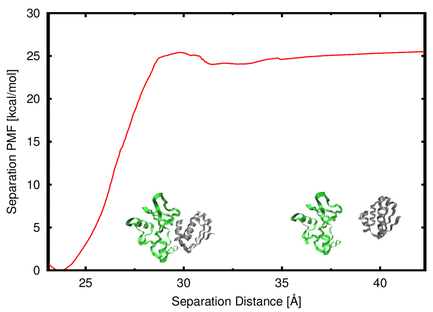
Enhanced Sampling Method with Hybrid non-equilibrium Molecular Dynamics / Monte Carlo

Donghyuk Suh[†], Brian Radak[*], Chris Chipot[¶], and Benoît Roux[‡]

[†] Department of Chemistry, and [‡] Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL
[*] Leadership Computing Facility, Argonne National Laboratory, Argonne, IL
[¶] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL

Motivation

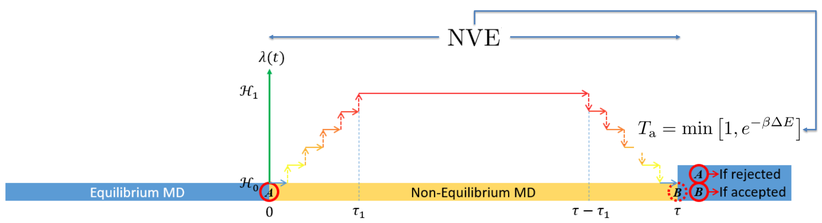
Protein-protein interactions govern many biological processes/functions and cell communications. High resolution structural data has become available, but many possible complexes they can form remain unknown. Some proteins bind together, while others do not despite of structural similarities. To tackle these problems, we may adopt computational approach to calculate binding free energy and find the governing factor.



Molecular Dynamics (MD) simulation has become essential tool for sampling complex multidimensional system. When further combined with biasing methods (i.e. Umbrella Sampling (US), Adaptive Biasing Force (ABF), etc), MD allows us to acquire important equilibrium properties for considerably large systems. A computational approach based on all atomistic MD simulation and free energy methodology is advantageous, but upon rugged free energy surface, its convergence could be painfully slow. Converging faster/overcoming free energy barrier/reducing the cost, are our goal to achieve here.

Hybrid Propagator Scheme

$$V^*(r) = V^*_{bonded}(r, E, \alpha) + V^*_{nonbonded}(r, \gamma)$$

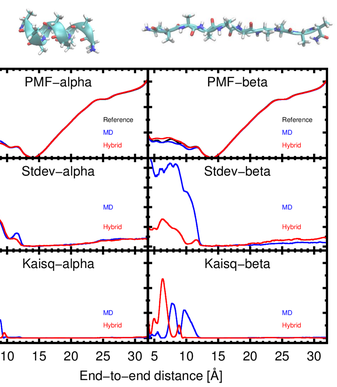


$$H(t) = H_0 + \Delta H[x, \lambda(t)]$$

$$\lambda(t;\tau,\tau_1) = \begin{cases} t/\tau_1, & (0 \le t < \tau_1) \\ 1, & (\tau_1 \le t < \tau - \tau_1) \\ (\tau - t)/\tau_1 & (\tau - \tau_1 \le t < \tau) \end{cases}$$

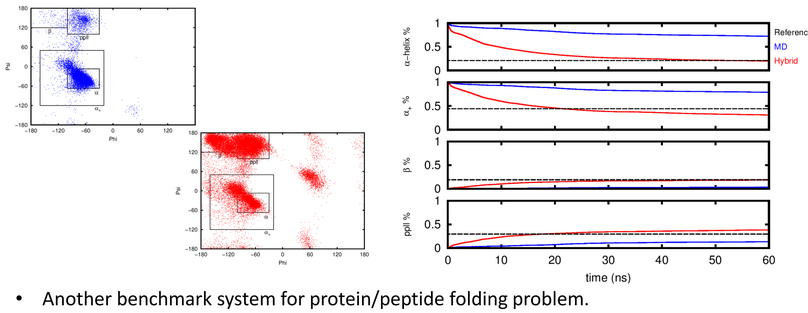
Result

Deca-alanine in vacuum



- Validated using double-well like system of Ala_{10} with $\epsilon = 20$ (Left).
- Brute-force Hybrid converges to the same PMF as MD, while randomly accepting with same mean acceptance ratio does not.
- Tested upon Ala10 in vacuum ($\epsilon = 1$) with ABF and Hybrid-ABF (Right).
- Well-studied benchmark system with many meta-stable states at < 8Å.
- Two initial structures were used: alpha helical and beta sheet.
- 8 independent trajectories, each 40 ns long.
- Reference was generated by multi-walker ABF, 2-μs long.
- Hybrid (eqMD = 7, $\tau = 3$, $\tau_1 = 1$ [ps/cycle]) has 30% less data points yet converges faster.

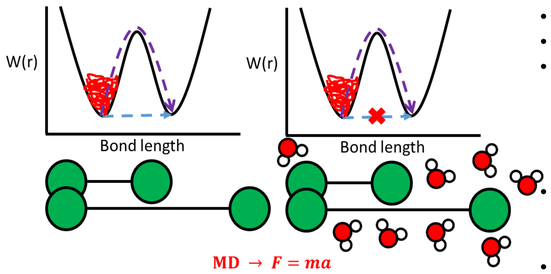
$AC - (AAQAA)_3 - NH_2$ in explicit solvent



- Another benchmark system for protein/peptide folding problem.
- Cumulative time-series for helicity converges faster towards the reference value for Hybrid (eqMD = 2, $\tau = 24$, $\tau_1 = 8$ [ps/cycle]) while MD got stuck in local minima, alpha-helical conformation.
- Reference was from Best et al paper, used T-REMD, 4.8-μs long total.

Hybrid Propagator with aMD & REST2

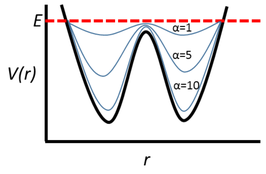
Hybrid neMD/MC



$$\mathbf{MD} \rightarrow F = ma$$
$$MC \rightarrow T_a = \min[1, e^{-\beta(E_f - E_i)}]$$
$$Hybrid\ neMD/MC \rightarrow x_i \xrightarrow{a_{i+1}} x^*_{i+1} \xrightarrow{K_{i+1}} x_{i+1} \xrightarrow{a_{i+2}} \dots \xrightarrow{a_f} x^*_f \xrightarrow{K_f} x_f\,; T_a$$

- Bi-stable dimer example.
- MD hardly hops high FE barrier.
- Metropolis Monte Carlo method (MC) introduces direct change in configuration space and accept/reject proposed configuration based on criterion.
- However, nearly all proposed configurations are rejected with explicit solvent present.
- Hybrid neMD/MC originated from constant-pH method avoids this issue.

Accelerated Molecular Dynamics (aMD)



$$\Delta V(r) = \begin{cases} 0 \\ \dfrac{\left(E - V(r)\right)^2}{\alpha + E - V(r)} \end{cases}$$

- aMD boosts potential energy below certain threshold E to reduce barrier.
- Two variables are used to control the boost : E and *a*.
- Reducing barrier induced from bonded interactions.

Replica Exchange Solute Tempering (REST2)

$$V_m^{REST2}(r) = \gamma^2 V_{ss}(r) + \gamma V_{sv}(r) + V_{vv}(r)$$

- REST2 scales potential energy between solute-solute and solute-solvent to reduce barrier.
- Single scaling variable $\gamma = \sqrt{\beta_0/\beta_m}$ is used to control the boost.
- Reducing barrier induced from non-bonded interactions.

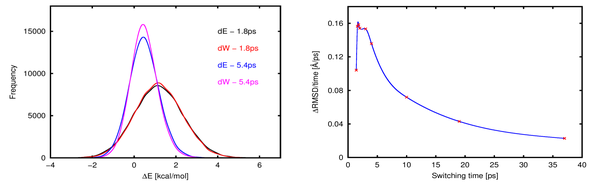
Switching Schedule Optimization

Perturbation schedule

- Parameter changes per ramp were set to equalize the perturbation among ramps as much as possible to maximize mean acceptance ratio.
- Thus, α was changed along 1/x and γ was changed linearly.

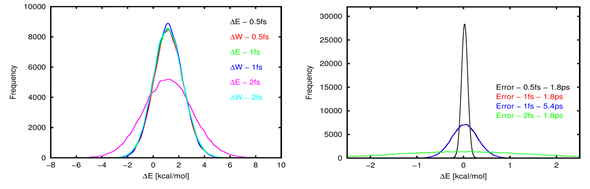
Switching time

- Switching time and the number of switches were chosen to maximize the efficiency, defined as ΔRMSD/t.
- The maximum number of switches for the schedule was limited to total switching time/10 to allow at least 5 time steps per each propagation kernel.



Shadow work

- Integrator error causes additional random Gaussian energy fluctuation that increases with time step.
- Using ΔW instead of ΔE for metropolis criterion allows us to use other than NVE ensemble for neMD trajectory and increase accuracy.



References

1. Nilmeier, et al. *Proc. Natl. Acad. Sci.* **2011**, *108*, E1009-E1017.
2. Hamelberg, et al. *J. Chem. Phys.* **2004**, *120(24)* 11919-11929.
3. Liu, et al. *Proc. Natl. Acad. Sci.* **2005**, *102*, 13749-13754.
4. Gumbart, et al. *J Comput Theory Chem.* **2013**, *9*, 3789-3798.
5. Wojtas-Niziurski, et al. *J Comput Theory Chem.* **2013**, *9(4)*, 1885-1895.
6. Torrie, G.M. and Valleau, J.P. *J. Comp. Phys.* **1977**, *23*, 187-199.
7. Chen, Y. and Roux, B. *J. Chem. Phys.* **2014**, *141*, 114107.
8. Chen, Y. and Roux, B. *J. Chem. Phys.* **2015**, *142(2)*, 024101.
9. Jo, S. and Jiang, W. *Comp. Phys. Comm.* **2015**, *197*, 304-311.
10. Darve, et al. *J. Chem. Phys.* **2008**, *128*, 144120.
11. Hazel, et al. *J Comput Theory Chem.* **2014**, *10(7)*, 2836-2844.
12. Best, et al. *J Comput Theory Chem.* **2012**, *8*, 3257-3273.
13. Radak, R. and Roux, B. *J. Chem. Phys.* **2016,** *145*, 134109.

NSF XSEDE Extreme Science and Engineering Discovery Environment