2. Helix-Coil Transitions

The regular $\alpha$ helix is among the most common secondary structures in folded proteins. This structure has $\sim 3.6$ residues per helix turn and is constrained by a series of hydrogen bonds between the carbonyl oxygen of residue $i$ and the amide proton of residue $i+4$. In aqueous solution, the hydrogen bonding pattern for the helix is competing with more random hydrogen bonding to water in an unstructured “coil” conformation.

The role of regular secondary structure in the overall formation of folded protein structure is still a matter of scientific inquiry. It is clear that there is some entropy loss in folding so there must be some compensating energy gain in the folded protein. One question is whether intraprotein hydrogen bonds are energetically more favorable than hydrogen bonds to water. If the two energies are about equal, then hydrogen bonding would not contribute to folding.

Despite these more theoretical questions, it has been known since the 1950’s that homopolymers of some amino acids would form helical structure under some conditions. Poly-[\(\gamma\)-benzyl-L-glutamate] is the classic example.

More recently, it was shown that helix formation can be accompanied by oligomer formation for certain sequences. These sequences are generally characterized by 7-residue “heptad repeat” sequences (abcdefg) for which there are hydrophobic residues at the first (“a”) and fourth (“d”) positions. To a lesser degree, there are residues of opposite charge at the fifth (“e”) and seventh (“g”) positions. Seven residues approximately corresponds to two turns of an $\alpha$ helix. For these sequences, the helix/oligomer formation is favorable because it places the sidechains of the hydrophobic residues at the interior of the oligomer. It also puts in close proximity the oppositely charged sidechains of the e residue of one molecule and the g residue of another molecule in the oligomer.

Formation of helical oligomer structure from these sequences can be of large biological consequence. Some of these structures bind DNA and are important in regulation of transcription (synthesis of RNA from DNA). In addition, this type of structure is thought to be important in mediating infection of some viruses.

Helix/coil transitions are also observed for sequences which have general “abcd” repeats with hydrophobic residues at the a positions and hydrophilic residues at the c positions. For these sequences, “coil” structure (i.e. a large ensemble of structures) is observed in aqueous solution while helical structure is observed for the peptide bound to a membrane surface. Helical structure is favored in membranes because it places the hydrophobic sidechains of a residues in the hydrophobic interior of the membrane and places the hydrophilic sidechains of residues c in aqueous environment. This type of helix is known as “amphipathic” with a hydrophobic face and a hydrophilic face.

Several peptides with these “abcd” repeats have antimicrobial activity, in particular if the hydrophilic residues are positively charged (e.g. lysine). Microbial membranes often have a significant fraction of lipid molecules with negatively charged headgroups (e.g. glycerol) and there is significant attraction between the positive and negative charges. These peptides can
disrupt the microbial membrane to a significant enough degree to induce membrane pores and microbe death. They are being considered as topical antibiotics because they have much less affinity for human cell membranes which have a much smaller fraction of negatively charged lipids.

In addition, there has been work to design peptide sequences which form helical structure in aqueous solution. One example from R. E. Baldwin’s group is: AEAAAKEAAAKEAAAKA. Alanine (A) is known to have some tendency to form helical structure and there is attraction between residues with oppositely charged sidechains at positions i and i + 4 (e.g. E-7 and K-11). In a α helical structure, these residues would be approximately above one another in the helix.

For examples of the peptides in aqueous solution, there is a narrow range of parameters over which the structure changes from helix to coil. For example, there will be a 10-20° temperature range over which the peptide goes from predominant helix to predominant coil structure. Similarly, there will be a fairly narrow range of “denaturant” concentration (e.g. urea or guanidine) over which the structural change occurs. This “cooperativity” of helix-coil transitions is similar to the cooperativity observed for unfolding of whole proteins. The statistical mechanics of helix-coil transitions is more straightforward than that of larger protein unfolding and understanding the helix-coil transition can provide insight into protein folding and unfolding.

The simplest model considers that each residue can either be in a helix (“h”) or coil “c” conformation and the conformations of residues are independent of one another. The free energy of the coil conformation is 0 and the free energy of the helix conformation is $\varepsilon$. The parameter $s$ is the equilibrium constant for a single residue and using Eq. 1.118:

$$s = \frac{[h]}{[c]} = e^{-\varepsilon/kT}$$  \hspace{1cm} (2.1)

Note that $\varepsilon$ is a free energy difference and equals $\Delta H – T\Delta S$ for a single residue; i.e. it is temperature-dependent. Because $S_c > S_h$, we might expect that $\varepsilon$ will be negative at low $T$ and will be positive at high $T$. This will lead to changes in $s$ from values > 1 at low $T$ to values close to 0 at high $T$.

Suppose that for a single residue, we define that the c contribution to the partition function is 1. Using Eq. 2.1, the h contribution will be $s$ and the partition function:

$$q = 1 + s$$  \hspace{1cm} (2.2)

For a peptide of length $N$ residues, the total partition function is the product of individual residue partition functions:

$$Q = (1 + s)^N = \sum_{n=0}^{N} \frac{N!}{[(N – n)!n!]} \times s^n = \sum_{n=0}^{N} a_n \times s^n$$  \hspace{1cm} (2.3)

These equations are very similar to Eq. 1.9. Each n term is proportional to the probability $p_n$ that n of the N residues are helical.
\[ p_n = (a_n \times s^n)/Q \]  
\hspace{2cm} (2.4) 

The typical experimentally measureable quantity is average helix content which in our model:

\[ f_h = \langle n \rangle / N = \sum_{n=0}^{N} (n \times p_n) = (s/\mathcal{Q}N) \times (\partial \mathcal{Q} / \partial s) = (1/N) \times (\partial \ln \mathcal{Q} / \partial \ln s) \]  
\hspace{2cm} (2.5) 

The last two equalities are derived in DB p. 502. For this model:

\[ f_h = s/(1 + s) \]  
\hspace{2cm} (2.6) 

The model predicts a transition from helix to coil over a broader temperature range than is observed experimentally. In addition, it predicts that the transition is independent of N whereas experimentally, the transition becomes sharper for larger N (i.e. longer peptides).

An alternative model would be two-state with either all helical or all coil residues. The general idea is that the loss in entropy in going from coil to helical conformation can only occur when there is a corresponding gain in energy of having complete helical structure. For this model:

\[ \mathcal{Q} = 1 + s^N \]  
\hspace{2cm} (2.7) 

Using Eq. 2.5:

\[ f_h = s^N/(1 + s^N) \]  
\hspace{2cm} (2.8) 

This two-state model predicts a much sharper transition from helical to coil conformation than the first “independent-residue” model. For example, for a peptide in the independent-residue model, a change in s from 1.01 to 0.99 leads to a change in \( f_h \) from 0.502 to 0.497. For a 50-residue peptide in the two-state model, the same change in s leads to a change in \( f_h \) from 0.728 to 0.377.

In general, the two-state model predicts a sharper transition than is observed experimentally. In addition, there is not a great justification for discounting states with intermediate helical content.

The “zipper model” is intermediate between the independent-residue and the two-state models. It is predicated on the idea that the relative probabilities of helical vs. coil conformation at a particular residue are dependent on the conformation of the preceding residue. The general rationale for this idea is that ~5 sequential helical residues are needed to form the first turn of the helix with accompanying stabilizing interactions such as a i, i+4 intrahelical hydrogen bond or electrostatic attraction between the sidechains of residues i and i+4. So, beginning or “nucleation” of a helix in coil structure will have positive free energy and low probability because of the large initial loss in entropy without commensurate gain in energy.

In some contrast, forming the sixth and subsequent sequential helical residues will have much less positive free energy and higher probability because addition of each of these helical residues
results in additional stabilizing interaction such as the hydrogen bond. Thus, there is low nucleation probability and much higher propagation probability (like a zipper).

In the zipper model, the probability of a particular structure with helical (h) and coil (c) residues is a product of the probabilities of the conformations of the residues. In the model, the probability for a c-residue is 1 and the probability for a h-residue which follows a h-residue is $s$. These probabilities are the same as those of the independent-residue model (see Eq. 2.1). For a h-residue at the N-terminus of the peptide or a h-residue which follows a c-residue, the probability is $\sigma s$ where $\sigma$ is a “nucleation parameter” and has order-of-magnitude $10^{-3}$-$10^{-4}$. For the sequence chhch:

$$\text{Probability} = 1 \times \sigma s \times s \times 1 \times \sigma s = \sigma^2 s^3$$  \hspace{1cm} (2.9)

DB pp. 505-507 gives an elegant description of the generation of the total partition function for a peptide of arbitrary length $N$. An even more complete description is found in Cantor and Schimmel, pp. 1060-1066. According to Cantor and Schimmel, for large $N$:

$$Q \approx (\lambda_1)^N$$

with:

$$\lambda_1 = \frac{1}{2} \times \{(1 + s) + [(1 - s)^2 + 4\sigma s]^{1/2}\}$$  \hspace{1cm} (2.10)

Using Eq. 2.5:

$$f_h = \frac{1}{2} \times (s/\lambda_1) \times \{1 + [(s - 1) + 2\sigma]/[(1 - s) + 4\sigma s]^{1/2}\}$$  \hspace{1cm} (2.11)

For $\sigma = 10^{-3}$, a value of $s = 1.01$ corresponds to $f_h = 0.578$ and a value of $s = 0.99$ corresponds to $f = 0.422$. For lower values of $\sigma$, there will be a sharper transition and for larger values of $\sigma$, there will be a broader transition.

The zipper model appears to fit experimental data fairly well and is also based on reasonable assumptions about helical structure.