BE/APh 161: Physical Biology of the Cell, Winter 2023
Homework #1
Due 2:30 PM, January 18, 2023.

Note from JB: This problem set is largely about estimation. You should be able to do these estimates on a cocktail napkin. Please try to complete this homework without the use of a calculator or computer for calculations. You may refer to any of the material in *PBoC2*, but avoid BioNumbers or *Cell Biology by the Numbers* or any other sources unless explicitly asked to. Remember that we're just estimating, so π is about 3, etc. Be sure to comment on what parts of your estimates are most suspect and why. After you come up with your estimates, you may look them up at BioNumbers or elsewhere. And of course, try to have fun!

Problem 1.1 (Exploring biology with numbers, 10 pts)**.**
Spend some time on the BioNumbers website (`http://www.bionumbers.org/`), looking at various numbers. Pick two that are particularly interesting to you, give their BNIDs, and write a few sentences about why you find each one interesting.

Problem 1.2 (Cellular eating and drinking, 12 pts)**.**
When growing up bacteria, we typically put a small number of bacterial cells in growth media, and the components of that solution are converted into more cells. (Think about that for a second. From a clear liquid solution, life makes life. After experience this profound moment of wonder, continue reading the problem statement.) *E. coli* can divide every twenty minutes under very good growth conditions.

a) Estimate the rate at which a growing *E. coli* cell drinks with a division time of twenty minutes. By "drinking" I mean estimate the number of water molecules that need to be imported into the cell per second to maintain the level of growth.

b) The growth media typically contains glucose as a carbon source. Estimate how much carbon is necessary to build the materials of a new cell (which is done in 20 minutes under very good conditions). Based on this estimate, at what rate must the bacterium "eat" glucose? It is also interesting to consider that there are about 1000 transmembrane proteins that transport suger in an *E. coli* cell. At what rate must sugar molecules come through each of these transmembrane proteins? For this estimate, we are considering only the carbon necessary to build the cellular material; we are completely neglecting the energy needed to build the cell.

Problem 1.3 (HIV estimates, problem 2.8 of *PBoC2*, 12 pts)**.**

a) Estimate the total mass of an HIV virion by comparing its volume with that of an *E. coli* cell and assuming they have the same density.

b) The HIV maturation process involves proteolytic clipping of the Gag polyprotein so that the capsid protein CA can form the shell surrounding the RNA genome and nucleocapsid NC can complex with the RNA itself. Using Figures 2.30 and 2.31 from *PBoC2* to obtain the capsid dimensions, estimate the number of CA proteins that are used to make the capsid and compare your result with the total number of Gag proteins.

Problem 1.4 (Concentrations and spacing, 30 pts)**.**
Use your skills of estimation to answer the following questions. Comment on the numbers you come up with.

a) Many biochemical studies in test tubes use nanomolar (nM) concentrations of purified proteins. If a protein species inside of an *E. coli* cell has concentration of 1 nM, how many total molecules of that species are there in the cell?

b) What is the typical intercell distance between *E. coli* cells in a saturated LB growth medium? *Hint*: According to BNID 104943, the saturation concentration in LB is 20.5 g/L.

c) It is estimated that there are of order 10^{30} prokaryotic cells on Earth (BNID ID 104960, see also the beautiful paper by Whitman, et al.). Roughly 10% of these are in the open ocean. Give a rough estimate for the concentration of bacteria in sea water. What is the approximate intercell spacing? *For fun*: If all the bacteria in the sea were lined up end-to-end, how long would the line be?

d) There are approximately 0.2 kg of bacteria in your large intestine. What is their intercell spacing?

e) Approximately how many hydronium ions are in an *E. coli* cell?

Problem 1.5 (Regeneration cost, 12 pts)**.**
Sender and Milo (*Nat. Med., 2021*) estimated that about 300 billion cells per day turn over (meaning 300 billion are lost and 300 billion are made each day) in a human body, with about 90% of the cells turning over being red blood cells. Estimate the daily energy requirement for this task as a fraction of the total energy budget of the body. *Hint:* Reading section 5.1.2 of *PBoC2* will help with this estimate.

Problem 1.6 (Tools of the model-building trade, 24 pts)**.**
In lecture, we mathematized a cartoon of flagellar growth and length control into differential equations. When we did so, we did not have particular physical principles in mind beyond keeping track of where the tubulin monomers are. (The physical principle here is conservation of mass.) We will learn about physical laws from statistical mechanics and continuum mechanics (and perhaps more) to build models, but for now, we will proceed to name a few useful tools for building models and put them to use.

a) The **law of mass action** is a *very* widely applicable rule in model building for cell biological systems. It states that the rate at which a chemical reaction proceeds is proportional to the product of the concentrations of the reactants. For example, for the reaction A + B $\longrightarrow$ C, the rate of the reaction is $kc_\mathrm{A}c_\mathrm{B}$, where c_i is the concentration of species i and the constant of proportionality k is referred to as a rate constant. If this is the only reaction happening, we can write a pair of differential equations describing the dynamics of the concentrations of species A and B.

$$\frac{\mathrm{d}c_\mathrm{A}}{\mathrm{d}t} = \frac{\mathrm{d}c_\mathrm{B}}{\mathrm{d}t} = -kc_\mathrm{A}c_\mathrm{B}. \tag{1.1}$$

Now, consider a more complex system of reactions that we might encounter when studying enzyme kinetics. Imagine that an enzyme E can bind reversibly to substrate S. That statement describes two reactions, S + E $\longrightarrow$ ES and ES $\longrightarrow$ S + E. (We typically write these reactions in one expression as S + E $\rightleftharpoons$ ES.) Define the rate constants for these two equations respectively as k_+ and k_-. Once the enzyme binds, it can convert the substrate to product in the reaction ES $\longrightarrow$ P + E. Let the rate constant for this reaction be k_2. Use the law

of mass action to write down a system, of differential equations for the concentrations of the respective species, S, E, ES, and P.

b) Another ubiquitous tool is a **Taylor series expansion**. When we do not know how one quantity depends on another, we often approximate it with a Taylor series. For example, we know that the force exerted by a spring is a function of the displacement of the spring end of the spring. That is, $F = F(x)$, x is the displacement. We don't know what $F(x)$ is, but if we expand it in a Taylor series, we can have an *approximate* expression for $F(x)$, at least for values of x close to the value we expanded about. So, we can write

$$F(x) = F_0 + F'(x = x_0)\,(x - x_0) + \frac{1}{2}F''(x = x_0)(x - x_0)^2 + \ldots. \tag{1.2}$$

If we take x_0 to be the position where the force vanishes such that $F_0 = 0$, and we set our coordinate axes such that $x_0 = 0$, we get

$$F(x) = F'(x = 0)\,x + \frac{1}{2}F''(x = 0)x^2 + \ldots. \tag{1.3}$$

If we define $k = F'(x = 0)$ and keep terms to only linear order, we get

$$F(x) \approx k\,x, \tag{1.4}$$

which is Hooke's law. The coefficient k in the Taylor series is a phenomenological parameter, in this case the spring constant.

Now, consider bacterial growth. We expect the rate of growth of bacteria to be a function of the number of bacteria present. If we have more bacteria, they are all dividing, leading to a faster increase in the number of total bacteria. However, if there are too many bacteria, their high density results in fouling of the media, so growth will slow down. The rate of change of bacterial concentration is $\mathrm{d}c/\mathrm{d}t$, where c is the concentration of bacteria. Write down an expression for $\mathrm{d}c/\mathrm{d}t$ by expanding it as a function of c to second order. You should get two phenomenological parameters, the *growth rate*, k, which has dimension of inverse time, and the *carrying capacity*, K, which has dimension of concentration and specifies the maximum concentration of bacteria before they stop growing. (You may need to manipulate the resulting expression a bit to identify these parameters.)

c) In many ways, probability is the mathematical language of biology. We will discuss this in depth when we introduce the tools of statistical mechanics. But many processes do not involve different energies for different states, so we do not need the machinery of statistical mechanics to model them. Rather, we can employ a coin toss, or a **Bernoulli trial**. A Bernoulli trial is an experiment or event that has two outcomes that can be encoded as success and failure (or heads/tails or true/false or 1/0 or . . .). We will define the probability of success of a Bernoulli trail as p. The probably of failure is therefore $1 - p$.

As an example, imagine that I have a collection of cells that can be in one of two states. In a given environment, a cell has probability p of being in state A and probability $1 - p$ of being in state B. We can use the idea of Bernoulli trials to work out the probability $P(n_\mathrm{A}; n)$ that n_A out of a total of n cells are in state A. Here, "success" is a given cell being in state A and "failure" is a given cell being in state B.

Now, say we have a bacterium that has n plasmids in it right before dividing. Write down a probability mass function for the number of plasmids its respective daughter cells get. That is, write down an expression for $P(n_\mathrm{A})$, the probability that daughter cell A gets n_A plasmids.