University of Washington
Department of Chemistry
Chemistry 453
Winter Quarter 2015

Lecture 8. 1/23/15

A. Statistics of Conformational Equilibria in Proteins

- It is known from experiments that some proteins, DNA, and RNA molecules undergo transitions between structured helical forms and essentially disordered or random coil forms.
- In proteins, the α–helix exemplifies such a structured form, which is constituted mainly by a system of hydrogen bonds between the amide proton (i.e. N-H) of an amino acid to the oxygen of a carbonyl group (i.e. C=O) four residues away. See Figure 8.1

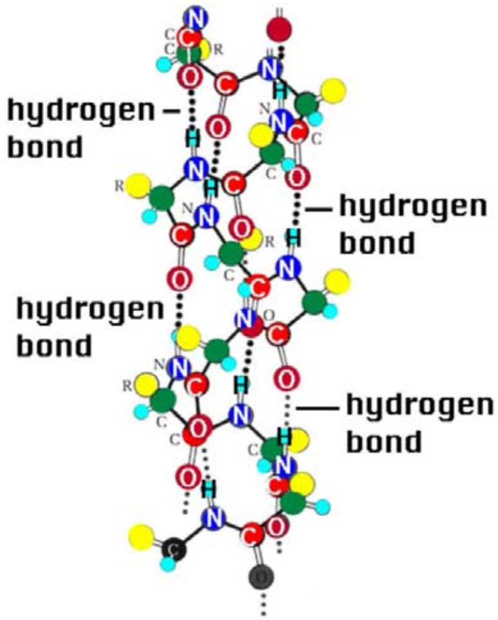
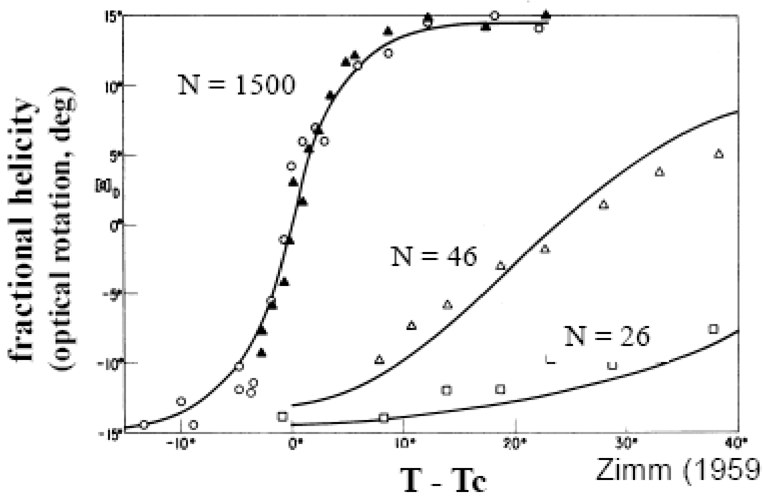


Figure 8.1: An alpha helix showing the system of hydrogen bonds between amide groups on the ith residue and the carbonyl groups on the i+4 residues.

- The helix form dominates at low temperatures. As temperature is increased, the random form dominates. Frequently the transition occurs over a very narrow range of temperatures, implying the entire helix unwinds very suddenly rather than sections of the helix unraveling gradually. See Figure 8.2.



Zimm (1959)

Figure 8.2: Fraction of helicity in peptide chains of different lengths N monitored as a function of temperature by optical rotation. T_c is the temperature in the midpoint of the transition when half the total chains are helical. Highly cooperative transitions have sigmoidal shapes
.

- As shown in Figure 8.2, when the fraction of chains that are helical is monitored as a function of temperature, the fraction changes rapidly at the melting temperature T_C and has a sigmoidal shape. Such a transition is said to be cooperative, because the chain sections convert from helix to random coil all together., as shown in Figure 8.3.

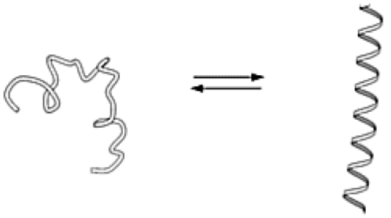


Figure 6.3: In a fully cooperative transition a helix suddenly unwinds into a random coil in an all or nothing fashion.There are no intermediate states composed of partially folding chains.

- The thermodynamics of helix-coil transitions can be modeled statistically. First let us assume the peptide chain is composed of N monomer units. Each monomer unit can either be in a helical state (H) or a coil state (C). So the structure of the chain can be coded in terms of H and C. A particular configuration of the chain might be HHCCCCHHHHCCHHH, for example. There are many possibilities.
- Our objective is to calculate the partition function of the chain assuming all possible configurations. To do this we have to assume a model for the transition. We will assume two models; non-cooperative, fully cooperative, and the zipper model.

B. Noncooperative Helix-Coil Transitions

- There are N monomers in the chain. The structural state of each monomer H or C is independent of it neighbors. We need to calculate the partition function for all configurations of such a chain
    - Suppose a chain has four monomeric units N=4. If such a chain is in the CCCC configuration the partition functions is q_0. If the configuration is HHHH, the partition function is q_4. There are four configurations with 3 C's and 1 H: CCCH, CCHC, CHCC, and HCCC. Call this partition function q_1. There are 6 configurations with 2 H's and 2C's with q_2. There are four configurations with 1 C and 3 H's for which the partition function is q_3. the total partition function is:

$$\begin{aligned} q &= q_0 + 4q_1 + 6q_2 + 4q_3 + q_4 \\ &= q_0\left(1 + 4\frac{q_1}{q_0} + 6\frac{q_2}{q_0} + 4\frac{q_3}{q_0} + \frac{q_4}{q_0}\right) \\ &= q_0\left(1 + 4k_1 + 6k_2 + 4k_3 + k_4\right) \end{aligned} \tag{8.1}$$

    - The parameters $k_n = \dfrac{q_n}{q_0}$ are called microscopic equilibrium constants. They represent equilibria between the CCCC configuration and each individual configuration containing n H-type monomers.
    - Now we make a big assumption about the energetics of these transitions. Regardless of which configurations are involved, every transition from C to H has the same energy change ΔG. Therefore the transition from CCCC to HCCC, CHCC, CCHC, or CCCH all have energy change ΔG and

$$k_1 = \frac{q_1}{q_0} = e^{-\Delta G^o/RT} = s \tag{8.2}$$

- Similarly the change from CCCC to any of the 6 configuations with 2H's has energy change 2ΔG so $k_2 = s^2$, and in general $k_n = s^n$ so

$$\begin{aligned} q &= q_0\left(1+4k_1+6k_2+4k_3+k_4\right) \\ &= q_0\left(1+4s+6s^2+4s^3+s^4\right) = q_0\left(1+s\right)^4 \end{aligned} \tag{8.3}$$

- Now each term in the expansion in 8.3 has a specific interpretation. When divided by q, they give the probabilities that particular configurations occur with particular numbers of H units:

$$\begin{aligned} &p_0 = \frac{q_0}{q} = \frac{1}{1+4s+6s^2+4s^3+s^4};\ p_1 = \frac{q_0 4s}{q} = \frac{4s}{1+4s+6s^2+4s^3+s^4} \\ &p_2 = \frac{q_0 6s^2}{q} = \frac{6s^2}{1+4s+6s^2+4s^3+s^4};\ p_3 = \frac{q_0 4s^3}{q} = \frac{4s^3}{1+4s+6s^2+4s^3+s^4} \\ &p_4 = \frac{q_0 s^4}{q} = \frac{s^4}{1+4s+6s^2+4s^3+s^4} \end{aligned} \tag{8.4}$$

- Equations 8.3 and 8.4 can be generalized to any chain with N monomers:

$$\begin{aligned} &q = q_0\left(1+Ns+\frac{N\left(N-1\right)}{2}s^2+\ldots\frac{N!}{n!\left(N-n\right)!}s^n+\ldots Ns^{N-1}+s^N\right) = q_0\left(1+s\right)^N \\ &p_n = \frac{q_0}{q}\frac{N!}{n!\left(N-n\right)!}s^n \end{aligned} \tag{8.5}$$

- To simulate data displayed in Figure 8.2, we need to calculate the fractional helicity, defined as:

$$f_H = \frac{\langle n\rangle}{N} = \frac{\sum\limits_{n=1,N} np_n}{N} \tag{8.6}$$

where <n> is the average number of helical units in a chain.

Example: Use equation 7.6 to calculate the fractional helicity for N=2.

$$\begin{aligned} f_H &= \frac{\langle n\rangle}{2} = \frac{\sum\limits_{n=1,2} np_n}{2} = \frac{1}{2}\left(p_1+2p_2\right) \\ &= \frac{1}{2}\left(\frac{2s}{1+2s+s^2}+\frac{2s^2}{1+2s+s^2}\right) = \frac{s\left(1+s\right)}{\left(1+s\right)^2} = \frac{s}{1+s} \end{aligned}$$

- This procedure is fairly simple for small values of N. But if N is large the series summation can become daunting. Fortunately, there is a really easy way to calculate <n> that does not use the p_n series expression. It can be shown:

$$\langle n\rangle = \frac{\partial \ln q}{\partial \ln s} = \frac{s}{q}\frac{\partial q}{\partial s} = \frac{sq_0}{q_0(1+s)^2}\frac{\partial}{\partial s}\left[(1+s)^2\right] = \frac{s}{(1+s)^2}2(1+s) = \frac{2s}{1+s}$$
$$\therefore\ f_H = \frac{\langle n\rangle}{2} = \frac{s}{1+s} \tag{8.7}$$

- In general, for N monomers:

$$f_H = \frac{\langle n\rangle}{N} = \frac{s}{1+s} \tag{8.8}$$

- Recall that the conversion from C to H occurs with a Gibbs energy change ΔG^o;

$$C \xrightarrow{\Delta G^o} H \tag{8.9}$$

- If s<1, then $\Delta G^o>0$, helix is not favored. This is the situation that prevails at high temperature.
- If s>1 then $\Delta G^o<0$, helix is favored which is the situation that prevails at low temperature. A plot of equation 8.7 for f_H versus is given in Figure 8.3:.

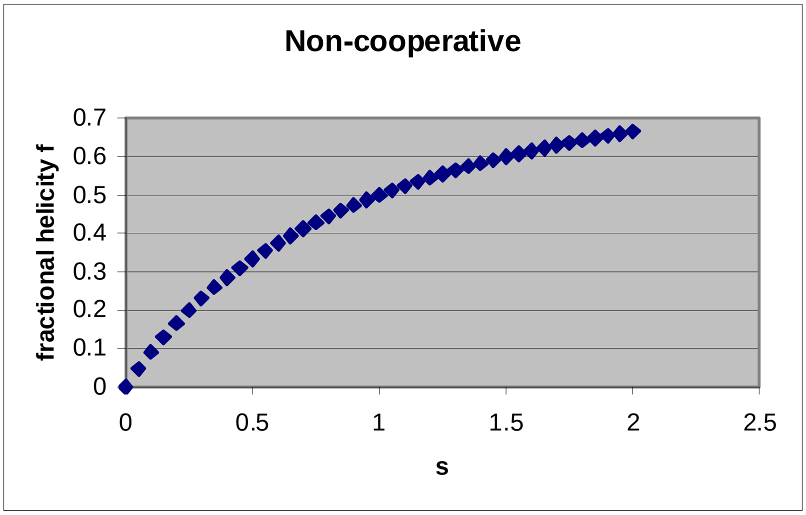


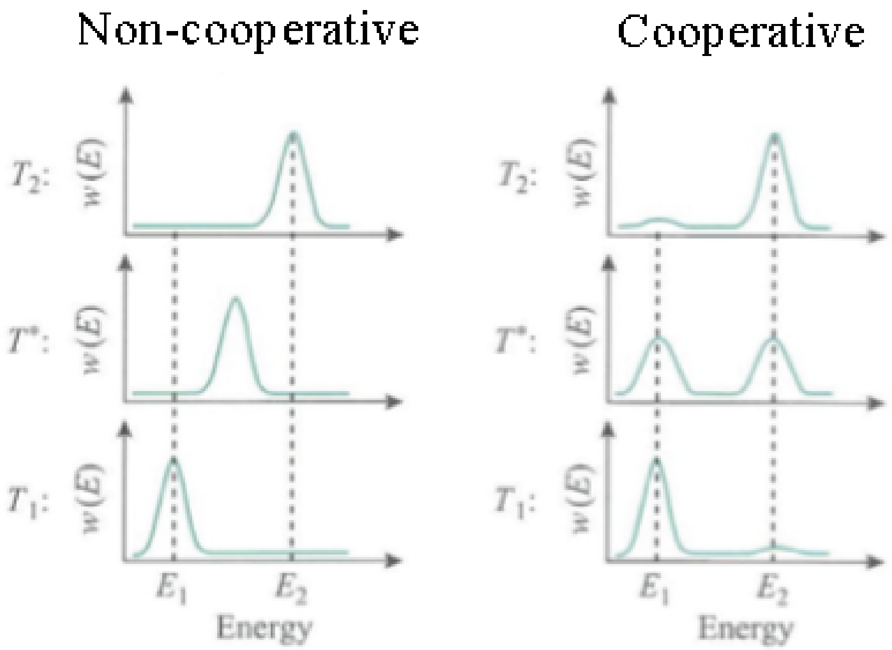
Figure 8.3: A plot of the fractional helicity as a function of s using equation 7.8.

- Equation 8.7 and Figure 8.3 indicate that in the absence of cooperativity, there is a smooth accumulation of helical monomers as s increases.

- In Figure 8.4 we can view the progress of non-cooperative and fully cooperative protein helix-coil transitions with temperature. Suppose the unstructured form of a protein with N monomers ...CCCC...has energy E_1 and the fully structured form ....HHHH.... has energy E_2. At temperature T_1 the number of unstructured molecules $w(E_1)$ is a large number and $\langle n\rangle \approx 0$. At temperature T_2 the number of fully structured molecules $w(E_2)$ is large and $\langle n\rangle \approx N$.

    - As the temperature is varied between T_1 and T_2, partially structured forms appear and the w maximum shifts gradually with temperature, and $\langle n \rangle$ increases gradually from 0 to N.

Figure 8.4: A non-cooperative model (left) has a gradual shift of the population maximum from T_1 where the unstructured form dominates to T_2 where the structured form dominates. At intermediate temperatures partially structured forms are most numerous. In the fully cooperative model (right) there are only two populations: fully structured and unstructured. At T_1 almost all the population is in the unstructured state and at T_2 almost all the population is in the structured state. At the melting temperature T*, the populations are equal.



C. Fully Cooperative Model

- In the fully cooperative model there are only two states:
    - The unstructured state: where the chain of length N has no helical monomers so has the configuration is CCCCCCCCC.... and the partition function is q_0.
    - The structured state: where the chain has all helical monomers which corresponds to HHHHHHHHH... and the partition function is $q_N=q_0s^4$.
    - Because there are only two forms in equilibrium: all C and all H, and assuming there are N monomers the equilibrium is

$$CCC\ldots CCC \xrightleftharpoons{k_N=s^N} HHH\ldots HHH \qquad (8.10)$$

    - To construct the partition function, we simplify equation 8.3. In this equilibrium only the constant $k_4=s^4$ is nonzero. Therefore the partition function simplifies to

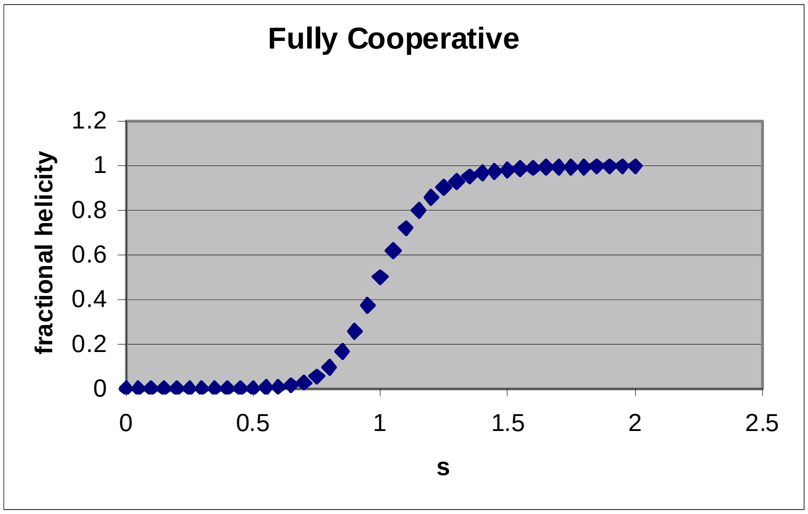
$$q = q_0\left(1 + s^N\right) \tag{8.11}$$

- Proceeding just as before:

$$\langle n \rangle = \frac{s}{q}\frac{\partial q}{\partial s} = \frac{s}{1+s^N}\frac{\partial}{\partial s}\left[1+s^N\right] = \frac{s}{1+s^N}Ns^{N-1} = \frac{Ns^N}{1+s^N} \Rightarrow f_H = \frac{\langle n \rangle}{N} = \frac{s^N}{1+s^N}$$

(8.12)

- Figure 8.5 A plot of equation 8.12 is shown to the right. It has the familiar sigmoidal dependence of the data in Figure 8.2.



- Figure 8.4, right and Figure 8.5 display the properties of a fully cooperative transition. In this model no intermediate forms exist and the fully structured and unstructured population change as a function of temperature. At the melting temperature T* the populations are equal so that f_H=0.5 and s=1.