Free Energy, Enthalpy, and Entropy ... Oh My!

You've probably seen this equation:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$$

But what does this even mean? And where does it come from? This is an intuitive derivation of this equation. We're going to analyze the equilibrium between a protein that can be unfolded or folded:

$$U \rightleftarrows F$$

We will find that:

- ΔG° (free energy) measures the relative probability the protein is folded vs. unfolded.
- ΔH° (enthalpy) measures the differences in the bonds formed in the folded and unfolded states.
- ΔS° (entropy) measures the difference in the number of microstates between the unfolded and folded state.

Derivation

Because each protein molecule is buffeted by solvent molecules, it will flip between a whole collection of slightly different structures over time. These individual structures form an *ensemble* of *microstates*. Some microstates will be favored over others because they have more favorable bonds and/or fewer unfavorable clashes. The probability of a microstate m is given by:

$$P_m = \frac{w_m}{\sum_{i=1}^{N} w_i} = \frac{w_m}{Q}$$

where w is the weight on each microstate—measuring things like bond favorability—and Q is the sum of the weights of all microstates.

Some of these microstates will be folded, others will be unfolded. The probability of the folded state is given by:

$$P_F = \frac{\sum_{i=1}^{N_F} w_i}{Q}$$

where where the sum on the top is the sum of the weights of all microstates that are folded. If we assume the weights follow a normal distribution, we can replace that sum with the following:

$$P_F = \frac{N_F \langle w_F \rangle}{Q}$$

where N_F is the number of microstates in the folded state and $\langle w_F \rangle$ is the average weight of the microstates in the folded state. We could write a similar equation for P_U.

One important question is the relative probability of finding a protein molecule folded or unfolded. We can capture this by taking the ratio of P_F and P_U:

$$\frac{P_F}{P_U} = \frac{N_F \langle w_F \rangle / Q}{N_U \langle w_U \rangle / Q} = \frac{N_F \langle w_F \rangle}{N_U \langle w_U \rangle}$$

Now let's do some algebra. Take the log of both sides:

$$ln\left(\frac{P_F}{P_U}\right) = ln\left(\frac{N_F \langle w_F \rangle}{N_U \langle w_U \rangle}\right)$$

Use the log rule $log(a \cdot b) = log(a) + log(b)$:

$$ln\left(\frac{P_F}{P_U}\right) = ln\left(\frac{N_F}{N_U}\right) + ln\left(\frac{\langle w_F \rangle}{\langle w_U \rangle}\right)$$

Multiply by $-RT$:

$$-RTln\left(\frac{P_F}{P_U}\right) = -RTln\left(\frac{N_F}{N_U}\right) + -RTln\left(\frac{\langle w_F \rangle}{\langle w_U \rangle}\right)$$

Make some definitions:

$$\Delta G^\circ \equiv -RTln\left(\frac{P_F}{P_U}\right)$$

$$\Delta S^\circ \equiv Rln\left(\frac{N_F}{N_U}\right)$$

$$\Delta H^\circ \equiv -RTln\left(\frac{\langle w_F \rangle}{\langle w_U \rangle}\right)$$

and you obtain the familiar:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$$

But now we can see the origins of these terms. ΔG° measures the relative probabilities of the two states, ΔH° measures the difference in the average weights (bonds, etc.) between states, and ΔS° measures the difference in the number of states.