Chapter 19
Performance Characteristics for Sensors and Circuits Used to Program *E. coli*

Jeffrey J. Tabor, Eli S. Groban, and Christopher A. Voigt

Contents

Abstract The behavior of *E. coli* can be reprogrammed by the introduction of foreign segments of DNA. Three classes of genetic parts, termed sensors, circuits and actuators comprise the DNA programs. Sensors are gene products which allow the cell to detect physical or chemical information in its environment. Genetic engineers can use sensors directly from nature, modify them in some manner, or design them *de novo* to control cellular processes with extracellular or intracellular signals. Genetic circuits act to process information from sensors in order to dictate the behavior of the cell. They can be designed with combinations of "off the shelf" regulatory parts such as transcription factors and promoters, or in some cases can be used "as is" from nature. Finally, genetic circuits govern the expression of actuators, genes whose products perform some physical function to alter the state or the environment within which the cell exists. Using recent DNA synthesis and assembly technologies, genetic sensors, circuits and actuators can be combined to create programs that

C.A. Voigt (✉)
Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143-2280 USA; Biophysics Program, San Francisco, CA 94158 USA
e-mail: cavoigt@picasso.ucsf.edu

S.Y. Lee (ed.), *Systems Biology and Biotechnology of Escherichia coli*,
DOI 10.1007/978-1-4020-9394-4_19, © Springer Science+Business Media B.V. 2009

command cells to perform a series of tasks. This approach will transform the way that genetic engineers approach problems in biotechnology. This review covers the construction of genetic sensors and circuits for use in *E. coli*, as well as genetic methods to perturb their performance features.

19.1 Introduction

To program novel behaviors into *E. coli*, handfuls of genetic parts, or segments of DNA with defined functions, are introduced into the cell. In the background, thousands of regulatory and metabolic reactions operate simultaneously and in direct physical contact with the heterologous parts. The engineered components can operate as insulated modules or can be functionally integrated with the preexisting networks of the host cell. Despite what would appear to be long odds, surprisingly complex behaviors with medical, industrial or academic relevance can be achieved.

In this chapter, we will discuss some of the principles which guide the programming of *E. coli*. We define biological programs as strings of genetic parts encoded on segments of DNA which are introduced to the cell on plasmid vectors or integrated into the genome. The designed DNA fragments carry three classes of parts which we will refer to as *sensors*, *circuits* and *actuators* (Voigt 2006). Each of these functions is encoded on a piece of DNA. When combined they create a genetic program that provides a set of instructions that the cell can read and execute. Though the sensor/circuit/actuator construction paradigm can be applied to program any number of genetically tractable organisms (Drubin et al. 2007, Greber and Fussenegger 2007, Sia et al. 2007), this chapter will be limited to a discussion of *E. coli* where much of the foundational work has been accomplished.

Sensors transmit information to genetic circuits. *Genetic circuits* are groups of regulatory molecules which control gene expression to program the cellular response to sensory inputs. Genetic circuits are ubiquitous in the genomes of natural organisms and the characterization of their input-output ranges and dynamic and steady-state responses, or *performance features*, can inform the construction of synthetic analogs with defined properties. In some cases the entire DNA segment encoding a natural circuit can be used "out of the box", or as found in nature, simply being connected to user defined sensors and actuators. Synthetic genetic circuits are built by designing a piece of DNA which carries a series of regulatory parts which interact in a defined manner.

Genetic circuits drive *actuators* which act to change the state or behavior of the host cell or its environment. Actuators range from simple reporters like Green Fluorescent Protein (GFP) to entire organelles. The programming of reliable and sophisticated behaviors in *E. coli* will require actuator expression and function to be tightly governed by environmental, physiological or metabolic signals which are transmitted through genetic circuits via sensors.

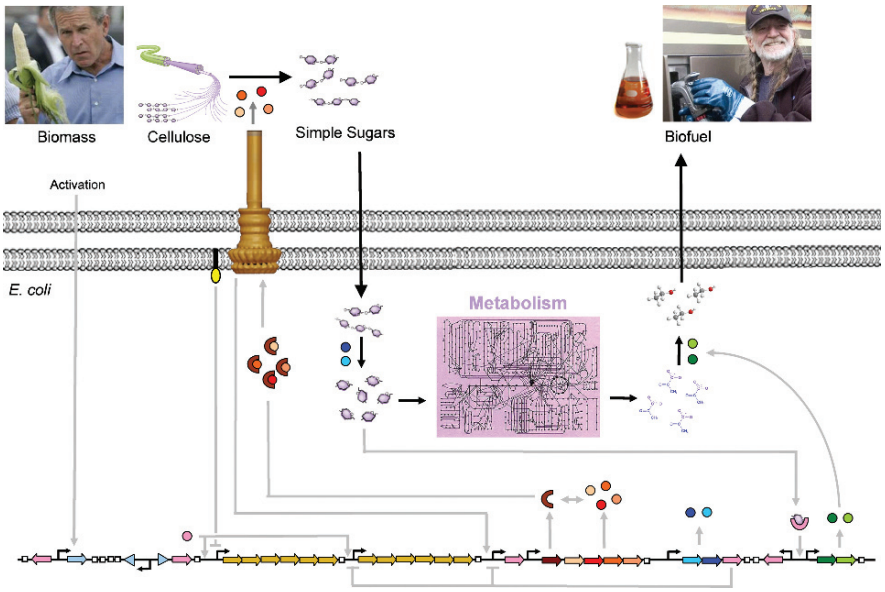

Fig. 19.1 A hypothetical *E. coli* program to convert biomass to liquid fuel. Complex plant material requires that multiple enzymes be exported in a timed sequence. The enzymes need to be exported from the cell, in this case using a type III secretion system imported from Salmonella. The build up of simple sugars induces a pathway to break them down into glucose and covert a metabolic product into a fuel. This is an example of integrated bioprocessing, where multiple steps of a manufacturing process are programmed into a single organism. This requires the combination of sensors, circuits, and actuators to control and respond to a sequence of events

Programs written from sensors circuits and actuators can coordinate sophisticated multistep behaviors with applications in biotechnology (Fig. 19.1). This type of integrated bioprocessing includes, for example, sensing, integrating and responding to media conditions or cell growth stages or densities within a fermenter for optimized yields of an industrially relevant natural product.

Historically limited to piecemeal stitching of naturally occurring DNA fragments, modern DNA synthesis and assembly methods allow the arbitrary connection of sensors, circuits and actuators. Very large (genome scale) biological programs can now be written *in silico* and constructed commercially (Endy 2008, Gibson et al. 2008). The reprogramming of genomes will enable streamlining of the cell through the wholesale addition, deletion or modification of regulatory and metabolic pathways. This will in turn increase the stability, efficiency and productivity (Posfai et al. 2006) of engineered cellular processes.

19.2 Sensors

Genetic sensors typically receive information from the extracellular environment or internal cell state, which is then transmitted to gene regulatory networks. Environmental sensing in *E. coli* largely comprises three strategies: classical regulation,

two-component sensing and riboregulation. We will discuss some of the best studied and most widely engineered examples of these sensors throughout this section.

Sensors can receive myriad physical and chemical inputs including small or macromolecules, pH, temperature, light and even signals from other cells. This chapter will focus only on small molecule signals which are the most widely used inputs for engineering *E. coli*.

19.2.1 Classical Regulation

Classical regulation is the control of promoter activity by ligand binding proteins (Fig. 19.2A). The sensor is a cytoplasmic transcription factor which receives an environmental signal by directly binding to a small molecule ligand. Ligand binding triggers a conformational rearrangement which results in increased or decreased affinity of the transcription factor for cognate DNA operator sequences. The sensory output can be transmitted in two ways, by activation or the relief of repression. Activation typically occurs by transcription factor-mediated recruitment of the RNA polymerase complex at the promoter while repression occurs by its occlusion (Wagner 2000).

Classical transcription factors are the most widely used sensors for programming *E. coli*. This is due to the simplicity of their components, their rapid output (strong transcriptional responses occur on the order of 1 minute (Guzman et al. 1995)), the ease with which their input and output specificities can be re-engineered, and the availability of their inducer compounds (Wagner 2000). Here, common strategies are outlined for re-engineering the specifities and performance features of classically regulated transcription factors. Throughout this section we will focus on a particularly well elaborated example, the tetracycline responsive TetR protein.

19.2.1.1 Re-engineering Classically Regulated Sensors

The steady-state quantitative relationships between the concentration of input signal and output gene expression, or transfer functions (Canton et al. 2008, Weiss et al. 1999, Yokobayashi et al. 2002), have been characterized for many classically regulated systems. The features of transfer functions arise from the rate of occupation of promoters by transcription factors and RNA polymerases at different input concentrations (Bintu et al. 2005b). The transfer function of a circuit can be measured by linking it to a sensor, varying the amount of input and measuring the output with a reporter gene (Fig. 19.3). Transfer functions are useful in the design of cellular behaviors because they define the minimal and maximal amount of sensory input which generate circuit responses, the magnitude of induction at any given input concentration and the sensitivity of the circuit to input (Bintu et al. 2005a).

The dynamic range of induction, or magnitude of output in the fully activated (ON) state divided by that of the inactive (OFF) state, is a critical feature of any sensor. In many cases, a large dynamic range of induction is desirable because it more clearly differentiates the absence and presence of an environmental input. Increased

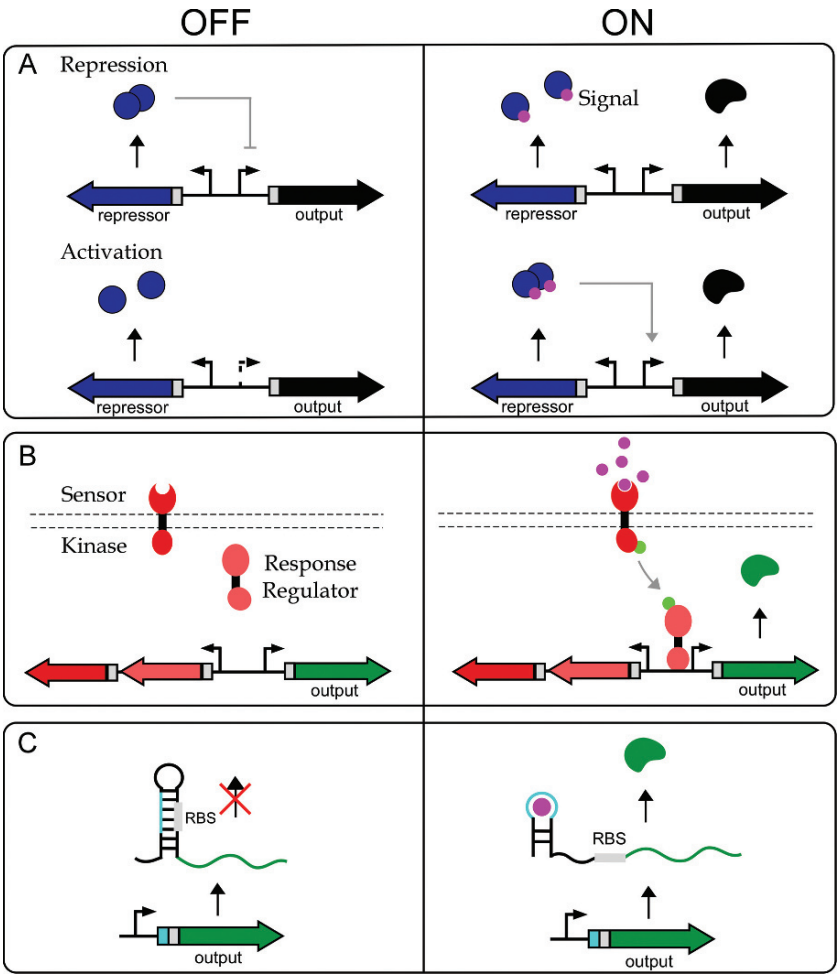

Fig. 19.2 (**A**) Classical transcriptional regulation. In classical systems a cytoplasmic transcription factor protein regulates the target genes in response to the presence of input ligand (*grey dots*). Classical regulation can occur in two forms, repression or activation. With repression, the transcription factor binds to the promoter in the absence of ligand (*left*) and undergoes a conformational change upon ligand binding which causes it to dissociate from the DNA, activating transcription (*right*). With activation, the transcription factor does not associate with the promoter in the absence of ligand (*dashed*), but does in its presence, increasing the rate of transcription. (**B**) Two-component sensing. A membrane associated sensor-kinase protein associates with an extracellular ligand at its sensor domain, which drives a structural change in its cytoplasmic kinase domain. This triggers autophosphorylation of the cytoplasmic domain. The phosphate group (*light grey dot*) is then transferred to the receiver domain of a cytoplasmic, diffusible response regulator protein. When phosphorylated, the response regulator changes conformation and binds to its cognate operator sites near promoters, activating or repressing gene expression. (**C**) An engineered riboswitch. A constitutive promoter drives the expression of a gene with an engineered RNA hairpin occluding its ribosome binding site (RBS, *grey*) and blocking translation. The hairpin also carries an aptamer sequence upstream of the RBS, which can bind to a cognate ligand (*large dot*), triggering a structural rearrangement which liberates the RBS for productive translation. Adapted from Topp and Gallivan, 2007

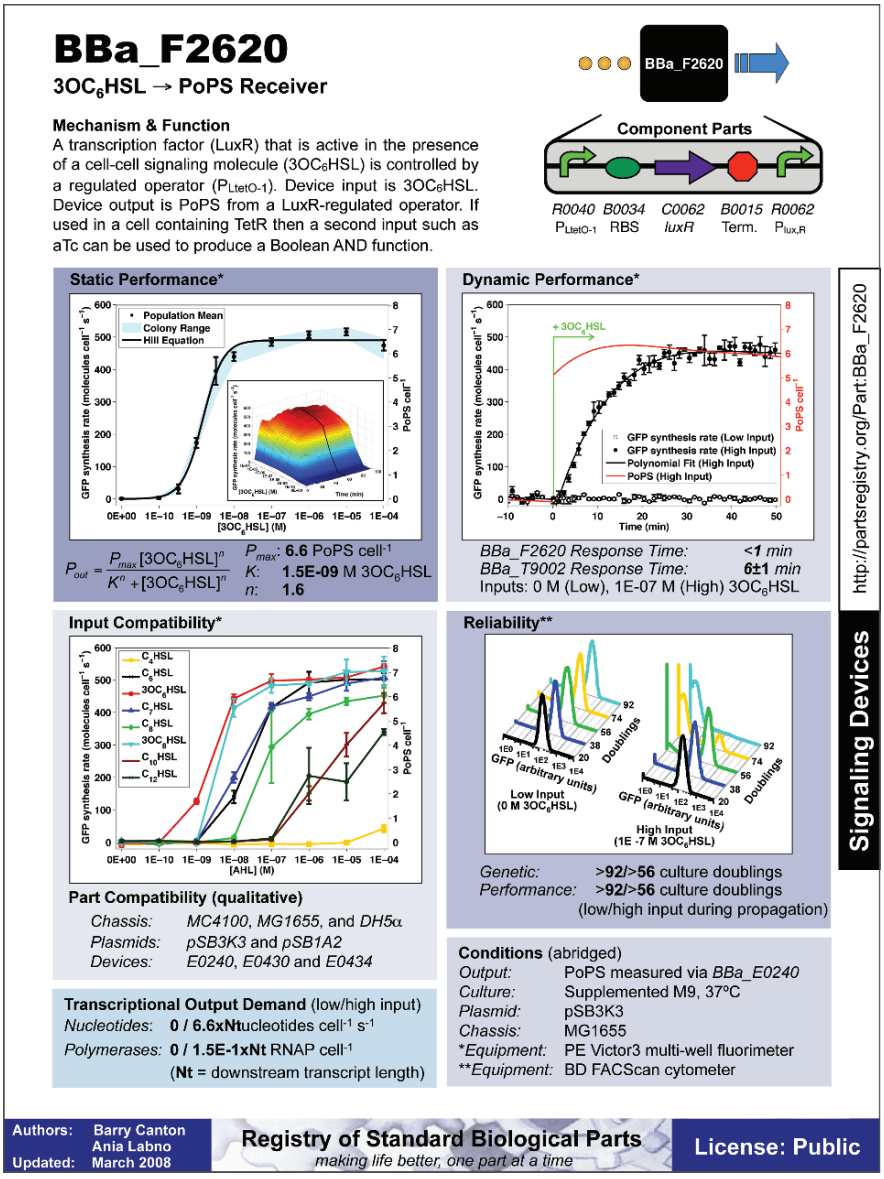

Fig. 19.3 Performance Feature Specification Sheet (from Canton et al. 2008)

dynamic ranges can be achieved by increasing the transcription rate of the ON state, decreasing the transcription rate of the OFF state, or both. The ON state can most easily be increased by strengthening the −35 and −10 RNA polymerase recognition sequences while the repressed state can be lowered by changing the configuration of operator sites around the promoter (Cox et al. 2007, Lutz and Bujard 1997,

Lutz et al. 2001). The *sensitivity*, or rate of increase in transcriptional output as a function of ligand concentration (Fig. 19.5) is largely proportional to the cooperativity of binding of the transcription factor at the promoter. We will discuss strategies for programming cooperativity in Section 19.3.

19.2.1.2 Increasing Dynamic Range

The dynamic range of classical transcription factor systems can be increased by changing the architecture of the output promoter. Traditionally this been accomplished by the addition, deletion or reorganization of the operator sites (Brosius et al. 1985, de Boer et al. 1983, Guzman et al. 1995, Lutz and Bujard 1997). In this section we will discuss efforts specific to the TetR protein.

TetR has been used as the basis for engineering a more tightly repressed and strongly inducible sensing system. To accomplish this two high affinity operator sites were added to an otherwise strong promoter. TetR was then constitutively expressed to repress the promoter in the absence of the input ligand. The system showed virtually no expression in the OFF state, was sensitive to very low levels of input and showed a ∼5000-fold dynamic range of induction (Lutz and Bujard 1997). The performance features of the re-engineered system were all marked improvements over the naturally occurring version from which it was derived (de la Torre et al. 1984, Kleckner et al. 1978) and as a result TetR has become one of the most widely used sensors for programming *E. coli*.

19.2.1.3 Changing Operator Specificity

Novel transcription factor:promoter pairs can also be derived from natural systems. The introduction of two mutations within the TetR operator sequence can reduce the affinity to levels insufficient for *in vivo* repression. Rational redesign of DNA binding domains or directed evolution can then be used to re-establish the affinity of the transcription factor for the mutant operators. Indeed, such methods have generated novel transcription factor:promoter pairs based on the TetR (Helbl and Hillen 1998, Helbl et al. 1998) LacR and lambda Cro (Backes et al. 1997) systems as well. Importantly, these novel specificities can be generated with very small numbers of amino acid substitutions in the transcription factors, allowing the rapid generation and screening of many new orthogonal sets in the cellular context. Similar strategies are likely to be amenable to virtually any classically regulated promoter system in *E. coli*.

19.2.1.4 Changing the Input Ligand

The input specificities of classically regulated systems can also be reprogrammed. This is typically accomplished by randomly mutating amino acid residues around the ligand binding pocket and screening variants in functional assays *in vivo* (Collins et al. 2005, 2006, Hawkins et al. 2007). We will discuss efforts to reprogram the ligand specificity of the TetR protein in this section.

TetR has been evolved to recognize an alternate ligand with strong preference over the natural ligand (Henssler et al. 2004). Importantly, the novel inducer is not recognized by the wild type TetR protein, a feature which gives rise to two orthogonal sensors. The combination of novel input and output specificities has the potential to generate completely orthogonal sensing systems which can be used in parallel with one another. Indeed, two TetR variants which sensed different ligands and activated different promoters were recently introduced into the same *E. coli* cell to control the expression of two separate genes (Kamionka et al. 2004). This work demonstrates the value of classically regulated sensing systems as a platform for the construction of genetic control elements with broad applications in biological design.

19.2.2 Two-Component Sensing

A common strategy for environmental sensing in bacteria is a process known as two-component sensing. The canonical two-component system consists of a membrane-bound sensor protein that receives an environmental signal at an extracellular sensory domain and passes the information to a cytoplasmic response regulator protein (Fig. 19.2B). This occurs via the transfer of a phosphate moiety from the cytoplasmic kinase domain of the sensor protein to the receiver domain of the response regulator protein, which can then bind to DNA operator sites at a DNA binding domain to activate or repress gene expression (Hoch and Silhavy 1995).

These sensors are slower to respond than their classically regulated counterparts. For example, the well studied EnvZ/OmpR system of *E. coli* reaches half maximal response to the presence of an input signal in about 5 minutes but requires much longer (on the order of 1 hour) to reach steady-state (Batchelor and Goulian 2006). This happens despite the fact that the phosphotransfer event occurs on a seconds time scale at most (Laub et al. 2007).

The re-engineering of two-component systems has been aided by the modularity of the protein structure. Modular systems are those that are composed of multiple interchangeable subcomponents, or modules. In two-component systems, the extracellular sensory domain of the sensor kinase protein can be replaced by the sensor module from a similar protein. Likewise, the kinase domain of a given sensor protein can be swapped with another to change its specificity for a response regulator (Fig. 19.4). Similar to the classically regulated systems, the specificity of the sensor kinase for its input signal can be altered by computational design methods.

19.2.2.1 Domain Swapping

Sensors can easily be rewired to new outputs by domain swapping. This involves fusing non-cognate sensor and kinase domains at a splice site in a linker region. Most two-component engineering efforts to date have been based on domain swapping, a design process by which chimeric proteins are built from the subdomains of two or more pre-existing proteins (Fig. 19.4). This type of engineering allows

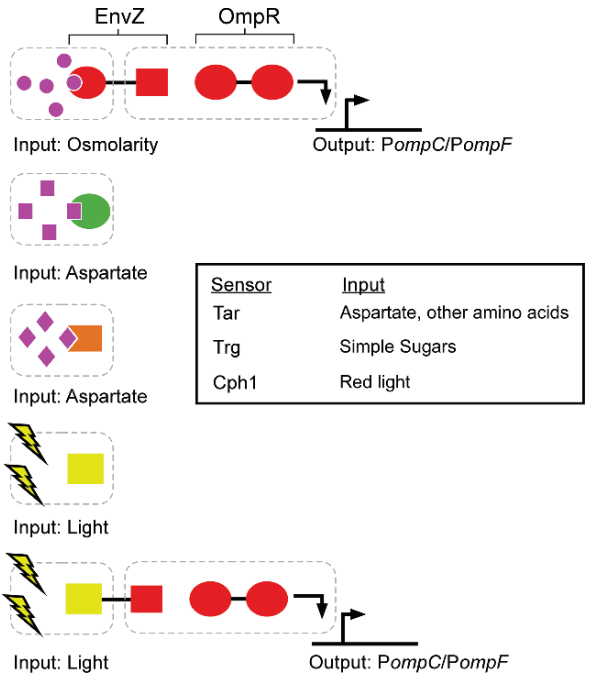

Fig. 19.4 Domain Swapping. The sensory domain of EnvZ receives inputs and transfers information to the response regulator OmpR through the kinase domain. OmpR then activates the expression of the output gene from a target promoter. Other sensory domains can replace the naturally occurring EnvZ sensory domain to create chimeric sensor proteins which activate output promoters in response to different inputs

the sensing pathway to be rewired such that, for example, the output promoter will respond to a completely different input ligand.

The early discovery of a convenient module boundary (Utsumi et al. 1989) made the osmo-responsive EnvZ/OmpR two-component system of *E. coli* a favorite target for many engineering efforts (Baumgartner et al. 1994, Levskaya et al. 2005, Looger et al. 2003). In the natural configuration, the sensor kinase EnvZ phosphorylates the response regulator OmpR in response to changes in osmolarity. Phosphorylated OmpR then binds to operator sites at a promoter, activating or repressing gene expression (Aiba et al. 1989, Aiba and Mizuno 1990, Forst et al. 1989). In the pioneering domain swapping effort, Inoyue and co-workers fused the cytoplasmic domain of EnvZ with the sensory domain of the transmembrane aspartate receptor (TAR), thus rewiring the EnvZ/OmpR pathway to be activated by the amino acid aspartate (Utsumi et al. 1989, Fig. 19.4).

The sensory domain of the chemoreceptor protein Trg has similarly been fused to the cytoplasmic domain of EnvZ (Baumgartner et al. 1994). The Trg sensory domain interacts with periplasmic sugar binding proteins only when they are bound to their ligands to direct *E. coli* chemotaxis. The hybrid Trg-EnvZ protein allowed control of the EnvZ/OmpR pathway with the unnatural ligand ribose via the ribose binding protein (RBP) (Baumgartner et al. 1994).

The sensory domain of a other sensor kinases have also been used to control a chemotactic signaling. NarX is a histidine kinase which senses nitrate and nitrite (Williams and Stewart 1997). Replacement of the sensory domain of the TAR

protein with the sensory domain of the NarX kinase has programmed *E. coli* to chemotax away from extracellular nitrate and nitrite (Ward et al. 2002).

In 2005, the osmosensing domain of EnvZ was replaced with a light sensing domain from the *Synechocystis* phytochrome protein Cph1 to program *E. coli* to respond to light (Levskaya et al. 2005). This also required the introduction of a two gene metabolic pathway to produce the chromophore PCB, which binds to the engineered sensor kinase (Gambetta and Lagarias 2001). A confluent lawn of the engineered *E. coli* could then be used as a high resolution film capable of directly converting a two-dimensional light input pattern into a pigment output pattern.

19.2.2.2 Redesigning Ligand Specificities

Other efforts have used computational methods to redesign of periplasmic sugar binding protiens to sense ligands as varied as trinitrotoluene (TNT), L-Lactate, (Looger et al. 2003) and Zn^{2+} (Dwyer et al. 2003) for control of gene expression through the Trg-EnvZ/OmpR pathway. Unlike domain swapping strategies, these studies required detailed knowledge of the three-dimensional structure of the parental proteins. The structural information guided the authors to consider between 5 and 17 amino acids residues as candidates for mutation, and the computational searches typically yielded small lists of candidate protein sequences which were directly amenable to experimental evaluation.

19.2.2.3 Designing the Histidine Kinase-Response Regulator Interface

There are at least 32 natural two-component systems in *E. coli*, all of which have similar structures at the sensor/response regulator interface (Ulrich et al. 2005). To maintain the fidelity of signal transmission through any one of these pathways the sensors and response regulators have evolved a great deal of pairwise molecular specificity (Skerker et al. 2005). Knowledge of the specificity determinants of the histidine kinase-response regulator interactions could allow rewiring of input-output relationships.

Bioinformatic algorithms have recently been used to elucidate regions of the histidine kinase proteins responsible for response regulator specificity. This information enabled the rewiring of two-component pathways by mutating sensor/response regulator interaction domains. The substitution of as few as three amino acid resides within a cytoplasmic subdomain of EnvZ reprogrammed its specificity away from OmpR to numerous other response regulators (Skerker et al. 2008). The ability to redesign protein/protein interfaces adds a valuable degree of freedom which will greatly increase the number of possible alternative two-component signaling pathways that can be constructed in *E. coli*.

19.2.3 Riboregulators

RNA molecules can sense inputs, often through interactions with small or macromolecular ligands, and transmit the information to control gene expression. This typically occurs via the formation of a ligand binding pocket within the regulatory

RNA (riboregulator) which triggers an overall change in its secondary structure. These structural rearrangements can hide or liberate regulatory domains which can then modulate gene expression *in cis* or *in trans* (on the same or another gene). To date, 16 *E. coli* genes have been shown to be subject to cis-acting regulation by ligand binding riboregulators termed riboswitches (Barrick and Breaker 2007).

Bacterial riboswitch sensors convert ligand binding into a change in the transcription or translation rate of the mRNA within which they are embedded (Winkler and Breaker 2003). Though not as widely utilized as their protein counterparts, the structural and functional simplicity of RNA makes it a very attractive platform for the engineering of sensing in bacteria (Isaacs et al. 2006). This is because secondary structure, which governs much of the overall shape and function of RNA, can be computationally predicted with good accuracy (Mathews et al. 1999) and experimentally verified much more rapidly than can three-dimensional protein structures (Soukup and Breaker 1999c). This allows realistic *in silico* design of riboregulators *de novo*, a monumentally difficult task in the protein world.

19.2.3.1 Reprogramming

Riboregulation is also compelling because simple base pairing rules and robust directed evolution methods allow the construction of many orthogonal regulators based on a single parent structure (Bayer and Smolke 2005, Isaacs et al. 2004, Jose et al. 2001, Koizumi et al. 1999, Soukup and Breaker 1999a, Soukup and Breaker 1999b, Soukup et al. 2001, Tang and Breaker 1997). The modular structure of riboregulators also allows them to be introduced into many different genes and even ported between vastly different organisms with surprising ease (Yen et al. 2004). Moreover, unlike in two-component engineering, the sensory domains of riboregulators need not bear any structural or evolutionary relationship to the regulatory domains to which they are fused (Bayer and Smolke 2005, Buskirk et al. 2004, Jose et al. 2001, Soukup and Breaker 1999b).

As a concise demonstration of the design advantages of riboregulators, a riboswitch was recently designed *de novo* to reprogram *E. coli* chemotaxis (Topp and Gallivan 2007). In this work an antisense RNA domain was engineered to base pair with and occlude a ribosome binding site (RBS) upstream of the open reading frame of a chemotaxis-dependent gene, inhibiting translation and subsequently chemotaxis (Fig. 19.2C). A ligand binding (aptamer) domain for the small molecule theophylline was included within the riboregulator such that when theophylline was present, a local base pairing rearrangement occurred which liberated the ribosome binding site, allowing translation. In this way, the engineered riboswitch guided *E. coli* to swim up a gradient of a chemical that does not normally function as an attractant. Though domain swapping and directed evolution have enabled the rewiring of chemotaxis at the protein level as well (Derr et al. 2006, Ward et al. 2002), the benefits of riboregulation are manifest in this example as high throughput efforts have allowed rapid increases in the dynamic range of induction in response to ligand (Lynch et al. 2007, Topp and Gallivan 2008).

19.2.4 Cell-Cell Communication

Cells also have the ability to sense the presence of other cells in the environment. In bacteria this often occurs through a process known as quorum sensing (Miller and Bassler 2001). In short, cells produce membrane-diffusible signals which diffuse into other cells and function as ligands for classical transcription factors. This type of sensing can drive coordinated decision making in cell communities, which enables more sophisticated behaviors.

Cell-cell communication sensors have been used in *E. coli* to control the density of a bacterial population (You et al. 2004), coordinate the timing and magnitude of gene expression between two different cell types (Brenner et al. 2007), drive multicellular pattern formation (Basu et al. 2004, 2005), coordinate the invasion of a malignant mammalian cell (Anderson et al. 2006) or even create a synthetic ecosystem (Balagadde et al. 2008). Each of these circuits was constructed from the Lux-type quorum sensing circuit of *V. fischeri*. A full review of the engineering applications of this type of cell-cell communication system is reviewed elsewhere (Salis et al. 2009).

19.3 Circuits: Processing Sensory Information

Genetic circuits, or networks of interacting regulatory molecules, can integrate one or more sensory inputs into logical and dynamic genetic outputs (Hasty et al. 2002, Kaern et al. 2003, Wall et al. 2004). Circuits have previously been constructed in *E. coli* which generate memory (Atkinson et al. 2003, Gardner et al. 2000), oscillations (Atkinson et al. 2003, Elowitz and Leibler 2000) or pulses (Basu et al. 2004) of gene expression. Other circuits have been designed to function as logic gates, capable of integrating information from multiple sensors to produce a single output (Anderson et al. 2007, Guet et al. 2002, Yokobayashi et al. 2002). Genetic circuits can also coordinate cell-cell communication and community-level decision making (Balagadde et al. 2008, Basu et al. 2005, Brenner et al. 2007, You et al. 2004) This section provides an overview of the performance features and engineering considerations for some of the best characterized and most useful genetic circuit motifs.

19.3.1 Classical Regulation

The simplest genetic circuits are the classical ligand-inducible transcription systems described in Section 19.2.1. In these simple circuits, the presence of input signal positively influences the transcription of an output gene. The transfer function of classically regulated circuits is important because it describes the level of gene expression out of the circuit in response to a given concentration of input signal. This is important when linking multiple circuits in series, because if the output of one circuit is not quantitatively matched with the input of another, then information transfer through the system breaks down. It is of particular interest to discuss the performance features of classically regulated circuits here as they constitute the foundation of many more complex circuit designs.

19.3.1.1 Simple Promoters

In classically regulated circuits the output abundance typically varies as a positive sigmoidal function of the input concentration (Bintu et al. 2005a) (Fig. 19.5A). This relationship arises because there are two input ranges where the system is non-responsive and one input range under which it is. At low input levels, well below the K_D of the transcription factor for the ligand, there is virtually no change in output. As the input ligand concentration approaches the K_D of the transcription factor, there is a monotonic increase in output protein abundance proportional to input.

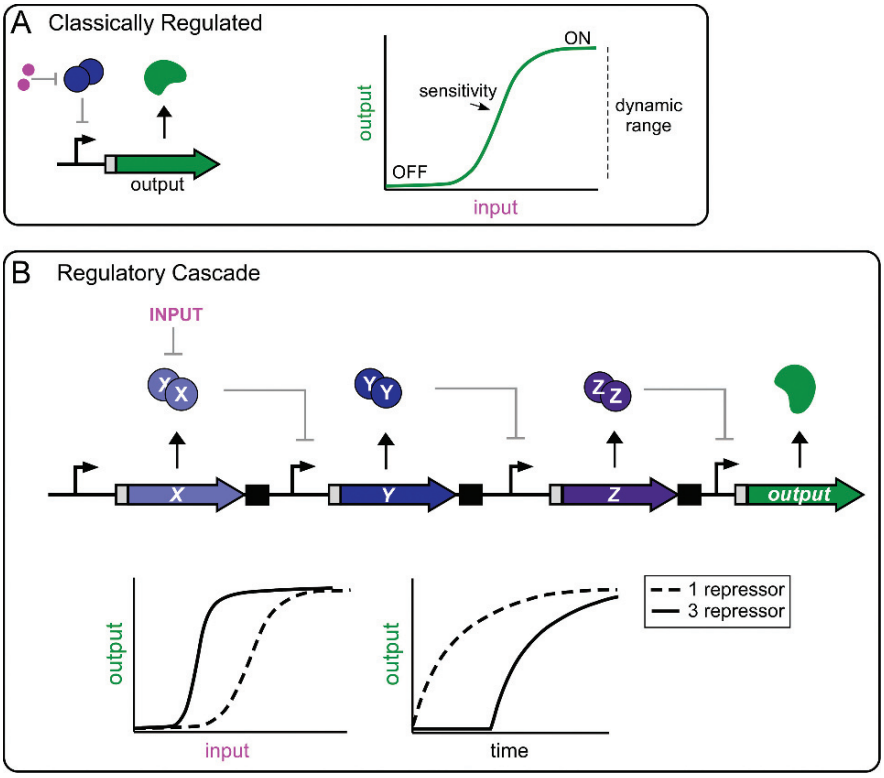

Fig. 19.5 (**A**) Transfer function. Classically regulated promoters typically show sigmoidal response profiles to the concentration of inducer. In the range of inducer (small dot) concentrations well below the K_D of the transcription factor, output changes little changes in input. In the responsive region the concentration of output increases steadily and continuously as a function of input concentration. At inducer concentrations well above the KD of the transcription factor there is no further increase in output. (B) Regulatory cascade. An input signal inactivates repressor protein X, resulting in the derepression of repressor Y. Upon accumulation of Y, repressor Z is repressed and its levels decline, increasing the concentration of the output. (*Lower left*) Cascading results in ultrasensitivity and lower sensing thresholds. A single repressor version of the above circuit (*dashed line*) shows a standard sigmoidal response. The 3 repressor cascade amplifies signal, reducing the absolute concentration of inducer required to activate the circuit and increasing the sensitivity of the response. (*Bottom right*) Cascading generates lags in response time. The single repressor circuit (*dashed line*) responds instantaneously to introduction of inducer while the 3 repressor cascade generates a significant latency in the response

The sensitivity (Wall et al. 2004), or slope of the response curve, in this range is largely determined by the cooperativity of binding of the transcription factor at the promoter of interest. Cooperativity refers to an effect where the affinity of a transcription factor for its DNA operator site increases as a consequence of a previous binding event by another transcription factor at a nearby operator (Bintu et al. 2005a, Ptashne and Gann 2002). This is often the result of protein-protein interaction domains which drive multimerization of the transcription factors on the DNA. Finally, as ligand concentrations increase well above the relevant K_D, the pool of transcription factors or relevant DNA operators become saturated and the output does not increase with further increases in input (Fig. 19.5A).

Certain features of the transfer function can be altered by changing the number and type of operator sites near the output promoter in a classically regulated system. For example the sensitivity, or log-log slope of the input-output function in the responsive range, is less than or equal to 1 for promoters with a single operator site. This is true whether the system is regulated by an activator or repressor (Bintu et al. 2005a,b). The addition of a second operator which enables cooperative binding can significantly increase sensitivity, typically ~2–4 fold (Bintu et al. 2005a,b). DNA looping can also be used to increase the sensitivity of the response (Vilar and Leibler 2003).

In activator systems, if binding is not cooperative, the sensitivity of the response remains the same with the introduction of a second operator, but the dynamic range of induction increases multiplicatively. In repressor based systems, additional operators which do not result in cooperative binding can still increase the sensitivity because the presence of a repressor at any the first site can significantly occlude the RNA polymerase, inherently facilitating binding of a repressor at the second site (Bintu et al. 2005a).

Continuous Response

Classically regulated transcriptional systems have the property of continuous responsivity. Continuous response means that the abundance of the output gene product in a single cell scales proportionally to the concentration of input signal in the environment. This allows the homogenous "fine-tuning" of output expression levels across an entire population. The fine-tuning of expression also allows the control of protein variance between individual cells, which has been shown to naturally decrease as protein abundance increase (Bar-Even et al. 2006). As we will discuss in Section 19.3.2.2, many natural genetic circuits lack continuous responsivity and some have even been intentionally modified to acquire it.

Speed of Response

An important performance feature of any circuit is the rise time, or time required after the addition of an input for the output to reach 50% of its steady-state value. This value, which has been measured for several systems in *E. coli* is approximately 1 cell cycle (45–135 minutes in these studies) (Mangan et al. 2006, Rosenfeld

et al. 2002). The time required for an *E. coli* cell to fully respond to an environmental stimulus via the classical mode of regulation is therefore greater than the time required for it to produce a complete copy of itself. The response time of classically regulated circuits can be increased by adding protease tags (Andersen et al. 1998) to speed degradation of regulatory proteins. The slow response times of classically regulated circuits will be compared with those of more complex circuits below.

19.3.1.2 Complex and Biphasic Promoters

Promoters bearing multiple operator sites which activate and repress gene expression can result in non-monotonic behavior in response to a monotonic increase in input signal. For example, the P_{RM} promoter of phage λ has three operator sites for the transcription factor CI. CI initially binds at two high affinity sites and has an activating effect on promoter output. When CI reaches higher concentrations, however, it binds to a low affinity site and functions as a repressor. A circuit wherein CI is expressed proportionally to an input can therefore result in an output which is OFF at both low and high input and ON only at intermediate inputs (Michalowski et al. 2004).

The operator for the AraC activator has been added to the *E. coli* lac promoter to generate a two-tiered activation response (Lutz and Bujard 1997). In this design, transcription increases proportional to the concentration of the first input IPTG but saturates at an intermediate level. This response is solely a function of promoter derepression. When provided saturating IPTG, the promoter can then undergo a second tier of activation proportional to the concentration of the activator arabinose. This occurs as a result of AraC mediated recruitment of RNA polymerase at the derepressed promoter. Many mutants of this promoter have also been constructed which offer different performance features as well (Lutz et al. 2001).

19.3.1.3 Regulatory Cascades

Multiple classically regulated circuits can be linked in series such that the output of one circuit serves as the input to the next (Fig. 19.5B). Cascades can be used to temporally order the expression of many different output genes in response to a single input stimulus (Kalir et al. 2001), allow cells to respond to increasingly small amounts of input (Hooshangi et al. 2005) and filter out transient or noisy input signals.

There are several inherent trade-offs in the use of regulatory cascades. For example, inducer sensitivity and signal amplification can be increased with the number of regulatory steps, but this occurs at a cost to response time. Moreover, lengthening can oftentimes require the redesign of upstream elements to ensure that the output ranges of the existing segment are matched to the input ranges over which the downstream segment can respond (Basu et al. 2004, Hooshangi et al. 2005, Yokobayashi et al. 2002).

Signal Amplification and Ultrasensitivity

To directly measure the performance features of genetic cascades, Weiss and coworkers constructed several synthetic genetic circuits that systematically increased the length of a cascade. This included circuits with 1, 2, and 3 repressors connected in series (Fig. 19.5B). As repressors were added to the cascade, the authors observed that the output reached half-maximal response at lower inducer concentrations; about 40% lower inducer per repressor added. Signal amplification allows cells to respond to inputs which are present in the environment at concentrations below the limit of detection of the natural sensory apparatus.

As with other circuit designs that we have discussed, regulatory cascades can increase sensitivity to the input (Hooshangi et al. 2005, Pedraza and van Oudenaarden 2005) (Fig. 19.5B). In the Weiss example, the range of inducer concentrations required to generate a full response decreased approximately 5-fold upon the addition of the second repressor and 8-fold upon addition of the third. Moreover, a mathematical model indicated that sensitivity would continue to increase as more than three repressors were added to the chain (Hooshangi et al. 2005).

Activation Delays

The relaying of an input signal through a multi-step regulatory cascade results in a temporal lag in response (Fig. 19.5B). Whereas a single repressor showed near immediate response and reached a steady-state output at two hours, the two repressor system took greater than six hours to reach steady-state (Hooshangi et al. 2005). The addition of the third repressor delayed signal transmission dramatically. This circuit showed no response to inputs at times less than two hours, and took 10 hours to reach steady-state. Furthermore, the model showed that with every two additional repressors added the rise time would continue to increase two-fold.

Cascade-Mediated Control of Complex Cellular Processes

The expression of many genes can be temporally ordered if they are regulated by cascades. The *E. coli* flagellum is encoded within 14 operons which contain its structural and regulatory genes. Upon induction, each operon is activated in an order commensurate with the sequence of assembly of the proteins which make up the flagellar apparatus (Kalir et al. 2001). The regulators in this cascade are able to activate each of their target operons in sequence with minutes long lag times in between. This highly regulated sequence of events is probably encoded at the DNA level by variable operator sequences at each promoter for which the regulators have slightly different binding affinities (Kalir et al. 2001). In this scenario, free floating cytoplasmic regulator proteins will occupy stronger operator sites before occupying any given lower affinity operator, allowing rank ordering of gene expression.

Quantitative measurements of gene expression in this system allowed the development of a rigorous computational model which could then be used to make predictive perturbations to circuit behavior (Kalir and Alon 2004). Similarly de-

tailed measurements of the regulatory interactions and their effect on gene expression will be invaluable in the troubleshooting, manipulation and optimization of forward engineered systems as well. Though synthetic biology is far from reliably designing structures as complex as the flagellum, one can envision many smaller scale applications where cascades could be used to time orders of expression in complex processes. For example, timed protein expression could facilitate the stepwise biosynthesis of novel antibiotics (Pfeifer et al. 2001), boost drug production (Keasling 2008) convert agricultural waste into fuel (Service 2007) or even coordinate the expression of existing complex cellular machines (Temme et al. 2008).

19.3.2 Feedback and Feed Forward Regulation

Linking the output of a classically regulated circuit back to its input or forward through intermediate regulators can dramatically alter its dynamic and steady-state properties. In this section we review the most common natural and engineered feedback and feed forward circuits, focusing on the impact of overall architecture and key parameters on circuit performance.

19.3.2.1 Negative Feedback

Negative feedback occurs when the output of a given circuit represses its own production (Fig. 19.6). Circuits controlled by negative feedback have unique response characteristics which are critical for certain biological design applications. Though negative feedback can be implemented as an inhibitory step at any point between production and decay of a gene product this section focuses on transcriptional feedback, which has been widely employed in the construction of synthetic circuits.

Response Accelerators

The response times of negative feedback circuits are markedly reduced compared to their analogous classically regulated counterparts (Savageau 1974). Using engineered variants of the *tet* system, Alon and coworkers experimentally demonstrated

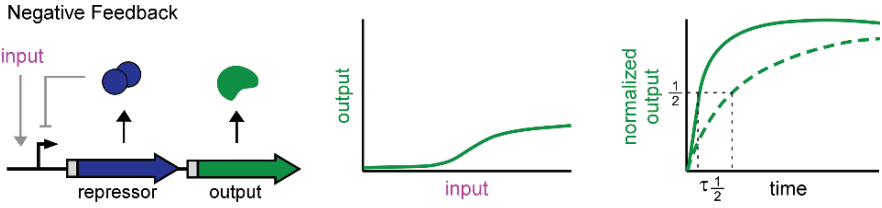

Fig. 19.6 Negative Transcriptional Feedback. A repressor protein is encoded under the control of the promoter which it regulates. The shape of the input/output curve is the same as in Fig. 19.5A above, but the system reaches equal or less output at any given concentration of input. The rise time (t(1/2)), or time required for the circuit to reach 50% of its steady-state output is significantly decreased in negative feedback (*solid line*) as compared to classically regulated (*dashed line*) systems

a reduction in rise time from over two hours to 15 minutes upon the introduction of negative feedback (Rosenfeld et al. 2002) (Fig. 19.6). The acceleration of the response is proportional to the strength of repression, a parameter which can be engineered by altering the number, strength or location of operator sites (Basu et al. 2004, Cox et al. 2007). Acceleration also increases with the cooperativity of binding of the repressor protein to the promoter (Rosenfeld et al. 2002, Savageau 1974). This term can be changed by the addition or removal of operator sites (Bintu et al. 2005a) or by the selection of repressor proteins with different oligomerization properties (Ninfa and Mayo 2004).

Though the negative feedback component reduces response time it also reduces the steady-state output of a circuit (Bashor et al. 2008, Rosenfeld et al. 2002). The rise time acceleration in negative feedback circuits occurs because shortly after induction the promoter is not repressed. Only after the accumulation of repressor does the activity of the promoter decrease to steady-state. This is in contrast to the classically regulated promoter which is active at a high level at all times after induction, resulting in a higher steady-state output which takes more time to achieve. The negative feedback circuit architecture is only useful, therefore, if the circuit output is operational at reduced steady-state expression levels.

Noise Buffering

Stochastic fluctuations, or noise, in gene expression is inevitable in genetic circuits and can reduce the fidelity of signal transmission and cellular decision making (McAdams and Arkin 1997). Moreover, as the number of components in an engineered circuit increases, the effects of noise in any one component can be compounded (Hooshangi et al. 2005, Pedraza and van Oudenaarden 2005).

It has long been recognized that negative feedback circuit architecture can reduce noise in output gene expression (El-Samad and Khammash 2006, Savageau 1974). To experimentally validate this effect, Becskei and Serrano constructed a synthetic circuit wherein a repressor protein inhibited its own transcription in *E. coli* (Becskei and Serrano 2000). Negative feedback reduced noise, measured as the coefficient of variation in protein expression across a population of cells, up to 70% over a circuit without feedback. Moreover, and as predicted (Savageau 1974), the magnitude of noise buffering was proportional to the strength of feedback.

The reason that negative feedback circuits buffer fluctuations is intuitive. In classically regulated transcriptional systems, fluctuations in any step of protein expression (transcription, mRNA decay, translation, etc.) are amplified by subsequent steps and cause variation in protein abundances between individual cells. In negative feedback circuits, fluctuations that cause increases in the output protein concentration are quickly dampened by increased repression while fluctuations that cause the output levels to decrease reduce repressor abundances and increase transcription rates. The result is that the system returns to steady state more rapidly after random fluctuations.

There is a caveat to the use of negative feedback as a safeguard against noise in engineered circuits. Though noise decreases proportional to feedback strength over

a large range of protein abundances (Becskei and Serrano 2000, Thattai and van Oudenaarden 2001), noise can actually increase if the strength of negative feedback becomes too strong (Shahrezaei et al. 2008). This is due to a phenomenon known as the "small number effect" where the impact of intrinsic fluctuations in chemical reactions increases rapidly as the concentration of reactants becomes very small (Bar-Even et al. 2006, Kaern et al. 2005). That is, at smaller protein concentrations each random protein production or decay event has a larger impact on the mean concentration. This highlights the general biological design principle that increasing the number of proteins in a cell reduces noise in protein abundance (Bar-Even et al. 2006).

19.3.2.2 Positive Feedback

Positive feedback occurs when the output of a circuit activates its own production (Fig. 19.7A). Circuits with positive feedback can have many features which are valuable in the engineering of more robust, decisive cellular behaviors including ultrasensitivity, bistability, hysteresis and memory (Fig. 19.7B–D). This section

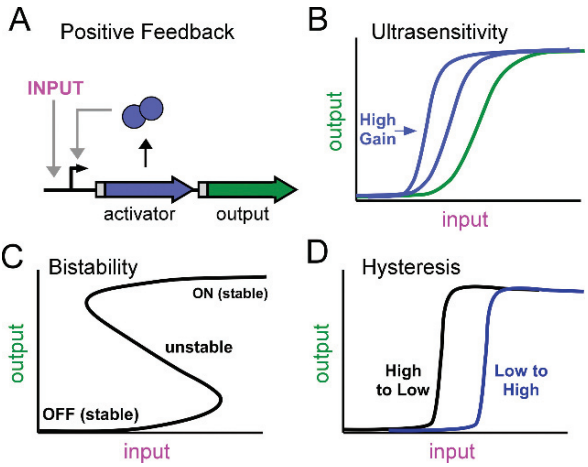

Fig. 19.7 Positive transcriptional feedback. (**A**) A self activating protein is expressed under control of the input. (**B**) Positive feedback circuits with lower kinetic orders of transcription factor binding and cooperativity result in ultrasensitive responses to inducer. Sensitivity is measured as the slope of the relationship between output and input. This increases from the classically regulated system (*right most line*) to 2 positive feedback systems with increasing kinetic constants of activation (*left most lines*). (**C**) Positive feedback circuits with very high kinetic orders of activation can achieve bistability. In these circuits, cells rest at low output levels or high output levels but never at intermediate output levels. (**D**) Hysteresis. When starting at low inducer concentrations and moving higher the circuit requires some amount of inducer to switch ON. When starting in the ON state and reducing the concentration of inducer available to the circuit, the switch happens at a significantly lower concentration

describes the performance features of positive feedback loops and how they can be changed by modifying the underlying circuit parameters.

Response Delays

In contrast to negative feedback circuits which accelerate response times, positive feedback circuits are thought to slow the rise to steady-state. Though it has not been measured in a well-controlled experimental setting, the magnitude of the rise time delay is predicted to be proportional to the strength of the positive feedback step (Savageau 1974). For a transcriptional circuit, feedback strength is governed by the binding affinity of the output transcription factor for its DNA operators, the mode by which the transcription factor interacts with RNA polymerase and its cooperativity (Bintu et al. 2005a, Ninfa and Mayo 2004).

The most direct strategy for manipulating the magnitude of delay in a positive feedback circuit is to vary the DNA operator sites at the promoter to which the activator binds. This can be done by varying the number, spacing and sequence of the operators. Single nucleotide mutations within operator sites can significantly reduce the affinity of a transcription factor for its operator (Basu et al. 2004, Falcon and Matthews 2000, Frank et al. 1997, Takeda et al. 1989). Increasing or decreasing the spacing between multiple operators can affect both binding affinity (Chen and Kadner 2000) and cooperativity of binding (Smith and Sauer 1995).

The introduction of positive feedback increases the steady-state output level of a classical transcriptional circuit. To compensate, one can decrease the production or increase the decay rate of the circuit output. For example, weakening the strength of the self-activating promoter or adding a degradation tag to transcription factor would reduce the steady-state and serve to more closely match the expression levels of a the two circuits.

Ultrasensitivity

It has been demonstrated that regulatory systems with positive feedback are more sensitive to inducer than systems without feedback (Fig. 19.7B) (Savageau 1974). Positive feedback has since been experimentally verified to impart ultrasensitivity in both natural and engineered circuits (Ferrell and Machleder 1998, Bashor et al. 2008). Ultrasensitivity occurs in positive feedback circuits where the strength of the feedback is not so large that the system loses the ability to occupy intermediate output states. The level of ultrasensitivity can be controlled by manipulating the strength of feedback. This can be achieved by changing the stability of the activator protein or its cooperativity or binding affinity at the promoter.

Bistability

Positive feedback can also create a bistable switch (Ferrell 2002). Bistable circuits can occupy only one of two states, canonically an OFF and an ON state, in response to a continuous range of input concentrations (Fig. 19.7C). This can be very useful

in circuit design and will be discussed further in Section 19.3.3. It is challenging to design bistable circuits based on positive feedback (Ajo-Franklin et al. 2007) because if either of the states is quantitatively off balance with the other, the system will only be able to occupy one state (Ferrell 2002). For example, leaky transcription of the positive feedback element is often sufficient to trip the switch and keep the circuit in a monostable ON state under all conditions.

A bistable switch based on positive transcriptional feedback has been constructed in *E. coli* (Isaacs et al. 2003). This circuit was composed of a temperature-sensitive transcriptional activator expressed under the control of the promoter which it activated. High kinetic constants of dimerization and transcriptional activation provided the non-linear responsivity required for bistability. At permissive temperatures, leaky transcription tripped the feedback switch driving all cells in the population to reach a stable ON state. At destabilizing temperatures, a lack of activator accumulation kept all cells in the OFF state. At intermediate temperatures the population bifurcated such that individual cells occupied either the ON or OFF state. This digital response occurred because intermediate protein expression states in any cell are unstable and small fluctuations are amplified to drive cells to quickly settle in either of the stable states (Ferrell 2002, Isaacs et al. 2003).

Bistable circuits have a unique property in that they can achieve different steady-state output responses under identical input conditions depending on their history (Fig. 19.7D) (Ferrell 2002, Ninfa and Mayo 2004). That is, if the circuit begins in the OFF state it requires a greater input concentration to switch than if it began in the ON state. This characteristic, known as hysteresis, is useful in the engineering of robust cellular decision making. This is because hysteresis makes circuits with switch-like behaviors less sensitive to fluctuations in input signal near the switch point.

Ninfa and coworkers designed a transcriptional positive feedback circuit with a dominant repressor protein to construct a bistable switch in *E. coli* (Atkinson et al. 2003). In the absence of inducer, the repressor inactivated the feedback loop and the switch was OFF. At activating concentrations of inducer the circuit rapidly switched to the ON state. If the circuit had previously been exposed to high levels of inducer, however, it switched ON at $\sim$ 70% lower inducer concentrations. Two key circuit parameters drove this system to exhibit hysteresis. First there was very high sensitivity within the switching range making intermediate expression states unstable. Second, the dynamic range of induction was large, on the order of 20-fold. These are the two most critical design requirements in the construction of positive feedback circuits with hysteretic properties (Angeli et al. 2004, Ferrell 2002, Ninfa and Mayo 2004).

Controlling Feedback Saturation

In a bistable switch, the magnitude of output gene expression in the ON state is determined by the protein production and decay parameters of the circuit. The level of gene expression in an activated bistable switch can therefore not be fine tuned. Because the steady-state output level is often an important design consideration in

genetic engineering applications, we will discuss several strategies for controlling the magnitude of the ON state, or point of feedback saturation here.

In a simple positive feedback circuit, where an activator protein drives its own promoter, the steady-state output of the fully activated circuit is determined by the maximal rate of production and decay rate of the protein. In the synthetic circuit constructed by Collins and coworkers, the per cell output of the fully activated switch decreased continuously as the activator protein was destabilized (Isaacs et al. 2003). It is likely though that other circuit parameters such as promoter or RBS strength, or mRNA stability could be modified to achieve a similar result.

Eliminating Bistability to Generate a Continuous Response

Sugar inducible systems like *lac* and *ara* are the most widely used elements for engineered genetic control in *E. coli*. They are bistable because sugar-mediated transcriptional activation increases the rate of sugar uptake from the environment, generating a positive feedback loop. For many engineering applications this bistability is undesirable. Bistability creates discontinuous jumps in output as inducer is added. This hampers the freedom of the genetic engineer to set the circuit at intermediate output phenotypes. Moreover, at intermediate inducer concentrations the population can bifurcate such that some cells occupy the OFF state, some the ON state and none occupy an intermediate state. In many applications in biotechnology it is beneficial for all cells in a population to behave identically. Fortunately, the bistable feedback circuits which nature provides can be modified for continuous input-output control and population homogeneity.

Several groups have shown that by expressing sugar uptake genes constitutively the positive feedback loops can be broken and bistability eliminated, allowing continuous induction over a large range of inducer (Khlebnikov et al. 2001, Khlebnikov and Keasling 2002, Khlebnikov et al. 2000, Morgan-Kiss et al. 2002). The deletion of the sugar catabolic genes from the host also aids in the homogeneity of the response (Morgan-Kiss et al. 2002).

19.3.2.3 Feed Forward Loops

A common genetic circuit in *E. coli* is the feed forward loop (FFL), where an input is split into two pathways, which then reconverge on an output (Milo et al. 2002, Shen-Orr et al. 2002). In its simplest form, an FFL consists of two regulatory genes (canonically X and Y) and one output gene (Z). Feed forward architecture results when X regulates the production of Y and both in turn regulate the production of Z (Fig. 19.8).

There are two major classes of FFLs. In the first class, termed coherent FFLs, the sign of the regulatory interaction remains the same all the way through the circuit. That is X regulates Y and Z in the same manner that Y regulates Z. Coherent FFLs therefore regulate outputs similarly to single transcription factors, but introduce several quantitative performance differences. In the second class of FFLs, termed incoherent FFLs, the regulatory effect changes after the circuit splits, resulting in

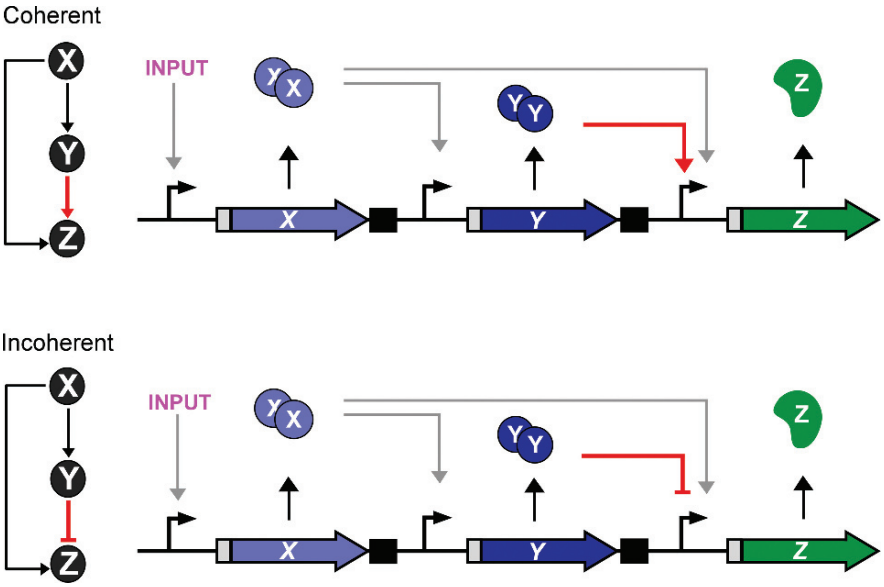

Fig. 19.8 Feed Forward Loops (FFLs). FFLs are genetic circuits composed of three proteins, X, Y and Z. X and Y are transcription factors which regulate the expression of Z. The "feed-forward" connectivity refers to the fact that X also regulates Y. Coherent FFLs result when the regulatory relationship between X and Z is the same as that between X and Y. Incoherent FFLs arise when these two relationships are opposite

opposing regulation at the output. As we will discuss below, this type of circuit can result in interesting dynamic behaviors such as overshoots or pulses of gene expression.

Coherent FFLs: Activation Delays

A FFL is coherent if the regulatory effect of X on Z is the same as the effect of X on Z through Y (Fig. 19.8). Coherent FFLs have been shown to act as sign-sensitive delays in *E. coli* signal processing networks (Kalir et al. 2005, Mangan and Alon 2003, Mangan et al. 2003). "Sign-sensitive" refers to the fact that the circuits generate a lag in the transcriptional response to either the introduction or removal of an environmental signal, but not both. Activation delays can function as noise filters in that they prevent the circuit from responding to transient pulses of signals. Coherent FFLs are useful tools then for the engineering of sense-response behaviors in which the cell must parse sustained signal from input noise in the environment.

The basis of the delay in this type of FFL is intuitive. Z depends on the presence of X and Y for expression. Though the presence of input signal immediately activates X, Y cannot be expressed until X first accumulates. From that point, Y must then accumulate to a concentration sufficient to activate Z. Indeed, increasing the basal expression level of Y decreases the length of the delay (Mangan et al. 2003).

The basal expression rate of an activator protein in a synthetic circuit is simple to tune with promoters or RBSs of different strengths, for example.

The natural arabinose responsive circuit of *E. coli* is a coherent FFL. This is not true, however, for the minimized pBAD circuit from which one of the natural regulators has been removed (Guzman et al. 1995). The natural arabinose FFL circuit generates a delay in the activation of transcription after induction (Mangan et al. 2003). In *E. coli*, the absence of glucose increases intracellular cyclic adenosine monophosphate (cAMP) levels which activate the transcription factor CRP (X). CRP activates the expression of the *araC* (Y) gene, the product of which is a transcription factor whose function is dependent upon arabinose. The output araBAD (Z) promoter functions as a logical AND gate, requiring the presence of cAMP:CRP and arabinose:AraC for productive transcription. This FFL results in a ∼0.2 cell cycle or 10–20 minute delay in activation of the Z promoter after the onset of inducing conditions. The delay is shown to be sign sensitive as the removal of the stimulus does not result in a delayed inactivation response as compared to a simple AND gate promoter without a feed forward connectivity between the two transcription factors.

Deactivation Delays

The sign sensitivity of a FFL mediated delay can be changed by changing the activation logic of the Z promoter from AND to OR (Mangan and Alon 2003). Alon and coworkers proofed this concept by demonstrating that part of the *E. coli* flagellar apparatus is expressed under the control of a coherent FFL in which Z is expressed as a SUM function of X and Y (Kalir et al. 2005). SUM is a modified OR where the influence of X and Y on Z output is additive. Moreover, SUM is a simple operation to engineer in *E. coli*. SUM can be achieved by simply placing two different promoters in series. In this configuration, the first or second promoter can drive expression of the output gene, and if both are active, the rate of production of mRNA is greater.

In the flagellar example, X activates Y and the two transcription factors additively activate the operons that produce the flagellar motor (Kalir et al. 2005). If the input signal is removed and X is transiently inactivated, the circuit prolongs flagellar expression because Y levels linger. The authors show that the delay occurs under a wide range of circuit parameters, and that manipulation of the kinetic parameters of regulation can alter the length of the delay (Kalir and Alon 2004). A similar effect was shown for the Salmonella SPI-1 Type III Secretion System, which contains both a feed forward and split positive feedback loop (Temme et al. 2008).

Incoherent FFLs

An incoherent FFL consists of a circuit where X activates Y and Z but Y represses Z (Fig. 19.8). There are over 100 examples of incoherent FFLs in the *E. coli* genome (Mangan et al. 2006). This circuit generates several interesting and unique dynamical outputs such as pulses of gene expression and time-derivative sensing (Basu et al. 2004). In this section we will discuss the performance features of incoherent FFLs in *E. coli*, the effect of key molecular parameters on their function, and their

application in the construction of some of the most sophisticated synthetic cellular behaviors to date.

ON Accelerators

Because X first activates and then indirectly represses the expression of Z, incoherent FFLs result in "overshoot dynamics" in the expression of Z (Mangan and Alon 2003). This means that Z temporarily reaches abundances greater than the final steady-state. Also, because the strength of a partially repressible promoter driving Z must be stronger than that of a non-repressible promoter capable of generating the same steady-state, the rise time of the output Z is necessarily increased in an incoherent FFL of this form (Mangan and Alon 2003). This property is similar to the accelerated response of negative feedback loops as described above. In a natural example, Alon and coworkers have demonstrated that the incoherent FFL in the galactose utilization network of *E. coli* results in a 1.75-fold overshoot of the steady-state output and an approximately 3-fold acceleration in rise time (Mangan et al. 2006).

In incoherent FFL circuits, important performance features such as the magnitude of response acceleration, the steady-state output and the size of the overshoot are particularly sensitive to the parameters associated with the repressor Y. In general, the higher the expression level of Y and the greater its repressive effects, the greater the acceleration of the circuit (Mangan and Alon 2003).

Pulse Generators

A pulse generator is a genetic circuit capable of activating and then completely repressing output gene expression in response to the addition of an input. Incoherent FFLs can generate pulses of gene expression if the repression of Z by Y is very strong. In 2004, Weiss and coworkers constructed a synthetic incoherent FFL in *E. coli*. In their design X was the transcription factor LuxR which is activated by the membrane permeable quorum sensing compound AI-1, Y was the strong transcriptional repressor CI and Z was the reporter gene *gfp*.

Because the circuit was constructed *de novo*, the authors could easily investigate the effects of genetic parameters such as the rate of synthesis of Y, and the strength of repression Z by Y. The authors noted that if either of these two parameters was too great, the circuit could never be activated by inducer (Basu et al. 2004). Under a range of permissive kinetic parameters, however, the circuit showed robust pulse generation after addition of inducer. The true pulse of gene expression occurred because the Y protein CI is a very strong repressor of its target promoter, capable of bringing output expression back to zero.

Critical pulse features such as amplitude and duration could be controlled by varying the kinetic parameters of the Y protein or the rate or concentration at which inducer was added. Specifically, the stronger the RBS or the affinity for the output promoter the shorter and smaller the resulting pulse. Furthermore, at fixed Y kinetic parameters, the pulse amplitude varied proportionally to both the absolute

concentration and the rate of increase of inducer. This synthetic circuit is an elegant demonstration of the level of behavioral sophistication that can be designed *de novo* and optimized to the specifications of the engineer.

19.3.2.4 Dynamic Circuits

Several genetic circuits have been engineered which drive dynamic responses. A striking example is the three protein transcriptional ring oscillator known as the "repressilator" (Fig. 19.9A) (Elowitz and Leibler 2000). In this circuit, protein A represses protein B, protein B represses protein C and protein C represses protein A. Oscillations occur because the addition of an input signal can cause one of the proteins, say A, to become abundant and repress the next protein in line (B). Because B is repressed, C begins to rise in abundance and can then in turn repress A. This process continues until A rises again, and in this manner the circuit encodes genetic oscillations. The repressilator was capable of generating three to four oscillations in a given cell, but showed a notable lack of uniformity across the population.

In another example, Ninfa and coworkers constructed a two-component transcriptional oscillator in which a transcription factor first activates itself and then activates its own repressor (Fig. 19.9A) (Atkinson et al. 2003). In this circuit an input triggers the activator to initially accumulate. After some time the activator is repressed by the accumulating repressor. As activator levels subsequently fall, so do repressor levels, triggering another round of activator accumulation. This circuit drove dampened oscillations over four periods, which spanned almost 60 hours.

A circuit based on cell-cell communication has been constructed to program population level oscillations in *E. coli* (Balagadde et al. 2005). In this design a gene

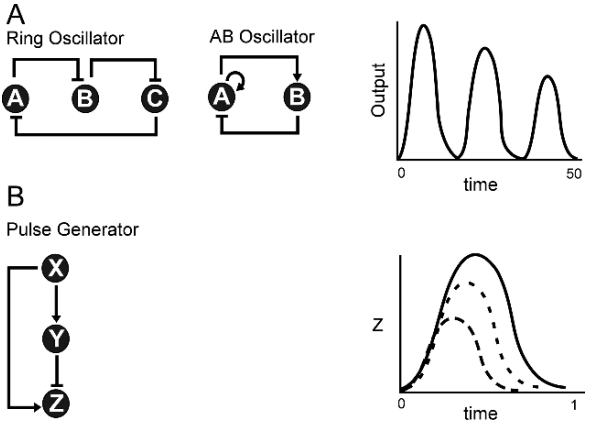

Fig. 19.9 Dynamic genetic circuits. (**A**) Genetic circuits composed of three transcriptional repressors in a closed loop or a self activating protein which also activates its own repressor can cause oscillations in gene expression. (**B**) Pulse Generator. An incoherent Feed Forward Loop produces temporal pulses of gene expression. The strength of expression or the kinetic order of repression of the repressor Y can change the duration and amplitude of the pulse (*dashed lines*)

which triggered cell death was expressed under the control of a quorum sensing circuit. The circuit was OFF at low cell densities but switched ON at high density. Microscopic monitoring demonstrated that *E. coli* expressing this circuit regularly oscillated in density from 1 to 3 cells per picoliter of media with a period of about 20 hours.

As discussed in Section 19.3.2.2, Weiss and coworkers also constructed a dynamic circuit capable of generating a temporal pulse of gene expression in response to a single, step introduction of input signal (Fig. 19.9B) (Basu et al. 2004). The amplitude and duration of the pulse could be programmed by changing the strength or production rate of the repressor in the circuit. Moreover, because the circuit input was a membrane diffusible quorum sensing compound, a cell could be triggered to pulse by production of the activator in a nearby cell.

19.3.3 Switches and Logic

Genetic switches are circuits which rapidly transition between discreet states in response to input. Logical devices are circuits which interpret the states of multiple switches to produce a single, unified output. Switches and logic are useful because they aid the programming of desirable IF/THEN behaviors in *E. coli*. Genetic logic is carried out by circuits which can be rationally designed or combinatorially screened.

Extensibility, or the ease with which a switch or logic device can be connected to a different input or output is a desirable trait in switches and logic devices. Extensibility requires knowledge of the transfer functions of the parts. For example, the output range of a given switch or switches must be matched to the input range of a given logic device in order for signal transmission to proceed properly through the circuit. If expression in the OFF state of a switch is leaky and it is interpreted as ON by the downstream logic gate, then the circuit will not properly respond to input signals. If the transfer functions of switches and logic gates are well documented, however, they can be used "off the shelf" and connected to other well characterized parts.

NOT Gate

One of the most useful and frequently constructed genetic logic operations is the Boolean NOT gate. Commonly referred to as an inverter, the NOT gate inverts the sign of the regulatory relationship between the input and output of the circuit. In the simplest system, this is accomplished by the introduction of a transcriptional repressor between the input and output (Fig. 19.10A). An input signal which would otherwise activate expression of the output therefore inactivates it via the activation of a repressor. Besides inverting the input/output logic, NOT gates are also known to increase input sensitivity (Hooshangi et al. 2005, Karig and Weiss 2005, Pedraza and van Oudenaarden 2005) and lower sensing thresholds (Karig and Weiss 2005).

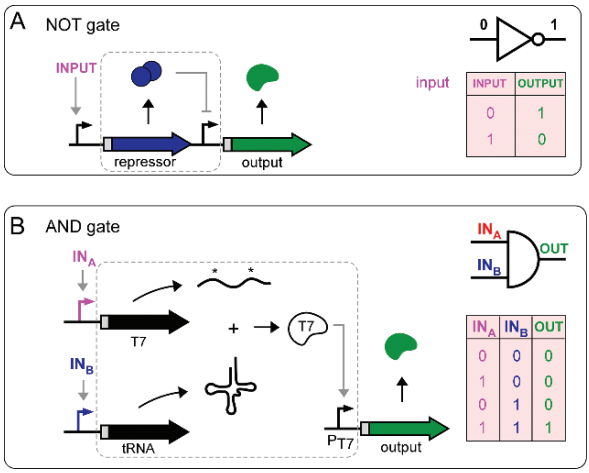

Fig. 19.10 Transcriptional Logic (**A**) NOT gate. Also known as a genetic inverter, the NOT gate encodes a repressor under the control of the environmental input. The repressor inactivates expression from an otherwise active output promoter. The inverter device (*dashed box*) comprising the repressor protein and the output promoter is an independent module which can be placed between any input promoter and output gene. The logic of the NOT circuit (*upper right*) is shown in the truth table (*bottom right*). (**B**) AND gate (*dashed box*) comprises an untranslatable T7 RNA polymerase mRNA bearing two stop codons (*asterisks*) in the open reading frame and a suppressor tRNA which incorporates amino acids at the stop codons to allow productive translation. Only when both halves are transcribed is T7 RNAP produced and does the output promoter become active. Each half of the AND gate can be driven by any inducible promoter, activated by its cognate input signal. Adapted from Anderson et al., 2007

Many genetic circuits containing NOT gates fail to function properly when constructed. This often occurs because basal expression of the repressor in the absence of input can be sufficient to inhibit the output promoter, constitutively trapping the inverter in the OFF state. The abundance of the repressor protein can then be reduced to match the relevant sensitivity of the output promoter. This can be accomplished by weakening the RBS on the repressor mRNA, weakening the operator sites at the output promoter (Hooshangi et al. 2005, Weiss 2001, Yokobayashi et al. 2002) or randomly mutating the repressor to reduce its strength (Yokobayashi et al. 2002)

Switches and Memory

Memory is required for many sophisticated functions in electronic systems and is also ubiquitous in molecular biology, forming the basis for the burgeoning field of epigenetics. One popular biological design goal which relies on memory is to construct cells that can count how old they are or how many times they and their ancestors have been exposed to some signal over a long period of time. Memory can be implemented as an extreme form of hysteresis in circuits with strong positive feedback. In such systems, previous exposure to high input signal triggers a circuit to remain active even when the signal goes to zero (Ferrell 2002).

In 2000, Collins and coworkers constructed a memory switch in *E. coli* (Gardner et al. 2000). The switch was comprised of two cross-inhibiting transcriptional repressors. If repressor A was expressed, it repressed B and the switch was OFF. If an input was added which inactivated A, B accumulated and in turn, repressed A. This turned the switch ON. This switch generated stable memory over at least 22 hours, allowing a cell many generations away from the ancestor which had received the signal to maintain a stable response. This switch required proper matching of the transfer functions of its two subcomponents. If the expression level of one repressor was too strong in the OFF state the system became monostable. This required the screening of several combinations of promoters and RBSs of different strengths.

Arkin and coworkers have also constructed a memory device based on a permanent genetic rearrangement event. This circuit makes use of the recombinase encoded by the *fimE* gene to flip an improperly oriented promoter into alignment with an output gene (Ham et al. 2006). The DNA reorganization event is permanent, resulting in stable long-term circuit memory. Moreover, because the *fimE* gene can be expressed as the output of any sensor, the *fimE* switch is modular and can potentially generate memory of any input stimulus capable of activating gene expression. An advantage of this circuit is that it produces virtually no basal expression when the promoter is in the opposite orientation from the gene it controls.

AND Gate

The logical AND operation, where the presence of two inputs (A and B) are required to activate output expression, is a useful concept for biological design and can be applied to the construction of many more sophisticated logical operations. The most parsimonious strategy for the construction of a genetic AND gate involves two interdependent genetic components which, when expressed simultaneously can initiate a downstream gene expression step. Such a system was recently implemented at the transcriptional level in *E. coli* (Anderson et al. 2007). In this setup, inducible promoter A drives the expression of an mRNA encoding the T7 RNA polymerase (RNAP) gene. The mRNA is non-functional, however, as two specific stop codons are introduced into the coding sequence. Inducible promoter B drives the expression of a tRNA which encodes an amino acid at those stop codons, rescuing translation of the RNAP. The circuit output is a promoter which is only transcribed by T7 RNAP protein such that it requires the presence of the two inputs A and B (Fig. 19.10B). Importantly, this system was designed to be modular such that any two inducible promoters could be used to drive the AND gate. This modularity allowed the circuit to integrate signals from four different promoters and drive two separate output genes.

In the initial circuit design, the two components of the AND were not properly matched. The basal, or leaky expression rate of the T7 mRNA was significantly high that the circuit produced positive output in the presence of only one input. To reduce leaky expression, the authors randomly mutated the RBS preceding the T7 open reading frame and screened a library for variants dependent upon both inputs for activation. A majority of the variants in the library showed significant

dependence on both inputs, indicating that the design was quite robust to variable expression levels. When the promoter driving the T7 mRNA was replaced, however, the new RBS failed to generate enough mRNA to activate the AND gate even when the promoter was fully active. To restore functionality an RBS library was again screened and again produced a viable circuit.

Other Logic

To construct other types of genetic logic, Leibler and coworkers randomly connected five promoters to three classical transcription factors which either activated or repressed them. Two ligands were chosen as inputs and one of the transcription factors was chosen to repress an output reporter gene. Several switch-like logical responses including NAND, NOR and NOT IF arose repeatedly from the circuit library (Guet et al. 2002). Interestingly, circuits with similar connectivities, or profiles of regulatory contacts between components, were capable of generating different logical responses while networks with different connectivities were capable of generate the same logic. Many of the constructed circuits also produced intermediate or "fuzzy" logic.

A large number of intermediate logical operations were also observed in a related study wherein four different transcription factor binding sites were randomly placed in three locations around a single promoter (Cox et al. 2007). This combinatorial approach revealed that activator sites function most effectively when placed directly upstream of the −35 site and function poorly if at all when placed downstream of it. Repressor sites are more tolerant to different locations but are most effective when placed between the −35 and −10 sites. These efforts demonstrate the power of screening random combinations of regulators to achieve a desired logical operation.

19.4 Actuators: Interfacing Cells with the Environment

A fundamental motivation for programming cells is that they have the ability to modify the chemistry and biology of their surrounding environments. Actuators are defined as gene products which carry out any type of cellular process or behavior from an enzyme capable of synthesizing drugs or fuels to the synthesis and control of entire organelles and molecular machines. This section is meant to only briefly outline some of the things that *E. coli* can do.

State Reporters

State reporters are molecules whose only function is to be observed or measured. When linked to genetic circuits, reporters can provide a "print-out" of information coming in from cellular sensors and circuits. In biosensing applications the acquisition of information about the presence, absence, concentration or temporal profile of an input signal in the environment or the cell is itself the goal of the system. Reporters can also provide a physical read out of the solution of computations performed by genetic circuits. The most common reporters are proteins such as

β-galactosidase or Green Fluorescent Protein (GFP), the abundances of which can easily be measured by standard molecular biological techniques.

Metabolic Engineering

Metabolic engineering involves the expression of enzymes which divert cellular metabolites into alternative pathways to produce desired output products (Lee and Papoutsakis 1999). The enzymes used in metabolic engineering are therefore actuators which can be expressed as the outputs of genetic circuits. A typical metabolic design might employ sensors which detect the presence of upstream metabolites to time the expression of the biosynthetic enzymes which act upon them.

One application of metabolic engineering is the production of liquid fuels (Jarboe et al. 2007, Keasling 2008, Mielenz 2001, Service 2007). To this end, Liao and coworkers recently re-engineered *E. coli* amino acid metabolism for the production of branched chain alcohols, compounds which have desirable fuel properties (Atsumi et al. 2008). This required the expression of one of several two-enzyme clusters which converted intermediate metabolites from amino acid biosynthetic pathways to the various alcohols. Endogenous amino acid metabolic genes could also be overexpressed as complementary actuators to increase flux through the pathways and bump up fuel yields.

Metabolic actuators can be used to reprogram *E. coli* for the production of therapeutic compounds as well (Pfeifer et al. 2001, Swartz 2001, Zhang et al. 2006). For example, Keasling and coworkers have introduced a large number of non-native isoprenoid biosynthetic enzymes into *E. coli* to efficiently convert the ubiquitous metabolite acetyl-CoA into artemisinic acid, a direct precursor to the potent and otherwise prohibitively expensive anti-malarial compound artemisinin. Optimization of enzyme expression levels and compensatory engineering to eliminate toxic byproducts has resulted in profound improvements in yield, approximately 1 million fold increase in a 4 years span (Keasling 2008). These efforts are likely to reduce the cost of artemisinin orders of magnitude, to prices compatible with its utilization in many underdeveloped countries with high malarial death rates.

Most metabolic engineering efforts to date have expressed the actuators under the control of classically regulated circuits. These have been chosen for their simplicity and the continuous fine-grained control that they offer over enzyme expression levels. The construction of more sophisticated sensor-circuit-actuator systems should facilitate the design of increasingly ambitious microbial factories and help to optimize yields.

Organelle Transfer

Clusters of genes encoding entire organelles can also be used as actuators. Historically, the ability to manipulate such large fragments of DNA has required the presence of fortuitous restriction sites in the natural organelle sequences or specialized polymerase chain reaction (PCR) based methods. Improved DNA synthesis technologies now allow the *de novo* fabrication of organelle scale fragments.

In the initial demonstration of organelle transfer, 11 genes responsible for the synthesis of cytoplasmic gas vesicles in *B. megaterium* were moved into *E. coli* (Li and Cannon 1998). Expression of this gene cluster from a classically regulated circuit on a standard expression plasmid resulted in the formation of functional vesicles capable of significantly increasing the buoyancy of *E. coli* in aqueous media.

Similar strategies have resulted in the transfer of the fully functional nitrogen fixation (*nif*) (Dixon et al. 1976) and O antigen lipopolysaccharide (Bastin et al. 1991) enzyme clusters from *Klebsiella* and enteropathogenic *E. coli*, as well as the Type III protein secretion organelle from *Salmonella* (Wilson et al. 2007) and the cryptic Type II organelle from *E. coli* itself (Francetic et al. 2000). These efforts used unmodified, contiguous genomic DNA fragments which were recombined into plasmids and introduced into *E. coli* "as is". These strategies therefore relied on expression from the natural promoters and RBSs of the relevant genes, and necessarily introduced the possibility of regulation by undefined control elements. The utility of organelle actuators will undoubtedly benefit from control through sensors and circuits.

Building Genetic Programs

Sensors, circuits and actuators are true modular engineering components only when they can easily and arbitrarily be linked together. Several methodologies have recently been developed which allow the combination of multiple stretches of DNA without the need for inherent restriction sequences. One example is a universal, iterative cloning method for the assembly of standardized "Biobrick" parts (Shetty et al. 2008). In this method, a DNA part is computationally designed to internally lack several specific restriction sites. These restriction sites are then added to the upstream and downstream regions of the part and used as universal handles for the iterative, arbitrary connection of any two Biobricks. This standardized strategy increases the efficiency and ease with which any two parts can be combined (composability).

A PCR-based strategy termed SLIC has recently been developed for the "one-pot" assembly of up to 10 unrelated stretches of DNA in a specific order (Li and Elledge 2007). This method uses oligonucleotide primers to add specific linker sequences to any piece of DNA which then guide the order of assembly. The benefits to this approach are that no specific restriction sites need be avoided in the internal sequence of any part and that more than two parts can be combined in one step. Other advanced assembly strategies based on large scale oligonucleotide synthesis and polymerase chain reaction (PCR) assembly have allowed the construction of complete viral (Cello et al. 2002, Smith et al. 2003, Tumpey et al. 2005) and even bacterial (Gibson et al. 2008) genomes from computationally designed DNA information.

Standardization and assembly technologies are already helping eliminate barriers between the design and physical construction of DNA (Endy 2005), a process which has been the historical rate limiting step in genetic engineering. A true leap in biological design will occur when these technologies become more widely

available and less expensive, allowing true modular assembly of sensors, circuits and actuators. In an early watershed example, Collins and coworkers linked a DNA damage sensor to a bistable genetic switch to drive an actuator which triggered biofilm formation in *E. coli*. In this bottom up programming effort, the *E. coli* could stably and strongly switch ON biofilm formation phenotype in the presence of DNA damaging environmental inputs such as UV light or a chemical mutagen (Kobayashi et al. 2004).

Finally, when genetic parts are linked together in a design, their quantitative input/output properties must be properly matched (Yokobayashi et al. 2002). As discussed in Section 19.3.3 above, if the OFF state of a sensor is sufficiently leaky to activate a downstream genetic circuit, the circuit will not be capable of receiving signaling information from the sensor. There are many strategies for matching the transfer functions of multiple parts, but until universal metrics of genetic activity can be established (Endy 2005, Canton et al. 2008) there will always be a significant troubleshooting component in the assembly of functional systems.

19.5 Conclusions

The vast molecular genetic literature on *E. coli* has made it the subject of choice for many early efforts in synthetic biology. Five decades of work have given genetic engineers a rich repository of parts, often sensors and actuators, which can be taken out of their natural context and used for new, user-defined purposes. More recent efforts have established useful circuit design principles that have further pushed the level of sophistication of behaviors that can be designed.

Complementing the scientific contributions, DNA synthesis and sequencing technologies have become increasingly high throughput and less expensive in the past few years. Further advances will bolster biological design by allowing researchers to bypass the arduous process of physically constructing designed DNA sequences. In the end, *E. coli* synthetic biology serves two major purposes. It enables the goal-oriented engineering of strains which can carry out novel functions of medical, industrial or academic interest and it serves as a bottom-up complement to top-down systems approaches for the elucidation of the molecular principles which govern cellular behavior.

References

Ajo-Franklin CM, Drubin DA, Eskin JA et al. (2007) Rational design of memory in eukaryotic cells. Genes Dev 21(18):2271–6

Andersen JB, Sternberg C, Poulsen LK et al. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64(6):2240–6

Anderson JC, Clarke EJ, Arkin AP et al. (2006) Environmentally controlled invasion of cancer cells by engineered bacteria. J Mol Biol 355(4):619–27

Anderson JC, Voigt CA, Arkin AP (2007) Environmental signal integration by a modular AND gate. Mol Syst Biol 3:133

Angeli D, Ferrell JE, Jr., Sontag ED (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. Proc Natl Acad Sci USA 101(7): 1822–7

Atkinson MR, Savageau MA, Myers JT et al. (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. Cell 113(5):597–607

Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451(7174):86–9

Backes H, Berens C, Helbl V et al. (1997) Combinations of the alpha-helix-turn-alpha-helix motif of TetR with respective residues from LacI or 434Cro: DNA recognition, inducer binding, and urea-dependent denaturation. Biochemistry 36(18):5311–22

Balagadde FK, Song H, Ozaki J et al. (2008) A synthetic *Escherichia coli* predator-prey ecosystem. Mol Syst Biol 4:187

Balagadde FK, You L, Hansen CL et al. (2005) Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. Science 309(5731):137–40

Bar-Even A, Paulsson J, Maheshri N et al. (2006) Noise in protein expression scales with natural protein abundance. Nat Genet 38(6):636–43

Bashor CJ, Helman NC, Yan S et al. (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science 319(5869):1539–43

Bastin DA, Romana LK, Reeves PR (1991) Molecular cloning and expression in *Escherichia coli* K-12 of the rfb gene cluster determining the O antigen of an *E. coli* O111 strain. Mol Microbiol 5(9):2223–31

Basu S, Gerchman Y, Collins CH et al. (2005) A synthetic multicellular system for programmed pattern formation. Nature 434(7037):1130–4

Basu S, Mehreja R, Thiberge S et al. (2004) Spatiotemporal control of gene expression with pulse-generating networks. Proc Natl Acad Sci U S A 101(17):6355–60

Batchelor E, Goulian M (2006) Imaging OmpR localization in *Escherichia coli*. Mol Microbiol 59(6):1767–78

Baumgartner JW, Kim C, Brissette RE et al. (1994) Transmembrane signalling by a hybrid protein: communication from the domain of chemoreceptor Trg that recognizes sugar-binding proteins to the kinase/phosphatase domain of osmosensor EnvZ. J Bacteriol 176(4):1157–63

Bayer TS, Smolke CD (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression. Nat Biotechnol 23(3):337–43

Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405(6786):590–3

Bintu L, Buchler NE, Garcia HG et al. (2005a) Transcriptional regulation by the numbers: applications. Current Opinion in Genetics & Development 15(2):125–135

Bintu L, Buchler NE, Garcia HG et al. (2005b) Transcriptional regulation by the numbers: models. Current Opinion in Genetics & Development 15(2):116–124

Brenner K, Karig DK, Weiss R et al. (2007) Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. Proc Natl Acad Sci USA 104(44):17300–4

Brosius J, Erfle M, Storella J (1985) Spacing of the −10 and −35 regions in the tac promoter. Effect on its in vivo activity. J Biol Chem 260(6):3539–41

Buskirk AR, Landrigan A, Liu DR (2004) Engineering a ligand-dependent RNA transcriptional activator. Chem Biol 11(8):1157–63

Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. Nat Biotechnol 26(7):787–93

Cello J, Paul AV, Wimmer E (2002) Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science 297(5583):1016–8

Chen Q, Kadner RJ (2000) Effect of altered spacing between uhpT promoter elements on transcription activation. J Bacteriol 182(16):4430–6

Collins CH, Arnold FH, Leadbetter JR (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. Mol Microbiol 55(3): 712–23

Collins CH, Leadbetter JR, Arnold FH (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. Nat Biotechnol 24(6): 708–12

Cox RS, 3rd, Surette MG, Elowitz MB (2007) Programming gene expression with combinatorial promoters. Mol Syst Biol 3:145

de Boer PA, Crossley RE, Rothfield LI (1983) Proc Natl Acad Sci USA 80:21–25

de la Torre JC, Ortin J, Domingo E et al. (1984) Plasmid vectors based on Tn10 DNA: gene expression regulated by tetracycline. Plasmid 12(2):103–10

Derr P, Boder E, Goulian M (2006) Changing the specificity of a bacterial chemoreceptor. J Mol Biol 355(5):923–32

Dixon R, Cannon F, Kondorosi A (1976) Construction of a P plasmid carrying nitrogen fixation genes from *Klebsiella pneumoniae*. Nature 260(5548):268–71

Drubin DA, Way JC, Silver PA (2007) Designing biological systems. Genes Dev 21(3):242–54

Dwyer MA, Looger LL, Hellinga HW (2003) Computational design of a Zn2+ receptor that controls bacterial gene expression. Proc Natl Acad Sci USA 100(20):11255–60

El-Samad H, Khammash M (2006) Regulated degradation is a mechanism for suppressing stochastic fluctuations in gene regulatory networks. Biophys J 90(10):3749–3761

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403(6767):335–8

Endy D (2005) Foundations for engineering biology. Nature 438(7067):449–53

Endy D (2008) Genomics. Reconstruction of the genomes. Science 319(5867):1196–7

Falcon CM, Matthews KS (2000) Operator DNA sequence variation enhances high affinity binding by hinge helix mutants of lactose repressor protein. Biochemistry 39(36):11074–83

Ferrell JE, Jr. (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Curr Opin Cell Biol 14(2):140–8

Ferrell JE, Jr., Machleder EM (1998) The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. Science 280(5365):895–8

Francetic O, Belin D, Badaut C et al. (2000) Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. Embo J 19(24):6697–703

Frank DE, Saecker RM, Bond JP et al. (1997) Thermodynamics of the interactions of lac repressor with variants of the symmetric lac operator: effects of converting a consensus site to a non-specific site. J Mol Biol 267(5):1186–206

Gambetta GA, Lagarias JC (2001) Genetic engineering of phytochrome biosynthesis in bacteria. Proc Natl Acad Sci USA 98(19):10566–71

Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. Nature 403(6767):339–42

Gibson DG, Benders GA, Andrews-Pfannkoch C et al. (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science 319(5867):1215–20

Greber D, Fussenegger M (2007) Mammalian synthetic biology: engineering of sophisticated gene networks. J Biotechnol 130(4):329–45

Guet CC, Elowitz MB, Hsing W et al. (2002) Combinatorial synthesis of genetic networks. Science 296(5572):1466–70

Guzman LM, Belin D, Carson MJ et al. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177(14):4121–30

Ham TS, Lee SK, Keasling JD et al. (2006) A tightly regulated inducible expression system utilizing the fim inversion recombination switch. Biotechnol Bioeng 94(1):1–4

Hasty J, McMillen D, Collins JJ (2002) Engineered gene circuits. Nature 420(6912):224–30

Hawkins AC, Arnold FH, Stuermer R et al. (2007) Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. Appl Environ Microbiol 73(18):5775–81

Helbl V, Hillen W (1998) Stepwise selection of TetR variants recognizing tet operator 4C with high affinity and specificity. J Mol Biol 276(2):313–8

Helbl V, Tiebel B, Hillen W (1998) Stepwise selection of TetR variants recognizing tet operator 6C with high affinity and specificity. J Mol Biol 276(2):319–24

Henssler EM, Scholz O, Lochner S et al. (2004) Structure-based design of Tet repressor to optimize a new inducer specificity. Biochemistry 43(29):9512–8
Hoch J, Silhavy T (1995) Two Component Signal Transduction. Washington, DC. ASM Press
Hooshangi S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc Natl Acad Sci USA 102(10):3581–6
Isaacs FJ, Dwyer DJ, Collins JJ (2006) RNA synthetic biology. Nat Biotechnol 24(5):545–54
Isaacs FJ, Dwyer DJ, Ding C et al. (2004) Engineered riboregulators enable post-transcriptional control of gene expression. Nat Biotechnol 22(7):841–7
Isaacs FJ, Hasty J, Cantor CR et al. (2003) Prediction and measurement of an autoregulatory genetic module. Proc Natl Acad Sci USA 100(13):7714–9
Jarboe LR, Grabar TB, Yomano LP et al. (2007) Development of ethanologenic bacteria. Adv Biochem Eng Biotechnol 108:237–61
Jose AM, Soukup GA, Breaker RR (2001) Cooperative binding of effectors by an allosteric ribozyme. Nucleic Acids Res 29(7):1631–7
Kaern M, Blake WJ, Collins JJ (2003) The engineering of gene regulatory networks. Annu Rev Biomed Eng 5:179–206
Kaern M, Elston TC, Blake WJ et al. (2005) Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 6(6):451–64
Kalir S, Alon U (2004) Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. Cell 117(6):713–20
Kalir S, Mangan S, Alon U (2005) A coherent feed-forward loop with a SUM input function prolongs flagella expression in *Escherichia coli*. Mol Syst Biol 1:2005 0006
Kalir S, McClure J, Pabbaraju K et al. (2001) Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Science 292(5524):2080–3
Kamionka A, Sehnal M, Scholz O et al. (2004) Independent regulation of two genes in *Escherichia coli* by tetracyclines and Tet repressor variants. J Bacteriol 186(13):4399–401
Karig D, Weiss R (2005) Signal-amplifying genetic circuit enables in vivo observation of weak promoter activation in the Rhl quorum sensing system. Biotechnol Bioeng 89(6):709–18
Keasling JD (2008) Synthetic biology for synthetic chemistry. ACS Chem Biol 3(1):64–76
Khlebnikov A, Datsenko KA, Skaug T et al. (2001) Homogeneous expression of the P(BAD) promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. Microbiology 147(Pt 12):3241–7
Khlebnikov A, Keasling JD (2002) Effect of lacY expression on homogeneity of induction from the P(tac) and P(trc) promoters by natural and synthetic inducers. Biotechnol Prog 18(3):672–4
Khlebnikov A, Risa O, Skaug T et al. (2000) Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. J Bacteriol 182(24):7029–34
Kleckner N, Barker DF, Ross DG et al. (1978) Properties of the translocatable tetracycline-resistance element Tn10 in *Escherichia coli* and bacteriophage lambda. Genetics 90(3):427–61
Kobayashi H, Kaern M, Araki M et al. (2004) Programmable cells: interfacing natural and engineered gene networks. Proc Natl Acad Sci USA 101(22):8414–9
Koizumi M, Soukup GA, Kerr JN et al. (1999) Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. Nat Struct Biol 6(11):1062–71
Laub MT, Biondi EG, Skerker JM (2007) Phosphotransfer profiling: systematic mapping of two-component signal transduction pathways and phosphorelays. Methods Enzymol 423:531–48
Lee SY, Papoutsakis ET (1999) Metabolic Engineering. Marcel Dekker, New York
Levskaya A, Chevalier AA, Tabor JJ et al. (2005) Synthetic biology: engineering *Escherichia coli* to see light. Nature 438(7067):441–2
Li MZ, Elledge SJ (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 4(3):251–6
Li N, Cannon MC (1998) Gas vesicle genes identified in *Bacillus megaterium* and functional expression in *Escherichia coli*. J Bacteriol 180(9):2450–8
Looger LL, Dwyer MA, Smith JJ et al. (2003) Computational design of receptor and sensor proteins with novel functions. Nature 423(6936):185–90

Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res 25(6):1203–10

Lutz R, Lozinski T, Ellinger T et al. (2001) Dissecting the functional program of *Escherichia coli* promoters: the combined mode of action of Lac repressor and AraC activator. Nucleic Acids Res 29(18):3873–81

Lynch SA, Desai SK, Sajja HK et al. (2007) A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function. Chem Biol 14(2):173–84

Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci USA 100(21):11980–5

Mangan S, Itzkovitz S, Zaslaver A et al. (2006) The incoherent feed-forward loop accelerates the response-time of the gal system of *Escherichia coli*. J Mol Biol 356(5):1073–81

Mangan S, Zaslaver A, Alon U (2003) The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J Mol Biol 334(2):197–204

Mathews DH, Sabina J, Zuker M et al. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288(5):911–40

McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA 94(3):814–9

Michalowski CB, Short MD, Little JW (2004) Sequence tolerance of the phage lambda PRM promoter: implications for evolution of gene regulatory circuitry. J Bacteriol 186(23):7988–99

Mielenz JR (2001) Ethanol production from biomass: technology and commercialization status. Curr Opin Microbiol 4(3):324–9

Miller MB, Bassler BL (2001) Quorum sensing in bacteria. Annu Rev Microbiol 55:165–99

Milo R, Shen-Orr S, Itzkovitz S et al. (2002) Network motifs: simple building blocks of complex networks. Science 298(5594):824–7

Morgan-Kiss RM, Wadler C, Cronan JE, Jr (2002) Long-term and homogeneous regulation of the *Escherichia coli* araBAD promoter by use of a lactose transporter of relaxed specificity. Proc Natl Acad Sci USA 99(11):7373–7

Ninfa AJ, Mayo AE (2004) Hysteresis vs. graded responses: the connections make all the difference. Sci STKE 2004(232):pe20

Pedraza JM, van Oudenaarden A (2005) Noise propagation in gene networks. Science 307(5717):1965–9

Pfeifer BA, Admiraal SJ, Gramajo H et al. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E-coli*. Science 291(5509):1790–1792

Posfai G, Plunkett G, 3rd, Feher T et al. (2006) Emergent properties of reduced-genome *Escherichia coli*. Science 312(5776):1044–6

Ptashne M, Gann A (2002) Genes & Signals. Cold Spring Harbor Laboratory Press, New York, 2002

Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. J Mol Biol 323(5):785–93

Salis H, Tamsir A, Voigt CA (2009) Engineering bacterial sensors and signals. Bacterial sensing and Signaling (in press, 2009)

Savageau MA (1974) Comparison of classical and autogenous systems of regulation in inducible operons. Nature 252(5484):546–9

Service RF (2007) Cellulosic ethanol. Biofuel researchers prepare to reap a new harvest. Science 315(5818):1488–91

Shahrezaei V, Ollivier JF, Swain PS (2008) Colored extrinsic fluctuations and stochastic gene expression. Mol Syst Biol 4:196

Shen-Orr SS, Milo R, Mangan S et al. (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat Genet 31(1):64–8

Shetty RP, Endy D, Knight TF, Jr (2008) Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2(1):5

Sia SK, Gillette BM, Yang GJ (2007) Synthetic tissue biology: tissue engineering meets synthetic biology. Birth Defects Res C Embryo Today 81(4):354–61

Skerker JM, Perchuk BS, Siryaporn A et al. (2008) Rewiring the specificity of two-component signal transduction systems. Cell 133(6):1043–54
Skerker JM, Prasol MS, Perchuk BS et al. (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLoS Biol 3(10):e334
Smith HO, Hutchison CA, 3rd, Pfannkoch C et al. (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. Proc Natl Acad Sci USA 100(26):15440–5
Smith TL, Sauer RT (1995) P22 Arc repressor: role of cooperativity in repression and binding to operators with altered half-site spacing. J Mol Biol 249(4):729–42
Soukup GA, Breaker RR (1999a) Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization. Structure 7(7):783–91
Soukup GA, Breaker RR (1999b) Engineering precision RNA molecular switches. Proc Natl Acad Sci USA 96(7):3584–9
Soukup GA, Breaker RR (1999c) Relationship between internucleotide linkage geometry and the stability of RNA. Rna 5(10):1308–25
Soukup GA, DeRose EC, Koizumi M et al. (2001) Generating new ligand-binding RNAs by affinity maturation and disintegration of allosteric ribozymes. Rna 7(4):524–36
Swartz JR (2001) Advances in *Escherichia coli* production of therapeutic proteins. Curr Opin Biotechnol 12(2):195–201
Takeda Y, Sarai A, Rivera VM (1989) Analysis of the sequence-specific interactions between Cro repressor and operator DNA by systematic base substitution experiments. Proc Natl Acad Sci USA 86(2):439–43
Tang J, Breaker RR (1997) Rational design of allosteric ribozymes. Chem Biol 4(6):453–9
Temme K, Salis H, Tullman-Ercek D et al. (2008) Induction and relaxation dynamics of the regulatory network controlling the type III secretion system encoded within *Salmonella* pathogenicity island 1. J Mol Biol 377(1):47–61
Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci USA 98(15):8614–9
Topp S, Gallivan JP (2008) Random walks to synthetic riboswitches–a high-throughput selection based on cell motility. Chembiochem 9(2):210–3
Tumpey TM, Basler CF, Aguilar PV et al. (2005) Characterization of the reconstructed 1918 Spanish influenza pandemic virus. Science 310(5745):77–80
Ulrich LE, Koonin EV, Zhulin IB (2005) One-component systems dominate signal transduction in prokaryotes. Trends Microbiol 13(2):52–6
Utsumi R, Brissette RE, Rampersaud A et al. (1989) Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate. Science 245(4923):1246–9
Vilar JM, Leibler S (2003) DNA looping and physical constraints on transcription regulation. J Mol Biol 331(5):981–9
Voigt CA (2006) Genetic parts to program bacteria. Curr Opin Biotechnol 17(5):548–57
Wagner R (2000) Transcription regulation in prokaryotes. Oxford University Press, Oxford, New York
Wall ME, Hlavacek WS, Savageau MA (2004) Design of gene circuits: lessons from bacteria. Nat Rev Genet 5(1):34–42
Ward SM, Delgado A, Gunsalus RP et al. (2002) A NarX-Tar chimera mediates repellent chemotaxis to nitrate and nitrite. Mol Microbiol 44(3):709–19
Weiss R (2001) Cellular Computation and Communications Using Engineered Genetic Regulatory Networks, Massachussetts Institute of Technology
Weiss R, Homsy GE, Knight TF, Jr (1999) Towards in vivo digital circuits. DIMACS Workshop on Evolution as Computation 1:1–18
Williams SB, Stewart V (1997) Discrimination between structurally related ligands nitrate and nitrite controls autokinase activity of the NarX transmembrane signal transducer of *Escherichia coli* K-12. Mol Microbiol 26(5):911–25

Wilson JW, Coleman C, Nickerson CA (2007) Cloning and transfer of the *Salmonella* pathogenicity island 2 type III secretion system for studies of a range of gram-negative genera. Appl Environ Microbiol 73(18):5911–8

Winkler WC, Breaker RR (2003) Genetic control by metabolite-binding riboswitches. Chembiochem 4(10):1024–32

Yen L, Svendsen J, Lee JS et al. (2004) Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. Nature 431(7007):471–6

Yokobayashi Y, Weiss R, Arnold FH (2002) Directed evolution of a genetic circuit. Proc Natl Acad Sci USA 99(26):16587–91

You L, Cox RS, 3rd, Weiss R et al. (2004) Programmed population control by cell-cell communication and regulated killing. Nature 428(6985):868–71

Zhang W, Ames BD, Tsai SC et al. (2006) Engineered biosynthesis of a novel amidated polyketide, using the malonamyl-specific initiation module from the oxytetracycline polyketide synthase. Appl Environ Microbiol 72(4):2573–80