# Chapter 1

# Introduction

Many instances of temporal patterns are found in biology. Examples span all levels of biological hierarchy, and recent studies include investigations of fluctuations within individual protein molecules [39], oscillations of protein concentrations inside cells [38], initiation and propagation of action potentials in specialized cells like neurons [4, 14], formation of spatial patterns of cells during organ development [49], spatiotemporal rhythms inside organs like the heart [21, 36], and population growth of species within ecosystems [3].

## 1.1 Genetic Circuits and Single-Cell Dynamics

As the smallest living unit that can self-replicate, the cell occupies a special position in the hierarchy of the biological systems. Indeed, living organisms can be unicellular, like bacteria or yeast, or multicellular, like plants and animals. In addition to self-replication, cells can exhibit other behaviors [1, 8–10, 64]: generating energy by metabolizing nutrients, searching for nutrients, maintaining circadian rhythms, responding to environmental changes. Many of these processes are inherently dynamical, and are observed to function reliably in a range of different environmental conditions. These behaviors are programmed inside cells in the form of their DNA, in structural units called genes. Genes produce proteins, which act individually or in combination with other proteins as molecular machines to perform specific functions: replicating DNA, digesting sugar, forming motors, regulating gene expression, modifying protein activity. Interactions between genes and proteins, called genetic circuits, form the basis of cellular behavior.

Reducing cellular behavior to interactions within their underlying genetic circuits is a fundamental problem in cell biology. While previous attempts to use mathematical mod-

els were considerably successful in shedding light on this problem [32, 40], they were only indirectly linked to molecular mechanisms, and also not equipped with direct methods to test model assumptions and predictions. However, two major developments in recent times have the potential to significantly overcome these limitations: First, significant effort has been invested in identification of genes, proteins, and their interactions, and this process has only accelerated over the last few years. These projects have uncovered a complex network of interactions in genetic circuits, rivaling the complexity in advanced engineering networks. Second, new tools like fluorescent proteins have been combined with time-lapse microscopy methods to usher in a new era of quantitative temporal measurements in individual cells, or movies. Such movies can provide measurements at the level of individual cells, revealing features that may have been previously averaged over in cell populations. Further, they have the potential to highlight the subset of regulatory links that are active during a cellular process, and provide a direct means to verify temporal patterns that typically arise from mathematical models. Efforts to combine these advances in genetic circuit identification and measurement techniques with mathematical models have the potential to offer more complete solutions to this problem. Additionally, such models can offer succinct summaries of cellular behavior, serving to highlight unifying features across other biological and engineering processes.

Recent instances of such efforts include investigations of bistability in cellular signaling [42], excitability in transient differentiation [58, 59], limit cycles underlying circadian rhythms [52], integral feedback control in bacterial signaling [54, 67], regulatory mechanisms during heat shock responses [18], pattern formation during embryonic segmentation [50], and spatiotemporal oscillations regulating cell division [30]. Complementarily, they can also serve as computational design aids in synthesizing genetic circuits to generate new and useful cellular dynamics, like switches [25], clocks [20, 57], and patterns [6, 16].

## 1.2 Phosphorelay: Circuit Structure and Signaling Architecture

Terminal differentiation is an important type of cellular behavior that recurs in numerous biological contexts. Examples in eukaryotes include neuron development [27], maturation of *Xenopus* oocytes [66], cell death by apoptosis [55], meiosis in yeast [41], and flowering

in plants [63]. Despite significant effort on identifying the molecular circuit controlling these processes, it often remains unclear how the approach to a terminal state plays out dynamically at the level of individual cells. One of the best-studied terminal differentiation processes is sporulation of the bacterium *Bacillus subtilis*, through which a vegetative cell under nutritional stress transforms into a stable, dormant spore [46] (see also Fig. 1.1). In some conditions, progression towards the terminal state spans several cell cycles, after which a precise sequence of molecular events remodels the cell into a spore [19, 35, 62]. The number of these pre-sporulation cell cycles can vary even among genetically identical cells in the same microenvironment. Although much of the genetic circuitry regulating sporulation initiation is known, it is still unclear how cells control the different timescales and the sequence of events leading to differentiation.

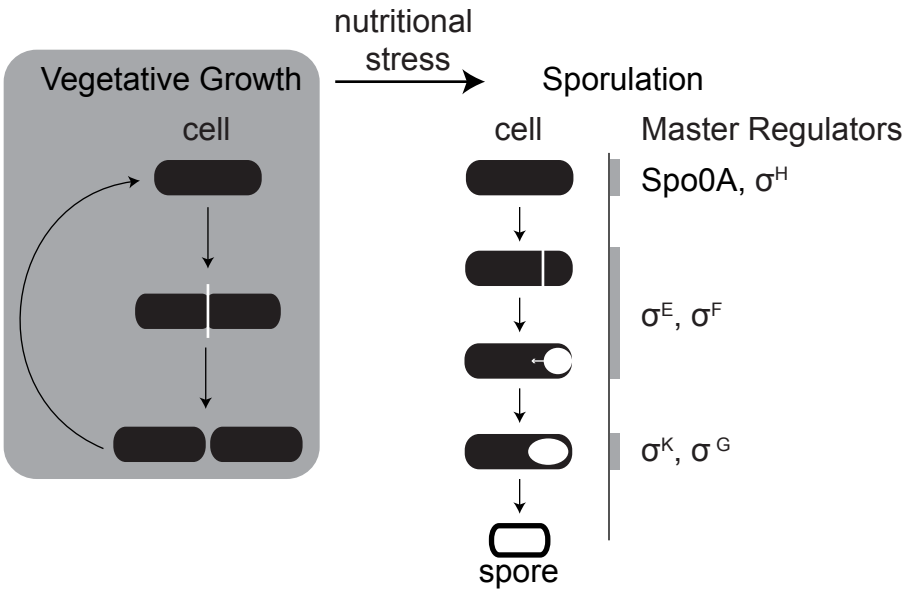


Figure 1.1: A vegetatively growing *B. subtilis* cell divides symmetrically into two daughter cells, each of which is capable of further growth and division. Under nutritional stress, such a cell can cease growth and transform into a stable, dormant spore. After the initiation of sporulation, the cell divides asymmetrically into two cells of different sizes. While both cells are required for spore formation, they have different fates. The smaller cell develops into the spore, whereas the larger cell engulfs the smaller cell, aids its development, and eventually lyses. This series of morphological changes is coupled to distinct programs of gene expression, orchestrated by a set of master regulators.

**Progression to sporulation is regulated by a phosphorelay signaling circuit em-**

**bedded in transcriptional feedback loops.** *B. subtilis* cells control sporulation by modulating the phosphorylation and expression of the master transcription factor Spo0A. The phosphorylation of Spo0A is controlled by a four-component phosphorelay, while its expression is controlled by Spo0A itself, as well as other regulators, some of which are also under the direct or indirect control of Spo0A (Fig. 1.2) [12, 43–45]. The inputs to this signaling circuit are five sporulation kinases, KinA–KinE (for simplicity, only KinA is shown in the figure), which autophosphorylate in response to nutrient limitation and other stresses, allowing them to transfer phosphates to Spo0F, which are then reversibly relayed via Spo0B to the master regulator Spo0A [12]. Additionally, Spo0F and Spo0A are dephosphorylated by the Rap and Spo0E family of phosphatases, respectively [43–45]. The KinA–KinE kinases can also act as phosphatases for Spo0F [11, 51]. The phosphorylated form of Spo0A, denoted Spo0A∼P, controls the expression of *kinA*, *spo0F*, and *spo0A* itself (but not *spo0B*) (Fig. 1.2), forming several feedback loops, which could be critical for the all-or-none nature of sporulation initiation [61].

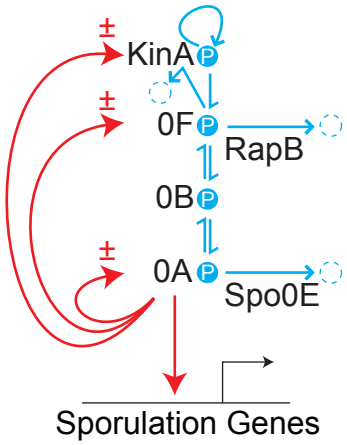


Figure 1.2: Diagram of the sporulation initiation circuit in *B. subtilis*. The main phosphorelay is embedded inside multiple transcriptional feedback loops (red arrows). Kinase autophosphorylation, phosphotransfers, and phosphatase activities are shown in blue arrows. RapB and Spo0E are examples of phosphatases that remove phosphate from the indicated proteins.

The feedback loops in this sporulation circuit have three striking features: First, the activities of the promoters controlling *spo0A*, *spo0F*, and *kinA* (referred to as $P_{0A}$, $P_{0F}$, and $P_{kinA}$ in what follows) respond in a "bandpass" manner to varying concentrations of

Spo0A∼P. That is, they are activated by low levels of Spo0A∼P, and repressed by high levels of Spo0A∼P [24], similar to the type of regulation shown to occur in the $P_{RM}$ promoter of phage lambda in response to CI [47]. Second, as discussed through a model in Chapter 2, there is an additional post-translational "bandpass" regulatory effect due to the dual role of Spo0F, which is required for Spo0A phosphorylation but can also lead to Spo0A dephosphorylation, due to reverse phosphotransfer and the activity of Spo0F phosphatases. These effects can cause net phosphorylation of Spo0A to first increase, and then decrease, as Spo0F levels rise [13]. Third, gene expression during progress to sporulation occurs in a pulse-like fashion every cell cycle [35, 62]. In particular, Spo0A's target promoters, including $P_{0A}$ and $P_{0F}$, pulse once per cell cycle, implying a periodic modulation of the phosphorelay activity, possibly driven by modulation of kinase activity. As a result of these features, models based only on continuous, monotonic interactions between components are inadequate to explain the dynamic behavior of this system.

**Phosphorelays are an architectural variant of a canonical two-component signaling system.** The genetic circuit described above is a classic example of a signaling circuit. In cells, signaling circuits provide a link between environmental variables, such as light intensity, pheromone levels, and availability of nutrients, and the activities of various transcriptional and post-transcriptional regulators. Signaling circuits exist in many different forms: The simplest circuits consist of enzymes like LacI whose transcriptional activity is modulated by the presence or absence of a bound sugar molecule [60]. Another large category of signaling circuits is based on modulating the activity of a protein by phosphorylation. These include mitogen-activated protein kinase (MAPK) signaling cascades that occur in numerous eukaryotic contexts from yeast to humans [48], and two-component signaling circuits ubiquitous in bacteria [29, 56].

A striking feature of two-component system circuits is that they exist in different architectures (Fig. 1.3). The simplest two-component system architecture has three reactions (Fig. 1.3a,b)— input-dependent autophosphorylation of a sensor kinase, phosphotransfer from the sensor kinase to a response regulator, and dephosphorylation of the response regulator [56]. This phosphotransfer takes place from a unique histidine residue on a domain of the sensor kinase (green block in Fig. 1.3a,b) to a unique aspartate residue on a domain of the response regulator (red block in Fig. 1.3a,b). In many cases, the kinase has

an additional role in dephosphorylating the response regulator, and this bifunctionality can make the response robust to the concentration of the two components via an integral feedback mechanism [7, 51, 54] (Fig. 1.3c). More complex two-component system architectures include a cascade of phosphotransfers across several protein components, such as in the phosphorelay [12] (Fig. 1.3d), which can integrate more inputs than a simple two-component system. Often, two or more components in a phosphorelay exist as separate domains of the same protein component (Fig. 1.3e) [26]. The potential for multiple phosphorylation in these hybrid proteins has been shown to increase the magnitude of response sensitivity [33].

An examination of the phosphorelay architecture and its operation during sporulation progression reveals three striking aspects: First, because it has more components than the simple two-component system architecture, it can integrate more inputs, even if multiple inputs act on the same component. Second, as stated above, it can exhibit a post-translational bandpass effect, which might play a critical role in its operation during sporulation progression. Third, it highlights how different contexts of the phosphotransfer reaction allow the generation of architectural diversity in two-component systems (Fig. 1.3). For example, a cascade of phosphotransfers gives rise to a phosphorelay. On the other hand, coupling phosphotransfer with phosphatase activity of the kinase generates the bifunctional two-component system architecture. Finally, constraining phosphotransfer to be intramolecular rather than intermolecular generates hybrid phosphorelays. These considerations offer opportunities to revisit questions of how cells can potentially utilize the phosphorelay architecture, and, more fundamentally, the phosphotransfer reaction.

## 1.3 Thesis Overview

In Chapter 2, we examine how pulsed bandpass feedback loops in the phosphorelay circuit operate in individual cells during sporulation progression. For this, we combine mathematical modeling of the phosphorelay circuit with single-cell measurements of key circuit interactions. We note the presence of an additional post-translational layer of bandpass regulation in the model. The model also predicts the appearance of a delayed phase shift in the pulsed activities of circuit components, which we verify experimentally using time-lapse fluorescence microscopy. The dynamical structure of the model contains an alternate cellu-

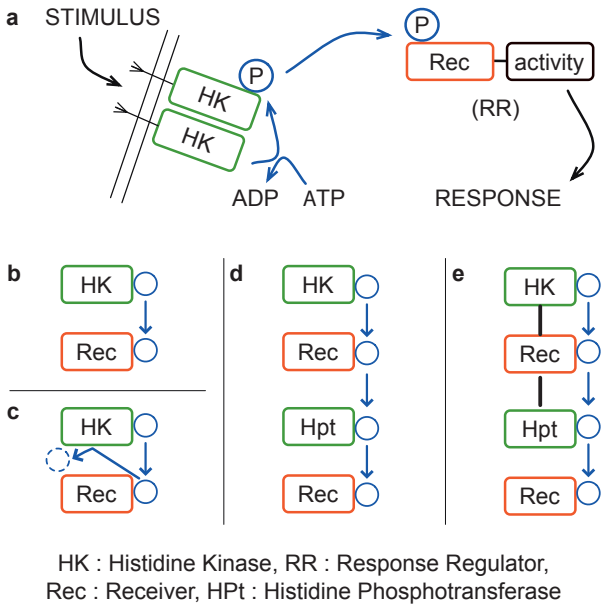


Figure 1.3: Signal transduction in two-component systems. (a) Stimulus levels are transduced into phosphates by a histidine kinase (HK). These phosphates are transferred from the histidine kinase to the receiver domain (Rec) of a response regulator protein. The phosphorylation status of the receiver domain regulates the activity of the response regulator. Red and green borders indicate protein domains where the phosphate is on the histidine and aspartate residues, respectively. (b) Schematic of a simple two-component system showing phosphotransfer from the histidine kinase to the receiver domain of the response regulator. (c) Architectural variant of the two-component system where the histidine kinase has an additional role in dephosphorylating the receiver domain of the response regulator. (d) Phosphorelay architecture where phosphate is transferred from the histidine kinase to the receiver domain of the response regulator via intermediatory phosphotransfer steps: First phosphotransfer is from the histidine kinase to a protein with only a receiver domain, second phosphotransfer is from this receiver domain to the histidine residue of another phosphotransferase protein (Hpt), and third phosphotransfer is from this residue to the receiver domain of the response regulator. (e) Two-component system architecture in which phosphotransfers occur between domains of the same protein. Thick black lines connect domains that exist on the same protein component.

lar state with no phase shift. This feature can be accessed with a circuit perturbation, and is also verified experimentally. These results illustrate the dynamics possible in a genetic circuit with pulsed bandpass feedback loops, and highlight the interplay between timescales of an external periodic input and the circuit components. We have recently submitted these results to the journal *PLoS ONE*.

Chapter 3 examines the architectural significance of the core phosphorelay against the backdrop of the family of two-component systems. Here, we address this issue computationally based on how the phosphorelay circuit structure operates in the physiological context of sporulation. Further, we explore how the phosphotransfer reaction, a basic building block of the phosphorelay and simpler two-component systems, affects the stimulus-response curve. We find that the two phosphorelay phosphatases can serve as tuning parameters to adjust the two thresholds of the post-translational bandpass response. We also find that the phosphotransfer reaction can be used to increase the range of sensitivity of response to stimulus, thereby allowing two-component systems to operate as signaling amplifiers. A similar conclusion to this result, also presented here [53], has been recently reported using a complementary approach [15]. These results emphasize signaling properties of two-component system architectures that may explain their role inside cells, and we discuss how these roles have broad correlations with the design of electrical circuits.

These results on the dynamical implications of the phosphorelay circuit structure and the signaling consequences of its architectural features are summarized in Chapter 4. In this chapter, we also propose theoretical and experimental directions for further investigation: First, we highlight molecular interactions and methods to identify their mechanisms that might be crucial to understand the dependence of sporulation initiation on circuit dynamics. Second, we suggest how the observed temporal patterns might assist mechanical processes immediately after initiation of sporulation. Third, we propose using the universal engineering constraint of power consumption to understand the role of architectural diversity in two-component systems. These directions have the potential to deepen our understanding of the relationship between phosphorelay dynamics and sporulation initiation, and transfer insights from engineering design to genetic circuits.

In this thesis, we make several primary contributions. First, we characterize, at the single-cell level, key bandpass input-output interactions in a phosphorelay circuit underlying a canonical terminal differentiation process. Second, we develop a simple mathematical

model based on these measurements, and analyze its operation under conditions prevailing in real cells. The model predicts the emergence of a delayed phase shift between circuit components, whose existence we verify experimentally using time-lapse fluorescence microscopy. Third, we predict an alternate cellular state in the model that can be accessed using a perturbation to the circuit, and show that similar behavior can be observed experimentally. Fourth, we explore the architectural potential of the core phosphorelay in the context of the superfamily of two-component system architectures, showing that two types of phosphorelay phosphatases enable independent tuning of a post-translational bandpass response, and that phosphorelay building blocks can be used to construct linear signaling amplifiers.