Understanding the impact of binding free energy and kinetics calculations in modern drug discovery

Victor A. Adediwura[1†], Kushal Koirala[1†], Hung N. Do[2,3], Jinan Wang[1], Yinglong Miao[1*]

[1]Department of Pharmacology and Computational Medicine Program, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[2]Center for Computational Biology, University of Kansas, Lawrence, KS, USA

[3]Present address: Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA

†These authors contributed equally.

*Correspondence: yinglong_miao@med.unc.edu

Abstract

Introduction:

For rational drug design, it is crucial to understand the receptor-drug binding processes and mechanisms. A new era for the use of computer simulations in predicting drug-receptor interactions at an atomic level has begun with remarkable advances in supercomputing and methodological breakthroughs.

Areas covered:

End-point free energy calculation methods such as Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) or Molecular-Mechanics/Generalized Born Surface Area (MM/GBSA), free energy perturbation (FEP), and thermodynamic integration (TI) are commonly used for

binding free energy calculations in drug discovery. In addition, kinetic dissociation, and association rate constants (k_{off} and k_{on}) play critical roles in the function of drugs. Nowadays, Molecular Dynamics (MD) and enhanced sampling simulations are increasingly being used in drug discovery. Here, we present a review of the computational techniques used in drug binding free energy and kinetics calculations.

Expert Opinions:

The applications of computational methods in drug discovery and design are expanding, thanks to improved predictions of the binding free energy and kinetic rates of drug molecules. Recent microsecond-timescale enhanced sampling simulations have made it possible to accurately capture repetitive ligand binding and dissociation, facilitating more efficient and accurate calculations of ligand binding free energy and kinetics.

Keywords: Computer-aided drug design, Molecular Dynamics, Enhanced sampling, Free Energy, Kinetics.

Article Highlights:

- Pharmacodynamics prediction using computer simulations is growing rapidly in the field of drug design and discovery.
- Accurate prediction of k_{on} and k_{off} using computational techniques is currently trending in the field of drug design.
- MM/PBSA, MM/GBSA, FEP and TI are common techniques used in free energy calculations.
- Enhanced sampling methods are advantageous in exploring drug binding and dissociation pathways and kinetics.

1 Introduction

Drug discovery and development is a time-consuming and expensive process. A new drug's development is expected to typically cost $2.6 billion and take 10–12 years to reach the consumer market (**1**). A poor understanding of pharmacodynamics of drug action has hindered effective design of drugs. However, considering recent advances in molecular modeling and simulation algorithms, pharmaceutical companies and scientists have increased their efforts to predict binding free energies and kinetics of drugs using computational approaches **(2-8)**.

The use of computer simulations to predict ligand binding kinetics and thermodynamics to therapeutic targets are gaining increasing recognition in recent years (**9**). At the forefront, molecular docking and scoring are crucial computational techniques in drug design (**10, 11**), used to predict drug binding modes and assess their binding affinities. While these methods are known for their efficiency, their accuracy is limited (**10**). Specifically, their accuracy wanes when distinguishing drugs with subtle differences in binding affinity, highlighting a crucial limitation in their applicability. A typical example where molecular docking might fall short is in estimating the binding affinities of congeneric ligands (ligands with the same core structure but differs in certain substituents or functional group attached to the core) and drugs with less than a tenfold difference in binding affinity (**10, 11**).

To address these limitations, molecular dynamics (MD) simulations offer valuable insights into molecular interactions under varied environmental conditions. Nonetheless, conventional MD simulations often suffer from insufficient sampling of high energy barriers in complex systems

(**12**). These limitations can lead to inefficiencies or inaccuracies in simulations, especially when dealing with dynamic systems exhibiting significant conformational changes (**13**).

In response to these challenges, advancement in computational technology (**14**), have paved the way for the adoption of alchemical free energy methods such as Free Energy Perturbation (FEP) and Thermodynamic Integration (TI) (**15)**. These methods and their variants offer rigorous approaches for calculating binding free energies, benefitting from recent innovations like those introduced by the York lab in 2023, which optimizes sampling of alchemical transformation pathways in AMBER software suite using a novel λ-dependent weight functions and softcore potential to increase sampling efficiency and ensure stable performance at critical points where λ is equal to 0 or 1 (**15,16**). Despite these advancements, the substantial computational demands of such detailed simulations make the adoption of more approximate methods such as Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) and Molecular-Mechanics/Generalized Born Surface Area (MM/GBSA) (**17,18,19,20,21,22**) more attractive.

MM/PB(GB)SA offer a compromise between computational demand and the depth of insight into ligand binding processes and as such, it may lead to limited precision in predicting binding energies (**10**). Furthermore, the emergence of enhanced sampling simulations marks a significant stride in understanding molecular systems' dynamic behavior, especially in studies of ligand binding kinetics **(5, 6)**. These techniques, particularly useful in ligand binding kinetics studies, enrich our comprehension of how molecules interact over time, providing a more complete picture of the binding process from docking predictions through to detailed kinetic and thermodynamic evaluations.

Enhanced sampling techniques using pre-defined collective variables (CVs) for effective simulations include, but are not limited, to umbrella sampling (**23**), adaptive biasing force (ABF)

(**24**), Metadynamics (MetaD) (**25**), conformational flooding **(26**), and variationally enhanced sampling (VES) (**27**). These techniques biases MD simulations to pinpoint various transition states along selected CVs. This is achieved through modifying the forces or employing external bias potentials. On the other hand, CV-free methods that have been applied to investigate ligand binding free energies and kinetics include Gaussian accelerated molecular dynamics variant (LiGaMD) (**5),** τ random accelerated molecular dynamics (τRAMD) (**28**), and dissipation-corrected targeted MD (dcTMD) (**29**), etc.

We will discuss principles of free energy methods including MM-PBSA, MM-GBSA, FEP and TI with their applications and limitations, as well as enhanced sampling methods for the calculations of ligand binding thermodynamics and kinetics.

2 Methods for binding free energy calculations

In this section, we explore various computational approaches for predicting binding free energies. Specifically, we delve into some methods, applications, and limitations of MM/PBAS, MM/GBSA, FEP and TI. These methods are relevant in quantitatively estimating the binding free energies of protein-ligand complexes.

2.1 Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area (MM/PBSA and MM/GBSA)

The MM/PBSA and MM/GBSA methods are popular for computing the ligand binding free energy (**10,30**). In principle, the MM/PB(GB)SA accounts for the energetic contribution per residue by decomposing the total system's free energy into various components, such as van der Waals, electrostatics, and the internal energies (bond, angle, and dihedral energies) (**10,31**). In practice, the MM/PB(GB)SA calculate the binding free energy between a protein and a ligand by estimating the free energy of the protein-ligand complex (PL), the free energy of the unbound protein (P), and

the free energy of the unbound ligand (L) separately. The binding free energy ΔG_{bind} is the calculated using the equation below:

$$\Delta G_{bind} = G_{PL} - (G_P + G_L),$$

where G_{PL} is the free energy of the protein-ligand complex, G_P is the free energy of the unbound protein, and G_L is the free energy of the unbound ligand (**10,31**). The rationale for MM/PB(GB)SA methods is rooted in a strategic compromise of optimizing computational efficiency while still maintaining a significant degree of accuracy. This approach allows simulations and analyses of molecular interactions within a feasible timeframe and with available computational resources, thus accelerating the pace of drug discovery. By judiciously balancing these two critical factors, MM/PB(GB)SA provides a powerful toolkit for enabling the exploration of complex molecular phenomena that were previously beyond reach due to computational limitations (**32,33,34**).

In most drug discovery efforts, the technique of virtual screening, when used in conjunction with MM/PB(GB)SA methods, has proven to be highly effective. This combination offers crucial insights by improving the ranking of binding affinities, accurately predicting how effectively different molecules will bind to a target and identifying the correct mode of ligand binding. Essentially, this approach enhances the precision of identifying promising compounds early in the drug discovery workflow (**35,36,37,38,39,40,41**). To make MM/PB(GB)SA more accessible, Wang et al. (**42**) developed a new webserver called fastDRH in 2022. This webserver integrates Autodock Vina and Autodock-GPU for the docking process, which predicts the optimal positioning of small molecules within the binding sites of target proteins. Additionally, it employs a streamlined, or 'structure-truncated' version of the MM/PB(GB)SA method to refine the docking calculations. This refined approach focuses on the most relevant parts of the molecules to efficiently predict the binding free energies. Thus, it ensures speed and less computational demand

without sacrificing accuracy (**42**). All these features are integrated into a platform that is easy to use for both experts and novices. Moreover, parameter tuning helps to improve the accuracy of MM/PB(GB)SA (**43**). For example, Wang et al. (**43**) recommend some parameters including a membrane dielectric constant of 7.0 and an internal dielectric constant of 20.0. These parameters have been found to recapitulate experimental binding affinity in soluble proteins and membrane-bound proteins (**43**). Rastelli et al. (**44)** showed that MM/PB(GB)SA techniques achieved larger Area Under Curve (AUC) and enrichment factor values than conventional docking approaches in virtual screening. Additionally, Zhang et al. (**39)** extended this evaluation to 38 drug targets in the Database for Useful Decoys (DUD) database and obtained a similar conclusion. In another example, Zhong et al. **(45)** used the interaction entropy (IE) technique in the MM/PBSA along with two MM/GBSA models (GB^{HCT} and GB^{OBC1}) to study the role of entropy and computed the binding affinities of 176 protein-ligand and protein-protein complexes within the Bcl-2 family. The results showed significant improvements in differentiating the native structure from decoys in both protein-ligand and protein-protein systems. According to their results, the GB^{HCT} model and IE technique combination had the best results, with an AUC of 0.97. Pan et al. **(40**) applied MM/PBSA to rank xanthine oxidase inhibitors in respect to their potency and their results correlated well with experimental results. Zhong et al. **(45)** demonstrated that incorporating interaction entropy in MM/PB(GB)SA calculations enhances the accuracy of binding affinity predictions of small molecules to Bcl-2 family targets. However, recent studies (**46,38)** have suggested that IE could reduce the accuracy of prediction for protein-small molecule interactions, specifically in the context of wild-type and mutant Epidermal Growth Factor Receptor (EGFR) systems (**46**). These disparities suggest that parameter tuning in MM/PB(GB)SA calculation is system specific and effort should be made by researchers in testing different parameters to obtain

reasonable accuracy of binding affinities. Another study by Crean et al. found IE to fall short in estimating accurate binding affinities in protein-protein interaction systems (**38**).

Although MM/PB(GB)SA has been effectively applied to recapitulate experimental data and enhance the outcomes of structure-based drug discovery, they often suffer from a number of limitations. First, the use of the implicit solvent continuum assumes a large approximation by neglecting water molecules. By extension, the energetic contributions of water enthalpy and entropy to the molecular system are thus neglected. The need for explicitly considering water molecules in these methods cannot be over-emphasized. Water molecules form hydrogen bonds with proteins or ligands and stabilize the overall structure of the protein. In case water molecules are tightly bound to the binding sites, the implicit water model will severely affect the prediction accuracy (**47,48**).

One of the major challenges in these methods is the precise calculation of conformational entropy, which is crucial for accurate binding free energy predictions. Normal mode analysis is employed to evaluate the vibrational frequencies of molecules, providing insights into molecular disorder (**49)**. However, this approach is sometimes overlooked in its tendency to introduce large statistical variations. Such variability can lead to significant approximations, potentially compromising the precision of binding affinity estimations (**49,50,51,52**).

2.2 Alchemical free energy perturbation

Alchemical free energy perturbation (FEP) is a computational technique used to calculate the free energy difference between two ligands by transforming one into the other through a series of transitional states while maintaining a seamless and uninterrupted route. Combining MD simulations and statistical mechanics, FEP measures the variations in free energy due to ligand

modifications, solvation effects, and molecular structure modifications. The popular Zwanzig exponential averaging equation (**53)** can be applied to the alchemical process to quantify the relative binding energy difference between the two ligands. The Zwanzig equation is given as:

$$e^{-\frac{1}{k_B T}\Delta F_{A\to B}} = \left\langle e^{-\frac{1}{k_B T}\Delta U_{A\to B}(x)} \right\rangle, \quad 1$$

where ΔF stands for the change in free energy required to change from ligand A to ligand B, and ΔU represents a mean estimate of the change in potential energy required to change from ligand A to ligand B as a function of reaction coordinate x .

Zwanzig's master equation enables the calculation of thermodynamic variances between two states, A and B. Prior to this, Kirkwood introduced a lambda (λ) parameter for improving the precision of free energy perturbation calculations in chemical transformation (**15,54)**. Lambda serves as a couple parameter, enabling the seamless transition between different states of a molecular system within simulations. It operates on a range between 0 and 1, where λ =0 denotes the system's initial state (A), and λ =1 represents the final state (B). Intermediates values of λ allow for a gradual modulation of interactions, facilitating a stepwise transformation between these states (**15,54**). This approach, by varying λ in small increments, permits the calculation of free energy difference between the initial and final states by integrating the energy changes associated with these incremental steps. The introduction of the λ parameter thus significantly improves the accuracy of FEP calculations, providing a nuanced tool for exploring the energetics behind molecular transformations, solvation effects, and binding interactions. Bennet later introduced a Bennet Acceptance Ratio (BAR) method, aimed at minimizing the squared error of the calculations (**55)**. In this context, the squared error refers to the average of the squared differences between the calculated and true values, a measure of the accuracy and reliability of computational predictions. By focusing on reducing this error, the BAR method enhances the precision and effectiveness of

free energy calculations (**55**). The BAR method underwent further refinement by adopting a statistically optimal evaluation of FEP simulations. This enhancement led to the creation of the Multistate Bennett Acceptance Ratio (MBAR), which expands BAR methodology to efficiently estimate free energy differences across multiple states, leveraging the collective data from overlapping simulations for improved accuracy and efficiency (**56,57**).

FEP calculations are shown to be highly effective in predicting the ligand binding affinities in many systems (**58-61**). For example, FEP has been applied to study the effect of mutation on protein-protein interactions particularly in the context of SARS-CoV-2 RBD: ACE2 binding affinity (**62**), identify novel allosteric inhibitor of human transcription factor (**63**) and discover Nirmatrelvir resistance mutations in SARS-CoV-2 3CLpro (**64**). Other variants of FEP have started to yield significant improvements in computing accurate relative binding free energy of congeneric ligands. Wang et al. (**65)** employed an improved forcefield, OPLS2.1 (**66)** and performed replica exchange molecular dynamics (REMD) simulation (**67-69)** to accurately predict binding affinities of 200 ligands across different targets. Similarly, Lenselink et al. (**70)** used the FEP+ approach as previously described by Wang et al. (**65**) to accurately predict the binding affinities of 45 ligands across 4 targets of GPCRs. Moreover, their findings outline a procedure for using FEP+ on GPCRs and offer practical implementation strategies for discovering potent compounds in lead optimization initiatives (**70**). FEP+ has also been applied in the hit-to-lead optimization campaign. Sun et al. utilized FEP+ during the hit-to-lead phase of a drug discovery initiative aimed at targeting soluble adenyl cyclase (**71**). They used FEP+ to discover a more favorable chemotype and enhance binding affinity to levels below sub-nanomolar, maintaining drug-like characteristics. Aside from the utility of FEP+ in protein-ligand studies, recent studies have demonstrated the use of FEP theory to predict the binding energies between antibody and antigen (**72-74**), thereby aiding

the design and optimization of antibodies (**75,76**). The FEP method and its variants are particularly attractive because of their high accuracy, usually within the 1.0 kcal/mol range (**70)**.

In addition, FEP and its variants have been demonstrated in many research papers to rigorously recapitulate binding free energies that correlate with experimental results (**58,59,61**). Chen et al. (**77**) implemented Absolute Protein-Ligand Binding Free Energy Perturbations (ABFEP) for a hit discovery. The binding free energies calculated in their study show a correlation with experimental data, achieving a weighted average correlation coefficient (R^2) of 0.55 across the whole entire dataset of 8 congeneric ligands and 8 targets (**77**). FEP calculations can be run independently in parallel. Hence, it is feasible to investigate whether the introduction of a functional group or the swapping of one atom for another might increase or decrease the ligand binding affinity **(78-80)**.

One of the associated issues with FEP is the force field. In classical force fields, the potential energy function of all atoms in the system is often approximated by training with experimental data and quantum mechanics. Lenselink et al. (**70**) successfully predicted binding affinities of congeneric ligands that correlate with the experimental values. The success of their published work could be attributed to improved forcefields like OPLS2.1 (improved nonbonded van der Waals interactions, partial charges, and torsional parameters) **(66)**. Even the most sophisticated free-energy methodologies will lead to an incorrect conclusion in the absence of an accurate molecular mechanical force field. In addition, FEP calculation can suffer from sampling issues. The presence of multiple-high energy wells may cause the system to become trapped in a local minimum, preventing rigorous sampling across the configuration phase space (**81,82,83).** It is incredibly challenging to account for significant protein flexibility in the FEP calculations **(84,85)**. This limitation prevents the use of FEP to investigate protein-ligand binding in the presence of substantial protein motions. The consideration of biologically relevant motions underlying

biomolecular recognition is made possible by the capacity of conformational changes to be tracked by MD simulations. Additionally, current implementations of Relative Binding Free Energy (RBFE) in drug research typically conduct only a few nanoseconds of MD simulations for each intermediate's λ window (**83**). This restricted duration hampers the comprehensive analysis of conformational changes. The primary challenges faced in RBFE simulations include understanding the conformational free-energy landscape associated with the target molecule and creating tailored approaches informed by this landscape. Additionally, there are specific hurdles such as effectively handling covalently bound ligands (**86)**, ensuring convergence of the simulations (**82,83)**, and accurately accounting for multiple binding poses (**87)**.

2.3 Thermodynamic Integration (TI)

Thermodynamic integration (TI) is a well-established technique to determine the binding free energy difference of two ligands. One can estimate this free energy difference by integrating over a range of a coupling parameter Lambda, λ. TI involves performing an alchemical transformation that connects the two ligands L_1 and L_2 through a series of intermediate states. By smoothly varying λ from 0 to 1, the system undergoes a gradual transition from the initial state (L_1) to the final state (L_2). When λ = 0, the system is assumed to represent L_1 state and when λ = 1, the system corresponds to L_2 (**88**). The potential energy of the system can be defined as:

$$V(\lambda, x) = (1 - \lambda)V_1(x) + V_2(x), \qquad 2$$

where, V_1 and V_2 represent the potential energies of L_1 and L_2, respectively, x denotes the system's coordinates, and λ serves as an interpolation parameter that linearly combines, V_1 and V_2 to transition the system from one state to another. Integrating the derivative of the potential energy

with respect to λ over the interval of 0 to 1 gives the free energy difference (ΔG) between the two states.

$$\Delta G = \int_0^1 \left\langle \frac{\partial V(\lambda, x)}{\partial \lambda} \right\rangle_\lambda . d\lambda \qquad 3$$

where $\left\langle \frac{\partial V(\lambda,x)}{\partial \lambda} \right\rangle_\lambda$ represents an ensemble average of the derivative of the potential energy with respect to λ at state λ. This equation calculates the free energy difference between the two states by considering the changes in potential energy as the system transition from one state to another. (**88**).

Zou et al. applied TI in Amber 18 to predict RBFE for a set of 39 ligands of Cathepsin S protein. The predicted values correlated well with experimental data (**89**). In 2020, He et al. (**90**) used CPU-based TI and GPU-TI to generate binding free energies of 134 ligands binding to four different proteins. TI method used in prediction of RBFE in antibody and antigen system of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) outperformed knowledge-based discrimination of beneficial and deleterious mutations and improved the binding affinity and neutralization potency of antibody (**91**).

TI calculations involve conducting a MD simulation for each specific λ window. As outlined in equation 3, the determination of the binding affinity hinges on the derivative of the potential energy with respect to λ. For the calculated change in free energy to be accurate, it is crucial that this derivative is measured precisely and reliably across each λ window. However, when sampling is biased on a single MD trajectory utilizing a Gaussian random process, challenges in reproducibility may arise, leading to inconsistencies in the calculated binding energy. A Gaussian random process, in this context, refers to a statistical method for predicting the outcomes of complex systems where variables evolve in a way that, at any given point, their distribution follows a Gaussian distribution,

characterized by its mean and variance. This approach is commonly used in simulation to model the randomness inherent in molecular movements, but its reliance on a single trajectory can introduce variability that undermines reproducibility (**92)**. To address these reproducibility concerns, an alternative method proposed by Bhat et al. (**92**) involves computing ensemble averages of MD trajectories. This method, known as ensemble-based thermodynamic integration (TIES), mitigates the issue of irreproducibility by averaging the outcomes over multiple MD simulations rather than relying on a single trajectory. By incorporating multiple simulation trajectories, TIES leverages the diversity of molecular configurations and interactions captured across different trajectories, thereby offering a more robust and reliable estimation of the free energy change (**92**).

3 Enhanced Sampling Techniques for Predicting Ligand Binding Kinetics

While free energy methods offer valuable insights into the thermodynamic aspects of molecular systems, enhanced sampling methods extend our capabilities to explore the dynamic behavior of these systems, offering a more complete picture of drug-target interactions. However, they often face challenges such as low efficacy for sampling rare events or overcoming high energy barriers in complex systems (**12,13**).

3.1 Metadynamics (MetaD)

Metadynamics (**25**) is designed to improve the sampling of rare events. A history-dependent external biasing potential is applied to the system. This potential is dependent on collective variables (CVs) and is expressed as an accumulation of Gaussian functions distributed over the

space of CVs (**25**). Gradual increase in the bias encourages the system to explore the previously unsampled regions and estimate the free energy surface.

In 2013, Infrequent metadynamics was introduced to compute the kinetic properties of molecular systems. Unlike standard metadynamics where biases are added frequently, infrequent metadynamics adopts a different approach. It extends the time interval between the addition of consecutive biasing potentials, effectively reducing their frequency (**93**). This modification is crucial in the context of transition states. Given the short time required to cross the transition states, extending the time interval between bias applications effectively minimizes potential bias in these regions, thereby facilitating more accurate and unbiased exploration of transition states (**93).**

Pramanik et al. (**94**) used infrequent metadynamics to benchmark the dissociation kinetics of two different millimolar fragments of FKBP protein and compared the data with unbiased MD simulations (**94**). For 4-hydroxy-2-butanone the residence time obtained from infrequent metadynamics was 27.3 ± 0.1 ns with 8 ps biasing frequency and 21.3 ± 0.2 ns from unbiased MD, demonstrating good agreement between the findings. For 4-diethylamino-2-butanone, infrequent metadynamics with 60 ps biasing frequency predicted the residence time of 1.83±0.03μs and unbiased MD predicted 0.54μs (**94**). Using the L99A variant of T4L as a model system, Wang et al. (**95**) applied infrequent metadynamics simulations to successfully capture repetitive ligand binding/unbinding. The total of ∼12 μs infrequent metadynamics simulations were used to capture 20 ligand dissociation and ligand association events. Based on these events, k_{on} and k_{off} were predicted to be 3.5 ± 2 × $10^4 M^{-1}s^{-1}$ and 7 ± 2 s^{-1}, respectively (**95**). One-to-two orders of magnitude difference were observed compared with experimental values of 0.8±1×$10^6 M^{-1}s^{-1}$and 800±20 s^{-1} (**95**).

Lamim et al. **(96**) used a combination of unbiased MD, machine learning (ML), and infrequent metadynamics to investigate the dissociation rates of two drugs (morphine and buprenorphine) from the μ-opioid receptor (**97**). The Automatic Mutual Information Noise Omission (AMINO) method (**98**) was employed to screen an extensive array of molecular features, with the objective of isolating a highly relevant subset. This technique leverages the principle of mutual information to evaluate and discern the interdependencies among various molecular features, thereby identifying those that are most significant and informative. By systematically excluding features deemed as noise or less relevance, the AMINO method streamlines the dataset to a more focused collection of features (**98**) This refined subset offers more insightful contributions to the subsequent analysis or predictive modeling. The Reweighted Autoencoded Variational Bayes for Enhanced Sampling (RAVE) (**99**) method was then used for predicting optimal reaction coordinates as CVs for infrequent metadynamics. A biasing frequency of 30 ps provided a good speed-accuracy trade-off to study ligand dissociation from the receptor. The k_{off} value of 3.19 min^{-1} calculated for morphine agreed well with experimentally obtained k_{off} value of 1.388 ± 0.1 min^{-1} (**100**). The k_{off} value of 1.27 min^{-1} was calculated for buprenorphine, which deferred by one order of magnitude with the experimentally calculated 0.106 ± 0.02 min^{-1} (**100**).

3.2 Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD)

Gaussian accelerated MD (GaMD) adds a harmonic boost potential to smooth the potential energy surface of the biomolecules. Cumulant expansion to the second-order aids in accurately reweighting the GaMD simulations. GaMD allows for unconstrained enhanced sampling without predefined CVs. Based on GaMD, various selective GaMD algorithms are developed to estimate biomolecular binding kinetics **(101**). LiGaMD enables us to efficiently simulate repetitive ligand

binding and unbinding processes in protein-ligand systems and thus characterize both kinetics and thermodynamics of ligand binding (**5**).

For a system of ligand L binding to protein P in biological environment E, the potential energy of the system could be decomposed as

$$V(r) = V_{L,b}(r_L) + V_{P,b}(r_P) + V_{E,b}(r_E) + V_{LL,nb}(r_L) + V_{PP,nb}(r_P) + V_{EE,nb}(r_E) + V_{PL,nb}(r_{PL}) + V_{LE,nb}(r_{LE}) + V_{PE,nb}(r_{PE}), \quad (4)$$

where $V_{L,b}$, $V_{P,b}$, $V_{E,b}$ represent the bonded potential energies in *L, P,* and *E* respectively. $V_{LL,nb}$, $V_{EE,nb}$, $V_{PP,nb}$ denote self-non-bonded potential energies and $V_{LE,nb}$, $V_{PE,nb}$, $V_{PL,nb}$ are non-bonded interactions between *L-E, P-E,* and *P-L,* respectively. Since ligand binding mostly involves non-bonded interaction energies of ligand, $V_{L,nb}(r) = V_{LL,nb}(r_L) + V_{LE,nb}(r_{LE}) + V_{PL,nb}(r_{PL})$, the LiGaMD selectively boosts these potential terms (**5**). To facilitate the ligand rebinding process, another boost is added to the remaining potential energy terms of the system. A recently developed LiGaMD2 method applies selective boost potential to both the ligand and protein residues in the binding pocket to facilitate the ligand binding and dissociation process in a closed (**6**).

LiGaMD was demonstrated on repetitive binding and dissociation of Nirmatrevlir drug in the 3CLpro binding domain with predicted k_{on} and k_{off} as $3.2 \pm 0.21 \times 10^5 M^{-1}s^{-1}, 2.92 \pm 0.37s^{-1}$, respectively (**5**). Since there were no experimental data for binding rates, $K_d = \left(\frac{k_{off}}{k_{on}}\right)$, was used to calculate the equilibrium dissociation constant. The predicted K_d rate of $9.10 \pm 0.29nM$ was found to be consistent with the experimental value of $7.3 \pm 3nM$ (**5**). In microsecond LiGaMD simulations, repetitive binding and dissociation of benzamidine in trypsin were observed. The benzamidine binding and dissociation rates were predicted as $1.5 \pm 0.79 \times 10^7 M^{-1}s^{-1}$ and $3.53 \pm 1.41s^{-1}$, respectively (**5**). The binding rate closely aligned with the experimental value of $2.9 \times 10^7 M^{-1}s^{-1}$, whereas the calculated dissociation rate differed by two orders of magnitude

in comparison with the experimental value of $600 \pm 300 s^{-1}$ (**102**). Similarly, the LiGaMD2 method predicted ligand kinetics, k_{on} and k_{off}, in four different complexes of small molecule bound to the L99A T4 lysozyme (T4L) mutants. In benzene-L99A T4L system, the predicted k_{on} and k_{off} were $7.42 \pm 4.81 \times 10^6 M^{-1}s^{-1}$ and $1440 \pm 880 s^{-1}$, respectively (**6**). These values agreed well with experimentally obtained k_{on} and k_{off} of $0.7 - 1.0 \times 10^6 M^{-1}s^{-1}$and $950 s^{-1}$ respectively. Similarly, k_{on} and k_{off} calculated for the benzene-M102A T4L system were $9.57 \pm 6.29 \times 10^6 M^{-1}s^{-1}$ and $2011 \pm 1606 s^{-1}$, respectively. These rates also agreed well with experimentally values of $3 - 5 \times 10^6 M^{-1}s^{-1}$and $3000 s^{-1}$, respectively. The predicted k_{on} and k_{off} values of the T4L:L99A-IND systems were $2.99 \pm 2.87 \times 10^6 M^{-1}s^{-1}$ and $3494 \pm 559 s^{-1}$, respectively, which were comparable to experimental values of $0.7 - 1.0 \times 10^6 M^{-1}s^{-1}$ and $325\ s^{-1}$, respectively (**6**).

3.3 τ Random accelerated molecular dynamics (τRAMD)

τRAMD (**103**) is based on the random accelerated molecular dynamics (RAMD) technique that is designed to investigate ligand dissociation pathways from deep binding pockets in proteins. In the RAMD technique, molecular simulation is enhanced by adding a small randomly directed force to facilitate ligand dissociation. If the movement of the ligand falls below a given threshold value within a defined time interval, the direction of the force is randomly reassigned to aid in the unbinding event. This process continues until the ligand's displacement exceeds a specified distance from its initial position. At this point, the ligand is assumed to dissociate from the proteins (**103**). τ-RAMD does not necessitate prior knowledge of the dissociation pathway nor requires extensive parameter fitting. Here, the magnitude of the randomly oriented force is specified by the user to facilitate the ligand dissociation from the protein pocket.

Kokh et al. (**103**) applied τRAMD to calculate the residence time of 70 diverse drug-like inhibitors of N-HSP90. The computationally computed residence time, τ_{comp}, was plotted against experimentally obtained residence time, τ_{expt}. Among different classes of drugs in the experiment, the τ_{comp} was systematically underestimated for 10 compounds that belong to amino-quinazoline and amino pyrrolopyrimidine class of drugs. Additional four drugs were identified as outliers based on Crook's distance method and one drug was omitted due to its failure to retain crystallographic binding pose during equilibration runs. By excluding the outliers, 78% of the compounds (55 out of 70) showed a good linear correlation coefficient R^2 value of 0.86 between the experimentally measured and computationally predicted residence times, with 36% of mean absolute error (MAE) and 2.3 τ mean of prediction uncertainty, (MPU), on average (**103**). In 2019, Kokh et al. performed τRAMD simulations on another 25 N-HS90 with newly reported binding kinetics (**104),** combined them with their previous simulations, and applied different ML approaches to identify the molecular determinants of drug-target residence times (**105**). For 80 out of 94 compounds, they observed a linear correlation coefficient R^2 value of 0.75, with MAE of 0.39 ± 0.06, and MPU of 3.1 τ on average (**105**). Nunes-Alves et al. (**106**) studied the relative residence times (τ) of ligand dissociation from different cavities in T4L mutants across a spectrum of temperatures using τRAMD. They found a good linear correlation coefficient value of 0.78, with MAE 38% between computed residence time and experimental residence time (**106**).

3.4 Dissipation-corrected targeted MD (dcTMD)

In dcTMD, (**107, 108**) an external steering force is applied to a subset of atoms in a molecular simulation, guiding them along a predefined pathway or reaction coordinate. The method introduces a holonomic constraint force that steers the atoms from an initial to a final state at a

constant velocity (**107, 108**). In a ligand-protein system, the steering coordinate corresponds to the center of mass distance between the ligand and the binding pocket. The theory is based on two main assumptions. First, the Langevin equation can be applied to the unbiased motion of the system and provides a proper description of nonequilibrium simulations. Second, it uses cumulant expansion to derive friction coefficient and thus ensures rapid convergence of Jarzynski's identity. Using the Langevin equation, dcTMD introduces a T-boosting term that is distinct from targeted MD (TMD). The key advantage of T-boosting is its ability to calculate free energy directly at the target temperature by avoiding the need for rescaling from high to low temperatures as in the TMD method (**108**).

By using high-temperature Langevin simulation, dcTMD predicted k_{on} and k_{off} for benzamidine in benzamidine-trypsin system as $8.7 \times 10^{6} M^{-1} s^{-1}$ and $2.7 \times 10^{2} s^{-1}$, respectively (**109**). The finding underestimates the experimentally predicted values of $2.9 \times 10^{7} M^{-1} s^{-1}$ and $600 s^{-1}$ for k_{on} and k_{off} by a factor of ~2-3, respectively (**109**). Similarly, the k_{on} and k_{off} of Hsp90-inhibitor complex were calculated using 5ms long dcTMD simulation. The simulation predicted a k_{on} of $9.0 \times 10^{4} M^{-1} s^{-1}$ and k_{off} as $1.6 \times 10^{2} s^{-1}$. However, these simulated values significantly underestimate the experimentally determined rates, with a k_{on} of $4.8 \pm 0.2 \times 10^{5} M^{-1} s^{-1}$ and k_{off} of $3.4 \pm 0.2 \times 10^{-2} s^{-1}$ by factor of 5-20 (**109**). This discrepancy highlights the potential limitations of the dcTMD simulation in accurately capturing the kinetics of Hsp90-inhibitor interactions as compared to experimental observations.

3.5 Milestoning

Milestoning uses a set of slowly changing variables, such as torsion angles, radius of gyration, or distances between chemical groups, to map out the multi-dimensional landscapes that represent all

possible configurations and states a molecular system can adopt during a chemical reaction or transformation in MD (**110, 111**). This approach entails constructing a mesh with cell boundaries known as 'milestones', and aids in capturing significant transitional states (**110**). The milestoning method assumes that variables not included within the defined reaction space rapidly equilibrate, allowing for their simplified treatment and analysis using standard. This approach enables a focused study on key transitional dynamics without the computational complexity of accounting for all system variables in detail **(110)**. A detailed mesh design allows for effective sampling of transitions between closely situated milestones, making sampling of local transitions and low-energy regions amenable to MD simulation (**110, 111**).

Milestoning was implemented in the simulation enabled estimation of kinetic rates (SEEKR) (**112**) approach. It integrates milestoning theory, MD, and Brownian dynamics (BD) to predict kinetic rates and mechanisms of ligand binding (**112**). SEEKR employs computation-intensive MD to model transitions between milestones near the binding site, and more computationally efficient BD for sampling transitions between broadly spaced milestones farther from the binding site. This strategy allows SEEKR to leverage the comprehensive flexibility of MD where molecular flexibility is crucial, while utilizing the less demanding BD in regions where molecular flexibility is of lesser significance (**112**). As SEEKR requires accurate determination of the first hitting point distribution for initializing new trajectories at each milestone, it creates an issue of high simulation cost of calculation and an issue of parallelizability of calculation (**113**). To address these problems the Markovian Milestoning with Voronoi tessellation was combined with SEEKR. This method bypasses the requirement to calculate the equilibrium distribution across all the milestones. Instead, milestones are identified as the boundaries of a Voronoi tessellation, and the paths of the trajectories are kept within a Voronoi cell by applying a reflective boundary condition (**113**).

SEEKR2, an updated version of SEEKR was introduced to use OpenMM (apart from previously supported NAMD) for MD (**114**). SEEKR2 also provides the user with the option of using either the conventional milestoning method or the MMVT technique (**114**).

The SEEKR, MMVT-SEEKR, and SEEKR2 methodologies were applied to estimate the k_{on} and k_{off} rates of benzamidine binding to trypsin. Utilizing the SEEKR approach, the k_{on} rate for the benzamidine-trypsin system was determined to be $2.1 \pm 0.3 \times 10^7 M^{-1}s^{-1}$ showing a deviation of approximately 1.5 times from the experimentally calculated k_{on} of $2.9 \times 10^7 M^{-1}s^{-1}$ (**112)**. Conversely, the estimated k_{off} rate was not notably lower, yet within an order of magnitude compared to the experimental value, with SEEKR predicting a k_{off} of $83 \pm 14s^{-1}$ against the experimental value of $600 \pm 300s^{-1}$. In an advancement, SEEKR2 incorporating hydrogen mass repartitioning (HRM) yielded a k_{on} of $2.4 \pm 0.2 \times 10^7 M^{-1}s^{-1}$, aligning more closely with the experimental k_{on}. However, it predicted a k_{off} of $900 \pm 130\ s^{-1}$, surpassing the experimental value of $600 \pm 300s^{-1}$ (**114)**. This indicates a refinement in predicting the k_{on} rate, though the k_{off} estimation still showed variability.

Contrastingly, MMVT-SEEKR's predictions deviated significantly from experimental results. It estimated the k_{on} as $12 \pm 0.5 \times 10^7 M^{-1}s^{-1}$ and k_{off} as $174 \pm 9s^{-1}$, marking a deviation by factors of approximately 6 and 3.5, respectively, from the experimental rates (**113)**. This highlights a substantial disparity in the accuracy of MMVT-SEEKR's predictions when compared to both SEEKR and SEEKR2, indicating a need for further refinement in its application to accurately model kinetics of the benzamidine-trypsin interaction.

4 Expert Opinion

Both MM/PBSA and MM/GBSA have been successfully applied in structure-based drug design. They are established methods with appropriate balance between computational cost and prediction accuracy. MM/PBSA is preferred to achieve higher accuracy, while MM/GBSA is preferred for computational efficiency (less computational demand). However, their computational efficiency is attained through contentious approximations to the sampling and energy calculation phases. These simplistic approximations could involve using an evenly distributed dielectric constant for the entire solute surrounded by a complex local microenvironment, disregarding ions or important water molecules in the binding site, neglecting or using simplistic computational techniques for computing conformational and solvation entropies, etc. Certain improvements have been made over the years. The conventional practice is the use of normal mode analysis for approximating the conformational entropy. Zhong and collaborators published an improved interaction entropy method that is computationally effective. It can measure the entropic component of the binding free energy using MD simulation without incurring extra expenses (**45**). Moreover, more accurate force fields such as OPLS2.1 and 3.0 **(115)** could improve MM/PB(GB)SA calculation performance.

FEP and its variants are increasingly used to accurately predict the selectivity and potency of compounds, with their reliability nearing experimental standards. This accuracy is evidenced by both retrospective testing, which validates predictions against known outcomes, and prospective testing, where the methods are used to forecast the results of future experiments. These approaches are proving critical for advancing the precision of computational predictions to levels comparable with actual laboratory results. It is possible to address difficult-to-drug targets by successfully completing studies analyzing tens to hundreds of thousands of prospective drug candidates using

FEP-enabled methods. The efficiency and range of applications of FEP-enabled drug discovery will improve with continued advancements in computational and experimental methods. Improvement needed from the experimental end will include quality protein-ligand complex structures. High-resolution protein-ligand structures obtained from cryo-EM or X-ray will be a good starting point for FEP calculations. If experimental structures are not available, Alphafold2 and homology modeling tools can be employed. Additionally, the accuracy of FEP calculations will be increased while progressively lowering the processing cost of each calculation, provided that FEP methods are coupled with enhanced sampling techniques like REST, improved GPU technology, and molecular mechanics force fields (e.g., OPLS 3.0) **(65,66,115)**. Moreover, significant progress has been achieved in enhancing the robustness and stability of alchemical transformation pathways within established free energy calculation methods such as FEP and TI **(116, 117).** Concurrently, it is worth acknowledging the exceptional GPU performance of certain academic codes like AMBER, which are not only advancing computational efficiency but also becoming routinely utilized in the industry for drug discovery efforts **(117**). A fascinating prospect for the FEP-enabled drug design is the increase of chemical space to hundreds of thousands of molecules and beyond. This opens an opportunity for *de novo* drug discovery. In the de novo drug discovery process, accurately estimating free energy is critical for the development of highly targeted compounds. This approach is fundamental in driving the innovation of small molecules, which are re-emerging as a key focus in the search for new therapeutic agents. It is projected that advancements in simulation technology, force field development, and quantum chemistry will lead to the emergence of accurate quantifiable predictive models in these associated spheres. FEP-enabled drug discovery applications are currently at a pivotal historical crossroads, with the chance for widespread validation in the clinic in the near future.

For reliable estimation of free energy difference, sufficient overlap in the phase space between two states is preferred. In case there is limited phase space overlap, free energy methods could struggle to provide accurate predictions. This often occurs for systems undergoing large conformational changes or when comparing vastly different molecular species (**13**). In such scenarios, enhanced sampling techniques could be employed to improve phase space sampling and ensure better overlap, thereby increasing the accuracy of free energy calculations.

Enhanced sampling techniques can be generally categorized into CV-based and CV-free methods, and both provide their own benefits. By overcoming the free energy barrier and exploring various transitional states, the enhanced sampling methods have greatly facilitated ligand binding studies. With increasing accuracy in the prediction of ligand binding free energy and kinetics, enhanced sampling techniques are more widely used for drug discovery **(118)**. Microsecond enhanced sampling simulations have been demonstrated to capture both ligand dissociation and binding in various model systems. Infrequent metadynamics, LiGaMD, dcTMD and τ RAMD have been shown to be very efficient in these studies. Incorporation of machine learning and artificial intelligence with various sampling techniques could make computational approaches to drug discovery more powerful and accurate.

Acknowledgments

This work was supported in part by the National Institutes of Health (R01GM132572), and National Science Foundation (2121063) and the startup funding at University of North Carolina-Chapel Hill.

References:

1. DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics, 2016; 47, 20-33.

2. Kollman, P. A., Massova, I., Reyes, C., et al. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Accounts of Chemical Research, 2000; 33(12), 889-897.

3. Lipinski, C. A., Lombardo, F., Dominy, B. W. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews, 2001; 46(1-3), 3-26.

4. Lusci, A., Pollastri, G., & Baldi, P. Deep architectures and deep learning in chemoinformatics: The prediction of aqueous solubility for drug-like molecules. Journal of Chemical Information and Modeling, 2016; 56(2), 256-269.

5. Miao, Y., Bhattarai, A., & Wang, J. Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD): Characterization of Ligand Binding Thermodynamics and Kinetics. Journal of Chemical Theory and Computation, 2020;6(9), 5526-5547. https://doi.org/10.1021/acs.jctc.0c00395

   **Introduction of the Ligand Gaussian accelerated molecular dynamcis (LiGaMD) method for effective ligand kinetics calculation.

6. Wang, J., & Miao, Y. Ligand Gaussian Accelerated Molecular Dynamics 2 (LiGaMD2): Improved Calculations of Ligand Binding Thermodynamics and Kinetics with Closed Protein Pocket. Journal of Chemical Theory and Computation, 2023; 19(3), 733-745. https://doi.org/10.1021/acs.jctc.2c01194

   ** Introduction of the Ligand Gaussian accelerated molecular dynamics 2(LiGaMD2) method to improve ligand kinetics and binding thermodynamics calculations with closed protein pockets.

7. Bhattarai, A., Pawnikar, S., & Miao, Y. Mechanism of Ligand Recognition by Human ACE2 Receptor. The Journal of Physical Chemistry Letters, 2021; 12(20), 4814-4822. https://doi.org/10.1021/acs.jpclett.1c01064

8. Lenselink, E. B., Louvel, J., Forti, A. F., et al. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. ACS Omega, 2016;1(2), 293-304. https://doi.org/10.1021/acsomega.6b00086

9. Cournia Z, Allen B, Sherman W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. J Chem Inf Model. 2017 Dec 26;57(12):2911-2937. doi: 10.1021/acs.jcim.7b00564

10. Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opinion on Drug Discovery, 10(5), 449-461. https://doi.org/10.1517/17460441.2015.1032936

11. Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Ang Chem Int Ed 2002;41:2644-76

12. Armacost KA, Riniker S, Cournia Z. Novel directions in free energy methods and applications. J Chem Inf Model. 2020;60(1):1–5. doi:10.1021/acs.jcim.9b01174.

13. Cournia Z et al. Free energy methods in drug discovery—introduction. Free Energy Methods in Drug Discovery: Current State and Future Directions. 2021;pp. 1–38. doi:10.1021/bk-2021-1397.ch001.

14. Bowman GR, Beauchamp KA, Boxer G, Pande VS. Progress and challenges in the automated construction of Markov state models for full protein systems. J Chem Phys. 2009;131(12):124101. doi: 10.1063/1.3216567

15. York MD. Modern Alchemical Free Energy Methods for Drug Discovery Explained. ACS Physical Chemistry Au. 2023;0. doi: 10.1021/acsphyschemau.3c00033

16. Lee TS, Tsai HC, Ganguly A, York DM. ACES: Optimized Alchemically Enhanced Sampling. J Chem Theory Comput. 2023. doi: 10.1021/acs.jctc.2c00697

17. Swanson JM, Henchman RH, McCammon JA. Revisiting Free Energy Calculations: A Theoretical Connection to MM/PBSA and Direct Calculation of the Association Free Energy. Biophys J. 2004;86:67–74.

18. Genheden S, Ryde U. Comparison of End-Point Continuum Solvation Methods for the Calculation of Protein−Ligand Binding Free Energies. Proteins: Struct, Funct, Genet. 2012;80:1326–1342.

19. Sham YY, Chu ZT, Tao H, Warshel A. Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLDLRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease. Proteins: Struct, Funct, Genet. 2000;39:393–407.

20. Aqvist J, Medina C, Samuelsson JE. A New Method for Predicting Binding Affinity in Computer-Aided Drug Design. Protein Eng, Des Sel. 1994;7:385–391.

21. Adekoya OA, Willassen NP, Sylte I. Molecular Insight into Pseudolysin Inhibition Using the MM-PBSA and LIE Methods. J Struct Biol. 2006;153:129–144.

22. Hansson T, Marelius J, Aqvist J. Ligand Binding Affinity Prediction by Linear Interaction Energy Methods. J Comput-Aided Mol Des. 1998;12:27–35.

23. Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella Sampling. J Comput Phys. 1977;23(2):187–199. doi:10.1016/0021-9991(77)90121-8.

24. Darve E, Rodríguez-Gómez D, Pohorille A. Adaptive biasing force method for scalar and vector free energy calculations. J Chem Phys. 2008;128(14). doi:10.1063/1.2829861.

25. Laio A, Parrinello M. Escaping free-energy minima. Proc Natl Acad Sci. 2002;99(20):12562–12566. doi:10.1073/pnas.202427399.

26. Grubmüller H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. Phys Rev E. 1995 Sep;52(3):2893-2906. doi: 10.1103/PhysRevE.52.2893.

27. Valsson O, Parrinello M. Variational approach to enhanced sampling and free energy calculations. Phys Rev Lett. 2014;113(9). doi:10.1103/physrevlett.113.090601.

28. Kokh DB, et al. Estimation of drug-target residence times by τ-random acceleration molecular dynamics simulations. J Chem Theory Comput. 2018;14(7):3859-3869. doi:10.1021/acs.jctc.8b00230.

29. Wolf S, Stock G. Targeted molecular dynamics calculations of free energy profiles using a nonequilibrium friction correction. J Chem Theory Comput. 2018;14(12):6175-6182. doi:10.1021/acs.jctc.8b00835.

30. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E. gmx_MMPBSA: A new tool to perform end-state free energy calculations with GROMACS. J Chem Theory Comput. 2021 Oct 12;17(10):6281-6291. doi: 10.1021/acs.jctc.1c00645.

31. Kollman PA, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33:889-97

32. Wang S, Sun X, Cui W, Yuan S. MM/PB(GB)SA benchmarks on soluble proteins and membrane proteins. Front Pharmacol. 2022 Dec 1;13:1018351. doi: 10.3389/fphar.2022.1018351.

33. Tuccinardi T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery? Expert Opin Drug Discov. 2021 Nov;16(11):1233-1237. doi: 10.1080/17460441.2021.1942836.

34. Virtanen SI, Niinivehmas SP, Pentikäinen OT. Case-specific performance of MM-PBSA, MM-GBSA, and SIE in virtual screening. J Mol Graph Model. 2015 Nov;62:303-318. doi: 10.1016/j.jmgm.2015.10.012.

35. Weis A, Katebzadeh K, Soderhjelm P, Nilsson I, Ryde U. Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field. J Med Chem. 2006;49:6596−6606.

36. Sahakyan H. Improving virtual screening results with MM/GBSA and MM/PBSA rescoring. J Comput Aided Mol Des. 2021 Jun;35(6):731-736. doi: 10.1007/s10822-021-00389-3.

37. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov. 2015 May;10(5):449-61. doi: 10.1517/17460441.2015.1032936.

38. Crean RM, Pudney CR, Cole DK, van der Kamp MW. Reliable in silico ranking of engineered therapeutic TCR binding affinities with MMPB/GBSA. J Chem Inf Model. 2022 Feb 14;62(3):577-590. doi: 10.1021/acs.jcim.1c00765.

39. Zhang X, Wong SE, Lightstone FC. Towards fully automated high-performance computing drug discovery: A massively parallel virtual screening pipeline for docking and MM/GBSA rescoring to improve enrichment. J Chem Inf Model. 2014;54:324−337.

40. Pan Y, Lu Z, Li C, Qi R, Chang H, Han L, Han W. Molecular dockings and molecular dynamics simulations reveal the potency of different inhibitors against xanthine oxidase. ACS Omega. 2021;6(17):11639-11649. doi: 10.1021/acsomega.1c00968.

41. Maffucci I, Hu X, Fumagalli V, Contini A. An Efficient Implementation of the Nwat-MMGBSA Method to Rescore Docking Results in Medium-Throughput Virtual Screenings. Front Chem. 2018;6. DOI: 10.3389/fchem.2018.00043

42. Wang Z, Pan H, Sun H, Kang Y, Liu H, Cao D, Hou T. fastDRH: a webserver to predict and analyze protein-ligand complexes based on molecular docking and MM/PB(GB)SA computation. Brief Bioinform. 2022 Sep 20;23(5):bbac201. doi: 10.1093/bib/bbac201.

43. Wang JM, Hou TJ, Xu XJ. Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. Curr Comput-Aided Drug Des. 2006;2:287−306.

44. Rastelli G, Del Rio A, Degliesposti G, Sgobba M. Fast and Accurate Predictions of Binding Free Energies Using MM-PBSA and MM-GBSA. J Comput Chem. 2009;31:797−810.

45. Zhong S, Huang K, Luo S, Dong S, Duan L. Improving the performance of the MM/PBSA and MM/GBSA methods in recognizing the native structure of the Bcl-2 family using the interaction entropy method. Phys Chem Chem Phys. 2020;22(7):4240-4251. doi: 10.1039/c9cp06459a

46. Bello M, Bandala C. Evaluating the ability of end-point methods to predict the binding affinity tendency of protein kinase inhibitors. RSC Adv. 2023 Aug 22;13(36):25118-25128. doi: 10.1039/d3ra04916g.

47. Homeyer N, Gohlke H. Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. Mol Inform. 2012 Feb;31(2):114-22. doi: 10.1002/minf.201100135.

48. Nurisso A, Blanchard B, Audfray A, Rydner L, Oscarson S, Varrot A, Imberty A. Role of water molecules in structure and energetics of Pseudomonas aeruginosa lectin I interacting with disaccharides. J Biol Chem. 2010 Jun 25;285(26):20316-27. doi: 10.1074/jbc.M110.108340

49. Foloppe N, Hubbard R. Towards Predictive Ligand Design With Free-Energy Based Computational Methods? Curr Med Chem. 2006;13(29):3583-3608.

50. Wang J, Hou T, Xu X. Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. Curr Comput-Aided Drug Des. 2006;2(3):287-306.

51. Homeyer N, Gohlke H. Free Energy Calculations by the Molecular Mechanics Poisson Boltzmann Surface Area Method. Mol Inform. 2012;31(2):114-122.

52. Ekberg V, Ryde U. On the Use of Interaction Entropy and Related Methods to Estimate Binding Entropies. J Chem Theory Comput. 2021 Aug 10;17(8):5379-5391. doi: 10.1021/acs.jctc.1c00374.

53. Zwanzig RW. High-Temperature Equation of State by a Perturbation Method. 1. Nonpolar Gases. J Chem Phys. 1954;22(8):1420-1426.

54. Kirkwood JG. Statistical Mechanics of Fluid Mixtures. J Chem Phys. 1935;3:300–313.

55. Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. J Comput Phys, 1976; 22(2), 245-268.

56. Shirts MR. Reweighting from the Mixture Distribution as a Better Way to Describe the Multistate Bennett Acceptance Ratio. arXiv preprint arXiv:1704.00891. 2017.

57. Matsunaga Y, Kamiya M, Oshima H, Jung J, Ito S, Sugita Y. Use of multistate Bennett acceptance ratio method for free-energy calculations from enhanced sampling and free-energy perturbation. Biophys Rev. 2022 Dec 14;14(6):1503-1512. doi: 10.1007/s12551-022-01030-9

58. Matricon P, Vo DD, Gao ZG, Kihlberg J, Jacobson KA, Carlsson J. Fragment-based design of selective GPCR ligands guided by free energy simulations. Chem Commun (Camb). 2021 Nov 19;57(92):12305-12308. doi: 10.1039/d1cc03202j.

59. Deflorian F, Perez-Benito L, Lenselink EB, Congreve M, van Vlijmen HWT, Mason JS, de Graaf C, Tresadern G. Accurate Prediction of GPCR Ligand Binding Affinity with Free Energy Perturbation. J Chem Inf Model. 2020;60(11):5563-5579.

60. Jespers W, Verdon G, Azuaje J, Majellaro M, Keränen H, García-Mera X, Congreve M, Deflorian F, de Graaf C, Zhukov A, Doré AS, Mason JS, Åqvist J, Cooke RM, Sotelo E, Gutiérrez-de-Terán H. X-Ray Crystallography and Free Energy Calculations Reveal the Binding Mechanism of A2A Adenosine Receptor Antagonists. Angew Chem Int Ed Engl. 2020 Sep 14;59(38):16536-16543. doi: 10.1002/anie.202003788.

61. Panel N, Vo DD, Kahlous NA, Hübner H, Tiedt S, Matricon P, Pacalon J, Fleetwood O, Kampen S, Luttens A, Delemotte L, Kihlberg J, Gmeiner P, Carlsson J. Design of Drug Efficacy Guided by Free Energy Simulations of the β2-Adrenoceptor. Angew Chem Int Ed Engl. 2023 May 22;62(22):e202218959.

62. Sergeeva AP, Katsamba PS, Liao J, Sampson JM, Bahna F, Mannepalli S, Morano NC, Shapiro L, Friesner RA, Honig B. Free Energy Perturbation Calculations of Mutation Effects on SARS-CoV-2 RBD:ACE2 Binding Affinity. J Mol Biol. 2023 Aug 1;435(15):168187.

63. Ibrahim MT, Verkhivker GM, Misra J, Tao P. Novel Allosteric Effectors Targeting Human Transcription Factor TEAD. Int J Mol Sci. 2023 May 19;24(10):9009. doi: 10.3390/ijms24109009.

64. Havranek B, Demissie R, Lee H, Lan S, Zhang H, Sarafianos S, Ayitou AJ, Islam SM. Discovery of Nirmatrelvir Resistance Mutations in SARS-CoV-2 3CLpro: A Computational-Experimental Approach. J Chem Inf Model. 2023 Nov 10. doi: 10.1021/acs.jcim.3c01269. Epub ahead of print.

65. Wang L, Wu YJ, Deng Y, et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. J Am Chem Soc. 2015;137(7):2695-2703.

66. Horton JT, Boothroyd S, Wagner J, Mitchell JA, Gokey T, Dotson DL, Behara PK, Ramaswamy VK, Mackey M, Chodera JD, Anwar J, Mobley DL, Cole DJ. Open Force Field BespokeFit: Automating Bespoke Torsion Parametrization at Scale. J Chem Inf Model. 2022;62(22):5622-5633. doi: 10.1021/acs.jcim.2c01153.

67. Liu P, Kim B, Friesner RA, Berne BJ. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc Natl Acad Sci U.S.A. 2005;102(39):13749-54. doi: 10.1073/pnas.0506346102.

68. Wang L, Friesner RA, Berne BJ. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). J Phys Chem B. 2011;115(30):9431-8. doi: 10.1021/jp204407d.

69. Wang L, Berne BJ, Friesner RA. On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities. Proc Natl Acad Sci U.S.A. 2012;109(6):1937-42. doi: 10.1073/pnas.1114017109.

70. Lenselink EB, Louvel J, Forti AF, van Veldhoven JPD, de Vries H, Mulder-Krieger T, McRobb FM, Negri A, Goose J, Abel R, van Vlijmen HWT, Wang L, Harder E, Sherman W, IJzerman AP, Beuming T. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. ACS Omega. 2016;1(2):293-304. doi: 10.1021/acsomega.6b00086.

71. Sun S, Fushimi M, Rossetti T, Kaur N, Ferreira J, Miller M, Quast J, van den Heuvel J, Steegborn C, Levin LR, Buck J, Myers RW, Kargman S, Liverton N, Meinke PT, Huggins DJ. Scaffold Hopping and Optimization of Small Molecule Soluble Adenyl Cyclase Inhibitors Led by Free Energy Perturbation. J Chem Inf Model. 2023 May 8;63(9):2828-2841.

72. Clark AJ, et al. Free energy perturbation calculation of relative binding free energy between broadly neutralizing antibodies and the gp120 glycoprotein of HIV-1. J Mol Biol. 2017;429:930–947.

73. Clark AJ, et al. Relative binding affinity prediction of charge-changing sequence mutations with FEP in protein–protein interfaces. J Mol Biol. 2019;431:1481–1493.

74. Yamashita T. Toward rational antibody design: Recent advancements in molecular dynamics simulations. Int Immunol. 2018;30:133–140.

75. Zhu F, Bourguet FA, Bennett WFD, Lau EY, Arrildt KT, Segelke BW, Zemla AT, Desautels TA, Faissol DM. Large-scale application of free energy perturbation calculations for antibody design. Sci Rep. 2022 Jul 21;12(1):12489.

76. Sormanni P, Aprile FA, Vendruscolo M. Third generation antibody discovery methods: In silico rational design. Chem Soc Rev. 2018;47:9137–9157.

77. Chen W, Cui D, Jerome SV, Michino M, Lenselink EB, Huggins DJ, Beautrait A, Vendome J, Abel R, Friesner RA, Wang L. Enhancing Hit Discovery in Virtual Screening through Absolute Protein-Ligand Binding Free-Energy Calculations. J Chem Inf Model. 2023 May 22;63(10):3171-3185.

78. Ghahremanpour MM, Saar A, Tirado-Rives J, Jorgensen WL. Computation of Absolute Binding Free Energies for Noncovalent Inhibitors with SARS-CoV-2 Main Protease. J Chem Inf Model. 2023 Aug 28;63(16):5309-5318. doi: 10.1021/acs.jcim.3c00874.

79. Lockhart C, Luo X, Olson A, Delfing BM, Laracuente XE, Foreman KW, Paige M, Kehn-Hall K, Klimov DK. Can Free Energy Perturbation Simulations Coupled with Replica-Exchange Molecular Dynamics Study Ligands with Distributed Binding Sites? J Chem Inf Model. 2023 Aug 14;63(15):4791-4802.

80. Jorgensen WL. Efficient drug lead discovery and optimization. Acc Chem Res. 2009;42(6):724-733.

81. Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical free energy methods for drug discovery: progress and challenges. Curr Opin Struct Biol. 2011 Apr;21(2):150-60. doi: 10.1016/j.sbi.2011.01.011.

82. Procacci P. Methodological uncertainties in drug-receptor binding free energy predictions based on classical molecular dynamics. Curr Opin Struct Biol. 2021 Apr;67:127-134. doi: 10.1016/j.sbi.2020.08.001.

83. Fu H, Chipot C, Shao X, Cai W. Standard Binding Free-Energy Calculations: How Far Are We from Automation? J Phys Chem B. 2023 Dec 14;127(49):10459-10468. doi: 10.1021/acs.jpcb.3c04370.

84. Shih, A. Y.; Hack, M.; Mirzadegan, T. Impact of Protein Preparation on Resulting Accuracy of FEP Calculations. J. Chem. Inf. Model. 2020, 60, 5287−5289

85. Muegge I, Hu Y. Recent Advances in Alchemical Binding Free Energy Calculations for Drug Discovery. ACS Med Chem Lett. 2023 Feb 16;14(3):244-250. doi: 10.1021/acsmedchemlett.2c00541.

86. Zhang H, Jiang W, Chatterjee P, Luo Y. Ranking Reversible Covalent Drugs: from Free Energy Perturbation to Fragment Docking. J Chem Inf Model. 2019;59:2093–2102.

87. Constantine, K. L., Mueller, L., Metzler, W. J., et al. Multiple and single binding modes of fragment-like kinase inhibitors revealed by molecular modeling, residue type-selective protonation, and nuclear overhauser effects. Journal of Medicinal Chemistry, 2008; 51(19), 6225-6229.

88. Shirts MR, Chodera JD. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. J Chem Phys. 2008;129:124105.

89. Zou J, Tian C, Simmerling C. Blinded prediction of protein–ligand binding affinity using amber thermodynamic integration for the 2018 D3R Grand Challenge 4. J Comput Aided Mol Des. 2019;33(12):1021–1029. doi:10.1007/s10822-019-00223-x.

90. He X, et al. Fast, accurate, and reliable protocols for routine calculations of protein–ligand binding affinities in drug design projects using Amber GPU-ti with FF14SB/GAFF. ACS Omega. 2020;5(9):4611–4619. doi:10.1021/acsomega.9b04233.

91. Sheng Z, et al. An optimized thermodynamics integration protocol for identifying beneficial mutations in antibody design. Front Immunol. 2023;14. doi:10.3389/fimmu.2023.1190416.

92. Bhati AP, et al. Rapid, accurate, precise, and reliable relative free energy prediction using ensemble based thermodynamic integration. J Chem Theory Comput. 2016;13(1):210–222. doi:10.1021/acs.jctc.6b00979.

93. Tiwary P, Parrinello M. From metadynamics to dynamics. Phys Rev Lett. 2013;111(23). doi:10.1103/physrevlett.111.230602

94. Pramanik D, et al. CAN one trust kinetic and thermodynamic observables from biased metadynamics simulations? Detailed quantitative benchmarks on millimolar drug fragment dissociation. J Phys Chem B. 2019;123(17):3672–3678. doi: 10.1021/acs.jpcb.9b01813.

95. Wang Y, Martins JM, Lindorff-Larsen K. Biomolecular conformational changes and ligand binding: from kinetics to thermodynamics. Chem Sci. 2017;8(9):6466–6473. https://doi.org/10.1039/c7sc01627a

96. Lamim Ribeiro JM, 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. The Journal of Chemical Physics. 2020;153(12). doi:10.1063/5.0019100.

97. Mahinthichaichan P, et al. Kinetics and mechanism of fentanyl dissociation from the μ-opioid receptor. JACS Au. 2021;1(12):2208–2215. doi:10.1021/jacsau.1c00341.

98. Ravindra P, Smith Z, Tiwary P. Automatic mutual information noise omission (AMINO): generating order parameters for molecular systems. Molecular Systems Design & Engineering. 2020;5(1):339–348. https://doi.org/10.1039/c9me00115h.

99. Ribeiro JML, Bravo P, Wang Y, Tiwary P. Reweighted autoencoded variational Bayes for enhanced sampling (RAVE). The Journal of Chemical Physics. 2018;149(7):072301. https://doi.org/10.1063/1.5025487.

100. Pedersen MF, Wróbel TM, Märcher-Rørsted E, Pedersen DS, Møller TC, Gabriele F, Pedersen H, Matosiuk D, Foster SR, Bouvier M, Bräuner-Osborne H. Biased agonism of clinically approved μ-opioid receptor agonists and TRV130 is not controlled by binding and signaling kinetics. Neuropharmacology. 2020;166:107718. https://doi.org/10.1016/j.neuropharm.2019.107718.

101. Miao Y, Feher VA, McCammon JA. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. Journal of Chemical Theory and Computation. 2015;11(8):3584–3595. https://doi.org/10.1021/acs.jctc.5b00436.

102. Florent Guillain, Thusius D. Use of proflavine as an indicator in temperature-jump studies of the binding of a competitive inhibitor to trypsin. Journal of the American Chemical Society. 1970;92(18):5534–5536. https://doi.org/10.1021/ja00721a051.

103. Kokh DB, Amaral M, Bomke J, Grädler U, Musil D, Buchstaller HP, Dreyer MK, Frech M, Lowinski M, Vallee F, Bianciotto M, Rak A, Wade RC. Estimation of Drug-Target Residence Times by τ-Random Acceleration Molecular Dynamics Simulations. J Chem Theory Comput. 2018;14(7):3859-3869. doi: 10.1021/acs.jctc.8b00230.

104. Schuetz DA, Richter L, Amaral M, Grandits M, Grädler U, Müsil D, Buchstaller HP, Eggenweiler HM, Frech M, Ecker GF. Ligand Desolvation Steers On-Rate and Impacts Drug Residence Time of Heat Shock Protein 90 (Hsp90) Inhibitors. Journal of Medicinal Chemistry. 2018;61(10):4397–4411. https://doi.org/10.1021/acs.jmedchem.8b00080.

105. Kokh DB, Kaufmann T, Kister B, Wade RC. Machine Learning Analysis of τRAMD Trajectories to Decipher Molecular Determinants of Drug-Target Residence Times. Frontiers in Molecular Biosciences. 2019;6. https://doi.org/10.3389/fmolb.2019.00036.

106. Nunes-Alves A, Kokh DB, Wade RC. Ligand unbinding mechanisms and kinetics for T4 lysozyme mutants from τRAMD simulations. Current Research in Structural Biology. 2021;3:106–111. https://doi.org/10.1016/j.crstbi.2021.04.001

107. Wolf, S. & Stock, G. Targeted molecular dynamics calculations of free energy profiles using a nonequilibrium friction correction. J. Chem. Theory Comput. 14, 6175—6182 (2018).

108. Schlitter J, Engels M, Krüger P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. Journal of Molecular Graphics. 1994;12(2):84–89. https://doi.org/10.1016/0263-7855(94)80072-3

109. Wolf S, Lickert B, Bray S, Stock G. Multisecond ligand dissociation dynamics from atomistic simulations. Nature Communications. 2020;11(1). https://doi.org/10.1038/s41467-020-16655-1

110. Bello-Rivas JM, Elber R. Exact milestoning. The Journal of Chemical Physics. 2015;142(9). https://doi.org/10.1063/1.4913399

111. Elber R. Milestoning: An Efficient Approach for Atomically Detailed Simulations of Kinetics in Biophysics. Annual Review of Biophysics. 2020;49(1):69–85. https://doi.org/10.1146/annurev-biophys-121219-081528

112. Votapka LW, Jagger BR, Heyneman AL, Amaro RE. SEEKR: Simulation Enabled Estimation of Kinetic Rates, A Computational Tool to Estimate Molecular Kinetics and Its Application to Trypsin–Benzamidine Binding. The Journal of Physical Chemistry B. 2017;121(15):3597–3606. https://doi.org/10.1021/acs.jpcb.6b09388

113. Jagger BR, Ojha AA, Amaro RE. Predicting Ligand Binding Kinetics Using a Markovian Milestoning with Voronoi Tessellations Multiscale Approach. Journal of Chemical Theory and Computation. 2020;16(8):5348–5357. https://doi.org/10.1021/acs.jctc.0c00495

114. Votapka LW, Stokely AM, Ojha AA, Amaro RE. SEEKR2: Versatile Multiscale Milestoning Utilizing the OpenMM Molecular Dynamics Engine. Journal of Chemical Information and Modeling. 2022;62(13):3253–3262. https://doi.org/10.1021/acs.jcim.2c00501

115. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, Hou T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. Chemical Reviews. 2019;119(16):9478-9508. doi: 10.1021/acs.chemrev

116. Tsai HC, Lee TS, Ganguly A, Giese TJ, Ebert MC, Labute P, Merz KM Jr, York DM. AMBER Free Energy Tools: A New Framework for the Design of Optimized Alchemical Transformation Pathways. J Chem Theory Comput. 2023 Jan 9:10.1021/acs.jctc.2c00725. doi: 10.1021/acs.jctc.2c00725

117. Lee TS, Allen BK, Giese TJ, Guo Z, Li P, Lin C, McGee TD Jr, Pearlman DA, Radak BK, Tao Y, Tsai HC, Xu H, Sherman W, York DM. Alchemical Binding Free Energy

Calculations in AMBER20: Advances and Best Practices for Drug Discovery. J Chem Inf Model. 2020 Nov 23;60(11):5595-5623. doi: 10.1021/acs.jcim.0c00613

118. Wang J, Do HN, Koirala K, Miao Y. Predicting biomolecular binding kinetics: A Review. Journal of Chemical Theory and Computation. 2023;19(8):2135–2148. https://doi.org/10.1021/acs.jctc.2c01085

** Presents detailed review on techniques used in biomolecular binding kinetics.

Figure 1. The illustrative representation of interaction between protein (P) and receptor (R) and schematic diagram for dissociation rate constant(k_{off}) and association rate constant(k_{on}). K_a represents equilibrium association constant.