HHS Public Access

Author manuscript

J Comput Chem. Author manuscript; available in PMC 2023 May 07.

Published in final edited form as:

J Comput Chem. 2023 March 30; 44(8): 935–947. doi:10.1002/jcc.27054.

Quality over quantity: Sampling high probability rare events with the weighted ensemble algorithm

Nicole M. Roussey[1], **Alex Dickson**[2]

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan, USA

[2]Department of Biochemistry and Molecular Biology, Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, Michigan, USA

Abstract

The prediction of (un)binding rates and free energies is of great significance to the drug design process. Although many enhanced sampling algorithms and approaches have been developed, there is not yet a reliable workflow to predict these quantities. Previously we have shown that free energies and transition rates can be calculated by directly simulating the binding and unbinding processes with our variant of the WE algorithm "Resampling of Ensembles by Variation Optimization", or "REVO". Here, we calculate binding free energies retrospectively for three SAMPL6 host-guest systems and prospectively for a SAMPL9 system to test a modification of REVO that restricts its cloning behavior in quasi-unbound states. Specifically, trajectories cannot clone if they meet a physical requirement that represents a high likelihood of unbinding, which in the case of this work is a center-of-mass to center-of-mass distance. The overall effect of this change was difficult to predict, as it results in fewer unbinding events each of which with a much higher statistical weight. For all four systems tested, this new strategy produced either more accurate unbinding free energies or more consistent results between simulations than the standard REVO algorithm. This approach is highly flexible, and any feature of interest for a system can be used to determine cloning eligibility. These findings thus constitute an important improvement in the calculation of transition rates and binding free energies with the weighted ensemble method.

Keywords

free energy; ligand binding; molecular dynamics; SAMPL; weighted ensemble

Correspondence Alex Dickson, Department of Biochemistry and Molecular Biology, Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI, USA. alexrd@msu.edu.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

1 | INTRODUCTION

The prediction of ligand (un)binding free energies and rates is important to make systematic improvements in potency during the preclinical stages of drug design.[1] Over the years, numerous computational methods have been developed to predict both relative and absolute binding free energies (and more recently, their binding and unbinding rates, k_{on} and k_{off}, as well).[2-10] These methods make different assumptions that affect accuracy[11,12] and methodological errors are compounded by more general sources of error such as approximations in force-fields or the choice of water model. The performance of free energy predictors can be rigorously examined by comparison with experimental quantities.[13,14] However, reported retrospective predictions are often overly optimistic. A more authentic way of assessing these tools is through participation in blind challenges such as the Statistical Assessment of Modeling of Proteins and Ligands (SAMPL): a series of challenges for the prediction of quantities such as unbinding free energies for small molecule (host-guest) pairs.[15,16] In this challenge, the computational predictions are made without knowledge of the experimental results. Methods for the determination of free energies used in the SAMPL challenges vary from path sampling algorithms such as weighted ensemble[15,17] to alchemical perturbation techniques such as double decoupling methods with thermodynamic integration or Hamiltonian replica exchange.[15]

While there have been significant improvements in the methods used to determine free energies, all methods come with limitations and drawbacks. Even for small systems including the host-guest pairs provided by the SAMPL challenge, computationally calculated values of ΔG for processes such as ligand (un)binding are often inconsistent, and can vary up to several kcal/mol from experimental reference values.[15,16] Efforts to determine “computational reference values” using extensive sampling for a given set of conditions (e.g., force field, binding pose, temperature) are important to isolate methodological sources of error for competing approaches. In the SAMPL6 challenge, the results of the OpenMM/HREX method—double decoupling with Hamiltonian Replica Exchange—were used as the computational reference value and were primarily used for the determination of relative efficiencies of different methods. It was found that even for simulations with seemingly converged predictions for ΔG and nearly identical simulation parameters, results varied up to ~1.0 kcal/mol, and while possible sources of methodological error were determined, the exact source of these discrepancies was unknown.

For alchemical methods, a major bottleneck is achieving adequate sampling of different binding modes.[18] In addition, inclusion of the residual charges of vanishing atoms in single topology methods or overlap of groups in the system in double topology models can also be significant sources of error.[11] Alchemical methods also have the major drawback of only determining free energy differences between the points of interest, and subsequently do not provide free energy profiles for the unbinding/transition pathway of interest. Physical sampling methods, such as metadynamics,[19,20] umbrella sampling,[21,22] multi-ensemble Markov Models,[23-25] and weighted ensemble (WE),[26-30] can determine these free energy profiles, but are limited by the determination of collective variables and convergence issues. In particular, the WE method allows for the generation of continuous trajectories between points of interest without employing any biasing forces. It does this by iteratively

"resampling" a collection of otherwise independent trajectories, referred to as "walkers," each of which carrying a statistical weight. The resampling process clones walkers in underrepresented regions of interest – dividing their weight evenly among the clones – and merges walkers in overrepresented regions – combining their weights in order to conserve probability. The use of the "REVO" resampler (Resampling of Ensembles by Variation Optimization)[28] builds upon these benefits as it resamples trajectories based upon their distances to one another, instead of constructing a set of bins. This makes it easier to sample high-dimensional spaces of collective variables, while avoiding exponential increases in simulation time associated with regularly-spaced bins.

Weighted ensemble methods are particularly useful for determining binding and unbinding rates.[17,31,32] This is done by measuring the transition flux into either the bound or unbound state as the sum of walker weights per unit time. These binding and unbinding rates are interesting in their own right, as they are necessary to model drug behavior in typical nonequilibrium conditions. In addition, the unbinding rate (or equivalently, its inverse, the residence time) is the crucial quantity in some systems to determine drug efficacy in vivo.[17,33,34] Although much progress has been made in the calculation of residence times for pharmaceutically-relevant ligands, as of now there are no blind challenges for the prediction of binding or unbinding kinetics. While researchers can compare to experimental data, as discussed above sampling errors are hard to isolate, and connecting computational and experimental models (for instance, in the definition of the bound and unbound states) is not trivial. For this reason our approach is to indirectly validate the transition rate calculations by using them to calculate binding affinities in the SAMPL challenges. Specifically, we calculate the binding free energy ΔG_b as a function of transition rates using: $\Delta G_b = -kT \ln k_{off} / (C_0 k_{on})$.

Although WE has been successful in simulating a wide variety of long-timescale events ranging from seconds, such as the opening of the SARS-CoV-2 spike protein,[35] to multi-minute long events, such as ligand unbinding,[36,37] a common problem is the high variation of walker weights between independent replicates, which causes large uncertainties in calculated observables such as free energies and rate constants.[17] REVO itself was developed as a way to lower the variation between independent replicates, however previous applications of REVO to the same systems used in this work still yielded rate calculations with large differences from one run to another.[17,38] This is caused by the rare occurance of high probability walkers crossing into the unbound state in unbinding simulations, which introduces large jumps in transition fluxes and unbinding rates. Interestingly, our previous work sought to identify microscopic determinants of these differences in probability and found that there was a correlation between guest-ion interactions and the unbinding weight of a given trajectory for several SAMPL systems.[39] Specifically, rare high-probability trajectories that contributed most significantly to the unbinding rate had a lower amount of guest-ion interaction than the far more numerous low-probability trajectories. By tracking the center of mass (COM)-to-COM distance (d_{COM}), we also found that the high- and low-probability trajectories could be distinguished even relatively early on in the unbinding pathway ($d_{COM} = 7$ Å) based on their guest-ion interaction, prior to there being a discernible difference in the trajectory weight. This d_{COM} distance also corresponded to a "commitment

to unbinding" point, where reactive trajectories left the binding pocket and no re-crossings were observed. This observation implied that the final weight of an unbinding trajectory was, to some extent, already determined early on. If so, it suggests that cloning trajectories with d_{COM} >7 Å is wasteful, as it is already known that these trajectories are going to unbind. Therefore, it may be beneficial to prevent cloning of trajectories that had already reached d_{COM} = 7 Å, as this leaves more spots available in the ensemble for other events to occur and will result in the generation of a high weight unbinding events.

Here, we determine unbinding free energies for four systems through binding and unbinding simulations with a new modification to the REVO algorithm. Three of these systems come from the SAMPL6 challenge[15,16]: "OA-G6," "OA-G3," and "CB8-G3." For these systems, we had prior knowledge of ΔG, however the binding free energy for a fourth system, β-CD-PMZ, was predicted prospectively as part of the SAMPL9 challenge. To determine these free energies, a modification of REVO[28] was used for unbinding simulations. In this method, which we refer to as cutoff-REVO, cloning is prevented for trajectories that have met a system-specific d_{COM} distance between the host and guest molecules. This cloning cutoff value is determined from previously run standard REVO simulation data. The use of a cloning cutoff produced ensembles with fewer unbinding events, but with much higher weights. Using these simulations, k_{off}, k_{on}, and ΔG are calculated for all systems for both standard and cutoff-REVO. The ΔG values are compared to computational (SAMPL6 systems) and experimental (all systems) reference values. Differences in resampling patterns between standard REVO and cutoff-REVO are also analyzed and visualized using a network algorithm. A comprehensive tutorial for running standard REVO unbinding simulations, cutoff-REVO unbinding simulations, and rebinding simulations (along with all relevant code) is provided along with this paper.[40]

2 | METHODOLOGY

2.1 | Host-guest systems

The SAMPL systems used here are small molecule pairs that come from the SAMPL6 and SAMPL9 challenges. These are artificial systems whose primary purpose is to test the accuracy and efficiency of new computational methods for the prediction of binding free energy. The systems used here include the OA-G6, OA-G3, and CB8-G3, systems from SAMPL6 and β-CD-PMZ from SAMPL9 (Figures 1 and S1). These are all host-guest systems, where the host molecules are referred to as: OA (octa-acid), CB8 (cucurbit[8]uril) and β-CD (β-cyclodextrin). The SAMPL6 guest molecules are named according to their index in the challenge: OA-G6 refers to 4-methyl pentanoic acid, OA-G3 refers to 5-hexenoic acid, and CB8-G3 is quinine. PMZ refers to promazine hydrochloride.

The OA host molecule has a four-fold symmetry along the vertical axis and a − 8 charge. Its guest molecules, G6 and G3 are both small molecules with an explicit charge of −1. The CB8 host has eight-fold symmetry along the vertical axis and two-fold symmetry along the horizontal axis. CB8, as well as its quinine guest are both neutral molecules. The β-CD host has no net charge and seven-fold symmetry along its vertical axis. Its guest, PMZ, has a + 1 explicit charge.

2.2 | The weighted ensemble method

The WE[26] method is an unbiased path sampling method that allows for the efficient generation of (un)binding paths. A generalized framework for this two-step method is as follows: in the first step an ensemble of trajectories, also called "walkers," are independently propagated forward in time by molecular dynamics (MD), and in the second step these are "resampled," using merging and cloning operations. The purpose of resampling is to clone walkers that are desirable and to merge walkers that are less desirable; this was originally guided by the populations of a set of bins that spanned the conformation space, but can be thought of more generally, as we discuss below. The goal of cloning, generally, is to help increase the chance of escaping local energy minima and observe a process of interest. The goal of merging is to reduce redundant trajectories in the ensemble and decrease the computational cost of the simulation.

In a WE simulation, all of the walkers have a probability, or a statistical weight. When a walker undergoes a cloning operation, two new independent trajectories are created with the same conformation as the original walker and half of its weight. When two walkers A and B undergo a merging operation they are combined to create walker C with weight $w_c = w_A + w_B$. Walker C takes on either the conformation of walker A or B, with a probability proportional to their weights. Generally, a resampling algorithm can be thought of as a function that takes in the ensemble of trajectories, guides the merging and cloning process, and returns a new ensemble where the conformations are taken from the input ensemble. In this new ensemble, while the number of returned walkers can vary, the sum of the walker weights (most often $\sum_{i=1}^{n} w_i = 1$), remains unchanged.

2.3 | REVO

The resampling algorithm used to perform both the binding and unbinding simulations in this work is the REVO (Resampling Ensembles by Variation Optimization) resampler.[28] After a cycle of dynamics has been run, REVO guides the merging and cloning operations for an ensemble by maximizing the "trajectory variation", V (Equation 1), which is a scaled sum of the all-to-all distances between the walkers in the ensemble.

$$V = \sum_i V_i = \sum_i \sum_j \left(\frac{d_{ij}}{d_\star}\right)^\alpha \phi_i \phi_j, \tag{1}$$

Where V_i is the "trajectory variation" for walker i, d_{ij}, is a distance calculated between walkers i and j, and ϕ_i is a "novelty" term that describes the significance of each individual walker in the ensemble. The distance metric d_{ij} is different for unbinding and rebinding REVO simulations, each of which is described below. The variable $d_\star$ is a "characteristic distance," which is used to make the variation function unit-less, and is equal to the mean of the distance metric after one cycle of MD. Note that this value does not have an impact on resampling behavior and is used for ease of comparison between varying distance metrics for the same system, α is a parameter that balances the "exploitation" (novelty) with the

"exploration" (distance) in the REVO algorithm. The novelty term, ϕ, is usually defined according to the walker weight, w_i, as follows:

$$\phi_i = \log(w_i) - \log\left(\frac{p_{\min}}{C}\right), \tag{2}$$

Where $p_{\min}$ is the predefined minimum walker weight allowed in the simulation, and C is a constant. The value of $p_{\min}$ should be set according to the anticipated probability of the events of interest, here we use 10^{-12} to be consistent with previous work.[17,38] We also enforce a maximum weight of 0.1, by preventing merging events that would exceed that weight. As weights range from 10^{-12} to 0.1, ϕ_i is in the range $[\log C, \log C + 25.3]$. The parameter C acts as an additive constant that raises the range of the novelties, reducing the ratio of the maximum to minimum novelty values. Here we use $C = 100$, as done in previous work,[17,38,41] resulting in a novelty range of [4.6, 29.9], and a max:min novelty ratio ≈ 6.

After a cycle of dynamics has been run and the distances have been calculated, REVO guides merging and cloning operations by first calculating the value of V. Walkers are then proposed for merging and cloning as follows. The walker i is proposed for cloning which has the highest V_i and a weight greater than $2p_{\min}$. A walker j, with the lowest value of V_j and a weight less than $p_{\max}$ (here, set to 0.1) is chosen for merging. The walker k that is closest to j with a weight such that $w_j + w_k < p_{\max}$ is identified as its merging partner. In order to proceed, their distance d_{jk} must be less than or equal to a predetermined value referred to as the "merge distance". Once these three walkers have been selected for resampling, V is recalculated as though the merging and cloning operations have been performed. If V increased, the proposed resampling operations are carried out and the process of proposing merging and cloning steps repeats until V no longer increases, or suitable walkers cannot be found. At that point we terminate the resampling process and run the next cycle of dynamics.

2.4 | The cutoff-REVO algorithm

Here we examine a modified version of the REVO resampler, referred to as cutoff-REVO. Overall, this algorithm works the same as the standard REVO algorithm described above, but with an additional criterion used to determine if a trajectory in the ensemble should be eligible for cloning. In this work, the eligibility function was a d_{COM} distance between the host and guest (with the value specifically-determined for each system). If a trajectory had a d_{COM} distance equal to or greater than the cutoff value, it was not eligible for cloning in that cycle. This cutoff value was determined from d_{COM} plots of unbinding events from standard REVO simulations (described further in Section 2).

Although the eligibility function was based on COM in this work, in general, it is highly flexible and can be customized on a system by system basis. For instance, this criterion can be based on physical features such as RMSD, number of interactions or contacts, solvation of binding pockets, or other non-structural features such as energies or work. We implemented this eligibility function in the `wepy` software package.[42] Similar to other

`wepy` objects, such as distance metrics or boundary conditions, the eligibility function is implemented in a modular fashion: users can implement and use their own customized eligibility function without changing the source code for the REVO resampler.

2.5 | Simulation details

All of the simulations here were run with wepy[42] and OpenMM.[43] The following parameters were used for all simulations: 3000 cycles, 10,000 dynamics steps per cycle with a 2 fs timestep with standard OpenMM implemented hydrogen bond constraints via SHAKE, and a temperature of 300 K. Langevin integration was used for all simulations with a friction coefficient of 10^{-1} picoseconds. The β-CD-PMZ system was built with CHARMM-GUI,[44-46] and the SAMPL6 systems were built with the Gromacs input files provided by the challenge organizers.[15] The β-CD-PMZ system was run at constant pressure of 1 bar, and OA-G6, OA-G3, and CB8-G3 were run at constant volume. Box sizes (in nanometers) for the systems are as follows: β-CD-PMZ – 4.88 × 4.88 × 4.88, OA-G6 – 4.27 × 4.32 × 4.33, OA-G3 −4.30 × 4.32 × 4.33, and CB8-G3 – 4.072 × 4.05 × 4.04.

For each SAMPL6 system, five potential binding poses were provided by the challenge organizers. In previous work we found that the choice of initial binding pose is inconsequential as all poses quickly interconvert with one another.[17] Here we chose pose 0 as the initial positions for all walkers in the REVO simulations for the OA-G6, OA-G3 and CB8-G3 systems. For β-CD-PMZ, we generated our own starting poses for the unbinding simulations. Three copies of the β-CD-PMZ system were built with the multicomponent assembler in CHARMM-GUI.[44-46] Following the 10 ns initial OpenMM simulation, a bound conformation was produced (Figure 2). The final frame of the bound run of the three initial simulations was used to initialize unbinding simulations.

For the SAMPL6 systems, five “standard” REVO unbinding simulations were run (without the eligibility criterion). The distance metric between walkers, d_{ij}, is calculated as the RMSD of the guest atoms after alignment of the hosts. Note that in contrast to previous work,[17,38] the distance calculations did not take into account the symmetries of the host (4-fold for OA and 8-fold for CB8), for simplicity. The distance metric was implemented using the ReceptorDistance class in wepy. We use the standard unbinding boundary condition UnbindingDistance where a trajectory is marked as “unbound” when the closest host-guest atomic distance is greater than 1 nm. In addition, five cutoff-REVO simulations were run, with a cutoff determined from the reactive trajectories in the standard REVO simulations (more information is available in Figure S2) and the same distance metric and unbinding boundary conditions as the standard REVO simulations. For the β-CD-PMZ system we ran 5 standard REVO simulations and 5 cutoff-REVO simulations, using the same approach.

In the rebinding simulations, all of which were run with standard REVO, starting poses were taken from the unbinding events of the standard REVO simulations. For simplicity, we used a different distance metric in the rebinding simulations for the β-CD-PMZ system, defined as follows:

$$d_{ij} = \frac{1}{d_{\mathrm{COM},i}} - \frac{1}{d_{\mathrm{COM},j}}. \quad (3)$$

This is similar to the `RebindingDistance` metric used in Dixon et al,[17] but is based on the host-to-guest d_{COM} instead of the RMSD to a native structure. The difference of the inverse is used in order to emphasize difference between smaller values of d_{COM}. In addition we use a binding boundary condition that is also based on the d_{COM}. For this boundary condition, binding events were triggered when $d_{\mathrm{COM}} < 0.4$ nm. This value was determined from Figure S2, which shows a subset of the COM distances for the whole bound to unbound transitions for all four systems. For OA-G6, OA-G3, and CB8-G3, we use rebinding data that was previously generated in Ref. 17.

Figure S2 was also used to determine the cutoff value for cloning eligibility. Coincidentally, the determined "commitment to unbinding" points for all four systems was 0.7 nm. While it appears that the cutoff COM distance chosen could have been reduced for the OA-G6 and OA-G3 systems, it was found in previous work (Ref. 39) that for these systems, rebinding can often occur below a 0.7 nm d_{COM} distance. The tutorial repository associated with this article[40] contains everything needed to reproduce the results here, including: code for the custom distance metric and boundary condition, scripts to run the binding and unbinding simulations, scripts to plot the d_{COM} for reactive trajectories, and scripts to extract the rebinding simulation starting poses.

2.6 | Calculation of rates and free energies

With the WE method, rates and subsequently free energies can be calculated through a method called "ensemble splitting".[17,47-50] In this method, the equilibrium ensemble is split into "binding" and "unbinding" ensembles. The binding ensemble is composed of the trajectories that have most recently been in the unbound basin, and the unbinding ensemble is composed of the trajectories that have most recently been in the bound basin. This technique can be used with any set of non-overlapping basins. Here, the unbound basin is defined as a set of structures where the minimum host-guest interatomic distance is greater than 1 nm, and the bound basin is either defined as a set of structures where the d_{COM} distance is below 0.4 nm (for β-CD-PMZ) or when the ligand RMSD to the native structure is <0.1 nm (OA-G6, OA-G3, and CB8-G3).

Here, simulations are conducted entirely in either the binding or unbinding ensemble. When a walker leaves its ensemble by entering the other basin, its weight is saved, and the structure, which is recorded as an "exit point" then undergoes a process called warping. When a trajectory warps, it is set back to its starting pose but its weight remains unchanged. In this work, for the unbinding ensemble, the starting structure is the initial bound pose, and for the binding ensemble, the starting structure is one of the initial unbound poses. Rates are calculated via the trajectory flux that leaves one ensemble for the other: k_{off} (4), k_{on}(5) and ΔG(6)can be determined as:

$$k_{\text{off}} = \overline{\phi_u} = \frac{\sum_{i \in U} w_i}{T}, \tag{4}$$

$$k_{\text{on}} = \overline{\phi_b} = \frac{\sum_{i \in R} w_i}{CT}, \tag{5}$$

$$\Delta G = kT \ln \frac{k_{\text{off}}}{C_0 K_{\text{on}}}, \tag{6}$$

Where, $\overline{\phi_b}$ and $\overline{\phi_u}$ are the binding and unbinding flux, T is the elapsed time, the sum is the sum of weights for the corresponding exit points for all unbinding events (U) and rebinding events (R), and C is the concentration of the guest molecule. C_0 is the reference concentration of 1 mol/L. The code for this analysis is provided in the associated tutorial.

3 | RESULTS

3.1 | Resampling trees show differences in merging and cloning behavior

For each host-guest system used, five unbinding simulations were run for both the standard REVO resampler and the cutoff-REVO resampler. In this Section, we focus on resampling differences in the β-CD-PMZ system, as these are representative of the other systems as well. We find that these resamplers have significant differences in their cloning and merging behavior, which we visualize using "resampling trees" (Figure 3). These are directed acyclic graphs where each cycle of each walker is represented by a node. The nodes for each cycle are evenly spaced along the vertical direction, with the earliest timepoints at the bottom of the resampling tree. Connections between nodes show the continuation of walkers over time: multiple connections between cycles show cloning events, and the absence of connections to the next cycle indicates that the walker was merged into one of the other walkers. The sizes of each node represent the weights of the walkers at a given timepoint, and in Figure 3 the nodes are colored according to their d_{COM} value. The code used to generate these trees is available as a tutorial.[51]

We find that standard REVO produces fewer cloning events per cycle ($\langle n_{wc} \rangle < 1$), with slightly fewer replicates ($\langle R \rangle$) produced at each of these events, whereas cutoff-REVO has more cloning events per cycle ($\langle n_{wc} \rangle = 1.9$) with slightly more replicates produced at each event, on average (Table 1). As the simulation progresses the number of cloning events for standard REVO significantly decreases. This is in contrast to cutoff-REVO, where $\langle n_{wc} \rangle$ stays consistant over the entire simulation (Figure 4). After the first walker leaves the binding site in standard REVO, we find that the majority of cloning events focus on high d_{COM} walkers. This is consistent with REVO's goal of maximizing the variation in the ensemble, but this limits our ability to clone walkers that are closer to the binding site and generate new, independent unbinding events. In contrast, for cutoff-REVO we see more consistent cloning of trajectories that are earlier in the unbinding process. This is due to the lack of mass cloning events that fragment the more unbound trajectories. Preventing this

fragmentation results in more slots being available for early-unbinding cloning events, as the trajectories of interest that are beyond the cutoff are not taking up these spaces. Preventing the fragmentation of unbound walkers results in the generation of fewer, but higher-weighted unbinding events.

The average d_{COM} distance for a cloning event for standard REVO across the five β-CD-PMZ runs was found to be 1.1 nm ± 0.008, whereas it was 0.35 ± 0.0008 nm for cutoff-REVO. As expected, the maximum d_{COM} of a cloning event was also much larger for REVO (2.3 ± 0.013 nm) than for cutoff-REVO (0.76 ± 0.004 nm). By shifting the emphasis of the cloning events to lower d_{COM} walkers, cutoff-REVO also shifts emphasis toward higher weighted walkers. The relationship between the d_{COM} and walker weights are shown in Figure S3. Figure S3A shows that walkers with a weight $< 10^{-7}$ are mostly in the unbound state, with an average d_{COM} >1.0 nm. As cutoff-REVO avoids the cloning of these walkers, the emphasis is dramatically shifted from low- to high-weight walkers (Figures S3C and S3D).

3.2 | Unbinding events, free energy, and rate analysis

All of these simulations generate numerous unbinding events, but with stark differences in both the probabilities of the unbinding events and their number as seen in Figure 5 (results for OA-G3 and CB8-G3 shown in Figure S4). For all four systems, standard REVO unbinding simulations produce numerous unbinding events, although most of these had extremely low probabilities. The highest weighted unbinding events are particularly important as they dominate the calculation of unbinding rates in Equation 4. For β-CD-PMZ, the highest weight event obtained was 10^{-2} for both standard REVO and cutoff-REVO. We observe differences in the highest weights for other systems: such as 10^{-7} and 10^{-3} (OA-G6), 10^{-6} and 10^{-3} (OA-G3), and 10^{-8} and 10^{-7} (CB8-G3), where the higher weights are always obtained by the cutoff-REVO algorithm. Interestingly, the number of unbinding events for cutoff-REVO did not increase as the exit point weight decreased; instead the unbinding event weights were roughly normally distributed. The number of unbinding events for all systems also significantly decreased for all four systems as summarized in Table 2.

The binding and unbinding rates for all systems were determined through the use of the "ensemble splitting" method as described in Section 2. The binding rate is found by dividing the rebinding trajectory flux by the guest concentration, where $C = \frac{1}{N_a V}$ and V is the box volume (Equation 5). Rolling estimates of the k_{off}, k_{on}, and ΔG are shown for all systems in Figure S5 for standard REVO results and Figure S6 for cutoff-REVO. We observe differences between the standard REVO and cutoff-REVO results in both the accuracy of the final ΔG obtained and the consistency between runs as seen for k_{off} values in Figure 6 and Table 3. For the SAMPL6 systems, these values are compared to computational reference values determined through OpenMM-Hamiltonian Replica Exchange[15,16] simulations that were run by the SAMPL organizers as a means to compare the effeciency of methods. No computational reference value is available for the β-CD-PMZ system. Experimental reference values are also available for all four systems. For the SAMPL6 systems, this value was determined with isothermal titration calorimetry.[16,52] The same method was used

to determine the free energy of unbinding for β-CD-PMZ, with more information on the method used at the SAMPL9 GitHub page.[53]

For cutoff-REVO, all systems have a final value of ΔG that is very close to either the computational (OA-G3 and CB8-G3) or experimental (β-CD-PMZ and OA-G6) reference value. Standard REVO obtained a very similar final ΔG to cutoff-REVO for the β-CD-PMZ system, however standard REVO has a large uncertainty, with k_{off} estimates varying between runs by up to six orders of magnitude. The final ΔGs obtained for CB8-G3 for both methods were very similar; however, cutoff-REVO produced highly consistent runs for this system with minimal error in ΔG (Figure 7), whereas standard REVO had an uncertainty of 1.01 kcal/mol. The final free energies for REVO and cutoff-REVO as well as all available reference values are shown in Table 4, with plots available in Figures S5 and S6. The errors shown in Table 4 are the standard error for the final values determined for ΔG across all five runs for each data set. Coincidentally, both OA-G6 and β-CD-PMZ have exactly the same experimental binding free energy. With this in mind, both REVO and cutoff-REVO provide the best rank ordering possible for these systems.

The root mean squared error for REVO and cutoff-REVO are also shown in Table 4 as compared to both the computational reference and experimental benchmarks. On average, the REVO calculations agree with the computational reference to within 0.74 kcal/mol, which is slightly lower than the cutoff-REVO RMSE of 1.2 kcal/mol. The difference between these RMSE values is comparable to the uncertainties of the free energy estimates. The lower computational reference RMSE for REVO is primarily due to better performance on the OA-G6 system, which is higher for cutoff-REVO (−5.21 kcal/mol). Conversely, the RMSE measured to experimental benchmarks is slightly lower for cutoff-REVO (2.5 kcal/mol) compared to standard REVO (2.9 kcal/mol). However, both of these values are driven higher by large over-estimates of the free energy for CB8-G3. As this is consistent with the computational reference, this suggests that force-field inaccuracies or some other feature of the computational model (e.g., protonation states, oligomerization states) is inconsistent with the experimental conditions. If this system is removed the RMSE drops to 1.7 kcal/mol for standard REVO and 0.77 kcal/mol for cutoff-REVO.

4 | DISCUSSION

The predicted free energies of unbinding from the cutoff-REVO algorithm improved upon those predicted by REVO in both their consistency across runs and their agreement with benchmark values. For OA-G3 and CB8-G3 the agreement with the computational reference values was improved, significantly in the case of OA-G3. For CB8-G3, both REVO and cutoff-REVO showed good agreement with computational reference values, although the consistency across runs was significantly improved for cutoff-REVO. For the OA-G6 system, although agreement with the computational reference decreased for cutoff-REVO, the ΔG moved closer to the experimental value. We are only cautiously optimistic about the agreement with experiment in this case, as we also observe the largest uncertainties in ΔG for this system. In general, we view agreement with computational reference values as more meaningful, as they separate out inaccuracies of the forcefield.

We also prospectively predicted the binding free energy for the β-CD-PMZ system and observe excellent agreement with the experimental quantity. This was achieved in a "blinded" fashion, where we did not have knowledge of the quantity at the time of prediction. Remarkably, the results in Figure 7A show that 4 of the 5 runs converged closely to the experimental quantity by the second half of the simulation, and the ensemble estimates of both REVO and cutoff-REVO agreed with the experimental value to within 0.1 kcal/mol. While we are encouraged by this agreement, we are again cautiously optimistic, as it is possible this resulted from fortuitous cancelation of error. However, a more complete assessment of REVO and cutoff-REVO requires prediction across a larger set of ligands, such as the complete ligand set in the SAMPL9 challenge, but this was not feasible in this work due to the computational costs involved.

Another factor to consider for the β-CD-PMZ system is the bound starting structure used to initialize the unbinding simulations. This was unknown beforehand and was generated using short, straight-forward simulations, initialized in the unbound state. While β-CD has a ring-like structure similar to CB8, it is not C2-symmetric around the horizontal axis, allowing for distinct bound poses in the top and bottom of the ring. In addition, the PMZ ligand has many feasible modes of interaction with β-CD in addition to that shown in Figure 2, with the dimethylamino group inserted into the ring. In previous work on SAMPL6 systems, we found that five different starting poses were "kinetically indistinguishable": they quickly interconverted with one another on a timescale that was much faster than the unbinding process.[17] In contrast, our previous work on the PK-11195 ligand dissociating from the TSPO membrane protein showed six poses that were not all able to interconvert before unbinding, and thus had pose-specific unbinding rates.[37] At this point, it is unclear which scenario applies to β-CD-PMZ, although we plan to address this in more detail in future work.

The cutoff-REVO approach is slightly more cumbersome to implement than standard REVO, due to the necessity of running a standard simulation to generate the data necessary to determine both a cutoff feature and its correspondingly value (if this information is not already known). However, we expect it to improve the prediction of free energies using complete transition paths in a wide variety of systems. The feature used for cloning eligibility is highly flexible and can be whatever physical aspect of the system is relevant to the process that the user wants to observe. For instance, in ligand unbinding examples that involve more flexible binding sites, it might be more appropriate to use the number of protein-ligand interactions in the eligibility function. Other examples could be the RMSD to a reference structure and interaction energies. We expect that this approach would be beneficial for any process that can be decomposed into two phases: (1) climbing a steep energetic barrier, where resampling is highly important, and (2) diffusion on a flatter energy landscape, where resampling is not important. This could generally include activities such as ligand (un)binding, protein-protein interactions, or large-scale protein conformational changes, among others.

5 | CONCLUSIONS

Restricting the cloning behavior in the REVO sampling method resulted in dramatic changes in the resampling behavior. As a result, the new “cutoff-REVO” approach heavily emphasized cloning of walkers at the early stages of unbinding. Although this affected the number of unbinding events observed by the algorithm, the total weight of the unbound trajectories was higher for all four systems examined here. As such, cutoff-REVO was able to generate unbinding events with weights up to 10^4 times higher than standard REVO for the same system, indicating that these unbinding pathways occur with higher probability. As cutoff-REVO consistently outperformed standard REVO, we determine that for unbinding events, quality matters more than overall quantity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by R01GM130794 from the National Institutes of Health and DMS 1761320 from the National Science Foundation.

DATA AVAILABILITY STATEMENT

The code used to collect and analyze this data are made freely available in a tutorial format [Ref. 40]. A separate tutorial on the visualization of trajectory trees is available in Ref. 51. Data collected by the authors that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]. Cournia Z, Allen B, Sherman W, J. Chem. Inf. Model 2017, 57, 2911. [PubMed: 29243483]
[2]. Zwanzig RW, J. Chem. Phys 1954, 22, 1420.
[3]. Lyubartsev A, Martsinovski A, Shevkunov S, Vorontsov-Velyaminov P, J. Chem. Phys 1992, 96, 1776.
[4]. Jarzynski C, Phys. Rev. Let 1997, 78, 2690.
[5]. Sugita Y, Kitao A, Okamoto Y, J. Chem. Phys 2000, 113, 6042.
[6]. Barducci A, Bussi G, Parrinello M, Phys. Rev. Let 2008, 100, 020603. [PubMed: 18232845]
[7]. Bansal N, Zheng Z, Cerutti DS, Merz KM, J. Comput. Aided. Mol. Des 2017, 1, 47.
[8]. Zuckerman DM, Chong LT, Annu. Rev. Biophys 2017, 46, 43. [PubMed: 28301772]
[9]. You W, Tang Z, Chang C. e. A., J. Chem. Theory Comput 2019, 15, 2433. [PubMed: 30811931]
[10]. Bussi G, Laio A, Nat. Rev. Phys 2020, 2, 200.
[11]. Pohorille A, Jarzynski C, Chipot C, J. Phys. Chem. B 2010, 114, 10235. [PubMed: 20701361]
[12]. de Ruiter A, Oostenbrink C, Curr. Opin. Chem. Biol 2011, 15, 547. [PubMed: 21684797]
[13]. Christ C, Fox T, J. Chem. Inf. Model 2013, 54, 108. [PubMed: 24256082]
[14]. Mobley D, Gilson M, Annu. Rev. Biophys 2017, 46, 531. [PubMed: 28399632]
[15]. Rizzi A, Jense T, Slochower DR, Aldeghi M, Gapsys V, Ntekoumes D, Bosisio S, Papadourkis M, Henriksen NM, de Groot BL, Cournia Z, Dickson A, Michel J, Gilson MK, Shirts MR, Mobley DL, Chodera JD, J. Comput.-Aided Mol. Des 2020, 34, 601. [PubMed: 31984465]
[16]. Rizzi A, Murkli S, McNeill JN, Yao W, Sullivan M, Gilson MK, Chiu MW, Isaacs L, Gibb BC, Mobley DL, Chodera JD, J. Comput.-Aided Mol. Des 2018, 32, 937. [PubMed: 30415285]

[17]. Dixon T, Lotz SD, Dickson A, J. Comput.-Aided Mol. Des 2018, 32, 1001. [PubMed: 30141102]
[18]. Mobley D, Klimovich P, J. Chem. Phys 2012, 137, 230901. [PubMed: 23267463]
[19]. Laio A, Parrinello M, Proc. Nat. Acad. Sci 2002, 99, 12562. [PubMed: 12271136]
[20]. Tiwary P, Mondal J, Berne BJ, Sci. Adv 2017, 3, e1700014. [PubMed: 28580424]
[21]. Nishikawa N, Han K, Wu X, Tofoleanu F, Brooks BR, J. Comput.-Aided Mol. Des 2018, 32, 1075. [PubMed: 30324304]
[22]. Kästner J, Wiley. Inter. Rev. Comput. Mol. Sci 2011, 1, 932.
[23]. Wu H, Fabian P, Wehmeyer C, Noé F, Proc. Nat. Acad. Sci 2016, 113, E3221. [PubMed: 27226302]
[24]. Ge Y, Voelz VA, J. Chem. Phys 2022, 156, 134115. [PubMed: 35395889]
[25]. Spiriti J, Noé F, Wong CF, J. Comput. Chem 2022, 43, 1911. [PubMed: 36073605]
[26]. Huber GA, Kim S, Biophys. J 1996, 70, 97. [PubMed: 8770190]
[27]. Dickson A, Brooks CL III, J. Phys. Chem. B 2014, 118, 3532. [PubMed: 24490961]
[28]. Donyapour N, Roussey N, Dickson A, J. Chem. Phys 2019, 150, 244112. [PubMed: 31255090]
[29]. Zwier MC, Adelman JL, Kaus JW, Pratt AJ, Wong KF, Rego NB, Suárez E, Lettieri S, Wang DW, Grabe M, Zuckerman DM, Chong LT, J. Chem. Theor. Comput 2015, 11, 800.
[30]. Roussey NM, Dickson A, J. Chem. Phys 2020, 153, 134116. [PubMed: 33032408]
[31]. Zwier MC, Pratt AJ, Adelman JL, Kaus JW, Zuckerman DM, Chong LT, J. Phys. Chem. Lett 2016, 7, 3440. [PubMed: 27532687]
[32]. Bhatt D, Zhang BW, Zuckerman DM, J. Chem. Phys 2010, 133, 014110. [PubMed: 20614962]
[33]. Copeland RA, Nat. Revi. Drug. Discov 2016, 15, 87.
[34]. Costa B, Da Pozzo E, Giacomelli C, Barresi E, Taliani S, Da Settimo F, Martini C, Sci. Rep 2016, 6, 18164. [PubMed: 26750656]
[35]. Sztain T, Ahn S-H, Bogetti AT, Casalino L, Goldsmith JA, Seitz E, McCool RS, Kearns FL, Acosta-Reyes F, Maji S, Mashayekhi G, McCammon JA, Ourmazd A, Frank J, McLellan JS, Chong LT, Amaro RE, Nat. Chem 2021, 13, 963. [PubMed: 34413500]
[36]. Lotz SD, Dickson A, J. Am. Chem. Soc 2018, 140, 618. [PubMed: 29303257]
[37]. Dixon T, Uyar A, Ferguson-Miller S, Dickson A, Biophys. J 2021, 120, 158. [PubMed: 33221248]
[38]. Hall R, Dixon T, Dickson A, Front. Mol. Biosci 2020, 7, 106. [PubMed: 32582764]
[39]. Roussey NM, Dickson A, Front. Mol. Biosci 2022, 9, 858316. [PubMed: 35558558]
[40]. CutoffREVO-Tutorial, https://gitlab.com/ADicksonLab/cutoff-revo-tutorial (2022).
[41]. Dixon T, MacPherson D, Mostofian B, Dauzhenka T, Lotz S, McGee D, Shechter S, Shrestha UR, Wiewiora R, McDargh ZA, Pei F, Pal R, Ribeiro JV, Wilkerson T, Sachdeva V, Gao N, Jain S, Sparks S, Li Y, Vinitsky A, Zhang X, Razavi AM, Kolossváry I, Imbriglio J, Evdokimov A, Bergeron L, Zhou W, Adhikari J, Ruprecht B, Dickson A, Xu H, Sherman W, Izaguirre JA, Nat. Commun 2022, 13, 5884. [PubMed: 36202813]
[42]. Lotz SD, Dickson A, ACS Omega 2020, 5, 31608. [PubMed: 33344813]
[43]. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP, PLoS. Comput. Bio 2017, 13, 1.
[44]. Jo S, Kim T, Iyer VG, Im W, J. Comput. Chem 2008, 29, 1859. [PubMed: 18351591]
[45]. Brooks BR, B. CL III, M. AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M, J. Comput. Chem 2009, 30, 1545. [PubMed: 19444816]
[46]. Lee J, Cheng X, Swails J, Yeom M, Eastman P, Lemkul J, Wei S, Buckner J, Jeong J, Qi Y, Jo S, Pande VS, Case DA, Brooks CL III, MacKerell AD Jr, Klauda JB, Im W, J. Chem. Theory Comput 2015, 12, 405. [PubMed: 26631602]
[47]. Dickson A, Warmflash A, Dinner AR, J. Chem. Phys 2009, 131, 154104. [PubMed: 20568844]
[48]. Dickson A, Maienschein-Cline M, Tovo-Dwyer A, Hammond JR, Dinner AR, J. Chem. Theory Comput 2011, 7, 2710. [PubMed: 26605464]

[49]. Vanden-Eijnden E, Venturoli M, J. Chem. Phys 2009, 131, 044120. [PubMed: 19655850]
[50]. Suárez E, Lettieri S, Zwier MC, Stringer CA, Subramanian SR, Chong LT, Zuckerman DM. J. Chem. Theory Comput 2014, 10, 2658. [PubMed: 25246856]
[51]. ResamplingTree-Tutorial, https://gitlab.com/nmroussey/resampling-tree-tutorial (2022).
[52]. Murkli S, McNeill JN, Isaacs L, Supra. Chem 2018, 31, 150.
[53]. SAMPL9-Github, https://github.com/samplchallenges/SAMPL9 (2022).
[54]. Humphrey W, Dalke A, Schulten K, J. Mol. Graph 1996, 14, 33. [PubMed: 8744570]

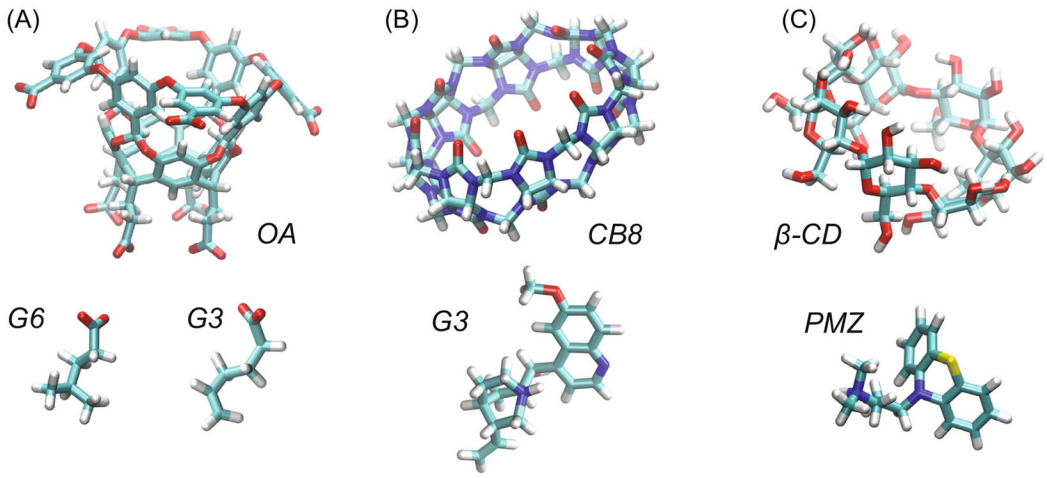

FIGURE 1.

Host-guest systems. Hosts are shown on top, and guests are shown on bottom. (A) The OA host and its guests, G6 (left) and G3 (right). (B) The CB8 host and its ligand, G3. (C) The *β*-CD host and its ligand, PMZ. All panels use the default “atom name” coloring scheme from VMD[54]: cyan = carbon, white = hydrogen, red = oxygen, blue = nitrogen, and yellow = sulfur.

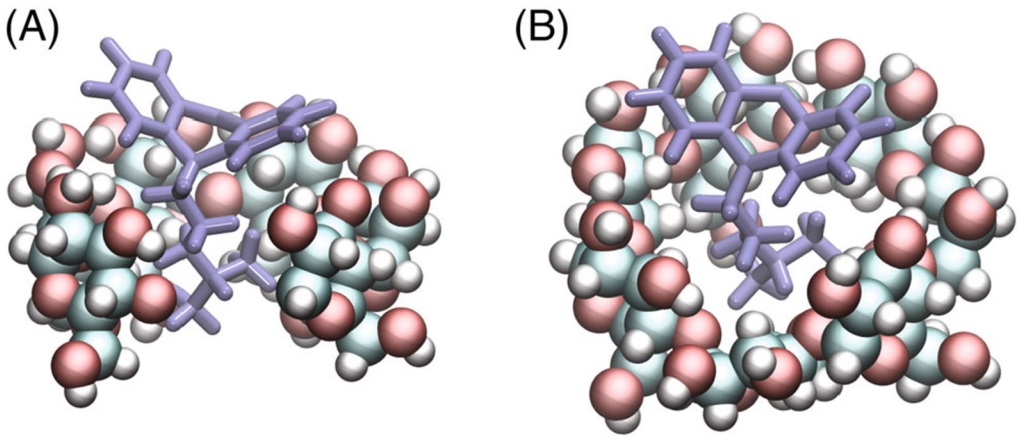

FIGURE 2.

The β-CD-PMZ starting pose. (A) and (B) Two different views of the binding pose obtained where the dimethylamino tail of PMZ is in the binding pocket of β-CD. Some atoms have been removed from β-CD in A for clarity. Both panels use the default "atom name" coloring scheme from VMD[54] for the host molecule: cyan = carbon, white = hydrogen, and red = oxygen.

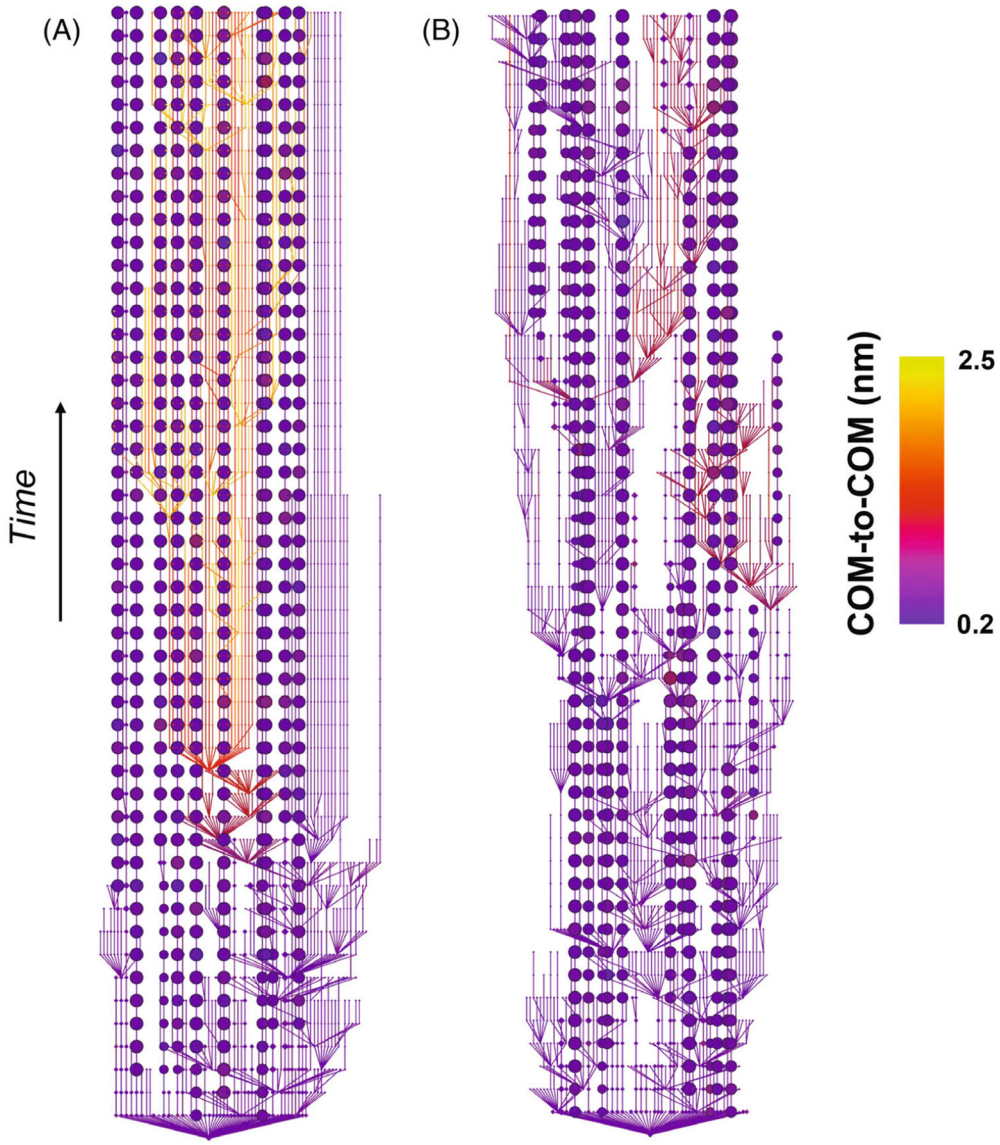

FIGURE 3.

Resampling trees for (A) REVO and (B) cutoff-REVO. Each tree shows the cloning and merging behavior during their respective unbinding simulation, where each node represents a walker at a given point in time. The trees move forward in time from bottom to top. Nodes and edges are colored according to d_{COM} distance. Data from β-CD-PMZ. For clarity, only the first 50 cycles are shown from a single run. Similar behavior was observed in other runs.

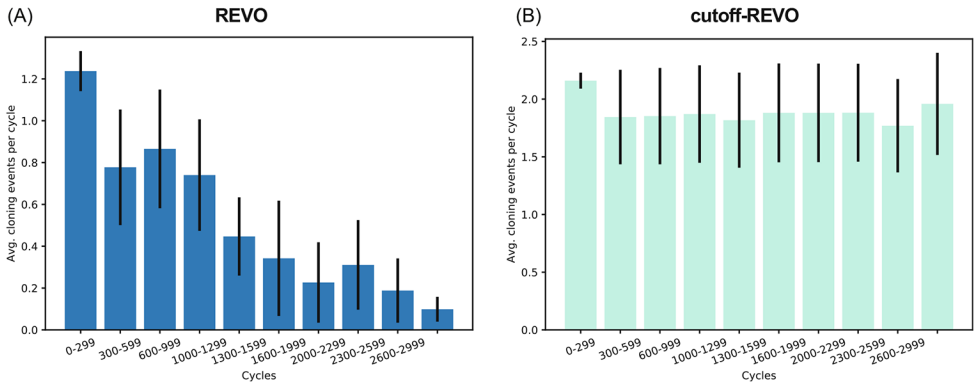

FIGURE 4.

$\langle n_{wc} \rangle$ over time for β-CD-PMZ using (A) REVO and (B) cutoff-REVO. The average number of cloning events per cycle ($\langle n_{wc} \rangle$) broken down by simulation time (300 cycles per bin). Values are averaged over 5 runs and error are the standard error of the mean.

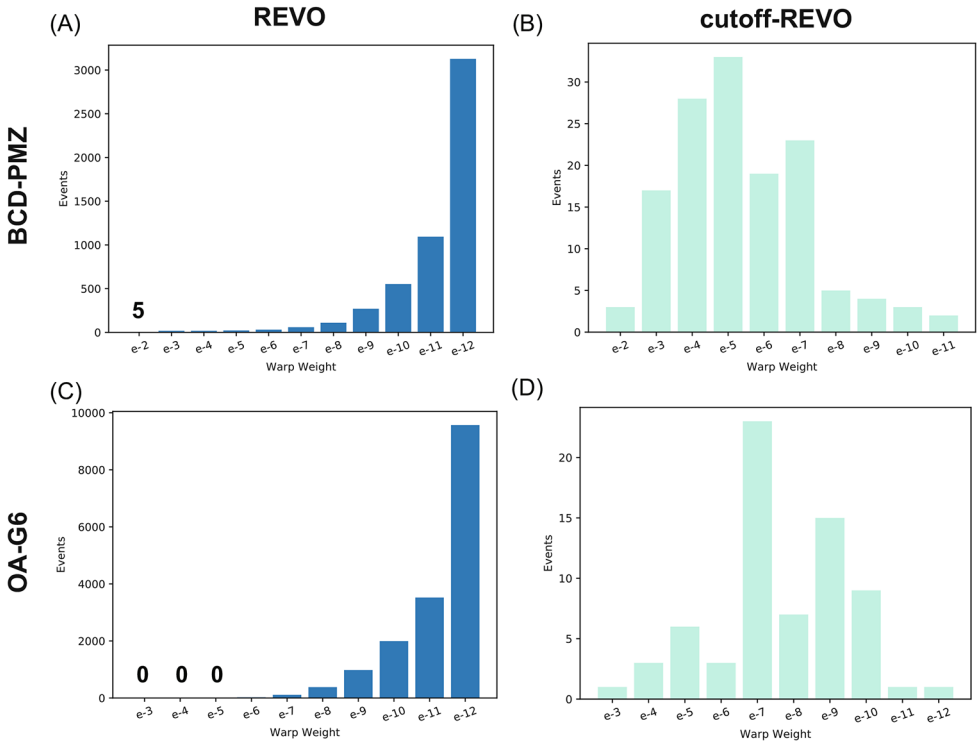

FIGURE 5.

Unbinding walker weights for both methods. Walker weights at the time of unbinding for standard REVO (left) and cutoff-REVO (right) for the β-CD-PMZ system (A and B) and the OA-G6 system (C and D). The standard REVO algorithm showed a much larger quantity of unbinding events. The total number of unbinding events in each panel is as follows: (A) 5303, (B) 137, (C) 16,597, and (D) 69.

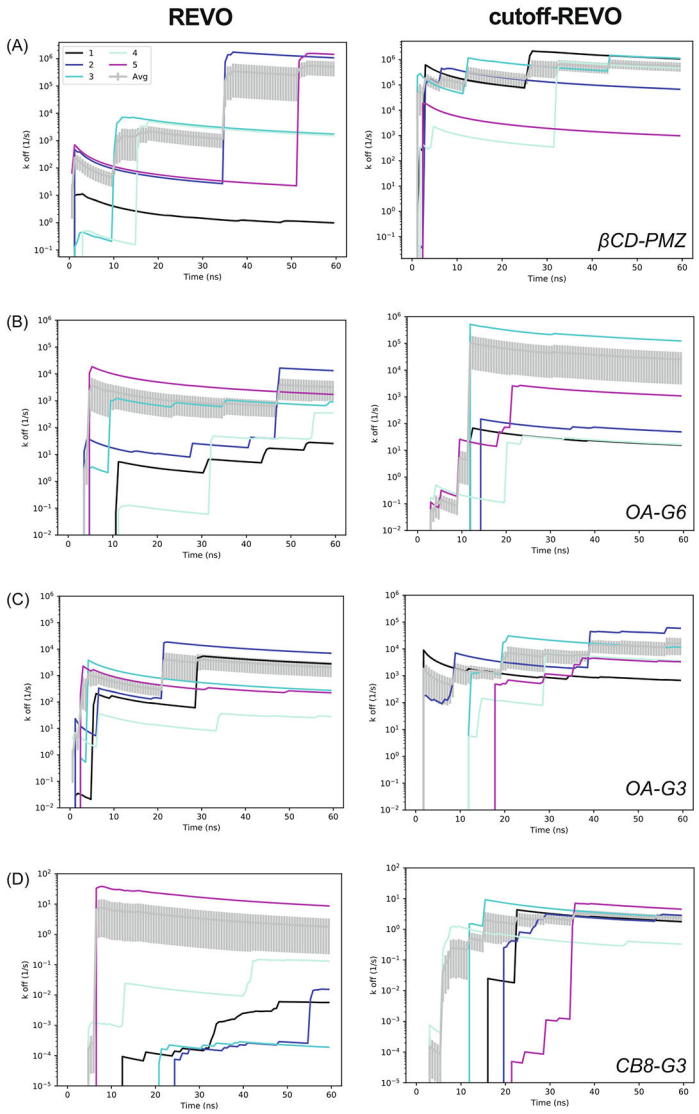

FIGURE 6.

Comparison of unbinding rates for REVO and cutoff-REVO for all systems. Individual runs are shown as thin lines. Average unbinding rates are shown as thick gray lines and uncertainties, computed using the standard error of the mean across the 5 runs, are shown as the shaded gray regions. (A) *β*-CD-PMZ. (B) OA-G6. (C) OA-G3. (D) CB8-G3. The horizontal axis represents trajectory length and not aggregate simulation time.

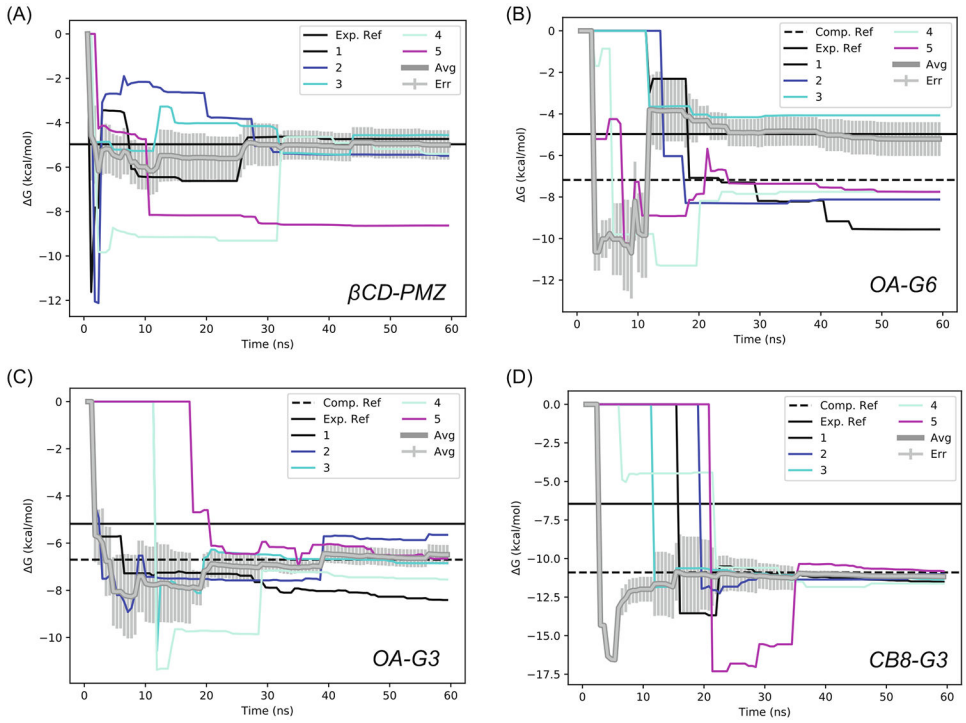

FIGURE 7.

Prediction of cutoff-REVO unbinding free energy for all systems. Individual runs are shown as thin lines. Average free energies are shown as thick gray lines and uncertainties, computed using the standard error of the mean across the 5 runs, are shown as shaded gray regions. (A) *β*-CD-PMZ, (B) OA-G6, (C) OA-G3, and (D) CB8-G3 systems. The horizontal axis represents trajectory length and not aggregate simulation time.

TABLE 1

The average number of walkers cloned per cycle ($\langle n_{wc} \rangle$) and the average number of replications made per cloning event ($\langle R \rangle$) is shown. In addition, we show the average ($\langle d_{COM} \rangle_c$), the minimum (Min . d_{COM}) and the maximum (Max . d_{COM}) value of d_{COM} in the set of cloned walkers. Values are shown separately for REVO and cutoff-REVO. These numbers are averaged across five runs and the errors are computed as the standard error of the mean across the set of runs.

Method	$\langle n_{wc} \rangle$	$\langle R \rangle$	$\langle d_{COM} \rangle_c$ **(nm)**	**Min .** d_{COM} **(nm)**	**Max .** d_{COM} **(nm)**
REVO	0.52 ± 0.14	2.71 ± 0.07	1.1 ± 0.008	0.14 ± 0.006	2.3 ± 0.013
cutoff-REVO	1.90 ± 0.38	3.25 ± 0.03	0.35 ± 0.0008	0.10 ± 0.005	0.76 ± 0.004

TABLE 2

The highest exit point weight obtained and the total number of unbinding events observed for all four systems for both standard REVO and cutoff-REVO over five simulations.

	REVO		Cutoff-REVO	
System	**Max. Wt.**	**Total**	**Max. Wt.**	**Total**
β-CD-PMZ	10^{-2}	5303	10^{-2}	137
OA-G6	10^{-7}	16,597	10^{-3}	69
OA-G3	10^{-6}	22,901	10^{-3}	224
CB8-G3	10^{-8}	1246	10^{-7}	93

TABLE 3

The off-rates from REVO and cutoff-REVO simulations. All k_{off} values are in s^{-1}. The mean first passage times are in ms. The uncertainties are calculated as the standard error of the mean of the log of k_{off}. They can be interpreted as the breadth of the uncertainty in orders of magnitude.

System	k_{off} (REVO)	REVO error	k_{off} (c-REVO)	c-REVO error
OA-G6	3.3×10^3	0.89	2.5×10^4	1.48
OA-G3	2.1×10^3	0.85	1.5×10^4	0.64
CB8-G3	1.77	1.54	2.38	0.39
β-CD-PMZ	5.0×10^5	2.26	5.5×10^5	1.15

TABLE 4

Binding free energies calculated for all systems, as well as experimental and computational reference values, where available. Computational reference values were calculated using extensive Hamiltonian Replica Exchange simulations (Ref. 15). All ΔG values are in kcal/mol.

System	ΔG (REVO)	ΔG (c-REVO)	ΔG (comp. Ref.)	ΔG (exp.)
OA-G6	−6.55 ± 0.53	−5.21 ± 0.81	−7.18 ± 0.06	−4.97 ± 0.02
OA-G3	−7.69 ± 0.55	−6.49 ± 0.41	−6.71 ± 0.05	−5.18 ± 0.02
CB8-G3	−11.35 ± 1.01	−11.17 ± 0.13	−10.80 ± 0.20	−6.45 ± 0.06
β-CD-PMZ	−5.05 ± 1.51	−5.00 ± 0.66	N/A	−4.97 ± 0.02
RMSE (comp.)	0.74	1.2	–	–
RMSE (exp.)	2.9	2.5	–	–