

Preface

“I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be.”

William Thomson (Lord Kelvin) [Popular lectures and addresses, Vol. 1, Electrical Units of Measurement, 1883]

Though Lord Kelvin was unaware of the great strides that one can make by looking at bands on gels without any recourse to numbers, his exaggerated quantitative philosophy focuses attention on the possible benefits of biological numeracy.

One of the great traditions in biology's more quantitative partner sciences, such as chemistry and physics, is the value placed on centralized, curated quantitative data. Whether thinking about the astronomical data that describes the motions of planets or the thermal and electrical conductivities of materials, the numbers themselves are a central part of the factual and conceptual backdrop for these fields. Indeed, often the act of trying to explain why numbers have the values they do ends up being an engine of discovery.

In our view, it is a good time to make a similar effort at providing definitive statements about the values of key numbers that describe the lives of cells. One of the central missions of our book is to serve as an entry point that invites the reader to explore some of the key numbers of cell biology. We hope to attract readers of all kinds—from seasoned researchers, who simply want to find the best values for some number of interest, to beginning biology students, who want to supplement their introductory course materials. In the pages that follow, we provide a broad collection of vignettes, each of which focuses on quantities that help us think about mass, concentrations, energies, rates, information content, and other key quantities that describe the living world.

However, there is more to our story than merely providing a compendium of important biological numbers. We have tried to find a balance between presenting the data itself and reasoning about these numbers on the basis of simple estimates that provide both surprises and sanity checks. With

each vignette, we play with the interaction of two mindsets when thinking about cell biology by the numbers. First, we focus on trying to present in one place the relevant numbers for some particular biological structure or process. A second thrust is to “reason out” the numbers—to try and think about what determines their values and what the biological repercussions of those numbers might be. We are inspired by the so-called “Fermi problems” made famous as a result of the simple estimates made by Enrico Fermi on subjects ranging from the number of piano tuners in a large American city to the advantages of having double windows for thermal insulation in winter. We were interested in the extent to which it is possible to gain insights from a Fermi-inspired order-of-magnitude biology in which simple order-of-magnitude estimates serve as a sanity check on our understanding of biological phenomena.

When our hypothetical readers page to an entry of interest, be it the rate of translation or the number of genes in their favorite organism, we hope to greet them with a vignette that is at once entertaining and surprising. Rather than a dry elucidation of the numbers, as captured in our many tables, we use each vignette as a chance to tell some story that relates to the topic in question. We consider our book to be a quantitative companion to classic textbooks on molecular and cell biology and a source of enrichment for introductory and advanced courses. We thus aim to supply a quantitative component, which we consider an important complementary way of organizing and viewing biological reality. We think that knowing the measure of things is a powerful and different way to get a “feel” for the organisms and their inner life.

Another reason for writing this book emerged from our own research. We often want to do “quick-and-dirty” analyses to estimate time scales, rates, energy scales, or other interesting biological parameters as a sanity check to see if some observation or claim makes sense. The issue is how to make it quick. Looking for key biological numbers using the internet or flipping through textbooks is laborious at best and often futile. It is a common experience that even after hours of searching, we are left either with no result at all or a value with no reference to the experimental conditions that gave rise to that number, hence providing no sense of either the uncertainty or variability in the reported values. Our aspirations are for a biology that can boast the same kind of consistency in its data as is revealed in **Figure P-1**, which shows how in the early twentieth century

CONCLUSIONS.

120.—THE AGREEMENT BETWEEN THE VARIOUS DETERMINATIONS.—In concluding this study, a review of various phenomena that have yielded values for the molecular magnitude enables us to draw up the following table :—

Phenomena observed.	N 10 ²³
Viscosity of gases (van der Waal's equation)	62
Distribution of grains	68.3
Displacements	68.8
Brownian movement { Rotations	65
Diffusion	69
Irregular molecular distribution { Critical opalescence	75
The blue of the sky	60 (?)
Black body spectrum	64
Charged spheres (in a gas)	68
Charges produced	62.5
Radioactivity { Helium engendered	64
Radium lost	71
Energy radiated	60

Our wonder is aroused at the very remarkable agreement found between values derived from the consideration of such

Figure P-1 The many measurements of Avogadro's number. The French physicist Jean Perrin in his book *Atoms* noted the broad diversity of ways to determine “atomic dimensions” and was justly proud of the consistent picture of the world that emerged from such different approaches.

a host of different methods yielded a surprisingly consistent set of values for Avogadro's number. Often in biology we are not measuring specific physical constants such as Avogadro's number, nevertheless, when measuring the same quantity under identical conditions we should find similar results. One of the points that will come up again in the first chapter is that reproducibility is required first as the basis for recognizing regularities. Then, once scientists are confident in their regularities, it becomes possible to recognize anomalies. Both regularities and anomalies provide a path to new scientific discoveries.

Our vision is that we need a sort of a "cheat sheet" for biology, just like those we got in high school for physical and chemical constants. We hope this book will serve as an extended cheat sheet or a brief version of the handbooks of the exact sciences—namely, those used prevalently in engineering, physics, and so on. Marc Kirschner, the head of the Systems Biology department at Harvard University, compared doing biology without knowing the numbers to learning history without knowing geography. Our aim is that our readers will find this book to be a useful atlas of important biological numbers with allied vignettes that put these numbers in context.

We are well aware that the particular list of topics we have chosen to consider is subjective and that others would have made different choices. We limited our vignettes to those case studies that were consistent with our mutual interests and to topics where we felt we either knew enough or could learn enough to make a first pass at characterizing the state of the art in quantifying the biological question of interest.

The organization of the various numbers in the pages that follow is based upon roughly five different physical axes rather than biological context. First, we provide a narrative introduction to both the mindset and methods that form the basis for the remainder of the book. We offer our views on why we should care about the numbers described here, how to make back-of-the-envelope estimates, and simple rules on using significant digits in writing out numbers. We then begin the "by-the-numbers" survey in earnest by examining the sizes of things in cell biology. This is followed by a number of vignettes whose aim is to tell us how many copies of the various structures of interest are found. Taking this kind of biological census is becoming increasingly important as we try to understand the biochemical linkages that make up the many pathways that have been discovered in cells. The third axis focuses on force and energy scales. The rates of processes in biology form the substance of the fourth section of the book, followed by different ways of capturing the information content of cells. As is often the case in biology, we found that our human effort at rational categorization did not fit nature's appetite for variety, and thus the last section

is a biological miscellany that includes some of our favorite examples that defy inclusion under the previous headings.

Unexpectedly to us, as our project evolved, it became ever more clear that there is a hierarchy of accuracy associated with the determination of the numbers we describe. For example, our first chapter deals with sizes of components in the cell, a relatively accurate and mature outgrowth of modern structural biology with its many different microscopies. Our second chapter on the cellular census ramps up the difficulty, with many of the numbers we report coming from very recent research literature, some of which show that calibrations of different methods, such as fluorescence techniques and those based upon antibodies, are not entirely consistent. Chapter 3, which deals with energy scales of various processes within the cell, suffers from challenges as severe as ambiguities in the definition of the quantities themselves. We thought hard about how to represent in writing the uncertainties associated with values that we collected from the literature. The guidelines we follow regarding how many significant digits to use are summarized in the opening chapter. It is our hope that attention to this issue of quantitative sanitation will become the norm among students and researchers in biology.

Inspiration for the approach taken here of “playing” with the numbers has come from many sources. Some of our favorites, which we encourage our readers to check out, include: *Guesstimation* by Lawrence Weinstein and John Adam; John Harte’s two books, *Consider a Spherical Cow* and *Consider a Cylindrical Cow*; Richard Burton’s *Physiology by Numbers* and *Biology by Numbers*; *Why Big Fierce Animals Are Rare* by Paul Colinvaux; and Sanjoy Mahajan’s fine books, *Street Fighting Mathematics* and *The Art of Insight in Science and Engineering: Mastering Complexity*. We are also big fans of the notes and homeworks from courses by Peter Goldreich, Dave Stevenson, and Stirl Phinney on “Order of Magnitude Physics.” What all of these sources have in common is the pleasure and value of playing with numbers. In some ways, our vignettes are modeled after the examples given in these other books, and if we have in some measure succeeded in inspiring our readers as much as these others have inspired us, our book will be a success.