A Minimal, Adaptive Binning Scheme for Weighted Ensemble Simulations

Paul A. Torrillo*[,1], Anthony T. Bogetti*[,1], and Lillian T. Chong[1]
*Equal contributors

[1]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

A promising approach for simulating rare events with rigorous kinetics is the weighted ensemble path sampling strategy. One challenge of this strategy is the division of configurational space into bins for sampling. Here we present the minimal adaptive binning (MAB) scheme for the automated, adaptive placement of bins along a progress coordinate within the framework of the weighted ensemble strategy. In this scheme, a fixed number of bins are evenly spaced between the positions of trailing and leading trajectories along a progress coordinate after each resampling interval. To further enhance the sampling of bottleneck regions, additional bins are dedicated to trajectories residing in the least visited regions of the free energy landscape. On average, bins are more finely spaced along steeper regions of the landscape and further apart along flatter regions of the landscape. We applied the scheme to simulations of the following processes, listed in order of increasing complexity: (i) transitions between states of a double-well toy potential; (ii) association of Na^+ and Cl^- ions; and (iii) conformational sampling of a tumor suppressor p53 peptide fragment. Results reveal that our binning scheme, despite its simplicity, efficiently estimates rate constants for rare-event processes with large free energy barriers and enhances conformational sampling.

1. Introduction

Over the past two decades, path sampling strategies have enabled the simulation of rare events such as large protein conformational transitions, protein (un)folding, and protein un(binding) by focusing computational effort on the functional transitions rather than on the stable states.[1] In particular, such strategies exploit the fact that for rare events, the time required to cross a free energy barrier (t_b) is much shorter than the dwell time in the preceding stable (or metastable) state ($t_b << t_{dwell}$) and can therefore be orders of magnitude more efficient than standard simulations in sampling rare events without introducing any bias into the dynamics.

A major challenge of path sampling strategies has been the division of configurational space, *e.g.* placement of bins along a chosen progress coordinate. Schemes that automate the adaptive placement of bins can therefore steamline the use of these strategies. A number of adaptive binning schemes have been developed for path sampling strategies. These schemes involve (i) interfaces as “bins” to improve flux through “bottleneck” regions in configurational space in forward flux sampling[2] and transition interface path sampling;[3] (ii) Voronoi bins in milestoning,[4] and the weighted ensemble (WE) strategy;[5,6] (iii) a history-augmented Markov State Model that yields microstates as “bins” in the context of the WE strategy;[7] and (iv) a variance-reduction approach to automate the placement of bins for the WE strategy.[8]

Here, we present the minimal adaptive binning (MAB) scheme for the WE strategy. This scheme places a fixed number of evenly spaced bins between "boundary" trajectories: the trailing and leading trajectories along the progress coordinate. After laying down these initial bins, the scheme identifies trajectories in "bottleneck" regions where their associated weights (probabilities) have plummeted. The scheme then assigns each boundary and bottleneck walker to a separate bin. To provide even coverage of configurational space and maintain a target number of trajectories per bin, the WE strategy replicates and prunes trajectories within the same bin at fixed time intervals (WE iterations). Therefore, on average, more trajectories are replicated in steeper regions of the free energy landscape.

The scheme can be used with high-dimensional progress coordinates and exhibits the following features: (i) no prior test simulations or training sets are required as the scheme relies only on the positions of the trailing and leading trajectories along the progress coordinate at chosen fixed time intervals; (ii) the scheme requires a fewer number of bins than required by a manual binning scheme due to earlier identification of bottlenecks along the progress coordinate; and (iii) the maximum number of CPUs or GPUs required is more easily estimated prior to running the simulation since a similar number of bins is used throughout the simulation; and (iv) the scheme is easily extensible to include more sophisticated schemes for adaptive binning.

To demonstrate the power of the adaptive binning scheme, we applied the algorithm to simulations of the following processes, in order of increasing complexity: (i) transitions between states in a double-well toy potential, (ii) molecular association of the Na^+ and Cl^- ions, and (iii) conformational transitions of an N-terminal peptide fragment of the p53 tumor suppressor.

2. Theory

The weighted ensemble (WE) strategy. The WE strategy involves running many trajectories in parallel and applying a resampling procedure at fixed time intervals, replicating trajectories that advance towards a target state.[9,10] The goal is to populate empty bins in configurational space -- typically along a progress coordinate. Importantly, the rigorous tracking of trajectory weights ensures that *no bias is introduced into the dynamics,* thereby enabling the calculation of non-equilibrium observables such as rate constants. Furthermore, since the trajectory weights are independent of the progress coordinate, the progress coordinate as well as bin positions can be adjusted "on-the-fly" during a WE simulation.[5]

Weighted ensemble simulations can be carried out under steady-state or equilibrium conditions.[11] Under steady-state conditions, trajectories that reach the target state are "recycled" by terminating the trajectories and starting a new trajectory from the initial state with the same statistical weight. Under equilibrium conditions, no trajectories are recycled, which means that target states need not be strictly defined in advance of the simulation.

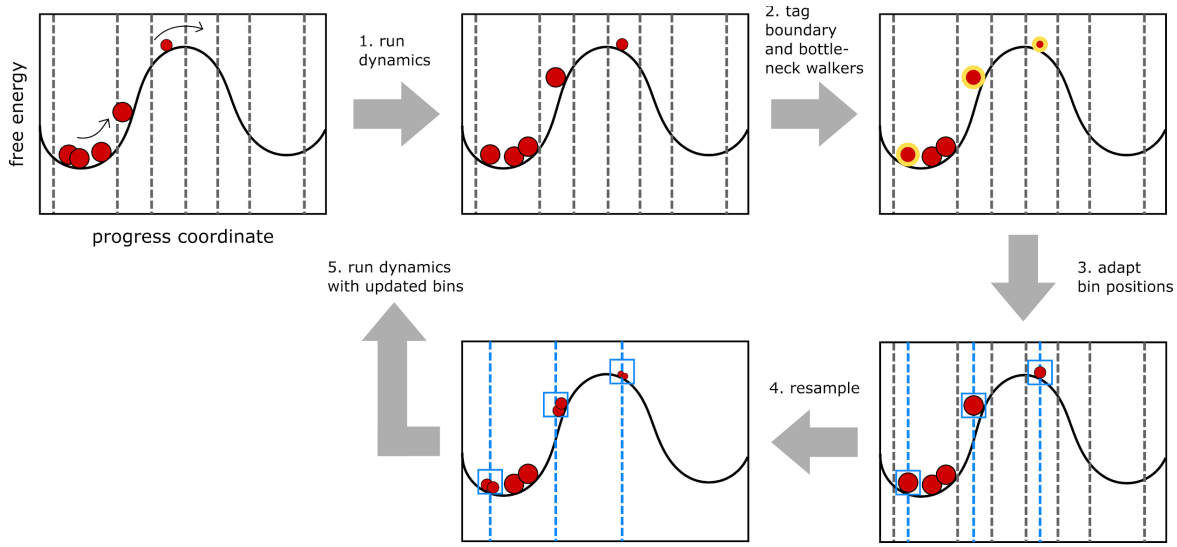

Figure 1: Illustration of the MAB Scheme for adaptive placement of bins along a one-dimensional progress coordinate. The scheme involves four steps. (1) Run dynamics for a short, fixed time interval τ using initially spaced bins indicated by gray vertical lines. Trajectories are represented by red circles with sizes that are proportional to their statistical weights. (2) Tag boundary and bottleneck trajectories (highlighted in gold). (3) Adapt bin positions (blue vertical lines) by placing a fixed number of bins evenly between the positions of the trailing and leading trajectories along the progress coordinate and assigning each boundary and bottleneck trajectory to a separate bin (blue boxes). (4) Replicate and prune trajectories to maintain a target number of trajectories per bin. Repeat steps 1-4 with updated bin positions until a desired amount of sampling is achieved.

The MAB Scheme. The minimal adaptive binning (MAB) scheme adapts the positions of bins along a progress coordinate after each WE iteration in two stages: 1) place a fixed number of evenly spaced bins between the "boundary" trajectories, *i.e.* trailing and leading trajectories along each dimension of the progress coordinate and 2) identify up to two "bottleneck" trajectories and assign each boundary and bottleneck trajectory to a separate bin. A bottleneck trajectory is defined as one that has reached a maximum in free energy units (-RTlnP where P is the probability, or statistical weight, of the trajectory) relative to the combined probabilities of all trajectories that are further along the progress coordinate. After each WE iteration, a resampling step is applied to maintain a target number of trajectories per bin. This step involves replicating or pruning trajectories within each bin, splitting or merging the trajectory weights, respectively, according to rigorous statistical rules.[5] On average, the scheme places bins more finely along steeper regions and further apart along flatter regions of the free energy landscape.

Figure 1 illustrates the steps of the MAB scheme for a one-dimensional progress coordinate:

1. Run dynamics for a fixed interval τ.
2. Tag each boundary and bottleneck trajectory.

3. Adapt bin positions by dividing the progress coordinate evenly into a specified number of bins between the positions of the trailing and leading trajectories identified in step 2; assign trailing, leading, and bottleneck trajectories to their own bins.
4. Complete the WE iteration by replicating or combining trajectories to yield a target number of trajectories in each bin.
5. Repeat steps 1-4.

For a multi-dimensional progress coordinate, steps 2 and 3 are carried out for each dimension of the progress coordinate. To avoid the replication of trajectories outside of the desired configurational space (*e.g.,* regions of unintentional protein unfolding), the MAB scheme includes the option to specify minimum and/or maximum limits of an additional dimension to the progress coordinate (or other observables) for the replication of trajectories.

3. Methods

3.1. WE simulations. All WE simulations were carried out using the open-source WESTPA software package.[12] For each benchmark system, we compared the efficiency of the MAB scheme for adaptive binning to a manual binning scheme that used the same number of bins. We present progress coordinates and binning schemes for each benchmark system below.

Double-well toy potential. The double-well toy potential consists of two equally stable states separated by a 34-kT free energy barrier. The potential was defined as

$$V = -60 * cos^2 x + \frac{3.75}{sin^2 x}$$

A one-dimensional progress coordinate was divided into 20 bins along a theoretical X position metric ranging from an initial state A at X=0.5 to a target state B at X=2.5. For each binning scheme, a single 2000-iteration WE simulation was run with a fixed time interval τ of 5×10^{-5} for each iteration and a target number of 5 trajectories/bin.

Dynamics were propagated according to the overdamped Langevin equation:

$$X(t + \delta t) = X(t) - \frac{\delta t}{\gamma} \nabla_X V + \delta X^G$$

where γ is the friction coefficient, δt is the time step, and δX^G is a random displacement with zero mean and variance $2\gamma kT\delta t$ with $\gamma = 0$, $\delta t = 5 \times 10^{-5}$, and $kT = 1$.

Na^+/Cl^- system. To sample Na^+/Cl^- associations in explicit solvent, five independent, steady-state WE simulations were carried out for each of the two binning schemes. A one-dimensional progress coordinate was used consisting of the Na^+/Cl^- distance ranging from a maximum value of 20 Å down to a target state at 2.6 Å with 28 bins. For both binning schemes, 1000 WE iterations

were carried out with a fixed time interval τ of 2 ps for each iteration and a target number of 4 trajectories/bin, yielding an aggregate simulation time of 0.2 μs.

Dynamics of the Na^+/Cl^- system were propagated using the AMBER18 software package[13] with the TIP3P water model[14] and corresponding Joung and Cheatham ion parameters.[15] Simulations were started from an unassociated state with a 12-Å separation between the Na^+ and Cl^- ions and a sufficiently large truncated octahedral box of explicit water molecules to provide a minimum 12 Å clearance between the ions and box walls. Temperature and pressure was maintained at 298 K and 1 atm using the Langevin thermostat (collision frequency of 1 ps^{-1}) and Monte Carlo barostat (with 100 fs between attempts to adjust the system volume), respectively. Non-bonded interactions were truncated at 10 Å and long-range electrostatics were treated using the particle mesh Ewald method.[16]

P53 peptide. To sample alternate conformation of the p53 peptide, a single equilibrium WE simulation was carried out using each of the two binning schemes and a two-dimensional progress coordinate that consisted of (i) heavy-atom RMSD of the peptide from its MDM2-bound, α-helical conformation, and (ii) end-to-end distance of the peptide. For both binning schemes, the WE simulations were carried out using a fixed time interval τ of 50 ps for each iteration and a target number of 4 trajectories/bin. A total simulation time of 2.0 μs was generated for each binning scheme (338 and 200 WE iterations for the MAB and manual binning schemes, respectively). The MAB scheme used a maximum of 44 bins while the manual binning scheme used a maximum of 294 bins that were evenly spaced between an RMSD of 0 and 20 Å and end-to-end distance of 0 to 26 Å. For the MAB scheme, no other limits were specified for the replication of trajectories.

Dynamics of the p53 peptide (residues 17-29) were propagated using the AMBER18 software package[13] with the Amber ff14SBonlysc force field[17] and a generalized Born implicit solvent model (GBneck2 and mbondi3 intrinsic radii).[18] Simulations were started from an energy-minimized conformation of the peptide that was based on the crystal structure of the MDM2-p53 peptide complex (PDB code: 1YCR).[19] Temperature was maintained at 298 K using the Langevin thermostat and a collision frequency of 80 ps^{-1} for water-like viscosity.

3.2. Standard simulations. A total of 5 independent 1-μs standard simulations were carried out for the Na^+/Cl^- system and one 2-μs simulation was carried out for the p53 peptide. Details of dynamics propagation and starting structures for these simulations are the same as those described above for the WE simulations.

3.3. Calculation of rate constants. The rate constant k_{AB} for the Na^+/Cl^- association process was directly calculated from the WE simulation using the following:[20,21]

$$k_{AB} = \frac{1}{MFPT(A \rightarrow B)} = \frac{<f_{AB}>}{<p_A>}$$

where $MFPT(A \longrightarrow B)$ is the mean first-passage time for the $A \rightarrow B$ process, $< f_{AB} >$ is the average flux from states A to B of the steady-state set of trajectories in the direction of interest

(*i.e.*, molecular association), and $< p_A >$ is the average steady-state population of the initial state A. The numerator $< f_{AB} >$ was calculated as a running average from states A to $B$of the steady-state set of trajectories in the direction of interest and the denominator $< p_A >$ was calculated as the running average of the steady-state population of the initial state A over all values at fixed time intervals τ. Uncertainties in the rate constant represent 95% confidence intervals. These uncertainties were estimated in two steps: (i) applying a Monte Carlo bootstrapping technique to determine the 95% confidence interval for the rate constant from each of the five simulations; and (ii) applying a one-sample t-test to determine the 95% confidence interval for the average rate constant among all five simulations. In step (i), each bootstrap involved drawing 1000 datasets (with replacement).

4. Results

We demonstrate the power of our minimal adaptive binning (MAB) scheme in sampling rare-event processes compared to fixed, manual binning schemes. We applied this scheme to the following processes which are listed in order of increasing complexity: (i) transitions between stable states in a double-well toy potential, (ii) molecular associations of the Na^+ and Cl^- ions, and (iii) conformational sampling of a peptide fragment of tumor suppressor p53.

4.1. Simulations with a double-well toy potential

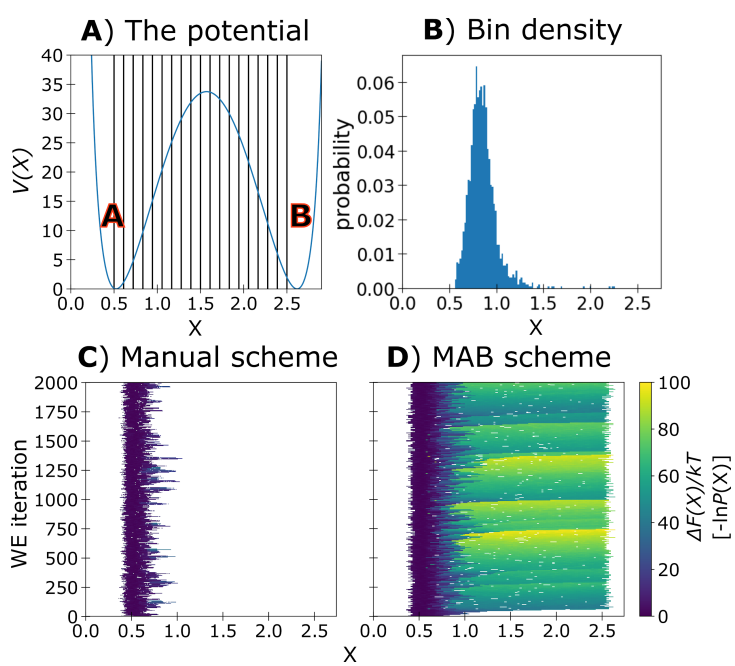

Figure 2. Transitions between stable states of a double-well toy potential: A) The double-well potential with the manual binning scheme indicated by vertical lines. B) Probability distribution of bottleneck walkers identified by the MAB scheme. C) Probability distributions as a function of the one-dimensional progress coordinate. D) Probability distributions as a function of the WE iteration for the MAB scheme.

To test how effectively the MAB scheme computes rate constants for a process with a large free energy barrier, we simulated a double-well potential in which the two equally stable states are separated by a 34 kT barrier. WE simulations with ac manual binning scheme (see **Fig. 2A** for bin positions) yielded no pathways from the initial state at X=0.5 to the target state at X=2.5 after 2000 WE iterations, occupying only 14% of the bins (**Fig. 2B**). In contrast, the MAB scheme generated pathways to the target state in 60 WE iterations, occupying 99% of the bins (**Fig. 2C**). This greater efficiency is due to the identification of bottleneck regions where the trajectory weights have sharply fallen. These regions correspond to the upward slope of the free energy barrier (**Fig. 2D**). In contrast, a variance-reduction strategy has determined that the set of bins that yields the lowest variance in an observable of interest consists of finer binning in transition-state regions.[8]

We also determined how the number of bins impacted the extent to which the MAB scheme sampled transitions from the initial state to the target state. **Table 1** shows computed rate constants and number of successful pathways for different numbers of bins (5, 10, 15, and 20). The number of bins does not affect the value or statistical precision of the rate constant. However, increasing the number of bins yields a greater number of successful pathways due to more complete coverage along the progress coordinate, particularly in the bottleneck regions (**Fig. S1**).

Table 1: Average probability flux from states A to B in the double well potential from WE simulations using the MAB scheme with different numbers of bins. Also indicated is the number of successful pathways from states A to B.

bins	**rate constant kAB(δt^{-1})**	**# successful pathways**
5	$(1.0 \pm 0.9) \times 10^{-16}$	1332
10	$(8.0 \pm 4.0) \times 10^{-16}$	3036
15	$(7.4 \pm 3.9) \times 10^{-17}$	4035
20	$(1.4 \pm 0.6) \times 10^{-16}$	5614

4.2. Simulations of the Na^+/Cl^- association process

To determine the effectiveness of the MAB scheme for a relatively fast process (ns timescale), we simulated the Na^+/Cl^- association process in explicit solvent (**Fig. 3A**). Given the modest free energy barrier for this process, it was feasible to compute association rate constants using standard simulations, providing validation of the rate constants computed using WE simulations and manual/MAB binning schemes.

Table 2 shows that the WE simulations yield rate constants that are within error of the value from standard simulations (see also **Fig. S2**). The WE simulations are twice as efficient as standard simulations in estimating these rate constants. Furthermore, for this process with a modest free energy barrier, there is no gain in efficiency for the MAB scheme relative to the manual binning scheme (see **Fig. 3B** for bin positions). Nonetheless, consistent with our results for the double-well toy potential, the majority (60%) of the bottleneck trajectories occupied bins along the upward slope of the free energy barrier; the remaining bottleneck trajectories (40%) occupied bins near the target state (**Fig. 3C**).

Although the MAB scheme had little impact on the efficiency of rate-constant estimation for the Na^+/Cl^- association process, the scheme increased the number of successful pathways for this process (**Table 2**). Relative to the manual binning scheme, the MAB scheme resulted in a 2-fold gain in the number of successful pathways.

Table 2. Computed rate constants for the Na^+/Cl^- association process using WE and standard simulations. Uncertainties represent 95% confidence intervals calculated from a one-parameter t-test. A total of five WE simulations with a total simulation time of 0.2 μs were carried out for each binning scheme; five 1-μs standard simulations were also carried out. The efficiency S_k of WE relative to standard simulations was calculated as described by others.[9,22]

type of simulation	**computed rate constant ($M^{-1}s^{-1}$)**	**total simulation time (μs)**	**S_k**	**# successful pathways**
WE simulations with manual binning scheme	$(3.8 \pm 0.6) \times 10^{10}$	1.02	1.65	1226
WE simulations with MAB scheme	$(3.9 \pm 0.5) \times 10^{10}$	0.98	2.32	2498
standard simulations	$(3.7 \pm 0.3) \times 10^{10}$	5.00		

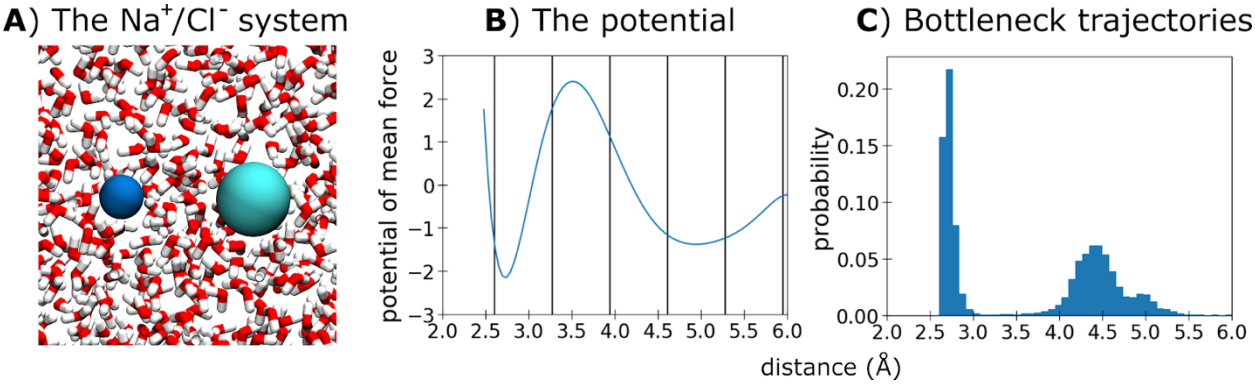

Figure 3. A) The Na+/Cl- system in explicit solvent. B) Potential of mean force for the Na+/Cl- association process with bin positions for the manual scheme indicated by vertical lines. C) Probability distribution of the positions of bottleneck trajectories tagged by the MAB scheme along the progress coordinate.

4.3 Conformational sampling of the p53 peptide

Given that the WE strategy has enhanced the conformational sampling of various systems,[23,24] we applied the MAB scheme to the conformational sampling of a p53 peptide. As expected, WE simulations using either the MAB scheme or the previously reported manual binning scheme[25] yielded greater coverage of configurational space than standard simulations with the same total computing time (**Fig. 4B-C**). The MAB scheme resulted in occupation of 66% of the specified bins (29 out of 44 bins) compared to only 17% (50 out of 294 bins) for the manual binning scheme. Notably, the MAB scheme sampled a “horn shaped” region with a relatively low probability; this region is not accessed by WE simulations using the manual binning scheme (or standard simulations) and consists of right-handed, helical conformations (**Fig. 4D-F**).

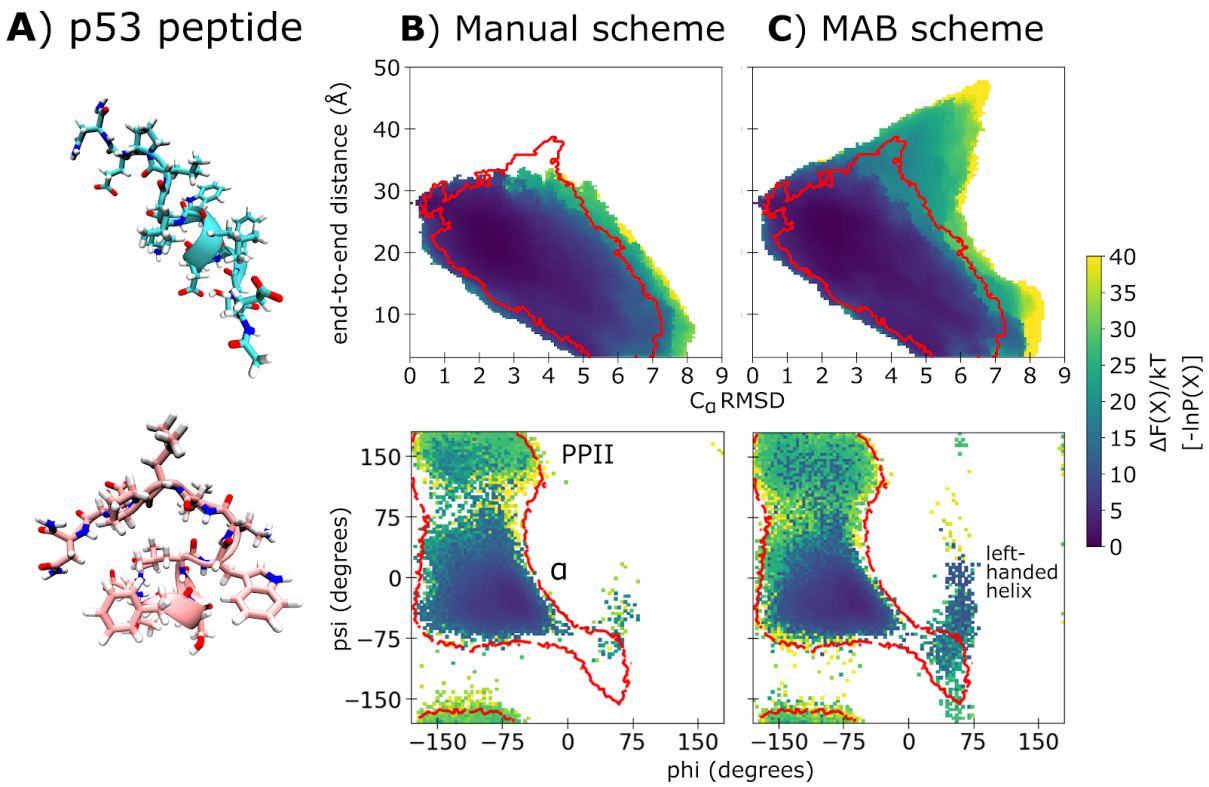

Figure 4. Conformational sampling of a p53 peptide. Probability distributions as a function of the two-dimensional WE progress coordinate from simulations of the p53 peptide (residues 19-23). **A)** starting conformation (cyan) of the WE simulation, extracted from the crystal structure of the MDM2-p53 peptide complex[19] and a representative, right-handed helical conformation sampled using the MAB scheme. **B)** WE simulations using the manual binning scheme. **C)** WE simulations using the MAB scheme. Also shown are Ramachandran plots of the p53 peptide for the manual binning scheme and MAB scheme. Regions sampled by standard simulations are delineated in red.

5. Conclusions

To streamline the execution of weighted ensemble simulations, we developed a minimal adaptive binning (MAB) scheme that adjusts the positions of bins along a progress coordinate based on the updated positions of trailing and leading trajectories as well as trajectories in “bottleneck” regions. Despite its simplicity, the MAB scheme results in greater coverage of configurational space relative to manual binning schemes for all three benchmark processes of this study: (i) transitions between states of a double-well toy potential; (ii) Na+/Cl- association; and (iii) conformational sampling of a peptide fragment of tumor suppressor p53. Furthermore, the MAB scheme enables the generation of pathways for processes involving free energy barriers that are prohibitively large for manual binning schemes (*i.e.* transitions involving the double-well potential) and the generation of more diverse pathways for processes with relatively modest free energy barriers. Due to the earlier identification of bottlenecks along the progress coordinate, the MAB scheme requires dramatically fewer numbers of bins than a manual binning scheme. Since the

number of bins is similar throughout the simulation, it is straightforward to estimate the number of CPUs or GPUs required for running the simulation. Taken together, the above results indicate that the MAB scheme is effective in automating the adaptive binning of progress coordinates for the weighted ensemble strategy. In addition, the scheme could be useful for other rare-event sampling strategies that require a progress coordinate and could be a quick means of obtaining an initial set of promising bins for further optimization by more sophisticated binning strategies.

Acknowledgements

This work was supported by the NIH (1R01GM115805-01) and NSF (CHE-1807301) to L.T.C., and the University of Pittsburgh to P.A.T. (Honors College Brackenridge Undergraduate Research Fellowship) and A.T.B. (Arts and Sciences Graduate Fellowship). Computational resources were provided by the University of Pittsburgh's Center for Research Computing.

References

(1) Chong, L. T.; Saglam, A. S.; Zuckerman, D. M. Path-Sampling Strategies for Simulating Rare Events in Biomolecular Systems. *Curr. Opin. Struct. Biol.* **2017**, *43*, 88–94.
(2) DeFever, R. S.; Sarupria, S. Contour Forward Flux Sampling: Sampling Rare Events along Multiple Collective Variables. *J. Chem. Phys.* **2019**, *150* (2), 024103.
(3) van Erp, T. S.; Moroni, D.; Bolhuis, P. G. A Novel Path Sampling Method for the Calculation of Rate Constants. *The Journal of Chemical Physics*. 2003, pp 7762–7774. https://doi.org/10.1063/1.1562614.
(4) Vanden-Eijnden, E.; Venturoli, M. Markovian Milestoning with Voronoi Tessellations. *J. Chem. Phys.* **2009**, *130* (19), 194101.
(5) Zhang, B. W.; Jasnow, D.; Zuckerman, D. M. The "Weighted Ensemble" Path Sampling Method Is Statistically Exact for a Broad Class of Stochastic Processes and Binning Procedures. *J. Chem. Phys.* **2010**, *132* (5), 054107.
(6) Dickson, A.; Brooks, C. L., 3rd. WExplore: Hierarchical Exploration of High-Dimensional Spaces Using the Weighted Ensemble Algorithm. *J. Phys. Chem. B* **2014**, *118* (13), 3532–3542.
(7) Copperman, J.; Zuckerman, D. M. Accelerated Estimation of Long-Timescale Kinetics from Weighted Ensemble Simulation via Non-Markovian "Microbin" Analysis. *J. Chem. Theory Comput.* **2020**. https://doi.org/10.1021/acs.jctc.0c00273.
(8) Aristoff, D.; Zuckerman, D. M. Optimizing Weighted Ensemble Sampling of Steady States. *Multiscale Modeling & Simulation*. 2020, pp 646–673. https://doi.org/10.1137/18m1212100.
(9) Huber, G. A.; Kim, S. Weighted-Ensemble Brownian Dynamics Simulations for Protein Association Reactions. *Biophysical Journal*. 1996, pp 97–110. https://doi.org/10.1016/s0006-3495(96)79552-8.
(10) Zuckerman, D. M.; Chong, L. T. Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annu. Rev. Biophys.* **2017**, *46*, 43–57.
(11) Suarez, E.; Lettieri, S.; Zwier, M. C.; Subramanian, S. R.; Chong, L. T.; Zuckerman, D. M. Simultaneous Computation of Dynamical and Equilibrium Information Using a Weighted Ensemble of Trajectories. *Biophysical Journal*. 2014, p 406a. https://doi.org/10.1016/j.bpj.2013.11.2285.
(12) Zwier, M. C.; Adelman, J. L.; Kaus, J. W.; Pratt, A. J.; Wong, K. F.; Rego, N. B.; Suárez, E.; Lettieri, S.; Wang, D. W.; Grabe, M.; Zuckerman, D. M.; Chong, L. T. WESTPA: An

Interoperable, Highly Scalable Software Package for Weighted Ensemble Simulation and Analysis. *J. Chem. Theory Comput.* **2015**, *11* (2), 800–809.
(13) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. *Amber 18*; University of California, San Francisco, 2018.
(14) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics*. 1983, pp 926–935. https://doi.org/10.1063/1.445869.
(15) Joung, I. S.; Cheatham, T. E., 3rd. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020–9041.
(16) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *The Journal of Chemical Physics*. 1995, pp 8577–8593. https://doi.org/10.1063/1.470117.
(17) Nguyen, H.; Maier, J.; Huang, H.; Perrone, V.; Simmerling, C. Folding Simulations for Proteins with Diverse Topologies Are Accessible in Days with a Physics-Based Force Field and Implicit Solvent. *J. Am. Chem. Soc.* **2014**, *136* (40), 13959–13962.
(18) Nguyen, H.; Roe, D. R.; Simmerling, C. Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J. Chem. Theory Comput.* **2013**, *9* (4), 2020–2034.
(19) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain. *Science* **1996**, *274* (5289), 948–953.
(20) Hill, T. L. Free Energy Transduction and Biochemical Cycle Kinetics. 1989. https://doi.org/10.1007/978-1-4612-3558-3.
(21) Suárez, E.; Pratt, A. J.; Chong, L. T.; Zuckerman, D. M. Estimating First-Passage Time Distributions from Weighted Ensemble Simulations and Non-Markovian Analyses. *Protein Sci.* **2016**, *25* (1), 67–78.
(22) Zwier, M. C.; Kaus, J. W.; Chong, L. T. Efficient Explicit-Solvent Molecular Dynamics Simulations of Molecular Association Kinetics: Methane/Methane, Na(+)/Cl(-), Methane/Benzene, and K(+)/18-Crown-6 Ether. *J. Chem. Theory Comput.* **2011**, *7* (4), 1189–1197.
(23) Zwier, M. C.; Pratt, A. J.; Adelman, J. L.; Kaus, J. W.; Zuckerman, D. M.; Chong, L. T. Efficient Atomistic Simulation of Pathways and Calculation of Rate Constants for a Protein-Peptide Binding Process: Application to the MDM2 Protein and an Intrinsically Disordered p53 Peptide. *J. Phys. Chem. Lett.* **2016**, *7* (17), 3440–3445.
(24) Dickson, A.; Mustoe, A. M.; Salmon, L.; Brooks, C. L., 3rd. Efficient in Silico Exploration of RNA Interhelical Conformations Using Euler Angles and WExplore. *Nucleic Acids Res.* **2014**, *42* (19), 12126–12137.
(25) Bogetti, A. T.; Mostofian, B.; Dickson, A.; Pratt, A. J.; Saglam, A. S.; Harrison, P. O.; Adelman, J. L. L.; Dudek, M.; Torrillo, P. A.; DeGrave, A. J.; Adhikari, U.; Zwier, M. C.; Zuckerman, D. M.; Chong, L. T. A Suite of Tutorials for the WESTPA Rare-Events Sampling Software [Article v1.0]. *Living Journal of Computational Molecular Science*. 2019. https://doi.org/10.33011/livecoms.1.2.10607.