PhySci/MiMG/CaSB M178

Homework 7

Due: 11/15/22 at 12:00PM PDT

Problems

In class, we talked about logic gates controlling gene expression. To begin this homework, we will start with a simple model of gene expression. In this model, a transcription factor, such as NFKB, binds to the promoter region of a gene and converts it from an inactive to active state. In this active state, transcription can occur, and mRNA is produced. Let *TF* denote the transcription factor or input signal, whose value can change over time. The effective rate of activation of the promoter region is not only a function of the abundance of *TF*, but also a function of the binding affinity of *TF* to the promoter region, K_d. We include this information in a Hill Equation:

$$H(t) = \frac{TF(t)}{TF(t) + K_d}$$

Assuming the rates of *TF* binding to and unbinding from the promoter region are fast so a steady state is quickly reached, *H* describes the proportion of binding sites that are occupied by the transcription factor. Hence if k_a is the maximum rate of promoter activation, the effective rate dependent on *TF* is given by the product $k_a * H(t)$.

In summary we have the following reactions:

Reactions	Description
$pr \xrightarrow{k\_a * H} pr\_a$	Activation of the promoter region
$pr\_a \xrightarrow{k\_d} pr$	Deactivation of the promoter region
$pr\_a \xrightarrow{k\_syn} pr\_a + tr$	Synthesis of mRNA transcript
$tr \xrightarrow{k\_deg}$	Degradation of mRNA transcript

1. (14 points) First write down the change equations for the model described above.

pr_prime = -k_a*H*pr + k_d*pr_a

pr_a_prime = k_a*H*pr - k_d*pr_a

tr_prime = -k_deg*tr + k_syn*pr_a

In the section of code called "simple gene expression model" implement the change equations to simulate the model. Note how the code defines the value of the Hill

Equation (H); use this value in your implementation. Once you have defined your model equations, run the cell containing the model code.

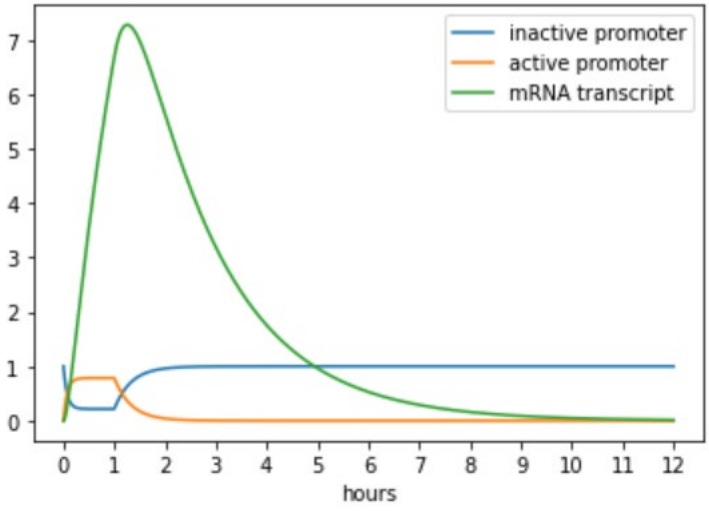
Now we will simulate the model in the section of code called "Simulation of simple model". In this section, we have defined our initial conditions and default parameter values which are also listed below for your reference.

Parameter	Value
k_a	0.2 min^{-1}
k_d	0.05 min^{-1}
Kd	0.1
k_syn	0.1 min^{-1}
k_deg	0.01 min^{-1}

For the simulations in this homework, the profile of activation of *TF* will be a box function (identical to that of IKK activation from HW 6). Hence to describe the temporal dynamics of *TF*, we have to specify an on time, off time, and amplitude (TF_on, TF_off, and TF_amp respectively). By default, TF_on = 0, TF_off = 60 (minutes), and TF_amp = 1.

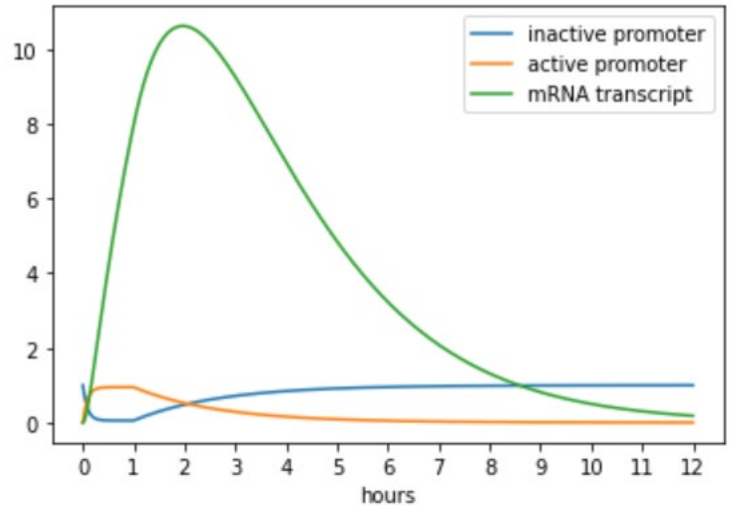
Run the section of code for "Simulation of simple model" and the section of code for "Checking Model Implementation". Check to see that the values on the right (from simulation of your model implementation) match the values on the left (from simulation of the correct model implementation). If the values don't match, double check your change equations and code before proceeding.

Run the section of code called "Plot Dynamics" and paste your graph here. Describe the profile of gene expression over time.



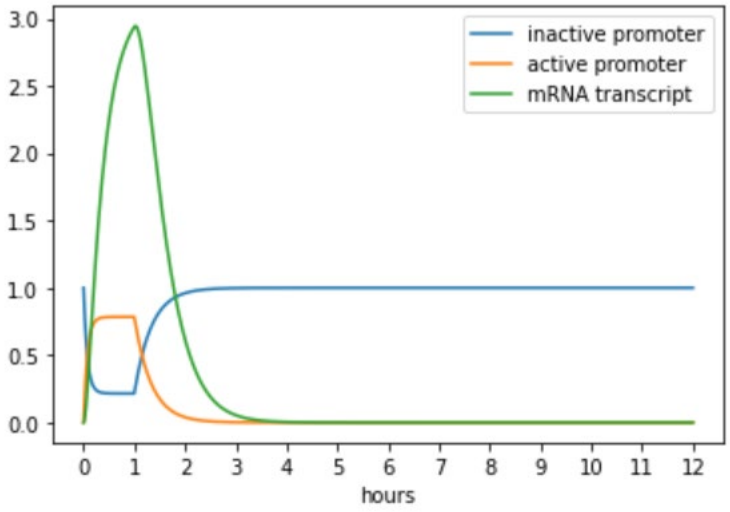
The profile of the gene expression is that the promoter is activated by the prescence of the TF being on from time 0 to 60 minutes, therefore there is a drop due to only degradation of mRNA transcript.

Modify the value for the parameter k_d in the section "Simulation of simple model" to 0.01 and rerun the section. Run the section of code called "Plot Dynamics" and paste your graph here. How does it affect the profile of gene expression over time?



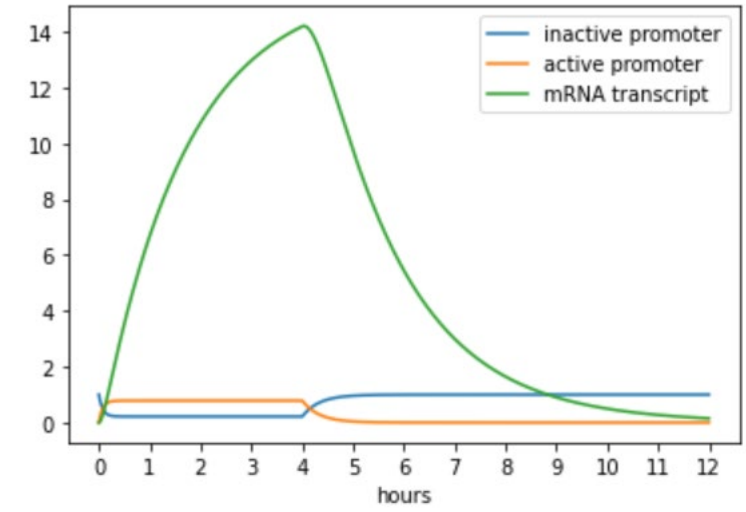
Changing the k_d, which is directly affecting this equation $pr\_a \xrightarrow{k\_d} pr$, means that we are not deactivating the promoter region at a higher rate of 0.05 but instead at 0.01. This means that we should see higher mRNA transcript being made, whereas the inactive promoter and active promoters remain unchanged. Here we notice that the peak mRNA transcript amount occurs at hour 2, around 11 as compared to at hour 1.5 around 7.

Set k_d back to its default value.
Modify the value for the parameter k_deg in the section "Simulation of simple model" to 0.05 and rerun the section. Run the section of code called "Plot Dynamics" and paste your graph here. How does it affect the profile of gene expression over time?



Over time, inactive promoter and active promoter abundance remains the same as previously. However, the k_deg is associated with the degradation of mRNA transcript, which means the abundance will be decreased. The peak of mRNA abundance occurs at 1.5 hours, at 3.

Set k_deg back to its default value.
Modify the value for the parameter TF_off in the section "Simulation of simple model" to 240 and rerun the section. Run the section of code called "Plot Dynamics" and paste your graph here. How does it affect the profile of gene expression over time?



The TF_off parameter is the time that the transcription factor is turned off, mRNA transcript starts tipping off (decreasing) around 4.5 hours instead of 1.5 hours. The promoter activity is correlated with the mRNA transcript activity.

2. (12 points) Now we will simulate a logical "OR" gate. We now have two transcription factors, TF_1 and TF_2, either of which can activate the promoter region. We enforce this logic by modifying our Hill Equation:

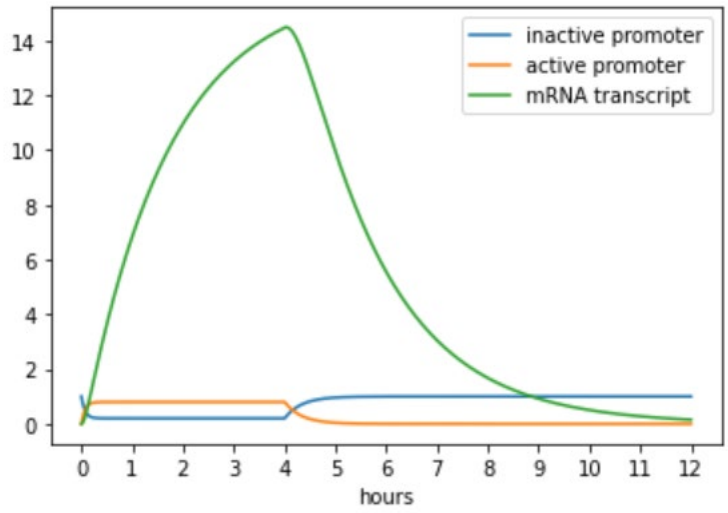
$$H(t) = 1 - \frac{K_{d1}}{TF_1 + K_{d1}} \frac{K_{d2}}{TF_2 + K_{d2}}$$

Now only if both TF_1 and TF_2 are small in value will the value of H be small. If either TF_1 or TF_2 is large in value, H will be large in value.

This model will be implemented in the section of code "OR gate gene expression model". Note that now, we again describe the dynamics of two transcription factors, TF_1 and TF_2. Additionally, we have defined our modified Hill Equation. Besides this change to the expression $H(t)$, the gene expression model is unmodified. **Take your change equations from the "simple_model" and place them into the "OR_model". Run the cell containing the model code.**

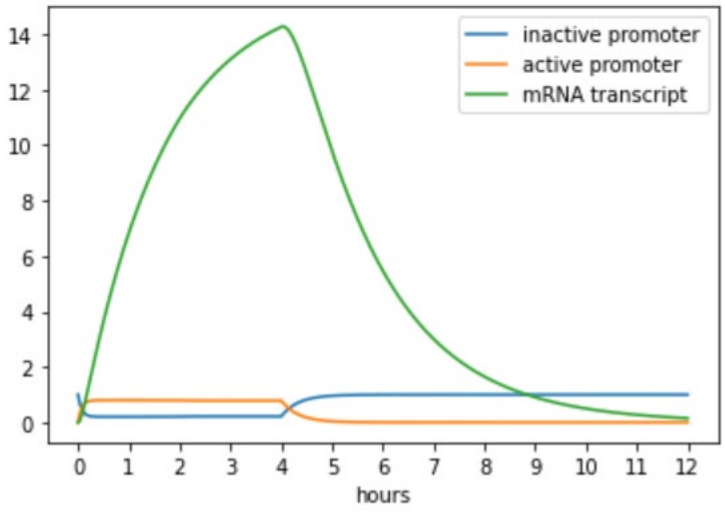
Run the section of code for "Simulation of OR model". Note that we have set the K_d of both transcription factors (Kd1 and Kd2) to the same value, 0.1. Additionally, we define the profiles of activation of TF_1 and TF_2. By default, TF_1 is active from time 0 to time 240 minutes with amplitude 1 and TF_2 is active from time 0 to time 240 minutes with amplitude 1.

Run the section of code called "Plot Dynamics – OR model" and paste your graph here. Describe the profile of gene expression over time. Compare to "simple_model" with TF_off = 240.



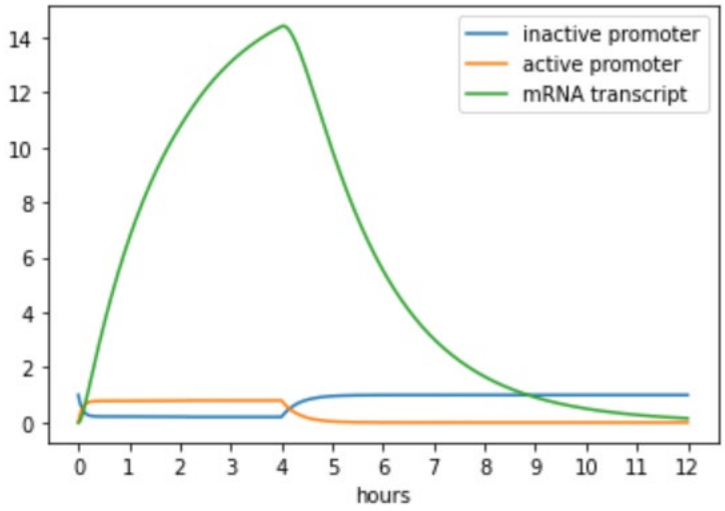
For the OR gate, two transcription factors with TF off at 240 would look like the same as the one transcription factor off at 240. This is because the signaling activation of the promoter only needs one transcription factor for mRNA transcription, and so two is the same effect as one. It is the same as the default values graph.

Modify the timing *of TF2* activation profile so that TF2_on = 0 and TF2_off = 120.
Rerun the section of code called "Plot Dynamics – OR model" and paste your graph here.



In these parameter settings, since the signaling does not depend on TF2 if TF1 is already present, changing the timing of TF2 being on for a shortened amount of time does not affect the activation of the promoter region, therefore, we will not see a change from default trends.

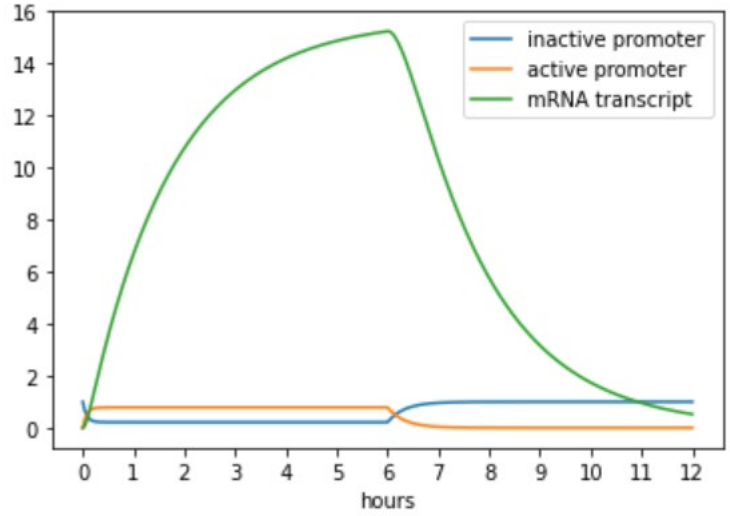
Modify the timing *of TF2* activation profile so that TF2_on = 120 and TF2_off = 240.
Rerun the section of code called "Plot Dynamics – OR model" and paste your graph here.



Explain any differences in gene expression over time in the last two plots compared to the OR model with default parameters. If there are no differences, explain why.

There are no differences because the system needs only one transcription factor for the OR model, there is no difference between these graphs.

Modify the timing *of TF2* activation profile so that TF2_on = 240 and TF2_off = 360. Rerun the section of code called "Plot Dynamics – OR model" and paste your graph here. Explain the difference in gene expression over time compared to the OR model with default parameters.



The plot is getting close the maximum saturation with a delayed mNRA transcript steady-state at 0. The superposition, the OR gate needs one TF to be fully on, and the

effects add up. One follows the other, and so after TF1 drives the promoter activation for hours, the TF2 effects carry over, seeing a longer duration of time of mRNA transcription.

Describe in general a biological context in which a logical "OR" gate controlling gene expression would be advantageous.

A "OR" gate would be advantageous in a biological context where there isn't enough signaling substrates where the system is less discriminatory. This is a more sensitive gene expression case, especially if two substrates are sometimes not present at the same time.

3. (12 points) Now we will simulate a logical "AND" gate. Again, we now have two transcription factors, *TF1* and *TF2*, which are both required for activation of the promoter region. We enforce this logic by modifying our Hill Equation:

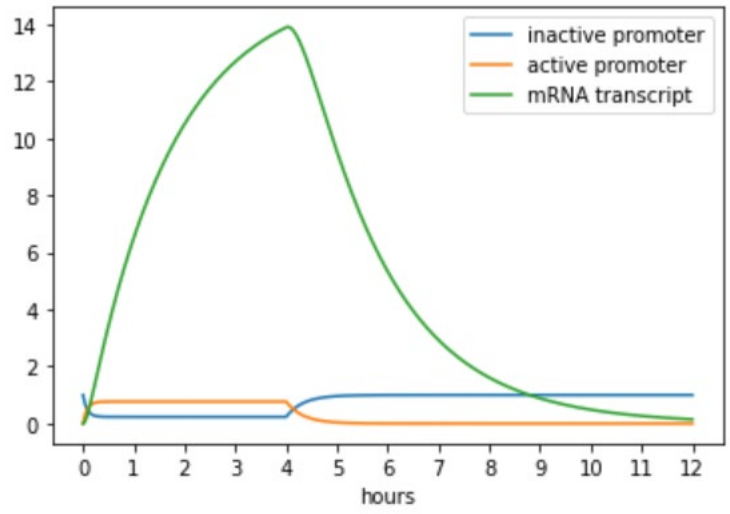
$$H(\mathrm{t}) = \frac{\mathrm{TF}_1}{\mathrm{TF}_1 + \mathrm{K}_{\mathrm{d1}}} \frac{\mathrm{TF}_2}{\mathrm{TF}_2 + \mathrm{K}_{\mathrm{d2}}}$$

Now only if both *TF1* and *TF2* are large in value will the value of *H* be large in value. If either *TF1* or *TF2* is small in value, H will be small in value.

This model will be implemented in the section of code "AND gate gene expression model". Note that now, we describe the dynamics of two transcription factors, *TF1* and *TF2*. Additionally, we have defined our modified Hill Equation. Besides this change to the expression for *H(t)*, the gene expression model is unmodified. **Take your change equations from the "simple_model" and place them into the "AND_model". Run the cell containing the model code.**

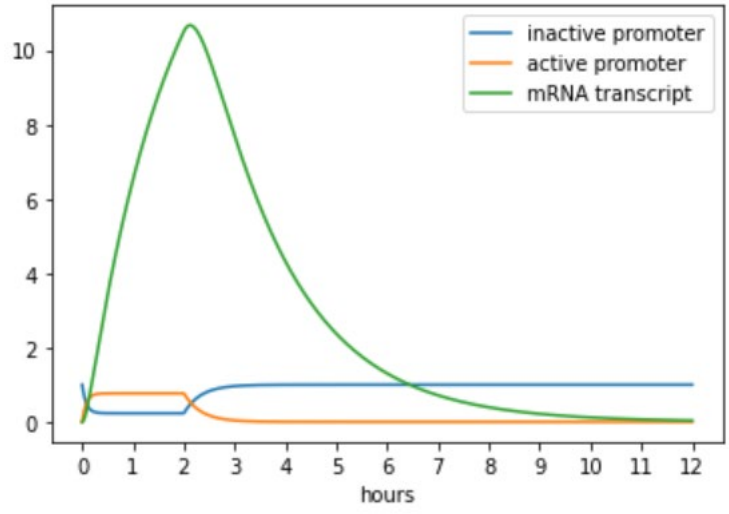
Run the section of code for "Simulation of AND model". Note that we have set the K_d of both transcription factors (Kd1 and Kd2) to the same value, 0.1. Additionally, we define the profiles of activation of *TF1* and *TF2*. By default, *TF1* is active from time 0 to time 240 minutes with amplitude 1 and *TF2* is active from time 0 to time 240 minutes with amplitude 1 (the same as the "OR" model default).

Run the section of code called "Plot Dynamics – AND model" and paste your graph here. Describe the profile of gene expression over time. Compare to "simple_model" with TF_off = 240.



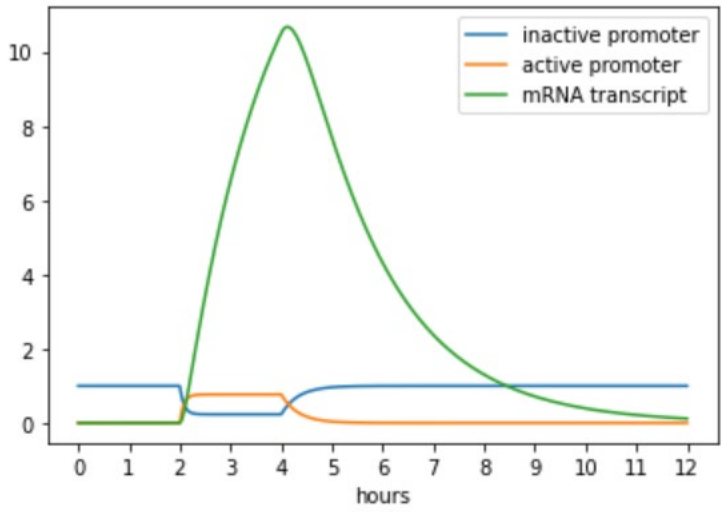
Like in the "simple model", the promoter is activated by the prescence of both TF1 and TF2. Being that both are present for 240 minutes, or 4 hours, the mRNA transcript abundance is heightened until the 4th hour.

Modify the timing *of TF2* activation profile so that TF2_on = 0 and TF2_off = 120. Rerun the section of code called "Plot Dynamics – AND model" and paste your graph here. Explain the difference in gene expression over time compared to the AND model with default parameters.



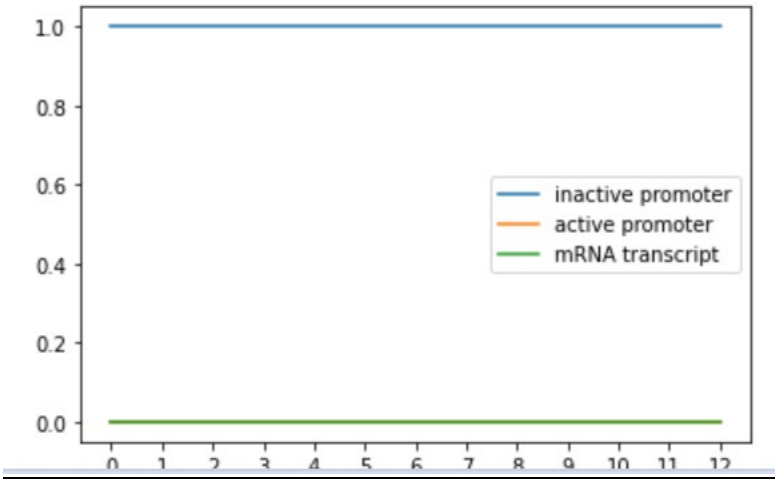
Here the difference is that even if the TF1 is active for 240 minutes, the fact that TF2 is only present for 2 hours makes the activation of the promoter region ongoing for two hours. As a behavior of the AND gate, this is why the activity of mRNA transcript drops after 120 minutes.

Modify the timing *of TF_2* activation profile so that TF2_on = 120 and TF2_off = 240. Rerun the section of code called "Plot Dynamics – AND model" and paste your graph here. Explain the difference in gene expression over time compared to the AND model with default parameters.



The delayed increase in mRNA transcript is because the promoter region is activated only when both TF1 and TF2 are present, and that is during 2 hours and 4 hours. So we notice the mRNA degradation from 4 hours onward.

Modify the timing *of TF_2* activation profile so that TF2_on = 240 and TF2_off = 360. Rerun the section of code called "Plot Dynamics – AND model" and paste your graph here. Explain the difference in gene expression over time compared to the AND model with default parameters.



We do not see mRNA transcript being made nor degraded because TF1 and TF2 do not coincide being present in the system at the same time, TF1 is present from 0 to 240 but TF2 is present from 240 to 360.

Describe in general a biological context in which a logical "AND" gate controlling gene expression would be advantageous.

T cells needs both dendritic and peptide signaling, an example of AND gate controlling. So in the effect that we don't have both components, T cells in the immune system will not have the nucleosome be unwinded. This is where a biological system is extremely restrictive, and needs two TFs.

4. (22 points) Now we will simulate a model in which the promoter region has to undergo two sequential transitions in order to reach its active state. The promoter region is first closed (inaccessible to RNA polymerase) and opens with the binding of a transcription factor, TF_1. The open promoter region is then converted to the active state by a second transcription factor, TF_2.

In summary we have the following reactions:

Reactions	Description
$pr\_c \xrightarrow{k\_a1 * H1} pr$	Opening of the promoter region
$pr \xrightarrow{k\_d1} pr\_c$	Closing of the promoter region
$pr \xrightarrow{k\_a2 * H2} pr\_a$	Activation of the promoter region
$pr\_a \xrightarrow{k\_d2} pr$	Deactivation of the promoter region
$pr\_a \xrightarrow{k\_syn} pr\_a + tr$	Synthesis of mRNA transcript

$tr \xrightarrow{k\_deg}$	Degradation of mRNA transcript

Where the hill equation terms are:

$$H_1(t) = \frac{TF_1(t)}{TF_1(t) + K_d}$$
$$H_2(t) = \frac{TF_{2(t)}}{TF_{2(t)} + K_d}$$

First write down the change equation for the new species, *pr_c*.
pr_c_prime = -k_a1*H1*pr_c + k_d1*pr

Do any of the other change equations from the simple model need to be modified? Write them here if so.

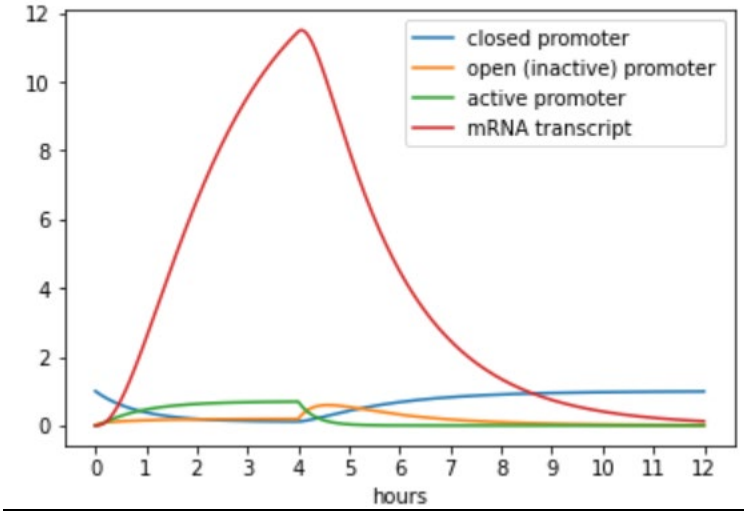
pr_prime = -k_d1*pr + k_a1*H1*pr_c - k_a2*H2*pr + k_d2*pr_a
pr_a_prime = k_a2*H2*pr - k_d2*pr_a
tr_prime = pr_a*k_syn - tr*k_deg

Implement the change equations for this model in the section called "Two step gene expression model". Note how the code defines the value of the Hill Equations (H1, H2); use these values in your implementation. Once you have defined your model equations, run the cell containing the model code.

Now we will simulate the model in the section of code called "Simulation of two step model". In this section, we have defined our initial conditions and default parameter values. Note that k_a1 is lower than ka_2, so the transition from closed to open is slower than the transition from open to active promoter region. By default, both TF_1 and TF_2 are turned on at time 0 with amplitude 1 and turned off at time 240 minutes.

Run the section of code for "Simulation of two step model" and the section of code for "Checking Model Implementation" below it. Check to see that the values on the right (from simulation of your model implementation) match the values on the left (from simulation of the correct model implementation). If the values don't match, double check your change equations and code before proceeding.

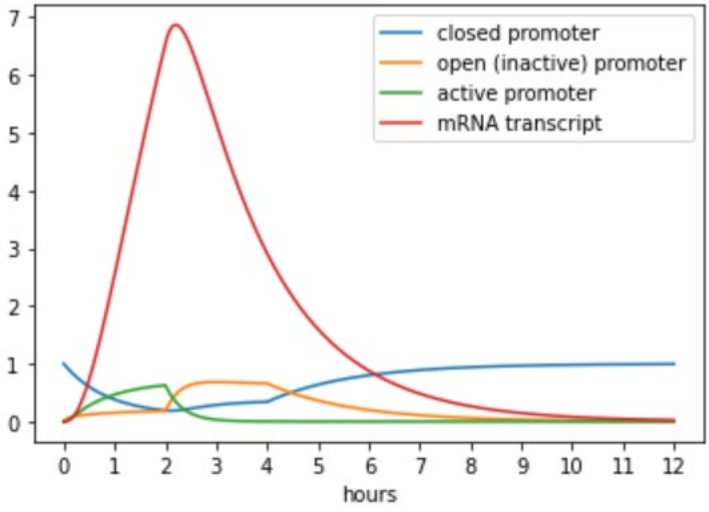
Run the section of code called "Plot Dynamics – two step model" and paste your graph here. Describe the profile of gene expression over time. How does it compare to the simple model results with TF off = 240?



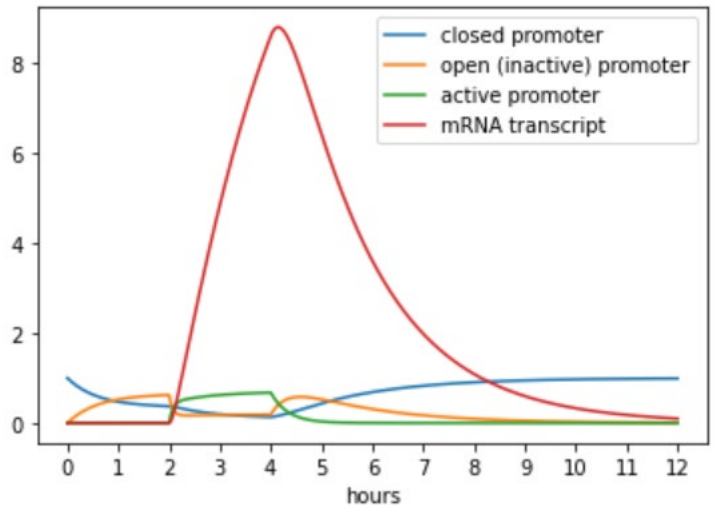
This graph is different because we now are looking at three different states, one is the closed promoter, the open but inactive promoter, and the activated promoter. The difference here is that the hill functions are controlled by one TF, and what we see in this graph is that the peak mRNA transcript abundance is achieved at hour 4, with a smaller peak at 12.

We will now run several simulations varying the on and off times for *TF2*. For each set of values, modify and run the code in "Simulation of two step model". Then paste the resulting graph from "Plot Dynamics – two step model".

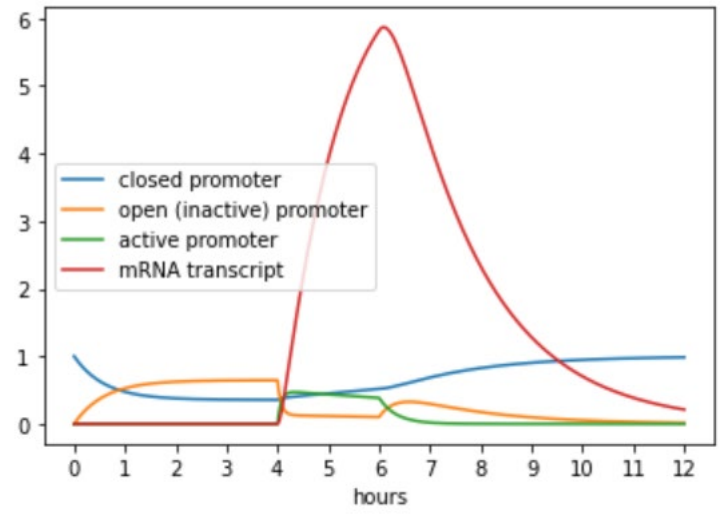
TF2_on = 0 TF2_off = 120



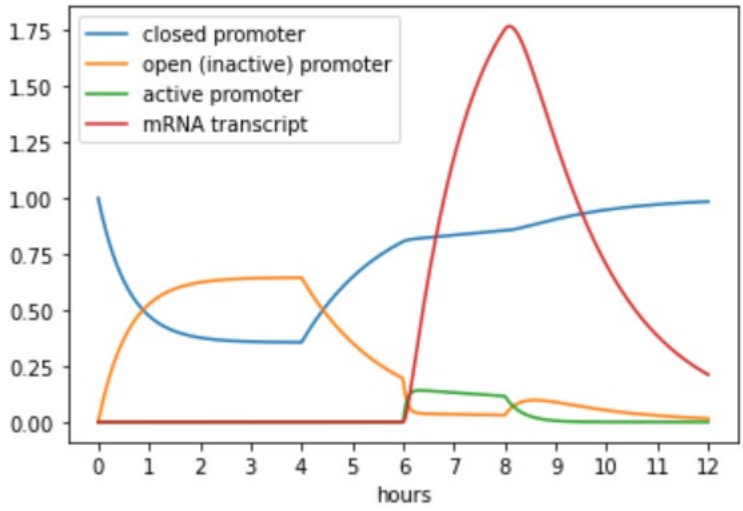
TF2_on = 120 TF2_off = 240



TF2_on = 240 TF2_off = 360



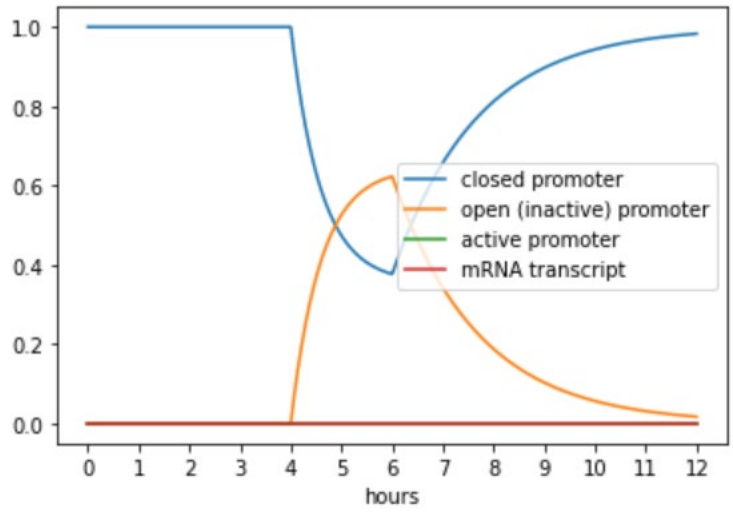
TF2_on = 360 TF2_off = 480



Explain the differences in the profile of gene expression over time for the prior four plots.

In each of these plots, we notice that when T2 is turned on, the mRNA transcription begins. Consequently, when T2 is turned off, mRNA transcription stops and mRNA abundance degrades. This is the difference between each profile of the gene expression over time for the four plots.

Unlike in the "OR" and "AND" models, the roles of TF_1 and TF_2 are distinct in the "two-step" model. Choosing to alter the timing of TF_1 instead of TF_2 can change the results. **In the section "Simulation of two step model" reset TF2_on = 0 and TF2_off = 240 and set TF1_on = 240 and TF1_off = 360. Run the code in "Simulation of two step model" and then paste the resulting graph from "Plot Dynamics – two step model".**



How does the profile of gene expression over time compare to that when TF2_on = 240 and TF2_off = 360 ? What does this tell you about the order of transcription factor activation required for gene expression in the "two-step" model if their activation is sequential?

When TF2_on = 240 and TF_off = 360, the two-step model works as expected because TF1 control PR_c to PR step, and TF1 is controlled by the first hill function H1. However, TF2 controls PR to PR_a, and TF2 is controlled by the second hill function H2. Each hill function is controlled by one TF, but both needs to happen for mRNA transcription. Here we see the AND gate situation, but in separate reactions. While this is similar to question 3, it is dissimilar because sequential activation requires certain TF to be activated first. In the case where TF1 is activated later, there is no activate promoter, therefore no mRNA transcription.

Finally run the section of code called "Plotting all results". This will return a plot summarizing all of your model results in one figure. So far, we have focused on comparing results within one model. Compare results across the models for different TF activation curves. How could you use this information in general to determine the regulatory logic controlling a promoter region?

The promoter regions are regulated by both TF1 and TF2, and from comparing the AND, OR, and two-step model we can conclude that the regulatory logic for controlling a promoter region is the AND model, but in two separate reactions for pr_c to go to pr to finally go to pr_a. The OR gate does not capture this regulatory logic control, because of the last column of graphs on the farthest right.