15

SYNTHETIC GENE NETWORKS

David Greber and Martin Fussenegger

Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, CH-4058 Basel, Switzerland

15.1 INTRODUCTION

Advances in molecular manipulation techniques, together with an ever-increasing accumulation of genetic information, are progressively opening new possibilities for gene therapy and biomedical engineering. By combining naturally occurring genetic components in unique ways, it has become possible to artificially engineer genetic networks that possess increasingly sophisticated functional capabilities. By analogy to electronic circuit engineering, the desired characteristics of such networks can be rationally designed and tested through predictive modeling. Similarly to electrical networks, genetic networks also possess "input" and "output" functionality such that they are capable of monitoring and responding in highly defined mechanisms. The creation of synthetic networks from well-defined modular components has enabled researchers to investigate and test many network characteristics found in natural genetic networks. It is from an applied perspective, however, that synthetic genetic networks represent a truly exciting innovation. It is not difficult to envisage applications where synthetic networks could be used to manipulate cellular behavior in a highly orchestrated way. While these concepts are still in their infancy, significant progress has been made in the creation of first-generation synthetic networks, which will one day enable the engineered control of cellular function to become a viable reality.

Systems Biology and Synthetic Biology Edited by Pengcheng Fu and Sven Panke

This chapter begins by describing the modular genetic components that form the building blocks of engineered genetic networks. It then describes the development of both simple and complex networks, many of which were initially developed in prokaryotic systems, but which have been subsequently extended to eukaryotic systems. The focus is upon describing networks that have been experimentally tested and validated. It does not cover the extensive modeling and computational work that has been conducted on either synthetic or natural genetic regulatory networks (readers are referred to Chapter 7). Advances in network functionality have been made on both the input and output dimensions. Examples of output functionality include the generation of stable behavior, such as bistable toggle and hysteric switches, and dynamic behavior such as an oscillatory network. From an input perspective developments include the creation of logical information "gates," where a range of input combinations produce highly defined outputs in a manner directly analogous to electrical circuits; the development of transcriptional cascades, which have enabled the range of inputs to a network to be greatly increased; and the development of novel sensory networks which, for example, can detect inputs within a defined concentration range, or respond precisely to a rising level of an input. The chapter concludes by presenting the initial first steps into the emerging field of semisynthetic networks. These are prosthetic genetic networks that are capable of responding to physiological cues so that they are effectively integrated into the host-cell's biology. Such networks, in response to acute or pathological cues, hold great promise for the controlled manipulation of cellular processes such as protein synthesis, metabolism, cell growth, and differentiation.

15.2 NETWORK BUILDING BLOCKS

While synthetic in the sense that they are artificially designed and created, synthetic genetic networks are actually engineered from naturally occurring genetic components. A discussion of these networks requires a basic understanding of these components and the manner in which they interact. While gene expression can be regulated and artificially manipulated at a number of levels, the networks described below have only utilized a limited number of transcriptional control elements. Hence, this overview is limited to the mechanisms and components that have been used in these systems. A comprehensive overview of other gene control systems and their application can be found in several recent reviews [1–3].

Transcriptional control operates at the level of mRNA synthesis through the use of inducible transcriptional activators and repressors that are capable of binding naturally occurring or specifically engineered promoters. The majority of systems utilize bacterial response regulators or activators that, upon binding to a target promoter, inhibit or activate transcription respectively. Binding of a specific molecule to the response regulator induces an allosteric change leading to disassociation of the regulator from its cognate promoter.

Prokaryotic gene control systems generally use inducible repressors and activators drawn from well-documented genetic operons such as the *lac* operon of *Escherichia*

coli [4], the tetracycline-resistance transposon Tn*10* [5], or the λcI repressor of bacteriophage lambda [6]. In each case, the respective response regulator binds to a DNA sequence, typically a short tandem repeat referred to as the "operator," located within or adjacent to a promoter where it either enhances transcription or sterically hinders the initiation of transcription. By substituting operators across different strength promoters it has been possible to generate inducible systems with varied induction characteristics [7].

Bacterial response regulators also form the basis of synthetic eukaryotic gene regulation systems although given transcriptional differences they require adaptation. This has been successfully achieved for many bacterial response regulators by placing the operator for the response regulator adjacent to an eukaryotic compatible promoter [8]. The response regulator thus acts as a heterologous DNA-binding protein (DBP) whose association with the desired promoter can be controlled through addition of an appropriate inducer. If the operator is placed close to an strong constitutive promoter (e.g., P_{CMV}, cytomegalovirus immediate early promoter), DBP binding can sterically prevent the initiation of transcription by RNA polymerase II machinery. Alternatively, transcription can be actively repressed by fusing a eukaryotic transcriptional silencer, such as the Kruppel-associated box protein (KRAB), to the DBP [9]. Such systems are referred to as ON-type systems, as the addition of an inducer leads to derepression of transcription (Fig. 15-1). In an OFF-type configuration, in which addition of inducer leads to transcriptional silencing, a transcriptional activation domain, such as the *Herpes simplex* virus VP16, is fused to the DBP [10]. By placing the corresponding operator site adjacent to a minimal promoter (e.g., $P_{hCMVmin}$, minimal version of the human cytomegalovirus immediate early promoter), DBP binding activates transcription from an otherwise silent minimal promoter. Addition of an inducer results in subsequent deactivation of transcription.

As many prokaryotic antibiotic response regulators have been well described, and given the low interference of many antibiotics with eukaryotic biology, they represent an ideal class of inducible DBPs for eukaryotic gene control. Using the aforementioned configurations, eukaryotic gene control systems responsive to tetracyclines [11], streptogramins [12], and macrolides [13] amongst others have been developed. As these gene control systems do not interfere with each other, they can be readily combined. For this reason, and their nonpleiotrophic effects, they have formed the basis of most eukaryotic synthetic gene networks. A list of the common transcriptional control elements used in the assembly of both prokaryotic and eukaryotic synthetic gene networks is provided in Table 15-1.

15.3 CHARACTERIZATION OF SIMPLE AND COMPLEX NETWORKS

The past decade has seen a progressive increase in the development and application of both prokaryotic and eukaryotic synthetic networks. In some cases, these networks have been relatively simple and have been used to test and investigate naturally occurring phenomena. In other cases, the networks exhibit far greater complexity as they seek to reproduce or create much more sophisticated functionality. When adopting

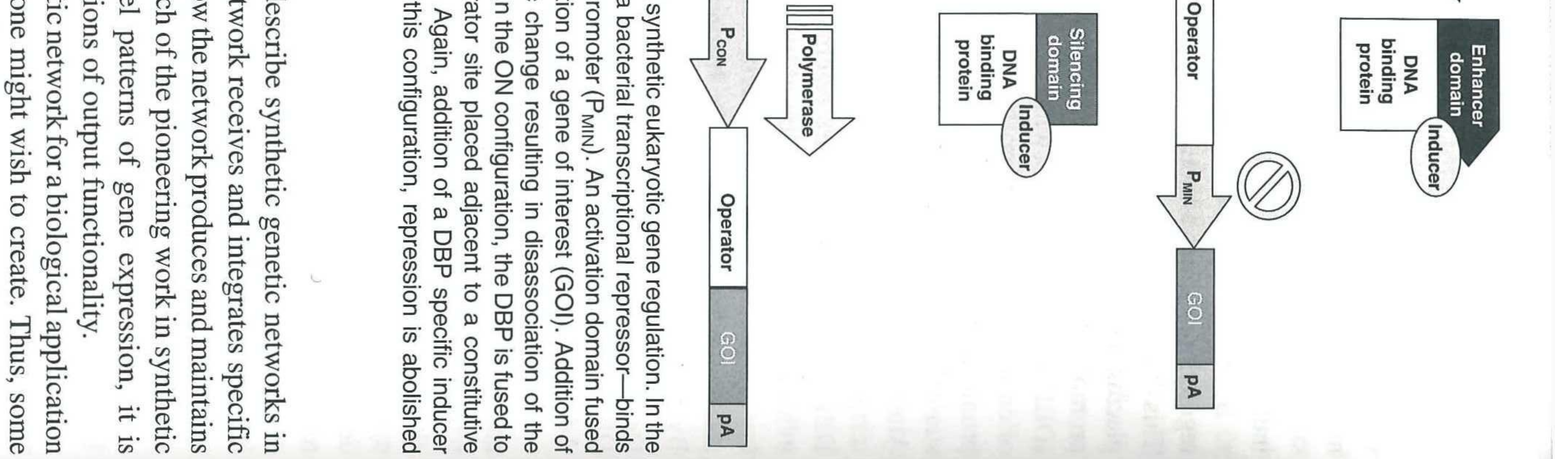

Figure 15-1 Molecular configuration of OFF and ON synthetic eukaryotic gene regulation. In the OFF configuration, a DNA-binding protein—typically a bacterial transcriptional repressor—binds a specific operator site placed adjacent to a minimal promoter (P_{MIN}). An activation domain fused to the DBP activates polymerase-mediated transcription of a gene of interest (GOI). Addition of an inducer specific to the DBP causes an allosteric change resulting in disassociation of the transactivator with subsequent transcriptional arrest. In the ON configuration, the DBP is fused to a repressor domain. Binding of DBP-TR to an operator site placed adjacent to a constitutive promoter (P_{CON}) represses transcription of the GOI. Again, addition of a DBP specific inducer results in transrepressor disassociation although in this configuration, repression is abolished resulting in expression of the GOI [8].

the electrical circuit analogy it is possible to describe synthetic genetic networks in terms of their input functionality—how the network receives and integrates specific signals as well as their output functionality—how the network produces and maintains a specific pattern of expression. Given that much of the pioneering work in synthetic circuits was directed toward producing novel patterns of gene expression, it is expedient to commence with network descriptions of output functionality.

In considering the design of a synthetic genetic network for a biological application it is useful to imagine what kind of functions one might wish to create. Thus, some applications may benefit from a mechanism that ensures a network produces a consistent and stable response even when there are considerable random fluctuations in either network components, inducer concentrations, or cellular components more broadly. For other applications, one may require a system that produces more than one

Table 15-1 Common genetic transcriptional components used in the reaction of synthetic genetic networks

DNA-Binding Protein	System Application	Engineered Regulatory Protein	Inducer	Response to Inducer	References
TetR	Prokaryotic	—	Doxycycline, aTc	Derepression	[5]
LacI	Prokaryotic	—	IPTG	Derepression	[4]
λcI	Prokaryotic	—	Temperature	Derepression	[6]
NRI	Prokaryotic	—	Phosphorylation	Activation	[94]
LuxR	Prokaryotic	—	Acyl-homoserine lactone	Activation	[95]
LacI	Eukaryotic	—	IPTG	Derepression	[96]
TetR	Eukaryotic	TetR-VP16 (tTA)	Tetracycline,	Deactivation	[11]
rTetR	Eukaryotic	rTetR-VP16	Doxycycline, aTc	Activation	[11]
Pip	Eukaryotic	Pip-KRAB	Streptogramins	Derepression	[12]
Pip	Eukaryotic	Pip-VP16	Streptogramins	Deactivation	[12]
E	Eukaryotic	E-KRAB	Macrolides	Derepression	[13]
E	Eukaryotic	E-VP16	Macrolides	Deactivation	[13]
ScbR	Eukaryotic	ScbR-VP16	Butyrolactones	Deactivation	[97]
Gal4	Eukaryotic	Gal4-VP16	Mifepristone	Deactivation	[56]
HIF-1α	Eukaryotic	–	Hypoxia	Activation	[98]

aTc, anhydrotetracycline; IPTG, isopropyl-β-D-thiogalactopyranoside; KRAB, Kruppel-associated box protein-derived transrepressor domain; VP16, *Herpes simplex* viral protein 16-derived transactivation domain.

discrete expression state. A mechanism that "remembers" what conditions the network has been exposed to may be useful in applications where only a transient pulse of an inducer is required or expected. A mechanism that not only remembers the past but also reacts differently to subsequent changes would also be desirable. Finally, a mechanism that produces continuous oscillations in expression readout may be highly practical where repeated temporal expression is required. All of these mechanisms have their counterpart in natural biological systems where they represent the molecular controls for numerous basic cellular functions ranging from cellular differentiation, cell-cycle control, and circadian rhythms. It is therefore not surprising that genetic engineers have applied considerable effort to synthetically reproduce these mechanisms. Apart from being useful tools, such synthetic networks also shed considerable light on how the equivalent mechanism occurs in a natural system.

15.3.7 Logic Gates

The expression output of many cell-based regulatory networks is often a logic response generated by one or more input signals. Due to their sigmoid-shaped dose–response curves, most gene control systems can be regarded as the genetic equivalent of an analog-to-digital converter. Their output is either ON or OFF across a wide range of inducer concentrations, except for a small concentration window where transitions between the two states occur. In this regard, the analogy between genetic networks and electronic circuitry is very compelling. This has led to the conceptualization of genetic networks as logic gates with switchboard-type truth-tables and schematic representations that directly mirror electronic circuit diagrams [60–62]. Adapting gene control systems to Boolean language, ON-type gene control systems represent IF type gates in the sense that expression results IF an input is present. Conversely, OFF-type gene control systems represent NOT type gates whereby expression results when an input is NOT present.

By utilizing several compatible heterologous gene control systems responsive to tetracycline, macrolide, streptogramin, and butyrolactone input signals, it has been

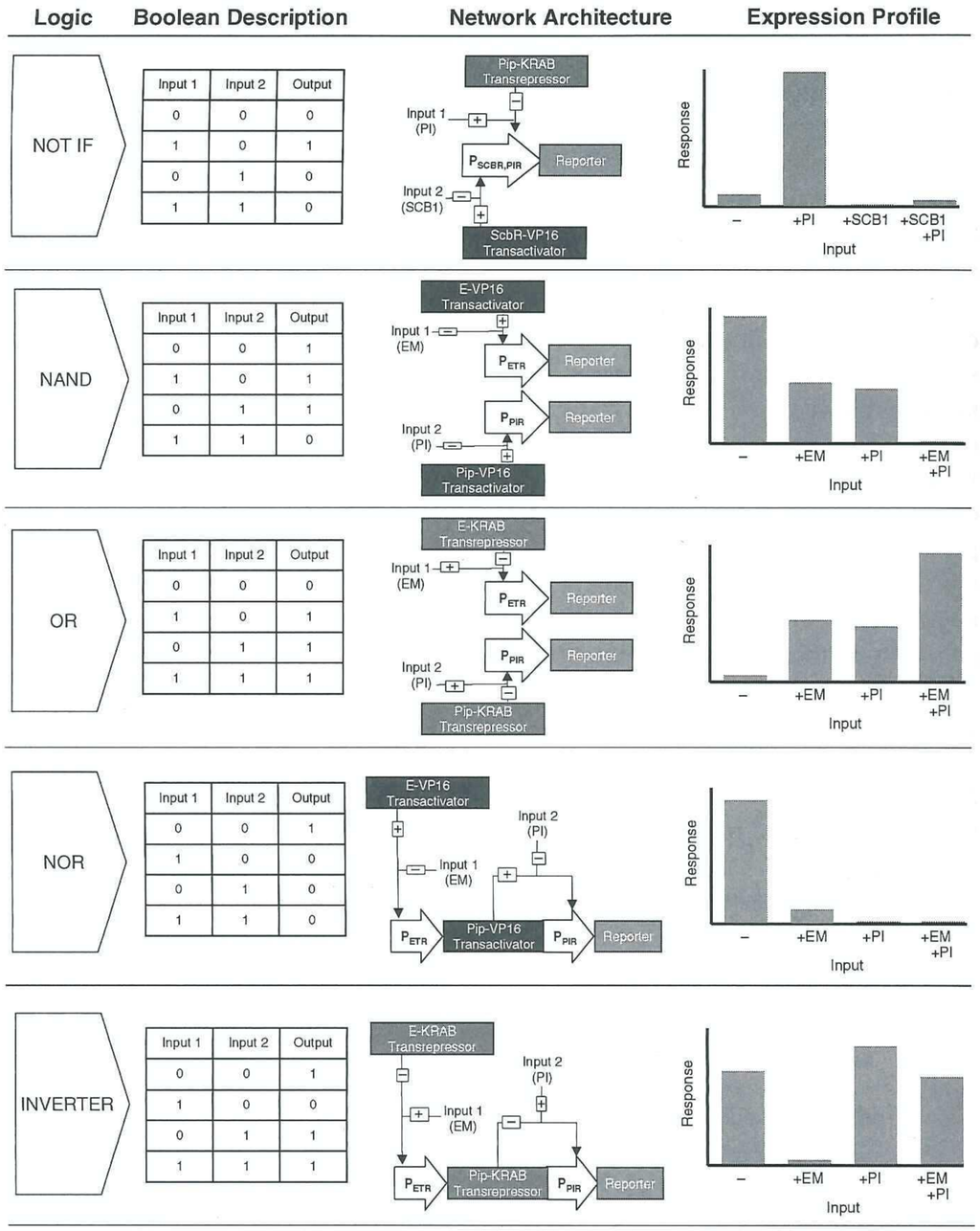

Figure 15-12 Boolean description, network architecture, and expression profile of five mammalian BioLogic Gates. All five mammalian logic gates were constructed from heterologous mammalian transcription systems. In the NOT IF gate, the butyrolactone-responsive transactivator (ScbR-VP16) and the streptogramin-responsive transrepressor (Pip-KRAB) are constitutively expressed and modulate expression of a reporter gene from a chimeric promoter ($P_{SCBR,PIR}$) containing operator sites for both ScbR-VP16 and Pip-KRAB. Input signals, 2-(1'-hydroxy-6-methylheptyl)-3-(hydroxymethyl) butanolide (SCB1) and/or PI result in disassociation of ScbR-VP16 and Pip-KRAB respectively. Expression only occurs when ScbR-VP16 is bound to the chimeric promoter and Pip-KRAB is disassociated therefore requiring the absence of SCB1 and presence of PI. For the NAND gate, both the macrolide-responsive transactivator (E-VP16) and the streptogramin-responsive transactivator (Pip-VP16) are constitutively expressed. Each transactivator binds its cognate promoter (P_{ETR} and P_{PIR}, respectively) which drive separate expression of two copies of

possible to design a range of eukaryotic logic circuits that follow strict Boolean logic in their integration of two input signals (Fig. 15-12) [63]. Hence, in the NOT IF gate, expression of a reporter gene occurs if and only if one specific input is present and the other input is absent. In the NAND gate, expression always occurs unless both inputs are present. The converse, where expression always occurs unless both inputs are absent, is reflected in the OR gate. The inverse, where expression occurs only when both inputs are absent is reflected in the NOR gate. Finally, the INVERTER gate represents the opposite of the NOT IF gate whereby expression always occurs unless one specific input is present and the other input is absent. Analogously to electronic circuit design some of these networks were constructed by linking elements in parallel while others were constructed by combining elements in series through the use of simple transcriptional cascades. These examples demonstrate that a considerable range of logical switches responding in unique ways to the same two input signals can be constructed from modular transcriptional control components. It is imaginable that such networks could be highly useful for gene therapy applications that require a particular response to highly specific inputs, which could vary depending upon the application.

Similar to electronic circuit design, the above switches were based on rational design principles. However, a number of other approaches have also been used to produce electronic-type circuit behavior, which produce a defined output in response to two inputs. Guet et al. used a combinatorial method involving prokaryotic transcriptional control systems that were randomly combined to generate a library of networks with varying connectivity [64]. From this library it was possible to isolate and characterize a range of diverse computational functions that produced unique phenotypes. While such an approach may yield unexpected network architectures for

the same reporter gene. Input signals, EM and/or PI, modulate transactivator activity respectively. Expression occurs when either or both transactivators are bound to their cognate reporter. The presence of both EM and PI are required to disassociate both transactivators to prevent expression. The OR gate is identical in design to the NAND gate but uses the transrepressor versions (i.e., E-KRAB and Pip-KRAB) of the macrolide- and streptogramin-responsive transcription control systems. Again, EM and/or PI modulate transrepressor activity respectively. In this case, expression is blocked only when both transrepressors are operator bound which only occurs when both EM and PI are absent. The NOR gate involves a short linear cascade between a constitutively produced macrolide-responsive transactivator (E-VP16) which drives the expression, via its cognate promoter (P_{ETR}), of the streptogramin-responsive transactivator (Pip-VP16) which in turn drives expression, via its cognate promoter (P_{PIR}), of a reporter gene. Modulation of transactivator activity is achieved through EM and PI, respectively. In this configuration, expression only occurs when E-VP16 is bound to its cognate operator and Pip-VP16 is disassociated from its cognate promoter therefore requiring the absence of both EM and PI. The final gate, the INVERTER, is identical in design to the NOR gate but uses the transrepressor versions (i.e., E-KRAB and Pip-KRAB) of the macrolide and streptogramin responsive transcription control systems. Again, EM and/or PI modulate transrepressor activity respectively. The only conditions under which expression will not occur are when E-KRAB is promoter disassociated and Pip-KRAB is promoter associated which occurs in the presence of EM and absence of PI. For each gate, the input and output characteristics of the Boolean description are reflected in the expression profile of the synthetic system [63].

a given function, the approach is not particularly amenable to forward engineering approaches that seek to design circuits that exhibit specifically required functions. In a related approach, Yokobayashi et al. combined rational design with an evolutionary approach to design specific circuits in *E. coli* [65]. Rational design based upon existing knowledge of well-characterized components was initially used to design a network with a specific function. Given that the synthesized network exhibited sub-optimal behavior, due to unexpected interactions and poor matching of network components, a directed-evolutionary approach was then used to fine-tune (or "debug") the system to obtain the required function. This was achieved through sequential rounds of localized random mutagenesis and recombination followed by phenotype screening. Subsequent sequence analysis of successful networks revealed that many changes, or "solutions", were capable of producing the desired phenotype. This could be manifested in changes which altered either protein-DNA or protein–protein interactions, but which nonetheless enabled superior biochemical matching of genetic components.