Minimal genetic device with multiple tunable functions

Sangram Bagh, Mahuya Mandal, and David R. McMillen*
Department of Chemical and Physical Sciences and Institute for Optical Sciences, University of Toronto Mississauga, 3359 Mississauga Rd N, Mississauga, Ontario, Canada L5L 1C6

(Received 29 April 2010; revised manuscript received 21 July 2010; published 12 August 2010)

The ability to design artificial genetic devices with predictable functions is critical to the development of synthetic biology. Given the highly variable requirements of biological designs, the ability to tune the behavior of a genetic device is also of key importance; such tuning will allow devices to be matched with other components into larger systems, and to be shifted into the correct parameter regimes to elicit desired behaviors. Here, we have developed a minimal synthetic genetic system that acts as a multifunction, tunable biodevice in the bacterium *Escherichia coli*. First, it acts as a biochemical AND gate, sensing the extracellular small molecules isopropyl β-D-1-thiogalactopyranoside and anhydrotetracycline as two input signals and expressing enhanced green fluorescent protein as an output signal. Next, the output signal of the AND gate can be amplified by the application of another extracellular chemical, arabinose. Further, the system can generate a wide range of chemically tunable single input-output response curves, without any genetic alteration of the circuit, by varying the concentrations of a set of extracellular small molecules. We have developed and parameterized a simple transfer function model for the system, and shown that the model successfully explains and predicts the quantitative relationships between input and output signals in the system.

DOI: 10.1103/PhysRevE.82.021911 PACS number(s): 87.18.Vf, 87.17.Aa, 87.85.md

I. INTRODUCTION

The field of synthetic biology includes an engineering effort to program living cells [1–9] much as computer scientists program a computer, by constructing and implementing artificial gene circuits that modify the cell's behavior. Although the design principles for engineering in a cellular context are much less well established than those in electrical engineering and circuit design, the rough analogy between electrical and genetic circuits proves helpful [1,6] in designing many artificial genetic circuits and devices. One of the immediate strategies of synthetic biology is to develop a toolbox [10,11] of well-characterized genetic circuits and devices that may be assembled in different ways to generate larger, more complex systems [3,7]. Recent efforts have yielded an ever-growing number of synthetic biological devices with varied functional capabilities, including Boolean logic gates [1,12,13], memory devices [14], switches [15,16], oscillators [16–19], amplifiers [20,21], receivers and senders [22], and measurement devices [23]. Generally, these devices have been developed and optimized to perform a single basic function.

The enterprise of connecting such synthetic genetic devices together into larger-scale systems will depend on the ability to interface the individual genetic components, match their input-output behaviors, and adjust their positions in parameter space so that they function as a whole. Adjusting devices by manipulating the system genetically is possible, but subject to severe limitations: the process is labor intensive and time consuming, making iterative testing difficult; and genetic alterations generally allow only discrete rather than continuous adjustments to parameter values. Though we do not offer a complete solution to this challenge, here, we propose a device with properties that will make it well suited to be integrated into larger systems, since its behavior can be adjusted over a wide range using external chemical inducers, without requiring genetic alterations. We have engineered a minimal synthetic genetic device that incorporates multiple tunable properties. It shows a logical AND behavior, integrating extracellular isopropyl β-D-1-thiogalactopyranoside (IPTG) and anhydrotetracycline (ATC) as input signals and expressing enhanced green fluorescent protein (EGFP) as its output signal. The output signal of this AND gate may be amplified by inducing with arabinose. The system may also be tuned to generate a range of chemically tunable single input-output signal-response curves (IPTG-EGFP and ATC-EGFP) by varying several extracellular chemicals.

II. RESULTS AND DISCUSSION

A. System design

The schematic design of the system is shown in Fig. 1(a). The DNA sequence of the operating sites of the proteins LacI, TetR, and AraC are fused together in a novel promoter we denote P_{LAT} [Fig. 1(a)] (a general principle adapted from Lutz and Bujard [24]). The promoter's rate of transcription initiation is regulated by three separate proteins, one (AraC) increasing the transcription rate and two others (LacI and TetR) decreasing the transcription rate. The two operating sites for protein AraC were placed upstream to a −33 sequence similar to the promoter $P_{lac/Ara-1}$ [24]; protein AraC binds to these two sites and in the presence of arabinose, it enhances transcription from our P_{LAT} promoter. The operating sites for TetR and LacI were placed between the −10 and −33 sites (similar to P_{TTL} [13]) and after the transcription start point, respectively. The proteins TetR and LacI bind with their respective operating sites and prevent RNA polymerase either from binding with the promoter (between −33

*david.mcmillen@utoronto.ca

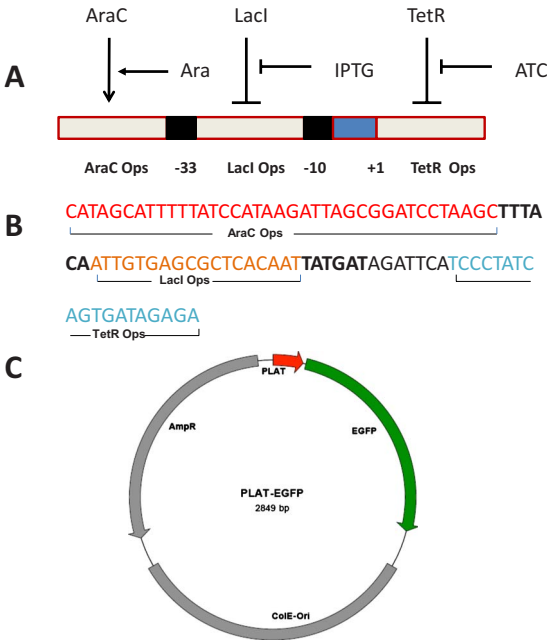

FIG. 1. (Color online) System design. (A) Schematic design of the synthetic promoter P_{LAT}. Two black squares represent the −33 and −10 hexamer sites, where RNA polymerase binds. The +1 site represents the transcription start point. The promoter contains operating sites for protein binding for the AraC, LacI and TetR proteins (from left to right). LacI and TetR repress transcription (represented by blunt-ended arrows). IPTG and ATC bind to LacI and TetR, respectively, and inactive the protein-DNA interactions. AraC and arabinose (Ara) together activate transcription (represented by arrows). (B) The DNA sequence (5′−3′) of the promoter. Operating sites for AraC, LacI, and TetR are indicated. The bolded sequences after the AraC and LacI operating sites denote the −33 and −10 hexamers. (C) Plasmid map for the PLAT-EGFP plasmid.

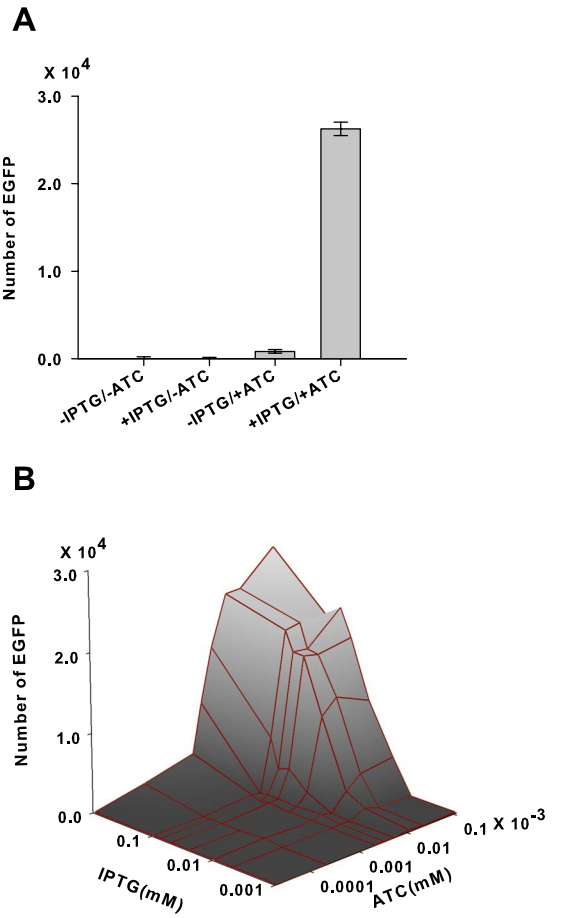

FIG. 2. (Color online) AND logic behavior. (a) Each bar represents the average EGFP number per cell grown from three independent colonies and measured in three independent experiments at four possible combinations of IPTG and ATC. +IPTG represents 1 mM concentration, −IPTG represents zero concentration. +ATC represents 1.08×10^{-4} mM concentration, −ATC represents zero concentration. Error bars correspond to one standard deviation. When both IPTG and ATC are present, high EGFP expression was observed (a digital ON state), while low EGFP expression is observed (digital OFF) under all other conditions; this demonstrates a logical AND operation is being carried out. The difference between the highest basal expression in the OFF state at −IPTG/+ATC and the ON state is over 30-fold. (b) Detailed EGFP expression levels as a function of IPTG and ATC concentrations. 81 different points in IPTG-ATC space were sampled experimentally, and mean EGFP expression levels are shown.

and −10) or from going forward with the transcription process (after the transcription start point). Thus, either TetR or LacI is able to repress transcription from promoter PLAT. Isopropyl β-D-1-thiogalactopyranoside (IPTG) and anhydrotetracycline (ATC) are two small molecules that bind with LacI and TetR respectively at their allosteric sites. This leads to a change in the conformation of the protein, in such a way that those proteins can no longer bind with the operating site of the promoter. Figure 1(b) gives the DNA sequence of the designed promoter. The EGFP is placed downstream of this promoter as a reporter through a strong ribosome binding site RBSII. The EGFP can be replaced with any protein encoding genes as required. This promoter-gene pair is inserted in a pZ based plasmid backbone with a ColE type origin of replication (maintaining approximately 50–70 copy numbers of the plasmid at log phase) and an ampicillin resistance gene as an antibiotic marker. The schematic of the plasmid is shown in Fig. 1(c). The plasmids reside in cells from the DH5αZ1 strain of *E. coli*, which natively produce the proteins LacI, AraC, and TetR [24].

B. Logical AND behavior and signal enhancement

The functional behavior of a genetic circuit is defined by the way it acts to transduce an input signal (arising from an extracellular stimulus or from some internal process) into an output response (usually the expression level of an output protein of interest). Figure 2(a) shows experimental results illustrating AND logic behavior. The EGFP expression inside the cell was monitored with ATC and IPTG at either zero concentration, or at saturating concentrations (1.08×10^{-4} mM for ATC, 1 mM for IPTG; saturation was confirmed by concentration sweeps similar to those shown in Fig. 6 (other experiments not shown). The cells showed a high EGFP expression (the ON state) when both ATC and IPTG were present, and a low basal level of expression (the OFF state) under any other combination of inputs; this implements the truth table of a logical AND between ATC

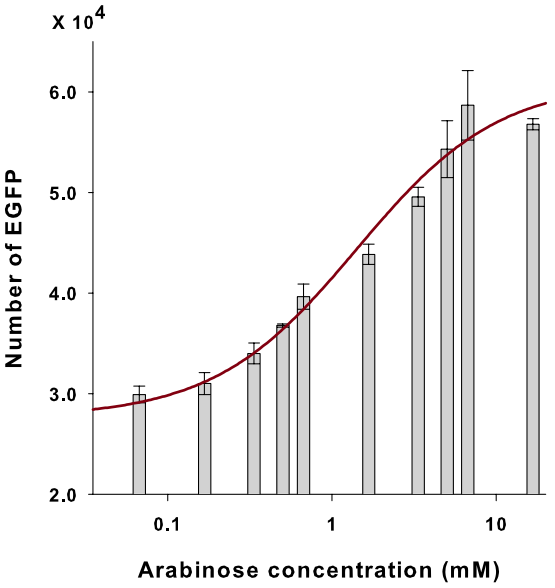

FIG. 3. (Color online) Varying output of the AND gate using arabinose. The arabinose concentration was varied systematically from 0 to 13.4 mM in the presence of saturating concentrations of IPTG (1 mM) and ATC (1.08×10^{-4} mM). Each bar represents the average EGFP number per cell at a particular arabinose concentration, with error bars representing one standard deviation from the mean over three experiments, performed on cell populations grown from three independent colonies. The solid line is the fit to Eq. (2); the numerical fit values are given in the text.

and IPTG. The minimum difference between the ON state (ATC and IPTG at saturated concentration) and the OFF state at 1.08×10^{-4} mM ATC, 0.0 mM IPTG (the highest basal level production) is over 30 times. To understand the detailed input-output relationship of the system, we measured the EGFP expression by systematically varying the concentrations of IPTG (0 to 1 mM) and ATC (0 to 1.08×10^{-4} mM) through 81 different combinations. The results are shown as a three dimensional surface plot [Fig. 2(b)], which shows a flat strip of OFF states parallel to IPTG axis until 2.16×10^{-6} mM of ATC, and another flat strip of OFF states parallel the ATC axis below 0.1 mM of IPTG. Beyond these regions there is a steep transition to a region of ON states; the sharpness of the transition represents a good approximation of digital behavior.

The output signal of this AND gate can tuned by applying another small diffusible chemical, arabinose. The increase of EGFP production with varying concentration of arabinose (at saturated concentration of IPTG and ATC; the ON state) is shown in Fig. 3. The magnitude of the AND gate's output can be changed by more than twofold as a function of arabinose concentration. This signal enhancement is continuous, easily implemented, and follows a hyperbolic relation with arabinose concentration, as described by the transfer function model, below.

C. Transfer function

The transfer function of a device (the transformation from its input to its output) is one of its most important quantitative features, and is widely used in engineering to characterize devices [11]. As for an electrical circuit, the transfer function of a genetic circuit can be defined as the input-output relationship of the system at steady state [3,5,25]. The transfer function of our system is a three dimensional function, where three input signals (IPTG, ATC, and arabinose) are transduced into a single output (expression of the protein EGFP). The Hill function description is a common method model the transfer function of a genetic regulatory module [5]. The rate of accumulation of EGFP inside the cell may be written as

$$\frac{d[\mathrm{EGFP}]}{dt} = k\left(b_1 + \frac{([\mathrm{IPTG}]/K_1)^{n_1}}{1+([\mathrm{IPTG}]/K_1)^{n_1}}\right) \times\left(b_2 + \frac{([\mathrm{ATC}]/K_2)^{n_2}}{1+([\mathrm{ATC}]/K_2)^{n_2}}\right) \times\left(b_3 + \frac{([\mathrm{Ara}]/K_3)^{n_3}}{1+([\mathrm{Ara}]/K_3)^{n_3}}\right) - k_d[\mathrm{EGFP}], \quad (1)$$

where: k is the scaling rate constant (this constant is modulated by the effects of the subsequent Hill functions representing the regulatory effects of each of the three chemical inducers, acting through their respective regulatory proteins); the K_i and n_i represent the Hill constants and coefficients; k_d represents the degradation rate of EGFP (including dilution due to cell growth); and the b_i are inducer-specific offsets representing the contribution of each operating site when its corresponding inducer is at a concentration of zero.

Setting Eq. (1) to zero to solve for steady state,

$$[\mathrm{EGFP}]_{ss} = \left(\frac{k}{k_d}\right)\left(b_1 + \frac{([\mathrm{IPTG}]/K_1)^{n_1}}{1+([\mathrm{IPTG}]/K_1)^{n_1}}\right) \times\left(b_2 + \frac{([\mathrm{ATC}]/K_2)^{n_2}}{1+([\mathrm{ATC}]/K_2)^{n_2}}\right) \times\left(b_3 + \frac{([\mathrm{Ara}]/K_3)^{n_3}}{1+([\mathrm{Ara}]/K_3)^{n_3}}\right). \quad (2)$$

To extract the parameter values of our system, we measured the steady-state EGFP number per cell by systematically varying the concentrations of each input chemical, arabinose (Fig. 3), IPTG [Fig. 4(a)], and ATC [Fig. 4(b)] separately while holding the other two constant at their saturated concentrations (1.1×10^{-4} mM for ATC, 1 mM for IPTG, and 6.7 mM for arabinose). (We note that in the literature, ATC is often reported in ng/mL, while arabinose is often reported in percentage weight by volume; the conversion factors are: 1 mM ATC$=4.63\times10^5$ ng/mL ATC; and 1 mM arabinose$=0.015\%$ weight/volume arabinose.) The results were fitted with Eq. (2), yielding the following parameters: $k/k_d=31\,000$ EGFP proteins per cell; $K_1=0.03$ mM; $K_2=3.56\times10^{-6}$ mM; $K_3=1.34$ mM; $n_1=3.7$ $n_2=2.0$; $n_3=1.0$; $b_1=0.09$; $b_2=0.005$; and $b_3=0.83$. After this initial fitting, we fixed the parameter values and tested the model's predictive capabilities under conditions not used for the fit, as described below.

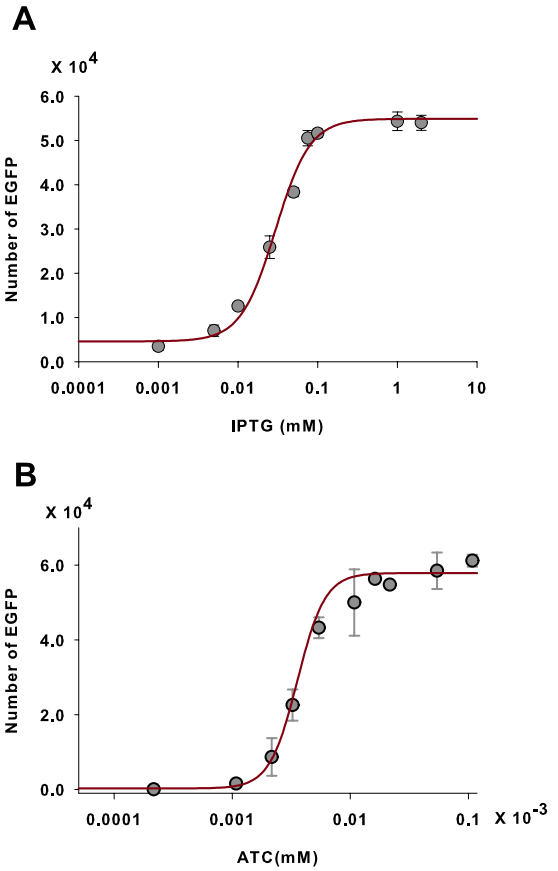

FIG. 4. (Color online) Steady-state measurements for parameter extraction. Solid circles represent the average number of EGFP per cell, measured over three experiments on cells grown from independent colonies, with error bars representing one standard deviation. Solid lines represent numerical fitting of Eq. (2), with parameter values as given in the text. (a) Steady-state EGFP expression as a function of IPTG concentration, at 1.1×10^{-4} mM ATC and 6.7 mM arabinose. (b) Steady-state EGFP expression as a function of ATC concentration, with 1 mM IPTG and 6.7 mM arabinose.

D. Chemically tunable signal-response curves

Genetic regulatory systems display sigmoidal signal-response curves as levels of protein expression or chemical inducer concentration are varied; such curves arise generically from the co-operative binding and saturation of a promoter by a regulatory protein. The ability to tune the signal-response curve of a circuit is a fundamental operation in synthetic biology: tuning enables designers of synthetic genetic systems to transform an input signal in ways appropriate to their particular application [3,5], before applying the transformed signal to the next step in their genetic circuit. A particular example is a two-step linear gene regulatory cascade, where the desired behavior of the second output module depends on the value of the gene expression rate [similar to the k' value in Eq. (3)] of the first module at saturated signal concentrations [3,26]. In the work reported in [3,26], the k value of the first module was changed (decreased) about tenfold by changing the DNA code of the ribosome binding sites. This mutation-based tuning is a standard technique in synthetic biology [3,5,27]. One limitation of this approach is that the parameter cannot be changed to any arbitrary, desired level, since there is no direct quantitative mapping available between DNA sequences and the resulting kinetic parameters. Another limitation is that the change in DNA sequence leads to discrete jumps in kinetic parameter values rather than continuous changes. In our system, Eq. (2) indicates that tuning of the k value (the scaling rate constant) in the Hill function can be achieved by using different combinations of chemical concentrations. If we want to study the signal-response curve of IPTG-EGFP from our system, at a constant ATC and arabinose concentration, Eq. (2) reduces to

$$[\mathrm{EGFP}]_{ss} = \left(\frac{k'}{k_d}\right)\left(b_1 + \frac{([\mathrm{IPTG}]/K_1)^{n_1}}{1+([\mathrm{IPTG}]/K_1)^{n_1}}\right), \tag{3}$$

where the k' (the scaling rate constant for EGFP expression at saturated IPTG concentrations) is

$$k' = k\left(b_2 + \frac{([\mathrm{ATC}]/K_2)^{n_2}}{1+([\mathrm{ATC}]/K_2)^{n_2}}\right)\left(b_3 + \frac{([\mathrm{Ara}]/K_3)^{n_3}}{1+([\mathrm{Ara}]/K_3)^{n_3}}\right). \tag{4}$$

Similarly, considering the ATC-EGFP signal-response curve, at constant IPTG and arabinose concentrations, Eq. (2) reduces to

$$[\mathrm{EGFP}]_{ss} = \left(\frac{k''}{k_d}\right)\left(b_2 + \frac{([\mathrm{ATC}]/K_2)^{n_2}}{1+([\mathrm{ATC}]/K_2)^{n_2}}\right), \tag{5}$$

where the k'' (the scaling rate constant for EGFP expression at saturated ATC concentrations) is

$$k'' = k\left(b_1 + \frac{([\mathrm{IPTG}]/K_1)^{n_1}}{1+([\mathrm{IPTG}]/K_1)^{n_1}}\right)\left(b_3 + \frac{([\mathrm{Ara}]/K_3)^{n_3}}{1+([\mathrm{Ara}]/K_3)^{n_3}}\right) \tag{6}$$

Figure 5 shows the model prediction for how these values, k' and k'', may be varied as functions of the chemical inducers. Figure 5(a) shows the value of k' as a function of arabinose and ATC concentration, while Fig. 5(b) shows the value of k'' as a function of arabinose and IPTG concentration.

The ability to vary these effective k values suggests that we should be able to vary the input-output characteristics of the system continuously over a wide range. To demonstrate this, and to test the model, we performed several experiments. Figure 6 shows the results of the transfer function model's prediction (lines) and experimental results (solid symbols) for: (a) IPTG-EGFP signal-response curves at different combinations of ATC and arabinose concentrations; and (b) ATC-EGFP signal-response curves at different IPTG and arabinose concentrations. The transfer function model provides a good quantitative match to the experimental results, in most cases; note that these predictions are being made under conditions not used in the initial parameter fitting. The two curves used for fitting in Figs. 3(a) and 3(b) are also included in Figs. 6(a) and 6(b), respectively, (solid triangles) for comparison.

Although the match between the experimental and predicted signal-response curves in Fig. 6 is generally good, there are cases in which the two curves diverge; note, for example, the curve with 3.24×10^{-6} mM ATC and zero ara-

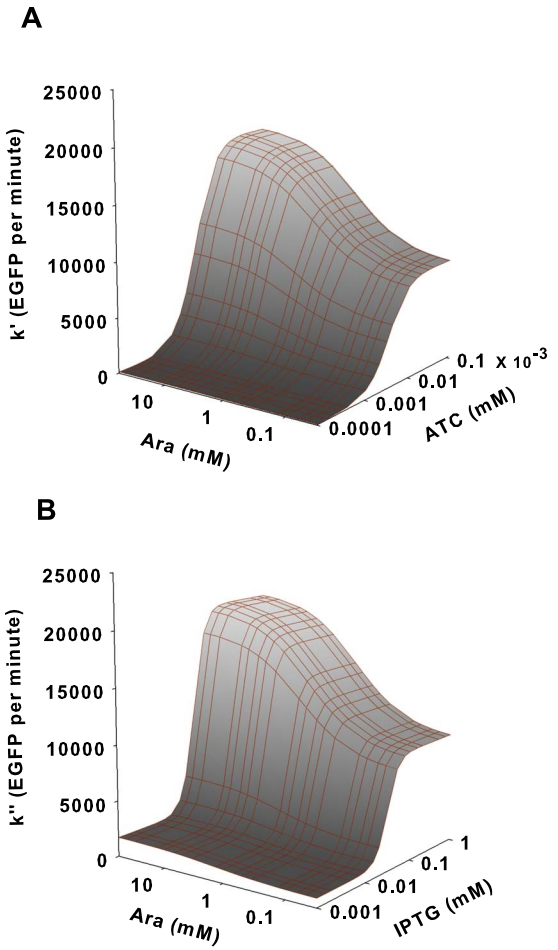

FIG. 5. (Color online) The theoretical calculation of the values of (a) k' (the scaling rate constant for EGFP expression at saturated IPTG concentrations) as a function of ATC and arabinose concentrations [from Eq. (4)], (b) k'' (the scaling rate constant for EGFP expression at saturated ATC concentrations) as a function of IPTG and arabinose concentration [from Eq. (6)].

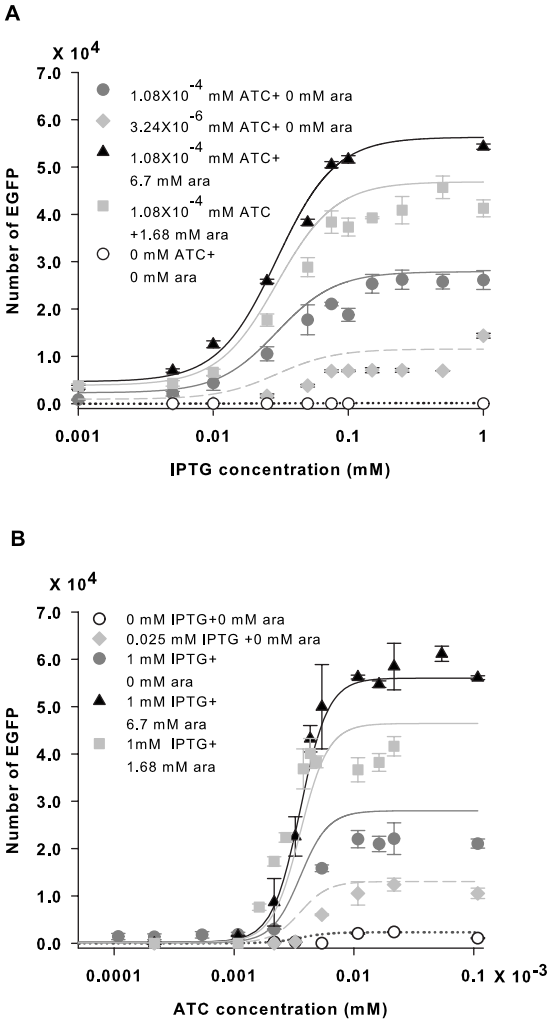

FIG. 6. Chemically tunable signal-response curves. Solid symbols represent experimental measurements of steady-state EGFP expression, averaged over three experimental runs from three independent colonies, with error bars representing one standard deviation. (a) EGFP vs IPTG signal-response curves, at varying ATC and arabinose (ara) concentrations. Curves represent model predictions from Eqs. (3) and (4) without fitting, for the following inducer concentrations:, 1.08×10^{-4} ATC+1.68 ara (light gray line), 1.08×10^{-4}+0 ara (dark gray line), 3.24×10^{-6} ATC+0 ara (dashed line), 0 ATC+0 ara (dotted line). The curve used for fitting [see Fig. 4(a)], 1.08×10^{-4} ATC+6.7 ara, is also shown (solid black line). The units of concentrations are mM in all cases. (b) EGFP vs. ATC signal-response curves, at varying ATC and arabinose (ara) concentrations. Curves represent model predictions from Eqs. (5) and (6) without fitting, for the following inducer concentrations: 1 IPTG+1.68 ara (light gray line), 1 IPTG+0 ara (dark gray line), 0.025 IPTG+0 ara (dashed line), 0 IPTG+0 ara (dotted line). The curve used in fitting [see Fig. 4(b)], 1 IPTG+6.7 ara, is also shown (black solid line). All concentrations are given in mM.

binose in Fig. 6(a), and the curve with 1 mM IPTG and 1.68 mM arabinose in Fig. 6(b). These discrepancies suggest that there are cases in which the assumptions in our model begin to break down. Several factors may underlie such discrepancies. The transfer function has been obtained, at a particular set of inducer concentrations, as a parameterized curve that does not incorporate any of the details of the underlying biochemical reactions, biochemical interactions, or the effect of cellular context. The values of the parameters in the transfer function equation are no doubt a complex function of protein-ligand (inducer) interactions, the resulting conformational changes in proteins, and protein-DNA operating site-ligand interactions, and effects operating on any of these levels have the potential to impair the accuracy of the model.

Perhaps the most significant assumption in the model is that we have chosen a form for our transfer function that assumes that the response to each inducer may be represented as an independent Hill function, with no "cross-talk" terms in which the effects of different operator sites, responding to their own inducers, influence the response to other inducers [28–34]. Inducers may influence the cooperative assembly of transcription factors, ligands, DNA operator sites, and other proteins and molecules, the details of which jointly determine the steepness of the dose-response curve [35] (as reflected in the Hill coefficient, n). Although the detailed mechanisms of cooperativity are unknown for most

transcription factors [32], recent studies have suggested a number of cases in which the effects of chemical inducers depend on the concentrations of other species: cooperative mechanisms in the tetracycline repressor, TetR, depend on tetracycline levels [30,31]; the central glycolytic genes repressor, CggR, is affected by fructose 1,6-bisphosphate concentration [32,33]; the response of the P_{lac} promoter to IPTG activation is influenced by levels of cAMP and the cAMP receptor protein [34]; and the PBAD and P_{lac} promoters can influence one another's responses when used in the same cell [29], with IPTG acting to inhibit expression from PBAD, where the inhibition depended on the relative levels of IPTG and arabinose present. Portions of our designed promoter, P_{LAT}, consist of operating sites similar to those in both the PBAD and P_{lac} promoters, and thus it is reasonable to expect that it may be subject to similar coupling between the effects of IPTG and arabinose. Any such cooperative effects would conflict with our assumption of independent response to each inducer, and cause deviations from the experimental observations. Correcting this would require either a significantly more detailed model, able to incorporate the mechanistic details of the interactions between inducers, or additional fitting of a modified model in the parameters in each of our previously independent Hill functions were permitted to be functions of the other inducer concentrations. We lack sufficient mechanistic information for the former approach, while the latter risks burying the general behavior under too much condition-specific detail. Given the reasonably good fit obtained with our relatively simple assumption of independence, for the moment we present only the noncoupled model, noting the issue of inducer cross-talk as an issue for future investigation.

Another factor that could have substantial effects on the gene circuit behavior is the coupling between gene circuits and the host cell's physiology. In recent work [36], the activation (by a single inducer) of a gene circuit in *E. coli* was shown to be capable of changing the rate of cell growth, which in turn affected the dilution rate of the expressed protein, directly affecting the steady-state behavior of a genetic circuit. In that study, bistable behavior was observed when a single step positive feedback gene regulatory cascade was coupled with the growth rate modulation originating from the activation of the same circuit. As different combinations of the three inducers (IPTG, ATC, and arabinose) activate our circuit in different ways, the growth rate of the cellular population (or subpopulation) could be modulated differently, which could affect the steady-state EGFP expression in ways not reflected in the model, with its assumption that background cellular behavior is independent of inducer concentration. Synthetic biology often relies on simple, phenomenological models [3,5] that can predict genetic circuit behavior with sufficient fidelity to guide the development of the system, but precise quantitative predictions may not not be possible without incorporating substantially more detail in terms of the underlying biochemical and cellular processes. We are far from a full quantitative description of the coupling between synthetic gene circuits and the global behavior of their host cells, but recent progress is encouraging [4,7,18,23,28,32,36,37].

The range of k values (effective scaling rate constant for EGFP expression) can be varied approximately 400-fold by changing the arabinose and IPTG or ATC concentrations (see Fig. 5). This change will allow us to produce ranges of different single input and output signal-response curves. By replacing the EGFP gene with any other gene of interest, the system may serve as a highly adjustable, continuously tunable module with predictable properties, ready to be interfaced with other biological devices.

III. CONCLUSIONS

We have developed a novel, minimal genetic device, which can perform multiple functions depending on where in its user-controllable parameter space it is placed. It shows a logical AND behavior, generating a strong output when two specific chemical signals are present, and producing outputs that are at least 30-fold lower under all other conditions. The magnitude of this logical output can be further amplified over a twofold range with the addition of the third chemical inducer. The same system can be used to generate a range of chemically tunable single-input signal-response curves. A simple transfer function model, populated using parameter values from one set of experiments, provides a generally good match to the observed experimental results under differing sets of conditions. The multifunctional nature of the device is an asset: by incorporating a range of behaviors into a single promoter, use of this system may reduce the number of nodes required in an engineered synthetic network. The chemical tuning of the system offers advantages over other techniques (such as altering the genetic makeup of a system) in terms of speed, predictability, and access to a continuous range of parameter values. This system will allow designers to explore different system behaviors with minimal time investment, and can serve as a multifunctional modular device in larger synthetic gene network designs.

IV. MATERIALS AND METHODS

Promoter, genes, plasmids, and cell strain. The designed promoter P_{LAT} (see results section for details of the design) was constructed by PCR amplification of two custom oligonucleotides (Eurofins MWG Operon, Huntsville, AL), which work both as primers and as templates. The sequence of the oligonuclides:

left primer

(5′ – 3′) GTCTTCACCTCGAGCATAGCATTTTTATCCATAAGATTAGCGGATCCTAAGCTTTACA

and right primer

(5′ – 3′) CTTTAATGAATTCTCTCTATCACTGATAGGGATGAATCTATCATAA
TTGTGAGCGCTCACAATTGTAAAGCTTAGGATCCGCTAA.

The oligonucleotides have 20 overlapping base pairs in the central region, as well as flanking regions. The flanking regions contain restriction sites *Kpn*I and XbaI. The plasmids were constructed using the pZE12-luc plasmid [24] as a backbone. The promoter pLlacO-1 and the luciferase gene were replaced by the P_{LAT} promoter (between the Xho-I and *Eco*RI restriction sites) and the EGFP gene (between the *Kpn*I and XbaI restriction sites). The EGFP gene was amplified by PCR from the pEGFP plasmid (Clontech, Mountain View, CA). PCR reactions were performed in a thermocycler (MJ Research, Waltham, MA) using Pfu Turbo Hotstart PCR Mastermix (Stratagene, La Jolla, CA). All restriction enzymes and T4 DNA ligase were obtained from New England Biolabs Canada (Pickering, ON). Plasmid extraction was carried out using QIAprep Spin Miniprep kits (Qiagen Canada, Mississauga, ON), and plasmid sequencing was done by The Centre for Applied Genomics (Toronto, ON). The plasmid PLAT-EGFP was transformed into the DH5αZ1 strain (Expressys, Ruelzheim, Germany) of *Escherichia coli* by electroporation (ECM 399 electroporator, BTX, San Diego, CA). The plasmid map is shown schematically in Fig. 1. The plasmid maintains 50–70 copies per cell [24] in the log phase of cell growth.

Cell growth and sample preparation. Cultures were grown overnight (16 h) in Lauria-Bertani (LB) medium (BioShop Canada, Burlington, ON) plus 100 ug/mL ampicillin (Amp) at 37 °C, from single colonies, then diluted 1:500 into fresh LB+Amp medium. Diluted cells were grown with different combinations of ATC, IPTG, and arabinose at varying concentrations (as described in the text) till mid log phase. Then the culture was spun down and washed twice with phosphate-buffered saline (PBS, *p*H 7.4) to minimize the background fluorescence from the medium. The washed cells were suspended in PBS and diluted to bring the cells into an appropriate concentration range (~2–5 times) before taking fluorimeter measurements.

Quantification of EGFP number. The fluorescence generated by EGFP expression was analyzed through a spectrofluorimeter by exciting the sample at 488 nm and collecting the spectra in the range of 500–560 nm. Correction for autofluorescence was performed by running cells without any plasmid. The cell concentration was measured by monitoring the optical density of the same sample at 600 nm and converting the value to cell concentration by using the relation previously developed in our laboratory [37]. The corrected fluorescence was converted into approximate number of EGFP per cell by comparing a calibration curve developed from standard EGFP solution (BioVision, Mountain View, CA) and dividing the EGFP concentration by cell concentration. The day to day instrumental variation of the spectrofluorimeter parameters was corrected by collecting spectra from a solid green standard. Further details of the EGFP quantification have been described previously [37].

Data analysis and modeling. All data analysis, curve fitting and transfer function model developed was carried out in Sigma Plot 10.0 (Systat, San Jose, CA).

ACKNOWLEDGMENTS

This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Ontario Research Fund (ORF), the Canada Foundation for Innovation (CFI), and the Canadian Institutes for Health Research (CIHR).

[1] R. Weiss, S. Basu, S. Hooshangi, A. Kalmbach, D. Karig, R. Mehreja, and I. Netravali, Nat. Comput. **2**, 47 (2003).
[2] J. Hasty, D. R. McMillen, and J. J. Collins, Nature (London) **420**, 224 (2002).
[3] M. Kaern, W. Blake, and J. J. Collins, Annu. Rev. Biomed. Eng. **5**, 179 (2003).
[4] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, Mol. Syst. Biol. **2**, 2006.0028 (2006).
[5] M. Kaern and R. Weiss, in *System Modeling in Cellular Biology*, edited by Z. Szallasi, J. Stelling, and V. Periwal (MIT Press, Cambridge, 2006), pp. 269–295.
[6] S. Mukherji and A. van Oudenaarden, Nat. Rev. Genet. **10**, 859 (2009).
[7] P. E. M. Purnick and R. Weiss, Nat. Rev. Mol. Cell Biol. **10**, 410 (2009).
[8] A. S. Khalil and J. J. Collins, Nat. Rev. Genet. **11**, 367 (2010).
[9] M. Pedersen and A. Phillips, J. R. Soc., Interface **6**, S437 (2009).
[10] P. M. Boyle and P. A. Silver, J. R. Soc., Interface **6**, S535 (2009).
[11] B. Canton, A. Labno, and D. Endy, Nat. Biotechnol. **26**, 787 (2008).
[12] J. C. Anderson, C. A. Voigt, and A. P. Arkin, Mol. Syst. Biol. **3**, 133 (2007).
[13] K. I. Ramalingam, J. R. Tomshine, J. A. Maynard, and Y. N. Kaznessis, Biochem. Eng. J. **47**, 38 (2009).
[14] C. M. Ajo-Franklin, D. A. Drubin, J. A. Eskin, E. P. Gee, D. Landgraf, I. Phillips, and P. A. Silver, Genes Dev. **21**, 2271 (2007).
[15] T. S. Gardner, C. R. Cantor, and J. J. Collins, Nature (London) **403**, 339 (2000).
[16] M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa,

Cell **113**, 597 (2003).
[17] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).
[18] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty, Nature (London) **456**, 516 (2008).
[19] T. Danino, O. Mondragon-Palomino, L. Tsimring, and J. Hasty, Nature (London) **463**, 326 (2010).
[20] D. Karig and R. Weiss, Biotechnol. Bioeng. **89**, 709 (2005).
[21] G. J. Nistala, K. Wu, C. V. Rao, and K. D. Bhalerao, J. Biol. Eng. **4**, 4 (2010).
[22] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, and R. Weiss, Nature (London) **434**, 1130 (2005).
[23] J. R. Kelly, A. J. Rubin, J. H. Davis, C. M. Ajo-Franklin, J. Cumbers, M. J. Czar, K. de Mora, A. L. Glieberman, D. D. Monie, and D. Endy, J. Biol. Eng. **3**, 4 (2009).
[24] R. Lutz and H. Bujard, Nucleic Acids Res. **25**, 1203 (1997).
[25] E. D. Conrad and J. J. Tyson, in *System Modeling in Cellular Biology*, edited by Z. Szallasi, V. Periwal, and S. Jorg (MIT Press, Cambridge, MA, 2006), pp. 97–123.
[26] R. Weiss and S. Basu, *First Workshop on Non-silicon Computation* (Cambridge, MA, 2002), pp. 54–61 (unpublished).
[27] J. Hasty, Proc. Natl. Acad. Sci. U.S.A. **99**, 16516 (2002).
[28] M. Hampf and M. Gossen, J. Mol. Biol. **365**, 911 (2007).
[29] S. Lee, H. Chou, B. Pfleger, J. Newman, Y. Yoshikuni, and J. Keasling, Appl. Environ. Microbiol. **73**, 5711 (2007).
[30] S. Reichheld and A. Davidson, J. Mol. Biol. **361**, 382 (2006).
[31] S. Reichheld, Z. Yu, and A. Davidson, Proc. Natl. Acad. Sci. U.S.A. **106**, 22263 (2009).
[32] D. Chaix, M. L. Ferguson, C. Atmanene, A. Van Dorsselaer, S. Sanglier-Cianferani, C. A. Royer, and N. Declerck, Nucleic Acids Res.10.1093/nar/gkq334 (2010).
[33] S. Zorrilla, T. Doan, C. Alfonso, E. Margeat, A. Ortega, G. Rivas, S. Aymerich, C. A. Royer, and N. Declerck, Biophys. J. **92**, 3215 (2007).
[34] T. Kuhlman, Z. Zhang, M. H. Saier, Jr., and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **104**, 6043 (2007).
[35] D. A. Burrill and P. A. Silver, Cell **140**, 13 (2010).
[36] C. Tan, P. Marguet, and L. You, Nat. Chem. Biol. **5**, 842 (2009).
[37] S. Bagh, M. Mazumder, T. Velauthapillai, V. Sardana, G. Q. Dong, A. B. Movva, L. H. Lim, and D. R. McMillen, Phys. Rev. E **77**, 021919 (2008).