4/7/2020 Week 12 Module 2
Analysis of Binding Free Energy Calculations 2

- This module will consist of
  - a mini-lecture describing the rationale behind binding pose prediction
  - a description of how to predict binding poses with YANK
- At the end of this module, you should also be able to predict a binding pose with YANK

Why use YANK for binding poses?

- Molecular docking scores are based on potential energy and don't include entropy
- The free energy of a pose depends on both
  - the enthalpy (proxy is potential energy) and
  - the entropy
- Molecular dynamics simulations of the bound pose are often trapped in a single minimum
- YANK is
  - likely to sample multiple bound conformations, due to Hamiltonian replica exchange
  - accounts for entropy

How to predict a binding pose

- In a molecular simulation, every structure is different
- Every structure has the same statistical weight
- To predict a pose, we need to group together similar structures: clustering
- For pose prediction, I implemented a method based on
  - aligning every protein structure according to alpha carbons
  - calculating a symmetry-corrected RMSD matrix of the ligand atoms
  - hierarchical clustering according to the symmetry-corrected RMSD
  - selecting a representative based on the medoid - the point closest to all other points - of each cluster
  - ranking poses based on the population of each cluster

Example pose prediction

- To extract the bound state from YANK results and align the protein to the first frame: https://github.com/daveminh/Chem456/blob/master/static_files/tutorials/bromodomains-YANK/getBoundState.ipynb
- To predict the ligand binding pose: https://github.com/daveminh/Chem456/blob/master/static_files/tutorials/bromodomains-YANK/ligandPosePrediction.ipynb
- Results are in poses.dcd