Física Biológica

Biología en números

1. (*tomado del libro "Physical Biology of the Cell"*) La figura 2.1 del libro (ver Fig. 1) muestra varias imágenes experimentales de una célula de *E. coli* y su esquematización en lo que llaman su "regla estándar". La imagen AFM y la micrografía electrónica de las Figuras 2.1 (A) y (B) muestran que

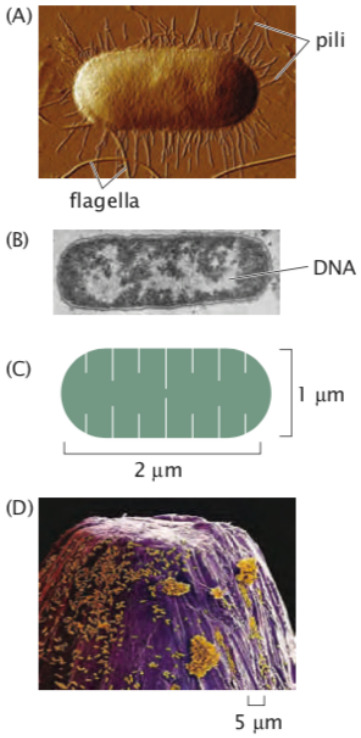

Figure 1. Fig. 2.1 del libro "Physical Biology of the Cell". A: cortesía de Ang Li; D: cortesía de Tony Brain.

estas bacterias tienen una morfología en forma de varilla con una longitud típica de entre 1 y 2 μm y un diámetro de entre 0,5 y 1 μm.

- (a)Estime el volumen de esta célula en fL y también su masa. Para esto último suponga que tiene la misma densidad que el agua.
- (b) Una estimación más razonable se obtiene considerando que la densidad de las macromoléculas de la célula es 1,3 veces la del agua. Como resultado, la estimación de la masa de una célula de *E. coli* se desvía un poco de la estimación anterior. Suponiendo que dos tercios de la masa corresponden al agua y que el tercio restante a macromoléculas, calcule el error porcentual cometido al tratar la densidad macromolecular igual a la del agua.

- (c) Aproximadamente de 2 a 3kg de bacterias se alojan en el intestino grueso. Use la masa calculada en (a) para estimar la cantidad total de bacterias que habitan su intestino. Estime la cantidad total de células humanas en su cuerpo y compare las dos cifras. Para esto último considere que las células tienen un diámetro característico de $10\mu m$, la misma densidad que el agua y que $\sim 30\%$ de la masa de una persona es debida a la masa de las células que contiene.
- (d) Se afirma que en los $200m$ superiores de los océanos del mundo hay aproximadamente 10^{28} procariotas. Calcule el volumen total ocupado por estas células en m^3 y km^3. Calcule su espaciado medio. ¿Cuántas de esas células hay por mililitro de agua del océano? Considere que $\sim 2/3$ de la superficie terrestre está cubierta por agua.

2. Reproduzca paso a paso los cálculos de la subsección "Estimate: Sizing Up E. coli", pp 39-42 del libro "Physical Biology of the Cell" que se incluyen al final de la guía.

3. (*tomado del libro "Physical Biology of the Cell"*) Haga una estimación de la composición de carbono, hidrógeno, oxígeno y nitrógeno en la masa seca de una bacteria. Utilizando el conocimiento del tamaño y la masa de una bacteria, la fracción de esa masa que es "masa seca" (es decir, $\sim 30\%$) y los componentes químicos de una célula, calcule los números enteros pequeños aproximados (< 10) para el compuesto $C_mH_nO_pN_q$, es decir, encuentre m, n, p y q. Ver: http://book.bionumbers.org/what-is-the-elemental-composition-of-a-cell/

4. (*tomado del libro "Physical Biology of the Cell"*)

   - (a) Estime el volumen de los aminoácidos en nm^3. Para ello estime el de la glicina aproximándolo por un cilindro de largo $0.7nm$ y radio $0.16nm$. Estime el resto teniendo en cuenta la relación entre las masas correspondientes.
   - (b) Estime la masa de un aminoácido "típico" en Daltons. Justifique su estimación explicando cuántos átomos de cada tipo eligió. Compare su estimación con la masa real de varios aminoácidos clave como glicina, prolina, arginina y triptófano.
   - (c) Sobre la base de su resultado para el inciso b), deduzca una regla empírica para convertir la masa de una proteína (expresada en kDa) a un número correspondiente de residuos. Aplique esta regla a algunas proteínas del banco de datos de proteínas (e.g., miosina: https://www.rcsb.org/structure/1br2, actina: https://www.rcsb.org/structure/4BQL, hemoglobina: https://www.rcsb.org/structure/1SI4, alguna otra que quiera buscar) y compare sus resultados con el número real de residuos en cada una de estas proteínas.

5. El medio de crecimiento mínimo para bacterias como *E. coli* incluye varias sales con concentraciones características en el rango del mM y una fuente de carbono. La fuente de carbono es típicamente la glucosa y se utiliza al 0,50,5 g / 100 mL). Para el nitrógeno, el medio mínimo contiene cloruro de amonio (NH_4Cl) con una concentración de 0.1 g / 100 mL.

   - (a) Haga una estimación del número de átomos de carbono que se necesitan para formar el contenido macromolecular de una bacteria como *E. coli*. De manera similar, haga una estimación de la cantidad de nitrógenos que se necesitan para formar el contenido macromolecular de una bacteria. ¿Qué sucede con el fosfato?

- (b) ¿Cuántas células se pueden cultivar en un cultivo de 5 mL utilizando un medio mínimo antes de que se agote el carbono? ¿Cuántas células se pueden cultivar en un cultivo de 5 mL utilizando un medio mínimo antes de que se agote el nitrógeno? Tenga en cuenta que esta estimación será errónea porque ignora el costo energético de sintetizar las macromoléculas de la célula.

Taking the molecular census is also important because we will use our molecular counts in Chapter 3 to estimate the rates of macromolecular synthesis during the cell cycle. How fast is a genome replicated? What is the average rate of protein synthesis during the cell cycle and, given what we know about ribosomes, how do they maintain this rate of synthesis? A prerequisite to beginning to answer these questions is the macromolecular census itself.

Ultimately, to understand many experiments in biology, it is important to realize that most experimentation is comparative. That is, we compare "normal" behavior to perturbed behavior to see if some measurable property has increased or decreased. To make these statements meaningful, we need first to understand the quantitative baseline relative to which such increases and decreases are compared. There is another sense in which numbers of molecules are particularly meaningful which will be explored in detail in subsequent chapters that has to do with whether we can describe a cell as having "a lot" or "a few" copies of some specific molecule. If a cell has a lot of some particular molecule, then it is appropriate to describe the concentration of that molecule as the basis for predicting cellular function. However, when a cell has only a few copies of a particular molecule, then we need to consider the influence of random chance (or stochasticity) on its function. In many cases, cells have an interesting intermediate number of molecules where it is not immediately clear which perspective is appropriate. However, knowing the absolute numbers always gives us a reality check for subsequent assumptions and approximations for modeling biological processes.

Because of these considerations, much effort among biological scientists has been focused on the development of quantitative techniques for measuring the molecular census of living cells (both bacteria and eukaryotes). In this chapter, we will rely primarily on order-of-magnitude estimates based on simple assumptions. These estimates are validated by comparison with measurements. In subsequent chapters, these estimates will be refined through explicit model building and direct comparison with quantitative experiments.

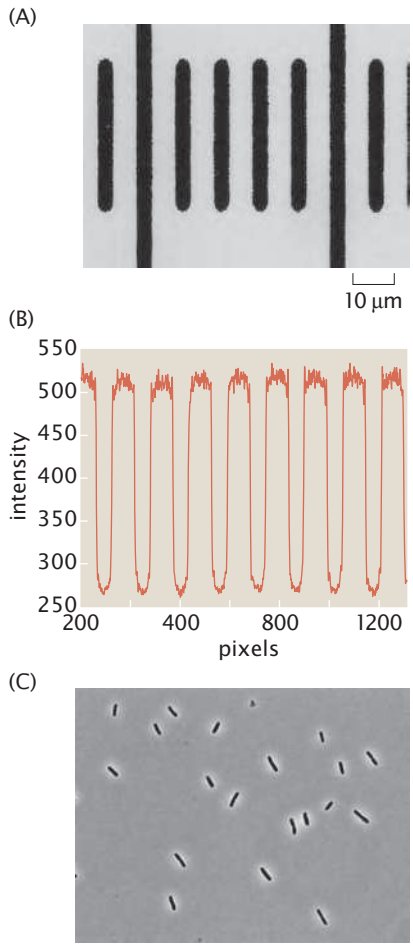

Figure 2.3: Sizing up *E. coli*. (A) Image of a graticule at 100× magnification. (B) Matlab plot of the intensity for a horizontal cut through the image. (C) Phase contrast image of a field of view with several *E. coli* cells taken at the same magnification as in (A).

Estimate: Sizing Up *E. coli* As already noted in the previous chapter, cells are made up of an array of different macromolecules as well as small molecules and ions. To estimate the number of proteins in an *E. coli* cell, we begin by noting that, with its 1 fL volume, the mass of such a cell is roughly 1 pg, where we have assumed that the density of the cell is that of water, which is 1 g/mL, though clearly Figure 2.2 shows that this assumption is not true. Measurements reveal that the dry weight of the cell is roughly 30% of its total and half of that mass is protein. As a result, the total protein mass within the cell is roughly 0.15 pg. We can also estimate the number of carbon atoms in a bacterium by considering the chemical composition of the macromolecules of the cell. This implies that roughly half the dry mass comes from the carbon content of these cells, a figure that reveals of the order of 10^{10} carbon atoms per cell. Two of the key sources that have served as a jumping off point for these estimates are Pedersen et al. (1978) and Zimmerman and Trach (1991), who describe the result of a molecular census of a bacterium.

ESTIMATE

As a first step toward revealing the extent of crowding within a bacterium, we can estimate the number of proteins

by assuming an average protein of 300 amino acids with each amino acid having a characteristic mass of 100 Da. These assumptions are further examined in the problems at the end of the chapter. Using these rules of thumb, we find that the mean protein has a mass of 30,000 Da. Using the conversion factor that $1\,\text{Da} \approx 1.6 \times 10^{-24}\,\text{g}$, we have that our typical protein has a mass of $5 \times 10^{-20}\,\text{g}$. The number of proteins per *E. coli* cell is estimated as

$$N_{\text{protein}} = \frac{\text{total protein mass}}{\text{mass per protein}} \approx \frac{15 \times 10^{-14}\,\text{g}}{5 \times 10^{-20}\,\text{g}} \approx 3 \times 10^{6}. \quad (2.1)$$

If we invoke the rough estimate that one-third of the proteins coded for in a typical genome correspond to membrane proteins, this implies that the number of cytoplasmic proteins is of the order of 2×10^6 and the number of membrane proteins is 10^6, although we note that not all of these membrane-associated proteins are strictly transmembrane proteins.

Another interesting use of this estimate is to get a rough impression of the number of ribosomes—the cellular machines that synthesize proteins. We can estimate the total number of ribosomes by first estimating the total mass of the ribosomes in the cell and then dividing by the mass per ribosome. To be concrete, we need one other fact, which is that roughly 20% of the protein complement of a cell is ribosomal protein. If we assume that all of this protein is tied up in assembled ribosomes, then we can estimate the number of ribosomes by noting that (a) the mass of an individual ribosome is roughly 2.5 MDa and (b) an individual ribosome is roughly one-third by mass protein and two-thirds by mass RNA, facts that can be directly confirmed by the reader by inspecting the structural biology of ribosomes. As a result, we have

$$\begin{aligned} N_{\text{ribosome}} &= \frac{0.2 \times 0.15 \times 10^{-12}\,\text{g}}{830{,}000\,\text{Da}} \times \frac{1\,\text{Da}}{1.6 \times 10^{-24}\,\text{g}} \\ &\approx 20{,}000 \text{ ribosomes.} \end{aligned} \quad (2.2)$$

The numerator of the first fraction has 0.2 as the fraction of protein that is ribosomal, 0.15 as the fraction of the total cell mass that is protein, and 1 pg as the cell mass. Our estimate for that part of the ribosomal mass that is protein is 830,000 Da. The size of a ribosome is roughly 20 nm (in "diameter") and hence the total volume taken up by these 20,000 ribosomes is roughly $10^8\,\text{nm}^3$. This is 10% of the total cell volume.

Idealizing an *E. coli* cell as a cube, sphere, or spherocylinder yields (see the problems) that the surface area of such cells is $A_{E.\,coli} \approx 6\,\mu\text{m}^2$. This number may be used in turn to estimate the number of lipid molecules associated with the inner and outer membranes of these cells as

$$N_{\text{lipid}} \approx \frac{4 \times 0.5 \times A_{E.\,coli}}{A_{\text{lipid}}} \approx \frac{4 \times 0.5 \times (6 \times 10^6\,\text{nm}^2)}{0.5\,\text{nm}^2} \approx 2 \times 10^7, \quad (2.3)$$

where the factor of 4 comes from the fact that the inner and outer membranes are each *bilayers*, implying that the lipids effectively cover the cell surface area four times. A lipid bilayer consists of two sheets of lipids with their tails pointing toward

each other. The factor of 0.5 is based on the crude estimate that roughly half of the surface area is covered by membrane proteins rather than lipids themselves. We have made the similarly crude estimate that the area per lipid is $0.5\,\mathrm{nm}^2$. The measured number of lipids is of the order of 2×10^7 as well.

In terms of sheer numbers, water molecules are by far the majority constituent of the cellular interior. One of the reasons this fact is intriguing is that during the process of cell division, a bacterium such as *E. coli* has to take on a very large number of new water molecules each second. The estimate we obtain here will be used to examine this transport problem in the next chapter. To estimate the number of water molecules, we exploit the fact that roughly 70% of the cellular mass (or volume) is water. As a result, the total mass of water is 0.7 pg. We can find the approximate number of water molecules as

$$\begin{aligned} N_{\mathrm{H_2O}} &\approx \frac{0.7 \times 10^{-12}\,\mathrm{g}}{18\ \mathrm{g/mol}} \times 6 \times 10^{23}\ \text{molecules/mol} \\ &\approx 2 \times 10^{10}\ \text{water molecules.} \end{aligned} \tag{2.4}$$

It is also of interest to gain an impression of the content of inorganic ions in a typical bacterial cell. To that end, we assume that a typical concentration of positively charged ions such as K^+ is 100 mM, resulting in the estimate

$$\begin{aligned} N_{\mathrm{ions}} &\approx \frac{(100 \times 10^{-3}\ \mathrm{mol}) \times (6 \times 10^{23}\ \text{molecules/mol})}{10^{15}\,\mu\mathrm{m}^3} \times 1\,\mu\mathrm{m}^3 \\ &= 6 \times 10^7. \end{aligned} \tag{2.5}$$

Here we use the fact that $1\,\mathrm{L} = 10^{15}\,\mu\mathrm{m}^3$. This result could have been obtained even more easily by noting yet another simple rule of thumb, namely, that one molecule per *E. coli* cell corresponds roughly to a concentration of 2 nM.

The outcome of our attempt to size up *E. coli* is illustrated schematically in summary form in Figure 2.4. A more complete census of an *E. coli* bacterium can be found in Neidhardt et al. (1990). The outcome of experimental investigations of the molecular census of an *E. coli* cell is summarized (for the purposes of comparing with our estimates) in Table 2.1.

How is the census of a cell taken experimentally? This is a question we will return to a number of different times, but will give a first answer here. For the case of *E. coli*, one important tool has been the use of gels like that shown in Figure 2.5. Such experiments work by breaking open cells and keeping only their protein components. The complex protein mixture is then separated into individual molecular species using a polyacrylamide gel matrix. First, the protein mixture is distributed through a tube-shaped polyacrylamide gel that has been polymerized to contain a stable pH gradient, and then an electric field is applied across the gel. The net charge on each protein depends on the pH and on the number and type of charged (protonatable) amino acid side chains that it contains. For example, the carboxylic acid group on aspartate will be negatively charged at high pH, but