1 Homework 1: Protein Structure and Energetics

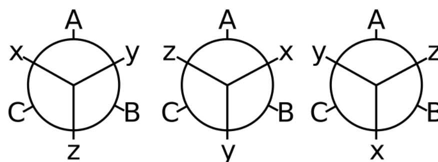
1. What do changes in free energy, enthalpy, and entropy measure?

2. You are studying a protein that favors the folded state by $-32.3\ kJ \cdot mol^{-1}$. You introduce a mutation that makes a new ion pair worth $-11\ kJ \cdot mol^{-1}$ formed only in the unfolded state.

   (a) What fraction of protein molecules are in the folded state for the wildtype protein?

   (b) What fraction of the protein molecules are in the folded state for the mutant protein?

3. A protein with 50 amino acids has 98 rotatable bonds in its backbone. To a first approximation, each bond can populate one of three configurations:



   The folded state is one unique combination of bond configurations; the unfolded state consists of all other combinations.

   (a) Given our definition of entropy from class, estimate the change in entropy for the folding of the protein backbone.

   (b) Do you think the assumptions above lead to an overestimate or understimate of the entropy change? (Why or why not?)

4. One interesting observation is that most proteins are *marginally stable*, meaning that their total stabilities are small, despite being the sum of large favorable and large unfavorable terms. On average, protein stabilities are about the same as 1-2 hydrogen bonds.

   (a) If there are 5,000 proteins in a cell, how many molecules of a protein are unfolded if it has a stability of:

      i. -10 kJ/mol?

      ii. -20 kJ/mol?

      iii. -30 kJ/mol?

      iv. -40 kJ/mol?

   (b) Why do you think organisms don't always evolve sequences that increase the stability of their proteins?

   (c) Can evolution "see" the difference between a stability of -20 kJ/mol and -30 kJ/mol for this protein? -30 kJ/mol and -40 kJ/mol? Does this change your answer in (b)?

5. You introduce an glycine to alanine mutation into a protein and find that it strongly stabilizes the folded state. You solve the crystal structure of both the wildtype and mutant protein and find their structures are essentially identical. How might this mutation be having its effect?

6. Consider the pentapeptide "DFPHR":

(a) Draw the peptide (including termini) with each titratable group in its most likely protonation state at pH 7.5.

(b) On the structure you drew, use an arrow to indicate which peptide bond is most likely to exist in the cis configuration. Explain why.

7. Open up "alpha-helix.pse" in PyMol.

   (a) Rotate the α-helix around. To orient yourself:

      i. Find the N-terminus.

      ii. Identify a single peptide bond.

   (b) Click "$A \rightarrow find \rightarrow polar\ contacts \rightarrow within\ selection$". This should create a set of yellow-dashed lines between hydrogen bond donors and acceptors.

      i. What atoms act as the primary hydrogen bond donors? What atoms act as acceptors?

      ii. How far away are the donors and acceptors of each hydrogen bond in primary structure? Put another way: how far apart in sequence are the interacting amino acids? (Is this fixed across the piece of secondary structure)?

   (c) What is the orientation of the sidechains relative to the piece of secondary structure:

      i. viewed lengthwise?

      ii. viewed looking down the helix from the N-terminus?

   (d) Helices have the sequence pattern $xHHxxHxxHHx$, where H are hydrophobic residues and x are polar/charged residues. What is significant about this pattern, given the structure of helices?

8. Open up "beta-sheet.pse". Reduced sidechains are shown as large rainbow-colored atoms.

   (a) Rotate the β-sheet around. To orient yourself:

      i. Find the N-terminus.

      ii. Identify a single peptide bond.

   (b) Click "$A \rightarrow find \rightarrow polar\ contacts \rightarrow within\ selection$". This should create a set of yellow-dashed lines between hydrogen bond donors and acceptors.

      i. What atoms act as the primary hydrogen bond donors? What atoms act as acceptors?

      ii. How far away are the donors and acceptors of each hydrogen bond in primary structure? Put another way: how far apart in sequence are the interacting amino acids? (Is this fixed across the piece of secondary structure)?

   (c) What is the orientation of the sidechains relative to the piece of secondary structure:

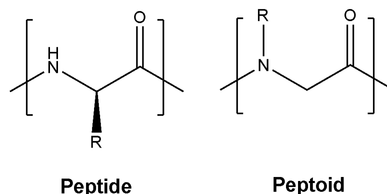
      i. viewed lengthwise?

      ii. viewed looking down the sheet from the N-terminus?

      iii. Many β-sheets form interaction surfaces that interact specifically with other large proteins. How might the orientations of these sidechains facilitate recognition?

   (d) Click $S \rightarrow$ *cartoon*. Is this a *parallel* or *anti-parallel* β-sheet?

9. Peptide analogs known as "peptoids" have been developed to overcome some of the problems with peptide-based drugs (e.g. protease sensitivity). Peptoids are N-substituted amino acids (see figure).



   (a) Would you expect a peptoid to be more or less flexible (i.e. have more available conformations) compared to a peptide with the same sequence? Please justify your answer.

   (b) Do you expect peptoids to form α-helices? Explain why or why not.

10. We're going to consider the structural and energetic effects of mutations to the protein lysozyme.

   (a) Which residues are within 3 Å of the residues I23, V74, and V121 in the wildtype structure (2BQA)?

   (b) The PDB files 2BQC, 2BQM, and 2BQI contain the structures for lysozyme with the mutations I23V, V74A, and V121A, respectively. Describe how the lysozyme structure responded to each mutation by comparing each mutant structure to the wild-type structure.

   (c) Based on your analysis, is protein structure sensitive or insensitive to sequence?

   (d) The PyMol command "get_area selection" calculates the surface area of the given selection. Use this command to find the surface area of the I23, V74, and V121 sidechains in the wildtype structure.

   (e) The I23V, V74A, and V121A mutations destabilize the protein (i.e. increase $\Delta G°$ folding) by 1.72, 1.80, and 7.28 $kJ \cdot mol^{-1}$, respectively. Give a structural explanation for their different effects on $\Delta G°$ folding.