Resource Competition Shapes the Response of Genetic Circuits

Yili Qian[1†], Hsin-Ho Huang[1†], José I. Jiménez[1,2] and Domitilla Del Vecchio[1,3,*]

[1] Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA02139, USA
[2] Faculty of Health of Medical Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK.
[3] Synthetic Biology Center, Massachusetts Institute of Technology, 500 Technology Square, Cambridge, MA02139, USA
[†] These authors contributed equally to this work.

∗ Correspondence:
Address: Room 3-469, Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA02139, USA.
Phone: +1-617-324-6108
Email: ddv@mit.edu

Abstract

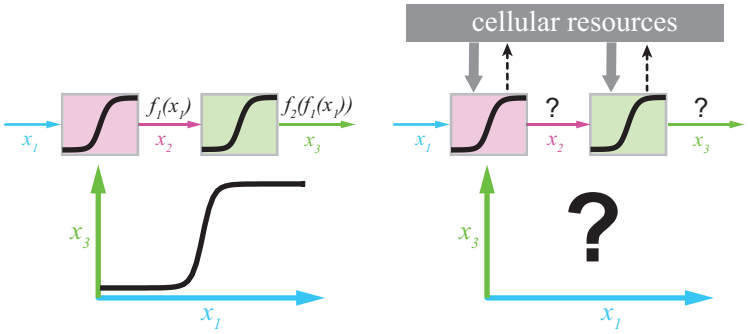
A common approach to design genetic circuits is to compose gene expression cassettes together. While appealing, this modular approach is challenged by the fact that expression of each gene depends on the availability of transcriptional/translational resources, which is in turn determined by the presence of other genes in the circuit. This raises the question of how competition for resources by different genes affects a circuit's behavior. Here, we create a library of genetic activation cascades in bacteria *E. coli*, where we explicitly tune the resource demand by each gene. We develop a general Hill-function-based model that incorporates resource competition effects through resource demand coefficients. These coefficients lead to non-regulatory interactions among genes that reshape circuit's behavior. For the activation cascade, such interactions result in surprising biphasic or monotonically decreasing responses. Finally, we use resource demand coefficients to guide the choice of ribosome binding site (RBS) and DNA copy number to restore the cascade's intended monotonically increasing response. Our results demonstrate how unintended circuit's behavior arises from resource competition and provide a model-guided methodology to minimize the resulting effects.

Keywords: genetic circuit, context dependence, modularity, resource competition, model-guided design, activation cascade

Resource Competition Shapes the Response of Genetic Circuits

Yili Qian, Hsin-Ho Huang, José I. Jiménez and Domitilla Del Vecchio

For Table of Contents Use Only



Introduction

Predicting the behavior of genetic circuits in living cells is a recurring challenge in synthetic biology [1]. Genetic circuits are often viewed as interconnections of gene expression cassettes, which we call *nodes*. Each cassette (node) is composed of core gene expression processes, chiefly transcription and translation. Here, we view each node as an input/output system that takes transcription factors (TFs) as input and gives a TF as output. The input TFs regulate the production of the output TF. Although in an ideal scenario we would like to predict the behavior of a circuit from that of its composing nodes characterized in isolation, in reality, a node's behavior often depends on its context, including other nodes in the same circuit and the host cell environment [2]. This fact significantly limits our current ability to design genetic circuits that behave as intended. There are a number of causes to context dependence, including unknown structural interactions between adjacent genetic sequences [3], loading of TFs by target DNA sites (retroactivity) [4, 5, 6], unintended coupling between synthetic genes and host cell growth (host-circuit interaction) [7, 8, 9], and competition among synthetic genes with each other for common transcriptional and translational resources [10, 11, 12, 13, 14]. Context dependence due to structural interactions and retroactivity has been addressed by engineering insulation parts and devices [15, 16, 6, 17, 18] and that due to host-circuit interaction may be mitigated to some extent by orthogonal RNA polymerases (RNAPs) and ribosomes [19, 20, 21, 22]. By contrast, the characterization and mitigation of competition for shared resources among synthetic genes remain largely unexplored.

Expression of all genes in a genetic circuit relies on a common pool of transcriptional and translational resources. In particular, the availability of RNAPs and ribosomes has been identified as a major bottleneck for gene expression in bacteria [23, 24, 25, 26]. When a node is activated, it depletes the pool of free RNAPs and ribosomes, reducing their availability to other nodes in the circuit. This can potentially affect the behavior of a circuit altogether. Recent experimental results have demonstrated that competition for transcriptional and translational resources can couple the expression of two synthetic genes that are otherwise unconnected [10, 12]. In particular, limitation in ribosome availability has been identified as the key player in this coupling phenomenon [12]. These works further demonstrate that upon induction of a synthetic gene, the expression level of a constitutively expressed gene on the same plasmid can be reduced by more than 60%. Similar trade-offs have been observed in cell-free systems [27] and in computational models [11, 13, 14, 28].

In this paper, we seek to determine how competition for RNAPs and ribosomes by the genes constituting a synthetic genetic circuit changes the intended circuit's behavior in bacteria *E. coli*. To address this question, we perform a combined modeling and experimental study. In particular, we develop a general mathematical model that explicitly includes competition for RNAPs and ribosomes in Hill-function models of gene expression. In our models, resource demand coefficients quantify the demand for resources by each node and shape the emergent dose response curve of a genetic circuit. We construct a library of synthetic genetic activation cascades in which we tune the resource demand coefficients by changing the RBS strength of the cascade's genes and DNA copy number. When the resource demand coefficients are large, the dose response curve of the cascade can either be biphasic or monotonically decreasing. When we decrease the resource demand coefficients, we restore the intended cascade's monotonically increasing dose response curve. For general circuits, our model reveals that due to non-zero resource demand coefficients, resource competition gives rise to non-regulatory interactions among nodes. We give a general rule for drawing the effective interaction graph of any genetic circuit that combines both regulatory and non-regulatory interactions.

Results

Surprising biphasic response of a genetic activation cascade

We built a genetic activation cascade composed of three nodes and two transcriptional activation stages. Two inducer-responsive TFs, LuxR from *Vibrio fischeri* [29] and NahR from *Pseudomonas putida* [30], activate gene expression in their active forms (i.e., *holo* forms) when their respective inducers *N*-hexanoyl-L-homoserine lactone (AHL) and salicylate (SAL) are present. Node 1 uses the *lac* promoter to constitutively express LuxR in a LacI-deficient host strain. By increasing inducer AHL concentration, the active form of LuxR increases, and it can transcriptionally activate the following node. We consider active LuxR as the output of node 1. Node 2 uses transcriptional activation by the active LuxR through the *lux* promoter (Figure 1A). To characterize the dose response curve of this node, we placed red fluorescent protein (RFP) under the control of the *lux* promoter. An increase in AHL concentration increases the active LuxR to promote the production of RFP (Figure 1A).

Node 3 employs transcriptional activation by active NahR and the *sal* promoter to express green fluorescent protein (GFP) as fluorescence output. Inactive NahR is first produced under the control of the *lux* promoter. We applied a saturating amount of AHL (100 nM) and expressed LuxR constitutively to produce

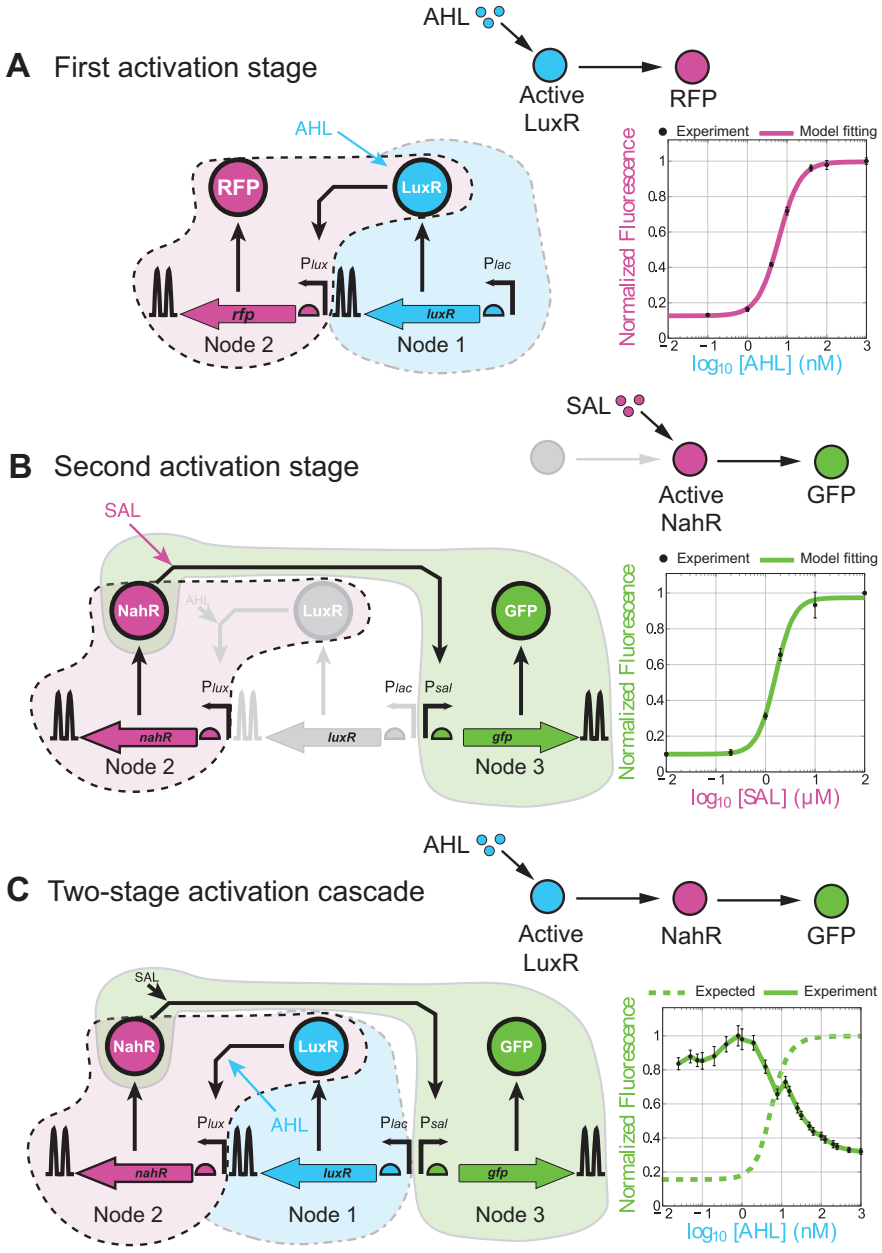


Figure 1: **Failure of modular composition in a simple two-stage activation cascade.** (A) The first activation stage consists of a node that takes as input the transcription activator LuxR to promote production of RFP as output in the presence of AHL, resulting in a monotonically increasing dose response curve. Upward arrows with leftward/rightward tips represent promoters, semicircles represent RBS, and double hairpins represent terminators. The illustrative diagram composed of nodes and edges at the upper-right corner represents regulatory interactions among species. (B) The second activation stage consists of a node that takes as input the transcription activator NahR to promote production of GFP as output in the presence of SAL, resulting in a monotonically increasing dose response curve. (C) The two-stage activation cascade CAS 1/30 was built by connecting the nodes in a cascade topology. Biphasic dose response curve (solid line) of the cascade was observed instead of the expected monotonically increasing dose response curve (dashed line), which is the composition of the two increasing Hill functions for the individual nodes according to equation (3). All experimental data represent mean values and standard deviations of populations in the steady state analyzed by flow cytometry in three independent experiments. Each plot is normalized to its maximum fluorescence value (see SI Section A7 for details).

a saturating amount of inactive NahR. By increasing the amount of inducer SAL, active NahR concentration increases, activating production of GFP (Figure 1B). We further confirm that the dose response curve of GFP activation by active NahR remains monotonically increasing under different AHL concentrations (see SI Section A6 and Figure S1).

To build a two-stage activation cascade (CAS 1/30), we connected the three nodes by replacing the RFP in node 2 by NahR. Active NahR can be regarded as the output of node 2 and the input to node 3. With a constant amount of SAL (1 mM), increased AHL concentration leads to increased active LuxR, and hence to increased concentration of active NahR, resulting in increased production of GFP (cascade output).

The expected behavior of this cascade is therefore a monotonically increasing GFP fluorescence as AHL is increased. This can be easily predicted by a standard Hill-function model of the circuit. Specifically, letting I_1 denote inducer AHL, I_2 denote inducer SAL, x_1 denote active LuxR, x_2 denote active NahR, and x_3 denote GFP output, and using I_1, I_2, x_1, x_2 and x_3 (*italics*) to represent their concentrations, we obtain the following ordinary differential equation (ODE) model (see SI Section B1 for details):

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= \underbrace{T_1 \cdot H_1(I_1)}_{\substack{\text{LuxR production}\\ \text{\& allosteric modulation}}} - \underbrace{\gamma x_1}_{\text{dilution}}, \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= \underbrace{T_2 \cdot F_2(x_1) \cdot H_2(I_2)}_{\substack{\text{NahR production}\\ \text{\& allosteric modulation}}} - \underbrace{\gamma x_2}_{\text{dilution}}, \\
\frac{\mathrm{d}x_3}{\mathrm{d}t} &= \underbrace{T_3 \cdot F_3(x_2)}_{\substack{\text{GFP}\\ \text{production}}} - \underbrace{\gamma x_3}_{\text{dilution}},
\end{aligned} \tag{1}$$

where T_i ($i = 1, 2, 3$) is a lumped parameter describing maximal production rate of node i (defined in equation (S15) in SI), γ is the dilution rate constant, and $H_1(I_1)$, $H_2(I_2)$, $F_2(x_1)$ and $F_3(x_2)$ are standard increasing Hill-functions whose maxima are normalized to 1. $H_1(I_1)$ and $H_2(I_2)$ describe allosteric modulation of LuxR by AHL, and of NahR by SAL. Since we applied a constant amount of SAL (i.e., I_2 = constant), $H_2(I_2)$ is a constant. Without loss of generality, we assume $H_2(I_2) \equiv 1$ in sequel, as any non-unity $H_2(I_2)$ can be absorbed into the lumped parameter T_2 without affecting our analysis (see SI Section B1). Hill functions $F_2(x_1)$ and $F_3(x_2)$ describe transcriptional regulations of NahR by active LuxR, and of GFP by active NahR, respectively. These Hill functions describing regulatory interactions are derived from the chemical reactions in SI Section B1, and are given by:

$$H_1(I_1) = \frac{(I_1/k_1)^{n_1}}{1 + (I_1/k_1)^{n_1}} \qquad F_2(x_1) = \frac{\beta_2 + (x_1/k_2)^{n_2}}{1 + (x_1/k_2)^{n_2}}, \qquad F_3(x_2) = \frac{\beta_3 + (x_2/k_3)^{n_3}}{1 + (x_2/k_3)^{n_3}}. \tag{2}$$

In equations (2), k_i are the dissociation constants between the regulators, I_1, x_1 and x_2, and their respective DNA/protein targets. Dimensionless parameters $\beta_i < 1$ characterize basal expressions, and n_i are Hill coefficients capturing cooperativities of the TF and promoter (or of the inducer and TF) bindings. Setting the time derivatives in (1) to zero, we obtain the dose response curve of the cascade as the composition of three increasing Hill-functions:

$$x_3 = \frac{T_3}{\gamma} F_3 \left[\frac{T_2}{\gamma} F_2 \left(\frac{T_1 H_1(I_1)}{\gamma} \right) \right]. \tag{3}$$

It is clear from (3), that independent of parameters, the steady state of x_3 (GFP concentration) always increases with I_1 (AHL concentration).

Surprisingly, the experimental results contradict this rather trivial prediction. In fact, although the input/output responses of both transcriptionally regulated nodes are monotonically increasing (Figure 1A-B), their cascade shows a biphasic dose response curve, in which the GFP fluorescence decreases with increased concentrations of AHL for higher AHL concentrations (Figure 1C). This fact clearly demonstrates that while the standard model well represents the activation behavior of each individual node, its predictive ability is lost when the nodes are connected and thus are simultaneously activated.

In the next section, we derive a new model, similar in form to that of model (1), which is able to predict the experimentally observed behavior.

A cascade model taking into account resource competition predicts non-regulatory interactions

An underlying assumption in the standard Hill-function model (1) is that the concentrations of free RNAPs and ribosomes can be regarded as constant parameters [31, 32] (refer to SI Section B1). In reality, because

their total availability is limited [23, 24, 26], their free concentrations should depend on the extent to which different nodes in a circuit demand them. With reference to Figure 1C, the biphasic response of x_3 can be explained by the following resource competition mechanism. When we increase I_1, the concentration of x_1 increases, promoting production of x_2. As node 2 sequesters more RNAPs and ribosomes, the amount of free resources decreases, which in turn result in decreased transcription and translation rates in node 3.

We therefore created a model that explicitly accounts for the limited concentrations of RNAPs and ribosomes and for their competition by the three nodes in the cascade. For a given growth rate, the total concentrations of RNAPs and ribosomes can be assumed constant parameters [9, 23]. Considering the conservation law for these resources and solving for their free concentrations (see SI Section B2), we obtain the following modified Hill-function model:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \underbrace{\frac{T_1 H_1(I_1)}{1 + J_1 + J_2F_2(x_1) + J_3F_3(x_2)}}_{G_1(x_1,x_2,I_1)\text{: effective node 1 production}} - \underbrace{\gamma x_1}_{\text{dilution}}, \tag{4}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \underbrace{\frac{T_2 F_2(x_1)}{1 + J_1 + J_2F_2(x_1) + J_3F_3(x_2)}}_{G_2(x_1,x_2)\text{: effective node 2 production}} - \underbrace{\gamma x_2}_{\text{dilution}}, \tag{5}$$

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = \underbrace{\frac{T_3 F_3(x_2)}{1 + J_1 + J_2F_2(x_1) + J_3F_3(x_2)}}_{G_3(x_1,x_2)\text{: effective node 3 production}} - \underbrace{\gamma x_3}_{\text{dilution}}, \tag{6}$$

in which $H_1(I_1)$, $F_2(x_1)$ and $F_3(x_2)$ are regulatory Hill functions defined in equations (2), and T_i $(i = 1, 2, 3)$ is the maximum production rate of node i defined in (S25) in SI. The lumped dimensionless parameter J_i can be understood as an indicator of maximal resource demand by node i, and we call it *resource demand coefficient*. It is defined as:

$$J_i := \frac{p_{i,T}}{K_i} \cdot \left(1 + \frac{\alpha_i}{\kappa_i \delta_i} y_T\right), \tag{7}$$

where $p_{i,T}$ is the DNA copy number of node i; α_i is its transcription elongation rate constant, describing the average number of mRNAs transcribed from a single DNA molecule in unit time; δ_i is mRNA decay rate constant, and y_T is the total concentration of RNAPs. The ability of each DNA molecule (mRNA molecule) to occupy free RNAPs (ribosomes) is characterized by lumped coefficient K_i (κ_i), defined in equations (S3) and (S8) in SI. They can be viewed as effective dissociation constants that decrease with (i) stronger affinity between activated promoter (RBS) in node i and free RNAPs (ribosomes), and (ii) lower transcription (translation) elongation rate constants. Physically, resource demand coefficient of node i (J_i) increases as (I) the total number of promoter sites ($p_{i,T}$) increases, (II) the total number of mRNA molecules ($p_{i,T}\alpha_i y_T / K_i \delta_i$) increases, (III) the ability of each DNA molecule to sequester free RNAPs ($1/K_i$) increases, or (IV) the ability of each mRNA molecule to sequester free ribosomes ($1/\kappa_i$) increases. For a given transcriptional activation level, the portion of resources allocated to each node is quantified by J_1, $J_2F_2(x_1)$ and $J_3F_3(x_2)$, respectively, and follow the conservation law (see SI Section B5 for derivation):

$$\underbrace{y_T \cdot z_T}_{\substack{\text{total available}\\ \text{resources}}} = \underbrace{y \cdot z}_{\substack{\text{free}\\ \text{resources}}} + \underbrace{y \cdot z \cdot J_1}_{\substack{\text{node 1}\\ \text{resource demand}}} + \underbrace{y \cdot z \cdot J_2F_2(x_1)}_{\substack{\text{node 2}\\ \text{resource demand}}} + \underbrace{y \cdot z \cdot J_3F_3(x_2)}_{\substack{\text{node 3}\\ \text{resource demand}}}, \tag{8}$$

where y_T (y) and z_T (z) are the total (free) amount of RNAPs and ribosomes, respectively.

The major difference between model (4)-(6) and the standard Hill-function model (1) is the common denominator $1 + J_1 + J_2F_2(x_1) + J_3F_3(x_2)$ in the effective node production rates $G_1(x_1, x_2, I_1)$, $G_2(x_1, x_2)$ and $G_3(x_1, x_2)$. In a resource-abundant situation where RNAPs and ribosomes bound to all nodes are much smaller than their free amounts ($y \approx y_T$ and $z \approx z_T$), we have $J_1, J_2, J_3 \ll 1$, and model (4)-(6) reduces to the standard Hill-function model (1). Detailed proof of this result is in SI Section B5. Because of the common denominator, the production of each node depends on all TFs present in the circuit as opposed to depending only on its own inputs as in equation (1). In particular, regardless of parameters, we always have the following effective interactions among the cascade nodes (see SI Section B3.1 for derivation):

$$\begin{aligned}
&\frac{\partial G_1}{\partial x_1} < 0 \Rightarrow \mathrm{x}_1 \dashv \mathrm{x}_1, \quad \frac{\partial G_1}{\partial x_2} < 0 \Rightarrow \mathrm{x}_2 \dashv \mathrm{x}_1, \quad \frac{\partial G_1}{\partial I_1} > 0 \Rightarrow \mathrm{I}_1 \to \mathrm{x}_1,\\
&\frac{\partial G_2}{\partial x_1} > 0 \Rightarrow \mathrm{x}_1 \to \mathrm{x}_2, \quad \frac{\partial G_2}{\partial x_2} < 0 \Rightarrow \mathrm{x}_2 \dashv \mathrm{x}_2,\\
&\frac{\partial G_3}{\partial x_1} < 0 \Rightarrow \mathrm{x}_1 \dashv \mathrm{x}_3, \quad \frac{\partial G_3}{\partial x_2} > 0 \Rightarrow \mathrm{x}_2 \to \mathrm{x}_3.
\end{aligned} \tag{9}$$

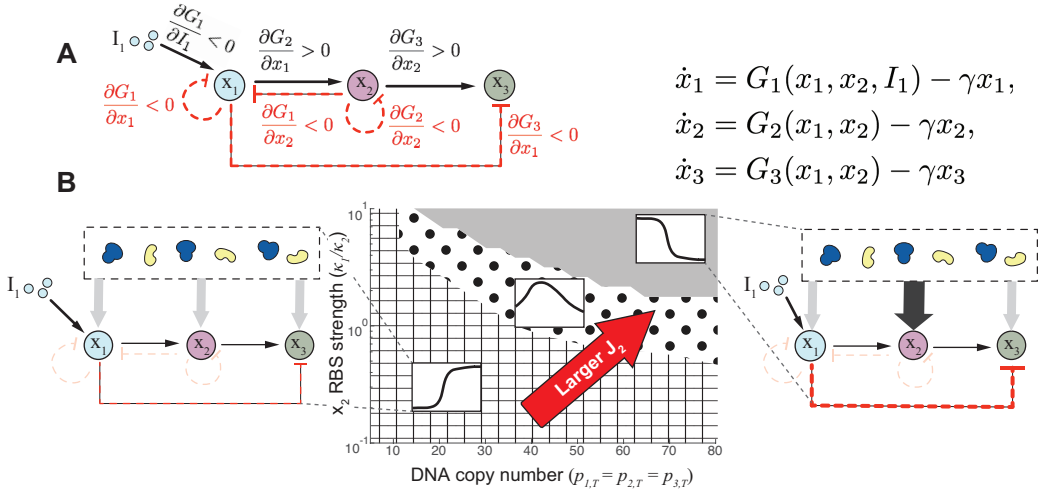


Figure 2: **Activation cascade becomes an effective type 3 IFFL due to resource competition.** (A) Effective interactions among nodes in a two-stage activation cascade with resource competition. Black solid edges are regulatory interactions, and red dashed edges represent emergent non-regulatory interactions due to resource competition. (B) Parameter space illustrating diverse cascade dose response curves obtained from numerical simulations when the resource demand coefficient J_2 is changed. The horizontal axis shows the DNA copy number, and the vertical axis shows the RBS strength of node 2. Numerical values on the vertical axis represent the ratio between the dissociation constant of node 1 (κ_1) between RBS and ribosomes (kept constant at 15 µM), and that of node 2 (κ_2). The cascade has monotonically decreasing, biphasic or monotonically increasing dose response curve depending on whether the parameters fall into the gray, dotted or grid shaded region in the parameter space, respectively. Simulations are based on a full reaction rate equation model corresponding to the chemical reactions in SI Section B1 and B2. Parameter values are listed in SI Table 2.

While interactions $I_1 \to x_1$, $x_1 \to x_2$ and $x_2 \to x_3$ are due to the intended allosteric modulation and transcriptional activations, the other interactions are not present in the standard model (1). They can be regarded as non-regulatory interactions arising from resource competition among nodes. In particular, the non-regulatory interactions $x_1 \dashv x_3$ and $x_1 \dashv x_1$ are due to the fact that as x_1 increases, production of x_2 is activated, depleting the pool of free resources, thus reducing the amount of resources available to initiate transcription and translation of x_3 and x_1, respectively. Similarly, an increase in x_2 activates production of x_3, reducing resources available to its own expression and that of x_1, leading to non-regulatory interactions $x_2 \dashv x_2$ and $x_2 \dashv x_1$.

Based on (9), the effective interactions among nodes in an activation cascade are shown in Figure 2A, where we use red dashed edges to represent emergent non-regulatory interactions due to resource competition. The non-regulatory interactions create a feed-forward edge $x_1 \dashv x_3$, a feedback edge $x_2 \dashv x_1$ and two negative auto-regulation edges: $x_1 \dashv x_1$, $x_2 \dashv x_2$. In SI Section B3.1, we demonstrate that regardless of emergent negative auto-regulation edges on x_1 and x_2, and the feedback edge $x_2 \dashv x_1$, x_1 and x_2 still increases with inducer input I_1 as expected. Therefore, the topology of this activation cascade effectively becomes a type 3 incoherent feed-forward loop (IFFL) [31], where x_3 production is jointly affected by regulatory activation from x_2 and non-regulatory repression from x_1. It is well-known that, depending on parameters, the dose response curve of an IFFL can be monotonically increasing, decreasing or biphasic [33, 34]. As we increase I_1 to increase x_1, if transcriptional activation $x_1 \to x_2 \to x_3$ is stronger than non-regulatory repression $x_1 \dashv x_3$, then the dose response curve is monotonically increasing. Conversely, if the non-regulatory repression is stronger than transcriptional activation, the dose response curve becomes monotonically decreasing. Biphasic responses can be expected when transcriptional activation dominates at lower inducer levels, and resource-competition-induced non-regulatory repression becomes more significant at higher inducer levels. A detailed analytical treatment is in SI Section B3.

The strength of the non-regulatory repression $x_1 \dashv x_3$ can be reduced by decreasing resource demand coefficient of node 2 (J_2). This is because, as a result, the dose response curve of an activation cascade is monotonically increasing when $J_2 \ll 1$ (see SI Section B3.2). Conversely, we expect the dose response curve to be monotonically decreasing when J_2 is large, and to be biphasic for intermediate values of J_2. Based on the definition of resource demand coefficient in (7), we can decrease J_2 by choosing weak node 2

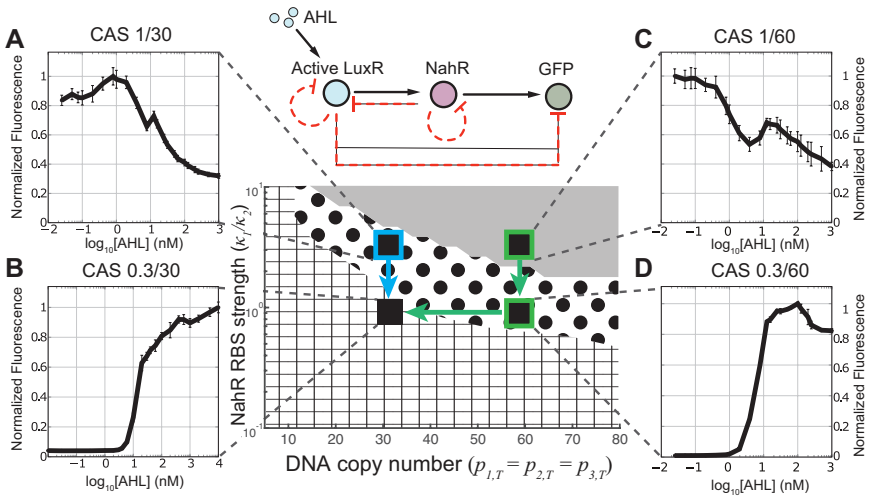


Figure 3: **Model-guided design restores the monotonically increasing dose response curve of the cascade.** The dose response curves of circuits CAS 1/30 (A) and CAS 1/60 (C) were biphasic and monotonically decreasing, respectively. By reducing the RBS strength of NahR, the dose response curve of CAS 0.3/30 (B) became monotonically increasing, and that of CAS 0.3/60 (D) became biphasic. Further decreasing the copy number of CAS 0.3/60 to CAS 0.3/30 restored the monotonically increasing dose response curve. Experimental results are presented on top of the parameter space created in Figure 2B by simulations. Blue and green arrows represent design actions to restore the monotonically increasing dose response curves starting from failed cascades CAS 1/30 and CAS 1/60, respectively. Mean values and standard deviations of fluorescence intensities at the steady state are calculated from three independent experiments analyzed by flow cytometry and normalized to the maximum value in each plot (see SI Section A7).

RBS strength and low DNA copy number. We simulated the dose response curves of activation cascades with different node 2 RBS strengths and DNA copy numbers, presented in the parameter space in Figure 2B. The lower left corner of the parameter space corresponds to the cascade with the smallest J_2, and the upper right corner corresponds to the largest J_2. In accordance with these predictions, simulations in Figure 2B confirms that smaller J_2 (weak x_2 RBS and low DNA copy number) results in monotonically increasing response (grid shaded region), while larger J_2 (strong x_2 RBS and high DNA copy number) results in monotonically decreasing response (gray region). The dotted region corresponds to intermediate values of J_2 which result in biphasic response.

Model-guided design recovers monotonically increasing response of the cascade

Based on the simulation map in Figure 2B and the mathematical analysis of model (4)-(6) described in the previous section, we created a library of activation cascades in which each cascade should result into one of the three different behaviors shown in Figure 2B. This library is composed of cascades that differ in the value of the resource demand coefficient of NahR (J_2), with the rationale that we can mitigate the strength of the key non-regulatory interaction $x_1 \dashv x_3$ to recover the intended monotonically increasing dose response curve of the cascade. In particular, starting from CAS 1/30, whose dose response curve is biphasic (Figure 3A), we designed circuit CAS 0.3/30 with about 30% RBS strength [12] of NahR compared to CAS 1/30, theoretically resulting in a reduction of J_2. We therefore expect a reduction of the $x_1 \dashv x_3$ interaction strength, leading to a monotonically increasing dose response curve, which is confirmed by the experiment (Figure 3B).

Similarly, we constructed another cascade circuit CAS 1/60 in which the DNA copy number is about twice as that of CAS 1/30 (about 60 vs 30). According to our model, resource demand coefficient of NahR J_2 in CAS 1/60 should double compared to that of circuit CAS 1/30. Therefore, we expect a possibly monotonically decreasing dose response curve. Experiments confirm this prediction (Figure 3C). A local increase in GFP fluorescence at about 10 nM AHL is due to the two-step multimerization of NahR proteins [35], which is detailed in SI Section A5. To obtain a monotonically increasing dose response curve from this circuit, we first reduced NahR resource demand coefficient J_2 by designing a circuit CAS 0.3/60, whose NahR RBS strength is 30% compared to that of CAS 1/60. Theoretically, depending on parameters, reduced J_2

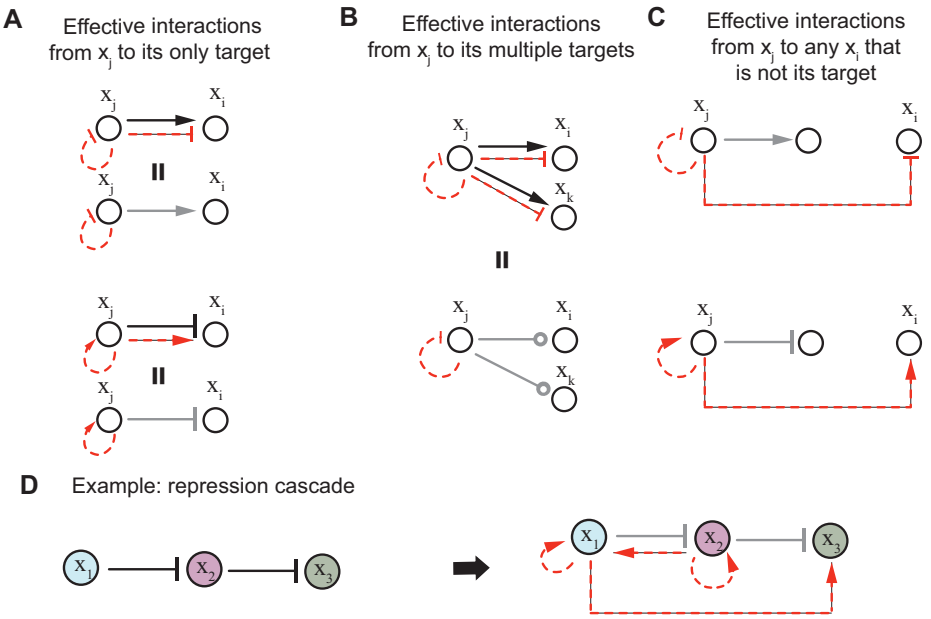


Figure 4: **Rules to determine effective interaction graphs arising from resource competition in a genetic circuit.** Black solid edges represent regulatory interactions; red dashed edges represent non-regulatory interactions arising from resource competition; if a black and a red edge have the same starting and ending nodes, we indicate their combined effect with a gray edge. (A) If TF x_j has only one target, then resource competition does not change the nature (activation or repression) of interaction from x_j to its target. However, it weakens the intended strength. (B) If TF x_j regulates multiple targets, then the effective interactions from x_j to its targets are undetermined. (C) If x_j is a transcriptional activator (repressor), then it becomes an effective repressor (activator) for all nodes that are not its target. (D) Applying the rules in A-C, we determine the effective interaction graph for a repression cascade.

can lead to either monotonically increasing or biphasic dose response curves (see Figure 2B). Our experiment show that the response of CAS 0.3/60 is indeed biphasic (Figure 3D). To restore a monotonically increasing dose response curve, we can further decrease J_2 by reducing DNA copy number to create circuit CAS 0.3/30, whose dose response curve is monotonically increasing (Figure 3B).

General rules to draw effective interactions in genetic circuits

Interaction graphs, which use directed edges to represent regulatory interactions, are a convenient graphical tool to design and/or analyze the qualitative behavior of a genetic circuit [31]. Here, we expand the concept of interaction graph to incorporate non-regulatory interactions due to resource competition. We call the resultant interaction graph an *effective interaction graph*, which includes both regulatory interactions and non-regulatory interactions due to resource competition. In an effective interaction graph, we draw $x \to y$ ($x \dashv y$) to represent effective activation (repression). We draw $x \multimap y$ if the interaction is undetermined, that is, it depends on parameters and/or x concentration.

The resource competition model (4)-(6) and the effective interaction graph identified in Figure 2A for the activation cascade can be generalized to any genetic circuit in a resource-limited environment. Each node i is a system that takes active TFs as inputs through the process of transcriptional regulation, and produce an active TF x_i as an output. Therefore, each node represents a dynamical process that can be captured by the ODE describing the rate of change of active x_i's concentration x_i. In a circuit with N nodes, we can write the dynamics of node i as (see SI Section B4 for derivation):

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \underbrace{\frac{T_i F_i(Q_i\mathbf{x}) \cdot H_i(I_i)}{1+\sum_{k=1}^{N} J_k F_k(Q_k\mathbf{x})}}_{G_i(\mathbf{x},I_i)\text{: effective node } i \text{ production}} - \underbrace{\gamma x_i}_{\text{dilution}}. \tag{10}$$

According to model (10), effective production rate of x_i, $G_i(\mathbf{x}, I_i)$, is jointly affected by transcriptional regulation $T_i F_i(Q_i\mathbf{x})$, allosteric modulation $H_i(I_i)$, and resource competition $R(\mathbf{x}) := [1+\sum_{k=1}^{N} J_k F_k(Q_k\mathbf{x})]$.

In particular, since resources are shared among all nodes, $R(\mathbf{x})$ is a common denominator to the effective production rate of every node. In model (10), T_i and J_i are lumped parameters that represent the maximum production rate of x_i and the maximum resource demand by node i, respectively (see equation (S70) and (7) for precise expressions). In particular, we gave the expression of J_i in equation (7) along with the explanation of its physical meaning. The binary matrix Q_i selects the TF inputs to node i (see SI Section B4.1 for precise definition), and the vector $\mathbf{x} = [x_1, \cdots, x_N]^T$ represents the concatenation of all active TFs' concentrations in the circuit. Active TFs are defined as proteins, either in inducer bound or unbound form, that can transcriptionally regulate a gene. Normalized Hill function $F_i(Q_i\mathbf{x})$ describes the transcriptional regulation of node i by its input TFs (see SI equation (S53) for precise expression). For those TFs whose activity can be allosterically modulated by an inducer I_i, normalized Hill function $H_i(I_i)$ represents the portion of TF that is active. If node i is not transcriptionally regulated, that is, x_i is constitutively expressed, then $F_i \equiv 1$. Similarly, if the activity of the TF x_i is not allosterically modulated by an inducer, then $H_i \equiv 1$. Parameter γ is a dilution rate constant that models cell growth.

Given model (10), a *regulatory interaction* from x_j to x_i is given by a non-zero $\partial F_i/\partial x_j$. It is an activation ($\rightarrow$) if $\partial F_i/\partial x_j > 0$ and a repression ($\dashv$) if $\partial F_i/\partial x_j < 0$. A *non-regulatory interaction* due to resource competition from x_j to any node x_i is present if $\partial R/\partial x_j$ is non-zero. Since $R(\mathbf{x})$ is in the denominator, conversely, the non-regulatory interaction is an activation ($\rightarrow$) if $\partial R/\partial x_j < 0$ and a repression ($\dashv$) if $\partial R/\partial x_j > 0$. The quantity $\partial R/\partial x_j$ captures the following physical phenomenon, which is responsible for non-regulatory interactions. As the concentration of active TF x_j increases, resource demand by the nodes that x_j activates/represses, which we call x_j's targets, increases/decreases; this in turn, reduces/increases the free amount of resources available to all nodes in the circuit. Therefore, the existence of a non-regulatory interaction originating from node j is exclusively dictated by the action (activation or repression) that TF x_j exerts on its targets; it is not dictated by the pure change in the concentration of TF x_j itself. *Effective interaction* from x_j to x_i represents the combined effect of regulatory and non-regulatory interaction from x_j to x_i, and is identified based on the sign of $\partial G_i/\partial x_j$.

Following the above, we list a set of immediate graphical rules to draw the effective interactions originating from x_j based on whether it is a transcriptional activator or repressor and based on the number of its targets (Figure 4A-C). These rules establish that when x_j transcriptionally regulates only one target, the nature of the effective interaction (i.e. activation vs. repression) from x_j to its target is unaffected by resource competition, but the strength of such interaction is weaker than the intended regulatory interaction. (Figure 4A). However, when x_j has multiple targets, the nature of effective interactions from x_j to its targets are undetermined (see Figure 4B and example in Figure S12). If x_j is a transcriptional activator (or repressor), then it is effectively repressing (or activating) all nodes that are not its targets, possibly including itself (Figure 4C). Detailed derivation of these graphical rules can be found in SI Section B6. Using these rules, the effective interaction graph of a two-stage activation cascade (Figure 2A) can be immediately identified. In Figure 4, we use black solid edges to represent regulatory interactions, and red dashed edges to represent non-regulatory interactions due to resource competition. If a black and a red edge have the same starting and ending nodes, we indicate their combined effect with a gray edge.

As an additional example of these graphical rules, we construct the effective interaction graph of a two-stage repression cascade in Figure 4D. Both x_1 and x_2 are repressors with only one target. Therefore, applying the rule in Figure 4A, we obtain $x_1 \dashv x_2$ and $x_2 \dashv x_3$. Since x_3 is not a target of x_1 and x_1 is a repressor, applying the rule in Figure 4C, we obtain $x_1 \rightarrow x_3$. Similarly, x_1 and x_2 are effectively activating themselves, and x_2 effectively activates x_1. Since x_3 does not transcriptionally activate or repress a target, there is no effective interaction originating from x_3. The resultant effective interaction graph in Figure 4D leads to a dose response curve that is monotonically increasing regardless of parameters (refer to SI Section B7). We can further use these interaction graphs to compare circuits with same functionality. Specifically, with a positive inducer input, the activation cascade of Figure 2A and the repression cascade of Figure 4D both are intended to have a monotonically increasing dose response curve. Since the repression cascade can keep this qualitative behavior in the face of resource competition, while the activation cascade may not, the former design is more robust to resource competition than the latter.

Discussion

Gene expression relies on transcriptional and translational resources, chiefly RNAPs and ribosomes. As all genes in a circuit compete for these limited resources, unintended non-regulatory interactions among genes arise. These interactions can dramatically change the intended behavior of a genetic circuit. In this paper, through a combined modeling and experimental study, we have characterized the extent to which resource competition affects a genetic circuit's behavior. We have incorporated resource competition into standard Hill-function models through resource demand coefficients, which can be readily tuned by key circuit parameters such as RBS strength and DNA copy number. These coefficients dictate the strengths

of non-regulatory interactions and can be effectively used to guide the design of a genetic circuit toward the intended behavior. Our mathematical model further provides a simple graphical tool to identify the nature of non-regulatory interactions (i.e. activation vs. repression) and to create the effective interaction graph of the circuit. Under the guidance of the model, we created a library of genetic activation cascades and demonstrated that, by tuning the resource demand coefficients of the cascade's nodes, the strengths of non-regulatory interactions can be predictably controlled and intended cascade's response can be restored.

Previous theoretical studies have analyzed how competition for shared resources affects gene expression. Using a stochastic model [13], Mather et al. found a strong anti-correlation of the proteins produced by ribosome-competing mRNAs. Rondelez [28] developed a general model to describe substrates competing for a limited pool of enzymes. De Vos et al. [11] analyzed the response of network flux toward changes in total competitors and common targets. More recently, Raveh et al. [14] developed a ribosome flow model to capture simultaneous mRNA translation and competition for a common pool of ribosomes. In [12], Gyorgy et al. developed a mechanistic resource competition model that gives rise to "isocost lines" describing tradeoffs in gene expression, which were experimentally validated. All these models, with the exception of [28], are restricted to circuits without regulatory links among competing nodes. In contrast, our general model explicitly accounts for regulatory interactions among nodes and reproduce the "isocost lines" of [12] as a special case (see SI Section B4.4). Furthermore, differently from [28], our model couples the resources' enzymatic reactions with the slower gene expression reactions to obtain a model for resource-limited genetic circuits.

Previous experimental studies have provided evidence that transcriptional and translational resources may be limited in the cell by showing that DNA copy number, mRNA concentration, and protein concentration do not always linearly correlate with each other [24, 26]. Accordingly, there has been extensive experimental evidence that synthetic genes' over-expression inhibits host cell growth [7, 36, 8, 9]. However, the effects of competition for shared resources on genetic circuits have only been recently addressed, mostly focusing on the single-gene effects as opposed to investigating the emergent effects at the network level [10, 12, 22, 37]. In this paper, we have theoretically predicted and experimentally demonstrated that significant network-level effects arise due to non-regulatory interactions dictated by resource competition. These interactions need to be accounted for in circuit design and optimization. Accordingly, we have provided a model-based approach to guide genetic circuit design to mitigate the effects of unintended interactions.

As a form of host-circuit interaction, previous studies have shown that overexpression of synthetic genes may retard host cell growth, which in turn affects dilution rates of the synthetic species [7, 8, 9, 38]. Our models (4)-(6) and (10) assume constant growth rate during circuit operation. This is consistent with our experiments, in which we limited the circuit plasmid copy number to only medium values to avoid substantial growth rate changes. In fact, in our experiments with CAS 0.3/30 and CAS 0.3/60, none to very modest changes in growth were observed. In experiments with CAS 1/60 and CAS 1/30, appreciable decrease in growth rates were observed only when AHL concentrations exceeded 10nM and 100nM, respectively (see Figure S4). However, all the unintended effects of resource competition investigated in this paper can be already observed for AHL concentrations lower than these.

As circuits grow in size and complexity, a "resource-aware" design approach needs to be adopted by synthetic biologists. While resource competition can be exploited in certain situations to our advantage [39, 40, 41], its global and nonlinear features largely hamper our capability to carry out predictive design. To alleviate the effects of resource competition, metabolic engineers down-regulate undesired gene expression to re-direct resources to the pathway of interest, thus increasing its yield [42, 43]. Similarly, in a genetic circuit, we can tune the resource demand coefficients of nodes by selecting appropriate RBS and DNA copy numbers to diminish the resource demand by certain nodes and hence make more resource available to other nodes. This tuning should be performed by keeping in mind other design specifications that the circuit may have, such as maximal output or sensitivity of the dose response curve [1]. A simulation example of how to relate easily tunable parameters, such as RBS strength and DNA copy number, to circuit's output is given in SI Section B3.3 for the genetic activation cascade. At the higher abstraction level of circuit topology, our model helps to identify topologies whose behavior is less sensitive to the effects of non-regulatory interactions. We provided an example of this with the two-stage activation and repression cascades. While the dose response curve of the former can be completely reshaped by non-regulatory interactions due to resource competition, the dose response curve of the latter is independent of resource competition.

Characterization of resource competition has deep implications in the field of systems biology, in which a major task is to reconstruct networks from data. In this case, it is critical to distinguish direct regulatory interactions from indirect ones [44], which may arise from non-regulatory interactions due to resource competition. In this sense, our model may provide deeper insights to guide the identification of natural networks from perturbation data.

Associated content

Methods and materials, detailed experimental data and mathematical models are described in Supporting Information.

Acknowledgements

We thank Fiona Chandra, Andras Gyorgy and Carlos Barajas for helpful discussions and suggestions. YQ and HHH were supported in part by AFOSR grant FA9550-14-1-0060 and ONR award N000141310074. JJ was supported by BBSRC grant BB/M009769/1.

Author contributions

All authors designed experiments and interpreted the data; HHH, JJ, YQ performed experiments; HHH and JJ cloned constructs; YQ developed the mathematical model and performed simulations; YQ and HHH wrote the paper; DDV edited the paper, designed the research and assisted with developing the mathematical model.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Brophy, J. A. N. and Voigt, C. A. (2014) Principles of genetic circuit design. Nat. Methods, **11**(5), 508–520.

[2] Cardinale, S. and Arkin, A. P. (2012) Contextualizing context for synthetic biology- identifying causes of failure of synthetic biological systems. Biotechnol. J., **7**, 856–866.

[3] Del Vecchio, D. (2015) Modularity, context-dependence, and insulation in engineered biological circuits. Trends Biotechnol., **33**(2), 111–119.

[4] Del Vecchio, D., Ninfa, A. J., and Sontag, E. D. (2008) Modular cell biology: retroactivity and insulation. Mol. Syst. Biol., **4**, 161.

[5] Jayanthi, S., Nilgiriwala, K. S., and Del Vecchio, D. (2013) Retroactivity Controls the Temporal Dynamics of Gene Transcription. ACS Synth. Biol., **2**(8), 431–441.

[6] Mishra, D., Rivera, P. M., Lin, A., Del Vecchio, D., and Weiss, R. (2014) A load driver device for engineering modularity in biological networks. Nat. Biotechnol., **32**, 1268–1275.

[7] Ceroni, F., Algar, R., Stan, G.-B., and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden.. Nat. Methods, **12**(5), 415–422.

[8] Gorochowski, T. E., Avcilar-kucukgoze, I., Bovenberg, R. A. L., Roubos, J. A., and Ignatova, Z. (2016) A Minimal Model of Ribosome Allocation Dynamics Captures Trade- offs in Expression between Endogenous and Synthetic Genes. ACS Synth. Biol., **5**(7), 710–720.

[9] Klumpp, S., Zhang, Z., and Hwa, T. (2009) Growth-rate Dependent Global Effect on Gene Expression in Bacteria. Cell, **139**, 1366–1375.

[10] Carbonell-Ballestero, M., Garcia-Ramallo, E., Montañez, R., Rodriguez-Caso, C., and Macía, J. (2015) Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. Nucleic Acids Res., **44**(1), 496–507.

[11] De Vos, D., Bruggeman, F. J., Westerhoff, H. V., and Bakker, B. M. (2011) How molecular competition influences fluxes in gene expression networks. PLoS One, **6**(12), e28494.

[12] Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H.-H., Chung, H., Weiss, R., and Del Vecchio, D. (2015) Isocost lines describe the cellular economy of gene circuits. Biophys. J., **109**(3), 639–646.

[13] Mather, W. H., Hasty, J., Tsimring, L. S., and Williams, R. J. (2013) Translation Cross Talk in Gene Networks. Biophys. J., **104**, 2564–2572.

[14] Raveh, A., Margaliot, M., Sontag, E. D., and Tuller, T. (2015) A Model for Competition for Ribosomes in the Cell. J. R. Soc. Interface, **13**, 20151062.

[15] Bashor, C. J. and Collins, J. J. (2012) Insulating gene circuits from context by RNA processing. Nat. Biotechnol., **30**, 1061–1062.

[16] Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., and Voigt, C. A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nat. Biotechnol., **30**(11), 1137–1142.

[17] Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q.-A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. Nat. Methods, **10**, 354–360.

[18] Nilgiriwala, K. S., Jiménez, J. I., Rivera, P. M., and Del Vecchio, D. (2015) Synthetic Tunable Amplifying Buffer Circuit in *E. coli*. ACS Synth. Biol., **4**(5), 577–584.

[19] An, W. and Chin, J. W. (2009) Synthesis of orthogonal transcription-translation networks. Proc. Natl. Acad. Sci. U. S. A., **106**(21), 8477–8482.

[20] Canton, B. Engineering the interface between cellular chassis and synthetic biological systems PhD thesis Massachusetts Institute of Technology (2008).

[21] Kushwaha, M. and Salis, H. M. (2015) A portable expression resource for engineering cross-species genetic circuits and pathways.. Nat. Commun., **6**, 7832.

[22] Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D., and Voigt, C. A. (2014) A "resource allocator" for transcription based on a highly fragmented T7 RNA polymerase. Mol. Syst. Biol., **10**, 742.

[23] Bremer, H. and Dennis, P. P. (1996) Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate. In Neidhardt, F. C., (ed.), Escherichia coli and Salmonella: Cellular and Molecular Biology, ASM Press.

[24] Churchward, G., Bremer, H., and Young, R. (1982) Transcription in Bacteria at Different DNA Concentrations. J. Bacteriol., **150**(2), 572–581.

[25] Stoebel, D. M., Dean, A. M., and Dykhuizen, D. E. (2008) The cost of expression of Escherichia coli lac operon proteins is in the process, not in the products. Genetics, **178**(3), 1653–1660.

[26] Vind, J., Sørensen, M. A., Rasmussen, M. D., and Pedersen, S. (1993) Synthesis of Proteins in Escherichia coli is Limited by the Concentration of Free Ribosomes: Expression from Reporter Gene does not always Reflect Functional mRNA Levels. J. Mol. Biol., **231**, 678–688.

[27] Siegal-gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V., and Murray, R. M. (2014) Gene circuit performance characterization and resource usage in a cell- free 'breadboard'. ACS Synth. Biol., **3**, 416–425.

[28] Rondelez, Y. (2012) Competition for catalytic resources alters biological network dynamics. Phys. Rev. Lett., **108**(1), 018102.

[29] Antunes, L. C. M., Ferreira, R. B. R., Lostroh, C. P., and Greenberg, E. P. (2008) A mutational analysis defines Vibrio fischeri LuxR binding sites. J. Bacteriol., **190**(13), 4392–4397.

[30] Hooshangi, S., Thiberge, S., and Weiss, R. (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc. Natl. Acad. Sci. U. S. A., **102**(10), 3584–3586.

[31] Alon, U. (2006) An Introduction to Systems Biology: Design Principles of Biological Circuits, Chapman & Hall/CRC Press, .

[32] Del Vecchio, D. and Murray, R. M. (2014) Biomolecular Feedback Systems, Princeton University Press, Princeton.

[33] Kaplan, S., Bren, A., Dekel, E., and Alon, U. (2008) The incoherent feed-forward loop can generate non-monotonic input functions for genes. Mol. Syst. Biol., **4**, 203.

[34] Kim, D., Kwon, Y.-K., and Cho, K.-H. (2008) The biphasic behavior of incoherent feed-forward loops in biomolecular regulatory networks. Bioessays, **30**(11-12), 1204–1211.

[35] Hoo, H. P., Woon, K. L., and Shin, H. J. (2005) In vitro binding of purified NahR regulatory protein with promoter Psal. Biochimica et Biophysica Acta - General Subjects, **1725**(2), 247–255.

[36] Dong, H., Nilsson, L., and Kurland, C. G. (1995) Gratuitous overexpression of gene in Escherichia coli leads to growth inhibition and ribosome destruction. J. Bacteriol., **177**(6), 1497–1504.

[37] Tabor, J. J., Bayer, T. S., Simpson, Z. B., Levy, M., and Ellington, A. D. (2008) Engineering Stochasticity in Gene Expression. Mol. Biosyst., **4**(7), 754–761.

[38] Tan, C., Marguet, P., and You, L. (2009) Emergent bistability by a growth-modulating positive feedback circuit. Nat. Chem. Biol., **5**(11), 842–848.

[39] Genot, A. J., Fujii, T., and Rondelez, Y. (2012) Computing with competition in biochemical networks. Phys. Rev. Lett., **109**(20), 208102.

[40] Kim, J., Hopfield, J., and Winfree, E. (2004) Neural Network Computation by *in vitro* Transcriptional Circuits. In Saul, L. K., Weiss, Y., and Bottou, L., (eds.), Advances in Neural Information Processing Systems Vol.17, pp. 681–688 MIT Press.

[41] Prindle, A., Selimkhanov, J., Li, H., Razinkov, I., Tsimring, L. S., and Hasty, J. (2014) Rapid and tunable post-translational coupling of genetic circuits.. Nature, **508**(7496), 387–391.

[42] Jones, K. L., Kim, S. W., and Keasling, J. D. (2000) Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab. Eng., **2**(4), 328–338.

[43] Solomon, K. V. and Prather, K. L. J. (2011) The zero-sum game of pathway optimization: Emerging paradigms for tuning gene expression. Biotechnol. J., **6**(9), 1064–1070.

[44] Kang, T., Moore, R., Li, Y., Sontag, E., and Bleris, L. (2015) Discriminating direct and indirect connectivities in biological networks. Proc. Natl. Acad. Sci. U. S. A., **112**(41), 12893–12898.