BT2042: Fundamentals of Biophysical Chemistry

- G Prashant (BS17B011)

Characterization of unfolding of a protein by following the change in signal at a single wavelength as a function of temperature

Introduction

Proteins perform most of the work of living cells. Understanding how cells work requires understanding how proteins function. In order to characterize the structure and functions of proteins, and to detect protein-protein interactions, various kinds of spectroscopic estimation techniques are used for analysing its 3-dimensional structure and folding properties. These techniques can also be used to determine if a mutation affects the conformational stability of the protein.

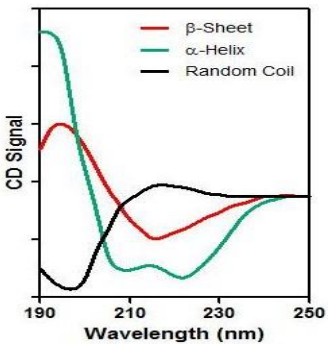
The most common methods used are listed below:

- X-Ray Crystallography of protein crystals
- Nuclear Magnetic Resonance (NMR) Spectroscopy
- Circular dichroism (CD) spectroscopy

For example, circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure result in a spectrum which can contain both positive and negative signals.

Secondary structure can be determined by CD spectroscopy in the far-UV spectral region (190-250 nm). At these wavelengths the chromophore is the peptide bond, and the signal arises when it is located in a regular, folded environment.

Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic shape and magnitude of CD spectrum, as shown below



In general, the changes in any signal as a function of temperature, at characteristic wavelengths, can be used to determine the thermodynamics of unfolding, i.e., the van't Hoff enthalpy (ΔH) and entropy (ΔS) of unfolding, the midpoint (inflection point) of the unfolding transition (T_M) and the free energy (ΔG) of unfolding. Analysis of the spectra obtained as a function of temperature may also be useful in determining whether a protein has unfolding or folding intermediates.

Equations

When a molecule undergoes an unfolding transition between two states: folded, F, and unfolded, U. At any temperature, T, the equilibrium constant of unfolding, K_{eq}, when folded state is taken as the reference, is

$$K_{eq} = [U]/[F]$$

The probability of folded state if given by,

$$P_f = 1/(1+K_{eq})$$

The probability of unfolded state is given by,

$$P_u = 1 - P_f = K_{eq}/(1 + K_{eq})$$

[F] and *[U]* are the concentrations of the folded and unfolded forms respectively.

The free energy of unfolding is,

$$\Delta G = -RT\ln K_{eq}$$

R is the Gas constant, and T is the absolute temperature (Kelvin).
The free energy of unfolding as a function of Temperature, in terms of the thermodynamic parameters, is given by,

$$\Delta G = \Delta H - T\,\Delta S$$
$$= \Delta H_m + \Delta C_p\,(T-T_m) - T\,[\Delta S_m + \Delta C_p \ln(T/T_m)]$$

Where T_m the melting temperature of the protein (the temperature at which there is equal concentration of both folded and unfolded states of the proteins), ΔC_p is the change in heat capacity going from the folded to the unfolded state and ΔH_m and ΔS_m are parameters and,

$$\Delta S_m = \Delta H_m/T_m$$

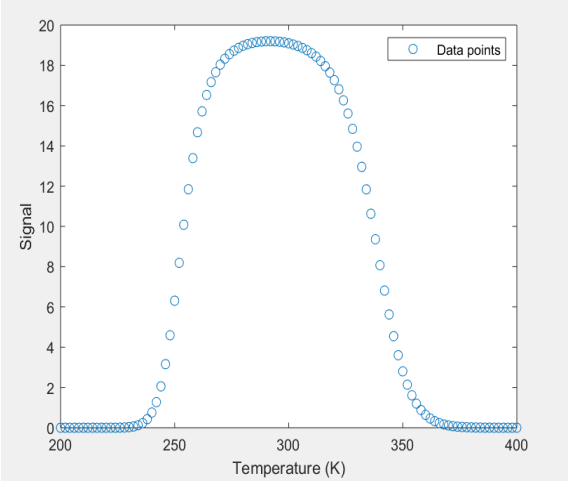
The observed ellipticity at any temperature θ_t, can therefore be given by the equation,

$$\theta_t = \theta_f . P_f + \theta_u . P_u$$

where θ_f *and* θ_u are the ellipticity of the fully folded and fully unfolded form respectively.

Curve Fitting

After experimental observations, the signals at different proteins with 50 residues were obtained and plotted, as shown below:



The initial parameters ΔH_m, T_m and ΔC_p were estimated as follows:

ΔH_m = 2.9 *N KJ/mol = 145 kJ/mol = 145000 J/mol
ΔC_p = 50*N J/mol = 2500 J/mol
T_m = corresponds to the point where the gradient of the signal is minimum

The expressions of θ_f and θ_u are dependent on temperature, which can be assumed to have a linear relationship with parameters m1, c1, m2 and c2.

$\theta_f = m_1 * T + c_1$
$\theta_u = m_2 * T + c_2$

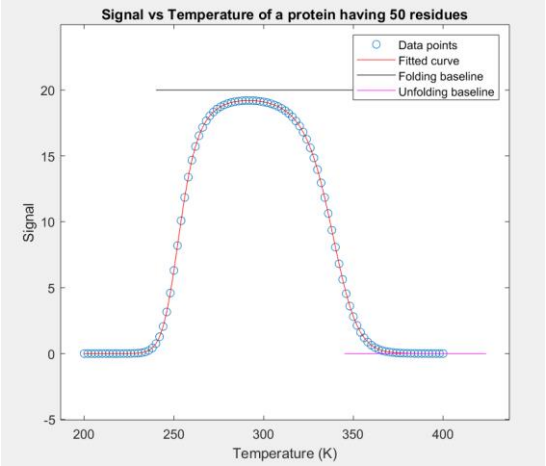
Parameters m_1, m_2, c_1 and c_2 are estimated by choosing few data points in the folded and unfolding regions of the plot and performing linear regression for both cases.

Nonlinear Regression if finally performed by feeding the data points and the initial guesses to the lsqcurvefit() function in MATLAB.

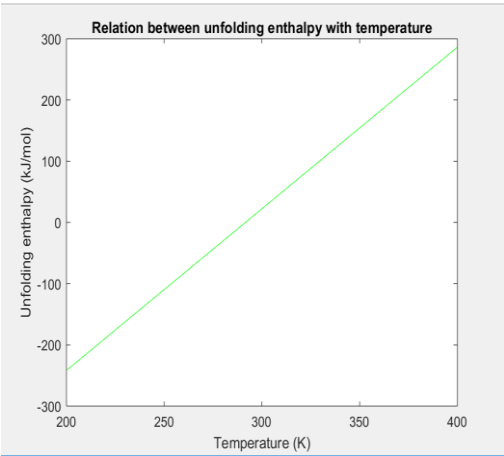
The function estimates the final value of the parameters such that the χ^2 loss function is minimum.

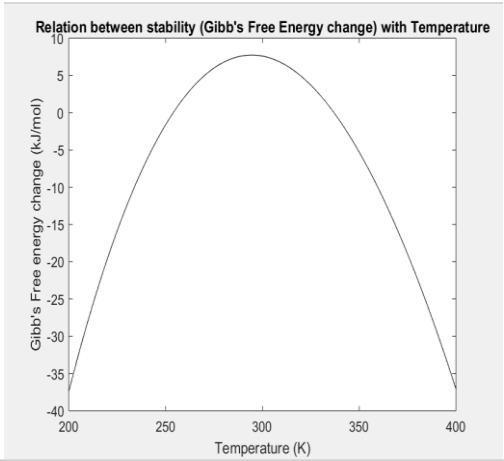
Observations

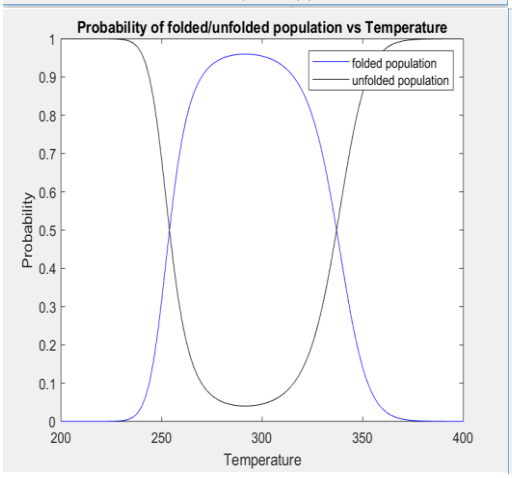
The fit obtained along with the baselines for folded and unfolded states is shown below:

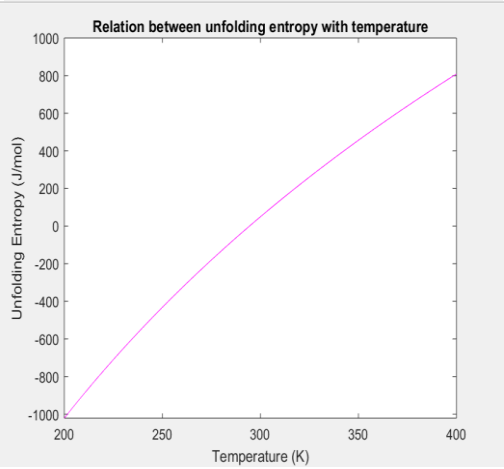


The Relations between the unfolding enthalpy, stability, unfolding entropy, probability of folded and unfolded states with Temperature are given below:









Inferences

- The predicted curve is found to fit the data almost perfectly.
- The unfolding Enthalpy (ΔH_m) linearly increases with temperature. The enthalpy is negative at lower temperatures, indicating that unfolding is exothermic at low temperatures and endothermic at higher temperatures.
- The unfolding entropy (ΔS_m) increases with temperature logarithmically. The entropy is negative at lower temperatures, indicating that unfolding leads to more order at low temperatures and higher order at higher temperatures.
- The slopes of the folding and unfolding baselines for purely folded or purely unfolded proteins are close to zero, suggesting that the signal for these states are not significantly dependent on temperature.
- The folded baseline lies above the maxima of the signal data. The predicted parameters are such that it minimizes the χ^2 loss function. This shift occurs as the predicted folded baseline corresponds to the population containing many such data points.
- The stability (change in Gibb's Free energy) has a parabolic relation with Temperature, where at certain Temperatures (below and above the two melting temperatures), the value of change in Gibb's Free energy is negative, which suggests that unfolding of a protein is thermodynamically favourable at these temperatures.
- On the other hand, at certain temperatures between the two melting temperatures, the value of change in Gibb's Free energy is positive, which suggests that unfolding of a protein is thermodynamically unfavourable at these temperatures.
- The probability of folded state initially increases with increase in temperature, and then decreases with temperature. The curve is symmetric about the maxima.
- The probability of unfolded state initially decreases with increase temperature, and then increases with temperature. The curve is symmetric about the minima.

Conclusion

From the above inferences, it can be clearly seen that a protein retains its folded state only for a particular range of temperatures. Denaturation or unfolding of a protein is thermodynamically favourable at temperatures outside this range. At lower temperatures, the Gibbs energy of hydration is negative and increases in magnitude at a temperature decrease. The unfolding process is enthalpically driven at low temperatures. As a result, the polypeptide chain, tightly packed in a compact native structure, unfolds at a sufficiently low temperature, exposing internal nonpolar groups to water. The folded versus unfolded state depends on two competing factors- entropy and stabilizing interactions. The folded state has low entropy, but many stabilizing interactions. The unfolded state has high entropy, but very few stabilizing interactions. As we increase temperature, the entropy component of free energy starts to dominate.

References

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2752288/
- https://www.ncbi.nlm.nih.gov/books/NBK26820/
- https://www.ncbi.nlm.nih.gov/pubmed/2225910