COMPUTATIONAL MODELING AND QUANTITATIVE ANALYSIS OF AN IN VITRO DNA TRANSCRIPTIONAL SWITCH SYSTEM BASED ON EXPERIMENTAL DATA

KENNETH L. HO

Abstract. We present a computational kinetic model of the dynamics of a bistable *in vitro* biochemical feedback system of two logical transcriptional inverters. Through the systematic introduction of hypothesized interactions, we show that incorporating the effects of enzyme competition, abortive transcription and partial degradation, and RNA polymerase (RNAP) initial burst produces an accurate, predictive model capable of explaining experimental data to within a normalized mean square error of 0.01. We also use simulation results to partially characterize preliminary aspects of RNAP initial burst. Parameter fittings indicate that the rate of pre-steady-state RNA synthesis is at maximum approximately an order of magnitude above that at steady state, with a characteristic time decay constant on the scale of minutes. Furthermore, we propose a reduced-order model of the system, which simplifies the governing differential equations and is thus more analytically tractable.

1. Introduction

Synthetic biology is an emerging field of research in which biological principles are investigated through the rational design and engineering of novel biological systems through the manipulation of their constituent parts.[1,2] Previous work in synthetic biology has exploited the modularity of biological function to construct biochemical networks implementing binary switches,[3] inverters, and ring oscillators.[4] Motivation for such work comes from the prospect of using such constructs as components to assemble systems performing arbitrarily complex computations.[5]

Such rationally designed circuits are difficult to characterize and often exhibit unexpected departures from theoretically predicted behaviors. In this paper, we consider the DNA transcriptional feedback circuit proposed in Kim, Hopfield, and Winfree (2004).[5] In theory, this system manifests bistability with fast and sharp steady-state transfer curves. Actual laboratory implementation, however, has shown that this bistability is less than ideal. Although previous work has established a qualitative explanation of this inconsistency,[6,7] quantitative verification of these results remains open.

Quantitative models have been used to great effect in biology to describe, analyze, and characterize a wide variety of biochemical phenomena.[8–14] Thus, it is with this focus that we consider the quantitative treatment of this circuit to be critical to its understanding. In particular, we use simulation results to analyze the relative contributions of enzyme competition, abortive transcription[15] and partial degradation,[16] and RNA polymerase (RNAP) initial burst[17] to the system dynamics. These quantitative data provide the precision necessary in such systems engineering.

Thus, the quantitative model presented here of such general biochemical effects inherent in many biological systems may lead both to an improved understanding of naturally occurring networks and to an enhanced capacity to use such quantitative tools in their rational synthetic design.

2. Switch dynamics

Kim, Hopfield, and Winfree (2004)[5] proposed a construct for a synthetic bistable *in vitro* feedback switch system consisting only of DNA and RNA species in addition to the enzymes RNAP and ribonuclease H (RNaseH) using genetic regulatory control. Bistability is achieved by connecting two transcriptional inverters to form a mutually inhibitory network. In this system, the nucleic acid species form the transcriptional switches and act as signals that control the circuit. The enzyme RNAP binds to the switches to produce RNA transcript, which RNaseH degrades to induce signal decay.

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA.

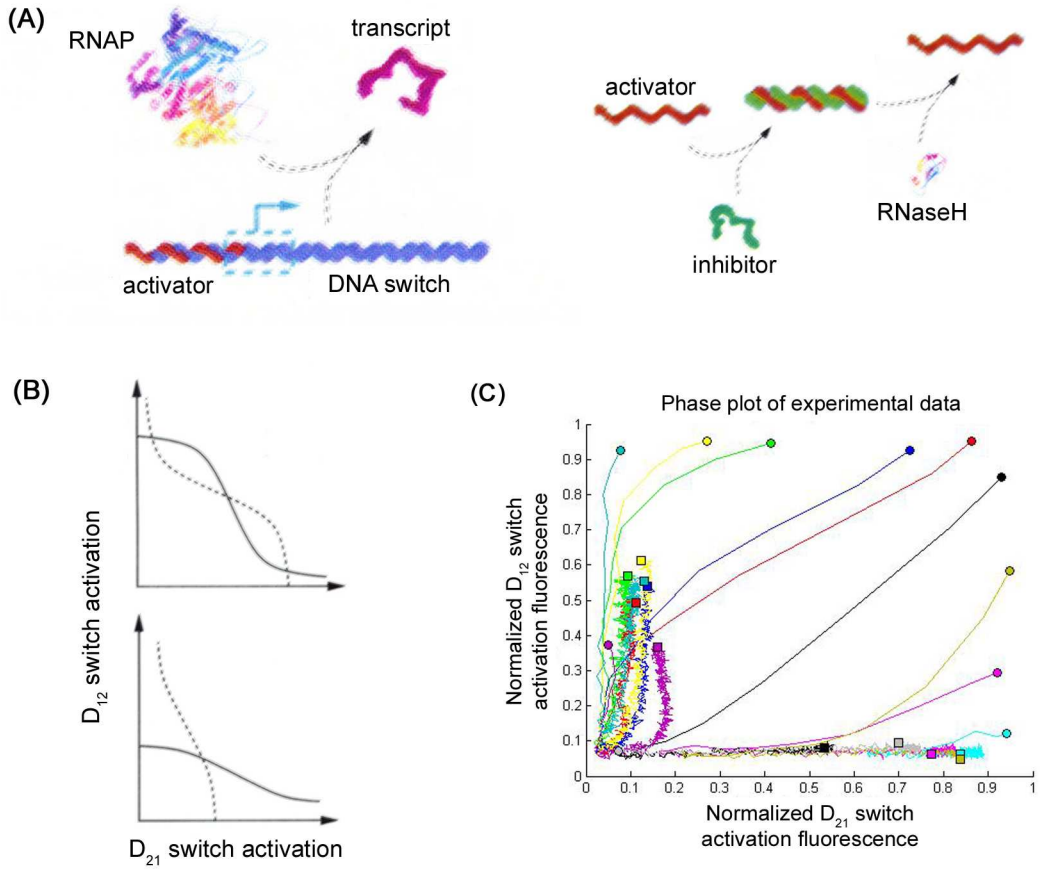

FIGURE 1. (A) DNA activator binds to a DNA switch, forming the activator-switch complex and completing the template promoter region, thus allowing RNAP to bind and synthesize the transcript. This process is disrupted by the introduction of an RNA inhibitor, which sequesters free activator by Watson-Crick base pairing. The resulting activator-inhibitor complex is degraded by RNaseH, which processes bound RNA, thus returning the original activator. Due to a designed toehold on the activator, free inhibitor can also displace bound activator from the switch by toehold binding followed by branch migration, resulting in an unbound DNA switch and an activator-inhibitor complex. Adapted from Kim, Hopfield, and Winfree (2004). (B) Hypothetical steady states of the bistable switch system. Solid curves trace out the nullclines of the first switch (D_{21}), and dashed curves trace out the nullclines of the second switch (D_{12}). The intersections of the two curves mark steady states of the system. Top, high RNAP concentration, resulting in three steady states, one of which is metastable, with each switch operating at an intermediate activity level. Bottom, decreasing RNAP concentrations cause a bifurcation in the system, resulting in three degenerate steady states. Reproduced from Dunlop *et al.* (2005). (C) Phase plot of experimental data. Eleven experimental runs are shown, with circles denoting initial system states and squares denoting final system states. Steady states appear to be located in a narrow strip of the dominant switch at $\hat{D}_{21} \approx 0.78$, $\hat{D}_{12} \approx 0.08$ and $\hat{D}_{21} \approx 0.1$, $\hat{D}_{12} \approx 0.55$ (where hat notation denotes normalized fluorescence).

The bistable switch system is constructed from two individual DNA transcriptional switches. Each switch is a synthetic double-stranded DNA template with an incomplete promoter region. Without bound single-stranded DNA activator, which is Watson-Crick complementary to the single-stranded extension of the switch promoter, RNAP cannot bind effectively and transcription cannot efficiently occur (Figure 1A). By design, however, single-stranded RNA inhibitor can displace bound activator from the switch. Thus, the switch assumes two operational states characterized by high and low transcription, respectively. The inhibitor can also bind to free activator, to which it is complementary, to form an activator-inhibitor complex, on which RNaseH acts to degrade bound inhibitor.

A feedback circuit is constructed out of two such DNA transcriptional switches by letting the transcript of the first switch be the inhibitor of the second, and vice versa. The bistability of this system is both theoretically and experimentally demonstrated (Figure 1B, C).

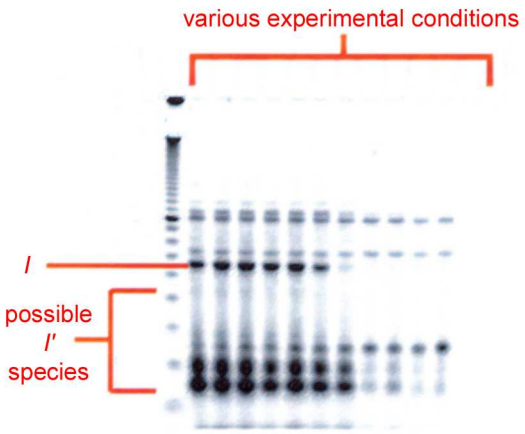

FIGURE 2. Electrophoresis gel showing the nucleic acids resulting from experimental runs of the bistable switch system. The dark bands near the bottom indicate the existence of significant amounts of partially transcribed or degraded inhibitor (I'), with a global length preference consistent across all lanes. Courtesy of Jongmin Kim.

3. Dynamical models

According to the interactions thus described, we develop several mathematical models for the dynamics of the bistable switch system. All are deterministic, continuous-time, differential equation models.

3.1. **Minimal model.** We begin by considering only the minimal set of essential biochemical reactions. Let D_{ij} be the DNA switch coding for inhibitor I_i and inhibited by inhibitor I_j, A_j be the activator that binds to I_j to form the activator-inhibitor complex AI_j, and DA_{ij} be the activator-switch complex formed upon binding of D_{ij} and A_j, where $(i,j) = (2,1)$ for the first switch and $(1,2)$ for the second. Assuming switch symmetry, the system can be described by the following reactions:

$$\begin{aligned}
A_j + I_j &\xrightarrow{k_1} AI_j \\
D_{ij} + A_j &\xrightarrow{k_2} DA_{ij} \\
DA_{ij} + I_j &\underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} D_{ij} + AI_j \\
DA_{ij} &\xrightarrow{k_p} DA_{ij} + I_i \\
D_{ij} &\xrightarrow{\alpha k_p} D_{ij} + I_i \\
AI_j &\xrightarrow{k_d} A_j,
\end{aligned}$$

where $0 < \alpha < 1$, for each set of (i,j). The system dynamics are then given by mass action:

$$\begin{aligned}
[\dot{D_{ij}}] &= -k_2[D_{ij}][A_j] + k_3[DA_{ij}][I_j] \\
&\quad - k_{-3}[D_{ij}][AI_j] \\
[\dot{A_j}] &= -k_1[A_j][I_j] - k_2[D_{ij}][A_j] + k_d[AI_j] \\
[\dot{I_j}] &= -k_1[A_j][I_j] - k_3[DA_{ij}][I_j] + k_{-3}[D_{ij}][AI_j] \\
&\quad + k_p\left([DA_{ji}] + \alpha[D_{ji}]\right) \\
[\dot{DA_{ij}}] &= k_2[D_{ij}][A_j] - k_3[DA_{ij}][I_j] \\
&\quad + k_{-3}[D_{ij}][AI_j] \\
[\dot{AI_j}] &= k_1[A_j][I_j] + k_3[DA_{ij}][I_j] - k_{-3}[D_{ij}][AI_j] \\
&\quad - k_d[AI_j],
\end{aligned}$$

where bracket notation denotes species concentration and dot notation denotes differentiation with respect to time.

3.2. **Full model.** Given that even the simplest of biological systems encompass a high degree of complexity, however, we should not expect the minimal model presented in the previous subsection to be an effective description of reality. In practice, switch kinetics are

asymmetric, transcription does not occur at a constant rate,[17] and key enzymes are often nonprocessive, generating significant amounts of partial products (Figure 2).[15,16] Indeed, previous research highlights the particular importance of considering the nonlinear effects induced by the competition among switches for a shared limited enzyme pool.[6,7] We therefore extend the minimal model in parts below. In this way, we implement functional modules which may be combined together to give a full model.

3.2.1. *Abortive transcription and partial degradation.* We introduce β_j, the probability of abortive transcription,[15] and the chemical species I'_j, AI'_j, and DAI'_{ij}, where I'_j collectively denotes all partial inhibitor strands which do not suppress transcription upon binding to DA_{ij}, with combinations of different species denoting complexes of the appropriate constituents. An inhibitory definition of I'_j was also used, where I'_j blocks the binding region of A_j so that AI'_j is prohibited from binding to D_{ij}. Simulation results under this definition, however, showed a gradual fall of switch activity instead of achieving steady state, a phenomenon which was not experimentally observed. Furthermore, we assume that RNaseH is strictly nonprocessive, hence RNA degradation is fully incomplete.[16] The complete set of reactions is then:

$$
\begin{aligned}
A_j + I_j &\xrightarrow{k_{1,j}} AI_j \\
D_{ij} + A_j &\xrightarrow{k_{2,j}} DA_{ij} \\
DA_{ij} + I_j &\underset{k_{-3,j}}{\overset{k_{3,j}}{\rightleftharpoons}} D_{ij} + AI_j \\
A_j + I'_j &\underset{k_{-4,j}}{\overset{k_{4,j}}{\rightleftharpoons}} AI'_j \\
DA_{ij} + I'_j &\underset{k_{-4,j}}{\overset{k_{4,j}}{\rightleftharpoons}} DAI'_{ij} \\
D_{ij} + AI'_j &\xrightarrow{k_{2,j}} DAI'_{ij} \\
DA_{ij} &\xrightarrow{(1-\beta_j)k_{p,j}} DA_{ij} + I_i \\
DA_{ij} &\xrightarrow{\beta_j k_{p,j}} DA_{ij} + I'_i \\
D_{ij} &\xrightarrow{\alpha_j(1-\beta_j)k_{p,j}} D_{ij} + I_i \\
D_{ij} &\xrightarrow{\alpha_j\beta_j k_{p,j}} D_{ij} + I'_i \\
DAI'_{ij} &\xrightarrow{(1-\beta_j)k_{p,j}} DAI'_{ij} + I_i \\
DAI'_{ij} &\xrightarrow{\beta_j k_{p,j}} DAI'_{ij} + I'_i \\
AI_j &\xrightarrow{k_{d,j}} AI'_j,
\end{aligned}
$$

with dynamics again given by mass action:

$$
\begin{aligned}
[\dot{D_{ij}}] &= -k_{2,j}[D_{ij}][A_j] + k_{3,j}[DA_{ij}][I_j] \\
&\quad - k_{-3,j}[D_{ij}][AI_j] - k_{2,j}[D_{ij}][AI'_j] \\
[\dot{A_j}] &= -k_{1,j}[A_j][I_j] - k_{2,j}[D_{ij}][A_j] - k_{4,j}[A_j][I'_j] \\
&\quad + k_{-4,j}[AI'_j] \\
[\dot{I_j}] &= -k_{1,j}[A_j][I_j] - k_{3,j}[DA_{ij}][I_j] \\
&\quad + k_{-3,j}[D_{ij}][AI_j] + (1-\beta_i)\,k_{p,i} \\
&\quad \left([DA_{ji}] + \alpha_i[D_{ji}] + [DAI'_{ji}]\right) \\
[\dot{DA_{ij}}] &= k_{2,j}[D_{ij}][A_j] - k_{3,j}[DA_{ij}][I_j] \\
&\quad + k_{-3,j}[D_{ij}][AI_j] - k_{4,j}[DA_{ij}][I'_j] \\
&\quad + k_{-4,j}[DAI'_{ij}] \\
[\dot{AI_j}] &= k_{1,j}[A_j][I_j] + k_{3,j}[DA_{ij}][I_j] \\
&\quad - k_{-3,j}[D_{ij}][AI_j] - k_{d,j}[AI_j] \\
[\dot{I'_j}] &= -k_{4,j}[A_j][I'_j] + k_{-4,j}[AI'_j] - k_{4,j}[DA_{ij}][I'_j] \\
&\quad + k_{-4,j}[DAI'_{ij}] + \beta_i k_{p,i} \\
&\quad \left([DA_{ji}] + \alpha_i[D_{ji}] + [DAI'_{ji}]\right) \\
[\dot{AI'_j}] &= k_{4,j}[A_j][I'_j] - k_{-4,j}[AI'_j] - k_{2,j}[D_{ij}][AI'_j] \\
&\quad + k_{d,j}[AI_j] \\
[\dot{DAI'_{ij}}] &= k_{4,j}[DA_{ij}][I'_j] - k_{-4,j}[DAI'_{ij}] \\
&\quad + k_{2,j}[D_{ij}][AI'_j].
\end{aligned}
$$

3.2.2. *Enzyme competition.* Michaelis-Menten enzyme kinetics[11,14] was then used to account for the nonlinear kinetic behavior induced by the saturation of a limited enzyme pool. We enumerate the loads on the two system enzymes RNAP and RNaseH, respectively:

$$
\begin{aligned}
L_{p,j} &\equiv \frac{[DA_{ij}] + [DAI'_{ij}]}{K_{Mp}} + \frac{[D_{ij}]}{\tilde{K}_{Mp}} \\
L_{d,j} &\equiv \frac{[AI_j]}{K_{Md}},
\end{aligned}
$$

where tilde notation denotes reference to inactivated switches, and define the corresponding enzymatic rates as rational functions of the system states:

$$
\begin{aligned}
\overline{k}_{p,j} &\equiv \frac{[\mathrm{RNAP}]k_{cp,j}/K_{Mp}}{1 + L_{p,j}} \\
k_{d,j} &\equiv \frac{[\mathrm{RNaseH}]k_{cd,j}/K_{Md}}{1 + L_{d,j}},
\end{aligned}
$$

where the constants k_c are the appropriate catalytic rate constants and the constants K_M are the corresponding Michaelis-Menten constants. In particular,

this gives an expression for α_j:

$$\alpha_j = \frac{\tilde{k}_{cp,j} K_{Mp}}{k_{cp,j} \tilde{K}_{Mp}},$$

which provides a measure of the basal transcription level of the DNA switches.

3.2.3. *RNAP initial burst.* Finally, we consider the RNAP initial burst effect, in which the rate of RNA synthesis is observed to decay from an initial peak of about thirty-fold that of steady-state levels.[17] This effect has not yet been well characterized, but we postulate that a reasonable mathematical formulation is

$$k_{p,j} \equiv \overline{k}_{p,j} \left(A_j e^{-t/\tau_j} + 1 \right).$$

Comparisons of simulation and experimental data allow fitting of A_j and τ_j, which is fairly straightforward as the two parameters control essentially independent aspects of the transcription function.

3.3. **Reduced-order model.** We now derive a reduced-order version of the full model discussed above, from which two conserved quantities are immediate for each (i, j):

$$\mu_j \equiv [D_{ij}] + [DA_{ij}] + [DAI'_{ij}]$$
$$\rho_j \equiv [A_j] + [DA_{ij}] + [AI_j] + [AI'_j] + [DAI'_{ij}],$$

hence

$$[DA_{ij}] = \mu_j - [D_{ij}] - [DAI'_{ij}]$$
$$[AI_j] = \rho_j - [A_j] - [DA_{ij}] - [AI'_j] - [DAI'_{ij}].$$

From parameter fittings, we observe that $k_{-3,j}$ and $k_{-4,j}$ are significantly smaller than the other rate constants (for numerical results, see section 4). This lends itself naturally to a multiple-time scale approach. We multiply the differential equation for $[D_{ij}]$ by $\epsilon = 10^{-2}$ and assume that all species concentrations are on the same order, thus allowing the omission of all terms with effective rate coefficients of magnitude less than ϵ. Using the conservation equation for $[DA_{ij}]$, we find that

$$[D_{ij}] = \frac{k_{3,j} \left(\mu_j - [DAI'_{ij}] \right) [I_j]}{k_{2,j} \left([A_j] + [AI'_j] \right) + k_{3,j} [I_j]}.$$

This reduces the full system of sixteen coupled differential equations to a differential-algebraic system of ten differential equations and six algebraic relations. Alternatively, we may make the approximation

$$\dot{[A_j]} = -k_{1,j}[A_j][I_j] - k_{2,j}[D_{ij}][A_j] - k_{4,j}[A_j][I'_j],$$

hence by direct integration,

$$[A_j(t)] = [A_j(t_0)] \exp\left[-f(t, t_0)\right],$$

where

$$f(t, t_0) \equiv k_{1,j} \int_{t_0}^{t} [I_j(\tau)]\, d\tau + k_{2,j} \int_{t_0}^{t} [D_{ij}(\tau)]\, d\tau + k_{4,j} \int_{t_0}^{t} [I'_j(\tau)]\, d\tau.$$

4. Simulation results

We use experimental data provided by Jongmin Kim and Erik Winfree. Simulation data were generated from initial experimental states and compared with experimental observations. Errors were calculated at each data point for each switch and then averaged. A normalized mean square error (NMSE) of 0.01 was chosen to correspond to a root mean square error of 0.1, which is 10% of the normalized signal span. By the NMSE of a data set, we will mean the average NMSE of the individual data in that set.

The minimal model derived in subsection 3.1 was initially simulated using the parameter values in Table 2. Maximal activity was observed for both switches, contrary to experimental observation. Consequently, primary rate constants (k_1, k_2) were decreased by several orders of magnitude and the assumption of symmetry was removed. Least-squares parameter fitting for the rate constants produced values which achieved a NMSE of 6.166×10^{-3}.

The hypothesized interactions of enzyme competition, abortive transcription and partial degradation, and RNAP initial burst were investigated next. Table 1 demonstrates the potential of modeling these effects. Each reported NMSE is the best error achieved by least-squares parameter fitting. We found that enzyme competition is the dominant effect in the system, capturing the general character

Table 2. Parameters used in the kinetic simulation of the minimal model.

Parameter	Value
k_1	10^6 M^{-1} s^{-1}
k_2	10^6 M^{-1} s^{-1}
k_3	10^5 M^{-1} s^{-1}
k_{-3}	1 M^{-1} s^{-1}
k_p	1.5×10^{-2} s^{-1}
k_d	1.7×10^{-3} s^{-1}
α	0.04

TABLE 1. Errors produced by simulations under models incorporating various combinations of enzyme competition, abortive transcription and partial degradation (enzyme nonprocessivity), and RNAP initial burst. These errors characterize the explanatory capacity of each model as the parameters used were obtained through least-squares parameter fitting against experimental data. Percent contributions were computed relative to the decrease in NMSE from the minimal to the full model.

Enzyme competition	Enzyme nonprocessivity	RNAP initial burst	NMSE	% contribution
Disabled	Disabled	Disabled	6.166×10^{-3}	0
Enabled	Disabled	Disabled	4.949×10^{-3}	52.389
Disabled	Enabled	Disabled	1.067×10^{-2}	−193.89
Disabled	Disabled	Enabled	5.644×10^{-3}	22.471
Enabled	Enabled	Disabled	4.719×10^{-3}	62.290
Enabled	Disabled	Enabled	3.923×10^{-3}	96.556
Disabled	Enabled	Enabled	5.732×10^{-3}	18.683
Enabled	Enabled	Enabled	3.843×10^{-3}	100

of the data and by itself accounting for 52.389% of the error drop from the minimal model to the full model (all features enabled). RNAP initial burst is second in importance, accurately tracing the starting trajectories though failing on intermediate time scales (22.471% error drop). Together, enzyme competition and RNAP initial burst accounted for 96.556% of the final drop.

Least-squares parameters fitted for the full model are given in Table 3. Characteristic plots of simulation versus experimental data using these parameters are shown in Figure 3 (NMSE $= 3.843 \times 10^{-3}$). When run against a second experimental data set provided midway through the work emphasizing higher switch concentrations and a lower RNAP concentration, however, these parameters failed in predicting the long-term behavior of the system (Figure 4A), resulting in a NMSE of 1.515×10^{-1}, an error increase of over two orders of magnitude. As the switch ratios remained essentially unchanged across the two data sets (see subsection 6.3), we hypothesized that the primary cause of disagreement is the reduction in transcription efficiency due to the lower RNAP concentration. To test this hypothesis, a simulation was run on the second data set assuming the RNAP concentration of the first, resulting in a NMSE of 2.097×10^{-2}. This value is misleading, however, as NMSEs for all experimental runs contained in this set were on the order of 10^{-3} or less, with the exception of experimental run 4 (NMSE $= 2.192 \times 10^{-1}$). With this data set omitted, the NMSE drops to an impressive 2.947×10^{-3}.

To further investigate the dynamics observed, we ran a least-squares parameter fit on the second data set to produce an alternate set of parameters (NMSE $= 2.188 \times 10^{-3}$).

TABLE 3. Parameters for the full model obtained by least-square parameter fitting.

Parameter	Value
$k_{1,1}$	2.353×10^3 M^{-1} s^{-1}
$k_{2,1}$	3.827×10^3 M^{-1} s^{-1}
$k_{3,1}$	1.144×10^4 M^{-1} s^{-1}
$k_{-3,1}$	1.014 M^{-1} s^{-1}
$k_{4,1}$	2.402×10^5 M^{-1} s^{-1}
$k_{-4,1}$	2.329×10^1 s^{-1}
$k_{1,2}$	4.675×10^3 M^{-1} s^{-1}
$k_{2,2}$	1.160×10^3 M^{-1} s^{-1}
$k_{3,2}$	2.046×10^4 M^{-1} s^{-1}
$k_{-3,2}$	1.006 M^{-1} s^{-1}
$k_{4,2}$	6.197×10^3 M^{-1} s^{-1}
$k_{-4,2}$	4.274 s^{-1}
$k_{cp,1}$	3.334×10^{-2} s^{-1}
$\tilde{k}_{cp,1}$	2.860×10^{-2} s^{-1}
$k_{cp,2}$	9.018×10^{-2} s^{-1}
$\tilde{k}_{cp,2}$	1.366×10^{-2} s^{-1}
K_{Mp}	1.401×10^{-8} M
$\tilde{K}_{Mp}$	5.588×10^{-8} M
$k_{cd,1}$	3.086×10^{-1} s^{-1}
$k_{cd,2}$	1.996×10^{-1} s^{-1}
K_{Md}	2.895×10^{-7} M
β_1	0.232
β_2	0.195
A_1	7.604, $A_2 = 8.096$
τ_1	87.375 s
τ_2	67.156 s

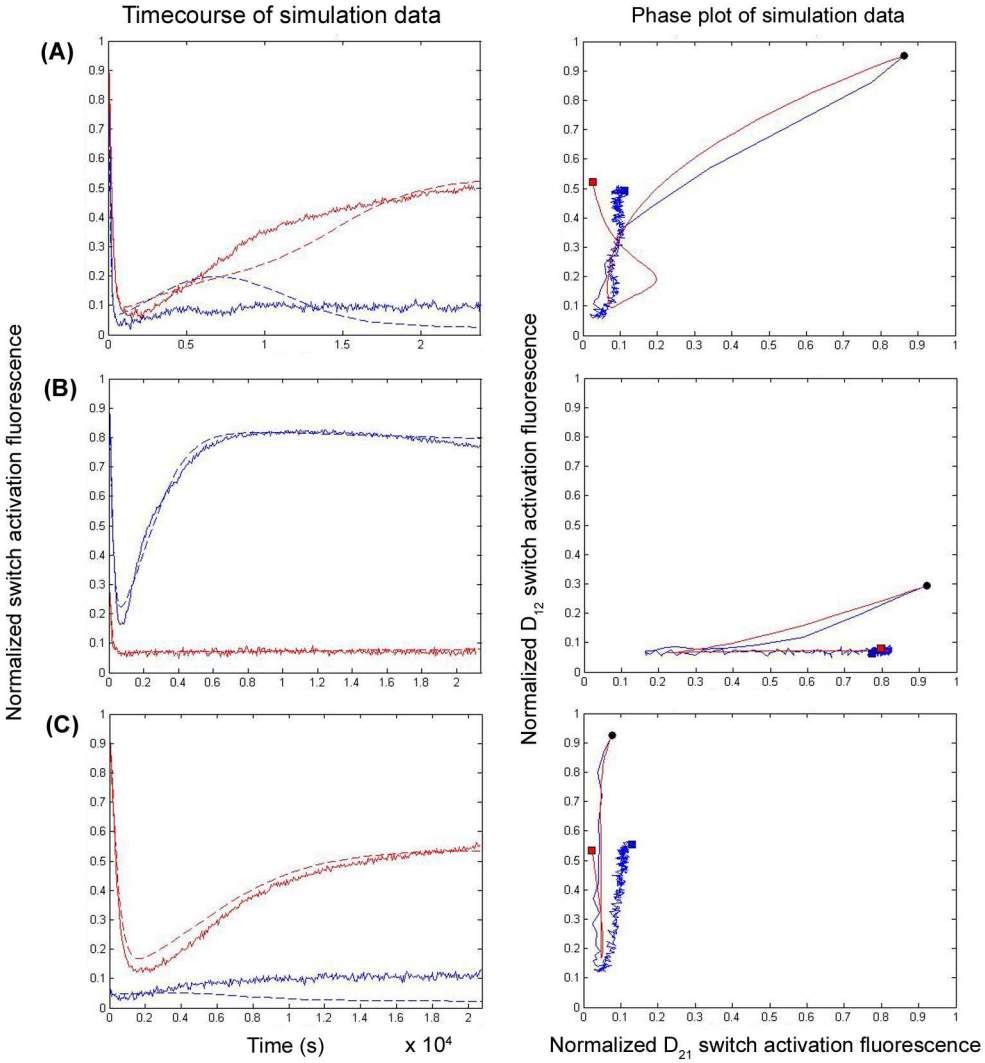

FIGURE 3. Characteristic plots of simulation versus experimental data for the original data set using the full model. On the timecourse plots, solid lines represent experimental data and dashed lines represent simulation data; blue, D_{21}; red, D_{12}. On the phase plots, blue represents experimental data and red represents simulation data. NMSE $= 3.843 \times 10^{-3}$. (A) Experimental run 1; (B) experimental run 5; (C) experimental run 10.

When run against the original data set, however, these parameters faced similar problems (Figure 4B), producing a NMSE of 6.298×10^{-2}, presumably caused by excess transcription. In this respect, we thus note that the part of parameter space probed is fairly sensitive. We then compromised between the two data sets by assuming an intermediate RNAP concentration for both. Another parameter fit was run, producing NMSEs of 7.203×10^{-3} and 3.922×10^{-3} for the first and second data set, respectively.

To establish possible explanatory power of the model, a final parameter fit was run on both data sets simultaneously, producing NMSEs of 5.748×10^{-3} and 3.536×10^{-3} for the first and second set, respectively. These are well under the established error threshold of 0.01 and thus indicate that both sets can be effectively described by a single model. As expected, the simulation operates on the correct time scales and accurately predicts the general character of the data on both sets very well, falling under expectations only when the initial states of the system

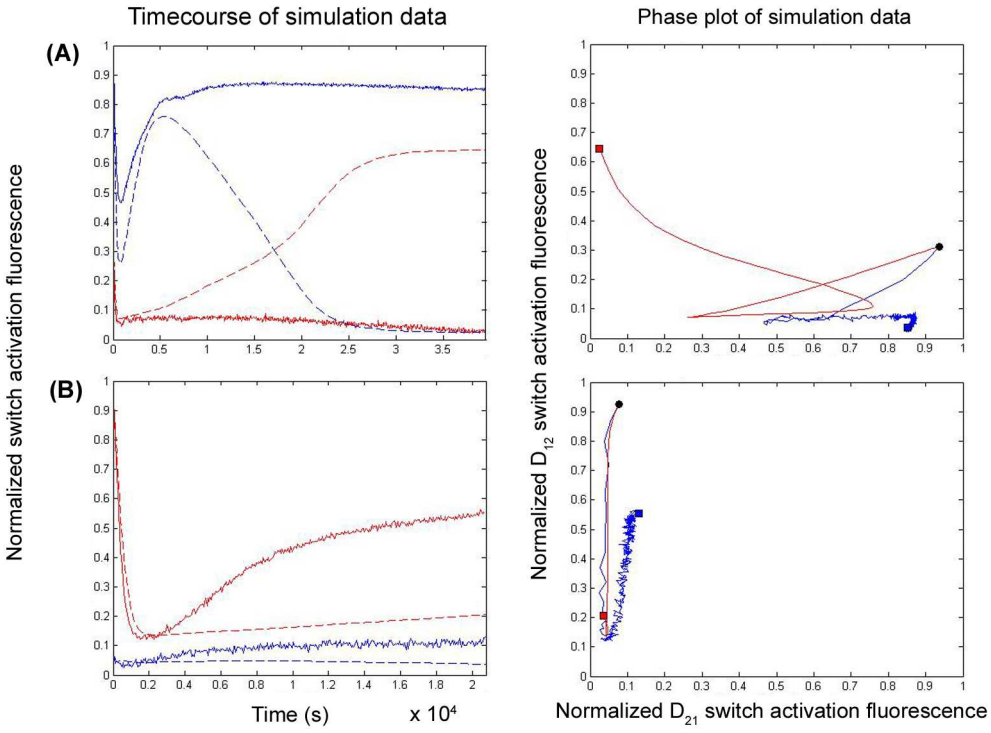

FIGURE 4. Characteristic plots of simulation versus experimental data using the full model with crossfitted parameters. On the timecourse plots, solid lines represent experimental data, and dashed lines represent simulation data; blue, D_{21}; red, D_{12}. On the phase plots, blue represents experimental data and red represents simulation data. (A) Characteristic failure of the model using the original fitted parameters simulated against the second experimental data set. The inability of the switches to tend toward the correct steady states was suspected to be caused by insufficient transcription. NMSE $= 1.515 \times 10^{-1}$. Data set 2, experimental run 9. (B) Characteristic failure of the model using least-squares parameters fit on the second experimental data set simulated on the first. Excess transcription suspected. NMSE $= 6.298 \times 10^{-2}$. Data set 1, experimental run 10.

demand a high level of interswitch competition on short time scales. We hypothesize that stochastic effects, which have not yet been explored, may play a significant role in this regime.

From the parameters reported, we see that abortive transcription and partial degradation play an active role in the system. The probabilities of abortive transcription, $\beta_1 = 0.232$ and $\beta_2 = 0.195$, are substantial, and the rate constants governing the partial products ($k_{4,j}$, $k_{-4,j}$) are strong. Also, the fitted RNAP initial burst parameters are consistent with previous results. The maximum pre-steady-state transcription efficiency amplitudes, $A_1 = 7.604$ and $A_2 = 8.096$, are both within an order of magnitude of a previously measured value of approximately 30.[17] Additionally, the fitted time constants, $\tau_1 = 87.375$ s and $\tau_2 = 67.156$ s, provide a brief characterization of the effective length of time over which transcription is significantly increased.

The performance of the reduced-order model derived in subsection 3.3 has not yet been evaluated, though we expect it to converge to approximately the same steady-state values as the full-order model.

5. Discussion

Our results confirm a significant influence of side reactions on the dynamics of the bistable switch system. In particular, the nonlinear kinetics of enzyme competition was found to be the primary determinant (of the three effects measured) of the experimental data. Simulation results demonstrate that such modeling induced the characteristic timecourse structure

observed. Indeed, without the modeling of this feature, numerical solutions were unable to match the slow activation curves obtained experimentally. This is consistent with previous analytic considerations which highlight the potential of enzyme competition in altering the qualitative behavior of the system.[6,7]

Moreover, our study corroborates qualitative results on the necessity of considering abortive transcription and partial degradation caused by enzyme nonprocessivity. Hodas *et al.* (2005)[7] showed that despite their prevalence during experimental trials, partial transcripts do not significantly influence the system dynamics. As shown in Table 1, this is precisely what our numerical results demonstrate. In no contingency was it observed that abortive transcription and partial degradation increased the explanatory power of the model by more than 10%. In fact, only under the conditions of enabled enzyme competition did the modeling of abortive transcription and partial degradation result in a positive percent contribution change (disabled RNAP initial burst, 9.901%; enabled, 3.444%); in all other conditions, such modeling actually resulted in a negative change.

Furthermore, while enzyme competition was sufficient to produce substantial explanatory power over long time scales, short time scales remained problematic. The introduction of the RNAP initial burst effect, however, seemed to provide a ready solution. By giving our model the freedom to temporally modulate transcriptional efficiency, we were able to match the steep initial switch deactivations observed with remarkable accuracy without suppressing the correct steady-state behaviors. This provides compelling support for the importance of the inclusion of the initial burst effect. Additionally, parameter fittings of the magnitude of the effect indicate consistency, to within a factor of five, with previous research.[17] These fittings also suggest that the effective time of operation is approximately one minute. Thus, the short time scales on which the effect acts suggest that long-term and, in particular, steady-state behaviors are not significantly affected (the large errors incurred in Table 1 are consequences of the large initial deviations), hence only full timecourse data need to be treated with such precision.

Both of the experimental data sets studied thus far have also only been concerned with the regime of low RNAP concentration, in the sense that we specifically induce, and indeed expect, enzyme competition effects. Therefore, it may be of particular interest to test whether our results still hold for data employing high enzyme concentrations. Given the model sensitivity already demonstrated across the two investigated data sets, it is reasonable to expect large inaccuracies upon moving to entirely new regions of parameter space. This indicates that we still do not yet fully understand the dynamics of the bistable switch system.

One possible explanation for the discrepancies observed is the fundamental stochasticity of discrete molecular components. Future directions will therefore certainly include the derivation of stochastic state models analogous to those discussed in section 3. Alternatively, we may approximate the assumption of deterministic continuity more closely by using higher concentrations of all relevant biomolecules. We are constrained here, however, by the laboratory implementation of such experiments. For example, temporal resolution may be limited as a result of the increased need for energy input.

Since transcription is the primary action of the system, future study may also benefit from a more detailed description of RNA synthesis.[8] The bistability of the system as a function of the parameter space also remains to be analyzed. Additionally, the performance of the reduced-order model with respect to the full model still needs to be evaluated. Given the complexity of current and, inevitably, of any future interactions that describe the system dynamics, it is of particular interest that we prune the model as much as possible while still retaining the essential interactions so as to facilitate qualitative insight. Computational demands may also be of concern, but this is not anticipated for a system of such minor complexity. A reduced-order model that is predictive in the sense that we have demonstrated may also prove to be a fast, reliable tool for computational design and analysis.[9,10,12,13]

By modeling such general biochemical phenomena as enzyme competition, abortive transcription and partial degradation, and RNAP initial burst, we seek to provide generally relevant characterizations applicable to a wide range of biological systems. Our results demonstrate strong progress toward the development of an accurate, predictive quantitative model implementing these effects. Given the computational power of genetic circuits,[5] we therefore hope that our work proves useful to future biological research, both in the understanding of naturally occuring networks and in the design and construction of novel systems (for directly related work on a follow-up system, we

direct the interested reader to Kim, White, and Winfree (2006)[22]). As synthetic biology becomes increasingly infused with strict engineering principles, such quantitative models will prove ever more valuable as tools for rational biological design.

6. Methods

6.1. **Numerical methods.** Kinetic simulations and parameter fittings were implemented in MATLAB, Release 14, Service Pack 2 (The MathWorks, Natick, Massachusetts). Differential equations were solved using built-in stiff solvers, while parameter fittings were performed using the Levenberg-Marquardt optimization algorithm.[18–21] Initial experimental concentrations were approximated from the data by assuming completeness of both activator-switch and activator-inhibitor reactions, with strict preference for the latter. Numerical bounds for parameters used in least-squares fittings are given in Table 4.

6.2. **Network implementation.** This work used experimental data taken by Jongmin Kim, whose laboratory implementation of the bistable switch system required the design of two different sets of switch elements. The first DNA transcriptional switch used is double-stranded downstream of the promoter with a 22-base-pair (bp) extension of the sense strand upstream.

Table 4. Numerical bounds used in least-squares parameter fittings.

Parameter	Lower bound	Upper bound
$k_{1,j}$	10^3	10^6
$k_{2,j}$	10^3	10^6
$k_{3,j}$	10^3	10^6
$k_{-3,j}$	10^{-2}	10^2
$k_{4,j}$	10^3	10^6
$k_{-4,j}$	10^{-1}	10^4
$k_{cp,j}$	10^{-3}	1
$\tilde{k}_{cp,j}$	10^{-3}	1
K_{Mp}	10^{-9}	10^{-6}
$\tilde{K}_{Mp}$	10^{-9}	10^{-5}
$k_{cd,j}$	10^{-3}	1
K_{Md}	10^{-9}	10^{-6}
β_j	0	1
A_j	5	50
τ_j	20	600

The antisense strand extends only partially (5 bp) into the promoter region. The activator has a free 8-bp toehold region that overhangs past the switch when bound. The second set of switch elements was designed around the same principles as the first, with the specific sequences chosen to minimize crosstalk in the system. Bacteriophage T7 RNAP and *Escherichia coli* ribonuclease H1 were used. Energy input was provided in the form of nucleoside triphosphate.

6.3. **Data acquisition and analysis.** This subsection details the techniques used by Jongmin Kim to obtain the experimental data used. Each DNA switch was tagged with an identifiable fluorescent molecule. Excitation and emission for D_{21} were at 559 nm and 580 nm, respectively; excitation and emission for D_{12} were at 597 nm and 615 nm, respectively. Fluorescence measurements were performed in a spectrofluorometer. Experiments began with both switches fully activated by excess activator ($\mu_1 = 45$ nm, $\mu_2 = 60$ nm for the first data set; $\mu_1 = 75$ nm, $\mu_2 = 100$ nm for the second; $\rho_1 = \rho_2 = 1$ μm for both; see section 4). Thus, the fluorescence was minimal. Between 0.88 μm and 1.03 μm of each RNA inhibitor was then added to calibrate initial switch states. We waited 30 min for hybridization to complete, at which point, RNAP and RNaseH were added ([RNAP] = 38 nm, [RNaseH] = 4.3 nm for the first data set; [RNAP] = 30 nm, [RNaseH] = 4.4 nm for the second). Fluorescence measurements were then taken at 60-second intervals. The data were normalized relative to their maximum and minimum values, and inverted to correspond signals with switch activation. The first and second data sets consisted of eleven and twelve experimental runs, respectively.

Acknowledgements. We thank Richard Murray and Erik Winfree for suggesting this project and for the valuable mentorship given, and Jongmin Kim for explaining and allowing the use of his experimental data. We further thank Richard Murray and Carol Readhead for critical readings of the manuscript. This work was supported by the California Institute of Technology Summer Undergraduate Research Fellowship program.

References

[1] J. Hasty, D. McMillen, and J. J. Collins, *Engineered gene circuits*, Nature, 420 (2002), pp. 224–230.

[2] R. McDaniel and R. Weiss, *Advances in synthetic biology: on the path from prototypes to applications*, Curr. Opn. Biotech., 16 (2005), pp. 476–483.

[3] T. S. Gardner, C. R. Cantor, and J. J. Collins, *Construction of a genetic toggle switch in Escherichia coli*, Nature, 403 (2000), pp. 339–342.

[4] M. B. Elowitz and S. Leibler, *A synthetic oscillatory network of transcriptional regulators*, Nature, 403 (2000), pp. 335–338.

[5] J. Kim, J. J. Hopfield, and E. Winfree, *Neural network computation by in vitro transcriptional circuits*, in *Advances in Neural Information Processing Systems 17*, L. K. Saul, Y. Weiss, and L. Bottou, eds., MIT Press, Cambridge, MA, 2004, pp. 681–688.

[6] M. Dunlop, T. Jones, R. Sayaman, and I. Shapiro, *Bistability in a Real In vitro Biochemical Circuit*, California Institute of Technology, Pasadena, CA, 2005.

[7] N. Hodas, Z. Jin, S. Waydo, K. Ho, J. Kim, and E. Winfree, *A Dynamic Analysis of Genetic Switches*, California Institute of Technology, Pasadena, CA, 2005.

[8] S. Arnold, M. Siemann, K. Scharnweber, M. Werner, S. Baumann, and M. Reuss, *Kinetic Modeling and Simulation of In Vitro Transcription by Phage T7 RNA Polymerase*, Biotechnol. Bioeng., 72 (2001), pp. 548–561.

[9] J. E. Bailey, *Mathematical Modeling and Analysis in Biochemical Engineering: Past Accomplishments and Future Opportunities*, Biotechnol. Prog., 14 (1998), pp. 8–20.

[10] J. M. Bower and H. Bolouri, eds., *Computational Modeling of Genetic and Biochemical Networks*, MIT Press, Cambridge, MA, 2001.

[11] L. Edelstein-Keshnet, *Mathematical Models in Biology*, SIAM, Philadelphia, PA, 2005.

[12] L. Glass and S. A. Kauffman, *The logical analysis of continuous, non-linear biochemical control networks*, J. Theor. Biol., 39 (1973), pp. 103–129.

[13] J. Hasty, D. McMillen, F. Isaacs, and J. J. Collins, *Computational studies of gene regulatory networks: in numero molecular biology*, Nat. Rev. Genet., 2 (2001), pp. 268–279.

[14] J. D. Murray, *Mathematical Biology I: An Introduction*, 3rd ed., Springer-Verlag, Berlin, 2002.

[15] R. Sousa, D. Patra, and E. M. Lafer, *Model for the mechanism of bacteriophage T7 RNAP transcription initiation and termination*, J. Mol. Biol., 224 (1992), pp. 319–334.

[16] D. R. Yazbeck, K. L. Min, and M. J. Damha, *Molecular requirements for degradation of a modified sense RNA strand by Escherichia coli ribonuclease H1*, Nucleic Acids Res., 30 (2002), pp. 3015–3025.

[17] Y. Jia, S. S. Patel, *Kinetic Mechanism of Transcription Initiation by Bacteriophage T7 RNA Polymerase*, Biochemistry, 36 (1997), pp. 4223–4232.

[18] K. Levenberg, *A method for the solution of certain problems in least squares*, Quart. Appl. Math., 2 (1944), pp. 164–168.

[19] D. Marquardt, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, SIAM J. Appl. Math., 11 (1963), pp. 431–441.

[20] J. J. More, *The Levenberg-Marquardt algorithm: implementation and theory*, in *Numerical Analysis: Lectures Notes in Mathematics 630*, G. A. Wilson, ed., Springer-Verlag, Berlin, 1977, pp. 105–116.

[21] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., John Wiley, New York, 1987.

[22] J. Kim, K. S. White, and E. Winfree, *Construction of an in vitro bistable circuit from synthetic transcriptional switches*, Mol. Syst. Biol., 2 (2006).