

# Computational design of RNA-based oscillatory circuits

J. Binysh

\* *University of Warwick, Complexity Department*

**Abstract**—genetic circuitry, RNA's offers an attractive alternative to more traditional methods, which typically involve using proteins to regulate DNA transcription. In comparison to proteins, it is relatively straightforward

## I. INTRODUCTION

The process of gene expression can be briefly summarised as: DNA is read, and a copy of it is made, in the form of an RNA molecule (this is called *transcription*). This RNA molecule (known as messenger RNA, or mRNA) makes its way to a piece of cellular machinery called the Ribosome, which reads it, and makes a protein - which protein is made depends on the DNA sequence originally read (*translation*).

The path from genetic transcription to protein expression is naturally regulated in many ways [?]. This regulation allows the cell to control protein expression, and so cell behaviour, in response to various environmental cues. The natural cell machinery which performs this regulation offers rich possibilities for modification, and an important goal within synthetic biology is to understand and manipulate it.

As well as acting as the intermediate between DNA and protein, RNA molecules play direct and important roles in regulating cell behaviour [1]. For the synthetic biologist looking to engineer regulation of genetic circuitry, RNA's offers an attractive alternative to more traditional methods, which typically involve using proteins to regulate DNA transcription. In comparison to proteins, it is relatively straightforward to predict the structure and function of an RNA from its sequence using physiochemical models. Recently, this has been exploited to computationally design DNA sequences encoding synthetic sRNA's - small RNA's which do not code for a protein, but rather have some direct regulatory function - with regulatory behaviour that can be predicted [2] [3].

This report will focus on one such system, introduced in [3]. It will extend existing understanding of the system beyond the qualitative, by proposing a quantitative model of gene expression, in the form of a set of ODE's, and fitting it to available time series data to estimate the model's unknown parameters.

The report is structured as follows. In the remainder of this section we review the sRNA regulatory system we will consider. In section II we propose a set of ODE's to model the system, and estimate its unknown parameters by fitting to time series data. Finally, in section III, we conclude, and suggest directions for further work.

## A. The sRNA regulatory system

In bacteria, one mechanism by which gene expression is regulated is as follows [4]: In order for a bacterial mRNA to be translated into a protein, the Ribosome must initially bind to the mRNA (Fig. 2). This occurs at the Ribosome Binding Site (RBS) [5], a specific nucleotide sequence found on the mRNA. In an mRNA there is an untranslated region of nucleotides at the 5' end of the molecule (the UTR), upstream of the RBS. Translation may be self repressed by this 'tail' of the mRNA folding over and binding across the RBS, forming a stem loop in the mRNA and preventing the Ribosome from binding (Fig. 2). This self repression may be released with an sRNA which binds to the same region on the mRNA - the new conformation of the sRNA:mRNA complex uncovers the RBS, allowing the Ribosome to bind. In summary, the presence of the sRNA positively regulates gene expression.

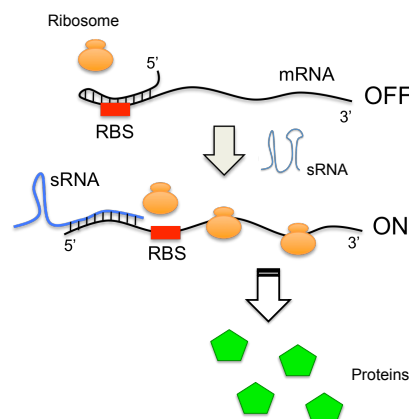


Fig. 2. A mechanism by which sRNA's can regulate gene expression. Initially, the 5' UTR of the mRNA is folded over the RBS, forming a loop and blocking Ribosome binding. The sRNA binds to this mRNA, causing a conformational change which uncovers the RBS, and allows translation to occur. Image reproduced from [3]

[3] proposed a computational methodology to design general genetic circuits based on RNA interactions, and as a case study of the methodology chose to design a synthetic sRNA- mRNA pair capable of acting in the manner described above. The algorithm assumed an interaction scheme between the RNA's as shown in Fig. 3. The two RNA's, originally in their own individually folded states, would initially interact via a small 'toehold' sequence of unpaired nucleotides to form an unstable transition state. This intermediate complex would

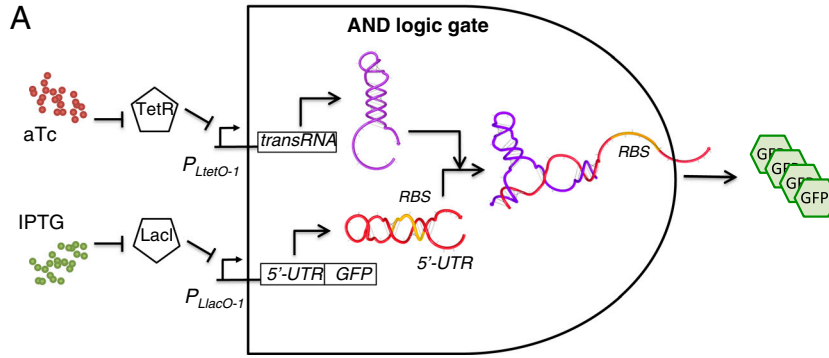


Fig. 1. A logical AND gate formed from a self repressed mRNA, and an sRNA which uncovers its RBS. In this system, transcription of the sRNA (transRNA) and mRNA (5'-UTR,GFP) are controlled by two promoter regions,  $P_{LtetO-1}$  and  $P_{LlacO-1}$ . These are disabled by the presence of two chemical repressors, TetR and LacI, found naturally in the strain of *E. coli* discussed. These chemical repressors are themselves disabled by two chemicals, aTC and IPTG. In the notation of the diagram, a barred line indicates repression, and an arrowed line indicates production. We see a 'double negative' in aTC repressing TetR, which itself represses transcription of the sRNA (likewise for IPTG and the mRNA). Thus presence of the sRNA and mRNA are controlled by the presence of aTC and IPTG, which can be experimentally introduced to the cell. Image reproduced from [3].

then rapidly form a final, stable complex with the desired conformation. By suggesting sRNA and mRNA sequences which optimised this energy landscape, [3] suggested several devices which would form a stable hybrid with the RBS free, and experimentally validated their function in *E. coli*, using an mRNA which codes for GFP for experimental ease (note the algorithm only optimises the 5' UTR of the mRNA, so the actual protein being coded for is unimportant).

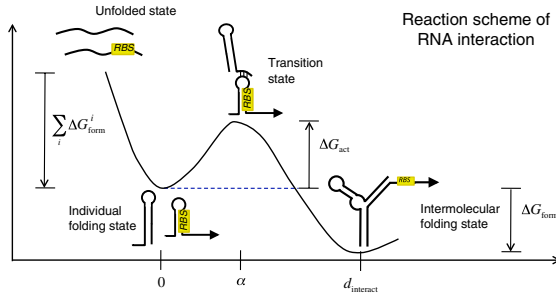


Fig. 3. A network with an intuitively clear community structure, which is captured by the partition chosen, shown in gray. Image reproduced from [3]

Further, by placing the concentrations of the sRNA and mRNA under the control of tuneable promoters, [3] constructs a logical AND gate from one of the proposed devices (RAJ11) *in vivo* (Fig. 1). In this system, transcription of the designed sRNA and mRNA are placed under the control of promoter regions,  $P_{LtetO-1}$  and  $P_{LlacO-1}$  [6]. These are in turn controlled by two transcriptional repressors, TetR and LacI, which are naturally present in the strain of *E. coli* considered. These repressors disable the promoter regions, and so by default transcription of the RNA's is turned off, and no protein is

produced. These repressors can themselves be disabled by the presence of two chemicals, aTC and IPTG, which can be introduced externally into the cell (Fig. 1). So transcription of the two RNA's is indirectly controlled by the presence of two chemicals - if neither is present, sRNA and mRNA transcription is repressed, and no protein is produced. If only one is present, the AND gate remains off, either because there is no mRNA to be translated into protein, or because the mRNA is self repressed. But when both are present, the conformational change discussed above occurs, and protein is produced.

promoter?

Although a qualitative understanding of this system exists [3], it is of interest to attempt a quantitative understanding of the genetic circuit involved. Such an understanding would allow, for example, tailoring of the system in response to design requirements, by altering the values of the important parameters of the model. By changing which sRNA-mRNA device is used in the system, it would also allow exploration of the relationship between the thermodynamic properties of each device, and the model's rate constants.

## II. RESULTS AND DISCUSSION

### A. Model Derivation

We propose a simple model, consisting of a set of ODE's with mass action kinetics [?], to describe the system. Tables I,II give complete descriptions of parameter and state variable meanings.

$$\frac{ds}{dt} = \frac{N\alpha_T}{f_T}y(t) - (\mu + \delta_s)s - k_{on}sm + k_{off}s : m \quad (1)$$

$$\frac{dm}{dt} = \frac{N\alpha_L}{f_L}x(t) - (\mu + \delta_m)m - k_{on}sm + k_{off}s : m \quad (2)$$

$$\frac{ds : m}{dt} = k_{on}sm - (k_{off} + k_{hyb})s : m - (\mu + \delta_{sm})s : m \quad (3)$$

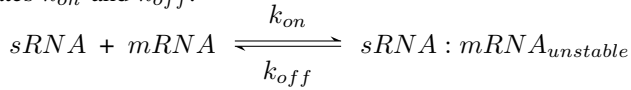
$$\frac{dc}{dt} = k_{hyb}s : m - (\mu + \delta_c)c \quad (4)$$

$$\frac{dp}{dt} = \beta m + f_s\beta c - (\gamma + \mu + \delta_g)p - \frac{v_z p}{K_z + p + g} \quad (5)$$

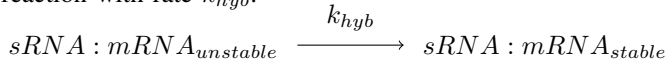
$$\frac{dg}{dt} = \gamma p - (\mu + \delta_g)g - \frac{v_z g}{K_z + p + g} \quad (6)$$

$$z = z_0 + \frac{g}{\theta} \quad (7)$$

Based on the reaction mechanism in Fig. 3, the hybridization of the sRNA and mRNA first into an unstable complex, then a stable one, is modelled in eqs.1 - 4. The initial binding is modelled as a reversible reaction with forward and backward rates  $k_{on}$  and  $k_{off}$ .



after which the stabilization is modelled as an irreversible reaction with rate  $k_{hyb}$ .



In addition, these complexes are given degradation rates,  $\delta_s$ ,  $\delta_m$ ,  $\delta_{sm}$ ,  $\delta_c$ , and dilution of chemical concentrations due to cell growth are modelled with a dilution rate  $\mu$ .

Control of the system by aTc and IPTG are modelled by  $y(t)$  and  $x(t)$  in eqs.1, 2.  $y(t)$  models the response of the sRNA transcription rate to a time varying aTc concentration - it is normalised to lie between 1 and  $f_T$ , and is typically sigmoid in response to aTc concentration [3].  $x(t)$  behaves analogously for the mRNA.

### B. Parameter Estimation

- basically the whole project. [1]

## III. CONCLUSIONS AND FURTHER WORK

### REFERENCES

- [1] F. J. Isaacs, D. J. Dwyer, and J. J. Collins, "RNA synthetic biology." *Nature biotechnology*, vol. 24, no. 5, pp. 545–554, 2006.
- [2] G. Rodrigo, T. E. Landrain, S. Shen, and A. Jaramillo, "A new frontier in synthetic biology: Automated design of small RNA devices in bacteria," pp. 529–536, 2013.
- [3] G. Rodrigo, T. E. Landrain, and A. Jaramillo, "De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells," *Proceedings of the National Academy of Sciences*, vol. 109, no. 38, pp. 15 271–15 276, 2012.

TABLE I. MODEL PARAMETERS

Parameter	Units	Definition
N		plasmid copy number
$z_0$	AFU	baseline fluorescence
$\alpha_L$	nM/min	Max translation rate of Plac01 promoter
$\alpha_T$	nM/min	Max translation rate of Ptet01 promoter
$f_L$		Max llacI fold
$f_T$		lltet fold
$\delta_g$	/min	gfp degradation rate
$\gamma$	/min	gfp maturation rate
$v_z$	nM/min	degradation constant of clpx
$K_z$	nM/min	Dissociation constant of clpx
$\theta$	nM/AFU	ratio between gfp amount and fluorensence
$\mu$	/min	dilution rate
$\delta_m$	/min	mRNA degradation rate
$\delta_s$	/min	sRNA degradation rate
$\delta_{sm}$	/min	unstable s:mRNA degradation rate
$\delta_c$	/min	stable s:m degradation rate
$k_{on}$	/min	s+m $\rightarrow$ s:m rate
$k_{off}$	/min	s+m $\leftarrow$ s:m rate
$k_{hyb}$	/min	s:m $\rightarrow$ c rate
$\beta$	/min	basal translation rate
$f_s$		sRNA repression fold

TABLE II. STATE VARIABLES

State variable	Definition
$s$	Definition
$m$	Definition
$s : m$	Definition
$c$	Definition
$p$	Definition
$g$	Definition
$z$	Definition
$y(t)$	Definition
$x(t)$	Definition

- [4] T. Soper, P. Mandin, N. Majdalani, S. Gottesman, and S. a. Woodson, "Positive regulation by small RNAs and the role of Hfq." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 21, pp. 9602–9607, 2010.
- [5] J. Shine and L. Dalgarno, "Identical 3'-terminal octanucleotide sequence in 18S ribosomal ribonucleic acid from different eukaryotes. A proposed role for this sequence in the recognition of terminator codons." *The Biochemical journal*, vol. 141, no. 3, pp. 609–615, 1974.
- [6] R. Lutz and H. Bujard, "Independent and tight regulation of transcriptional units in escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements," *Nucleic Acids Research*, vol. 25, no. 6, pp. 1203–1210, 1997.