Characterization and Application of trans-cleaving hammerhead ribozymes in genetic circuits

Ryan Tsoi *, **Zackery Field** *

*University of California, Berkeley

Submitted to Adam Arkin and Ron Weiss for review

Our current ability to engineer biological circuits is hindered by the inability to design and create systems without unintended crosstalk or side interactions from a limited set of inducer-promoter pairs. This creates poor scalability of complex designs that are restricted by the number of well-characterized DNA regulators, and limits the overall complexity that can be encoded into a genetic circuit. To address these problems, we propose the use of RNA based logic as the building block for genetic engineering and replace inducer-promoter pairings with riboregulators such as ribozymes. RNA should be superior to DNA with regards to the goal of synthetic biology to produce safe, reliable, predictable, composable, tunable, and orthogonal genetic circuits. Two well-characterized circuits, the toggle switch and the repressilator, were investigated, and RNA-based equivalents were constructed from trans-cleaving hammerhead ribozymes. The outputs of the riboregulated toggle and repressilator were compared to their DNA counterparts, and it was found that the ribozyme systems did not achieve the desired output due to the fast rates of transcription and degradation of mRNA. This work introduces the prospect of using trans-cleaving ribozymes as a replacement for promoter and repressor protein pairs, and presents challenges that must be confronted in the future if genetic regulation by riboregulatory elements is to replace genetic regulation by protein regulatory elements.

Synthetic Biology | Orthogonal part sets | Rational Logic Gate Design

Introduction

The world is currently facing a large variety of problems ranging from accessibility of environmentally friendly energy to curing disease, and researchers are looking to microorganisms as a possible solution due to the diversity in microbial species as well as genetic codes. It is believed that once these genetic codes are identified and understood, they can be manipulated to help solve these problems. [1]

Over the past few decades, bacteria have been genetically engineered to produce a variety of compounds ranging from commodity chemicals to pharmaceuticals. They have been used in agriculture, medicine, bioremediation, and even treatments and therapies for complicated diseases like cancer.

However, synthetic biology is still far from being an engineering science, and we can look to its macroscale analog electrical engineering to see why. Electrical engineers can use the same transistor or wire multiple times within the same circuit and achieve the desired output without undesired side interactions. [2] In contrast, current methods of genetic circuit design are limited by the inability to produce orthogonal systems from similar parts.

Without physical barriers separating individual molecules or reactions, it is impossible to prevent crosstalk within the circuit and severely inhibits our ability to regulate these systems. As a result, effective circuit designs must contain all unique promoter-inducer combinationsthis hinders the overall complexity that can be achieved but also is limited by the total number of identified and well-characterized pairings.

We seek to address these concerns by proposing that researchers shift their focus within the central dogma of biology from protein-based regulators to RNA-based regulators for circuit design. While proteins are the most commonly investigated for synthetic biology applications, in many ways RNA is more suited to meet these design principles than proteins.

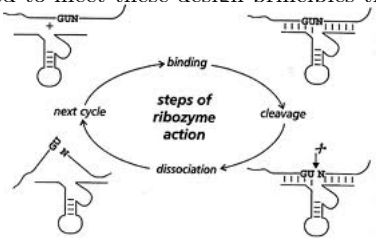

Fig. 1. Depiction of the action of a trans-cleaving ribozyme

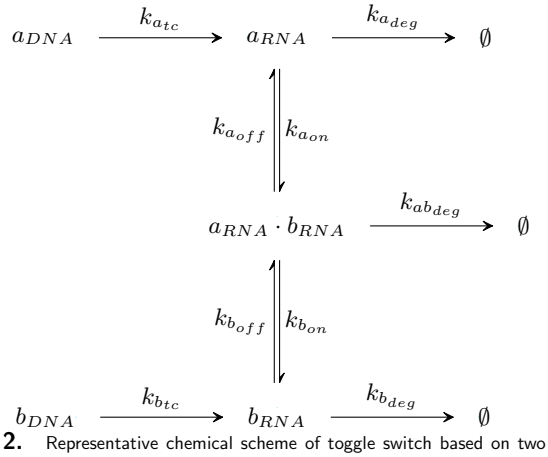

Fig. 2. Representative chemical scheme of toggle switch based on two trans-cleaving hammerhead ribozymes. Ribozyme A cleaves ribozyme B and vice versa after reversibly binding to one another. Translation is not depicted and is assumed to have no effect on transcription nor riboregulation in the circuit.

Reaction Constant	Chemical Equation	Parameter Value (molecules/s)
k_{atc}	$A_{DNA} \rightarrow A_{RNA}$	0.6897
k_{btc}	$B_{DNA} \rightarrow B_{RNA}$	0.6897
$k_{a,deg}$	$A_{RNA} \rightarrow \emptyset$	0.0033
$k_{b,deg}$	$B_{RNA} \rightarrow \emptyset$	0.0033
$k_{a,on}$	$A_{RNA} + B_{RNA} \rightarrow A_{RNA} \cdot B_{RNA}$	101.4
$k_{b,on}$	$A_{RNA} + B_{RNA} \rightarrow A_{RNA} \cdot B_{RNA}$	101.4
$k_{a,off}$	$A_{RNA} \cdot B_{RNA} \rightarrow A_{RNA} + B_{RNA}$	0.05
$k_{b,off}$	$A_{RNA} \cdot B_{RNA} \rightarrow A_{RNA} + B_{RNA}$	0.05
$k_{ab,deg}$	$A_{RNA} \cdot B_{RNA} \rightarrow \emptyset$	0.0033

Fig. 3. Table 1. Approximated parameters of individual chemical reactions within the riboregulated RNA toggle switch. Values represent rate constants for the reactions and are averaged in units of molecules/s for purposes of Gillespie stochastic modeling. Constants are calculated based on constants for Escherichia coli bacteria.

Reserved for Publication Footnotes

For example, RNA is easier to design and tune because we can take advantage of Watson-Crick base pairing, whereas there has been minimal success in designing novel proteins. [3] [4] In addition, a wider variety of targets are available by simply knowing the DNA sequence of the target protein. Production and degradation of mRNA are faster than equivalent processes for protein, and mRNA production puts less energetic strain on the cell than protein production. [5] [6] Finally, physical properties such as open complex formation [3], elongation, translation [7], and degradation rates [8] can all be controlled in the mRNA itself.

In this paper, we compare the advantages and disadvantages between riboregulators and protein regulators by looking at two equivalent versions of two different genetic circuitsa toggle switch [9], and a repressilator. One uses hammerhead ribozymes regulating an mRNA circuit, and the other uses traditional promoter-repressor pairings. The repressilator is a synthetic genetic regulator network, which exhibits a stable oscillation similar to an electrical oscillator system with set time periods [10]. We model both circuits stochastically, and optimized tunable parameters to give the best output. Using both models, we test hypotheses concerning the ideal design principles mentioned previously and reach conclusions for the advantages of each type of circuit.

Results

Implementation of RNA Toggle Switch Using trans-cleaving hammerhead ribozymes. Two hammerhead ribozymes with corresponding cleavage sites on the opposing ribozymes were computationally designed for transcription in Escherichia coli and stochastically modeled in Java using the Gillespie algorithm in order to simulate an RNA logic toggle switch. The final system did not display switch-like behavior however, and the steady states of the ribozyme toggle switch were not consistent with the classic toggle circuit. In the RNA circuit (See Fig. 2,3(Table 1),4) , steady state concentrations of all species were consistently zero despite changing inlet parameters. At longer timescales, concentration of the species was dominated by degradation, as new ribozyme constructs produced by transcription participated immediately in binding and trans-cleaving of one another. As a result, the ribozyme in higher concentration would cleave all of the ribozyme in lower concentration, but inevitably degrade over the course of the reaction due to high RNA production and degradation rates relative to proteins. The system also lacked memory, a key characteristic of the toggle switch, which limited our designed circuit from being a true analog to the toggle circuit.

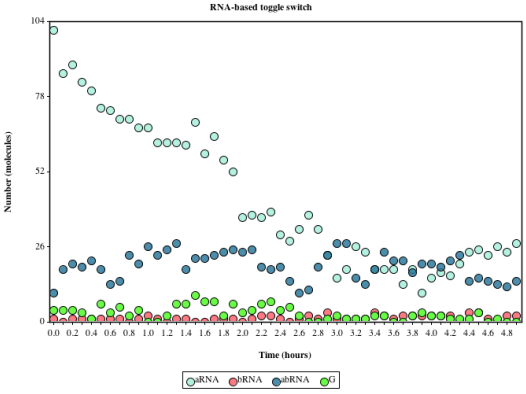

Fig. 4. Timecourse of RNA-based toggle switch with two ribozymes from Gillespie algorithm. Given an initial concentration of ARNA, concentration does not reach some steady state value and eventually drops to zero molecules after 10 hours. BRNA remains low and ABRNA seems to stay constant throughout the reaction despite high initial concentratinos of ARNA, implying high amounts of degradation of ARNA at early times. G is concentration of reporter protein GFP which is encoded on a separate mRNA strand that is cleaved by ribozyme A. When A is not present, GFPRNA is free to be translated and thus fluoresce.

Implementation of RNA Repressilator Using Ribozymes. We also looked at incoporating three hammerhead ribozymes with corresponding cleavage sites into a circuit in a rock-paper-scissors fashion in order to closely simulate the repressilator by [10] in E. coli. This circuit was modeled stochastically in

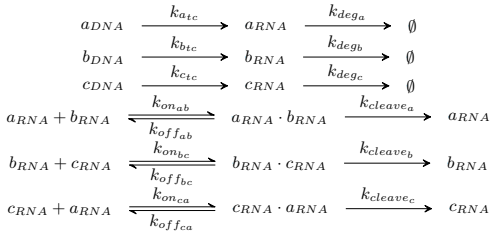

Fig. 5. Representative chemical scheme of repressilator based on three trans-cleaving hammerhead ribozymes in a rock-paper-scissor interaction. Ribozyme A binds to and cleaves ribozyme B, ribozyme B binds to and cleaves ribozyme C, and ribozyme C binds to and cleaves ribozyme A. Translation is not depicted and is assumed to have no effect on transcription nor riboregulation in the circuit.

Reaction Constant	Chemical Equation	Parameter Value (molecules/s)
k_{atc}	$A_{DNA} \rightarrow A_{RNA}$	0.6897
k_{btc}	$B_{DNA} \rightarrow B_{RNA}$	0.6897
k_{ctc}	$C_{DNA} \rightarrow C_{RNA}$	0.6897
$k_{a,deg}$	$A_{RNA} \rightarrow \emptyset$	0.0033
$k_{b,deg}$	$B_{RNA} \rightarrow \emptyset$	0.0033
$k_{c,deg}$	$C_{RNA} \rightarrow \emptyset$	0.0033
$k_{ab,on}$	$A_{RNA} + B_{RNA} \rightarrow A_{RNA} \cdot B_{RNA}$	101.4
$k_{bc,on}$	$B_{RNA} + C_{RNA} \rightarrow B_{RNA} \cdot C_{RNA}$	101.4
$k_{ca,on}$	$C_{RNA} \cdot A_{RNA} \rightarrow C_{RNA} + A_{RNA}$	101.4
$k_{ab,off}$	$A_{RNA} \cdot B_{RNA} \rightarrow A_{RNA} + B_{RNA}$	0.05
$k_{bc,off}$	$B_{RNA} \cdot C_{RNA} \rightarrow B_{RNA} + C_{RNA}$	0.05
$k_{ca,off}$	$C_{RNA} \cdot A_{RNA} \rightarrow C_{RNA} + A_{RNA}$	0.05
$k_{cleave,a}$	$A_{RNA} \cdot B_{RNA} \rightarrow A_{RNA}$	0.5
$k_{cleave,b}$	$B_{RNA} \cdot C_{RNA} \rightarrow B_{RNA}$	0.5
$k_{cleave,c}$	$C_{RNA} \cdot A_{RNA} \rightarrow C_{RNA}$	0.5

Fig. 6. Table 2. Approximated parameters of individual chemical reactions within the riboregulated RNA repressilator. Values represent rate constants for the reactions and are averaged in units of molecules/s for purposes of Gillespie stochastic modeling. Constants are calculated based on constants for Escherichia coli bacteria.

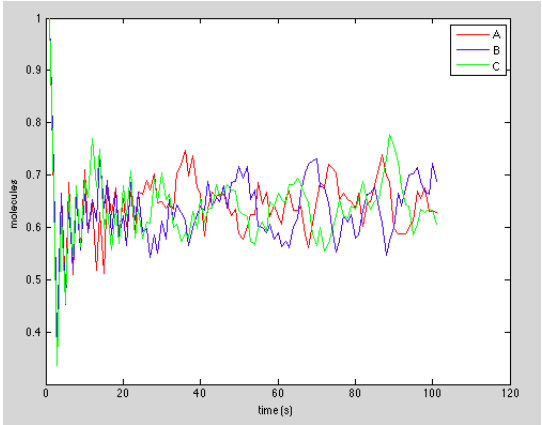

Fig. 7. Average trace over 100 seconds of RNA-based repressilator with three ribozymes from Gillespie algorithm. Graph is based on an average sample of 500 repetitions of timecourses from stochastic model. Time and molecules are both kept at low levels in order to clearly observe the oscillations in concentrations. While oscillations are observed, they are irregular and unstable, presumably due to high rates of transcription and degradation of mRNA.

Matlab using the Gillespie algorithm in order to observe how natural variation would effect the stability of the system.

For the purposes of the model, the total amount of molecules and the time of reaction were kept low (biologically-unrealistic) in order to observe the oscillations of the RNA repressilator more clearly. In addition, the amplitude and the period of the oscillations were both tunable based on the rate of binding, rate of cleaving, and rate of production of our hammerhead ribozyme constructs.

Comparing the RNA repressilator to the classic repressilator (see Figure 5,6(Table 2),7), the time for one cycle of oscillations to complete for the ribozyme circuit (on the order of seconds) was significantly shorter than that of the DNA circuit (on the order of minutes). This difference in period of oscillations is not only a key difference but also explains why the oscillations minimum value in the ribozyme circuit is nonzero compared to the DNA repressilator. As the system varies between high concentrations of species A, B, and C, on average none of the ribozymes are not completely cleaved and removed from the system before the behavior of the system shifts to the next state. As a result, we see the oscillations shown in the figure where the number of molecules of each ribozyme varies between approximately 0.5 and 0.8 molecules on average.

While RNA is still a superior system than DNA in terms of its safety, composability, and tunability, our repressilator was extremely sensitive to noise caused by high degradation and production rates. The system failed to operate as desired, producing unstable, irregular oscillations at low concentrations of molecules. As a result, the circuit failed to meet most of the design criteria for biological circuit design, and additional parameters would need to be adjusted in order obtain an effective repressilator.

Discussion

As tertiary structure does for proteins, 2D and 3D structures of RNA molecules play a major role in the function of these molecules [12]. There are many tools available the predict these higher order structures in RNA that are based on free-energy minimzation through WC base-pairing [13]. The predictability of inter- and intra-RNA interactions plays a key role in the ability to tune the relative ribozyme activity (binding interactions), as well as the relative stability of a given ribozyme (introduce slow-acting cleaving sites). With all of these characteristics of RNA in mind, we assumed that the use of a self-contained RNA regulator would be far better than the traditional use of promoter-repressor pairs in the modeling of logic circuits biologically. This initial hypothesis was further supported by the fact that RNA regulation would be on smaller timescales, which ramps up the rate of signal transduction.

One of the major challenges that we faced when modeling systems built with trans-cleaving hammerhead ribozymes is the lack of cooperativity in these systems. There were many unknowns that came with designing and characterizing a novel set of rationally designable parts. The design of a new hammerhead ribozyme that dimerized before

Instead, MichaelisMenten kinetics best described the relationship between the ribozyme and its substrate RNA (also a ribozyme in our cases). The lack of cooperativity also resulted in the lack of bistable modes in the modeled toggle switch. In the toggle switch the two ribozymes cleaved each other a_{RNA} b_{RNA} As discussed above, any

Materials and Methods

Methods.

The modeling of the ribozyme toggle switch, and the ribozyme repressilator used the same set of tools, and used the same process. The representative chemical reactions of each of these circuits designed by hand in a way that was equivalent to the original toggle switch (Fig 2.) and repressilator design (Fig 3.). Each of these chemical reactions given its own definition of likelihood based on its subtrate(s) and associated constants. These equations are then added to a Markov Chain simulator object. When run, this object simulates random behavior by selecting two random numbers. The first of which is used to portray the length of the next time step. After taking into account the relaitve concentrations of each of the species and the resulting likelihood of each of the reactions, the second random number is used to select the next reaction to take place. So on each iteration the time in the model increases, and the chain terminates upon the arrival at the endtime.

Literature search revealed the likely rates at which each of the reactions was to occur. Necessary parameters that we had to determine included: transcription rates, binding and diffusion rates of partially complimentary RNA strands, rate of cleavage of hammerhead ribozymes, degradation of before and after cleavage.

Source is available here: https://github.com/zacatac/ribozyme-circuit-modeling

Modeling Tools.

Java implementation of the Gillespie algorithm created by Anderson for his Genetic Devices course [11].

Matlab implementation of the Gillespie algorithm created by Ryan Tsoi.

ACKNOWLEDGMENTS. Thanks to Jeremy and Sergey for their feedback.

1. **Keasling J., The Promise of Synthetic Biology, The Bridge, 15 (2005),** http://www.nae.edu/Publications/Bridge/Cutting-EdgeResearchinEngineering/ThePromiseofSyntheticBiology.aspx.
2. **Madec, M.; Lallement, C.; Karstens, K.; Dittman, S.; Gersbacher, M.; Sorg, R.; Wild, M.; Muller, M.; Bourgine, P.; Donzeau, M.; Haiech, J., Synthetic biology and microelectronics: A similar design flow, Circuits and Systems and TAISA Conference, 2009. NEWCAS-TAISA '09. Joint IEEE North-East Workshop on , vol., no., pp.1,4, June 28 2009-July 1 2009 doi: 10.1109/NEWCAS.2009.5290444**
3. **Ignacio Tinoco Jr, Carlos Bustamante, How RNA folds, Journal of Molecular Biology, Volume 293, Issue 2, 22 October 1999, Pages 271-281, ISSN 0022-2836, http://dx.doi.org/10.1006/jmbi.1999.3001.**
4. **Onuchic, J.N. et al., Theory of Protein Folding: The Energy Landscape perspective, Annu. Rev. Phys. Chem. 1997.48:545-600,** http://www.embnet.qb.fcen.uba.ar/embnet/references/frustra_ref1.pdf
5. **Diethard Mattanovich, Brigitte Gasser, Hubertus Hohenblum, Michael Sauer, Stress in recombinant protein producing yeasts, Journal of Biotechnology, Volume 113, Issues 13, 30 September 2004, Pages 121-135, ISSN 0168-1656,** http://dx.doi.org/10.1016/j.jbiotec.2004.04.035.
6. **Paulo P. Amaral, Marcel E. Dinger, and John S. Mattick Non-coding RNAs in homeostasis, disease and stress responses: an evolutionary perspective Briefings in Functional Genomics (2013) 12 (3): 254-278 doi:10.1093/bfgp/elt016**
7. **Salis Lab, RBS Calculator,** http://salis.psu.edu/RBS_Calculator.shtml
8. **Newbury S.F., Control of mRNA stability in eukaryotes, Biochemical Society Transactions (2006) 34, (3034) (Printed in Great Britain) doi:10.1042/BST20060030**
9. **Gardner T.S. et al., Construction of a genetic toggle switch in E. coli, Nature (2000) 403,** http://ender.bu.edu/~tgardner/be209/lectures/17/articles/gardner.ts_nature_1.20.00.pdf
10. **Elowitz, Michael B.; Leibler, Stanislas, A synthetic oscillatory network of transcriptional regulators Nature 2000/01/20/print 403 335-338,** http://dx.doi.org/10.1038/35002125
11. **Anderson J.C., Java implementation of Gillespie, Fall 2013 Devices Course**
12. **Rebecca K. Montange and Robert T. Batey, Riboswitches: Emerging Themes in RNA Structure and Function Annual Review of Biophysics, Vol. 37: 117 -133 (Volume publication date June 2008)**
13. http://en.wikipedia.org/wiki/List_of_RNA_structure_prediction_software