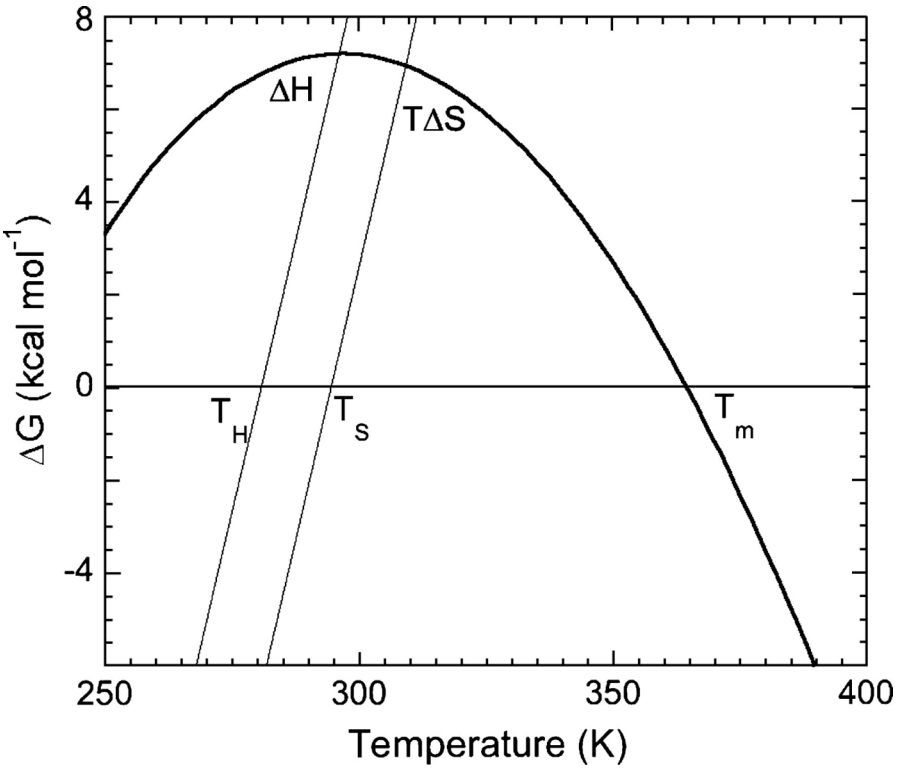
Problem Set #4:

1. Protein stability curves, first introduced by Becktel and Schellman, use a modified version of the Gibbs-Helmholtz equation to plot free energy of stabilization as a function of temperature. The equation used and a stability curve for a hypothetical protein are shown below:

$$\Delta G(T) = \Delta H_m (1 - T/T_m) - \Delta C_P \times [(T_m - T) + T \ln (T/T_m)]$$

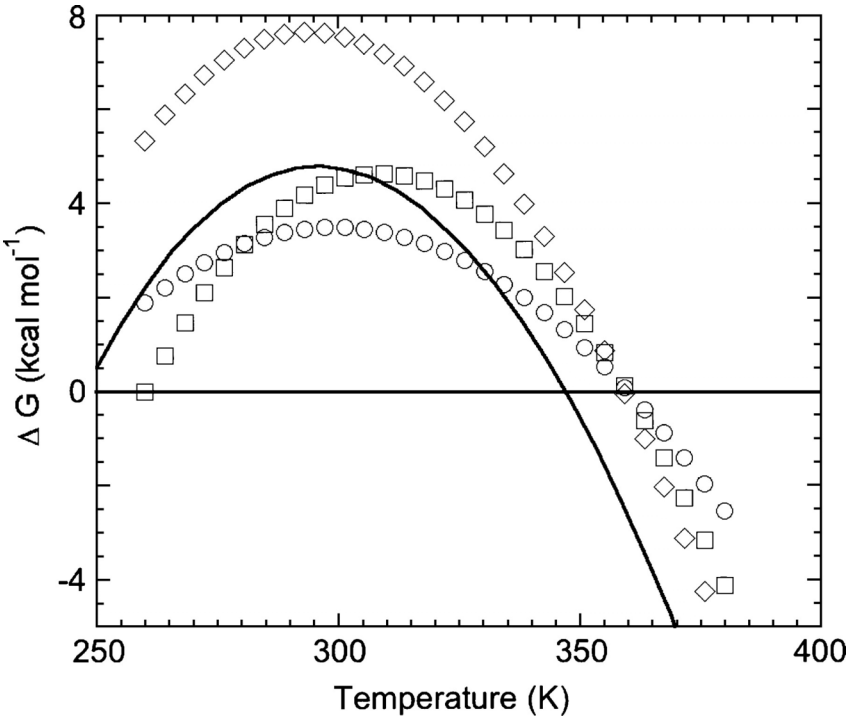


A. Show that the folding equilibrium constant, K = [N] / [U], is equal to one at T_m. Similarly, prove that protein stability is maximal at T_S, and that K is maximal at T_H. How could you determine each of these temperature values experimentally?

B. The form of the stability plot suggests analysis of thermal unfolding curves via a simple linear extrapolation of ΔG *vs.* T from the transition region to room temperature will lead to large errors. Will linear extrapolation underestimate or overestimate the stability of the folded *vs.*

unfolded protein at room temperature? In terms of the given equation, what has to be true for linear extrapolation to be rigorously correct?

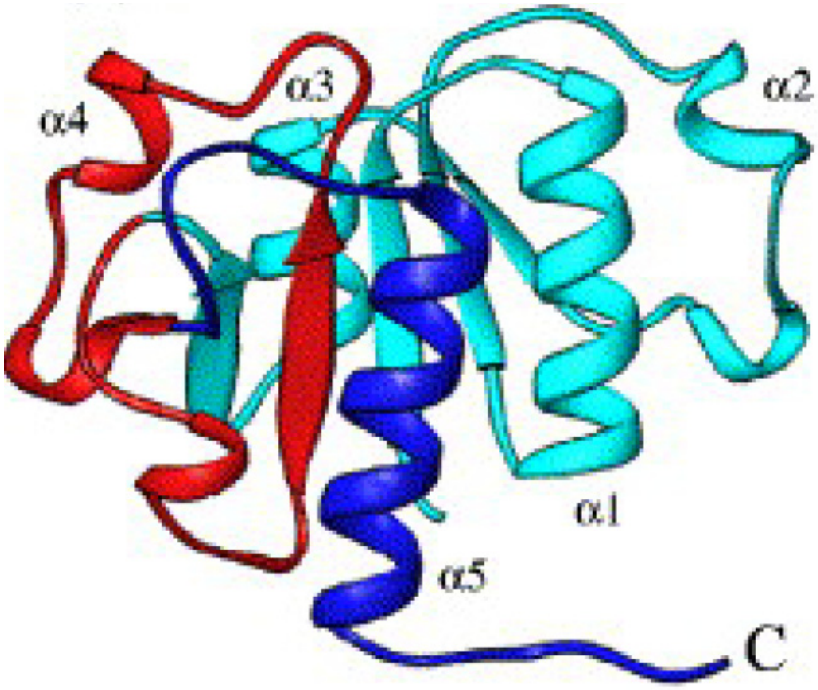
2. Thermophilic organisms survive at high temperatures, and their proteins must remain folded under these extreme temperatures. Nature uses three thermodynamic strategies to stabilize thermophilic proteins compared to homologous proteins of mesophiles: (a) increasing the value of ΔH_S, the change in enthalpy measured at T_S, without compensating changes in ΔS, (b) reducing the value of ΔC_P, the heat capacity change associated with the folding transition, and (c) lowering the value of ΔS, the change in entropy for the folding transition. Shown below are stability curves for a hypothetical mesophilic protein (solid line) and three homologous thermophilic proteins.



A. Associate each of the thermophilic stability curves with one of the three thermodynamic stabilization strategies described above.

B. The figure suggests proteins may become unstable at low temperatures, in addition to the typical instability at high temperature. Such "cold denaturation" is indeed observed for some proteins. What physical phenomenon, involved in protein folding, exhibits a similar temperature dependence? Explain in molecular terms how this phenomenon might be related to cold denaturation.

3. The 160-residue protein YibK from *Haemophilus influenzae* is a methyltransferase (MTase) characterized by a distinct fold. A ribbon diagram derived from the crystallographic structure contains five α-helices and six β-strands as shown below:

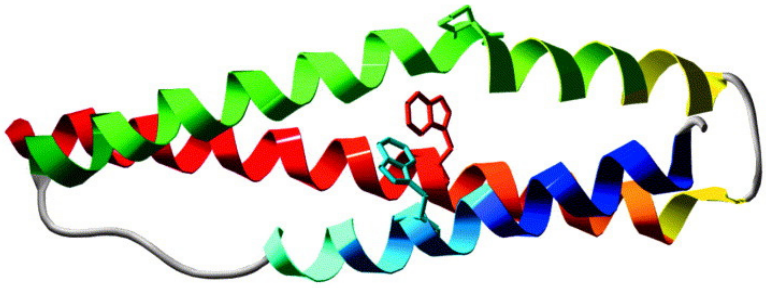


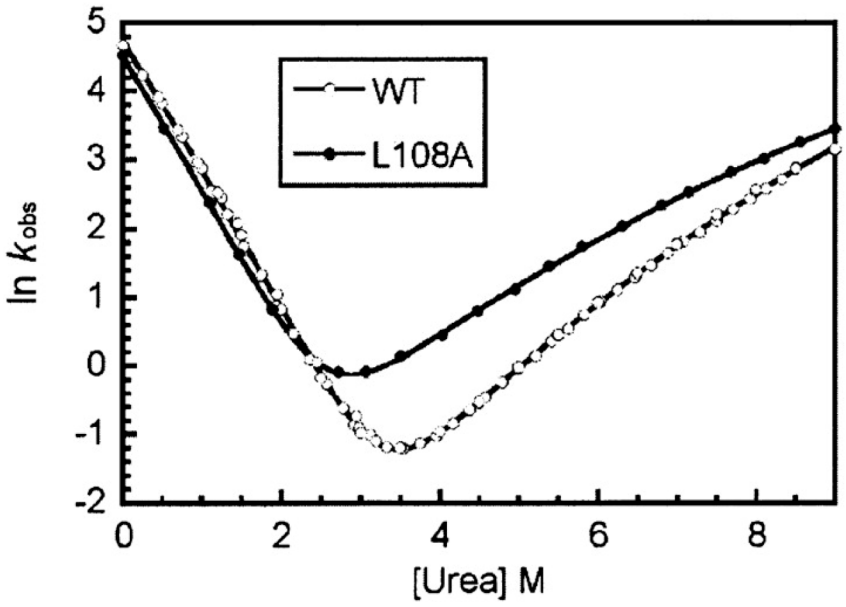
A. What motif(s) are present in the YibK structure? To what protein fold sub-family does the structure belong?

B. Draw a topological diagram of the YibK fold. The overall fold of the backbone chain of YibK contains a unique topological feature that poses a challenge to current protein folding models. What is this feature?

4. Draw the full chemical structure of acetyl-L-arginine-*N*-methylamide, a "capped" Arg residue, as it would exist at physiological pH. Show correct stereochemistry at chiral centers. Label on your structure each of the phi, psi and chi angles used to specify the conformation of this molecule.

5. Spectrin domains are found as multiple tandem repeats in a number of cytoskeletal proteins such as the spectrins, α-actinin and dystrophin. The repeat unit folds into a coiled coil with three antiparallel helices. The structure of the 16th repeat from chicken brain α-spectrin (R16) has been determined and is shown as a ribbon diagram. Chevron plots derived from stopped-flow fluorescence experiments on wild-type R16 (WT) and the L108A mutant are also shown below:





A. Estimate the difference in stability between the WT and L108A mutant. What assumptions do you make in giving this estimate?

B. The unfolding arms of these chevron plots show small but noticable curvature, indicating a deviation from pure 2-state behavior. Possible explanations include a broad transition state region, or presence of a high-energy on-path intermediate. Is the effect of the mutation L108A present early in the folding pathway, or is this likely a residue that makes many of its interactions late in folding? Explain.

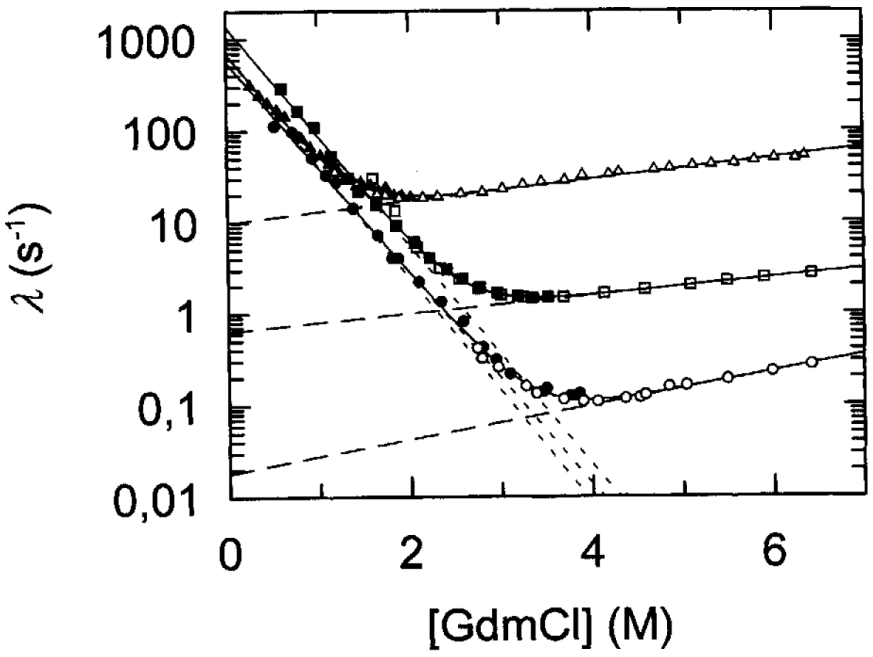
6. Let T_m be the temperature at the midpoint of a protein unfolding curve. Results from the analysis of thermal unfolding curves for a wild-type protein and two mutants are given below. Use this data to compute the T_m value for each protein, and then to estimate the $\Delta\Delta G$ of the unfolding reaction for each mutant relative to the wild-type.

Protein	**ΔH_m** (kcal/mol)	**ΔS_m** (cal/mol/deg)
Wild-type	95	295
Mutant 1	110	335
Mutant 2	100	315

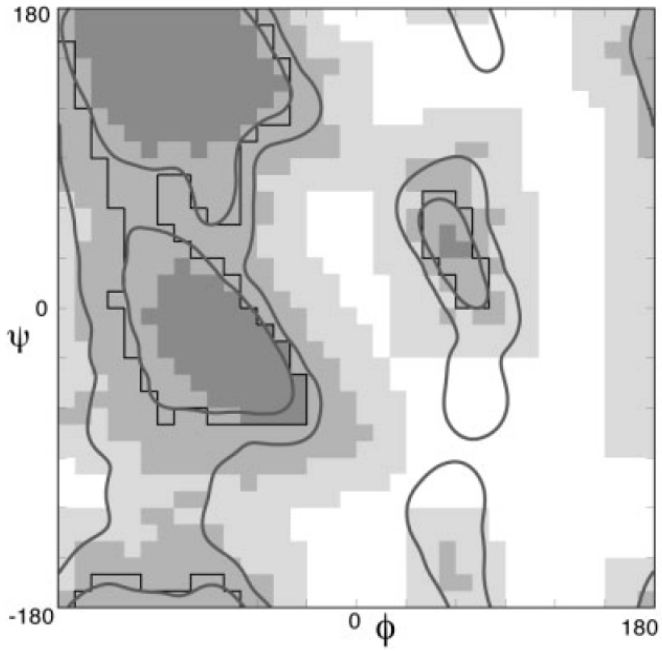
7. Shown below are data from guanidinium chloride-induced kinetic unfolding/refolding experiments on homologous cold-shock proteins from *B. subtilis* (▲/Δ), the thermophile *B. caldolyticus* (■/□), and the hyperthermophile *T. maritime* (●/○). The apparent rate constant, λ, is plotted as a function of denaturant concentration.

A. Estimate the rate constants for folding and unfolding of each of these proteins. What does the data suggest about the mechanism used by thermophilic organisms to achieve protein stability?

B. Estimate the $\Delta G_{U \rightarrow N}$ for the cold shock protein from each of these three organisms. What assumptions are made in arriving at your answers?



8. The consensus Ramachandran plot for L-alanine, as derived from a representative subset of the Protein Data Bank, is as shown below:



A. Draw the expected Ramachandran map for the uncommon, but naturally occurring, amino acid Aib (*ie*, α-aminoisobutyric acid; an amino acid with two methyl groups at the α-carbon atom). What is the relationship between your answer and the alanine data shown above?

B. What type of secondary structure do you think would be favored for a peptide sequence containing a mixture of several Aib residues with some of the 20 standard L-amino acids? Explain.