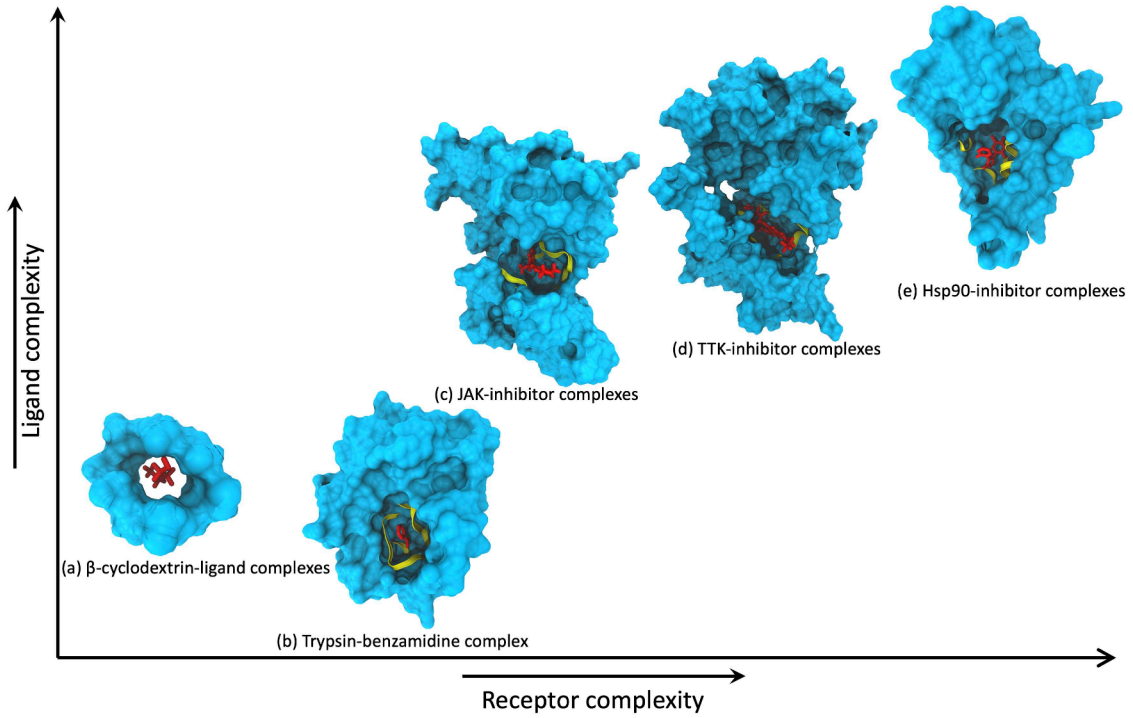


Figure 2. Hierarchical mapping of receptor and ligand complexity in benchmark systems for validating existing and new milestoning simulation algorithms for (a) β-cyclodextrin−ligand complexes, (b) trypsin−benzamidine complex, (c) JAK−inhibitor complexes, (d) TTK−inhibitors complexes, and (e) Hsp90−inhibitor complexes. The x axis represents receptor complexity, incorporating factors such as the conformational diversity of the protein, allosteric regulation sites, protein−ligand interaction strength, the number of amino acid residues, residence time, and the flexibility of the binding site and associated loops/hinges. The y axis represents the ligand complexity, considering molecular size, chemical diversity, stereochemistry, dynamic behavior within the binding site, and structural flexibility.

While this adaptation is not explicitly discussed in terms of the Markovian property, it suggests that the method adjusts the sampling strategy based on the current state of the system, which aligns with the idea of maintaining the Markovian property. Another aspect of M-WEM is the elimination of trajectory interruptions at milestone interfaces. Traditional methods often require stopping trajectories and initiating new ones at milestone boundaries. By avoiding this interruption, M-WEM maintains the continuity of the trajectory, ensuring that the behavior of the system remains Markovian. The M-WEM method has been applied to study the two-dimensional toy model of the Müller-Brown potential, conformational transitions in alanine dipeptide, and the study of protein−ligand (un)binding kinetics for the trypsin−benzamidine complex. While the M-WEM method has shown promise in diverse applications, addressing the complexities of drug−target interactions characterized by a higher degree of freedom remains an area for further investigation. The applicability and limitations of M-WEM in such scenarios require additional exploration and validation.

3. DRUG−TARGET KINETICS STUDIED BY THE MILESTONING APPROACH

3.1. Simulation-Enabled Estimation of Kinetic Rates (SEEKR): A Multiscale Milestoning Framework for Estimating k_{on} and k_{off}. Simulation Enabled Estimation of Kinetic Rates (SEEKR) estimates drug−target binding and unbinding kinetic rate constants using atomistic MD simulations and less detailed but computationally efficient Brownian dynamics (BD) simulations.[36,51] The SEEKR workflow starts with a three-dimensional structure of a bound-state receptor−ligand complex.[52] A collective variable (CV) is defined, which, in most occurrences, is the center of mass (COM)−COM distance between the ligand and the alpha carbons of the bound state. Based on the CV, the phase space of the complex is divided into MD and BD regions. The MD region is further partitioned into Voronoi cells between milestones. Starting structures for simulations are obtained by running steered MD (SMD) or metadynamics simulations from the bound to the unbound state, where copies of the receptor−ligand complex are saved for each Voronoi cell as the ligand slowly moves out of the binding pocket. Independent and parallel MD simulations are performed in the space between milestones, while BD simulations are performed in the BD region. Reflective boundary conditions are imposed to confine the trajectories between respective milestones. Transition counts and time spent in each Voronoi cell account for the transition matrix, which is then solved to calculate the drug−target unbinding rate constant (k_{off}). BD simulations offer a computationally efficient method for calculating the binding rate constant (k_{on}) for drug−target complexes. Unlike MD simulations, BD simulations use simplified models and approximations, such as treating molecules as rigid bodies and using implicit solvents, allowing larger time steps in the simulations. Despite these simplifications, BD can accurately model the initial stages of drug-binding events and account for explicit electrostatic interactions. The SEEKR framework incorporates the Browndye simulation package, which employs the Luty−McCammon−Zhou algorithm to generate multiple trajectories of ligands around receptors to estimate the association rate.[53] These trajectories either end in a molecular encounter or escape, and the probabilities of such encounters

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

is used to compute the binding rate constant. The method is particularly useful when electrostatic forces are significant in molecular recognition and binding. The SEEKR framework demonstrated close-to-experiment k_{off} rates for a range of receptor−ligand complexes, encompassing residence times that vary from microseconds to hours. These applications span from relatively straightforward systems such as β-cyclodextrin−ligand[54] and trypsin−benzamidine complexes[55] to more challenging cases such as JAK−inhibitor complexes,[37] a series of Hsp90−inhibitor complexes,[38] and a series of threonine-tyrosine kinase (TTK) inhibitors.[39] One valid concern about drug−target kinetics relates to whether milestoning can effectively capture the effects of flexibility and conformational changes in complex biological systems, which are not explicitly represented in the milestoning model. In such a case, if the time scales of flexible motions are short relative to the durations of individual milestoning simulations, such that these motions are sufficiently sampled within the span of these simulations, then these motions will be automatically accounted for within the milestoning model. However, if these motions are slow enough and not adequately sampled in a single milestoning simulation, and if these slow, flexible motions would directly impact the time scales or energetics of the process under consideration (such as binding/unbinding), then the motions themselves ought to be explicitly characterized with a CV and milestoned to be correctly accounted for. If more than one CV is required to examine a system under question, then multidimensional milestoning must be employed using, for example, a Voronoi tessellation. This capability readily exists within the SEEKR software.

3.2. Independent Kinetic Studies with a Milestoning Simulation Approach. While the SEEKR framework streamlines the method of calculating k_{on} and k_{off} rates using multiscale approaches, it is worth noting that several independent milestoning studies have been conducted specifically to calculate k_{off} rate constants. These studies have added depth to our understanding of drug−target (un)binding kinetics by exploring various drug−target complexes and scenarios, further establishing these approaches as an efficient method for estimating k_{off} rates. Figure 2 illustrates the diverse complexity levels encountered in receptors and their corresponding inhibitors across various benchmark systems used in milestoning simulations. These systems include β-cyclodextrin−ligand, trypsin−benzamidine, JAK−inhibitor, TTK−inhibitor, and Hsp90−inhibitor complexes. Several factors contribute to the complexity of the receptor−ligand complexes and underscore the necessity of novel simulation approaches to capture such dynamics accurately. While the complexity of ligands encompasses factors such as molecular size, chemical diversity, stereochemistry, dynamic behavior, structural flexibility, and their specificity and selectivity in interactions, the complexity of receptors depends on their conformational diversity, allosteric regulation sites, protein−ligand interaction strength, and the flexibility of binding sites and associated loops or hinges.

A recent study used milestoning to investigate the molecular dissociation pathway and absolute dissociation rate of Gleevec from Abl kinase, a key target in cancer therapy.[56] Using the milestoning simulation approach, an average MFPT ($1/k_{off}$) was estimated to be 0.055 s, against the experimentally estimated residence time of 0.04 ± 0.01 s, with a total simulation time of 1 μs. The milestones were decided on a progressive approach where SMD simulations were employed to generate the initial 43 milestones. As unconstrained trajectories starting from these initial milestones deviated from the SMD path, new milestones were subsequently discovered, and this process was iterated until a connected transition matrix and a finite MFPT were obtained. The MFPT was determined by summing all possible paths leading from the reactant (bound state) to the product (unbound state) and considering the adjusted transition probabilities between milestones.

Another study estimated the dissociation rate constant of a ligand from the serine-threonine kinase, glycogen synthase kinase 3β (GSK-3β), which is a potential drug target for the treatment of various diseases, including neurodegenerative disorders and diabetes.[57] Similar milestoning protocols were employed as in the previous study. An absolute dissociation rate (k_{off}) was estimated to be 15.4 ± 0.3 s, against the experimentally estimated k_{off} of 18.4 s, with a total simulation time of less than 1 μs.

4. CHALLENGES AND FUTURE DIRECTIONS

4.1. Force Field Accuracy. The interaction of a drug molecule at its binding site is a complex interplay of molecular forces involving electrostatics, van der Waals forces, hydrogen bonding, and polarization. Polarization at the bound state, i.e., electron-density redistribution in response to an external electric field, generally affects drug−target interactions. Thus, an important and active area of research involves creating better fixed charge force fields for use in drug discovery and design (e.g., see the efforts of the Open Force Field Initiative[58−60]). While the drug molecule and the binding site residues can undergo electronic polarization, leading to changes in their charge distributions, it is worth noting that nonpolarizable force fields can still perform well in specific milestoning calculations, depending on the system being studied. A precise estimation of such polarization effects is crucial because they influence the binding affinity, kinetics, and, ultimately, the therapeutic efficacy of a drug. Polarizable force fields, such as AMOEBA[61] and CHARMM Drude force fields,[62] are tailored to capture electronic polarization in receptor−ligand complexes, providing a more realistic description of polarization-induced effects on (un)binding kinetics. A recent study focused on the quantum mechanical reparameterization of ligand charges at its binding site, refining the potential energy landscape within the bound state of the receptor−ligand complex and offering a more accurate representation of intermolecular interactions and polarization effects at the bound state.[38] This approach, named QMrebind (Quantum Mechanical force field reparameterization at the receptor−ligand binding site), has been integrated into the multiscale milestoning simulation approach, i.e., SEEKR2, and has been successful in accurately estimating and rank-ordering the drug−target unbinding rate constants for a series of Hsp90−inhibitor complexes. Early work on quantum mechanical corrections has also shown the importance of efficient QM/MM free energy methods for capturing polarization effects.[63] Charge transfer is another phenomenon influencing the estimation of drug−target kinetics, involving the transfer of electrons from one molecular entity to another, leading to changes in their charge states. Augmentation of force fields with charge transfer models can explicitly account for such events during (un)binding kinetics. Milestoning simulations coupled with force fields that accommodate the charge transfer models can potentially provide a dynamic view of electron

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

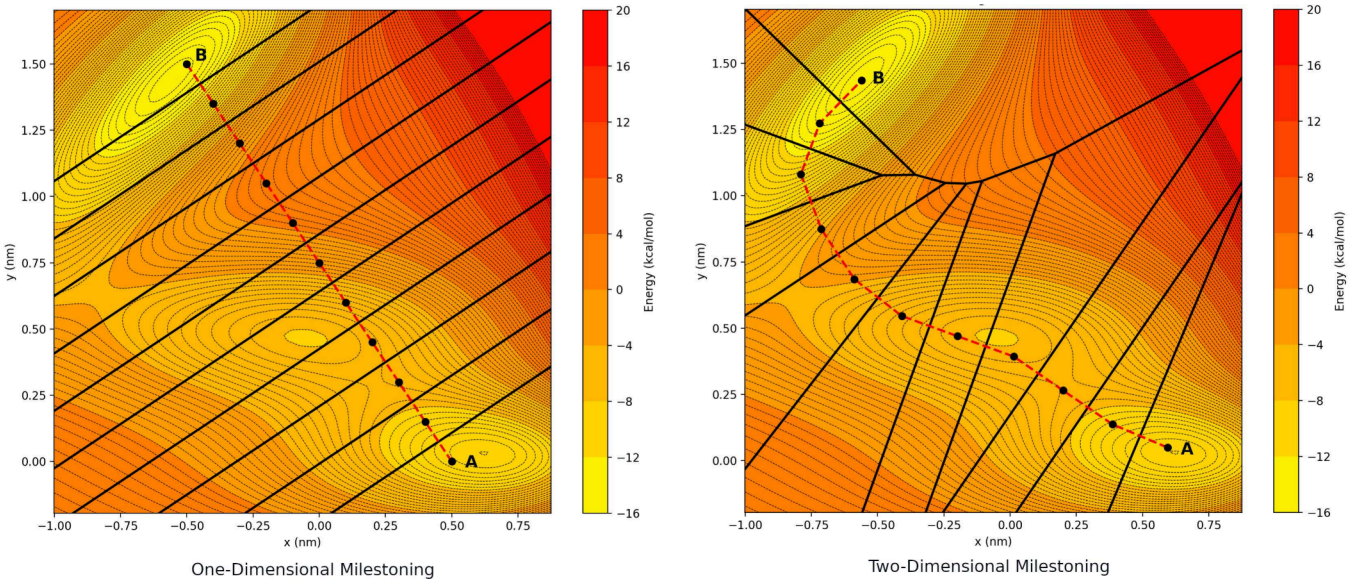


Figure 3. (left) Energy landscape of the Muller potential with transition path illustrating the potential energy surface between the equilibrium states A and B marked. A one-dimensional collective variable is defined as a reaction function (typically linear) along which anchors are placed, with only two neighbors each, along the transition from state A to state B. (right) Optimized transition path through a Muller potential landscape using a Voronoi tessellation approach, representing the most probable pathway of the evolution of the system from state A to state B within a two-dimensional space. A Voronoi diagram is generated by a set of input points that serve as anchors, each defining a Voronoi cell, essentially the nearest set of points to its respective anchor. The cell boundaries determine the transition between regions closest to adjacent anchor points. Within this framework, a two-dimensional Voronoi tessellation is generated from these anchor points. Milestones, positioned at equal intervals between anchors, further segment the space into distinct Voronoi cells.

transfer processes within the binding pocket, shedding light on their influence on (un)binding kinetics.[64]

4.2. Determining the Best Estimate of the Free Energy Profile of the Ligand Dissociation. In the context of milestoning simulations, determining initial conditions, particularly the starting milestones, is crucial. Traditional approaches employ SMD simulations, potentially leading to strained receptor−ligand dynamics as the ligand moves out of the binding pocket. One of the essential aspects of milestoning simulations, therefore, is obtaining an accurate estimate of the free energy profile for a dissociating ligand from its binding site.

Dissipation-corrected targeted MD[65] (dcTMD) is a novel approach for calculating free energy profiles for slow processes. It addresses the challenge of accurately simulating complex processes, such as the ligand dissociation from a binding site, by incorporating corrections for energy dissipation, providing a more accurate estimate of the dynamics of the system. dcTMD corrects for energy dissipation by considering the cumulative effect of friction during the simulation, ensuring that the calculated free energy profiles accurately account for energy losses due to dissipation. dcTMD applies a biasing force to a specific set of atoms, typically the ligand in a receptor−ligand complex, to drive the ligand unbinding process as a steering mechanism, guiding the system along a predefined reaction coordinate. It leverages Jarzynski's identity,[66] which relates the exponential average of work done on the system to the free energy change associated with an unbinding mechanism. A unique feature of dcTMD is its ability to separate the total reaction flux into multiple pathways, each with its own energy curves and friction factors. This separation allows dcTMD to handle complex free energy landscapes, which can be particularly beneficial when dealing with systems where the energy landscape is divided into multiple pathways (e.g., T4 lysozyme, HSP90, among others). By providing optimized initial conditions for simulations, dcTMD ensures that milestoning simulations start from a more realistic and representative state, improving their accuracy.

Random acceleration MD (RAMD) is another enhanced sampling approach that applies randomly oriented forces to the ligand to accelerate protein−ligand dissociation events.[67] A small and randomly oriented force of constant magnitude is applied to the ligand during MD simulations. The orientation of this force is initially chosen randomly and may change based on the motion of the ligand relative to the receptor protein. The procedure involves several key parameters, including the time interval for inspecting ligand motion, ligand−receptor displacement threshold, maximum ligand displacement, and the magnitude of the applied force. A free energy profile can then be generated employing RAMD simulations that efficiently sample the ligand unbinding pathway. Snapshots representing the ligand unbinding simulation pathway can be saved as starting structures for milestoning simulations. Recently, a modified version of RAMD, i.e., τ-RAMD, has been employed to estimate relative drug−target unbinding kinetics and rank order a series of Hsp90−inhibitor complexes based on their residence times.[68] Incorporating RAMD or τ-RAMD in milestoning simulations for predicting initial milestones could help estimate accurate kinetic rate constants.

The metadynamics approach, specifically well-tempered metadynamics, offers an alternative method for estimating the free energy profile of ligand dissociation in milestoning simulations.[69] Metadynamics facilitates the escape of the ligand from its energy minima by iteratively applying a history-dependent biasing potential to the system, thereby accelerating the sampling of unbinding events. A ligand dissociation free energy profile is generated, providing insights into the potential barriers and stable states involved in ligand dissociation. This

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

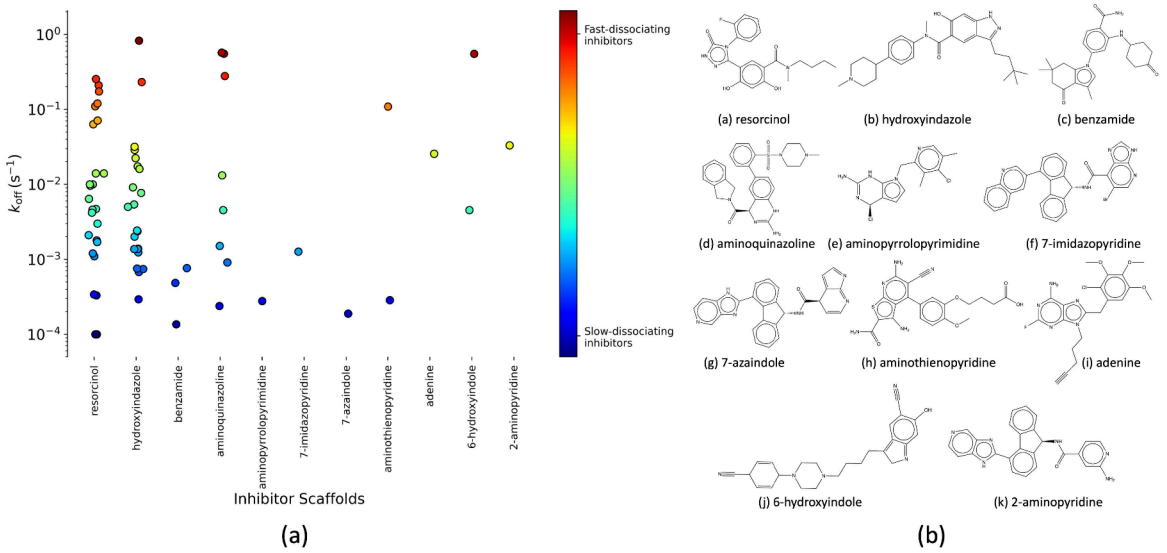


Figure 4. (a) Scatter plot illustrating the dissociation rate constants (k_{off}) for 70 Hsp90 inhibitors, with each point representing an inhibitor, with its corresponding k_{off} rate. The inhibitors are categorized based on their dissociation rate constants into fast-dissociating ($k_{off} > 10^{-1}$ s^{-1}) and slow-dissociating ($k_{off} < 10^{-3}$ s^{-1}). (b) Categorizing 70 Hsp90 inhibitors based on their structural diversity, scaffold, and chemical motifs, namely, resorcinol, hydroxyindazole, benzamide, aminoquinazoline, 7-imidazopyridine, adenine, 7-azaindole, aminothienopyridine, aminopyrrolopyrimidine, 6-hydroxyindole, and 2-aminopyramidine.

iterative process not only aids in escaping local minima but also ensures a smoother convergence toward the true free energy surface. Incorporating metadynamics into the milestoning framework has provided a successful approach by efficiently identifying initial structures and enhancing the accuracy of kinetic rate constant predictions.[39]

4.3. Finding Accurate Collective Variable Estimates and Employing Multidimensional Milestoning Schemes. Complex biological processes, such as drug−target (un)binding events, often undergo multiple conformational changes, and employing a one-dimensional CV may oversimplify the process, potentially missing higher dimensional conformational configurations and transitions in the system. In other words, milestoning may result in biased or inaccurate kinetics if the selected CV does not represent the slowest process of the system. There are potential avenues where multidimensional milestoning would be advantageous. Employing multidimensional CVs in milestoning enables the description of more complex and curved reaction pathways. By combining two or more CVs within a multidimensional grid, we can represent a higher-dimensional free energy landscape, allowing for a more accurate and detailed exploration of the energy landscape. Moreover, certain energy barriers may only be evident when employing multiple dimensions in the reaction pathway. Additionally, it facilitates the observation of processes involving coordinated movements across various dimensions, offering reliable insights into reaction pathways, transition rates, and energy landscapes. Figure 3 illustrates the energy landscape of the Muller potential. It compares a one-dimensional CV approach, marked along the minimum energy pathway between states A and B, with a two-dimensional Voronoi tessellation, which offers a more probable and detailed pathway representation within a complex energy landscape.

4.4. Scope of QM/MM MD Simulations in the Milestoning Scheme. Quantum Mechanics/Molecular Mechanics MD (QM/MM MD) simulations combine the accuracy of QM calculations with the efficiency of MM simulations to study biological processes in complex environments, where a subset of atoms, often the active site or a region of interest, is treated quantum mechanically, while the remaining atoms or residues are treated classically using molecular mechanics.[70] Certain mechanisms during the drug−target (un)binding processes, such as proton transfer, electronic rearrangements, bond formation, and cleavage, can be quantitatively estimated using QM/MM MD simulations near the binding site regions. QM/MM MD simulations are particularly useful in determining transition states, i.e., the high energy regions along the CV, providing information about the rate-limiting steps for the unbinding processes, which is crucial for understanding the kinetics of a process. The initial innermost milestones along the CV are situated near the drug−target binding site, presenting an opportunity to apply QM/MM MD simulations. While QM/MM simulations within these milestones require substantial computational resources, exploring the leverage of enhanced accuracy on the simulation time scale would be interesting.

4.5. Expanding the Breadth of Benchmarking Studies and Potential ML Application of Milestoning-Generated Trajectories. Existing enhanced sampling methods for predicting drug−target kinetics have been applied to diverse target proteins, many of which share little structural complexity and chemical composition overlap. The trypsin−benzamidine complex has gained widespread recognition as a benchmark system well-suited for method development.[50,51,71] Nevertheless, its simplicity as a drug−target complex, lacking hidden CVs, poses a limitation.[20] Within this system, the unbinding process is straightforward, involving a transition out of the binding pocket, sometimes followed by surface diffusion, leading the benzamidine into the solvent. A recent study, however, highlighted additional complexities in the trypsin−benzamidine complex, particularly regarding the role of water molecules, where the CVs used primarily focused on the behavior and arrangement of water molecules in the binding cavity.[72] This revelation contests the previous understanding of trypsin−benzamidine as a simple drug−target complex, underscoring the necessity of considering these additional

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

variables for more accurate predictions of drug−target kinetics. To thoroughly assess new enhanced sampling methods, the benchmark systems must feature a set of diverse ligands with varying chemical scaffolds. Moreover, these systems must be supported by high-resolution experimental structures and kinetics data to facilitate comprehensive evaluations. Recent studies have made significant strides focusing on a series of Hsp90−inhibitor complexes, where ligands exhibited diverse scaffold structures, varying residence times, and multiple (un)binding pathways. Notably, these complexes also featured substantial conformational changes during ligand dissociation. Such characteristics position the Hsp90−inhibitor complexes as one of the most comprehensive and robust benchmarking systems available for the rigorous evaluation of new enhanced sampling methods in drug−target kinetics prediction. The set of eight inhibitors of the TTK system also provides a medically relevant, although smaller, set of benchmarking compounds kinetics predictions.[39] Figure 4 depicts a scatter plot of dissociation rate constants for 70 Hsp90 inhibitors, categorized by their rate constants and scaffold diversity, demonstrating the variety and complexity of this benchmark system.

Machine learning (ML) offers multiple potential applications within milestoning simulations, enhancing various aspects of the method, from applying ML-driven force fields to identifying key collective variables. One promising approach involves postprocessing SEEKR-generated trajectories, where ML models analyze simulation data sets to uncover hidden patterns or correlations between specific molecular features, such as ligand functional groups, and their interactions with protein side chains, directly linking them to kinetic or thermodynamic properties. Variational autoencoders (VAEs)[73] and generative neural networks (GNNs)[74] can analyze simulation trajectories to identify potential ligand-unbinding exit pathways, uncovering alternative (un)binding pathways that may be missed in standard simulations. Training on SEEKR-generated trajectories can also be employed for ligand design, where such models identify structural modifications that optimize binding kinetics or selectivity. ML-driven force fields,[75,76] trained on high-level quantum mechanical data further enhance simulation accuracy by dynamically adjusting force field parameters in individual Voronoi cells during SEEKR runs, leading to improved predictions of on/off rates and free energy landscapes.

5. CONCLUSION

Milestoning simulations have emerged as a powerful and versatile approach for studying the kinetics of complex biological processes, particularly in drug−target interactions. This perspective attempts a comprehensive overview of the history, applications, challenges, and future directions of milestoning simulations in drug−target kinetics. Several advancements in milestoning simulations have been presented, including MMVT, which addresses traditional milestoning challenges. The introduction of multiscale approaches such as SEEKR has provided a framework for accurately estimating drug−target binding and unbinding rate constants. Additionally, incorporating WEM and M-WEM approaches has further improved the efficiency and convergence of milestoning simulations. However, several challenges lie ahead in the field. Key challenges and opportunities in milestoning simulations include improving force field accuracy to account for polarization and charge transfer effects, ensuring precise determination of initial simulation conditions, exploring multidimensional milestoning for complex processes, and enhancing accuracy through the integration of QM/MM simulations at the innermost milestones, all while expanding benchmarking studies to encompass diverse ligands and complex target proteins. Optimizing starting structures and protonation states for milestoning simulations is essential for accurate kinetic estimates, and, as mentioned previously, both kinetics and thermodynamics can be highly representative of the efficacy of a drug candidate. With the goal of a computational method that can efficiently and accurately predict the on-rate, off-rate, and free energy of binding, it can be used to evaluate new lead compounds confidently, and tools such as SEEKR can provide atomic-level details that might suggest potential modifications to lead compounds to optimize the desired quantities. These refinements are essential for capturing the true dynamics of drug−target (un)binding processes and ensuring simulations mirror physiological conditions, thereby improving the reproducibility and reliability of milestoning simulations. Milestoning simulations, while a powerful enhanced sampling method, are computationally intensive in large-scale drug discovery projects where multiple inhibitors must be screened efficiently. Recent advancements in computational power and parallel computing have reduced computational loads. Algorithmic developments, including the incorporation of machine learning models and a better choice of collective variables, coupled with the integration of adaptive algorithms that can dynamically refine milestone placement based on real-time simulation data, are expected to make significant strides in efficient phase space exploration. This approach would allow for a more focused sampling in regions with high transition probabilities or kinetic significance. Using cloud computing resources and distributed computing platforms would enable the scaling up of simulations, facilitating the simultaneous processing of multiple trajectories, which is especially advantageous for high-throughput screening in drug discovery. The impact of milestoning simulations extends beyond academic research into the practical world of drug discovery and development. In the pharmaceutical industry, where the average time to bring a new drug to market can exceed a decade and cost billions, milestoning simulations offer a particularly promising avenue for accelerating the drug discovery and development process.

AUTHOR INFORMATION

Corresponding Authors

Anupam Anand Ojha − *Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States; Center for Computational Biology and Center for Computational Mathematics, Flatiron Institute, New York, New York 10010, United States;* orcid.org/0000-0001-6588-3092; Email: aojha@flatironinstitute.org

Rommie E. Amaro − *Department of Molecular Biology, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0002-9275-9553; Email: ramaro@ucsd.edu

Author

Lane W. Votapka − *Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0002-0865-5867

Complete contact information is available at:

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769

https://pubs.acs.org/10.1021/acs.jctc.4c01108

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.A.O. acknowledges the support of the Molecular Sciences Software Institute (MolSSI) Fellowship under NSF Grant OAC-1547580. R.E.A. acknowledges support from NSF Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) CHE060063 and NIHGM132826. The authors thank Ron Elber for his valuable suggestions for improving this Perspective. The authors also thank Javier Sanlley-Hernandez, Shiksha Dutta, and Ambuj Srivastava for going through the manuscript and providing valuable feedback.

■ REFERENCES

(1) Hénin, J.; Lelièvre, T.; Shirts, M. R.; Valsson, O.; Delemotte, L. Enhanced Sampling Methods for Molecular Dynamics Simulations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2022**, *4*, No. 1583.
(2) Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. Enhanced sampling in molecular dynamics. *J. Chem. Phys.* **2019**, *151*, No. 070902.
(3) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters* **1999**, *314*, 141−151.
(4) Yang, L.; Qin Gao, Y. A selective integrated tempering method. *J. Chem. Phys.* **2009**, *131*, No. 214109.
(5) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. *J. Chem. Theory Comput.* **2015**, *11*, 3584−3595.
(6) Bussi, G.; Laio, A. Using metadynamics to explore complex free-energy landscapes. *Nat. Rev. Phys.* **2020**, *2*, 200−212.
(7) Valsson, O.; Parrinello, M. Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* **2014**, *113*, 090601.
(8) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition path sampling. *Adv, Chem. Phys.* **2002**, *123*, 1−78.
(9) Zuckerman, D. M.; Chong, L. T. Weighted ensemble simulation: review of methodology, applications, and software. *Annu. Rev. Biophys.* **2017**, *46*, 43−57.
(10) Van Erp, T. S.; Bolhuis, P. G. Elaborating transition interface sampling methods. *J. Comput. Phys.* **2005**, *205*, 157−181.
(11) Elber, R.; Fathizadeh, A.; Ma, P.; Wang, H. Modeling Molecular Kinetics with Milestoning. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, e1512.
(12) Lu, H.; Tonge, P. J. Drug−target residence time: critical information for lead optimization. *Curr. Opin. Chem. Biol.* **2010**, *14*, 467−474.
(13) Copeland, R. A. The drug−target residence time model: A 10-year retrospective. *Nat. Rev. Drug Discovery* **2016**, *15*, 87−95.
(14) Huggins, D. J.; Sherman, W.; Tidor, B. Rational Approaches to Improving Selectivity in Drug Design. *J. Med. Chem.* **2012**, *55*, 1424−1444. PMID: 22239221.
(15) Fang, Y. Ligand−receptor interaction platforms and their applications for drug discovery. *Expert Opin. Drug Discovery* **2012**, *7*, 969−988.
(16) Tang, Z.; Roberts, C. C.; Chia-en, A. C. Understanding ligand-receptor non-covalent binding kinetics using molecular modeling. *Front. Biosci.-Landmark* **2017**, *22*, 960−981.
(17) Liu, W.; Jiang, J.; Lin, Y.; You, Q.; Wang, L. Insight into thermodynamic and kinetic profiles in small-molecule optimization. *J. Med. Chem.* **2022**, *65*, 10809−10847.
(18) Tonge, P. J. Drug−target kinetics in drug discovery. *ACS Chem. Neurosci.* **2018**, *9*, 29−39.
(19) Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C. New approaches for computing ligand−receptor binding kinetics. *Curr. Opin. Struct. Biol.* **2018**, *49*, 1−10.
(20) Wolf, S. Predicting Protein−Ligand Binding and Unbinding Kinetics with Biased MD Simulations and Coarse-Graining of Dynamics: Current State and Challenges. *J. Chem. Inf. Model.* **2023**, *63*, 2902−2910.
(21) Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of protein−ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E386−E391.
(22) Bernetti, M.; Masetti, M.; Recanatini, M.; Amaro, R. E.; Cavalli, A. An integrated Markov state model and path metadynamics approach to characterize drug binding processes. *J. Chem. Theory Comput.* **2019**, *15*, 5689−5702.
(23) Schuetz, D. A.; Bernetti, M.; Bertazzo, M.; Musil, D.; Eggenweiler, H.-M.; Recanatini, M.; Masetti, M.; Ecker, G. F.; Cavalli, A. Predicting residence time and drug unbinding pathway through scaled molecular dynamics. *J. Chem. Inf. Model.* **2019**, *59*, 535−549.
(24) Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative characterization of the binding and unbinding of millimolar drug fragments with molecular dynamics simulations. *J. Chem. Theory Comput.* **2017**, *13*, 3372−3377.
(25) Lamim Ribeiro, J. M.; Provasi, D.; Filizola, M. A combination of machine learning and infrequent metadynamics to efficiently predict kinetic rates, transition states, and molecular determinants of drug dissociation from G protein-coupled receptors. *J. Chem. Phys.* **2020**, *153*, No. 124105.
(26) Lee, S.; Wang, D.; Seeliger, M. A.; Tiwary, P. Calculating Protein−Ligand Residence Times through State Predictive Information Bottleneck Based Enhanced Sampling. *J. Chem. Theory Comput.* **2024**, *20*, 6341−6349.
(27) Wang, J.; Miao, Y. Ligand Gaussian Accelerated Molecular Dynamics 3 (LiGaMD3): Improved Calculations of Binding Thermodynamics and Kinetics of Both Small Molecules and Flexible Peptides. *J. Chem. Theory Comput.* **2024**, *20*, 5829−5841.
(28) Elber, R. A new paradigm for atomically detailed simulations of kinetics in biophysical systems. *Q. Rev. Biophys.* **2017**, *50*, e8.
(29) Elber, R. Milestoning: An efficient approach for atomically detailed simulations of kinetics in biophysics. *Annu. Rev. Biophys.* **2020**, *49*, 69−85.
(30) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **2004**, *120*, 10880−10889.
(31) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. On the assumptions underlying milestoning. *J. Chem. Phys.* **2008**, *129*, No. 174102.
(32) Aristoff, D.; Bello-Rivas, J. M.; Elber, R. A mathematical framework for exact milestoning. *Multiscale Model. Simul.* **2016**, *14*, 301−322.
(33) Cardenas, A. E.; Elber, R. Modeling kinetics and equilibrium of membranes with fields: Milestoning analysis and implication to permeation. *J. Chem. Phys.* **2014**, *141*, No. 054101.
(34) Lee, B. L.; Kuczera, K. Simulating the free energy of passive membrane permeation for small molecules. *Mol. Simul.* **2018**, *44*, 1147−1157.
(35) Votapka, L. W.; Amaro, R. E. Multiscale estimation of binding kinetics using Brownian dynamics, molecular dynamics and milestoning. *PLoS Computat. Biol.* **2015**, *11*, e1004381.
(36) Votapka, L. W.; Jagger, B. R.; Heyneman, A. L.; Amaro, R. E. SEEKR: simulation enabled estimation of kinetic rates, a computational tool to estimate molecular kinetics and its application to trypsin−benzamidine binding. *J. Phys. Chem. B* **2017**, *121*, 3597−3606.
(37) Ojha, A. A.; Srivastava, A.; Votapka, L. W.; Amaro, R. E. Selectivity and ranking of tight-binding JAK-STAT inhibitors using Markovian milestoning with Voronoi tessellations. *J. Chem. Inf. Model.* **2023**, *63*, 2469−2482.

(38) Ojha, A. A.; Votapka, L. W.; Amaro, R. E. QMrebind: incorporating quantum mechanical force field reparameterization at the ligand binding site for improved drug−target kinetics through milestoning simulations. *Chem. Sci.* **2023**, *14*, 13159−13175.
(39) Votapka, L. W.; Ojha, A. A.; Asada, N.; Amaro, R. E. Prediction of Threonine-Tyrosine Kinase Receptor−Ligand Unbinding Kinetics with Multiscale Milestoning and Metadynamics. *J. Phys. Chem. Lett.* **2024**, *15*, 10473−10478.
(40) Narayan, B.; Yuan, Y.; Fathizadeh, A.; Elber, R.; Buchete, N.-V. Long-time methods for molecular dynamics simulations: Markov State Models and Milestoning. *Prog. Mol. Biol. Transl. Sci.* **2020**, *170*, 215−237.
(41) Husic, B. E.; Pande, V. S. Markov state models: From an art to a science. *J. Am. Chem. Soc.* **2018**, *140*, 2386−2396.
(42) Chodera, J. D.; Noé, F. Markov state models of biomolecular conformational dynamics. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135−144.
(43) Allen, R. J.; Valeriani, C.; ten Wolde, P. R. Forward flux sampling for rare event simulations. *J. Phys.: Condens. Matter* **2009**, *21*, No. 463102.
(44) Goutelle, S.; Maurin, M.; Rougier, F.; Barbaut, X.; Bourguignon, L.; Ducher, M.; Maire, P. The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam. Clin. Pharmacol.* **2008**, *22*, 633−648.
(45) Ojha, A. A.; Thakur, S.; Ahn, S.-H.; Amaro, R. E. DeepWEST: Deep learning of kinetic models with the Weighted Ensemble Simulation Toolkit for enhanced sampling. *J. Chem. Theory Comput.* **2023**, *19*, 1342−1359.
(46) Ahn, S.-H.; Ojha, A. A.; Amaro, R. E.; McCammon, J. A. Gaussian-Accelerated molecular dynamics with the weighted ensemble method: A hybrid method improves thermodynamic and kinetic sampling. *J. Chem. Theory Comput.* **2021**, *17*, 7938−7951.
(47) Bello-Rivas, J. M.; Elber, R. Exact milestoning. *J. Chem. Phys.* **2015**, *142*, No. 094102.
(48) Vanden-Eijnden, E.; Venturoli, M. Markovian milestoning with Voronoi tessellations. *J. Chem. Phys.* **2009**, *130*, No. 194101.
(49) Ray, D.; Andricioaei, I. Weighted ensemble milestoning (WEM): A combined approach for rare event simulations. *J. Chem. Phys.* **2020**, *152*, No. 234114.
(50) Ray, D.; Stone, S. E.; Andricioaei, I. Markovian weighted ensemble milestoning (M-WEM): Long-time kinetics from short trajectories. *J. Chem. Theory Comput.* **2022**, *18*, 79−95.
(51) Votapka, L. W.; Stokely, A. M.; Ojha, A. A.; Amaro, R. E. SEEKR2: Versatile multiscale milestoning utilizing the OpenMM molecular dynamics engine. *J. Chem. Inf. Model.* **2022**, *62*, 3253−3262.
(52) Ojha, A. A.; Votapka, L. W.; Huber, G. A.; Gao, S.; Amaro, R. E. An introductory tutorial to the SEEKR2 (Simulation enabled estimation of kinetic rates v. 2) multiscale milestoning software [Article v1.0]. *Living J. Comput. Mol. Sci.* **2023**, *5*, No. 2359.
(53) Luty, B. A.; McCammon, J. A.; Zhou, H.-X. Diffusive reaction rates from Brownian dynamics simulations: Replacing the outer cutoff surface by an analytical treatment. *J. Chem. Phys.* **1992**, *97*, 5682−5686.
(54) Ahn, S.-H.; Jagger, B. R.; Amaro, R. E. Ranking of ligand binding kinetics using a weighted ensemble approach and comparison with a multiscale milestoning approach. *J. Chem. Inf. Model.* **2020**, *60*, 5340−5352.
(55) Jagger, B. R.; Ojha, A. A.; Amaro, R. E. Predicting ligand binding kinetics using a Markovian milestoning with voronoi tessellations multiscale approach. *J. Chem. Theory Comput.* **2020**, *16*, 5348−5357.
(56) Narayan, B.; Buchete, N.-V.; Elber, R. Computer simulations of the dissociation mechanism of Gleevec from Abl Kinase with milestoning. *J. Phys. Chem. B* **2021**, *125*, 5706−5715.
(57) Rathnayake, S.; Narayan, B.; Elber, R.; Wong, C. F. Milestoning simulation of ligand dissociation from the glycogen synthase kinase 3β. *Proteins: Struct., Funct., Bioinf.* **2023**, *91*, 209−217.
(58) Boothroyd, S.; Behara, P. K.; Madin, O. C.; Hahn, D. F.; Jang, H.; Gapsys, V.; Wagner, J. R.; Horton, J. T.; Dotson, D. L.; Thompson, M. W.; et al. Development and benchmarking of open force field 2.0. 0: the Sage small molecule force field. *J. Chem. Theory Comput.* **2023**, *19*, 3251−3275.
(59) Wang, L.; Behara, P. K.; Thompson, M. W.; Gokey, T.; Wang, Y.; Wagner, J. R.; Cole, D. J.; Gilson, M. K.; Shirts, M. R.; Mobley, D. L. The Open Force Field Initiative: Open Software and Open Science for Molecular Modeling. *J. Phys. Chem. B* **2024**, *128*, 7043−7067.
(60) Qiu, Y.; Smith, D. G.; Boothroyd, S.; Jang, H.; Hahn, D. F.; Wagner, J.; Bannan, C. C.; Gokey, T.; Lim, V. T.; Stern, C. D.; et al. Development and benchmarking of open force field v1. 0.0−the parsley small-molecule force field. *J. Chem. Theory Comput.* **2021**, *17*, 6262−6280.
(61) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr; et al. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(62) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* **2011**, *115*, 580−596.
(63) Woods, C. J.; Manby, F. R.; Mulholland, A. J. An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J. Chem. Phys.* **2008**, *128*, 014109.
(64) Dinur, U.; Hagler, A. T. Geometry-dependent atomic charges: Methodology and application to alkanes, aldehydes, ketones, and amides. *J. Comput. Chem.* **1995**, *16*, 154−170.
(65) Wolf, S.; Post, M.; Stock, G. Path separation of dissipation-corrected targeted molecular dynamics simulations of protein−ligand unbinding. *J. Chem. Phys.* **2023**, *158*, No. 124106.
(66) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690.
(67) Kokh, D. B.; Doser, B.; Richter, S.; Ormersbach, F.; Cheng, X.; Wade, R. C. A workflow for exploring ligand dissociation from a macromolecule: Efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories. *J. Chem. Phys.* **2020**, *153*, No. 125102.
(68) Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H.-P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; et al. Estimation of drug−target residence times by τ-random acceleration molecular dynamics simulations. *J. Chem. Theory Comput.* **2018**, *14*, 3859−3869.
(69) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(70) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(71) Miao, Y.; Bhattarai, A.; Wang, J. Ligand Gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *J. Chem. Theory Comput.* **2020**, *16*, 5526−5547.
(72) Ansari, N.; Rizzi, V.; Parrinello, M. Water regulates the residence time of Benzamidine in Trypsin. *Nat. Commun.* **2022**, *13*, 5438.
(73) Tian, H.; Jiang, X.; Trozzi, F.; Xiao, S.; Larson, E. C.; Tao, P. Explore protein conformational space with variational autoencoder. *Front. Mol. Bio.* **2021**, *8*, 781635.
(74) Jing, B.; Stärk, H.; Jaakkola, T.; Berger, B. Generative modeling of molecular dynamics trajectories. *arXiv (Quantitative Biology.Biomolecules)*, September 26, 2024, 2409.17808, ver. 1. https://arxiv.org/abs/2409.17808.
(75) Takaba, K.; Friedman, A. J.; Cavender, C. E.; Behara, P. K.; Pulido, I.; Henry, M. M.; MacDermott-Opeskin, H.; Iacovella, C. R.; Nagle, A. M.; Payne, A. M.; et al. Machine-learned molecular mechanics force fields from large-scale quantum chemical data. *Chem. Sci.* **2024**, *15*, 12861−12878.
(76) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine learning force fields. *Chem. Rev.* **2021**, *121*, 10142−10186.

https://doi.org/10.1021/acs.jctc.4c01108
J. Chem. Theory Comput. 2024, 20, 9759−9769