1 Life from a physics perspective

In this first chapter we review some basic concepts of organization and dynamics in living systems, with emphasis on what makes them different from systems that one normally studies with the tools of physics. Subsequent chapters use model systems to illustrate how interactions between proteins, metabolites and DNA govern strategic decisions on the scales ranging from a living cell, to organisms to populations.

Life has provided scientists with a vast number of stories [1–9] reflecting facets of the dynamic interplay between material and memory, an interplay that is formed by dynamical processes from the noise on the molecular scale, to the exceptionally long memory of replicating DNA, to the exponential growth of populations [10, 11].

In short, life is self-reproducing, persistent and robust (we are nearly 4 billion years old). Living systems are open systems that have "memorized" how to channel energy into self-reproducing networks. Living cells are complex (after all more than 1000 different types of molecules are needed to make even the simplest cell work), "more" than the sum of its parts (arbitrarily dividing an organism kills it). Life harvests energy and it evolves. The essential processes for the workings of life take place from the scale of a single water molecule to balancing the atmosphere of the entire planet [7, 12].

Biology has provided us with some fundamental/universal mechanisms that one meets in various disguises and variations [13].

1.1 Life copies from itself and from other life

Life copies and repeats on all scales, DNA, cells, individuals, young animal copying their parents, to us humans who copy each other's behavior, organizations [14] and inventions (see Fig. 1.1). Copying is often good as it opens up the possibility of "conquering the world" by exponential (Malthusian [15]) growth (see Fig. 1.2):

$$1 \rightarrow 2 \rightarrow 4 \rightarrow 8 \rightarrow 16 \rightarrow \ldots\ldots$$

Copying is functional, especially when it replicates what has already been proven to work. Learning also includes copying, see Fig. 1.1. Learning is often thought of as perfect copying, but without mistakes and failed attempts at copying there would be little progress in life. Failed attempts are necessary when life is struggling to keep up

Figure 1.1 Copying is important in the learning process where children often copy behavior from their older peers.

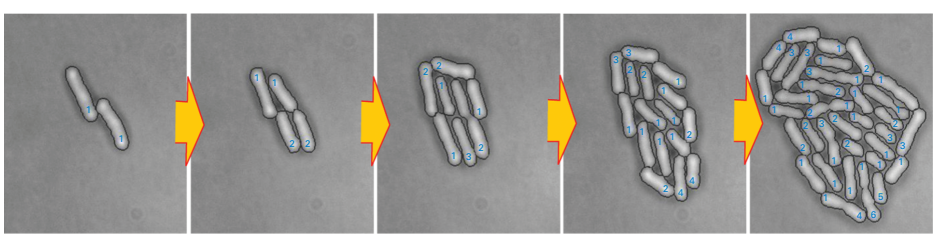

Figure 1.2 Copying and growth of E. coli cells that replicate every 30 minutes (micrographs courtesy of A. Trusina). The numbers refer to the age of cells in the number of cell generations.

with changes in its environment. A consequence of the power of exponential growth is the ability of any life form to explore scale, from the simplest single-celled organisms to algae bloom on the scale of an ocean; organisms replicate until they meet boundaries, set by other species or ultimately by the planet. The down side of exponential growth is that it is so powerful that it will often induce crashes, thereby making risk management essential for long-term sustainability.

1.2 Biology is the result of a very long historical process [1, 4, 16, 17, 18, 19]

A view highlighted first by Darwin, based on his observation of differences between species in similar but spatially separated habitats. Dependence on history implies that

it is not possible to "explain" a biological system by applying a few fundamental laws in the same way it is done in physics. If one replays the tape of history, the hydrogen atom could not be different from what it is, but an *E. coli* cell could. In evolution, it is much easier to modify already existing mechanisms than to invent completely new ones [20]. On evolutionary timescales nearly everything changes by cutting, pasting and reshuffling of already working subsections. As a consequence, all living organisms on Earth are part of the same 3.6 billion-year-old "family" that share ribosomes and the genetic code.

Evolution is self-organization on a long timescale, a stochastic but biased process where survivors are primarily selected to best deal with the contemporary environment formed in part by other life forms. Such a relatively short-term optimization often conflicts with long-term survival when the environment experiences catastrophic changes. However, there may be selection for the ability to evolve fast enough to adapt to changes in their surroundings. Thus, evolution may select the ability to evolve, for example through modulated single nucleotide mutations or large-scale DNA reshuffling using recombinations and transposons [9].

Perhaps the most mind-boggling aspect of evolution is the enormous time spans involved. Understanding the evolutionary processes is often hampered by our lack of perspective of how vast time is. Enormous time spans can apparently be hugely creative, provided that selection and evolvability is combined with memories of earlier successes:

$$Time \approx God \tag{1.1}$$

an "equation" that primarily serves to highlight the power of time in self-organizing systems. Figure 1.3 illustrates evolution and the randomness of history, by following the births and deaths of ammonite species over a quarter of a billion years.

1.3 Self-assembly and the genetic code

Self-assembly is an inherent property of life, and is, in particular, essential to the production of new cells using information from their ancestors. On the smallest scale it includes the equilibrium relaxation of highly non-random chains of amino acids into their minimum-energy state, which is the native shape of a functional protein [23]. Another larger-scale self-assembly process is the development of a full-scale adult animal from a single fertilized cell. In general, self-assembly is the replay of a pre-designed plan according to some fixed rules. Rules that in life are set by the information stored in the genetic code, the macromolecules it encodes and various processes and feedbacks that these molecules execute together in biomolecular circuits and networks.

Self-assembly relies on information stored on a one-dimensional, double-stranded DNA molecule (Fig. 1.4). Its linear nature mirrors the one-dimensional "backbone" of other polymers that make life work. The complementarity of two DNA strands allows *copying* of each of them, by virtue of separating the double-stranded DNA into two single-stranded DNAs that each carry the full information. The actual copying is done

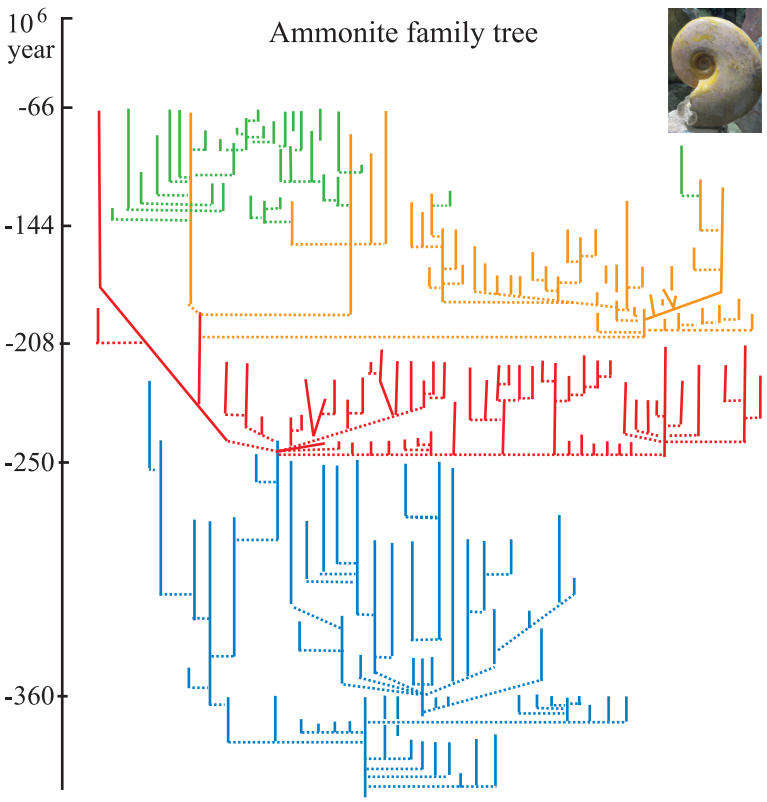

Figure 1.3 Ammonite family tree redrawn from [16]. Ammonites lived in water and fossilized well, leaving a fossil record of about 7000 ammonoid species. During their evolutionary history spanning 400 to 66 million years ago, one observes periods with fast speciation and simultaneous extinction of many different species. Today there are no living ammonites. Remarkably, in spite of a copying ability that would allow any species to take over the world in a few years, the individual species coexisted for million of years. A simple model showing cascades of co-evolutionary extinctions can be found in [17].

by DNA polymerase, a specialized protein machine that can copy 1000 base pairs in about one second.

DNA is transcribed into complementary RNA by RNA polymerase and subsequently "translated" into the amino-acid backbone of a protein by the ribosome. This last translation step again uses matching of complementary base pairs on mRNA to the corresponding tRNA, thereby translating codons consisting of three consecutive base pairs to the corresponding amino acid. Each codon can code for one of the 20 possible amino acids used by Life.[1] Inside the ribosome, these amino acids are subsequently added to an amino-acid chain that eventually will fold into the structure of a protein.

[1] In *E. coli* the translation takes of the order of 0.05 seconds per amino acid, and thus production of a typical 300 amino-acid protein takes of the order of 15 seconds.

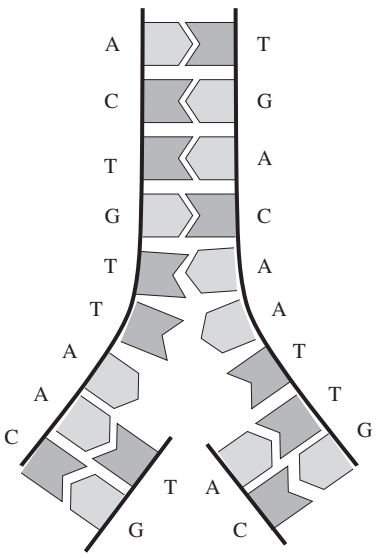

Figure 1.4 Information in life is maintained one-dimensionally by a double-stranded polymer called DNA. Each polymer strand in the DNA contains exactly the same information, coded in the form of a sequence of four different bases, creating pairs formed by bases that are complementary to each other. This complementarity is constructed using hydrogen bonds, with each base providing a particular pattern of hydrogen donors and hydrogen acceptors. Duplication is achieved by separating the strands and copying each one. This interplay between memory and replication allowed billions of years of complex history.

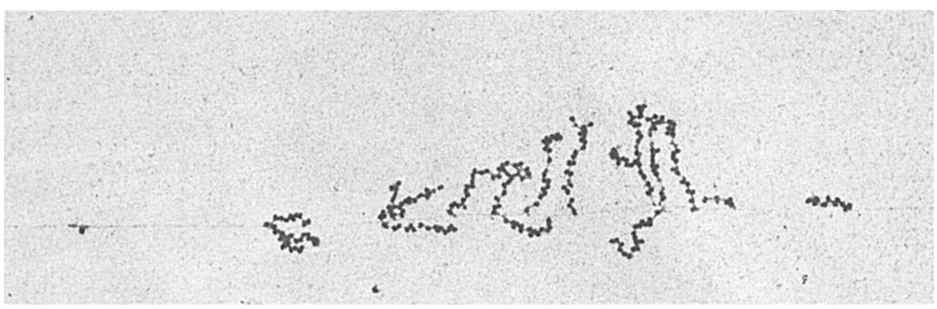

Figure 1.5 Transcription of messenger RNA (mRNA) and its subsequent translation into multiple identical proteins by ribosomes. The picture illustrates the flow of information from DNA, to mRNA, to proteins, and was taken using an electron microscope by Miller *et al.* [21] (reused with permission). Beautiful drawings of the interior of a living cell can also be found in [22].

The sequence of events leading to the assembly of a protein is usually summarized in terms of the central dogma for flow of information, from "memory," to "carbon copy," to "machine" (see Fig. 1.5.):

$$\text{DNA} \rightarrow \text{RNA} \rightarrow \text{protein} \tag{1.2}$$

This represents a simplified view of the actual situation. That is, if there were only DNA inside a cell then nothing would happen. To facilitate copying one needs several proteins to convert DNA to RNA, as well as an ample supply of energy in the form of ATP and GTP. A sustainable copying process in practice involves a complicated molecular network with feedback.

1.4 Biological molecules have long backbones

At the molecular scale life is made of polymers: DNA, RNA and proteins. Even membranes are built of nearly one-dimensional molecules with large aspect ratios. Perhaps mechanics at the nanoscale can only work with polymers, molecules that are kept together by strong (covalent) bonds along their backbones, while having the ability to form specific structures by utilizing much weaker bonds perpendicular to the backbone. These weak bonds are typically hydrogen bonds, where a hydrogen atom is shared in the same way as water is held together by sharing H atoms. The interplay between robustness and flexibility of polymers is illustrated in Fig. 1.6. Flexibility also allows folded polymers to work as complex molecular machines, in spite of their nano-scale size. Thus, functional molecules in life extensively use a seperation of energy scales: strong binding to maintain their identity, and weak binding to perform work.

1.5 Computation

A living cell is an incredible information-processing machine: a single *E. coli* cell can make about 3 000 000 proteins in 30 minutes, corresponding to a rate of about

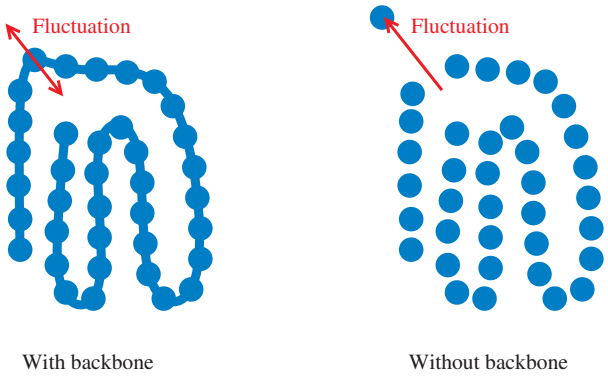

Figure 1.6 Illustration of the self-healing properties of a device with a polymer backbone. Thermal or other fluctuations may dislodge a single element, but if it is part of a polymer it typically moves back into the correct position.

3Mb/second information transfer, comparable to the speed of a typical internet connection today. And all this computation takes place within a very small volume of only $1\,\mu m^3 = (0.001\,mm)^3$, containing DNA with about 5 000 000 base pairs. The information density far outnumbers any computer chip created by humans, and even one million *E. coli* cells occupy much less space than a modern CPU, thus beating it on computation speed as well. The information management of copying the genome is supplemented by "computation" associated with organisms responding to shifts in environment, conditioned by specialized regulatory proteins. Only the simplest organisms (such as the prokaryote *Mycoplasma pneumoniae* with 677 genes) can essentially manage without transcription control. More complex prokaryotes typically need a larger number of control units, a number that in fact grows with the square of the number of genes [24].

1.6 Information gathering and communication

Information gathering and communication is used when organisms have to prepare for changes in food or for sudden exposure to predators. *E.coli*, for example, senses nutrient gradients by use of receptors, and uses this information to navigate towards better conditions [546, 548, 550, 551]. Other bacteria sense the presence of each other through signaling molecules [390, 552] and sometimes use this to decide when to form protective biofilms [391] which involves the cooperative effort of billions of bacteria. In wider perspective, communication opens for a large scale coordination that may reduce dangers and improve the usage of available resources. Absence of information about eventual calamities favors timely bet-hedging, where part of the population prepares for the worst at the cost of growing slower when conditions are good [705, 706, 708, 709]. Interestingly, the long-term fitness gain by bet-hedging can be expressed in terms of the severity and probability of calamities (see Section 12.3).

1.7 Feedback

Feedback is a near-essential part of life, allowing biological organisms to adapt to a number of environmental changes by altering their behavior (see Fig. 1.8). Biological networks are complex systems with multiple feedbacks. Often these involve proteins called transcription factors that regulate production of other proteins, which in turn perform some functional task. When the task changes, a signal goes back to the transcription factor, which then alters the protein production rate, for example as illustrated in Fig. 1.7. Remarkably, feedback nearly always combines several types of network connections, for example regulation of protein production by binding of transcriptional regulators to small metabolic molecules. Negative feedback loops are sometimes combined with positive feedback loops. This is, for example, seen in temporal and

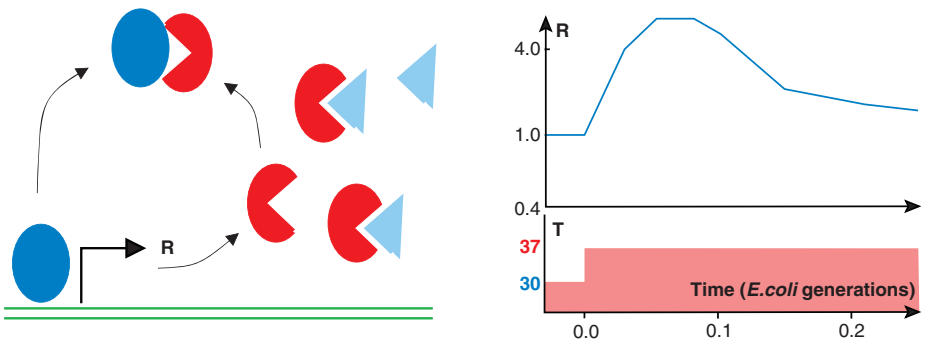

Figure 1.7 Basic negative feedback in biological networks where a regulatory protein (blue), directs production of many "worker proteins" (red). These "red" proteins bind to the "triangle molecule." If some "red" proteins are still available they bind to the "blue" protein and prevent it from initiating production of more "red" proteins. The right-hand panel shows an example of such a response, quantifying the activity of a heat-shock promoter after a 7 °C temperature shift in *E. coli* [25]. This response implies that the protein content of the cell changes and adapts to the new, higher temperature.

spatial decisions during embryonic development, as well as in spatio-temporal wave propagation [26].

1.8 Push–pull

Sustainability and stability of biological systems, from a folded protein, to cellular homeostasis, to the population of a species, often relies on pairs of counteracting "forces." In protein folding, the total binding energy of hundreds of kcal mol^{-1} is counteracted by almost equally strong entropic forces, resulting in the net stability of a folded protein of only about 10 kcal mol^{-1}. To maintain the cell in its optimally functioning state, proteins are constantly activated and de-activated to keep the cell primed for responding to external changes on a fast timescale. To achieve sensitivity, enzymes and signaling proteins are constantly phosphorylated and de-phosphorylated [27]. Maintaining epigenetic memory of differentiated cells involves hundreds of molecules that are reshuffled or changed on a fast timescale, while maintaining the overall state of the system.[2] Even pathogen–host interactions often result in an exceedingly long lasting battle, in spite of a fast turnover of individual viruses in the host, as seen, for example, with HIV [28] or trypanosomes [29]. Perhaps many of our other diseases also reflect a disturbed balance between the large and nearly equally strong opposing drives of repair and decay; for example, is Parkinson's a disease associated with aggregation that occurs within days outside the cell, but takes decades to unfold when cellular machinery can

[2] Other examples include toxin–anti-toxin systems in bacteria, as well as restriction-modification systems in bacteria, both examples that allow their "host" to prime themselves against external challenges.

counteract it [30]? An example of a push-pull in regulation is presented by phage λ, where the decision to integrate into the *E. coli* genome involves a sequence of promoters that alternately favors each of the two pathways: PRM − PR + PRE − PR′ + PAQ (Fig. 4.5).

1.9a Life is modular

Life is built of parts that in their turn are built of even smaller parts, over a wide range of scales. This facilitates robustness: if a process doesn't work, there are many ways in which parts can be modified or replaced in order to fix it [32]. Molecular-scale examples include secondary, tertiary and quaternary protein structures (the latter being complexes of multiple proteins). Network modules that each facilitate an appropriate response to external stimuli include subcellular compartments such as the nucleus, cytoplasm, etc. Most importantly, the minimal independent (or nearly independent) module of life is a single-celled organism.

The living cell is in itself a compact, complex system, with a large number of macromolecules squeezed into a minimum volume. These components are connected by networks orchestrating cellular behavior and response. As with computers, the smaller the cell, the faster the response. Therefore small cells are often competitive, with some macromolecules with copy numbers as low as one; a small number for sure, but one that allows the cell to respond sensibly to changes in external conditions.

1.9b Life is not modular

Life is more than the sum of its parts. Removing a single gene often leads to the death of an organism. Another observation is that the number of regulatory genes for prokaryotes increases as the square of the number of genes that should be regulated [488, 489]. In other words, doubling the size of a system requires four times more regulators. Thus the regulatory network of a living cell is a well-integrated system and not modular in any simple sense of this word (see Fig. 1.8).

1.10 Stochastic processes play an important role in life

Randomness influences processes from molecules [33, 34] to cells [35, 36, 37] to life and death of individuals or entire ecosystems [3]. At the smallest scale, stochasticity in life encompasses Brownian motion of molecules, mutations, trial-and-error strategies and individuality of genetically identical cells in bacterial populations. An example of the trial-and-error mechanism is microtubule growth, attachment and collapse [34]. This particular molecular process is important in the construction of connections between neurons in our brain.

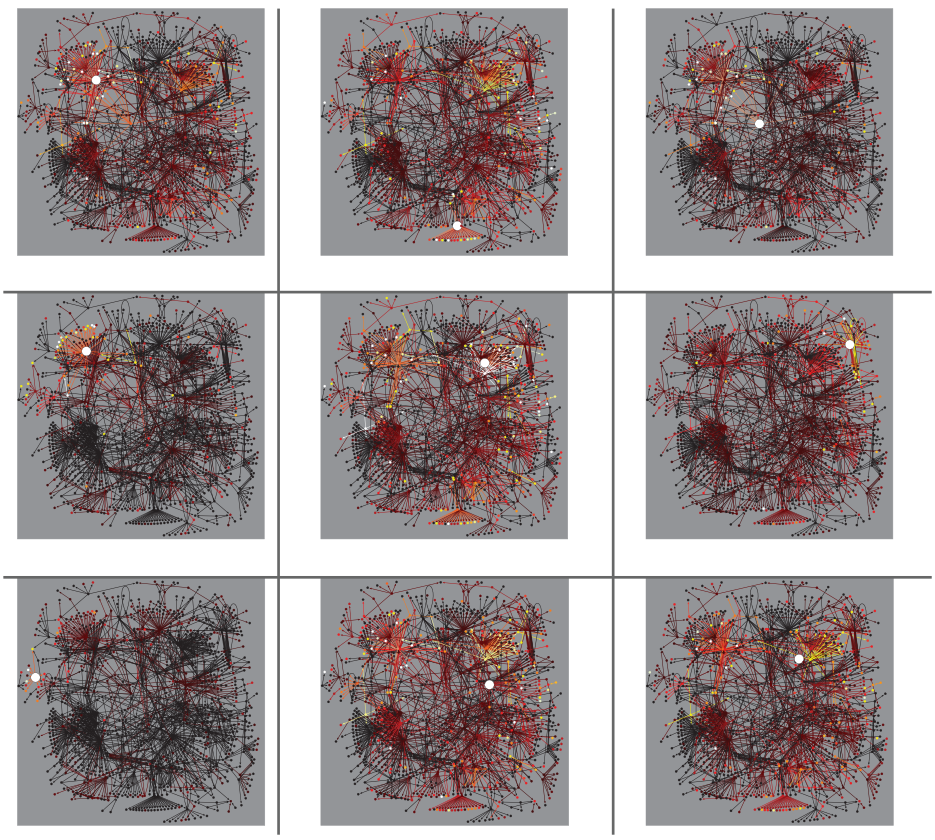

Figure 1.8 Responses in protein expression of the gene regulatory network of *S. cerevisiae* to various types of external environments/shocks [31]. One can see that the response is mostly localized (light color means large change in gene expression). The network only includes transcription regulation, and thus systematically misses the feedback links.

The individuality of cells has been probed by single-cell measurements of gene expression. Variability of cell fate has been attributed to these fluctuations (see Fig. 1.9). In fact, large variability in gene expression from individual to individual is often a good strategy to hedge against sudden environmental changes.[3] Without phenotypic and genetopic variations evolution could not have taken place.

1.11 Life harvests energy and converts it to cyclic motion

Energy allows life to maintain itself in an excited state, thereby allowing control of directional processes, including mechanical work and sorting of useful molecules from

[3] Examples of bet-hedging include persistent state of bacteria [40], sporulation [37] and lysogenic state of temperate phages [38, 39].