

BE/APh161: Physical Biology of the Cell

Homework 2

Due Date: Wednesday, January 17, 2024

Solutions by: Tom Röschinger

“Whatever you can do, or dream you can do, begin it. Boldness has genius, power and magic in it.” - Goethe

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?

Solution: We convert pH to concentration of H^+ ions from the relationship

$$[H^+] = 10^{-\text{pH}} M. \quad (1)$$

So for a pH of 7.0, $[H^+] = 10^{-7}$ M, or 10² nM. Thus, since we know that 1 molecule per bacterium is equivalent to a concentration of ≈ 1 nM, a pH of 7.0 corresponds with 100 H^+ ions in the cell.

(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.

Solution: Our goal is to determine the average spacing between molecules given a specified concentration. From the molar concentration, we can express the molecular density as

$$\frac{c \text{ mol}}{L} \cdot \frac{6 \times 10^{23} \text{ molecules}}{\text{mol}} \cdot \frac{L}{10^{24} \text{ nm}^3} = c \cdot 0.6 \text{ molecules/nm}^3. \quad (2)$$

Inverting this result, we generate the volume of solution occupied by each molecule at molar concentration C M,

$$V = 1.66 \text{ nm}^3/\text{molecule } 1/c. \quad (3)$$

The cubed root of this volume thus indicates the average separation between molecules, such that we may conclude

$$d \propto V^{1/3} = \frac{1.18}{c^{1/3}} \text{ nm}. \quad (4)$$

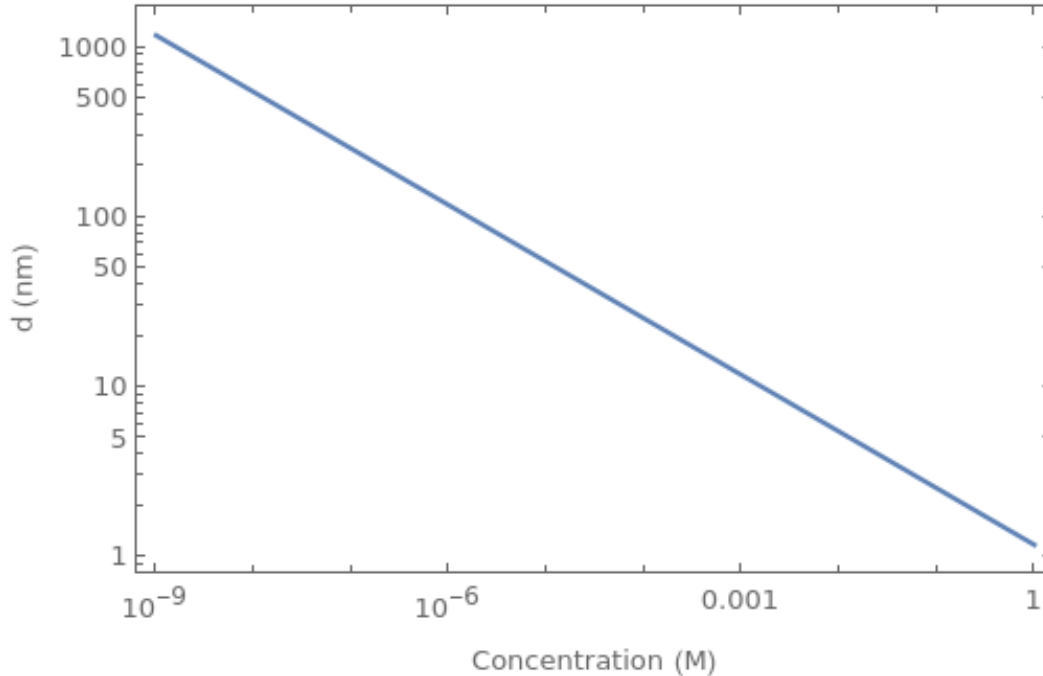


Figure 1: Plot of average separation d between molecules against concentration ranging from nM to M.

(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.

Solution: Polymerization in solution can be considered under a simple model: after a nucleation phase where 3-4 monomers randomly interact to form a nucleus which can then be elongated. This elongation is then governed by monomer capture events (monomers polymerize onto one end of the polymer) and monomer escape events (a monomer leaves one end of the polymer). Monomer capture is dependent on the interaction between one end of the polymer and a monomer in solution, and is therefore likelier the smaller the mean separation between monomers in solution - we can capture this by setting the rate of monomer capture to be proportional to the concentration of monomers. Monomer escape, however, does not require interaction with monomers in solution and is therefore independent of the monomer concentration. This can be written as

$$\frac{dn}{dt} = k_{on}C - k_{off} \quad (5)$$

where n is the number of monomers that constitute the polymer we are considering, k_{on} (resp. k_{off}) are rate constants for monomer capture (resp. escape), and C is the monomer concentration in solution.

We can then see that the system is in steady state when $C = C_{crit} \equiv \frac{k_{off}}{k_{on}}$. This is referred to as the critical concentration - below this concentration the filaments depolymerize and above this concentration the filaments keep polymerizing, pulling monomers out of solution until the critical concentration is reached.

Actin polymerization, however, is more complicated than this draft model captures. It turns out that the ends of actin filaments behave asymmetrically - i.e. actin filaments have a ‘plus’ end and a ‘minus’ end with the ‘plus’ end having a higher growth (and shrinkage) rate. A better first-order model is given by

$$\frac{dn}{dt} = k_{on}^+ C + k_{on}^- C - k_{off}^+ - k_{off}^- \quad (6)$$

where the ‘+’ and ‘-’ superscripts in the rate constants denote the ‘plus’ and ‘minus’ end of the filament respectively. (Note: There are more subtleties involved in actin polymerization - e.g. the monomer capture rates also depend on whether the monomer is ADP or ATP bound - but the above is sufficient as a first-order model.)

This system now has three critical concentrations: $C_+ \equiv \frac{k_{off}^+}{k_{on}^+}$ below which both ends shrink, $C_- \equiv \frac{k_{off}^-}{k_{on}^-}$ above which both ends grow, and $C_{TM} \equiv \frac{k_{off}^+ + k_{off}^-}{k_{on}^+ + k_{on}^-}$ at which the system reaches steady-state (referred to as ‘treadmilling’) whereby the ‘plus’ end grows at the same rate at which the ‘minus’ end shrinks. For concentrations between C_- and C_+ , the ‘plus’ end grows and the ‘minus’ end shrinks, with the relative rates of the two processes determining whether or not the filament elongates.

As per BNID 112788, $C_+ \sim 0.06 \mu\text{M}$ and $C_- \sim 0.6 \mu\text{M}$, corresponding to a mean separation of $\sim 300 \text{ nm}$ and $\sim 100 \text{ nm}$ respectively. C_{TM} lies somewhere between the two, which we estimate to be $\sim 0.2 \mu\text{M}$ and which corresponds to a mean separation of $\sim 200 \text{ nm}$.

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.)

Solution: We have discussed measurements showing that the dry weight of an *E. coli* cell is roughly 30% of its total mass, half of which is protein. The total mass of the cell is estimated to be a picogram from the assumption that its density is nearly that of water and its volume roughly $1 \mu\text{m}^3$. It is given in the problem that the β and β' subunits together have a mass of 300 kDa and comprise 0.5% of the protein mass. The number of β , β' subunit pairs is assumed to equal the number of RNA polymerases (RNAP). Putting all this together yields

$$\# \text{ of RNAP} = \frac{\text{total mass of RNAP in cell}}{\text{mass per RNAP}} = \frac{0.15 \times 10^{-12} \text{ g} \times 0.005}{3 \times 10^5 \text{ Da} \times 1.6 \times 10^{-24} \text{ g/Da}} = 1.5 \times 10^3. \quad (7)$$

(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 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.

Solution: Comparing Figure 1(A) and Figure 1(B), one sees that 40 seconds after rifampin addition roughly 1.5 kb of the DNA from the start site has become free of RNAP. The micrographs are aligned well enough that one can assume the left edge in all of them is the start site. Assuming that the last RNAP to initiate transcription did so at nearly the same time as rifampin addition, one can infer that this RNAP transcribed 1.5 kb of DNA in 40 seconds, implying an elongation rate of

$$\text{Elongation Rate} = \frac{1.5 \text{ kb}}{40 \text{ seconds}} \approx 0.04 \text{ kb/sec} \quad (8)$$

Making the same comparison of Figure 1(A) and 1(C), indicates an elongation rate of $3.5 \text{ kb}/70 \text{ sec} = 0.05 \text{ kb/sec}$, or roughly 50 nucleotides/sec.

(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*.

Solution: The operon is roughly 6 kb long. Given the elongation rate in (b), one RNAP would require $\frac{6 \text{ kb}}{0.05 \text{ kb/sec}} = 120 \text{ seconds}$ to complete a transcript.

To estimate the rate at which transcripts of the operon are made, one needs the number of RNAP on the operon at any one time. Looking at the micrograph, one can make a rough

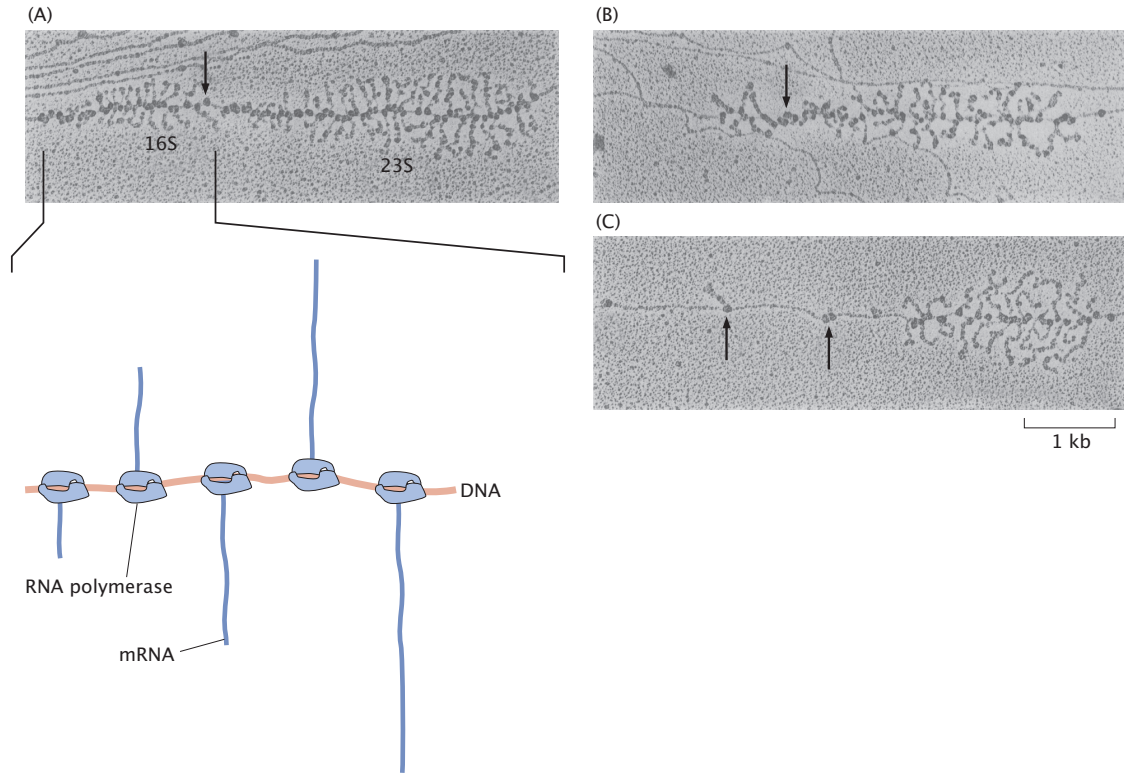


Figure 2: Effect of rifampin on transcription initiation. Electron micrographs of *E. coli* rRNA operons: (A) before adding rifampin, (B) 40 s after addition of rifampin, and (C) 70 s after exposure. No new transcripts have been initiated, but those already initiated are carrying on elongation. In parts (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.)

count that under normal conditions there are 10 – 20 RNAP per kilobase and that the operon is roughly 6 kb long. This implies roughly

$$6 \text{ kb} \times 15 \text{ RNAP/kb} = 90 \text{ RNAP} \quad (9)$$

on the operon, and if each RNAP requires 120 seconds to complete transcription, then the initiation frequency of the operon is $90/120\text{s} = 0.75\text{s}^{-1}$. This corresponds to a production rate that just over ~ 1 transcript per second.

Alternatively, we can notice that the mean spacing between RNAPs is roughly $1000\text{nt}/15 \approx 67\text{nt}$. The average speed of each such RNAP (from the previous part of the problem) is 50 nt/s . Hence, the initiation frequency is

$$\frac{50 \text{ nt/s}}{67 \text{ nt}} \approx 0.75 \text{ s}^{-1} \quad (10)$$

which is the same as our first result.

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.

Solution: From the table, we see that a typical red blood cell count is around 5×10^6 cells per μL . Scaling this up to the 5 L of blood, we get a total number of red blood cells,

$$5\text{L} \times \frac{10^6 \mu\text{L}}{\text{L}} \times 5 \times 10^6 \frac{\text{cells}}{\mu\text{L}} \approx 2 \times 10^{13} \text{ cells}$$

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

Solution: Red blood cells have a volume of around 100 fL (BNID:110805). This means the total volume of all 20 trillion red blood cells is

$$100 \frac{\text{fL}}{\text{cell}} \times 2 \times 10^{13} \times \frac{\text{L}}{10^{15} \text{fL}} = 2 \text{ L}$$

This means that the red blood cells take up two-fifths, or about 40% of the total blood volume. Coincidentally, the hematocrit value from the table corresponds to this value as empirically determined, and we see that 40% is right on the money for the actual value.

(c) their mean spacing.

Solution: Converting our 5 L of blood into a length scale more meaningful for cells, we get that we have $5 \times 10^{15} \mu\text{m}^3$ of blood. Dividing by the number of cells, we get

$$\frac{5 \times 10^{15} \mu\text{m}^3 \text{ blood volume}}{2 \times 10^{13} \text{ cells}} = 250 \mu\text{m}^3$$

blood volume that each cell is “allotted”. We can alternatively think of this as each cell getting an $\approx 6 \times 6 \times 6 \mu\text{m}$ box to call its own, meaning there is $\approx 6 \mu\text{m}$ spacing between cells.

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

Solution: Reading off the table, we see that hemoglobin is around 15g/dL . Scaling up to 5 L of blood, we get a total hemoglobin mass of

$$15 \frac{\text{g}}{\text{dL}} \times 10 \frac{\text{dL}}{\text{L}} \times 5\text{L} = 750\text{g}$$

(e) the number of hemoglobin molecules per cell.

Solution: To convert the mass of hemoglobin to a number of hemoglobin molecules we simply need to divide by the mass of a hemoglobin molecule. To estimate this, we harken back to Problem 2, where the typical amino acid is 100 Da and the typical protein, meaning the typical protein is around 30 kDa in mass. If we recall that hemoglobin is actually made of 4 subunits, we might adjust our estimate of hemoglobin mass to be four times larger, or around 120 kDa. (It turns out that this is a bit of an over estimate, since each subunit isn't as big as a "typical" protein, but this is sufficient for an order-of-magnitudes estimate). This gives us

$$\frac{750 \text{ g of hemoglobin}}{120 \text{ kDa}} \times \frac{6 \times 10^{20} \text{ kDa}}{\text{g}} \approx 4 \times 10^{21} \text{ hemoglobin molecules}$$

Finally, the number of hemoglobin per cell is

$$\frac{4 \times 10^{21} \text{ hemoglobin}}{2 \times 10^{13} \text{ cells}} = 2 \times 10^8 \text{ hemoglobin/cell}$$

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

Solution: Again reading off the table, we see that a typical value for white blood cells is 8×10^3 per μL . Scaling up to 5 L of blood, we get a total number

$$5\text{L} \times \frac{10^6 \mu\text{L}}{\text{L}} \times 8 \times 10^3 \frac{\text{cells}}{\mu\text{L}} \approx 4 \times 10^{10} \text{ cells}$$

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 taken by happy tourists in modern long-distance jetliners as shown in Figure 3. During a visit to New Zealand's South Island, one of us had the chance to see these amazing birds in Okarito Lagoon with a naturalist guide who claimed that over the course of their ten-day,

| 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.)

ten-thousand kilometer trip, these 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.

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

$$F_{drag} = \text{const.} \rho^\alpha v^\beta L^\gamma. \quad (11)$$

Solution: Because the units of force are $\text{kg} \cdot \text{m}/\text{s}^2$, and the quantity ρ is the only parameter on the right of Eq. 11 that features the mass, we can conclude that $\alpha = 1$. Once we have determined α , the whole structure falls like dominoes. Using similar reasoning, we see that $\beta = 2$ since v is the only quantity featuring the time, and we need two powers of time in the denominator to match the units on the left side of the equation. Dimensional consistency then requires that $\gamma = 2$ as well, signaling that the drag force scales with the area of the moving object.

(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

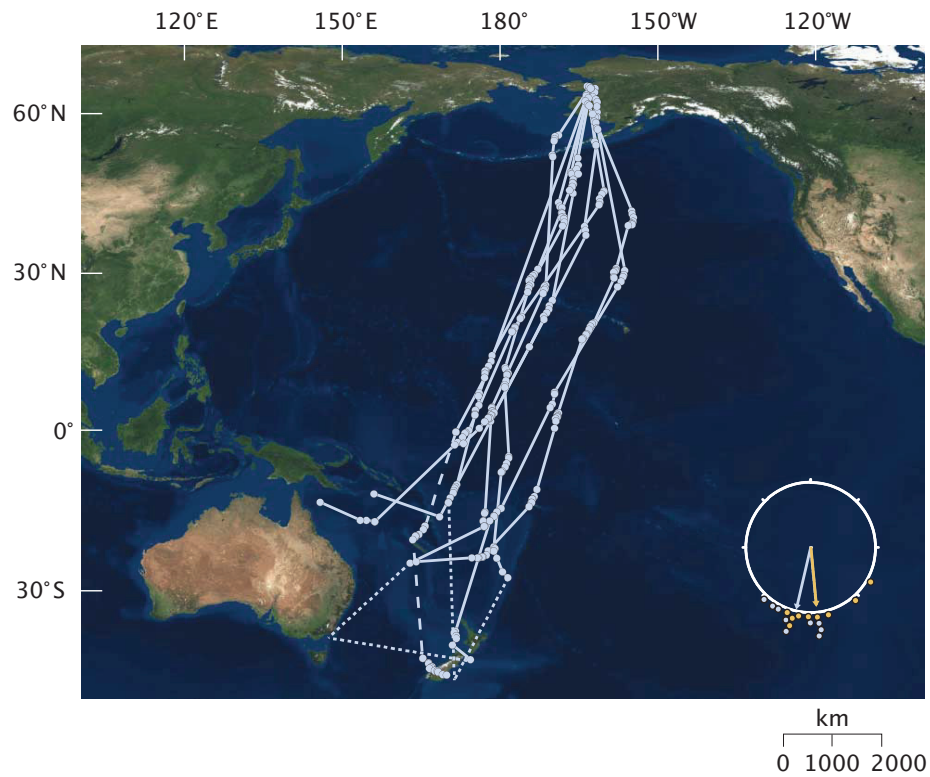


Figure 3: Map showing the migration pattern of the bar-tailed godwit. Adapted from Gill *et al.*, Extreme endurance flights by landbirds crossing the Pacific Ocean: ecological corridor rather than barrier?, Proc Biol Sci. 2009 Feb 7; 276(1656): 447-457.

force.

Solution: We can compute

$$\text{power during flight, } P = F_{drag}v = \text{const.} \cdot \rho v^3 L^2 \quad (12)$$

and the total energy expended is

$$\text{energy expended, } E = Pt_{tot} \quad (13)$$

where t_{tot} is the total time of flight. For the purpose of an order of magnitude estimate, we will assume that the dimensionless constant in (12) is of order unity. The average velocity of the bird during its trip can be estimated as

$$v = \frac{\text{total migratory distance}}{\text{time of flight}} \quad (14)$$

$$\approx \frac{10^4 \text{ km}}{10 \text{ days}} \quad (15)$$

$$\approx \frac{10^7 \text{ m}}{10 \times 10^5 \text{ s}} \quad (16)$$

$$\approx 10 \text{ ms}^{-1} \quad (17)$$

By looking up pictures of satellite trackers being put on bar-tailed godwits, we can see that the bird is roughly the size of a human palm, i.e. $L \approx 10^{-1} \text{ m}$. The density of air is $\rho \approx 1 \text{ kg m}^{-3}$. Then we have

$$P \approx (1 \text{ kgm}^{-3}) \times (10 \text{ ms}^{-1})^3 (10^{-1} \text{ m})^2 \approx 10W \quad (18)$$

and

$$E = Pt_{tot} \approx (10 \text{ W}) \times (10 \times 10^5 \text{ s}) \approx 10^7 \text{ J} \quad (19)$$

(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.

Solution: Using our result from the previous part,

$$\text{total fat required} = \frac{10^7 \text{ J}}{9 \text{ kcal/g}} \quad (20)$$

$$\approx \frac{10^7 \text{ J}}{9 \times \text{kcal/g} \times 4000\text{J/kcal}} \quad (21)$$

$$\approx 250 \text{ g} \quad (22)$$

Taking the birds to be spherical with the density of water, we can estimate a body mass of ~ 500 g. Therefore, the claim that the bar-tailed godwit loses 1/3 of its body mass over its migration is order-of-magnitude correct.

Note: Our estimate for the energy expenditure is an overestimate by a factor of a few - much of this error is due to the rough way we have estimated the drag force experienced by the bird, from which we should know to only expect order-of-magnitude results. The body mass of the bird is also slightly overestimated here, with bar-tailed godwits weighing in closer to ~ 300 g.

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 (also at Harvard) 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”?

Solution: There are many ways in which adding a chemical group can affect the structure of a protein. For instance, it could promote dimerization by providing an energetically favorable surface for two binding events. Alternatively, adding a charged chemical group could cause increased electrostatic repulsion within the protein. This could cause the protein to “open up,” which, among other effects, could alter function by allowing access to previously occluded binding pockets.

(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?

Solution: The state of the protein will be determined by which residues are phosphorylated, not merely how many. Therefore, for N residues, there will be 2^N states. For bacteria, the most commonly phosphorylated amino acids are histidine, serine, threonine, and tyrosine, while for eukaryotes they are serine, threonine, and tyrosine. A phosphate group is composed of a phosphorus and four oxygen atoms. Its mass is thus roughly 100 Daltons, which is approximately the same as that of an amino acid.

A typical protein is composed of 300 amino acids, and so, the fraction of the total mass that one represents is $100 \text{ Da}/30,000 \text{ Da} = 5 \times 10^{-3}$. A phosphate group has a charge of $-2e^-$ when bound to a residue, which can be very important for protein's function. The addition of a phosphate group will greatly affect mass to charge ratio, so the use of mass spectrometry would be a very powerful technique for measuring the number of phosphorylated amino acids.

(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.

Solution: In our toy model we can assume that each residue is itself a sphere. The crude scaling of the radius of our spherical protein will go as $N^{1/3}$. However, we can do a little better. If the amino acid spheres are maximally close packed, then the volume of the protein will be roughly $0.75N \times V_{aa}$. Therefore, the radius of our protein will be given by $R = (0.75N)^{1/3}R_{aa}$. Given that a typical protein has 400 amino acids and a radius of 2.5 nm, the radius of a single amino acid becomes $R_{aa} \approx 0.4 \text{ nm}$. Therefore, the scaling of the protein radius with the number of amino acids will go approximately as

$$R(N) \approx 0.4 \times (0.75N)^{1/3} \text{ nm}. \quad (23)$$

As a very crude estimate for the total number of residues on the surface of the protein, we can take the total surface area and divide it by the 2D projection area of a single residue, that is

$$S_{\text{array}} = 4\pi R^2, \quad (24)$$

$$S_{aa}^{\text{proj}} = \pi R_{aa}^2, \quad (25)$$

$$N_{\text{surface}} = \frac{S_{\text{protein}}}{S_{aa}^{\text{proj}}} = 4 \times (0.75N)^{2/3} \approx 3N^{2/3}. \quad (26)$$

Taking $N = 400$ for a typical protein, we find that the number of surface residues to be

$$N_{\text{surface}} \approx 150. \quad (27)$$

As discussed in part (b), 4 of 20 amino acids are commonly phosphorylated in bacteria. Making the crude assumption that all amino acids are equally represented on the surface,

only 20% of the residues, or $150 \times 0.2 \approx 30$, will on average be phosphorylatable. This means that a total of $2^{30} \approx 10^6$ phosphorylation states are available to proteins.

(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.

Solution: If each protein is distinguishable, the total number of cells that could be created will be

$$N_{\text{cells}} = (2^{30})^{3 \times 10^6} = 10^{2 \times 10^7}, \quad (28)$$

since there are effectively 10^8 sites available for phosphorylation. However, if we assume instead that each protein is indistinguishable from any other, then the situation is identical to choosing 3×10^6 proteins to make a cell from the possible 2^{30} phosphorylation states (with replacement). The number of different cells in this case becomes (see the “Combinations with Repetition” section here: <https://www.mathsisfun.com/combinatorics/combinations-permutations.html>)

$$\begin{aligned} N_{\text{cells}} &= \binom{2^{30} + 3 \times 10^6 - 1}{3 \times 10^6} \approx \frac{(10^6 + 3 \times 10^6)!}{(10^6)!(3 \times 10^6)!} \\ &\approx \frac{(4 \times 10^6)^{10^6}}{(10^6)!} \\ &\approx 10^{10^6}. \end{aligned} \quad (29)$$

Taking the geometric mean of the upper and lower limits, we obtain

$$(10^{2 \times 10^7} \times 10^{10^6})^{1/2} \approx 10^{10^7}, \quad (30)$$

which is an “astronomically large” number.

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

$$\text{power density} = 10^6 \frac{\text{ATP}}{\mu\text{m}^2 \text{ s}}. \quad (31)$$

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.

Solution: Here we use the estimate that one *E. coli* cell weighs about 1pg, 30% of which is dry mass and of which half is protein. This would give us a total of 0.1pg of protein weight. Using the fact that one amino acid weighs about 100 Daltons, and 1 Dalton equals 1g/mol, we can compute the number of amino acids to be $0.1\text{pg} \times \text{mol}/100\text{g} \approx 10^9$. Assuming that all these amino acids connected by polypeptide bonds, that is a total of 4×10^9 ATPs that are required to make all proteins in the cell. Given the division time of 1000s, rate of ATP consumption is 4×10^6 ATP/s. To get the power usage, we use that *E. coli* has a volume of about $1\mu\text{m}^3$, which gives us a power usage of

$$\text{power density} = 4 \times 10^6 \text{ATP}/\mu\text{m}^3 \text{s}.$$

(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 $10^6 \text{ATP}/(\mu\text{m}^2 \text{ s})$ 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.

Solution: The total power of the cell is given by its power density multiplied by the volume of the cell with radius R , which is given by

$$\text{power} = 4 \times 10^6 \text{ATP}/\mu\text{m}^3 \text{s} \times 4R_{\text{cell}}^3 = 10^7 \times \text{ATP } R^3 \text{cell/s } \mu\text{m}^3,$$

where we assumed the cell is a sphere and $\pi/3 \approx 1$. Now, in the problem it is stated that ATP synthase turns at 6000rpm, producing few ATPs per rotation. This gives us an ATP production rate of

$$\text{ATP production rate} = 6000 \text{rotations}/\text{min} \times \text{min}/60\text{s} \times f \times \text{ATP}/\text{rotation} = f \times 10^2 \text{ATP}/\text{s}.$$

To find out how many synthases can be fit into the membrane, we have to compute the cross-section of a single protein. Assuming the protein is a sphere containing of 300 amino acids, the radius is approximately the third root of the volume, $R_{\text{protein}} \sim V^{(1/3)} = 6$. An amino-acid has a diameter on the order of 1nm, so the radius of our protein sphere is approximately 6nm. The cross-section of the protein is then given by

$$\text{cross section} = \pi R_{\text{protein}}^2 \approx 100\text{nm}^2.$$

To compute how many synthases can possibly fit into the membrane, we compute the surface area of the cell, assuming it is a sphere, and divide by the cross-section of one synthase,

$$\text{synthases in membrane} = 4\pi R_{\text{cell}}^2 / 100\text{nm}^2.$$

For the cell to be functional, the ATP production of a fully packed membrane has to exceed the ATP required inside the cell, therefore,

$$f \times 10^2 \text{ATP/s} \times 4\pi R_{\text{cell}}^2 / 100\text{nm}^2 > 10^7 \times \text{ATP } R_{\text{cell}}^3 / \text{s } \mu\text{m}^3,$$

which we can solve for the radius of the cell,

$$R_{\text{cell}} < f \mu\text{m}.$$

Here we see, that for cells to be able to produce their entire ATP need from ATP synthases in the cell membrane, they can only be the size of a few microns. Eukaryotes solve this issue by having mitochondria, which have are organelle within the cell that have an increased surface area by folding its inner membrane, allowing for more ATP synthases to be included in its membrane.

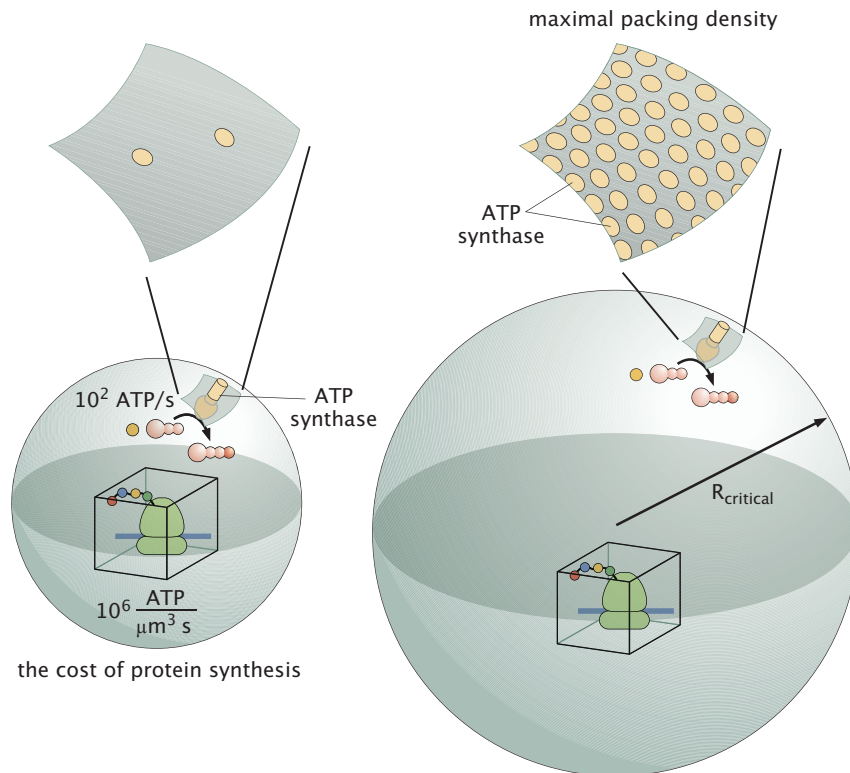


Figure 4: Surface coverage of bacterial cells with ATP synthase. For small cells, the demands of the cytoplasmic power consumption can be met by ATP synthases on the plasma membrane. However, for larger cells, there is not enough surface area to keep up with the demands of the power needs of the cellular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