Use of site specific proteases to make the toggle switch robust to temperature increases

Clair Travis

May 20, 2021

1 Introduction

Multistable gene regulatory networks are ubiquitously found in nature[1]. The first synthetic biological toggle switch, reported by Gardner, Cantor & Collins (2000), was a simple bistable system that can alternate between two different states by using two repressors that repress the transcription of each other shown in Figure 1a [2]. Depending on the starting state and the amount of inducers, this system can have 2 stable equilibrium as show in Figure 1b, where when the stable states get closer to the separatrix, stochasitc variation causes them to degrade into a single state. Additionally a high amount of one inducer can lead to only one stable state also shown in Figure 1b. For this system the ratio of the overall expression rate to the overall degradation rate or β/γ determines bistability (Figure 1c).

2 Problem: Increasing temperature disrupts toggle switch bistability

2.1 Temperature effects on protein-ligand association and dissociation

As temperature increases the number of molecule collisions increases which increases the association rate of proteins with their target ligands [3]. Similarly, the increase of temperature also causes the dissociation rate to increase, as higher energy molecules are more likely to break bonds or intermolecular interactions formed upon complex formation [3]. In antibodies, the increase in dissociation rate is much larger than the increase in association rate upon temperature increases as can be seen in Figure 2 [3]. This means that the dissociation constant $K_d = (dissociation\ rate)/(association\ rate)$ will also increase with temperature and therefore the binding strength of relevant species will decrease. We can generalize this to other protein-ligand interactions given that the mechanism of binding is the same, particularly DNA binding proteins like the repressors used in this study.

2.2 Assumptions and premises

In this study several assumptions are made for the modeling of this system. First it is assumed that there is no protein unfolding due to the increase in temperature. This is a reasonable assumption for mild temperature increases but breaks down at higher temperatures and as the cell goes more into the heat shock response[4]. We will assume that this circuit is in E.coli and that the heat shock response of the cells has no effect on the circuit. It is also assumed that the only change in the toggle switch circuit due to temperature is the increase of K_d with respect to temperature. Furthermore this increase with respect to temperature will only be modeled as an increase in dissociation rate for stochastic simulations, which has the greater change. The change of K_d with respect to temperature will be modeled as an increase of $0.1\ nM/^{\circ}C$ for an initial $K_d = 1nM$. This is an arbitrary assignment for ease of modeling and is likely a lower estimate of

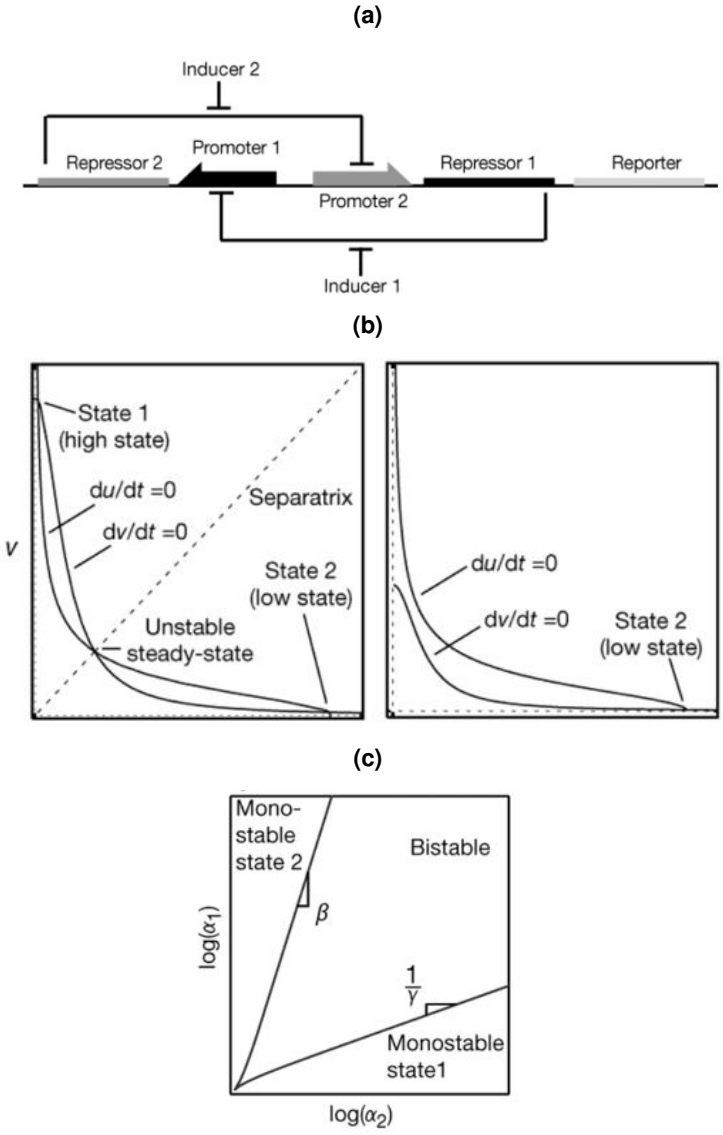

Figure 1: Toggle switch reported by Gardner et al. (2000) [2]. a) Schematic of the toggle switch. b) Nullclines of a bistable system (left) and a monostable system (right). c) Parameter space of mono vs bi stable states. β is the transcription/translation rate of the proteins and γ the overall degradation/dilution rate.

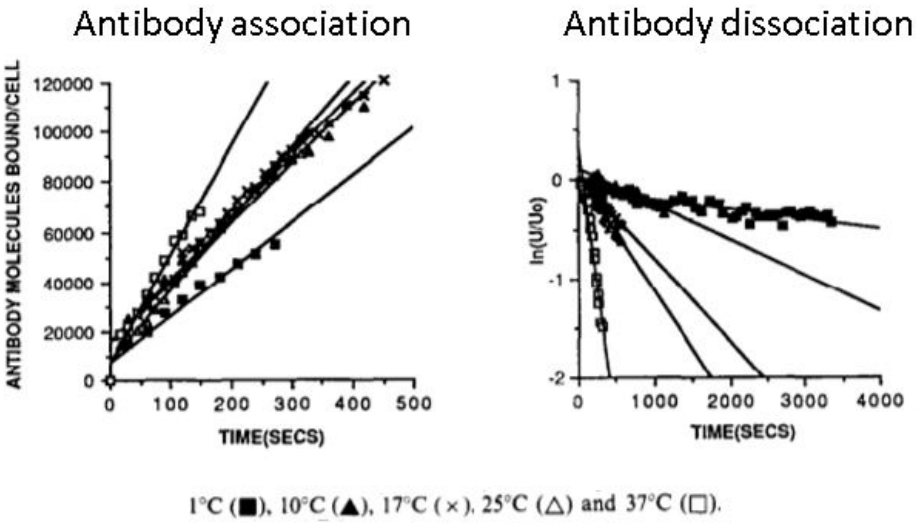

Figure 2: Effect of temperature on the association and dissociation rates of antibodies. Increasing the temperature increase the association rate (slope) of antibodies but to a lesser extent than the dissociation rate. Figures taken from Johnstone et al (1990) [3]

this phenomenon. The temperature change will only be considered in the range of $30°C$ to $40°C$, as this is within the growth range of E. coli and does not reach temperatures that are too high and would induce major changes in protein folding or stress responses (ie above $42°C$).

For the gene expression it will be assumed that mRNA production is at steady state, so the translation of repressors termed X_1 and X_2 will be modeled as lumped parameter β. Additionally the degradation and dilution of these repressors will be modeled as lumped parameter γ. These are assumed to be identical for both X_1 and X_2, and the K_d for the target DNA is assumed to be the same for both of these repressors. These assumptions lead to the symmetric differential equations:

$$\frac{dX_1}{dt} = \frac{\beta}{1 + (X_2/K_d)^2} - \gamma X_1$$

$$\frac{dX_2}{dt} = \frac{\beta}{1 + (X_1/K_d)^2} - \gamma X_2$$

describing the change in species X_1 and X_2 over time which interact as shown in Figure 3. The parameter values will arbitrarily be set to:

$$\beta = 5\ nMs^{-1},\ \gamma = 1\ s^{-1},\ K_d = 1\ nM\ ,\ with\ \beta/\gamma = 5$$

unless stated otherwise, where these parameters result in a bistable distribution at low temperature (ie $T = 30°C$). Finally all enzymes will be assumed to follow Michalis-Menten kinetics and it will be assumed that a concentration of $1\ nM$ corresponds to 1 molecule per cell for stochastic simulations.

2.3 Toggle switch simulation upon an increase of temperature

Nullcline analysis was applied to the toggle switch for no change in temperature (baseline at T=$30°C$), an increase of $5°C$ and an increase of $10°C$ which were plotted as shown in Figure 4 (Top). For this system

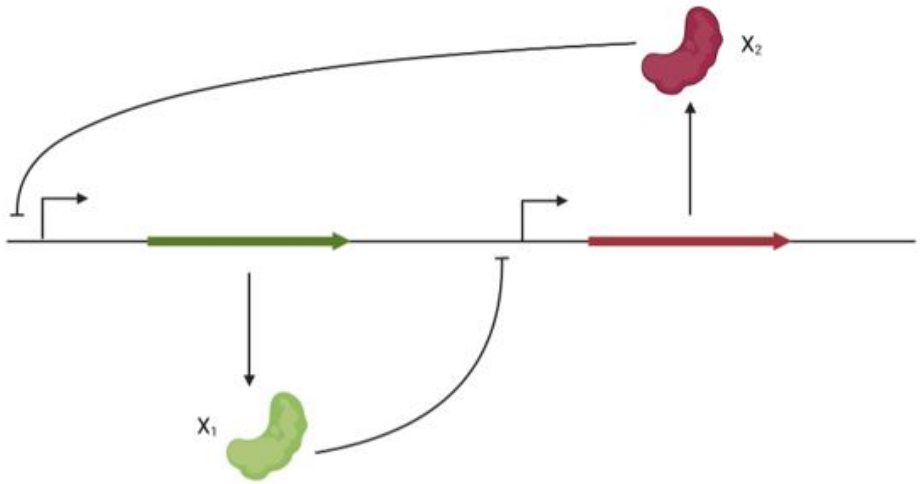

Figure 3: Toggle switch schematic used in this study

there are 3 steady states, of which the outer two are stable. As temperature and hence K_d increases the two stable states get closer and closer until they are almost indistinguishable.

Using the above parameters the toggle switch can be simulated stochastically via the Gillespie algorithm [5]. For the toggle switch, stochastic simulations are particularly useful as opposed to deterministic simulations as they allow for switching to occur so only stochastic simulations will be shown. After simulation the distribution of molecule counts for X_1 and X_2 were plotted in a 2D histogram as shown in Figure 4 (Bottom) for increasing temperature. At T=$30°C$ (baseline) there are two distinct states centered around 20 nM. As the temperature increases these states get closer together and ultimately become indistinguishable at T=$40°C$, which mirrors the results of the nullcline analysis.

These results demonstrate that an increase of temperature can disrupt the bistability of the toggle switch. Therefore to maintain stability in this temperature range there is a need for some method of control.

3 Approach: Site Specific Protease for feed forward control

Feed-forward control is a viable mechanism to make this circuit robust to temperature fluctuations. To implement this, I propose the use of a site specific protease to degrade the X_1 and X_2 that decreases in activity at higher temperatures due to protein unfolding. A site specific protease is desireable, since a non specific protease would likely target some number of cellular proteins and could effect the viability of cells or global processes. For this protease, the tobacco etch virus (TEV) protease was chosen as its temperature sensitivity is well known, it has relatively high efficiency, and it is highly specific for its cleavage site (ENLYFQ'G/S) [6].

As an initial step X_1 and X_2 would need to be engineered to contain a TEV cleavage site. The site could be added in a solvent accessible loop region of the proteins such that it minimally affects the DNA binding function of these proteins. Additionally linker amino acids like G/S repeats may be useful to insert on either side of the TEV cleavage site to increase flexibility of the insert and help maintain structure. With this modification X_1 and X_2 should be sensitive to TEV cleavage, and after cleavage shouldn't bind their DNA targets and get degraded (Figure 5). Here we will assume that this design is possible and that upon cleavage X_1 and X_2 no longer bind their targets, so have no downstream effect. A model of this modified circuit can be seen in Figure 6, where TEV is constiutively expressed such that it is at a constant level. This could be done by genomically modifying the bacterial strain used to express TEV or simply putting it under

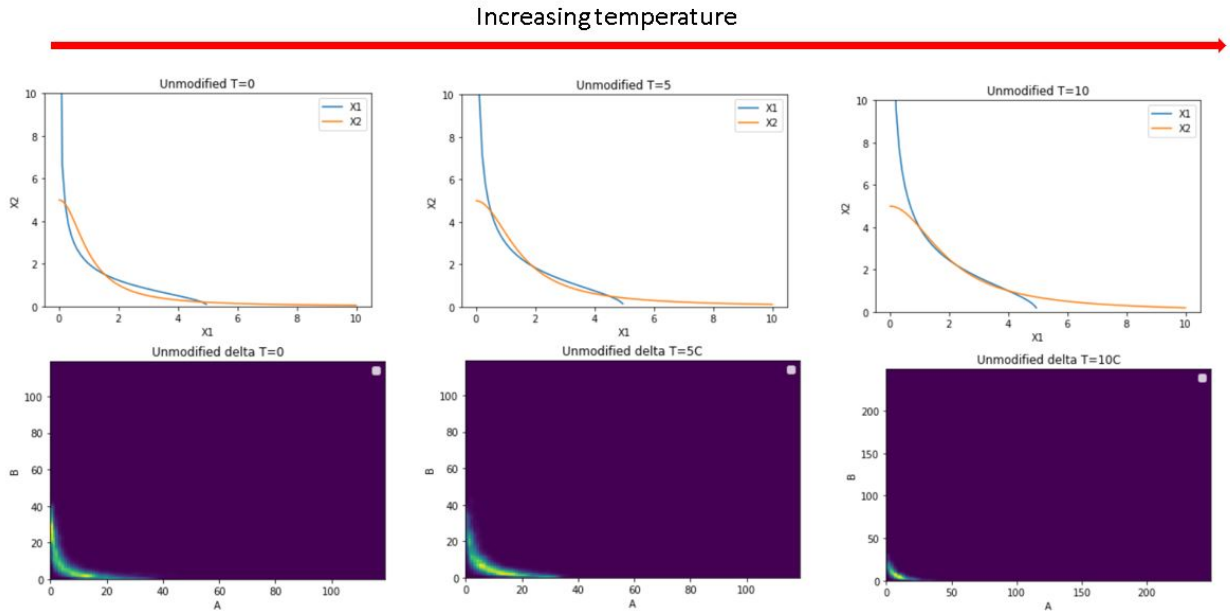

Figure 4: Top: Nullcline analysis of the unmodified toggle switch at increasing temperature. As temperature increases the stable equilibrium become less separated. Bottom: Population of concentrations using stochastic simulation. At low temperature there are two stable states, but as temperature increases they merge into one. [3]

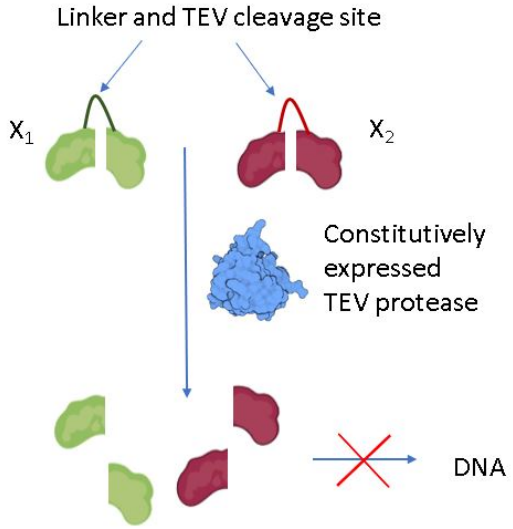

Figure 5: Proposed modification to X_1 and X_2 to allow for clevage by TEV protease

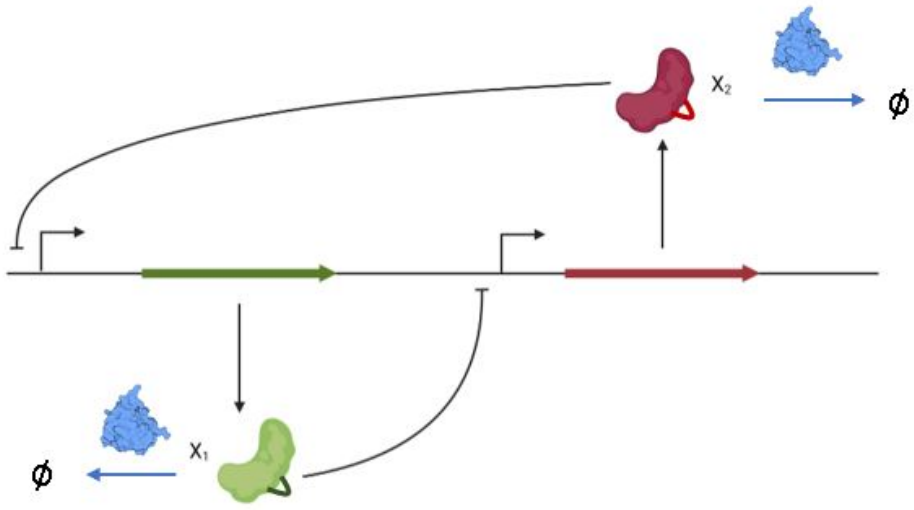

Figure 6: Modified toggle switch circuit to include feed forward control by the TEV protease

a constitutive promoter and inducing the expression of the repressors after the level of TEV protease has stabilized.

To model the TEV cleavage we'll use Michalis-Menten kinectis such that:

$$v = \frac{d[P]}{dt} = \frac{-d[S]}{dt} = \frac{k_{cat}[E][S]}{K_M + [S]}$$

where [S] would be our substrate X_1 or X_2. The K_M of TEV at $30°C$ is $87\ \mu M$ and for this system we are assuming that concentrations are in the nanomolar range with the highest simulations of X_1 or X_2 reaching 100nM [6]. So it is evident that $K_M >> [S]$ so we can make the approximation:

$$\frac{d[S]}{dt} \approx -\frac{k_{cat}[E][S]}{K_M} = -K'[S],\ where\ K' = \frac{k_{cat}[E]}{K_M}$$

Since the level of TEV is constant by design, K' is a constant that can be tuned by changing the level of TEV protease expressed. Additionally TEV is known to be temperature sensitive as shown in Figure 7, where it decreases in overall activity (K') per μg of substrate. We can then assume that we can tune the expression such that the activity of TEV protease at $30°C$ is $K' = 1.0s^{-1}$, and will model the response between $30°C$ and $40°C$ using linear approximation in that region. With some scaling this results in:

$$K'_T = K'_{30} - (\Delta T) \times (-0.075\ s^{-1}{}^{\circ}C^{-1}),\ for\ T \geq 30°C\ and\ K'_{30} = 1\ s^{-1}$$

When this is combined with the toggle switch differential equations the following reaction rate equations are obtained:

$$\frac{dX_1}{dt} = \frac{\beta}{1 + (X_2/K_d)^2} - (\gamma + K'_T)X_1,\ \ \frac{dX_2}{dt} = \frac{\beta}{1 + (X_1/K_d)^2} - (\gamma + K'_T)X_2$$

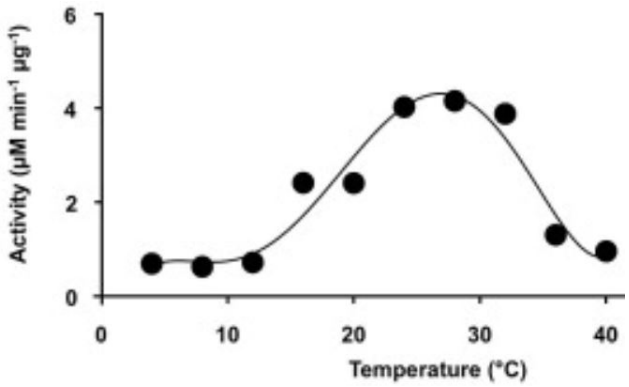

Figure 7: Activity of TEV with respect to temperature. Taken from Raran-Kurussi et al. (2013) [6]

To maintain the same β/γ ratio as in the unmodified toggle switch presented earlier at $T = 30°C$, we need to increase β since γ was effectively increased by adding an additional K' term. Increasing β is relatively simple as a stronger promoter or higher copy number of plasmid could be used. This gives us:

$$\gamma' = \gamma + K'_{30} = 1s^{-1} + 1s^{-1} = 2s^{-1} = 2\gamma$$

Hence β must also be doubled. So the conditions for this system are:

$$\beta = 10nMs^{-1},\ \gamma = 1s^{-1},\ K'_{30} = 1s^{-1},\ K_d = 1nM,\ with\ \beta/\gamma' = 5$$

4 Results

To evaluate if using TEV protease as a feed forward controller is valid for this system nullcline analysis was performed using the above differential equations (Figure 8 (Top)). In the Nullcline analysis, as temperature increases in the range $30°C$ to $40°C$ the stable equilibrium states actually move further apart and hence stabilize. This is a sharp departure from what was observed in the unmodified system in Figure 4, as by $40°C$ the stable states had almost completely disappeared. This indicates that using a temperature dependent protease for a feed forward controller is a powerful method of control and significantly overcompensates for temperature increases. With further tuning, most easily done by changing the amount of TEV protease used, the the stable equilibrium states could likely remain relatively constant over this range.

Next stochastic simulations were employed to confirm that for this feedforward controller that the states were indeed stabilized. Simulations were done as above with the parameters as detailed in the design. Example time course simulations can be seen in Figure 9 and the 2D histograms of the timecourses are shown in Figure 8 (Bottom). From the stochastic simulations it can be seen that at any point in time for all temperatures only two states occur: high X_1 and low X_2 or vice versa (though at early timepoints as expression is increasing this is less pronounced and the level of "high" protein amounts is lower). This is maintained even as temperature increases. Additionally as temperature increases there is less switching between the two states. This is supported by the nullcline analysis which shows that the stable states get further apart, so they require larger stochastic variations to induce a switch. This also explains why in the 2D histograms at higher temperatures only one stable state is visible per stochastic simulation. It is important to note, that in different simulations either of the above mentioned states can occur (data not shown), but due to their stability there is no switching observed (except during of the initial increase in expression) within the time frame simulated.

To address the overcompensation a lower value of K' was analyzed to create a feedforward controller that better maintained the system without any temperature changes. To do so the following parameters were

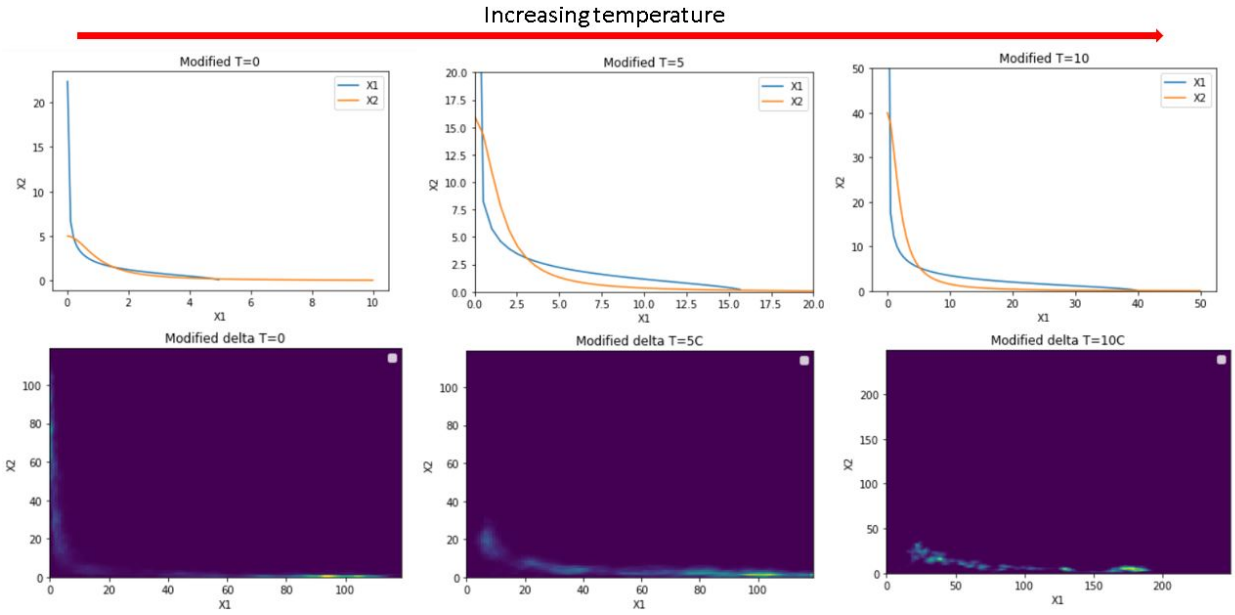

Figure 8: Nullcline analysis of the modified feed forward system (Top) As temperature increases the stable equilibrium get further apart. (Bottom) 2D histogram of stochastic simulated states for a single simulation. Due to computational resource limitations only 200,000 steps were calculated so for higher temperature only one of the stable states could be seen.

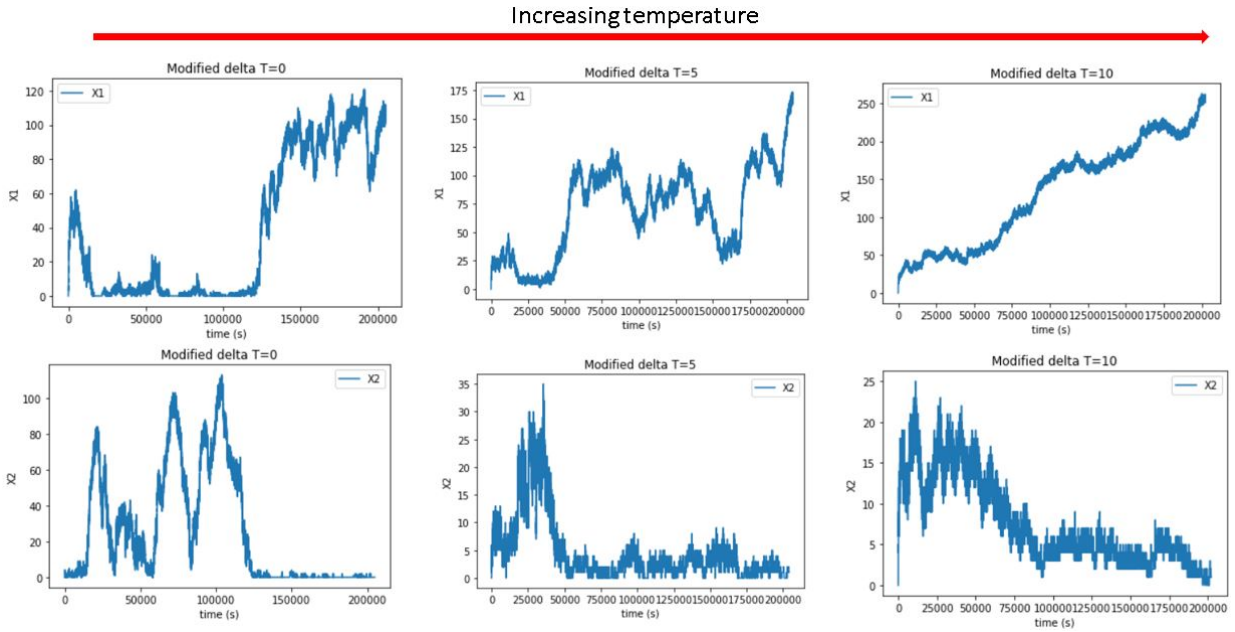

Figure 9: Depiction of sample stochastic simulations over time of X_1 (top) and X_2 (bottom). For all time courses a switching behavior is observed indicative of two stable states. As temperature increases, switching between state becomes less likely and the systems stays in one state. This state is random as other time courses (not shown) have different state chosen for higher temperatures.

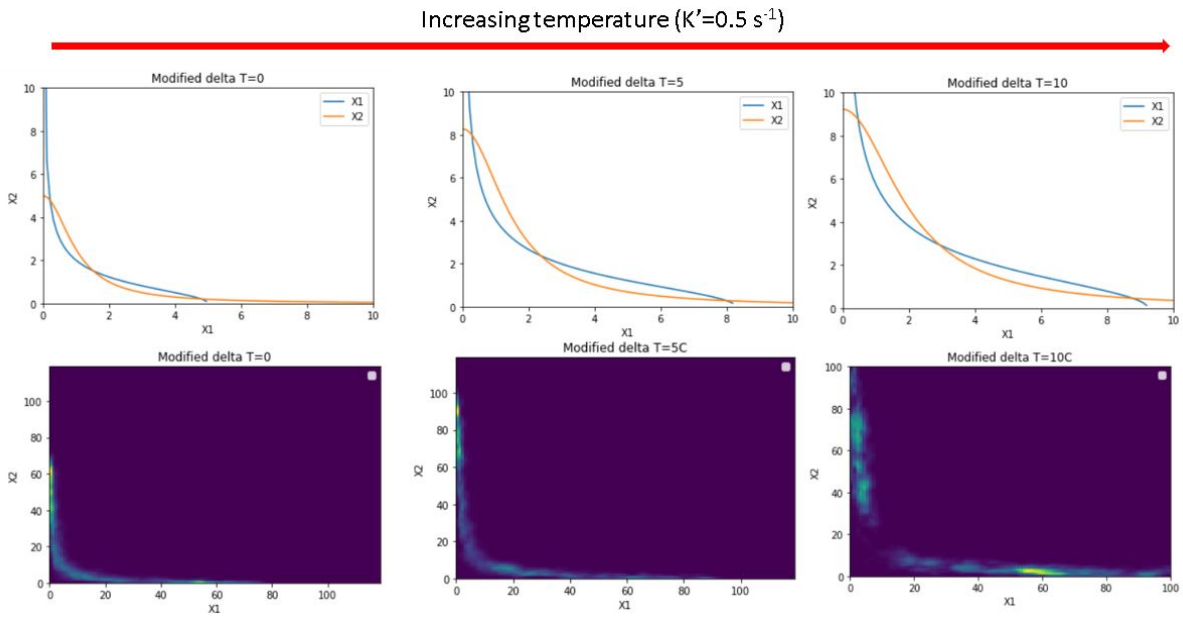

Figure 10: Nullcline analysis of the modified feed forward system for $K' = 0.5s^{-1}$ and $\beta = 1.5\ nMs^{-1}$ (Top) As temperature increases the stable equilibrium still get further apart but to a lesser extent than in the inital case with $K' = 1.0s^{-1}$ and $\beta = 2.0\ nMs^{-1}$ (Bottom) 2D histogram of stochastic simulated states for a single simulation for $K' = 0.5s^{-1}$ and $\beta = 1.5\ nMs^{-1}$. Using the lower level of K' two stable states are observed in all conditions though they do become more distinct as temperature increases.

changed such that:

$$K' = 0.5s^{-1}\ and\ \beta = 1.5\ nMs^{-1}$$

Where β was also changed to maintain the same effective β/γ ratio as the unregulated state at baseline temperature. Using these parameters nullcline analysis was performed as can be seen in Figure 10 (Top). Using these parameters, the two stable states still got further apart as temperature increased, but to a much lesser extent than the original values chosen. Additionally, in stochastic simulations two stable states were observed in all simulations (Figure 10 (Bottom)), meaning that with this set up the switch is still sensitive enough to change occasionally. This is further validated by looking at the timecourses for the simulation as seen in Figure 9. These show distinct switching behavior regardless of temperature, though at higher temperature there appears to be less switching. Overall this demonstrates that tuning the level of TEV protease and hence K' can be used to maintain the initial state to a satisfactory level. It is important to note that becuase this controller acts on γ and not K_d, and has a different repsonse to temperature it will never maintain the initial state exactly. Regardless it can maintain the desired behavior (ie bistable states with a toggle ability within some range).

5 Future directions and potential problems

There are several places where this feed forward controller might fail. On the more biological/chemical side, the design depends on the construction of X_1 and X_2 with an added TEV protease site. This may not be simple to do as there may not be a suitable, solvent accessible loop region to put the cleavage site without disrupting binding or protein folding. Additionally upon cleavage the repressors may still bind the DNA as it is possible the two halves could be held together by intermolecular interactions. On a similar note, the cleavage of the TEV cleavage site may differ between X_1 and X_2, as the structure of these repressors might interfere sterically to different extents with cleavage. Another possible point of failure is the constitutive

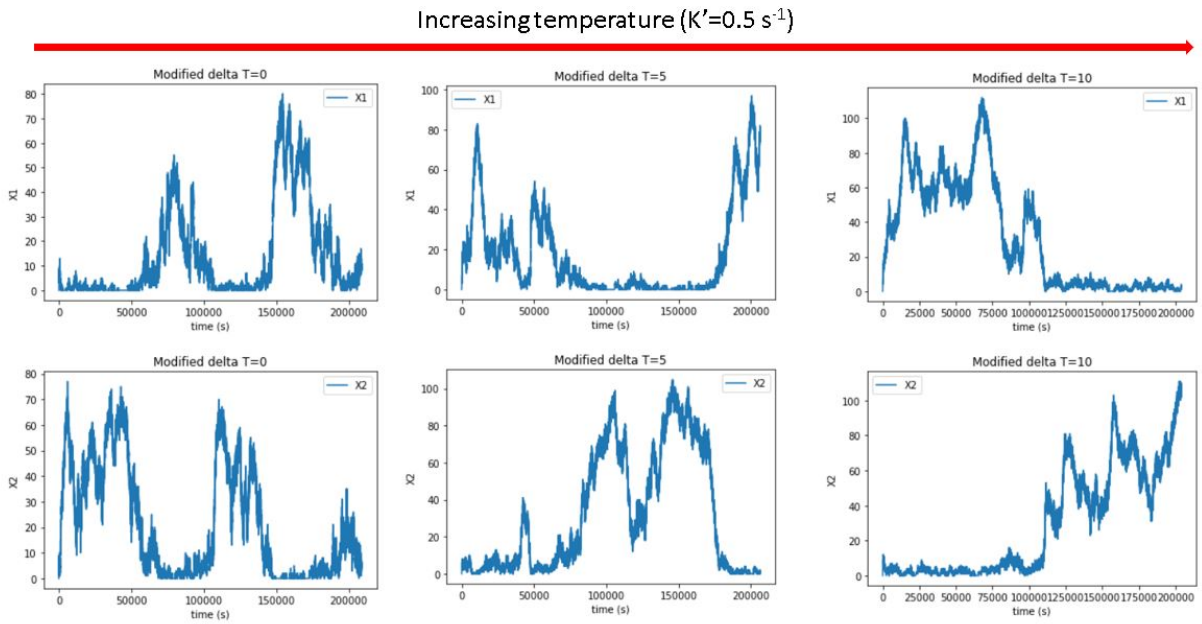

Figure 11: Depiction of sample stochastic simulations over time of X_1 (top) and X_2 (bottom) $K' = 0.5s^{-1}$ and $\beta = 1.5\ nMs^{-1}$. For all time courses a switching behavior is observed indicative of two stable states. Using the new conditions switching is more likely and all conditions demonstrate switching behavior and hence reflect the baseline temperature situation.

expression of TEV protease as an active enzyme in bacteria. This protein is not a native bacterial enzyme, so it may not be correctly folded or expressed well in vitro.

This feed forward controller is also only valid in the range of $30°C$ to $40°C$, as above $40°C$ the TEV protease is almost completely inactive. For this particular set up there is some leeway, as there is a significant amount of over compensation at higher temperatures, namely that the stable states get further apart and become more stable. So temperatures above $40°C$ would remain stable to a certain point but there would be no additional compensation. This isn't too significant of an issue as E. coli doesn't do well in temperatures much above $42°C$, so it is unlikely that the system will need to be used in temperatures much higher than $40°C$ [4]. If this control system were to be used in a thermophilic organism at higher temperatures then a different protease would need to be used, namely one that becomes inactive in the temperature region being used.

It is also important to note that with the setup described here that the TEV based feed forward controller significantly overcompensates for temperature. This is likely because there is a large decrease in activity over the temperature range used in this study and only a mild increase in the K_d. The temperature dependence of K_d was chosen somewhat arbitrily, though should be within an order of magnitude of the actual dependence based on antibody binding changes with temperature. So the actual K_d difference might be higher in which case the system would overcompensate less. Additionally the expression level of TEV could be further decreased to balance it's stabilization effect as demonstrated above or the cleavage site in X_1 and X_2 could be modified to make cleavage less efficent. This system could also be modified by using a different protease with a more shallow response to temperature like the Rhinovirus 3C protease as a way to further tune the system [6]. Finally, changing the expression of TEV could be used as a way to tune the system, where the TEV promoter element could be activated by X_1 or X_2, so when both are low there is less TEV (pushing the stable states apart) while when one is high there is more TEV (destabilizing the stable states).

Overall the use of a site specific protease seems to be a simple and promising method for feedforward

control of the toggle switch. This method could easily be applied to other, more complex systems to enforce multi-state stability, given the above issues are taken into account.

5.1 Availability

All code is available at `https://github.com/cstravis/2.180_Project`

References

[1] J. B. Lugagne, S. Sosa Carrillo, M. Kirch, A. Köhler, G. Batt, and P. Hersen. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. *Nat Commun*, 8(1):1671, 11 2017.

[2] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342, Jan 2000.

[3] R. W. Johnstone, S. M. Andrew, M. P. Hogarth, G. A. Pietersz, and I. F. McKenzie. The effect of temperature on the binding kinetics and equilibrium constants of monoclonal antibodies to cell surface antigens. *Mol Immunol*, 27(4):327–333, Apr 1990.

[4] F. Arsène, T. Tomoyasu, and B. Bukau. The heat shock response of Escherichia coli. *Int J Food Microbiol*, 55(1-3):3–9, Apr 2000.

[5] D. B. Gillespie. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *Journal of Computational Physics*, 22:403–434, 1976.

[6] S. Raran-Kurussi, J. Tözsér, S. Cherry, J. E. Tropea, and D. S. Waugh. Differential temperature dependence of tobacco etch virus and rhinovirus 3C proteases. *Anal Biochem*, 436(2):142–144, May 2013.