Problem answers, Assignment #3 – Proteins

1. To analyze these experiments, you first have to transform the raw spectroscopic data into fraction of protein in the unfolded state, \( f_u \) and then into \( K_{eq} \) and \( \Delta G^o \). Just to be ornery, I am going to consider the reaction going in the unfolding direction; so \( \Delta G^o \) will refer to the free energy of unfolding. You should realize that this is the same problem as our usual way of treating the reaction going in the folding direction. Only the sign is changed!

\[
f_u = 2 \ (A - 0.5) \quad \text{[make sure you see why!!]}.
\]

This is tabulated and graphed below.

<table>
<thead>
<tr>
<th>( T, ^\circ C )</th>
<th>( A_{222}, \text{WT} )</th>
<th>( f_u, \text{WT} )</th>
<th>( K, \text{WT} )</th>
<th>( A_{222}, \text{A25S} )</th>
<th>( f_u, \text{A25S} )</th>
<th>( K, \text{A25S} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.50</td>
<td>~0</td>
<td>-</td>
<td>0.50</td>
<td>~0</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>0.49</td>
<td>0.02</td>
<td>0.0204</td>
<td>0.50</td>
<td>0.04</td>
<td>0.0417</td>
</tr>
<tr>
<td>20</td>
<td>0.53</td>
<td>0.06</td>
<td>0.064</td>
<td>0.52</td>
<td>0.04</td>
<td>0.0417</td>
</tr>
<tr>
<td>30</td>
<td>0.60</td>
<td>0.20</td>
<td>0.250</td>
<td>0.55</td>
<td>0.10</td>
<td>0.111</td>
</tr>
<tr>
<td>40</td>
<td>0.76</td>
<td>0.52</td>
<td>1.083</td>
<td>0.67</td>
<td>0.34</td>
<td>0.515</td>
</tr>
<tr>
<td>50</td>
<td>0.88</td>
<td>0.76</td>
<td>3.80</td>
<td>0.81</td>
<td>0.62</td>
<td>1.632</td>
</tr>
<tr>
<td>60</td>
<td>0.96</td>
<td>0.92</td>
<td>11.5</td>
<td>0.93</td>
<td>0.86</td>
<td>6.143</td>
</tr>
<tr>
<td>70</td>
<td>1.00</td>
<td>~1</td>
<td>-</td>
<td>0.97</td>
<td>0.94</td>
<td>15.67</td>
</tr>
<tr>
<td>80</td>
<td>1.00</td>
<td>~1</td>
<td>-</td>
<td>0.99</td>
<td>0.98</td>
<td>49</td>
</tr>
</tbody>
</table>

a. \( T_m \) may be read right off the \( f_u \) curve at the half-point (see graph).

b. We get \( \Delta G^o(T) \) from the values of \( f_u \):
\[
\Delta G^o(T) = -RT \ln(f_u/(1-f_u)).
\]

These are plotted as a function of \( T \), below, to derive \( \Delta S^o \). Note that all the thermodynamic functions refer to the unfolding process. From curve, \( \Delta S^o = 83 \text{ cal/mol-K} \) for wt and mut.

c. Then, \( \Delta H^o = T_m \Delta S^o = +25.9 \text{ kcal/mol for wt, and } +26.5 \text{ kcal/mol for mutant.} \)

d. Read the answer right off the plots above.

e. At 15\(^\circ\) we can use: \( \Delta G^o(T) = \Delta H^o - T \Delta S^o \), and our previously determined values of \( \Delta H^o \) and \( \Delta S^o \) (remember to use Kelvin!):
\[
\begin{align*}
\Delta G^o_{\text{wt, kcal/mol}} &= +2 \\
\Delta G^o_{\text{mut, kcal/mol}} &= +2.6
\end{align*}
\]

(Positive for UNfolding, right!)

Make sure that these numbers make sense to you, given the raw data!
2. Look it up in Vandenburg et al. PNAS 95, 2056, 1998. All of these manipulations reduce the
number of allowable configurations of the polypeptide chain in the unfolded state, but they don’t
affect the folded state. They therefore “entropically disfavor” the unfolded state, which is
equivalent to favoring the folded state. They thus make the folded protein more stable.

3. The key to this problem is to remember two things:
   * $\Delta G^o$ is a function of $T$: $\Delta G^o(T) = \Delta H^o - T\Delta S^o$ (where $H, S$ are NOT $T$-dependent)
   * $T_m = \Delta H^o/\Delta S^o$

   Thus, at a given temperature $T$, we can write the $\Delta G^o$ for each protein, 1 and 2, as:

   $\Delta G^o_1(T) = \Delta H^o_1 - T\Delta S^o$
   $\Delta G^o_2(T) = \Delta H^o_2 - T\Delta S^o$ (note same $\Delta S^o$ for each, by assumption)

   Also, $\Delta H^o_1 = T_{m1}\Delta S^o$
   $\Delta H^o_2 = T_{m2}\Delta S^o$

   Making the substitution, we get:
   $\Delta G^o_1(T) = (T_{m1} - T)\Delta S^o$
   $\Delta G^o_2(T) = (T_{m2} - T)\Delta S^o$

   So $\Delta G^o(T) = \Delta G^o_2(T) - \Delta G^o_1(T) = (T_{m2} - T_{m1})\Delta S^o$, as was to be proven.

   This means that the difference in free energy of folding can simply be read off the denaturation
curve from the shift in $T_m$, if you know $\Delta S^o$. Check with problem #1 to see if this works!

4. This is a similar answer as in problem 2. Near a surface, the unfolded polypeptide chain has a
restricted amount of space that it can wiggle around in (half of a sphere, in fact). This reduces the
number of configurations — and thus the entropy — of the unfolded protein, but it doesn’t affect
the folded protein at all. The number of configurations of the unfolded protein is reduced by a
factor of $\sim 2$, so you might expect the entropy of unfolding to be less favorable by the amount
$R\ln 2$, which is equivalent to a free-energy stabilization of the folded state by:

   $RT\ln 2 \sim 0.4$ kcal/mol.
5. KcsA problem

By manipulating the KcsA backbone in RasMol (residues 75-78), I estimate the $\phi, \psi$ angles:

<table>
<thead>
<tr>
<th>Res</th>
<th>$\phi,^\circ$</th>
<th>$\psi,^\circ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T75</td>
<td>-45</td>
<td>-10</td>
</tr>
<tr>
<td>V76</td>
<td>+50</td>
<td>+40</td>
</tr>
<tr>
<td>G77</td>
<td>+90</td>
<td>+70</td>
</tr>
<tr>
<td>Y78</td>
<td>-60</td>
<td>-60</td>
</tr>
</tbody>
</table>

Note how V76 and G77 have nonstandard dihedral angles that fall into the “forbidden” first quadrant of the Ramachandran plot. It’s this feature that allows 3 carbonyls in a row to all be pointing in the same direction. Thus could not happen if there were not a glycine at position 77; the glycine here also allows the funny angles for the preceding valine. Note what would happen with steric clashes if there were an Ala at position 77!

6. DNA triplex problem

You throw equal molar concentrations, $m$, of W, C, and F into a reaction, and allow equilibrium to take place. Using appropriate spectroscopic tools, you find that for $m=10$ nM, 10% of the oligos are in the triple-helical form.

**a.** What is the equilibrium constant for triple-helix formation?

**b.** If you performed this same experiment with $m = 5$ nM, what would you find?

The equil const is: $K = h/(wcf)$ (K has units M$^{-2}$)

And for this problem: $w + c + f + 3h = 3m$ and $w=c=f$

So: $w + h = m$

$K = h/(m-h)^3$

Now define the fraction of oligos in helical form.

$f = (\text{number of strands in helix}) / (\text{number of total strands}) = 3h/3m = h/m$

So now we have the equil const in terms of f:

$K = f / m^3(1-f)^3$ (*make sure you get this!*)

So, if $m = 10^{-8}$ M, and $f = 0.1$, $K = 1.4 \times 10^{-15}$ M$^{-2}$

Now, you could plug this value of K into the cubic equ above and try to solve for f, but that would involve programming a numerical solution to the cubic. Instead, just think it out: at m=10 nM, f is pretty small, and if you cut m in half, f will be even smaller, so in both cases, (1-f) is close to 1. Therefore, rearrange the above equation:

$f/(1-f)^3 = m^3K \sim f$

So if you lower m by a factor of 2, you will lower $f/(1-f)^3$ by a factor of 4. So to a first approximation, f goes down 4-fold, to 0.025. At 5 nM of each oligo, about 2.5% of the strands will be in helical form.
7. The problem is asking you to estimate the perturbation of pKₐ of a “central” phosphate by the neighboring negatively charged phosphate backbone groups. For simplicity, assume that there are only two nearby groups to be considered – one on each side of the central group on the same RNA strand. From the structure of RNA, each of these is 7 Å distant from the central group.

Hey -- we’ve done this problem before, in a protein context – the Russell & Fersht paper! Go back and remind yourself why, if you want to calculate the pKₐ change due to a change in charge at a nearby residue:

\[ \Delta pK_a = \frac{-1}{(\ln 10) (4\pi \epsilon_r)} \Delta z / Dr = -244 \Delta z / Dr \]

That’s for a single charge nearby, but we are considering two nearby charges, one on each side. So:

\[ pK_a(\text{in RNA}) = pK_a(\text{isolated}) - 488 \ \Delta z / Dr \]

\[ pK_a(\text{isolated}) = 2.5 \quad \Delta z = -1 \quad r = 5.5 \ \text{Å} \ (\text{check out a structure like tRNA}) \quad \text{Let D=80} \]

Do the arithmetic:

\[ pK_a(\text{in RNA}) = 2.5 + 1.1 = 3.6. \]

This makes sense: the phosphate group binds the proton more tightly, since the binding is stabilized by the negative electrostatic potential set up by the nearby negative charges. Note that this is a crude calculation, since we have not taken into account the fact that the effective charges of the neighboring groups will themselves be altered. At very high ionic strength, the electrostatic influence of nearby groups will be suppressed, and the pKₐ will not be much perturbed.


We want to find the number, N, of residues in a polypeptide, such that half the sidechains are surface-exposed and half are buried in a folded spherical protein. Let:

- \( V \) = total volume of protein
- \( S \) = total surface area
- \( r \) = radius of the folded protein
- \( v_o \) = volume per residue
- \( s_o \) = surface area per residue

Let’s just write the total volume and surface area:

\[ V = Nv_o = \frac{4}{3} \pi r^3 \]
\[ S = \frac{1}{2} Ns_o = 4 \pi r^2 \ (\text{factor of } \frac{1}{2} \text{ because only half of the residues are on surface}) \]

Now we have 2 equations in two unknowns, N and r, and all we have to do to find N is to eliminate the other unknown – r:

\[ (3Nv_o/4\pi)^2 = r^6 \]
\[ (Ns_o/8\pi)^3 = r^6 \]
\[ N \sim 288 \pi v_o^2 / s_o^3 \]

Does this simplistic estimate make any sense? Let’s try some numbers.

\[ s_o = 50 \text{ Å}^2 \text{ (given)} \]

How can we estimate \( v_o \)? Easy! We know typical values for the density of protein (a little more dense than water, right? Let’s make up a reasonable number):

\[ d \sim 1.3 \text{ gm/cm}^3 \]

Since the average molecular weight of an amino acid is \( \sim 100 \text{ gm/mol} \), we can write:

- volume / mole of residues = \( 1 \text{ cm}^3 / 0.013 \text{ mole} \sim 80 \text{ cm}^3 / \text{mole} \)
- So all we have to do is divide by Avogadro’s number and change to Å\(^3\), to get:

\[ v_o \sim 130 \text{ Å}^3 \text{ per residue} \]

So \( N \sim 120 \)

Does this work? Find in the database a small, roughly spherical protein, Photoactive yellow protein (PYP) – 120 residues. Spacefill the image in RASMOL, and then mark all the beta-carbons with a funny color. How many beta-carbons are visible, hence, surface-exposed? (This would be a bad way to do it if there were lots of glycines in the protein, but there are not too many, so that won’t make much error.) It turns out to be about 66, or 55% of the total. This is a pretty good estimate for such a moronically simple approach.