Bistable responses in bacterial genetic networks: Designs and dynamical consequences

Abhinav Tiwaria, J. Christian J. Ray, Jatin Narula, Oleg A. Igoshin

Department of Systems Biology, Unit 950, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77054, USA

Department of Bioengineering, Rice University, MS-142, 6100 Main St., Houston, TX 77005, USA

Article history:
Available online 6 March 2011

A key property of living cells is their ability to react to stimuli with specific biochemical responses. These responses can be understood through the dynamics of underlying biochemical and genetic networks. Evolutionary design principles have been well studied in networks that display graded responses, with a continuous relationship between input signal and system output. Alternatively, biochemical networks can exhibit bistable responses so that over a range of signals the network possesses two stable steady states. In this review, we discuss several conceptual examples illustrating network designs that can result in a bistable response of the biochemical network. Next, we examine manifestations of these designs in bacterial master-regulatory genetic circuits. In particular, we discuss mechanisms and dynamic consequences of bistability in three circuits: two-component systems, sigma-factor networks, and a multistep phosphorelay. Analyzing these examples allows us to expand our knowledge of evolutionary design principles networks with bistable responses.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Living cells react to external stimuli by mediating specific responses that are governed by the dynamics of underlying biochemical and genetic networks. Evolutionary design principles have been well studied in networks that display graded responses, with a continuous relationship between input signal and system output. Alternatively, biochemical networks can exhibit bistable responses such that the network possesses two stable steady states over a range of signals.

The possibility of bistability in simple genetic and metabolic networks has been realized for quite some time. One of the first experimental observations of bistability dates back more than 50 years to Novick and Weiner, who characterized induction of the lactose (lac) operon with a gratuitous inducer [1]. They showed the existence of a range of inducer concentrations for which cells can be in either an ‘off’ state, in which the lac operon is not expressed, or an ‘on’ state, in which the lac operon is fully induced. In this intermediate range of inducer concentrations, the composition of the cell population will depend on its history: initially fully induced cells will remain in the ‘on’ state for many generations, whereas initially uninduced cells will remain mostly ‘off’ and will have a small probability of switching to the ‘on’ state. Later, single-cell experiments confirmed the conclusions of Novick and Weiner [1] and related stochastic switching between states to the underlying stochasticity in bacterial gene expression [2–5].

As classical mechanisms of gene regulation were being discovered, researchers realized that certain circuits can display multiple steady states. In 1961, Jacob and Monod [6] proposed several such circuits based on the known regulatory elements contained within a positive feedback, which can either be direct or result from a combination of two negative interactions (a double-negative feedback). Even without experimental evidence, Jacob and Monod realized that these or similar circuits might explain cell differentiation. Since these early studies, many examples of bistable developmental switches have been identified. Among these genetic switches are those controlling the alternative ‘lifestyles’ of phage λ [7–9], the induction of maturation in Xenopus laevis oocytes [10,11], cell cycle progression [12–14], and cell fate determination in the sea urchin [15–17] and hematopoietic stem cells [18,19]. In addition, several synthetic bistable switches have been constructed [20–24].

What are the characteristics of a bistable switch? First, the steady-state signal–response curve (mathematically speaking, a one-parameter bifurcation diagram of the underlying dynamical system) contains a range of signals at which two different steady-state
responses are possible. This curve consists of three branches; two of them represent the stable steady states, and the intermediate branch represents the unstable steady state (Fig. 1a). As the intermediate branch is unstable, a signal corresponding to Point 2 (which lies within the range of bistability) will result in either of the two stable branches, depending on the initial conditions. Such curves can be easily computed from a deterministic mathematical model of the underlying network. In the case of the lac operon, the two steady states correspond to two levels of lac operon expression (response) at the same level of extracellular inducer (signal). At the boundaries of the bistable signal range, the steady-state response of the system discontinuously jumps from one state to the other (arrows in Fig. 1a). Note that this discontinuous jump in the steady state does not indicate a fast dynamic response to a signal that crosses the threshold. In fact, the second characteristic of a bistable switch is a slow response to a signal near the switching threshold (Fig. 1b). In addition, stochastic models of bistable switches can reveal other dynamic properties. In single cells, slow switching in response to an above-threshold signal will lead to a very noisy response with heterogeneous switching times in the population (Fig. 1b). This heterogeneity may manifest as a transient bimodal distribution in the population. A bimodal distribution is also expected in populations responding to a signal in the bistable range (Fig. 1c).

In this review, we discuss some conceptual network designs that produce bistable behavior. Later, we present examples of how these designs are used in bacterial master-regulatory circuits. We discuss mechanisms of bistability in two-component systems, sigma-factor networks, and a multistep phosphorelay. For each example, we point out physiologically relevant dynamical consequences of bistability. Analyzing these examples allows us to expand the knowledge of evolutionary design principles of biochemical networks with bistable responses.

2. Conceptual network designs of bistable mechanisms

2.1. Positive feedback with cooperativity

One of the most widely accepted and studied mechanisms through which bistability can be attained in a genetic circuit is a direct or indirect transcriptional positive feedback characterized by a kinetic order greater than one (cooperativity), so that the dependence of the expression rate on the transcription factor (TF) is superlinear. This mechanism is sufficient to produce bistability for a wide range of parameter values. Fig. 2a illustrates one of the simplest examples of such a mechanism. Protein $A$ is expressed from a promoter autogenously regulated by its own homodimer, $A_2$. A simple model for this system has the following kinetic equations:

\[
\frac{dA}{dt} = \beta + \frac{vA_2}{K + A_2} - 2k_dA^2 + 2k_dA_2 - k_{deg}A
\]

(1)

and

\[
\frac{dA_2}{dt} = 2k_dA^2 - 2k_dA_2 - k_{deg}A_2,
\]

(2)

where $A$ and $A_2$ are the concentrations of monomer $A$ and activator dimer $A_2$, respectively; $\beta$ and $v$ are the basal and maximal synthesis rates of monomer $A$, respectively; $K$ is the equilibrium dissociation constant of dimer $A_2$ from the promoter; $k_d$ and $k_{deg}$ are the rate constants for dimer association and dissociation, respectively; and $k_{deg}$ is the protein degradation rate (for stable proteins in bacteria, this degradation is dominated by dilution due to growth and thus reflects the doubling time).

Assuming the quasi-steady-state approximation for the kinetics of dimer formation in Eq. (2) and using the result obtained in Eq. (1), the rate of change of $A$ (dA/dt) can be plotted as a function of $A$ (Fig. 2b). The quasi-steady-state assumption is justified biologically as protein production and degradation processes are slower than the post-translational reactions. This assumption is used here to graphically illustrate the existence of bistability, but the resulting conclusions can be generalized beyond this approximation. The intersections with the dashed line (dA/dt = 0) define the steady states of the network. The two filled circles represent the stable steady states, and the open circle represents the unstable steady state. The existence of bistability depends on the kinetic parameters of the network: for some parameter values, the inflection points of the curve fall on opposite sides of the dashed line, whereas for others, this is not the case and the system possesses only one (physically meaningful) steady state.

2.2. Positive feedback without cooperativity: post-translationally generated ultrasensitivity

In the previous example, dimerization of the activator is necessary to produce the superlinear transcriptional input that is required for bistability. However, for TFs that do not undergo dimerization and therefore function as monomers, positive transcriptional feedback does not lead to bistability in the system (dashed gray curve in Fig. 2d). Not all transcriptional activators function as high-cooperativity multimers; what mechanisms can provide superlinearity in these cases? One way to achieve superlinearity is by activating the TF via a post-translational network that is ultrasensitive, in which a sharp transition occurs between inactive and active forms of the TF. For example, ‘zero-order ultrasensitivity’ can be observed in multistep or reversible covalent modification cascades as long as one of the enzymes involved operates near saturation (zero kinetic order) [25–27].

---

**Fig. 1.** Characteristics of a bistable switch. (a) The steady-state signal–response curve shows a range of signals for which two different steady-state responses are possible. At the boundaries of the range of bistability, the steady-state response of the system discontinuously jumps from one state to the other (arrows). The two solid curves represent the stable steady states, which are separated by the unstable steady state (dotted curve); results in a noisy response with switching-time heterogeneity in the population. The black curve corresponds to the deterministic response, whereas the gray curves correspond to simulations of the stochastic model. (c) Deterministic bistability in the system gives rise to a bimodal population distribution at steady state. Distributions are computed from the long-time limit of the Gillespie simulations at the signal corresponding to Point 2 in panel a. The two peaks correspond to the low (‘off’) and high (‘on’) steady-state responses of panel a, respectively.
Another way to achieve ultrasensitivity is via stoichiometric sequestration [28–30]. The design presented in Fig. 2c shows that if the regulatory protein B sequesters the activator A in a transcriptionally inactive complex AB, then the cooperativity of autogenous activation is not needed. The kinetic equations for this system can be formulated in terms of the total concentration of activator A, A_T = A + AB and the heterodimer concentration, AB, as follows:

\[
\frac{dAB}{dt} = k_0 A \cdot B - k_d A B
\]

and

\[
\frac{dA_T}{dt} = \beta + \frac{vA}{K + A} - k_{deg} A_T,
\]

where the parameters have the same meanings as in Eqs. (1) and (2). In Eqs. (3) and (4) the total concentration of protein B (B_T = B + AB) is assumed to be constant, which allows it to be treated as a parameter. This assumption is biologically justified when transcription of B is not regulated by activator A.

If the formation or dissociation of the AB complex is faster than its degradation, we can assume that Eq. (3) is in a quasi-steady state and use the conservation laws for the total concentrations of the network proteins A and B to express A as a function of A_T. As a result, the right-hand side of Eq. (4) can be plotted as a function of A_T (black curve in Fig. 2d). The intersections of this curve with the dashed black line (dA_T/dt = 0) are the steady states of the network. The two filled circles represent the stable steady states, and the open circle represents the unstable steady state.

Another way of demonstrating bistability is through a one-parameter bifurcation diagram (or, biologically, a steady-state dose–response curve), as discussed in Section 1 and Fig. 1a.

In this system, the existence of bistability depends on the interaction strength of the complex measured by its dissociation constant (K_D = k_d/k_0) and the total concentration of protein B (B_T). The ratio K_D/B_T is used as a bifurcation parameter in Fig. 2g. This bifurcation diagram is drawn by obtaining the steady state for A_T by equating the expression for dA_T/dt (derived for Fig. 2d from Eq. (4)) to zero and expressing A_T in terms of the composite parameter K_D/B_T. The solid and dashed curves correspond to the steady-state levels of A_T for two different sets of initial conditions.

An ultrasensitive response in active versus total TF concentrations can also be generated by alternative mechanisms. These mechanisms, including saturated degradation [31], inhibition of cell growth [32], and multisite phosphorylation [33], may play an important role in generating bistability in networks involving transcriptional regulators, such as bacterial sigma factors, that work only as monomers.

2.3. Implicit feedback: dead-end complex formation

In the preceding two examples, the bistable response was associated with the existence of an explicit transcriptional feedback in the network. However, the existence of a positive feedback is not always obvious from the network architecture, and it can result from complex interactions or conservation laws among network components (implicit positive feedback). Craciun et al. illustrated...
this point by reviewing several classical enzyme mechanisms that are capable of bistable behavior and proving a theorem stating the necessary conditions for bistability [34]. In addition, earlier work by Feinberg et al. described a rigorous theory providing sufficient conditions for multiple steady states in mass-action networks [35]. In Fig. 2e, we show another example of bistability in a simple enzymatic reaction. TF A is enzymatically degraded by enzyme E via the catalytically active complex EA. The complex can bind another molecule of A (substrate inhibition) to form the ‘dead-end’ complex EA2, from which no catalytic conversion is possible. The system is described by the following kinetic equations:

\[
\frac{dE}{dt} = \beta - k_{1}A \cdot E + \frac{k_{1}E}{C_{1}} \cdot A - k_{2}A \cdot EA + \frac{k_{2}A}{C_{2}} \cdot EA_{2} - k_{a}A,
\]

\[
\frac{dE.A}{dt} = k_{a}A \cdot E - k_{2}E.A - k_{d}E.A_{2} - k_{c}E.A,
\]

\[
\frac{dE.A_{2}}{dt} = k_{d}A \cdot E.A - k_{d}E.A_{2},
\]

where E, EA, and EA2 are the concentrations of the enzyme E, complex EA, and complex EA2, respectively; \( \beta \) is the basal synthesis rate of protein A; \( k_{0} \) is the non-specific degradation/dilution rate of protein A; \( k_{a} \) and \( k_{c} \) are the rate constants for association and dissociation, respectively, of complex EA; \( k_{d} \) and \( k_{2} \) are the rate constants for association and dissociation, respectively, of complex EA2; and \( k_{c} \) is the catalytic degradation rate. Note that, unlike the previous two examples (Fig. 2a and c), synthesis of TF A is not auto-regulated and no transcriptional feedback is present in the system. Nevertheless, as shown below, enzymatic degradation of A can lead to an implicit positive feedback and bistability.

Similar to the situation in Section 2.1, assuming a quasi-steady state for the enzyme-containing complexes in Eqs. (6) and (7) and using the conservation law for the total concentration of enzyme E, we obtain expressions for EA and EA2 as functions of A. Substituting these expressions into Eq. (5), we plot \( dA/dt \) as a function of A to demonstrate bistability (Fig. 2f). The intersections with the dashed line (\( dA/dt = 0 \)) define the steady states of the network. The two filled circles represent the stable steady states, and the open circle represents the unstable steady state. One of the steady states corresponds to low A with no dead-end complex formation and fast catalytic degradation, and the other corresponds to high A with inhibited catalytic degradation because of the abundance of dead-end complex.

In addition to the graphical solution of \( dA/dt = 0 \) (Fig. 2b, d, and f) and the bifurcation diagram (Figs. 1a and 2g), yet another graphical representation of bistability involves the investigation of nullclines. In a two-dimensional system of ordinary differential equations, nullclines represent the curves \( dx/dt = 0 \) for all i, where \( x_i \) (i \( \in \) (1, 2)) are the two variables. To obtain the equation for nullclines in this system, we first use the quasi-steady-state approximation for Eq. (6) and the conservation law for the total concentration of enzyme E to express both \( dA/dt \) and \( dE.A_2/dt \) as functions of A and EA2. Next, these expressions are equated to zero to obtain the nullclines, which are simultaneously plotted in a phase plane of EA2 versus A (Fig. 2h). The dashed curve represents the nullcline \( dA/dt = 0 \), whereas the solid curve represents the nullcline \( dE.A_2/dt = 0 \). The two filled circles represent the stable steady states (one has low concentrations of A and EA2; the other has high concentrations of both), and the empty circle represents the unstable steady state.

Although there is no apparent positive feedback in the kinetic scheme, it still exists because different enzyme forms are conserved. Formation of the dead-end complex EA2 is self-enhancing: its formation inhibits catalytic degradation and thus increases the concentration of TF A. This leads to a further increase in the concentration of dead-end complex. In summary, bistability results from the formation of a dead-end complex between the enzyme and two substrate molecules: classical substrate inhibition [36]. Generally, any mechanism of substrate inhibition will result in the same effect [37].

3. Bistability in two-component systems

Two-component systems (TCSs) are a major class of bacterial sensory apparatus that respond to specific physical or chemical stresses [38]. In a TCS, a sensor histidine kinase (SHK) responds to environmental stress by modulating the phosphorylation of a cognate response regulator (RR), which then dimersizes and becomes transcriptionally active [38]. The typical genetic structure of a TCS is a single operon with RR and SHK transcribed in that order from an initiation site just upstream of RR (Fig. 3a).

3.1. Biochemical interactions and genetic regulation in a two-component system

Typically, bacterial species have many TCSs; for example, Escherichia coli has approximately 30 distinct systems [39], most with a single SHK-RR cognate pair that is kinetically preferred [38,40]. SHKs typically respond to a set of specific environmental stresses, such as osmotic stress, shifts in extracellular acidity or ion content, and phosphate changes [41]. The mechanisms for sensing specific stresses are poorly understood, and some sensors can respond to multiple types of signal. For example, the well-characterized Salmonella typhimurium TCS PhoP/PhoQ responds primarily to Mg\(^{2+}\) depletion but also responds to depletion of other divalent cations, such as Ca\(^{2+}\) [42]. However, the overall pattern is for a single stress to result in a single response by altering a context-appropriate regulon controlled by the transcriptionally active RR [40]. A pattern of distinct signals acting in parallel (relatively well buffered from one another) permits evolutionary selection that may fine-tune responses for optimal context-dependent dynamics. Thus, a TCS may exhibit graded monostable or bistable steady-state responses without substantially altering other stress-response systems.

The prototypical SHK is a homodimeric transmembrane protein that responds to an environmental stress by autophosphorylating a histidine residue on its cytoplasmic tail [38]. Classic SHKs have a single phosphorylation domain that transfers the phosphate to an aspartate residue in the unphosphorylated RR. However, the interactions between SHK and RR are more complex than this simple model suggests (Fig. 3b). Notably, SHK phosphatase activity, a separate catalytic event that dephosphorylates phosphorylated RR (RR–P), has been recognized to play an important role in TCS response dynamics [41]. Phosphatase activity may speed up responses by increasing RR–P turnover rates but may also be an important buffer against non-cognate SHK–RR signaling [43]. Kinase and phosphatase activity may be modulated separately [e.g. 44,45], or simultaneously [46] by stress signals. The strength of SHK–RR binding and the balance of phosphotransfer to phosphatase activity of SHK often differ from system to system, with important dynamical consequences.

In most known cases, expression of RR is much higher than that of SHK [47], resulting in a large difference in their intracellular concentrations that may stem from differential processing of RR and SHK mRNA [47]. In many cases, the transcriptionally active form of RR regulates the TCS operon; this autoregulation is usually, but not always, positive [48–53]. Such apparently positive feedback loops may lead to bistability in some TCSs, as the resulting design resembles that in Fig. 2a. For instance, the Agrobacterium tumefaciens TCS VtrA/VirG displays a bimodal signal response
However, the mechanism that produces bimodality in this system is unknown. The *Mycobacterium tuberculosis* TCS MprA/MprB is similarly positively autoregulated and plays a role in inducing bistability but does so only through interactions with a stress-response sigma factor (Section 4.2; Ref. [56]). Nonetheless, bistability is possible in a positively autoregulated MprA/MprB TCS (or in TCSs in general) in a certain parameter regime, provided that fold change for gene activation is large or a dominant dephosphorylation flux is present as a result of an exogenous phosphatase ([54,55]. However, most TCSs are not characterized by these scenarios, and thus TCSs tend to be monostable.

Therefore, although the most common TCS operon genetic architecture (Fig. 3a) bears a superficial resemblance to cooperative positive feedback (Fig. 2a), this architecture is unlikely to play a role in bistability, for at least two reasons. First, feedback proportionally upregulates expression of both RR and SHK proteins. This may lead to a negative feedback effect if SHK is bifunctional and dephosphorylates RR–P ([57]. Second, many TCSs also have a signal–response relation that is nearly independent of the level of RR and SHK expression; thus, the effects of feedback on signal level are negated and TCSs function robustly regardless of gene-expression fluctuations ([58,59]). Thus, we must beware of jumping to the conclusion that positive feedback loops imply bistability.

The evolution of TCSs has favored the maintenance of cognate SHK–RR pairs at a single genetic locus, so in principle genetic or biochemical modifications could arise from the genetic background of the prototypical TCS architecture. This genetic background may then favor several possible small modifications in the course of evolution. Thus, although evidence is lacking that the prototypical TCS architecture produces bistability, several variant TCS architectures that may have arisen from this prototype can result in bistability. In Sections 3.2 and 3.3, we discuss examples that demonstrate mechanisms for attaining bistability, and the functional and evolutionary consequences of bistable TCSs.

### 3.2. Bistability in the Bacillus subtilis TCS DegU/DegS drives heterogeneous response times

In the *B. subtilis* TCS DegU/DegS, the positions of the SHK and RR genes are swapped compared with the prototypical TCS operon, so that the SHK, DegS, is upstream of the RR, DegU. Moreover, the DegU–P upregulates the *degU* gene alone from a separate promoter ([60]. The DegU/DegS TCS is a master regulator that activates, among other genes, two extracellular proteases, *aprE* (subtilisin) and *bpr* (bacillopeptidase) ([61–63]. These two proteases are also regulated by multiple repressors of gene expression and are repressed by activation from the master sporulation regulator Spo0A ([64]. Therefore, the protease expression is under the control of two input signals. Because of this regulatory logic, the stationary-phase *B. subtilis* population has three distinct subpopulations: endospore formers (i.e., sporulated cells), protease ‘on’, and protease ‘off’ ([65]. The population heterogeneity in this system may be a consequence of system bistability, which itself may arise from network architecture.

Indeed, a deterministic mathematical model for the DegU/DegS TCS shows that the positive transcriptional feedback in the network architecture can give rise to bistability in DegU–P ([65]. The bistability-generating mechanism in this network makes use of the design discussed in Section 2.1 (Fig. 3d). For a range of signal levels (DegU phosphorylation rates), two steady states are possible. The DegU ‘off’ state corresponds to inactivated levels of *degU* expression, and the DegU ‘on’ state corresponds to a nearly maximal activation of the positive feedback. Furthermore, stochastic simulations show that the time needed for the transition from the ‘off’ to the ‘on’ state can be highly variable (a characteristic property of bistable switch; Fig. 1b), as indicated by the sample trajectories in Fig. 3e ([65]. Thus, the positive transcriptional feedback and the resultant bistability in the DegU/DegS TCS appear to induce a noisy response with heterogeneous relaxation times.
Bistability in Spo0A activation (Section 5) may further increase the heterogeneity of protease production.

### 3.3. Dead-end complex between TCS sensory kinase and response regulator

Can bistability in TCSs arise independently of transcriptional feedback, as in the design discussed in Section 2.3? A biochemical TCS model developed in Ref. [66] shows how this can occur with conserved total concentrations of RR and SHK. The results indicate that bistable behavior can arise from a dead-end complex between SHK and RR (Fig. 3c). A structural analysis of the well-studied E. coli TCS EnvZ/OmpR suggests that the EnvZ.OmpR complex is likely to be dead-end because the histidine on the EnvZ SHK is not accessible to ATP for autophosphorylation [66]. If this is the case, the reaction scheme will have a dead-end complex with an implicit positive feedback.

We note that alternative views exist on the role of EnvZ in the dephosphorylation of OmpR–P, which governs the existence of the EnvZ.OmpR dead-end complex. Mattison and Kenney found evidence that OmpR–P has a low affinity for EnvZ [67]. In contrast, a subsequent finding presented evidence that the affinities of the phosphorylated and unphosphorylated forms of OmpR for EnvZ are not very different [68]. This result justifies the inclusion of a reversible reaction in the model in which the unphosphorylated forms of RR and SK form a dead-end complex (SHK.RR).

A SHK.RR dead-end complex is necessary but not sufficient for bistability. One of the following two conditions must also hold [66]: (i) SHK is monofunctional; or (ii) if SHK is bifunctional, RR–P has an alternate phosphatase independent of SHK. Intuitively, these criteria apply because SHK phosphatase activity can act as a negative interaction that ‘cancels out’ the positive feedback arising from the dead-end complex. In this case, for bistability a large fraction of the phosphatase activity must be achieved via a mechanism not involved in the implicit feedback.

Transient dynamic responses strongly depend on feedback [66], so the implicit positive feedback arising from dead-end complex formation may affect response times. Possible dynamic responses of such systems include large variability, bimodal population distributions of outputs, slow response times, and high signal capacities [66,69]. Stochastic simulations predict faster responses in a bistable model than a monostable one as long as the signal intensity is counteracted by a negative feedback from anti-sigma auto-regulation, making bistability unlikely [80]. For example, the general stress response of B. subtilis is controlled by the alternative sigma factor σ^B, which is transcribed from an autoregulated operon along with its anti-sigma factor and anti-anti-sigma factor [75]. This network is triggered by deleterious energy and environmental stimuli and is expected to display a graded response [80]. It does not exhibit a bistable response because of insufficient cooperativity in the transcriptional feedback loops. Investigation of the steady-state performance of this network revealed that the negative feedback loop associated with the upregulation of anti-sigma-factor transcription is essential to ensure gradual increase of σ^B, whereas the positive feedback loops are required for a wide range of responses [80].

Nonetheless, post-translational interactions of alternative sigma factors with their network partners may lead to bistability. In Sections 4.1 and 4.2, we present two examples of alternative-sigma-factor networks, from B. subtilis and M. tuberculosis that use variations of the designs discussed in Sections 2.2 and 2.3 to produce bistable responses. In the B. subtilis network, which controls the activity of the alternative sigma factor σ^B, bistability arises from the formation of a long-lived dead-end complex without explicit transcriptional feedback, as was the case in Section 3.3. In contrast, in the network present in M. tuberculosis and its non-pathogenic cousin Mycobacterium smegmatis, positive transcriptional feedback in the production of the mycobacterial sigma factor σ^R coupled with the stoichiometric sequestration of σ^B by the anti-sigma factor RseA leads to bistability.

### 4. Bistability in sigma-factor networks

A sigma factor is a subunit of the bacterial RNA polymerase complex that is necessary for promoter recognition and transcription initiation [71–73]. In addition to the (usually single) housekeeping sigma factor associated with simple growth and division of cells in rich environments, many alternative sigma factors play a crucial role in the specificity of transcription initiation under a variety of conditions. Each sigma factor specifically interacts with the nucleotide sequence of various promoters to regulate the expression of multiple genes. Alternative sigma factors mediate the response to various environmental changes, such as stresses arising from nutrient depletion, heat shock, or oxidative damage [74].

Controlling the level and activity of an alternative sigma factor determines the expression of the corresponding regulon. At the transcriptional level, many alternative sigma factors are autoregulated (i.e., their gene is part of their own regulon) [74]. This autoregulation results in a positive feedback that may result in bistability. However, as sigma factors bind to core RNA polymerase as monomers, no cooperativity is present in an autoregulated sigma-factor network and bistability would not be expected unless the post-translational sigma-factor activation network is ultrasensitive.

A common mechanism that post-translationally controls activity is the sequestration of a sigma factor by binding of its antagonists, called anti-sigma factors [75]. This sequestration prevents the sigma factor from binding to core RNA polymerase. This design resembles the one discussed in Section 2.2 (Fig. 2c and d) and may result in bistability. However, in many cases sigma and anti-sigma factors are co-transcribed from the same operon, which is preceded by a promoter positively autoregulated by the sigma factor [76–79]. In this case, the positive feedback from sigma autoregulation is counteracted by a negative feedback from anti-sigma autoregulation, making bistability unlikely [80]. For example, the general stress response of B. subtilis is controlled by the alternative sigma factor σ^B, which is transcribed from an autoregulated operon along with its anti-sigma factor and anti-anti-sigma factor [75]. This network is triggered by deleterious energy and environmental stimuli and is expected to display a graded response [80]. It does not exhibit a bistable response because of insufficient cooperativity in the transcriptional feedback loops. Investigation of the steady-state performance of this network revealed that the negative feedback loop associated with the upregulation of anti-sigma-factor transcription is essential to ensure gradual increase of σ^B, whereas the positive feedback loops are required for a wide range of responses [80].

Nonetheless, post-translational interactions of alternative sigma factors with their network partners may lead to bistability. In Sections 4.1 and 4.2, we present two examples of alternative-sigma-factor networks, from B. subtilis and M. tuberculosis that use variations of the designs discussed in Sections 2.2 and 2.3 to produce bistable responses. In the B. subtilis network, which controls the activity of the alternative sigma factor σ^B, bistability arises from the formation of a long-lived dead-end complex without explicit transcriptional feedback, as was the case in Section 3.3. In contrast, in the network present in M. tuberculosis and its non-pathogenic cousin M. smegmatis, positive transcriptional feedback in the production of the mycobacterial sigma factor σ^R coupled with the stoichiometric sequestration of σ^B by the anti-sigma factor RseA leads to bistability.

#### 4.1. Dead-end complex in the σ^R network results in bistability

Asymmetric cell division during sporulation in B. subtilis leads to differential gene expression in the two progeny cells with separate developmental fates [81]. This process is governed by the sequential activation of sigma factors in the two compartments formed by an asymmetric sporulation septum: a larger mother cell and a smaller forespore. The first sigma factor to be activated, σ^A, is present in both cell compartments before cell division. However,
σF is activated only in the forespore and only after the asymmetric septum is formed.

The activity of σF is controlled by a biochemical network that functions via a partner-switching mechanism first proposed by Alper et al. [82]. The essential features of this pathway and mechanism are reviewed in Refs. [81,83] and are depicted in Fig. 4a. Free σF can associate with the RNA polymerase complex (not shown in Fig. 4a) to turn on the transcription of σF-dependent genes. However, in the predivisional cells, σF is deactivated by the binding of the anti-sigma factor SpoIAB (AB). In contrast, the unphosphorylated anti-anti-sigma factor SpoIIE (AA) can bind AB and prevent it from binding and inactivating σF. Unphosphorylated AA can also attack the AB. σF complex, causing the release of σF. In the presence of ATP in its catalytic site, the AB anti-sigma factor phosphorylates the serine-58 residue of its antagonist, AA. The phosphorylated form of AA has a low affinity for AB and rapidly dissociates. A specific phosphatase, SpoIIE (denoted ‘signal’ in Fig. 4a), activates AA by dephosphorylating it. Therefore, the level of σF activity is determined by the balance between phosphorylation and dephosphorylation of the AA anti-anti-sigma factor. The dephosphorylation rate, determined by the activity or the concentration of SpoIIE phosphatase, serves as an important signal to activate σF [84]. Before the septum is formed, most of the AA is phosphorylated and is therefore incapable of interfering with the AB.σF complex. After septation, AA is mostly unphosphorylated in the forespore and induces the release of σF, thus allowing σF to initiate forespore-specific transcription.

An important feature of the σF network is the formation of a long-lived complex between AA and AB with ADP in the catalytic site (dashed box in Fig. 4a). Without the γ-phosphate of ATP, phosphorylation of AA is not possible and ATP-to-ADP exchange requires dissociation of the complex. Therefore, this is a dead-end complex as its only fate is to dissociate and resume the phosphorylation–dephosphorylation cycle. We now briefly summarize a mathematical modeling study of the σF network that revealed how the formation of the AA.AB.ADP dead-end complex gives rise to bistability [69].

The network responds to two signals that couple asymmetric septum formation to σF activation in the forespore. These morphological signals are (i) an increased rate of AA dephosphorylation by the SpoIIE phosphatase in the forespore, resulting from the association of SpoIIE with cell-division proteins and the subsequent localization of SpoIIE to the asymmetric septum [85–88] and (ii) depletion of the AB anti-sigma factor in the forespore by means of a transient genetic asymmetry between the two compartments, coupled with a fundamental instability of AB [89–91]. When the steady-state concentration of σF is plotted against either of these signals, a bistable response is observed for a range of parameter values. Fig. 4b shows the steady-state solutions for free σF concentration as a function of the dephosphorylation rate corresponding to different initial conditions: inactivated or fully activated. At low and high dephosphorylation rates, the solutions coincide. However, at intermediate dephosphorylation rates, the system is bistable. The solid curve shows the steady state that corresponds to an initially low dephosphorylation rate, so that very little dead-end complex is formed and most of the anti-sigma factor AB is in a complex with σF, as would be the case in predivisional cells. In contrast, the dashed curve shows the steady state corresponding to an initially high dephosphorylation rate, so that most of the AB is in the dead-end complex with AA and most of the σF is free, as would be the case after activation in the forespore. At the boundaries of the bistable range, the concentration of free σF jumps from one branch of the solution to the other.

The possibility of a bistable response is not obvious from the σF diagram shown in Fig. 4a as no explicit positive feedback is apparent. The existence of a bistable response was shown to be associated with the self-enhancing formation of the dead-end complex [69], a design principle discussed in Sections 2.3 and 3.3 that gives rise to an implicit positive feedback. The dead-end complex is formed after the phosphorylation step, when the fate of complex between AB and ADP depends on a competition between the rate of exchange of ADP for ATP and the rate of AA binding. Note that the self-enhancing formation of the dead-end complex serves as an intuitive explanation for the bistable response, but it cannot replace the mathematical model in predicting whether this behavior can be observed. In other words, the implicit positive feedback due to self-enhancing formation of the dead-end complex is required for bistability and was shown by modeling to be sufficient for a specific range of parameter values [69]. In particular, formation of the dead-end complex crucially depends on the presence of excess anti-anti-sigma factor AA relative to anti-sigma factor AB [69]. One way to achieve AA excess is via the proposed depletion of AB in the forespore [89,91]. Indeed, genetic and biochemical experiments have shown that AB depletion can partially compensate for decreased SpoIIE activity caused by mislocalization of the phosphatase [89,91].

In addition to affecting the steady-state network performance, bistability associated with the dead-end complex also influences network dynamics. A sudden increase or decrease in the dephosphorylation rate affects the dynamic behavior of free σF, such that

![Fig. 4](image-url)

**Fig. 4.** Implicit feedback due to the dead-end complex leads to bistability in the σF network. (a) σF activity is regulated by a network that functions via a partner-switching mechanism. σF is deactivated by the binding of anti-sigma factor AB, whereas unphosphorylated anti-anti-sigma factor AA can bind AB and prevent it from binding and inactivating σF. Unphosphorylated AA can also attack the AB.σF complex, causing the release of σF. In the presence of ATP, the anti-sigma factor AB can phosphorylate its antagonist AA, which has low affinity for AB and rapidly dissociates. AA is phosphorylated via dephosphorylation by an external phosphatase, SpoIIE (signal). An important feature of the σF network is the formation of a self-enhancing dead-end complex (dashed box) of AA, AB, and ADP. The only fate of this complex is to dissociate and resume the phosphorylation–dephosphorylation cycle. (b) The network displays bistable behavior as the system possesses two stable steady states over a range of dephosphorylation rates (signals). Solid and dashed curves correspond to different initial conditions. (c) Dynamic response of free σF concentration to an instantaneous increase (solid curve) or decrease (dashed curve) in signal. Graph shows that the bistable switch is much slower to turn ‘off’ than to turn ‘on’.
turning $\sigma^E$ activity ‘on’ is much faster than turning it ‘off’ (Fig 4c). The time needed to turn ‘off’ $\sigma^E$ activity is determined by the dissociation of the dead-end complex; therefore, the formation of this complex contributes to the slow ‘off’ response. Computing the transient response of a hypothetical mutant that lacks the dead-end complex demonstrates this (see [69] for details). As one would expect, turning $\sigma^E$ activity ‘off’ is faster in the hypothetical mutant than in the wild-type cells.

4.2. Large interaction cooperativity in the $\sigma^E$-RseA module leads to bistability

Mycobacterial stress responses are triggered by various environmental conditions, such as nutrient or oxygen depletion, heat shock, and exposure to oxidizing agents [92–94]. Such stress responses have recently been associated with the phenomenon of bacterial persistence, in which a fraction of a genetically identical population can survive exposure to stress by reduction or cessation of growth [5,95]. Persistence exhibited by the pathogen M. tuberculosis allows it to successfully colonize host cells by avoiding elimination via the immune response or drugs [96]. With very slow or no replication, these pathogens can persist in a latent state for years [97]. Although several genes in M. tuberculosis have been associated with persistence [98–100], the precise mechanism governing the phenotypic switch to persistence remains unknown.

A part of the mycobacterial stress-response network consisting of the MprA/MprB TCS that interacts with the alternative sigma factor $\sigma^E$ and its anti-sigma factor RseA (Fig. 5a) has recently been modeled [56]. In this network, external stress serves as an input signal that triggers the MprA/MprB TCS. Activation of the MprA/MprB TCS uses a biochemical mechanism characteristic of TCSs (Section 3.1 and Fig. 3b): phosphorylation of MprA produces transcriptionally active MprA-P, which controls the operon encoding the sigma factor $\sigma^E$. Apart from this regulation, $\sigma^E$ activity is post-transitionally regulated by the anti-sigma factor RseA, which is encoded in the operon downstream of $\sigma^E$ and is not controlled by MprA-P. Like any other anti-sigma factor, RseA binds to $\sigma^E$ and prevents its association with the RNA polymerase. The various transcriptional interactions present in the network give rise to two positive feedback loops. The first feedback is part of the MprA/MprB TCS and originates when MprA-P positively autoregulates the mprAB operon (not shown in Fig. 5a). The second feedback arises through transcriptional regulation of the sigE operon by MprA-P and subsequent regulation of the mprAB operon by $\sigma^E$.

Evidence also exists of possible direct but weak autoregulation of $\sigma^E$ on its own transcription [94].

$\sigma^E$ is a master regulator that controls several downstream genes involved in transcriptional control, translation, electron transport, and oxidative stress response [79]. Note that $\sigma^E$ mediates the transcription of the stringent-response regulator relA, which regulates the expression of antigenic and enzymatic factors required for mycobacterial persistence [101]. Recently, experiments with single M. smegmatis cells revealed that the distributions of expression levels of mprA, sigE and relA genes in a population is bimodal [102,103]. This bimodality suggests that the stress-response network is bistable.

To understand the role of transcriptional and post-translational interactions in generating bistability in the $\sigma^E$ network, a comprehensive mathematical model was built [56]. Subsequently, the various modules of this network were systematically analyzed to identify interactions facilitating bistability. This analysis revealed that bistability in $\sigma^E$ targets can be associated with the post-translational regulation of $\sigma^E$ by RseA coupled with the positive transcriptional feedback of $\sigma^E$ to its own transcription via the MprA/MprB TCS [56]. Notably, the other feedback loop in the system – transcriptional autoregulation of MprA/MprB production by MprA-P – is not sufficient to produce bistability (Section 3.1). This network design – an autoregulated sigma factor subject to sequestration – is similar to the bistability-inducing ultrasensitive mechanism discussed in Section 2.2. Because the feedback to $\sigma^E$ is indirect, bistability is only possible when the loop is active (i.e., when the MprA/MprB TCS is activated). The model also predicts the key role of RseA in bistability; bistability is eliminated if RseA is deleted or overexpressed [56].

Why is RseA so important for ultrasensitivity? In this network, RseA production is constitutive and is not regulated by $\sigma^E$ or MprA-P, whereas the concentration of total $\sigma^E$ present in the system increases with an increase in the concentration of active (free) $\sigma^E$ as a result of indirect feedback via the MprA/MprB TCS. As long as the concentration of total $\sigma^E$ is less than that of total RseA, most of $\sigma^E$ is bound to RseA in an inactive complex and very little free $\sigma^E$ is present in the system. However, when the concentration of total $\sigma^E$ exceeds that of total RseA, the excess $\sigma^E$ cannot be bound by RseA, resulting in a drastic increase in free $\sigma^E$. Hence, near the concentration of total $\sigma^E$ equal to the concentration of total RseA, a small change in the concentration of total $\sigma^E$ can lead to a very large change in the concentration of free $\sigma^E$ [56]. This ultrasensitive mechanism can generate very large effective cooperativity.

Fig. 5. Positive feedback with ultrasensitivity results in bistability in the $\sigma^E$ network. (a) Schematic depicting transcriptional and post-translational interactions in the mycobacterial stress-response network, which consists of the MprA/MprB TCS and the $\sigma^E$-RseA sigma/anti-sigma network. Dashed arrows represent post-translational interactions. Free $\sigma^E$ regulates the transcription of TCS and RelA whereas the transcription of $\sigma^E$ is controlled by TCS. The activity of $\sigma^E$ is post-transitionally regulated by binding of its anti-sigma factor RseA (R), which sequesters $\sigma^E$ into an inactive complex. (b) Steady-state signal-response curve depicts a range of MprB autophosphorylation rates (the external signal which represents the application of stress to the network) for which two different steady states are possible for the concentration of free $\sigma^E$. The two solid curves represent the stable steady states, which are separated by the unstable steady state (dotted curve). Inset: Above a certain threshold level, the network cooperativity (defined as the ratio of change in system output [free $\sigma^E$] to change in system input [total $\sigma^E$]) on a logarithmic scale) increases tremendously with larger interaction strength (defined as the ratio of the total RseA concentration to the dissociation constant for the $\sigma^E$-RseA complex), thus exhibiting an ultrasensitive response. (c) Bistability in the stress-response network results in a time-dependent bimodal distribution of reporter-gene expression.
when the $\sigma^2$.RseA interaction strength is large (Fig. 5b, inset). This large interaction cooperativity, as a result of stoichiometric inactivation of $\sigma^2$ by RseA, makes the complete network bistable for a range of MprB autophosphorylation rates (MprB autophosphorylation rate is a surrogate for the stress signal to the system). This is evident from the bifurcation diagram in Fig. 5b, where for a range of MprB autophosphorylation rates the network has two stable steady states (solid curves) for the system output (concentration of free $\sigma^2$). Bistability manifests itself in the bimodal distribution of a downstream target (such as the relA gene) even when the signal exceeds the bistability threshold because of slow and noisy switching times (Fig. 1b and c). As a result, a stochastic formulation of the model qualitatively reproduces the experimentally observed bimodal distribution of relA gene expression (Fig. 5c) [56]. The heights of the two peaks, which correspond to the basal and maximally expressed levels of relA reporter respectively, change with time, indicating that the number of cells showing induced expression slowly increases over time. Thus, the $\sigma^2$–RseA module is the key network component responsible for bistability in the mycobacterial stress response [56].

5. Bistability in the sporulation phosphorelay

5.1. Design of the sporulation initiation phosphorelay in B. subtilis

B. subtilis cells respond to nutrient starvation by producing spores that do not replicate but can survive extreme conditions [104]. The process of sporulation is energy-intensive and irreversible. A central challenge in understanding the developmental biology of B. subtilis cells is elucidating the decision process that commits cells to sporulation. In B. subtilis, this decision is made by the sporulation phosphorelay shown in Fig. 6a [104–106]. This phosphorelay is the integration point for environmental and metabolic signals that track nutrient availability, cell population density, and DNA damage [107,108]. Deletion mutants of phosphorelay components cannot initiate the sporulation process [109]. At the top of the phosphorelay are five SHKs (KinA–KinE), which act as sensors of nutrient limitation and other environmental signals and activate the phosphorelay [110–112]. The phosphorylated kinases transfer their phosphate group to Spo0F, and this phosphate group is relayed to the master regulator Spo0A via Spo0B. The transfer of this phosphate group down the Spo0 cascade is antagonized by several phosphatases that dephosphorylate Spo0 proteins and modulate the activation dynamics of the phosphorelay. For example, RapA dephosphorylates Spo0F–P and acts to integrate environmental signals about cell density into the phosphorelay. Another phosphatase, Spo0E, dephosphorylates the RR Spo0A–P and is part of a genomically encoded negative feedback that controls phosphorelay dynamics. Spo0A tetramerizes and acts as a TF in its phosphorylated form; it controls the expression of more than 500 genes and is the master regulator of the sporulation response in B. subtilis [113]. These genes have a broad range of affinities for Spo0A–P, which acts as an activator for some genes and a repressor for others [114]. Among the targets of transcriptional regulation by Spo0A–P are several proteins involved in the phosphorelay itself. These include Spo0F, which is regulated by Spo0A–P both via a direct positive interaction and via the sinR operon; Spo0E, which is indirectly activated by Spo0A–P through the repression of AbrB; and RapA, which is activated by Spo0A–P.

These feedback loops (Fig. 6b) play an important role in the dynamics of sporulation initiation via the phosphorelay. Note that the phosphorelay is a combination of two of the prototypical designs for bistability discussed in Sections 2.1 and 2.2, namely the cooperative multimeric positive feedback system and the ultrasensitive post-translational modification system. As Spo0A–P regulates gene expression as a tetramer, its effect is highly significant.
cooperative. Moreover, the phosphorylation and dephosphorylation of Spo0A–P can generate an ultrasensitive response to the phosphorelay signals. Does this imply that the phosphorelay produces a bistable response? We now, review experimental and modeling studies that address this question.

### 5.2. Insights into steady state and dynamical properties of the phosphorelay from mathematical modeling

The intricate design of the phosphorelay makes it difficult to make outright predictions about its response to environmental signals. However, the previously reported qualitative and quantitative experimental results about the sporulation initiation process can guide the development of mathematical models that capture the dynamics of the sporulation commitment process. In addition to information about the structure of the phosphorelay and associated feedback loops, mathematical models of the sporulation initiation process need to be based on the following key qualitative results: (1) commitment to sporulation is associated with increases in Spo0A activity and transcription of the downstream targets of Spo0A [104]; (2) even when conditions are favorable for sporulation, not all cells sporulate. This variability is expected to be associated with the heterogeneity of Spo0A–P activity that exists even in isogenic populations of *B. subtilis* [65, 115, 116]; (3) deletion mutants of Spo0E and RapA have reduced Spo0A–P heterogeneity and increased sporulation efficiency [117, 118]; (4) the phosphotransfer down the phosphorelay is the rate-limiting step in the increase of Spo0A–P [119, 120]; (5) artificial induction of KinA expression above a threshold concentration leads to sporulation even in nutrient-rich conditions [121]; (6) a slow increase in the Spo0A–P concentration is essential for efficient sporulation [122].

Observations 1 and 2, together with the combination of cooperative positive feedback and a potentially ultrasensitive post-translational modification cycle in the phosphorelay, support the hypothesis that the phosphorelay acts as a bistable switch. Mathematical models of the phosphorelay can indeed easily show a bistable signal–response relationship (Fig. 6c). According to this hypothesis, an increase in KinA can push Spo0A–P over a threshold concentration and activate the positive feedback loops to switch the cell into a state of high Spo0A–P activity that resists deactivation. The bimodal population that results from such a bistable switch can explain the simultaneous existence of sporulating and non-sporulating cells in an isogenic population. Several experimental studies demonstrated that the gene expression patterns of Spo0A–P targets, including spo0A and spo011A, are bimodal, suggesting that the positive feedback in the phosphorelay leads to bistability [116, 118].

However, more recent experimental studies show that during sporulation initiation, gene expression of Spo0A–P targets is highly heterogeneous but not bimodal [120, 123, 124]. These studies suggest the hypothesis that the phosphorelay is not the switch that commits cells to sporulation, instead serves to integrate environmental and metabolic signals and acts as a noise generator to produce large variability in the Spo0A–P activity within a population [120, 125]. However, it is difficult to integrate this hypothesis with the evidence that sporulation is ultrasensitive to artificial induction of KinA [121]. One possible way to resolve the controversy is to note that a bimodal distribution of the TF Spo0A–P can result in a long-tailed unimodal distribution of target genes like spoilA if Spo0A–P regulates these genes in highly cooperative fashion (Fig. 6d).

Several modeling studies have used the experimental results discussed above to decipher the design and function of the sporulation phosphorelay. From stability analysis of a simple model of the phosphorelay, Morohashi et al. showed that the system is bistable and that the phosphatase Spo0E can modulate the threshold for the bistable response [126]. These results agree with the experimental findings of Veening et al. [118], Jabbari et al. [127] described a detailed model of the relationship between the phosphorelay inputs – environmental and metabolic signals – and the resulting output decision to activate sporulation pathways. Their model predicts multiple steady-state solutions for Spo0A–P and also predicts that intercellular communication signals from PhrA and nutrient starvation signals from KinA can overpower other phosphorelay inputs, which tend to prevent sporulation. De Jong et al. [128] used a qualitative modeling framework to build a model that can match the available experimental data for various sporulation mutants. These authors also predict that two steady-state solutions are possible for Spo0A–P levels inside the cell and that the system is highly sensitive to noise in gene transcription. Bischofs et al. [125] focused on the design of the feedback structure of the phosphorelay and concluded that the phosphorelay would not be able to integrate information about nutrient availability (from KinA) and cell density (from RapA–PhrA) if a feedback from Spo0A–P acted on Spo0B. Schultz et al. [129] used a modular approach to offer greater insight into the cross-talk of different *B. subtilis* differentiation pathways. For the sporulation pathways, they showed that the mutual-inhibition network of Spo0A–P, AbrB, and Spo0E can lead to repressilator-like oscillations in their concentrations, which might explain the phenotypic heterogeneity associated with the sporulation response. Chastanet et al. [120] described a model for the phosphorelay to match their experimental observation that Spo0A–P activity is not bimodal. However, to allow sufficient accumulation of Spo0A–P, they were forced to either increase the concentrations of phosphorelay components or introduce a feedback loop from Spo0A–P to Spo0B. Both assumptions seem untenable given that Spo0B concentrations stay constant during sporulation initiation [123]. Moreover, Chastanet et al. only provided results of time-course simulations for sporulation initiation and no signal dose–response curves, making it difficult to judge whether the system response is bistable or not. Nevertheless, several design properties of the phosphorelay can be deduced from these modeling results. First, the phosphorelay steady-state response can be bistable. Second, the nutrient starvation and cell density signals are the most important determinants of the sporulation decision. Third, the competition between the phosphorylation signals through KinA and the phosphatases RapA and Spo0E both sensitizes the phosphorelay output to environmental inputs and generates a great deal of variability, which is responsible for the phenotypic heterogeneity associated with sporulation.

Notably, most of these studies focused on the magnitude of the change in the steady-state activity of Spo0A–P in response to nutrient starvation rather than on the dynamics of this change [120, 125, 127, 130]. But according to observation 6, the dynamics of Spo0A–P accumulation may also affect the success of sporulation. Under nutrient-limited conditions, wild-type cells exit the exponential growth phase and enter a transition stage where the accumulation of low levels of Spo0A–P is enough to suppress high-affinity genes like *abrB* but not enough to affect the transcription of low-affinity genes like *spolIA* [104]. Eventually, feedback from Spo0A to the phosphorelay leads to an increase in phosphate flux, pushing the cell out of the transition stage and activating late-stage sporulation genes. The transition stage provides the cells with a time window in which the cells can collect enough energy for sporulation through the degradation of complex carbohydrates and neighboring cells (Fig. 6e). Sporulation efficiency was significantly reduced in mutants in which the wild-type Spo0A had been replaced with a constitutively active form, Spo0A* [122]. Spo0A* functions as a TF, and artificial induction of Spo0A* led to a rapid increase in the expression of downstream target genes, practically without any transition phase (Fig. 6f). Without the extended time window of the transition phase, sporulation efficiency is low.
despite high levels of *spoll* sporulation genes. Therefore, the often-neglected dynamics of the phosphorelay switch clearly play a very important role in determining sporulation efficiency.

6. Discussion

Design principles of biochemical networks may represent two overall response strategies, assuming sufficient conditions for fine-tuned evolutionary optimization of the network responses. Classical designs tend to favor homeostatic responses: monotonicity and fast response times. However, as we have outlined in this review, a major class of networks conform to a second design paradigm with different criteria for functional effectiveness. Networks evolving to meet this paradigm favor heterogeneity (particularly bistability) and slower transient responses.

6.1. Mechanisms of bistable responses

The examples presented in this review use multiple mechanisms to produce bistability in a network. But do any general necessary and sufficient conditions exist for this behavior? The presence of a positive feedback is a necessary, but not a sufficient, condition for a biochemical network to have multiple steady states and, thus, a bistable response [131]. What plausible biochemical interactions lead to a positive feedback? One possibility is an autogenous positive or double-negative transcriptional regulation. Such loops are usually apparent from the network diagram. Several theoretical networks of this type were proposed by Monod and Jacob [6], and since then the analogues of naturally occurring and artificially constructed networks have been shown to be bistable [21,132–134].

However, in many cases the positive feedback necessary for bistability is only apparent after analysis of the network’s Jacobian matrix [131,135]. The examples summarized in this review emphasize that the existence of a positive feedback may not be apparent from the kinetic scheme. Indeed, a wide class of systems described by mass-action equations can display multistability even though the analysis of their network diagram does not reveal an explicit positive feedback. In these systems, an implicit positive feedback exists as a result of complex interactions or conservation laws among network components. Craciun et al. formulated a necessary condition for such systems to attain multistability [34]. In this review, we emphasize one such mechanism for achieving positive feedback: the self-enhancing formation of a dead-end complex. As shown in Sections 2.3, 3.3 and 4.1, an increase in the concentration of the dead-end complex decreases the concentration of the catalytically active complex, leading to substrate accumulation and, thus, an even greater increase in dead-end complex formation. This positive feedback is not obvious from the network structure but may still lead to a bistable response, as in the cases of *σF* network and post-translational interactions in TCSs.

Positive feedback is necessary for multistability, but can one deduce general conditions that guarantee a bistable response? Angeli et al. [136] developed methods to deduce bistability for networks of arbitrary complexity that do not contain negative feedback loops (monotonic systems). They showed that a sigmoidal response of an open-loop system (i.e., a system with a broken positive feedback loop) is sufficient to ensure bistability for a range of feedback strengths. For the explicit positive feedback loops acting at the level of transcriptional control (such as the examples in Sections 2.1 and 2.2), the loop can often be broken by replacing the native promoter with an exogenously inducible one. In this case, the open-loop response can be measured experimentally and the method of Angeli et al. [136] can be readily applied. However, designing an experimental way of creating an open-loop system for an implicit feedback loop is not trivial (Section 2.3, and as predicted for the *σF* network and the TCS network). The non-linear amplification or sigmoidality required for bistability may arise from a variety of network interactions, including cooperative or allosteric effects, a non-linear saturating expression rate, and differences in concentrations of interaction partners (as in the *σF–RseA* network) [56].

6.2. Physiological interpretations and design principles

The above examples show that bistable responses can be changed into graded responses, and vice versa, by adjusting either network structure or kinetic parameters. In the course of evolution, networks in large populations tend to adapt to physiological demands by adopting context-appropriate responses. Can we define design principles that predict this selection? If it is costly or even deleterious for a cell to fully induce expression of a certain gene unless environmental conditions demand it, then a graded response is advantageous; it is important to continuously adjust the response to achieve intermediate levels of response for intermediate stimuli. In contrast, for binary outcomes such as cell differentiation or division, no intermediate fates exist and a bistable switch is beneficial. Another difference is seen in robustness to variations in the stimulus or other parameter values: bistable switches are usually more robust. Indeed, once a bistable switch is fully ‘on’ or ‘off’ (i.e., away from the bistable range), substantial changes in the stimulus or other parameters are needed to switch it back. This property again shows usefulness of bistable responses in making cell-fate decisions.

By examining the transient responses of bistable switches, we conclude that bistable response is often associated with a slowdown in response (see also [137]). Thus, fast responses in graded switches allow cells to quickly adapt to dynamically changing environments, but slower responses of bistable switches can be advantageous in filtering out such transient signals. Analysis of the partner-switching signaling networks controlling *B. subtilis* sigma factor *σE* further illustrates these points. The predicted bistable response for the *σF* network is consistent with its role in commitment to cell differentiation [138]. Our mathematical model predicts that *σE* activity can be induced even by transient signals and that once fully induced, the system will stay this way for a wide range of parameter values. The implicit biochemical feedback loop (involving self-enhancing formation of a dead-end complex rather than transcriptional control) allows a fast ‘on’ switch on a biochemical time scale, which is faster than gene induction. However, the slow dissociation of a dead-end complex ensures a slow ‘off’ switch, so that transient fluctuations of kinetic parameters will not result in *σE* deactivation.

6.3. Stochastic effects in bistable switches

The differential equations modeling the biochemical networks reviewed here represent a deterministic approximation, which is only capable of describing quantities of the system averaged over large populations of cells. However, many important dynamic consequences arise at the scale of single cells, where fluctuations in the numbers of molecules present at low concentrations, as well as the random nature of the binding kinetics at the promoter, lead to stochastic effects. Indeed, stochastic effects often result in intracellular protein concentrations that are different from those predicted by deterministic simulations for small system sizes or slow promoter kinetics [139,140]. Stochastic simulations predict such phenotypic variability within an isogenic population [139,141].

Deterministic simulations of a bistable switch predict that only one of the two stable states will be achieved, depending on initial conditions, and that transitions between the two steady states are not possible. Examples of bistability in TCSs suggest that noise in bistable slow-response regions may induce heterogeneous
Evolutionary selection toward specific dynamic properties in biochemical networks tends to convergent evolution of networks, so that similar network architectures are repeatedly found in evolutionarily unrelated networks. Foundational studies on design principles by Savageau et al. determined a set of functional criteria based on the idea that biochemical networks constantly evolve to simultaneously optimize response times, robustness, monostability, and other criteria that favor homeostatic circuits [137,145–147]. In contrast, circuits displaying a bistable response are subject to different evolutionary selective pressures and therefore have several unique functional properties. In deterministic systems, these include a binary ‘on’/‘off’-type response that is robust to signal fluctuations and a slow switching rate that filters out transient signals. Moreover, stochastic transitions between two stable steady states give rise to phenotypic heterogeneity, which may manifest as bimodality in the population distribution. These functional attributes are usually favored in networks that control cell-fate determination because they allow the system to make reliable decisions and to use the phenotypic heterogeneity as a bet-hedging strategy.

In this review, we have presented some conceptual designs of bistable switches. However, we emphasize that the unique features of most biochemical networks warrant detailed experimental and modeling studies before the presence of bistability can be confirmed and these design principles can be applied.

Acknowledgments

This work was supported by awards MCB-0920463 from the National Science Foundation and R01-GM096189-01 from the National Institutes of Health.

References


7. Concluding remarks

Evolutionary selection toward specific dynamic properties in biochemical networks tends to convergent evolution of networks, so that similar network architectures are repeatedly found in...
88

A. Tiwari et al./Mathematical Biosciences 231 (2011) 76–89


