

MCB137L/237L: Physical Biology of the Cell
Spring 2022
Homework 2: Scales and Scaling in Biology
(Due 2/3/22 at 11:00am)

Hernan G. Garcia

“Trying to understand molecular biology without numbers is like studying History without knowing Geography.” - Prof. Marc Kirschner

1. The concentration rule of thumb.

In the last homework, we worked out the rule of thumb that one molecule per *E. coli* cell corresponds to a concentration of ≈ 1 nM.

(a) As an application of this idea, how many H^+ ions are there in a bacterial cell if the pH is 7.0?

(b) It is very useful to have a sense of how far molecules are apart at a given concentration. Work out a formula that relates the spacing between molecules d to the concentration c . Then, make a plot that shows the distance between molecules as a function of the concentration for concentrations ranging from nM to M.

(c) As an application of your thinking from part (b), explain what the concept of the “critical concentration” is for the polymerization of actin filaments. Then, provide a rough estimate of the mean spacing between actin monomers in a solution at the critical concentration.

2. RNA Polymerase and Rate of Transcription.

One of the ways in which we are trying to cultivate a “feeling for the organism” is by exploring the processes of the central dogma. Specifically, I want you to have a sense of the number of copies of the key molecular players in the central dogma as well as the rates at which they operate. Further, I argue that it is critical you have a sense of *how* we know these numbers.

(a) If RNA polymerase subunits β and β' together constitute approximately 0.5% of the total mass of protein in an *E. coli* cell, how many RNA polymerase molecules are there per cell, assuming each β and β' subunit within the cell is found in a complete RNA polymerase

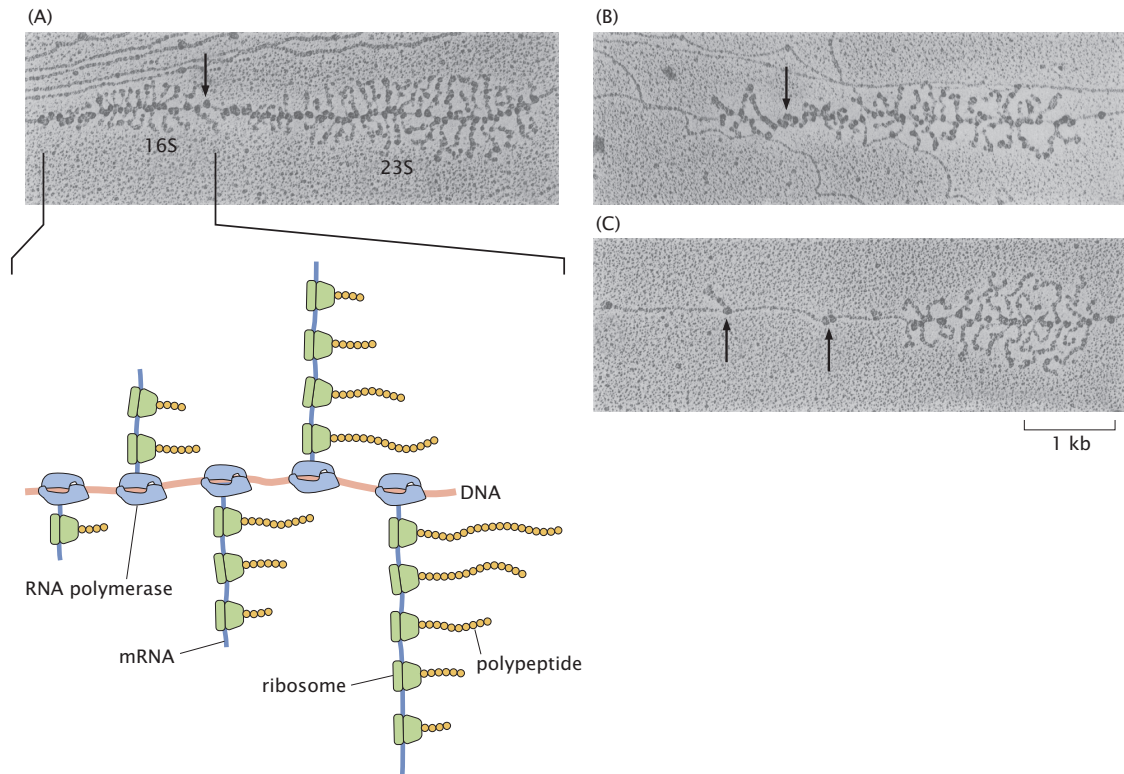


Figure 1: Effect of rifampin on transcription initiation. Electron micrographs of *E. coli* rRNA operons: (A) before adding rifampin, (B) 40 s after addition of rifampin, and (C) 70 s after exposure. No new transcripts have been initiated, but those already initiated are carrying on elongation. In parts (A) and (B) the arrow signifies the site where RNaseIII cleaves the nascent RNA molecule producing 16S and 23S ribosomal subunits. RNA polymerase molecules that have not been affected by the antibiotic are marked by the arrows in part (C). (Adapted from L. S. Gotta et al., *J. Bacteriol.* 20:6647, 1991.)

molecule? The subunits have a mass of 150 kDa each. (Adapted from problem 4.1 of Schleif, 1993.)

(b) Rifampin is an antibiotic used to treat *Mycobacterium* infections such as tuberculosis. It inhibits the initiation of transcription, but not the elongation of RNA transcripts. The time evolution of an *E. coli* ribosomal RNA (rRNA) operon after addition of rifampin is shown in Figure 3.29(A)–(C). An operon is a collection of genes transcribed as a single unit. Use the figure to estimate the rate of transcript elongation. Use the beginning of the “Christmas-tree” morphology on the left of Figure 3.29(A) as the starting point for transcription.

(c) Using the calculated elongation rate estimate the frequency of initiation off of the rRNA operon. These genes are amongst the most transcribed in *E. coli*.

3. A feeling for the complete blood count (CBC) test.

Test	Value
Red blood cell count (RBC)	Men: $\approx(4.3\text{--}5.7) \times 10^6$ cells/ μL Women: $\approx(3.8\text{--}5.1) \times 10^6$ cells/ μL
Hematocrit (HCT)	Men: $\approx(39\text{--}49)\%$ Women: $\approx(35\text{--}45)\%$
Hemoglobin (HGB)	Men: $\approx(13.5\text{--}17.5)$ g/dL Women: $\approx(12.0\text{--}16.0)$ g/dL
Mean corpuscular hemoglobin (MCH)	$\approx(26\text{--}34)$ pg/cell
MCH concentration (MCHC)	$\approx(31\text{--}37)\%$
Mean corpuscular volume (MCV)	$\approx(80\text{--}100)$ fL
White blood cell count (WBC)	$\approx(4.5\text{--}11) \times 10^3$ cells/ μL
Differential (% of WBC):	
Neutrophils	$\approx(57\text{--}67)$
Lymphocytes	$\approx(23\text{--}33)$
Monocytes	$\approx(3\text{--}7)$
Eosinophils	$\approx(1\text{--}3)$
Basophils	$\approx(0\text{--}1)$
Platelets	$\approx(150\text{--}450) \times 10^3$ cell/ μL

Table 1: Typical values from a CBC. (Adapted from R. W. Maxwell, Maxwell Quick Medical Reference, Tulsa, Maxwell Publishing Company, 2002.)

Typical results for a complete blood count (CBC) are shown in Table 1. Assume that an adult has roughly 5 L of blood in his or her body. Based on these values estimate:

- the number of red blood cells.
- the percentage in volume they represent in blood.
- their mean spacing.
- the total amount of hemoglobin in the blood.
- the number of hemoglobin molecules per cell.
- the number of white blood cells in the blood.

4. Testing the model of nucleolus scaling.

In class, we discussed the scale and scaling of various cellular structures and processes. In particular, we talked about the scaling of the size of the nucleolus with the size of the nucleus itself in the *C. elegans* embryo. Using a simple model, we derived an expression for

the number of molecules that make the nucleolus (in this case FIB-1 molecules) given by

$$M = \left(\frac{N}{V} - \frac{k_{off}}{k_{on}} \right) V, \quad (1)$$

where N is the total number of FIB-1 molecules inside the nucleus, k_{on} is the rate of FIB-1 incorporation into the nucleolus, k_{off} is the rate with which FIB-1 molecules detach from the nucleolus, and V is the nuclear volume.

We explored two types of experiments. First, we discussed an experiment in which the normal course of development leads to the progressive reduction of cell—and nuclear—size. In this case, the FIB-1 concentration within each nucleus remains constant such that we can rewrite Equation 1 as

$$M = (c_{tot} - c_*) V, \quad (2)$$

where $c_{tot} = N/V$ is the FIB-1 concentration and $c_* = k_{off}/k_{on}$ is the critical concentration at which the nucleolus forms. A second experiment relied on altering the expression of genes that lead to the formation of *C. elegans* embryos with larger or smaller cells. The assumption is that, in these mutants, the total number of FIB-1 molecules N will not change, but its nuclear concentration will. As a result, we can write Equation 1 for this case of constant number as

$$M = N - c_* V. \quad (3)$$

(a) Read the paper by Weber and Brangwynne (provided on the course website) and, in one short paragraph, explain how they managed to change the size of cells within the embryo and how they ensured that, for all embryo sizes, the total number of FIB-1 molecules remained constant.

The two types of experiments captured by Equations 2 and 3 give us an opportunity to test the predictive power of our model. Specifically, note that Equation 3 predicts that, for the fixed FIB-1 number experiment, the y-intercept of the scaling of the nucleolus with nuclear volume will be given by N while the slope will be $-c_*$.

(b) Write Python code to plot the data (provided on the course website) and perform a manual linear fit to the data in order to estimate the value of N and c_* .

(c) Now, use the parameters inferred in (b) to predict the scaling of nucleolar size versus nucleus volume for the fixed FIB-1 concentration experiment. Specifically, draw a plot where you overlay the experimental data with your theoretical prediction.

5. Post-Translational Modifications and “nature’s escape from genetic imprisonment”

In a very interesting article (“Post-translational modification: nature’s escape from genetic imprisonment and the basis for dynamic information encoding”), Prof. Jeremy Gunawardena

discusses how we should think about post-translational modifications as a way of expanding the natural repertoire of the 20-letter amino acid alphabet. Similarly, Prof. Christopher Walsh wrote a whole book entitled “Posttranslational Modifications of Proteins: Expanding Nature’s Inventory,” again making the point that by adding chemical groups to proteins we can significantly change their properties.

(a) Provide at least one mechanistic idea about how adding a chemical group to a protein can alter its structure or function. Your answer should be offered in less than a paragraph, but should be concrete in its assertions about how these modifications change the protein. Why does Gunawardena refer to this process of post-translational modification as “escape from genetic imprisonment”?

(b) As a toy model of the combinatorial complexity offered by post-translational modifications, let’s imagine that a protein has N residues that are able to be phosphorylated (NOTE: please comment on which residues these are - the answer is different for bacteria and eukaryotes). How many distinct states of the protein are there as a result of these different phosphorylated states? Make an approximate estimate of the mass associated with a phosphate group and what fraction of the total mass this group represents. Similarly, give some indication of the charge associated with a phosphate group. What ideas do you have about how we can go about measuring these different states of phosphorylation?

(c) In this part of the problem, we make a very crude estimate of the number of sites on a protein that are subject to phosphorylation. To do so, imagine that the protein is a sphere with N residues. How does the radius of that sphere depend upon the number of residues in the protein? Given that estimate, what is the number of residues that are on the surface? Given that number, what fraction of those are phosphorylatable? Remember, these are crude estimates. Work out these results for a concrete case of a typical protein with roughly 400 amino acids.

(d) Let’s close out these estimates by thinking about a bacterial cell. If all 3×10^6 proteins in such a cell can be phosphorylated with the number of different phosphorylation states that you estimated above, how many distinct cells could we make with all of these different states of phosphorylation.

6. The pandemic elephant in the room.

(a) What is the information density of the SARS-CoV-2 virus? What I mean is that there are a certain number of bits of information contained in the viral genome, so you can report a density of bits/nm³.

(b) What is the information density of a typical hard drive for backing up our laptops?

(c) Given your answer to part (a), how many SARS-CoV-2 viruses would it take to capture all the information in the Library of Congress? How much volume would such a “library” take up? Could the whole Library of Congress fit into one 5 mL tube, a one L flask, one

shelf of a -80 freezer?