Building the components for a biomolecular computer

Clint Morgan
Darko Stefanovic
Cristopher Moore

Department of Computer Science
University of New Mexico

Milan N. Stojanovic

Department of Medicine
Columbia University

November 2004

Abstract

We propose a new method for amorphous bio-compatible computing using deoxyribozyme logic gates [1] in which oligonucleotides act as enzymes on other oligonucleotides, yielding oligonucleotide products. Moreover, these reactions can be controlled by inputs that are also oligonucleotides. We interpret these reactions as logic gates, and the concentrations of chemical species as signals. Since these reactions are *homogeneous*, i.e., they use oligonucleotides as both inputs and outputs, we can compose them to construct complex logic circuits. Thus, our system for chemical computation offers functionality similar to conventional electronic circuits with the potential for deployment inside of living cells. Previously, this technology was demonstrated in closed-system batch reactions, which limited its computational ability to simple feed-forward circuits. In this work, we go beyond closed systems, and show how to use thermodynamically open reactors to build biomolecular circuits with feedback. The behavior of an open chemical system is determined both by its chemical reaction network and by the influx and efflux of chemical species. This motivates a change in design process from that used with closed systems. Rather than focusing solely on the stoichiometry of the chemical reactions, we must carefully examine their kinetics. Systems of differential equations and the theory of dynamical systems

1

become the appropriate tools for designing and analyzing such systems. Using these tools, we present an *inverter*. Next, by introducing feedback into the reaction network, we construct devices with a sense of state. We show how a combination of analytical approximation techniques and numerical methods allows us to tune the dynamics of these systems. We demonstrate a flip-flop which exhibits behavior similar to the RS flip-flop of electronic computation. It has two states in which the concentration of one oligonucleotide is high and the other is low or vice versa. We describe how to control the state of the flip-flop by varying the concentration of the substrates. Moreover, there are large regions of parameter space in which this behavior is robust, and we show how to tune the influx rates as a function of the chemical reaction rates in a way that ensures bistability.

1   Introduction

We use deoxyribozymes (nucleic acid enzymes) as gates to transform input and substrate signals (molecular concentrations) into product signals and thereby perform simple computation. Since the inputs are of the same type as the outputs, viz. oligonucleotides, gates may, in principle, be connected in complex circuits, with the output of one gate acting as the input of another. Thus, we may design chemical systems that perform complex computations from simple boolean primitives in much the same way electronic computers are built from simple logic gates. These devices could operate without macroscopic intervention in a biological environment, and the goal of this technology is autonomous *in vivo* computation for diagnostic and therapeutic purposes. We have reported gates with a single layer of logic, and no inter-gate communication [1]. Devices that function as a half-adder [2] and a tic-tac-toe automaton [3] have been built and tested in the laboratory.

These gates have been deployed in a closed reactor, which effectively limits this technology to one-shot boolean computations. To overcome this limitation, we explore using this chemistry in an open reactor, in which gates could be re-used many times and connected in recurrent, rather than feed-forward, circuits. This adds a level of complexity to the engineering task, but we develop a process that may be used to engineer these devices. We apply methods of dynamical systems to construct reaction networks in open reactors that implement rudimentary elements of digital chemical computation. This allows us to investigate complex reaction networks that make use of inter-gate communication and feedback.

2 The Chemical Kinetics of Deoxyribozyme Logic Gates

The four components of our deoxyribozyme system are inputs, gates, substrates, and products. Under certain input conditions a gate is an active enzyme [1]. The effect of input molecules on the catalytic activity of the gate defines the logic operation that the gate performs. A gate requires the presence and/or absence of certain inputs to be active. When active, the enzymatic gate is a phosphodiesterase: it catalyzes an oligonucleotide cleavage reaction. A substrate molecule is cleaved into two product molecules. The product molecules represent the output signal of the gate. Computations are carried out in solution, where gates communicate by diffusion of oligonucleotides. Logic signals, true or false, are expressed by high or low concentrations of specific oligonucleotides. Oligonucleotides transmit information by participating in the reactions of multiple gates. The simplest example is an oligonucleotide that is a product of one gate and an input to another; serving as a substrate would suffice as well.

The mechanism of a deoxyribozyme gate is as follows. Input molecules bind to the designated locations on the gate molecules. The binding of an input to a gate affects the conformation of the gate, which in turn affects catalytic activity. Under appropriate circumstances, the gate is an active enzyme, in which case it binds to a substrate molecule, cleaves it into two molecules of product, and separates into two molecules of product and one active gate complex. Active gates continue to operate as long as there is substrate remaining to be cleaved.

In order to design larger circuits, we must first understand the dynamic behavior of individual logic gates. We set up the experiment as follows. We prepare a solution with a concentration of $G = 250\,\text{nM}$ of a specific YES gate (which becomes active in the presence of input), a certain concentration I of the matching input, and a concentration $S = 2500\,\text{nM}$ of the substrate cleaved by the gate. At 900s intervals we record the instrumentally measured fluorescence. We repeat the experiment varying I, starting with $I = S$ and repeatedly halving it. The measured fluorescence level of a molecular species is proportional to its concentration. The specific fluorescences of the product and the substrate have been established separately and are in a ratio of 8:1. Therefore the increase of total fluorescence is proportional to the amount of product, which allows us to convert measured fluorescence into product concentration, shown in Figure 1.

For small values of the input concentration I, product concentration P rises linearly with time, with slope proportional to I. For larger values of I, the growth soon reaches a plateau defined by the initial substrate concentration S: when all of the substrate has been converted to product, the reaction stops. Note also the saturating behavior: whereas the slope of $t \mapsto P$ increases with I for small I, it remains roughly constant for $I > G$.
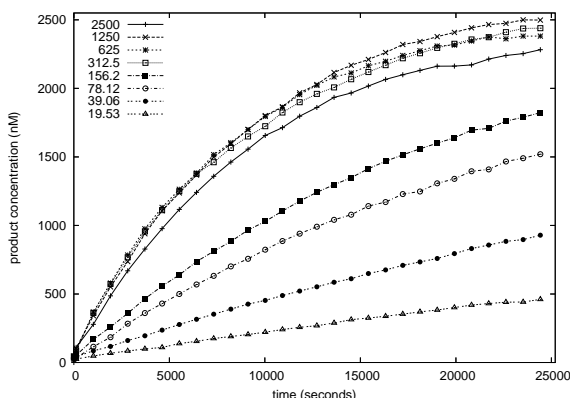
Figure 1: Measured kinetics of a deoxyribozyme gate for different input concentrations I (nM).

We will model the kinetics of the deoxyribozyme gates as follows. First we note that the cleavage and separation of substrate molecules is the slowest of the reactions—it is the rate-limiting process. We will assume that bonding between gate and input molecules is instantaneous and complete. Thus, the number of active gates at a given time is a simple calculation depending solely on the number of gates and inputs.

Cleavage of substrate requires both substrate and an active gate complex. Experiments have shown that the rate of production is proportional to the concentration of both reactants. Hence, a model for the rate at which product is produced is: $\frac{dP}{dt} = \beta S G_A$, where P is the product concentration, β is the reaction rate constant, S is the substrate concentration, and G_A is the concentration of *active* gates. In the experiments shown in Figure 1, I (and thus G_A) was held constant. Therefore, the solution is an exponential decay of the substrate S. This model agrees well with observed measurements in Figure 1, and analysis of the measured data gives a rough estimate of the reaction rate constant $\beta = 5 \cdot 10^{-7}\, \mathrm{nM}^{-1}\mathrm{s}^{-1}$. This value will be used to model the chemical gates in the circuit designs presented herein.

3   The Reactor

Chemical reactors may be divided into closed systems, where reactants are added to a solution and the reaction is allowed to proceed toward equilibrium, and open systems, where reactants are continuously supplied and excess solution is removed. We explore the benefits and design considerations associated with using our chemistry in an open reactor. Previous theoretical work has described circuits created

from hypothetical enzymatic transistors in open reactors [4]. We expand on this work by presenting circuit designs based on the chemical technology described above. Reactions in a closed environment are subject to the Second Law of Thermodynamics, which posits that the free energy in a closed system will continuously decrease and implies that the system will move toward an equilibrium. This does not rule out interesting behavior, such as oscillations on the way to an equilibrium [5], but it implies a finite number of cycles through these oscillations. Thus it would be impossible to implement a recurrent digital circuit in which gates could switch on and off an arbitrary number of times.

Instead, we use a thermodynamically open system; material is continuously supplied and removed, as in a living cell. The circuit may be reused and produce many outputs over its lifetime, so that it is recurrent rather than feed-forward. While a long-term goal of this technology is deployment inside of (thermodynamically open) living cells, the first step toward that end is testing and verification in a laboratory setting. A model open environment is the continuous-flow stirred tank reactor. It delivers reactants into a reaction chamber, stirred to maintain a uniform distribution of chemical species. An outflow removes solution from the reactor to maintain constant volume. The inputs can be varied in terms of their concentrations in the input solution and the flow rate into the reactor. Both the concentration and the *volumetric* influx of a solution can be varied while still maintaining the same total *molecular* influx rate. Thus we can manipulate total efflux while maintaining desired concentrations of chemical species inside the reactor.

The decay rate of the reactor (k) is equal to the efflux rate (E) divided by the volume (V). As the decay rate is increased, material spends less time inside the reactor. Because the reactor state changes faster, the circuit speeds up. However, this increases the amount of chemical species needed to maintain the same concentrations. In the specific design below, the reactor will have a total efflux of $5 \cdot 10^{-8}\,\mathrm{m^3 s^{-1}}$ and a volume of $5 \cdot 10^{-4}\,\mathrm{m^3}$. This results in a decay rate of $10^{-4}\,\mathrm{s^{-1}}$. While the resulting circuits will operate *very* slowly, these values were chosen as design points corresponding to equipment that can be found in a traditional chemistry laboratory.

4   A Simple Inverter

We begin by examining a simple computational device: the digital inverter. It is built from a single type of NOT gate G operating in an open reactor. The reactor is supplied with a constant influx of gate and substrate molecules. In addition, input I is supplied to provide an external drive. The output of this inverter is expressed by the concentration of product P cleaved from substrate S. The behavior of the system

can be modeled with the following system of four coupled differential equations:

$$\frac{dG}{dT} = \frac{G^m - EG(T)}{V} \tag{1}$$

$$\frac{dI}{dT} = \frac{I^m - EI(T)}{V} \tag{2}$$

$$\frac{dP}{dT} = \beta S(T)\max(0, G(T) - I(T)) - \frac{EP(T)}{V} \tag{3}$$

$$\frac{dS}{dt} = \frac{S^m}{V} - \beta S(T)\max(0, G(T) - I(T)) - \frac{ES(T)}{V} \tag{4}$$

where I^m, G^m, and S^m are the constant rates of *molar* influx of the respective chemical species, V is the volume of the reactor, E is the rate of volume efflux, and β is the reaction rate constant. The max terms in (3) and (4) come from our assumption that the binding of input to gate molecule is both instantaneous and complete.

Clearly, this system can function as an inverter. If there are no inputs in the reactor, all of the gates are active and produce product—this is the high signal. As input is added gates become inhibited, and the product concentration falls. As the input concentration reaches the gate concentration, all gates become inhibited and the product concentration falls to zero—this is the low signal.

To explore the equilibrium behavior of the inverter we first assume that the input concentration never exceeds the gate concentration; we can then eliminate the max functions from equations (1)-(4). We can now set the derivatives to zero and solve for P. This produces the following relation between input concentration and output (product) concentration:

$$P = \frac{\beta S^m V(\frac{G^m}{E} - I)}{E^2 + \beta V E(\frac{G^m}{E} - I)} = \frac{\frac{S^m}{E}(\frac{G^m}{E} - I)}{\frac{E}{\beta V} + \frac{G^m}{E} - I}$$

Introducing rescaling parameters $\alpha = \frac{S^m}{E}$, $\gamma = \frac{G^m}{E}$, and $\delta = \frac{E}{\beta V}$ allows further simplification. Thus we arrive at the following equation for the static transfer curve:

$$P = \frac{\alpha(\gamma - I)}{\delta + (\gamma - I)}$$

This shows how the output concentration P depends on the input concentration I.

Further constraints must be introduced to create an inverter with well-defined signal levels. First, the concentration corresponding to a high logic value is defined as H. We require that $P = H$ when $I = 0$ and $P = 0$ when $I = H$. These conditions yield the constraints $\gamma = H$ and $\alpha = H + \delta$. Thus, to alter the static transfer curve, we may vary the parameters δ and α, while maintaining the relationship: $\alpha = H + \delta$. Figure 2 shows the transfer functions obtained by setting δ to a range of
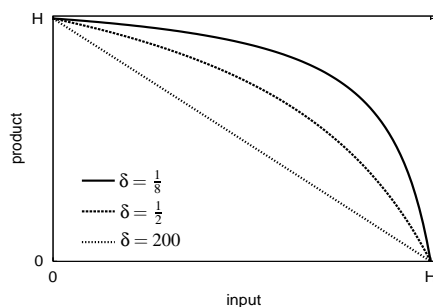
Figure 2: The static transfer curve for an inverter constructed from a NOT gate in an open system with several different values for δ: $\frac{1}{8}$, $\frac{1}{2}$, and 200. As δ increases the transfer curve approaches a straight line from high product and low input to low product and high input concentrations.

values. As δ is increased, the curve flattens out and becomes close to linear. As δ is decreased the curve becomes more bowed out (i.e., a large derivative).

This transfer curve is far from the sigmoid shape desired in digital computing. Any noise that moves the input concentration away from its digital value will propagate through to the output, possibly resulting in computational errors. However, we may construct inverters with differing static characteristics by concatenating several gates in a cascade. Details of this construction are given in [6].

While the static behavior addresses equilibrium characteristics, this analysis neglects the dynamic behavior of the system. Ultimately, the static transfer characteristic depends on only one ratio, but the dynamic behavior is less restricted and depends on several variables. However, this solution space is narrowed by the physical restrictions of our technology. We define a logical high value to be a concentration of 250 nM and calculate the rest of the parameters for the system. Figure 3 shows the results of numerical integration of the inverter working under such conditions. The reactor was started with zero concentration of all chemical species. The propagation delay of the inverter is $\approx 7.9 \cdot 10^3$ s, with a $t_{PHL} \approx 12.5 \cdot 10^3$ s, and $t_{PLH} \approx 3.2 \cdot 10^3$ s.

5   A Chemical Flip-Flop

Moving to an open system allows us to construct recurrent chemical and logical circuits—circuits with a lasting internal state or memory, which can change and be accessed over time. The simplest such system in digital logic is a *flip-flop*. This is
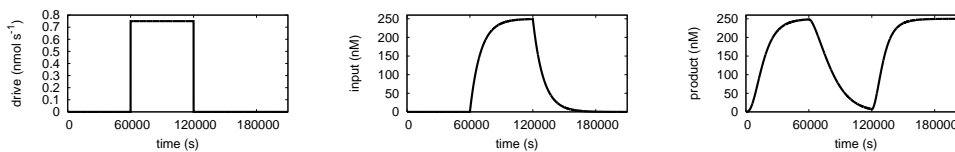
Figure 3: Left: external drive (molecular influx), as input to the circuit. Middle: concentration of input I inside the reactor. Right: product concentration P, the output of the inverter. The input concentration is moved from low, to high, and back to low at $6 \cdot 10^4$ s intervals.

simply a bistable system, which exhibits three behaviors depending on its inputs, commonly called *hold*, *set*, and *reset*. In the hold behavior, there are two stable states, which represent high and low outputs of the system. Set forces the system into its high stable state regardless of its previous state; similarly, reset forces it to its low stable state. Thus a flip-flop represents a single bit of memory, which can be stored (hold) or overwritten (set or reset).

A system that functions as a flip-flop can be constructed with a network of two NOT gates connected in a cycle of inhibition. A gate G_1 cleaves substrate S_1 to produce product P_1, which inhibits the catalytic activity of gate G_2; gate G_2 cleaves S_2 to produce P_2, which inhibits G_1 to complete the cycle. Output from the flip-flop is in terms of the concentration of the cleaved product P_2, with high or low concentration corresponding to a logical one or zero. The flip-flop is controlled by varying the influx of substrates 1 and 2 to the reactor, while gates 1 and 2 are continuously supplied.

We define constants G_1^m and G_2^m to be the rates of molecular influx of gate solutions. The external control is modeled by the functions $S_1^m(T)$ and $S_2^m(T)$, which describe the variable *molecular* influx of substrates at time T. The rate of efflux of the system is given by E. We define $P_1(T), P_2(T), S_1(T), S_2(T), G_1(T)$, and $G_2(T)$ to be the concentrations within the reactor at time T of product 1, product 2, substrate 1, substrate 2, gate 1, and gate 2, respectively. The system's dynamics

8

are modeled by the following system of six coupled differential equations:

$$\frac{dG_1}{dT} = \frac{G_1^m - EG_1(T)}{V}$$

$$\frac{dG_2}{dT} = \frac{G_2^m - EG_2(T)}{V}$$

$$\frac{dP_1}{dT} = \beta_1 S_1(T)\max(0, G_1(T) - P_2(T)) - \frac{EP_1(T)}{V}$$

$$\frac{dP_2}{dT} = \beta_2 S_2(T)\max(0, G_2(T) - P_1(T)) - \frac{EP_2(T)}{V}$$

$$\frac{dS_1}{dT} = \frac{S_1^m(T)}{V} - \beta_1 S_1(T)\max(0, G_1(T) - P_2(T)) - \frac{ES_1(T)}{V}$$

$$\frac{dS_2}{dT} = \frac{S_2^m(T)}{V} - \beta_2 S_2(T)\max(0, G_2(T) - P_1(T)) - \frac{ES_2(T)}{V}$$

where β_1 and β_2 are the reaction rate constants and V is the volume of the reactor.

We now examine the model to determine the conditions under which it will function as a flip-flop. Since we control it using substrate concentrations, we must determine how its output depends on these. To simplify the analysis, we take the gate concentrations to be constant (depending only on the efflux of the system). Hence we view the substrate concentrations as parameters of the system rather than dynamic variables. Now the behavior of the flip-flop can be understood using the following two-dimensional system:

$$\frac{dp_1}{dt} = r_1\max(0, g_1 - p_2(t)) - kp_1(t) \tag{5}$$

$$\frac{dp_2}{dt} = r_2\max(0, g_2 - p_1(t)) - kp_2(t) \tag{6}$$

where p_1 and p_2 are the product concentrations, r_1 and r_2 are lumped parameters representing substrate concentration and the reaction rate constant for the two reactions, g_1 and g_2 are the gate concentrations, and k is the decay constant. In order to function as a flip-flop, the system must have two stable states: a *set* state with a high value of p_2, and a *reset* state with a low value of p_2. Additionally there should exist control mechanisms to switch between states, in this case by varying the substrate concentrations. High concentration of both substrates is used to hold the flip-flop state, while the absence of one is used to set or reset the flip-flop. We now examine how the system will behave for various values of the parameters r_1 and r_2.

We begin our analysis by examining the *nullclines* of the system, i.e., the curves along which the time derivatives of the variables are constant. Setting equations (5) and (6) to zero yields:

$$\frac{dp_1}{dt} = 0 \quad \Rightarrow \quad p_1 = \frac{r_1\max(0, g_1 - p_2)}{k}$$

$$\frac{dp_2}{dt} = 0 \quad \Rightarrow \quad p_2 = \frac{r_2\max(0, g_2 - p_1)}{k}$$

9

The derivative on the nullclines will have components in a single direction, along either the p_1 or p_2 axis. This observation can be further qualified by inspecting equations (5)-(6). Notice that as we increase the value of p_1 the value of $\frac{dp_1}{dt}$ decreases, and as we decrease the value of p_1 the value of $\frac{dp_1}{dt}$ increases. This same relationship holds between p_2 and $\frac{dp_2}{dt}$. Thus a point above the p_1 nullcline will have a derivative in the negative p_1 direction and a point below the p_1 nullcline will have a derivative in the positive p_1 direction. We can use these facts to partition the phase space into regions where the signs of the derivatives are known [7].
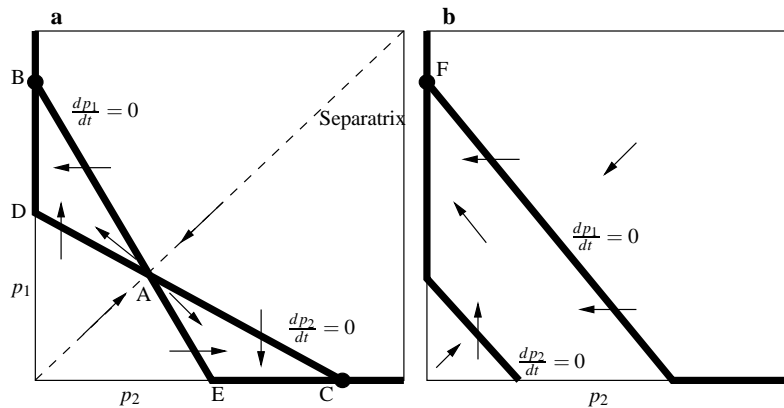

Figure 4: The geometric structure of the flip-flop equations. Dark lines indicate the nullclines, where p_1 or p_2 is unchanging. The arrows indicate the sign of the derivatives at various regions in the phase space. **a** The hold state: a bistable system with balanced substrate concentrations. Points B and C are the stable points of the system. Point A is an unstable saddle node. **b** The reset state: a mono-stable system caused by a low concentration of substrate 2. The nullclines intersect once at point F, which all locations will be attracted to.

Figure 4 shows two possible configurations of the nullclines along with the signs of the derivatives in different regions of the phase space. In the case shown in Figure 4**a** the two nullclines intersect (to create fixed points) at points A, B, and C. The stability of these points can be investigated by examining the slopes of the surrounding regions. The intersection at point A has four adjacent regions. Two of the regions adjacent to A have derivatives pointing toward A, but two of the regions have derivatives away from A. This reveals that A is a *saddle point*: perturbations in one direction will return back to A, while perturbations in another direction will fall away from A. B and C, on the other hand, are stable. The separatrix shown in Figure 4**a** divides the phase space into two basins of attraction. All trajectories

10

starting above the separatrix will be attracted to B; all trajectories starting below it will be attracted to C. This configuration represents a hold state of the flip-flop as the system will evolve to either B or C. B is the low state of the flip-flop, in which the system has a low concentration of p_2 and a high concentration of p_1. Similarly, C is the high state, with a high concentration of p_2 and a low concentration of p_1. Thus we have two stable states and a valid flip-flop.

If the nullclines do not intersect in the positive quadrant then there will be a single intersection on either the p_1 or the p_2 axis. Such a situation is shown in Figure 4**b**. The point F is a stable node with a low p_2 value and a high p_1 value. An examination of the three regions of the phase space reveals that all points will be attracted to point F. Thus this state will reset the flip-flop. For bistability to occur the two nullclines must intersect in the orientation shown in Figure 4**a**. This requires that point B be above point D and that point C be to the right of point E, and translates into the following conditions: $kg_2 < r_1g_1$ and $kg_1 < r_2g_2$. If we wish to maintain symmetry, we can set $g_1 = g_2$ and $r_1 = r_2 = r$ and the conditions become $k < r$. When this condition is met the flip-flop will be in a hold state. The flip-flop is set or reset when the two nullclines do not intersect in the positive quadrant. Then there will exist one stable fixed point at their intersection (point F in Figure 4**b**). This will be the case if either r_1 or r_2 takes the value of zero. Suppose that the flip-flop is in the hold state with a high concentration of p_2 at point C in Figure 4**a**. As the parameter r_2 is reduced, the slope of the $\frac{dp_2}{dt}$ nullcline will decrease and the stable point C will slide down the p_2 axis toward point E. As C passes E the fixed point is annihilated, and the system will jump to the only fixed point: point F in Figure 4**b**. After the system has gone to point F the parameter r_2 can be returned to its original value and the flip-flop will remain in its low state. This lack of reversibility as the parameter is varied is called *hysteresis* in dynamical systems [7].

This stability analysis reveals the constraints we need to satisfy for the system to function as a flip-flop. First, in order to maintain output symmetry we require that $g_1 = g_2$. This ensures that both the output of the flip-flop (p_2) and the negated output (p_1) have the same value for a logical high. Next, in order to maintain a symmetric separatrix along the line $p_1 = p_2$ we require that $r_1 = r_2 = r$. Thus, during the hold state of the flip-flop, the phase-space is equally divided between values which are attracted to the high state and values which are attracted to the low state. Finally, the constraint for bistability is $k < r$. This means that the rate at which concentrations decay in the reactor (by means of outflow) must be less than the rate at which the enzymatic gates can create product. In other words, the gates must be capable of creating product faster than product is being removed.

We then convert these constraints on the dimensionless model to specifications for a physical system. Figure 5 shows a numerical integration of the system over

a period of $9.6 \cdot 10^4$ s. At $2.4 \cdot 10^4$ s intervals the system is moved between set, hold, reset, and hold operations by controlling the influx of substrates. The top two traces on the left show substrate molecular influx rates. The top two traces on the right show the corresponding concentrations of substrates. The bottom two traces show concentrations of products 2 and 1. These represent the output and the negated output of the system, respectively.
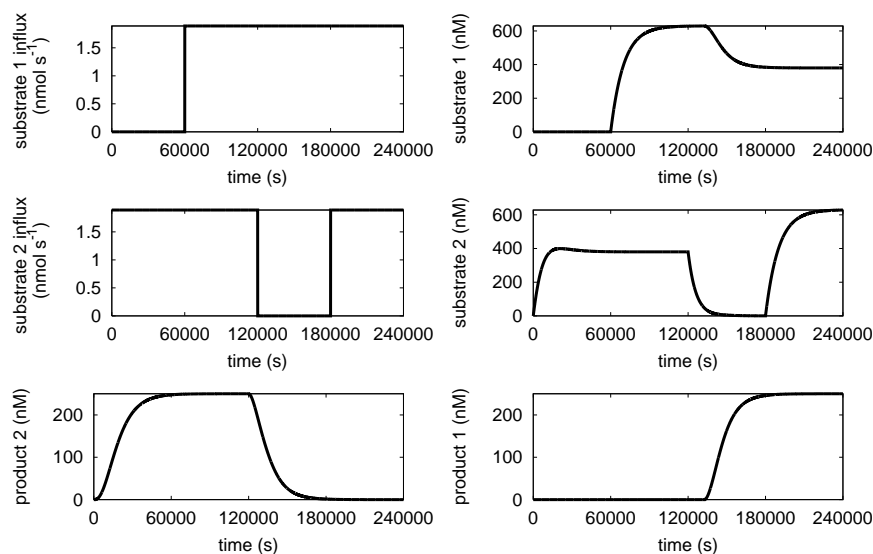

Figure 5: Exercising control over the flip-flop system.

6 Conclusions

Deoxyribozyme logic gates may be used to construct a biomolecular computer. By moving to open reactors, we increase the computational abilities of the underlying logic gates by making it possible to build recurrent circuits and devices with feedback. Techniques from dynamical systems offer qualitative and quantitative insights about the behavior of these chemical networks. Using these techniques, we have designed two fundamental components of a biomolecular computer: a single-bit inverter and a flip-flop that provides a single bit of memory. Compared to electronic computers, this technology is slow (about 1 mHz) but the possibility it offers of amorphous computation inside living cells is extremely exciting.

References

[1] M. N. Stojanovic, T. E. Mitchell, and D. Stefanovic, "Deoxyribozyme-based logic gates," *Journal of the American Chemical Society*, vol. 124, pp. 3555–3561, Apr. 2002.

[2] M. N. Stojanovic and D. Stefanovic, "Deoxyribozyme-based half adder," *Journal of the American Chemical Society*, vol. 125, pp. 6673–6676, May 2003.

[3] M. N. Stojanovic and D. Stefanovic, "A deoxyribozyme-based molecular automaton," *Nature Biotechnology*, vol. 21, pp. 1069–1074, September 2003.

[4] M. Hiratsuka, T. Aoki, and T. Higuchi, "Enzyme transistor circuits for reaction-diffusion computing," *IEEE Transactions on Circuits and systems—I: Fundamental Theory and Applications*, vol. 46, pp. 294–303, Feb. 1999.

[5] I. R. Epstein and J. A. Pojman, *An Introduction to Nonlinear Chemical Dynamics*. New York: Oxford University Press, 1998.

[6] C. Morgan, "Units of a biomolecular computer constructed from networks of modular chemical gates in an open reactor," Undergraduate thesis, University of New Mexico, 2003.

[7] S. H. Strogatz, *Nonlinear Dynamics and Chaos (With Applications to Physics, Biology, Chemistry, and Engineering*. Addison-Wesley, 1994.