Text

Olivia Fernflores

October 16, 2024

1 Introduction

Transcription regulation is of key importance biological systems because at the most basic level, it controls gene expression. By having a mechanism to control gene expression, organisms can differentially express genes under specific conditions such as receipt of an external signal, being in a particular stage of development, or being exposed to a pathogen. Because of its applications across most things that an organism would ever need to do, being able to differentially express genes is essential for survival and is a highly conserved feature. One common way of modeling transcription regulation is through networks, where each node is a gene or protein and each edge is the relationship between two nodes (activation or repression). Once you have a few unique nodes, there are many possibilities for how the nodes will be connected, but the possibilities are limited. Because there are a finite number of ways the nodes can be connected, we can think of each possibility as a motif. This allows us to recognize when the same pattern of node connections arises more than once despite the context being different.

In biological systems, we would expect the simplest network motif to be conserved unless there is a biological advantage to a more complicated network. Milo et al. explored this hypothesis and found that some transcription regulation networks, such as feed forward loops, are observed about 10 standard deviations more in biological systems than we would expect if this network motif occurred due to chance alone [?]. In this project, we seek to understand why these feed forward loops are biologically advantageous (as indicated by their high prevalence) by modeling two types of feed forward loops and comparing to simpler networks of direct and indirect transcription regulation.

2 Model Development

2.1 Direct Regulation of Transcription

To model transcription regulation, we started with a simple system of direct regulation. Under a model of direct regulation, a signal turns on a single transcription factor, *TF1*, and that transcription figure binds to a gene, *Gene i* with a binding constant *Kd*. When *TF1* is bound, *Gene i* is transcribed at a rate *k1*, producing *RNAi*. *RNAi* is also degraded, which happens at a rate *k2*. This is shown in Figure 1A. Together, these parameters can describe the rate at which the $[RNAi]$ is changing over time with the following equation:

$$\frac{d[RNAi]}{dt} = k_1 * f_b - k_2 * [RNAi]$$

This equation quantifies the change in $[RNAi]$ over time using the *k1* and *k2* parameters previously mentioned and includes an additional factor, *fb*, that describes what fraction of transcription factor is bound to our gene of interest. This is calculated as 1 minus the fraction unbound, which is a classic result in biochemistry called the *Hill Equation* and is written like this [?]:

$$f_b = \left(1 - \frac{1}{1 + \left(\frac{TF_1}{Kd_1} \right)} \right)$$

If we combine the two equations above, we can build our full mathematical model for direct regulation.

$$\frac{d[RNAi]}{dt} = k_1 * \left(1 - \frac{1}{1 + \left(\frac{TF_1}{Kd_1} \right)} \right) - k_2 * [RNAi]$$

3 Results

Because of their prevalence in biological systems, we hypothesized that the "and gate" and "or gate" models for regulating transcription would have unique properties that are not possible to achieve with a simple direct or indirect regulation. To test this hypothesis, we continuously solved for $\frac{d[RNAi]}{dt}$ using our system of coupled differential equations at very small time intervals under a wide range of parameter value for *k1, k2, k3, k4, k5, k6,* and *Kd1, Kd2, Kd11.* We find that the "and gate" model provides unique capabilities to have a delay when initiating production of *RNAi* and a quick stop in the production of *RNAi.* We also find that the "or gate" model behaves in the opposite manner, showing a quick initiation of production of *RNAi* and a delay in stopping production of *RNAi.* For comparison purposes, all results described here were done with all *k* and *Kd* values equal to one, although our findings appear to be robust to variation in these parameter values.

To examine the behavior when initiating *RNAi* production, we allowed all four systems to reach their steady state (achieved when the derivative is equal to zero and indicates that the concentration has reached its maximum value) and then measured the time it takes to reach half of the steady state amount and the full steady state amount (Table 1).

Table 1: Time to Half and Full Steady States

Regulation Type	Time to Half Steady State	Time to Steady State
Direct Regulation	0.9091	22.7273
Indirect Regulation	2.4242	27.8788
And Gate	2.4242	27.2727
Or Gate	1.2121	27.2727

As shown by the values in the table and visualized in Figure 2, the "and gate" model behaves most similarly to the indirect model in terms of time to half and full steady state, which both show an initial delay to reach steady state that isn't captured in the "or gate" and direct models. Contrastingly, the "or gate" model initiates production very quickly, almost mimicking the direct regulation model, which can also be seen in Figure 2.

When we just examine the behavior of initiating $RNAi$ production, it seems like the "and gate" and "or gate" behave very similarly to the indirect and direct models, respectively. This prompted us to examine the behavior of all four models when $TF1$ is "turned off", which we mathematically represent from changing the binary value of $TF1$ from 1 ("on state") to 0 ("off state") (Table 2).

Table 2: Time to Return to [RNAi] = 0

Regulation Type	Time to Return
Direct Regulation	23.33
Indirect Regulation	28.18
And Gate	26.36
Or Gate	28.79

When $TF1$ is set to zero after all four models have had a chance to reach their steady state concentrations of $RNAi$, we see that the "and gate" turns off earlier than the indirect circuit while the "or gate" turns off later than the direct circuit. These results are opposite of the results for turning on $TF1$, revealing the unique behavior of the "and gate" and "or gate" feed forward loops. The unique property of the "and gate" is to initiate production with a delay and rapidly stop production. The "or gate" behaves oppositely, as it can delay stopping production while rapidly initiating production. The behavior after "turning off" $TF1$ is also shown in Figure 3.

4   Discussion

In this project, we have determined that both the "and gate" and "or gate" feed forward loop models of regulating transcription have unique properties that are not captured in a simple direct or indirect regulation system. Although you can adjust parameters such as $k2$, $k4$, and $k6$ to increase the rate of degradation or production in the direct or indirect models, you cannot adjust the parameters such that only the production rate increases without the degradation rate increasing or vice versa. For example, doubling $k2$ will increase the rate of production and increase the rate of degradation. Similarly, doubling $k1$ will affect how the concentration of $RNAi$ at steady state, but will not change the time to reach steady state because this is determined by the degradation rates. In a feed forward loop like the "and gate" or "or gate", you can have a quick switch to production or degradation and the and the slow response to the opposite, which is not mathematically possible under the direct or indirect model.

The "and gate" model shows a delay in initiating RNA production and a rapid decline in RNA production when signal is lost. Biologically, this could be relevant in stress response. The delay we see with the "and gate" happens because it takes time to build up enough of both transcription factors so they can simultaneously bind to the gene and initiate transcription. Because of the time delay, any noise in signaling will not be enough to initiate full levels transcription but an immediate stop in signaling will abruptly stop transcription. In a stress response, you need a way to filter out noise. If the cell responds to every small amount of noise and immediately activates all stress response pathways, the cell will frequently engage in stress response behavior despite not being under conditions of stress. With the "and gate" feed forward loop, this dysfunction isn't possible due to the time delay. So far, this property could be achieved by the indirect model but not the direct model. Where the "and gate" model stands out is that it can also rapidly stop RNA production, which the indirect model cannot do under any parameter combinations. If the RNA this process regulates is part of a stress response pathway, we need production to stop as soon after the stress signal stops as possible so that the cell isn't behaving in the stress condition for longer than necessary. This is possible with the "and gate" and also the direct regulation model, but not the indirect model. Thus, the "and gate" model provides unique functionality that could be useful in conditions like stress response.

The "or gate" model showed a rapid initiation of RNA production and delay in declining RNA production. Biologically, this could be relevant in an immune response. During an immune response, you want a cell to rapidly activate genes that will help with immune function and once the signal that starts the immune cascade is turned off, a delay will help make sure the cell can still respond to any remaining pathogenic activity. The "or gate" model works perfectly for this, as it allows a very quick initiation of RNA production, which is also seen in the direct regulation model but not the indirect regulation model. What sets the "or gate" apart is that when signal is lost, the "or gate" shows a delay that's only seen at the same time point as the indirect model. This means that the "or gate" is the only model that would allow a cell to respond rapidly to a signal to activate an immune response

and also delay turning off the immune pathway when signal is lost, increasing the cell's immune function at initial exposure to a pathogen and as the pathogen clears.