MCB137L/237L: Physical Biology of the Cell
Spring 2024
Homework 2
(Due 2/1/24 at 2:00pm)

Hernan G. Garcia

Last updated January 22, 2024

"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 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 Figures 1(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 1(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.

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:

(a) the number of red blood cells.

(b) the percentage in volume they represent in blood.

(c) their mean spacing.

(d) the total amount of hemoglobin in the blood.

(e) the number of hemoglobin molecules per cell.

(f) the number of white blood cells in the blood.

4 Migration of the bar-tailed godwit

Animal migrations are one of the greatest of interdisciplinary subjects, bringing together diverse topics ranging from animal behavior to the physics of navigation to the metabolism required for sustained long-distance travel. The bar-tailed godwit is a small bird that each year travels between Alaska and New Zealand on the same kind of incredible nonstop voyage

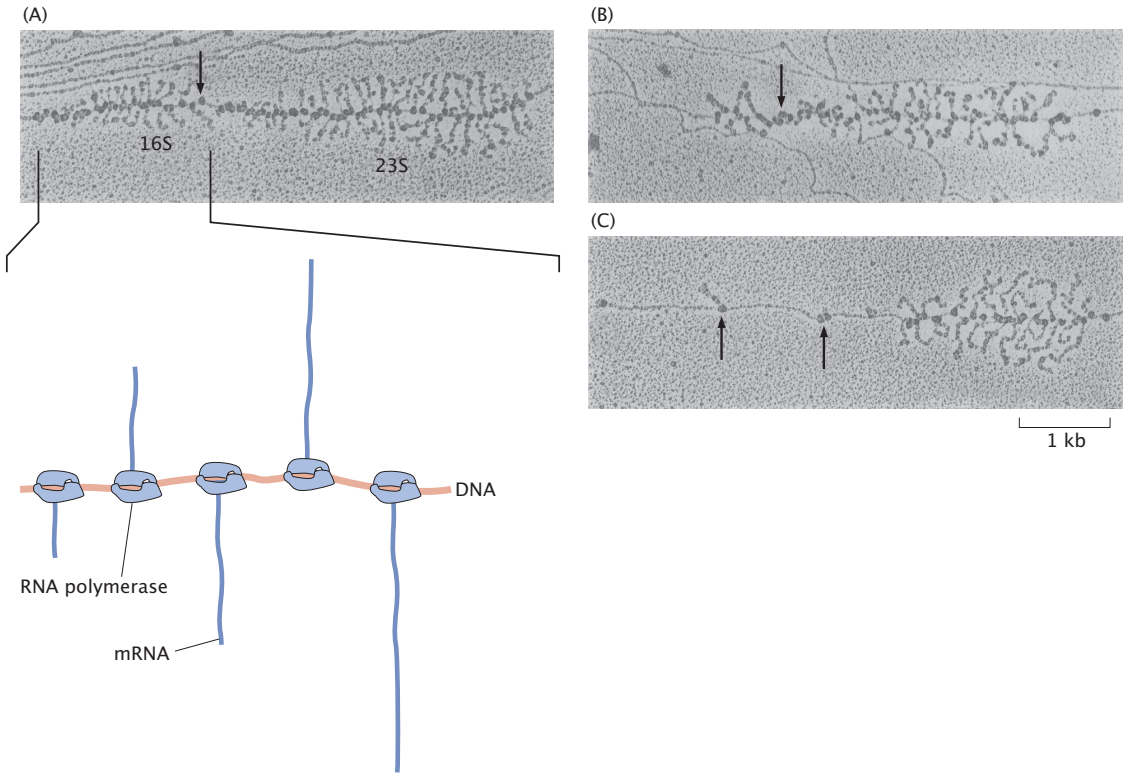

Figure 1: Effect of rifampin on transcription initiation. Electron micrographs of *E. coli* rRNA operons: (A) before adding rifampin, (B) 40 s after addition of rifampin, and (C) 70 s after exposure. No new transcripts have been initiated, but those already initiated are carrying on elongation. In parts (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.)

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.)

taken by happy tourists in modern long-distance jetliners as shown in Figure 2. A naturalist guide in the Okarito Lagoon in New Zealand's South Island once claimed that over the course of their ten-day, ten-thousand kilometer trip, these amazing migratory birds lose 1/3 of their body mass. In this problem, we make a series of simple divide-and-conquer estimates to see whether this claim might be true.

One of the most powerful tools for estimation is dimensional analysis. Here, we will use this type of analysis to estimate the drag force experienced by flying godwits as they migrate. This sort of analysis makes it possible to quickly answer questions such as whether the drag force increases linearly or quadratically with the length of the birds. In dimensional analysis, we amass the various physical parameters that we imagine will dictate the drag force with their attendant units on the right hand side of the equation. In this case, we will consider the density of air (ρ), the speed of the birds (v) and the size of the birds (L, representing the size of the cross-section of the bird). On the left hand side, we have the drag force, resulting in

$$F_{drag} = \mathrm{C}\rho^{\alpha} v^{\beta} L^{\gamma}, \tag{1}$$

where C is a dimensionless constant that we will not consider further. Note that in the equation we have proposed a set of exponents α, β and γ for each variable. The idea of dimensional analysis is to find the exponents that balance the units on each side of the equation as a means to uncover the scaling of the drag force with the various physical parameters we proposed.

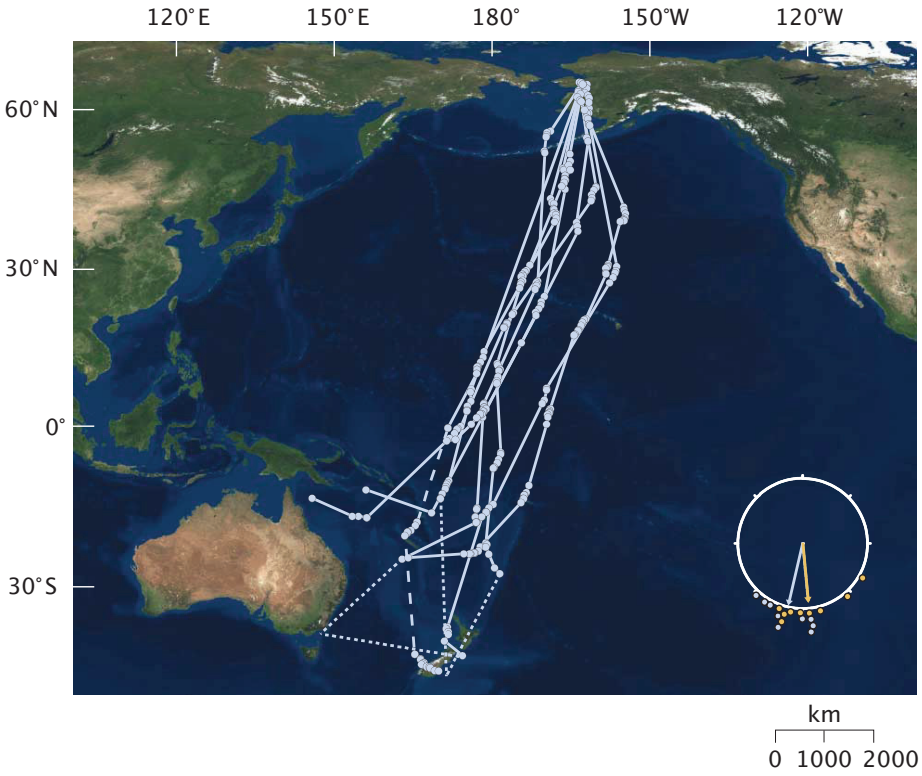

Figure 2: Map showing the migration pattern of the bar-tailed godwit. Adapted from Gill *et al.*, Extreme endurance flights by landbirds crossing the Pacific Ocean: ecological corridor rather than barrier?, Proc Biol Sci. 2009 Feb 7; 276(1656): 447-457.

(a) Using dimensional-analysis arguments, work out how the drag force experienced by the flying godwits depends upon the density of air (ρ), the speed of the birds (v) and the size of the birds (L). Specifically, work out the coefficients α, β and γ in Equation 1.

(b) Work out the power expended by the bar-tailed godwit to overcome the drag force. Then, work out the total energy expended during the ten-day migration in overcoming this drag force.

(c) Given that burning fat yields 9 kcal/g, work out the number of grams of fat that would need to be burned to sustain the ten day flight of the bar-tailed godwit. What fraction of the bird's body mass would be loss during such a migration based on these estimates?

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

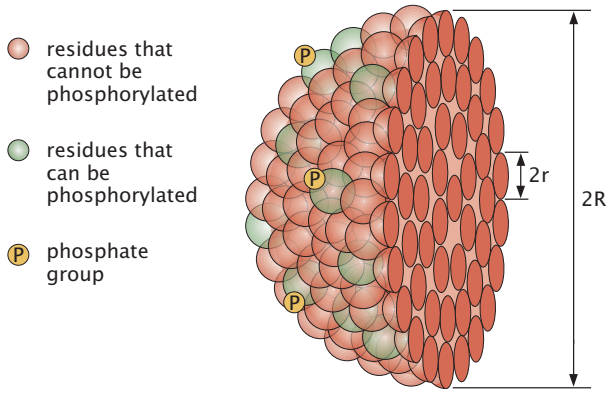

Figure 3: Schematic of a protein showing the surface residues that are available for phosphorylation.

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 Real Estate for the Factories of ATP Synthesis

We are captivated by the tension between those things about living organisms that are universal and those things that are baroque and specific to a given organism. One of the nearly universal features of living organisms on our planet is their use of ATP hydrolysis as an energy source for a huge variety of processes. Further, as explained below, there is a nearly constant power density to run cellular life across the entire span of the tree of life. Where does all of this ATP come from? Cells have tiny molecular machines known as ATP synthase in the membrane which use an ion gradient to drive the 6000 rpm rotation of these machines to produce a few ATPs each rotation. However, the ATP is consumed within the volume of cells, but is produced on membranes. This leads to the possibility that as cells get bigger, there may be a point at which the surface area is insufficient to keep up with the demands of the cytoplasmic volume. Indeed, this problem explores the hypothesis that for cells above a certain size, the synthesis of ATP at the plasma membrane (such as in bacteria) no longer sufficed and that a new specialized energy factory (i.e. the mitochondria) was required.

(a) By considering the cost of protein synthesis for a dividing bacterium with a 1000 s division time, justify the assertion that the power usage is

$$\text{power density} = 10^6 \frac{\text{ATP}}{\mu\text{m}^3 \text{ s}}. \tag{2}$$

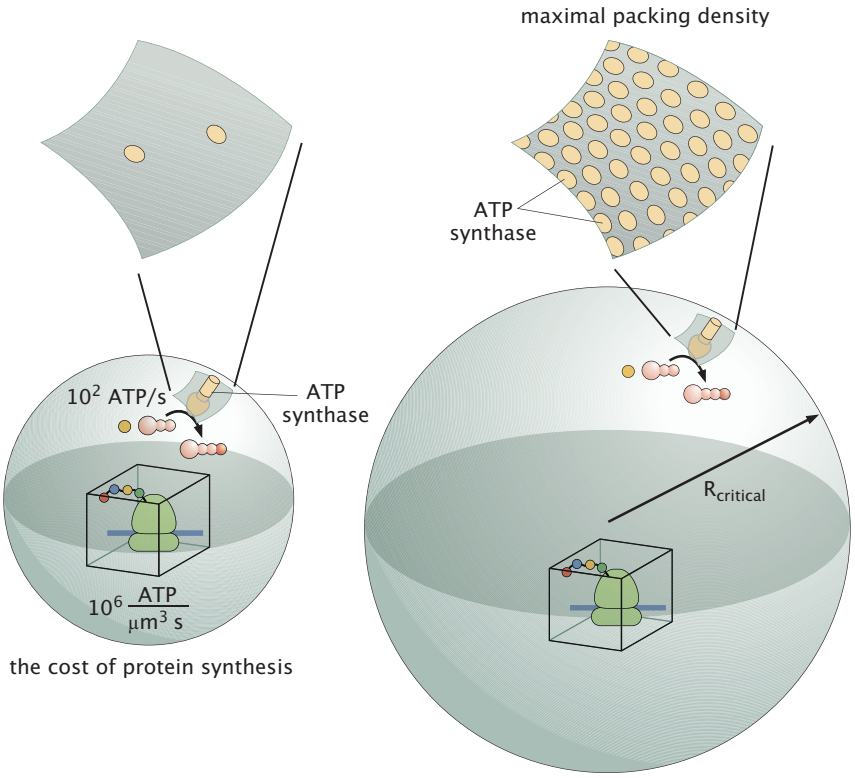

Figure 4: Surface coverage of bacterial cells with ATP synthase. For small cells, the demands of the cytoplasmic power consumption can be met by ATP synthases on the plasma membrane. However, for larger cells, there is not enough surface area to keep up with the demands of the power needs of the cellular interior.

As a reminder, one way to do this estimate is to figure out how many proteins there are in an *E. coli* cell and how many amino acids there are per protein, noting that it costs 4 ATP equivalents for every polypeptide bond.

(b) As shown in Figure 4, compute the maximum radius of a spherical cell that could sustain the demands of ATP synthesis (i.e. the $10^6 \text{ATP}/(\mu\text{m}^3\ \text{s})$ required to run the cellular economy) by the presence of ATP synthases on its surface. Use your results to comment on the way prokaryotes and eukaryotes generate ATP and how large eukaryotes get around this conundrum.