1 Robust Self-replicating Machines Shaped by Evolution

Biological cells are wonderfully robust self-replicating machines that selfishly propagate their DNA until environmental factors such as nutrients limit them. Cells have mastered the use of nanoscale devices in a water environment to fabricate their progeny and perform many critical functions for survival. To understand how cellular machines work, we need to understand the identity and structure of molecular components as well as the network of interactions between these molecules. We also need to study the complex cell functions in which an orchestra of molecules interacts transiently in space and time to cause an observable cellular behavior. Such an integrated systems engineering view of the cell is currently far from our reach; however, there are sufficient tools and techniques in place to provide us with many useful insights into the complex functions of the wonderful machines that are cells.

In this book, we will mainly discuss eukaryotic cells and where possible the details of mouse embryo fibroblasts (3T3 cells) and a human cancer line (HeLa). Because there are over 300 different cell types with important specializations, we have chosen a relatively general cell type to describe what cells do. Many complex functions are missing in fibroblasts and there will be some necessary extrapolations at certain points. 3T3 and HeLa cells do not form multicellular tissues; cell–cell interactions and the higher-order functions of tissues will be discussed in later chapters. The level of integration and complexity needed for the control of multicellular complex functions is at least an order of magnitude greater than for single cell complex functions. We first focus on dissecting the individual complex functions in single cells, because it is then possible to analyze multicellular complex functions in a modular fashion.

In this chapter, we will define the biological cell that we wish to understand and discuss how evolution has made it more robust over the last two billion years (Huxley, 1958). A cutaway diagram of a cell in suspension (Figure 1.1) provides an illustration of the basic organelle compartments that are found in mammalian cells. Compartmentalization of functions is critical for a complex machine like a cell. Each compartment performs some basic function(s) and communicates bidirectionally with the rest of the cell. Often compartments, like the cell itself,

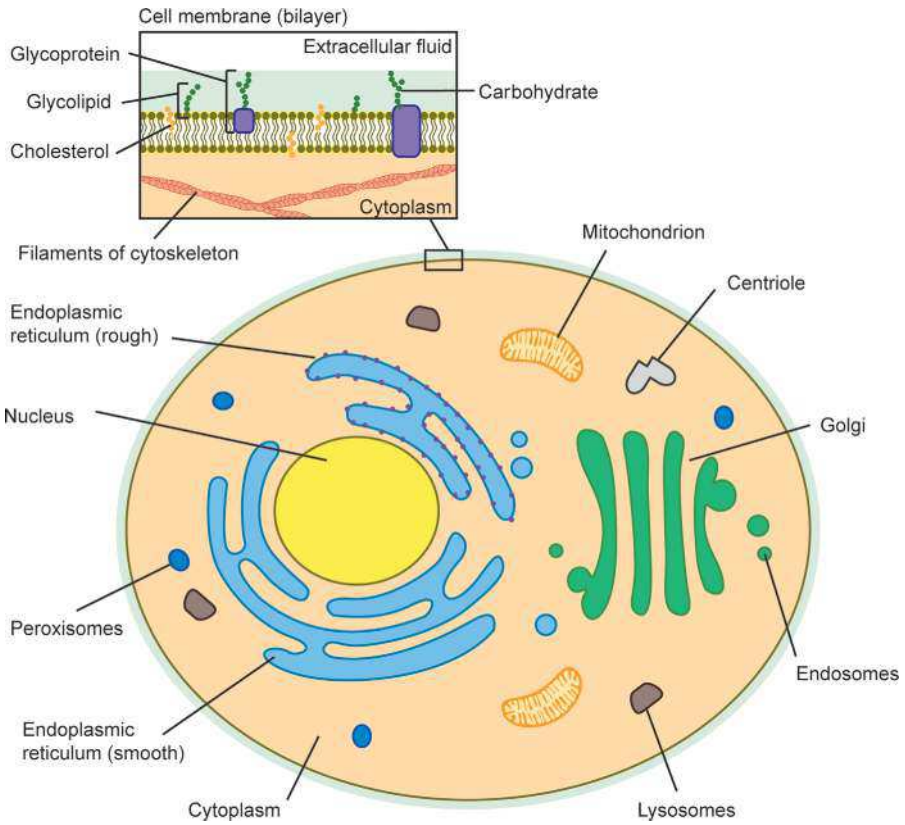

Figure 1.1
Diagram of cell with plasma membrane.

are bounded by lipid bilayers, whose hydrophobic cores effectively prevent open movements of ions and charged materials into or out of the compartments. The nucleus contains the DNA that encodes the plan for the cell that is read out by RNA polymerases. The mRNAs produced in the nucleus are transported through nuclear pores to the cytoplasm where they can be translated into protein by ribosomes. Proteins destined for extracellular secretion or the outer surface of the plasma membrane (the boundary of the cell) are translated by ribosomes on the rough endoplasmic reticulum. Carbohydrates are added to the proteins (making them glycoproteins) in the endoplasmic reticulum and the Golgi before movement to the plasma membrane in secretory vesicles (not shown). The plasma membrane barrier is primarily formed by a lipid bilayer with cholesterol and embedded glycoproteins. Actin filaments support the plasma membrane mechanically and attach it to the cell cytoskeleton that structures the cytoplasm. In cell mitosis, the centrioles organize the microtubule arrays that carry the chromosomes to the daughter cells and specify the location of the cleavage furrow for forming the two

daughter cells. Invaginations of the plasma membrane form endocytic vesicles that pass inactive proteins on to the lysosomes where they are degraded. ATP to power the cell comes primarily from the mitochondria. This is only a partial list of the cellular compartments, as many other cell functions are performed by localized complexes that could be considered as compartments even when they aren't surrounded by a membrane bilayer. This machine is able to grow and divide once a day.

1.1 The Cell is a Self-contained and Self-replicating Machine

A defining aspect of cellular life is the ability of a cell to grow and produce two daughter cells, i.e. to replicate itself. Thus, our working definition of a eukaryotic cell is a self-contained unit that can replicate itself in an organism. With proper nutrients and a growth-stimulating environment, a cell can make new proteins, nucleic acids, carbohydrates and lipids in a controlled way so that two daughter cells can be formed after 24 hours. This is not a continuous process and the cell goes through at least three distinct phases of the cell cycle: growth phase one, synthesis of DNA, and growth phase two. Further, the synthesis of at least proteins and ribonucleic acids occurs in bursts followed by periods of inactivity. As in any complex process, a series of steps needs to be completed in the proper sequence to replicate the original cell. Because the cell has many functional compartments, parallel processing can occur, but there will be check points where all of the necessary steps need to be completed before the cell can transition to the next phase in the cycle of self-replication.

An important aspect of the self-replication process is that each cell of an organism contains the information needed to form that organism in its nucleus. Thus, the cell replication also replicates the genetic material needed to direct the continued propagation of the organism. Recent studies have demonstrated that an adult cell can be used to create a new organism either through transplantation of an adult cell nucleus into an egg or by formation of induced pluripotent stem cells. With the exponential growth of the number of cells, the mass of the cells can rapidly exceeds the supply of nutrients required for further growth. Controls on growth are also needed to regulate the size of the organism. Thus, the self-replication aspect of cells is limited in most organisms and even bacteria have phases with limited growth. The organism and not the individual cell is critical for the propagation of the DNA, and exponential growth of organisms is limited only by resources.

The notion of self-replicating machines is not new, and several have considered the theoretical aspects of what is needed in a self-replication machine. With the

recent development of rapid-prototyping printers that can build three-dimensional structures, it is conceivable that a machine could be developed to replicate itself with a few raw materials. The critical issue is how to encode the information needed to direct the self-replication process in a form that the machine could reproduce. This aspect of bacterial cells has been addressed in the field of synthetic biology, Gibson and colleagues have synthesized the DNA of a bacterium and succeeded in putting that DNA into the cytoplasm of another bacterium. The bacterium grew with the synthesized DNA and the previous genome was degraded, making a synthesized bacterium (Gibson et al., 2008). Thus, it is possible to reprogram a bacterial self-replicating machine to produce a new bacterium with a totally defined set of genes. Understanding the problems that must be overcome to enable an artificial organism to replicate itself will help us further understand cellular life, which will in turn serve as a foundation for the engineering of drugs and biologics to treat diseases and facilitate tissue regeneration for biomedical applications.

1.2 Cell Functions have Evolved According to Darwinian Selection

There is geological evidence of cellular life in fossilized remains of organisms over the past 2 billion years on earth. Whether the first cells formed directly from some primordial soup, were created by God or came to Earth from some other planet is not considered here. Instead, we are concerned about how the cell functions observed today were selectively preserved through an evolutionary process known as Darwinian selection. As Julian Huxley wrote in the introduction to the 150th Anniversary edition of Darwin's *On the Origin of the Species*, "Natural selection was seen, not as involving the sharp alternative of life or death, but as the result of the differential survival of variants; and it was established that even slight advantages, of one-half of one percent or less, could have important evolutionary effects." Thus, over such a long period, countless individual cells have lived and the process of evolution has modified the cells in any given organism for robustness (see textbox for analogy to selective evolution of automobiles). Basically, mutations in the genome will modify the organism's ability to survive to produce offspring. As the environment changes over time, life vs. extinction should be considered in a stochastic context where greater probability of survival is key and small advantages over large populations and many generations will dominate. The critical message is that we only find the successes alive. No moral message is intended to be drawn from the fact that some survived and others did not; rather, it is important to understand how optimization of certain functions increased survivability.

Selection in the Marketplace and Relation to Performance

Darwinian selection is analogous in many ways to the process of evolution in the commercial marketplace. Since the first automobile was developed in the late nineteenth century, billions of vehicles have been produced with increasing sophistication. Refinements such as fuel injection, seatbelts and even cup-holders have been incrementally introduced into most modern vehicles according to buyers' preferences. Features that resulted from the fashion-of-the-day or whimsical fads were not often retained in later models. However, features that improve performance, efficiency and comfort are selectively preserved, and stand the test of time. Small differences in the perceived advantage of features are often sufficient for those features to be incorporated into almost all vehicles. For example, cup-holders are a small convenience making it marginally more likely that a person will buy a car with them than without. However, virtually all cars sold in the USA have cup-holders.

This analogy can help us to understand some of the difficulties faced by scientists who commonly try to understand how cells work by removing or mutating proteins in those cells. They assume that the proteins have functional advantages for the cell and altering the proteins will alter the relevant function. However, many proteins have been deleted from cells or organisms without any obvious change in function. If those proteins confer a small selective advantage, they can be selected for over many generations; however, their removal may not cause a major change in cell function. Similarly, a naïve person would find it very difficult to understand the advantage of a cup-holder unless they tested the driving performance of people drinking hot coffee. The point is that the selective advantages needed for proteins or mutations of proteins to be incorporated in all cells over thousands of generations are only in the 1% range that would not be readily observed by most assays. Darwinian selection will result in highly sophisticated machines robust to many perturbations that have finely tuned features.

Time to Extinction: It is useful to consider how long the genes coding for less-competitive proteins will remain in the gene pool.

Steady state system: In a stable ecosystem, the population of any given species will assume some average level over cyclic fluctuations. For example, the mouse population on a hypothetical island is on average 100,000 and they have a generation time of 6 months. Assume that a mutation in a protein gives the mice a 1% survival advantage as a heterozygote and 2% as a homozygote over the original protein in one generation. This means for the sake of calculation that for every 100 homozygous mutant mice in the first generation 102 will reach the second

(*continued*)

generation at the expense of the original mice. After 1000 generations and several cycles of population growth and contraction, the population of animals with the original gene will be very small (less than 3% of the total) and the number of homozygous animals will be even less (it should be pointed out that inbreeding of homozygous animals will increase the rate of decline).

Feast–famine cycles: More common in history is the expansion of the population of a given species and then a dramatic decline in that population due to the expansion of a predator population, change in habitat or a disease. In those cases, the selection pressures can be much more severe and over 90% of a given population can be lost in a brief period. Those situations are difficult to model because they select for markedly different factors, e.g. the decimation of buffalo populations by early settlers and of native Hawaiians by European diseases. The high selection pressures in those situations can often select for mutant genes in the population that would otherwise not be favorable, e.g. malaria selection of G6PD deficiency and sickle hemoglobin. In such cases, there may be clonal selection of individuals who have particular genes.

It is difficult to do evolutionary studies on mammals because of their long lifetimes; however, in bacterial systems it is possible to apply selective pressures and see mutations. For example, many recent studies have looked at the mutations in rate-limiting proteins in metabolic pathways that are produced upon dramatic changes in the temperature. Often, reproducible mutations in the key enzymes are found because those changes confer better performance at the new temperature. These experiments take a month or so because the bacterial doubling time is about 30 minutes and nearly 1000 generations will occur in that period (reviewed in Lindsay, 1995). As noted in the text box, if a mutation causes only a small decrease in the generation time, then in 1000 generations that mutation will dominate in the cell population.

In the same way, we can look at a cell or an organism that has a reasonable selective advantage over another cell type in the population, and with a few assumptions calculate the number of generations that must pass before the disadvantaged species is lost. If generation times of small mammals are on the order of months, several hundred million generations could have occurred in the 60 million years since dinosaurs walked the earth. Over that number of generations, even a gene that confers a weak selective advantage could displace other related genes. Thus, the organisms have tested many mutations of proteins and have shared DNA with other species through infections and other gene transfer mechanisms. For this reason, we will often need to have very quantitative assays of function to understand where a given protein plays a role in that function.

1.3 Robustness is Strongly Favored in the Evolution of Cell Functions

Cellular functions that enable species to survive and carry forward the genetic materials must be preserved through many different adverse conditions. The property of robustness is therefore strongly favored in cells. Robustness is defined as *the ability to adapt to and tolerate a variety of conditions*, such as the changes in: (1) number of proteins per cell, (2) salinity and pH, (3) temperature, (4) nutrient level, and (5) other environmental factors such as disease or predation. Robustness has a cost and very specialized systems adapted to an unusual environment can at times outcompete a more generally adaptable organism. For example, in caves where there are limited food resources, salamanders have evolved to become blind and to have no pigment in their skin. Producing pigment and eyes requires valuable energy and the animals without those features have a slight selective advantage. However, if the cave was suddenly opened to the light, the specialized salamanders would not compete well because they would be seen more easily by predators and would not be able to see predators to evade them. Therefore, a robust organism might not dominate in every situation, but over time with many different challenges and subtle environmental changes, robustness is favored over efficiency to survive.

Robustness is often dependent upon changes in the protein composition of cells through changes in gene expression. The mechanisms of regulated gene expression enable cells to express new proteins as needed. Specialized proteins can tune-down sensitivity to environmental perturbations, aid in repair processes after injury or other types of insults, or transform a cell from one phenotype to another. Different cell types will exhibit different expression programs, e.g. fibroblast vs. epithelial cell, and many of these programs are known. However, depending on the phase of the cell cycle, the same cell can also exhibit different expression patterns (Figure 1.2). For example, a fibroblast in division at M phase shuts down many complex functions, including growth, and drastically increases its ability to change shape – as though it were a different cell. In the case of fibroblasts growing in wounded tissue where damage to the environment is often severe, they must express many new proteins to enable them to rapidly divide and initiate healing.

The ability of single cells to change phenotype to survive an environmental challenge is critical for the propagation of the DNA. An example of this is the ability of cells to express heat shock proteins at high temperatures. Heat shock proteins aid in the refolding of heat-denatured proteins and thereby reduce the heat damage to the cell. The various environmental challenges that a robust cell should be able to live through (under normal circumstances) are outlined in detail below (Figure 1.3).

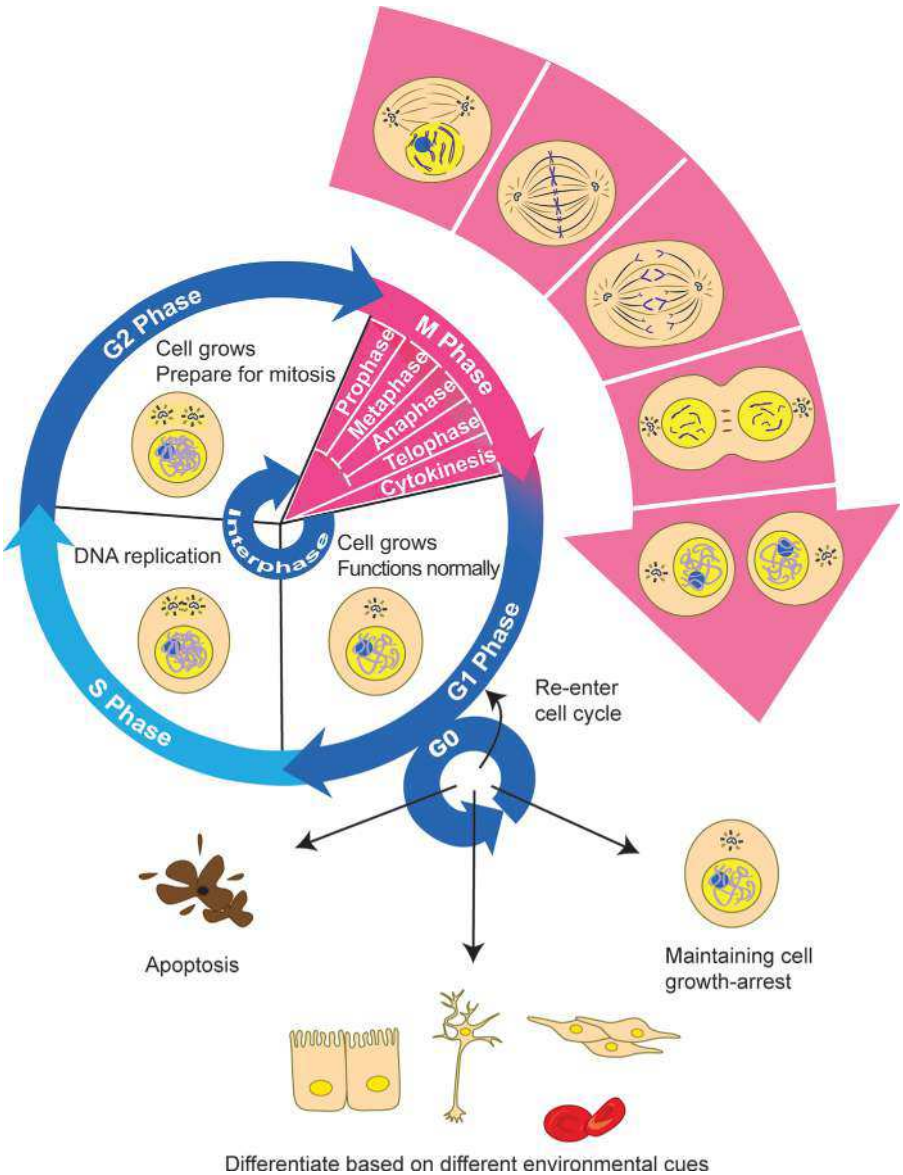

Figure 1.2
Cell phases and differentiation with different sets of complex functions on and off.

1.3.1 Number of Proteins per Cell

It seems that the number of molecules per cell of many proteins involved in a specific function can vary over a wide range without altering that function significantly. For example, during bacterial chemotaxis (the ability of bacteria to move up a concentration gradient to a food source), the concentration of many

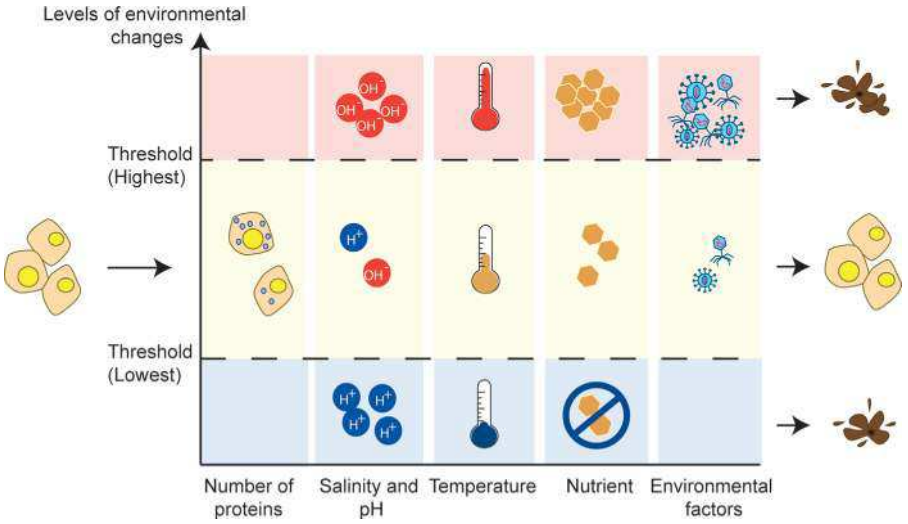

Figure 1.3
Robust cell adaptation in different environmental fluctuations.

proteins in the chemotactic pathway can vary over a thousandfold without altering the ability of the cells to undergo chemotaxis. There is a compromise in the rate of the response, however, when certain proteins are depleted. During development where cells are changing to a new type, the change in the cell protein composition can be incomplete and yet it will be able to function as the new cell. Further, there can be a major loss of many proteins during starvation without a major compromise of cell function. Although there are exceptions where the level of functional activity is directly proportional to the number of proteins per cell, in most cases the cell is robust to large changes in protein number. A rationale for this is that the production of protein is slow and often stochastic plus many proteins are concentrated at the site of function so that the concentration in cytoplasm is not particularly relevant. To turn off functions, cells often rely upon targeted proteases that can rapidly degrade critical proteins in a functional complex. Thus, the concentration of a given protein that is involved in an important function is not a good indication of the level of functional activity.

1.3.2 Salinity and pH

For bacteria and many fishes, the ability to withstand osmotic shock, as well as changes in ion content or pH of the surrounding medium, is critical to their survival. A sudden flood can decrease the salinity in a saltmarsh by over twofold and the organisms must therefore adapt rapidly, otherwise they will undergo lysis (rupture of the plasma membrane) when water moves into the cells. Along with

osmotic pressure, additional factors such as mechanical tears and detergents can also cause cells to lyse. Mammalian cells possess a large reservoir of internal lipid that can be used to seal over a leak in the plasma membrane that results during lysis and this prevents the critical contents of the cell from leaking out. Surprisingly, a large fraction of lysed cells will reseal their membranes and survive that trauma. A consequence of this observation is that many cell functions are performed by proteins that are anchored to the membrane or to cytoskeletal structures and these will not diffuse out of the lysed cell. We will discuss cell volume control and osmotic pressure as a regulator of ion channels in later chapters.

Calculations of Numbers of Molecules/Cell and Concentration

Eukaryotic cells range in size from about 4 μm to millimeters (for larger syncytial cells). An average size for the common cells used in cell culture (mouse 3T3 cells) is 2700 μm^3 for a suspended cell (volume $4/3\pi r^3$ = 2700 μm^3 or 2.7 × 10^{-9} cm^3). In contrast, bacteria (prokaryotic cells) such as *Eschericia coli* are cylinders of about 2 μm in length and 0.8 μm in diameter (volume $l\pi r^2$ ~1 μm^3 or 1 × 10^{-12} cm^3). The concentration of cytoplasmic proteins is about 180 mg/ml. If we assume that the average molecular weight of proteins is 50 kDa, then the overall concentration is 3.6 mM or 2 × 10^{18} molecules per ml (alternatively, 8 × 10^9 molecules per eukaryotic cell or 2 × 10^6 molecules per prokaryotic cell). If we carry this approximate calculation further and introduce the number of different protein molecules in a cell (10,000 for the eukaryotic and 2000 for the prokaryotic), then the number of molecules of any given protein will be on the order of 10^5 and 10^3, respectively. When we divide these numbers by Avagadro's number and the cell volume in liters, the concentrations are about 10^{-7} and 10^{-6} M, respectively. These numbers have important implications for the functional organization of cells that will be discussed later.

1.3.3 Temperature

Freeze–thaw cycles are common for organisms in polar climates and cells must be able to survive freezing or organisms will be lost. At low temperature some filament systems such as microtubules will disassemble which will block mitosis (cell division to form two daughter cells). In tropical climates, excess heat poses a different set of problems. Denaturation, or unfolding, of proteins may result from excessive heat, and this can lead to an excess of non-functional protein and insoluble aggregates that the cell must clear. Expression of a family of proteins known as heat shock proteins (HSPs) is stimulated at high temperatures and under conditions of stress to the organism. This highlights the need for cells to have