Chapter 1

Introduction

> ...the probability of any one of us being here is so small that you'd think the mere fact of existing would keep us all in a contented dazzlement of surprise... The normal, predictable state of matter throughout the universe is randomness, a relaxed sort of equilibrium, with atoms and their particles scattered around in an amorphous muddle. We, in brilliant contrast, are completely organized structures, squirming with information at every covalent bond... You'd think we'd never stop dancing.
>
> –*Lewis Thomas*, in *The Lives of a Cell*

1.1 Systems Biology and Synthetic Biology

Life is the most potent technology on the planet. It is also the most complex. This staggering complexity presents a fantastic puzzle to those studying its mysteries; more importantly, it offers a wealth of opportunities to those seeking to use our knowledge of biology to improve the quality of life for humanity.

Biology—the study of life—has a long and distinguished history dating back millennia, but our understanding of the *mechanisms* by which living things operate is fairly recent, and is still developing. We are, of course, intimately familiar with the behaviour of multi-cellular organisms (such as ourselves!), but the mechanisms by which living organisms function remained obscure until the 1950s. At that time the nascent field of molecular biology began to reveal the networks of interacting molecules that drive all cellular behaviour (and hence all life). These discoveries were made possible by experimental advances that allowed researchers to make observations on the tiny spatial scales of biomolecular processes. Over the last half-century, the molecular biology community has continued to uncover the details of this molecular domain. The painstaking effort involved in these nano-scale experiments necessitated a so-called 'reductionist' approach, in which research projects often addressed individual molecules or molecular interactions.

At the turn of the 21$^{\text{st}}$ century, further breakthroughs in experimental techniques set the stage for a shift in focus. The advent of so-called 'high-throughput' approaches allowed researchers to simultaneously observe the behaviour of large numbers of distinct molecular species. A cornerstone for these developments was the sequencing of the human genome (the first draft of which appeared in the year 2000). As a result, current molecular biology efforts have been dubbed 'post-genomic.' This 'modern' activity is characterized by experiments that reveal the behaviour of entire molecular systems, and so came to be called **systems biology**.

A key feature of present-day biological studies is a reliance on computation. The human genome project could not have been completed without advances in bioinformatics that allowed the processing and interpretation of vast amounts of sequencing data. In this book, we will take up a complementary use of computers in the study of molecular biology: the investigation of intracellular processes as *dynamic systems*. We will carry out these investigations by analysing mathematical models that mimic the behaviour of intracellular networks. Such modeling efforts have facilitated tremendous advances in other scientific disciplines. The use of such models in molecular biology has been, in the past, hampered by the absence of experimental observations of system behaviour; that is no longer the case.

In addition to their use in scientific investigation, dynamic mathematical models are used in engineering, where they play a central role in the design and analysis of engineered constructs. Biology shares several features with *engineering science*—defined as the application of the scientific method to the 'made world' of engineered artifacts. Because engineered objects have express reasons for existing, engineering scientists are able to use *performance measures* to assess the efficiency and robustness of their function. Although biological systems are part of the natural world, they exist (that is, they have been selected) because they carry out specific functions. Consequently, performance measures can be used to assess their behaviour, and biological 'design principles' can be identified. There are limits to this analogy between biology and engineering; natural selection is nothing like rational engineering design. Nevertheless, there are instances in which we can be reasonably confident of a biomolecular network's primary function. In these cases, biology straddles the line between natural science and engineering science, and can be described as *reverse engineering*—the unraveling (and ultimately reconstruction) of the products of an unfamiliar technology. (Historical examples of reverse engineering are primarily from wartime, e.g. the reconstruction of enemy aircraft.)

The construction, or *forward-engineering*, of biomolecular networks is an aspect of **synthetic biology**. This field is focused, in part, on the construction of designed genetic networks. The first engineered gene circuits were announced in the year 2000. Since then, the field of has synthetic biology grown rapidly. One of its most prominent activities is the international Genetically Engineered Machine (iGEM) competition, in which undergraduate student teams design, construct, and test genetic networks of their own imagining.*

Systems and synthetic biology represent unprecedented opportunities. In health and disease, agriculture, manufacturing, energy production, and environmental remediation, the use of biological technologies is leading to rapid progress in a wide range of human endeavours.

1.2 What is a Dynamic Mathematical Model?

This book addresses dynamic mathematical models of biochemical and genetic networks. These models, like all models, are abstractions of reality. Models are designed to focus on certain aspects of the object of study; other aspects are abstracted away. For instance, the familiar ball-and-stick model of chemical structure focuses on a molecule's chemical bonds. It does not capture, for example, the resulting polarity in the molecule's atoms.

Biologists regularly make use of tangible 'real-world' models. These can be simple, such as the molecular ball-and-stick, or complex, such as model organisms, or animal disease models. Biologists also use conceptual models. These typically take the form of verbal descriptions of systems, and

*The competition's website is `www.igem.org`

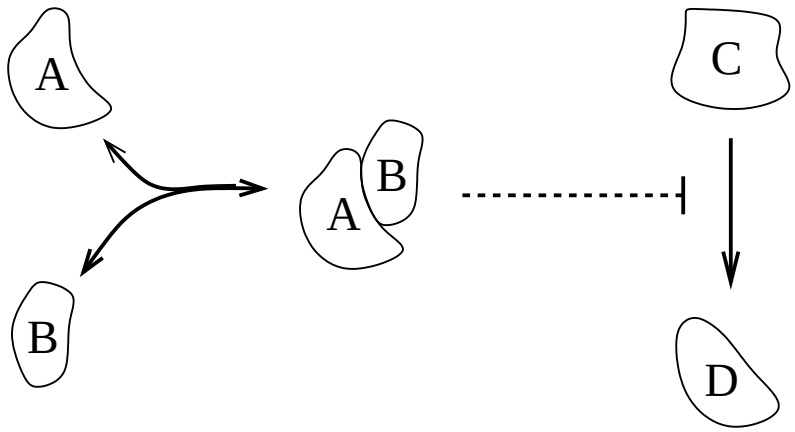

Figure 1.1: An interaction diagram, or 'cartoon' model. Molecular species A and B bind reversibly to form a molecular complex. This complex inhibits the rate at which molecules of species C are converted to species D. (The blunt-ended arrow signifies inhibition or repression. The dashed line indicates that this is a regulatory interaction in which the complex is not consumed.)

are communicated by diagrams that illustrate a set of components and the ways in which they interact (e.g. Figure 1.1). These interaction diagrams, or 'cartoon' models, play a central role in representing our knowledge of cellular processes.

A drawback of these cartoon models is that they can leave significant ambiguity regarding system behaviour, especially when the interaction network involves feedback. By using a mathematical description of the system, we can eliminate uncertainty in model behaviour, at the cost of demanding a quantitative representation of each of the interactions in the cartoon model.

As an example, suppose that, as in Figure 1.1, molecular species A and B bind to form a complex. In order to quantify that interaction, a numerical description of the process must be provided. In some instances, it may be sufficient to provide the equilibrium constant for the reaction. In other cases, separate rates of binding (association) and unbinding (dissociation) are needed. For a great many cellular processes, our current level of knowledge cannot support a quantitative description: we have only a qualitative understanding of the relevant molecular interactions. However, for a growing number of well-studied mechanisms, sufficient data has been collected to allow this sort of quantitative characterization.

When the relevant quantitative data is known, the interaction diagram can be used to formulate a dynamic mathematical model. The model-development process will be presented in Chapter 2. The resulting model consists of a set of equations that describe how the system changes over time—the system's dynamic behaviour.

Quantitative descriptions of molecular interactions typically invoke the laws of physics and chemistry. The resulting models are thus *mechanistic*—they describe the mechanisms that drive the observed behaviour. Each component of a mechanistic model represents some aspect of the system being studied; modifications to model components thus mimic modifications to the real system. (Mechanistic models can be contrasted with so-called descriptive models that seek only to summarize given data sets. Descriptive models provide limited insight into system behaviour.)

Investigation of mechanistic models follows two complementary paths. The more direct ap-

proach is model **simulation**, in which the model is used to predict system behaviour (under given conditions). Simulations are sometimes referred to as *in silico* experiments, because use computers to mimic the behaviour of biological systems. Simulations are carried out by numerical software packages, and will be used heavily in this book.

Alternatively, models can be investigated directly, yielding general insight into their potential behaviours. These **model analysis** approaches sometimes involve sophisticated mathematical techniques. The pay-off for mastering these techniques is an insight into system behaviour that cannot be reached through simulation. While simulations indicate how a system behaves, model analysis reveals *why* a system behaves as it does. This analysis can reveal non-intuitive connections between the structure of a system and its consequent behaviour. Chapter 4 presents model analysis techniques that are useful in molecular systems biology.

1.3 Why are Dynamic Mathematical Models Needed?

As mentioned above, interaction diagrams typically leave ambiguities with respect to system behaviour, especially when feedback is involved. Moreover, as the number of components and interactions in a network grows, it becomes increasingly difficult to maintain an intuitive understanding of the overall behaviour. This is the challenge of systems biology, and is often summarized by saying that "cells are complex systems." We can unpack this statement by providing some definitions.

The term **system** is often used without formal definition (as in the previous section!). Its meaning is somewhat context-dependent, but it typically refers to a collection of interacting components. In his book *Out of Control* (Kelly, 1995), Kevin Kelly defines a system as "anything that talks to itself." For example, an isolated stone is not considered a system, but an avalanche of stones is; the stones in the avalanche "talk" by pushing one another around.

Besides multiple interacting components, the other defining feature of a system is a *boundary*. A system consists of a set of components; anything that is not one of those components is not part of the system, and so is part of the 'external environment.' For example, a cell's membrane defines a boundary between the cell as a system and the extracellular environment. In certain contexts, a system is defined exclusively in terms of its interaction with this 'outside world,' and is then called an *input-output system*.

The term **complexity** also means different things to different people. Most would agree that a system qualifies as complex if the overall behaviour of the system cannot be intuitively understood in terms of the individual components or interactions. A defining feature of complex systems is that the *qualitative* nature of their behaviour can depend on *quantitative* differences in their structure. That is, behaviour can be drastically altered by seemingly insignificant changes in system features. Analytical methods for the investigation of complex behaviour will be presented in Chapter 4.

Two essential features of complex systems are nonlinear interactions and feedback loops. Feedback can be classified as negative or positive:

Negative feedback is exhibited when system components inhibit their own activity. (A familiar example is a household thermostat that corrects for deviation of temperature from a set-point.) These feedback loops generally stabilize system behaviour; they are the key feature of self-regulation and homeostasis. We will see, however, that instability and oscillations can arise when there is a lag in the action of a negative feedback loop.

Positive feedback is typically associated with unstable divergent behaviour. (Think of the runaway screech that occurs when a microphone and amplifier are connected in a positive feedback

loop.) However, when constrained by saturation effects, positive feedback can serve as a mechanism to 'locked in' a system's long-term behaviour—thus allowing a cell to retain a memory of past conditions.

1.4 How are Dynamic Mathematical Models Used?

Dynamic mathematical models serve as aids to biological investigation in a number of ways. The act of constructing a model demands a critical consideration of the mechanisms that underlie a biological process. This rigorous, reflective process can reveal inconsistencies in a 'cartoon' model and highlight previously unnoticed gaps in knowledge. Once a model has been constructed, it serves as a transparent description of the system, and can be unequivocally communicated. Moreover, a model recapitulates system behaviour; it concisely summarizes of all of the data it was constructed to replicate.

Both a cartoon model and a mathematical model are manifestations of a hypothesis: they correspond to putative descriptions of a system and its behaviour. The advantage of a mathematical model is that it is a 'working hypothesis,' in the sense that its dynamic behaviour can be unambiguously investigated. Although model simulations will never replace laboratory experiments, a model allows one to probe system behaviour in ways that would no be possible in the lab. Simulations can be carried out quickly (often in seconds) and incur no real cost. Model behaviour can be explored in conditions that could never be achieved in a laboratory. Every aspect of model behaviour can be observed at all time-points. Furthermore, model analysis yields insights into why a system behaves the way it does, thus providing links between network structure and behaviour.

Since a model is a hypothesis, the results of model investigation are themselves hypotheses. Simulations cannot definitively predict cellular behaviour, but they can serve as valuable guides to experimental design, by indicating promising avenues for investigation, or by revealing inconsistencies between our understanding of a system (embodied in the model) and laboratory observations. In fact, the identification of such inconsistencies is a key benefit of modelling. Because a model can be exhaustively investigated, it follows that a negative result—the inability of a model to replicate experimental observations—can be taken as a falsification of the hypotheses on which the model was built. This can lead to a refinement of the biological hypotheses, and subsequently a refined model, which can then be tested against additional experiments. This iterative process leads to a continually improving understanding of the system in what has been called a 'virtuous cycle.'

The end goal of most modelling efforts is a fully predictive description; simulations are then guaranteed to be accurate representations of real behaviour. Today's models of intracellular networks fall short of this goal, but examples abound in other sciences, and in engineering. Engineers make use of accurate predictive models for *model-based design*, resulting in faster and more efficient development of engineered constructs. The Boeing 777 jet provides a compelling example; it was designed and tested extensively in computer simulations before any physical construction began.

Model-based design is also being used in synthetic biology. Although models of cellular networks have only limited predictive power, they are useful for guiding the choice of components and suggesting the most effective experiments for testing system performance. The use of model-based design in the construction of engineered genetic networks will be illustrated briefly in Section 1.6.3 and discussed in more detail in Chapter 7.

1.5 Basic Features of Dynamic Mathematical Models

This section introduces some fundamental concepts in dynamic mathematical modelling.

State variables and model parameters

The primary components of a dynamic mathematical model correspond to the molecular species involved in the system (which are represented in the corresponding interaction diagram). The abundance of each species is assigned to a **state variable** within the model. The collection of all of these state variables is called the *state* of the system. It provides a complete description of the system's condition at any given time. The model's dynamic behaviour is the time-course for the collection of state variables.

Besides variables of state, models also include **parameters**, whose values are fixed. Model parameters characterize interactions among system components and with the environment. Examples of model parameters are: association constants, maximal expression rates, degradation rates, and buffered molecular concentrations. A change in the value of a model parameter corresponds to a change in an environmental conditions or in the system itself. Consequently, model parameter are typically held at constant values during simulation; these values cab be varied to explore system behaviour under perturbations or in altered environments (e.g. under different experimental conditions).

For any given model, the distinction between state variables and model parameters is clear-cut. However, this distinction depends on the model's context and on the time-scale over which simulations run. For instance, in Chapter 5, we will focus on models of metabolism, in which enzyme catalysts provide a fixed 'background.' In that context—and on the relevant time-scale of seconds to minutes—we will treat enzyme abundance as fixed model parameters. In contrast, the models in Chapter 7 describe gene regulatory networks, which are responsible for the regulation of enzyme abundance (on a slower time-scale). In those models, enzyme concentrations will be time-varying state variables.

Steady-State Behaviour and Transient Behaviour

Simulations of dynamic models represent time-varying system behaviour. Models of biological processes almost always arrive, in the long run, at steady behaviours. Most commonly, models exhibit a persistent operating state, called a **steady state**; some systems display sustained oscillations. The time-course that leads from the initial state to the long-time (or *asymptotic*) behaviour is referred to as the **transient**. In some cases, we will focus on transient behaviour, as it reflects the immediate response of a system to perturbation. In other cases, our analysis will concern only the steady-state behaviour, as it reflects the prevailing condition of the system over significant stretches of time.

Linearity and nonlinearity

A relationship is called **linear** if it is a direct proportionality. For example, the variables x and y are linearly related by the equation $x = ky$, where k is a fixed constant. Linearity allows for effortless extrapolation: a doubling of x leads to a doubling of y, regardless of their values. Linear relationships involving more than two variables are similarly transparent, e.g. $x = k_1 y + k_2 z$. A dynamic mathematical model is called linear if all interactions among its components are linear

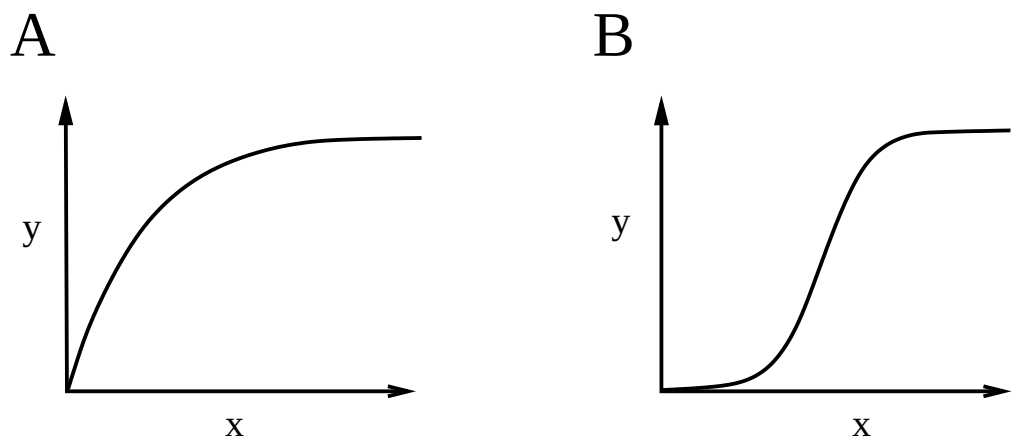

Figure 1.2: Common nonlinear relationships in cell biological processes. **A.** Hyperbolic saturation. As x increases, y also increases, but at an ever-diminishing rate. The value of y thus approaches a limiting, or *asymptotic*, value. **B.** Sigmoidal nonlinearity. The values of y show a slow rate of increase for small values of x, followed by a rapid 'switch-like' rise toward the limiting value.

relationships. This is a highly restrictive condition, and consequently linear models display only a limited range of behaviours.

Any relationship that is not linear is referred to (unsurprisingly) as **nonlinear**. Nonlinear relations need not follow any specific pattern, and so are difficult to address with any generality. The nonlinearities that appear most often in biochemical and genetic interactions are saturations, in which one variable increases with another at a diminishing rate, so that the dependent variable tends to a limiting, or *asymptotic* value. Two kinds of saturating relationships that we will encounter repeatedly in this text are shown in Figure 1.2. Panel A shows a *hyperbolic saturation*, in which the rate of increase of y declines continuously as the value of x increases. Panel B shows a *sigmoidal saturation*, in which y initially grows very slowly with x, then passes through a phase of rapid growth before saturating as the rate of growth drops.

Global and local behaviour

Nonlinear dynamic models exhibit a wide range of behaviours. In most cases, a detailed analysis of the overall, **global**, behaviour of such models would be overwhelming. Instead, attention can be focused on specific aspects of system behaviour. In particular, by limiting our attention to the behaviour near particular operating points, we can take advantage of the fact that, over small domains, nonlinearities can always be approximated by linear relationships (e.g. a tangent line approximation to a curve). This **local** approximation allows one to apply linear analysis tools in this limited purview. Intuition might suggest that this approach is too handicapped to be of much use. However, the global behaviour of systems is often tightly constrained by their behaviour around a handful of nominal operating points; local analysis at these points can then provide comprehensive insight into global behaviour. Local approximations are of particular use in biological modelling because self-regulating (e.g. homeostatic) systems spend much of their time operating around specific nominal conditions.

Deterministic models and stochastic models

The notion of determinism—reproducibility of behaviour—is a foundation for much of scientific investigation. A mathematical model is called **deterministic** if its behaviour is exactly reproducible. Although the behaviour of a deterministic model is dependent on a specified set of conditions, *no* other forces have any influence, so that repeated simulations under the same conditions are always in perfect agreement. (To make an experimental analogy, they are perfect replicates.)

In contrast, **stochastic** models allow for randomness in their behaviour. The behaviour of a stochastic model is influenced both by specified conditions and by unpredictable forces. Repeated stochastic simulations thus yield distinct samples of system behaviour.

Deterministic models are far more tractable than stochastic models, for both simulation and model analysis. In this text, our focus will be on deterministic models. However, stochastic models are often called for, particularly in studies of gene regulatory networks, where thermal agitation of individual molecules is a significant source of randomness. A stochastic modelling framework is introduced in Section 7.6.

1.6 Dynamic Mathematical Models in Molecular Cell Biology

A great many mathematical models of cellular phenomena have been published in the scientific literature. These are routinely archived in model repositories, such as the Biomodels database, the CellML model repository, and the JWS online repository.*

Before beginning our discussion of model construction in Chapter 2, we briefly present four examples of modelling projects. These short case-studies illustrate the range of biological domains that will be explored in this text and demonstrate a number of uses for dynamic mathematical modelling in systems biology. Readers unfamiliar with molecular biology may find it useful to consult Appendix A before continuing.

1.6.1 Drug target prediction in *Trypanosoma brucei* metabolism

The single-cell parasite *Trypanosoma brucei* infects the bloodstream, causing sleeping sickness. It is a single-celled eukaryote, so is not susceptible to bacterial antibiotics. The search for efficient treatments is ongoing. In 1999, Barbara Bakker and her colleagues published a study of glycolysis in *Trypanosoma brucei* (Bakker *et al.*, 1999). Glycolysis, an energy-producing metabolic pathway, is crucial to both the parasite's and the host's metabolism. Fortunately, the mammalian enzymes responsible for catalysing the reactions in the host pathway are significantly different from those of *Trypanosoma*. Thus the enzymes of the parasite can be inhibited by drugs that have little effect on the host. Bakker and her colleagues sought to identify which enzymes control the rate of glycolysis in *Trypanosoma*, with the aim of predicting ideal targets for growth-inhibiting drugs. The interaction diagram for their model, reproduced as Figure 1.3, shows the metabolic reactions in the network.

Using data from previously published studies of *Trypanosoma* and from their own experiments, Bakker and her colleagues formulated a dynamic mathematical model of the pathway. Their focus was on the steady-state behaviour of the system, and particularly on the rate of energy production. In order to predict the effects of enzyme-inhibiting drugs on the energy-production rate, they applied a *sensitivity analysis* to the model. This technique determines how sensitive model behaviour is to

*These and other repositories can be accessed from `systems-biology.org/resources/model-repositories`.

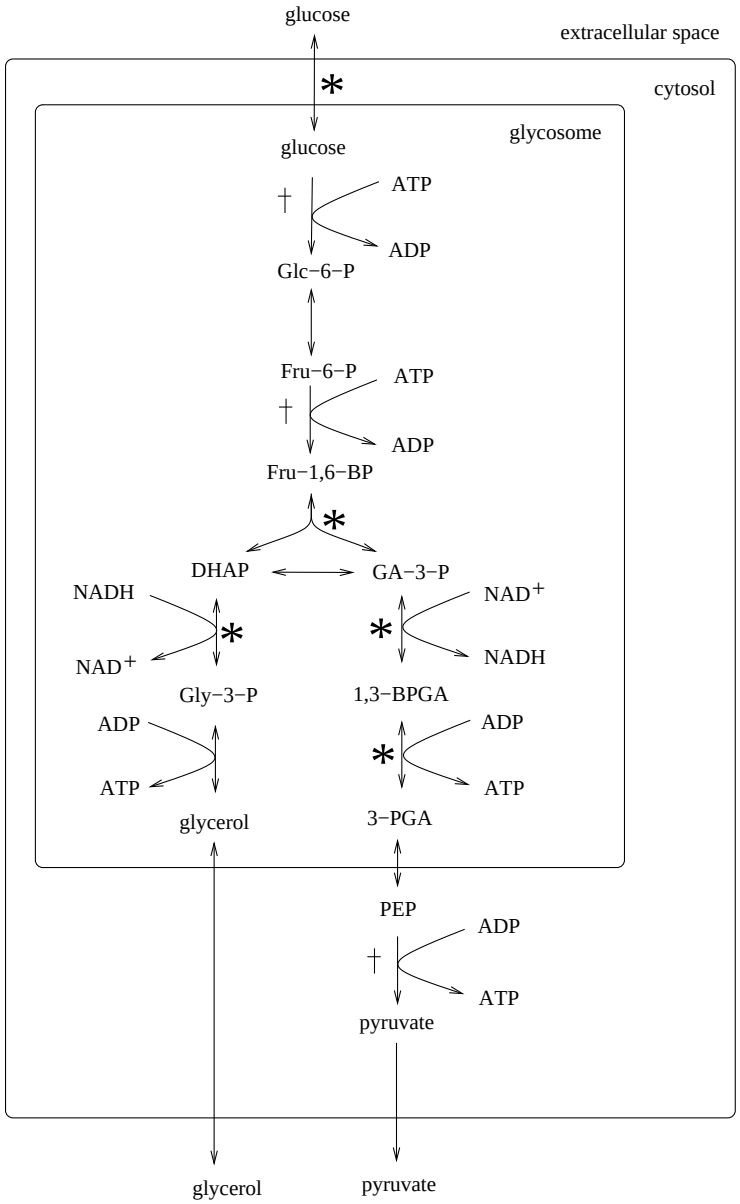

Figure 1.3: The glycolytic pathway of *Trypanosoma brucei* (simplified from Figure 1 of Bakker *et al.*, 1999). The reactions occur primarily in the glycosome—a specialized organelle. The analysis performed by Bakker and her colleagues indicated that the best targets for inhibition are those steps marked by an asterisk (*). The three steps marked with a dagger(†) were commonly believed to have significant influence over the pathway. However, Bakker and colleagues found that these are poor targets—their model predicted that inhibition of those reaction steps has little effect on the overall pathway flux. Abbreviations: Glc-6-P, glucose 6-phosphate; Fru-6-P, fructose 6-phosphate; Fru-1,6-BP, fructose 1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; GA-3-P, glyceraldehyde 3-phosphate; Gly-3-P, glycerol 3-phosphate; 1,3-BPGA, 1,3-bisphosphoglycerate; 3-PGA, 3-phosphoglyceric acid; PEP, phosphoenolpyruvate; ATP, adenosine triphosphate; ADP adenosine diphosphate; NAD+, NADH, nicotinamide adenine dinucleotide. Adapted from Figure 1 of (Bakker *et al.*, 1999).

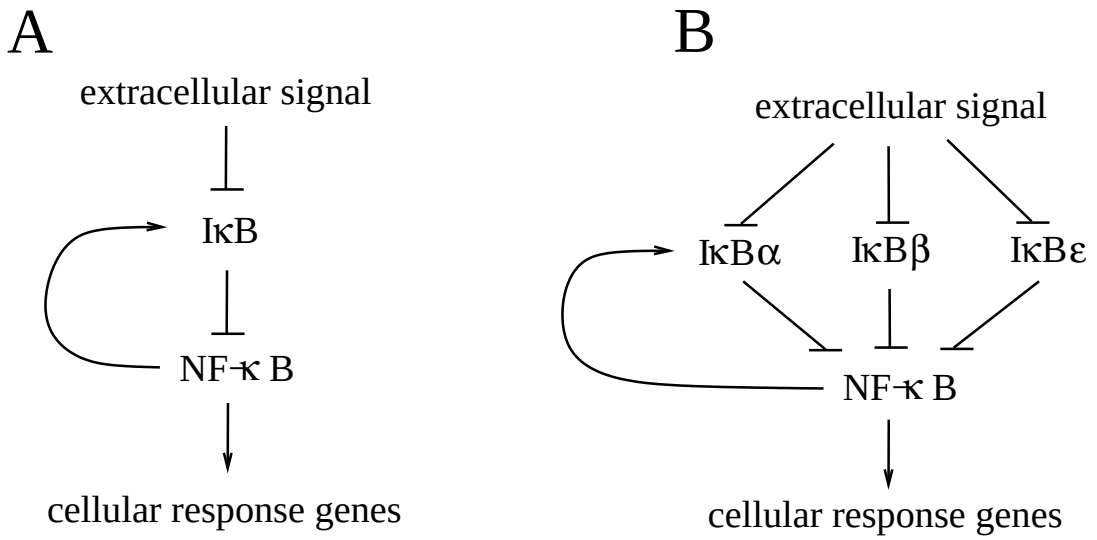

Figure 1.4: **A.** The main components of the NF-κB signalling pathway. Extracellular events trigger degradation of the inhibitor IκB. Once free of this inhibitor, NF-κB proteins trigger an appropriate cellular response. At the same time, NF-κB stimulates production of IκB proteins, leading to restored inhibition of NF-κB. **B.** A more detailed diagram, showing a family of IκB proteins. All three forms act as inhibitors of NF-κB. IκBβ and IκBε are unaffected by NF-κB; only the IκBα form is stimulated by NF-κB activity.

perturbations in the model parameters. Bakker and her colleagues identified five enzymes whose inhibition would have a significant impact on pathway activity. Moreover, they demonstrated that three other enzymes that had previously been proposed as effective drug targets are in fact poor targets; inhibition of these enzymes has little impact on pathway flux.

Later, Bakker and colleagues provided experimental confirmation of several of their model predictions (Albert *et al.*, 2005). Their experiments also provided evidence for regulatory interactions that were not included in the original model, thus opening an avenue for further model development.

We will address modeling of metabolic networks in Chapter 5. Sensitivity analysis, to be introduced in Section 4.5, plays a key role in the investigation of metabolism.

1.6.2 Identifying the source of oscillatory behaviour in NF-κB signalling

The protein NF-κB is involved in a number of animal cell responses, including the regulation of cell division, inflammation, and programmed cell death. The NF-κB pathway plays a role in inflammatory diseases and has been implicated in the development of cancer.

In the absence of stimuli, NF-κB is inhibited by proteins called IκB (Inhibitor of κB), as shown in Figure 1.4A. Extracellular stimuli (such as hormones) trigger protein activity in the cell that leads to a decrease in IκB levels. This frees NF-κB from the action of its inhibitor, allowing it to stimulate a cellular response (through changes in gene expression). In addition, NF-κB activity causes IκB levels to rise, resulting in renewed inhibition of NF-κB itself. Thus the pathway response is self-limiting. This sort of negative feedback loop typically leads to robust steady-state behaviour, but it can also lead to persistent oscillations.

Indeed, NF-κB pathways exhibit a range of behaviours upon stimulation, including both damped and persistent oscillations, as sketched in Figure 1.5. In 2002, Alexander Hoffmann and his col-

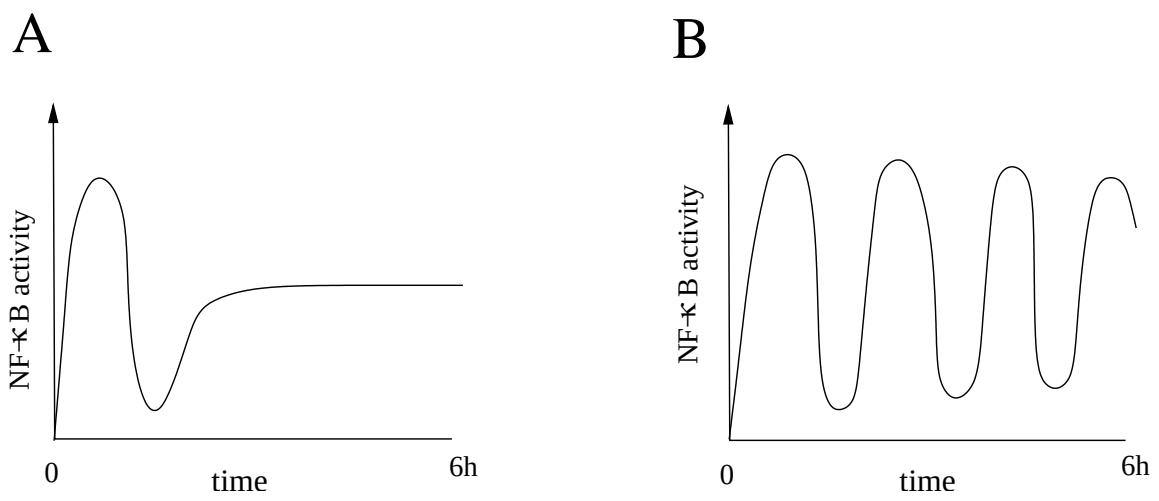

Figure 1.5: Oscillatory behaviour in the NF-κB pathway. **A.** The normal (wild-type) response to persistent signalling. The oscillations are quickly damped, resulting in a persistently active response. **B.** The behaviour of a cell in which IκBβ and IκBε are absent. The oscillations persist. Adapted from Figure 1 of (Cheong *et al.*, 2008).

leagues presented a mathematical model of NF-κB signalling that sheds light on the system's dynamic response (Hoffmann *et al.*, 2002). The model focuses on the roles of three distinct forms of the inhibitory IκB proteins, called IκBα, IκBβ, and IκBε. When all three of these forms are present in the cell, the pathway exhibits damped oscillations in response to stimulation (Figure 1.5A). However, when cells are modified so that certain IκB proteins are absent, the response changes. When IκBα is absent, cells show pathologically high activity. Alternatively, when both IκBβ and IκBε are absent, cells respond to stimuli with sustained oscillations in NF-κB activity (Figure 1.5B). This difference in behaviour is a consequence of the fact that, of the three IκB forms, only IκBα production is enhanced by NF-κB activity (Figure 1.4B).

From a design perspective, an ideal response would be a quick rise to a steady activated level. This ideal response is closely approximated by the damped oscillations normally displayed by the cells (Figure 1.5A). Hoffman and his colleagues used their model to determine that this response is generated by the combined behaviour of the IκB proteins. IκBα provides a negative feedback that quenches the quick initial rise, resulting in an appropriate steady level. However, a fast response demands a quenching signal so strong that oscillations would arise unless there were a secondary persistent quenching signal to damp the response. This secondary quenching is provided by the steady activity of IκBβ and IκBε.

Hoffmann and colleagues generated model simulations that describe the response of the pathway to stimuli of varying strengths and durations. These model predictions, verified by experiments, show that the complementary roles of the IκB proteins generate qualitatively different responses to stimuli of different durations, and that these differences in signalling activity lead to distinct cellular responses. (The paper (Cheong *et al.*, 2008) describes the results of additional NF-κB modelling efforts.)

Chapter 6 is devoted to intracellular signalling pathways. Tools to address oscillatory behaviour are presented in Section 4.3. A model of NF-κB activation is explored in Problem 7.8.15.

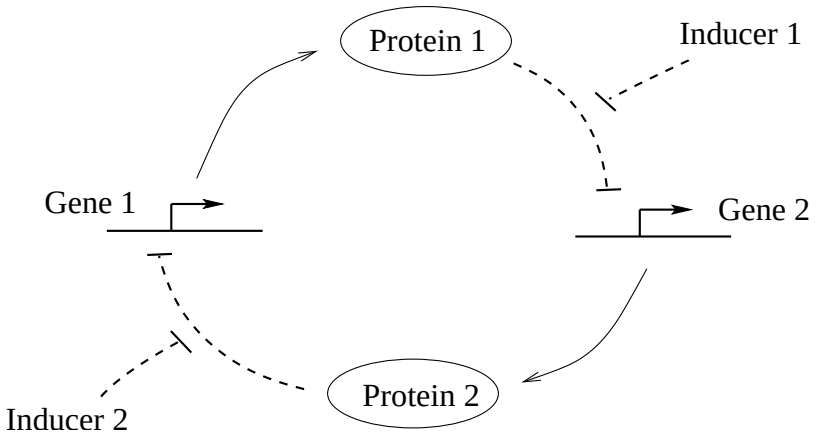

Figure 1.6: Engineered genetic toggle switch. Each gene's protein product represses production of the other, resulting in two modes of persistent operation: one protein is abundant while the other is repressed. The proteins are chosen so that intervention by the experimenter can deactivate the abundant protein, inducing a switch in the protein levels.

1.6.3 Model-based design of an engineered genetic toggle switch

In the year 2000, the emerging field of synthetic biology was heralded by the simultaneous announcement of two engineered genetic networks. Both of these devices will be covered in depth in Chapter 7. For now, we will briefly introduce one of these devices—a genetic toggle switch—and describe how modelling played a key role in its design.

A toggle switch is a device that transitions between two states in a user-controlled manner. Such a system is called *bistable*, meaning that the two states are persistent—the transition occurs only under intervention by the user.

The first engineered genetic toggle switch was constructed by Timothy Gardner, Charles Cantor and Jim Collins, and is commonly known as the Collins toggle switch (Gardner *et al.*, 2000). An interaction diagram for the gene network is shown in Figure 1.6. Two genes are involved. Each gene's protein product represses production of the other. This mutual antagonism results in a bistable system: in one stable condition, protein 1 is abundant and production of protein 2 is repressed; the roles are reversed in the other steady state. Gardner and his colleagues chose to employ two proteins whose activity could be specifically inhibited by laboratory interventions. Inactivation of the abundant protein induces an increase in the concentration of the other, resulting in a transition between the two stable states.

If the two genes and their products have symmetric properties, intuition suggests that the switch will operate as described above. However, if it is asymmetric, the network may fail to exhibit bistability—the 'stronger' protein might always dominate in their competition of mutual repression. This fact posed a challenge to Gardner and his colleagues. They could only select genetic components from the small collection of suitable genes that had been characterized in existing organisms. Whichever components they chose, there would be asymmetry in the network, and so bistability could not be assured.

Instead of carrying out an exhaustive experimental search for combinations of component that would function properly, Gardner and his colleagues developed a dynamic mathematical model to guide their design choices. Rather than fit the model to a particular instance of the design, they

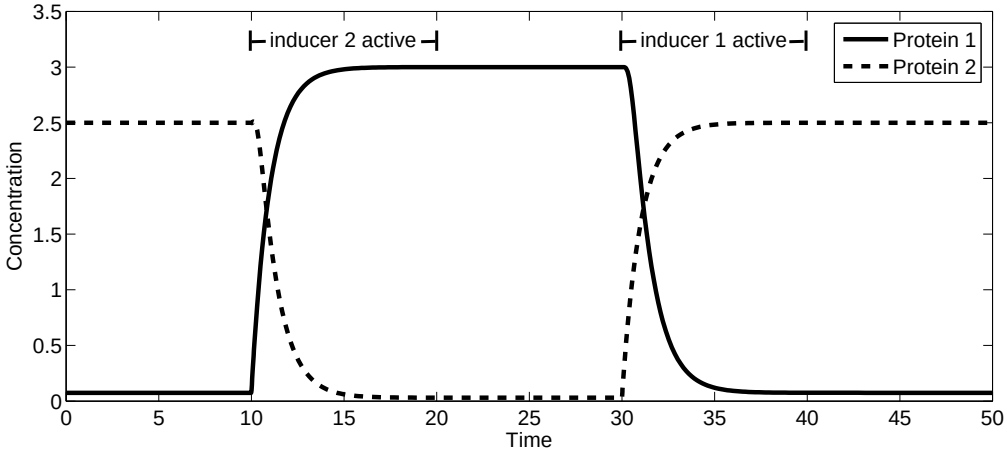

Figure 1.7: Simulation of the toggle switch model. At time zero, the system is in steady state, with protein 2 abundant and protein 1 at a low concentration. At time $t = 10$, inducer 2 is introduced, repressing protein 2 and allowing protein 1 to be expressed. Protein 1 then rises to dominance and inhibits production of protein 2. When the inducer is removed (at $t = 20$) the system remains in the high-protein-1, low-protein-2 state. At time $t = 30$, inducer 1 is introduced, causing a return to the original stable state. Units are arbitrary.

constructed a simple generic model and used it to investigate the behaviour of a wide variety of potential designs. Figure 1.7 shows a simulation of their model displaying the desired behaviour.

Their analysis led to two useful conclusions. Firstly, the model demonstrated that bistability cannot be achieved (even in the symmetric case) if the rates of gene expression are too low. Secondly, the model indicated that nonlinearity in the protein-DNA interactions could compensate for asymmetry in the genes and proteins. Some degree of nonlinearity is critical in a functioning switch; the more significant the nonlinearity, the more forgiving the constraints on symmetry. Guided by these insights, Gardner and his colleagues were able to construct a number of successful instances of the genetic toggle switch.

Since the year 2000, a wide range of synthetic biological devices have been constructed. (See (Khalil and Collins, 2010) for a review.) Model-based design is a common feature of these projects. This analysis often begins with a simple generic model that play an exploratory role in addressing possible designs. Such 'toy' models are often used for the same purpose in scientific exploration—they provide 'proof of principle' hypotheses that can then be improved as further experimental evidence becomes available.

The Collins toggle switch will be covered in more depth in Section 7.2.3. Tools for analysing bistable systems will be introduced in Section 4.2. The nonlinearities in the toggle switch design are a result of cooperative binding effects, discussed in Section 3.3.

1.6.4 Establishing the mechanism for neuronal action potential generation

Neurons, the primary cells in the animal nervous system, encode information in the electrical potential (i.e. voltage difference) across the cell membrane. This information is communicated by *action potentials*—sweeping changes in membrane potential that propagate along the length of the cell.

In a series of papers published in 1952, Alan Hodgkin and Andrew Huxley, along with Bernard

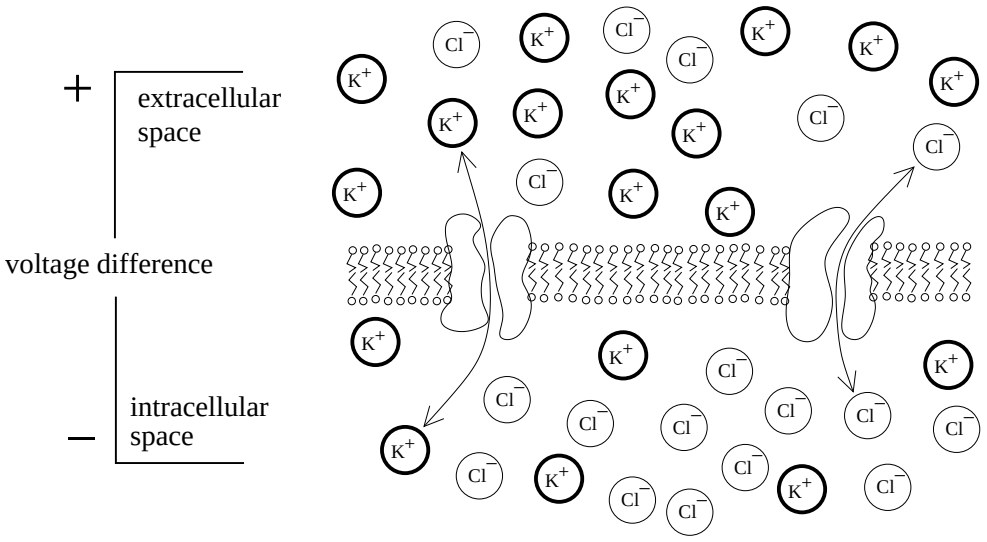

Figure 1.8: Electrical activity at the neuronal membrane. The difference in ionic charge across the membrane results in a voltage difference. The membrane contains ion-specific channels, each allowing only one type of ion to pass. These channels are voltage-sensitive—they open or close depending on the transmembrane voltage. This voltage-sensitivity sets up a feedback loop: changes in voltage lead to changes in ion-transfer rates, which lead to subsequent changes in voltage, and so on. The interplay of these feedback loops can generate action potentials.

Katz, described a biophysical mechanism for the generation of action potentials. They confirmed the behaviour of their proposed mechanism with a dynamic mathematical model (reviewed in (Rinzel, 1990)).

Prior to their work, it had been established that individual ions, such as Na^+ and Cl^-, are the primary carriers of electrical charge at the cell membrane, and that cells maintain very different concentrations of these ions in the intracellular and extracellular spaces. Moreover, it had been hypothesized that changes in the transmembrane voltage cause changes in the permeability of the membrane to these ions. These changes in permeability can result in significant ion flux across the membrane and thus produce further changes in transmembrane potential. Hodgkin and Huxley, using newly developed laboratory techniques, carried out a series of experiments showing that membrane permeability is ion-specific. They (correctly) hypothesized that this specificity is a result of ion-specific channels that are lodged in the membrane (as illustrated in Figure 1.8) and that these channels are sensitive to membrane potential.

To verify that their hypothetical mechanism was capable of generating action potentials, Hodgkin and Huxley developed a dynamic model of membrane voltage and ion transport. Simulations of their model replicated neuronal behaviour in a range of conditions, providing significant support for their hypothetical mechanism.

The generation of action potentials is referred to as *excitable* behaviour, and depends on a specific voltage threshold, as shown in Figure 1.9. Small electrical perturbations cause no significant response—voltage quickly relaxes to the pre-stimulus level. However, for perturbations that exceed a certain threshold, the response is a dramatic pulse in transmembrane voltage: an action potential. Excitability is caused by coupled positive and negative feedback loops that act between the transmembrane voltage and the voltage-gated channels in the membrane. Complex dynamic

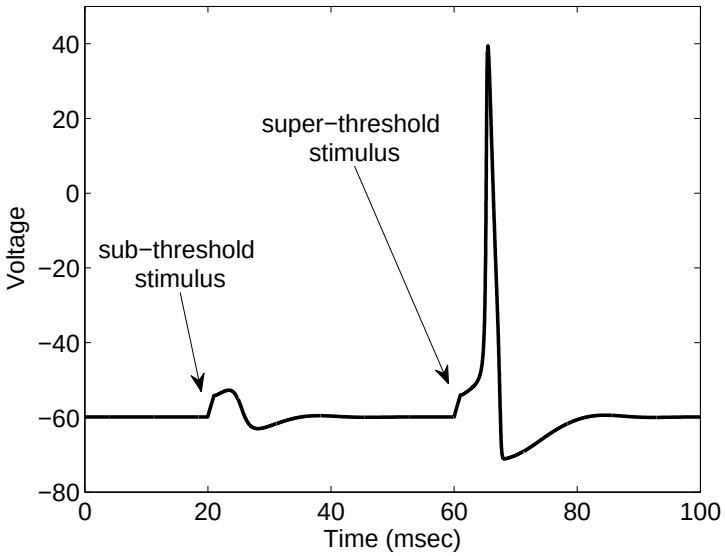

Figure 1.9: Simulation of the Hodgkin-Huxley model. The membrane voltage, originally at rest, is perturbed at time 20 milliseconds. The disturbance lasts one millisecond; once removed, its effects are quickly washed away as the voltage relaxes to its pre-disturbance level. A second disturbance, of the same form, occurs 40 milliseconds later, and again lasts for one millisecond. This disturbance, which is only 3% stronger than the first, elicits a dramatically different response; this wide excursion in voltage is characteristic of an action potential.

behaviours such as this can only be rigorously studied through mathematical modelling.

The Hodgkin-Huxley framework for modelling neuronal behaviour is introduced Chapter 8. The original Hodgkin-Huxley model is presented in Problem 8.6.4.

1.7 Suggestions for Further Reading

- **Computational Systems Biology:** This book focuses on a few fundamental modelling approaches in systems biology. Wider surveys of the tools used in computational systems biology can be found in *Systems Biology: a textbook* (Klipp *et al.*, 2009), *System Modelling in Cellular Biology* (Szallasi *et al.*, 2006), *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Alon, 2007), and *A First Course in Systems Biology* (Voit, 2012).

- **Dynamic Modelling in Molecular Cell Biology:** Several texts focus on modelling of particular biological domains. The books *The Regulation of Cellular Systems* (Heinrich and Schuster, 1996), and *Kinetic Modelling in Systems Biology* (Demin and Goryanin, 2009) focus on modelling in metabolism. *Computational Modeling of Gene Regulatory Networks* (Bolouri, 2008) addresses modelling formalisms used to study genetic networks. The use of modelling in synthetic biology is addressed in *Engineering Genetic Circuits* (Myers, 2010). Modelling of neuronal systems is surveyed in *Mathematical Foundations of Neuroscience* (Ermentrout and Terman, 2010).

- **Mathematical Modelling:** *Modeling the Dynamics of Life* (Adler, 2004) is an introductory

calculus text with an emphasis on dynamic modelling in biology. A more advanced treatment of differential equations in this context is provided in *Differential Equations and Mathematical Biology* (Jones *et al.*, 2009). Nonlinear dynamics is introduced in the text *Nonlinear Dynamics and Chaos: with applications to physics, biology, chemistry, and engineering* (Strogatz, 2001).

- **Mathematical Biology:** Texts in Mathematical Biology often cover intracellular processes, and typically introduce a range of modelling tools used in the field. These include *Computational Cell Biology* (Fall *et al.*, 2002), *Mathematical Models in Biology* (Edelstein-Keshet, 2005), and *Mathematical Physiology* (Keener and Sneyd, 1998).