5. Thermodynamics & Protein Conformation

Getting Thermodynamic Data[1]

The goal of this lecture is to describe methods for obtaining structure-stability relationships in protein structure. In previous lectures we've discussed the size and shape side of things, but have not addressed in any quantitative way the interplay of enthalpy and entropy in shaping proteins. Studies directed at such information typically require systems that operate in a two-state equilibrium. In general, the two states will be an unfolded (U) state in equilibrium with a native (N), folded state:

$$N \Leftrightarrow U \qquad K_{unfold} = [U]/[N] \qquad \text{(Eq. 5.1)}$$

As will be discussed later, it is not always easy to set up experiments in which the folded an unfolded states are in a two-state equilibrium, but for our initial purposes, we will discuss two techniques that provide thermodynamic data derived from such equilibrium studies.

van't Hoff Analysis

Equilibrium constants are temperature dependent. At a simple level (naively, and incorrectly, assuming that ΔH and ΔS are constant within the temperature range of interest), the relationship can be captured in the van't Hoff equation (Eq. 5.3), which is readily derived from the relationship between equilibrium constants and free energy (Eq. 5.2):

$$\ln(K_{eq}) = -(\Delta G^\circ/RT) \qquad \text{(Eq. 5.2)}$$

$$\ln(K_{eq}) = \frac{-\Delta H^\circ}{R} \cdot \left(\frac{1}{T}\right) + \frac{\Delta S^\circ}{R} \qquad \text{(Eq. 5.3)}$$

By obtaining equilibrium data for a two-state process at multiple temperatures, one may plot the natural logarithm of the equilibrium constants vs. inverse temperature (measured in Kelvins) to obtain ΔH° and ΔS° from the slope and intercept, respectively.

It is common in structural equilibrium studies to focus on a spectroscopic signal (such as circular dichroism or fluorescent emission intensity) that correlates to a structural change. Such a signal (S) will have two extreme values, S_U and S_N, for the unfolded and native forms. In the course of measuring the structural equilibrium at multiple temperatures, the signal will evolve as the stability of one state gradually degrades and the other begins to predominate (Figure 5.1). The equilibrium constant for folding (for example) may then be expressed as follows (Eq. 5.4):

---

[1] For an accessible, but more mathematically explicit treatment of this subject, see: www.biophysics.org/education/bloomfieldv.pdf (accessed 9/9/07).

$$K_{unfold} = \frac{[U]}{[N]} = \frac{\text{fraction unfolded}}{\text{fraction folded}} = \frac{S_N - S}{S - S_U} \qquad \text{(Eq. 5.4)}$$

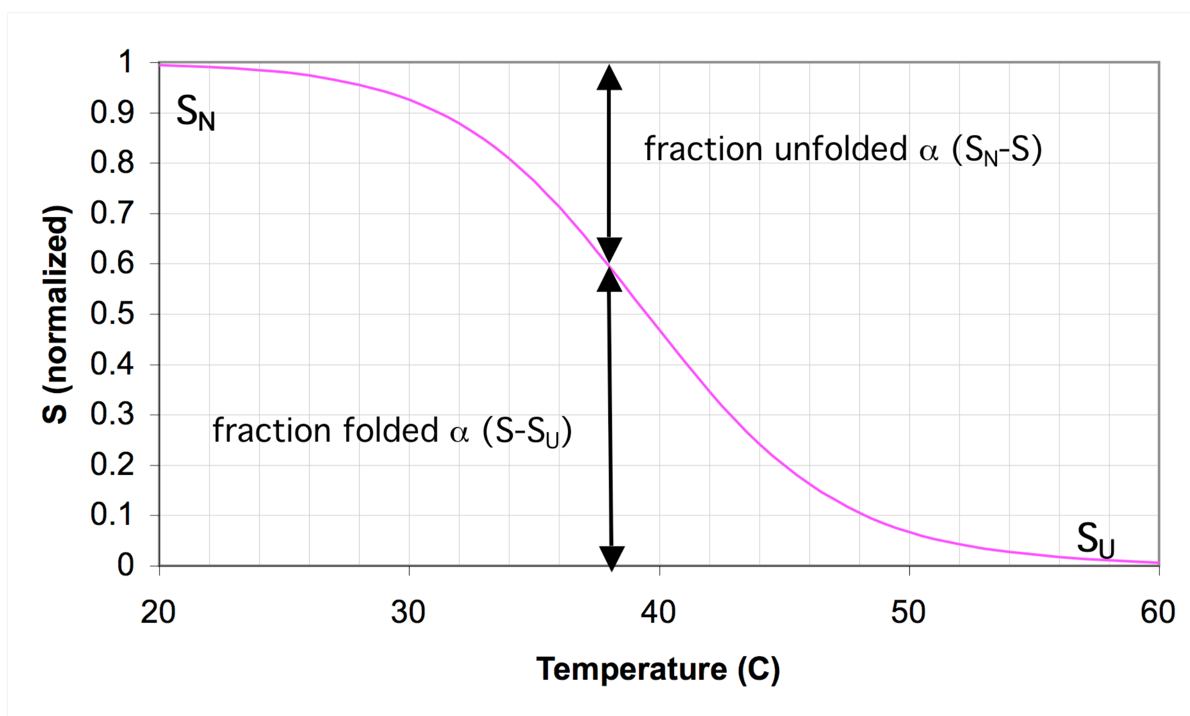


Figure 5.1. Temperature dependent equilibrium study. Note that the maximum signal is obtained for the folded (native) form and the unfolded form gives the minimal signal in this instance. Experimental data is never this clean, and often baselines must be obtained to cope with temperature dependence of the signal that is independent of the chemical system in question.

Working with denaturants

Sometimes, however, a two-state equilibrium cannot be obtained under conditions that might be deemed physiological. This is particularly a problem for protein unfolding studies. Often, unfolded proteins spontaneously and irreversibly aggregate (think of an egg white hitting a frying pan). Thus, it is sometimes necessary to study protein unfolding under **denaturing** conditions, where the unfolded protein is both stable and soluble.

Figure 5.2. Two common chaotropic agents. Urea is on the left and guanidinium chloride (a salt) is on the right.

The two principle denaturants used in these studies are urea and guanidinium chloride (Figure 5.2). These **denaturing** species are **chaotropes**, meaning that they diminish the

hydrophobic effect. For example, olive oil which separates spontaneously from water is actually soluble in concentrated solutions of urea. The chemical basis for chaotropic behavior is that these species disrupt hydrogen-bonding networks in water absent any non-polar solute. Thus, upon addition of a non-polar species, there is not enthalpic drive for water to form clathrates in order to maintain H-bonding networks around the solute (since it didn't have any to begin with), and there is not entropic cost to the water molecules. Likewise, chaotropic agents interact favorably with non-polar surface area, providing an enthalpic benefit for the dissolution of non-polar solutes in concentrated solutions of urea and guanidinium. Proteins **denature** under such circumstances and adopt an unfolded conformation.

A common method of employing chaotropes in unfolding studies is to measure the equilibrium constant for unfolding as a function of [denaturant] instead of temperature. At each concentration, and equilibrium constant is found, and used to determine a free energy of unfolding at that denaturant concentration, which may be plotted vs. denaturant concentration to exploit the following relationship (Eq. 5.5, assuming urea):[2]

$$\Delta G_{unfold}^{[urea]} = \Delta G_{unfold}^{H_2O} - m[urea] \qquad \text{Eq. 5.5}$$

As urea concentration increases, the free energy of unfolding becomes more spontaneous, and thus more negative. The value of the slope, -**m,** is generally a good indication of the amount on non-polar surface area exposed upon unfolding. The greater the dependence of unfolding on urea concentration, the more positive **m** value and the more non-polar surface area becoming exposed. The truly interesting value to arise from this type of study is the free energy of unfolding in water, which is the y-intercept obtained by plotting equation 5.5. To use this data to obtain ΔH° and ΔS° for unfolding, simply perform the study at multiple temperatures.

Differential Scanning Calorimetry

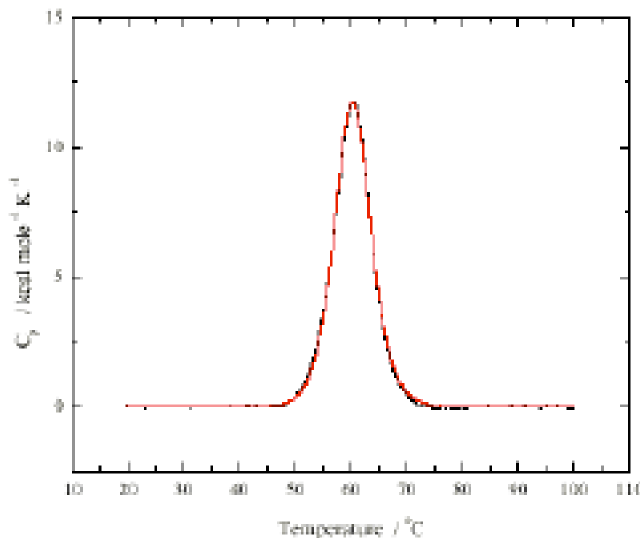
Differential scanning calorimetry (DFC) is an alternate technique for obtaining thermodynamic data involves measuring the amount of heat that must be applied to a sample to raise the temperature 1°C. This is the heat capacity of the sample. Most transitions from the folded to unfolded state are endothermic. As one raises the temperature from a low starting point, one approaches a temperature at which the equilibrium constant between the folded and denatured state is roughly unity (~40°C in Figure 5.1). This is the so-called **melting temperature**, **T_m.** The heat capacity of a protein solution becomes larger during the melting transition because most of the heat being applied to the solution is going into the unfolding of the protein, rather than into raising the temperature of the solution.[3]

---

[2] An inverse relationship exists for ΔG_{fold}, whereby the slope is +m.

[3] This is analogous to melting ice. One may apply a great deal of heat to a mixture of ice and water, but it will remain at 0°C until all the ice has melted. The applied heat simply goes into converting ice to water.

In a simple DFC experiment, a solution of a folded protein sample is gradually warmed, while monitoring the heat input to the system. As the protein unfolds, the endotherm is registered by the increased flow of heat, leading to a peak in the plot of heat capacity (C_P) vs. temperature. To obtain ΔH_{unfold}, one integrates the area under the curve, relative to the baseline, and the excess heat absorbed is the enthalpy of unfolding at the melting temperature (Eq. 5.6). Similarly, one can obtain the ΔS_{unfold} by integrating the excess heat applied, divided by temperature (Eq. 5.7). The melting temperature, in which ΔG_{unfold} is zero, is the ratio of ΔH to ΔS (Eq. 5.8).



$$\Delta H^\circ = \int_{T_1}^{T_2} \Delta C_P dT \qquad \text{(Eq. 5.6)}$$

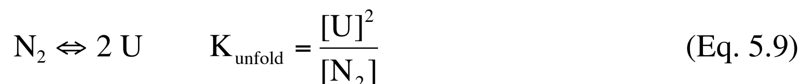
$$\Delta S^\circ = \int_{T_1}^{T_2} \frac{\Delta C_P}{T} dT \qquad \text{(Eq. 5.7)}$$

$$T_m = \frac{\Delta H^\circ}{\Delta S^\circ} \qquad \text{(Eq. 5.8)}$$

Figure 5.2. Idealized DSC scan of a protein unfolding event. The sharp increase in C_P (the heat capacity – or the amount of energy needed to raise the sample temperature by a given amount) is due to absorption of heat by the unfolding process. In reality, changes in heat capacity before and after unfolding and buffer effects would generally cause a more complicated plot.

Comparison of van't Hoff and Calorimetric Results

Often one will see ΔH_{vH} and ΔH_{cal} explicitly distinguished from one another. This is because the methods have the potential to yield differing results based on the system being studied. Where one has a simple one-to-one, two state equilibrium (Eq. 5.1), the two values will be identical. But often, there are additional interactions in one of the two states that complicate matters. For example, if the folded protein exists as a dimer, the two-state equilibrium will be:



$$N_2 \Leftrightarrow 2\,U \qquad K_{unfold} = \frac{[U]^2}{[N_2]} \qquad \text{(Eq. 5.9)}$$

However, one might not necessarily know that the folded state is dimeric and assume that the ΔH_{vH} obtained reflects unfolding of a single chain. Calorimetric measurements allow direct determination of the enthalpy absorbed on a per chain basis, since one has a direct measurement of the heat absorbed and one knows the total number of chains. If ΔH_{cal} is smaller than ΔH_{vH}, that

is evidence that some additional interactions are taking place in the folded state that give a larger enthalpy of unfolding than is detected calorimetrically.

On the Stability of α-Helices

Despite the prevalence of secondary structure observed in folded proteins, isolated secondary structure elements tend to be unstable. For example, peptides adopting an isolated α-helical conformation are unknown in nature (as are β-strands, but because of the need for sheet formation, we'll leave them alone). However, special circumstances allow the stabilization of helical conformation in isolated peptides. The following examples will provide some useful information related to the stabilization of secondary structure.

To a first approximation, helical stability is determined by the relative magnitude of the enthalpic stabilization achieved by the helix (favorable ϕ, ψ angles, good vdW interactions and internal H-bonding) versus the entropic costs of freezing out backbone and side chain entropy in the helical conformation. For the equilibrium between helix and coil, one expects positive enthalpic and entropic terms (Eq. 5.10).

$$\text{Helix} \Leftrightarrow \text{Coil} \qquad \Delta H_{unfold} > 0, \Delta S_{unfold} > 0 \qquad \text{(Eq. 5.10)}$$

In creating a stable helix, one must maximize the enthalpic term and minimize the entropic term, as seen in the following examples.

Helix Stabilization by Organic Solvent

Hydrogen-bonding between amide groups is mildly favorable in aqueous solution, but the enthalpic pay-off is mitigated by the loss of H-bonds to solvent that accompany the formation of H-bonds between amide groups. One way to maximize the enthalpic contribution of internal H-bonds to the stability of helices is to reduce H-bonding opportunities to solvent. The classic solvent used for this purpose is 2,2,2-trifluoroethanol (TFE). A polar solvent, TFE solubilizes peptides well, but the non-interactive nature of the C-F bond reduces the quality of intermolecular interactions between the peptide and the solvent.

Jasanoff and Fersht (1994)[4] performed a study of the effects of TFE on helix stability using short peptides of 10-14 residues. Using circular dichroism to measure helix content (θ_{222}), the study used similar methods to those described above, in which the equilibrium constant of unfolding is monitored at different denaturant concentrations. However, in this case, the equilibrium constant of unfolding *decreased* with the increase in the stabilizing mole fraction of TFE that was present (Figure 5.3). The free energy of helix unfolding is thus related to mol fraction of TFE by equation 5.11:

$$\Delta G^{TFE}_{unfold} = \Delta G^{H_2O}_{unfold} + m \cdot \chi_{TFE} \qquad \text{(Eq. 5.11)}$$

---

[4] Jasanoff and Fersht (1994) Quantitative Determination of Helical Propensities from Trifluoroethanol Titration Curves. *Biochem.* **33**, 2129-2135.

In this instance the positive slope reflects the growing "reluctance" of the helix to unfold with increasing TFE. However, because it was difficult to observe a fully unfolded state for the peptides, the equilibrium constants obtained are inherently limited. Nevertheless, it can be shown that simple 12 residue peptides can adopt significant helical structure in water, with perhaps as much as 25% of the peptide in the helical conformation at 25˚C.

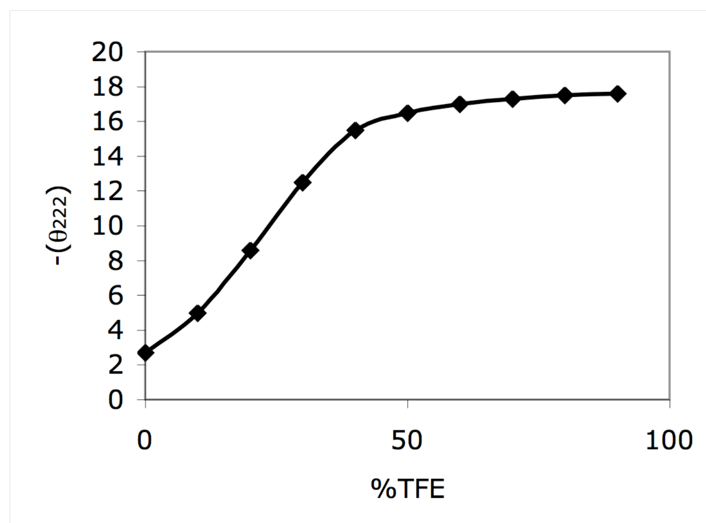


Figure 5.3. Titration of a 12-residue peptide (EPAEAAKAAAGR) with trifluoroethanol. The y-axis is directly proportional to helix content as the -θ values become more positive with greater helix content. Note that even at low concentrations of TFE it appears that the lower baseline ellipticity has yet to be attained.

Enthalpic Contributions via the Helix Macrodipole

In 1968, Wim Hol noted that the alignment of peptide bonds in the helix could have the effect of summing up the individual dipoles associated with each peptide bond to create an overall dipole on the α-helix that places an excess of positive charge at the N-terminus of the helix and an excess of negative charge on the C-terminus (Figure 5.4). Alternatively, one could imagine that the helix is destabilized by unpaired amide protons at the N-terminus and unpaired carbonyl oxygens at the C-terminus, which would have the same effect of creating a zone of partial positive charge at the N-terminus and negative charge at the C-terminus (Figure 5.4).

In 1987, Buzz Baldwin's lab investigated the contribution of amino acid side chains on helix stability through interaction with the helix "macrodipole". In principle, negatively charged residues at the N-terminus would be stabilizing while positively charged residues at the C-terminus would likewise have a stabilizing effect. Using a modified peptide based on a naturally occurring α-helical peptide segment from a protein, the Baldwin group prepared a helix with modest stability in solution and with titratable groups at either end of the peptide:

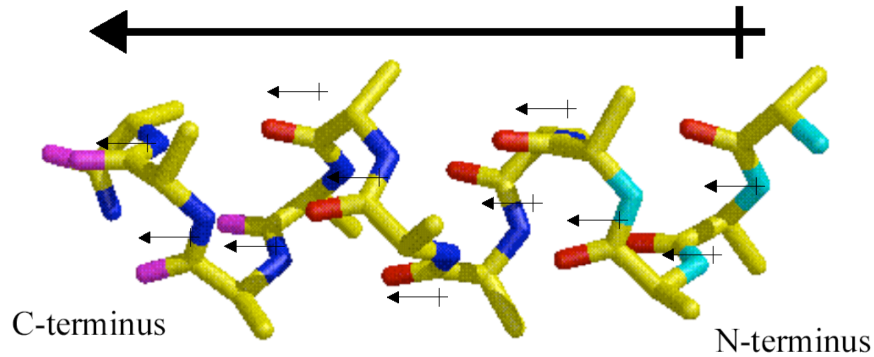


Figure 5.4. The macrodipole is proposed to form via the summation of the many smaller dipoles associated with each peptide bond. However, an alternative argument claims that it is the unpaired carbonyl oxygens (in magenta) at the C-terminus and the unpaired amide nitrogens at the N-terminus (in cyan) that give rise to the negative and positive character of the two ends, respectively.

OK, this is as far as I got. I'll get back to it on Wednesday and have something more complete soon. I hope.