A Kinetic Model Based On Saturation for Biological Circuit Speed Control

William and Mary IGEM 2017

November 2, 2017

1 Summary

One of the main goals of synthetic biology is to create a modular genetic basis for the independent control of circuit behavior properties. Much progress has been made in achieving this aim for properties like gene expression strength (where well-characterized ribosome binding sites (RBSs) can be swapped within a genetic part), circuit architecture (where promoters can be swapped out to introduce connections and feedback architectures), and even gene expression noise (through a combination of the above two modulations).

However, one property of gene expression has been less amenable to this approach: the speed or rate of gene expression. Current control strategies require either a rewiring of the circuit architecture to achieve different time-dependent dynamics [1, 2] or a complete circumvention of transcriptional circuitry altogether, relying on post-translational dynamics like phosphorylation [3] or protein-protein interactions [4] to rush information through a circuit. These approaches are often inaccessible to iGEM teams because they require too drastic an overhaul of existing circuit implementations.

In this paper, we will begin by first introducing protein degradation tags, and then the concept of saturation. We then show that by accounting for saturation, the results of our single-gene characterization should be directly applicable to the circuit context. For proof of evidence, we will show some larger-scale circuits case studies in the future sections to show the validity of our model when the circuits are operating under unsaturated regime. In such case, we are able provide a model to elegantly control gene expression strength using modular, 'plug-and-play' methods requiring little modification to the protein of interest.

1.1 The Protease Saturation Model

1.1.1 Introduction

In the past months, we have been able to perform a full characterization in the context of a single gene's expression, but our intention is to create a model that can be used on a circuit-level context. To adapt to such a situation, we introduce the idea of protease saturation: **Protease saturation is the state where additional proteins create a load on protease.** From our findings, we are able to note that protease saturation effects activate through an ultra-sensitive transition from unsaturated to maximally saturated, regardless of the kinetic parameter used to cross the saturation threshold. These is mainly because protease has reached its maximal rate of degrading tagged proteins. Since effective degradation rate decreases, the magnitude of the speed change will decrease also. Since there is possibility for the tagged proteins to exceed the maximal effective degradation amount for a given amount of protease, when there exist multiple components tagged with pdts, a saturation effect should exist.

To further ensure that our degradation and speed measurements are translatable in large circuit context, we developed the protease saturation model which is further described in the following section.

1.1.2 The Protease Saturation Model

In this section, we provide a model that is based on literature[5] to describe how the quantity of tagged proteins varies with time. Instead of using it to analyze the steady-state concentration of proteins, our focus on speed-to-steady-state distinguishes our work from that of McBride and DelVeccio.

The following pages summarizes the model.

- Our protein x is produced at a constant rate and diluted with a dilution constant δ:

$$\emptyset \underset{\delta}{\overset{H_i}{\rightleftharpoons}} x_i$$

- And similarly, the protease is produced and diluted:

$$\emptyset \underset{\delta}{\overset{a_0}{\rightleftharpoons}} P_{\text{free}}$$

- The protease and the protein associate to form a complex, C_i. This reaction is reversible, and the complex may dissociate to return both its components.

$$x_i + P_{\text{free}} \underset{d_i}{\overset{a_i}{\rightleftharpoons}} C_i$$

- And the complex-ed protease can also become free by consuming its substrate.

$$C_i \xrightarrow{k} \emptyset + P_{\text{free}}$$

- Finally, the complex is also subjected to the same dilution experienced by x and P

$$C_i \xrightarrow{\delta} \emptyset$$

Each of these reaction follow mass-action kinetics, with rate constants written near the arrow for each reaction.

From the reactions given above, our base model paper derives the following equations, numbered as eq. set (3) in McBride's paper:

$$\begin{aligned}
\dot{x_i} &= H_i - a_i x_i P_{\text{free}} + d_i C_i - \delta x_i \\
\dot{C_i} &= + a_i x_i P_{\text{free}} - d_i C_i - \delta C_i - k C_i \\
\dot{P_{\text{free}}} &= a_0 - \sum_i a_i x_i P_{\text{free}} + \sum_i d_i C_i - \delta P_{\text{free}} + \sum_i k C_i
\end{aligned}$$

Which can be reduced to this one equation $f(x)$[5](Derivation available in the original paper) :

$$\dot{x_i} = \left(H_i(x) - \frac{P(k_i + \delta)\frac{x_i}{K_i}}{1 + \sum_{k=1}^{n} \frac{x_k}{K_k}} - \delta x_i \right) \Bigg/ \left(1 + \frac{P/K_i(1 + \sum_{i \neq j}(x_j/K_j))}{(1 + \sum x_k/K_k)^2} \right) \tag{1}$$

By assuming:

- The total protease concentration, free plus bound, is at equilibrium; i.e. the $P = P_{\text{free}} + \sum_i C_i$ is at equilibrium. It is easy to see that $\dot{P}_{\text{free}} = a_0 - \delta P_{\text{free}}$ by computing the sum over equations above, giving a steady-state P of $\frac{a_0}{\delta}$.
- The reactions $x_i + P_{\text{free}} \rightleftharpoons C_i$ are in rapid equilibrium; i.e. the free and bound species equilibrate fast enough that we can say $a P_{\text{free}} x_i = d C_i$ at all times. Mathematically, $a, d \gg \delta$. This is a reasonable assumption biologically.

The numerator in the overall fraction represents the model's main dynamics, with three terms: production, a nonlinear removal term representing degradation by protease, and a linear removal term representing dilution. The denumerator 'tones down' any change in the protein concentration and accounts for the popluation of complexed protein buffering the free-protein concentration. It has no relevance to steady-state analyses but is very important for finding speed-to-steady-state, since it is always greater than 1 and multiplied onto the whole equation.

In our model, $-a x_i P_{free}$ explicitly accounts for the protease complexing behavior. In such case, x_i is the amount of load-proteins, and $-a x_i P_{free}$ is the amount of free proteases that are sequestered.

The parameters are listed in the following table:

Parameter	Meaning
H	production rate
a	association rate
d	dissociation rate
$K = \frac{d}{a}$	Michealis-Menten dissociation constant
δ	dilution rate
k	degradation rate
P	Total protease concentration

This concludes the summary for the McBride-DelVeccio model.

1.2 Goal: the speed measure and how to minimize it

Again, our overarching project goal is to tunably rush the flow of information in genetic circuits. To formalize this concept,we define a variable called τ. For any protein, τ measures the time it takes for protein, starting with an initial concentration of 0, to reach half of steady-state concentration. In this document, we are only concerned with measuring τ for proteins whose concentrations start at zero, so τ is equal to the time taken for the concentration to rise from 0 to half the steady state concentration.

If we could adjust any parameter we want as much as we want, lowering τ would be extremely easy: increased speed to steady state merely requires high proportionate degradation rates.

Consider an illustration of this in a grossly simplified model for protein degradation, where the amount of protein degraded per unit time is simply proportionate to the the total amount. Here, H is the protein production rate and k is the proportionate degradation rate.

$$\frac{dx}{dt} = H - k \cdot x$$

From the solution of the following equation, we can verify that increasing k would cause a proportionate decrease in τ:

$$\frac{H}{k} - x = (\frac{H}{k} - x_0)e^{-kt}$$

Here, x tends towards its steady state $\frac{H}{k}$ with an exponential rate of k, and we have $\tau = \frac{\ln(2)}{k}$. To decrease τ, simply increase k. The catch here is that increasing k would also decrease the steady state $\bar{x} = \frac{H}{k}$. But in this simplified model the drop is easily compensated by increasing the H value, as evident from the expression of the steady state.

Note that the proportionate-degradation model does not account for the saturation effect in enzyme-catalyzed degradation. In the simplified model, the term $-k \cdot x$ grows without bounds as x grows, always proportionate to x. When x is close to 0, the enzyme-catalyzed complex-forming model has a similar behavior. However, as x becomes large, the term $-\frac{P(k+\delta)}{K+x}x$ becomes less than proportionate to x. No matter what x becomes, it cannot exceed $V_{\max}$. This is because the protease gets saturated and becomes a limiting factor in the degradation reaction. As we shall see, saturation would become a central theme to our investigation around this model, preventing us from simply using potent proteases and high production rates to minimize τ.

Thus, our goal is to minimize τ while maintaining the steady state at a constant or near-constant level, operating under metabolic constraint and the limit on degradation rate imposed by protease saturation. We have three parameters which we can vary experimentally: we can change the protein production rate H and the total protease concentration P by switching in different promoters and/or RBSs, and the inverse affinity ratio K can be altered by switching protein degradation tags. The remaining parameters, k and δ, are not considered for manipulation: the literature suggests little variability in these parameters, and they are tied into the intricacies of peptide chain cleaving and cell division respectively, and wouldn't be easy to manipulate.

2 Model Characterization: Methods

In this section, we explain the methods that we used in analyzing our model.

2.1 Analytical Solving

By doing analytical solving, we are able to provide a method to find the speed to half steady-state by algebraic integration. As it is mentioned in the previous section, τ is equal to the time takes for the concentration to travel from 0 to half the steady state level, $\bar{x}/2$. In the single-protein case, an analytical solution can be obtained for τ.

Since we have $\dot{x} = f(x)$, we can solve for $\dot{x} = 0$ to find all the steady states x may tend to. The steadystates are solved with the assistance of the *Mathematica* computer algebra system:

$$\bar{x}_1 = -K$$
$$\bar{x}_2 = \frac{\sqrt{(H - kP - \delta K - \delta P)^2 + 4\delta HK} + H - kP - \delta K - \delta P}{2\delta}$$
$$\bar{x}_3 = \frac{-\sqrt{(H - kP - \delta K - \delta P)^2 + 4\delta HK} + H - kP - \delta K - \delta P}{2\delta}$$

Only $\bar{x}_2$ has biological relevance (is positive and real) when all the parameters are positive and real. The sign of $\bar{x}_1$ warrants immediate elimination. Pulling out $\Delta = H - kP - \delta K - \delta P$, we also have

$$\bar{x}_{2,3} = \frac{1}{2\delta}\left(\Delta \pm \sqrt{\Delta^2 + 4\delta HK}\right) \tag{2}$$

. In this form, we can see that as long as $4\delta HK$ is positive, $x_2 > 0 > x_3$ no matter the sign of Δ. Under biological conditions where all parameters are positive and real, only $\bar{x} = \frac{1}{2\delta}\left(\Delta + \sqrt{\Delta^2 + 4\delta HK}\right)$ is relevant.

Thus, we now know that

$$\bar{x} = \bar{x}_2 = \frac{\sqrt{(H - kP - \delta K - \delta P)^2 + 4\delta HK} + H - kP - \delta K - \delta P}{2\delta}$$

.

To make sure that the equation is correct, we consider the impact of different parameters on the steady state we just solved. We expect higher production rates and higher inverse-affinity K to result in more protein at steady state, and this is indeed the case: for positive, real parameters,

$$\frac{\partial \bar{x}}{\partial H}, \frac{\partial \bar{x}}{\partial K} > 0$$

, which means that we expect the steady state to always increase as we increase K and H. Also, since

$$\frac{\partial \bar{x}}{\partial P} < 0$$

for all valid parameters, increasing the protease concentration must always lower the protein's steady state concentration, as expected.

Now, with $\bar{x}$ in hand, we can derive a formula for τ in the following manner: Consider the reciprocal of $\dot{x}$, which is $\frac{dt}{dx}$. At any given level of x, it stands for the time it takes for the concentration to pass through an infinitesimal quantity of x near that lever of x. Thus, if we integrate it over x between 0 and $\bar{x}/2$, we should have

$$\tau = \int_0^{\bar{x}/2} \frac{dt}{dx} \cdot dx$$

.

Using symbolic integration of the indefinite integral in *Mathematica* gives us this expression for τ:

$$\tau = -\frac{\Delta + \Psi}{2(\delta + k)}\Bigg|_{x=0}^{x=x_{ss}/2}$$

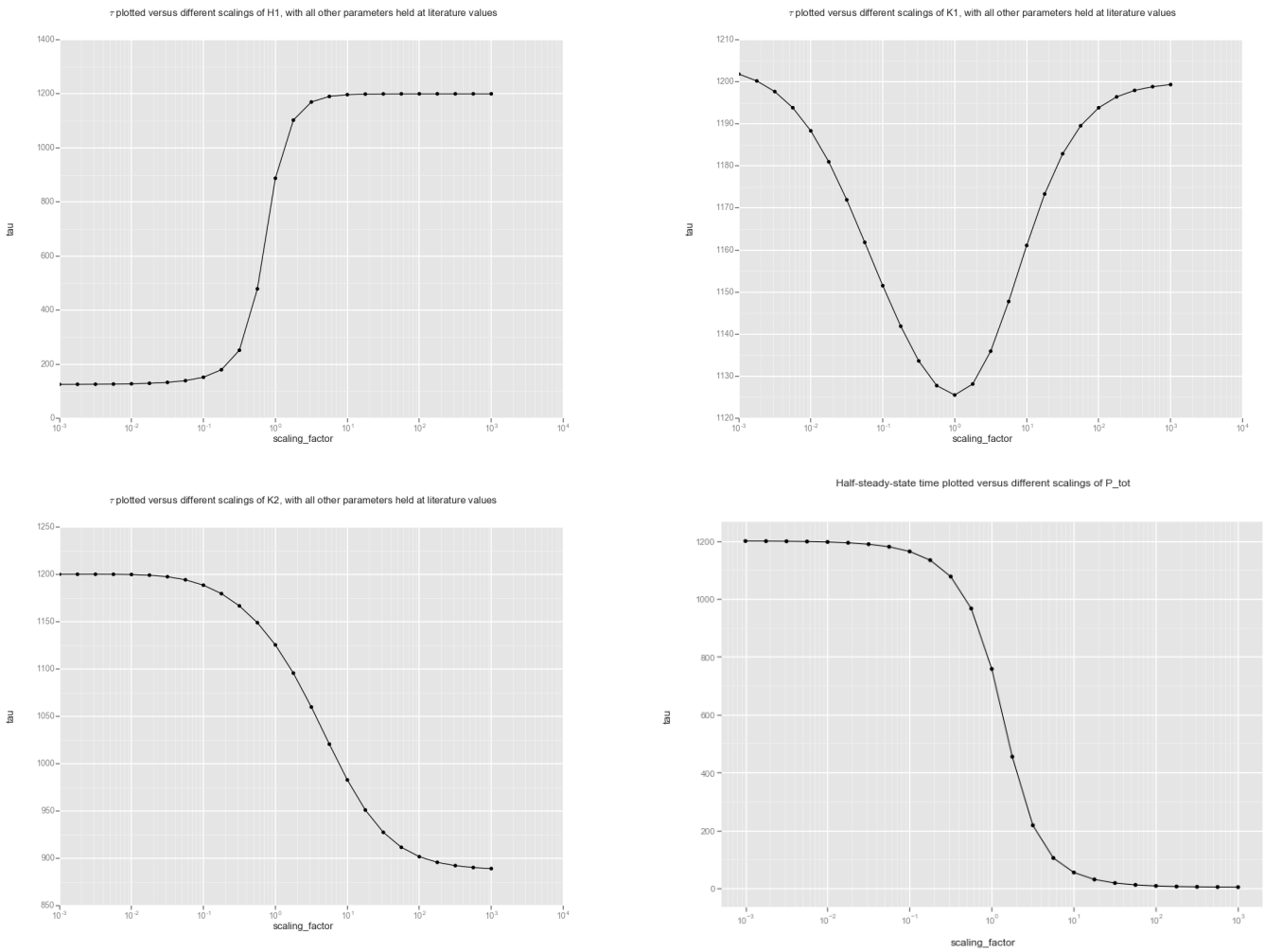

Figure 1: Impact of parameters on τ

Where

$$\Delta = \frac{2k(H - kP + \delta(K - P)) \tan^{-1}\left(\frac{-H+kP+\delta(K+P+2x)}{\sqrt{-H^2+2H(kP+\delta(P-K))-(kP+\delta(K+P))^2}}\right)}{\delta\sqrt{-H^2 + 2H(kP + \delta(P - K)) - (kP + \delta(K + P))^2}}$$

$$\Psi = \frac{(2\delta + k)\log(x(kP + \delta(K + P + x)) - H(K + x))}{\delta} - 2\log(K + x)$$

2.2 Simulation

In systems when there is more than one protein, the same calculus techniques are no longer sufficient to find such a neat solution. Hence, to study speed in systems where proteins compete with each other for degradation or even regulate each other's production, we abandoned analytical techniques in favor of simulations with numerical integrators: we take in a system of coupled differential equations and their respective initial conditions. Using a numerical integrator we integrate over $t \to \infty$ and confirm that each $\frac{dx_i}{dt} \to 0$.

With the help of package **ggplot**, we were able to obtain the above graphs (Figure1)

3 Results

By using the techniques introduced in the previous section, we are able to obtain the following insights from our model:

I Protease degradation behaves as an additional term to dilution when protein concentration is small, but when protein concentration is comparatively large, it can be considered as a reduction from the protein production constant. This effect can be considered in terms of the saturation effect, which is the case when the protease has a reduced impact on the concentration of proteins due to the great amount of proteins present.

II The speed is positively correlated with the rate of protein production under the domain of all positive real numbers. If the promoter and ribosome is changed to express a higher level of gene expression, then the time to half steady-state will increase accordingly.

III The speed of one main protein is negatively associated with the production rate of other proteins.

IV The relationship between K, michealis-menten constant, and τ depends on the production rate of protein. If the rate of expression for protein is significantly lower than the normal protein expression rate, then K will have a positive correlation with τ for all positive values. However, for slightly lower expression, regular expression, or high expression rates, the relationship between τ and K depends on the strength of protein-protease affinity. If the affinity rate of pdt tag is at the same or higher magnitude compared to our parameters, the positive correlation still exists between the half-time to steady-state and K. In such case, there is an optimal value corresponding to τ_{min}.

V Speed will decrease if protease concentration increases. In addition, it will take more time for protein to reach steady-state level given that the protease is at a steady-state concentration compared to the case where the concentration of protease is initially at 0 and reaches steady-state at a timescale similar or smaller than that of protein.

VI The time it takes for protein to reach half steady-state is tunable. To achieve a faster speed, the production rate of protein should be lowered, protease concentration should be increased, or under realistic setting, have a higher protein-protease affinity rate. However, the impact that tuning the parameters have on speed decreases significantly as the magnitude of the parameters changes exponentially.

4 Plug-And-Play For Tuning Speed

We provide a general guideline for the reader to tune speed of a gene circuit. The toolbox takes in the current parameters and speed as an input and outputs the protein-protease affinity rate and protein degradation tag that may result in such speed. Due to the limited number of tags and limitation of the impact of parameters on speed, the speed may not be reachable.In such case, the tag that will make the speed closest to the user input value will be shown. In addition, the speed given the degradation tag, and the difference between the outputted tag speed and the speed that user inputted will be shown. The user can also choose whether to output graph for a given set of pdt tag for the main protein, the michealis-menten values can be chosen by selecting from the dropdown menus generated with respect to the number of proteins. For more information about the toolbox, please consult the following link: `https://github.com/TBigJ/iGEM-2017-toolbox/blob/master/Final_files/Main(ClickMe).ipynb`

All the graphs above are plotted using logrithmic scale, the parameters used to generate the graphs are shown as the following:

Parameters	Value	Dimension	Reference
δ	0.00058	s^{-1}	division of e.coli
K	3.7	µmol	*cameron, collins, 2014*
H	2	µmol min^{-1}	*technion hs israel*
P_{tot}	.25	µmol	*mcbride, 2017*
k	11.5	min^{-1}	*gur sauer, 2008*

References

[1] *Team:technion hs israel/constants database.*

[2] D. E. Cameron and J. J. Collins, *Tunable protein degradation in bacteria*, Nature Biotechnology, 32 (2014), p. 12761281.

[3] E. Gur and R. T. Sauer, *Evolution of the ssra degradation tag in mycoplasma: Specificity switch to a different protease*, Proceedings of the National Academy of Sciences, 105 (2008), p. 1611316118.

[4] D. Huang, W. J. Holtz, and M. M. Maharbiz, *A genetic bistable switch utilizing nonlinear protein degradation*, Journal of Biological Engineering, 6 (2012).

[5] D. McBride, *Analyzing and exploiting the effects of protease sharing in genetic circuits*, (2017).