UNIVERSITY OF CALIFORNIA,
IRVINE

Enhanced sampling of ligand binding modes through BLUES and Molecular Darting

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Samuel Charles Gill

Dissertation Committee:
Professor David Mobley, Chair
Professor Ioan Andricioaei
Professor Douglas Tobias

2020

DEDICATION

I would like to thank my parents for their unwavering support for everything that I do in my life–including my pursuit of a doctorate in chemistry.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGMENTS

I want to acknowledge the support of the many people who have contributed to my scientific growth and journey. Firstly, I would like to thank David Mobley for his continual support and scientific insight throughout my graduate career, as well as being a wonderful advisor and person. I would also like to thank the members of my lab who contributed to a positive, happy, and rewarding working environment during my time: Shaui Liu, Pavel Klimovich, Nathan Lim, Caitlin Bannan, Guilhereme D.R. Matos, Hanh Ngyuen, Victoria Lim, Kalistyn Burley, David Wych, Jessica Maat, Danielle Bergazin, Chris Zhang, Ohan Tran, Hannah Baumann, Trevor Gokey, Léa el Khoury, Sukanya Sasmal, and Gaetano Calabrò, thank you for making the day-to-day experience a great one. And I would especially like to thank Nathan Lim, who I worked with extensively with on BLUES.

VITA

Samuel Charles Gill

EDUCATION

Ph.D. in Chemistry — **April 2020**
University of California, Irvine — Adviser: David L. Mobley

B.S. in Chemistry — **2014**
Occidental College

RESEARCH EXPERIENCE

Graduate Student Researcher — **Jan. 2015– April 2020**
University of California, Irvine — *Irvine, CA*

Computational Chemistry Co-op — **June 2018– August 2018**
GlaxoSmithKline — *Collegeville, PA*

TEACHING EXPERIENCE

Teaching Assistant — **2014–2016**
University of California, Irvine — *Irvine, CA*

Software

BLUES `https://github.com/MobleyLab/blues`
NCMC/MC + MD approach to sample ligand binding modes

ABSTRACT OF THE DISSERTATION

Enhanced sampling of ligand binding modes through BLUES and Molecular Darting

By

Samuel Charles Gill

Doctor of Philosophy in Chemistry

University of California, Irvine, 2020

Professor David Mobley, Chair

Free energy perturbation methods serve an important role in drug discovery by providing accurate predictions of binding affinity, solubility, and other quantities. However, in order for the free energy estimates to be accurate, the system must be able to sample all the relevant low energy states during the course of a simulation. This proves to be challenging for binding affinity calculations in particular, since there can be many different potential binding modes, and binding modes are slow to interconvert at simulation timescales. It is possible to treat each binding mode separately, perform a free energy calculation on each binding mode, and then combine the results into a total free energy prediction, but the computational cost of free energy calculations does make this parallelization approach feasible. Another alternative approach would be to use a method that could sample between the different binding modes efficiently and produce accurate estimates of the populations of each of those binding modes. This information, along with a binding free energy calculation on one of the binding modes, would allow the estimation of the overall binding free energy.

In order to sample between binding modes efficiently, I helped develop a new method that uses nonequilibrium candidate Monte Carlo (NCMC) to remove the ligand and reinsert it in a new binding site with a center of mass rotation to further improve sampling. I validated this methodology on T4 lysozyme L99A, a model protein for binding, and was able to show that

it enhanced the sampling of the binding modes of toluene and 3-iodotolune. Following these results, I helped create the BLUES (Binding modes of Ligands Using Enhanced Sampling) software package to facilitate the use of this technique.

Originally BLUES only could further improve binding mode sampling with a center of mass rotation Monte Carlo (MC) move, which limited it's applicability to small, rigid ligands. To further improve binding mode sampling, I also developed a new type of MC move called molecular Darting (MolDarting) to sample specific binding modes. Through MolDarting it is possible to sample predefined conformations–obtained by docking, for example–and reversibly sample them in a MC framework. MolDarting also opens up the ability to even sample ligand binding modes in separate binding sites. We validated this move on an alanine-valine dipeptide system, as well the previouly explored T4 lysozyme and attempted to sample all the potential binding sites in HIV integrase.

Chapter 1

Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo

Accurately predicting protein-ligand binding affinities and binding modes is a major goal in computational chemistry, but even the prediction of ligand binding modes in proteins poses major challenges. Here, we focus on solving the binding mode prediction problem for rigid fragments. That is, we focus on computing the dominant placement, conformation, and orientations of a relatively rigid, fragment-like ligand in a receptor, and the populations of the multiple binding modes which may be relevant. This problem is important in its own right, but is even more timely given the recent success of alchemical free energy calculations [127, 4]. Alchemical calculations are increasingly used to predict binding free energies of ligands to

receptors. However, the accuracy of these calculations is dependent on proper sampling of the relevant ligand binding modes. Unfortunately, ligand binding modes may often be uncertain, hard to predict, and/or slow to interconvert on simulation timescales, so proper sampling with current techniques can require prohibitively long simulations. We need new methods which dramatically improve sampling of ligand binding modes. Here, we develop and apply a nonequilibrium candidate Monte Carlo (NCMC) method to improve sampling of ligand binding modes. In this technique, the ligand is rotated and subsequently allowed to relax in its new position through alchemical perturbation before accepting or rejecting the rotation and relaxation as a nonequilibrium Monte Carlo move. When applied to a T4 lysozyme model binding system, this NCMC method shows over two orders of magnitude improvement in binding mode sampling efficiency compared to a brute force molecular dynamics simulation. This is a first step towards applying this methodology to pharmaceutically-relevant binding of fragments and, eventually, drug-like molecules. We are making this approach available via our new Binding Modes of Ligands using Enhanced Sampling (BLUES) package which is freely available on GitHub.

1.1 Introduction

1.1.1 Ligand binding modes are important, but difficult to predict

One of the motivations of computer aided drug design is to understand and predict what factors increase ligand binding affinity to allow for better design of new ligands for further drug development. Successfully predicting binding affinity depends on many factors, including the determination of the most favorable or relevant binding mode, or modes, of the ligand. Medicinal chemists often use knowledge of the likely binding mode or binding modes to attempt rational improvements upon the scaffold, as well as giving insight into the

important interactions driving binding. The binding mode or binding modes also provide a fundamental input for many calculations that can predict binding affinities, such as free energy calculations [83].

As important as binding modes are, actually determining them can be difficult. The standard experimental techniques for binding mode determination, X-ray crystallography and nuclear magnetic resonance, can be time-consuming, difficult, or costly, and are not suitable for all targets (membrane proteins can be particularly challenging, for instance). Additionally, experimental methods do not always clearly resolve the binding mode. For example, in the case of fragment-based drug discovery, small, relatively rigid ligands can often have some ambiguity in their binding modes because of internal pseudosymmetry, or other issues [94, 89]. Additionally, methods to make X-ray diffraction data easier to collect—such as cryocooling crystals—potentially stabilize binding modes that are not observed under the conditions of interest [37]. Multiple binding modes may also contribute substantially to a ligand's affinity [82, 11, 80, 125, 57], therefore knowledge of a single experimental binding mode may be misleading or provide an incomplete picture.

Computationally determining binding modes is similarly difficult. One of the most widely used computational methods for binding mode determination is docking, which scores a variety of ligand poses in the binding site [118, 62]. Docking has been shown to perform well in generating candidate binding poses from the top scoring poses; however, the top scoring poses from docking tend not to be the ones found experimentally [130, 84]. This is partly because docking was designed to have a particularly low computational cost (usually seconds or less per molecule) in order to allow application to large databases [25, 103]. Thus, docking necessarily makes many approximations in order to achieve this speed.

In a recent D3R challenge, which consisted of predicting binding modes of HSP90 ligands, different docking studies had varying levels of accuracy–even within submissions using the same docking software–but human screening of the structures seemed to help identify the

correct binding mode [41]. Four of the 11 top scoring methods used visual inspection of the computationally predicted poses, while the less successful methods did not, indicating how it remains extremely challenging to predict binding modes [41]. Another study by Warren *et al.* looked at how well different docking programs performed across a variety of different protein targets [130]. They found that docking methods could explore the conformational space of the ligand sufficiently, but the top scoring pose often did not correspond to the observed crystallographic pose [130].

In fact, humans tend to outperform automated methods at predicting binding modes in blind challenges [112, 120], further showing that there are still many aspects of binding mode prediction that could benefit from improvement. In the SAMPL4 HIV integrase challenge, for example, determining the binding mode or even binding site of a set of ligands to HIV integrase was shown to be very difficult for many methods [84]. A human expert with more than 10 years working on the target provided the best submission, in large part guided by his expertise [120]. The best purely computational method in this challenge used docking followed by free energy calculations to predict whether compounds would bind to HIV integrase. In this study, the majority of false negative binding results used starting poses vastly dissimilar from crystallographic observations [39], indicating that many of the failures resulted from incorrect binding modes.

An alternative to docking which is more rigorous, but computationally expensive, is to apply free energy calculations based on molecular simulations to predict populations of possible stable binding modes. For example, the "confine and release" approach allows multiple binding modes to be treated separately by distinct free energy calculations, and then subsequently combining the individual binding free energies to yield a total binding free energy [79] (Figure 1.1(a)). Specifically, the overall binding free energy of a ligand to a protein can be decomposed into a particular type of average over the effective binding free energies of different metastable binding modes [78]. As long as these metastable binding modes are defined a

way that they cover the full bound state including all the relevant binding modes, and they do not overlap in phase space, this approach is rigorous. However, the number of required binding free energy calculations scales linearly with the number of binding modes for this already computationally demanding approach, making it unappealing to consider multiple candidate binding modes separately in this manner [83].

1.1.2 Other approaches exist to determine binding modes

Another option is to sample over the binding modes within a given binding free calculation, as reviewed elsewhere [83]. Many binding free energy calculations use *alchemical* techniques [119, 44, 60] where binding free energies are computed by turning off interactions between the ligand and receptor (controlled by an alchemical parameter λ), taking the ligand through a nonphysical pathway that allows it to be moved from the binding site to solution, yielding the binding free energy. The Binding Energy Distribution Analysis Method [40] (BEDAM) is one such alchemical method which includes multiple binding modes by allowing the non-interacting or weakly-ligand to rearrange and reorient in the binding site before turning back on interactions, thus allowing relatively easy interchange between binding modes in a single set of simulations.

A similar approach is taken by Wang *et al.* in the application of Hamiltonian replica exchange molecular dynamics to ligand binding [125]. In their work, multiple replicas of a protein-ligand system were simulated in implicit solvent with varying λ couplings of the steric and electrostatics between replicas. To enhance conformational sampling, translational and rotational Monte Carlo moves were applied before exchange attempts. This potentially allows efficient sampling across binding modes in a single binding free energy calculation, though the use of implicit solvent was critical to the success of their instantaneous Monte Carlo moves. The POPFEP work of Jayachandran *et al.* proposed an alternative approach, cor-

recting for poorly mixing sampling that resulted in highly erroneous binding pose populations by decomposing the sampled configurations into distinct poses with a Markov state model and independently computing alchemical binding free energies with respect to a common noninteracting state [53].

1.1.3 Efficiently sampling binding modes in a simulation would greatly increase free energy calculation performance

Our goal in this paper is a computational method which can reproduce equilibrium binding mode populations with much less computational time than treating binding modes separately. Specifically, each alchemical binding free energy calculation requires simulation at N different alchemical intermediate states (where N is typically at least 12-20 [78, 40, 125]), where each alchemical state is associated with a vector of alchemical parameters λ (which we will refer to as λ values). If we consider M different binding modes, the total cost of a binding free energy calculation that covers all binding modes separately is MNx where x is the cost of a single simulation (Figure 1.1(A) and (C)). This becomes impractical as the number of potential binding modes grows. Instead, the approach we envision is one where we calculate an absolute binding free energy for a single, reasonably populated binding mode and, in a separate calculation that can be run concurrently, efficiently determine the relative free energies (or equilibrium populations) of all M potential binding modes (Figure 1.1(B)). Then, we can combine the populations of the individual binding modes and the free energy estimate for a single binding mode into a binding free energy that includes all of the possible binding modes [83]. Thus this approach would have a computational cost of $Nx + y$, where y is the simulation time to determine the binding mode populations.

Such an approach could work as outlined in Ref. [83], providing a way to compute interconversion free energies between different metastable binding modes. This would have implications

for both absolute and relative binding free energy calculations. For absolute calculations, depending on how many binding modes are being considered, such an approach could drastically reduce the total amount of simulation time, as long as $y \ll Nx$ (see Figure 1.1(C)) (and the wallclock time as long as $y \ll x$ since the N independent alchemical simulations could be run in parallel). This is not currently feasible because we have no suitable, general-purpose method for efficiently sampling binding modes, and thus the cost of these calculations (y) is far too expensive in terms of both human effort and computational time. Our focus here is on developing a method for obtaining binding mode populations which has a cost y which is relatively favorable compared to Nx, or ideally even x so that a parallel calculation could complete in at most x wall-clock time.

A method that allows efficient sampling of binding modes would have broad uses in free energy applications, but would also aid in predictions of binding modes for structure-based design, fragment-based discovery, and other applications [42, 100, 88]. Here, our primary focus is on different "binding modes": defined as different metastable conformations of a fragment-like ligand within a single relatively rigid protein cavity. Metastable binding modes are thus those which are slow to interconvert on a simulation timescale x. Generally, if binding modes interconvert at a timescale slower than $x/10$ then proper sampling is a major concern; different metastable conformations may have different binding free energies but will be sampled in incorrect proportions, resulting in highly biased results. Moreover, the concept of a binding mode can include multiple ligand conformations in the same site, binding to in multiple sites, or even multiple protein conformations [83], though we do not specifically address enhanced sampling of protein motion here.

1.2 Theory and computational methods

1.2.1 Various sampling methods can be applied

There are a number of common sampling methods which can be used so that simulations sample the equilibrium distribution of populations. The efficiency of these methods can vary dramatically depending on which particular system or class of problem they are applied to, and a method that works best for one class of problem not necessarily most suitable for another class. Thus it is often nontrivial to determine which sampling method is best suited to a particular problem, or whether there is even a suitable method. Here, our particular interest is in accelerating sampling across ligand binding modes while still sampling the correct distribution of populations. Our goal is to develop a general method that can efficiently determine binding mode populations, in part by reducing the time it takes for simulations to switch between binding modes relative to other methods. This section will discuss some common sampling methods and the difficulties they encounter when applied to the ligand binding mode sampling problem.

Molecular dynamics (MD) is limited by the metastability of ligand binding modes

MD is typically used to simulate the dynamics of biomolecular systems by application of a force field which gives the forces between the atoms in the system as a function of their positions. With enough simulation time, MD should sample different metastable states with populations that are correct for a given choice of force field and ensemble, assuming that other simulation details—such as the integrator used to propagate dynamics—do not introduce errors. However, in practice sampling transitions between binding modes using MD is typically inefficient because of large energy barriers (and hence slow timescales) separating binding modes [78, 83, 24, 49].

Some free energy calculations attempt to get around this problem by assuming that similar ligands will have similar binding modes, so if a bound structure of a related ligand is available, it is assumed that new related ligands will share the same binding mode. However, this is not necessarily the case – even closely related ligands can have disparate binding modes that are slow to interconvert [49, 80, 83] Perhaps this is one reason why the accuracy of relative free energy calculations based on MD still falls short of what is desired for pharmaceutical applications [110], therefore, adequate binding mode sampling via direct MD simulation requires considerable computational expense and can necessitate specialized simulation hardware [109]. This inefficiency is compounded further in free energy calculations, as detailed above, where it is often necessary to adequately sample all relevant binding modes at each λ value to obtain correct binding free energies. In some cases, it is possible to sample long enough at the physical end states to cover all binding modes and then apply restraints to restrict the space treated at intermediate λ values, then compute the free energy of imposing and removing the restraints at the end states. This can improve efficiency modestly [78], but still requires simulations on timescales substantially longer than the timescales of the relevant motions.

Markov State Models (MSMs) can predict long timescale behavior efficiently, but are not ideally suited to our problem

The MSM approach assumes that a trajectory is generated from a Markov process. This assumption allows a statistical interpretation of MD trajectories. Specifically, a Markov State Model (MSM) is a matrix containing the transition probabilities between defined microstates, which can be used to predict the long timescale behavior of a system. The resulting model approximates the temporally coarse-grained dynamics with a Markovian surrogate model, which has certain properties that can used to predict the kinetics and equilibrium populations of each state [99]. Because a MSM is concerned only with the transitions between

states, multiple simulations can be used to generate the model, leading to more efficient use of computational resources. Specifically, rather than running a single very long simulation to adequately sample all binding modes, many shorter simulations can be used with substantially less wallclock time, at least if parallel resources are available [91].

The MSM framework also works to predict the long-timescale behavior of a system even before global equilibrium is reached, as long as local equilibrium is achieved, allowing a smaller total amount of simulation time to be used to estimate the equilibrium populations of of all states rather than having to fully converge to the global equilibrium [93]. Sampling at different thermodynamic states can also be employed with MSMs to improve transition estimates and sampling [53, 132].

However, the mathematics and assumptions behind MSMs unfortunately make this method difficult to use without expert knowledge and considerable care, and a maximum increase in efficiency is obtained only with prior knowledge of all potential binding modes. A sufficient number of transitions between states is necessary to properly estimate equilibrium populations from the MSM robustly. It is difficult to know *a priori* how much simulation data will be required to reach this stage, and it can require careful checking to know when this has been achieved [22]. There are also many parameter choices (order parameters, lag-times, clustering methods) which make constructing MSMs difficult to generalize, although recent developments such as GMRQ [73] and tICA [96, 108] help to reduce dependence on parameter choices.

Effective Monte Carlo proposals can accelerate sampling, but are difficult to construct for condensed-phase systems

In some cases, sampling can be dramatically accelerated by introducing Monte Carlo proposals informed by prior information, but this becomes increasingly difficult in condensed

phase systems. As a running example, consider a bistable dimer [65]. If we know approximately the relative locations of free energy minima (i.e. how far apart are the minima of the bond-length term), we might construct proposals that instantaneously hop from the vicinity of one minimum to the vicinity of the other. In a vacuum, this will dramatically accelerate mixing between the two metastable states of the dimer [90]. However, in a densely solvated environment it can be difficult to construct nontrivial proposals that avoid having near-universal rejection since instantaneously perturbing the coordinate of interest is likely to introduce clashes with solvent. While high acceptance rates could be achieved with extremely small perturbations, the long correlation times resulting from these small perturbations leads to highly inefficient sampling. Thus, while MC techniques have seen substantial use for biomolecular systems [54, 76], much of the field has moved towards using MD as a more general sampling engine and MC has to some extent fallen out of favor, partly because naïve MC moves in densely packed systems tend to overwhelmingly be rejected.

However, there have been some successes at combining MC and MD. For example the common replica [70] and Hamiltonian replica [117] exchange approaches use MC moves (involving swaps between replica simulations run under different conditions) to allow increased sampling in a variety of systems. MD itself can also be used as a MCMC proposal move as in hybrid Monte Carlo [30]. Additionally, in prior work in the YANK package, MC rotational and translational moves have been combined with MD to help with rapid ligand positional/orientational decorrelation while doing binding free energy calculations in implicit solvent [20, 125]. In general, however, designing MC moves that fully exploit available knowledge (to make nonlocal proposals) while retaining reasonable acceptance rates is difficult in the condensed-phase.

Nonequilibrium Candidate Monte Carlo (NCMC) couples MD with the benefits of MC and yields more efficient sampling

Nonequilibrium Candidate Monte Carlo (NCMC) provides a framework for translating insight about the system (in the form of a naïve Monte Carlo proposals) into practical algorithms [90] that retain some of the advantages of Monte Carlo while providing dramatically higher acceptance. The motivation is that it can be easier to construct a *finite-time proposal process* (a nonequilibrium "protocol") that achieves high acceptance rates with short correlation times than to construct a successful *instantaneous proposal*. In the dimer example above, instead of instantaneously proposing a single large dimer extension move, we may construct a nonequilibrium process including a sequence of small dimer extension increments. If, after every incremental "perturbation," the rest of the system is allowed to "relax"/"propagate," then we might end up with an acceptable proposal that has crossed a free energy barrier.

NCMC was originally presented in a very general setting, where (1) the target distribution is an *expanded ensemble* of configurations and thermodynamic states, (2) the protocols may mix arbitrary sequences of steps, and (3) each proposal is drawn from a distribution over protocols. Here, we consider a special case where we have only a single thermodynamic state, a single time-symmetric protocol, and a simple "perturbation kernel," so many of these terms cancel out and leave a simpler exact expression for the acceptance criterion. We make a further approximation in the acceptance criterion to improve performance, as we discuss further below.

NCMC permits nonequilibrium relaxation of most of the system while part of the system is being driven over (or around) kinetic or energetic barriers prior to acceptance or rejection of the NCMC move. Instead of proposing large instantaneous perturbations to the system, NCMC divides a target large perturbation into a series of steps consisting of smaller instantaneous perturbations, where each perturbation is followed by propagated dynamics. After

this series of perturbation and propagation steps, the whole sequence is accepted or rejected as an NCMC move. The intermediate states are always discarded and do not count towards any equilibrium averages or other properties as they are transiently out of equilibrium.

The NCMC procedure is performed via a protocol Λ, which utilizes a sequence of *perturbation* kernels $a_t(,\mathbf{y})$ and *propagation* kernels $K_t(,\mathbf{y})$. By "kernels" we mean conditional probability distributions, p, that we can evaluate pointwise and draw samples from. Furthermore, each kernel p must satisfy the requirement that if $p(,\mathbf{y}) > 0$ then $p(\mathbf{y},) > 0$, for all pairs $,\mathbf{y}$.

Here, we use a symmetric protocol consisting of T steps, where the perturbation and propagation steps are alternated with either a or K appearing at both the beginning and the end, so that $\Lambda = (a_1, K_1, a_2, K_2, \ldots, K_T, a_{T+1}) = \tilde{\Lambda}$, where $\tilde{\Lambda}$ is the reverse protocol. This protocol produces a trajectory $X \equiv (_{0,1}, \ldots, _T)$. To generate the appropriate acceptance for an NCMC move to maintain detailed balance, we also have to consider the probability of observing a time-reversed trajectory $\tilde{X} \equiv (\tilde{}_T, ...\tilde{}_1, \tilde{}_0)$ under the reverse protocol $\tilde{\Lambda} \equiv (K_T, a_T, K_{T-1}, a_{T-1}..., K_1, a_1)$, where $\tilde{}_t$ is the microstate $_t$ with reversed momenta. Because we are considering a symmetric protocol, however, the forward and reverse protocol are identical, thus simplifying the acceptance criterion.

The protocol used in BLUES is thus a symmetric sequence of perturbation and propagation events, starting and ending with perturbation. The perturbation typically consists "thermodynamic perturbation" — modifying the potential energy function to change interactions between the ligand and the protein. However, the central perturbation step in each NCMC cycle is an instantaneous perturbation of the ligand coordinates. These perturbation (thermodynamic or instantaneous) events are interspersed with propagation via Langevin dynamics.

For *perturbation*, we alchemically modify the potential energy function (described in detail below) to slowly annihilate and then restore ligand interactions with the environment, result-

ing in a sequence of reduced potentials u_t that incorporate the time-dependent interactions of the ligand with its environment. In the middle of the protocol, when the ligand is no longer interacting with the environment, we rotate the ligand into a new orientation in an operation that does not change the potential energy of the system. This is done by translating the center of mass of the ligand to the origin, applying a rotation matrix to its coordinates, and reversing the translation to restore the ligand's original center of mass. The rotation matrix is drawn uniformly over the space of all rotations using a quaternion approach, in which a random quaternion is generated uniformly over a 4D hypersphere and converted to a rotation matrix. This ensures the probabilities of generating a particular rotation matrix and its inverse are equal so that the overall proposals are time-symmetric.

For *propagation*, we use a Langevin integrator with specific properties (described below).

A variety of acceptance criteria $A[X]$ applied at this point would restore the system to equilibrium. For the case of symmetric protocols $\Lambda = \tilde{\Lambda}$ where all perturbation operations are symplectic (preserve phase space volume), the acceptance probabilities for a proposed NCMC trajectory X given protocol Λ must satisfy

$$\frac{A[X]}{A[\tilde{X}]} \equiv e^{-\Delta\mathcal{S}[X]} = e^{-w[X]}, \tag{1.1}$$

where $\Delta\mathcal{S}(X)$ is the conditional path action difference, which is equivalent to the total work $w[X]$. The total work includes both protocol work and "shadow work" [111],

$$w[X] \equiv w_{\text{protocol}}[X] + w_{\text{shadow}}[X] \tag{1.2}$$

where the protocol work depends on the changes in reduced potential energy during each of

the perturbation steps,

$$w_{\text{protocol}}[X] \equiv \sum_{t=1}^{T} [u_t(\overset{*}{t}) - u_{t-1}(t^*)] \tag{1.3}$$

By contrast, the $w_{\text{shadow}}(X)$ depends on internal details of the propagation scheme used [111].

While neglect of the shadow work can lead to large errors in general [111], we select a specific Langevin integrator that preserves the configurational distribution to very high accuracy, the BAOAB integrator of Matthews and Leimkuhler [63, 64], allowing us to neglect this contribution without introducing large error. We justify this approximation by observing that the sequence of Langevin propagation kernels are *nearly exact* Markov kernels, each preserving the distribution $\pi_t() \propto e^{-u_t()}$ with high fidelity. Recall that, due to discretization error, the invariant distribution $\rho_t()$ sampled by a numerical algorithm for Langevin dynamics will differ slightly from the target (i.e. $\rho_t \approx \pi_t$), and the magnitude of this difference increases with the integrator step size. This may introduce bias. We neglect this bias, since the specific integrator employed here is thought to preserve the configurational distribution to very high accuracy [63, 64]. This conclusion is based on analytical results showing that the integrator approximates configurational averages to fourth-order in the timestep (as opposed to second-order for competing integrators), and extensive numerical evidence examining particular biomolecular observables [63, 64]. Note that we would also use this criterion if each propagation step were a reversible MCMC move.

In practice, using an exact MCMC kernel (such as Generalized Hamiltonian Monte Carlo) for propagation would substantially increase the computational expense of a protocol by (a) introducing costly energy evaluations during accept-reject steps, (b) reducing the feasible integration timestep, and (c) dramatically increasing correlation times if the acceptance rate is even slightly less than 1 [121]. Including the shadow work contribution would also

substantially reduce the acceptance rate of long protocols. In future work, we will examine the bias vs. efficiency trade-offs of this approximation, and the extent to which these can be mitigated by choice of Langevin integrator, or by using reduced-momentum-flipping variants of Hamiltonian Monte Carlo [121, 115, 14]. While it has been argued more generically that the contribution of "shadow work" in nonequilibrium simulations can be neglected without introducing much bias [19], this is likely highly dependent on the specific choice of integrator used, so we recommend caution if other integrators are considered.

We combine NCMC with random ligand rotational moves

Here, we provide details of our NCMC move proposals for ligand binding mode sampling. We combine thermodynamic perturbation of the ligand (alchemically changing its interactions with the protein) with uniform random rotation around the ligand center of mass (COM). Specifically, we scale λ over a series of n NCMC steps until the ligand no longer interacts with the protein (removing its steric and electrostatic interactions). λ scales the interactions between the ligand and the rest of the system; to elaborate further, λ controls the strength of interactions between the ligand and its environment, with $\lambda = 1$ corresponding to the fully interacting state, and $\lambda = 0$ corresponding to the non-interacting state (as discussed in 1.2.5. The ligand is then randomly rotated to a new orientation in the binding site around it's center of mass. Then its interactions are turned back on by scaling λ over a series of another n NCMC steps, as conceptually shown in Figure 1.2. Finally, we use the analogue of the Metropolis-Hastings acceptance criteria [47] that satisfies Eq. 1.4 to accept or reject the resulting move.

$$A[X] = \min\left\{1, e^{-w_{\text{protocol}}(X)}\right\} \tag{1.4}$$

Ligand rotation does not strictly need to be around the COM; it could be around a randomly selected heavy atom, or a point chosen within a Gaussian distribution around the COM, or various other options; we use the COM here for simplicity.

Figure 1.2 shows a cartoon of how these NCMC moves can work for exploring ligand binding modes. The ligand starts fully interacting (Figure 1.2(A)) and its interactions with the rest of the system are slowly turned off through alchemical λ coupling over a series of NCMC steps (Figure 1.2(B,C)). When the ligand is fully non-interacting, a random rotation (see Section 1.2.1) around the ligand's center of mass is performed (Figure 1.2(D)). Then the ligand's interactions are subsequently turned back on until it is once again fully interacting, potentially allowing it to find an alternate favorable orientation in the binding site (Figure 1.2(E,F)). We then accept or reject the move based on the acceptance criteria in Equation 1.4. In order to preserve detailed balance, the momenta of the system must be reversed after acceptance or rejection of proposed moves, or the momenta must be reassigned randomly from a Boltzmann distribution after the move [51, 59, 121]; in this work, we take the latter approach.

1.2.2 We study a T4 lysozyme cavity mutant which binds simple ligands

Here, we test several methods, including our new NCMC rotational method, on a T4 lysozyme cavity mutant which binds simple ligands. The T4 lysozyme L99A cavity mutant studied here has a buried binding site which readily binds non-polar molecules, making it a common model system for free energy calculations [81].

Toluene, a T4 lysozyme L99A binder, was chosen as the initial ligand for testing this method for a number of reasons. One is that toluene's symmetry allows for a convenient check of correctness; symmetric binding modes should have equivalent populations with adequate

sampling. Also, the different potential binding modes for toluene differ primarily based on rigid body rotation of the ligand in the binding site, so rotational moves should increase sampling of the relevant binding mode(s). In addition, previous conventional MD simulations we ran of toluene bound to T4 lysozyme suggest two distinct stable binding modes are present. Adequate sampling of even these two simple binding modes poses significant challenges for conventional MD [78]. After testing our NCMC rotational method on toluene, we also explored its capacities on 3-iodotoluene, a more bulky ligand. 3-iodotoluene does not have the same symmetry as toluene, meaning that we cannot take advantage of symmetry as a convenient check for convergence of populations. However, it is still valuable as a test for efficiency on bulkier molecules.

1.2.3 System preparation

Generic T4 lysozyme/toluene system setup

The T4 lysozyme L99A structure with toluene bound was taken from the 4W53 protein structure from the Protein Data Bank. Hydrogens were added to the protein using `tleap` from AmberTools14 [17, 15]. Hydrogens were added to toluene using Maestro and parameterized using GAFF version 1.7 [124] and AM1-BCC charges [52]. Hydrogens and missing atoms of the protein were added using `tleap` in AmberTools14, and parameterized using `ff99sbildn` [67]. A TIP3P rectangular box was added with 10Å padding from the protein to the nearest box edge, and Cl^- atoms were added to neutralize the charge of the system. The resulting `.prmtop` and `.inpcrd` files were converted to the equivalent GROMACS formats using `ACPYPE` [28].

The system was then minimized using steepest descent running for 2500 steps. The system was then equilibrated in GROMACS 5.1 for 25000 2 fs steps with constant volume, then equilibrated under constant pressure for the same number of steps using a Parrinello-Rahman

barostat to maintain a pressure of 1 atm. Long range dispersion corrections were used for calculating the energy and pressure. These preparatory simulations were performed at 300 K and `v-rescale` with `tau_t` set to 0.1 ps was used to perform temperature coupling.

Full details of the simulation setup can be found in the `.mdp` files in the Supporting Information (SI).

T4 lysozyme/3-iodotoluene system setup

The T4 lysozyme L99A structure was taken from the 4W53 protein structure from the Protein Data Bank, with the toluene ligand removed. 3-iodotoluene was then docked using OpenEye's FRED (ver 3.2.0.2) and we retained the top scoring generated pose. Preparation of the system using `tleap` was done the same as in 1.2.3. The system's energy was then locally minimized with a tolerance of 10 kJ/mol. The system was subsequently equilibrated in OpenMM under constant pressure at 1 bar with and 300 K using a Monte Carlo barostat for 25 ns using a Langevin integrator with a 1 ps timestep and 1/ps collision rate.

Setup for OpenMM NCMC simulations

OpenMM 7.1.0 was used [31]. The OpenMM simulations used the same systems loaded from the .prmtop and .inpcrd files as prepared in Section 1.2.3. For the MD portions of the protocol a Langevin integrator was used with a 2fs timestep and 1/picosecond collision rate. No barostat was used for these simulations after equilibration (and thus simulations were done in the NVT ensemble).

1.2.4 We built Markov state models of toluene binding to lysozyme

The T4 lysozyme system with toluene bound (as described in 1.2.3) was minimized in GROMACS 5.1 [7, 2] via steepest-descent, followed by 1 ns of NVT simulation and then 5 ns of NPT simulation at 1 atm and 300 K for equilibration. The leapfrog integrator was used with a 2 fs timestep and the bonds involving hydrogen constrained with LINCS [50]. The system was then simulated for a total of 806 ns under the same NPT conditions, saving configurations to a trajectory file every 30ps. tICA was performed on the pairwise-distances of the toluene heavy-atoms and the alpha carbons of the binding site of the trajectory with a lagtime of 0.6ns to generate order parameters for MSM construction. Of the 210 initial dimensions, 22 dimensions were retained—enough to account for 95% of the kinetic variance in the data, and were scaled by the kinetic map scheme. These large number of initial dimensions were used to help ensure all relevant binding modes were separated. An initial MSM was estimated from the order parameters computed from the trajectory using PyEMMA [106], using 1000 microstates generated from k-means clustering and a lagtime of 6 ns. The MSM was coarse-grained into four macrostates using Perron-Cluster Cluster Analysis ++(PCCA+) [13, 29, 101]; full details are available in scripts deposited in the SI. Two random trajectory frames from each of the four macrostates were then used as the starting point for new simulations to further sample each identified binding mode and potentially generate additional transitions. These eight additional simulations were each run for 60 ns and combined with the longer run above to re-estimate a MSM, following the same sequence of steps, with these additional simulations added to better explore transitions out of each macrostate. The total amount of aggregated simulation time used for the final MSM was 1.286 μs spread across nine trajectories. Additional simulation details can be found in the SI.

1.2.5 We use Nonequilibrium candidate Monte Carlo (NCMC) to study toluene binding to lysozyme

Our NCMC procedure uses random rotational moves to enhance binding mode sampling

Here we use NCMC to enhance sampling of ligand binding modes in the T4 lysozyme binding site. As discussed in Section 1.2.1, coupling thermodynamic perturbation with rotational move proposals can allow the ligand to cross energy barriers between binding modes while allowing some amount of relaxation to improve the acceptance of proposed moves. In our procedure (Figure 1.2) interactions between the protein and ligand are on at the beginning of a move proposal. Then the interactions are turned off by scaling λ from 1 to 0 (where 1 corresponds to full interactions and 0 corresponds to no interactions) over a series of n steps, following the scheme shown in Figure 2.4. Soft core potentials were used to avoid numerical instabilities related to scaling the steric and electrostatic interactions, with a 1-1-6 potential with $\alpha = 0.5$ [9]. As the NCMC protocol progresses, we first turn off the electrostatics of the ligand by scaling its potential energy contribution linearly with lambda so that the electrostatics are completely removed as we go from $\lambda = 0$ to $\lambda = 0.2$. Then we decouple the Lennard-Jones interactions using soft core potentials from $\lambda = 0.2$ to $\lambda = 0.5$ so that the ligand is now fully non-interacting.

Then a random rotation of the ligand is performed (as described in Sec 1.2.1), with the random quaterion generated using `mdtraj` [71]. Finally, the interactions are subsequently turned back on via a reverse of the original procedure.

The work done during this process is accumulated and used to accept or reject the move (consisting of the full decoupling, rotation, and recoupling procedure) via Equation 1.4.

After the NCMC move is accepted or rejected, velocities are randomized by drawing from the

Maxwell-Boltzmann distribution appropriate for the temperature and then a phase of conventional MD is performed to better sample the other (protein/solvent) degrees of freedom. This full procedure consisting of NCMC moves plus MD steps is then repeated many times until convergence, and the populations can be then estimated from clustering the resulting trajectory and computing the time spent in each state.

We implemented this approach via our BLUES package for binding mode sampling

We constructed a package called Binding modes of Ligands Using Enhanced Sampling (`BLUES`) to facilitate the use of NCMC to enhance ligand sampling. `BLUES` implements the approach outlined in Section 1.2.5 and switches between sampling the system via normal MD and NCMC alchemical perturbation. The BLUES package allows straightforward control of the number of MD steps between each NCMC move, the number of alchemical steps used within each NCMC move, and the total simulation time and number of MD/NCMC cycles.

In `BLUES` the `alchemicalfactory` module of `openmmtools` [23] version `0.11.2` was used to allow annihilation and restoration of toluene's steric and electrostatic interactions. The MD portions of these simulations used `OpenMM`'s Langevin integrator. The NCMC portion of the `OpenMM` simulations used an implementation of the BAOAB integrator of Langevin dynamics [64] During the NCMC propagation steps we also froze the positions of protein residues outside of 5Å from the ligand, and the solvent molecules. The selection of the frozen water and protein residues was not updated during the simulation; this was appropriate in this case as the binding site is a buried, non-polar binding site with no water nearby and the ligands remain stably in the binding site, so no updates were needed. The long range dispersion correction was turned off during the NCMC integration steps due to computational costs recalculating the correction while scaling λ, but was accounted for by taking into account the differences in energy between the alchemical and normal systems at the initial and final

states.

Full details of the implementation and a class diagram are available on GitHub at the link below.

`BLUES` is also an extensible framework in that it allows general MC moves to be performed during the NCMC portion. Here we consider only random rigid-body rotations around the ligand center of mass as described in Section 1.2.5. However, other moves which might be of interest to explore later could include translations of subsets of a given system, moves involving ligand internal coordinates, or sidechain MC moves.

`BLUES` is freely available on GitHub under the MIT license at `https://github.com/MobleyLab/blues`. We used `BLUES` version 0.1.0 to obtain the data found in this paper. The same systems from Sections 1.2.3-1.2.3 were used.

Variations in the NCMC protocol dramatically impact move acceptance

With NCMC in BLUES, we can vary the number of perturbation and propagation steps relative to the amount of standard MD in order to allow an adequate amount of relaxation in order to ensure reasonable acceptance without using so much relaxation that the approach becomes tremendously inefficient. Here, we tested this by measuring the acceptance ratio as a function of the amount of relaxation (Figure 1.4), and found that the acceptance probability increases rapidly from around 100 NCMC switching steps up to 10000 NCMC switching steps, then begins increasing more slowly with further relaxation. Thus, here, we selected 10000 NCMC steps per cycle as a reasonable choice in order to determine how much enhancement in sampling (and therefore efficiency) NCMC can provide relative to standard MD or MD with MC. Corresponding work distributions and standard deviations of work distributions are shown in SI Figures 1–2.

1.2.6 For reference, we compare NCMC with conventional MD and MD/MC

To compare the efficiency in sampling with NCMC versus more traditional forms of sampling, we also ran normal MD and MD with MC rotational moves on toluene with the same T4 lysozyme/toluene system. In order to make a fair comparison between methods we kept the number of force evaluations consistent across methods.

For MD, we ran 20000 integration steps of MD during one iteration to mimic the number of NCMC force evaluations per iteration. For MD with MC we ran 20000 integration steps followed by 10 MC random rigid body ligand rotations using the Metropolis criteria for each iteration. We ran each of these methods for 5000 iterations and then compared the trajectories and binding mode populations found in our NCMC simulation (Figure 1.9). The code used to perform these calculations can be found in the SI.

We observed very few transitions between binding modes in both the normal MD and MC with MD simulations. Because there are so few transitions in these cases we cannot expect the binding mode populations in these simulations to be converged to the equilibrium populations.

1.2.7 We analyze our binding mode sampling using a dihedral angle which discriminates between the stable binding modes

Originally, we used many pairwise-distances as an order parameter when constructing the MSM to identify toluene's binding modes (Sec 1.2.4). Once those binding modes were identified however, further analysis found that a simple 1-dimensional progress coordinate could separate and identify them. To monitor the binding mode of toluene in the cavity, we picked a dihedral angle which clearly discriminates between toluene's four distinct binding modes.

Specifically, toluene's binding mode was monitored via calculation of the dihedral formed by the alpha carbon of ARG118 and the C1, C5, and C7 atoms of toluene (Figure 1.5).

These angles were then used for construction of histograms to monitor populations of the observed binding modes and the number of transitions between binding modes (Figure 1.9). The populations of the different binding modes were monitored by using the following different bin boundaries [-π, -1.5), [-1.5, 0), [0, 1.5), [1.5, π). We assigned the following state labels according to the bin locations: [-π, -1.5) is B1, [-1.5, 0) is A1 [-1.5, 0) is B2, and [1.5, π) is A2, where the A labels correspond to the crystallographic binding mode and B labels correspond to an in-plane rotated binding mode. To determine the uncertainty in the computed population as a function of simulation time, the populations of each binding mode were determined using fractions of the total simulation data in a blocking approach [45, 38]. The uncertainty in the computed populations were determined based on breaking each 10% of the simulation into a set of smaller blocks. The number of blocks used was the amount that maximized the standard deviations of the populations between blocks. For NCMC/MD this was 8 blocks. For MD and MC/MD, the standard deviation in the mean across blocks via bootstrapping failed to reach a maximum even with only one block per 10% of the simulation. The error bars in the plots of the MD and MC/MD simulations are shown based on one block per 10% of the data, but are likely to severely underestimate the true error.

1.2.8 We generated synthetic data to compare MD and NCMC transition efficiency

To get a better sense of the efficiency gains of NCMC compared to standard MD we constructed statistical models from the data of the MSM and the NCMC simulations. For the MSM we took the estimated MSM transition matrix directly after clustering the 1.286 μs into four clusters corresponding to the four binding modes. For the NCMC simulation, we

assigned each iteration to a particular macrostate using the dihedral order parameter as defined in Sec 1.2.7. We used those state assignments from those 5000 iterations to generate a transition matrix. To generate the synthetic data, we started from state A2 and iteratively applied the transition matrices to get trajectories of states. A new state could be the same as the previous, depending on the transition probabilities given by that particular transition matrix. This process was repeated, and the total state populations at each iteration were recorded. We performed 5000 propogation iterations for the MD and NCMC transition matrices with 1000 trials each. This allowed us to cheaply estimate uncertainties in the populations of each state at each time point by taking the standard deviation in the estimated population across all trials. Overall, this analysis facilitated an assessment of the rate of convergence of the populations.

1.2.9 We examined rotational distributions and added the case of 3-iodotoluene as an example of a bulkier ligand

To better compare the performance of NCMC and standard MC move proposals (as discussed in Results 1.3.2), we chose the bulky lysozyme ligand 3-iodotoluene and compared the efficiency of running a large number of standard MC move proposals with the efficiency of running many NCMC move proposals. Because our BLUES tool is not designed for efficient MC performance (since it has additional overhead to facilitate relaxation with NCMC) we did this test outside of BLUES, instead running standard MD simulations and then selecting snapshots from these to compare acceptance of MC and NCMC move proposals. Thus this test is not a benchmark of NCMC against MC, but an assessment of the performance of NCMC and MC move proposals starting from the same ensemble of MD snapshots.

As preparation, we simulated the T4 lysozyme system with 3-iodotoluene for 90 ns, saving the positions every 0.2 ns, thereby saving a total of 450 trajectory snapshots. We used these

snapshots to facilitate our comparison of the acceptance of MC moves and NCMC moves. For 3-iodotoluene, our goal is to assess overall acceptance, and see whether substantial rotational moves are being accepted with reasonable frequency – in contrast to our work on toluene (Section 1.2.6) where we were interested in estimating populations in order to ensure that our approach converges to the correct populations. Thus at the start of each MC or NCMC move attempt we randomly pick a starting trajectory snapshot to use as a starting point for a new move proposal. This allowed us to have a variety of starting points for our move proposals, while also ensuring that we were assessing the performance of move proposals with equivalent environments. For MC we performed 10 trials, each consisting of 2,000,000 move attempts where each move is instantaneous. For NCMC we performed 7 trials, each consisting of 2000 move attempts; each move consisted of 6500 NCMC switching steps.

Since we were interested in comparing not just acceptance rate but how these moves fared at *substantially* decorrelating ligand orientation within the binding site by sampling across different binding modes, we also monitored the angle by which each move rotated the ligand. Specifically, when a move was accepted, we calculated the angle of rotation by first calculating the rotation matrix needed to generate the final ligand positions from the initial ligand positions [8], then calculating the angle of rotation θ by Eq 2.1, where R is the rotation matrix.

$$\theta = \arccos(\frac{Tr(R) - 1}{2}) \tag{1.5}$$

We also performed a similar routine to determine the rotation distributions in the toluene case, except we performed MD/MC or MD/NCMC sampling as previously described in Sec 1.2.6. For toluene, we ran 5 MD/MC trials consisting of 10000 iterations, with each iteration consisting of 10 MC moves attempts followed by 1000 steps of MD. For MD/NCMC we ran one trial consisting of 10000 iterations, with each iteration consisting of a NCMC

move of 10000 NCMC switching steps and 1000 steps of MD.

In all cases, the resulting rotational distributions were histogrammed using 8 bins of 22.5 degrees; the error for each bin was determined using the standard error in the mean estimated across trials, except for the MD/NCMC T4 lysozyme/3-iodotoluene simulation, in which we estimated the error by dividing up the accepted frames of the trajectory into 8 blocks, which maximized the standard deviation, then computing the standard error in the mean by bootstrapping over the accepted blocks.

To monitor the binding mode, we found a dihedral order parameter that separated the 3-iodotoluene binding modes observed during the simulations (see SI); this involved the C1, C5, and I8 atoms of 3-iodotoluene and the alpha carbon of VAL111 of T4 lysozyme. This was used to monitor the overall orientation of the ligand in the binding site, e.g. in SI Figure 3.

Results are given in Section 1.3.2.

1.3 Results and Discussion

1.3.1 Kinetics and populations of binding modes through MD and Markov State Modeling

We constructed a MSM from approximately 1 μs of simulation data on the T4 lyszosyme/toluene system (see Sec 1.2.4) to estimate equilibrium populations of the binding modes and timescales of interconversion. From the implied timescales of the MSM (Figure 1.8) we identified 4 kinetically separated binding modes of toluene as expected from the gaps between the third and fourth timescales.

Perron cluster cluster analysis+ (PCCA+) was then used to separate the trajectory frames of the MSM into 4 clusters. Visual inspection of the resulting macrostate clusters from PCCA+ revealed that there were two clusters, each with a symmetry equivalent partner (Figure 1.6). The populations estimated from the MSM show the populations of the symmetric states to be roughly equal, with 32±8% and 26±6% for the two symmetric binding modes corresponding to the crystallographic binding mode. The other binding mode showed 18±5% and 23±6% populations for the symmetric equivalents (SI Figure 4).

We find that timescales for binding mode interconversion are extremely slow, both from analyzing our long conventional MD simulation directly, and from the implied timescales of the MSM. Direct analysis of our long single 806 ns trajectory (Figure 1.7) showed that certain binding mode transitions are quite rare.

Additionally, the slow kinetics involved in sampling are highlighted by the implied timescales of the MSM (Figure 1.8). The slowest transition–switching between symmetric binding modes–occurs on a timescale of 100 ns, while the faster transitions occur approximately every 4ns.

It is important to note that even with the simplicity of both this binding site and toluene (which is a rather small ligand compared to the size of the site), slow transitions are still observed, consistent with earlier observations that binding mode interconversion is quite slow in the buried lysozyme binding site[11].

Given the timescale of the slowest binding mode transitions observed here, obtaining accurate ligand binding mode populations from brute force MD or even MSMs seems particularly costly in this case. Specifically, to generate an accurate representation of the populations either approach will need to observe multiple transitions between binding modes. Especially in the MD case, this would require simulations which are at least 10 times longer than the 100ns timescale for the slower binding mode interconversion events—equivalent to at least 1μs

of simulation. While toluene's symmetry could be used to obtain correct populations without adequately sampling the symmetric ring flip, any new ligand differing by a substitution breaking this symmetry (and there are many such ligands which bind in this site, such as the 3-iodotoluene case considered later in this work [82]) would require adequate sampling of these previously symmetry-equivalent binding modes. In other words, adequate binding mode population estimates would likely require multiple microseconds of simulation time.

1.3.2 BLUES rapidly samples binding modes

Initial simulations indicate BLUES samples more rapidly than MD

Here, to compare the efficiency of our NCMC approach, BLUES, to that of brute force molecular dynamics and MSMs, we applied BLUES to the T4 lysozyme/toluene system. We applied the NCMC protocol of thermodynamic perturbation with random rigid-body ligand rotations (as described in Section 1.2.1) to observe the protocol's efficiency in sampling binding mode interconversions. The NCMC protocol was applied over 5000 iterations, each consisting of 10000 MD steps separated by NCMC move proposals consisting of turning off and restoring ligand interactions over 10000 steps, with random rotations while the ligand is noninteracting.

For reference, we also performed standard MD and MD/MC simulations using the same total number of force evaluations. In the MD case, this meant 5000 iterations of 20000 MD steps. And in the MD/MC case, this meant 5000 iterations of 20000 MD steps interspersed by 10 conventional MC move proposals involving random ligand rotation. The number of force evaluations at each iteration was kept constant between methods. Thus, with a 2fs timestep we simulated for an equivalent of 200ns total with each method.

Figure 1.9 shows the dihedral angle (indicating the binding mode) sampled versus time over

each method's iterations, along with the resulting histogram of the binding mode populations. Compared to MD or MD coupled with traditional MC moves, this NCMC method allows rapid interconversion between all four binding modes.

This allows BLUES to reproduce the correct equilibrium populations (Figure 1.9(C), right), and on a force evaluation basis, NCMC was approximately 17 times more efficient than brute force MD with Markov State Modeling.

This is also evidenced by the fact that over the same number of iterations, BLUES converges rapidly to the correct equilibrium populations within 2000 iterations (Figure 1.10), whereas MD and MD/MC still have significant errors. For MD after 5000 iterations the major binding mode populations differ from the equilibrium populations by as much as 45%; for MC they differs by as much as 26%.

Although this NCMC implementation is more efficient on a force evaluation basis compared to MD, there is some computational overhead for alchemically modifying the sterics and electrostatics during integration in OpenMM for GPU simulations, causing the wall-clock time per NCMC iteration to be about three times longer than a MD iteration. Specifically, our calculations shown in Figure 1.10 took 2249 minutes for MD, 2189 minutes for MD/MC, and 5413 minutes MD/NCMC. Convergence to within uncertainty of the correct population, and to within 5% of the correct population, appears to occur for this system (Figure 1.10) well before 40% of the total simulation time (80ns); with a factor of three in additional cost, this takes about as long to run as 240ns of conventional MD simulation. Thus the savings of NCMC in terms overall wall-clock time is still about a factor of five compared to the MSM approach for this system, which required roughly 1.3μs of aggregate simulation data.

Statistical analysis confirms the efficiency of BLUES compared to MD

To validate and further assess the relative efficiency of BLUES compared to MD, we built a statistical model of convergence of populations in these two cases. Specifically, we wanted to use transition matrices between the four states (in both cases) to propagate the populations over a long time in order to analyze convergence properties. We obtained the MSM transition matrix for the MD case. For the NCMC case we constructed a transition matrix from our BLUES simulation (as described in Section 1.2.8). We then built a model of convergence using these two matrices as a starting point. In each case, we started a hypothetical simulation from binding mode A1 and used the transition matrix to propagate the populations, at each timestep choosing a new state to transition to based on the probabilities in the transition matrix. In the MD case the transition matrix was constructed using a lagtime of 6 ns so our simulation timesteps corresponded to 6 ns, whereas in the BLUES case the transition matrix was constructed for a 40 ps MD/NCMC iteration so timesteps were 40 ps. We then repeated 1000 such simulations for each case and examined the mean population as a function of time and the standard deviation over trials. These are shown in Figure 1.11. In this case we find that NCMC converges much more quickly than MD, specifically, for MD it takes approximately 12000 nanoseconds for the standard deviation in the slowest converging population to drop below 5%, indicating that typical simulations would have a population error of no more than 5%. In contrast, for NCMC the standard deviation drops below 5% for the slowest converging population by 60 nanoseconds, a factor of 200 reduction in total simulation time compared to MD.

It is important to note that in this model the transition matrices are only estimates of the true transition matrices, so populations will eventually converge to a stationary distribution as seen in Figure 1.11, but the final populations will have some amount of error. Here we are more interested in examining the rate of convergence than the populations as our goal is to measure the relative efficiency of both techniques.

Ultimately, the difference in performance between MD/NCMC and the benchmark MD and MD/MC approaches is fairly simple to understand. Conventional MD cannot cross kinetic barriers any faster than their inherent timescales, so, since timescales for interconversion between the slowest binding modes here are around 100ns (Section 1.3.1), convergence in conventional MD will necessarily take many times longer than 100ns. The MD/MC approach here couples conventional MD with occasional random ligand rotational moves which are accepted or rejected via conventional Monte Carlo, but because the binding site is relatively densely packed—even though it is not solvent exposed—the vast majority of these are rejected for toluene (giving an acceptance rates of 0.091±0.004%). Thus, this approach performs almost equivalently to standard MD. Our NCMC approach implemented in BLUES converges much more rapidly because ligand rotational move proposals can relax before being accepted or rejected, thus dramatically enhancing the acceptance rate to approximately 11%. These acceptance rates are reflected in the transitions between states, (Figure 1.12), which show that MD/NCMC produced 497 transitions. This is more than twice as many transitions than the other methods employed; during the same number of iterations MD produced 242 transitions and MD/MC produced 230 transitions. Also, MD/NCMC produced high transition counts from any given binding mode to any other binding mode, while the other methods primarily produced transitions from a given binding mode to a subset of all the binding modes. Specifically, the other methods had the most transitions between in-plane binding modes (A1-B1 or A2-B2 transitions, Figure 1.12(a) and (b)) which are relatively fast to interconvert in normal dynamics, where BLUES had significant numbers of transitions between all binding modes, even the out-of-plane flip, which has a characteristic timescale of roughly 100 ns for conventional MD (Section 1.3.1). For example, the A1 to A2 transition occurred only twice in standard MD, and once in MD/MC, but 48 times in BLUES. This is also clearly apparent from Figure 1.9 (left panels), where the MD and MD/MC approaches have few transitions between the top pair of states and the bottom pair of states, but BLUES has a very large number.

NCMC does not compare as favorably to MC for toluene, but NCMC moves have clear advantages for bulkier iodotoluene

While BLUES compares favorably to standard MD in the case of toluene bound to lysozyme, and has a better acceptance ratio (10 ± 1 %) than standard MC (0.091 ± 0.004 %), the difference in acceptance ratio between BLUES and standard MC is not actually enough to justify the additional computational expense of the NCMC relaxation. Specifically, instead of doing 10000 NCMC steps (as in BLUES) to achieve a reasonable acceptance rate, we could simply do a very large number of MC trials (e.g. 10000, for a similar computational expense) with the low acceptance rate and still see a reasonable number of moves accepted. We believe this relative advantage of standard MC may be a peculiarity of toluene in this particular binding site (which is relatively large compared to the size of toluene, and known to be especially rigid [126, 82, 86, 81]), and not representative of typical MC performance in condensed phase systems [90].

To further test the relative performance of MC and NCMC, we examined performance of NCMC versus MC move proposals for a larger ligand in the lysozyme binding site, 3-iodotoluene; we find that the presence of the bulky iodo substituent dramatically impairs acceptance of MC moves, presumably due to the larger size of the ligand relative to the size of the binding site (see Section 1.2.9 for methods). 3-iodotoluene is another known binder in the lysozyme L99A site; however, due to its lack of symmetry, we are unable to take advantage of ligand symmetry to provide a simple metric for convergence of binding mode populations. It is nevertheless useful here as a good example of a larger ligand which should have several different metastable binding modes in this site.

For 3-iodotoluene, our comparison focuses just on acceptance of MC and NCMC *move proposals* given a fixed ensemble of MD snapshots as a starting point, and we find that the acceptance rate of standard MC is $(1.2 \pm 0.2) \times 10^{-2}$ % while for NCMC, it is 0.8 ± 0.1%.

Thus, standard MC results in an order of magnitude lower acceptance for 3-iodotoluene than toluene, meaning that standard MC is closer to NCMC's performance on 3-iodotoluene rather than outperforming it (in terms of acceptance rate) as in the case of toluene.

However, the overall acceptance probability is not the only consideration – we are also interested in how each technique improves the acceptance of *substantial* moves that significantly alter the ligand binding mode. After all, when proposing random ligand rotational moves, a rotation of zero, or of very nearly zero, is a valid move proposal, so potentially many of the moves being accepted are in fact very small ligand rearrangements. To examine this, we determined the fraction of rotational moves accepted as a function of amount of rotation (Figure 1.14), both for toluene and for 3-iodotoluene.

For toluene, due to its relatively small size, MC and NCMC result in relatively similar acceptance profiles as a function of the amount of rotation (Figure 1.14(a) and (b)), except perhaps that NCMC yields improved acceptance of intermediate amounts of rotation between 0 and 180 degrees. However, for 3-iodotoluene (Figure 1.14(b) and (c)), standard MC results in virtually no acceptance of moves larger than 20 degrees ($(5 \pm 2) \times 10^{-5}$ %), whereas NCMC retains good acceptance of such moves (0.68 ± 0.07 %). (For the MC case on the 3-iodotoluene system, we performed a total of ten trials of 2,000,000 MC attempts and in each trial, typically saw at most one or two accepted moves consisting of significant rotations.) Accounting for the 6500 relaxation steps used in NCMC, MC would have a $(3 \pm 1) \times 10^{-1}$ % acceptance of significant rotations indicating that NCMC is still approximately twice as efficient than MC in this case.

This is further supported by SI Figure 5, where we examine how effective MC versus NCMC moves are at sampling new binding modes not represented in the MD trajectory providing the starting points for our move attempts. For MC, only very few moves outside the starting region are accepted, whereas NCMC is quite effective at sampling new binding modes. Additionally, MC gives the apparently false impression that the initial set of binding modes

is by far the most favorable (because it is so difficult to find a combination of ligand orientation and protein conformation which can allow a rotational move to be accepted with no relaxation), whereas NCMC suggests that there are alternate binding modes that may be considerably more favorable.

Thus, the acceptance ratio only gives a small part of the overall picture; NCMC does dramatically better than MC at sampling significant binding mode transitions, enough so that even in this relatively simple 3-iodotoluene test system, NCMC outperforms simply performing a very large number of MC trials (with an equivalent number of energy evaluations) by roughly a factor of 2.

We also examined the performance of NCMC move proposals for 3-iodotoluene as a function of the amount of relaxation, as shown in Figure 1.13. In keeping with the analysis just prior, we find that NCMC move proposals perform nearly as well for substantial rotations as for all rotations, whereas MC move proposals do not.

Overall, our tests on 3-iodotoluene indicate that for larger ligands and/or more confined environments, standard MC move proposals perform dramatically worse than for toluene, in keeping with what might be expected for large moves in condensed-phase systems [90]. This, combined with the overall flexibility of the NCMC approach in combining some of the advantages of MD with those of MC, indicates that this approach may be a good general strategy for ligand binding mode sampling. Additionally, this work highlights how important it is not just to monitor the overall acceptance rate of moves, but how the acceptance rate of moves is coupled to the size of moves; here, NCMC results in high acceptance of large moves, while MC does not.

1.4 Conclusions

1.4.1 Summary

Overall, we find that NCMC with random ligand rotational moves dramatically enhances sampling of ligand binding modes compared to the other more conventional methods employed here.

Particularly, we have shown that NCMC can greatly enhance move acceptance for exploring ligand binding modes by allowing for relaxation during attempts. NCMC also allows dramatically faster sampling than standard MD because of its ability to cross steric barriers; advantages over standard MC are less clear but grow with the size of the ligand relative to the amount of space it has in the binding site. The generality of this method is particularly appealing. We did not use any prior information about the binding modes in generating our move proposals, which involve random rigid-body rotations, thus this type of move shows promise in broadly sampling different potential binding modes without any prior knowledge. Extensions of this approach however, could potentially make use of other information–for instance from docking–to perform guided rotations targeting specific binding modes. Although NCMC rotational moves can help sample potentially slow binding mode transitions, there are some factors which can pose challenges for this approach. The acceptance rate will likely decrease as the ligand size grows, since a larger percentage of possible random rotations will lead to particularly significant clashes that cannot relax in the span of the move, and favorable binding modes will become correspondingly harder to find by random exploration. Additionally, rotational moves alone will not cover all binding mode possibilities in some cases, but the addition of other Monte Carlo moves (such as translation) could perhaps help address this. Also, rigid body random rotations of a ligand will likely not be as effective for flexible ligands, whose binding modes can be dependent on changes to the internal degrees of freedom such as torsional rotations.

While toluene and iodotoluene binding to T4 lysozyme might not seem to be particularly relevant to drug discovery problems, the problem confronted here actually has considerable similarity to problems encountered in fragment based drug discovery (FBDD). FBDD attempts to find promising leads for early stage drug discovery by studying the binding of very small, often relatively rigid, ligands[46, 87, 33]. These ligands can in fact be of relatively similar size and rigidity to toluene in some cases [89]. Thus, prediction of binding modes of small rigid fragments is in fact of considerable interest. Additionally, even when structural data is available for the binding of fragments, the X-ray crystal structures obtained from FBDD campaigns sometimes have ambiguous electron density for the ligand, making the binding mode difficult to determine [42]. Applying this NCMC method to cases involving rigid ligands could help determine the major binding mode(s) and/or resolve ambiguity in experimental structural data.

1.4.2 Future Work

Future work will focus on exploring other degrees of freedom not just of the ligand, but also the protein. For example, previous studies of T4 lysozyme with p-xylene have shown the VAL101 sidechain orientation greatly impacts which of p-xylene's binding modes are favorable. That valine sidechain is, however, slow to sample and thus would make an excellent test case for NCMC sidechain rotational sampling.

We are also interested in exploring the internal degrees of freedom of the ligand. Performing random rotations of ligand rotatable bonds might be one way to explore the internal degrees of freedom. T4 lysozyme with n-propylbenzene might be suitable for such a test, as the crystal structure shows multiple binding modes due to rotations of the ligand's alkyl tail. Also, rings within a molecule can often be pseudosymmetric, thus necessitating sampling of each ring conformation. These ring flips can be similarly treated by rotating the internal

bonds of the molecule.

The NCMC framework in BLUES has been written to allow straightforward extension to other types of move proposals, such as protein sidechain or ligand torsion rotations as noted above. Even more ambitious move types may be of interest as well. For example, techniques like smart darting [5] could potentially be used to allow ligand hops between different candidate binding sites or binding modes that have been determined in advance.

1.4.3 BLUES

The BLUES package introduced in this work as available free and open-source at `https://github.com/MobleyLab/blues`. We believe this approach shows considerable promise for enhanced sampling of protein and ligand motion and will be useful for a wide range of applications.

1.5 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at (details) and includes a PDF file containing SI Figure 1 (showing work distributions for rotating toluene in lysozyme as a function of the amount of NCMC relaxation), Figure 2 (showing the work standard deviation for toluene in lysozyme as a function of the amount of switching), Figure 3 (showing the dihedral progress coordinate used for 3-iodotoluene), Figure 4 (showing the estimated MSM transition matrix for toluene in lysozyme), and Figure 5 (showing acceptance of NCMC vs standard MC move proposals as a function of dihedral angle/binding mode, given a fixed ensemble of MD snapshots); a `.tar.gz` file containing a set of scripts for running MD and MD/MC and MD/NCMC, simulation run input files for the GROMACS simulations described here, scripts for the OpenMM simulations described, input

topology and coordinate files for all simulations, a README.md file detailing organization, and scripts for MSM construction and PCCA analysis as noted in Methods, as well as a copy of the BLUES code used to generate the data presented here. These figures can also be found in Appendix A

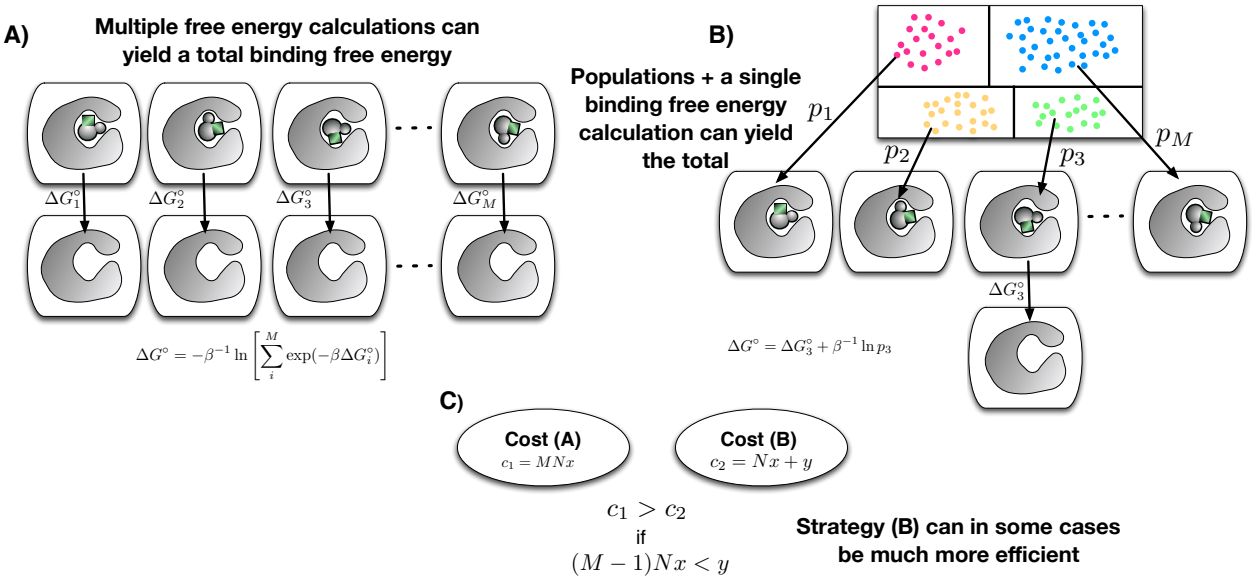


Figure 1.1: **Potential free energy efficiency gains using binding mode populations.** (A) shows calculations of M different effective binding free energy values (ΔG_i°) for each different metastable binding mode of a ligand in a receptor; these effective binding free energies can be rigorously combined to recover the total binding free energy [78]. However, the total computational cost (C) will be MNx where M is the number of binding modes considered, N is the number of intermediate alchemical states used, and x is the length of the simulation used at each alchemical state (assuming each alchemical state uses an equally long simulation). Alternatively, (B) shows how if relative populations (p_i) of different metastable binding modes can be recovered from end state simulations (colored circles, top; each circle represents an amount of simulation time spent in the binding mode, so the populations can be determined from counting time in each mode, with binding modes separated by clustering techniques or any reasonable decomposition of state space [83]), then the full binding free energy can be recovered from the calculation of a single effective binding free energy (here, ΔG_3° is selected for convenience) and the populations of the different binding modes. This approach has a computational cost (shown in (C)) of $Nx + y$, where y is the cost of determining the binding mode populations, which, to be more cost effective than approach (A), requires that $(M-1)Nx > y$.

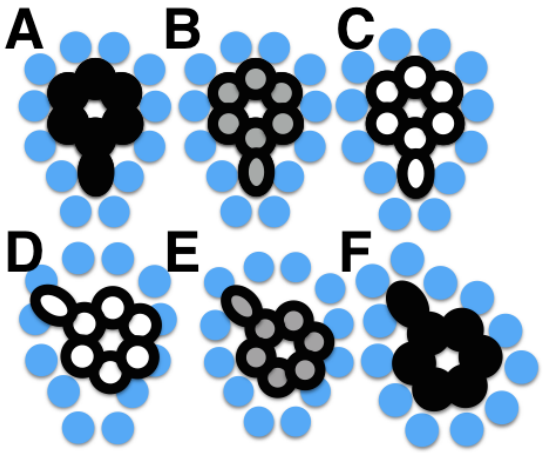


Figure 1.2: **NCMC moves for ligand binding modes.** The blue circles represent the atoms in the binding site, black circles represent the fully interacting ligand, white circles represent the fully non-interacting ligand, and gray circles indicate intermediate levels of interaction. A) The ligand is fully interacting in the binding site. B) The ligand's interactions are partially off, allowing the protein to modestly relax the binding site. C) The ligand's interactions are fully turned off. D) The ligand is randomly rotated around its center of mass; its interactions remain off. E) The ligand's interactions are partially turned on and the propagation steps of NCMC allow relaxation of the rotated binding mode to resolve clashes. F) At the end of the NCMC protocol the ligand is again fully interacting in a new orientation. The NCMC move is then accepted or rejected based on the work performed via Equation 1.4.

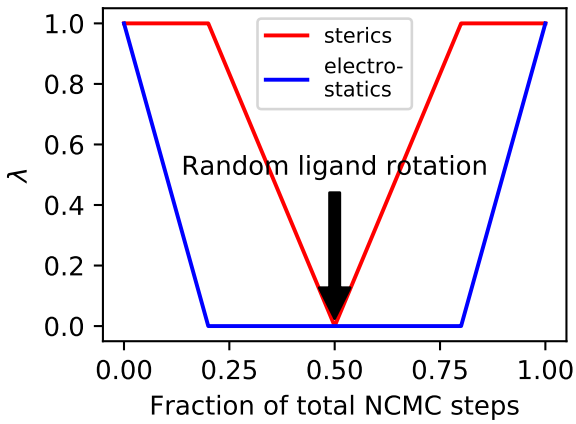


Figure 1.3: **Lambda scaling over the course of our NCMC steps.** The ligand's electrostatic interactions are first turned off, followed by the sterics, until the halfway point (where $n = n_{\text{total}}/2$). The interactions are then turned on in reverse order. This protocol resembles what is typically done for efficient alchemical free energy calculations, such as binding free energy calculations. In particular, the electrostatics are the first to turn off and the last to turn on because having electrostatic interactions present without first turning off the steric interactions can lead to numerical instabilities [123].

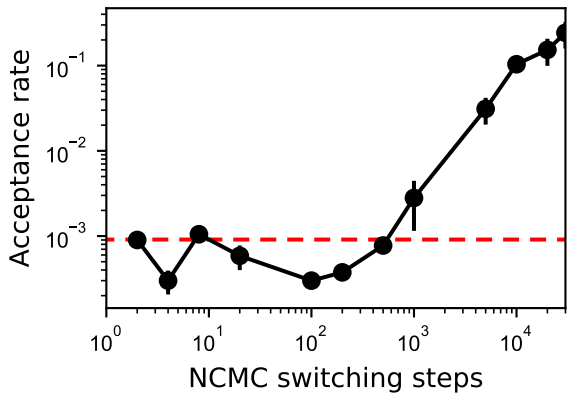


Figure 1.4: **Acceptance probability for toluene as a function of the amount of NCMC relaxation.** The acceptance probability—also referred to as the acceptance rate—is shown on a log scale as a function of the number of NCMC switching steps per cycle, for toluene in the L99A site of T4 lysozyme. It increases dramatically up to 10000 NCMC switching steps per cycle, then increases more slowly, so here we focus on comparing efficiency with other approaches at 10000 steps per cycle. The red dashed line marks the acceptance probability of the instantaneous MC rotation. Error bars are the standard error in the acceptance rate. For trials using 1000 NCMC switching steps and more, the uncertainty was calculated based on blocking [45, 38]. The number of blocks used was the amount that maximized the standard deviations of the acceptance rate across blocks. For trials using fewer than 1000 NCMC switching steps, accepted moves were rare enough that we took the standard deviation across four trials and computed the standard error from that.

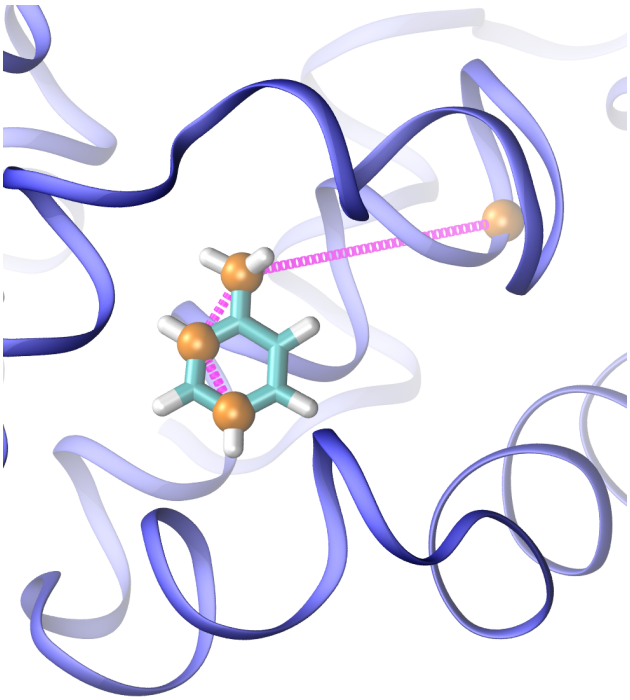


Figure 1.5: **Order parameter used for identifying binding modes of toluene.** Shown is a depiction of the dihedral order parameter used to differentiate toluene's binding modes. The dihedral which we monitor is defined by the alpha carbon of ARG118 and the C1, C5, and C7 toluene atoms, shown in orange in CPK representation. In the image, the atoms involved in the dihedral are connected by a purple line, and the dihedral angle measures rotation around the central dashed purple line. The protein is shown in a blue ribbon representation, and toluene is shown in cyan.

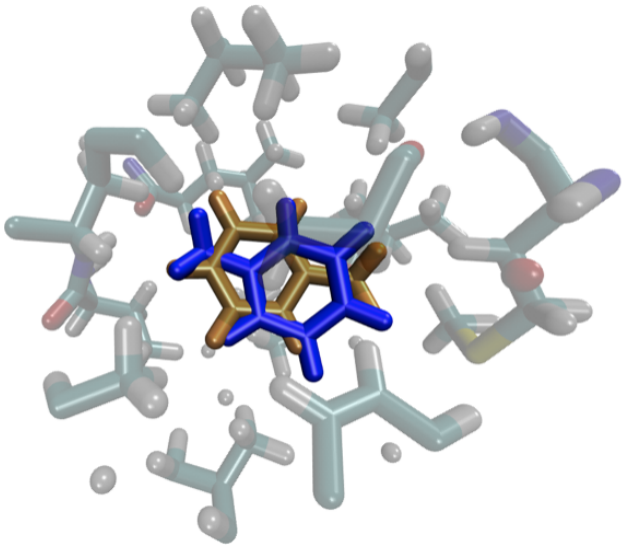
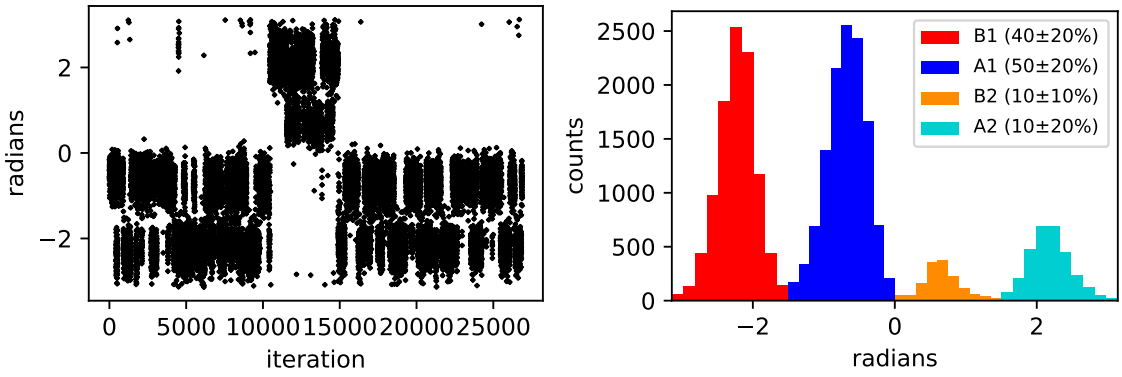


Figure 1.6: **Toluene binding modes.** Toluene exhibits four binding modes. The toluene molecule shown in orange corresponds to the crystallographic binding mode, while toluene in blue corresponds to another binding mode. The other two binding modes come about from the symmetric equivalents of these two binding modes, where the molecule is flipped in the plane of the ring.



(a) Dihedral angle versus time

(b) Population histogram

Figure 1.7: **Toluene binding mode populations from a long trajectory.** (*a*) Dihedral angle (corresponding to binding modes) observed in the initial long trajectory as a function of simulation time (see Sec). (*b*) A histogram plot of the selected dihedral order parameter computed from the trajectory (as shown in Figure 1.5). Labels A1 and A2 correspond to the two different, but symmetry-equivalent populations of the more favorable binding mode. Labels B1 and B2 correspond to the two different symmetry-equivalent populations of the less favorable binding mode. The binding mode fraction of the total population is denoted by the numbers in parentheses in the legend. With enough simulation time the symmetric binding modes should have equivalent populations, which is not the case after over 800 ns of simulation, partly because out-of-plane flips between symmetry equivalent modes are so rarely observed (here, primarily around 350 and 450ns; the A2 and B2 states are at the top in panel (a)). Thus, A2 and B2 end up underpopulated relative to their symmetry equivalent partners A1 and B1. The bootstrapped errors were calculated by breaking the simulation into 5 blocks and calculating the standard error between the populations in each of the 5 blocks.

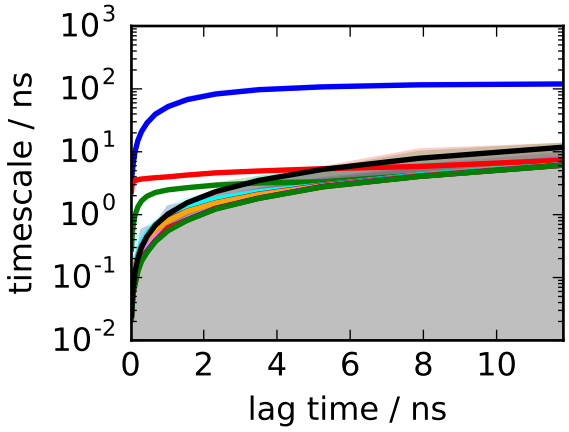


Figure 1.8: **Implied timescales of binding mode transitions.** The implied timescales shown here were calculated from an MSM utilizing all of our MD simulation data of toluene in T4 lysozyme L99A. The black line denotes when the lagtime is equal to the implied timescale; timescales below this line have already relaxed and cannot be estimated accurately; shown here are the 10 slowest implied timescales. Overall, this shows that the slowest timescale in this system (in this case the out-of-plane flip of the ring) has an implied timescale of roughly 100 ns. The gray below the black line indicates when the lagtime is greater than the implied timescale, at which point information about that implied timescale is lost.

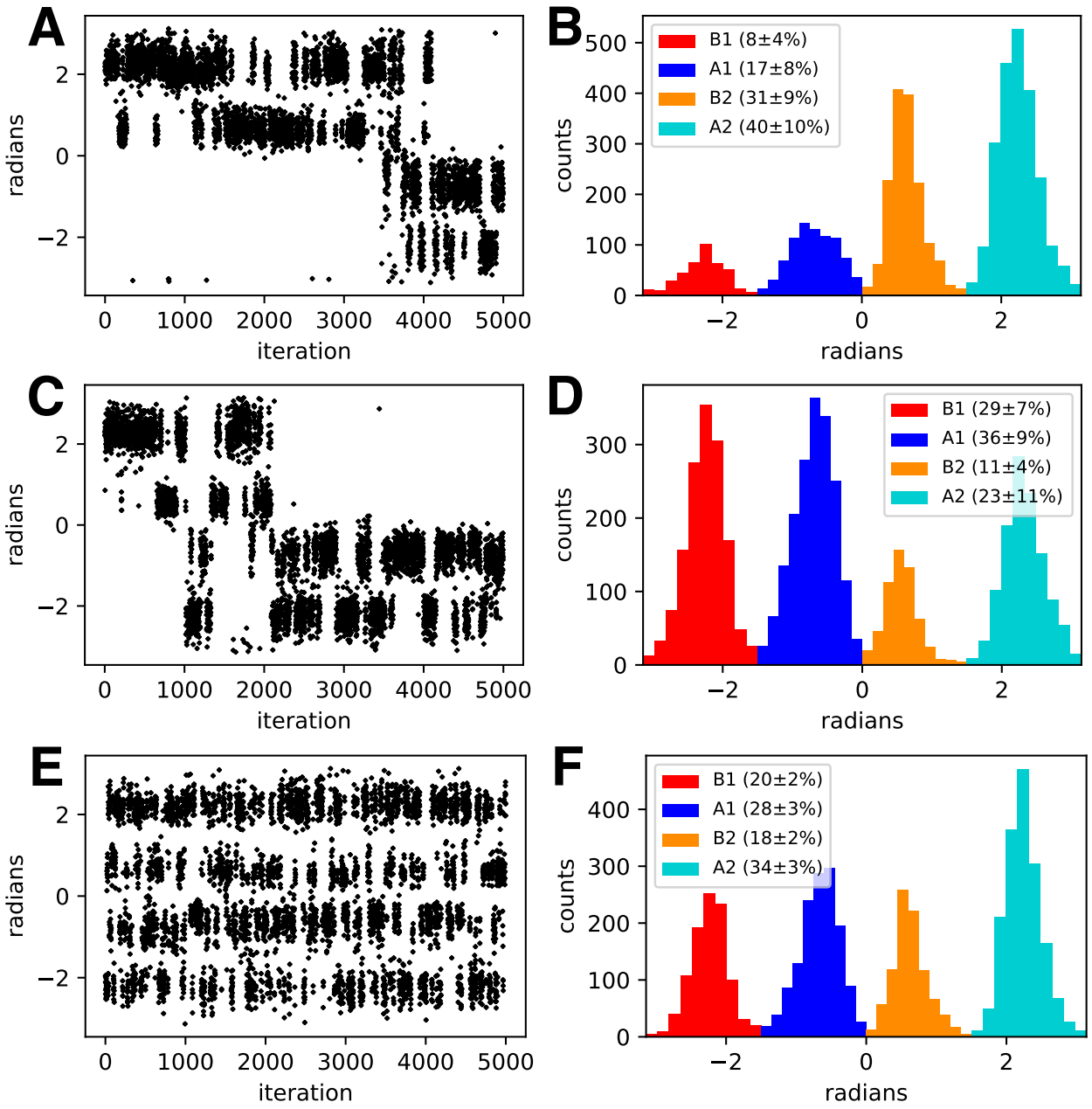
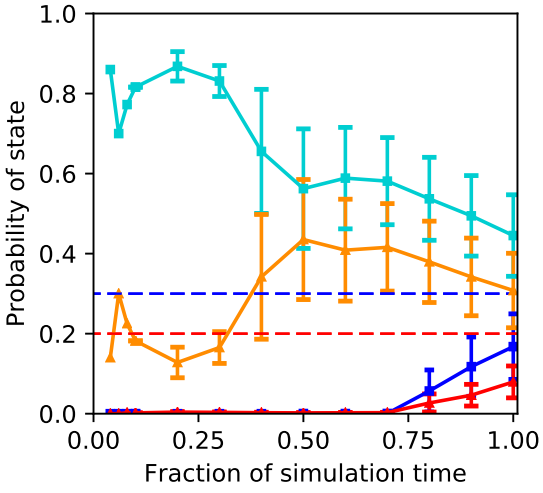
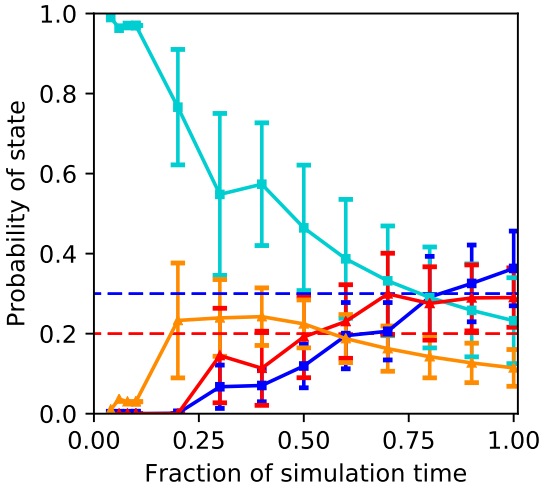


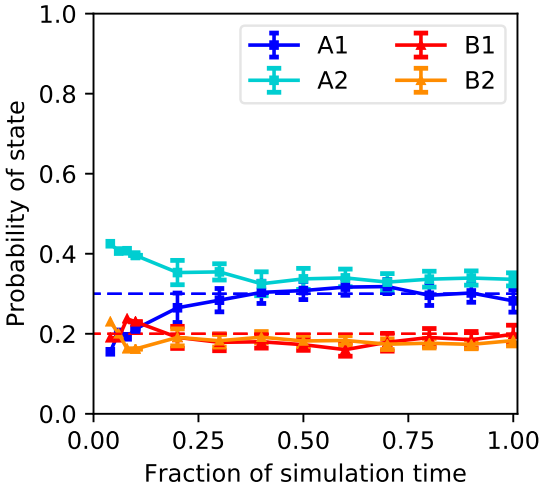
Figure 1.9: **Binding mode sampling of toluene in T4 lysozyme with various methods over 5000 iterations.** This compares the performance of various methods for sampling the four binding modes of toluene in T4 lysozyme over a comparable number of iterations; each iteration corresponds to the same number of force evaluations (20000) for each method. The dihedral angle plotted (on the vertical axis in the left column) discriminates between binding modes, so rapid transitions in this value denote transitions between binding modes. (*A,C,E*) The trajectories from the simulations, showing the the dihedral order parameter plotted as a function of iteration number (loosely, simulation time). The slow out-of-plane flip of toluene results in a transition between the top two states and the bottom two states; relatively few such transitions can be seen in (A) and (C), though more can be seen in (E). (*B,D,F*) Histogram plots of dihedral angles observed in the trajectories, colored by binding mode. Each binding mode's fraction of the total population is denoted by the numbers in parentheses in the legend. Labels A1 and A2 correspond to the two different, but symmetry-equivalent populations of the more favorable binding mode. Labels B1 and B2 correspond to the two different symmetry-equivalent populations of the less favorable binding mode. (*A,B*) MD sampling of toluene in T4 lysozyme. (*C,D*) MC with MD sampling of toluene



(a) MD

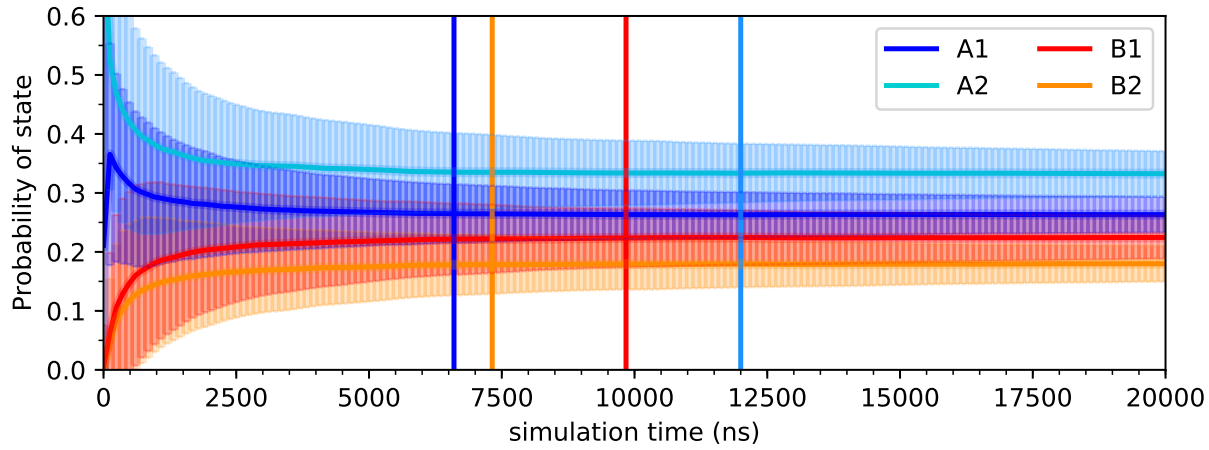


(b) MD/MC

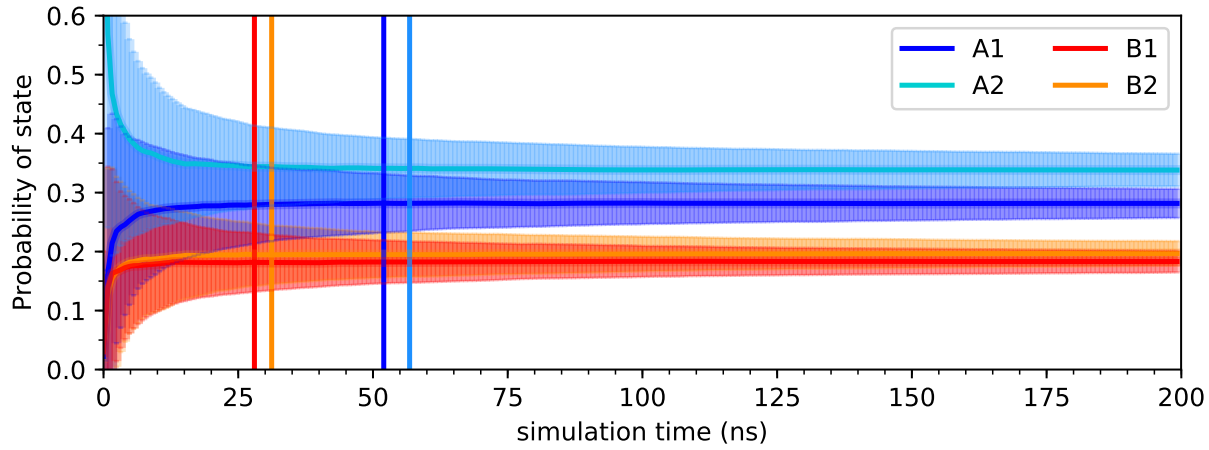


(c) MD/NCMC

Figure 1.10: **Convergence of binding mode populations for toluene.** Shown is convergence of the computed binding mode populations over 5000 iterations (200ns) for toluene in T4 lysozyme L99A. Labels A1 and A2 correspond to the two different, but symmetry-equivalent populations of the more favorable binding mode; each should converge to 0.30, marked by the dashed blue line. Labels B1 and B2 correspond to the two different symmetry-equivalent populations of the less favorable binding mode; each should converge to 0.20, marked by the dashed red line. Over the course of the simulation, the MD/NCMC approach much more quickly to the correct equilibrium distribution of populations than the other approaches. The populations computed by BLUES are within uncertainty of the true result well before 10% of the total simulation time, whereas with MD and MC the populations are not until much later if at all.

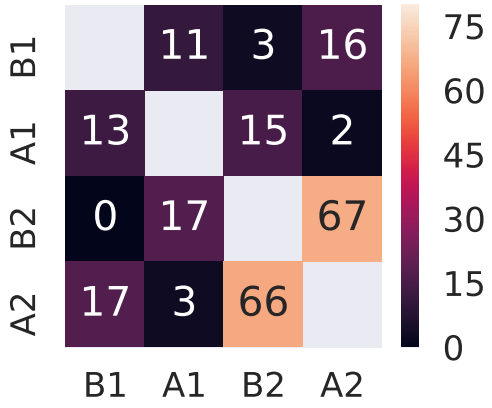


(a) MD

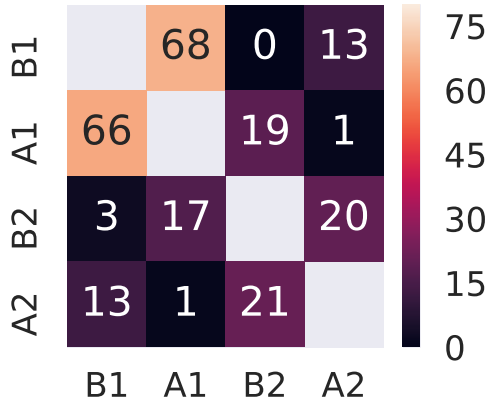


(b) MD/NCMC

Figure 1.11: **A model of the convergence of binding mode populations for toluene in T4 lysozyme L99A.** The transition matrices from the MSM and MD/NCMC simulation were used to estimate the convergence of binding mode populations as a function of time for a hypothetical simulation starting in state A1. We ran 1000 trials in each case. For each trial we propagated the transition matrix by selecting a new state to transition to at each timestep with probabilities given by the transition matrix as described in the text. Heavy lines show the mean population estimated over the trials, and the lighter shaded regions give the standard deviation over trials, indicating the region within which a typical single simulation would usually fall. Vertical bars denote the point at which the standard deviation of each estimated population first falls below 5%. (*a*) The statistical model estimated from the MSM which shows that it takes approximately 12000 ns for the standard deviation in the slowest converging population to get below 5%. (*b*) The statistical model estimated from the MD/NCMC simulation which shows that it takes approximately 60 ns for the standard deviation in the slowest converging population to get below 5%. In both cases, because the transition matrices were estimated from relatively short simulations, the populations converge to a steady state but have some error due to the underlying transition matrices. Together, (a) and (b) demonstrate that MD/NCMC results in dramatically faster (more than two orders of magnitude) convergence of populations as a function of simulation time compared to MD alone.



(a) MD



(b) MD/MC

(c) MD/NCMC

Figure 1.12: **Binding mode transitions for toluene.** Shown is the transition matrix counting the number of transitions between binding modes for toluene in T4 lysozyme L99A over 5000 iterations (200ns), for the different sampling methods. Labels A1 and A2 correspond to the more favorable binding mode. Labels B1 and B2 correspond to the less favorable binding mode. A1 and A2 comprise a symmetry-equivalent pair, as do B1 and B2, but to transition between states in a symmetry-equivalent pair (A1 to A2, or B1 to B2) requires an out-of-plane flip. Transition counts to the same binding mode (the main diagonal of the matrix) are omitted for clarity. Here, in general, hotter colors are better as they indicate more transitions between binding modes. (*a*) Transitions of the MD simulation. The total number of transitions is 242. (*b*) Transitions of the MD/MC simulation. The total number of transitions is 230. (*c*) Transitions of the MD/NCMC simulation. The total number of transitions is 497. Here, it can be seen that in the MD case, only the A2 to B2 and B2 to A2 cases have more than 30 transitions, because the simulation mostly remained stuck in these two states without flipping out-of-plane (Figure 1.9) and a similar effect happened in the MD/MC case but for A1 to B1. In contrast, in the NCMC case, all transitions occur more than 30 times because out-of-plane transitions are also relatively frequent.

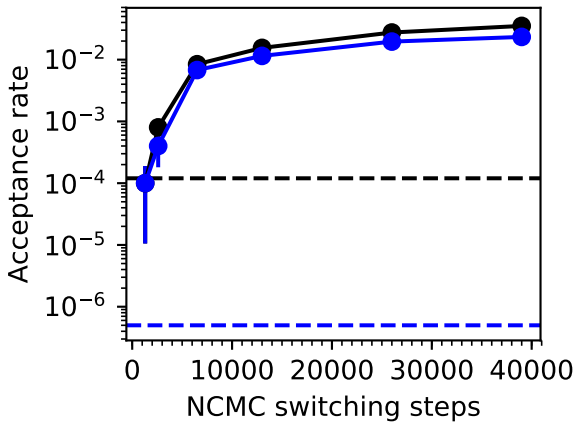
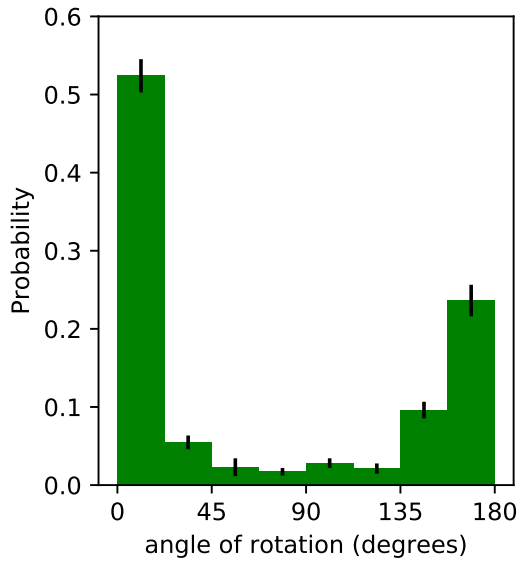
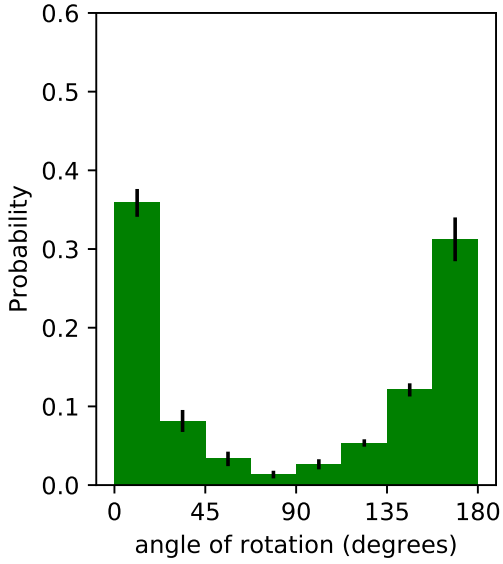


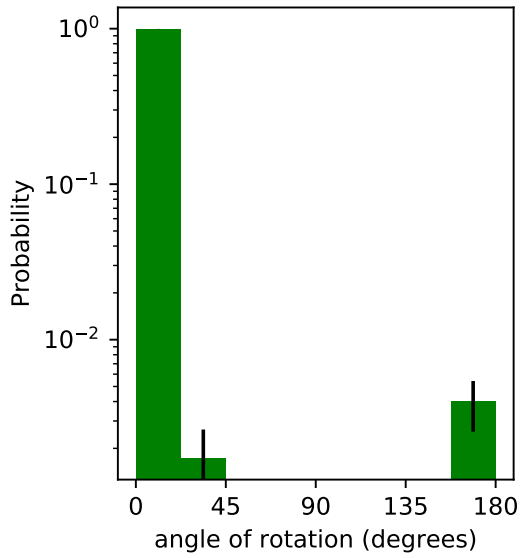
Figure 1.13: **Acceptance probability for iodotoluene as a function of the amount of NCMC relaxation.** Shown is the acceptance probability for rotational moves of 3-iodotoluene in the L99A site of T4 lysozyme, as a function of the number of NCMC switching steps, analogous to Figure 1.4 except that this test uses a fixed set of MD snapshots as a basis for move proposals, as described in the text. Here, we observe that overall acceptance (black line) increases dramatically up to 10000 NCMC switching steps per cycle, then increases more slowly. The black dashed line marks the acceptance probability of instantaneous MC rotations, given the same set of MD snapshots as starting points. The solid blue line denotes the acceptance probability of *substantial* rotations, those larger than 45 degrees, and the dashed blue line indicates the overall acceptance probability of instantaneous MC rotations from the same set of snapshots. Thus, NCMC does only modestly worse at sampling substantial rearrangements than sampling all rearrangements, whereas MC has orders of magnitude lower acceptance of substantial rearrangements.



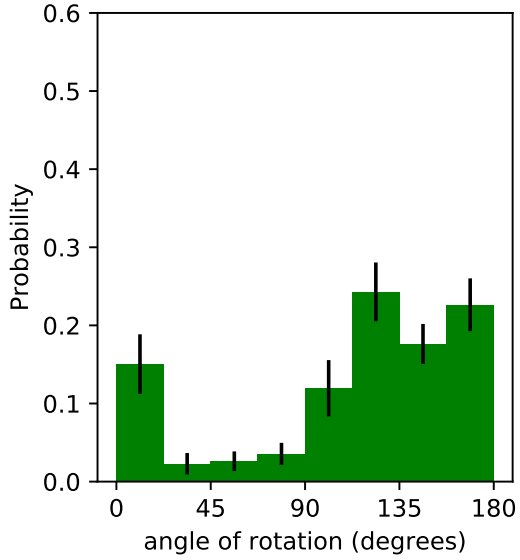
(a) Toluene with MC



(b) Toluene with MD/NCMC



(c) 3-iodotoluene with MC



(d) 3-iodotoluene with NCMC

Figure 1.14: **Rotational distribution of accepted moves for toluene and 3-iodotoluene in T4 lysozyme.** Shown are the distribution probabilities of accepted rotational moves, with standard Monte Carlo and with NCMC, for toluene (top) and the bulkier iodotoluene (bottom). Results come from 10000 MC iterations of 10 attempts each (a and c) or 10000 NCMC iterations (b and d). With NCMC and BLUES, we are interested in improving the decorrelation time of ligand binding modes, so an important metric is not just the acceptance ratio, but how many *substantial* rotational moves are accepted. For toluene, which is relatively small compared to the available volume of the binding site, standard Monte Carlo (a) and NCMC (b) yield relatively similar numbers of large moves accepted (though NCMC has better acceptance of intermediate moves, presumably due to the additional relaxation). However, iodotoluene is substantially bulkier, and it is difficult to rotate it in the binding site without at least some amount of relaxation, so the acceptance rate for MC moves is lower (Section 1.3.2) and the number of *significant* rotations is dramatically lower (c), with virtually no rotations larger than 22.5 degrees observed; for panel (c) we use a log scale to make it apparent that *some* significant rotations were observed.

Chapter 2

Sampling Conformations Using Molecular Darting

Sampling multiple binding modes of a ligand in a single molecular dynamics simulation is difficult. A given ligand may have many internal degrees of freedom, along with many different ways it might orient itself a binding site or across several binding sites, all of which might be separated by large energy barriers. We have developed a novel Monte Carlo move called Molecular Darting (MolDarting) to reversibly sample between predefined binding modes of a ligand. Here, we couple this with nonequilibrium candidate Monte Carlo (NCMC) to improve acceptance of moves. We apply this technique to a simple dipeptide system, a ligand binding to T4 Lysozyme L99A, and ligand binding to HIV integrase in order to test this new method. We observe significant increases in acceptance compared to uniformly sampling the internal, and rotational/translational degrees of freedom in these systems.

2.1 Introduction

Structure-based drug design allows for rational design of ligands, as computational methods can help predict desired qualities of a potential ligand prior to its synthesis [36, 55, 113, 68]. However, an understanding of ligand binding modes is often viewed as critical for structure-based design [133, 74, 56] yet binding modes are not necessarily well known before compounds are made and tested [85, 129, 27].

Thus, many computational methods seek to predict ligand binding modes. Several such methods for binding mode prediction are available, but overall computational prediction of binding modes is a difficult problem [61, 85]. One of the most commonly used methods for binding mode prediction, docking, is able to sift through millions of compounds efficiently, however, docking does not tend to do well at predicting the true binding mode [129]. On the other end of the spectrum of computational cost are free energy simulation-based methods, which are very promising for structure-based design and are attracting tremendous interest from industry [75, 26, 107, 122].

However, computational methods for studying binding have their limitations. Free energy methods for predicting binding affinity need to start close to, or sample the correct binding mode in order to offer accurate free energy predictions [3, 77, 26, 58]. This reliance on the starting position can cause issues; since the binding mode of a novel ligand has to be predicted and is typically slow to sample in a simulation [109], adequate sampling of the ligand's motion in the binding site can be challenging. Even in the case of a congeneric series of molecules binding to the same target, the binding mode of the ligands can differ [85, 66].

In order to circumvent some of these shortcomings of MD-based methods, we previously developed a mixed MD/nonequilibrium candidate Monte Carlo (NCMC) based method, and implemented it in a package called Binding modes of Ligands Using Enhanced Sampling (BLUES) [43]. Typically, Monte Carlo (MC) moves have difficulty achieving high accep-

tance rates in condensed-phase systems because of tight packing, allowing for only small perturbations to be performed on a system. NCMC provides a framework where a larger, instantaneous MC move can be broken up into a series of smaller perturbations. Between each perturbation the system is allowed to relax by applying dynamics. This process is repeated a number of times and the whole move is accepted or rejected based on the total work done during the perturbation steps. In BLUES we use NCMC moves to alchemically remove the interactions of a ligand and then reinstate them over the course of some number of steps (N). At the start of reinserting the ligand, a MC move can also be performed to further improve binding mode sampling. By slowly removing and regrowing the ligand, we can insert the ligand into a new binding mode and allow the rest of the system to slowly relax in response to the ligand's motion, potentially leading to higher rates of acceptance compared to instantaneous MC moves.

As noted, an MC move can be performed at the midpoint of the NCMC protocol. In our original paper describing the BLUES method, the only such move offered was a center of mass rotation of the ligand. In subsequent work, the MC moves available were further expanded to include protein side-chain torsions [12] as well as selected torsions of the ligand. [12, 105]

These types of moves are helpful in generating small perturbations of the ligand's binding mode, but ideally we would like to be able to generate binding mode predictions and sample between those directly. Generally, proposing reasonable candidate binding modes is a relatively easy task, since docking methods tend to do a good job at generating plausible binding modes, but are poor at ranking these binding modes [18, 129, 128]. In many cases, such poses can be equilibrated via MD simulations to find a variety of different stable or metastable binding mode candidates. [66, 35, 104, 69].

While some methods can improve sampling of a ligand's internal degrees of freedom, we are not aware of any current MC method which can efficiently hop between potentially disparate predefined ligand binding modes in a way that preserves detailed balance.

Techniques such as Rosenbluth sampling [102], or configurational bias Monte Carlo [34] are sampling methods originally applied to flexible molecules to grow and arrange polymers favorably, but these methods do not offer a way to directly sample between two specific conformations of a molecule.

Distance Geometry is another technique used to perform conformational analysis of ligands. methods [116]. In this technique the atoms of a molecule are randomly placed and then minimized to generate a new structure. Like configurational bias MC, however, distance geometry methods do not satisfy detailed balance since they depend on a minimization step.

To more efficiently sample binding moves, we have developed a new Monte Carlo based method to directly sample transitions between candidate poses–which may even be in different binding sites. Furthermore, we have implemented this method in the BLUES package in connection with our previous BLUES NCMC-based method in an attempt to directly sample multiple binding modes in protein systems.

2.2 Theory and computational methods

Here, we first describe the background and motivation of the method we implement here, then move on to discuss technical details of its implementation and how it was tested.

2.2.1 Smart Darting allows for selective sampling between minima

Our novel Monte Carlo method is a logical descendant of another Monte Carlo sampling method called Smart Darting Monte Carlo [6]. The general process of Smart Darting involves defining two key pieces of information. The first piece we need to specify is a set of "darts", which represent different configurations of the system that are of interest. The second piece

we need to specify is a set of parameters (and their ranges), which correspond to and define each of those darts, in order to specify the boundaries associated with each conformation.

To explain Smart Darting in more technical terms, a set of darts $d_0, d_1 ... d_j$ are first specified. Each of those darts corresponds to a particular set of microstates (i.e. a metastable binding mode which was given as input) each of which is defined by a set of parameters k_0, k_1,...k_n, with each parameter k_i having an associated range r_{k_i}. Each parameter refers to a quantity that defines that microstate–such as a torsion angle, or some distance measurement, such as the distance between two atoms. The range should be the same for each parameter k_i, (which is necessary to preserve detailed balance, or the acceptance criterion needs to be altered). When a given parameter is within its associated range, we refer to it as being within that parameter region. These parameters (and the size of the parameter range) are user-defined input and should be designed to cover the typical value ranges of those parameters, which can be determined by example by running short exploratory/equilibration simulations. When attempting to make a Smart Darting Monte Carlo move, the parameters are evaluated (the current value of that parameter is checked) for each dart. When the parameter is evaluated, if the current configuration is within the parameter regions r_{k_i} for all r_k of a given dart—which we refer to as being within the dart—then the system can jump to another set of parameters with equal probability. In the process of jumping to the new configuration, a new k_0, k_1,...k_n are each generated–either uniformly between the ranges for a given r_{k_i} or deterministically through some one-to-one mapping from the old k_0 to the new k_0. Additionally, to maintain detailed balance, no Smart Darting move can be performed on a system if the system is within the range of multiple darts.

2.2.2 Molecular darting moves use internal coordinates as part of move proposals

In our novel Smart Darting-inspired methodology, called Molecular Darting (MolDarting), the parameters that define a dart are defined by the internal torsions of the molecule, as well as a translational and rotational distance to a given configuration. The internal coordinates are described by a Z-matrix, which describes the molecule's configuration in terms of internal bond distances, angles, and dihedrals. For this case of MolDarting, we assume the bond and angle internal coordinates are invariant between ligand conformations, and that the dihedral internal coordinates are independent of one another. The translational distance is defined by the Euclidean distance between the first atom of the Z-matrix of the current configuration and the corresponding atomic positions of the given dart. The translational distance is defined by the Euclidean distance of a given configuration to each reference position of the first atom in the Z-matrix. We used Chemcoords [131] to generate the internal coordinates for our molecules of interest. The rotation matrix of the first three Z-matrix atoms of the ligand is calculated to each of the first three Z-matrix atoms of the references. The rotational distance is calculated by Eq 2.1, where R is the rotation matrix.

$$\theta = \arccos(\frac{Tr(R) - 1}{2}) \tag{2.1}$$

When using MolDarting on a protein-ligand system, it's necessary to first account for the overall rotation and translational changes for the protein-ligand complex in regards to the reference darts. To account for those rotational and translational changes, heavy atoms of the residues around the binding site are chosen. When checking if the current configuration is within the rotational and translational regions, the chosen binding site residues of the selected dart are superposed to the same binding site residues of the current pose, then the rotational and translational distances are calculated.

When MolDarting between binding modes, the proposed internal coordinates from MolDarting are uniformly chosen anywhere inside the newly selected internal coordinate region (Figure 2.1). The rotational and translational motions are deterministically updated by assessing the displacement from the starting pose to the center of each of their respective regions and then applying those same displacements again after it is MolDarted (Figure 2.2, Figure 2.3).

When combining MolDarting with BLUES, an additional step is added to the MolDarting procedure. We found that these restraints were needed because when the ligand steric interactions are diminished, it is more labile inside the binding pocket and can frequently end up outside the darts. To reduce the lability of the ligand, an orientational restraint, also known as a Boresch-style restraint [10] is applied to the first three ligand Z-matrix atoms, relative to three reference atoms in the protein. This restraint restricts the orientation relative to the binding site via restricting one distance, two angles and three torsions, and involves three reference atoms in the ligand and three in the receptor. Here, we scale this restraint with the lambda parameter that controls the electrostatics and sterics; when the ligand is fully non-interacting, the restraints are in full effect (Figure 2.4). To maintain detailed balance when applying restraints, before the NCMC move occurs the we check if the ligand is currently within a dart; if it is then the orientational restraints associated with that pose will be turned on over the first half of the NCMC move. If the ligand is not within the same dart as at the start of the move, then the move is rejected.

Subsequently, after the MolDarting move is performed, the restraints corresponding the new pose are turned on, and the previous pose's restraints are turned off. Finally, after the NCMC move occurs, the parameters are evaluated again to see if they are within any dart. The modulation of steric, electrostatic and restraint interactions over the course of the NCMC move are illustrated in Figure 2.4, and the overall procedure is illustrated in Figure 2.5.

If the ligand is in a different pose than the pose the ending restraints were associated with, then the move is automatically rejected, since such a move would not be reversible. Otherwise

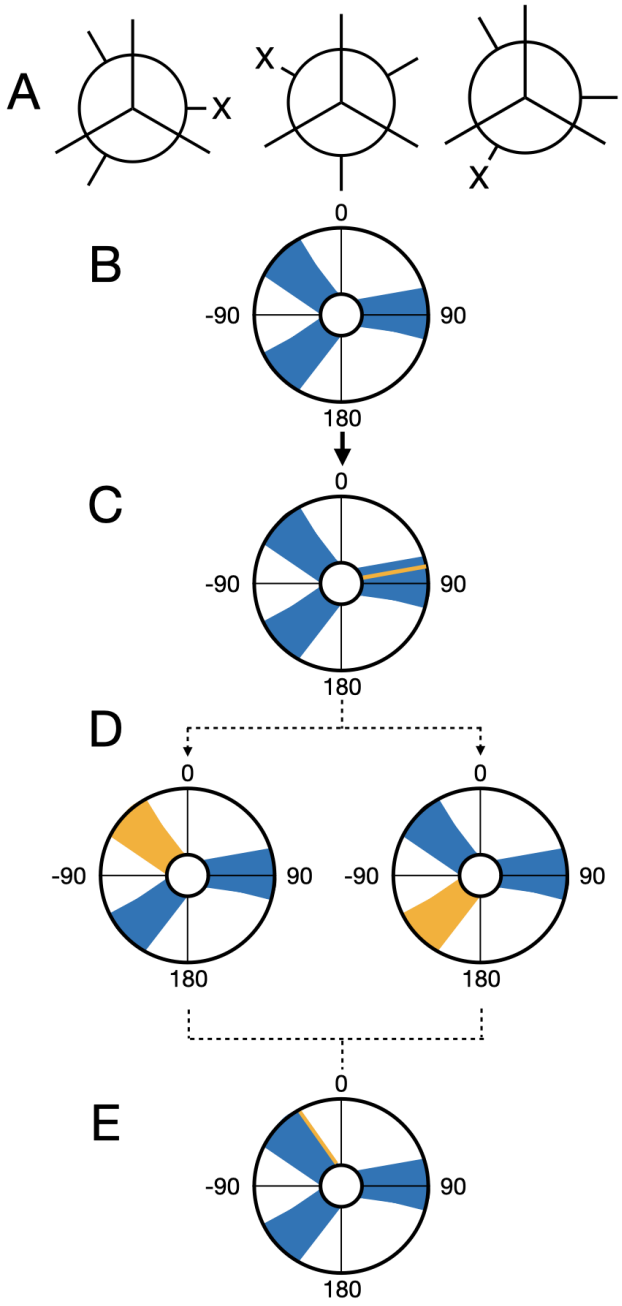


Figure 2.1: **Dihedrals are uniformly sampled during MolDarting.** We illustrate how we perform our rotational darting moves using a rose plot representation of a dihedral angle (in degrees) as an example. The dihedral regions are represented by the blue areas, and the current dihedral angle is represented by the yellow line/areas. In this example, there are three total darts, each with an associated region. (A) The Newman projection of a hypothetical ligand illustrating three different stable conformations. (B) A representation of the three dihedral regions for the three conformations. (C) When a particle is within a dihedral region then a darting move can be performed. (D) When MolDarting the dihedrals, the new dihedral is selected uniformly from a region the dihedral is not currently in (shown in yellow). The arrows refer to the two potential outcomes of the MolDarting move in which the ligand is darted to a new configuration. (E) One of the other dihedral regions are chosen randomly (with equal probability) to be MolDarted, and then a new dihedral is chosen randomly from the chosen region, resulting in a new configuration.

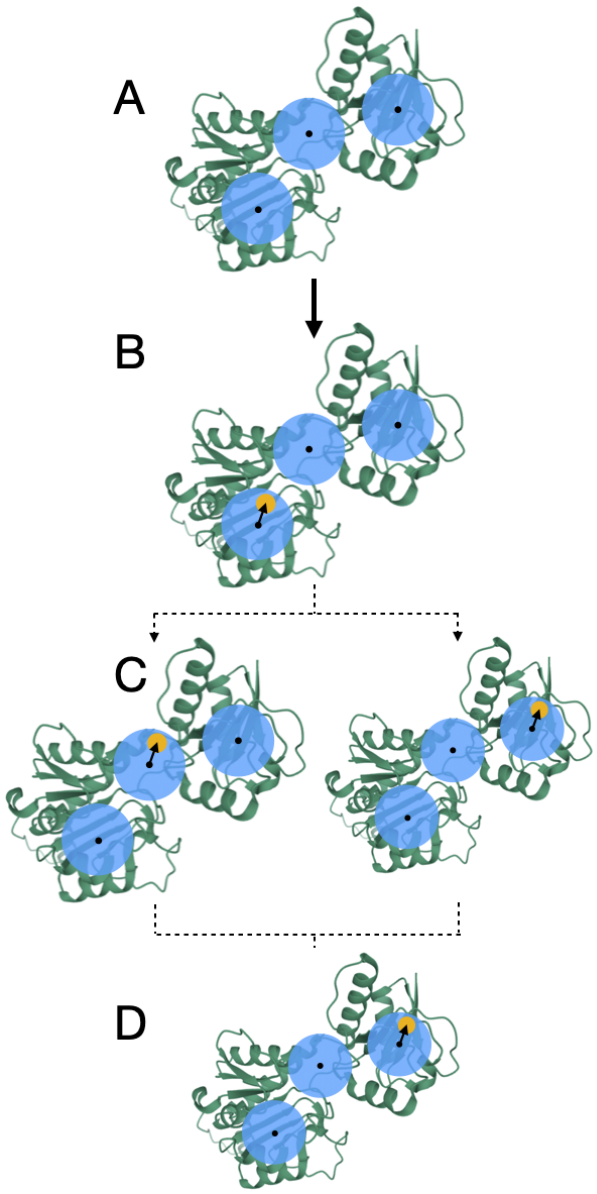


Figure 2.2: **Translations are handled deterministically during MolDarting.** We illustrate how we perform our translational darting moves using a 2-dimensional translational region as an example, with a single particle, (that can represent an atom of a ligand, for example) that will be Moldarted. The translational regions are represented by the blue circle, with the center of each translational region represented by a black dot, and simplified molecule represented by yellow circles. In this example, there are three total darts. (A) A representation of the three rotational regions used. (B) When a particle is within a translational region, the vector from the particle's center, to the translational region's center is calculated (represented by the arrow). (C) When MolDarting the vector calculated in (B) is applied to the center of each other translational region to determine the particle's new position. The dotted arrows refer to the two potential outcomes of the MolDarting move in which the ligand is darted to a new configuration. (D) One of the new reference regions are chosen randomly (with equal probability) to be MolDarted, resulting in a new configuration.

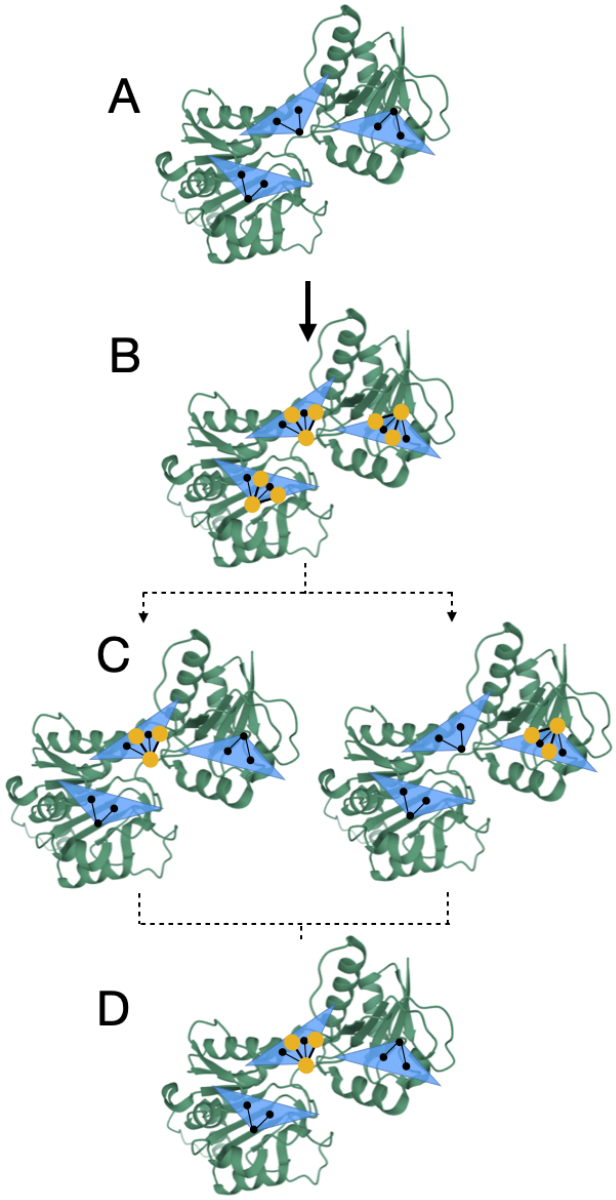


Figure 2.3: **Rotations are handled deterministically during MolDarting.** We illustrate how we perform our rotational darting moves using a 2-dimensional rotational region as an example, with a single molecule that will be moved via MolDarting. The rotational regions are represented by the blue triangle, with the center of each rotational region (which was defined by some reference pose) represented by the three black circles connected by black lines, and the ligand in our simulations represented by the yellow circles connected by yellow lines. In this example, there are three total darts, each with an associated rotational region. (A) A representation of the three rotational regions used. (B) When a particle is within a rotational region the rotation matrix is calculated from the current positions to the reference positions. (C) When MolDarting, the rotation matrix calculated in (B) is applied to the reference positions of each other rotational region to determine the molecule's new position. The dotted arrows refer to the two potential outcomes of the MolDarting move in which the ligand is darted to a new configuration. (D) One of the new reference regions are chosen randomly (with equal probability) to be MolDarted, resulting in a new configuration.

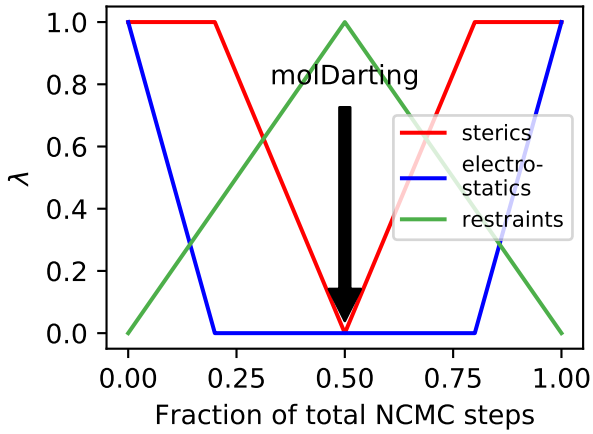


Figure 2.4: **Restraints are included in the NCMC switching protocol. In order to keep the ligand in the binding site while the ligand's interactions are off, an orientational restraint is used which corresponds to the dart that the ligand is in at the beginning of an NCMC move proposal. At the middle of the NCMC protocol, a MolDarting move is performed, and the restraint switches to a new orientational restraint corresponding to the new dart, which is subsequently turned off throughout the rest of the protocol.**

the protocol work (the work that is done over the course of the NCMC move) determines whether the NCMC move is accepted or rejected. The application of the restraint is taken into account in the work done during the course of the NCMC move.

Taking into account the major degrees of freedom of the molecule allows reversible MolDarting moves between different potential ligand binding modes, not only with different ligand conformations, but potentially even in separate binding pockets.

2.2.3 We tested Molecular Darting on three different systems

To validate and explore the potential of MolDarting, we look at three different system with different requirements needed to sample binding modes. The first system explored is an alanine-valine dipeptide. While not typically considered a ligand, this peptide is a simple model system which exhibits three different stable conformations that vary by an internal

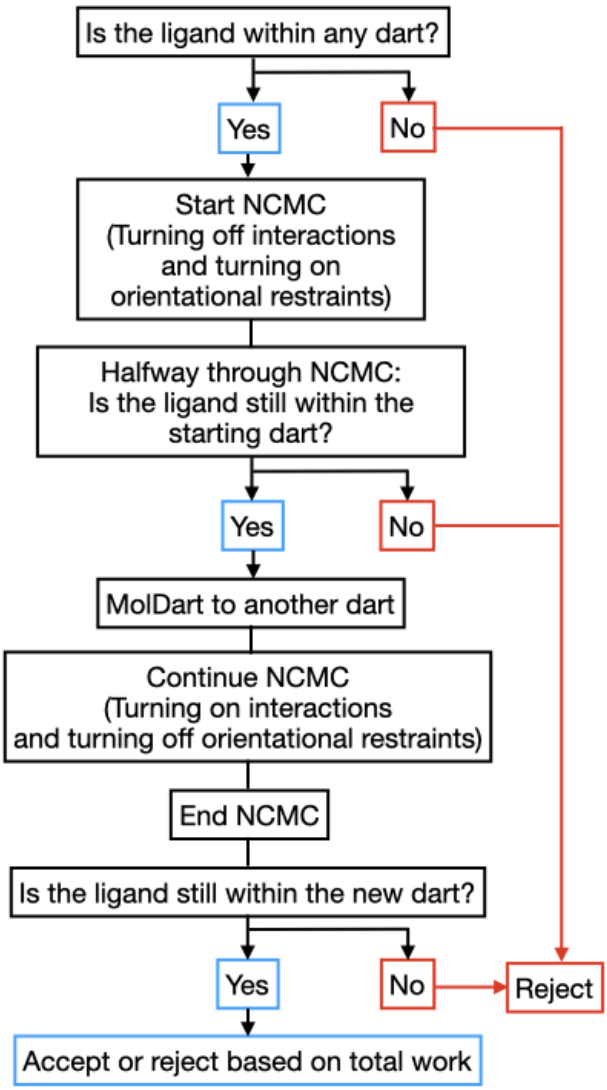


Figure 2.5: **Adding restraints with NCMC and MolDarting requires additional consideration. When restraints are used alongside NCMC and MolDarting, it's necessary to take into account several additional factors, which are illustrated by this flowchart and elaborated further in Section 2.2.2.**

torsion and can be slow to sample through plain MD [12]. It also is a good test system for the darting approach we develop here, as MolDarting can be applied to any selected object in our system, not just a ligand. Here, since sidechains play an important role in ligand binding it is also important to be able to sample the rotamers in a binding site. The second system we look at with MolDarting is T4 lysozyme L99A with toluene bound, where the binding modes varies by rotation and translation. The final system we look at is HIV integrase with a variety of ligands bound. HIV integrase is an interesting test system because it has multiple binding sites where ligands can bind, and has proven difficult for binding mode predictions in a previous blind challenge [85], and we would like to test whether MolDarting can directly sample the binding modes in each binding site.

2.3 Methodology

2.3.1 System preparation

Alanine-valine dipeptide system setup

An alanine-valine dipeptide system was created using tleap from AmberTools 16 [16]. The amber99SBILDN forcefield was used for the protein parameters. Simulations were carried out at 300K with a Langevin integrator using a 0.002ps step size in implict OBC2 solvent [92] using OpenMM version 7.3 [32]. Nonperiodic cutoffs were used, with the hydrogen bonds constrained and a 1/ps friction applied. The peptide's CA, N, and O backbone atoms were restrained using a restraint of 25 kcal/(mol·angstrom^2) based on their starting conformation.

To prepare for MolDarting between the different stable rotameric states for this dipeptide, we initially ran a 100 ns simulation to identify the dihedral minima of the system. From this simulation, we found three stable valine rotamers, with dihedral maxima at approximately

-170, -65, and 53 degrees. These dihedrals was calculated by measuring the dihedral angle between the CA, CB and CG1 atoms of valine on the alanine-valine dipeptide atom using `MDTraj 1.9.3` [72].

From the three maxima, regions were chosen so that the region size encompassed 95% of the probability density associated with that dihedral maximum, estimated from a kernel density approximation with a 0.2 bandwidth and a Gaussian kernel.

Simulations of the alanine-valine system were performed using BLUES for 150000 iterations, with each iteration consisting of 1000 steps of MD and an instantaneous MC move consisting of either a sidechain rotation using the `SideChainMove` class or a MolDarting move using the `MolDartMove` class. The code used to run these simulations can be found in the SI.

Populations of the three dihedral maximums were separated based on the following bin definitions: from (-120,-40] defined one bin (with a maximum at 68 degrees), from [20,100] defined another bin (with a maximum at 68 degrees) and a third bin is discontinuous and is defined between [-180, -120] and [115,180] (with a maximum at 180 degrees).

T4 lysozyme/toluene system and simulation setup

Here, we used the same T4 lysozyme and toluene system and parameters for NCMC from our previous work [43]. The only difference in our simulation protocol was that now a MolDarting move was performed instead a random center of mass rotation. For the MolDarting move, a rotational dart of 40 degrees was defined, using two poses of the non-symmetrically equivalent binding poses as a reference. A Boresch restraint with a force constant of $3kcal/(mol * angstrom^2)$ for the radial component and $3kcal/(mol * rad * *2)$ for the angular and dihedral components was used with the first three internal coordinate atoms of toluene as chosen by ChemCoords (being the C6, C4, and C5 atoms respectively of the toluene molecule) and the CA atoms of PRO85, ALA98, and LEU117 using the Yank's

`BoreschRestraint` class to implement the restraints with the provided atoms from the receptor as the `restrained_receptor_atoms` and the ligand atoms as the `restrained_ligand_atoms` arguments for the class [21].

HIV integrase system setup

We used the 4CHY pdb file as the basis structure for our study to serve as a uniform starting point for docking and equilibration. Omega from Openeye [48] was used to generate the conformers for the 4 ligands from the pdb files of 4CHY, 4CGD, 4CHZ, and 4CJV [95], and Fred was used to dock the compounds in the three different binding sites [1]. The highest scoring poses from docking was used, and to generate a diverse set of structures, root-mean-square deviation (RMSD) centroid clustering was performed on the poses, and the most diverse poses retained, to promote pose diversity. To further elaborate on the clustering procedure, the first centroid was defined using the top-scoring docking pose, and the subsequent centroids were chosen which were the greatest RMSD distance away from the other existing centroids for that binding site. Clustering of poses were done separately for each binding site, and the two poses with the centroids furthest from the top scoring pose were used as reference poses for use with MolDarting, for a total of three poses per binding site. Antechamber was then used with the AM1-BCC method [52, 16] to assign partial charges to the molecules.

Finally, Amber was used to add missing sidechains, heavy atoms, and hydrogens to the protein, with the parameter set from ff14SB used for the protein [16]. Because the binding sites of HIV integrase are solvent exposed, we chose to use OBC2 implicit solvent model [92] to reduce the amount the system has to respond to solvating and desolvating the binding sites in response to the ligand being removed and inserted when MolDarting the ligand. Equilibration MD simulations were performed at 300K for 1 ns for each binding pose. The positions of this equilibration trajectory were saved every 10,000 steps.

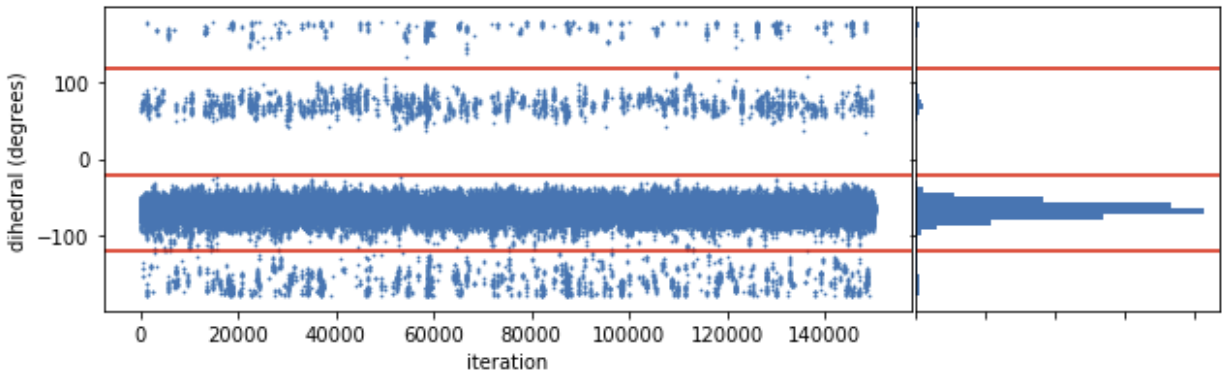
Unless otherwise noted, the simulation settings were the same as alanine-valine dipeptide system. The equilibration simulation trajectories were also used to define the dihedral regions. Kernel density estimation (KDE) was performed on the dihedral internal coordinates from the trajectory with a bandwidth of 0.5. From this, the maxima in the dihedral KDEs were identified. The maxima that the dihedral was closest to at the end of equilibration was used to determine the start of the region for that dihedral. The width of the dihedral regions were determined were made account for 95% of the probability density estimated by KDE. The width was calculated by first finding the total probability density contained within a maximum, and then expanding the width of the region starting at the maximum until 95% of that maximum's probability density was covered by the region. During MolDarting simulations, restraint atoms were automatically chosen from the heavy atoms within 10 angstroms of the ligand using Yank [21].

2.4 Results

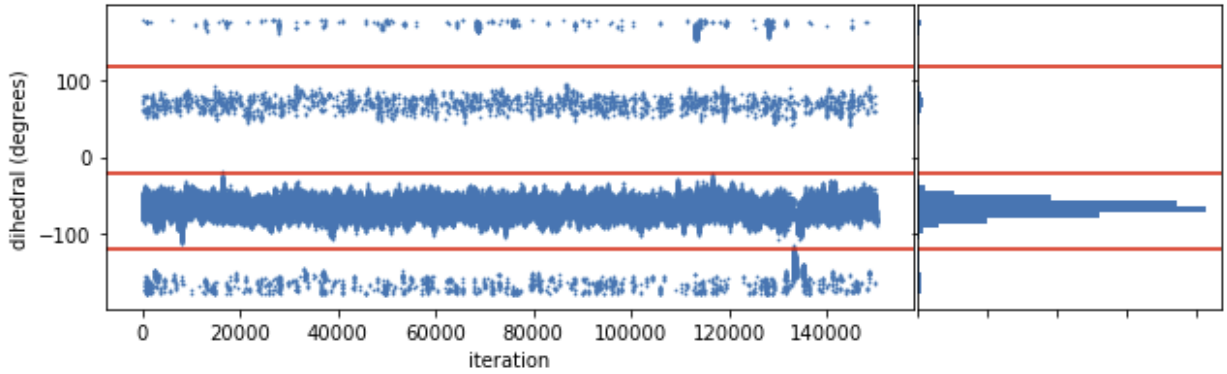
2.4.1 We validated the internal coordinate sampling of our method against uniform dihedral sampling of the valine-alanine dipeptide.

Molecular Darting samples the three dihedrals of valine-alanine dipeptide efficiently

We assessed the ability of MolDarting to sample the sidechain torsion of the valine-alanine dipeptide in implicit solvent. We also the compare the sampling efficiency of Moldarting to that of uniform sampling of the torsion. We applied the MolDarting procedure described in Section 2.3.1 to validate this MC move correctly samples the correct population distributions,



(a) Uniform sidechain sampling



(b) Moldarting

Figure 2.6: **MolDarting efficiently samples the conformations of valine-alanine.** (a) (top) A trajectory consisting of MD+MC uniform rotations of the valine sidechain, with the histogram of the data (right). (b) (bottom) A trajectory consisting of MD+MC MolDarting moves of the valine sidechain. Molecular darting converges to the same distribution as uniform torsion rotations. However, MolDarting ends up being about twice as efficient at generating torsion transitions in this system. The red horizontal lines are included to help visually separate the three binding modes.

and to compare the sampling efficiency of MolDarting to a traditional MC method. Both methods converged to the same values for the three dihedral populations (Figure 2.6). Across seven simulations replicates using MolDarting, the acceptance rate of MolDarting moves was only $2.23\% \pm 0.6\%$, compared to the acceptance rate of uniform sampling at $8.04\% \pm 0.5\%$. Although the acceptance rate for molecular darting was lower, the number of transitions generated between dihedral populations was nearly doubled compared to uniform dihedral sampling, with an average of approximately 3400 transitions generated with MolDarting compared to approximately 1400 transitions on average with uniform dihedral sampling.

Thus, because of the targeted nature of MolDarting, the number of transitions between conformations is higher than the uniform sampling case, despite the lower number of accepted moves.

2.4.2 We applied Molecular Darting to a T4 lyosozyme L99A system

Molecular Darting selectively the rotational and translational degrees of freedom in a binding site

We further evaluated our method to sample rotational and translational degrees of freedom by applying MolDarting to sampling the binding modes of toluene bound to T4 lysozme L99A. The are four binding modes of toluene when bound to T4 Lysozyme L99A. These binding modes vary by rotational and translational degrees of freedom; two are distinct and vary by a rotation, and the other two binding modes are symmetry-equivalent to the first pair [43]. We applied MolDarting sampling with BLUES to the non-symmetric binding modes of toluene. The populations of the two binding modes were selectively sampled using MolDarting, without sampling the non-symmetric binding modes (Figure 2.7). MolDarting also was able to recover the correct populations of the binding modes, with the correct population split being 60:40, and our triplicate runs giving 58%±3% for the dominant binding mode and 42% ± 3% for the less populated binding mode. The acceptance rate for these moves over these trials was approximately 22%, which is roughly two times the acceptance rate for random center of mass moves we explored in the original BLUES paper [43], which further shows the benefit of targeted moves.

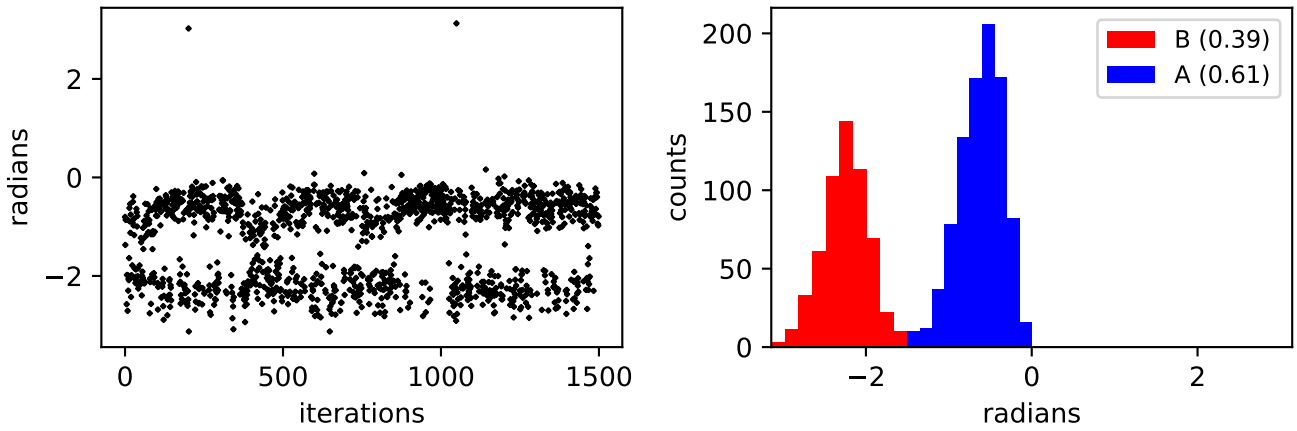


Figure 2.7: **MolDarting generates selective transitions between binding modes** Toluene has four binding modes in the binding site, but only two of the binding modes are sampled here, due to the targeted nature of MolDarting. MolDarting is able to reproduce the correct relative probabilities of both binding modes, which are approximately 60% for binding mode A (the crystallographic binding mode), and 40% for the noncrystallographic pose.

2.4.3 Molecular Darting does not accelerate sampling when outside the dart

Sometimes, running longer simulations on the T4 lysozyme/toluene system resulted in toluene switching to the symmetry-equivalent binding mode (Figure 2.8). When this occurs, the ligand ends up being outside the pre-specified darts we defined in this test, and thus MolDarting moves cannot be attempted. We could have instead included all four ligand binding modes (two symmetry-equivalent pairs) as darts, but we elected not to here as we wanted to focus on non-redundant sampling. This issue highlights a key point: while MolDarting can be used to accelerate sampling, it is only effective when the system is within the selected darts; when outside the darts, we are effectively running plain MD. Thus, to maximize the applicability of MolDarting moves, care should be taken when defining the regions used for MolDarting.

Essentially, MolDarting attempts to trade bias for efficiency. More random procedures, like our initial translational moves in BLUES, allow enhanced exploration of binding mode

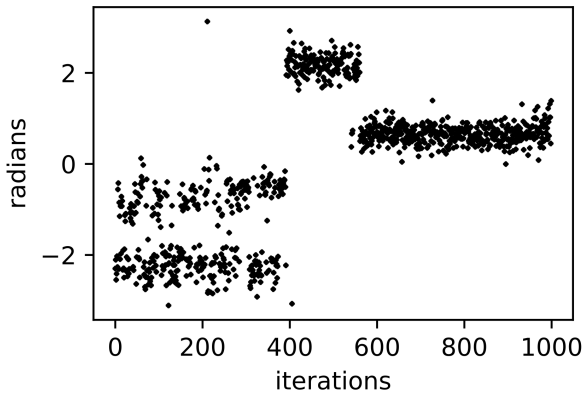


Figure 2.8: **MolDarting does not improve sampling when the simulation moves outside the darts.** Here, the initial binding modes of toluene between 0 and π radians are well sampled (in the first 400 iterations), since these are covered by the rotational regions from MolDarting. However if the simulation leaves that region, then a MolDarting move cannot take place, and thus the simulation becomes just a normal MD simulation. In this particular simulation, around the 400th iteration toluene flips to the symmetric equivalent binding mode, which is not covered by the rotational regions, greatly reducing sampling.

transitions regardless of what pose the ligand is in, but do so rather inefficiently since so many proposed moves are to unfavorable binding modes. MolDarting requires more advance input or bias – selection of a set of potential binding modes to focus sampling on – and thus is able to ensure that proposed moves focus near those binding modes, potentially enhancing efficiency, but when the simulation strays from pre-defined binding modes, no enhanced sampling is possible.

2.4.4 We attempt to use Molecular Darting to explore multiple binding modes of HIV integrase Ligands

We applied Molecular Darting to an HIV integrase system with a set of diverse ligands. We chose HIV integrase in this study since this protein has three distinct binding sites ligands potentially bind to, leading to a plethora of potential binding modes which were hard for methods to discriminate between in a previous blind challenge [85]. By using MolDarting we

aimed to sample the various binding modes in the three binding sites in a single simulation.

The ligands we tested were chosen from the SAMPL4 dataset to include a diverse set of ligands as well as a diverse set of three poses in each binding site, for a total of 9 different binding modes (Section 2.3.1).

We attempted to use MolDarting to sample between binding sites. However, in all the cases with the ligands we studied, the acceptance rate for the moves was 0, thus no moves were accepted.

We looked at two possible sources that could lead to these MolDarting moves being rejected. One possible source of rejection is that the ligand falls outside the regions when MolDarting is being attempted, leading to these moves being rejected.

Another possible source of rejection is the protocol work produced during the move is high, so these moves are rejected by the acceptance criteria.

We first looked at the distribution of attempted MolDarting moves for the ligands (Figure 2.9). We found that although some moves did end up outside the defined regions (indicated by the ligand staying in the initial binding mode, shown in red), the majority of times, the ligand is being proposed to a new binding mode. While our handling of the regions could be improved, it does not appear to be the major cause of MolDarting moves being rejected.

We then looked at the protocol work distributions that are accumulated throughout the NCMC MolDarting move attempts (Figure 2.10).

From the work distributions, we can see that there is that the protocol work accumulation is very large. Even for 50,000 NCMC switching steps, most of the moves attempted aren't close to being favorable (near 0). To investigate further into these high protocol work values, we looked at the instantaneous derivative throughout the NCMC switching protocol (Figure 2.11). If there were infinite switching steps, then we would expect to see the instantaneous

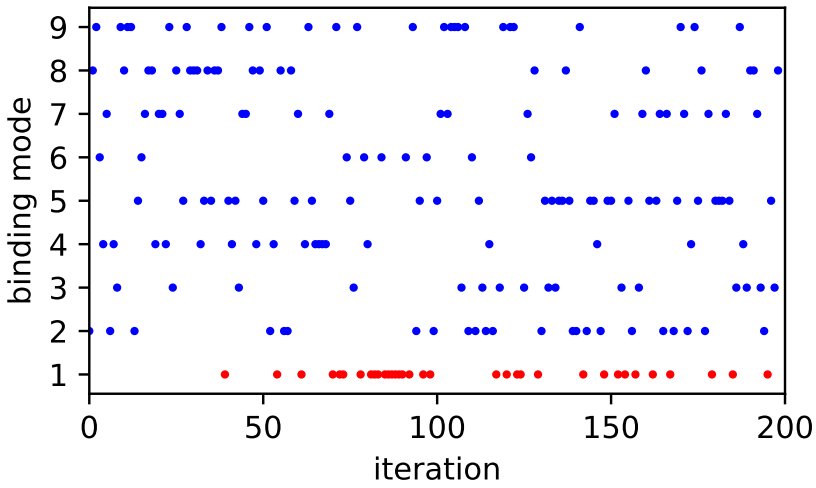
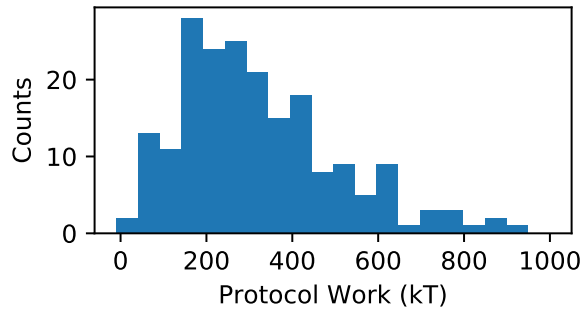
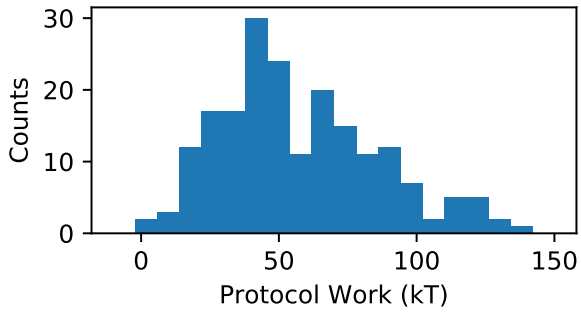


Figure 2.9: **MolDarting attempts sample all the defined binding modes.** We looked at the binding modes sampled by MolDarting moves attempts. All 9 binding modes that were used for MolDarting with this ligand (4CGD) were sampled over the 200 iterations performed. The ligand started in binding mode 1. The points in blue indicate MolDarting move attempts which were successful at sampling new binding modes, while the red indicates that the ligand was outside the defined regions, so no darting move was attempted.

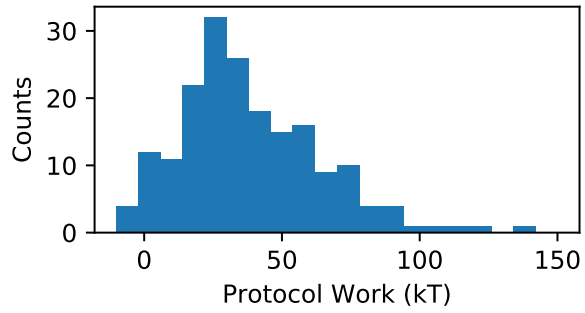
derivative being roughly inversely symmetric around the middle of the protocol. Instead, what we see is that when the ligand's steric interactions are being turned back on, there is a huge spike of protocol work being accumulated. On the other hand, the electrostatics for the system are well-behaved when both turning off and turning on those interactions. These pieces of data suggest that the moves we propose introduce the steric interactions too quickly or in a way which causes clashes that are too severe. We therefore could potentially improve MolDarting move acceptance rates by altering our NCMC switching protocol. Specifically, one route we can take to improve the switching protocol is to increase the proportion of steric NCMC switching steps to the electrostatic NCMC switching steps. Another potential way to increase the acceptance rates is to minimize the variance of the protocol work [98]. As seen in Figure 2.11, the protocol work variance is not constant and changes over the course of the switching steps, so modification of how we change the sterics and, to a lesser extent, the electrostatics (Figure 2.4) during our NCMC protocol could improve our acceptance rates of these MolDarting moves–and NCMC moves in general.



(a) 1,000 NCMC switching steps

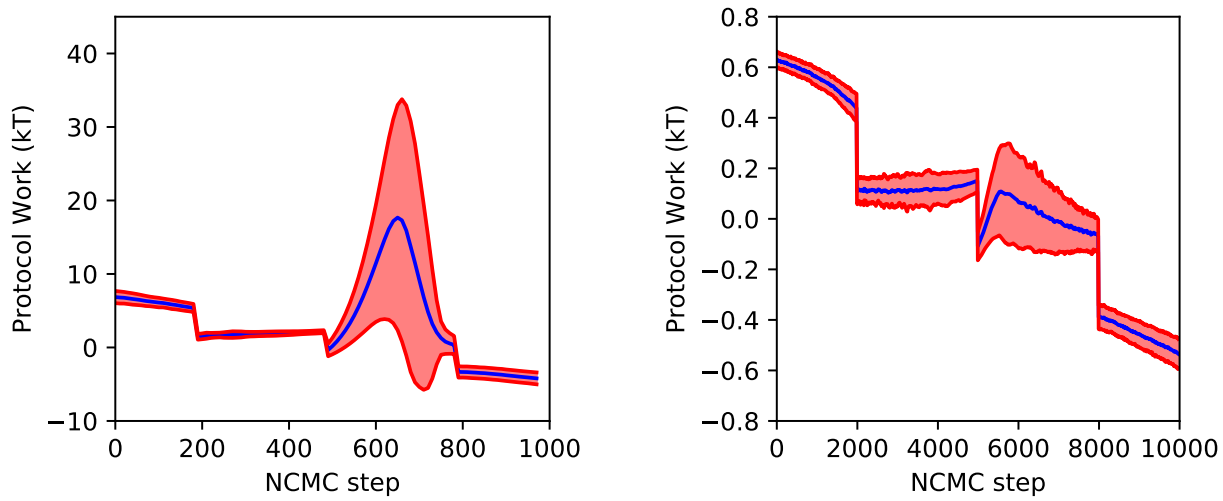


(b) 10,000 NCMC switching steps



(c) 50,000 NCMC switching steps

Figure 2.10: **High protocol work leads to rejection for MolDarting moves.** (a) The protocol work distribution of NCMC with MolDarting move attempts with 1,000 (a), 10,000 ((b), and 50,000 ((c) NCMC switching steps with the HIV integrase and the ligand found in 4CGD. The protocol work done over the course of the NCMC moves generally is highly positive (unfavorable), leading those moves to be rejected by the acceptance criteria. There are a small number of cases when the work values approach zero or are negative, but these were still rejected. In these cases, rejection was due to the ligand ending up outside the defined regions at one of the checks during the course of the move.



(a) 1,000 NCMC switching steps

(b) 10,000 NCMC switching steps

Figure 2.11: **Turning on the steric interactions leads to unfavorable accumulation of protocol work.** (*a*) (left) The instantaneous difference of protocol work accumulation over 1000 switching steps. (*b*)The instantaneous difference of protocol work accumulation over 10,000 switching steps. From 200 iterations of NCMC and MolDarting simulation, we took the average values of the protocol work at each step for 1000 and 10,000 switching steps. From these average values, we calculated the instantaneous difference between the work values, shown by the blue line. The standard deviation of these differences are shown in red. We can see that there is a large accumulation of protocol work when the ligand's interactions are being turned back on (after the halfway point of the NCMC steps).

2.5 Conclusion/Discussion

2.5.1 MolDarting allows sampling of specific binding modes

We have shown that our newly developed Monte Carlo method — Molecular Darting — allows reversible sampling of specific binding modes/conformations by constructing darting moves based on the internal and external degrees of freedom of a ligand. This allows reversible hops between pre-defined metastable binding modes or conformations, opening up exciting new possibilities. Molecular Darting worked well in improving sampling of the different binding modes/conformations in the simpler model systems we considered, and notably showed marked improvements in sampling compared to uniform Monte Carlo sampling

methods and plain Molecular Dynamics.

We did experience challenges, however, in getting acceptance of MolDarting moves in combination with NCMC in the HIV integrase system. Even though the NCMC/MolDarting moves were not accepted, we did find that the attempted MolDarting move proposals were into the intended binding sites/binding modes.

More work can be done in regards to improving move acceptance with NCMC. Potential areas to be explored could be to look into more efficient paths of turning off and on the electrostatics and sterics of the system. Different soft-core potentials could potentially be used as well, to further decrease the accumulated protocol work while turning on the ligand's interactions by minimize the variance of this process [98, 97].

Molecular Darting also has potential applications in combination with other methods, which can be further explored. For instance, MolDarting could find use in equilibrium or expanded ensemble simulations to improve sampling. In the non-interacting states, MolDarting moves should have significant acceptance rates; since there are no clashes with the surrounding atoms of the ligand acceptance will just depend on the ligand's internal degrees of freedom.

Further work can be also be done on generalizing Molecular Darting. One aspect of MolDarting to improve would be allowing regions of arbitrary sizes. While our original implementation of MolDarting only handles regions of the same size, different sized regions can be used instead if they are factored into the acceptance criterion [114]. Similarly, instead of uniform sampling the dihedral regions, we could sample using a Gaussian distribution centered at the maximum of the dihedral, which would favor lower energy conformations of the ligand and thus potentially yield higher acceptance.

Overall, we are excited of the potential applications of Molecular Darting, and its ability to sample phase space in combination with other sampling techniques.

2.6 Acknowledgments

D.L.M. and S.C.G. appreciate the financial support from the National Science Foundation (CHE 1352608) and the National Institutes of Health (1R01GM108889-01) and computing support from the UCI GreenPlanet cluster, supported in part by NSF Grant CHE-0840513. We would like to thank Nathan M. Lim for helping design and maintain the core BLUES code infrastructure and documentation. We also would like to acknowledge Christopher I. Bayly (OpenEye Scientific Software) and Ioan Andricioaei (UC Irvine) for their helpful scientific discussions and insights.

2.7 Disclosures

DLM is a member of the Scientific Advisory Board of OpenEye Scientific Software and an Open Science Fellow with Silicon Therapeutics.

2.8 Supporting information

The Supporting Information is available free of charge. The SI contains the set of scripts used to run the BLUES simulations with MolDarting on the systems described in this paper. Also included are the parameter and coordinate files for the systems used, as well as the analysis scripts used to interpret the output. In addition, a copy of the BLUES version used here is included, along with a README.md file detailing the files present in this SI. The SI can be found at `https://doi.org/10.26434/chemrxiv.12670676.v1`

- Set of scripts for running BLUES simulations with MolDarting,
- Parameter and coordinate files for the systems used

- Analysis scripts for interpreting the output
- A copy of the BLUES version used
- A README.md file detailing the layout of these files

BLUES is also available at `https://github.com/mobleylab/BLUES`.

Bibliography

[1] OEDOCKING.

[2] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, Sept. 2015.

[3] M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp, and P. C. Biggin. Accurate calculation of the absolute free energy of binding for drug molecules. 7(1):207–218.

[4] M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp, and P. C. Biggin. Accurate calculation of the absolute free energy of binding for drug molecules. *Chem. Sci.*, 7(1):207–218, 2016.

[5] I. Andricioaei, J. E. Straub, and A. F. Voter. Smart Darting Monte Carlo. *The Journal of Chemical Physics*, 114(16):6994–7000, Apr. 2001.

[6] I. Andricioaei, J. E. Straub, and A. F. Voter. Smart Darting Monte Carlo. *J. Chem. Phys.*, 114(16):6994–7000, Apr. 2001.

[7] H. J. C. Berendsen, D. Van Der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Comm.*, 91(1-3):43–56, Sept. 1995.

[8] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb. 1992.

[9] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical Physics Letters*, 222(6):529–539, 1994.

[10] S. Boresch, F. Tettinger, M. Leitgeb, and M. Karplus. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. 107(35):9535–9551.

[11] S. E. Boyce, D. L. Mobley, G. J. Rocklin, A. P. Graves, K. A. Dill, and B. K. Shoichet. Predicting Ligand Binding Affinity with Alchemical Free Energy Methods in a Polar Model Binding Site. *Journal of Molecular Biology*, 394(4):747–763, Dec. 2009.

[12] K. H. Burley, S. C. Gill, N. M. Lim, and D. L. Mobley. Enhancing Side Chain Rotamer Sampling Using Nonequilibrium Candidate Monte Carlo. *J. Chem. Theory Comput.*, 15(3):1848–1862, Mar. 2019.

[13] C Schutte, A. Fischer, W. Huisinga, and P. Deuflhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *Journal of Computational Physics*, 151(1):146–168, 1999.

[14] C. M. Campos and J. Sanz-Serna. Extra Chance Generalized Hybrid Monte Carlo. *Journal of Computational Physics*, 281:365–374, Jan. 2015.

[15] D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, Walker, J. Wang, R. M. Wolf, X. Wu, and P. A. Kollman. *Amber 14*. University of California, San Francisco, 2014.

[16] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. The Amber biomolecular simulation programs. 26(16):1668–1688.

[17] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16):1668–1688, Dec. 2005.

[18] H. Chen, P. D. Lyne, F. Giordanetto, T. Lovell, and J. Li. On Evaluating Molecular-Docking Methods for Pose Prediction and Enrichment Factors. 46(1):401–415.

[19] Y. Chen and B. Roux. Constant-pH Hybrid Nonequilibrium Molecular Dynamics-Monte Carlo Simulation Method. *Journal of Chemical Theory and Computation*, 11(8):3919–3931, Aug. 2015.

[20] J. D. Chodera. Yank.

[21] J. D. Chodera. Yank.

[22] J. D. Chodera and F. Noé. Markov state models of biomolecular conformational dynamics. *Curr Opin Struct Biol*, 25:135–144, Apr. 2014.

[23] J. D. Chodera and A. Rizzi. Openmmtools.

[24] A. J. Clark, P. Tiwary, K. Borrelli, S. Feng, E. B. Miller, R. Abel, R. A. Friesner, and B. J. Berne. Prediction of protein-ligand binding poses via a combination of induced fit docking and metadynamics simulations. *J. Chem. Theory Comput.*, 12(6):2990–2998, June 2016.

[25] R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, and B. K. Shoichet. Ligand Pose and Orientational Sampling in Molecular Docking. *PLOS ONE*, 8(10):e75992, Oct. 2013.

[26] Z. Cournia, B. Allen, and W. Sherman. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. 57(12):2911–2937.

[27] J. B. Cross, D. C. Thompson, B. K. Rai, J. C. Baber, K. Y. Fan, Y. Hu, and C. Humblet. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. 49(6):1455–1474.

[28] A. W. S. da Silva and W. F. Vranken. ACPYPE-Antechamber python parser interface. *BMC research notes*, 5(1):367, 2012.

[29] P. Deuflhard and M. Weber. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra and its Applications*, 398:161–184, Mar. 2005.

[30] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, 1987.

[31] P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, and V. S. Pande. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *Journal of Chemical Theory and Computation*, 9(1):461–469, Jan. 2013.

[32] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. 13(7):1–17.

[33] D. A. Erlanson, R. S. McDowell, and T. O'Brien. Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry*, 47(14):3463–3482, July 2004.

[34] F. A. Escobedo and J. J. de Pablo. Extended continuum configurational bias Monte Carlo methods for simulation of flexible molecules. *J. Chem. Phys.*, 102(6):2636–2652, Feb. 1995.

[35] S. Evoli, D. L. Mobley, R. Guzzi, and B. Rizzuti. Multiple binding modes of ibuprofen in human serum albumin identified by absolute binding free energy calculations. 18(47):32358–32368.

[36] L. G. Ferreira, R. N. Dos Santos, G. Oliva, and A. D. Andricopulo. Molecular Docking and Structure-Based Drug Design Strategies. 20(7):13384–13421.

[37] M. Fischer, B. K. Shoichet, and J. S. Fraser. One Crystal, Two Temperatures: Cryocooling Penalties Alter Ligand Binding to Transient Protein Sites. *ChemBioChem*, 16(11):1560–1564, July 2015.

[38] H. Flyvbjerg and H. G. Petersen. Error estimates on averages of correlated data. *The Journal of Chemical Physics*, 91(1):461–466, July 1989.

[39] E. Gallicchio, N. Deng, P. He, L. Wickstrom, A. L. Perryman, D. N. Santiago, S. Forli, A. J. Olson, and R. M. Levy. Virtual screening of integrase inhibitors by large scale binding free energy calculations: The SAMPL4 challenge. *Journal of Computer-Aided Molecular Design*, 28(4):475–490, Apr. 2014.

[40] E. Gallicchio, M. Lapelosa, and R. M. Levy. Binding Energy Distribution Analysis Method (BEDAM) for Estimation of Protein-Ligand Binding Affinities. *Journal of Chemical Theory and Computation*, 6(9):2961–2977, Sept. 2010.

[41] S. Gathiaka, S. Liu, M. Chiu, H. Yang, J. A. Stuckey, Y. N. Kang, J. Delproposto, G. Kubish, J. B. Dunbar, H. A. Carlson, S. K. Burley, W. P. Walters, R. E. Amaro, V. A. Feher, and M. K. Gilson. D3R grand challenge 2015: Evaluation of protein-ligand pose and affinity predictions. *J Comput Aided Mol Des*, 30(9):651–668, Sept. 2016.

[42] C. Georgiou, I. McNae, M. Wear, H. Ioannidis, J. Michel, and M. Walkinshaw. Pushing the Limits of Detection of Weak Binding Using Fragment-Based Drug Discovery: Identification of New Cyclophilin Binders. 429(16):2556–2570.

[43] S. C. Gill, N. M. Lim, P. B. Grinaway, A. S. Rustenburg, J. Fass, G. A. Ross, J. D. Chodera, and D. L. Mobley. Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo. *The Journal of Physical Chemistry B*, 122(21):5579–5598, May 2018.

[44] M. Gilson, J. Given, B. Bush, and J. McCammon. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophysical Journal*, 72(3):1047–1069, Mar. 1997.

[45] A. Grossfield and D. M. Zuckerman. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annu Rep Comput Chem*, 5:23–48, Jan. 2009.

[46] P. J. Hajduk and J. Greer. A decade of fragment-based drug design: Strategic advances and lessons learned. *Nature Reviews Drug Discovery*, 6(3):211–219, Mar. 2007.

[47] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[48] P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, and M. T. Stahl. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. 50(4):572–584.

[49] G. Heinzelmann, N. M. Henriksen, and M. K. Gilson. Attach-pull-release calculations of ligand binding and conformational changes on the first BRD4 bromodomain. *J. Chem. Theory Comput.*, May 2017.

[50] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.*, 18(12):1463–1472, Sept. 1997.

[51] A. Horowitz. A generalized guided Monte Carlo algorithm. *Physics Letters B*, 268(2):247–252, Oct. 1991.

[52] A. Jakalian, D. B. Jack, and C. I. Bayly. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.*, 23(16):1623–1641, Dec. 2002.

[53] G. Jayachandran, M. R. Shirts, S. Park, and V. S. Pande. Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics. *The Journal of Chemical Physics*, 125(8):084901, Aug. 2006.

[54] W. L. Jorgensen and J. Tirado-Rives. Molecular modeling of organic and biomolecular systems using BOSS and MCPRO. *J. Comput. Chem.*, 26(16):1689–1700, Dec. 2005.

[55] S. Kalyaanamoorthy and Y.-P. P. Chen. Modelling and enhanced molecular dynamics to steer structure-based drug discovery. 114(3):123–136.

[56] S. Kalyaanamoorthy and Y.-P. P. Chen. Structure-based drug design to augment hit discovery. 16(17):831–839.

[57] J. W. Kaus, E. Harder, T. Lin, R. Abel, J. A. McCammon, and L. Wang. How To Deal with Multiple Binding Poses in Alchemical Relative Protein Ligand Binding Free Energy Calculations. *Journal of Chemical Theory and Computation*, 11(6):2670–2679, June 2015.

[58] K. Kellett, S. A. Kantonen, B. M. Duggan, and M. K. Gilson. Toward Expanded Diversity of Host–Guest Interactions via Synthesis and Characterization of Cyclodextrin Derivatives. 47(10):1597–1608.

[59] A. D. Kennedy, R. Edwards, H. Mino, and B. Pendleton. Tuning the generalized hybrid monte carlo algorithm. *Nuclear Physics B-Proceedings Supplements*, 47(1-3):781–784, 1996.

[60] P. Kollman. Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews*, 93(7):2395–2417, 1993.

[61] P. I. Koukos, L. C. Xue, and A. M. J. J. Bonvin. Protein–ligand pose and affinity prediction: Lessons from D3R Grand Challenge 3. 33(1):83–91.

[62] A. R. Leach, B. K. Shoichet, and C. E. Peishoff. Prediction of Protein-Ligand Interactions. Docking and Scoring: Successes and Gaps. *Journal of Medicinal Chemistry*, 49(20):5851–5855, Oct. 2006.

[63] B. Leimkuhler and C. Matthews. Rational Construction of Stochastic Numerical Methods for Molecular Sampling. *Applied Mathematics Research Express*, June 2012.

[64] B. Leimkuhler and C. Matthews. Robust and efficient configurational molecular sampling via Langevin dynamics. *The Journal of Chemical Physics*, 138(17):174102, May 2013.

[65] T. Lelièvre, M. Rousset, and G. Stoltz. Thermodynamic integration and sampling with constraints. In *Free Energy Computations: A Mathematical Perspective*, pages 149–258. Imperial College Press, June 2010.

[66] N. M. Lim, M. Osato, G. L. Warren, and D. L. Mobley. Fragment Pose Prediction Using Non-equilibrium Candidate Monte Carlo and Molecular Dynamics Simulations. 16(4):2778–2794.

[67] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 78(8):1950–1958, June 2010.

[68] E. Lionta, G. Spyrou, D. K. Vassilatis, and Z. Cournia. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. 14(16):1923–1938.

[69] K. Liu and H. Kokubo. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-docking Study. 57(10):2514–2522.

[70] E. Marinari and G. Parisi. Simulated Tempering: A New Monte Carlo Scheme. *Europhysics Letters (EPL)*, 19(6):451–458, July 1992.

[71] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, and V. S. Pande. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.*, 109(8):1528–1532, Oct. 2015.

[72] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, and V. S. Pande. Mdtraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophysical Journal*, 109(8):1528 – 1532, 2015.

[73] R. T. McGibbon and V. S. Pande. Variational cross-validation of slow dynamical modes in molecular kinetics. *The Journal of Chemical Physics*, 142(12):124105, Mar. 2015.

[74] J. Michel and J. W. Essex. Prediction of protein–ligand binding affinity by free energy simulations: Assumptions, pitfalls and expectations. 24(8):639–658.

[75] J. Michel, N. Foloppe, and J. W. Essex. Rigorous Free Energy Calculations in Structure-Based Drug Design. 29(8-9):570–578.

[76] J. Michel, R. D. Taylor, and J. W. Essex. Efficient Generalized Born Models for Monte Carlo Simulations. *Journal of Chemical Theory and Computation*, 2(3):732–739, May 2006.

[77] D. L. Mobley, J. D. Chodera, and K. A. Dill. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. 125(8):084902.

[78] D. L. Mobley, J. D. Chodera, and K. A. Dill. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *The Journal of Chemical Physics*, 125(8):084902, Aug. 2006.

[79] D. L. Mobley, J. D. Chodera, and K. A. Dill. Confine-and-release method: Obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.*, 3(4):1231–1235, July 2007.

[80] D. L. Mobley and K. A. Dill. Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". *Structure*, 17(4):489–498, Apr. 2009.

[81] D. L. Mobley and M. K. Gilson. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annual Review of Biophysics*, 46(1):531–558, 2017.

[82] D. L. Mobley, A. P. Graves, J. D. Chodera, A. C. McReynolds, B. K. Shoichet, and K. A. Dill. Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. *Journal of Molecular Biology*, 371(4):1118–1134, Aug. 2007.

[83] D. L. Mobley and P. V. Klimovich. Perspective: Alchemical free energy calculations for drug discovery. *J. Chem. Phys.*, 137(23):230901, Dec. 2012.

[84] D. L. Mobley, S. Liu, N. M. Lim, K. L. Wymer, A. L. Perryman, S. Forli, N. Deng, J. Su, K. Branson, and A. J. Olson. Blind prediction of HIV integrase binding from the SAMPL4 challenge. *J. Comput. Aided Mol. Des.*, 28(4):327–345, Apr. 2014.

[85] D. L. Mobley, S. Liu, N. M. Lim, K. L. Wymer, A. L. Perryman, S. Forli, N. Deng, J. Su, K. Branson, and A. J. Olson. Blind prediction of HIV integrase binding from the SAMPL4 challenge. *J Comput Aided Mol Des*, 28(4):327–345, Apr. 2014.

[86] A. Morton, W. A. Baase, and B. W. Matthews. Energetic origins of specificity of ligand binding in an interior nonpolar cavity of T4 lysozyme. *Biochemistry*, 34(27):8564–8575, 1995.

[87] C. W. Murray and D. C. Rees. The rise of fragment-based drug discovery. *Nature Chemistry*, 1(3):187–192, June 2009.

[88] C. W. Murray, M. L. Verdonk, and D. C. Rees. Experiences in fragment-based drug discovery. *Trends in Pharmacological Sciences*, 33(5):224–232, May 2012.

[89] P. C. Nair, A. K. Malde, N. Drinkwater, and A. E. Mark. Missing Fragments: Detecting Cooperative Binding in Fragment-Based Drug Design. *ACS Medicinal Chemistry Letters*, 3(4):322–326, Apr. 2012.

[90] J. P. Nilmeier, G. E. Crooks, D. D. L. Minh, and J. D. Chodera. Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation. *Proceedings of the National Academy of Sciences*, 108(45):E1009–E1018, Aug. 2011.

[91] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proceedings of the National Academy of Sciences*, 106(45):19011–19016, 2009.

[92] A. Onufriev, D. Bashford, and D. A. Case. Exploring protein native states and large-scale conformational changes with a modified generalized born model. 55(2):383–394.

[93] V. S. Pande, K. Beauchamp, and G. R. Bowman. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods*, 52(1):99–105, Sept. 2010.

[94] N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, and F. von Delft. A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nature Communications*, 8:15123, Apr. 2017.

[95] T. S. Peat, O. Dolezal, J. Newman, D. L. Mobley, and J. J. Deadman. Interrogating HIV integrase for compounds that bind- a SAMPL challenge. 28(4):347–362.

[96] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé. Identification of slow molecular order parameters for Markov model construction. *The Journal of Chemical Physics*, 139(1):015102, July 2013.

[97] T. T. Pham and M. R. Shirts. Identifying low variance pathways for free energy calculations of molecular transformations in solution phase. 135(3):034114.

[98] T. T. Pham and M. R. Shirts. Optimal pairwise and non-pairwise alchemical pathways for free energy calculations of molecular transformation in solution phase. 136(12):124120.

[99] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, and F. Noé. Markov models of molecular kinetics: Generation and validation. *The Journal of Chemical Physics*, 134(17):174105, 2011.

[100] D. C. Rees, M. Congreve, C. W. Murray, and R. Carr. Fragment-based lead discovery. *Nature Reviews Drug Discovery*, 3(8):660–672, Aug. 2004.

[101] S. Röblitz and M. Weber. Fuzzy spectral clustering by PCCA+: Application to Markov state models and data classification. *Advances in Data Analysis and Classification*, 7(2):147–179, June 2013.

[102] M. N. Rosenbluth and A. W. Rosenbluth. Monte Carlo Calculation of the Average Extension of Molecular Chains. 23(2):356–359.

[103] M. Rueda and R. Abayan. Best Practices in Docking and Activity Prediction. *bioRxiv*, Feb. 2016.

[104] T. Sakano, M. I. Mahamood, T. Yamashita, and H. Fujitani. Molecular dynamics analysis to evaluate docking pose prediction. 13(0):181–194.

[105] S. Sasmal, S. C. Gill, N. M. Lim, and D. L. Mobley. Sampling Conformational Changes of Bound Ligands Using Nonequilibrium Candidate Monte Carlo and Molecular Dynamics. *J. Chem. Theory Comput.*, 16(3):1854–1865, Mar. 2020.

[106] M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, and F. Noé. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *Journal of Chemical Theory and Computation*, 11(11):5525–5542, Nov. 2015.

[107] C. Schindler, H. Baumann, A. Blum, D. Böse, H.-P. Buchstaller, L. Burgdorf, D. Cappel, E. Chekler, P. Czodrowski, D. Dorsch, M. Eguida, B. Follows, T. Fuchß, U. Grädler, J. Gunera, T. Johnson, C. Jorand Lebrun, S. Karra, M. Klein, L. Kötzner, T. Knehans, M. Krier, M. Leiendecker, B. Leuthner, L. Li, I. Mochalkin, D. Musil, C. Neagu, F. Rippmann, K. Schiemann, R. Schulz, T. Steinbrecher, E.-M. Tanzer, A. Unzue Lopez, A. Viacava Follis, A. Wegener, and D. Kuhn. Large-Scale Assessment of Binding Free Energy Calculations in Active Drug Discovery Projects.

[108] C. R. Schwantes and V. S. Pande. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *Journal of Chemical Theory and Computation*, 9(4):2000–2009, Apr. 2013.

[109] Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw. How Does a Drug Molecule Find Its Target Binding Site? *Journal of the American Chemical Society*, 133(24):9181–9183, June 2011.

[110] B. Sherborne, V. Shanmugasundaram, A. C. Cheng, C. D. Christ, R. L. DesJarlais, J. S. Duca, R. A. Lewis, D. A. Loughney, E. S. Manas, G. B. McGaughey, C. E. Peishoff, and H. van Vlijmen. Collaborating to improve the use of free-energy and other quantitative methods in drug discovery. *J Comput Aided Mol Des*, 30(12):1139–1141, Dec. 2016.

[111] D. A. Sivak, J. D. Chodera, and G. E. Crooks. Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium discrete Langevin dynamics simulations. *Physical Review X*, 3(1), Jan. 2013.

[112] A. G. Skillman, G. L. Warren, and A. Nicholls. Sampl at first glance: So much data, so little time... http://www.eyesopen.com/2008_cup_presentations/CUP9_Skillman.pdf, Jan. 2008.

[113] G. Sliwoski, S. Kothiwale, J. Meiler, and E. W. Lowe. Computational Methods in Drug Discovery. 66(1):334–395.

[114] C. Sminchisescu and M. Welling. Generalized darting Monte Carlo. *Pattern Recognition*, 44(10):2738–2748, Oct. 2011.

[115] J. Sohl-Dickstein, M. Mudigonda, and M. R. DeWeese. Hamiltonian monte carlo without detailed balance. In *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ICML'14, pages I–719–I–726. JMLR.org, 2014.

[116] D. C. Spellmeyer, A. K. Wong, M. J. Bower, and J. M. Blaney. Conformational analysis using distance geometry methods. *Journal of Molecular Graphics and Modelling*, 15(1):18–36, Feb. 1997.

[117] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1-2):141–151, Nov. 1999.

[118] R. D. Taylor, P. J. Jewsbury, and J. W. Essex. A review of protein-small molecule docking methods. *J Comput Aided Mol Des*, 16(3):151–166, 2002.

[119] B. L. Tembe and J. A. McCammon. Ligand Receptor Interactions. *Comput Chem*, 8(4):281–283, Jan. 1984.

[120] A. R. D. Voet, A. Kumar, F. Berenger, and K. Y. J. Zhang. Combining in silico and in cerebro approaches for virtual screening and pose prediction in SAMPL4. *Journal of Computer-Aided Molecular Design*, 28(4):363–373, Apr. 2014.

[121] J. A. Wagoner and V. S. Pande. Reducing the effect of Metropolization on mixing times in molecular dynamics simulations. *The Journal of Chemical Physics*, 137(21):214105, Dec. 2012.

[122] E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang, and T. Hou. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. 119(16):9478–9508.

[123] J. Wang, Y. Deng, and B. Roux. Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophysical Journal*, 91(8):2798–2814, Oct. 2006.

[124] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9):1157–1174, July 2004.

[125] K. Wang, J. D. Chodera, Y. Yang, and M. R. Shirts. Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. *Journal of Computer-Aided Molecular Design*, 27(12):989–1007, Dec. 2013.

[126] L. Wang, B. J. Berne, and R. A. Friesner. On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities. *Proceedings of the National Academy of Sciences*, 109(6):1937–1942, 2012.

[127] L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, D. L. Romero, C. Masse, J. L. Knight, T. Steinbrecher, T. Beuming, W. Damm, E. Harder, W. Sherman, M. Brewer, R. Wester, M. Murcko,

L. Frye, R. Farid, T. Lin, D. L. Mobley, W. L. Jorgensen, B. J. Berne, R. A. Friesner, and R. Abel. 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. *Journal of the American Chemical Society*, 137(7):2695–2703, Feb. 2015.

[128] Z. Wang, H. Sun, X. Yao, D. Li, L. Xu, Y. Li, S. Tian, and T. Hou. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: The prediction accuracy of sampling power and scoring power. 18(18):12964–12975.

[129] G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, and S. Senger. A Critical Assessment of Docking Programs and Scoring Functions. 49:5912.

[130] G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff, and M. S. Head. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.*, 49(20):5912–5931, Oct. 2006.

[131] O. Weser. An efficient and general library for the definition and use of internal coordinates in large molecular systems. Master's thesis, Georg August Universität Göttingen, November 2017.

[132] H. Wu, A. S. J. S. Mey, E. Rosta, and F. Noé. Statistically optimal analysis of state-discretized trajectory data from multiple thermodynamic states. *The Journal of Chemical Physics*, 141(21):214106, Dec. 2014.

[133] P. Śledź and A. Caflisch. Protein structure-based drug design: From docking to molecular dynamics. 48:93–102.

Appendix A

Supporting Information: Binding Modes of Ligands Using Enhanced Sampling (BLUES): Rapid Decorrelation of Ligand Binding Modes via Nonequilibrium Candidate Monte Carlo

This is the supporting information for Chapter 1 (SI Figures 1–5) Specifically, Figure S1 shows work distributions for rotating toluene in lysozyme as a function of the amount of NCMC relaxation. Figure S2 shows the work standard deviation for toluene in lysozyme as a function of the amount of switching. Figure S3 shows the dihedral progress coordinate used for 3-iodotoluene. Figure S4 shows the estimated MSM transition matrix for toluene in lysozyme. Figure S5 shows acceptance of NCMC vs standard MC move proposals as a function of dihedral angle/binding mode, given a fixed ensemble of MD snapshots.

A separate supporting `.tar.gz` file is available, containing an extensive set of input files, scripts, and code which can be used to reproduce the calculations described in this work.

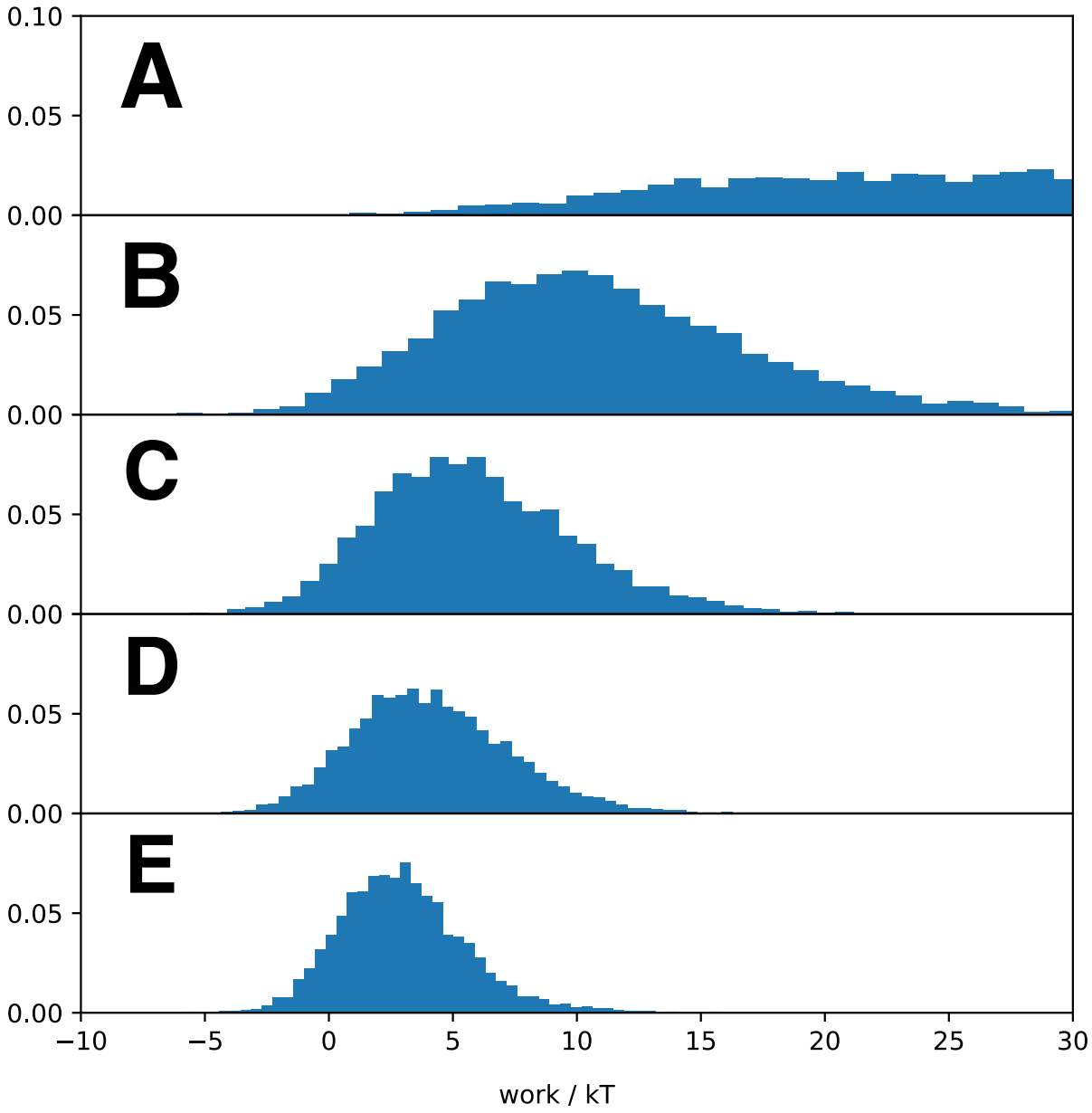
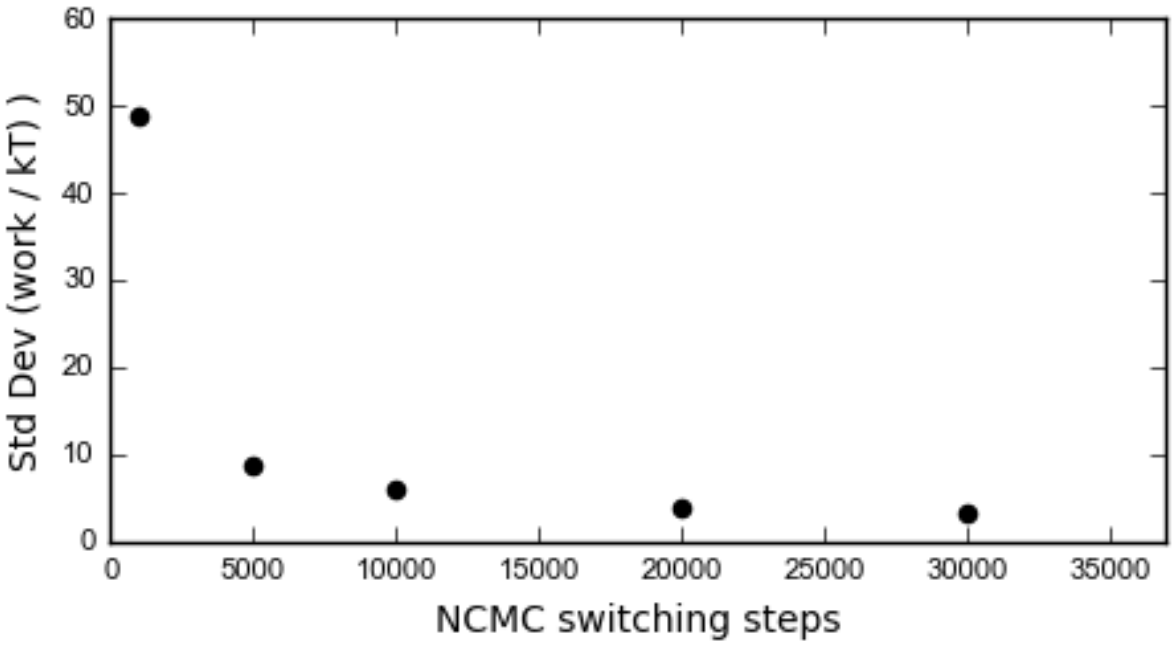


Figure A.1: **Work distributions from increasing NCMC relaxation for rotation of toluene in T4 lysozyme.** The work distributions from 5000 NCMC+MD iterations of varying NCMC relaxation steps are plotted as a histogram over the range `[-10,30]`. A given histogram is over all the counts from that protocol. (*A*) Work distribution from 1000 NCMC steps. (*B*) Work distribution from 5000 NCMC steps. (*C*) Work distribution from 10000 NCMC steps. (*D*) Work distribution from 20000 NCMC steps. (*E*) Work distribution from 30000 NCMC steps. Increasing the number of relaxation steps increases the likelihood that a move will be accepted.



(a) MD/MC

Figure A.2: **Work standard deviations from increasing NCMC relaxation for rotation of toluene in T4 lysozyme.** The standard deviation of the work distributions from 1000 NCMC+MD iterations of varying NCMC relaxation steps. As the number of relaxation steps increase the standard deviation also decreases, which is correlated with the probability of NCMC move acceptance.

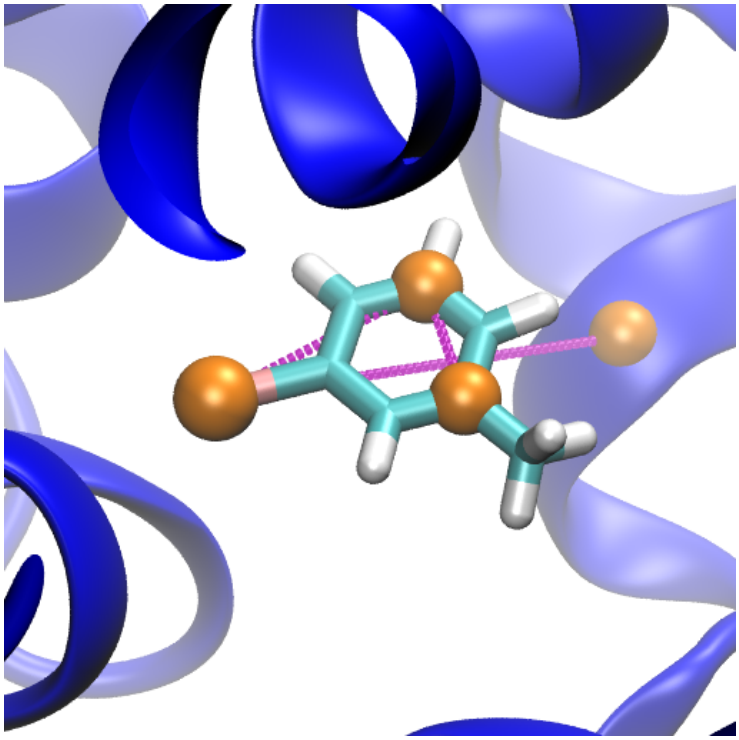


Figure A.3: **Order parameters used for identifying binding modes for 3-iodotoluene in T4 lysozyme.** Shown is a depiction of the dihedral order parameter used to differentiate toluene's binding modes. The dihedral which we monitor is defined by the C1, C5, and I8 atoms of 3-iodotoluene and the alpha carbon of VAL111, shown in orange in CPK representation in orange. In the image, the atoms involved in the dihedral are connected by a purple line, and the dihedral angle measures rotation around the central dashed purple line. The protein is shown in a blue cartoon representation, and 3-iodotoluene is shown in cyan.

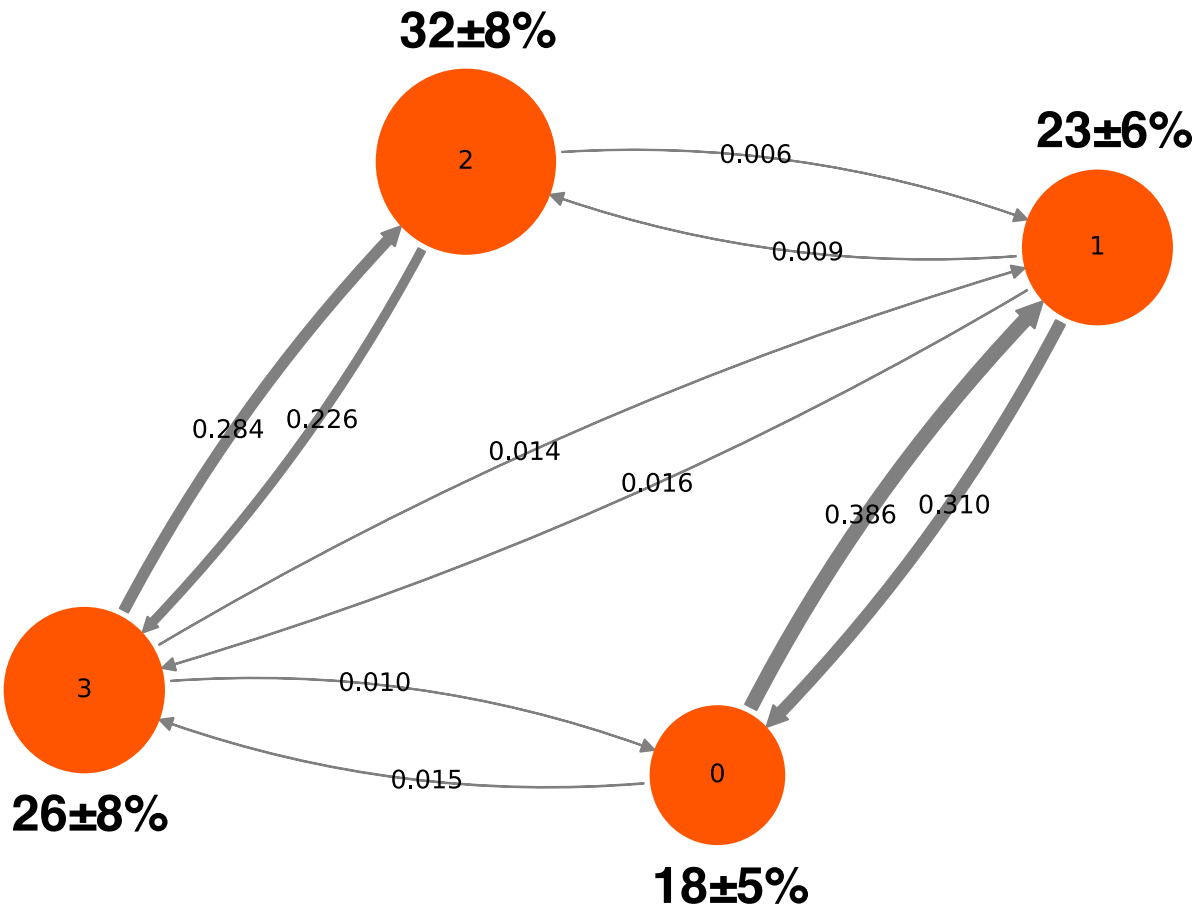
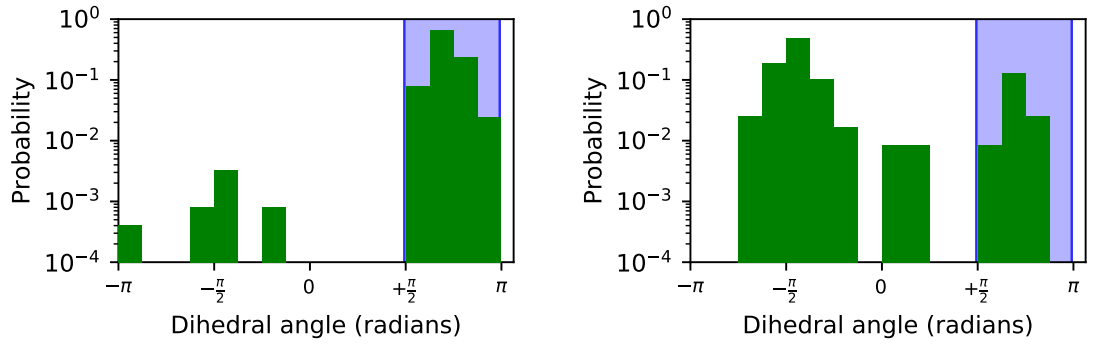


Figure A.4: **Populations and transitions between MSM macrostates for toluene in lysozyme.** Visual representation of the MSM transition matrix and populations generated from Section 2.4 of the main text. The circles labeled with numbers represent separate macrostates, with the populations of each state given by the percentages above and below each circle. The arrows between them represent transition probabilities. Representative binding modes from a macrostate are pictured next to that macrostate.



(a) 3-iodotoluene rotations via MC

(b) 3-iodotoluene rotations via NCMC

Figure A.5: **Acceptance of conventional MC move proposals versus NCMC move proposals proposed from a fixed set of configurations of 3-iodotoluene in T4 lysozyme L99A.** Shown is a comparison of the probability of accepting moves sampling a given ligand binding mode (monitored by a dihedral angle progress coordinate) for 3-iodotoluene in T4 lysozyme L99A for MC move proposals (left plot) and NCMC move proposals (right plot), giving the same ensemble of MD snapshots as a starting point for move proposals. Starting MD snapshots have the ligand in a binding mode in the blue region, but accepted moves involve random rotations of the ligand and thus can be to any binding mode/dihedral angle. In other words, each MC or NCMC trial starts from a selected MD snapshot from within the blue region and, if the move is accepted, the final dihedral angle is computed and a counter is incremented which is used to compute the probabilities on the vertical axis. The MC panel (a) shows data from ten trials of 2,000,000 MC attempts with an overall acceptance rate of $(1.2 \pm 0.2) \times 10^{-2}$ %, but significant moves (larger than 20 degrees) accepted at a rate of only ($(5 \pm 2) \times 10^{-5}$ %). The NCMC panel (b) shows data from seven trials (denoted by dashed vertical lines) of 2000 move attempts, with each move attempt consisting of 6500 NCMC switching steps. Here, the overall acceptance rate is 0.8 ± 0.1%, with moves larger than 20 degrees accepted at a rate of 0.68 ± 0.07 %. In (a), for MC, we observed a total of 13 significant rotations, whereas in (b), in the equivalent number of force evaluations we observed 24 (though the data shown here represents a much larger number of force evaluations for better statistics). In the MC case, because moves are instantaneous, very few significant moves outside the blue region are accepted, giving an (apparently false) impression that the binding mode in the initial green region is by far the more favorable binding mode. In contrast, in the NCMC case, because NCMC allows relaxation of the environment, *most* of the accepted moves are significant and outside the green region, indicating that the alternate binding mode is in fact likely to be more favorable.

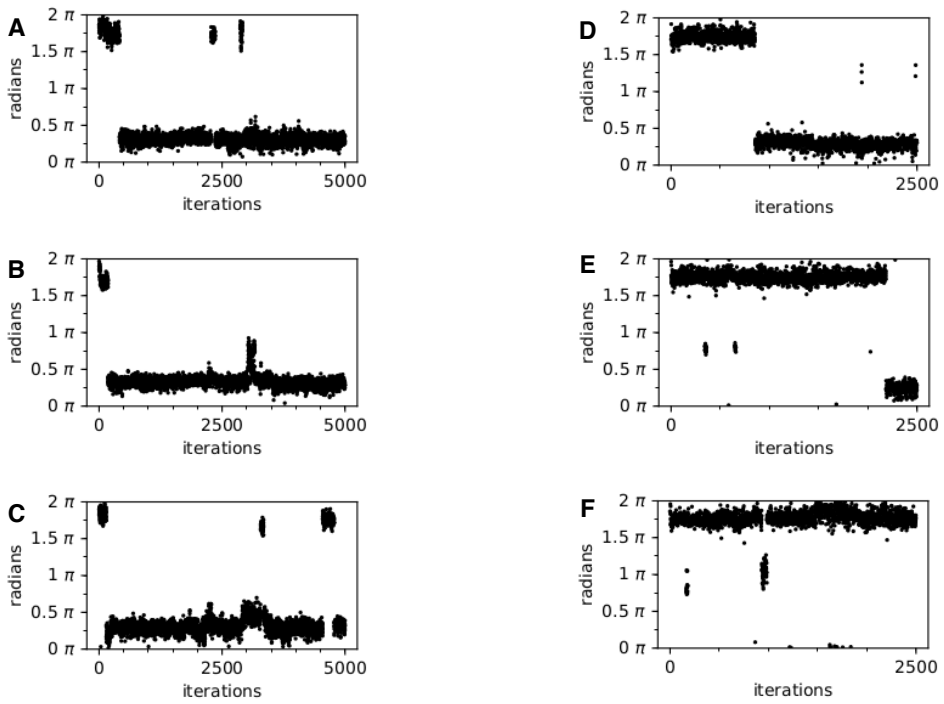


Figure A.6: **3-iodotoluene trajectory comparisons between NCMC and MC as a function of a dihedral progress coordinate.** (*a,b,c*) Dihedral angle (corresponding to binding modes) observed with NCMC as a function of simulation time. (*d,e,f*) Dihedral angle (corresponding to binding modes) observed in the MC as a function of simulation time. Each iteration consists of the same number of energy evaluations, with either 6500 NCMC switching steps or 6500 MC attempts, followed by 10,000 steps of MD. For MC we simulated for 2500 iterations, while for NCMC we simulated for 5000 iterations The periodic dihedral plotted here spans from [0,2π]. The NCMC simulations show more consistent behavior between simulations compared with MC; each NCMC simulation transitions to another stable binding mode (around 1 rad) within 1000 iterations, while the MC simulations sometimes fail to transition to this binding mode at all within 2500 iterations. The overall success rate (per energy evaluation) for transitioning to alternate binding modes appears roughly comparable between the two cases.

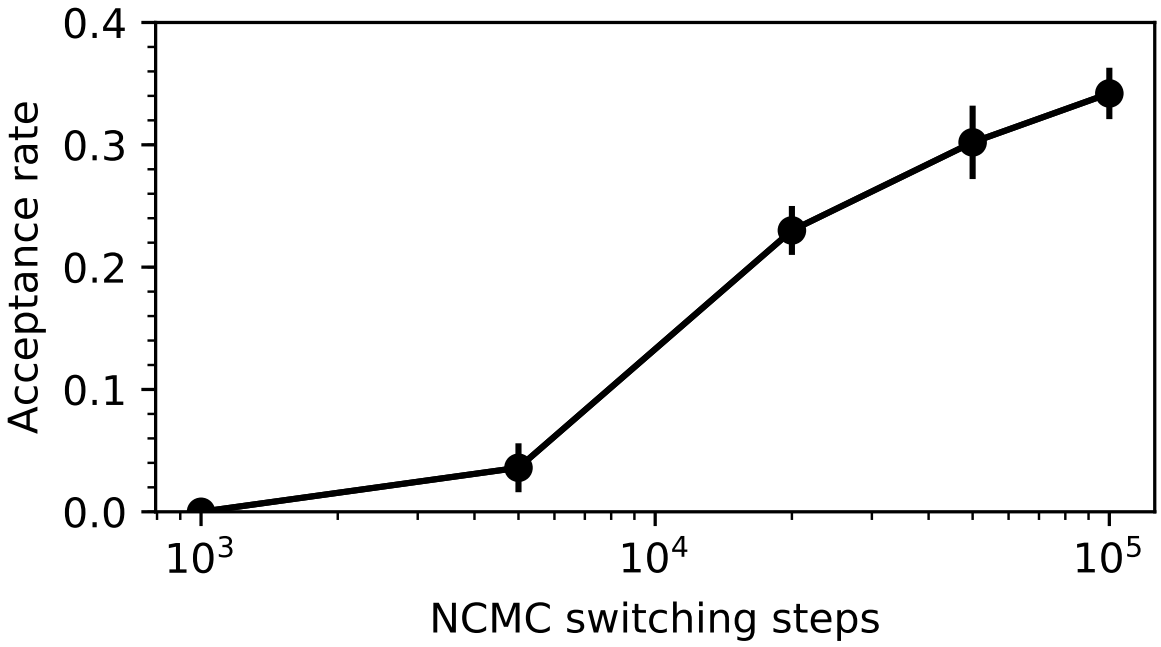


Figure A.7: **Acceptance probability for toluene in solution as a function of the amount of NCMC relaxation.** The acceptance probability of turning toluene's steric and electrostatics off, followed by a random rotation and turning back on toluene's interactions is shown here as a function of the NCMC switching steps. At 1000 switching steps these moves show no acceptance, but the acceptance rates increase with increasing switching steps, up to 32%±2% for 100,000 steps. For each switching step, the uncertainty was calculated based on blocking of 500 BLUES iterations, consisting of the number of switching steps and 100 steps of MD. The number of blocks used was the amount that maximized the standard deviations of the acceptance rate across blocks.