BT5340: Protein Folding and Stability
Assignment 02

Ipsita Manthan, Harikaran S., Aadit Mahajan

Consider a 15-residue long contiguous helix that you have synthesized via solid-phase peptide synthesis. Now, assume that the unfolding thermodynamics of this helix is well described by the Zimm-Bragg theory of helix-coil transition with a σ of 10-3.

Q1. Calculate the free-energy profile as a function of number of helical residues when the residue-level equilibrium constant s is significantly greater than 1 (assume a range of values, if you prefer), for a range of temperatures from 278 to 368 K (scenario a). Similarly, do the same when $s < 1$ (scenario b). Compare the free-energy profiles. What do you observe when comparing the two different scenarios as a function of temperature? What do the differences means? In both these scenarios, the parameters are independent of the sequence.

Ans:

According to the Zimm-Bragg Model, s is an equilibrium constant for converting a completely random coil conformation to a helix conformation. It is represented as a ratio of no. of helical conformations over the no. of coil conformations:

$$s = [H]/[C] \qquad \text{eq. 1}$$

For a value of s greater than 1, it favours a residue to attain a helical conformation and for the values of s smaller than 1, a residue attaining a helical conformation is unfavourable.

In terms of free energy profiles, residue with coil conformation has much more energy due to more degrees of freedom and increased entropy but as the residue attains a helical like conformation its free energy is decreased due to bonded state and decrease in the degrees of freedom.

Since, k is an equilibrium constant and represented as:

$$k = [\text{Folded state}]/[\text{Unfolded state}] = e^{-\Delta G/RT} \qquad \text{eq. 2}$$

On comparing eq. 1 & 2: expression of s in terms of free energy state of a particular residue can be written as:

$$s = e^{-\Delta G_{(C \rightarrow H)}/RT}$$

Taking logarithm both sides the expression becomes:

$$\ln s = -\Delta G_{(C \rightarrow H)}/RT \qquad \text{eq. 3}$$

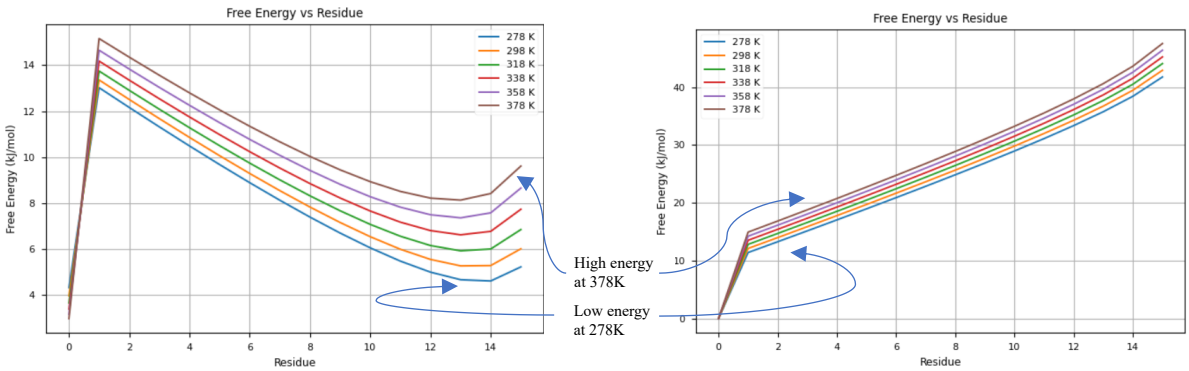

Figure 1 Free energy plots for Scenario A (s >1) and scenario B (s<1)

So, according to the equation 3, as temperature increases the equilibrium constant of a residue to convert from C to H decreases which means the free energy associated with the residues increases that is why as the temperature increases the free energy profile of 15-residue sequence increases e.g. the profile obtained at 378K is higher than profile obtained at 278K (in both scenario A & B).

According to the graph obtained in scenario A & B, it has been observed that the free energy associated with the first residue to convert from random coil to helix is very much high. This is because the attainment of first helix turn is difficult due to massive decrease in entropy hence requirement of more energy (*Zimm and Bragg 1959*). This is called nucleation and crucial for the helix formation. There is sharp increase in energy for nucleation is due to probability of the first residue to attain helix like conformation is very less can be stated as:

$$p = k\, e^{-\Delta G/RT} \qquad \text{(where } k \text{ is proportionality constant)}$$

Now, in scenario A, the energy associated with each residue decreases after nucleation and the 15-residue sequence attains helical conformation, a stable conformation with low energy state whereas in scenario B, after nucleation the subsequent residue attaining helical-like conformation is high energy state. This is because in scenario B the equilibrium constant s is smaller than 1 which means the situation is not in favor to convert a random coil to helix, thus to obtain a residue a helical-like conformation required high energy. Therefore, scenario B is more unlikely to occur.

Again, in both scenarios we can see, 15th residue is associated with high energy this can be explained as first four residues of the and last four residues of the peptide lack intrahelical hydrogen bonds and free energies associated with these positions also differ and these they get stabilized by forming side-chain-main chain hydrogen bonds, electrostatic interactions with helix dipole, exposure to solvent, conformational entropy or hydrophobic groups capping (Aurora and Rosee 1998, Iqbalsyah and Doig 2004).

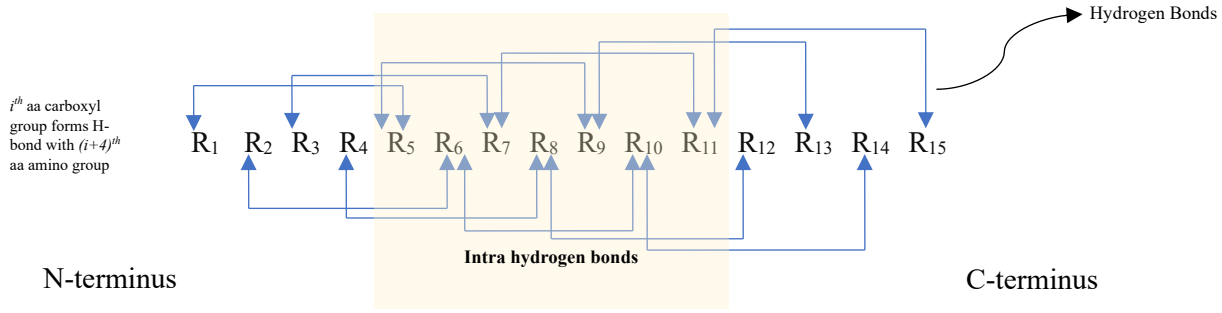

Figure 2 A Representation of ith aa carboxyl group forms H-bond with (i+4)th aa amino group

Q2. Consider another scenario wherein the 15-residue sequence has two stretches of helices in the folded state – i.e. residues 3 – 7 and 10 – 14 are helical ($s > 1$), while the rest are coil-like even at 298 K ($s < 1$). How does the free-energy profile now vary when compared with scenario a.

Ans: i) ii)

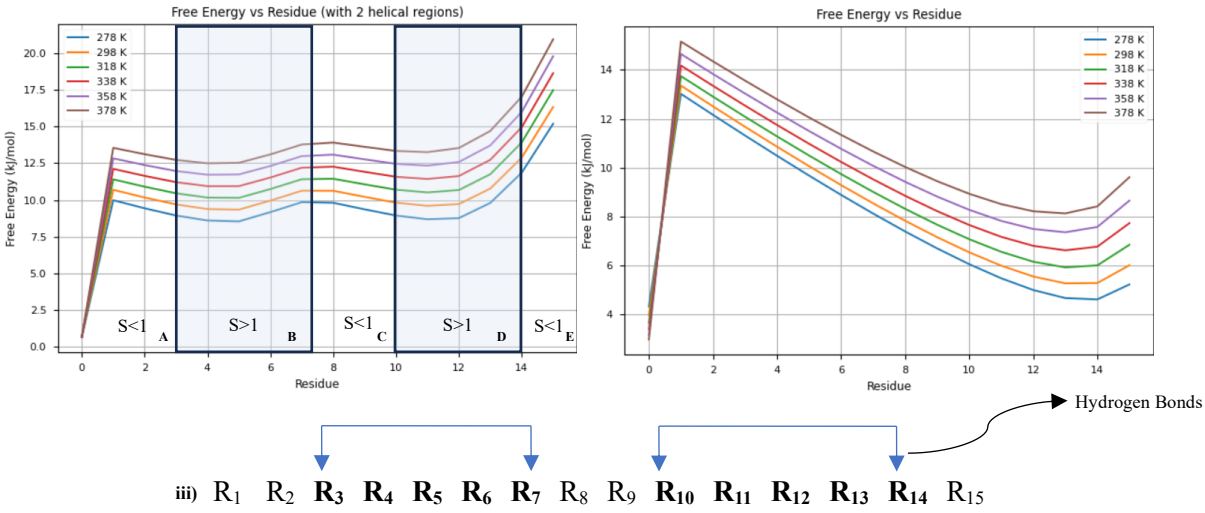

Figure 3 Comparison between sequence having two helical stretches interrupted by a random coil stretch (i) and a sequence in which all residues are helical (ii) (scenario C). Representation of the sequence (iii)

The graph obtained *(i)*, regions A, C and E having residues with s<1, i.e. coil conformation, whereas regions B and D have residues with s>1, i.e. helical conformation.

The free energy profile for regions A, C and E are higher as compared to regions B and D. As discussed earlier, A, C and E have s<1, which means helix formation in these regions is unfavourable thus high energy states as compared to the region whose s>1 which are low energy state.

When the helical stretches are separated by a random coil stretch, it requires more energy to achieve the native state because, as we can see from the two graphs, (i) is less stable than (ii) also nucleation energy of (i) is less than (ii). Additionally, the hydrogen bond is formed between R_3 and R_7 (first stretch); R_{10} and R_{14} (in the second stretch) may be due to which R_4, R_5, R_6, R_{11}, R_{12}, R_{13} attain helical conformation, which is thermodynamically more stable.

As compared to scenario A, scenario C is a high-energy microstate and thermodynamically less stable as the average energy of scenario C is more (contributed by the random coil stretches) than the average energy of scenario A. In addition to this, scenario A is stabilised by more no. od hydrogen bonds, and scenario C is stabilized by two hydrogen bonds (Fig 3 and Fig 4(iii)) more the hydrogen bonds, the more energy is utilized to form the bonds; hence, the system is stabilized.

Q3. Predict the thermal unfolding curves, i.e. the fraction helicity, for these three scenarios and identify the melting temperature.

Ans:

i) Scenario A **ii) Scenario C**

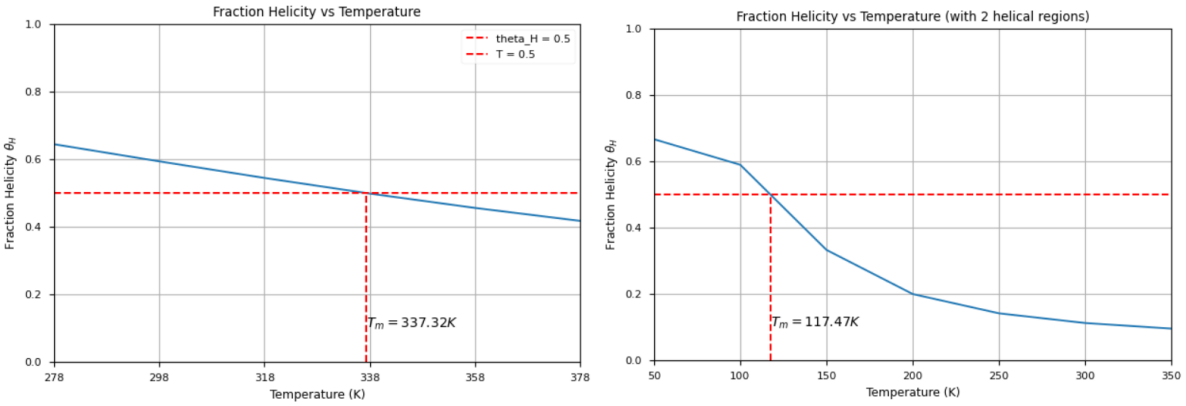

Figure 4 Comparison between the fraction helicity between scenario A and scenario C and Tm value calculated.

Melting temperature is a point at which 50% of the peptide has been unfolded, which is the indirect measure of the energy required to unfold or destabilise a peptide. As we can see from the graphs obtained T_m (=337.3K) of scenario A is greater than T_m(=117.4K) of scenario C as discussed earlier peptide in scenario A is more stabilised due to s>1, fraction helicity is more, which implies more hydrogen bonds are contributed for its stabilisation thus it requires more energy to destabilise.

In scenario C, two stretches of helices are present, which contributes to some kind of stability but has a very low Tm value.

Scenario B cannot have a valid fraction helicity curve because the s value is less than 1, in which case the helix is more unfavorable than the coil, preventing the helix formation altogether.

Q4. Predict the number of microscopic routes available for the helix in scenario A to fold

(starting from the fully unfolded state) at 298 K using a Monte-Carlo (MC) scheme and

the Metropolis criterion.

Ans:

For a 15 amino acid resides, from fully unfolded to folded, there are 2^n (32768) microstates in between for helix to coil transition. By using Monte-Carlo simulation and metropolis criterion, the process from fully unfolded (000000000000000) to folded (111111111111111) was repeated 10000 times to get the microstates that are frequently occurring. Here, the table of microstates and its frequency were listed (only forward conversion is shown).

Here, we do observe that the microstate that occurs most frequently at the start of the transition from completely unfolded to completely folded involves the microstate that has the nucleation near the central part of the peptide, indicating that the propagation is bidirectional in this model. This is corroborated in the model by the absence of a propagation direction constraint in the model. The central nature of the nucleating residues may also be preferred due to the larger number of accessible microstates to bidirectionally complete the transition from unfolded to folded state.

Table 1: Frequency of microstates (Folded and unfolded states were not shown).

Microstate	Number of Helices	Frequency
“0000000010000000”	1	4855
“0000000011000000”	2	5179
“0000000011100000”	3	5277
“0000000111100000”	4	5389
“0000001111100000”	5	5596
“0000011111110000”	6	5827
“0000111111110000”	7	6064
“0001111111110000”	8	6362
“0001111111111000”	9	6589
“0011111111111000”	10	6841
“0011111111111100”	11	7176
“0011111111111110”	12	7319
“0111111111111110”	13	7710
“0111111111111111”	14	6916

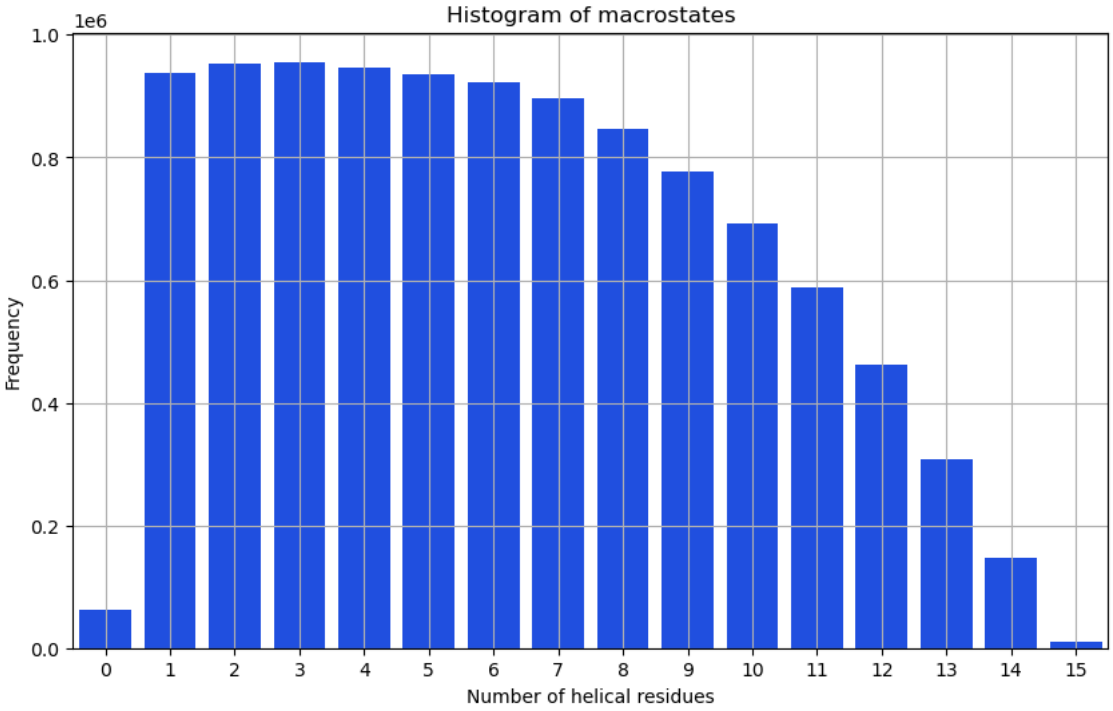

Figure 5 Histogram of macrostates

This plot shows evidence of the amount of time a peptide spends in each macrostate. Since the multiplicity of the macrostates is proportional the time spent by the peptide in that state (due to probability values derived from statistical mechanics and ensemble models), most paths can be said to have this correlation between the macrostate corresponding to a given microstate and the amount of time it spends in that microstate.

As for the folding paths that the peptide tends to follow, the residues that are involved in the transition pathways are the central ones with a high probability. This information is in the graphs below, which show the frequency of the microstates in the folding paths, where the microstates have been plotted separately according to their corresponding macrostate. At a low number of helical residues in the folding pathway, the central residues are preferred to be folded over the residues on the side due to their access to more microstates. As the number of helical residues increases, the distribution becomes more evenly spread.

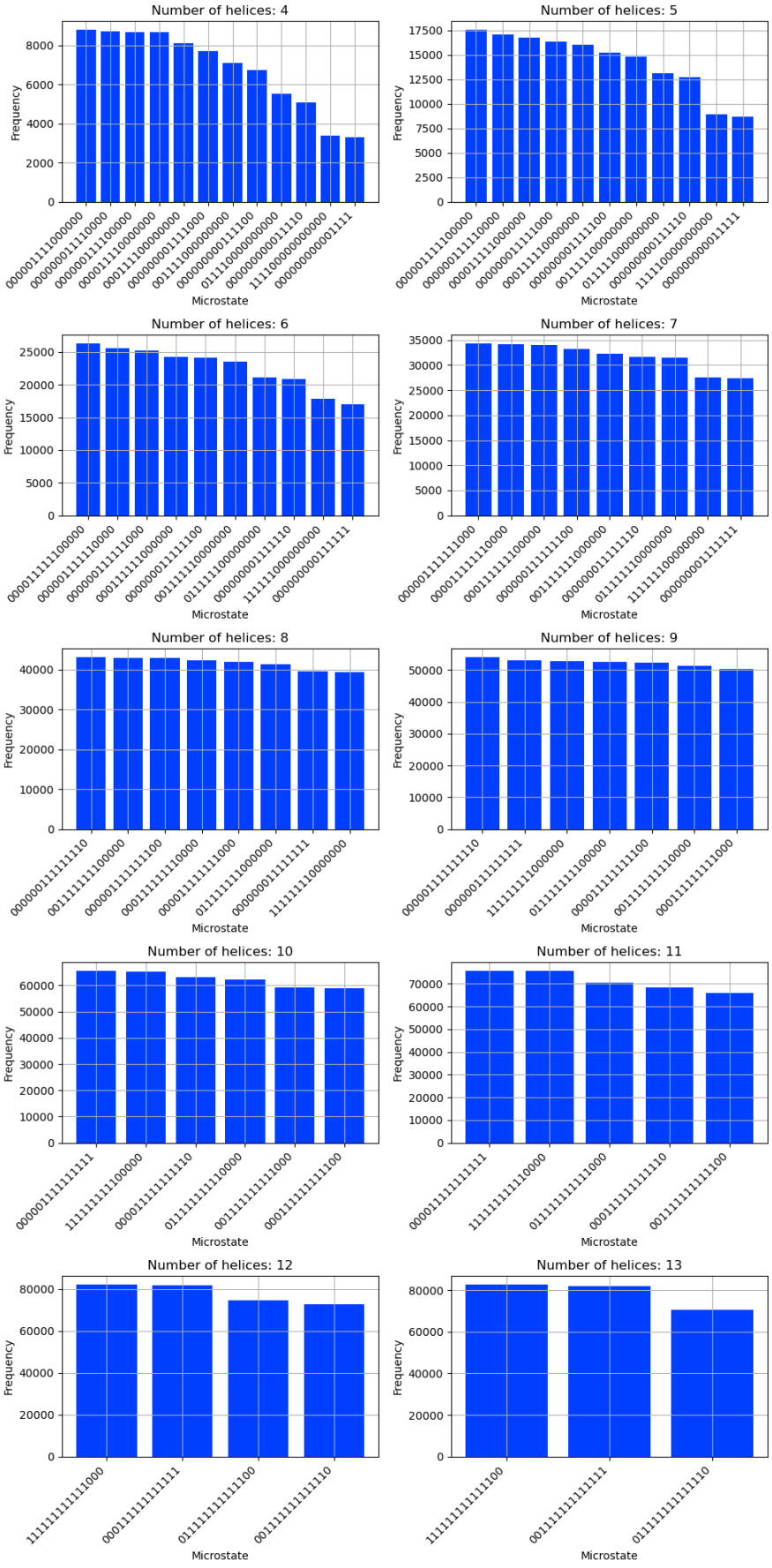

Figure 6 Frequency of each microstate in the folding paths

REFERENCES

1. Aurora, R. and G. D. Rosee (1998). "Helix capping." Protein Science **7**(1): 21-38.

2. Iqbalsyah, T. M. and A. J. Doig (2004). "Effect of the N3 residue on the stability of the α-helix." Protein Science **13**(1): 32-39.

3. Zimm, B. H. and J. Bragg (1959). "Theory of the phase transition between helix and random coil in polypeptide chains." The journal of chemical physics **31**(2): 526-535.