Binding Free Energies of Host-Guest Complexes Using 3D RISM

February 16, 2018

1. Abstract
   (a) Abstract goes here! I have a draft for this, but I moved it so it will be the base of the introduction.
2. Introduction
   (a) Binding information is very important to drug design because drugs must bind to structures in the cell to be effective. Simulation of protein bindings yields valuable information to drug designers, since it helps map the areas where drugs can stick well. However, explicitly simulating every atom is massively expensive because of the size of typical proteins and the number of water molecules in the system. To overcome this computational cost, we used the 3D reference interaction site model (RISM), rather than explicitly simulating each water molecule. We simulated the small host molecule cucurbit[7]uril CB7 with 16 guest molecules in order to calculate their binding energies.
3. Theory
   (a) Binding Energies
      i. The statistical mechanics and thermodynamics of the binding process.
   (b) 3D-RISM
      i. General Information
         A. The implicit solvent model used which models water as a continuum.
      ii. The Universal Correction
         A. This is used to correct for the partial molar volume of the molecules which 3D-RISM can't handle.
      iii. Enclosures
         A. These are part of the RISM parameters, and I still need to do some reading.

(c) Effective Potential

i. 3D-RISM gives us some potentials that we use to find the effective potential needed in binding claculations.

$$\overline{E} = \langle E_{MM} + \mu_{excess} \rangle \tag{1}$$

(d) Entropy Contributions

i. By using an end-state analysis script called MMPBSA.py, we calculate the entropy through nomral-mode analysis of an ensemble of snapshots throughout the simulation. This calculation gives us the translational, rotational, and vibrational entropic contributions. We are now working on calculating the entropies for our data sets. This can be quite computationally expensive.

ii. However it is vastly important to also consider the conformational entropy involved in the dynamics process. The conformational entropy is directly proportional to the average of the second power of the fluctuations of the effective potential. CITE CHONG AND HAM HERE.

$$T\Delta S_{Conf} = \frac{1}{k_B T}\overline{\delta E^2} \tag{2}$$

(e) Exponential Averaging

i. Essential part of the analysis. Uses partition functions to find the average energy. We are not presently at this stage of the analysis. CITE EXPONENTIAL AVERAGING PAPER HERE.

$$\Delta G = \beta^{-1} \ln \langle e^{-\beta \Delta U(\vec{q})} \rangle_0 \tag{3}$$

4. Methods

(a) Molecule Parameterization

i. Used TLEAP to parameterize molecules for simulation, define parameters used and the minimization/equilibration process.

(b) Simulation and Data Acquisition

i. Through the use of Dr. Luchko's cluster, we simulated the molecule complexes through a batch system. Simulations were performed each on a single CPU and completed within 30 hours of submission. Once the molecules are properly parameterized, the simulation is performed in three steps.

ii. Minimization

A. Here the energies of the system are allowed to move from their original parameterized position to find a minumum configuration.

iii. Equilibration

A. This is no different from a production run in the sense that the system is allowed to move and dynamics are recorded. The purpose of equilibration is to allow the system a short window to run before full molecular dynamics is performed.

iv. Production

A. Production runs are where data are collected from.

v. We parsed through the output files to find the relevent data, and placed them in a data frame which can be maipulated easily. Most of the data selection was done through the use of the 'pandas' module.

5. Results

(a) Here will be the figures of the data. Still working on the analysis, so I do not have a placeholder yet.

6. Discussion

(a) Discussion of important figures/tables. Will look at the comparison between our results and experiment.

7. Conclusions

(a) Here we make claims about the effectivness of the 3D-RISM when calculating the binding free energies of host-guest pairs.