Four Steady State Switch Using Transcriptional Logic

Jessie Hsu
University of Washington
Department of Biology

WenTao Luo
University of Notre Dame
Department of Chemical Engineering

Khang Tran
University of California Berkeley
Department of Electrical Engineering

Abstract

From networks of simple genetic regulatory elements, scientists have executed some genetic circuits such as the circadian oscillator[1] and the toggle switch[2]. These circuits utilize the property of multi-stability as provided by regulatory elements. The multi-stability in genetic regulatory networks inspired our group to theorize possible multiple steady states models that could later be carried out in vivo. The motivation is to eventually use these steady states to store information in a reliable way. Ideally, a transient input will lead to a permanent storage unit. Furthermore, the multiple steady states implement the finite state machine. Relevant applications in bio-computing and gene therapy[3]-[4] can also benefit from this type of research. Simple genetic regulatory network can be modeled by ordinary differential equations. By analysis of the equations, one can gain insight to the different behaviors of a genetic circuit. This paper theorizes and discusses two different models of four-steady state genetic regulatory networks and analyzes their equations.

1 Introduction

What is a steady state? In the context of genetic regulatory networks, a steady state[5]-[6] is a condition where there is no net change in the system. The system is stable if slight perturbations bring the system back to its steady state. More significant perturbations may move the system to a new steady state. Using four repressible promoters arranged in a mutually inhibitory network, four steady states can be realized. One can use transient chemical or thermal induction to move from state to state. The toggle switch experiment exhibited a nearly ideal switching threshold; hence, with the right inducers, achieving a four steady states system should be possible.

2 The Model

2.1 First model

In the first model, the four steady states system in this case is composed of four promoters and four corresponding repressors (Fig. 1). Each promoter transcribes a gene that codes for a repressor, which in turn inhibits the other three promoters. A steady state is measured by the concentration of protein product. There are four possible stable states: promoter A transcribes repressor A, promoter B transcribes repressor B, promoter C transcribes repressor C, and promoter D transcribes repressor D. By transiently introducing an inducer, a repressor is maximally transcribed until it stably represses the other three promoters.

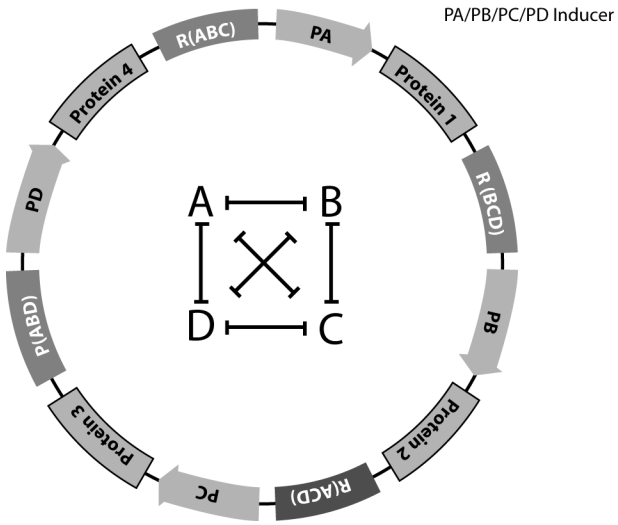


Figure 1: Mutually repressive states

Since the proposed four steady states system is a direct extension from the toggle switch experiment, we varied the steady state equations found in Gardner's research to correspond to our proposal. We can mathematically visualize the behavior of the system by understanding the following equations modeling for the network:

$$\begin{cases} \dfrac{dA}{dt} = \dfrac{\alpha_A}{1+B^{\beta_b}} + \dfrac{\alpha_C}{1+D^{\beta_d}} + \dfrac{\alpha_D}{1+C^{\beta_c}} - A \\ \dfrac{dB}{dt} = \dfrac{\alpha_B}{1+A^{\beta_a}} + \dfrac{\alpha_C}{1+D^{\beta_d}} + \dfrac{\alpha_D}{1+C^{\beta_c}} - B \\ \dfrac{dC}{dt} = \dfrac{\alpha_C}{1+D^{\beta_d}} + \dfrac{\alpha_B}{1+A^{\beta_a}} + \dfrac{\alpha_A}{1+B^{\beta_b}} - C \\ \dfrac{dD}{dt} = \dfrac{\alpha_D}{1+C^{\beta_c}} + \dfrac{\alpha_B}{1+A^{\beta_a}} + \dfrac{\alpha_A}{1+B^{\beta_b}} - D \end{cases} \tag{1}$$

A is the concentration of repressor A, B is the concentration of repressor B, C is the concentration of repressor C, D is the concentration of repressor D. α_A is the effective rate of synthesis of repressor A which represses promoter B, C, and D; β_A is the cooperativity of repression of promoter A, etc.

After the critical values are found, one can determine the stability through linear analysis. Since the steady state equation is a nonlinear equation, one must linearize the equations to form the Jacobian matrix and then examining its eigenvalues. The Jacobian matrix is:

The first three terms in every equation are the cooperative repression of constitutively transcribed promoters, and the last term in every equation represents the degradation of the repressors. The parameters α_A, α_B, α_C, α_D factor into the equation the net effect of RNA polymerase binding, open-complex formation, transcript elongation, transcript termination, repressor binding, ribosome binding and polypeptide elongation. The parameters β_A, β_B, β_C, β_D come from the multimerization of the repressor proteins and the cooperative binding of repressor multimers to multiple operator sites in the promoter.

To reach steady state, dA/dt, dB/dt, dC/dt, and dD/dt must approach zero. dA/dt means change in concentration of the repressor A over time. Setting the four steady states equation equal to zero, one can find the critical values where the equation is not differentiable or the derivative of a function is equal to zero.

After the critical values are found, one can determine the stability through linear analysis. Since the steady state equation is a nonlinear equation, one must linearize the equations to form the Jacobian matrix and then examining its eigenvalues. The Jacobian matrix is:

$$J = \begin{bmatrix} -1 & -\frac{\alpha_A B^{\beta_B}\beta_B}{\left(1+B^{\beta_B}\right)^2 B} & -\frac{\alpha_D C^{\beta_C}\beta_C}{\left(1+C^{\beta_C}\right)^2 C} & -\frac{\alpha_B A^{\beta_A}\beta_A}{\left(1+A^{\beta_A}\right)^2 A} \\ -\frac{\alpha_C D^{\beta_D}\beta_D}{\left(1+D^{\beta_D}\right)^2 D} & -1 & -\frac{\alpha_D C^{\beta_C}\beta_C}{\left(1+C^{\beta_C}\right)^2 C} & -\frac{\alpha_B A^{\beta_A}\beta_A}{\left(1+A^{\beta_A}\right)^2 A} \\ -\frac{\alpha_C D^{\beta_D}\beta_D}{\left(1+D^{\beta_D}\right)^2 D} & \frac{\alpha_A B^{\beta_B}\beta_B}{\left(1+B^{\beta_B}\right)^2 B} & -1 & -\frac{\alpha_B A^{\beta_A}\beta_A}{\left(1+A^{\beta_A}\right)^2 A} \\ -\frac{\alpha_C D^{\beta_D}\beta_D}{\left(1+D^{\beta_D}\right)^2 D} & \frac{\alpha_A B^{\beta_B}\beta_B}{\left(1+B^{\beta_B}\right)^2 B} & -\frac{\alpha_D C^{\beta_C}\beta_C}{\left(1+C^{\beta_C}\right)^2 C} & -1 \end{bmatrix} \tag{2}$$

Now the steady-state equations are linearize into the form $x' = Px$

$$x' = \begin{bmatrix} J_{11} & \dots & & J_{14} \\ \vdots & \ddots & & \\ & & \ddots & \vdots \\ J_{41} & & \dots & J_{44} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} \tag{3}$$

The Eigenvalues λ satisfy the equation $(\lambda I - 1)\, x = 0$.

$$\begin{bmatrix} \lambda - J_{11} & \dots & \dots & -J_{14} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ -J_{41} & \dots & \dots & \lambda - J_{44} \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = 0 \tag{4}$$

The eigenvalues are the roots of the equation, which can be calculated by taking the determinant of $(\lambda I - 1)$. Linear stability analysis states that an equilibrium state exists if all the eigenvalues of the Jacobian result in negative real part.

To demonstrate how this works, α_A and α_C of 156.25 α_B and α_D of 15.6 and β_A and β_C of 2.5 β_B and β_D of 1 were chosen from Gardner's paper[3]. Equation (1) yields 2 positive real solutions. Imaginary solutions and negative solutions were discarded because they do not have a physical meaning. The two solutions were:

A = 24.015 B= 12.016 C= 24.015 D= 12.016
A= 318.337 B= 317.867 C= 0.098 D= 317.867

By substituting the solution into the Jacobian and taking the determinate, the corresponding eigenvalues are λ = -0.192, -0.993,-0.993, -0.077 and λ = -0.999, -1.176, -0.826, -0.998. Since all the eigenvalues are negative, both solutions are at a stable steady state.

Equation (1) was plotted numerically using ODE45 in Matlab (Mathworks)

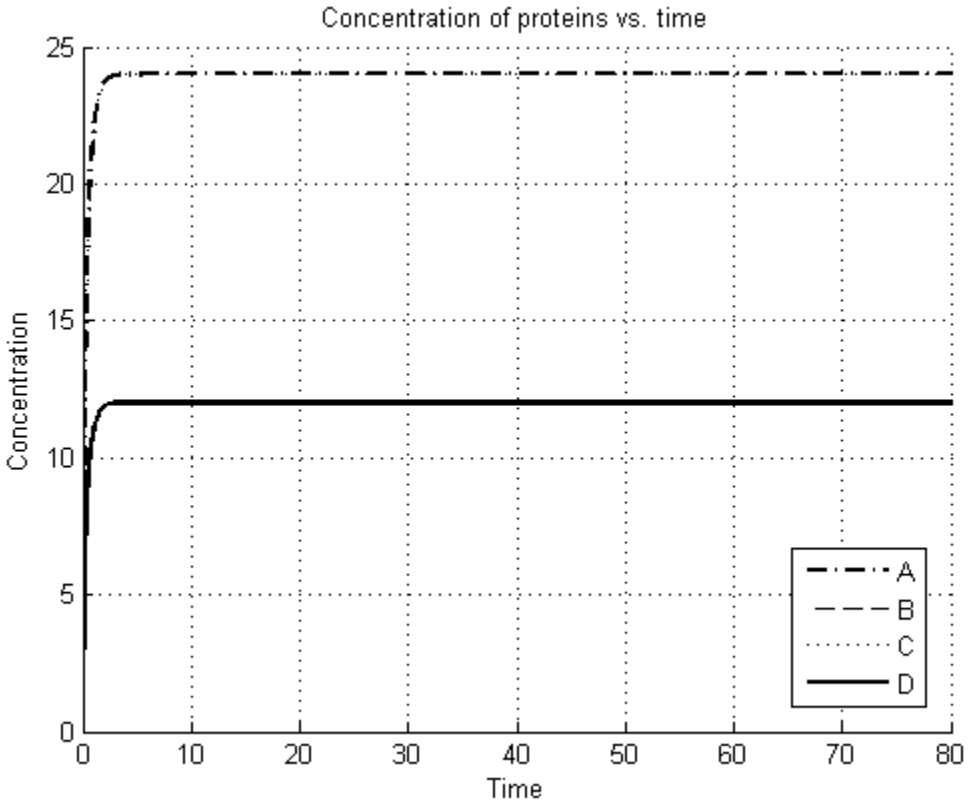


Figure 2: Numeric solution of equation (1) using setting α_A and α_C of 156.25, α_B and α_D of 15.6, and β_A and β_B of 2.5, β_C and β_D of 1, initial condition A,B,C,D= 0.

In the numeric solution, the concentration of A and C converge 24.015 and concentration of B and D converge to 12.016. Those concentration matches with the first set of critical values, A = 24.015, B= 12.016, C= 24.015, D= 12.016. The solution shows a stable behavior as time increases.

Several interesting things emerge from simulating equation 1. First, the equations are extremely parameter sensitive. An insignificant change to the parameter will result in drastic consequences. For example, changing α_A to 160, a 2.4 percent increase. The numeric solution changes to A = 24.436, B = 10.373, C = 24.436, D = 14.078, figure 3. Second, the equations are depended on initial condition. For example, changing the initial condition from 0 for all species to D = 100, it takes approximately 20 times longer to reach stability, figure 4. Therefore, time to reach stability can be controlled by initial concentration of species. However, the effect of initial condition depends on the parameter too. For example, changing the initial condition of A to 100, it only takes approximately 2 times longer to reach stability, figure 5.

Figure 2 with Modified α_A

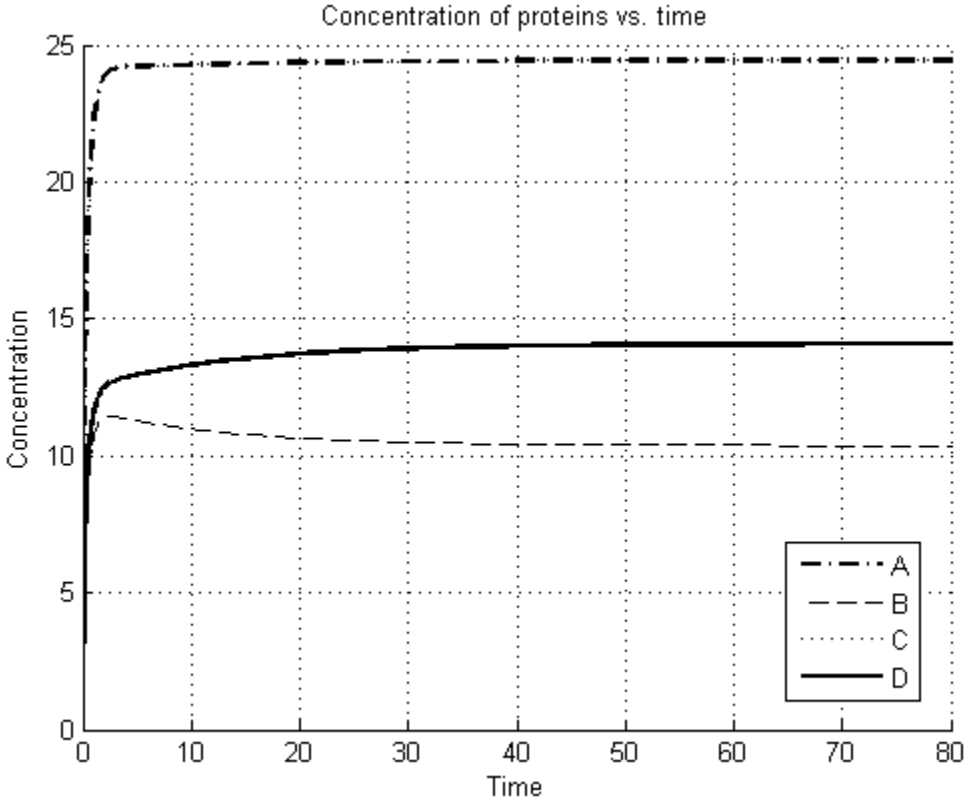


Figure 3: Numeric solution of equation (1) using setting α_A of 160, and α_C of 156.25, α_B and α_D of 15.6, and β_A and β_B of 2.5, β_C and β_D of 1, initial condition A,B,C,D= 0.

Figure 2 with Modified Initial Condition of D

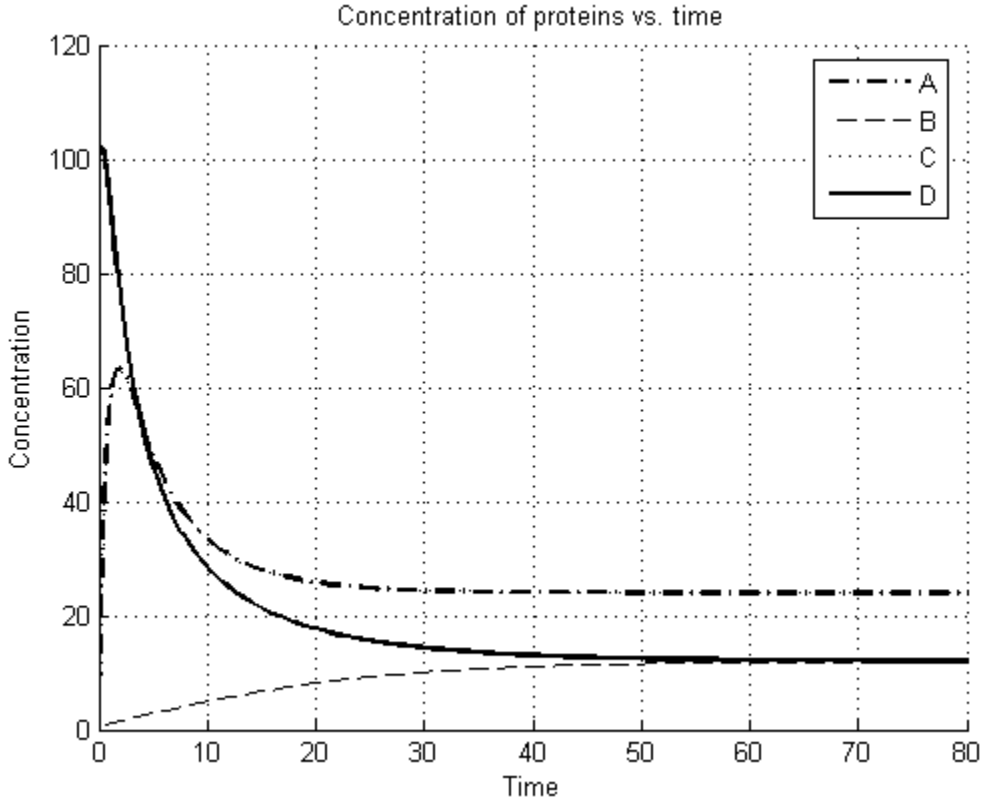


Figure 4: Numeric solution of equation (1) using setting α_A and α_C of 156.25, α_B and α_D of 15.6, and β_A and β_B of 2.5, β_C and β_D of 1, initial condition A,B,C = 0, and $D = 0$.

Figure 2 with Modified Initial Condition of A

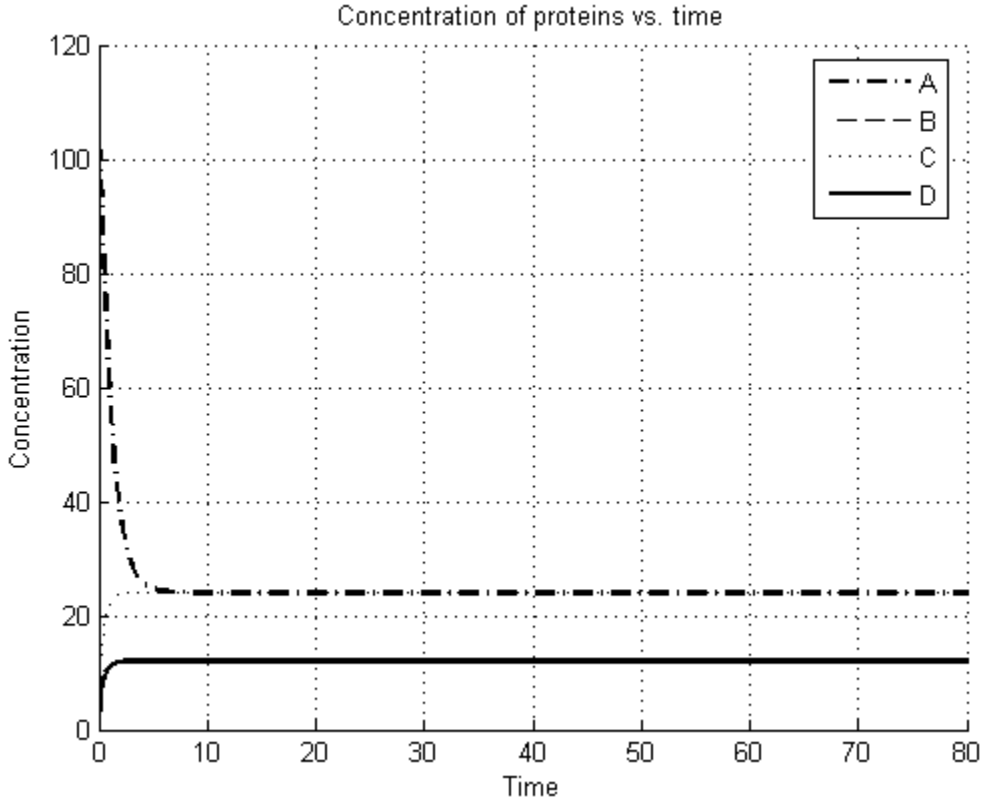


Figure 5: Numeric solution of equation (1) using setting α_A and α_C of 156.25, α_B and α_D of 15.6, and β_A and β_B of 2.5, β_C and β_D of 1, initial condition A = 100, and B,C,D = 0

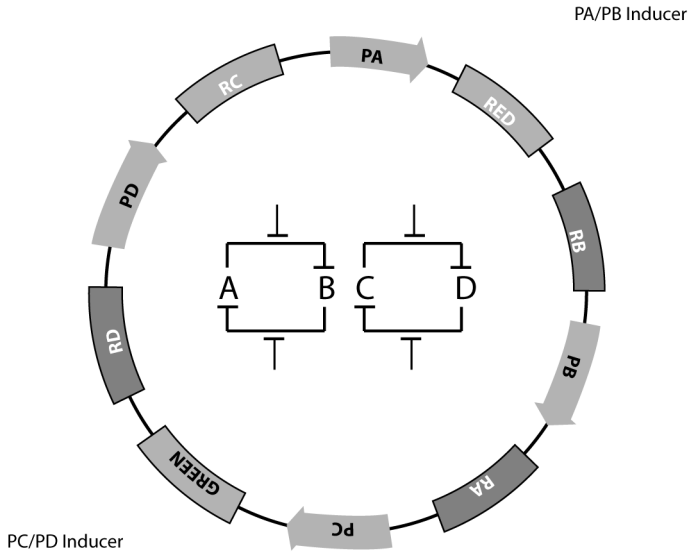
One can use stability data over a range of parameters to generate solution branch diagrams. Using the solution branch diagram, one can locate the shift from stable to unstable steady state and bifurcation point (split into stable and unstable steady state). Whether the equation is chaotic or stable, such findings will enhance one's comprehension on the different regimes of behavior within a particular genetic regulatory circuit. With a solution branch diagram, one can narrow down the physical constraints of the genetic circuit and optimize its performance by picking the best conditions or ingredients for the circuit to operate. Another use of the equation would be if a genetic circuit must meet a concentration condition, one can use the data to fit the equation and find the parameters. Once parameters are found, one can then find repressors and promoters which fit these ranges.

As simple as it might sound, there are problems that could surface along the way. First of all, the steady state equations are quite complex. For a four steady states system, there are eight unknown parameters and four unknown variables in four coupled differential equations. This can be solved by taking into account the symmetry of the equations. Finding genes that have same α and β will decrease the complexity of the equation. This model does present some problems. First, the system may still not be biologically feasible for several reasons. It is difficult to find promoters and repressors that fit into the ranges specified by the equations, plus four rates have to be considered simultaneously. Since one of the applications of the multiple steady states system is implementing the finite state machine[7], there will be timing issues. The rates will vary, meaning transferring from state to state requires different amounts of time, which is a set back since state transition is based on clock edge trigger in the current technology. Information storage can be

complicated when the accessing of one unit takes longer than another. The problem does not stop here. The complexity will scale up exponentially when the number of states increases. After realizing such problems, another method was proposed to reduce the complexity.

2.2 Second model

The second model for a four steady state system is a direct extension of the bi-stable toggle switch[2]. This model is the superposition of two sets of toggle switches, each constituted of two genes that mutually repress the promoter of the other gene in the set. In order to move between states and reach the desired state, each promoter has a corresponding inducer that activates the promoter. By combining the states of two sets, one can achieve four states, figure 6.



Combinatorial bi-stable toggle switches

Inducer PA causes Promoter A to be turned on, consequently transcribing the Red gene and Repressor B. Both gene products are then translated into protein. For the remainder of the paper, stating that a gene is colored means the protein product fluoresces the specified color. Repressor B binds Promoter B, preventing transcription factors from binding Promoter B and initiating transcription of genes downstream of Promoter B. Conversely, when Promoter B is induced by some external factor, Repressor A is transcribed and binds Promoter A, preventing transcription of the Red gene. Similarly, the exact situation is found for the other set of genes (C and D). Transcription of the Green gene and Repressor D results in the repression of Promoter D, while transcription of Repressor C causes repression of Promoter C, preventing production of the Green gene. From two independent toggle switches, one has a system that can achieve four different steady states. The four states are as follows:

PA (Red)	PB	PC (Green)	PD	Color = State
0	1	0	1	None
0	1	1	0	Green
1	0	0	1	Red
1	0	1	0	Yellow

0 represents an inactive promoter, while 1 represents an active promoter. The combination of the Red and Green genes gives a Yellow gene. We found some potential promoter-repressor pairs from reading the literature of similar experiments and from the BioBricks database based on comparable activation and repression rates. These include Ptrc-2/LacI, Pslcon/ λcI, PtetO-1/TetR, and P2/P22 c2 [8]. Potential inducers include heat, isopropyl-b-D-thiogalactopyranoside (IPTG), and anhydrotetracycline (aTc). The equation is similar to the first model, but less complicated.

$$\begin{aligned}
\frac{dA}{dt} &= \frac{\alpha_A}{1+B^{\beta_b}} - A \\
\frac{dB}{dt} &= \frac{\alpha_B}{1+B^{\beta_a}} - B \\
\frac{dC}{dt} &= \frac{\alpha_C}{1+D^{\beta_d}} - C \\
\frac{dD}{dt} &= \frac{\alpha_D}{1+C^{\beta_c}} - D
\end{aligned} \tag{5}$$

A is the concentration of repressor A, B is the concentration of repressor B, C is the concentration of repressor C, D is the concentration of repressor D. α_A is the effective rate of synthesis of repressor A which represses promoter B; β_A is the cooperativity of repression of promoter A, etc.

The mathematical analysis of equation (5) from model 2 is same as model 1. Since the equation is non-linear, one must find the jacobian and calculate the eigenvalue in order to determind the stability of the critical values.

Since a two-gene-promoter toggle switch has been proven to work, using two of these toggle switches should also work provided they do not interact with each other. Some issues that we must resolve include the difference in reaction rate, which must be set in such a way that there is no false transient state. If we want to achieve more states, this model is still has many drawbacks. The factor of greatest concern is the often uncontrollable promiscuous interaction between proteins.

3 Future Work

Future work will include testing out the models in vivo to compare with the theoretical results. Other interesting work might be to use direct evolution to evolve genes to fit the specific context of a genetic circuit. To generate bifurcation diagram and locate the equilibrium states and bifurcation of equilibria, Interval-Newton/generalized-bisection algorithm can be implemented to increase accuracy and ensure that all roots of a non-linear equation system are enclosed. There may also be more ways to implement the goal of achieving multiple steady states using genetic regulatory circuits which are worth pursuing. We hope to find a more efficient design to realize multiple steady states without complications exponentially scaling up as the number of states increases.

Acknowledgement
Yi-Chun Chen, Cornell University
Kenneth L. Ho, California Institute of Technology
Dr. Michael Elowitz
Dr. Erik Winfree
Marcia Punsalan
Synthetic Biology group
CBSSS

References

[1] Elowitz MB, Leibler S. "A Synthetic Oscillatory Network of Transcriptional Regulators." Nature. Vol. 403. 335-338. 2000.

[2] Gardner Timothy S., Cantor Charles R., and Collins James J. "Construction of a Genetic Toggle Switch in Escherichia coli." Letters to Nature. Vol. 203. 339-342. 20 January 2000.

[3] Benenson Yaakov, Gil Binyamin, Ben-Dor Uri, Adar Rivka, and Shapiro Ehud. "An autonomous molecular computer for logical control of gene expression." Letters to Nature. Vol. 429. 423-429. 27 May 2004.

[4] Buchler Nicolas E., Gerland Ulrich, and Hwa Terence. "On Schemes of Combinatorial Transcription Logic." PNAS. Vol. 100. No. 9. 5136-5141. 29 April 2003.

[5] Campbell Neil A. Biology. Fourth Edition. 297-348. The Benjamin/Cummings Publishing Company. Menlo Park, CA 94025. 1996.

[6] Alberts Bruce, Bray Dennis, Lewis Julian, Reff Martin, Roberts Kuth, and Watson James D. Molecular Biology of the Cell. Third Edition. 417-426. Garland Publishing. New York, NY 10003. 1994.

[7] Benenson Yaakov, Paz-Elizur Tamar, Adar Rivka, Keinan Ehud, Livneh Zvi, Shapiro Ehud. "Programmable and Autonomous Computing Machine Made of Biomolecules." Letters to Nature. Vol. 414. 430-434. 22 November 2001.

[8] Olmec Group, http://biobricks.ai.mit.edu/IAP_Projects/Olmec.