MCB137L/237L: Physical Biology of the Cell
Spring 2024
Homework 4
(Due 2/20/24 at 2:00pm)

Hernan G. Garcia

"We are not students of some subject matter, but students of problems. And problems may cut right across the boundaries of any subject matter or discipline" - Karl Popper

1 Protein Sequences: The Frances Arnold Estimate Problem

In a 2001 Bioengineering seminar at Caltech, Professor Frances Arnold made a startling remark that it is the aim of the present problem to examine. The basic point is to try and generate some intuition for the **HUGE**, **ASTRONOMICAL** number of ways of choosing amino acid sequences. To drive home the point, she noted that if we consider a protein with 300 amino acids, there will be a huge number of different possible sequences.

(a) How many different sequences are there for a 300 amino acid protein?

But that wasn't the provocative remark. The provocative remark was that if we took only one molecule of each of these different possible proteins, it would take a volume equal to five of our universes to contain all of these different *distinct* molecules.

(b) Estimate the size of a protein with 300 amino acids. Justify your result, but remember it is an estimate. Next, find an estimate of the size of the universe and figure out whether Frances was guilty of hyperbole or if her statement was on the money.

2 Solving Ligand-Receptor Multiple Ways

In class we solved for the dynamics of mRNA production and degradation using the dynamics protocol. In this problem, we are going to use that analysis as a jumping off point

for thinking about one of the most ubiquitous problems in biology: ligand-receptor binding. This ligand-receptor binding problem is a paradigm for a broad swath of biological processes ranging from neuroscience, to physiology, to gene regulation.

(a) Imagine a situation in which we have a receptor fixed at some point in space as shown in the top right panel of Figure 1. Write a rate equation for the concentration of ligand-receptor pairs in terms of the concentration of ligands and receptors. Now, assume steady state and, given that equation, derive an expression for the dissociation constant

$$K_d = \frac{[L][R]}{[LR]} \tag{1}$$

in terms of the on and off rates. Make sure you explain the dimensions of your on and off rates and hence, the dimensions of K_d.

(b) A second route to considering ligand-receptor interactions is to think of binding probabilistically with the probability that the receptor is occupied given by

$$p_{bound} = \frac{[LR]}{[R] + [LR]}. \tag{2}$$

Given the definition of the dissociation constant introduced in the previous part of the problem, find a simple expression for $p_{bound}([L])$ that is only a function of the concentration of ligand. (NOTE: for now, we are ignoring the subtlety that the amount of total ligand and free ligand are not actually the same, though in the case considered here with a single receptor we have somewhat finessed that point.) Make a plot of $p_{bound}([L])$ as a function of $[L]$ and comment on where K_d belongs on the axes. Later on in the course, we will solve this problem in yet another way, by using statistical mechanics.

3 Dynamics of the constitutive promoter

In class, we determined that the rate of mRNA decay γ, and not the production rate r, dictates the time it takes for the mean mRNA number to reach its steady state value. Here, we further explore this conclusion that could be at odds with our initial expectations about the dynamics of the constitutive promoter.

(a) If r does not dictate the time to reach steady state, what aspect of the promoter dynamics does it determine? Solve for the mRNA concentration as a function of time for two different values of r such as 10 mRNA/min and 20 mRNA/min using an initial condition $m(t = 0) = 0$. Use $\gamma = 1$ /min. Plot mRNA number vs. time and show that r controls the initial slope.

(b) If r determines this initial slope, how come both curves take the same time to reach their steady state value? Plot the phase diagrams corresponding to both choices of model

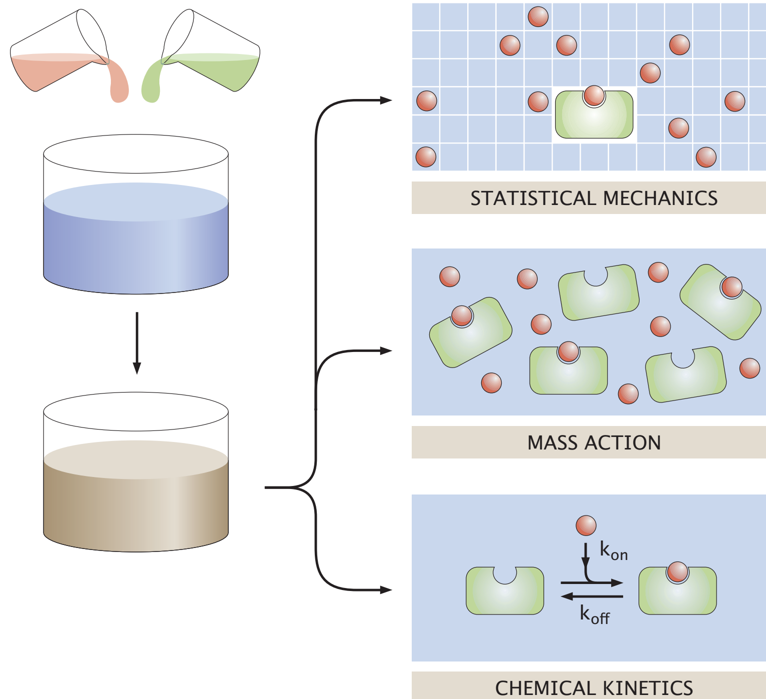

Figure 1: Three treatments of ligand-receptor binding.

parameters and show that, while $r = 20$ mRNA/min has a faster initial increase in mRNA number, its steady state value is also larger such that the time it takes to reach steady state remains unaltered.

4 Protein-mRNA Ratio

In this problem we go beyond the calculation on mRNA production we did in class, and think about how transcription and translation shape the protein-to-mRNA ratio inside cells.

(a) In class, we described the temporal evolution of the number of mRNA molecules using the equation

$$m(t + \Delta t) = m(t) + r_m \Delta t - \gamma_m m(t) \Delta t. \tag{3}$$

Here, $m(t)$ is the number of mRNA at time t, r_m is the rate of mRNA production, and γ_m is the mRNA decay rate. Write the corresponding equation for the number of protein molecules given a rate of protein production *per mRNA* of r_p and a protein decay rate γ_p. Make sure to incorporate the fact that the number of mRNA molecules present will determine how many proteins are produced in a time interval Δt.

(b) Calculate the ratio of protein to mRNA in steady state, p_{SS}/m_{SS} and show that it is given by r_p/γ_p. Find typical values for the various model parameters in *E. coli* and estimate the ratio of proteins to mRNA molecules. How do your numbers compare to those measured

in Figure 3C of Taniguchi *et al.*, which is provided on the course website?

We can also obtain this protein-mRNA ratio in the context of fruit flies.

(c) Using flies with different dosages of Bicoid-GFP, Petkova *et al.* measured the relation between the number of *bicoid* mRNA molecules deposited by the mother, and the resulting number of Bicoid proteins. Read their paper (available on the course website) and write a short paragraph about how their Figure 3 is generated.

(d) The paper by Drocco *et al.* (available on the course website) uses a photoactivatable fluorescent protein to measure the lifetime of the Bicoid protein. Read the paper (available on the course website) and explain the technique in one paragraph. You might find it useful to draw a schematic plot such as shown in Figure 1f of the paper.

(e) Assuming that Bicoid-GFP is in steady state, use what you learned about r_p and γ_p for the Bicoid protein in order to calculate its protein-mRNA ratio r_p/γ_p.