Adaptive sampling for alchemical free energy calculations and applications for drug design

Bruce Macdonald, H. E.[1], Rufa, D. A.[2], Grinaway P. B.[3], Chodera, J. D.[1]

[1]Memorial Sloan Kettering Cancer Center, New York, NY, US

[2]Tri-Institutional Program in Computational Biology and Medicine, New York, NY, US

[3]Onai, 180 Varick St., New York, NY, US

1. Introduction

Binding free energy calculations (BFE's) are routinely used in drug design to accurately predict the binding free energy of small molecules to drug targets,[1] however the cost of simulation often prohibits their application to smaller sets of molecules. Groups of molecules are typically compared through free energy maps, where each ligand may be compared to at least two other small molecules, however the decision process involved in the generation of this map is un-rigorous. Certain calculations between pairs of ligands or even a given atom-mapping protocol will converge faster than others, proportional to the thermodynamic length of the specific transformations. Using perturbations with large thermodynamic lengths is inefficient,[2] however it is not possible to calculate thermodynamic length *a priori*, and can only be established post simulation.[3] Typically, the generation of a ligand free energy map involves naïve reasoning over which pairs of ligands to compare, while the efficiency of the chosen pairings, and therefore the overall quality of the map is only apparent post simulation. Each perturbation is generally simulated using 'equal allocation' whereby the same length of simulation is used for each perturbation.

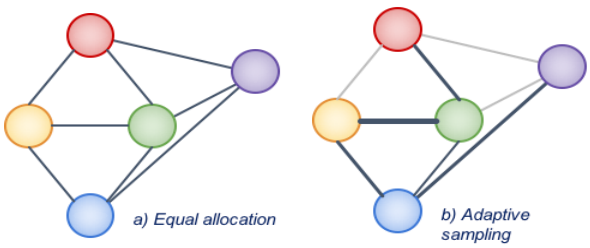

Figure 1 *free energy cycles for a set of ligands using: (a) equal allocation of computational resources and (b) the proposed method of adaptive sampling*

We demonstrate the application of Bayesian bandit framework as an alternative to equal allocation sampling. Bayesian bandits allow for optimal sampling of a given problem following an explore vs. exploit paradigm,[4] whereby initially many options are considered at random, and as information is gained—such as about the thermodynamic length of a perturbation—the sampling of the system can be adjusted accordingly to exploit those that achieve faster convergence. A common example of a Bayesian bandit is A/B testing on websites, whereby the frequency of a given advert can be adapted on the fly proportional to its click through rate, allowing for continual development, rather than a classical test-then-implement method. Bayesian bandits provide the

opportunity for optimal sampling to minimize regret; whether that be minimizing the variance of perturbations, prioritizing sampling of highest affinity ligands or any definable reward function.

We illustrate the benefits of adaptive sampling using a range of reward functions appropriate for drug design. Relative alchemical calculations have been performed for both hydration and binding free energy energies and are shown to be consistent to both experimental values and alternative computational methods, using the publicly available Freesolv[4] and Schrodinger[5] datasets. Reduction in the variance of these simulations and the required computational expense relative to typical equal allocation efforts for a variety of small-molecule free energy maps.

2. References

[1] Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57* (12), 2911–2937.

[2] Crooks, G. E. Measuring Thermodynamic Length. *Phys. Rev. Lett.* **2007**, *99* (10), 100602.

[3] Shenfeld, D. K.; Xu, H.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Minimizing Thermodynamic Length to Select Intermediate States for Free-Energy Calculations and Replica-Exchange Simulations. *Phys. Rev. E* **2009**, *80* (4), 046705.

[4] Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput. Aided Mol. Des.* **2014**, *28* (7), 711–720.

[5] Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137* (7), 2695–2703.