Biophysical Chemistry – CH 4403 01
Assignment 6
Due Friday, October 16 at 4:30 pm

Please complete the answers to this assignment on a separate page (or pages), showing your work and sources (if you referred elsewhere for constants, enthalpies, etc.).

1. In class we related the folding free energy to the fraction of protein that is folded.

$$N \leftrightarrow U, \text{ where } f = \frac{1}{1+K} \qquad \text{(eq. 1)}$$

We also discussed in class several weeks ago how $\Delta\bar{G}^0$ (and thus *K*) for chemical reactions depends on temperature. To a first approximation, $\Delta\bar{G}^0$ vs. temperature can be determined using the following equation if enthalpy and entropy are known at a reference temperature:

$$\Delta\bar{G}^0(T) = \Delta\bar{H}^0(T_{ref}) - T\Delta\bar{S}^0(T_{ref}) \qquad \text{(eq. 2)}$$

Alternatively, the Van't Hoff Equation can be used if only the enthalpy is known Both of these expressions assume enthalpy and entropy are constant over a range of T. In Assignment 4, we examined the consequences if that assumption didn't hold, but for this assignment, we will assume that $\Delta\bar{H}^0$ and $\Delta\bar{S}^0$ are constant *vs.* T (and eq. 2 applies).

(a) Derive an expression for the fraction of folded protein *vs.* temperature given $\Delta\bar{G}^0(298\ K) = 5$ kcal mol^{-1} and $\Delta\bar{H}^0 = 168$ kcal mol^{-1}. You will need to express eq. 1 in terms of the variables in eq. 2. (4 points)

(b) Using Excel (or any other plotting program) plot your expression for f from (a) vs. temperature for temperatures from 20 ^{o}C to 90 ^{o}C in increments of 2 ^{o}C (remember that T in eq. 1 is in K, not ^{o}C!). On your plot, indicate the position where $\Delta\bar{G}^0$ is zero. Where is $\Delta\bar{G}$ zero? (4 points)

(c) Thermophiles are organisms that live in very hot environments, such as volcanic vents under the ocean or in hot springs. When examined, many of the proteins in thermophilic organisms have similar stabilities at room temperature to those in their "normal" mesophilic counterparts. Assume your results from (a) and (b) are from a mesophilic bacteria that that thrives under normal temperatures. If $\Delta\bar{G}^0(298\ K)$ is the same for a thermophilic organism, what must be true about $\Delta\bar{H}^0$ if such an organism is to thrive at much higher temperatures? (2 points)

2. It is not always possible to do equilibrium dialysis on a system. Sometimes measuring the concentrations is not straightforward, or the binding may be sufficiently weak or tight that it may not be detectable. In this problem, we will consider the case where one can still measure $\bar{\nu}$, but where one only knows $[L]_{total}$ and cannot measure [L] directly. The reaction of interest is single-site binding:

$$P \cdot L \leftrightarrow P + L$$

(a) Given K and initial values of $[P]_{total}$ and $[L]_{total}$, set up an equilibrium table to determine general expressions for the concentrations of [PL], [P], and [L] at equilibrium. You will have to solve a quadratic equation. (4 points)

(b) If $[P]_{total}$ is 5 μM and K is 5×10^{-7}, make a plot of $\bar{\nu}$ vs. x, where $x = [L]_{\text{total}}$. Your values for x should range from 0.0 μM to 10 μM in increments of 0.1 μM. For comparison, on the same plot, plot $\bar{\nu}$ vs. x, where $x = [L]$. Recall that, for dissociation, $\bar{\nu} = \frac{[L]}{K+[L]}$. (3 points)

(c) Repeat part (b) for $[P]_{total}$ of 0.1 μM. What happens to the difference between the $[L]_{total}$ vs. [L] curves? If you could not use equilibrium dialysis, would you want your $[P]_{total}$ to be greater than, less than, or approximately equal to the K? Why? Note that for this problem K can be thought of as $5 \times 10^{-7}\text{M} = 0.5$ μM. (3 points)

3. A transcription factor is a protein that binds to a short section of DNA and regulates gene expression at a nearby site. It is estimated that approximately 13% of commercial drugs target a specific class of transcription factor, nuclear receptors (Overtington, *et al*. *Nat. Rev. Drug Disc.* **5**: 993). You can take a look at a structure of a nuclear receptor transcription factor in the PDB by looking up 1CIT. In this problem, we will examine how cooperativity can modulate gene expression in transcription factors.

Let ${}^{o}\text{DNA}^{o}$ represent the DNA with no transcription factor bound at either site A or B. Then, ${}^{P_A}\text{DNA}^{o}$ and ${}^{o}\text{DNA}^{P_B}$ represent the DNA when sites A and B are occupied. When site B is occupied, transcription can occur. Both sites A and B bind to transcription factors P_A and P_B with their respective intrinsic affinities. However, when both P_A and P_B are bound, a cooperative interaction provides additional stabilizing energy (cooperativity), represented by the factor τ. The binding reactions at each site are:

$$P_A + {}^{o}\text{DNA}^{o} \rightleftarrows {}^{P_A}\text{DNA}^{o}, K_{eq} = K_A$$
$$P_B + {}^{o}\text{DNA}^{o} \rightleftarrows {}^{o}\text{DNA}^{P_B}, K_{eq} = K_B$$
$$P_B + {}^{P_A}\text{DNA}^{o} \rightleftarrows {}^{P_A}\text{DNA}^{P_B}, K_{eq} = \tau K_B$$
$$P_A + {}^{o}\text{DNA}^{P_B} \rightleftarrows {}^{P_A}\text{DNA}^{P_A}, K_{eq} = \tau K_A$$

(a) Show that this system satisfies path independence. In other words, show that the total ΔG is the same whether you bind P_A first, then P_B, or whether you bind P_B, then P_A. What is the energy associated with the cooperativity factor τ, or the energy you gain when τ is present (> 1)? *Hint:* This "extra" energy should be zero when there is no cooperativity ($\tau = 1$). (2 points)

(b) Using the methods outlined in class, write an expression for the degree of binding to the DNA for P_A ($\bar{\nu}_A$) and P_B ($\bar{\nu}_B$). Your expression will depend on P_A, P_B, τ, K_A, and K_B. Then, write expressions for the concentration of $[\,{}^{o}\text{DNA}^{o}\,]$, $[\,{}^{P_A}\text{DNA}^{o}\,]$, $[\,{}^{o}\text{DNA}^{P_B}]$ and

[$^{P_A}DNA^{P_B}$] in terms of K_A, K_B, τ, the free protein concentrations [P_A] and [P_B], and the total DNA concentration [DNA_{tot}]. (5 points)

(c) Suppose that $K_A = 1 \times 10^6$, and $K_B = 8 \times 10^6$, and $\tau = 100$. Remember that, in our model, any time site B is occupied, transcription will occur. On a single graph, plot the transcription level as a function of [P_B] for two conditions: 0 μM [P_A] and 0.1 μM [P_A]. To make your plot smooth, plot [P_B] from 0 to 200 nM in increments of 1 nM. (3 points)

(d) Given your plot from (c), why do many transcription factors exhibit cooperativity? (2 points)

(e) While we have ignored the potential effects of drug binding in this problem, our model could be extended to include drug interactions with either P_A or P_B. If you were designing a drug to inhibit transcription in the system above, which protein do you think would be a better drug target and why? *Hint:* What would happen if we could reduce either K_A or K_B by 10%? (2 points)

For more information on modeling transcription regulation, you may wish to check out He, X., *et al*. *PLoS Comp. Biol.* **6**: e1000935. The seminal paper on these ideas was written by Madeline Shea and Gary Ackers (*J. Mol. Biol.* **181**: 211) and is a must-read for any budding biophysicist.

4. A scientist is using equilibrium dialysis to measure zinc binding to a protein. Using atomic absorption spectroscopy, the scientist is able to measure the total concentration of zinc at equilibrium inside and outside of the bag:

Data point	Protein (μM)	Zinc Inside (μM)	Zinc Outside (μM)
1	20	20.1	5.1
2	20	74.8	29.3
3	20	125.9	65.6
4	20	173.9	110.5
5	20	200.2	131.9

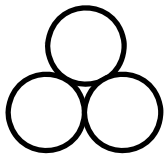
(a) Using the data above, determine the equilibrium constant for binding as well as the number of total binding sites, assuming each site is equivalent and independent. Submit a copy of your Scatchard plot. (4 points)

(b) The above data were collected at 25 ^{0}C; the scientist repeated the experiment at 15 ^{0}C and got the following results:

Data point	Protein (μM)	Zinc Inside (μM)	Zinc Outside (μM)
1	20	20.0	2.9
2	20	75.5	22.2
3	20	124.8	58.1
4	20	171.7	101.8
5	20	199.2	126.5

Estimate $\Delta\bar{H}^0$ and $\Delta\bar{S}^0$ of binding assuming that both are constant as a function of temperature. (6 points)

5. A homotrimeric protein is a multimeric protein consisting of three copies of an identical protein subunit. For many proteins, it is reasonable that the homotrimer will be symmetric, with each subunit arranged around a C_3 symmetry axis:



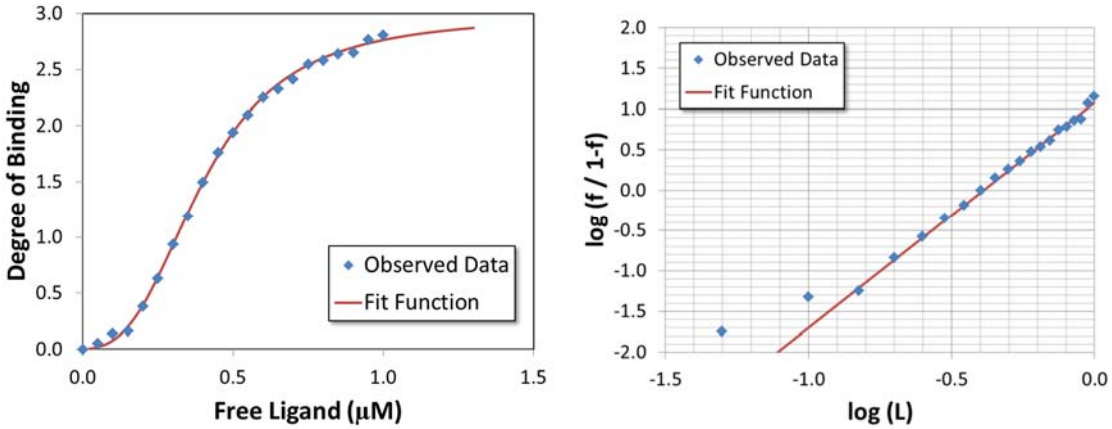
An example of a homotrimeric protein in biology is Tumor Necrosis Factor α (TNF-α). TNF-α is a protein involved in immune response, and it can trigger cell death in response to catastrophic events (e.g. cancer and bacterial infection). It can also trigger inflammation, and thus there are times when it is useful to interfere with TNF-α. Using drugs that bind TNF-α, it is possible to prevent it from binding to receptors on the cell surface.

A reasonable model for TNF-α binding is a nearest-neighbor binding model, with an allosteric term τ and a single (identical) site binding constant K. Unlike the linear chain model we discussed in class, every subunit has two nearest neighbors. Rather than a linear chain, this binding can be envisioned as a circular loop.

(a) There are a total of eight states for binding to this system. List these states and give the statistical weight for each in terms of K and τ, and [L]. (8 points)

(b) Write an expression $\bar{\nu}$, the degree of binding as a function of τ, K, and [L]. (4 points)

(c) You have developed a small-molecule that binds to TNF-α. You perform a series of binding experiments to determine $\bar{\nu}$ vs. [L]. Two representations of the data are below. Using the model from part (a) and (b), you optimize the parameters and find that, for the best fit, K is 8.3×10^4 and τ is 30.



Estimate the Hill coefficient from this data and comment how it compares to the maximally cooperative value (all-or-none) expected for a three-site system. Do you think the model (the "Fit Function" in the left-hand plot) is a good one? (3 points)