the thickness is given by $h = V/A$ one finds that the height of a lipid molecule is a few nanometers. Irving Langmuir took these ideas much further in the early 20th century by measuring the relationship between the tension it takes to squish this surface layer and the area itself. These impressive measurements and inferences permitted Langmuir to figure out not only the length of different lipid molecules but also their cross-sectional area, the number of lipid tails per molecule and whether the bonds in those tails are saturated or unsaturated, inducing kinks.

Our "feeling for the organism" in the form of an intuition for molecular scales can be built around Figure 1.1 (p. 28) where we see the key classes of macromolecules of cells all depicted on a similar scale, illustrating the place of prominence of the nanometer scale when thinking about these macromolecules. Figure 1.3 (p. 31) gave us an impression of the molecular scales associated with the great molecule of heredity. Similarly, Figure 1.4 (p. 32) shows the amino acids, the building blocks of proteins with their characteristic nanometer scale dimensions and ≈ 100 Da mass. We see that all of the amino acids are essentially the same size.

Besides the scales set by the sizes of the molecules of life, perhaps the most important concept we can carry with us a scalpel of intuition is the idea of the thermal energy scale. The point is that as we shrink the physical dimensions of objects to ever smaller scales, their interactions with the surrounding medium (i.e. the solution that serves as their "thermal reservoir") become ever more important. Indeed, at molecular scales, the scaling rule of thumb that we will carry throughout the book is that each degree of freedom has an average of $k_BT \approx 4$ pN nm of energy. It is the random kicks to molecules due to this thermal energy to which we owe the Brownian motion (or diffusive motions) that will serve as biology's dynamical null hypothesis as will be explored in chap. 5 (p. 277).

2.4.2 Cellular Scaling

More than a century of cell biology, bolstered each generation by imaginative new microscopies with higher and higher resolution, has revealed a variety of quantitative trends relating the size of cells to the organelles within them. In this section, we explore a variety of different examples that give a sense of the kind of empirical scaling relations that have emerged from these studies, followed by a first pass at delving more deeply into the scaling of the nucleolus size in *C. elegans* embryos.

One of the signature features of cells is compartmentalization, a topic so important that we have devoted an entire chapter to it (chap. 14, p. 1161). One of the predominant examples of cellular compartmentalization is the organelles seen in Figure 2.43. Familiar names such as the nucleus, the endoplasmic reticulum, the Golgi apparatus and the mitochondria indicate distinct spatial compartments that are the seat of a myriad of different processes, including the ATP synthesis so central to the cellular economy found in the mitochondria.

Figure 2.44 provides a gallery of examples of how organellar size scales with

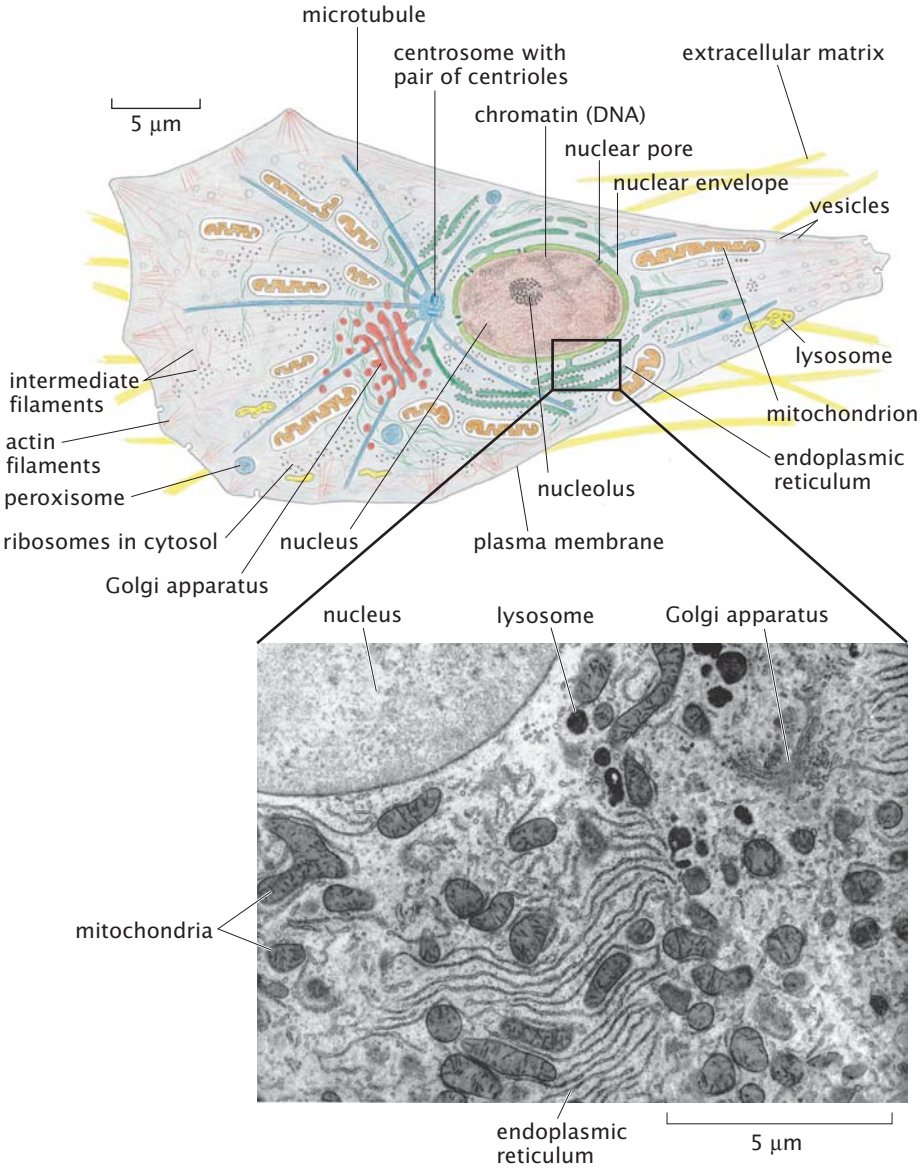


Figure 2.43: Eukaryotic cell and its organelles. The schematic shows a eukaryotic cell and a variety of membrane-bound organelles. A thin-section electron microscopy image shows a portion of a rat liver cell approximately equivalent to the boxed area on the schematic. A portion of the nucleus can be seen in the upper left-hand corner. The most prominent organelles visible in the image are mitochondria, lysosomes, the rough endoplasmic reticulum, and the Golgi apparatus. (Electron micrograph from D. W. Fawcett, The Cell, Its Organelles and lnclusions: An Atlas of Fine Structure. Philadelphia, W. B. Saunders & Co., 1966.)

cell size. For example, Figures 2.44(A) and (B) show how cell volume and the nuclear size in fission yeast are correlated. Those same kinds of scaling relations are shown in Figures 2.44(C) and (D) where we see how the mitotic spindle size scales with the size of the volume within which it is enclosed. In this case, the experiments were done in a biochemical setting within droplets, but similar scaling is observed in the in vivo context as well. Figures 2.44(E) and (F) bring similar sensibilities to bear on the question of the size of mitochondria in budding yeast. As with the previous two examples, we see that mitochondrial volume scales with cell size. The final example shown in Figures 2.44(G) and (H) reveal how the centrosome volume scales with the cell size. All of these examples reveal a compelling, reproducible example of cellular scaling that demand some deeper explanation. Rather, given our dictum that quantitative data demands quantitative models, these scaling results leave important theoretical challenges in their wake.

Not all organelles are membrane bound. Indeed, as we will see throughout the book, one of the most important discoveries in cell biology in the last few decades is the realization that there are whole classes of membraneless organelles. In the case of these organelles, such as the nucleoli shown in Figure 2.45, the compartmentalization emerges as a result of phase separation, a topic we take up in detail several times throughout the book. Before exploring the phenomenology of scaling for these organelles, we first remind ourselves of their important role in the lives of eukaryotic cells. Already in this chapter, we have seen the central role of ribosomes in bacterial cell biology and these important molecular machines are central to the behavior of eukaryotes too. The nucleolus can be thought of as a little localized ribosome factory within the nucleus.

Figure 2.45(A), shows that, during the early development of *C. elegans*, cells divide repeatedly without any change to the overall size of the embryo. Throughout development, the size of the nucleolus can be measured by fluorescently labeling a protein known as a well conserved part of the nucleolus such as fibrillarin-1 (FIB-1). The simultaneous measurements of cell, nuclear and nucleolar size reveals that, with each cell division, the nuclei become smaller as do their nucleoli, indicating a scaling of nucleolar size with nuclear size. This scaling is quantified in Figure 2.45(B), where see a linear relation between these two quantities.

A key feature of the experiment shown in Figure 2.45(A) and (B) is the fact that, while nuclear size changes, the concentration of nucleolar constituents such as FIB-1 remains constant. A second very clever experiment explores the question of nucleolar size by tuning different knobs. In the case shown in Figure 2.46(A), RNAi was used to knockdown specific genes that result in worm embryos of different overall sizes but having the same number of cells. Interestingly, while these genetic changes affect the overall embryo size, they leave the total number of nucleolar proteins per cell unchanged. As a result, the number of FIB-1 molecules (as opposed to its total concentration) is constant in all embryo mutants. As shown in Figure 2.46(B), in this case it is seen that the size of the nucleolus decreases with increasing nuclear volume.

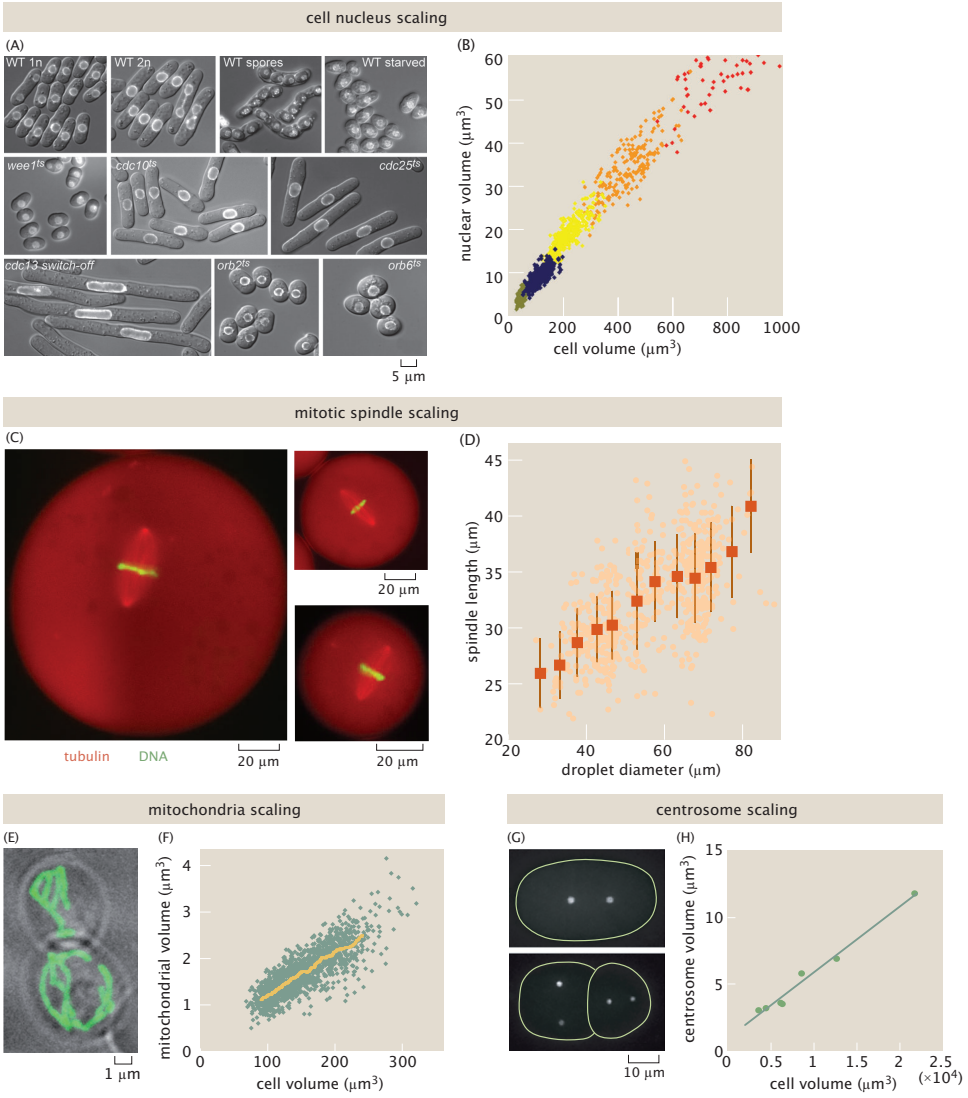


Figure 2.44: Scaling of organelle size with cell size. (A) Wild-type fission yeast cells grown under different conditions and mutant cells have different sizes, which can be quantified by the cell volume. (B) The volume of the nucleus of fission yeast cells is proportional to the cell volume; different colors correspond to different growth conditions and different mutants. (C) Mitotic spindles assembled in *Xenopus* egg extract confined to droplets of different size. (D) The size of the spindle, characterized by its length, is proportional to the diameter of the droplet. (E) Labeled mitochondria in a yeast cell. (F) Scaling of mitochondrial size with cell size. (G) Centrosomes in RP cells. (H) Scaling of centrosome volume with cell volume. **HG to RP: Need to add references**

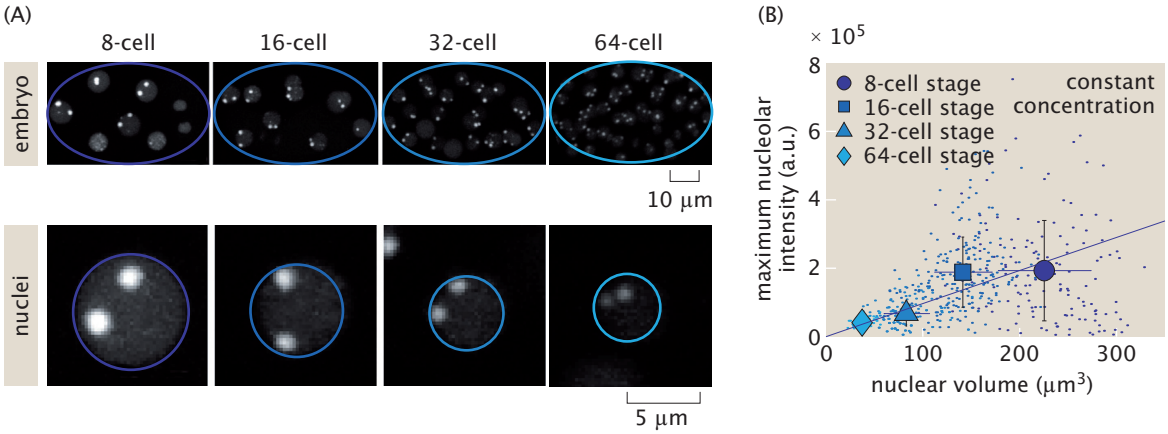


Figure 2.45: Scaling of nucleolar size in developing *C. elegans* embryos. (A) During the early development of the worm *C. elegans*, cells divide without any growth of the embryo. Therefore, cells become smaller with each cell cycle as shown at the top. The bottom row shows that the nuclei become smaller with each round of cell division and furthermore, that the fluorescence intensity of the nucleolar protein fibrallarin also decreases. (B) Plotting measurements for many cells from many embryos on the same axes we can appreciate that there is a roughly linear scaling between the volume of the nucleus and the intensity of the labeled nucleolar protein. (Adapted from S. Weber and C. Brangwynne, Curr. Biol. 25:641, 2015.) **HG: Add cartoon or snapshots of developing worms.**

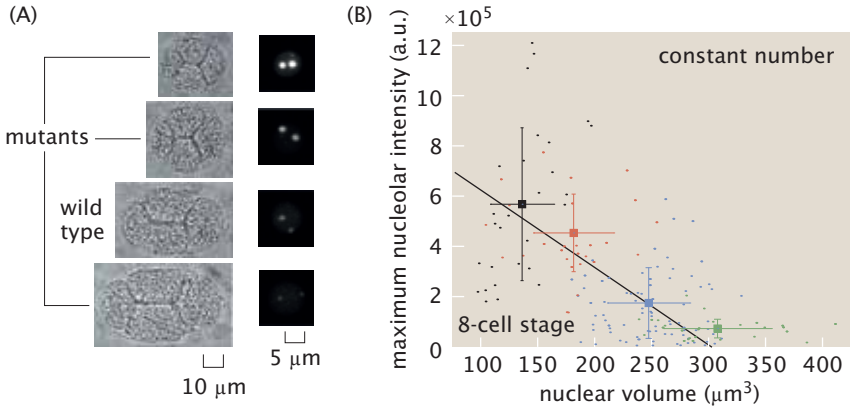


Figure 2.46: Engineering nucleolar scaling. (A) As an alternative to exploring how nucleolar size changes as cell size decreases during *C. elegans* development, cell size can be engineered by disruption of particular genes using RNAi. The images on the left show wild-type and mutant embryos at the 8-cell stage. Next to the transmitted light image of each embryo is the corresponding fluorescence image showing the intensity of fibrallarin in a single nucleus at that stage. (B) Plotting many measurements from many individual embryos, this time all at the 8-cell stage, but in these four different genetic backgrounds, we now see an inverse relationship between nuclear volume and nucleolar intensity. (Adapted from S. Weber and C. Brangwynne, Curr. Biol. 25:641, 2015.)

Figures 2.45 and 2.46 show quite distinct scaling results, both of which can be explained by a simple model of phase separation as will be shown later in the book in Chapter 7 (p. 521). For now, we content ourselves with a simple first introduction not only to nucleolar size, but also provide a first look at how we will use rate equations to describe a myriad of different phenomena.

We begin invoking a simple cartoon model of nucleolar formation which is shown schematically in Figure 2.47. Here, we assume that the nucleus of volume V contains a total of N FIB-1 molecules. M of these molecules are assembled into the nucleolus. FIB-1 molecules are constantly coming onto and off of the nucleolus, processes that we can describe using chemical rate equations. Specifically, individual FIB-1 molecules are incorporated into the nucleolus at a rate $k_{on}c_{free}$, where we define the concentration of free FIB-1 molecules as

$$c_{free} = \frac{(N - M)}{V}, \tag{2.17}$$

while they leave the nucleolus at a rate k_{off}. Note that as we will see repeatedly, k_{on} has units of $M^{-1}\ s^{-1}$ and k_{off} has units of s^{-1}. Given these definitions, we develop a simple intuitive kinetic model of the rate of growth of the nucleolus in the form

$$\frac{dM}{dt} = k_{on}\frac{(N - M)}{V} - k_{off}, \tag{2.18}$$

where the essence of the first term on the right is the idea that the rate of addition of molecules to the nucleolus is given by the on rate times the concentration of free nucleolar proteins which, in turn, is given by $(N - M)/V$. If we now solve for the steady-state value of the nucleolar size, by setting $dM/dt = 0$, we find the simple result

$$M = N - \frac{k_{off}}{k_{on}}V, \tag{2.19}$$

which can be rewritten as

$$M = \left(\frac{N}{V} - \frac{k_{off}}{k_{on}}\right)V. \tag{2.20}$$

We find it convenient to define the total concentration as $c_{tot} = N/V$ and the critical concentration above which one forms the nucleolus as $c_* = k_{off}/k_{on}$. In light of these definitions, the number of molecules in the nucleolus itself is given by

$$M = (c_{tot} - c_*)V. \tag{2.21}$$

We are now ready to bring our simple model to bear against the experimental data shown in Figures 2.45 and 2.46. For the case in which the total concentration is fixed, the nucleolar size is given by eqn. 2.20, revealing a linear scaling of that size with the volume of the nucleus. Interestingly, the slope of that linear relationship is dictated by the excess of the total concentration above the critical concentration c_*. Figure 2.45(B) shows the experimental relationship between nucleolus size and nuclear volume. In this case, with each successive

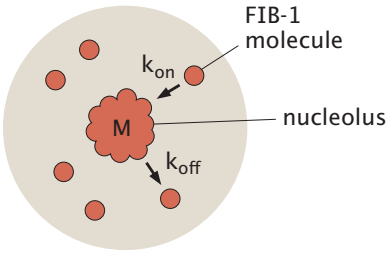


Figure 2.47: Simple model of nucleolar assembly. A nucleus of volume V contains a total of N FIB-1 molecules, M of which are assembled into the nucleolus. Single FIB-1 molecules are incorporated into the nucleolus at a rate k_{on}, and are separated from the nucleolus at a rate k_{off}.

cell division, the size of the cells and nuclei become smaller and hence the nucleoli become smaller as seen in the images of Figure 2.45(A) and as predicted by eqn. 2.20.

A more subtle effect was seen in the experiments reported in Figure 2.46(A). Here using RNAi to knockdown genes that control embryo size, an inverse nucleolar size scaling was found as revealed by Figure 2.46(B). This effect can also be understood by appealing to eqn. 2.19, where now what is fixed is N, meaning that for larger embryos, c_{tot} will be smaller implying smaller nucleoli as the nuclei get bigger. Specifically, this result becomes

$$M = N - c_* V, \tag{2.22}$$

implying the inverse scaling slope is $-c_*$ and the y-intercept is N.

The qualitative agreement between our model and the data shown in Figures 2.45 and 2.46 is certainly encouraging. However, we can go beyond these qualitative comparisons and explore how well theory and experiment agree quantitatively. Specifically, as noted above, eqn. 2.21 predicts that, in the case of fixed FIB-1 number, the y-intercept of the scaling relation will be given by the total number of molecules N, while the slope will be determined by the critical concentration c_*. By fitting a line to the experimental data, we can thus extract these two key model parameters, as shown in Figure 2.48(A), obtaining a y-intercept that implied $N = 7.5$ arbitrary fluorescence units, and a slope that indicates $c_* = 0.025$ a.u./μm^3.

Now, using these two parameters inferred for the case of fixed number of FIB-1 molecules, we seek to predict the outcome of the experiment performed at fixed FIB-1 concentration using eqn. 2.20. To make this possible, we need to know the values of c_{tot} and c_*. c_* was already determined directly form the slope of the fit shown in Figure 2.48(A). To calculate c_{tot}, we use the value of the total number of FIB-1 molecules given by N and inferred from the fit shown in Figure 2.48(A) and divide by the nuclear volume V at the 8-cell stage at which the measurements at fixed FIB-1 number were performed. This nuclear volume can be directly read out from the x-axis value of the last data point

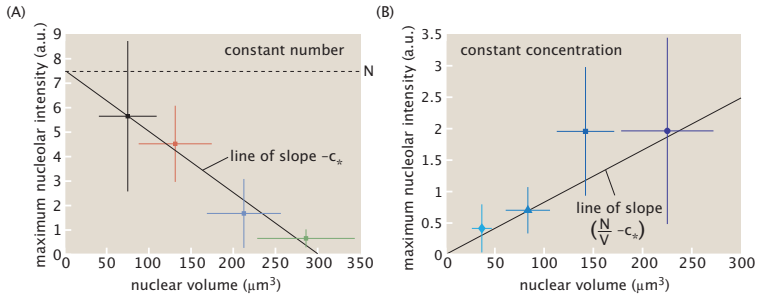


Figure 2.48: Testing the predictive power of the simple model of nucleolar assembly. By fitting the model to the data corresponding to a constant FIB-1 number, we obtain the total number of FIB-1 molecules N and the critical concentration c_*. (B) These parameters can be used to calculate $c_{tot} - c_*$ for the experiment performed at fixed FIB-1 concentration in order to predict the nucleolar scaling with nuclear volume without any free parameters.

in Figure 2.45(B). As a result, we now have the two parameters necessary to predict nucleolar volume as a function of nuclear volume without the need to invoke any other free parameters. The outcome of this calculation is shown in Figure 2.48(B) where we see that, indeed, this model can predict the nucleolar scaling under the conditions of fixed FIB-1 concentration. Thus, the case study of the scaling of nucleolar volume with nuclear volume not only gives a flavor of the interesting phenomenon of organellar size scaling, but also demonstrates how we can write down simple models that help us develop intuition for, and predictive power over, the of cellular structures scaling.

RP: extra topics to come back to. same size salamander, but different cell sizes RP: smallest bacterium can't be smaller than such and such cell size and DNA content ATP synthase scaling - this is where we then introduce surface to volume Scaling of spindle size - Rebecca Scaling of amphibian - Evy and Heald 2015

Our discussion thus far focused on the way that organelles within cells scale with cell size. Another fascinating kind of scaling shows up in thinking about embryogenesis as we will discuss in greater detail in chap. 15. Specifically, as seen in Figure 2.49, the body plans of different species of *Drosopholids* (i.e., fruit flies) scale proportionally so that the structures of the body plan occur in the right places. One of the key mechanisms for establishing body plans that we will discuss later in the book is positional information, the idea that some absolute coordinate system is laid down in the embryo giving cells an "address" within the embryo. Examples of these morphogen gradients in several species of fruit fly are shown in Figure 2.49. Mathematically, the results of these experiments can be visualized as shown in Figure 2.50. This figure provides a first introduction to an approach that we will use again and again throughout the book, namely, the method of data collapse. Here we see that if we use the "natural variables" of the problem, the concentration normalized by its maximum value and the position normalized by the length of the embryo, then all the morphogen gradients collapse onto a single master function.

RP: this is an orphan figure that needs to be moved somewhere else

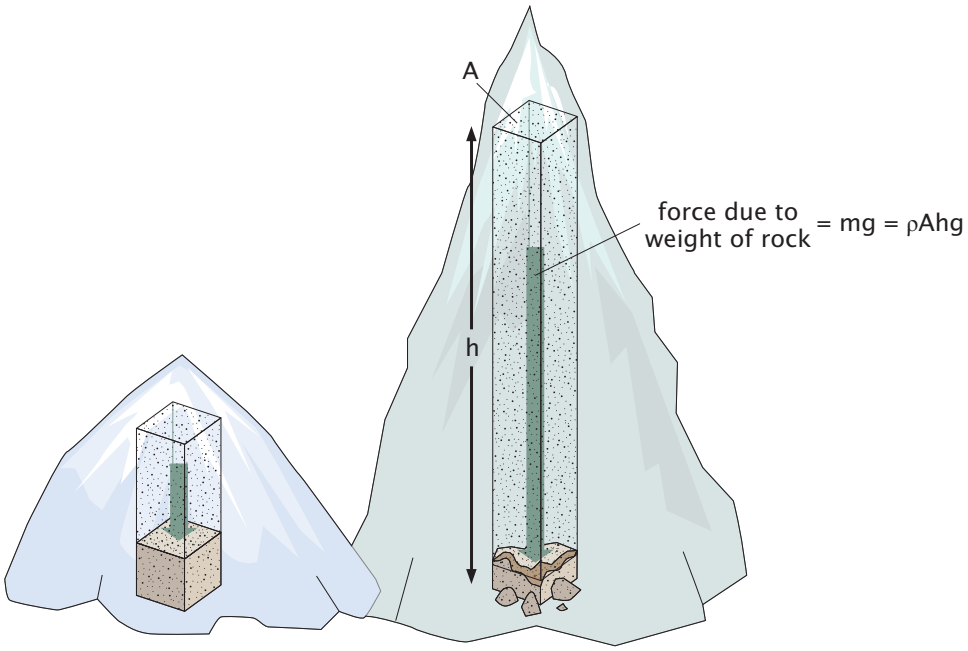


Figure 2.55: Estimating the height of mountains. A mountain can be thought of as a column of rock that is a weight sitting on its base. When the mountain gets sufficiently high, the weight of that column of rock is sufficient to crush the rock at its base.

2.4.4 Mechanics and Organism Size

Like all material objects, biological systems have mechanical limits beyond which they will fail. A beautiful and instructive example that will help us understand mammalian body plans and teach us about the kind of scaling arguments that are one of the central preoccupations of this chapter is an estimate of the height of mountains. How might such a thing help us understand mammalian body plans? Because for legs and mountains alike, there is a critical weight of material that will induce the structures that support them to fail as shown in Figure 2.55.

Let's consider the case of a mountain. We can think of the mountain as a giant weight that is resting on the rock at its base. A crude estimate of the force due to that weight is given by

$$\text{weight of column of rock} = mg = \rho Ahg. \tag{2.23}$$

To turn this weight into a criterion for failure we now need to invoke a fundamental material parameter that has emerged from generations of experimentation and phenomenological thinking, namely, the compressive yield stress. The compressive yield stress attempts to capture in a single number how much stress is

needed to cause the material to permanently deform. Anyone who has stepped on a piece of chalk has seen the compressive yield stress in action, and clearly this material parameter is different for chalk than for granite! A characteristic scale for rock is $\sigma_Y \approx 100$ MPa.

If our readers are anything like us, they may have a residue of confusion about the difference between stress and strain. As we will see in later chapters, strain is a geometric idea that tells us how a body has been deformed. Specifically, strain tells us how different parts of a material move relative to each other upon deformation of the material. By way of contrast, stress tells us about the force acting on a unit area of some surface within the material, hence with units of force/area $= N/m^2$.

To implement the failure criterion for the rock, we search for the condition at which the stress due to the weight of the rock is equal to the failure stress of the rock material itself, namely,

$$\sigma_Y = \frac{\rho A h g}{A} = \rho h g \tag{2.24}$$

implying that the height of the mountains is

$$h = \frac{\sigma_Y}{\rho g}. \tag{2.25}$$

Given this result, we can now estimate the height of mountains on planet Earth (and other planets as well if we replace our g with the gravitational acceleration at the surface of those planets). Specifically, we have

$$h = \frac{\sigma_Y}{\rho g} \approx \frac{100 \times 10^6 Pa}{3000 kg/m^3 \times 10 m/s^2} \approx \frac{10^4}{3} m! \tag{2.26}$$

For such a crude estimate, we have arrived to within roughly a factor of two of the observed maximum height of mountains on planet earth, and further, learned how the heights of mountains *scale* with the strength of gravity on a given planet. Note that our result can be recast in the form of a dimensionless number,

$$Cr = \text{crushing number} = \frac{\sigma_Y}{\rho h g}. \tag{2.27}$$

When $Cr > 1$, the rock can sustain the load and when $Cr < 1$, it can't.

Our estimate of the height of mountains offers more than a fun diversion in geophysics. Exactly this same kind of thinking comes into play in trying to understand animal body plans as shown in Figure 2.56. What is it that makes the legs of an antelope or gazelle so different from those of an elephant? As with the height of mountains, the answer lies in the analysis of mechanical instability. In the case of the mountains, the critical stress corresponded to the compressive yield stress. For the case of the leg, we will evaluate the propensity for a one-dimensional rod to undergo a buckling instability.

Figure 2.56 presents the key parameters that we suspect will dictate the buckling force for an elastic rod. The physics behind these choices will unfold

in detail in chap. 10. First, we see that the critical force should depend upon the elastic properties of the one-dimensional rod as measured by the Young modulus. The Young modulus tells us how the stress (σ = force/area) and the strain (ε= change in length/length) relate to each other via Hooke's law, $\sigma = E\varepsilon$. All of us have the intuition that the thicker a rod is, the harder it is to bend. As we will explain in chap. 10, that intuition is formalized in the areal moment of inertia, I, a geometric quantity with dimensions L^4 that reveals that the force required to achieve beam bending scales as the fourth power of the radius of that beam. The other critical parameter that will govern the buckling force is the length of the beam. Here too, our intuition playing with straws should remind us that if we create a double straw that is twice as long as an ordinary straw and stand it on a table, the force required to buckle the straw changes (and is reduced).

As shown in the right column of Figure 2.56, we are interested in the buckling force. Recall through Newton's second law of motion ($F = ma$) that the dimensions of force are $[F] = ML/T^2$. We hypothesize that the buckling force takes the form

$$F_{buckle} = E^{\alpha} I^{\beta} \ell^{\gamma}. \tag{2.28}$$

We see directly that E is the only quantity that features both M and T, requiring that $\alpha = 1$. Given this choice, we then require that $2 = 4\beta + \gamma$. We expect that $\gamma < 0$ since larger ℓ corresponds to smaller buckling force. The simplest choice corresponds to $\beta = 1$ which implies that $\gamma = -2$. Intuitively, the resulting buckling force scales as we might expect: larger cross sectional radius of legs (as imposed by I) leads to larger buckling forces and the longer the leg (as measured by ℓ), the smaller the buckling force.

RP: need to figure out how this whole story depends upon animal size. Galileo made estimates. See Steven Vogel on Biomechanics.

RP: decide later if we keep here or make a problem

Animals are not the only living organisms that construct structures that must support mechanical loads. Indeed, the vast majority of biomass on planet Earth is associated with plants, and giant Sequoias, redwoods and baobabs remind us of the mechanical stresses faced by these wooden giants.

Our scaling estimates tell us that, as with the problem of the heights of mountains, the essential comparison is between the weight of the structure itself and the critical force that will engender an instability. Here we perform a naive estimate of the buckling of a tree trunk by imagining that *all* of the weight of the trunk is applied at the top of the trunk, leading to a criterion for buckling of the form

$$\text{trunk weight} = \text{buckling force} \tag{2.29}$$

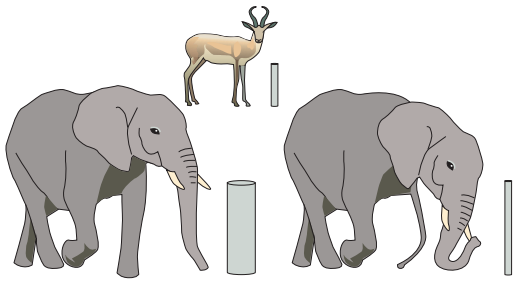
which we can rewrite as

$$\rho h r^2 g = E\frac{r^4}{h^2}, \tag{2.30}$$

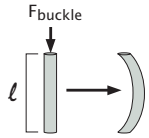
where we have suppressed all numerical factors due to the cylindrical geometry of the tree trunk. Solving this equation for the tree height yields

$$h = \left(\frac{Er^2}{\rho g}\right)^{1/3} = \left(\frac{10^{10}\text{N/m}^2(5\ m^2)}{600\ kg/m^3 10\ m/s^2}\right)^{1/3} \approx 215\ m, \tag{2.31}$$

PARAMETERS INFLUENCING THE BUCKLING FORCE

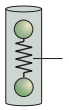


how much force does it take to buckle a beam of length ℓ?



relevant parameters

1. beam material, described by Young modulus E



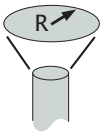
Young modulus E
units of force/area, $\frac{ML}{T^2}$

2. beam geometry, described by areal moment of inertia I

radius has units of L

$I = \frac{\pi R^4}{4}$

areal moment of inertia has units of L^4



3. beam length ℓ, units of L

ESTIMATING THE BUCKLING FORCE BY DIMENSIONAL ANALYSIS

unknown exponents

$$\underbrace{F_{buckle}}_{\frac{ML}{T^2}} = \underbrace{E^{\alpha}}_{\left(\frac{M}{LT^2}\right)^{\alpha}} \underbrace{I^{\beta}}_{L^{4\beta}} \underbrace{\ell^{\gamma}}_{L^{\gamma}}$$

only E includes mass, so $\alpha = 1$

$$\underbrace{F_{buckle}}_{\frac{ML}{T^2}} = \underbrace{E}_{\frac{M}{LT^2}} \underbrace{I^{\beta}}_{L^{4\beta}} \underbrace{\ell^{\gamma}}_{L^{\gamma}}$$

to balance units $\beta = 1$, $\gamma = -2$

$$F_{buckle} = \frac{EI}{\ell^2}$$

Figure 2.56: Estimating the buckling force of animal legs. The left panel hypothesizes the parameters that influence the buckling force of a one-dimensional beam subjected to a compressive load. The right panel shows how scaling arguments can be used to estimate the buckling force. **HG: Note that, as presented, β and γ are not uniquely defined.**

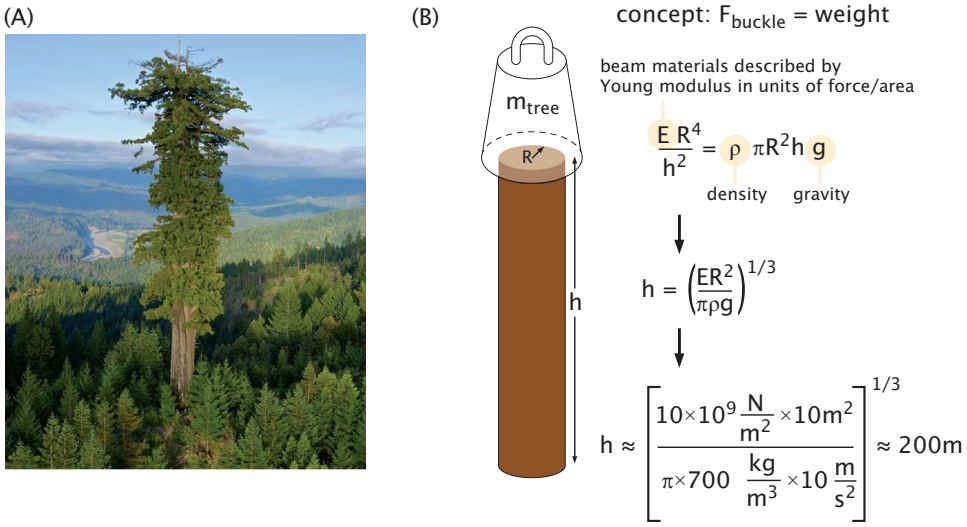


Figure 2.57: Estimating the buckling force of tree trunks under their own weight. (A) The tallest tree in the world, *Hyperion*. (B) Estimating the buckling force for a tree under the force of its own weight on the naive assumption that the entirety of its mass is concentrated at the top of the trunk.

where we used a diameter for a Redwood of roughly 4 m. This naive estimate is to be compared with the height of the tallest tree known on Earth, a lonely redwood named Hyperion, with a towering height of 115 m. Note that this analysis reflects but one possible hypothesis for what limits the height of trees in much the same spirit that earlier we outlined different hypotheses for what controls the growth rate of bacterial cells. An alternative hypothesis for tree heights is that there is a certain height above which water could not be brought to the top of the trees.

RP: scaling across species Thomas Gregor stuff, Cassandra Extavour the rules of scaling in body plans - paper 2019

RP: Species-area relationships island biogeography, the great human experiment

2.4.5 Scaling in Ecosystems

RP: not yet started

2.5 Dimensional Analysis

The examples above of estimating the heights of mountains and the buckling force for animal legs are specific examples of a much broader principle exploited in physics, namely, dimensional analysis. One of the most brilliant practitioners of the scaling arts was Lord Rayleigh, who used scaling arguments to work out the strength of bridges, the velocity of surface waves on water, vibration of tuning forks and drops of falling water, the color of the sky, the decay of charge of an electrical circuit and many others. On the subject of dimensional analysis,

he had this offering: "I have often been impressed with the scanty attention paid even by original workers to the great principle of similitude. It happens not infrequently that results in the form of laws are put forth as novelties on the basis of elaborate experiments, which might have been predicted a priori after a few minutes consideration." In this part of the chapter, we explore this great principle of similitude.

2.5.1 Bird Migrations and Energy Expenditure

Long-distance migrations are one of the greatest wonders of the natural world. The issues of navigation that arise in contexts ranging from monarch butterflies to salmon to the bar-tailed godwit with its 10,000 km voyage from Alaska to New Zealand offer a picture of nearly zero tolerance for error. In addition to the navigational challenges, unlike humans, many animals are subject to huge changes in their body mass as a normal part of their lifestyle. For the bar-tailed godwit, there have been claims that these migratory birds lose 1/3rd of their body mass during their 10-day uninterrupted journey. In this section, we set the stage for understanding such loss of body mass by exploring the work these birds need to do to overcome the drag force due to the surrounding air.

Figure 2.60 sketches how we can use arguments from dimensional analysis to evaluate the drag force. In the left column of the figure, we amass the various physical parameters that we imagine will dictate the drag force with their attendant units. The right column then does exponent balancing to try to figure out what combination of density, speed and bird length scale together conspire to give rise to the drag force. As seen in the figure, the dimensions of force are $M\ L/T^2$. Since on the right side the only quantity that features the mass is the density, we see that the density must enter the problem to the first power. Now like falling dominoes, we see that this in turn implies that the drag force must scale as the speed squared since speed is the only quantity involving time which appears in the denominator to the second power. The final domino to fall then is the recognition that the drag force must scale as the second power of the length scale of the bird.

2.5.2 The Length Scale of Morphogen Gradients

The point of the kinds of estimates we are exploring in this section is to develop facility with dimensional analysis and order-of-magnitude thinking. As a second example, we consider the way that animal body plans are laid down through gradients of transcription factors. The classic example of this phenomenon is offered by the Bicoid gradient in the early fruit fly embryo as hinted at in Figure 2.61. As seen in the figure, mechanistically, the morphogen gradient is set up as a competition between two distinct effects. At the anterior end of the embryo, the mother has deposited the mRNA that will be used to translate the Bicoid protein. Once synthesized, these proteins can diffuse, exercising a random walk which we idealize as a strictly one-dimensional process. However, at each instant, the protein is also subject to degradation with a rate of approximately

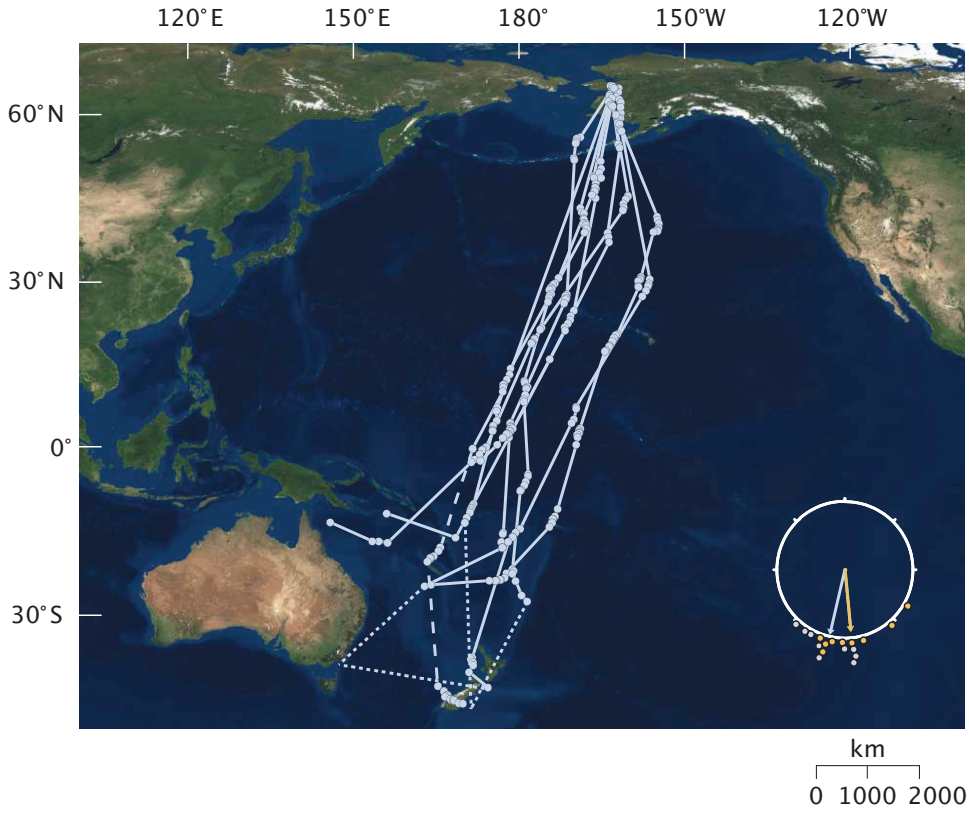


Figure 2.58: GotwidFlightMap.pdf. (Adapted from R. E. Gill Jr et al., *Proc. R. Soc. B* 276:447–457, 2009.)

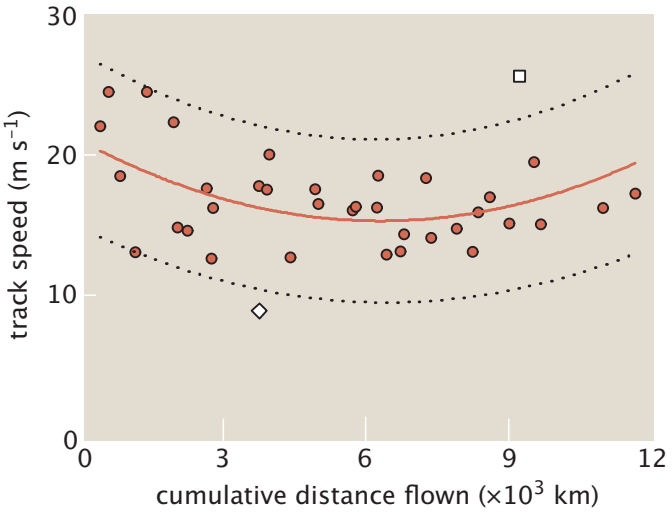


Figure 2.59: GotwidFlightSpeed.pdf. (Adapted from R. E. Gill Jr et al., *Proc. R. Soc. B* 276:447–457, 2009.)

$1/\tau$, where τ is the degradation time. This means that the farther towards the anterior a protein is, the older it is and thus those proteins will have had longer to be degraded. This interplay between diffusion and degradation sets a characteristic length scale which we can find through dimensional analysis, though later in the book we will work it out by solving the relevant partial differential equation known as the reaction-diffusion equation.

We now follow the procedure in Figure 2.61 to work out the length scale of the morphogen gradient. Once again, in the left column of the figure, we amass the various physical parameters that we imagine will dictate the length scale adopted by the morphogen gradient, namely, the diffusion constant which determines how quickly the proteins wander from their point of translation at the anterior end of the embryo and the degradation time, which provides a measure of the average survival time of a protein after synthesis. We posit then that the length scale is given by $\ell_{morphogen} = D^{\alpha}\tau^{beta}$ and use the process of elimination to figure out these exponents. For example, since D is the only quantity harboring a length, we must choose the exponent $\alpha = 1/2$. Like in our previous example, then like falling dominoes, we have no choice but to choose $\beta = 1/2$ as well. As advertised by the Lord Rayleigh quote about the great principle of similitude, a very modest effort has taken us all the way to the answer of the critical "what sets the scale of X?" problem.

2.5.3 The Buckingham Pi Theorem

The strategy that unfolded in the case studies given above are all specific examples of a much more general idea known as the Buckingham Pi theorem. This

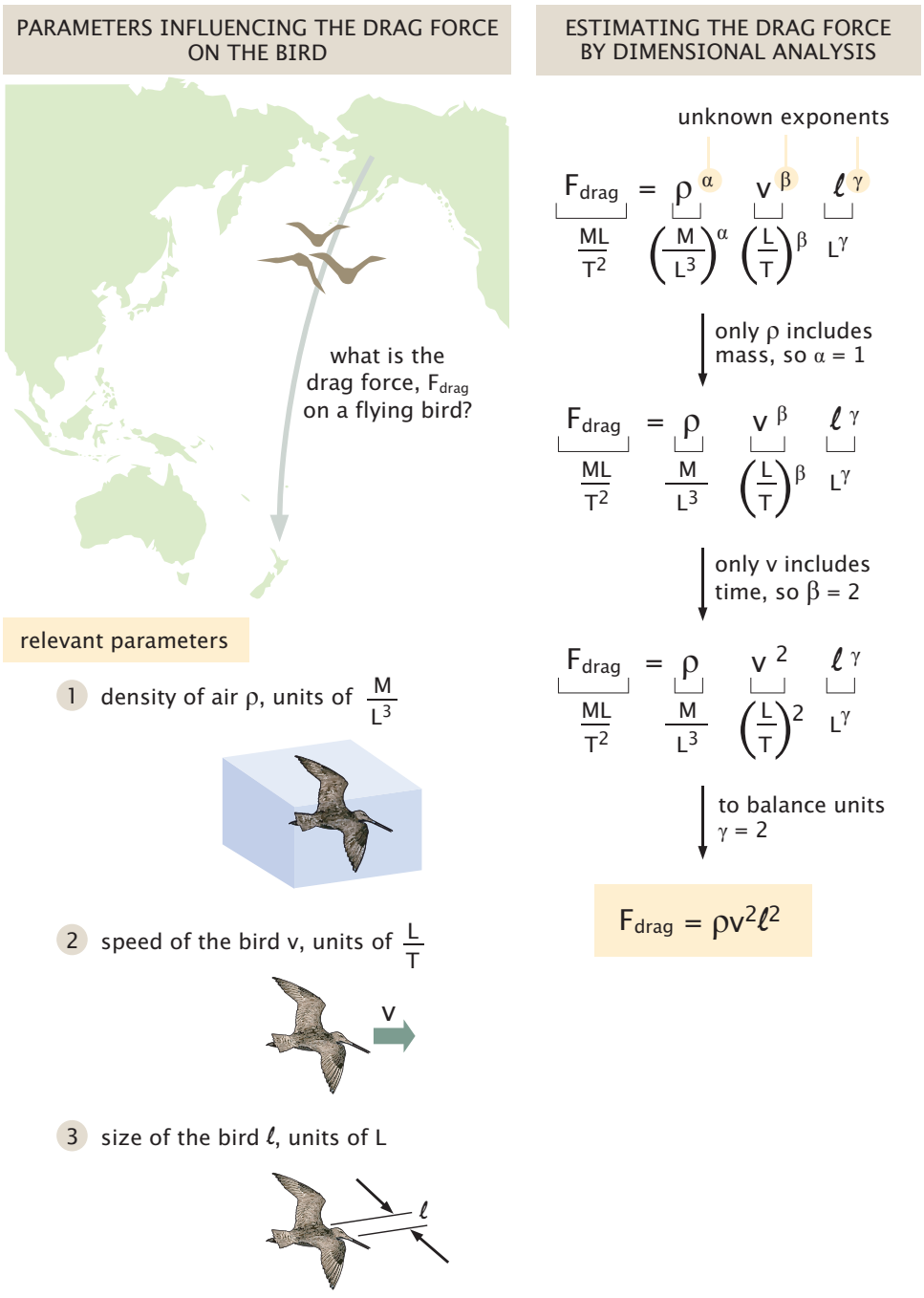


Figure 2.60: Dimensional analysis argument to determine the drag force on migrating birds. The left column shows the parameters we hypothesize determine the drag force on a flying bird. Each such parameter has a particular set of units. The right column shows how the relevant parameters can be assembled to differing powers resulting in a quantity with units of force.

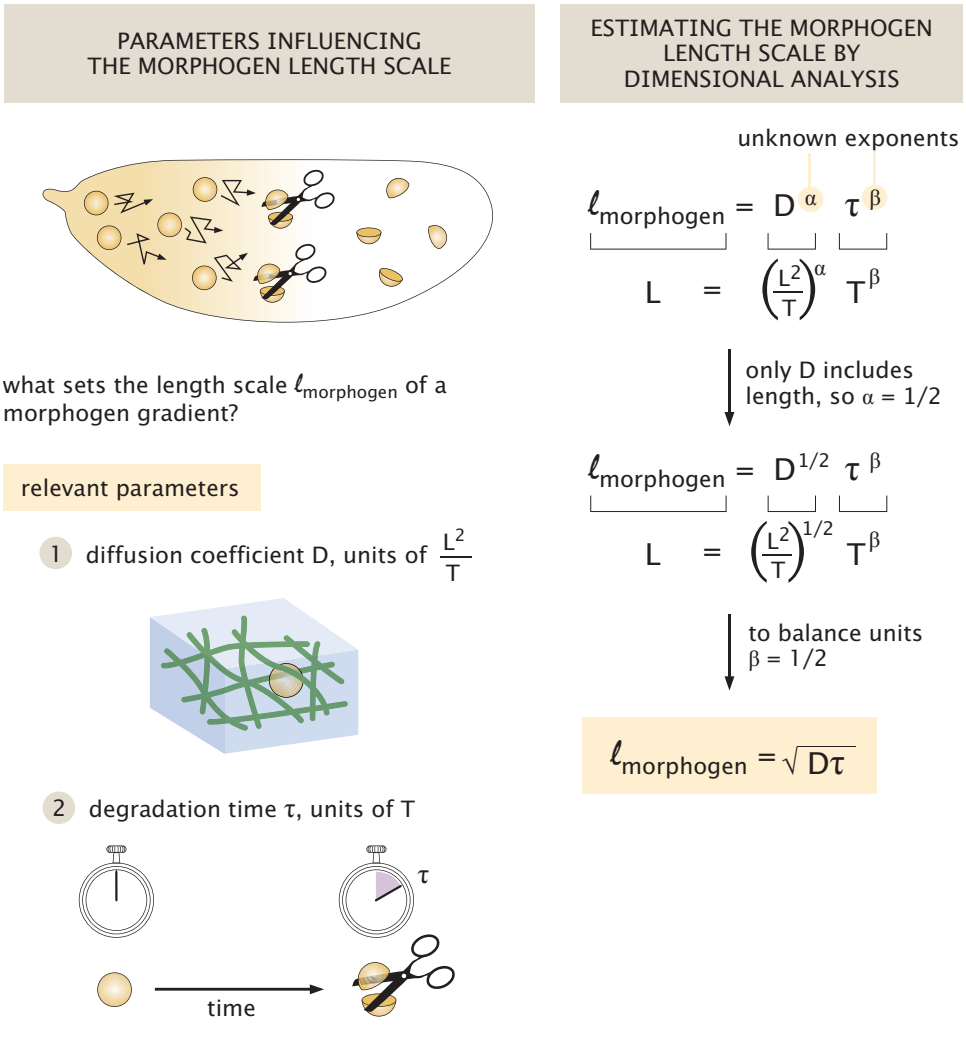


Figure 2.61: The length scale of morphogen gradients. The left column shows the relevant governing variables that impact the morphogen length scale. The right column shows how balancing dimensions determines the exponents in the scaling relation. **HG: Shall we include an argument for why the production rate cancels out? I think something could be done in terms of the production terms being number of molecules/time.**

theorem is in some ways a mathematical statement from linear algebra about matrices and vectors. However, for our purposes we are not going to think of it as a mathematical result, but rather focus on the intuition it brings as a method of reasoning about the parameters that set the scale of a given process or phenomenon and how they combine together to yield observables. At the outset, we need to make sure the reader understands the difference between *units* and *dimensions*. The units refer to the particular subjective choice a given scientist makes about their units of mass (i.e., kilograms or grams, etc), length (i.e., meters or feet, etc.), time (i.e., seconds or minutes), temperature (i.e., degrees Kelvin or centigrade) and so on. The dimensions, on the other hand, tell us about the intrinsic physical makeup of a given quantity. Acceleration has dimensions of L/T^2, no matter what units we decide to use. The Buckingham Pi theorem allows us to formalize the relation between the different parameters that appear in our analysis of a given problem strictly on the grounds of making sure that the dimensions are consistent.

Part of the art of determining the relevant dimensionless parameters that govern a given phenomenon is making the right guesses about what governing variables influence the process. For example, in Figure 2.61, we made the guess that only two governing variables influence the length scale of morphogens in developing embryos, namely, the diffusion coefficient and the protein degradation time. To give the reader an idea of other examples, Figure 2.62(C) shows a list of governing variables that influence phenomena ranging from the drag experienced by migratory birds to the buckling force of animal legs and tree trunks. The art of doing dimensional analysis is to make the right guesses about the parameters in the central column (governing variables) that determine our process of interest. For the flying bird, in Figure 2.60, we made the guess that the bird's speed, the bird's size and the density of the air were the governing variables of interest. However, for the motion of *E. coli* when swimming, our guess will be different and instead will argue that the drag force depends upon the size of the bacterium, the viscosity of the medium through which it is swimming and its speed. Once we have assembled this set of governing variables, the Buckingham Pi theorem tells us how many dimensionless parameters it takes to describe the given problem of interest.

Specifically, the theorem of Figure 2.62 tells us that if we want to figure out how many different dimensionless variables (N_p) characterize a given problem of interest, we can determine this through the relation

$$N_p = N_v - N_D, \tag{2.32}$$

where N_D is the number of physical dimensions (i.e. M, L, T or θ (temperature)) that appear in the problem and N_v is the number of control variables that appear in the problem. To get a sense of the power and insight that derive from this theorem, we appeal to the revolutionary studies of Galileo that helped usher in the modern era in science. While still a very young man, he apparently entered a cathedral in Pisa and was intrigued by the idea of figuring out the period of oscillation of a chandelier. Later, while laboring intensively to understand mechanics, focused on the "toy problem" of projectile motion as shown

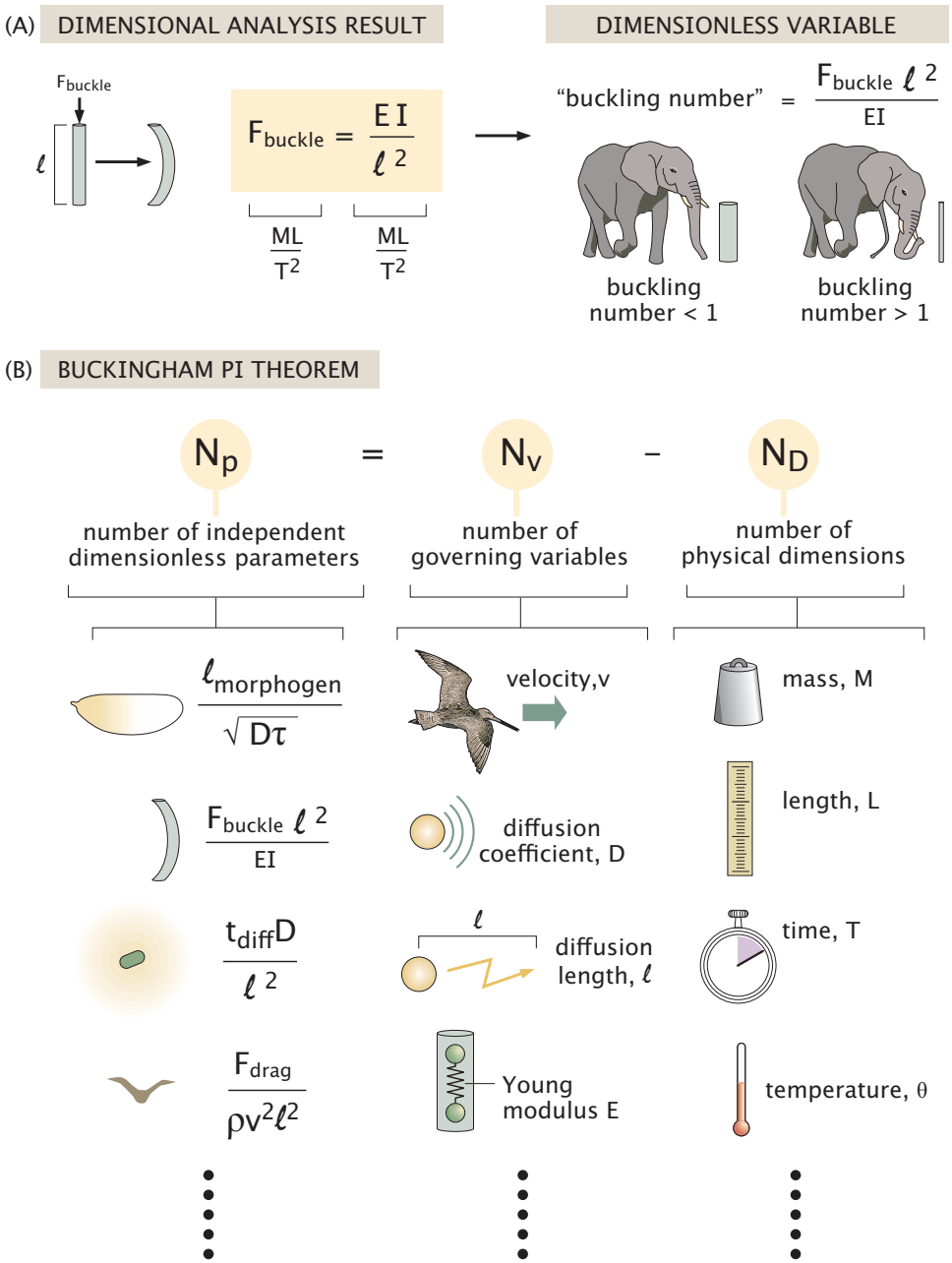


Figure 2.62: The Buckingham Pi Theorem. (A) Dimensional analysis result for the critical force for buckling. Construction of a dimensionless ratio to characterize the propensity for a one-dimensional rod to buckle as applied to the case of animal legs. (B) The Buckingham Pi theorem relates the number of independent dimensionless parameters to the number of variables needed to describe the problem of interest and the number of distinct physical dimensions. Each column gives examples of each of these categories. There is no significance to the particular rows. Only the columns in the diagram have a classification significance.

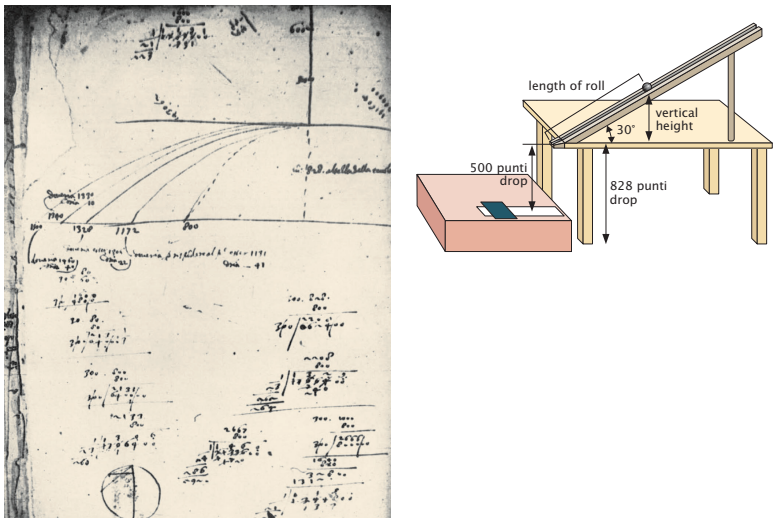


Figure 2.63: Galileo's studies on parabolic motion. The page from Galileo's notebook reveals him discovering the famed $s = 1/2gt^2$ result. The schematic on the right illustrates how Galileo was able to stop the projectile at various points in its trajectory by adjusting the height of the platform which the projectile landed on.

in Figure 2.63, he worked out the laws that relate the speed of the projectile, the distance it travels and the acceleration due to gravity, g. These two classic case studies which were at the center of the birth of modern mechanics can both be explored in the context of the Buckingham Pi theorem. For the case of the pendulum, we are interested in how it's period depends upon the length of the pendulum, ℓ, the acceleration due to gravity g (i.e., we would speculate that the period would differ on the moon for the same pendulum), the initial angle θ_0 and the mass m. According to the Buckingham Pi theorem, we have five governing variables (τ, ℓ, g, m and θ_0) and three (M, L and T) physical dimensions, so we expect two dimensionless governing groups. For the pendulum, we can directly see that the only parameter involving dimensions of time is g and thus we can guess the relation

$$\tau = \sqrt{\frac{\ell}{g}} f(\theta_0). \tag{2.33}$$

Similar reasoning can be applied to the question of the distance traveled by a projectile like that in Galileo's experiment shown in Figure 2.63. We leave the detailed analysis of both of these problems using the Buckingham Pi theorem to the reader in the problems at the end of the chapter. On dimensional grounds alone, we can already say much about both the mechanics of the pendulum and projectile motion without any detailed mechanics calculations.

RP: put problem on both pendulum and projectile

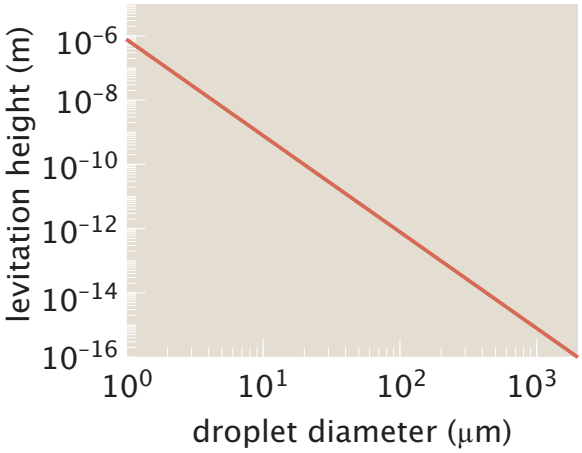


Figure 2.64: DropletBuddhaNumber.pdf

Another way of thinking about the insights born from the Buckingham Pi theorem will be of great value through the entirety of the book and indeed represent one of the foundational objectives of the physical biology mindset, namely, the quest to find what we will call the "natural variables" of a problem. The way we will explain this sometimes will be through the distinction between what we will refer to as "pipettor's variables" and "natural variables." For example, when we consider ligand-receptor binding, we can, of course, consider the pipetting of our ligand at some concentration c. However, this is not the concentration that the receptor itself "cares" about. The receptor's response is based on the dimensionless variable c/K_d, the normalized concentration. When the occupancy probability is plotted as a function of this dimensionless concentration for simple binding, all receptors fall onto one universal curve, belying the idea that each receptor type is its own new intellectual adventure. Over and over, we will attempt to understand what the natural variables are for a given problem, and often, that act will lead to a much more sophisticated intuition.

2.6 Surface-to-Volume Problems

2.6.1 ATP Synthesis

An abiding example of Rayleigh's principle of similitude that shows up in scientific contexts of all kinds is the interplay between surface area and the volume it encloses. The surface-to-volume ratio permits us to understand diverse phenomena such as the critical size of a volume of material in order to overcome a nucleation barrier, the relationship between the volumetric metabolic processes and the ability of the surface to carry away heat and so on. Here we use such scaling ideas to formulate a hypothesis about energy production in cells and how such production might imply the necessity for mitochondria once cells get

above a certain critical size.

As we saw earlier in the chapter in the context of bacterial growth, rapidly dividing bacteria require a power density of roughly 10^6 ATP/$\mu m^3 s$. But recall, all of this ATP is synthesized on the bacterial surface through the action of the ATP synthases as shown in Figure 2.65. The figure raises a very interesting question. How large would a spherical cell have to be before the ATP synthases on the surface are not able to keep up with the demands of the interior volume for ATP?

The key point for the purposes of a scaling estimate is that each ATP synthase contributes roughly 100 ATP/s. We can formalize our thinking by introducing the power density

$$P_v = 10^6 \frac{\text{ATP}}{\mu\text{m}^3\text{s}}, \tag{2.34}$$

which is the power consumed per unit volume. The corresponding maximal power produced per unit area of membrane is given by

$$P_s = 3 \times 10^6 \frac{\text{ATP}}{\mu\text{m}^2\text{s}}, \tag{2.35}$$

where as shown in the right panel of Figure 2.65, this production rate corresponds to packing the membrane as fully as possible with ATP synthases, an unrealistic proposition given the need for other membrane proteins as well. We can use these two numbers to find the critical cell size at which the surface production can't keep up with the volume consumption as

$$\frac{4}{3}\pi R^3 P_v = 4\pi R^2 P_s, \tag{2.36}$$

which implies that the critical size is given by

$$R = 3\frac{P_s}{P_v} = 3\frac{3 \times 10^6 \frac{\text{ATP}}{\mu m^2 \text{S}}}{10^6 \frac{\text{ATP}}{\mu m^3 \text{S}}} \approx 10\ \mu\text{m}. \tag{2.37}$$

This very interesting estimate argues that any cell greater than 10 μm in radius (and actually less because the membrane has to accommodate other membrane proteins as well) cannot produce ATP fast enough to keep up with the demands of its volume. Figure 2.66 shows a resolution to the conundrum in the form of moving ATP production off of the exterior cell membrane and onto convoluted membrane structures in the cellular interior, namely, the mitochondria.

2.7 Combinatorial Scaling

RP: RP: this section is going to be on how a variety of important biological processes are subject to a combinatorial explosion and how that explosion scales with number of objects.

One of the most famous types of scaling of all will take center stage in chap. 7 where we will explore the entropy, a quantity that reflects the way that the number of different ways of configuring a system scales with the size of the system. What we will find is that biological systems feature numbers that are

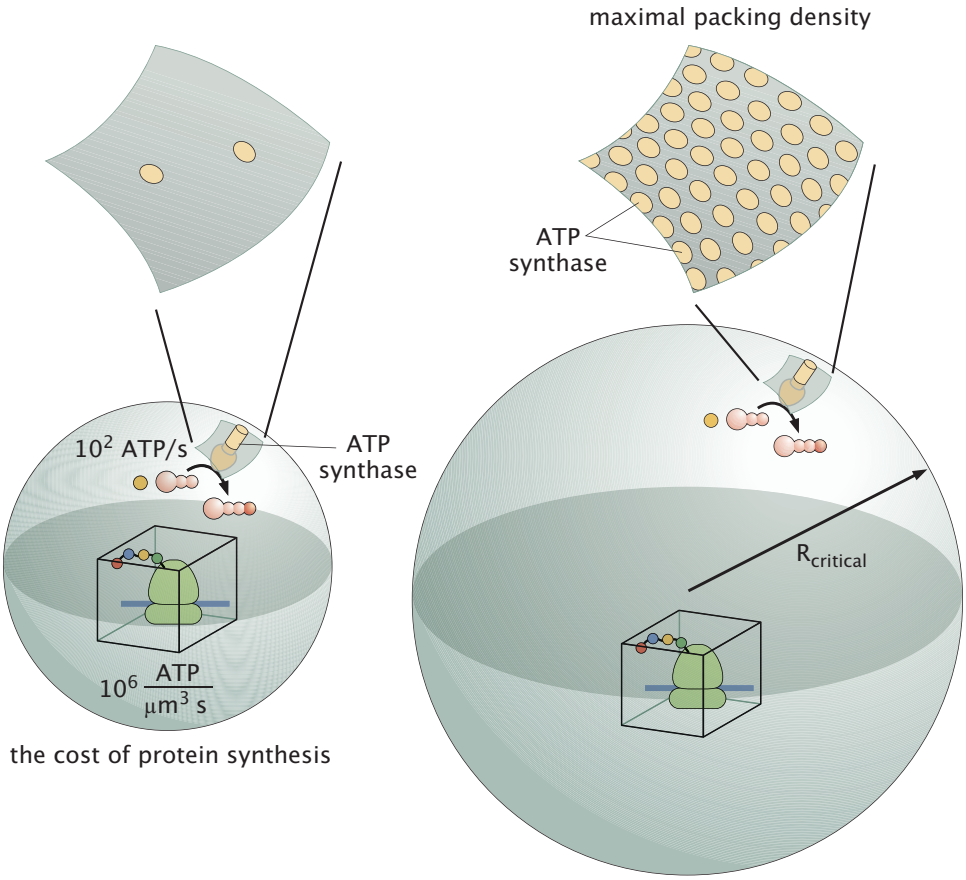


Figure 2.65: Examination of the ability of ATP synthase in the plasma membrane to keep up with bulk consumption of ATP. A toy model of a spherical cell imagines the plasma membrane decorated with a density σ of ATP synthases per unit area. For the small cell, the ATP synthases are sufficient to keep up with the demands of ATP synthesis.

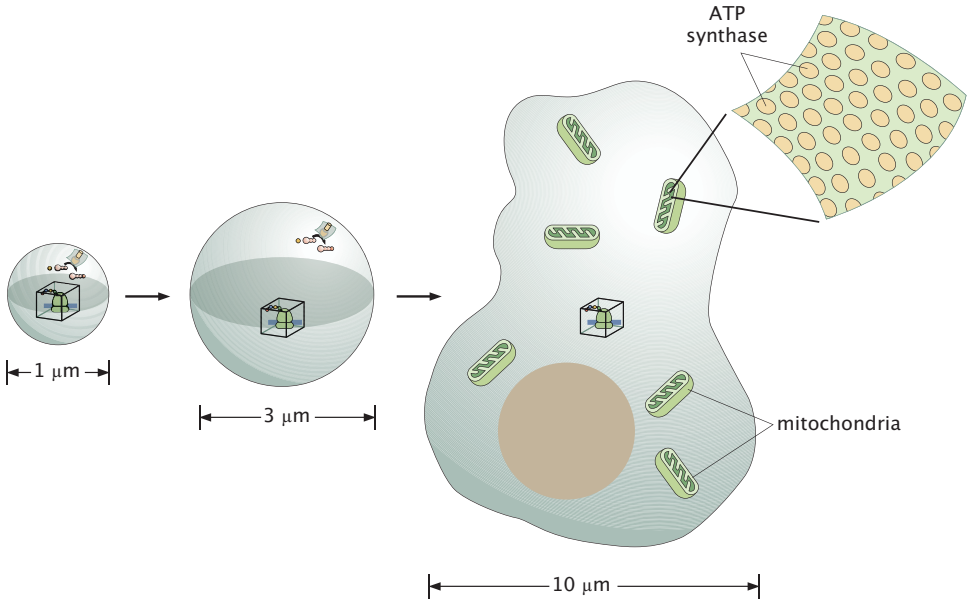


Figure 2.66: Membrane area for ATP synthases and mitochondria. Once the cell reaches a certain critical size, internal structures such as mitochondria are required to maintain ATP production at a high enough rate to keep up with the demands of the cytoplasm.

bigger than astronomical. They are so large, as we will show below, that we have to find new ways of representing huge biological numbers and to that end, we will introduce the Slog scale.

Perhaps the easiest ways to see the power of combinatorial scaling is with reference to the size of biological conformation and sequence spaces. Several molecular examples that are of great significance concern the space of different post-translational modifications available to a collection of proteins or the space of different antibody sequences available to the human immune system as a result of the process of V(D)J recombination. Or, perhaps most dizzying of all is to consider the total number of genomes that have been explored in the history of life and how that number compares to the total number of possible genomes.

To get a feeling for the problem, let's consider what has been referred to poetically as nature's way to escape from genetic imprisonment, namely, the amendment of proteins by the addition of chemical groups such as phosphate groups. The point is that proteins are all made up of the same 20 amino acids. But once those proteins are synthesized, some residues such as threonine, serine and tyrosine, can be modified through the addition of phosphate groups (there are a host of other post-translational modifications such as acetylation, palmitoylation and methylation, to name a few). If we consider amino acid usage, we find that roughly 10% of the residues are available for phosphorylation. Our aim here is to make a simple estimate of the number of distinct states of phosphorylation there are for a typical protein. Such a typical protein has roughly 300 amino acids and a simple estimate (see the problems at the end of the chapter) tells us that 2/3s of them are available close enough to the protein surface to suffer post-translational modification. Out of these 200 amino acids, if 10% of them are targets for phosphorylation, that means there are 20 potential phosphorylation sites on any given protein. Each such site can be in one of two states and hence there are $2^{20} \approx 10^6$ distinct states of phosphorylation. Given that there are several million proteins in a given *E. coli* cell, this means that there are

$$\text{number of possible proteins in one bacterium} \approx (10^6)^{f \times 10^6}, \tag{2.38}$$

distinct possible proteins in one bacterium! We really don't have any good way to think about these kinds of huge numbers.

We playfully illustrate some examples of both physical and biological significance in Figure 2.68 which introduces the slog scale, a way of capturing numbers that defy even the logarithmic scale. For example, if our number of interest has a slog of 4 that means that we can write out number in the form

$$2^{2^{2^2}} = 65,536 \tag{2.39}$$

or if it has a slog of 5, the we can write the number as

$$2^{2^{2^{2^2}}} \approx 2 \times 10^{19728}. \tag{2.40}$$

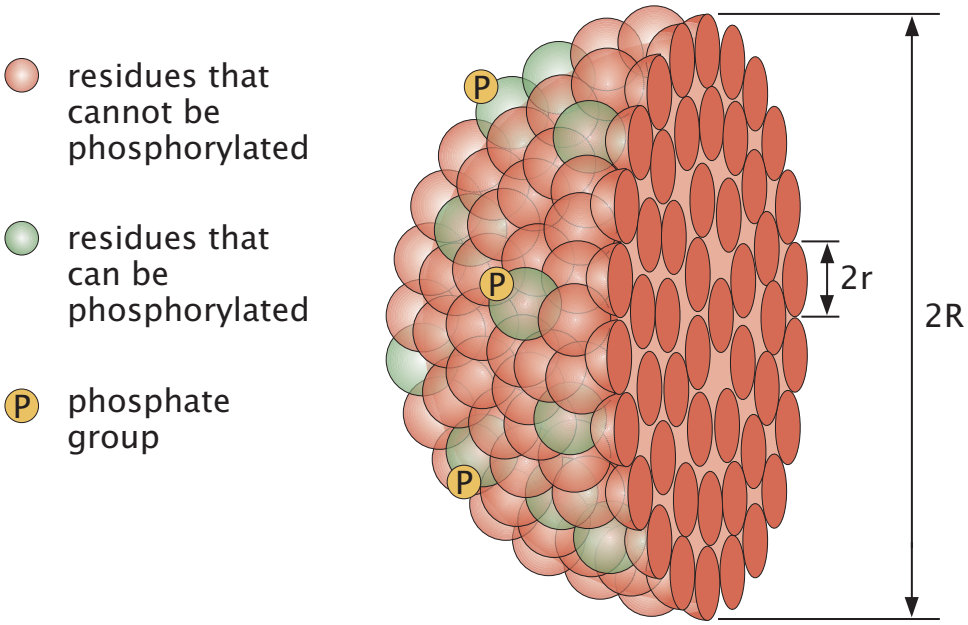


Figure 2.67: Estimate of number of distinct states of phosphorylation for a "typical" globular protein.

One thing is certain. Each step up the slog ladder represents a hallucinatingly large increase in the size of our numbers of interest. As seen in the figure, a great variety of interesting physical and biological numbers fall between Slog 5 and Slog 6, illustrating the weight of combinatorial scaling, that is, the rapid scaling of our sequence spaces and post-translational modification spaces with the addition of more nucleotides, or phosphorylation sites, respectively.

- Lucas Pelkmans, Jan Ellenberg - JT
- Dan Needleman evolution stuff

Concepts to raise:

- Different definitions of models: we need to draw the distinction to statistical models

2.8 Problems

Key to the problem categories: • Model refinements and derivations, • Estimates, • Data interpretation