Characterization of a Biomolecular Fuel Delivery Device Under Load

Joshua D. Bishop
Department of Electrical Engineering
University of Washington
Seattle, WA 98195, USA
jdbishop@uw.edu

Eric Klavins
Department of Electrical Engineering
University of Washington
Seattle, WA 98195, USA
klavins@uw.edu

Abstract—We describe the design, implementation, and characterization of a biomolecular RNA fuel delivery device under load. The device is based on genelet technology, and is capable of driving dynamic behavior in complex nucleic acid circuits. Here we closely examine the behavior of a system consisting of the device driving the operation of a simple, fluorescently-labeled DNA probe. We use a simple model to characterize the stability region of the system and the effect of changing load conditions. Finally, we validate these analytical observations with basic experiments.

I. INTRODUCTION

The field of DNA nanotechnology has leveraged the simple mechanism of Watson-Crick base pairing to implement an impressively wide array of behaviors in vitro, enabling incredible control of matter at the nanoscale. For example, DNA has been used to implement catalytic and feedback circuits [1], [2], computation systems [3], [4], linear systems [5], digital circuits [6], [7], self-assembling systems [8], [9], triggered amplification circuits [10], and a host of autonomous nanomachines [11], among many examples (also see [12] for a thorough review of this literature). Theoretical work has suggested that DNA may be the ultimate programmable molecule, a “universal substrate” for the implementation of arbitrary chemical kinetics [13].

In vitro DNA systems typically require the concurrent design of “fuel” strands that drive their operation to completion (i.e. thermodynamic equilibrium). Although the ability to design specific, targeted fuels from a common substrate can be an advantage, changing system dynamics due to the irreversible consumption of these fuels is a major disadvantage to most in vitro DNA systems [5].

In vitro “genelet” technology, first introduced by Kim, et al. [14], provides the ability to dynamically regulate the production of sequence-programmable RNA outputs using common metabolic precursors and two protein enzymes. The RNA outputs may dynamically regulate the behavior of other genelets, resulting in circuits with regulatory network topologies, e.g. a bistable switch [15], oscillators [16], or a transcriptional rate regulator [17]; or they may be used to dynamically fuel the operation of other DNA circuits. For example, various components of a genelet oscillator were recently tested for their ability to drive the operation of DNA tweezers [18].

The aim of this paper is to quantitatively describe the ability of a basic genelet device to dynamically fuel the operation of a simple DNA-based device. The organization of the paper is as follows. First, we describe the design and implementation of a system consisting of the fuel delivery device under load. Next, we develop a descriptive model of the system from the chemical kinetics of the underlying biochemical reactions. We present the results of basic experimental tests and make quantitative observations guided by the model. Finally, we discuss future opportunities to extend the work.

II. DESIGN AND IMPLEMENTATION

A high-level system design specification is illustrated in Fig. 1. This high-level specification describes a chemical reaction network, where arrows separate left-hand reactants from right-hand products. Our device should catalytically generate fuel (R1) that binds to and powers a downstream load (R2), which consumes the fuel and reverts to its initial state (R3). This section describes the process of turning this reaction network specification into a nucleic acid implementation based on genelet technology.

Our implementation is described in Fig. 2, which shows the component nucleic acid strands subdivided into labeled domains, a common design abstraction employed in DNA circuit design. A domain is a contiguous sequence of nucleic acid bases that act as a single functional subunit. Thus, a domain specifies that no two bases in the same domain will independently interact with other molecular species. We label complementary domains \( x \) and \( x^* \).
**A. Nucleic Acid and Protein Enzyme Components**

1) **Device Produces Fuel (R1):** We use a simple genelet device to produce an RNA fuel strand. Genelets are input/output-programmable DNA templates for *in vitro* transcription catalyzed by the protein enzyme T7 RNA polymerase (\(R_P\)). Genelet design is focused on three overlapping regions: an input region (Fig. 2, domains 1, 2), an \(R_P\) promoter site (Fig. 2, domains 2, 3, 4), and an output region (Fig. 2, domains 5, 6, 1, 2, 7, 8, 7*, 8).

   The device combines an inactive genelet, \(g\), with a short DNA “activator” strand, \(s\), (Fig. 2, domains 2*, 1*), that is exactly complementary to the input region and therefore binds irreversibly to an inactive genelet, \(g\). An activated genelet generates fuel strands, \(r\) (Fig. 2, domains 4, 5, 6, 1, 2, 7, 8, 7*, 8), from its output region, catalyzed by \(R_P\) binding to the double-stranded promoter site (Fig. 2, domains 2, 3, 4); an inactivated genelet generates very little or no RNA output because the promoter site is partially single-stranded (Fig. 2, domains 2, 3, 4).

   The irreversible binding of \(s\) to \(g\) meets our specification and allows us to interrogate changing production rates by choosing the initial amount of \(s\). A removable activator would allow us to embed the device in a larger regulatory framework, but is unnecessary in this simple setting.

2) **Fuel Powers Load (R2):** We use a single-stranded DNA probe, \(p\) (Fig. 2, domain 6*), as the load. The probe is a seventeen-base oligonucleotide that is fully complementary to a region of \(r\) (Fig. 2, domain 6) and has a fluorophore and quencher FRET pair (18) incorporated at opposing ends. Alone, \(p\) coils on itself, which co-localizes the attached fluorophore/quencher pair and yields a low fluorescence output: the OFF state. When \(p\) binds to \(r\) they form a rigid, double-stranded, RNA/DNA hybrid complex, \(p \cdot r\) (Fig. 2, domain 4, 5, 6, 1, 2, 7, 8, 7*, 8), that separates the attached fluorophore/quencher pair and yields a high fluorescence output: the ON state.

3) **Load Consumes Fuel (R3):** The load consumes the fuel via a degradation reaction catalyzed by the protein enzyme Ribonuclease H (\(R_H\)), an enzyme that selectively degrades RNA bound to DNA. At any given time, the ensemble load exists in some ratio of ON to OFF states and thus dynamically reports the percentage of load fueled by the device and, more indirectly, on the quantity of fuel.

**B. Sequence Design**

The physical implementation of the system (device, fuel, and load) requires choosing the sequences of the nucleic acid components so they determine all – and only – the interactions that are specified by our design specifications shown in Fig. 1 and Fig. 2 and described above. We input the domains, strands, desired secondary structures, and sequence constraints illustrated in Fig. 2 to NUPACK, a nucleic acid sequence design and thermodynamic analysis software package [19], [20], [21], [22], [23].

The software produces a set of sequence designs from an algorithmic attempt to minimize the average number of nucleotides paired at equilibrium relative to the input specifications [20]. The best result, illustrated in Fig. 2 and confirmed by analysis, in NUPACK, of predicted minimum free energy structures at thermodynamic equilibrium, was ordered as a set of commercially-synthesized, PAGE-purified, synthetic oligonucleotides. These strands give us a reasonable degree of confidence that the system will behave as expected in experiments.
### III. MODEL

We derive a relatively simple model of the system from the high-level chemical kinetics of the design specification. Since the specified reactions are the production of fuel (R1), binding of fuel to probe (R2), and degradation of fuel from probe (R3), we model the dynamics with the equations

\[
\begin{align*}
[r] &= f_p ([g], [s]) - \lambda [p] [r], \\
[p-r] &= \lambda [p] [r] - f_d ([p-r]),
\end{align*}
\]

(1)

where \( f_p \) and \( f_d \) are functions that describe the rates of R1 and R3, respectively, and \( \lambda \) is a reaction rate constant associated with R2. The notation \( [x] \) indicates the concentration of molecular specie \( x \), which is considered a function of time, and the overdot notation indicates the time derivative.

While \( r \) is RNA and is produced and degraded via enzymatic reactions, all other species in the system are DNA and are therefore conserved, so that we can define the mass-conservation constants

\[
\begin{align*}
p_{\text{tot}} &= [p] + [p-r], \\
g_{\text{tot}} &= [g] + [g\cdot s], \\
s_{\text{tot}} &= [s] + [g\cdot s].
\end{align*}
\]

(2)

Observe that at steady-state the rates of production and degradation of \( r \) must balance since from (1), \( f_p ([g]^*, [s]^*) = f_d ([p-r]^*) \) in equilibrium, where \( [x]^* \) indicates the equilibrium concentration of \( x \). Our intuition confirms that too much production will overwhelm degradation, saturating the output at \( [pr] = p_{\text{tot}} \) and leading to an unstable \([r]\), but we also examine this intuition mathematically, below.

This property is independent of all forms of \( f_p ([g], [s]) \) and \( f_d ([p-r]) \), and so the system appears, in general, to be sensitive to the concentration of enzymes. We next examine this and other properties of the model given specific forms of these functions.

#### A. Michaelis-Menten Approximations

There are many candidate forms for the functions \( f_p \) and \( f_d \) that are true to the underlying chemical kinetics. However, since these functions represent the enzyme kinetics of the system, we choose to employ the well-known Michaelis-Menten approximation [24] for each, such that

\[
\begin{align*}
f_p ([g], [s]) &= \frac{k_1 [g\cdot s]}{K_1 \left(1 + \frac{[g]}{K_1} + \frac{[s]}{K_2}\right)} + \frac{k_2 [g]}{K_2 \left(1 + \frac{[g]}{K_1} + \frac{[s]}{K_2}\right)}, \\
f_d ([pr]) &= \frac{k_3 [p-r]}{K_3 + [p-r]}.
\end{align*}
\]

(3)

The function \( f_p \) now approximates the rate of production of \( r \), which is due to the interaction of \( R_p \) with one of two competing substrate complexes, \( g\cdot s \) and \( g \). (We include \( g \) as a substrate to account for leaky expression from the incomplete promoter site.) These reactions have catalysis rates \( k_1 \) and \( k_2 \), and Michaelis-Menten constants \( K_1 \) and \( K_2 \), respectively. Similarly, the function \( f_d \) now approximates the rate of degradation \( r \), which is due to the interaction of \( R_H \) with its substrate complex, \( p\cdot r \). This reaction has catalysis rate \( k_3 \) and Michaelis-Menten constant \( K_3 \).

In general, the Michaelis-Menten approximation defines the catalysis rate as the maximum rate of product formation, \( v_{\text{max}} \), times the concentration of enzyme; and the Michaelis-Menten constant as the concentration of substrate that yields half the maximum rate, \( v_{\text{max}}/2 \).

We can now rewrite (1) in full, using (3) and the mass-conservation relationships in (2). We then have the full system model

\[
\begin{align*}
[r] &= \frac{c_1 + c_2 s_{\text{tot}}}{c_3 + s_{\text{tot}}} - \lambda (p_{\text{tot}} - [p-r]) [r], \\
[p-r] &= \lambda (p_{\text{tot}} - [p-r]) [r] - \frac{k_3 [p-r]}{K_3 + [p-r]},
\end{align*}
\]

(4)

where

\[
\begin{align*}
c_1 &= \frac{K_1 k_2 g_{\text{tot}}}{K_2 - K_1}, \\
c_2 &= \frac{k_1 K_2 - k_2 K_1}{K_2 - K_1}, \\
c_3 &= \frac{K_1 K_2 + K_1 g_{\text{tot}}}{K_2 - K_1}.
\end{align*}
\]

#### B. Equilibrium Analysis

We now examine the equilibrium behavior of our full model. At equilibrium,

\[
0 = \frac{c_1 + c_2 s_{\text{tot}}}{c_3 + s_{\text{tot}}} - \lambda (p_{\text{tot}} - [p-r]^*) [r]^* \quad \text{and} \\
0 = \lambda (p_{\text{tot}} - [p-r]^*) [r]^* - \frac{k_3 [p-r]^*}{K_3 + [p-r]^*}.
\]

(5)

We can algebraically solve (5) for the equilibrium values \([r]^*\) and \([p-r]^*\), which are

\[
\begin{align*}
[r]^* &= \frac{c_4 (c_4 - c_5 k_3)}{\lambda c_5 (c_4 K_3 + (c_4 - c_5 k_3) p_{\text{tot}})} \quad \text{and} \\
[p-r]^* &= \frac{c_4 K_3}{c_5 k_3 - c_4}.
\end{align*}
\]

(6)
where \(c_4 = c_1 + c_2 \, s^{tot}\) and \(c_5 = c_3 + s^{tot}\).

The existence of a stable steady-state requires positive values of \([r]^*\) and \([p \cdot r]^*\), which depends on the system parameters and initial conditions. However, for a given set of parameters, we can predict the stable steady-states of the system using the equilibrium model (6). To illustrate this property, we take as an example a system with specific parameters and initial conditions.

**1) Example:** The values in Table I, Set I represent realistic, but hand-picked, parameters based on previous genetet studies [14] and initial conditions based on current experimental design. This set is of interest due to the saturation of the system output, \([p \cdot r]\), at \(s^{tot}\) when increasing the initial condition \(s^{tot}\), as illustrated in Fig. 3 and discussed below.

First, we observe in the parameterization of (6) by \(s^{tot}\) plotted in Fig. 3(a) that the output saturation of \([p \cdot r]^*\) at \(p^{tot}\) coincides with a point of discontinuity in \([r]^*\). We can calculate from (6) that when \(([r]^*, [p \cdot r]^*) = (\infty, p^{tot})\), then

\[
s^{tot} = \frac{c_1 (K_3 + p^{tot}) - c_3 \, k_3 \, p^{tot}}{k_3 \, p^{tot} - c_2 \, (K_3 + p^{tot})},
\]

which for this example is \(s^{tot} = 69.4444\) nM. This bifurcation between stable and unstable behavior occurs in the physical system only when the expression (7) is positive, as in this example.

Second, we plot several simulated system trajectories with varied initial condition \(s^{tot}\) in Fig. 3(b) to show the effect of this bifurcation more clearly. Fig 3(c) charts endpoint measurements of \([r]\) and \([p \cdot r]\) for these trajectories for two time points: 1) realistic experimental \(t_f = 24\) h and 2) equilibrium \(t_f \to \infty\); note that the former set of points approach the latter. The seven trajectories with initial condition \(s^{tot}\) in the interval \([0, 69.4444]\) have stable steady-states predicted by (6) while those four with initial condition \(s^{tot}\) in the interval \([69.4444, 100]\) have \([r]\) increasing without bound, while \([p \cdot r]\) saturates at \(p^{tot}\).

Third, as a final characterization of the equilibrium behavior of this example, we examine the nullclines of the full model (4)

\[
[p \cdot r] = p^{tot} - \frac{c_4}{\lambda \, [r] \, c_5},
\]

\[
[p \cdot r] = \frac{-k_3 - \lambda \, (K_3 - p^{tot}) \, [r]}{2 \lambda \, [r]} \\
\pm \frac{\sqrt{4 \, K_3 \, p^{tot} \, \lambda^2 \, [r]^2 + (k_3 + \lambda \, (K_3 - p^{tot}) \, [r])^2}}{2 \lambda \, [r]},
\]

which correspond to \(\dot{[r]} = 0\) and \(\dot{[p \cdot r]} = 0\), respectively. The steady-states of the stable trajectories never meet, and the inset of Fig. 3(d) shows the lack of intersection even at large \([r]\) for the four unstable trajectories, as expected.
C. Disturbance Rejection

We can clearly see from (6) that the steady-state output of the system, \([p\cdot r]^*\), is independent of the term \(p^{tot}\). Thus, the output of the system \([p\cdot r]\), when a stable steady-state exists, will demonstrate a behavior known to biologists as “perfect adaptation,” alternately known as disturbance rejection, with respect to disturbances in \(p^{tot}\). Experiments, the results of which we describe in the next section, confirm that this behavior does occur in the physical system.

IV. EXPERIMENTS

We experimentally tested the behavior of the system to validate the descriptive capabilities of our model. The test consisted of observing the initial response of eight genelet systems, each with a different initial condition \(s^{tot}\), and the second response of each system to an induced change in \(p^{tot}\) at \(t = 8\) h. As in our simulated example, we can explain observed changes in equilibrium behavior, with respect to the initial condition \(s^{tot}\), using our model. Additionally, we expect to see the output signal \([p\cdot r]\) reject the disturbance in \(p^{tot}\) that occurs at \(t = 8\) h.

A. Methods and Materials

The total volume of each test solution was 100 µL. The solutions were first prepared with buffer components at final concentrations of 1x Transcription Buffer (NEB), 10 mM rNTPs (NEB), and 6 mM MgCl\(_2\) (Sigma). DNA strands (IDT) were added at final concentrations of \(g^{tot} = 100\) nM, \(s^{tot} = (0, 5, 10, 15, 20, 30, 50, 100)\) nM, and \(p^{tot} = 100\) nM.

Enzymes were added to each solution at \(t = 0\) h, at final concentrations \([R_P] = 1\) U/µL and \([R_H] = 0.01\) U/µL.

Solutions were prepared in eight wells (one column) of a 96-well plate (Corning) and covered with 50 µL of mineral oil to prevent evaporation. The fluorescence response of each solution was observed in a fluorescence plate reader (Biotek Synergy HT) with 575/15 nm excitation and 620/15 nm emission wavelength/bandpass filters installed, and the internal temperature controller set to 37°C. At \(t = 8\) h, \(p^{tot}\) was manually increased from 100 nM to 200 nM.

B. Results

We converted the raw fluorescence test data to \([p\cdot r]\) using the empirically-derived function

\[
[p\cdot r] = \frac{397.033 + 13.8776 p^{tot} - \text{fluorescence}}{13.8776 - 373.775},
\]

which is based on separate calibration experiments (data not shown).

The converted test data is plotted in the left panel of Fig. 4(b). We fit our model to this data using the nonlinear optimization function FindFit in Mathematica and the parameters in Table I, Set I as an initial guess. The resulting best-fit parameters are found in Table I, Set II, with which we performed the system characterization illustrated in Fig. 4(a-c).

First, the equilibrium model with these best-fit parameters has no stability bifurcation in the right-half plane, as shown in Fig. 4(a). We calculate from (6), using the best-fit,
parameters, that \(( [r]^*, [p \cdot r]^* ) \) will asymptotically approach \((22.1688, 18.6813) \) nM when \( p^{\text{tot}} = 100 \) nM and \( (9.94236, 18.6813) \) nM when \( p^{\text{tot}} = 200 \) nM, since

\[
\lim_{s^{\text{tot}} \to +\infty} \left( [r]^*, [p \cdot r]^* \right) = \frac{c_2 (k_3 - c_2)}{\lambda (k_3 p^{\text{tot}} - c_2 (K_3 + p^{\text{tot}}))} \cdot \frac{c_2 K_3}{k_3 - c_2}.
\]

(10)

Thus, the model indicates that the system never saturates given the parameters and initial conditions in Table I, Set II and any positive \( s^{\text{tot}} \). Accordingly, none of the test data trajectories show saturation in the output, and all exhibit stable steady-states, as expected.

Second, test data and simulated data agree at the endpoints at \( t = 8 \) h and \( t = 16 \) h, which are plotted in Fig. 4(c). The plot shows these points overlaid on the parameterization of (6) by \( s^{\text{tot}} \). However, the fit is poor with respect to the transient dynamics. A possible cause is that few of the trajectories in the first eight-hour period, and none in the second, appear to be at a true steady-state. The model, fitted to this data, thus overestimates the time required for output of the test systems to reject the disturbance in \( p^{\text{tot}} \).

Third, regardless of the speed of the transient dynamics, the test data do show the physical system rejects the disturbance in \( p^{\text{tot}} \) at \( t = 8 \) h, returning to approximately the same steady-state level by \( t = 16 \) h. Again, Fig 4(c) shows that \([p \cdot r]^*\) endpoints are close for test and simulated output data at both \( t = 8 \) h and \( t = 16 \) h.

V. CONCLUSIONS AND FUTURE WORK

The genelet device clearly demonstrates the ability to dynamically drive the operation of a simple DNA-based device. We have developed a basic model of this system that allows us to quantitatively characterize behaviors given kinetic parameters and initial conditions. We have shown, through a simulated example and experimental tests, how we can use the model to describe regions of stability, asymptotic equilibrium behavior, and disturbance rejection in the output. The experiments validate important features of the model, since we were able to characterize the physical system by tuning \( s^{\text{tot}} \) and to observe rejection of a disturbance in \( p^{\text{tot}} \) by the output.

If disturbance rejection in the output is not the preferred behavior, however, we may be able to regulate the device in the same way that a voltage regulator regulates the available power in an electrical circuit. Ideally, this regulation would result in the free supply of fuel, rather than the output state, being independent of the amount of load. The device described in this paper can easily be modified to use an activator that is removed by \( r \), resulting in a simple negative feedback loop that might provide the type of regulation required.

We therefore plan to investigate the effect of negative feedback on the regulation of fuel in this simple setting. We believe that feedback regulation will result in a robust, stable fuel source, which may be useful in powering larger, more complex DNA circuits.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful conversations with E. Winfree, G. Seelig, J. Kim, and E. Franco.

REFERENCES