Extracting thermodynamic information from Differential Scanning Calorimetry measurements

We are interested in learning about the temperature dependence of protein folding thermodynamics. As we saw in class:

$$\Delta G(T) = \Delta H(T) - T\Delta S(T)$$

$$\Delta H(T) = \Delta C_p \left(T - T_{ref,H}\right) + \Delta H_{ref}$$

$$\Delta S(T) = \Delta C_p ln\left(\frac{T}{T_{ref,S}}\right) + \Delta S_{ref}$$

How do we get this stuff out of a DSC trace?

- ΔC_p is the difference in the folded and unfolded baselines.

- We use T_m as our $T_{ref,H}$ and $T_{ref,S}$. It is the midpoint of the unfolding curve.

- ΔH_{Tm} is the area of the peak. How do we get this? We need to do some quick math. From the definition of heat capacity, we know:

$$\delta q = C_p dT$$

We also know that:
$$dH = \delta q + dw.$$

If the system is not doing work, we get:

$$dH = \delta q$$

We can combine with our expression for δq above:

$$dH = C_p dT$$

Finally, we can integrate.

$$\int dH = \Delta H = \int C_p dT$$

We have no idea (at this point) what the functional form of $C_p(T)$ is, but it doesn't matter. In a DSC experiment, we experimetnally *measure* $C_p(T)$ and can thus integrate numerically.

- ΔS_{Tm} is $\Delta H_{Tm}/T_m$. We can get there by considering the following:

$$\Delta G(T_m) = 0$$

This means that:

$$\Delta G(T) = \Delta H(T_m) - T_m \Delta S(T_m) = 0$$

Rearrange, yielding:

$$\Delta S(T_m) = \Delta H(T_m)/T_m$$

Finally, we should note that when experimentalists analyze DSC curves, they usually do so by fitting a statistical model to the data rather than integrating directly. One such model is:

$$C_p(T) = B_0 + B_1 T + \left[\frac{K(T)\Delta C_p}{1 + K(T)} + \frac{K(T)\Delta H(T)^2}{(1 + K(T))^2 RT^2} \right]$$

In this model B_0 and B_1 are the intercept and slope of $C_{p,folded}$[1] $K(T) = exp(-\Delta G(T)/RT)$, using $\Delta G(T)$ as defined above. $\Delta H(T)$ is the function we defined above. If you really want to be a nerd, you can use nonlienar regression to fit this model to the data below and extract the relevant parameters. Otherwise, you can do the entire analysis using Excel.

The data included in this exercise are derived from thermodynamic parameters reported by:

- Hart KM, Harms MJ, Schmid BH, Elya C, Thornton JW, Marqusee S "Thermodynamic System Drift in Protein Evolution." (2014) *PLOS Biology.* https://doi.org/10.1371/journal.pbio.1001994

---

[1]Yes, C_p can change with temperature, making things slightly more complicated.

1. The file **ec.csv** contains experimental DSC data for the unfolding of the protein RNase H taken from *E. coli*. There are three series in the file.

   (a) The "temperature" series reports the temperature of the calorimeter cell in kelvin.

   (b) The "buffer" series reports the measured C_p as a function of temperature for buffer alone ($20\ mM\ NaOAc$, $50\ mM\ KCl$, and $1\ mM\ TCEP$, pH 5.5).

   (c) The "protein" series reports the measured C_p as a function of temperature for the 50 $\mu g/mL$ protein dissolved in the buffer.

   Figure out how to extract T_m, ΔC_p, ΔH_{Tm}, and ΔS_{Tm} from these data.

2. On separate graphs, plot $\Delta H(T)$, $-T\Delta S(T)$, and $\Delta G(T)$ vs. T using the values for T_m, ΔC_p, ΔH_{Tm}, and ΔS_{Tm} you determined above. Use a wider T range than was accessed experimentally: 200 K- 500 K.

   (a) What do you observe about the $\Delta G(T)$ curve?

   (b) Can you justify the shape of the $\Delta G(T)$ curve using the shapes of the $\Delta H(T)$ and $-T\Delta S(T)$ curves? It might help to consider two points on the curves: What is happening at low temperature? At high temperature?

   (c) Can you justify the shape of the $\Delta G(T)$ curve in molecular terms? As before: What is happening at low temperature? At high temperature?

3. Create a plot of the fraction of protein in the folded state versus T. What does this imply happens to the protein at low temperatures? Why might this be?

4. The file **tt.csv** contains experimental DSC data for the unfolding of the protein *RNase H* taken from *T. thermophilus*. This organism grows at 65 °C, while *E. coli* grows at 37 °C.

   (a) Extract T_m, ΔC_p, ΔH_{Tm}, and ΔS_{Tm} from these data.

   (b) How, mechanistically, does it appear that the *T. thermophilus* protein resists unfolding at high temperatures? (Hint, go through the exercise in #2, comparing your answers for *E. coli* and *T. thermophilus* RNase H).