# 20.320, notes for 11/7

Thursday, November 08, 2012
9:35 AM

## Last time

We were talking about how to mutate a protein to change its specificity. How do you predict which mutations would be useful? ΔΔG is the difference between the interaction energies of two different interactions. In the computer we can compute alchemical transformations, which are conversions between regular and mutated proteins. This is not something we can do experimentally, but we can do in the computer. We are still left with the problem, though, of how to know what the mutant form is, so that we can calculate its potential energy. This is our method:

1. Draw the free energy (thermodynamic) cycle. We are looking for ΔΔG, calculated from the vertical (alchemical)reactions.
2. Predict the mutant structure, starting from wild type and making mostly local changes.
    a. This is not trivial, for every single amino acid we change moves around all the AAs in the vicinity and around the protein. Every change has far repercussions.
3. Use the Metropolis Algorithm to find the structure with the minimum interaction energy.

## The Metropolis Algorithm

1. Start wit state $S_N$.
2. Randomly choose neighbor state $S_{test}$.
    a. By 'neighbor', we mean a state that is only a small change away from $S_N$.
3. Compute the acceptance ratio, which is a ratio of probabilities. Probability here is a proxy for energy. Greater probability means lower energy, or a more favorable state.

    $$a = \frac{P(S_{test})}{P(S_N)}$$
4. IF $a > 1$ , $S_{N+1} = S_{test}$
5. ELSE: Accept $S_{test}$ with probability $a$, as defined before.
    a. We always accept the new state if it's more favorable than the previous, but we don't always discard it otherwise. Instead, we give it a small chance to be accepted which is proportional to its favorability. If it has a high probability, then we accept it.

The precise meaning of probability is something we need to define. To give it physical meaning, we define it using the Boltzmann distribution. This gives us a probability (for a particular mutation) proportional to the energy difference, but also adds in temperature effects. At higher temperatures, the probability of improbable structures goes up. We accept less favorable stuff.

$$\frac{P(S_{test})}{P(S_N)} = e^{-\frac{[E_T - E_W]}{kT}}$$

The advantage of this is that we can perform a simulated annealing, where we start at a high temperature and gradually go cooler. We effectively increase the stringency of our selection. Starting at high temperature allows us to move away from our starting position, maybe crossing regions of relatively high energy to get to distant, deeper minima.

For all these calculations, we to know how to calculate the potential energy of any given conformation. For that we can measure force fields, we can perform knowledge-based energy allotments, or we can determine the energy empirically.

For knowledge-based energy calculations, we use the number of atoms at a given distance from ours. As we start our calculation, we need to collect good protein structures, categorize all atoms, and compute the number of pairs at a given distance (r = [1.5, 2.5, 3.5…]).

$$P_{ij}(r) = observed\ number\ of\ atoms\ of\ type\ i,j\ at\ distance\ r.$$

$$P_{ij}^{*}(r) = expected$$

And from that,

$$U_{ij}(r) = -k_B T ln\left[\frac{P_{ij}(r)}{P_{ij}^{*}(r)}\right]$$

The empirical analysis, on the other hand, tries to break down the problem into several terms, including both effects that have a well known and characterized physical basis (like VdW forces), and others that we know exist but for which we only have empirical relationships (like the hydrophobic effect). We add them all up.

$$\Delta G_{bind} = \Delta G_{VdW} + \Delta G_{Hbond} + \Delta G_{deform} + \Delta G_{hydrophobic} + \Delta G\_\varphi$$

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012