Towards Logical Designs In Biology

Garima Agarwal

This article highlights an emerging field known as synthetic biology that envisions integrating designed circuits into living organisms in order to instruct them to make logical decisions based on the prevailing intracellular and extracellular conditions and produce a reliable behavior. The attempt is to design cells capable of sensing a bioweapon or an environmental pollutant, activate its degradation pathway and perform bioremediation or carry out programmed cell death or synthesize complex biomaterials. Within the living cell, a complex interplay of networks formed by genes and proteins mediate all cellular processes. The networks in a system can be dissected into small regulatory gene circuit modules. Synthetic biology attempts to construct and assemble such modules step by step, plug the modules together and modify them, in order to generate a desired behavior. The review gives an insight into the creation of gene circuits and discusses the potential applications in the field of biotechnology, medicine and environmental sciences.

Garima Agarwal is a research student in the Molecular Biophysics Unit, Indian Institute of Science. Her areas of research interest include structural biology, bioinformatics and systems biology.

Introduction

Logic gates are the basic building blocks in electronic circuits that perform logical operations. These have input and output signals in the form of 0's and 1's; '0' signifies the absence of signal while '1' signifies its presence. Similar to the electronic logic gates, cellular components can serve as logic gates. A typical biological circuit consists of i) a coding region, ii) its promoter, iii) RNA polymerase and the iv) regulatory proteins with their v) DNA binding elements, and vi) small signaling molecules that interact with the regulatory proteins (see *Box* 1). Messenger RNA or their translation products can serve as input and output

Keywords
Synthetic biology, logic gates, regulatory modules, gene circuits.

Box 1.

There are three steps in a simple gate:i) Translation of the input mRNA signal, ii) cooperative binding of the protein (repressor) to the DNA (operator) and iii) regulated gene expression to generate the output signal. Therefore, the output signal is influenced by the amino acid synthesis occurring in the cell, the affinities of the ribosome binding sites on the mRNA and its stability. Other factors like dimerization of the repressor proteins and the affinities to their respective operators also play a role in determining the amount of the output signal [1].

signals to the logic gates formed by genes with which these gene products interact. The concentration of the gene product determines the strength of the signal. High concentration indicates the presence of signal (=1) whereas low concentration indicates its absence (=0). The functioning of a biological circuit can be read out by the expression of a marker gene (for example, Green Fluorescent Protein (GFP) that fluoresces in response to ultraviolet irradiation) in the circuit. The marker is expressed as a consequence of the interactions between various components in the circuit.

The Need for Designing Circuits

Genetic engineers play around with genes or metabolic pathways and observe the effects on the system. This approach may not always bring the desired cellular behavior because of the complexity of the system. On the other hand, synthetic biologists try to build the system from first principles. They try to build models of the gene regulatory circuit, simulate the models, observe the behavior, make changes in the models and implement the model *in vivo* to generate the desired effect. Once the designed principles are set, it becomes easy to control the behavior. These manipulations help synthetic biologists to elucidate and understand the system much better. The following section describes how the gene regulatory components can be directed to perform some simple logic functions.

Logic Circuits or Regulatory Modules

The NOT Gate: The NOT gate is the simplest biochemical circuit. The gate has a single input signal. The NOT gate 'inverts' the input signal, hence known as the inverter gate (*Figure* 1). The enzyme RNA polymerase binds to DNA elements called promoters to carry out transcription (the conversion of the information in DNA to an intermediate, mRNA). The gate is used to determine the intracellular state of the cell [1, 2].

The NAND Gate: The NAND gate consists of two NOT gates connected in a manner such that the gates have different input

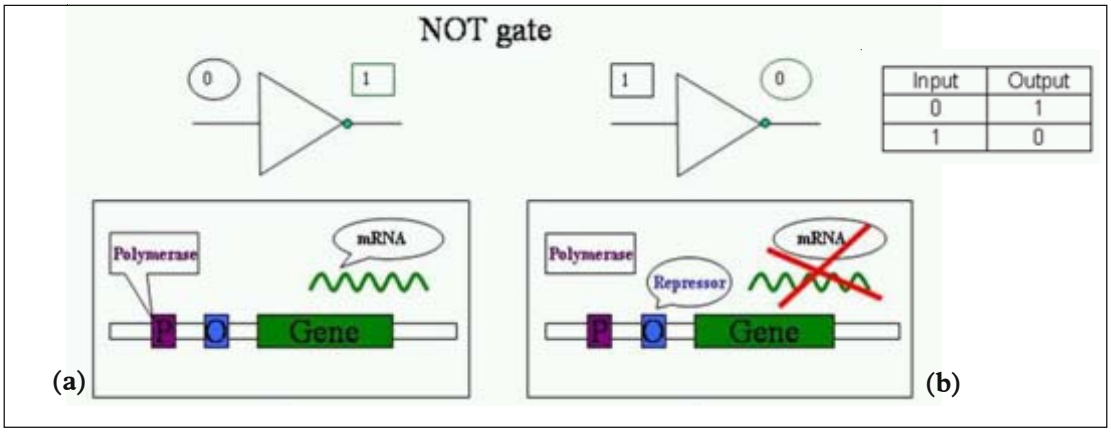


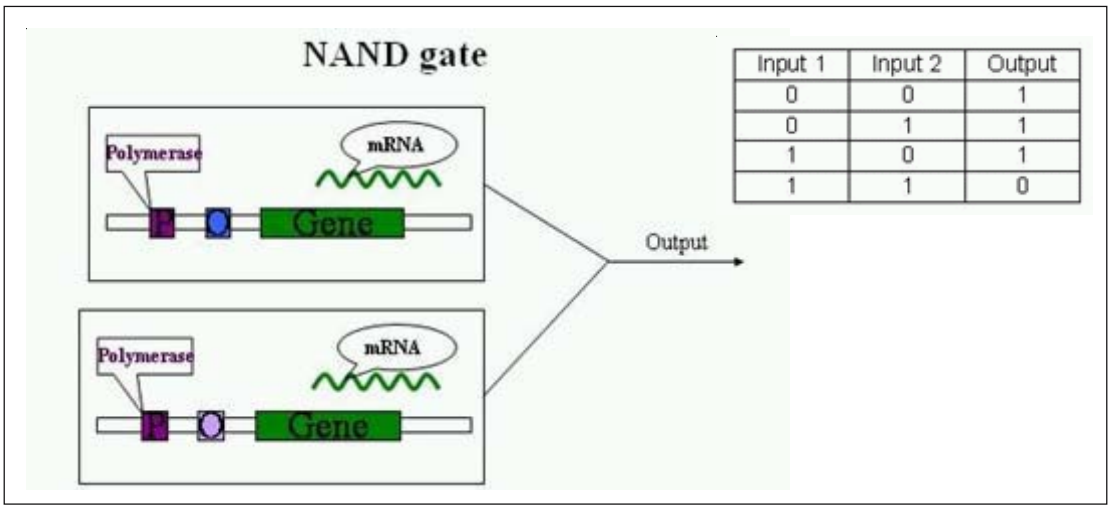
Figure 1. NOT gate. '0' implies absence of repressor protein. As a result polymerase binds to promoter and causes gene expression. (a) and (b) depict the possible situations of the presence and absence, respectively, of repressor.

signals but the same output signal. Only when both the input signals are present, the transcription of the gene is turned off (*Figure* 2).

The AND Gate: The AND gate has two input signals and only when both the signals are present, an output signal is generated. The polymerase intrinsically has low affinity for promoters, hence there is no transcription. The activator and inducer together result in turning on a gene (*Figure* 3). The gate can be used for cell-cell communication [1, 3].

The Implies Gate: The Implies gate requires two input signals that can be a repressor protein and an inducer molecule. If both

Figure 2. NAND gate. The output of the gate depends on the output generated from the two NOT gates, as depicted in the truth table.



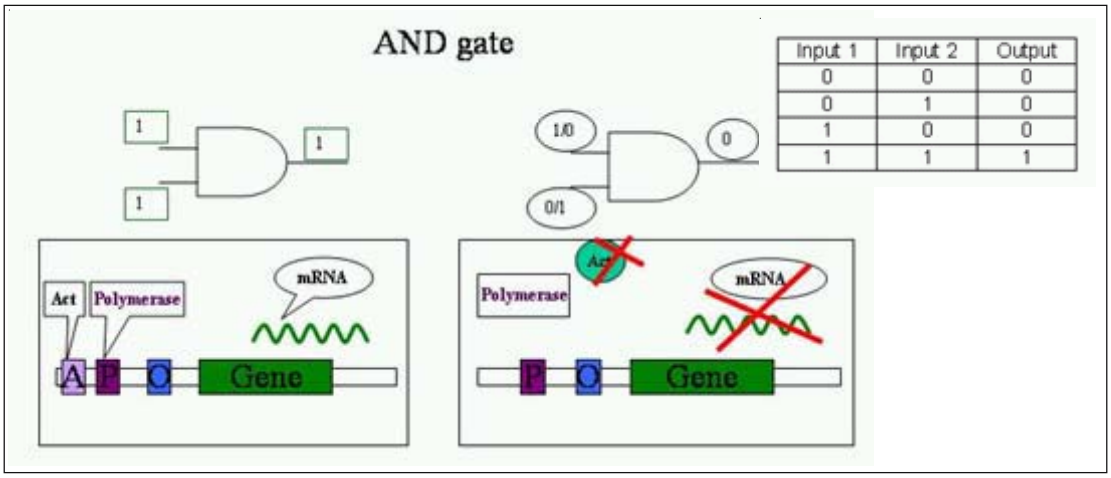
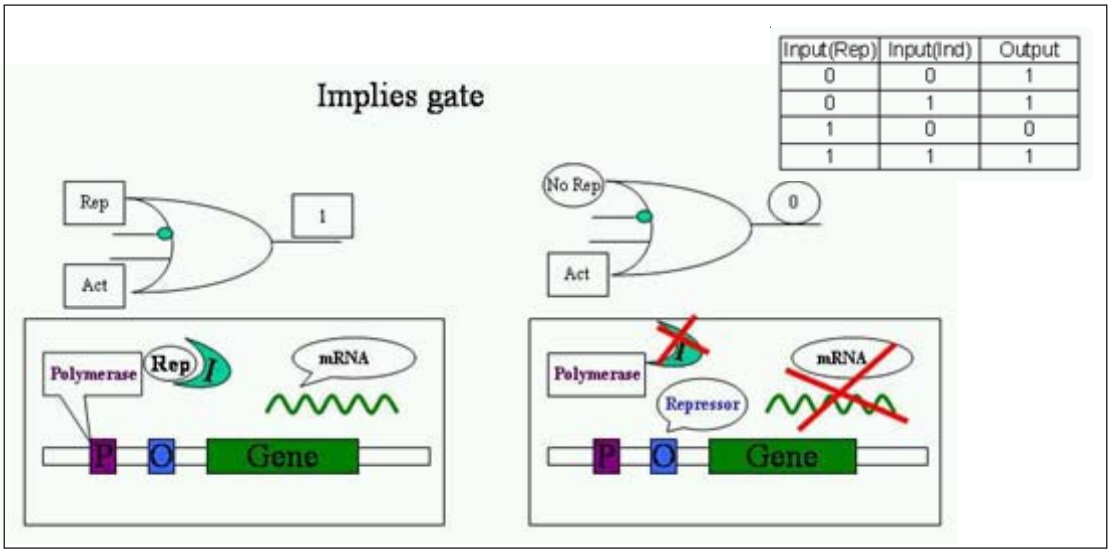


Figure 3. AND gate. It requires two input signals, namely, polymerase and activator. The presence of both proteins is required to generate an output.

the input signals are absent, the gene is transcribed. If only repressor is present, the circuit behaves like a NOT gate. If both the input signals are present, the output signal is generated (*Figure* 4). Since the inducers are freely diffusible molecules, the gate can be used to sense external conditions by the cell [1, 3].

***The Toggle Switch*:** The Toggle Switch circuit comprises two stable states and can be induced by an external signal to switch from one state to another [4]. Once induced, the system remains in that state until another signal flips the switch. The circuit consists of two genes coding for two repressors. The two repres-

Figure 4. The truth table of Implies gate depicts the possible input signal combinations and the output signals.



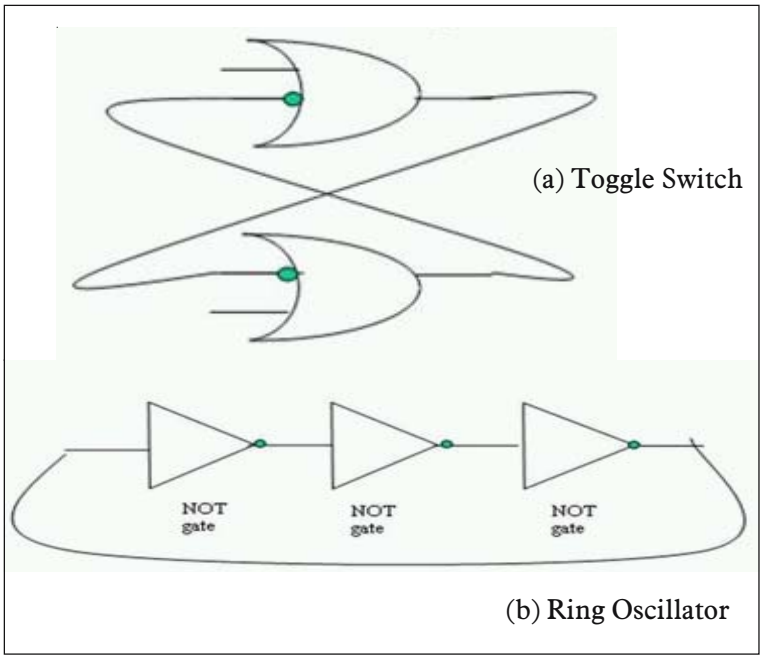


Figure 5.
(a) A Toggle Switch where two Implies gates are connected such that the output of one represses the other. The Switch gets flipped in the presence of an inducer.
(b) An oscillator where the final output signal represses the first NOT gate. This causes the circuit to switch between output and no output states.

sors mutually repress each other such that a high concentration of one protein inhibits the transcription of the other gene. Any signal that causes the breakdown of the existing repressor protein molecules or increases the transcription of the other repressor protein would cause the flip (*Figure* 5).

***The Ring Oscillator*:** The Ring Oscillator is constructed by integrating an odd number of NOT gates in a circular fashion such that the output of the last gate is the input of the first one (*Figure* 5). The circuit is designed so as to make the last output signal, the logical NOT of the first signal. This is possible through the integration of odd number of inverter gates. The importance of an odd number of gates can be understood from the following example. Consider a high output signal generated from the last inverter gate, which is fed to the first inverter gate. If the number of inverters is even, the output generated from the last gate will be high. However, if the number is odd, the output from the last inverter gate will be low and when fed to the first gate generates a high output from the last inverter. Hence the last output signal oscillates between low and high alternately [5].The circuit has been used in generating blinking bacterial cells.

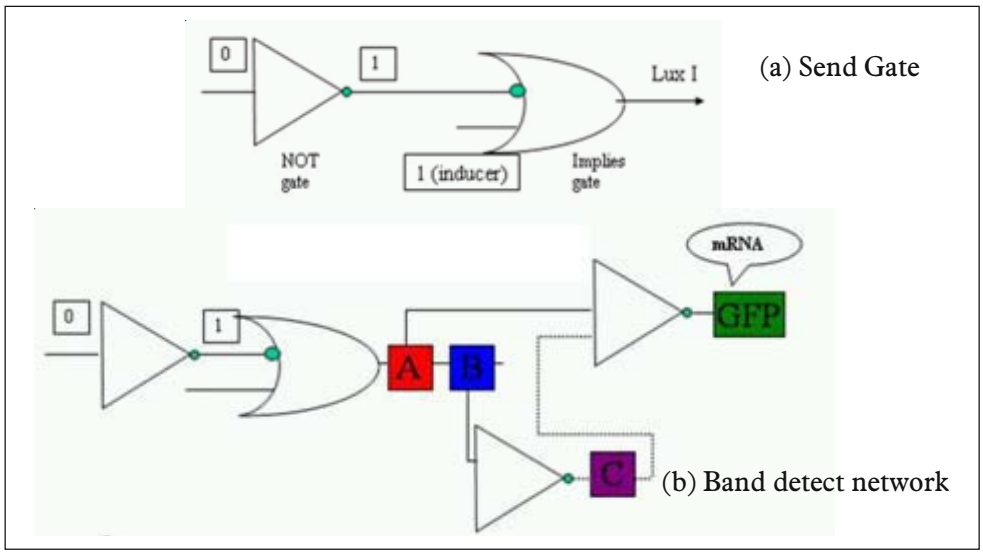
Box 2.

Quorum sensing is a cell density–dependent intracellular communication system used by many microorganisms. *Vibrio fischeri* uses it to turn on the genes responsible for luminescence [6]. The genes involved are organized into two divergently transcribed operons. The *luxR* gene that encodes a transcriptional activator is transcribed to the left and the gene *luxI* and other genes involved in luminescence are transcribed to the right. At low cell densities, basal levels of LuxI in the cell catalyze the synthesis of the auto-inducer Homoserine Lactone (HSL) which is secreted. As cell density increases, HSL builds up in the extracellular medium and can diffuse back into the cell. Once a threshold concentration is reached, it binds to LuxR and activates it. This results in enhanced transcription of the right operon, leading to the synthesis of more LuxI (and hence more HSL) and the gene products needed for luminescence.

Apart from the development in single cell networks, multicellular networks have also been engineered. These multicellular networks use cell-cell communication to achieve a coordinated behavior. Cells can talk among themselves through secretion of small molecules that can pass from one cell to another. To mediate cell-cell communication, the elements of quorum sensing, a diffusion controlled mechanism used by luminescent micro-organisms, have been implemented in designing the circuits [1, 7] (see *Box* 2). The elements of the sensing mechanism are partitioned in two circuits, the SEND gate and Band detect network.

The SEND Gate: The SEND gate takes the mRNA coding for a catalytic enzyme LuxI as the input signal (*Figure* 6a). The enzyme uses the metabolic precursors of the cell to synthesize inducer molecules. The output signal is the synthesized inducer [1, 3].

Band Detect Network: Its characteristic feature is the generation of an output signal only when the concentration of the input signal is in a specific range. The output signal of SEND gate is fed into it. The network consists of three sub-circuits: a low threshold component, a high threshold component and a NOT logic function. The cells constitutively express LuxR, an activator protein. When the inducer is present, the transcription of



genes A and B gets activated (*Figure* 6b). A is repressor of the reporter GFP. B is repressor for another gene C. The C protein is a repressor of GFP. The circuit through A acts as low detect component and that through B, a high detect component. The two components are connected to NOT gates to generate an output signal as GFP expression (*Figure* 6b). The differing repression capabilities of the repressors are the determinants of low and high threshold levels of the input signal detected [1, 3].

Figure 6.
(a) The SEND gate circuit used for the synthesis of LuxI.
(b) The circuit of band detect network, generates output signal for a certain range of concentration of the input signal.

Designing Circuits

The original components of gene regulatory machinery are well optimized through evolution to perform a task efficiently, for example the LexA repressor or the RecA protein that are involved in SOS regulation (see *Box* 3). The arbitrary assemblage

Box 3.

The SOS regulon is a tightly regulated repair system that gets activated upon DNA damage of the bacterial genome caused by agents such as ultraviolet irradiation. In the inactive state of the system, the LexA repressor binds to the operators of the SOS genes including the gene *recA*, repressing their transcription. Upon damage to the genome, the resulting single stranded DNA fragments activate the basal level of RecA and convert it into a protease. The activated RecA protease cleaves the LexA repressor, thus turning on the repair genes and *recA*.

of components from various sources need not generate the desired output and efficiency. One of the approaches towards this end is rational design where the targeting of gene mutations to generate a functional circuit is guided by models and simulations [3]. Modeling requires accurate knowledge of rate constants. Each of the steps in a circuit, such as translation of the input mRNA and cooperative interactions of molecules, is associated with a kinetic rate constant. Rate equations for each of these steps are used to obtain differential equations. The set of differential equations can then be used to simulate the model circuit. The dynamic behavior of the cells is observed through simulations. Mutations can then be generated *in vivo* in the genetic components of the circuit based on the simulation results. The mutations may lie (i) in the repressor or operator region to alter their affinities, (ii) in the promoters altering its strength or (iii) in the ribosome binding sites [2, 3]. Another approach used is directed evolution where mutations are randomly made and cells with desired behavior are selected and the mutation sites identified by sequencing. The above two approaches have been used in improving the components of the circuit.

The important point to be considered while assembling the gates to generate more complex machinery is the matching of kinetic characteristics of various components. For example, consider that two inverters are connected such that the output of X gate is the input of Y gate. If the low output of X gate is still high enough to repress the output of Y gate, the gate will not generate an output and will thus be non-functional [8]. It is important to note that the signals generated from the circuits do not interfere with the normal functioning of the cell.

Applications

Achievements have been made recently in directing organisms to perform novel tasks. Adenoviruses have been engineered to detect the levels of p53 protein, a tumor suppressor. The protein is present in low amounts in a tumor cell. If the protein concen-

tration is low, the circuit is designed so as to derepress the replication cycle of the virus. The activation of lytic cycle of the virus causes the lysis of tumor cells [9].

The SEND gate and Band detect network were integrated in two different strains of *E.coli*. The cells have been engineered to detect a range of concentrations of the inducer molecules and produce fluorescence in that range. The discs of sender cells containing the SEND gate were placed in the center of an agar plate and the receiver cells containing the band detect network were plated on agar. The inducer molecules diffuse into the medium. The receiver cells could sense the concentration and produce concentric rings of fluorescence [7]. This has implications in understanding development and differentiation where concentration gradients of chemicals known as morphogens guide pattern formation.

A toggle switch, with LacI and cI as the repressors, has been incorporated to create organisms that glow in response to DNA damage [10]. The SOS pathway gets activated in response to DNA damage (see *Box* 3). The activated RecA cleaves cI protein that is now unable to repress the synthesis of LacI and GFP. As a result, the organism glows.

Conclusions

Synthetic biology thus offers the ability to understand complex living systems and produce reliable behaviour of an organism through the generation of simple genetic circuit modules. One of the major challenges is to design large and complex networks, consisting of diverse modules that can carry out a plethora of functions. One critical rate limiting step is the incorporation of designed genetic circuits in the organism. Despite all odds, research is on to make the dream come true.

Acknowledgements

The author would like to thank Prof. Dipankar Chatterji and Prof. N Srinivasan for valuable suggestions and encouragement.

Suggested Reading

[1] R Weiss, *Cellular Computation and Communications using Engineered Genetic Regulatory Networks*, PhD thesis, MIT, Cambridge, MA, 2001.

[2] R Weiss, S Basu, *Device Physics of Cellular Logic Gates*, First workshop on non-silicon computing, Boston, MA, 2002.

[3] R Weiss, S Basu, S Hooshangi, A Kalmbach, D Karig, R Mehreja and I Netravali, Genetic circuit building blocks for cellular computation, communications, and signal processing, *Natural Computing,* pp.47–84, 2003.

[4] T S Gardner, CR Cantor and J J Collins, Construction of genetic toggle switch in Escherischia coli, *Nature*, Vol.403, pp. 339–342, 2000.

[5] MB Elowitz, S Leibler, A synthetic oscillatory network of transcription regulators, *Nature*, Vol.403, pp. 335–338, 2000.

[6] B L Bassler, How bacteria talk to each other: regulation of gene expression by quorum sensing, *Current Opinion in Microbiology*, Vol.2, pp. 582–587, 1999.

[7] S Basu, Y Gerchman, C H Collins, F H Arnold and R Weiss, A synthetic multicellular system for programmed pattern formation, *Nature*, Vol.434, pp. 1130–1134, 2005.

[8] X Feng, S Hooshangi, D Chen,G Li, R Weiss and H Rabitz, Optimizing Genetic Circuits by Global Sensitivity Analysis, *Biophysics Journal*, Vol.87, pp. 2195–2202, 2004.

[9] M Ramachandra, A Rahman, A Zou, M Vaillancourt, J A Howe, D Antelman, B Sugarman, GW Demers, H Engler, D Johnson and P Shabram, Re-engineering adenovirus regulatory pathways to enhance oncolytic specificity and efficacy, *Nature Biotechnology*, Vol.19, pp.1035–1041, 2001.

[10] H Kobayashi, M Kaern, M Araki, K Chung, TS Gardner,CR Cantor, J J Collins, Programmable cells: Interfacing natural and engineered gene networks, *PNAS*,Vol.101, pp.8414–8419, 2004.

Address for Correspondence
Garima Agarwal
Molecular Biophysics Unit
Indian Institute of Science
Bangalore 560012, India
Email:
garima@mbu.iisc.ernet.in