Problem 1

2-3 What is the magnitude of the Coulombic interaction energy between an oxygen atom and a hydrogen atom participating in a hydrogen bond in water, assuming that the partial charge on the oxygen is – 0.8e, and the partial charge on the hydrogen is + 0.4e? What is the interaction energy of this hydrogen bond in the interior of a protein? (Inspired by Tinoco et al. [75].)

Water

$$U = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{\epsilon_r R} = \frac{1}{4\pi\epsilon_0}\frac{-0.8(e)\cdot 0.4(e)}{80\cdot 1.8\times 10^{-10}(m)} =$$

$$\frac{8.99\times 10^9 Nm^2}{C^2}\frac{(-0.8\cdot 1.602\times 10^{-19}C)(0.4\cdot 1.602\times 10^{-19}C)}{80\cdot 1.8\times 10^{-10}(m)}\frac{1kcal}{4186.8Nm}(6.022\times 10^{23} charges/mol) = -0.737 kcal/mol$$

$\therefore$ attractive since magnitude is negative

Protein

$$U = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{\epsilon_r R} = \frac{1}{4\pi\epsilon_0}\frac{-0.8(e)\cdot 0.4(e)}{4\cdot 1.8\times 10^{-10}(m)} =$$

$$\frac{8.99\times 10^9 Nm^2}{C^2}\frac{(-0.8\cdot 1.602\times 10^{-19}C)(0.4\cdot 1.602\times 10^{-19}C)}{4\cdot 1.8\times 10^{-10}(m)}\frac{1kcal}{4186.8Nm}(6.022\times 10^{23} charges/mol) = -10.208 kcal/mol$$

Problem 2

2-8 A protein and a ligand mixed in a solution can bind to form a complex. After waiting a sufficiently long time for equilibrium to be reached, the concentrations of the protein, the ligand, and the complex in the 1 mL volume are measured to be 0.3 μM, 0.8 μM, and 6.5 μM, respectively. The association rate constant is 3 × 10^6 M–1s–1. (Contributed by L. Jeng.)

a. What is the rate of complex formation?

$k_{on}[P][L] = (3 \times 10^6)(0.3 \times 10^{-6})(0.8 \times 10^{-6}) = 7.2 \times 10^-7 M/s$

b. What is the rate of complex dissociation?

$k_{off}[PL] = k_{on}[P][L] \implies k_{off}[PL] = 7.2 \times 10^-7 M/s$

c. What is the dissociation rate constant?

$k_{off}[PL] = 7.2 \times 10^-7 M/s \implies k_{off} = \frac{=7.2 \times 10^-7 M/s}{6.5 \times 10^{-6}} = 0.111 s^{-1}$

d. What is the equilibrium dissociation constant?

$K_d \equiv \frac{k_{off}}{k_{[}on]} = \frac{0.111 s^{-1}}{3 \times 10^6 m^{-1} s^{-1}} = 3.7 \times 10^{-8} M$

e. In a negligible additional volume, 100 picomoles of ligand are spiked into the mixture. What will be the new equilibrium complex concentration?

$[L]_{spike} = \frac{100 \times 10^{-12}}{1 \times 10^- 3} = 100 \times 10^{-7} = 0.1 \mu M$

$[L]_{total} = 0.1 \mu M + 0.8 \mu M = 0.9 \mu M$

$[PL]_{eq} = \frac{[P]_{total}[L]_{free}}{K_D + [L]_{free}} = \frac{6.8 \cdot 0.9}{0.037 + 0.9} = 6.53 \mu M$

Problem 3

Gabe Rocklin's group at Northwestern published a paper in Nature in 2023 (Here; 'Rocklin_Nature2023.pdf' on Canvas), in which they developed an assay for the high throughput determination of protein stability. They were able to identify the standard Gibbs free energy of folding at 22^o C by this equation (this is equation (5) shown in Figure 1): $\Delta G^o_{fold} = RTln\frac{([F])}{([U])}$

a. Describe, in your own words, the experimental protocol the Rocklin group used to identify the ΔG^o_{unf} for different proteins.

A. Create DNA library encoding protein variants. B. Digest protein with protease at different concentrations. C. Determine cleaved and intact proteins from cDNA tag. D. Model kinetics for (un)folded $\forall$ proteins to determine kinetic constants in each

instance. E. Plug in necessary information into their model to determine $\Delta G^0_{unf} \& \Delta G^0_{fold}$

Folded proteins resist proteolysis. Unfolded proteins get cleaved. The protease concentration at half-maximal cleavage (K_{50}) reflects stability. They model the equilibrium between F and U states, where each has different protease susceptibility (K_{50},F and K_{50},U), to back-calculate Δ G.

b. In the paper, they claim that the dynamic range for measuring the standard Gibbs free energy of folding is 0-5 kcal/mol. To understand the upper limit of this range, convert this range to folding percentages. Use this conversion to explain the upper limit of the dynamic range.

$$5kcal/mol = RTln(\frac{[F]}{[U]}) = 1.987 \times 10^{-3}(298)ln(\frac{[F]}{[U]}) \implies 4648[U] = [F]$$

$$Since [F] + [U] = [S_{tot}] \implies 4648[U] + [U] = [S_{tot}] = 4649[U] \implies$$

$$\frac{[U]}{[S_{tot}]} = 2 \times 10^{-4}$$

$$0kcal/mol = RTln(\frac{[F]}{[U]}) \implies 1 = \frac{[F]}{[U]} \implies [U] = [F]$$

$$\implies 2[U] = [S_{tot}] \implies \frac{[U]}{[S_{tot}]} = 0.5$$

ΔG	% Folded
0	50%
5	99.9%.

At Δ G > 5 kcal/mol, proteins are >99.98% folded. Since proteolysis only occurs from the unfolded state, there's insufficient unfolded population to measure differences in K_{50} between very stable variants, they all resist proteolysis equally.

c. In Figure 3a, they show environmental factors that contribute to certain amino acid substitution tendencies. Using your own words, describe what this plot shows. Then, explain this plot using some of the energetic contributions to proteins discussed in Lecture 4.

Figure 3a shows a PCA of amino acid substitution effects on protein stability across many sites, highlighting which residue properties dominate stability in different environments (PC1–PC5 explain ~71% variance).

- PC1 (31%): Hydrophobic effect / solvation — favors nonpolar residues in buried sites
- PC2 (15%): Secondary structure preferences (e.g., helix constraints)
- PC3 (12%): Van der Waals packing / sterics — aromatic vs. aliphatic residues
- PC4 (7%): Electrostatics — charged residues and salt bridge formation
- PC5 (6%): Size and burial constraints — small residues in tightly packed regions

These components correspond to energetic contributions discussed in Lecture 4: hydrophobic stabilization, packing efficiency, electrostatic interactions, and structural constraints.