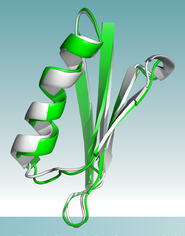
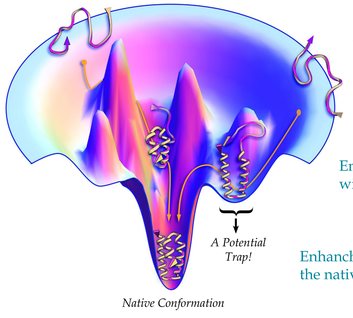
Data Guided Protein Folding Employing Hamiltonian- and Temperature-Annealed Importance Sampling

Mir Ishruna Muniyat and Justin MacCallum | Center for Molecular Simulations, Department of Chemistry, University of Clagary, Calgary, Canada
E-mail: mirishruna.muniyat@ucalgary.ca, justin.maccallum@ucalgary.ca



Conformational sampling of biomolecules remains a challenge



Sampling the native state in a complex landscape: Looking for "a needle in a haystack"

Energy landscapes of biomolecules are rugged with large energetic and entropic barriers.

Enhanced sampling methods guide states towards the native state, but often gets trapped!

Replica-Exchange Molecular Dynamics

√ Boltzmann Distribution
x Poor kinetic behavior
x Parrallel operation requires multiple GPUs

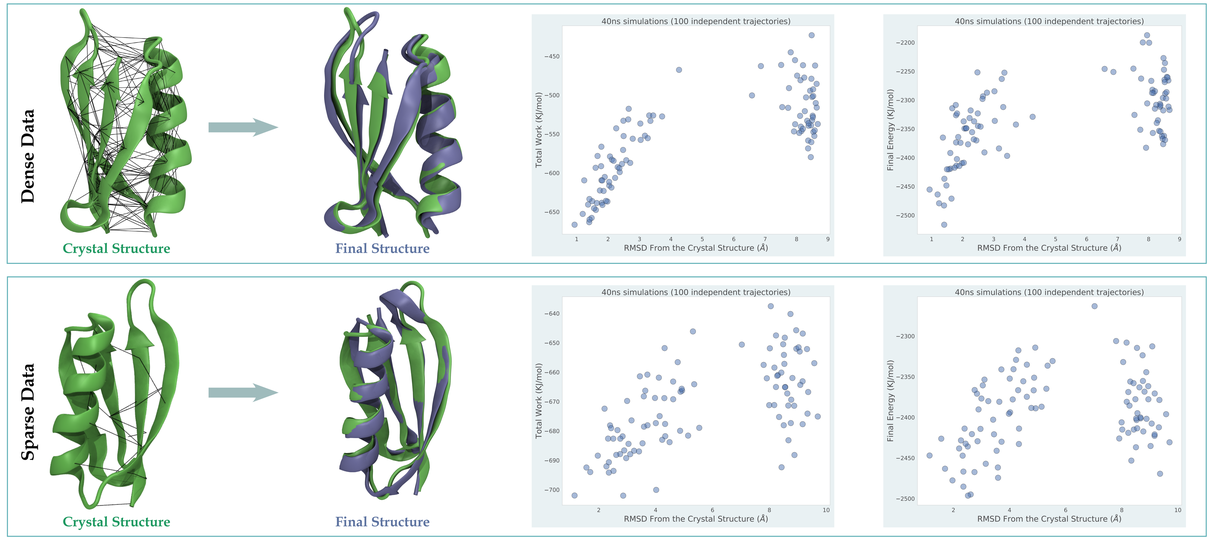
Simulated Annealing

√ Computationally efficient
x Boltzmann Distribution
x Non-equilibrium algorithm

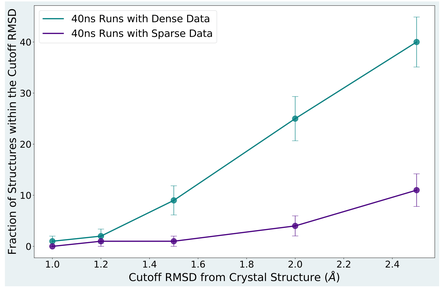
Annealed Importance Sampling (AIS)

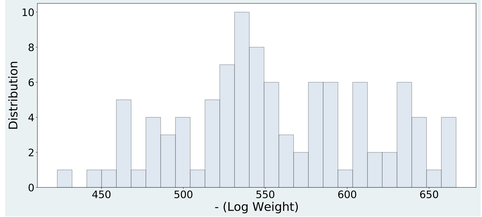
√ Boltzmann Distribution
√ Computationally efficient

AIS can predict accurate models of proteins even when the data is sparse

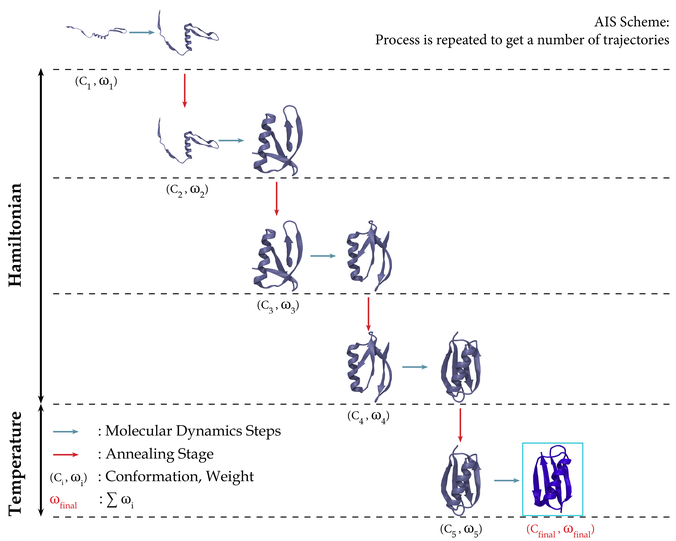


Distribution generated with AIS shows high variance in weight

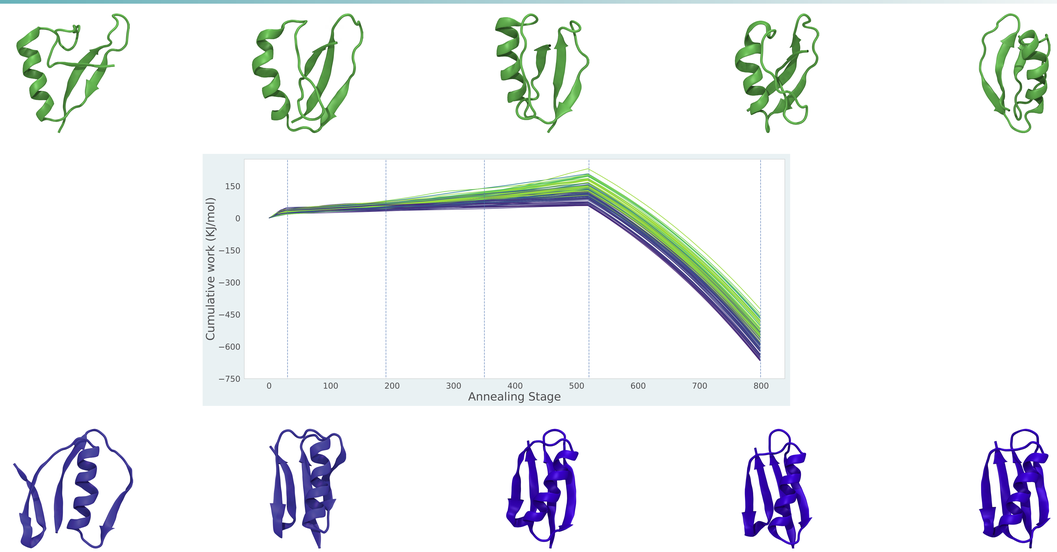




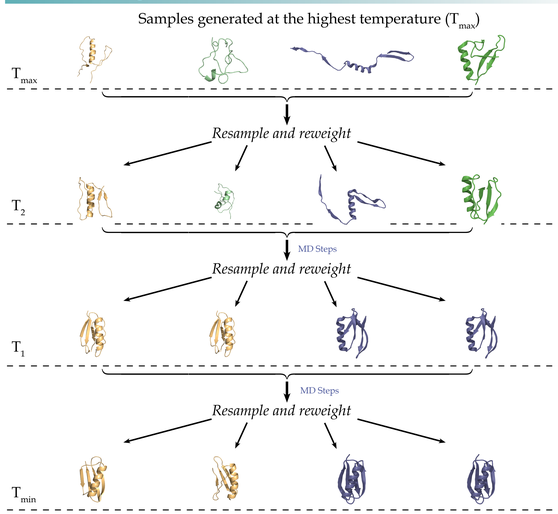
AIS generates a Boltzmann distribution at a lower computational cost



Structures with low RMSD require lower cumulative work



AIS combined with resampling and reweighting can minimize the variance in weight



We maintain a Boltzmann distribution at each annealing stage.

References:
Dill, Ken A., and Justin L. MacCallum. "The protein-folding problem, 50 years on." science 338.6110 (2012): 1042-1046; **Neal, Radford M.** "Annealed importance sampling." Statistics and computing 11.2 (2001): 125-139; **Lyman, Edward, and Daniel M. Zuckerman**. "Annealed importance sampling of peptides." The Journal of chemical physics 127.6 (2007): 08B605