MCB137L/237L: Physical Biology of the Cell
Spring 2022
Homework 10
(Due 4/14/22 at 11:00am)

Hernan G. Garcia

NOTE FOR MCB237L STUDENTS: You don't need to do problem (1). Please do problem (2) and then choose to do (3) or (4). You can do more problems for extra credit (15% of total possible score per extra problem).

1. Dimoglobin: A Toy Model of Hemoglobin

In Homework 9, you derived the probability of a receptor being bound by a ligand using a lattice model from the statistical mechanics perspective. This resulted in

$$p_{bound} = \frac{\frac{L}{\Omega} e^{-\beta \Delta\varepsilon}}{1 + \frac{L}{\Omega} e^{-\beta \Delta\varepsilon}}, \tag{1}$$

where L is the number of ligands in the solution and $\Delta\varepsilon = \varepsilon_b - \varepsilon_{sol}$ with ε_b being the binding energy of a ligand to the receptor and ε_{sol} the energy of a ligand when in the lattice. Further, Ω is the number of lattice sites.

(a) Write p_{bound} in terms of the concentration of ligands $[L] = \frac{L}{\Omega v}$, where v is the volume of a lattice box. Now, note that we can think of the inverse of v as a concentration c_0 corresponding to each lattice site being occupied by a ligand such that $v = 1/c_0$. If the volume of a lattice site is 1 nm^3, what is the corresponding c_0? In biochemistry this c_0 is called the concentration of the standard state. How does this concentration compare to those you'd usually pipette in an experiment? What do you conclude about how dilute the solutions you usually deal with in the lab are?

In class, we discussed how cooperativity in oxygen binding to hemoglobin makes it possible for the binding curve to be switch-like. Now that we are experts at ligand-receptor binding, we want to mathematically explore the consequences of cooperativity in the context of a toy model of hemoglobin: dimoglobin. Unlike hemoglobin, which binds four oxygen molecules, dimoglobin binds only to two oxygen molecules.

Figure 1 features a lattice model of dimoglobin. Here, oxygen molecules in solution have an energy ε_{sol}, oxygen binds to either dimoglobin site with energy ε_b. Finally, when two oxygen

molecules are bound, they also interact with energy ε_{int}.

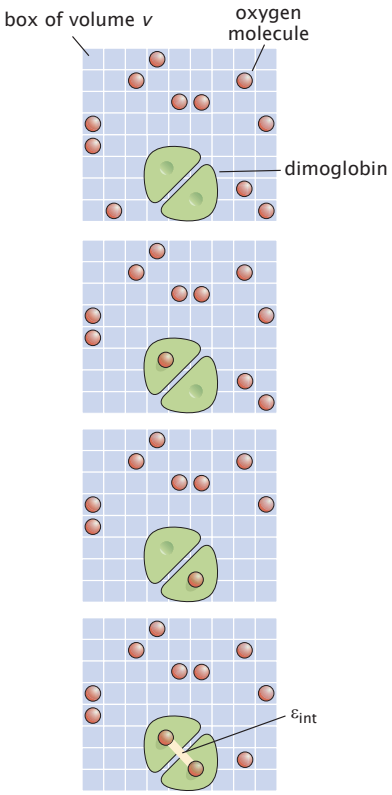

Figure 1: Cooperativity model of dimoglobin in a lattice. Different states the dimoglobin molecule and the oxygen molecules in the lattice can be found in. An oxygen molecule in solution has energy ε_{sol}, while it has a binding energy to dimoglobin of ε_b. Two oxygen molecules bound to dimoglobin interact with an energy ε_{int}.

(b) Use the statistical mechanics protocol to calculate p_0, p_1 and p_2, the probabilities of having no, one, or two oxygen molecules bound to dimoglobin. Use these probabilities to show that the average fraction of molecules bound is given by

$$\langle N \rangle = \frac{2\frac{[L]}{c_0}e^{-\beta\Delta\varepsilon} + 2\left(\frac{[L]}{c_0}\right)^2 e^{-\beta(2\Delta\varepsilon+\varepsilon_{int})}}{1 + 2\frac{[L]}{c_0}e^{-\beta\Delta\varepsilon} + \left(\frac{[L]}{c_0}\right)^2 e^{-\beta(2\Delta\varepsilon+\varepsilon_{int})}}, \tag{2}$$

where $[L]$ is the oxygen partial pressure (which is a measure of concentration) and $c_0 = 760$ mmHg is the standard state partial pressure. Make sure to include and explain all steps in your derivation.

(c) Plot the average number of bound molecules as a function of oxygen partial pressure for $\varepsilon_{int} = -5\ K_B T$ and for $\varepsilon_{int} = 0$ on a linear-log plot in order to show the effect of ε_{int} on the sharpness of the occupancy curve. Use $\Delta\varepsilon = -5\ K_B T$ for both curves.

(d) Plot p_0, p_1 and p_2 as a function of oxygen partial pressure. Make one plot for $\varepsilon_{int} = -5\ K_BT$ and one for $\varepsilon_{int} = 0$ in order to show sharpness is achieved through ε_{int} by draining probability from p_1.

2. A minimal genetic switch

In class, we introduced how genetic switches can be constructed using two repressors that repress each other's gene expression. Here, we consider a simpler regulatory architecture that can also result in a genetic switch. Specifically, we will model the self-activation of an activator molecule. For this problem, you might find it useful, to review the phase portrait concept we described in class for the case of mRNA production and degradation, as well as the phase diagram of the logistic equation you had to draw earlier on in the semester.

Figure 2(A) presents a regulatory architecture, where an activator activates its own production. Figure 2(B) shows the states and weights for our model. In this problem, we will ignore mRNA and associate the rate indicated in the figure with the rate of protein production. Here, a promoter has two activator binding sites in its vicinity. In the absence of activators, or in the presence of only one activator, the rate of protein production is r_0. When both activators are bound, the rate is r. The activators bind with a dissociation constant K_d and interact with a cooperativity factor $\omega = e^{-\beta\varepsilon_{int}}$, where ε_{int} is the interaction energy between activators. A is the concentration of activator.

(a) Write down an equation describing the temporal evolution of the number of activators. Consider the rate of production stemming from the model shown in Figure 2, as well as a rate of protein degradation γ. Hint: Remember that the overall rate of production $\langle r \rangle$ of a system is given by

$$\langle r \rangle = \sum_i p_i r_i, \tag{3}$$

where p_i is the probability of the system being in state i, and r_i is the rate of production when the system is in that state.

(b) Plot the phase diagram for this equation in order to find how many equilibria the system can support. Namely, plot the rates of production and degradation as a function of the activator concentration. Use $K_d = 5$ nM, $\gamma = 0.1$/min, $r_0 = 0.01$ nM/min, and $r = 0.5$ nM/min. Make plots for $\omega = 1$ and $\omega = 10$.

(c) Draw vectors indicating the direction of the concentration change under your plots as we did in class for the mRNA production and degradation case, and as you explored in the homework in the context of the logistic equation. How many equilibrium points do you find? Indicate whether these points correspond to stable or unstable equilibria. You can review the concept of phase portraits by reading "Computational Exploration: Growth Curves and the Logistic Equation" on page 103 of PBoC2, paying special attention to vectors drawn on the lower part of Figure 3.10.

(d) Solve the equation you derived in (a) for different initial conditions, and plot all of

them on the same graph. Choose initial conditions that help illustrate how the system can converge to different levels of activator in steady state.

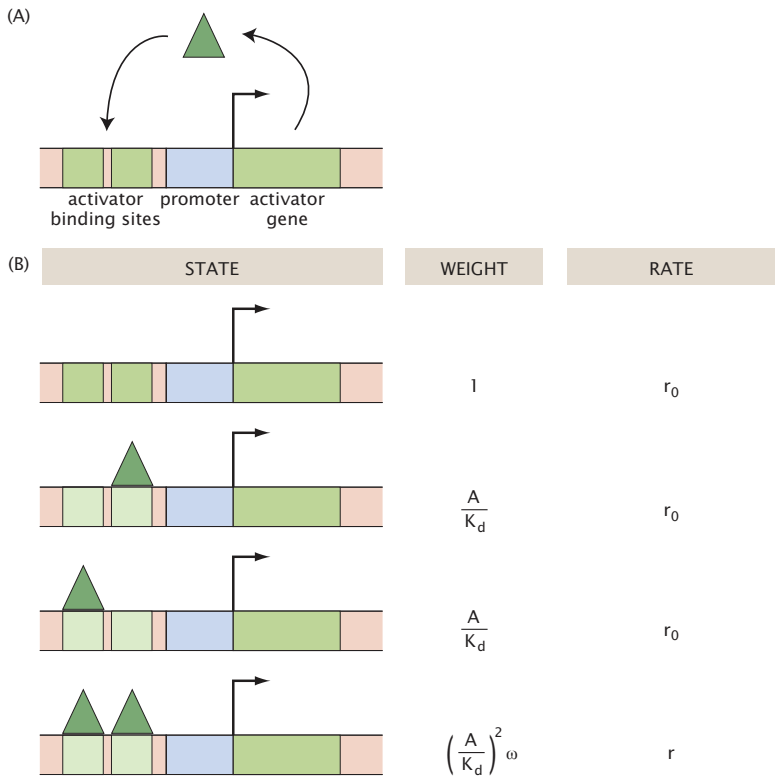

Figure 2: A simple autoactivation switch model. (A) Cartoon of the autoactivation switch. (B) States, weights and rates for the autoactivating genetic switch model.

3. Mutation correlation and physical proximity on the gene

(a) Read the section "Flies and the Rise of Modern Genetics" starting on page 170 of PBoC2.

(b) Do problem 4.4 from PBoC2.

• 4.4 Mutation correlation and physical proximity on the gene

In Section 4.6.1, we briefly described Sturtevant's analysis of mutant flies that culminated in the generation of the first chromosome map. In Table 4.2, we show the crossover data associated with the different mutations that he used to draw the map. A crossover refers to a chromosomal rearrangement in which parts of two chromosomes exchange DNA. An illustration of the process is shown in Figure 4.26. The six factors looked at by Sturtevant are B, C, O, P, R, and M. Flies recessive in B, the black factor, have a yellow body color. Factors C and O are completely linked, they always go together and flies recessive in both of these factors have white eyes. A fly recessive in factor P has vermilion eyes instead of the ordinary red eyes. Finally, flies recessive in R have rudimentary wings and those recessive in M have miniature wings. For example, the fraction of flies that presented a crossover of the B and P factors is denoted as BP. Assume that the frequency of recombination is proportional to the distance between loci on the chromosome.

Reproduce Sturtevant's conclusions by drawing your own map using the first seven data points from Table 4.2.

Keep in mind that shorter "distances" are more reliable than longer ones because the latter are more prone to double crossings. Are distances additive? For example, can you predict the distance between B and P from looking at the distances B(C,O) and (C,O)P? What is the interpretation of the two last data points from Table 4.2?

Table 4.2: Fraction of crossovers of six sex-linked factors in *Drosophila*. (Adapted from A. H. Sturtevant, *J. Exp. Zool.* 14:43, 1913.)

Factors	Fraction of crossovers
BR	115/324
B(C,O)	214/21736
(C,O)P	471/1584
(C,O)R	2062/6116
(C,O)M	406/898
PR	17/573
PM	109/458
BP	1464/4551
BM	260/693

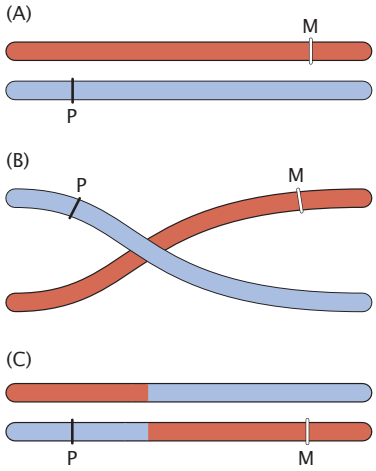

Figure 4.26: Crossing over of chromosomes. (A) Chromosomes before crossing over showing two loci labeled P and M. (B) Illustration of the crossing-over event. (C) Chromosomes after crossover.

Figure 3: Problem 4.4 from PBoC.

4. Bacterial foraging

Bacteria use swimming to seek out food. Imagine that the bacterium is in a region of low food concentration. For the bacterium to profit from swimming to a region with more food, it has to reach there before diffusion of food molecules makes the concentrations in the two regions the same. Here we find the smallest distance that a bacterium needs to swim so it can outrun diffusion.

(a) Make a plot in which you sketch the distance traveled by a bacterium swimming at a constant velocity v as a function of time t, and the distance over which a food molecule will diffuse in that same time. Indicate on the plot the smallest time and the smallest distance that the bacterium needs to swim to outrun diffusion. You don't need to use Python here, just make the plot by hand and show the two curves schematically.

(b) Calculate these minimum times and distances for an *E. coli* swimming at a speed of $30\,\mu$m/s. The diffusion constant of a typical food molecule is roughly $500\,\mu\text{m}^2/\text{s}$.

(c) Estimate the number of ATP molecules the bacterium must consume (hydrolyze) per second in order to travel at this speed, assuming that all of the energy usage goes into overcoming fluid drag. The drag force felt by the bacterium is given by

$$F = 6\pi\eta R v, \tag{4}$$

where R is the typical size of an *E. coli*, η is the viscosity (we can assume it's swimming in water) and v is the speed of the bacterium. The power necessary to move the bacterium at a speed v against this viscous drag is

$$P = Fv. \tag{5}$$

The amount of energy released from one ATP molecule is approximately $20\,k_{\text{B}}T$. Note that the bacterial flagellar motor is actually powered by a proton gradient and this estimate focuses on the ATP equivalents associated with overcoming fluid drag.