INVITED PAPER

Cybergenetics: Theory and Applications of Genetic Control Systems

This expository article presents an introduction to the exciting field of genetic control systems (cybergenetics). It covers the basic theory, implementation, and applications of this nascent field.

By MUSTAFA H. KHAMMASH, *Fellow IEEE*

ABSTRACT | There is no design template more important than DNA. Within the sequences of this exquisite substance lie the design plans for each of us and for every living organism. Shaped over billions of years by the creative machinations of evolution, this design template encodes the most complex dynamical systems known to us. Yet, it is only in our lifetimes that we are able to directly edit this template and engineer our own designs. The story that I tell in this article is about our early attempts to design and commission our own control systems in living cells. Guided by what we have learned from controlling man-made systems, we are beginning to develop the theory and methodologies needed to build control systems at the molecular level, an endeavor that is as challenging as it is rewarding. If carried out responsibly, this new ability to reshape the DNA template can have a tremendous benefit for our health and well-being, and will drive major advances in basic science, industrial biotechnology, and medical therapy. In this article, I will take the readers of the proceedings on a journey through the new and promising world of rationally designed genetic control systems. Using a minimum of jargon, I will introduce them to the biological concepts needed to develop an understanding and appreciation of the main design concepts emerging in this nascent area of research. My goal is to convey my own sense of excitement about the possibilities, but it is also to impart a feeling of the opportunities that lay ahead for members of the IEEE to contribute with their own creative ideas to the shaping of this most versatile of design templates, the DNA.

KEYWORDS | Biological control; control theory; cybergenetics; genetic control systems; synthetic biology.

I. INTRODUCTION

Norbert Wiener's celebrated 1948 book [1], *Cybernetics*, presented a far-reaching vision—one in which the study of control and communication in the animal and the machine are combined, with feedback as a unifying core concept. He argued that biological processes can be viewed through the lens of the systems approach, and that they can be broken down into interconnections of black box subsystems with inputs and outputs, which can be explored using known concepts of information processing, communication, feedback, stability, and noise. In essence, Wiener brought the engineering approach to biology, adding a systems perspective that led to a deeper understanding of concepts of biological adaptation that had been previously articulated by physiologists. These concepts include Claude Bernarde's "La Fixité du milieu intérieur" and Walter Canon's "homeostasis."

In what follows, I will present a tutorial introduction of the theory and methods for genetic control systems–one that is colored by my personal perspective. The presentation will be introductory, and I will attempt to cover all the biological backgrounds needed to understand and appreciate the challenges and opportunities in this exciting area of research that lies at the interface between engineering and biology. Inevitably, I had to make some choices regarding the covered topics, and some topics could not be covered. In particular, cell-free genetic control systems were not covered all in this article. Nevertheless, I hope that this introduction will motivate the interested

Manuscript received October 5, 2021; revised March 16, 2022; accepted April 19, 2022. Date of current version May 19, 2022. This work was supported by the European Research Council (ERC) through the European Union's Horizon 2020 Research and Innovation Programme (CyberGenetics) under Grant Agreement 743269.

The author is with the Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology (ETHZ), 4058 Basel, Switzerland (e-mail: mustafa.khammash@bsse.ethz.ch).

Digital Object Identifier 10.1109/JPROC.2022.3170599

reader to explore this important and fascinating area of genetic control systems, a recent review of which can be found in [2]. Other reviews on the broad topic of control systems in synthetic biology can be found in [3], [4].

This article is organized as follows. In Section II, I motivate the main topic of this article by giving an example of a simple genetic control circuit and compare its function to that of an electronic amplifier. In section III, I present a basic background of the biology of living cells and genetic circuits. In Section IV, I give an introduction to genetic engineering methods in synthetic biology, which I hope will give the reader an understanding of how genetic circuits are designed and introduced to living cells. In Section V, I begin exploring genetic control systems, starting with circuits that use feedforward and negative feedback loops, and then moving on to circuits that employ more advanced control systems, such as integral (I), proportional–integral (PI), and proportional–integral–derivative (PID) controllers. In Section VI, I describe some of the most promising applications of genetic control systems, highlighting some of the opportunities in this new field. In Section VII, I provide some concluding remarks and offer an outlook for the future.

II. CONTROL SYSTEMS IN LIVING CELLS

Living systems employ regulatory strategies at all levels of organization, from the molecular level to the whole organism. Inside living cells, regulatory circuits abound, with feedforward and feedback circuits being some of the most common. As an introduction to cellular regulatory circuits, I will present a simple genetic feedback circuit that exemplifies a common theme in living cells: gene autoregulation. Consider the gene circuit schematic in Fig. 1(a). A single gene (a stretch of DNA that carries the instructions to synthesize a protein) is repeatedly copied into messenger RNA (mRNA) molecules. In turn, the mRNA molecules, whose concentration is denoted by $r(t)$, are used as templates to synthesize the target protein at a rate $kr(t)$. Both the mRNA and the protein degrade at rates $\gamma_r m(t)$ and $\gamma_p y(t)$, respectively. Viewing y as the system output, when the mRNA production rate is constant, say u, the output protein will be unregulated, which corresponds to an open-loop circuit. However, when the mRNA production rate is a function of the protein concentration, then the output will be regulated, resulting in an autoregulatory gene circuit. Here, we shall take the mRNA production rate to be $u/(1 + y(t)^n)$, which corresponds to *negative* feedback. The dynamics of the feedback-regulated system can be easily written

$$\dot{r} = u/\left(1 + y^n\right) - \gamma_1 r$$
$$\dot{y} = kr - \gamma_2\, y.$$

At steady state, the output y is given by the solution of the algebraic equation $y = h(y)$, where $h(y) = Au/(1 + y^n)$ with $A := k/(\gamma_1\gamma_2)$. In contrast, in the unregulated case, the steady-state value of the output, y, is given by $y = Au$ [see Fig. 1(a)].

The added robustness through negative feedback regulation is one of several benefits of this type of feedback. Indeed, comparing the open-loop with the closed-loop circuit when A is perturbed by 50%, it is clear that perturbation will lead to far smaller changes in the output y in the negative feedback circuit than in the unregulated open-loop circuit (see Fig. 1).

Fundamentally, the function of the autoregulatory gene circuit is not unlike that of the negative feedback amplifier. To see this, consider the electronic negative feedback amplifier circuit in Fig. 1(b). As in the gene-regulatory circuit, the electronic circuit has a gain A in its forward path, which feeds a negative feedback path leading to a negative feedback closed-loop interconnection. For a constant input u, the output y is relatively insensitive to variations in the forward loop gain A, which is typically very large. In contrast, the same circuit becomes quite sensitive to variations in A in the open-loop case. This highlights one of the most fundamental benefits of negative feedback—robustness. The gain of the closed-loop input–output system is determined largely by the feedback loop parameters: β for the amplifier and n for the gene circuit. These parameters remain fairly constant during the operation of the circuits: β is realized with precision elements (e.g., resistors), and in the case of the gene circuit, n depends on biochemical properties that are not expected to change during the operation (see [5] for a more detailed treatment of these two circuits). Thus, the autoregulating genetic circuit is, in fact, not much different from the negative feedback amplifier in spite of the vast differences in their substrates and timescales.

While the autoregulatory gene circuit represents a control system that is present naturally in many cells, it is now possible to design novel gene regulatory circuits with desirable properties that control natural or newly introduced cellular species, such as enzymes. This is enabled by genetic engineering methods, which I will describe shortly. However, first, a brief introduction to cell biology is in order. Readers familiar with the basics of cell biology can safely skip this section.

III. BASICS OF GENETIC CIRCUITS

In this section, I will explain what genetic circuits are, how they function, and how synthetic ones are engineered into living cells. This will allow us to focus on genetic control circuit and their properties. I will start with a brief introduction to cell biology, which will provide an appreciation of the biological substrate that underlies genetic circuits.

A. Biology of the Living Cell: A Brief Introduction

The basic unit of life is the cell [6]. A living cell exists either as a separate entity, as does a single bacterium, or as part of a multicellular organism, where it tends to be highly

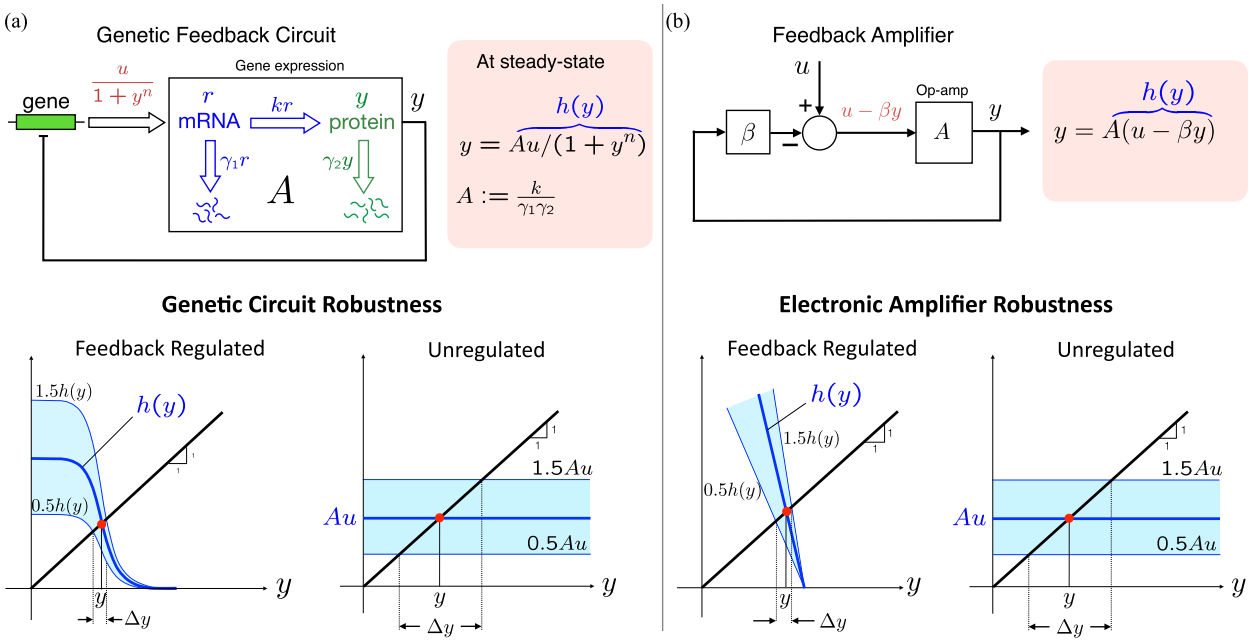

Fig. 1. *Autoregulatory gene circuit and the feedback amplifier. In spite of their vast differences in substrate and time scale, the autoregulatory gene circuit (a) and the electronic negative feedback amplifier (b) share much in common. In the self-regulating gene circuit, a single gene produces many mRNA molecules, each of which gets translated to many proteins, making gene expression effectively a high-gain "amplifier." Using the output y to repress the mRNA synthesis introduces negative feedback that results in a robust closed-loop system. Indeed, for the regulated gene circuit, ±50% changes in A lead to small variations in the output y compared to the unregulated circuit. In the negative feedback amplifier, high gain A is exploited to achieve robustness to parameter variations in a similar fashion.*

specialized, as is the case with a single neuron. In spite of the fact that cells can look drastically different, they share many common features that allow them to function and survive in diverse environments while achieving drastically different functional roles. For example, all cells use the same general strategies to generate energy, replicate themselves, and regulate substance traffic into and out of their membranes.

Cells are separated from their environments by membranes made from phospholipids [6], [7]. Aside from playing a structural role and as a natural barrier, the membrane also serves as a scaffold to hold proteins that have many different functions: some act as identity markers, while others act as connectors that bring cells together. Yet, other membrane proteins specialize as receptors that send and receive signals from the environment and from other cells—a role that is critically important for cellular decision-making.

Inside the cellular membrane, one finds organic molecules of different types [6], [7]. These include nucleic acids, proteins, carbohydrates, and lipids. Nucleic acids are the molecules that carry the cell's blueprint—its genetic code. DNA is one type of nucleic acid that encodes the information needed to build the cell; RNA, on the other hand, has multiple roles, one of the most important being its central involvement in the process of reading the genetic information encoded in the DNA and using it to synthesize cellular proteins (see Section III-B1 on gene expression). Proteins, organic substances consisting of chains of amino acids, are the workhorses of the cell. Proteins fulfill several structural, catalytic, and regulatory roles. For example, proteins called enzymes act as catalysts of chemical reactions, allowing the cell to rapidly produce molecules needed for cellular function. Other proteins act as dynamic regulatory agents that determine which proteins are synthesized, as well as the timing and rate of their synthesis.

Besides nucleic acids and proteins, other key cellular molecules are carbohydrates [6], [7]. While simple carbohydrates are broken down to produce energy for the cell, complex carbohydrates are used for cellular energy storage. Finally, lipids are also molecules that cells use for energy storage. They can alternatively play a structural role, as they make up the membrane of the cell and the membranes of organelles within the cell. Such organelles are compartments within the cell that performs highly specialized functions. An example of such an organelle is the mitochondrion, a compartment that maintains the necessary cellular machinery needed for producing ATP, the cell's main energy currency. Another important example of an organelle is the nucleus, where the genetic material is stored.

Biologists categorize cells by the way they package their DNA. If the DNA is stored in a nucleus, the cell is categorized as a eukaryotic cell. Examples of eukaryotic cells include mammalian cells. However, if the cell has no

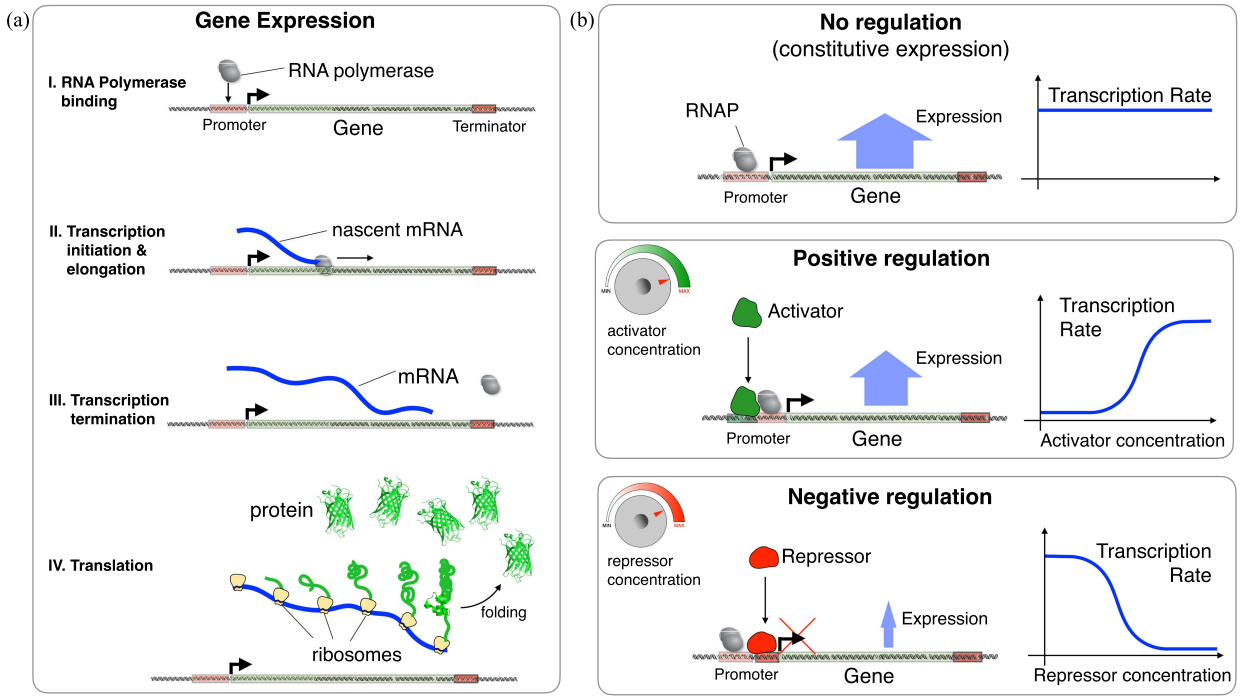

Fig. 2. *Gene expression and gene regulation. (a) Main steps of gene expression. Gene expression begins with the enzyme RNA polymerase (RNAP) binding a sequence upstream of the gene, called the promoter. RNAP then moves along the DNA, acting like a "reading head" that reads a ticker tape and transcribes its sequence into another linear sequence of nucleotides, called mRNA. This transcription process ends when RNAP reaches a specific terminator (STOP) sequence, directing it to be released from the DNA along with the fully transcribed mRNA molecule. In the next step, complex cellular machines, called ribosomes, act simultaneously to read the mRNA sequentially and translate the information within it to a sequence of amino acids (polypeptide). In this process, each sequence of 3 nucleotides is translated to one specific amino acid. When the mRNA translation process is complete, the polypeptide assumes a 3-D folding state and is called a protein. (b) Key ways gene expression is regulated. Some genes are unregulated and are always on, expressing proteins at a fixed rate. Many proteins are regulated, either positively or negatively, by proteins called transcription factors, which may act as activators or repressors. For positively regulated genes, activators bind to the DNA and recruit RNAP to express the gene. For negatively regulated genes, repressors bind to the DNA and prevent RNAP from initiating transcription, leading to low or no gene expression. Transcription factors often work cooperatively, leading to transcription rate functions that are sigmoidal in shape.*

nucleus in which to hold its DNA, it is categorized as a prokaryotic cell. Examples include bacteria and archaea. All prokaryotic cells and some eukaryotic cells are single-celled organisms, capable of living freely on their own. On the other hand, multicellular organisms, such as plants and animals, are entirely made up of eukaryotic cells.

When studying the molecular biology of living cells, biologists focus on a relatively small number of cell types and study those in great detail. A huge amount of information is available on these model cells, which considerably speeds up the pace of scientific understanding of the basic underlying biological processes, many of which are shared by all living cells. For example, *Escherichia coli* (*E. coli*) has been the most studied and best understood prokaryotic model organism, while *S. cerevisiae* (budding yeast), along with several mammalian cell lines, such as immortal cervical cancer cells (HeLa cells), human embryonic kidney cells (HEK293 cells), and Chinese hamster ovarian cells (CHO cells), are among the most studied eukaryotic model cells.

Bacterial, yeast, and mammalian cells differ drastically in size and in the number of different cellular constituents. To get an idea about these differences, I list some key comparative numbers that were compiled in [8]. An *E. coli* cell has a volume of 1 μm^3 and is 1 μm long, while a yeast cell is typically 30 times as large in volume and five times as long, and a mammalian cell is typically 3000 times as large in volume and 20 times as long. An *E. coli* cell has approximately 2×10^3 mRNA molecules and 3×10^6 protein molecules. In contrast, a budding yeast cell contains approximately 3×10^4 mRNA molecules and 10^8 protein molecules, and an HeLa cell has approximately 2×10^5 mRNA molecules and 10^{10} protein molecules. All three cell types are commonly used in biological studies, including synthetic biology.

B. Basic Unit of Genetic Circuits: The Gene

A gene is a stretch of DNA that encodes the synthesis of a gene product. Most genes code for cellular proteins, usually one gene per protein, but some genes code for RNA. The DNA itself consists of a linear sequence of nucleotides, each made up of a sugar molecule attached

to a phosphate group and a nitrogen-containing base. Importantly, only four different bases are found in the DNA of all living cells: adenine (A), guanine (G), cytosine (C), and thymine (T). Hence, a strand of DNA can be thought of as a quaternary linear sequence made of the four letters, A, G, C, and T, and a gene is a finite subsequence thereof. The bases G and C on two different DNA strands can bond forming a G-C complementary base pair; the bases A and T can similarly form an A-T complementary base pair. Consequently, a strand of DNA can bind tightly to a complementary DNA strand where the sequence of bases on one strand is matched perfectly with a sequence of complementary bases on the other strand. The result is a very stable double-stranded DNA, which assumes the well-known helical structure.

1) Gene Expression and Its Regulation: The process of synthesizing a protein from its encoded DNA sequence is called *gene expression*. Gene expression is a central process in gene–gene circuits [see Fig. 1(a)]. It consists of two key steps: 1) reading the DNA and using the information to synthesize mRNA, a process called *transcription* and 2) using the information in the transcribed mRNA to synthesize the protein encoded in the mRNA sequence, a process called *translation*. The central dogma of molecular biology summarizes the direction of information flow involved in gene expression: DNA $\longrightarrow$ mRNA $\longrightarrow$ protein. Each of these processes is dynamic in nature and involves multiple steps that require complex cellular molecules, such as RNAP and ribosomes [see Fig. 2(a)]. Consequently, the proper execution of gene expression depends on the abundant availability of these molecules in the cell and sufficient energy sources that power their function, e.g., ATP. This biological context must be factored in when designing synthetic gene circuits that use gene expression.

Some genes are continuously ON (constitutive expression), while others are regulated by proteins called transcription factors, which come in two varieties: activators and repressors. As their names suggest, activators turn on gene expression when their cellular concentration is sufficiently high, while repressors have the opposite effect [see Fig. 2(b)]. Transcription factors often function cooperatively so that the binding of one transcription factor to the DNA increases the binding affinity of another. Some genes are controlled by both activators and repressors, making their transcription rate a function of more than one transcription factor.

2) Genetic Circuits Are Gene Networks: The fact that gene expression can be regulated with transcription factors that are themselves gene expression products allows the formation of complex gene regulatory networks, in which genes interact with one another and with other substances in the cells to execute various regulation, information processing, and decision-making processes. Such networks are naturally found in living cells, but they can also be designed and engineered *de novo* into living cells. Often referred to as gene circuits, such networks can be designed to perform digital logic or analog computations. An example of a gene circuit that implements a simple negative feedback loop was shown in Fig. 1(a). A two-gene genetic feedback circuit is shown in Fig. 3(b). Here, the protein product of gene A is an activator that positively controls the expression of gene B. The product of gene B, in turn, is a repressor that inhibits the expression of gene A, thereby closing the negative feedback loop. The modeling, analysis, design, and implementation of such regulatory gene circuits are at the heart of cybergenetics and will be explored next.

IV. ENGINEERING GENETIC CIRCUITS IN LIVING CELLS

We now turn to the question of how new genetic circuits, such as the one shown in Figs. 1(a) and 3(b), can be designed and implemented inside a living host cell. The ability to introduce novel functioning biological circuits is made possible by a technology that began nearly 50 years ago. Genetic engineering, also known as gene cloning or recombinant DNA technology, is an enabling technology that arose in the 1970s from the field of microbial and molecular genetics [9], [10]. It involves manipulating DNA by selectively cutting it, modifying it, and joining its pieces together using specialized enzymes. It is akin to the process of cutting and pasting text, but the sequence that is being manipulated consists of the nucleotides that make up DNA. So successful has this technology been that, today, it is difficult to imagine doing basic biological science, biotechnology, or medicine without it.

Engineering genes into living cells requires a few basic steps: generation of the desired DNA fragment, joining the fragment to a carrier molecule, transferring the carrier into host cells, and then selecting those host cells that have taken up the desired DNA sequence (see Fig. 3 for details). In this way, a desired piece of DNA that codes for a gene of interest can be introduced into the cell, and then, the cell will express this gene using its native gene expression machinery. This gene may come from another organism or may be a completely novel gene that was designed on a computer. For example, one gene could code for a transcription factor that activates an endogenous gene. Alternatively, it could activate a second gene that was introduced into the cell with its promoter designed to be activated by the transcription factor from the first gene. By carefully designing genes and their promoter sequences, one can, thus, imagine that an entire gene network may be introduced into a living cell. This allows the design and engineering of genetic components, devices, and system modules that can be programmed to achieve desired functions once introduced into living host cells. This is what synthetic biology is all about, and its birth in the year 2000 was a natural outcome of the maturation of the technology behind genetic engineering. In that year, three synthetic genetic circuits published in the same journal ushered in the new field: a synthetic genetic oscillator [11], a genetic toggle switch [12], and genetic feedback circuit [13].

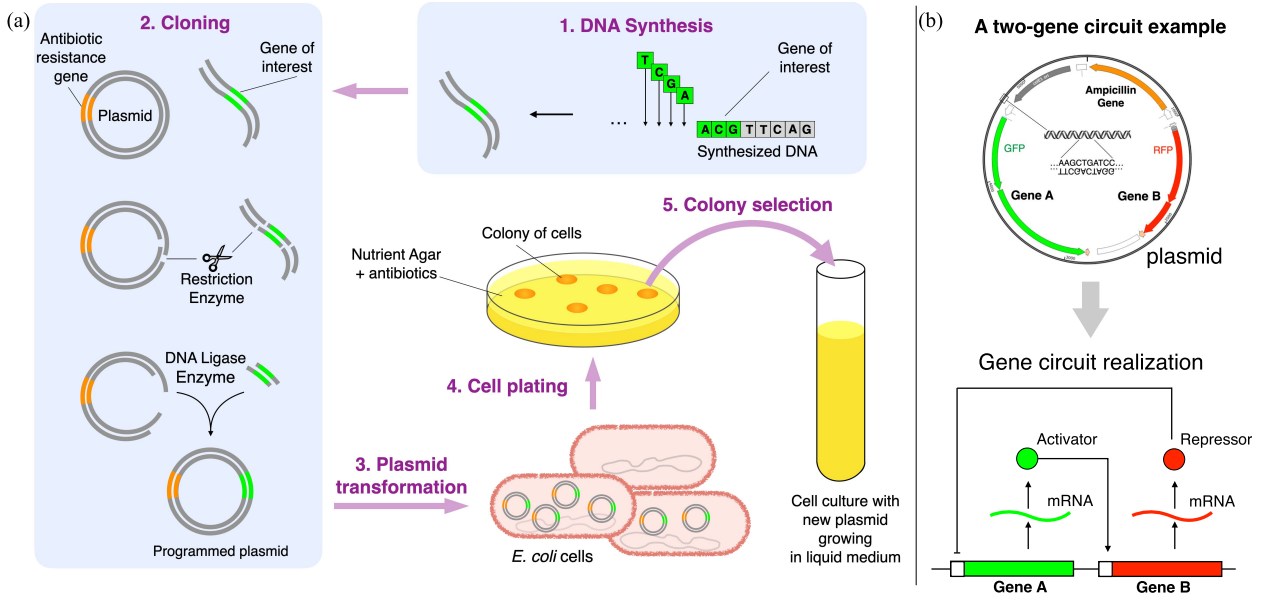

Fig. 3. *(a) Basic steps of genetic engineering. In step 1, a gene of interest consisting of a finite sequence of nucleotides synthesized one nucleotide at a time resulting in a double-stranded DNA linear sequence, called a gene fragment. The current cost for DNA synthesis is less than $0.1 per nucleotide pair. In step 2, the gene fragment is integrated into a circular piece of DNA, called a plasmid, which is the main delivery vehicle into the cell and contains an antibiotic resistance gene for plasmid selection. This process is referred to as cloning. It uses restriction enzymes to cut the plasmid and the gene fragment at sequence-specific locations that depend on the enzyme used. Then, enzymes called ligases "paste" the cut gene of interest into the cut plasmid. The result of step 2 is a plasmid containing the gene of interest. In step 3, the plasmid is introduced into living cells, e.g., E. coli. Different methods can be used for this step. One such method is called electroporation in which an electric field is applied to the cells making their membrane more porous and allowing plasmids to be taken up through these pores. At the end of this step, different cells will have taken up different numbers of plasmids, with some not taking up any. In step 4, the cells are placed on an agar plate containing nutrients and an antibiotic. The role of the antibiotic is to stop the growth of all the cells that did not take up the plasmid, as these cells do not have the antibiotic resistance gene present only on the plasmid. The colonies that grow on the plate contain cells that have the plasmids, and then, they are taken and grown in liquid media for circuit testing. (b) Example of a simple synthetic gene feedback network. The plasmid was engineered to contain three genes, one constitutively expressed gene coding for an antibiotic resistance protein, and two genes coding for a transcriptional activator and a transcriptional repressor. The promoter region in each of these two genes is designed such that it is the target of the transcription factor expressed by the other gene. In the figure, a line with an arrow pointing to the promoter indicates activation, while a line with a bar indicates repression. When this plasmid is transformed into a living E. coli cell, the cell's gene expression machinery expresses the two genes and realizes the negative feedback genetic circuit shown.*

I now provide a specific example showing how genetic engineering can be used to design a simple two-gene synthetic feedback control circuit in *E. coli*. Fig. 3(b) shows a piece of circular DNA (called a plasmid) that has been genetically engineered to realize the control system. The plasmid itself is made up of a circular piece of DNA, which carries three genes that have been synthesized and cloned into the plasmid. One gene codes for ampicillin-resistance (ampR), an antibiotic resistance that is expressed constitutively in the host cell, making it resistant to antibiotics. This allows for the selection of only those cells that have taken up the plasmid, as cells lacking it will be killed by the antibiotic that is included in the agar plate on which the cells are placed. The remaining two genes make up the feedback control circuit. Gene A codes for a transcription factor that is designed to bind the promoter of gene B and activate its expression. Gene B codes for a different transcription factor that acts as a repressor of gene A when bound to its promoter [see Fig. 3(b)]. Once the plasmid is introduced into the host cell, the arrangement just described realizes a negative feedback loop. It is often desired to modulate the transcription rate for one or both genes using an external input. One possible reason to do this is to tune the steady-state value of the output of interest. Such modulation can be achieved using small molecules that are added externally to the cells. Once they diffuse into the cell's cytoplasm, these small molecules interact with the transcription factor and modulate their function, thereby providing a tuning knob for the transcription rates of a given gene. One example of this is the repressor TetR, which can bind to the molecule anhydrotetracycline (aTc) [14]. When aTc binds a TetR molecule, the complex can no longer bind to the promoter. This reduces the overall repression of the gene, thus increasing its expression rate. The amount of increase in the expression rate depends on the external concentration of aTc, making it an ideal external input for modulating the output level of the circuit, thereby obviating the need to "hardwire" the output level of the genetic circuit.

The quantitative study of gene circuits requires models that capture their stochastic dynamic behavior. It turns

out that mathematical models of biochemical reaction networks (CRNs) provide a very suitable framework for modeling gene expression and other complex cellular processes. In Section V, I will provide a basic introduction to such models.

V. CHEMICAL REACTION NETWORKS: A LANGUAGE FOR MODELING CELLULAR SYSTEMS

Consider a system of chemical reactions with n species, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ that interact with each other through m chemical reactions

$$\begin{aligned} &\nu_{11}\mathbf{X}_1 + \cdots + \nu_{1n}\mathbf{X}_n \xrightarrow{c_1} \nu'_{11}\mathbf{X}_1 + \cdots + \nu'_{1n}\mathbf{X}_n \quad (R_1) \\ &\qquad \vdots \\ &\nu_{m1}\mathbf{X}_1 + \cdots + \nu_{mn}\mathbf{X}_n \xrightarrow{c_m} \nu'_{m1}\mathbf{X}_1 + \cdots + \nu'_{mn}\mathbf{X}_n \quad (R_m)\,. \end{aligned}$$

Note that, when all the products of a reaction are not important to model and do not themselves enter as reactants, they are often replaced with a generic empty set symbol, $\emptyset$. Similarly, when reactants in a given reaction have concentrations that do not change over time, they can be replaced with $\emptyset$. For example,

$$\mathbf{X}_1 + \mathbf{X}_2 \longrightarrow \emptyset \quad \text{and} \quad \emptyset \longrightarrow \mathbf{X}_1.$$

In the case when reactants are replaced by $\emptyset$, their constant concentrations will need to be reflected in the reaction rate.

Given the above reaction system, whenever the kth reaction fires, the number of molecules of the species $(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ changes by

$$s_k = \left(\nu'_{k1} - \nu_{k1}, \ldots, \nu'_{kn} - \nu_{kn}\right)^T.$$

s_k is referred to as the stoichiometry vector of the kth reaction, and all such vectors for the reaction system can be arranged in a matrix, called the stoichiometry matrix

$$S = \begin{bmatrix} s_1 & \ldots & s_m \end{bmatrix}.$$

The stoichiometry matrix summarizes the change in the chemical system whenever any of the reactions takes place. Together with the reaction rates, it determines the dynamics of the kinetic system, as we shall see shortly.

A. Deterministic Models of Chemical Reaction Networks

In deterministic models of CRNs, the state of the chemical system is described by the concentration of its constituents. The concentration evolves deterministically and continuously in time. Let $X(t) = \left(X_1(t), \ldots, X_n(t)\right)^T$ be the vector of concentrations of the species $\mathbf{X}_1, \ldots, \mathbf{X}_n$ at time t. Let $v(t) = \left(v_1(X(t)), \ldots, v_m(X(t))\right)^T$ be the vector of reaction rates at time t. The vector of species concentrations $X(t)$ evolves according to the differential equation

$$\dot{X}(t) = Sv\left(X(t)\right).$$

I will assume that the reactions follow mass-action kinetics [15], [16], which implies that their rate is given by

$$v_k(X) = c_k \prod_{i=1}^{n} X_i^{\nu_{ki}}$$

where c_k is the reaction rate constant. For the kth reaction, the sum $\sum_i \nu_{ki}$ is the order of the reaction. Therefore, for a zero-order reaction, the reaction rate is independent of the concentration of the reactants; for a first-order reaction, it depends only on the concentration of one of the reactants; and so on.

In mass-action formulations of chemical kinetics, one often considers elementary reactions consisting of at most two reactants, i.e., at most second-order. Reactions with more than two reactants can always be broken into multiple reactions consisting of two reactants each. Table 1 summarizes the reaction rates for these simple reactions.

Using these ideas, deterministic models of gene expression can be easily obtained (see Fig. 4).

Table 1 Reaction Rates for Zero-, First-, and Second-Order Reactions That Follow Mass-Action Kinetics

Reaction type	Reaction	Reaction rate
zero-order	$\emptyset \xrightarrow{c}$ Products	c
monomolecular	$\mathbf{X}_i \xrightarrow{c}$ Products	cX_i
bimolecular $(i \neq j)$	$\mathbf{X}_i + \mathbf{X}_j \xrightarrow{c}$ Products	cX_iX_j
bimolecular	$2\mathbf{X}_i \xrightarrow{c}$ Products	cX_i^2

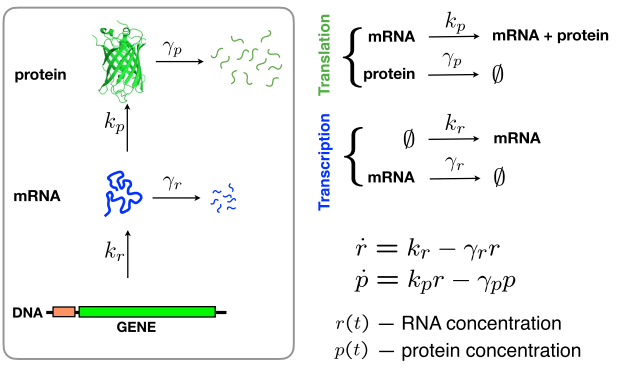

Fig. 4. *Simple model of gene expression. The basic steps of gene expression can be captured by four chemical reactions: one for transcription, one for translation, one for mRNA degradation, and one for protein degradation. These reactions can in turn be translated into dynamic equations that describe the time evolution of the mRNA and the protein species. For this simple gene expression model, the dynamics of the mRNA and protein concentrations are described by linear ODEs.*

In synthetic biology, it is often convenient to combine multiple mass-action reactions into a single reaction that is derived from the law of mass-action after making certain time-scale assumptions. One example is an enzyme catalyzed reaction [17]

$$\mathbf{S} + \mathbf{E} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathbf{C} \xrightarrow{k} \mathbf{E} + \mathbf{P}$$

where **S** is a substrate, **E** is the enzyme, **C** is the substrate-enzyme complex, and **P** is the product. These reactions can be approximated by a single reaction

$$\mathbf{S} \longrightarrow \mathbf{P}$$

with a rate equaling $v = (V_m S)/(K_m + S)$, where S is the concentration of **S**, $K_m = (k + k_{-1})/k_1$, and $V_m = kE_T$, with E_T being the total enzyme concentration.

Another important example is a transcription reaction that involves the cooperativity of the transcription factor, whereby the binding of one transcription factor molecule to the promoter increases the rate of binding of another. Alternatively, more than one transcription factor may bind together before the complex can bind the promoter to regulate gene expression. Such reactions are very common for both transcriptional activators and repressors. They can be expressed using mass-action kinetics, but they are often approximated by Hill-type kinetics [17]

$$\emptyset \longrightarrow \mathbf{mRNA}$$

with a reaction rate given by $v = cA^n/(K^n + A^n)$ for an activator **A** and $v = cK^n/(K^n + R^n)$ for a repressor **R**, where K is a constant and n is called the Hill coefficient, which reflects the degree of cooperativity of the transcription factor.

B. Stochastic Models of CRNs

The living cell is abuzz with noise. One of the main sources of this noise is the random nature of chemical reactions at the molecular scale, which can be traced back to the thermal motion of molecular reactants [18]. This implies that the timing and order of chemical reactions have a random character. At very large volumes with a huge number of reacting molecules, e.g., a test tube, such randomness averages out, and the concentrations of reactants can be modeled as continuous variables that evolve continuously over time. When the number of reacting molecules is relatively small, as is often the case inside a living cell, the discreteness and low abundance of reactants and the randomness of chemical reactions acting on them can become very important [19], [20]. This randomness manifests as appreciable stochastic fluctuations in the abundance of molecular reactants within the same cell over time. As a result, genetically identical cells exhibit cell-to-cell variability in their species abundances, a fact that can be observed and quantified experimentally.

Random fluctuations in living cells (commonly referred to as noise) propagate in a biochemical reaction network and can impact events and processes based on the network topology and the dynamics of the reactions that define it [21], [22]. Cellular noise has been classified based on its source [23], [24]: *intrinsic noise* refers to stochastic fluctuations originating from the inherent discrete random nature of the underlying chemical reactions described above, while *extrinsic noise* refers to fluctuations that have more global origins, such as cell-to-cell variability in plasmid copy numbers, RNAP copy number, ribosome copy numbers, local environment, and cell-cycle stage.

I now describe how stochastic models of biochemical reaction networks can be used to capture the discrete and random nature of chemical reactions in living cells [18], [25], [26]. Consider the same set of n species and m reactions, and the corresponding stoichiometry matrix S discussed previously for deterministic kinetics. In the stochastic formulation of chemical kinetics, the abundance of the reacting species is not modeled as a continuously changing concentration variable but rather as a discrete random variable $X_i(t)$ that measures the number of molecules of the ith species $\mathbf{X}_i$ at time t. The chemical system is, thus, described by a vector stochastic process $\{X(t)\}_{t\geq 0}$ with a state space $\mathbb{Z}_+^n$. When the kth reactions fires, the state of the process changes from $X(t) = x$ at time t to $X(t^+) = x + s_k$, where s_k is the stoichiometry vector associated with the kth reaction. To fully describe this stochastic process, one must also specify when the different reactions take place. The firing time of each reaction will be an exponential random variable with parameter λ_i that depends on the state of the system $X(t)$ and the reaction rate constant. More precisely, given $X(t) = x$, the probability that the kth reaction fires in the time interval $[t, t+h)$ is given by $\lambda_k(x)h + o(h)$. For mass-action kinetics, the transition intensity $\lambda_k(x)$ is given by

$$\lambda_k(x) = c_k' \prod_{i=1}^{n} \binom{x_i}{\nu_{ki}}$$

where c_k' is a reaction rate constant, and the product term describes the number of different combinations in which the individual reacting molecules can come together to form the proper number and type of reactants needed for reaction i to fire. The transition intensities $\lambda_k(x), k = 1, \ldots, m$, are referred to as the propensity functions in the chemical literature.

As with the deterministic case, I will focus on elementary reactions of order at most 2. Table 2 shows the propensity functions for each of these reactions.

For a given reaction, the stochastic reaction rate constant c' is generally not equal to the deterministic reaction rate constant c although they are closely related. This is because, in the stochastic case, $X(t)$ is a vector

Table 2 Propensity Function $\lambda(x)$ for Zero-, First-, and Second-Order Reactions That Follow Mass-Action Kinetics

Reaction type	Reaction	Propensity $\lambda(x)$
zero-order	$\emptyset \xrightarrow{c'}$ Products	c'
monomolecular	$\mathbf{X}_i \xrightarrow{c'}$ Products	$c'x_i$
bimolecular $(i \neq j)$	$\mathbf{X}_i + \mathbf{X}_j \xrightarrow{c'}$ Products	$c'x_ix_j$
bimolecular	$2\mathbf{X}_i \xrightarrow{c'}$ Products	$c'x_i(x_i-1)/2$

of integers representing molecule counts in a reaction volume, whereas, in the deterministic case, $X(t)$ is a vector of *concentrations*, so the reaction volume is already accounted for. Table 3 shows the relationship between the deterministic and stochastic reaction rate constants.

As defined, the vector stochastic process $\{X(t)\}_{t\geq 0}$ is a continuous-time discrete-state Markov process. It can be compactly described using independent unit-rate Poisson processes $\{Y_k(t)\}_{k=1}^m$, following Kurtz's random time-change formulation [27]

$$X(t) = X(0) + \sum_{k=0}^{m} Y_k\left(\int_0^t \lambda_k\left(X(s)\right) ds\right) \cdot s_k.$$

This formula suggest one way to generate sample trajectories of the stochastic process $\{X(t)\}_{t\geq 0}$ [28]. Other equivalent methods for generating sample trajectories include Gillespie's stochastic simulation algorithm (SSA) [18], [25]. Aside from generating sample paths, one is often interested in computing the probability mass function $\{p(x,t) : x \in \mathbb{Z}_+^n\}$ for different times, t. This function evolves according to Kolmogorov's forward equation, which is usually referred to in the chemical literature as the chemical master equation (CME) [29]

$$\dot{p}(x,t) = \sum_{k=1}^{m} p(x - s_k, t)\,\lambda_k(x - s_k) - p(x,t)\,\lambda_k(x)$$

which holds $\forall x \in \mathbb{Z}_+^n$. Here, $p(x,t)$ is taken to be 0 whenever $x \notin \mathbb{Z}_+^n$. Note that the CME defines one *linear* differential equation for each x in the state space, $\mathbb{Z}_+^n$. Since the subset of $\mathbb{Z}_+^n$ where $p(x,t) \neq 0$ is often infinite, one has to solve an infinite set of linear differential equations to compute the probability mass function. One can judiciously truncate the state space to get a finite linear system that can be solved. If done properly, the truncated system can give error bounds to the truncation error. The finite state projection (FSP) method shows how this can be done [30], [31].

Alternatively, one may be interested in computing the statistical moments, $\mathbb{E}[X^i(t)]$, of the process $\{X(t)\}_{t\geq 0}$. Such moments evolve deterministically, according to ODEs that capture the moment dynamic equation. The dynamics of the first moments are given by

$$\dot{\mathbb{E}}\left[X(t)\right] = S\mathbb{E}\left[\lambda\left(X(t)\right)\right]$$

where $\lambda(x) = (\lambda_1(x), \ldots, \lambda_m(x))^T$ is the vector of reaction propensity functions. For the second moments, the dynamics are described by

$$\begin{aligned}\dot{\mathbb{E}}\left[X(t)X^T(t)\right] &= \mathbb{E}\left[S\lambda(X(t))X^T(t)\right] + \mathbb{E}\left[X(t)\lambda^T(X(t))S^T\right] \\ &\quad + \mathbb{E}\left[S\,\mathrm{diag}\left(\lambda(X(t))\right)S^T\right].\end{aligned}$$

Note that these equations are generally not closed, i.e., the first-order moments will depend on the second-order moments, which will, in turn, depend on the third-order moments, and so on. This makes these equations difficult to solve. To address this problem, several moment-closure methods have been suggested in the literature [32]–[34]. These methods express higher order moments as functions of lower order ones based on certain assumptions about the statistics of the process being described. The resulting moment equations will be closed and can be solved numerically. However, the solution remains approximate, and the error depends largely on the validity of the assumptions made. Fortunately, when the propensity functions are affine, $\lambda(x) = \lambda_0 + \Lambda x$, as would be the case when the reaction network has only zero-order and monomolecular reactions, the moment equations simplify considerably: they become closed and linear. Indeed, for affine propensities [35]

$$\dot{\mathbb{E}}\left[X(t)\right] = S\lambda_0 + S\Lambda\mathbb{E}\left[X(t)\right].$$

The second moment equations simplify in a similar manner. Instead of giving these here, I will instead show the equations of the covariance matrix $\Sigma = \mathbb{E}(X - \mathbb{E}[X])\,\mathbb{E}(X - \mathbb{E}[X])^T$

$$\dot{\Sigma}(t) = S\Lambda\Sigma(t) + \Sigma(t)\,\Lambda^T S^T + S\,\mathrm{diag}\left(\lambda_0 + \Lambda\mathbb{E}\left[X(t)\right]\right)S^T.$$

Note that the mean and covariance equations above are linear and can be readily solved analytically.

VI. BIOLOGICAL CHASSIS

As the basic unit of life, a cell can be viewed as a factory that processes energy and materials (see Fig. 5). It takes in nutrients in the form of sugars, amino acids, and fatty acids and then breaks them down, through the process

Table 3 Relation Between the Deterministic Reaction Rate Constant, c, and the Stochastic Rate Constant c'. Ω Is the Reaction Volume, e.g., Volume of the Cell

Reaction type	Relation between c and c'
zero-order	$c' = c \cdot \Omega$
monomolecular	$c' = c$
bimolecular $(i \neq j)$	$c' = c/\Omega$
bimolecular	$c' = 2c/\Omega$

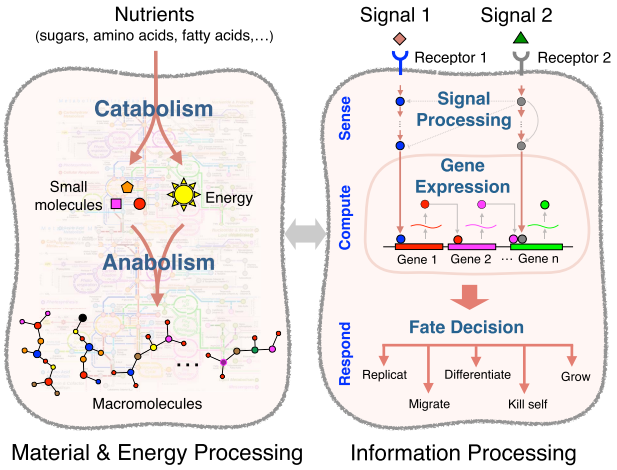

Fig. 5. ***Different facets of the living cell. A cell can be viewed as an efficient factory, taking in raw materials in the form of nutrients, breaking them down into small molecules, and generating energy that is subsequently used for building complex large molecules. The living cell can also be viewed as a sophisticated command-and-control center: it senses a multitude of dynamic signals from its environment, transduces these signals, processes them through signaling pathways, and then passes the processed signals to gene regulatory networks for information integration and decision-making. Making the correct decision is essential for the survival of the cell itself or the multicellular organism that it may be part of. The two facets are strongly interrelated, and together, they present a unique biological context within which synthetic circuits function. This, in turn, introduces design considerations that are quite unique to biological circuits.***

of catabolism, into small molecules and energy sources, e.g., ATP . Then, through the process of anabolism, the cell expends energy to combine some of the small molecules into macromolecules that perform the cell's most complex functions. The combination of catabolism and anabolism defines the cell's metabolism, a process that is highly regulated using a multitude of enzymes controlled by gene regulatory networks. It is interesting to reflect on the architectural features of metabolism. The process takes in a very large variety of nutrients and converts them into a much smaller variety of small molecules and energy carriers, which are subsequently converted into a huge variety of macromolecules. This architecture has been compared to a "bowtie," which is large at both ends but small in the middle at the "knot." This is an important architectural feature of metabolism where control is heavily focused on the middle part [36].

The cell can alternatively be viewed as an information processing unit that senses information from the outside and transduces that information into chemical signals that travel through signaling pathways where they are processed and used to drive complex gene expression networks (see Fig. 5). The resulting gene expression pattern determines the set of actions that the cell takes in response to the external and internal states. Possible actions include moving, dividing, growing, differentiating, or even committing suicide (apoptosis). Yet, other cellular decisions are made through the faster protein–protein interaction networks without requiring gene expression that takes place at longer time scales. This is not unlike layered control system architectures in engineering, which employs slow and global control at high layers and fast, local control at lower layers.

The two facets of the cell just described are far from being independent of each other. Information processing, as well as gene expression in general, hinges on the availability of sufficient energy and materials made available by cellular metabolism. At the same time, cellular metabolism relies on the cell's signal processing and gene expression networks for finding nutrients and producing the necessary enzymes needed to metabolize them. The engineering of novel genetic circuits must contend with this biological context and its underlying layered architecture. Such context is quite unique to biology, and it brings challenges that must be addressed with new solutions that are forged specifically to deal with the biological substrate.

A. Challenges for Building Synthetic Circuits

In spite of the availability of efficient genetic engineering tools, the practical realization of even simple controller designs demands that many practical challenges peculiar to the cellular environment be confronted.

Indeed, the environment within the living cell is quite distinct from those in which more classical engineering control systems operate. This calls for the creation of novel control methods and tools that are designed specifically for living cells. Next, I will briefly outline some of the main difficulties facing genetic control system design and implementation. I must hasten to add that, while many of these challenges are unique to biological circuits, others are not restricted to genetic controller design but are also faced by designers of nonbiological circuits and systems.

1) Positivity of Variables: Biomolecular reactions and the networks that they form are the dynamical systems that underlie cellular function. In such systems, the variables are species concentrations, which are always nonnegative quantities. This is a fact that any genetic circuit, whether natural or synthetic, must contend with. Simple and complex computations alike must be realized using positive variables, which adds a major constraint that has to be respected by any engineered circuits. We will see this firsthand later on in this article, when I discuss synthetic integral feedback systems.

2) Unknown and Uncertain Controlled Network: Despite our ever-increasing understanding of cellular network interconnections, it is inevitable that key players and interactions will be missing. This results in models of cellular behaviors that are incomplete, both in terms of topology and parameters. While model and parameter uncertainty are common in typical control engineering applications, the extent of this uncertainty is typically much bigger in models of cellular systems. This is due to their sheer

complexity and the lack of measurement devices of the resolution and accuracy of those available for nonliving engineered systems. Nevertheless, similar to more traditional control applications, model and parameter uncertainty make feedback control an effective strategy for achieving robust regulation [37].

3) Stochastic Cellular Noise: As mentioned earlier, chemical reactions in a living cell are inherently stochastic [38]. Within the cell, the times when chemical reactions fire are random, a fact that becomes significant when the reacting species (e.g., metabolites, DNA, mRNA, and proteins) are present in low copy numbers. In the presence of such scarcity, the resulting randomness cannot be overcome by the law of large numbers, and the consequences are both measurable and consequential. As mentioned previously, this intrinsic noise manifests as random fluctuations in the abundances of reacting species over time and is a key contributor to cell-to-cell variability among otherwise identical cells [23].

4) Context Dependence: Context dependence refers to a host of conditions that are responsible for differences in the specific state in individual cells. These may include the cell's local microenvironment, internal cellular resources (e.g., ATP, RNAP, and ribosome abundances), amount of genes delivered (gene dosage), growth conditions, specific state of the cell cycle, and so on [39]. This context dependence presents challenges that are quite unique to biological circuits. Overcoming these challenges requires new design strategies that often depart from those used in the design of electronic circuits.

5) Resource Sharing: When synthetic circuits are genetically engineered into living cells, such circuits must share the host cell's resources to express their genes and realize the designed synthetic network. Such resource sharing introduces a dynamic coupling between the designed synthetic components themselves and the host cell's natural circuits [40]. Such a coupling, if significant, can interfere with the cell's behavior and lead to unpredictable functionality of the engineered gene circuit. Genetic engineers must pay close attention to such effects and must design their circuits so that they do not have a large "resource footprint." As we will see later, regulatory circuits can also be designed to mitigate some of the burden resulting from resource sharing [41], [42].

6) Loading Effects and Modularity: When interconnecting molecular circuits and devices, the constituting molecules of these systems interact creating a loading effect, much like that seen when interconnecting two electric circuits. The effective influence of the interconnection on the dynamics of the individual circuits has been referred to as retroactivity [43]. As with electric circuits that do not have infinite-input/zero-output impedances, retroactivity makes the dynamic behavior of the connected devices different from their behavior in isolation. This, in turn, makes systematic modular design and scale-up more challenging. To recover modularity, biological insulator circuits and load drivers have been proposed [44]. However, it remains to be seen whether strict modularity should be a goal of biocircuit designers. A yet-to-developed alternative approach that embraces retroactivity and works with it may lead to more flexible and less costly designs. This seems to be more in line with the way natural biological circuits function.

7) Crosstalk: Unlike electric circuits where signals can be sent through wires that act as conduits for information transmission, targeted communication between different species must rely on the specificity of their interaction. Nonspecific interactions are a recurring theme inside the living cell, leading to crosstalk effects between different signaling pathways and subsystems. In the face of such crosstalk, biomolecular circuit design becomes more challenging. Careful selection of parts and robust designs that explicitly account for extraneous signals can help overcome the effect of crosstalk.

8) Mutational Escape: This is one of those challenges that have no analog among electric circuits. *De novo* engineered genetic circuits inevitably add some burden to the host cell. If this burden is big enough, it will end up slowing the growth rate of the host cell. Starting with an entire population of such engineered cells, the problem will not be apparent at first. As cells grow and divide, random genetic mutations take place, and occasionally, one of those mutations will affect the genetic circuit of one cell, knocking its designed circuit out of function. Free from the added burden of the designed circuit, this cell along with its progeny will grow slightly faster than the rest of the population and, hence, will gain a small fitness advantage. Given the exponential growth rate of cells, even a slight growth advantage will quickly make this mutant subpopulation take over the entire population. The result is a cellular population devoid of the designed circuit. Strategies for overcoming mutational escape and designing evolutionarily robust circuits are beginning to emerge [45].

VII. DESIGNING SYNTHETIC CONTROL SYSTEMS

A. Synthetic Negative Feedback Circuits

There are many benefits of negative feedback besides robustness to parameter variations. Negative feedback endows a system with robustness to dynamic uncertainty. It can also be used to linearize the response characteristics of a nonlinear system. This is particularly evident when the feedback causes the output of the circuit to track its input perfectly, as we will see later with integral feedback. The linearizing effect of negative feedback was demonstrated experimentally in [46], where it was shown that autoregulation linearizes the steady-state input–output map (dose response) of otherwise nonlinear response. This is not unlike the action of negative feedback in an amplifier, where the feedback reduces the effect of nonlinear gain

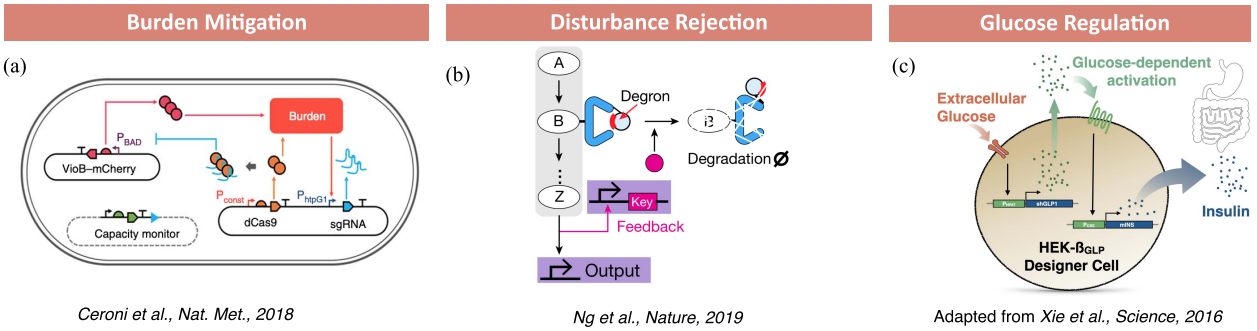

Fig. 6. ***Selected examples of genetically engineered negative feedback circuits. (a) Engineered negative feedback control system in E. coli designed to counteract gene expression burden. When burden is sensed, target genes are repressed, bringing down the burden level. (b) Negative feedback control circuit in yeast that uses a nonnatural protein in the feedback loop. The circuit was demonstrated to reject disturbances, e.g., perturbations to protein Z. (c) Mammalian cell that has been engineered to sense glucose and secrete insulin in proportion to sensed glucose. When implanted in mice, the closed-loop negative feedback system was shown to mitigate glycemia in Type I diabetic mice.***

distortions, making the input–output characteristics more linear [47].

Negative feedback can yield improved step response dynamics, resulting in closed-loop circuits that respond faster to a step input than ones that reach the same steady-state output level with no feedback [48]. This is achieved by exchanging high gain with speed. In [49], genetically engineered metabolic feedback dramatically shortened the rise time of metabolites, decreasing it by as much as 12-fold.

Yet, another benefit of negative feedback for genetic circuits derives from its ability to reduce the impact of intrinsic cellular noise. This was demonstrated experimentally in the year 2000 by the first synthetic negative feedback circuit [13]. The gene circuit consisted of a self-regulated gene (its output repressed its own input) whose output can be measured by using a fluorescent protein tag. Using such measurements, it was shown that negative feedback reduces the variance of the output, as can be seen when comparing the output distribution of a large number of genetically identically autoregulated cells with that of a different set of cells lacking autoregulation. The mechanism of noise suppression using feedback was also theoretically addressed in [38], [50], and [51].

Negative feedback can also facilitate effective reference tracking. Hsiao *et al.* [52] used synthetic protein scaffolds, for achieving concentration tracking through negative feedback. The input to the circuit triggered scaffold production, which, in turn, resulted in the production of an inhibitory antiscaffold protein, closing the feedback loop. This tracker circuit was demonstrated to achieve protein concentration tracking in *E. coli*.

The transactivator (TA) function as input actuator is critical for activating gene expression in mammalian cells. When the number of plasmids carrying these TAs varies, so will the protein output of the gene being activated. If the plasmid number reaches high levels, the cell will become burdened by the TA expression, and the maximal output achievable will be reduced. Using microRNAs to impose negative feedback at the posttranscriptional level, it was shown in [53] that such feedback provides robustness to TA dosage variability, reducing the burden and increasing the maximum achievable protein output.

Gene expression burden is a frequent problem facing genetic circuit designers, and methods for its mitigation often rely on the use of negative feedback. Fig. 6(a) shows one such circuit, which was reported in [54]. The authors used a promoter that was activated by burden and used it to function as a burden biosensor. When the burden is sensed, the promoter drives the expression of a CRISPR single-guide RNA (sgRNA) [55], which was designed to bind to a specific target region on the promoter of a gene thought to be burdensome. The expressed sgRNA binds to a constitutively expressed repressor protein, called dCas9, and "guides" it to bind to the target promoter region, thereby repressing the expression of the burdensome gene. When the burden state of the cell improves, the sgRNA will be downregulated, and the expression of the target gene will be relaxed. Thus, negative feedback mitigates the impact of the burden by regulating the expression of burdensome genes. Different concepts for mitigating burden using negative feedback circuits were also reported in [53] and [56]–[58], among others.

Rejection of exogenous disturbances is a hallmark of robust systems—a desirable feature that can be realized with negative feedback. Fig. 6(b) shows one such negative feedback applied to the mating pathway of yeast cells [59]. The pathway activity, which is regulated by a protein regulator (B in the figure), is measured by an output fluorescent protein. The same transcription factor, Z, which activates the output also activates the expression of a protein called "Key." The regulator protein B is fused to a cage protein hiding a degron. When exposed, degrons recruit the cell's machinery to degrade themselves and the proteins that they are fused to. The expressed Key protein unlocks the hidden degron, exposing the

regulator-cage-degron complex to degradation and, in the process, downregulating the expression of the output (and Key) in a negative feedback fashion. This negative feedback circuit was shown to reject disturbances in the form of externally induced changes in the creation/removal of Z. An interesting aspect of this circuit is that the key-cage proteins were not natural to a living cell; they were instead designed using protein engineering methods [60].

A vast number of human diseases are caused by a breakdown in homeostasis, the regulation achieved by the body's own control systems. One example is Type I diabetes, where the patient's own immune system attacks the body's blood glucose controllers, the beta cells. In this next example shown in Fig. 6(c) and reported in [61], the authors engineered mammalian designer cells to provide closed-loop glycemic control to restore homeostasis for Type I diabetic mice. Unlike the negative feedback circuits discussed so far where the closed-loop systems were contained in their entirety within the confines of single cells, here, only the feedback controller was engineered within the cell, with the remaining part of the system being the entire blood volume of the animal. The implanted cells were engineered to sense the extracellular glucose concentration. In response to high glucose concentrations, the cells activated the expression of a synthetic insulin gene, whose insulin product was then secreted out of the cell. Once in the blood, insulin acted on the body's glucose metabolism to bring down the glucose concentration. This negative feedback loop was shown to restore glucose homeostasis in diabetic mice.

The examples I have selected above demonstrate the diversity of benefits that genetically engineered negative feedback can confer, as well as the wealth of different problems that it has solved. The list is not comprehensive and serves only as a sampling of the growing literature on synthetic negative feedback circuits (see [62]–[66] among several others for additional implementations).

B. Engineering Robust Controllers With Feedforward Control

One control motif that has been discovered repeatedly in natural systems is the so-called feedforward motif [48], [67]. There are several types of feedforward motifs, but they can generally be divided into two classes: coherent and incoherent. Feedforward motifs include two paths that start from the input and converge to produce the output, and the difference between the two classes of feedforward motifs is related to the coherence of the sign of the action of each. If an increase in the input leads to activation (positive action) of the output by both paths, then the feedforward is said to be coherent. If, on the other hand, an increase in the signal leads to activation of the output by one path and repression by the other, then the two paths have opposing (incoherent) effects on the output, and the underlying topology is called incoherent feedforward. Here, I will focus on one class of motifs, the type 1 incoherent feedforward loop (iFFL) motif. A prototype of this regulatory motif is shown in Fig. 7. As can be seen in the figure, the input is u, and the output is y. A direct path positively connects u to y, while a second path positively connects u to an intermediate variable x, which then inhibits y, resulting in a negative connection from u to y. Of course, there are many different implementations of this motif, depending on the species that realize u, x, and y, and the mechanisms by which activation and repression take place.

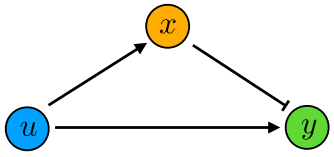

Fig. 7. *Network topology of an iFFL. The node u acts as an input node, which activates the output node y. u also activates an intermediate node x, which, in turn, represses y. The activation and repression actions on y are usually delayed, so the effect of the topology on the input–output behavior can be better seen by looking at the dynamics of y. If the opposing actions are balanced at steady state, the output adapts to changes in the input u.*

At a first glance, it may seem that having two incoherent paths connecting the input to the output is, well, incoherent. The resolution of this paradoxical implementation lies in the dynamics, where some paths take their action faster than others, with the delay serving a useful role. The antagonistic actions of both paths also serve to mitigate the effect of input changes on the output at a steady state, a useful property, as we will see in the following. This is not unlike forward control loops used by control engineers in the process industry, where measurements of a disturbance that adversely affects the process output are reintroduced into the forward path in the negative phase in order to counteract the direct effect of the disturbance on the output.

One of the benefits of the iFFL and all the other feedforward motifs is that, when realized with dynamically stable subsystems, dynamic stability is not an issue of concern, as the overall system will be stable. This is in contrast to feedback motifs in which dynamic stability must be carefully considered, as the closed-loop feedback system can become unstable even when all its subsystems are stable. I will next examine some of the regulatory benefits of this iFFL motif, which could explain why this motif is so prevalent in natural circuits and why it has been the subject of several synthetic implementations.

1) iFFL Can Speed Up the System Response: Consider the iFFL circuit in Fig. 8. The input u can be thought of as resulting from an external signal from outside the cell that has been sensed by a receptor on the cell surface and transduced by a signaling pathway leading to the activation of the input molecule u that acts as an activator. For the purpose of the analysis, we can think of u as the

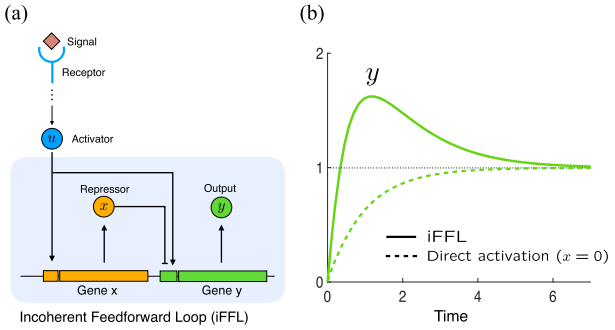

Fig. 8. *(a) Simple iFFL circuit. An external signal is sensed by the receptor and transduced to activate a transcription factor, u, which is considered as the circuit input. Itself an activator, u turns on expression of Gene x and Gene y. The product of Gene x is a repressor that counteracts the effect of u on the expression of y, which reaches a steady state that depends on both activation and repression strengths. Since the activation by u is immediate, while repression by x is delayed, the expression of y proceeds rapidly in the beginning and then comes back down after the concentration of the repressor y builds up. (b) Speed up in response rise time in comparison to a direct activation by y where Gene x is not present. To reach the same steady-state value in both cases, the transcription rate must be higher in the case of iFFL circuit ($\rho = 3.78$ for iFFL versus $\rho = 1$ for simple activation). The remaining parameters are $\alpha = 1$, $\gamma_1 = \gamma_2 = 1$, $x = 1$, and $k = 0.36$.*

independent input of the network. u activates two genes: Gene-x and gene-y. Gene-y expresses a protein y, which is the output of the system, while gene-x expresses an intermediate protein, x, which functions as a repressor of gene-y. The circuit clearly functions as an iFFL. A dynamical model of this circuit is given by

$$\dot{x} = \alpha u - \gamma_1 x$$
$$\dot{y} = \rho \frac{uk}{k + x^2} - \gamma_2\, y.$$

Here, gene expression was modeled in one step, for simplicity. It can be seen that u increases the expression rate, while x decreases it. Upon a change of the input from 0 to 1, it can be seen that expression rises quickly due to the direct activation by u, while the repression by x initially remains low. Once x builds up sufficiently, it starts repressing gene expression, bringing it back to a steady state that is a balance of activation and repression. In contrast, one can imagine direct activation of y by u with no intermediate species x. For y to reach the same steady state as in the iFFL case, the transcription rate needs to be lower, which means that the initial rate of increase of y must necessarily be slower. In this way, the iFFL can be used to speed up the response of y to input u. An experimental study of this speedup can be found in [68].

2) iFFL Motif Exhibits Adaptation Properties: Since the iFFL combines activation with repression, it has the potential to achieve adaptation by compensating for changes in the input or other parameters shared by its two opposing paths. Consider the iFFL circuits in Fig. 9, which represents synthetic gene circuits that have been built and tested. The circuit in Fig. 9(a) depicts a simplified microRNA-based iFFL circuit. Such circuits were engineered in [41], [53], [69], and [70] to demonstrate different aspects of adaptation. A microRNA is a small single-stranded RNA that binds to a specific mRNA with a complementary sequence and induces its degradation. MicroRNAs are also transcribed by RNAPs, but, unlike mRNA, they do not code for proteins; therefore, instead of being translated, they are processed and prepared for their function as posttranscriptional regulators. The circuit in Fig. 9(b) represents a similar iFFL motif, except that the protein product, x, of the first gene is an enzyme, called riboendonuclease, which binds to the mRNA coding for protein y and cleaves it, thereby preventing its translation. This synthetic circuit was recently engineered in [42]. To understand some of the adaptation features of these two circuits, I will examine a simplified model of the first; the analysis of the second circuit is similar.

For the microRNA circuit, a basic model is given here

$$\dot{x} = f(u) - \gamma_x x$$
$$\dot{m}_y = f(u) - \gamma_{m_y} m_y - \eta x m_y$$
$$\dot{y} = k_y m_y - \gamma_y y$$

where $u(t)$ is the concentration of the input protein u, $x(t)$ and $y(t)$ are the concentrations of microRNA x and protein y, and m_y is the mRNA for y. The transcription rate $f(u)$ is taken to be a monotonically increasing function of the input, u. $f(\cdot)$ will also depend on other parameters, such as the amount of genetic material encoding the genes (e.g., plasmid copy number). This amount frequently varies from one cell to the next, leading to cell-to-cell variability in the transcription rate. To see the impact of such variability on

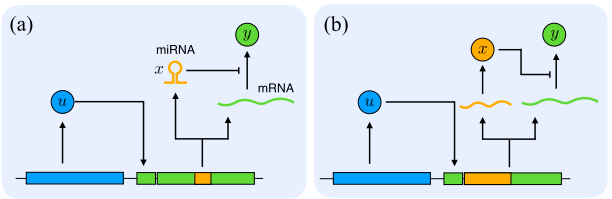

Fig. 9. *iFFL implementations in mammalian cells. (a) Network that implements an iFFL using a microRNA as the intermediate variable x. The microRNA binds to the mRNA of y and degrades it. This circuit was demonstrated to mitigate burden and plasmid copy number variability [41], [53], [69]. (b) Similar network, except that the intermediate node is a protein called riboendonuclease that degrades the mRNA of y. The network was shown to mitigate gene expression burden [42]. Because it is the protein that performs the degradation, one expects that this circuit should help mitigate variations in the translational resources (e.g., variation in ribosome copy numbers).*

the output, let us examine the steady-state value of y

$$y_{ss} = \frac{\frac{k_y}{\gamma_y} f(u)}{\frac{\eta}{\gamma_x} f(u) + \gamma_{m_y}}.$$

From this, it can be seen that, when $f(u)$ is sufficiently large (e.g., $f(u) \gg \gamma_{m_y}\gamma_x/\eta$), the output y_{ss} will be insensitive to the transcription rate $f(u)$ and, hence, to variations in the genetic template. This was indeed observed in synthetic circuits, implementing this iFFL loop [53], [69].

Another source of variability is in gene expression resources. For example, in the case of transcription, the amount of RNAP may vary from one cell to the next, or it may vary over time in the same cell, as other genes are turned on siphoning RNAP away for this circuit. The problem of limited resource sharing is one of the common problems facing synthetic circuit designers, as it introduces unwanted couplings between the expressions of different genes. In this circuit, the function $f(\cdot)$ captures the dependence of transcription on transcriptional resources. As with the genetic template variability, changes in the transcriptional resources will be buffered by this circuit for large $f(u)$. This is the rationale for building this circuit and using it for burden mitigation in mammalian cells [41].

In light of the above discussion, it is not difficult to see that the iFFL motif can achieve robust perfect adaptation (RPA). In fact, if $\gamma_{m_y} = 0$, then y_{ss} will be independent of u, and the previous two designs can achieve RPA with $y_{ss} = (k_y\gamma_x)/(\eta\gamma_y)$ for the first circuit and $y_{ss} = (k_y\gamma_x\gamma_{m_x})/(\eta k_x\gamma_y)$ for the second. Other iFFL circuits that achieve RPA have been studied in the literature. Two examples include the sniffer network [71] and a fold-change detection circuit [72]–[74]. These circuits differ in their mechanism of inhibition, with one circuit repressing the expression of the output, while the other enhancing its degradation, but they converge in that their outputs depend only on network parameters but not on their input. Finally, I mention that, depending on its implementation, the iFFL motif can exhibit noise buffering properties as well [75]–[77].

3) Other Implementations: Aside from the synthetic iFFL implementations mentioned above, in [78], iFFLs were engineered into *E. coli* promoters using transcription-activator-like effectors (TALEs), resulting in similar gene expression in different genome locations and plasmids in spite of gene copy number variability. A synthetic implementation of the iFFL motif was also successfully built and tested in a cell-free system (*in vitro*) and living cells [79]–[82]. The *in vitro* circuit in [79] exhibited perfect adaptation and another property called, fold-change detection, whereby the output signal is invariant to positive scalar multiples of the input—a property that some living cells use to respond to relative, rather than absolute, change in their input, which gives them a large response range.

C. Engineering Robust Controllers

Thus far, the negative feedback examples that I discussed employed finite-gain in their feedback path. While such feedback has tremendous benefits as seen previously, it generally does not result in zero steady-state tracking errors to constant reference inputs nor does it guarantee that constant disturbances are rejected perfectly. In both cases, a steady-state error will persist. Furthermore, this error will depend on the network topology, changing in either direction as the network or controller parameters change. One can make the error smaller, by increasing the feedback gain, but, aside from being difficult to achieve at the molecular level, high-gain feedback can have undesirable consequences, such as instability and poor dynamic performance. In control engineering, these problems are solved by introducing an integrator in the controller. The integral controller can be viewed as delivering infinite gain at zero frequency, which completely gets rid of the steady-state error. At the same time, the introduced gain decreases with increased frequency and can be arranged to be sufficiently small at higher frequencies, so it does not promote instability.

Integral feedback has been discovered in many biological systems functioning at different levels and time scales. The first paper to identify integral feedback in living cells was [83], which showed that such feedback was the main mechanism for achieving robust adaptation in bacterial chemotaxis in *E. coli* [84], [85]. Since then, integral feedback was discovered in the regulation of intracellular osmolarity [86], [87], sigma-70 activity [88], ammonia uptake [89], temperature control [90], copper and zinc homeostasis [89], iron homeostasis [90], calcium homeostasis [91], and glucose uptake in growing cancer cells [92]. The main question that I will focus on next is how to genetically engineer integral feedback so that one can exploit its remarkable robustness benefits in synthetic biology. I will start by posing the robust tracking and disturbance rejection problem at the molecular level, and then, I will show how a form of integral feedback provides a solution to it.

1) Robust Steady-State Tracking and Disturbance Rejection: Given an uncertain intracellular system (the plant network) shown in Fig. 10, the system evolves according to stochastic dynamics arising from the stochastic biochemical interactions of its molecular constituents (see Section V-B). It is assumed that the network can be actuated only by activating $\mathbf{X}_1$, the control input species, using only measurements of the output species $\mathbf{X}_L$. Find a feedback interconnection of stochastically interacting species (controller network) so that the following requirements are met.

1) The closed-loop interconnection of the intracellular network and the controller network is ergodic (ergodicity is a notion of stochastic stability).
2) The regulated variable, X_L, tends asymptotically to the reference setpoint r, i.e., $\mathbb{E}[X_L(t)] \longrightarrow r$ as

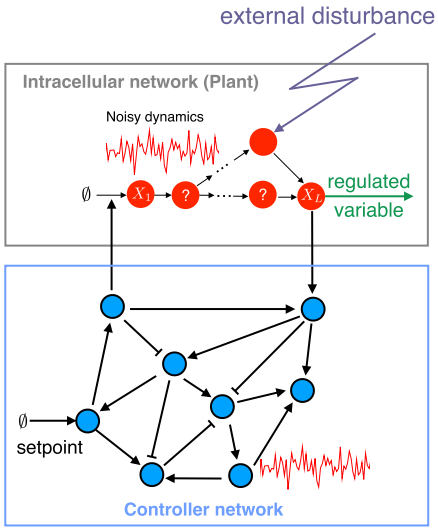

Fig. 10. *RPA problem setup. Given (a possibly uncertain) network of stochastically interacting cellular species $X_1, \ldots, X_L$, which we call the plant. Find a controller network of interacting species that take the regulated variable X_L and a reference signal r as inputs and feedback an output signal that actuates the plant through its input X_1 such that the closed-loop system achieves stability (suitably defined), and the mean value of the regulated variable follows the set point asymptotically in spite of constant disturbances, parameter shifts, or network perturbations that maintain the stability of the interconnection.*

$t \longrightarrow \infty$, in spite of constant unknown external disturbances and network perturbations that leave the closed-loop system ergodic.

The state space of our system is $\mathbb{Z}_+^N$, where N is the sum of the number of controller species and plant species. Here, ergodicity of the network simply means that the closed-loop system has a unique stationary distribution over the state space to which all initial distributions converge [93].

In the systems biology literature, when the output of a system of interest that is subjected to a constant input (e.g., external disturbance) responds to a sudden change in the value of that input by returning to its prechange level, the system is said to achieve perfect adaptation. When perfect adaptation is robust to certain perturbations, the system is said to achieve RPA. It is important to note that the perturbations to which the adaptation property is robust must be specified. With this in mind, I will take robust perfect adaption to mean robust steady-state tracking and disturbance rejection as defined above, and I will use this terminology going forward.

2) Antithetic Integral Control: A solution to the above problem was given in [88], where a simple controller topology was shown to provide the desired RPA. This controller consists of two species and four reactions, which together perform four different functions (see Fig. 11). These reactions are

$$\begin{aligned}
\mathbf{X}_L &\xrightarrow{\theta} \mathbf{X}_L + \mathbf{Z}_2 \quad \text{(sensing reaction)} \\
\mathbf{Z}_1 &\xrightarrow{k} \mathbf{Z}_1 + \mathbf{X}_1 \quad \text{(actuation reaction)} \\
\emptyset &\xrightarrow{\mu} \mathbf{Z}_1 \quad \text{(setpoint reaction)} \\
\mathbf{Z}_1 + \mathbf{Z}_2 &\xrightarrow{\eta} \emptyset \quad \text{(computation reaction).}
\end{aligned}$$

The corresponding dynamical equations for the mean of Z_1 and Z_2 of the controller are given as follows:

$$\begin{aligned}
\dot{\mathbb{E}}[Z_1] &= \mu - \eta \mathbb{E}[Z_1 Z_2] \\
\dot{\mathbb{E}}[Z_2] &= \theta \mathbb{E}[X_L] - \eta \mathbb{E}[Z_1 Z_2].
\end{aligned}$$

Subtracting the two equations gives

$$\dot{\mathbb{E}}[Z_1] - \dot{\mathbb{E}}[Z_2] = \mu - \theta \mathbb{E}[X_L]$$

or equivalently

$$\mathbb{E}[(Z_1 - Z_2)(t)] = \int_{-\infty}^{t} (\mu - \theta \mathbb{E}[X_L(s)])\, ds$$

and hence, this network topology realizes integral feedback control. The topology was called antithetic integral control (AIC) owing to the antagonistic role that the two species $\mathbf{Z}_1$ and $\mathbf{Z}_2$ play. Note that, given ergodicity, at steady state, we have

$$\mathbb{E}[X_L] = \frac{\mu}{\theta}$$

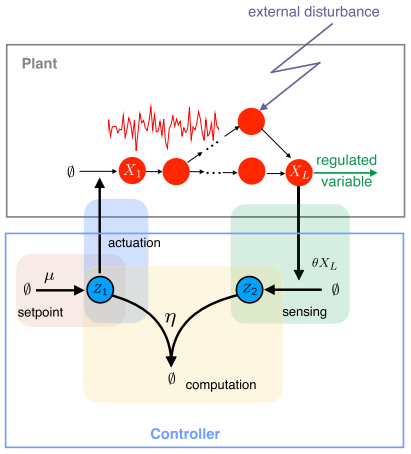

Fig. 11. *Antithetic integral feedback controller. The figure shows the simplest network that solves the RPA under the closed-loop stability assumption. The controller network has two species: Z_1 and Z_2. The rate of production of Z_2 is a proportional function of X_L, and hence, Z_2 acts as a sensor of the abundance of the regulated variable. Z_1 is produced at a rate μ, and it serves as an actuator by activating X_1. The loop is closed when Z_1 and Z_2 mutually sequester (or annihilate) each other through the reaction $Z_1 + Z_2 \longrightarrow \emptyset$. The set point to which the output will converge is given by μ/θ. The network is an implementation of integral feedback control.*

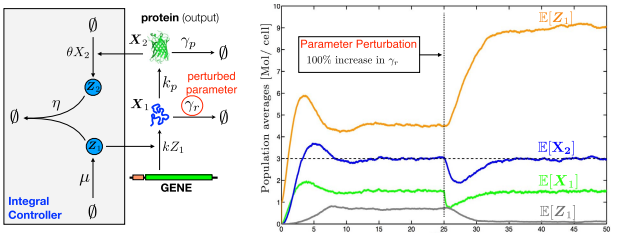

Fig. 12. ***Left: antithetic integral controller in feedback with a gene expression process. The average protein levels are guaranteed to converge to μ/θ in the steady state. Right: results of numerical simulations showing the dynamics of the mean of all closed-loop states before and after a perturbation is applied. It can be clearly seen that the output adapts perfectly to a 100% increase in the parameter γ_r. The plotted trajectories are estimates of the true mean values computed by taking the mean of 8000 sample paths generated using Gillespie's SSA. The parameters used in the simulation are given as follows: $\mu = 3$, $\theta = 1$, $k_r = 1$, $k_p = 2$, $\gamma_r = 3$, $\gamma_p = 1$, $k = 1$, and $\eta = 50$.***

independent of the network topology or external constant inputs. The robustness of this antithetic integral controller can be seen in Fig. 12 where it is used in feedback to control a simple gene expression network. For the parameters chosen, the setpoint is $r = 3$, and it can be seen that the output of interest $\mathbb{E}[X_2]$ goes asymptotically to r from its initial value of zero. When the parameter γ_r is suddenly increased by a factor of 2 at time $t = 25$ and maintained at this higher value, the output $\mathbb{E}[X_2]$ returns to r after a short transient. At the same time, the other system variables automatically assume different values in order to maintain the output at its desired level. The same robustness to parameter variations can be seen if any other parameters in the system are changed, with the exception of μ/θ, which determines the setpoint value, and hence, the output cannot (and should not) be robust to it.

3) Necessity of the Antithetic Motif: While the antithetic integral controller can successfully achieve RPA, one question that arises is what other controller architectures can do the same. Indeed, it will be very desirable to provide conditions for a candidate controller to achieve RPA. Going further, it would be highly desirable to characterize all molecular controllers that achieve RPA. It turns out these questions can be answered in the case of stochastic dynamics. In [94], simple necessary and sufficient conditions for any controller to achieve RPA were derived. These conditions were used to arrive at the striking conclusion that any controller that achieves RPA in the stochastic setting has to embed within its topology an antithetic motif. More precisely, the molecular species of any such controller can always be partitioned into three distinct classes: $\mathscr{C}^+$, $\mathscr{C}^-$, and $\mathscr{C}^\emptyset$, and there will be a reaction that combines one of the species in class $\mathscr{C}^+$ with another in class $\mathscr{C}^-$ to produce a product in the class $\mathscr{C}^\emptyset$. This latter class always contains the empty set $\emptyset$ (see Fig. 13). In this way, all integrators that achieve RPA when the dynamics are stochastic must be of the antithetic type. Furthermore, the minimal such controller contains two species, which corresponds to the controller topology in Fig. 11. Finally, it must be stressed that this result applies only when the network has stochastic dynamics. If the dynamics of the closed-loop interconnection were deterministic, the antithetic motif is *not* necessary for RPA, as there will be other integrator topologies that can also achieve this property. In fact, I will shortly show several such deterministic integrators. Nevertheless, the antithetic motif remains *sufficient* for RPA in the case of deterministic dynamics.

4) Integral Feedback Controllers for Deterministic Networks: When the dynamics of the plant-controller interconnection network can be assumed to be deterministic, several motifs are known to realize biomolecular integral controllers. Fig. 14 provides a family of such controllers. Some of these require only one species **Z** to be realized, such as the P- and N-type zero-order-outflow controllers (see [90] and the references therein) and the N- and P-type autocatalytic controllers [74], [95], [96], while others require two species, such as the N- and P-type antithetic controllers [88], [94]. I use the descriptor N-type to refer to a controller with negative gain at a steady state. Similarly, P-type controllers are those that have positive steady-

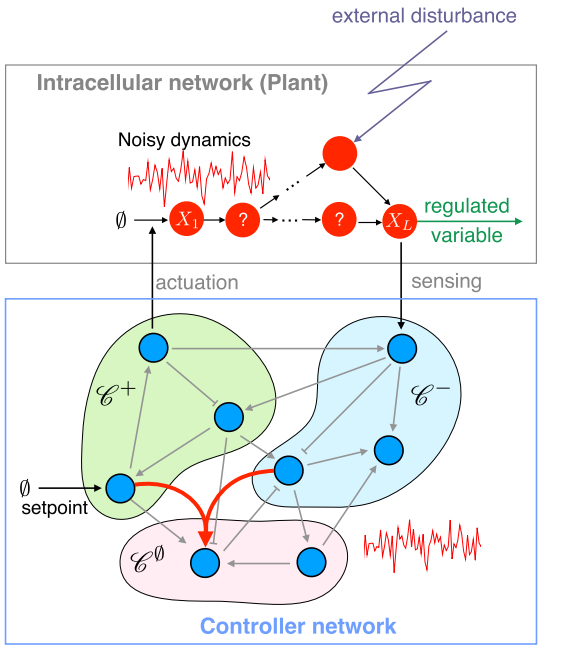

Fig. 13. ***When the dynamics are stochastic, every controller that achieves RPA must be of the antithetic type. All such controllers have the topology shown in the figure, whereby the species can be partitioned into three distinct molecular classes: "positively charged," "negatively charged," or "neutrally charged" species. Furthermore, a chemical reaction must exist that takes a positively charged species and a negatively charged species into a neutrally charged product species. The latter may be the empty set, in which case the reaction is as an annihilation reaction, e.g., $Z_1 + Z_2 \longrightarrow \emptyset$.***

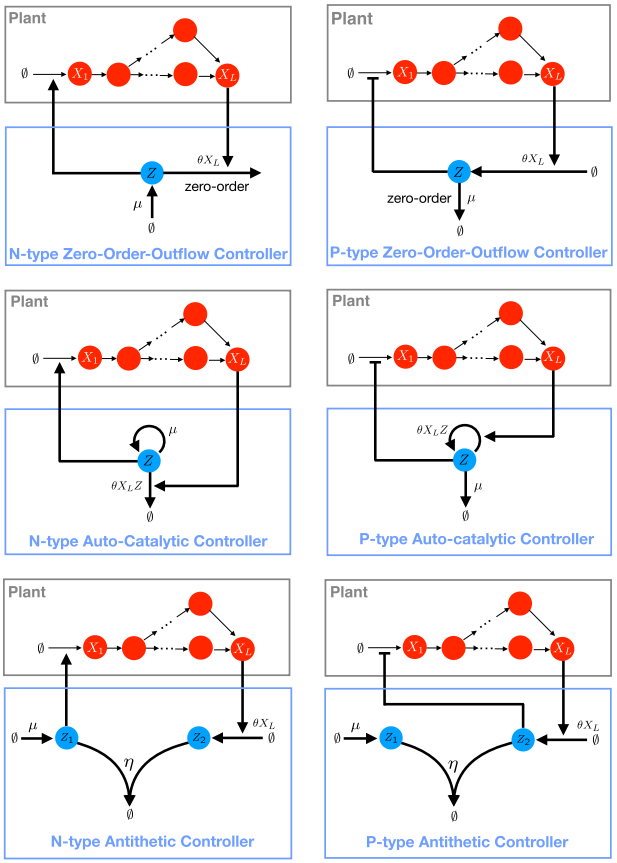

Fig. 14. ***Several biochemical reaction motifs that achieve integral feedback. Provided a positive fixed point at which the closed loop is stable exists, the interconnection will achieve RPA with the output always converging to μ/θ in the limit. External constant disturbances to the plant as well perturbations to its dynamics that do not change the stability of the loop will be rejected perfectly at the output. Two classes of controllers are shown, depending on the sign of their gain. N-type controllers have negative gain so that the actuating variable moves in the opposite direction of the sensed variable at a steady state. P-type controllers have positive gains.***

state gain. Stated differently, if an increase in the input of the controller results in an increase in the abundance of the controller's output species (actuating species), I consider this to be a P-type; otherwise, if the actuating species abundance decreases, then the controller is designated N-type. The sign of the gain of the controller is relevant to the stability of the closed-loop network, as one would expect that an overall negative loop gain is needed for stability. For example, if the plant has a positive gain then the feedback controller needs to be N-type to achieve stability. However, a P-type can also be used in this case if the controller output acts as a repressor on the plant input, thereby ensuring that the loop has negative gain [see Fig. 14 (right column)]. The integral motifs collected here have been developed independently or discovered in various natural systems. In [97], a systematic approach for generating integral variables that are realizable with CRNs was developed. I now show why these controllers provide integral action and achieve RPA in the deterministic setting. It goes without saying that the plant must have a positive fixed point at the desired reference values. I also assume that the closed-loop system is stable when the integrator is interconnected with μ and θ chosen so that μ/θ equals the desired reference values. If this were not the case, feedback in addition to the integrator should be introduced to ensure stability, and I shall assume that such feedback has already been absorbed into the plant dynamics. I will focus on the P-type zero-order-outflow controller; the others follow similarly. Here, the controller has two reactions

$$\mathbf{X}_L \xrightarrow{\theta} \mathbf{X}_L + \mathbf{Z}$$
$$\mathbf{Z} \xrightarrow{\mu} \emptyset \quad \text{(zero-order)}.$$

The dynamic equations can be expressed as follows:

$$\dot{Z} = \theta X_L - \mu$$

which clearly shows that Z acts as an integrator. Indeed,

$$Z(t) = \int_{-\infty}^{t} (\theta X_L(s) - \mu)\, ds$$

which, in turn, implies that $X_L(t) \longrightarrow \mu/\theta$ as $t \longrightarrow \infty$, as desired. The zero-order kinetics reaction is an idealization that breaks down for small Z, as it assumes that the rate of removal of $\mathbf{Z}$ is constant regardless of its concentration, Z—an assumption that can lead to negative Z. Instead, the removal of $\mathbf{Z}$ is usually accomplished with an enzyme following Michaelis–Menten kinetics, as follows:

$$\mathbf{Z} \xrightarrow{\frac{\mu}{k+Z}} \emptyset$$

so that the corresponding ODE becomes

$$\dot{Z} = \theta X_L - \frac{\mu Z}{k+Z}.$$

When the concentration of $\mathbf{Z}$ is large relative to k, the Z dynamics approximates the zero-order dynamics and, hence, the ideal action of the zero-order-outflow integral controller.

The reactions for the P-type autocatalytic controller are

$$\mathbf{X}_L + \mathbf{Z} \xrightarrow{\theta} 2\mathbf{Z}$$
$$\mathbf{Z} \xrightarrow{\mu} \emptyset.$$

Following the law of mass-action, the dynamics of the species concentrations can be expressed as follows:

$$\dot{Z} = \theta X_L Z - \mu Z.$$

As long as the closed-loop has no stable fixed point with zero Z-component, this will deliver integral action, and at a steady state, $X_L = \mu/\theta$.

In this section and the one before it, I have asserted that the antithetic feedback motif is sufficient for realizing integral feedback in both deterministic and stochastic networks, but it is only necessary for stochastic networks. While we have proved this rigorously [94], here, I will attempt to give an intuitive explanation for this. In molecular circuits that use molecular concentrations/abundances as state variables, these variables are restricted to remain nonnegative. This is in contrast to electric circuits, where voltages and currents can assume positive and negative values. Inevitably, negative quantities can be represented as the difference between two positive quantities. The antithetic reaction $\mathbf{Z}_1 + \mathbf{Z}_2 \longrightarrow \emptyset$ acts like a differencing device. Indeed, if each of the species $\mathbf{Z}_i$ is being produced at a rate $\alpha_i(t)$, the difference in concentration of the two species, $Z_1(t) - Z_2(t)$, is a proxy for the integral of $\alpha_1(t) - \alpha_2(t)$, which can be either positive or negative. Importantly, this continues to hold even if one of the species' concentrations goes to zero. Keeping this in mind, let us now look at the other two mechanisms for integration used in the deterministic setting: zero order and autocatalytic. In the zero-order mechanism, a species $\mathbf{Z}$ is being created at a rate $\alpha(t)$ and degraded at a rate $\gamma(t)$. Thus, the concentration of that species, $Z(t)$, is the integral of $\alpha(t) - \gamma(t)$. Note that this arrangement only allows positive integral values, as the concentration cannot go negative. Even if it can be arranged in the deterministic setting (e.g., through dynamic stabilization) that $\alpha(t) - \gamma(t)$ does not remain negative long enough for the integral, and hence $Z(t)$, to go negative, this can never be achieved in the stochastic setting. Instead, the population will go negative with probability one for each trajectory owing to the inevitable fluctuations inherent in the stochastic process. In the autocatalytic mechanism, the production rate of $\mathbf{Z}$ is $\alpha(t)Z(t)$ (positive feedback), while its degradation is $\gamma(t)Z(t)$. Hence, $Z(t)$ is the integral of $Z(t)(\alpha(t) - \gamma(t))$. In this setting, the population can never go negative and will always stay positive in the deterministic system. In contrast, in the stochastic setting, the population of $\mathbf{Z}$ for each trajectory will go to zero and stay there with probability one, breaking the integrator.

In biology as in engineering, designing integral feedback controllers necessarily involves design tradeoffs. These were studied in [98]–[100], where tools from control theory, such as Bode's integral formula, were employed to shed light on the fundamental limitations in closed-loop performance for molecular control systems that use the antithetic integral feedback controllers alone. These studies also highlighted the inevitable design tradeoffs between competing performance requirements, such as tracking error, response speed, robustness, leakiness, and stability.

5) PID Controllers: While RPA through integral feedback is an important and highly desirable feature in the regulatory synthetic circuit, it is not the only one. Good transient response and noise reduction are features that can be added on top of RPA to further enhance the performance of a closed-loop system. As shown in [99] and [100], there are, in fact, fundamental limits to what can be achieved with antithetic integral feedback alone. In [98], it was shown that augmenting the antithetic integral feedback with proportional feedback improves the dynamic performance of the circuit and results in reduced noise variance. In [101] and [102], molecular PID controllers were proposed and analyzed. These studies showed that such controllers can enhance the transient response and reduce the effect of cellular noise while still achieving RPA. In [102], a hierarchy of molecular PID controllers of different degrees of complexity was presented and analyzed. Fig. 15 shows a fourth-order molecular PID design studied in [102].

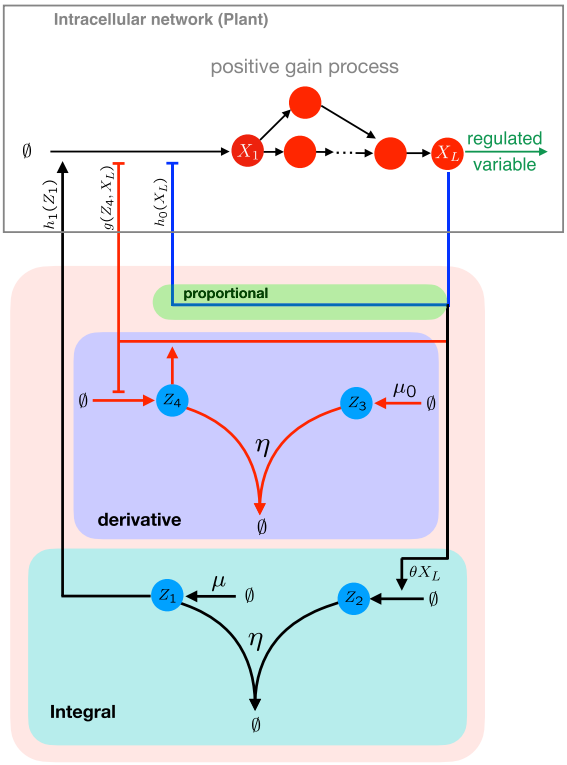

Fig. 15. ***Molecular PID controller. When an integrator is placed in the feedback loop of the unity system in the forward loop, a filtered derivative can be realized. This can be augmented to a PI feedback system to yield a PID controller with no constraints on the PID parameters. See [102] for details on this and other PID designs.***

6) Synthetic Implementations: Ever since my student and I discovered in the late 1990s that integral feedback is the main mechanism for calcium homeostasis in mammals [91], I have been intrigued with the possibility of synthetically engineering integral controllers in living cells using molecular interactions. The opportunity presented itself when I started my own wet lab at ETH Zürich in

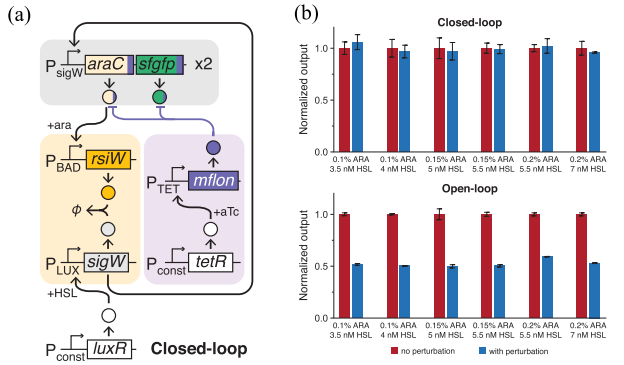

Fig. 16. ***Antithetic integral feedback circuit in E. coli. The first synthetic realization of an integral feedback controller in a living cell [94], [103]. (a) Pair sigma/antisigma factors, SigW and rsiW, which bind tightly realize the antithetic pair Z_1 and Z_2. The expression of the sigma factor SigW is controlled by the external inducer HSL, which modulates the activity of the activator luxR. The expressed sigW controls the production of the regulated output, the protein araC. To quantify the abundance of araC, a fluorescent protein is also expressed at the same rate. Being itself an activator, araC controls the expression of the antisigma factor rsiW, which, in turn, binds SigW closing the feedback loop. A test circuit was also engineered to introduce a persistent disturbance in the form of a protease that degrades the regulated output along with its fluorescent reporter. (b) Steady-state value of the regulated output of the closed-loop system remains unchanged after increasing the rate of output degradation by the protease. In sharp contrast, in the open-loop circuit, the protease had a big effect on the output of interest, degrading it to the point where its steady-state value was about half its value before the protease perturbation was introduced.***

2011. In our first attempt, we set out to reconstitute the natural calcium homeostasis integrator, which is a zero-order-outflow integrator (see Fig. 14), but we failed. As discussed in [104] where the physical constraints on biomolecular integrator design were studied, the engineering of genetic integral feedback controllers faces several challenges. Synthetic realizations of zero-order integrators are confronted with difficulties related to the stringent tuning requirements needed to implement zero-order kinetic reactions in the desired dynamic range. At some point in 2014, we thought that we had a functioning integral controller, but, upon further analysis, we discovered that we were seeing an artifact of crosstalk that was acting in a feedforward fashion to reject just the right amount of disturbance to make it appear as though our feedback controller was achieving perfect adaptation—it was not. The problem resisted a good solution, and frustration was building up.

In the meantime, theoretical studies in my group resulted in a very different mechanism for integral feedback, which appeared to have the added feature of being resistant to noise [88]. This motif, which I called antithetic feedback (see Fig. 16), presented a blueprint for integral feedback implementations that did not rely on zero-order kinetics. However, new challenges needed to be overcome: we needed to engineer molecular species that realize the function of the antithetic pair $\mathbf{Z}_1$ and $\mathbf{Z}_2$. In particular, $\mathbf{Z}_1$ had to function as an activator (or repressor depending on the design), and together, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ should annihilate, inactivate, or sequester each other in a 1:1 stoichiometric ratio. Equally important, if the engineered integrator were to be minimally leaky, these two molecular species should have very low degradation rates in their free form. This would eliminate the steady-state error that could result from leaky integration. Moreover, degradation is not the only way to get leaky integrators. Fast cell growth causes dilution of all cellular species and has an effect on steady-state error similar to degradation [94], [104], [105]. These effects can be ameliorated by increasing the gain of the feedback loop through suitable parameter choices or through additional circuitry [66], [94], [105], [106].

In bacteria, mRNA typically has a short half-life (in the order of minutes), making it much less suitable than proteins for realizing the controller species $\mathbf{Z}_1$ and $\mathbf{Z}_2$. In contrast, proteins are far more stable, so it made good sense to use proteins. Our search for suitable protein pairs settled on two complementary proteins, called sigma/antisigma factors [88], [107], which can be found in the bacterium *Bacillus Subtilis*. SigW is a sigma factor that functions as a transcription factor capable of activating gene expression when not bound to its antisigma factor RsiW (see Fig. 16). Thus, the two protein pairs make ideal candidates for realizing the reaction $\mathbf{Z}_1 + \mathbf{Z}_2 \longrightarrow \emptyset$. Using this protein pair, we, finally, had a functioning synthetic integrator that achieved RPA with experimentally undetectable steady-state errors in response to step disturbances. We reported these results in bioRxiv in 2017 [103]. To the best of our knowledge, this was the first synthetic implementation of a biomolecular integral feedback controller in a living cell. We then demonstrated the experimental utility of our integral circuit by using it to achieve robust growth control. These results, together with rigorous proof of the universality of the antithetic motif in the presence of stochastic dynamics, appeared in the journal *Nature* in 2019 [94].

Another circuit implementation used small RNA (sRNA) in *E. coli* to implement the sequestration reaction in the antithetic motif [56]. Since mRNA has a significantly shorter half-life than protein, small steady-state errors could not be achieved without using high gain to counteract the impact of leaky integration, resulting in "quasi-integral" feedback. Yet, another circuit used sigma factors to realize the antithetic integral motif in a cell-free system, providing an *in vitro* demonstration of robust adaptation using sigma factors [108].

Mammalian implementations of integral feedback controllers are expected to have significant uses in mammalian applications, such as in precise drug delivery. In 2020, we reported on the design of a genetic antithetic integral feedback controller in mammalian cells [57]. The controlled system exhibited robustness to disturbances

in the form of a drug and to dynamic perturbations that changed the plant's topology. Proportional feedback was subsequently added to this circuit and shown to retain the adaptation properties and reduce the output variance [109].

Finally, I mention that the sequestration/annihilation reaction that is at the heart of the antithetic motif has been used in synthetic circuits for purposes other than realizing integral feedback. For example, it was proposed for circuit dose-response linearization [46], latch design [107], concentration tracking [110], oscillations and bistability [111], and as a comparator using sigma factors [112], among other uses.

D. Control of Cell Consortia

Thus far, the network to be controlled and the controller network resided in the same cell. Another approach is to separate the two networks such that each resides in a different cell strain. Thus, the cells in a population are divided into "plant" cells and "controller cells." The two cell types communicate by excreting molecules that diffuse in the medium where the cells grow and are subsequently sensed by the cells of the other cell and often by the same cell type as well. This enables the cells to close the feedback loop (see Fig. 17). One advantage of this strategy is to minimize the metabolic burden on individual cells, as the closed-loop circuitry is divided between controller cells and plant cells. The prospect of achieving feedback regulation in multicellular consortia was analyzed computationally in [113] and [114]. In [113], a computer analysis of the use of antithetic feedback motif using sigma/antisigma pair showed the feasibility of regulating the population on an output species of interest. The controller part of the circuit was subsequently tested experimentally [112], while the entire closed-loop system was tested using a hybrid cell/computer interface in [115]. The full closed-loop control using two cell populations has not yet been achieved. In contrast, the regulation objective in [114] was to explore the use of toxin/antitoxin pairs to regulate the total cellular population and the relative ratio between cell strains—an objective that was successfully demonstrated using computational modeling. Subsequent work demonstrated the role of interaction network topology in bacterial population control [116]. Early experiments have demonstrated limited control over cell density and composition achieved with narrow robustness margins, prompting further analysis and improved design [117].

More generally, one need not think of the two cell populations shown in Fig. 17 as plants and controllers but as agents that dynamically interact through diffusing chemical signals. From this perspective, tasks other than regulation can be achieved. These include genetic oscillation generation [118], robust coexistence [119], creation of spatiotemporal dynamics [120], enhancing genetic stability [121], and majority sensing [122].

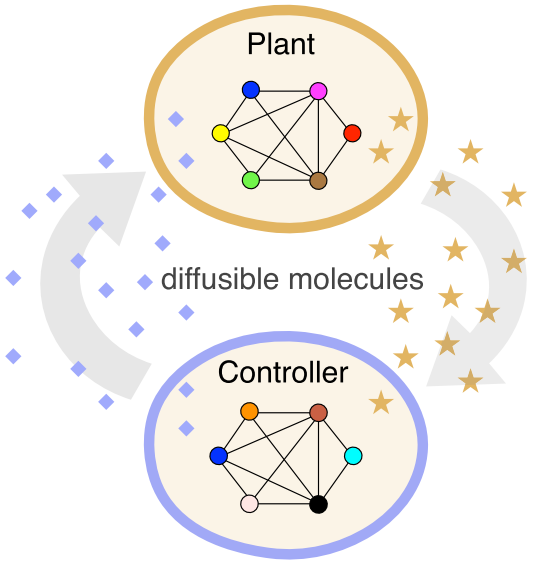

Fig. 17. *How multicellular control is realized. Two populations of cells are grown together with each population secreting diffusible molecules that can be received and interpreted by the other population. In this way, a feedback loop can be established among the populations. This facilitates the regulation of the sizes of the different populations and their respective behavior.*

VIII. INTERFACING CELLS WITH COMPUTERS: *IN SILICO* CONTROL OF GENETIC CIRCUITS

Instead of realizing controllers in the living cell using cellular components, an alternative approach is to interface the cells with digital computers where the control algorithm is executed and used to drive cellular inputs to close the feedback loop. For this to work, one needs the means to measure the controlled quantity in the living cell in real time, as well as ways to externally actuate cellular processes that can drive the output of interest. The technology for doing both has become available in the last decade of the 20th century. The first was the discovery of fluorescent protein and its synthetic expression, and the use of a fluorescent biomarker, which gives a direct way to measure protein abundance in real time [123]. The second was the development of optogenetic methods, which enabled the engineering of cells that are light-responsive, whereby the light drives cellular processes, such as activation of ion channels [124] or gene expression [125], which, in turn, drives the controlled variable of interest.

A computer-controlled cellular system functions by measuring the output of interest, e.g., through instruments that measure the fluorescence marker of the variable of interest, and then passing these measurements to a computer where the controller is running in real time. The output of the controller, the control input, is converted into an electric signal that drives a light source (e.g., LED) that shines light onto the controlled cells, which have been engineered to be light-sensitive (see Fig. 18). Depending on where the cells

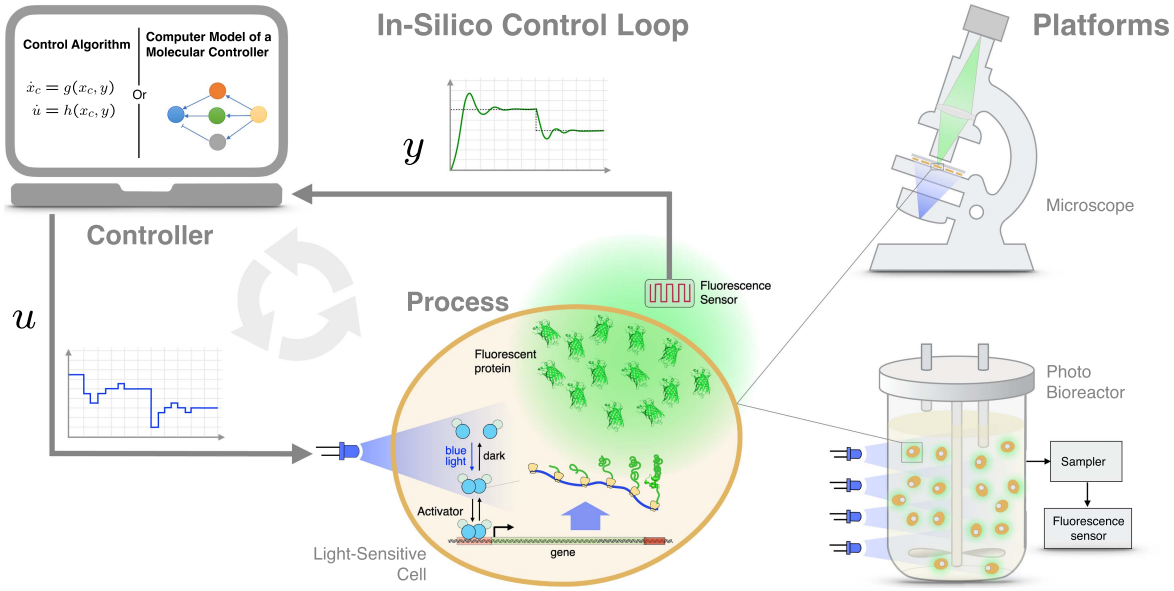

Fig. 18. ***Computer control of living cells. A feedback loop can be established to control living cells in real time by a digital computer. The input to the cell can either be light that is delivered by a laser or an LED, or a chemical inducer delivered by a microfluidic device (not shown). The measured output is typically fluorescence whose intensity is indicative of the controlled variable, e.g., protein concentration. The cells are either controlled in bulk in a reaction compartment (e.g., a test tube or a bioreactor) or under the microscope, which enables their independent control.***

are kept during the control experiment, there are generally two different approaches for cellular control. In the first approach, the cells are grown in liquid culture in a test tube or a bioreactor. In this case, light input is delivered to the entire cell population, and measurement samples represent the statistics of the entire population, e.g., average protein level. In this approach, effective feedback control of the cell population, but not single cells, is possible (see Fig. 18). In the second approach, the cells to be controlled are kept under a fluorescent microscope that continuously measures the fluorescence of individual cells in the field of view. In this scenario, micromirror devices can be designed to shine light precisely and independently on different cells. This makes possible the simultaneous and independent closed-loop control of a large number of single cells.

In a microscopy platform, such as the one described above, the cells are typically grown in microfluidics devices that deliver the nutrients throughout the experiment. Since microfluidics technologies enable the controlled flow of fluids over cell populations, it is possible to use small chemical inducers to actuate the controlled cells, instead of light. In this way, cells that can respond to small chemical inducers can be more readily controlled in feedback, as the engineering of a light actuator is not required. However, due to the limited spatial resolution by which chemical inducers can be delivered, independent single-cell control becomes more difficult with chemical inducer control, and one is limited to population control.

Computer control of living cells is how I personally got involved in experimental research in biology. After a sabbatical at ETH Zürich with the group of J. Lygeros during which we designed and tested by simulation a hypothetical population-level control system, I proposed that we work with a former student of mine (H. El Samad) who had started an experimental lab at UCSF to build this closed-loop system in the lab. We found the beautiful paper of Shimizu-Sato *et al.* [125] who had developed a light-induced gene expression system, which seemed perfect for closed-loop control. Initial closed-loop experiments were implemented with computers in Zurich controlling, in real time, a population of yeast cells grown in San Francisco! Further improvements led to the first successful computer control of gene expression. The work was published in 2011 in [126] and quickly became a media sensation (see [127]).

Around the same time in a neighboring lab, another group showed that light can be used to control intracellular signal dynamics [128]. Subsequent work demonstrated the use of microfluidics-based control of gene expression in yeast cells [129]. In [130], it was shown that light-inducible gene expression can be used to guide iterative experiments that aim to optimally characterize the dynamics of the underlying gene system. In 2016, we demonstrated that precise and robust population control can be achieved using either PI or model predictive control, and we applied this to controlling growth in *E. coli* populations [131]. Meanwhile, studies using microfluidics-based closed-loop control work explored several different control strategies in yeast [132]. Feedback control was also used to control a population of *E. coli* cells around an unstable equilibrium of genetic toggle switch circuit [133].

Using digital micromirror devices (DMDs) that use the same technology as modern projectors, it was possible to use real-time optogenetic feedback to control the gene expression of individual *E. coli* cells [134] and the mRNA level of individual yeast cells [135]. In the latter study, it was demonstrated that even highly stochastic and bursty transcription can be effectively regulated with feedback control. In [136] and [137], computer control was applied to mammalian cells, where it was shown that gene expression and signaling pathways can be regulated using microfluidics-based control.

One of the exciting applications of computer control of living cells is their potential use for getting a deeper understanding of the underlying biology. In [138], a closed-loop optogenetic compensation (CLOC) strategy used the output of a pathway deleted for a feedback regulator to deliver a dynamic light-enabled transcriptional input designed to compensate for the effects of the feedback deletion, thereby shedding light on the dynamics of the natural feedback regulator that it replaces. In [139], the ability to optogenetically control different cells independently allowed the emulation of cell-to-cell signaling, whereby the activation of each cell depends on the state of its neighbors. This cell-in-the-loop concept was then used to emulate a mutual inhibition communication strategy that led to the reconstitution of checkerboard-like cellular patterning commonly seen in natural development. Another example concerns the study of the cell cycle in budding yeast [140] through the forced synchronization of multiple cells via microfluidics-based closed-loop feedback control.

Finally, a recent study from my lab showed how a single-cell feedback control platform can be used to study the performance of newly designed biomolecular controller candidates [141]. This is achieved by feeding cellular outputs to stochastic biochemical network models of these controllers in a computer and using the simulated outputs to drive in real-time the dynamics of the living cells to be controlled. This closed-loop platform, called the *Cyberloop*, enables the rapid prototyping and tuning of novel biomolecular control systems before they are genetically engineered into the living cell–a much more resource- and time-demanding process.

More detailed reviews of computer control of living cells and their applications can be found in [142]–[149].

IX. APPLICATIONS OF GENETIC CONTROL SYSTEMS

1) Better Tools for Probing and Engineering Biology: Whether one is interested in understanding biological networks through basic research or engineering new ones, having robust genetic tools with favorable characteristics is highly desirable. As a simple example, inducing gene expression is a standard step in the modern biology research toolkit, allowing a researcher to modulate the concentration of a certain endogenous protein to study its effect on the cell. Similarly, when engineering biology, gene expression is commonly a part of the designed module or system. Having gene expression systems with good characteristics may be the key to a successful implementation. Such characteristics include linear induction over a wide dynamic range, fast response, precise expression, robustness to sloppy uncertain parts, robustness to gene copy number variability, robustness to growth conditions, low burden, noise resistance, reduced cell-to-cell variability, and so on. We have seen that, by adding circuitry for genetic control, these features can be realized, leading to high-performance gene induction systems. The cost of these improvements is more complex circuitry. As with electrical circuits, in many applications, this may be a small price to pay for the realized improvements.

2) Paradigm for Exploring Natural Regulation Motifs in Biology: Living cells are very complex. Much of this complexity is due to the demands of regulation. Indeed, regulation is a running theme throughout biology, and cells use exquisite mechanisms to achieve effective regulation, which is essential for cellular homeostasis and survival. As a result, much of biological research concerns dissecting this regulation by learning its mechanisms, discovering its common motifs, and understanding its impact on cellular function. To reverse engineer the natural regulatory processes and their function, a researcher must contend with nonlinear dynamic processes operating in the messy environment of the cell—one that is replete with noise, crosstalk, and uncertainty. How would a researcher test a proposed mechanism of regulation, ascertain its plausibility, if not confirm its presence? After all, many regulatory mechanisms that work well in a computer simulation may not be viable in the environment of the cell simply because a computer model can never account for all the factors. An attractive alternative is to build proposed regulatory circuits in the complex environment of the cell and to test their function in a realistic setting [150]–[152]. Synthetic biology and control theory come together to test regulation hypotheses and propose new ones, allowing us to get a deeper understanding of biological regulation and its complexity in the cellular environment. This is in keeping with Richard Feynman's dictum, "what I cannot create, I do not understand."

3) Industrial Biotechnology: Living cells, including bacteria, yeasts, and mammalian cells, are exploited in modern biotechnology as factories to produce useful chemicals, metabolites (products and intermediates of cellular metabolisms, such as biofuels, vitamins, and antibiotics), and biologics (large molecule medicines, such as vaccines, allergenics, and antibodies). In metabolic engineering, the natural metabolic pathways of these cells is genetically modified or rewired to direct the cell to synthesize these products. The optimal production and yield of these bioproducts require good regulation of growth rate and production speed. It also requires tight control of various enzyme and metabolite levels in the metabolic pathway: too much production is at best wasteful but often leads to toxicity for the host cell and/or degradation in the final product (e.g., misfolded proteins) [154], [155]. Genetic control systems are being used to achieve better regulation

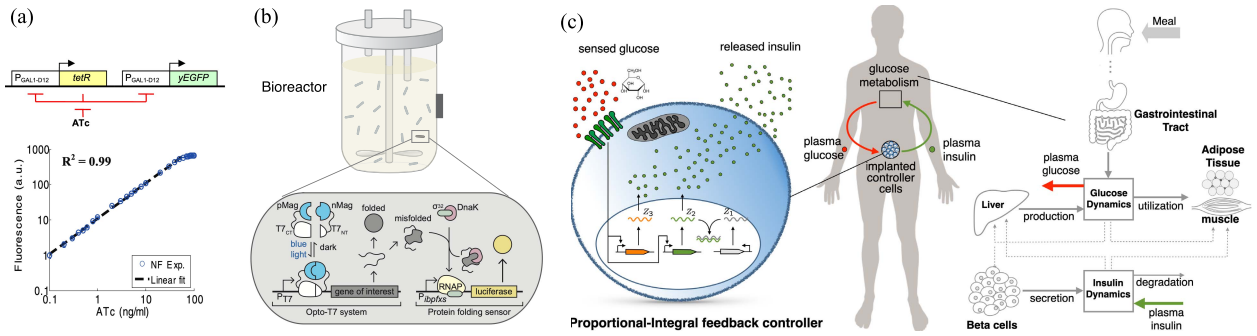

Fig. 19. ***Some areas of application of engineered genetic control system. (a) Synthetic gene control circuits can be used to make better genetic devices, as can be seen from the linearized response [46] of an induced gene. Genetic control systems can be used in industrial biotechnology to enhance the efficiency and yield of bioproducts, such as chemicals and biopharmaceuticals. (b) One example is where the level of misfolded proteins in the cell is sensed, and the corresponding signal is used to adjust the expression level of the bioproduct, keeping it high enough for a good yield, but not too high, so as to induce significant misfolding. (c) Example of using genetic controllers for human therapy. Here, the controller interacts with the body's glucose metabolism, secreting insulin in response to high glucose and restoring homeostasis. Using proportional feedback, engineered cells using this concept were shown to treat hypoglycemia in diabetic mice [61], [153]. In [109], an antithetic PI genetic feedback controller was shown to achieve much improved performance and perfect set-point tracking in a simulation study that utilized an FDA approved computer model of type I diabetes.***

and, hence, increase the product yield. Fig. 19(b) shows a genetic negative feedback controller that reduces the production rate of a protein of interest in response to stress induced by an increase in the level of misfolded proteins. Another direction where genetic controllers are expected to play an important role is maintaining microbial communities, whereby different cell types communicate with each other through small molecules and jointly regulate their growth rates to maintain a certain cell-type population ratio—a problem referred to as ratiometric control [114], [156], [157]. Genetic regulation in industrial biotechnology has become an active area of research with huge promise (see [158]–[172] and the references therein).

4) Medical Therapy: One of the most exciting applications of genetic control systems is in the area of biotherapeutics. One promising direction involves the genetic engineering of 'smart' human cells that, when implanted into the body of a patient with a chronic regulatory disease, sense the disease state and respond by releasing therapeutic agents in a closed-loop fashion, thereby automatically managing the disease and bringing the system back into homeostasis. One example of such cell therapy is the treatment of Type I diabetes [61], where cells encapsulated in alginate beads and implanted in a diabetic mouse were engineered to detect glucose levels and activate the expression of insulin in proportion to the sensed glucose. The insulin is released into the blood to downregulate glucose and restore homeostasis. This cell therapy concept is shown in Fig. 19(c). Although this idea has not yet been applied in human therapy, it has been shown to work in mice using an engineered proportional feedback system. In [109], a genetic PI feedback controller was proposed to improve the performance of the controller by robustly bringing the glucose level to a normal set-point value for all subjects [see Fig. 19(c)]. The same idea for cell therapy can be applied to other metabolic diseases, such as gout [173]. Controlling chronic metabolic diseases is just one area of application of genetic control systems in medicine. Others include immunotherapy, stem cell differentiation, and tissue engineering.

X. FUTURE OUTLOOK

Synthetic biology is a transformative field of research that enables the engineering of living systems from natural and designed parts with applications in science, human health, industrial production, and environmental remediation. Up until now, the design of synthetic gene circuits has followed similar design principles that electrical engineers have successfully used to make electronic devices.

Synthetic biology is first and foremost an engineering discipline. As such, ideas and principles developed in engineering, particularly circuit design in electrical engineering, have had an outsize impact on this field since its infancy. Yet, genetic circuits differ from electric circuits in important ways. Being hosted in living cells, genetic circuits must contend with molecular cellular noise, cellular growth, cell division, biological mutations, and a large amount of inevitable coupling with the natural circuits of the host. These and other factors introduce novel challenges to genetic circuit designers [174]. While some of the robustness principles of control theory that I have described in this article have certainly helped in understanding and mediating some of these challenges, addressing the many remaining ones will require the creation of novel theoretical and practical tools that are customized for the unique environment of the living cell. In other words, we need a control theory for living systems. In addition to such a theory, I believe that an information theory and a signal-processing theory for living systems will also be required for establishing a sound foundation for synthetic biology and systems biology. Such theories must not be a superficial adaptation of what has already been developed for man-made systems. Instead, they need

to genuinely account for the many characteristics peculiar to living cells. Furthermore, theoretical developments should take place in close cooperation with synthetic biology practitioners whose challenging problems should motivate the theoretical formulations and who will then test proposed solutions.

The future of synthetic biology is full of potential, with many challenges laying ahead before this potential is fully realized. This presents many exciting opportunities for electrical engineers to work at the interface of synthetic biology and electrical engineering and to be part of what promises to be one of the most important scientific and engineering developments in the 21st century. ■

Acknowledgment

The author would like to acknowledge Stephanie Aoki, Maurice Filo, Ching-Hsiang Chang, Timothy Frei, and Ankit Gupta for their comments on the manuscript draft.

REFERENCES

[1] N. Wiener *et al.*, *Cybernetics*. New York, NY, USA: JSTOR, 1948.

[2] V. Noireaux and A. P. Liu, "The new age of cell-free biology," *Annu. Rev. Biomed. Eng.*, vol. 22, no. 1, pp. 51–77, Jun. 2020.

[3] D. D. Vecchio, A. J. Dy, and Y. Qian, "Control theory meets synthetic biology," *J. Roy. Soc. Interface*, vol. 13, no. 120, Jul. 2016, Art. no. 20160380.

[4] V. Hsiao, A. Swaminathan, and R. M. Murray, "Control theory for synthetic biology: Recent advances in system characterization, control design, and controller implementation for synthetic biology," *IEEE Control Syst. Mag.*, vol. 38, no. 3, pp. 32–62, Jun. 2018.

[5] M. Khammash, "An engineering viewpoint on biological robustness," *BMC Biol.*, vol. 14, no. 1, pp. 1–11, Dec. 2016.

[6] C. M. O'Connor, J. U. Adams, and J. Fairman, *Essentials of Cell Biology*, vol. 1. Cambridge, MA, USA: NPG Education, 2010, p. 54.

[7] B. Alberts *et al.*, *Essential Cell Biology*. New York, NY, USA: Garland Science, 2015.

[8] R. Milo and R. Phillips, *Cell Biology by the Numbers*. New York, NY, USA: Garland Science, 2015.

[9] D. A. Jackson, R. H. Symons, and P. Berg, "Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*," *Proc. Nat. Acad. Sci. USA*, vol. 69, no. 10, pp. 2904–2909, Oct. 1972.

[10] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, and R. B. Helling, "Construction of biologically functional bacterial plasmids *in vitro*," *Proc. Nat. Acad. Sci. USA*, vol. 70, no. 11, pp. 3240–3244, Nov. 1973.

[11] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, Jan. 2000.

[12] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in *Escherichia coli*," *Nature*, vol. 403, no. 6767, pp. 339–342, Jan. 2000.

[13] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, no. 6786, p. 590, 2000.

[14] R. Bertram and W. Hillen, "The application of Tet repressor in prokaryotic gene regulation and expression," *Microbial Biotechnol.*, vol. 1, no. 1, pp. 2–16, Jan. 2008.

[15] F. Horn and R. Jackson, "General mass action kinetics," *Arch. Rational Mech. Anal.*, vol. 47, no. 2, pp. 81–116, Jan. 1972.

[16] P. Atkins and J. De Paula, *Physical Chemistry for the Life Sciences*. London, U.K.: Oxford Univ. Press, 2011.

[17] B. P. Ingalls, *Mathematical Modeling in Systems Biology: An Introduction*. Cambridge, MA, USA: MIT Press, 2013.

[18] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, 1977.

[19] H. H. McAdams and A. Arkin, "Stochastic mechanisms in gene expression," *Proc. Nat. Acad. Sci. USA*, vol. 94, no. 3, pp. 814–819, Feb. 1997.

[20] A. Raj and A. van Oudenaarden, "Nature, nurture, or chance: Stochastic gene expression and its consequences," *Cell*, vol. 135, no. 2, pp. 216–226, Oct. 2008.

[21] J. M. Pedraza and A. van Oudenaarden, "Noise propagation in gene networks," *Science*, vol. 307, no. 5717, pp. 1965–1969, Mar. 2005.

[22] J. Paulsson, "Summing up the noise in gene networks," *Nature*, vol. 427, no. 6973, pp. 415–418, Jan. 2004.

[23] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, no. 5584, pp. 1183–1186, Aug. 2002.

[24] P. S. Swain, M. B. Elowitz, and E. D. Siggia, "Intrinsic and extrinsic contributions to stochasticity in gene expression," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 20, pp. 12795–12800, Oct. 2002.

[25] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *J. Comput. Phys.*, vol. 22, no. 4, pp. 403–434, Dec. 1976.

[26] D. J. Wilkinson, *Stochastic Modelling for Systems Biology*. London, U.K.: Chapman & Hall, 2018.

[27] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, vol. 282. Hoboken, NJ, USA: Wiley, 2009.

[28] D. F. Anderson, "A modified next reaction method for simulating chemical systems with time dependent propensities and delays," *J. Chem. Phys.*, vol. 127, no. 21, Dec. 2007, Art. no. 214107.

[29] D. T. Gillespie, "A rigorous derivation of the chemical master equation," *Phys. A, Stat. Mech. Appl.*, vol. 188, nos. 1–3, pp. 404–425, Sep. 1992.

[30] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *J. Chem. Phys.*, vol. 124, no. 4, Jan. 2006, Art. no. 044104.

[31] A. Gupta, J. Mikelson, and M. Khammash, "A finite state projection algorithm for the stationary solution of the chemical master equation," *J. Chem. Phys.*, vol. 147, no. 15, Oct. 2017, Art. no. 154101.

[32] C. A. Gómez-Uribe and G. C. Verghese, "Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations," *J. Chem. Phys.*, vol. 126, no. 2, Jan. 2007, Art. no. 024109.

[33] A. Singh and J. P. Hespanha, "Approximate moment dynamics for chemically reacting systems," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 414–418, Feb. 2011.

[34] D. Schnoerr, G. Sanguinetti, and R. Grima, "Comparison of different moment-closure approximations for stochastic chemical kinetics," *J. Chem. Phys.*, vol. 143, no. 18, Nov. 2015, Art. no. 185101.

[35] M. Khammash, "Modeling and analysis of stochastic biochemical networks," in *Control Theory and Systems Biology*, P. A. Iglesias and B. P. Ingalls, Eds. Cambridge, MA, USA: MIT Press, 2010.

[36] M. Csete and J. Doyle, "Bow ties, metabolism and disease," *Trends Biotechnol.*, vol. 22, no. 9, pp. 446–450, Sep. 2004.

[37] M. E. Csete and J. C. Doyle, "Reverse engineering of biological complexity," *Science*, vol. 295, no. 5560, pp. 1664–1669, Mar. 2002.

[38] M. Thattai and A. Van Oudenaarden, "Intrinsic noise in gene regulatory networks," *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 15, pp. 8614–8619, Jul. 2001.

[39] S. Cardinale and A. P. Arkin, "Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems," *Biotechnol. J.*, vol. 7, no. 7, pp. 856–866, Jul. 2012.

[40] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. D. Vecchio, "Resource competition shapes the response of genetic circuits," *ACS Synth. Biol.*, vol. 6, pp. 1263–1272, Apr. 2017.

[41] T. Frei *et al.*, "Characterization and mitigation of gene expression burden in mammalian cells," *Nature Commun.*, vol. 11, no. 1, pp. 1–14, Dec. 2020.

[42] R. D. Jones *et al.*, "An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells," *Nature Commun.*, vol. 11, no. 1, pp. 1–16, Dec. 2020.

[43] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag, "Modular cell biology: Retroactivity and insulation," *Mol. Syst. Biol.*, vol. 4, no. 1, p. 161, 2008.

[44] D. Mishra, P. M. Rivera, A. Lin, D. D. Vecchio, and R. Weiss, "A load driver device for engineering modularity in biological networks," *Nature Biotechnol.*, vol. 32, no. 12, pp. 1268–1275, Dec. 2014.

[45] H.-I. Son, A. Weiss, and L. You, "Design patterns for engineering genetic stability," *Current Opinion Biomed. Eng.*, vol. 19, Sep. 2021, Art. no. 100297.

[46] D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josić, and G. Balázsi, "Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 13, pp. 5123–5128, Mar. 2009.

[47] A. S. Sedra, D. E. A. S. Sedra, and K. C. Smith, *Microelectronic Circuits*. New York, NY, USA: Oxford Univ. Press, 1998.

[48] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Boca Raton, FL, USA: CRC Press, 2019.

[49] D. Liu and F. Zhang, "Metabolic feedback circuits provide rapid control of metabolite dynamics," *ACS Synth. Biol.*, vol. 7, no. 2, pp. 347–356, Feb. 2018.

[50] P. S. Swain, "Efficient attenuation of stochasticity in gene expression through post-transcriptional control," *J. Mol. Biol.*, vol. 344, no. 4, pp. 965–976, Dec. 2004.

[51] A. Singh and J. P. Hespanha, "Optimal feedback strength for noise suppression in autoregulatory gene networks," *Biophys. J.*, vol. 96, no. 10, pp. 4013–4023, 2009.

[52] V. Hsiao, E. L. C. de los Santos, W. R. Whitaker, J. E. Dueber, and R. M. Murray, "Design and implementation of a biomolecular concentration tracker," *ACS Synth. Biol.*, vol. 4, no. 2, pp. 150–161, Feb. 2015.

[53] G. Lillacci, Y. Benenson, and M. Khammash, "Synthetic control systems for high performance gene expression in mammalian cells," *Nucleic Acids Res.*, vol. 46, no. 18, pp. 9855–9863,

Oct. 2018.
[54] F. Ceroni *et al.*, "Burden-driven feedback control of gene expression," *Nature Methods*, vol. 15, no. 5, pp. 387–393, May 2018.
[55] L. S. Qi *et al.*, "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression," *Cell*, vol. 152, no. 5, pp. 1173–1183, Feb. 2013.
[56] H.-H. Huang, Y. Qian, and D. D. Vecchio, "A quasi-integral controller for adaptation of genetic modules to variable ribosome demand," *Nature Commun.*, vol. 9, no. 1, pp. 1–12, Dec. 2018.
[57] T. Frei, C.-H. Chang, M. Filo, A. Arampatis, and M. Khammash, "Genetically engineered integral feedback controllers for robust perfect adaptation in mammalian cells," *BioRxiv*, 2021.
[58] R. D. Jones *et al.*, "Robust and tunable signal processing in mammalian cells via engineered covalent modification cycles," *BioRxiv*, 2021.
[59] A. H. Ng *et al.*, "Modular and tunable biological feedback control using a de novo protein switch," *Nature*, vol. 572, no. 7768, pp. 265–269, Aug. 2019.
[60] R. A. Langan *et al.*, "De novo design of bioactive protein switches," *Nature*, vol. 572, no. 7768, pp. 205–210, Aug. 2019.
[61] M. Xie *et al.*, "β-cell–mimetic designer cells provide closed-loop glycemic control," *Science*, vol. 354, no. 6317, pp. 1296–1301, Dec. 2016.
[62] R. J. Bloom, S. M. Winkler, and C. D. Smolke, "Synthetic feedback control using an RNAi-based gene-regulatory device," *J. Biol. Eng.*, vol. 9, no. 1, pp. 1–13, 2015.
[63] T. Folliard, H. Steel, T. P. Prescott, G. Wadhams, L. J. Rothschild, and A. Papachristodoulou, "A synthetic recombinase-based feedback loop results in robust expression," *ACS Synth. Biol.*, vol. 6, no. 9, pp. 1663–1671, Sep. 2017.
[64] C. L. Kelly, A. W. K. Harris, H. Steel, E. J. Hancock, J. T. Heap, and A. Papachristodoulou, "Synthetic negative feedback circuits using engineered small RNAs," *Nucleic Acids Res.*, vol. 46, no. 18, pp. 9875–9889, Oct. 2018.
[65] P. Saxena, G. C.-E. Hamri, M. Folcher, H. Zulewski, and M. Fussenegger, "Synthetic gene network restoring endogenous pituitary–thyroid feedback control in experimental Graves' disease," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 5, pp. 1244–1249, Feb. 2016.
[66] C. C. Samaniego and E. Franco, "Ultrasensitive molecular controllers for quasi-integral feedback," *Cell Syst.*, vol. 12, no. 3, pp. 272–288, 2021.
[67] T. Frei and M. Khammash, "Adaptive circuits in synthetic biology," *Current Opinion Syst. Biol.*, vol. 28, Dec. 2021, Art. no. 100399.
[68] S. Mangan and U. Alon, "Structure and function of the feed-forward loop network motif," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 21, pp. 11980–11985, Oct. 2003.
[69] L. Bleris, Z. Xie, D. Glass, A. Adadey, E. Sontag, and Y. Benenson, "Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template," *Mol. Syst. Biol.*, vol. 7, no. 1, p. 519, Jan. 2011.
[70] T. J. Strovas, A. B. Rosenberg, B. E. Kuypers, R. A. Muscat, and G. Seelig, "MicroRNA-based single-gene circuits buffer protein synthesis rates against perturbations," *ACS Synth. Biol.*, vol. 3, no. 5, pp. 324–331, May 2014.
[71] J. J. Tyson, K. C. Chen, and B. Novak, "Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell," *Current Opinion Cell Biol.*, vol. 15, no. 2, pp. 221–231, Apr. 2003.
[72] L. Goentoro, O. Shoval, M. W. Kirschner, and U. Alon, "The incoherent feedforward loop can provide fold-change detection in gene regulation," *Mol. Cell*, vol. 36, no. 5, pp. 894–899, Dec. 2009.
[73] O. Shoval, L. Goentoro, Y. Hart, A. Mayo, E. Sontag, and U. Alon, "Fold-change detection and scalar symmetry of sensory input fields," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 36, pp. 15995–16000, 2010.
[74] O. Shoval, U. Alon, and E. Sontag, "Symmetry invariance for adapting biological systems," *SIAM J. Appl. Dyn. Syst.*, vol. 10, no. 3, pp. 857–886, Jan. 2011.
[75] M. Osella, C. Bosia, D. Corá, and M. Caselle, "The role of incoherent MicroRNA-mediated feedforward loops in noise buffering," *PLoS Comput. Biol.*, vol. 7, no. 3, Mar. 2011, Art. no. e1001101.
[76] G. Buzi and M. Khammash, "Implementation considerations, not topological differences, are the main determinants of noise suppression properties in feedback and incoherent feedforward circuits," *PLOS Comput. Biol.*, vol. 12, no. 6, Jun. 2016, Art. no. e1004958.
[77] A. Carignano, S. Mukherjee, A. Singh, and G. Seelig, "Extrinsic noise suppression in micro RNA mediated incoherent feedforward loops," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 4353–4359.
[78] T. H. Segall-Shapiro, E. D. Sontag, and C. A. Voigt, "Engineered promoters enable constant gene expression at any copy number in bacteria," *Nature Biotechnol.*, vol. 36, no. 4, pp. 352–358, Apr. 2018.
[79] J. Kim, I. Khetarpal, S. Sen, and R. M. Murray, "Synthetic circuit for exact adaptation and fold-change detection," *Nucleic Acids Res.*, vol. 42, no. 9, pp. 6078–6089, May 2014.
[80] M. K. Takahashi *et al.*, "Characterizing and prototyping genetic networks with cell-free transcription–translation reactions," *Methods*, vol. 86, pp. 60–72, Sep. 2015.
[81] S. Guo and R. M. Murray, "Construction of incoherent feedforward loop circuits in a cell-free system and in cells," *ACS Synth. Biol.*, vol. 8, no. 3, pp. 606–610, Mar. 2019.
[82] S. Hong, D. Jeong, J. Ryan, M. Foo, X. Tang, and J. Kim, "Design and evaluation of synthetic rna-based incoherent feed-forward loop circuits," *Biomolecules*, vol. 11, no. 8, p. 1182, 2021.
[83] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Nat. Acad. Sci. USA*, vol. 97, no. 9, pp. 4649–4653, 2000.
[84] N. Barkai and S. Leibler, "Robustness in simple biochemical networks," *Nature*, vol. 387, no. 6636, pp. 913–917, Jun. 1997.
[85] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, "Robustness in bacterial chemotaxis," *Nature*, vol. 397, no. 6715, pp. 168–171, 1999.
[86] D. Muzzey, C. A. Gómez-Uribe, J. T. Mettetal, and A. van Oudenaarden, "A systems-level analysis of perfect adaptation in yeast osmoregulation," *Cell*, vol. 138, no. 1, pp. 160–171, Jul. 2009.
[87] T. You *et al.*, "A systems biology analysis of long and short-term memories of osmotic stress adaptation in fungi," *BMC Res. Notes*, vol. 5, no. 1, p. 1–16, Dec. 2012.
[88] C. Briat, A. Gupta, and M. Khammash, "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks," *Cell Syst.*, vol. 2, pp. 15–26, Jan. 2016.
[89] F. Mairet, "A biomolecular proportional integral controller based on feedback regulations of protein level and activity," *Roy. Soc. Open Sci.*, vol. 5, no. 2, Feb. 2018, Art. no. 171966.
[90] X. Y. Ni, T. Drengstig, and P. Ruoff, "The control of the controller: Molecular mechanisms for robust perfect adaptation and temperature compensation," *Biophys. J.*, vol. 97, no. 5, pp. 1244–1253, Sep. 2009.
[91] H. El-Samad, J. P. Goff, and M. Khammash, "Calcium homeostasis and parturient hypocalcemia: An integral feedback perspective," *J. Theor. Biol.*, vol. 214, no. 1, pp. 17–29, Jan. 2002.
[92] D. M. Tveit, G. Fjeld, T. Drengstig, F. V. Filipp, P. Ruoff, and K. Thorsen, "Exploring mechanisms of glucose uptake regulation and dilution resistance in growing cancer cells," *BioRxiv*, 2020.
[93] J. R. Norris, *Markov Chains*, no. 2. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[94] S. K. Aoki, G. Lillacci, A. Gupta, A. Baumschlager, D. Schweingruber, and M. Khammash, "A universal biomolecular integral feedback controller for robust perfect adaptation," *Nature*, vol. 570, no. 7762, pp. 533–537, Jun. 2019.
[95] T. Drengstig, X. Y. Ni, K. Thorsen, I. W. Jolma, and P. Ruoff, "Robust adaptation and homeostasis by autocatalysis," *J. Phys. Chem. B*, vol. 116, no. 18, pp. 5355–5363, May 2012.
[96] C. Briat, C. Zechner, and M. Khammash, "Design of a synthetic integral feedback circuit: Dynamic analysis and DNA implementation," *ACS Synth. Biol.*, vol. 5, pp. 1108–1116, Jul. 2016.
[97] F. Xiao and J. C. Doyle, "Robust perfect adaptation in biomolecular reaction networks," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 4345–4352.
[98] C. Briat, A. Gupta, and M. Khammash, "Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks," *J. Roy. Soc. Interface*, vol. 15, no. 143, Jun. 2018, Art. no. 20180079.
[99] N. Olsman, A.-A. Baetica, F. Xiao, Y. P. Leong, R. M. Murray, and J. C. Doyle, "Hard limits and performance tradeoffs in a class of antithetic integral feedback networks," *Cell Syst.*, vol. 9, no. 1, pp. 49–63, 2019.
[100] N. Olsman, F. Xiao, and J. C. Doyle, "Architectural principles for characterizing the performance of antithetic integral feedback networks," *iScience*, vol. 14, pp. 277–291, Apr. 2019.
[101] M. Chevalier, M. Gómez-Schiavon, A. H. Ng, and H. El-Samad, "Design and analysis of a proportional-integral-derivative controller with biological molecules," *Cell Syst.*, vol. 9, no. 4, pp. 338–353, 2019.
[102] M. Filo, S. Kumar, and M. Khammash, "A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance," *Nature Commun.*, vol. 13, no. 1, pp. 1–19, Dec. 2022.
[103] G. Lillacci, S. Aoki, D. Schweingruber, and M. Khammash, "A synthetic integral feedback controller for robust tunable regulation in bacteria," *BioRxiv*, 2017.
[104] J. Ang and D. R. McMillen, "Physical constraints on biological integral control design for homeostasis and sensory adaptation," *Biophys. J.*, vol. 104, pp. 505–515, Jan. 2013.
[105] Y. Qian and D. D. Vecchio, "Realizing 'integral control' in living cells: How to overcome leaky integration due to dilution?" *J. Roy. Soc. Interface*, vol. 15, no. 139, Feb. 2018, Art. no. 20170902.
[106] D. A. Oyarzún and E. Hancock, "Stabilisation of antithetic control via molecular buffering," *J. Roy. Soc. Interface*, vol. 19, p. 2021076, 2022.
[107] D. Chen and A. P. Arkin, "Sequestration-based bistability enables tuning of the switching boundaries and design of a latch," *Mol. Syst. Biol.*, vol. 8, no. 1, p. 620, Jan. 2012.
[108] D. K. Agrawal, R. Marshall, V. Noireaux, and E. D. Sontag, "In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller," *Nature Commun.*, vol. 10, no. 1, pp. 1–12, 2019.
[109] T. Frei, C.-H. Chang, M. Filo, A. Arampatzis, and M. Khammash, "A genetic mammalian proportional-integral feedback control circuit for robust and precise gene regulation," *Proc. Nat. Acad. Sci. USA*, to be published.
[110] V. Hsiao, E. L. C. de los Santos, W. R. Whitaker, J. E. Dueber, and R. M. Murray, "Design and implementation of a biomolecular concentration tracker," *ACS Synth. Biol.*, vol. 4, no. 2, pp. 150–161, Feb. 2015.
[111] C. Cuba Samaniego, G. Giordano, J. Kim, F. Blanchini, and E. Franco, "Molecular titration promotes oscillations and bistability in minimal network models with monomeric regulators," *ACS Synth. Biol.*, vol. 5, no. 4, pp. 321–333, Apr. 2016.
[112] F. Annunziata *et al.*, "An orthogonal multi-input integration system to control gene expression in *Escherichia coli*," *ACS Synth. Biol.*, vol. 6, no. 10, pp. 1816–1824, Oct. 2017.
[113] G. Fiore *et al.*, "In-silico analysis and implementation of a multicellular feedback

control strategy in a synthetic bacterial consortium," *ACS Synth. Biol.*, vol. 6, no. 3, pp. 507–517, Mar. 2017.
[114] X. Ren, A.-A. Baetica, A. Swaminathan, and R. M. Murray, "Population regulation in microbial consortia using dual feedback control," in *Proc. IEEE 56th Annu. Conf. Decis. Control (CDC)*, Dec. 2017, pp. 5341–5347.
[115] B. Shannon *et al.*, "*In vivo* feedback control of an antithetic molecular-titration motif in *Escherichia coli* using microfluidics," *ACS Synth. Biol.*, vol. 9, no. 10, pp. 2617–2624, Oct. 2020.
[116] X. Ren and R. M. Murray, "Role of interaction network topology in controlling microbial population in consortia," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 2691–2697.
[117] R. D. McCardell, A. Pandey, and R. M. Murray, "Control of density and composition in an engineered two-member bacterial community," *BioRxiv*, to be published.
[118] Y. Chen, J. K. Kim, A. J. Hirning, K. Josić, and M. R. Bennett, "Emergent genetic oscillations in a synthetic microbial consortium," *Science*, vol. 349, no. 6251, pp. 986–989, Aug. 2015.
[119] X. Ren and R. M. Murray, "Cooperation enhances robustness of coexistence in spatially structured consortia," in *Proc. 18th Eur. Control Conf. (ECC)*, Jun. 2019, pp. 2651–2656.
[120] R. N. Alnahhas *et al.*, "Spatiotemporal dynamics of synthetic microbial consortia in microfluidic devices," *ACS Synth. Biol.*, vol. 8, no. 9, pp. 2051–2058, Sep. 2019.
[121] M. J. Liao, M. O. Din, L. Tsimring, and J. Hasty, "Rock-paper-scissors: Engineered population dynamics increase genetic stability," *Science*, vol. 365, no. 6457, pp. 1045–1049, Sep. 2019.
[122] R. N. Alnahhas *et al.*, "Majority sensing in synthetic microbial consortia," *Nature Commun.*, vol. 11, no. 1, pp. 1–10, Dec. 2020.
[123] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, "Green fluorescent protein as a marker for gene expression," *Science*, vol. 263, no. 5148, pp. 802–805, Feb. 1994.
[124] G. Nagel *et al.*, "Channelrhodopsin-2, a directly light-gated cation-selective membrane channel," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 24, pp. 13940–13945, Nov. 2003.
[125] S. Shimizu-Sato, E. Huq, J. M. Tepperman, and P. H. Quail, "A light-switchable gene promoter system," *Nature Biotechnol.*, vol. 20, no. 10, pp. 1041–1044, Oct. 2002.
[126] A. Milias-Argeitis *et al.*, "In silico feedback for *in vivo* regulation of a gene expression circuit," *Nature Biotechnol.*, vol. 29, no. 12, pp. 1114–1116, Dec. 2011.
[127] J. Palmer, *Cyborg Yeast Genes Run by Computer*. London, U.K.: BBC News, Nov. 2011.
[128] J. E. Toettcher, D. Gong, W. A. Lim, and O. D. Weiner, "Light-based feedback for controlling intracellular signaling dynamics," *Nature Methods*, vol. 8, no. 10, pp. 837–839, Oct. 2011.
[129] J. Uhlendorf *et al.*, "Long-term model predictive control of gene expression at the population and single-cell levels," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 35, pp. 14271–14276, Aug. 2012.
[130] J. Ruess, F. Parise, A. Milias-Argeitis, M. Khammash, and J. Lygeros, "Iterative experiment design guides the characterization of a light-inducible gene expression circuit," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 26, pp. 8148–8153, Jun. 2015.
[131] A. Milias-Argeitis, M. Rullan, S. K. Aoki, P. Buchmann, and M. Khammash, "Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth," *Nature Commun.*, vol. 7, no. 1, pp. 1–11, Nov. 2016.
[132] G. Fiore, G. Perrino, M. di Bernardo, and D. di Bernardo, "*In vivo* real-time control of gene expression: A comparative analysis of feedback control strategies in yeast," *ACS Synth. Biol.*, vol. 5, no. 2, pp. 154–162, Feb. 2016.
[133] J.-B. Lugagne, S. S. Carrillo, M. Kirch, A. Köhler, G. Batt, and P. Hersen, "Balancing a genetic toggle switch by real-time feedback control and periodic forcing," *Nature Commun.*, vol. 8, no. 1, pp. 1–8, Dec. 2017.
[134] R. Chait, J. Ruess, T. Bergmiller, G. Tkačik, and C. C. Guet, "Shaping bacterial population behavior through computer-interfaced control of individual cells," *Nature Commun.*, vol. 8, no. 1, pp. 1–11, Dec. 2017.
[135] M. Rullan, D. Benzinger, G. W. Schmidt, A. Milias-Argeitis, and M. Khammash, "An optogenetic platform for real-time, single-cell interrogation of stochastic transcriptional regulation," *Mol. cell*, vol. 70, no. 4, pp. 745–756, 2018.
[136] L. Postiglione *et al.*, "Regulation of gene expression and signaling pathway activity in mammalian cells by automated microfluidics feedback control," *ACS Synth. Biol.*, vol. 7, no. 11, pp. 2558–2565, Nov. 2018.
[137] E. Pedone *et al.*, "A tunable dual-input system for on-demand dynamic gene expression regulation," *Nature Commun.*, vol. 10, no. 1, pp. 1–13, Dec. 2019.
[138] P. Harrigan, H. D. Madhani, and H. El-Samad, "Real-time genetic compensation defines the dynamic demands of feedback control," *Cell*, vol. 175, no. 3, pp. 877–886, 2018.
[139] M. L. Perkins, D. Benzinger, M. Arcak, and M. Khammash, "Cell-in-the-loop pattern formation with optogenetically emulated cell-to-cell signaling," *Nature Commun.*, vol. 11, no. 1, pp. 1–10, Dec. 2020.
[140] G. Perrino *et al.*, "Automatic synchronisation of the cell cycle in budding yeast through closed-loop feedback control," *Nature Commun.*, vol. 12, no. 1, pp. 1–12, Dec. 2021.
[141] S. Kumar, M. Rullan, and M. Khammash, "Rapid prototyping and design of cybergenetic single-cell controllers," *Nature Commun.*, vol. 12, no. 1, pp. 1–13, Dec. 2021.
[142] M. Khammash, M. Di Bernardo, and D. D. Bernardo, "Cybergenetics: Theory and methods for genetic control system," in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, Dec. 2019, pp. 916–926.
[143] J.-B. Lugagne and M. J. Dunlop, "Cell-machine interfaces for characterizing gene regulatory network dynamics," *Current Opinion Syst. Biol.*, vol. 14, pp. 1–8, Apr. 2019.
[144] A. Banderas, M. Le Bec, C. Cordier, and P. Hersen, "Autonomous and assisted control for synthetic microbiology," *Int. J. Mol. Sci.*, vol. 21, no. 23, p. 9223, Dec. 2020.
[145] S. Pouzet, A. Banderas, M. L. Bec, T. Lautier, G. Truan, and P. Hersen, "The promise of optogenetics for bioproduction: Dynamic control strategies and scale-up instruments," *Bioengineering*, vol. 7, no. 4, p. 151, Nov. 2020.
[146] C. Carrasco-López, S. A. García-Echauri, T. Kichuk, and J. L. Avalos, "Optogenetics and biosensors set the stage for metabolic cybergenetics," *Current Opinion Biotechnol.*, vol. 65, pp. 296–309, Oct. 2020.
[147] F. Bertaux, J. Ruess, and G. Batt, "External control of microbial populations for bioproduction: A modeling and optimization viewpoint," *Current Opinion Syst. Biol.*, vol. 28, Dec. 2021, Art. no. 100394.
[148] I. Ruolo, S. Napolitano, D. Salzano, M. di Bernardo, and D. di Bernardo, "Control engineering meets synthetic biology: Foundations and applications," *Current Opinion Syst. Biol.*, vol. 28, Dec. 2021, Art. no. 100397.
[149] G. Perrino, A. Hadjimitsis, R. Ledesma-Amaro, and G.-B. Stan, "Control engineering and synthetic biology: Working in synergy for the analysis and control of microbial systems," *Current Opinion Microbiol.*, vol. 62, pp. 68–75, Aug. 2021.
[150] S. Mukherji and A. van Oudenaarden, "Synthetic biology: Understanding biological design from synthetic circuits," *Nature Rev. Genet.*, vol. 10, no. 12, pp. 859–871, Dec. 2009.
[151] Y. Y. Chen, K. E. Galloway, and C. D. Smolke, "Synthetic biology: Advancing biological frontiers by building synthetic systems," *Genome Biol.*, vol. 13, no. 2, pp. 1–10, 2012.
[152] C. J. Bashor and J. J. Collins, "Understanding biological regulation through synthetic biology," *Annu. Rev. Biophys.*, vol. 47, no. 1, pp. 399–423, May 2018.
[153] H. Ye *et al.*, "Self-adjusting synthetic gene circuit for correcting insulin resistance," *Nature Biomed. Eng.*, vol. 1, no. 1, pp. 1–9, Jan. 2017.
[154] B. Gasser *et al.*, "Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview," *Microbial Cell Factories*, vol. 7, no. 1, pp. 1–18, Oct. 2008.
[155] P. M. Boyle and P. A. Silver, "Parts plus pipes: Synthetic biology approaches to metabolic engineering," *Metabolic Eng.*, vol. 14, no. 3, pp. 223–232, May 2012.
[156] N. S. McCarty and R. Ledesma-Amaro, "Synthetic biology tools to engineer microbial communities for biotechnology," *Trends Biotechnol.*, vol. 37, no. 2, pp. 181–197, Feb. 2019.
[157] D. Salzano, D. Fiore, and M. D. Bernardo, "Ratiometric control for differentiation of cell populations endowed with synthetic toggle switches," in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, Dec. 2019, pp. 927–932.
[158] A. Zaslaver *et al.*, "Just-in-time transcription program in metabolic pathways," *Nature Genet.*, vol. 36, no. 5, pp. 486–491, May 2004.
[159] C.-S. Chin, V. Chubukov, E. R. Jolly, J. DeRisi, and H. Li, "Dynamics and design principles of a basic regulatory architecture controlling metabolic pathways," *PLoS Biol.*, vol. 6, no. 6, p. e146, Jun. 2008.
[160] W. J. Holtz and J. D. Keasling, "Engineering static and dynamic control of synthetic pathways," *Cell*, vol. 140, no. 1, pp. 19–23, Jan. 2010.
[161] M. J. Dunlop, J. D. Keasling, and A. Mukhopadhyay, "A model for improving microbial biofuel production using a synthetic feedback loop," *Syst. Synth. Biol.*, vol. 4, no. 2, pp. 95–104, Jun. 2010.
[162] M. J. Dunlop *et al.*, "Engineering microbial biofuel tolerance and export using efflux pumps," *Mol. Syst. Biol.*, vol. 7, no. 1, p. 487, Jan. 2011.
[163] F. Zhang, J. M. Carothers, and J. D. Keasling, "Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids," *Nature Biotechnol.*, vol. 30, pp. 354–359, Apr. 2012.
[164] V. Chubukov, I. A. Zuleta, and H. Li, "Regulatory architecture determines optimal regulation of gene expression in metabolic pathways," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 13, pp. 5127–5132, Mar. 2012.
[165] D. A. Oyarzún and G.-B.-V. Stan, "Synthetic gene circuits for metabolic control: Design trade-offs and constraints," *J. Roy. Soc. Interface*, vol. 10, no. 78, Jan. 2013, Art. no. 20120671.
[166] R. H. Dahl *et al.*, "Engineering dynamic pathway regulation using stress-response promoters," *Nature Biotechnol.*, vol. 31, no. 11, pp. 1039–1046, Nov. 2013.
[167] P. Xu, L. Li, F. Zhang, G. Stephanopoulos, and M. Koffas, "Improving fatty acids production by engineering dynamic pathway regulation and metabolic control," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 31, pp. 11299–11304, Aug. 2014.
[168] D. Liu, Y. Xiao, B. S. Evans, and F. Zhang, "Negative feedback regulation of fatty acid production based on a malonyl-CoA sensor–actuator," *ACS Synth. Biol.*, vol. 4, no. 2, pp. 132–140, Feb. 2015.
[169] A. Gupta, I. M. B. Reizman, C. R. Reisch, and K. L. J. Prather, "Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit," *Nature Biotechnol.*, vol. 35, no. 3, pp. 273–279, Mar. 2017.
[170] Y. Siu, J. Fenno, J. M. Lindle, and M. J. Dunlop, "Design and selection of a synthetic feedback loop for optimizing biofuel tolerance," *ACS Synth. Biol.*, vol. 7, no. 1, pp. 16–23, Jan. 2018.

[171] C. Briat and M. Khammash, "Perfect adaptation and optimal equilibrium productivity in a simple microbial biofuel metabolic pathway using dynamic integral control," *ACS Synth. Biol.*, vol. 7, no. 2, pp. 419–431, Feb. 2018.
[172] Y. Boada, A. Vignoni, J. Picó, and P. Carbonell, "Extended metabolic biosensor design for dynamic pathway regulation of cell factories," *iScience*, vol. 23, no. 7, Jul. 2020, Art. no. 101305.
[173] C. Kemmer, M. Gitzinger, M. D.-E. Baba, V. Djonov, J. Stelling, and M. Fussenegger, "Self-sufficient control of urate homeostasis in mice by a synthetic circuit," *Nature Biotechnol.*, vol. 28, no. 4, pp. 355–360, Apr. 2010.
[174] D. Del Vecchio, Y. Qian, R. M. Murray, and E. D. Sontag, "Future systems and control research in synthetic biology," *Annu. Rev. Control*, vol. 45, pp. 5–17, Jan. 2018.

ABOUT THE AUTHOR

Mustafa H. Khammash (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rice University, Houston, TX, USA, in 1990.

He joined the Electrical Engineering Department, Iowa State University (ISU), Ames, IA, USA. While at ISU, he created the Dynamics and Control Program and led that control group until 2002, when he joined the Faculty of the University of California at Santa Barbara, Santa Barbara, CA, USA. He has served as the Vice-Chair of the Mechanical Engineering Department and the Director of the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, from 2003 to 2006 and 2005 to 2011, respectively. In 2011, he moved with his group to Switzerland, joining the Department of Biosystems Science and Engineering, ETH Zürich, Zürich, Switzerland. At ETH Zürich, he served as the Department Chair from 2015 to 2017. He is currently a Professor of control theory and systems biology with the Department of Biosystems Science and Engineering (D-BSSE), ETH Zürich. He works at the interface of the areas of control theory, systems biology, and synthetic biology. His research aims to understand the role of dynamics, feedback, and randomness in biology and to develop the tools needed to aid in this understanding. His group is currently developing the theory, computational methods, and experimental tools for controlling living cells using both computer and molecular control systems.

Dr. Khammash has been a Fellow of the International Federation of Automatic Control since 2010 and the Japan Society for the Promotion of Science since 2003. He is a winner of the ISU Early achievement in Research and Scholarship Award, the Young Engineering Research Faculty Award, the ETH Golden Owl Award, and the European Research Council Advanced Grant.