MCB137L/237L: Physical Biology of the Cell
Spring 2025
Homework 12
(Due 4/22/25 at 2:00pm)

Hernan G. Garcia

"How can the events in *space and time* which take place within the spatial boundary of a living organism be accounted for by physics and chemistry?" - Erwin Schrödinger **What is Life?**

1 Dimoglobin: A Toy Model of Hemoglobin

In the previous problem, 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}.

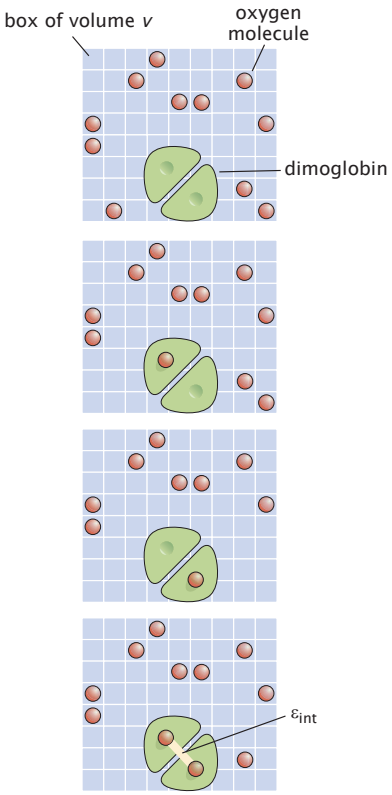

Figure 1: Cooperativity model of dimoglobin in a lattice. Different states the dimoglobin molecule and the oxygen molecules in the lattice can be found in. An oxygen molecule in solution has energy ε_{sol}, while it has a binding energy to dimoglobin of ε_b. Two oxygen molecules bound to dimoglobin interact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

bf (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 number of bound molecules 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 $\varepsilon_{int} = -5\ K_BT$ 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 $\Delta\varepsilon = -5\ K_BT$ 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 Simple repression

In class, we derived a mathematical model of how simple repression dictates gene expression levels. In particular, we showed that the fold-change in gene expression is given by

$$\text{fold-change} = \frac{1}{1 + \frac{R}{N_{NS}} e^{-\beta\Delta\varepsilon}}, \tag{3}$$

where R is the number of repressors, N_{NS} is the number of non-specific sites, and $\Delta\varepsilon$ is the repressor binding energy. Experimentally, we can create bacterial strains that express a YFP reporter under the control of the repressor. If the YFP fluorescence signal is $F_{reporter}$, then the fold-change is measured by calculating

$$\text{fold-change} = \frac{F_{reporter}(R)}{F_{reporter}(R=0)}. \tag{4}$$

However, there is an extra subtlety that has to be taken into account when obtaining such fluorescence measurements. The fluorescence we measure under the microscope is the total fluorescence F_{total}. This magnitude contains the YFP reporter signal as well as the intrinsic fluorescence of the cells such that

$$F_{total} = F_{reporter} + F_{auto}, \tag{5}$$

where F_{auto} is the autofluorescence of the cell. As a result, we need to be able to subtract the cells' average autofluorescence if we want to report on $F_{reporter}$. Thus, the fold-change is obtained by calculating

$$\text{fold-change} = \frac{F_{total}(R) - F_{auto}}{F_{total}(R=0) - F_{auto}}. \tag{6}$$

During discussions, we already learned how to extract fluorescence levels from microscopy images of bacteria using Python (see Colab notebook provided on course website). In this problem, you will use the code you wrote in class in order to test the prediction made by Equation 3. You can get our final version of that code as well as the full bacterial gene expression data set from website. Note that you can also learn more about this protocol from "Computational Exploration: Extracting Level of Gene Expression from Microscopy Images" in chapter 19 of PBoC2.

(a) You can find `lacI_titration` folder under "Data Folder" shared on the course website. This data set corresponds to bacteria containing varying Lac repressor copy numbers, and a reporter containing one of three Lac repressor binding sites (called O1, O2 and O3) that control YFP expression. The repressor copy numbers for each strain are given in the following table.

Strain	Repressor number
Delta	0
R22	22 ± 4
R60	60 ± 20
R124	124 ± 30
R260	260 ± 40
R1220	1220 ± 160

If you want to learn more about the measurements of the repressor copy number, you can refer to Garcia2011c, which is provided on the website. In addition, a strain called "Auto" lacking a YFP reporter is provided with the data set. This strain can be used to measure F_{auto}. Another strain called "Delta" can be used to measure $F_{total}(R = 0)$.

The name of the files can be interpreted as follows:

- Lac repressor binding site: O1/O2/O3
- Repressor copy number: Delta, R22, R60, R124, etc.
- Type of image: phase/yfp corresponding to the phase contrast and fluorescence snapshots of the cells, respectively.

For example, a file named "O2_R22_yfp_pos_01.tif" is a YFP image of a reporter containing the O2 Lac repressor binding site, and repressor copy number 22 ± 4 (see the table above). Finally, each file name ends with "pos_XX.tif". This suffix corresponds to the position on the microscope slide that each snapshot was taken at. For each combination of Lac repressor binding site and number of Lac repressor molecules you will have 21 unique snapshots whose analysis you will combine to obtain the corresponding average F_{total}.

Use your code to calculate the fold-change in gene expression as function of repressor copy number by performing the various fluorescence measurements prescribed by Equation 6. Plot your results on a log-log plot.

(b) Using the data you just analyzed, estimate the binding energy of O1, O2 and O3 by fitting Equation 3 "by eye" by trying a reasonable range of parameters for $\Delta\varepsilon$. How do these compare to the binding energies obtained in Garcia2011c using an enzymatic assay instead of fluorescence as a report of gene expression? EXTRA CREDIT: Alternatively, you can perform a least-squares minimization as discussed in class. If you want to go down this route, you might find it better to fit your data to the fold-change expressed in the language of dissociation constants

$$\text{fold-change} = \frac{1}{1 + \frac{[R]}{K_d}}. \tag{7}$$

Here, $[R]$ is the concentration of repressors inside the cell, and K_d is the dissociation constant of repressor to the DNA.