Basic Synthetic Biology circuits

Note: these practices were obtained from the *Computer Modelling Practicals* lecture by Vincent Rouilly and Geoff Baldwin at Imperial College's course of *Introduction to* Syn*thetic Biology*. Please refer to Vincent Rouilly and Geoff Baldwin as the authors of this material.

Practical 1

Objectives

- To learn how to use a computational modeling tool for biochemical reaction simulations.
    - To build biochemical networks
    - To simulate the time evolution of the reactions k
- To explore the properties of simple biochemical reactions.
    - A --> B --> C model
    - Michaelis-Menten model

Practice

1. ***Building the first model.***

Analysing the following network

$A \xrightarrow{k_1} B \xrightarrow{k_2} C$

Following the Law of Mass action, the dynamic of the system is described as:

$$\frac{d[A]}{dt} = -k_1[A]$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B]$$

$$\frac{d[C]}{dt} = k_2[C]$$

Where k_1=1 and k_2=10 and the initial value of a is 10. Simulate the process during 10 s.

Questions:

1. How do A, B and C, change with time using these default parameters?
2. Now swap the values of k1 and k2 (k1=10 and k2=1)under the parameters tab
3. How does this alter the formation of C?
4. How does B change?
5. Explain these results
6. If you had real life data showing the accumulation of C for an A-B-C reaction you could fit the data using this model and two rate constants would be returned. Could you assign these rate constants to k1 or k2 (yes or no)?

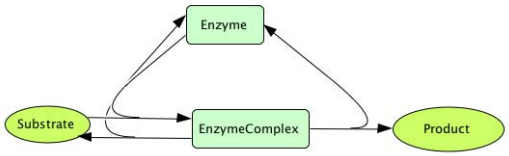
7. What additional data would you need to assign k1 and k2?

2.Michaelis-menten model

An **enzyme** converts a **substrate** into a **product**, this is usually an irreversible reaction and is treated as such in the Michaelis-Menten model. An enzyme reaction constitutes a dynamic process and can be studied as such. One may look at the time courses of the reactants, or look at the steady-states and their stability properties. This part of the practice deals with well-known Michaelis-Menten formula. Here, we will focus on comparing the Michaelis-Menten approximation to the full enzymatic reaction network.

$$E + S \xrightarrow{K_1} ES \xrightarrow{K_3} E + P$$

$$\frac{d[E]}{dt} = k_2[ES] - k_1[E][S] + k_3[ES]$$
$$\frac{d[S]}{dt} = k_2[ES] - k_1[E][S]$$
$$\frac{d[ES]}{dt} = k_1[E][S] - k_2[ES] - k_3[ES]$$
$$\frac{d[P]}{dt} = k_3[ES]$$



Questions

1. From the ODE system description, create all the necessary kinetics reactions in the network provided. We will be considering $k_1 = 10^5 M^{-1}s^{-1}$,$k_2 = 1000s^{-1}$,$k_3 = 10^{-1}s^{-1}$,$[E]_{t=0} = 0.01M$,$[S]_{t=0} = 0.1M$,$[P]_{t=0} = 0$
2. Open the Simulation Panel, set Time=150, NbPoints=1000.
3. Run a simulation, and comment on the different phases during the product formation. Pay special attention to the formation and decay of the [ES] complex. Note that this is a full simulation of the reaction scheme and so does not rely on any assumptions.

$$S \xrightarrow{E} P$$

4. We want now to investigate the Michaelis-Menten approximation. Show that under the assumption that the complex [ES] is at steady-state ($\frac{d[ES]}{dt} = 0$), we can write: $\frac{d[P]}{dt} = \frac{Vmax[S]}{Km + [S]}$. (Note that $[E]_{t=0} = [E]_t$

+ $[ES]_t$). Also, make sure that the concentration of the substrate is at least 10 fold greater than the concentration of the enzyme.

5. Express (Km and Vmax) with regards to k_1, k_2, k_3 and $[E]_0$;
6. Now create a new reaction in CellDesigner with an Enzyme that acts on the reaction. Define the maths for this reaction based on the above form of the Michaelis Menten equation. Make sure that both models are equivalent with regards to their parameters.
7. Run simulations, and comment on the differences observed between the full model, and the Michaelis-Menten approximation.

http://en.wikipedia.org/wiki/Michaelis-Menten_kinetics

http://en.wikipedia.org/wiki/Steady_state_%28chemistry%29

Practical 2

Objectives

To explore computationally some simple genetic motifs:

- Synthesis Degradation Model.
- Constitutive Gene Expression.
- Activated and Repressed Gene Expression.
- Positive and Negative Feedback Gene Expression.

Practice

1. Synthesis degradation model

In this section, we investigate a very common motif in biochemistry. It models the continuous and constant synthesis of a compound, and its natural degradation. From a Mathematical point of view, the model is described as a first-order linear ordinary differential equation.

0->A->0

$$\frac{d[A]}{dt} = k_1 - k_2[A]$$

(k_1=1.0 and k_2=0.01)

Questions:

1. Run a simulation over t=1000s, nb points=1000. Comment on the time evolution of 'A'. (illustration needed).
2. Using the dynamical system definition, what is the steady state level of 'A' with regards to the parameters k1 and k2 ? (Steady state means that $\frac{d[A]}{dt} = 0$
3. Using the 'Parameter Scan' feature, illustrate the influence of both parameters (k_1 and k_2), on the steady state level of 'A' (illustration needed).
4. **Bonus**: Give the analytical solution of the ODE system.

5. Now, consider that k_1=0, and $[A]_{t=0} = A_0 > 0$. Keep k_2=0.01. Illustrate the concept of half-life for the compound 'A'.

http://en.wikipedia.org/wiki/Exponential_decay

2. *Constitutive gene expression*

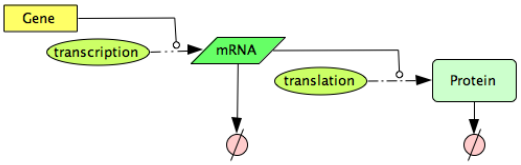
In this section, we explore a computational model to describe a constitutively expressed gene. The model is based on a simple interpretation of the central dogma: **Gene -(transcription)-> mRNA -(translation)-> Protein**, with both the mRNA molecules and the protein molecules being naturally degraded.

The modelling parameters, used throughout this practical, are characteristic of the E.Coli bacteria.

Gene --> mRNA → Protein

$$\frac{d[mRNA]}{dt} = k_1 - d_1[mRNA]$$

$$\frac{d[\mathrm{Protein}]}{dt} = k_2[mRNA] - d_2[\mathrm{Protein}]$$



Where:

k_1 is the **transcription rate**. It is considered to be constant, and it represents the number of mRNA molecules produced per gene, and per unit of time.

- d_1 is the **mRNA degradation rate** of the mRNA molecule. The typical half-life for the mRNAs, in E.Coli, has been measured to be between 2min and 8min (average 5min).
- k_2 is the **translation rate**. It is considered to be constant, and it represents the number of protein molecules produced per mRNA molecule, and per unit of time.
- d_2 is the **protein degradation rate**. In this practical, we will only consider very stable proteins, i.e. not engaged in any active degradation pathways. In that case, we can approximate the degradation of the protein to be only due to the dilution effect caused by the cell division. Cell division will be 40min.

Questions:

1. From the ODE system given, write down the steady-state expression of the [mRNA] concentration and the [Protein] concentration, with regards to

k_1,k_2,d_1,d_2. *Remember that steady-state means that we consider d[mRNA]/dt=0 and d[protein]/dt=0*.

2. Using the previously found equations, and knowing that average number of mRNA molecules per gene is 2.5 in E.Coli, what is the average transcription rate ? *Keep in mind that the problem only gives the mRNA half life, not the actual degradation rate.*
3. In the same way, knowing that the average number of proteins per gene is 1000 in E.Coli, what is the average translation rate ? *Once again, remember that the problem gives the protein half-life, and not the actual protein degradation rate.*
4. In CellDesigner, define all the necessary kinetics laws for the model, and create all the appropriate parameters.
5. Run a simulation, and comment on the simulation outputs (mRNA and Protein), with regards to the transient phase and the steady-state.
6. From a Synthetic Biology point of view, this motif can be seen as a 'Protein Generator'. One might be interested in controlling the steady-state protein output level of this device. Using the 'parameter scan' function, run a simple sensitivity analysis on each of the 4 parameters, within a 10% range of their default value. Illustrate, and describe briefly, how each parameter impacts the protein steady state.

Next, we want to explore the quasi-steady-state assumption on the mRNA molecules expression. From the previous simulations, you might have noticed that the concentration of mRNA reaches steady-state very quickly, compared to the protein concentration. In that case, we want to explore a model where we would consider that the [mRNA] concentration is always considered at steady state , i.e. $\frac{d[mRNA]}{dt} = 0$ all the time. This model means that we want to apply a quasi-steady-state assumption on the [mRNA] molecules.

Gene→ Protein

$$\frac{d[\mathrm{Protein}]}{dt} = s - d[\mathrm{Protein}]$$

Questions:

7. Taking into account the quasi-steady-state assumption on the [mRNA], work-out the value of 's', and 'd' with regards to k_1,k_2,d_1,d_2, so that the two models are equivalent.

8. Simulate the full model, alongside with the quasi-steady state approximation. When using the parameters from the previous section, comment on how good this approximation seems to be.
9. Why do you think such a reduced model would be useful?

Useful data for *e.coli:*

http://redpoll.pharmacy.ualberta.ca/CCDB/cgi-bin/STAT_NEW.cgi

Definition of steady state

http://en.wikipedia.org/wiki/Steady_state_(chemistry)

3. Activated and repressed gene expression

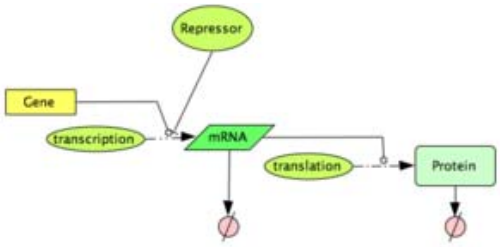
Very few genes are known to have a purely constitutive expression, most genes have their expression controlled by some outside signals (DNA-binding proteins, Temperature, metabolites, RNA molecules ...). In this section, we will particularly focus on the study of DNA-binding proteins, called transcription factors. These proteins, when binding to a promoter region, can either have an activation effect on the gene (positive control), or a repression effect (negative control). In prokaryotes, control of transcriptional initiation is considered to be the major point of regulation. In this part of the practice, we investigate one of the most common model used to describe this type of interactions.

Let's first consider the case of a transcription factor acting as a **repressor**. A repressor will bind to the DNA so that it prevents the initiation of transcription. Typically, we expect the transcription rate to decrease as the concentration of repressor increases. A very useful family of functions to describe this effect is the Hill function: $f(R) = \frac{\beta . {K_m}^n}{{K_m}^n + R^n}$. The Hill function can be derived from considering the transcription factor binding/unbinding to the promoter region to be at equilibrium (similar to the enzyme-substrate assumption in the Michaelis-Menten formula). This function has 3 parameters: β,n,K_m:

1. β is the maximal expression rate when there is no repressor, i.e. $f(R = 0) = \beta$.

2. K_m is the repression coefficient (units of concentration), it is equal to the concentration of repressor needed to repressed by 50% the overall expression, i.e $f(K_m) = \frac{\beta}{2}$
3. *n* is the Hill Coefficient. It controls the steepness of the switch between no-repression to full-repression.

$$Repressor \perp Gene \rightarrow mRNA \rightarrow Protein$$



$$\frac{d[mRNA]}{dt} = \frac{k_1 K_m^n}{K_m^n + R^n} - d_1[mRNA]$$

$$\frac{d[\mathrm{Pr}otein]}{dt} = k_2[mRNA] - d_2[\mathrm{Pr}otein]$$

Consider K_m=100, n=2. k_1, k_2, d_1, d_2 as found previously.

Questions:

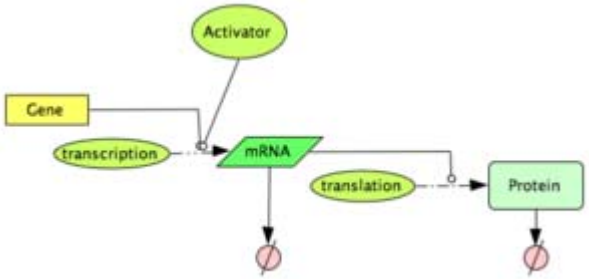
4. We want to establish the transfer function between the [repressor] concentration and the protein steady-state level. To do so plot $[Protein]_{steady-state} = F([Repressor])$ using the results of the simulations where the repressor concentration varies from [0 to 1000] by steps of 10.
5. Suggest an application where this genetic circuit might be useful.

Now, let's consider the case of a transcription factor acting as an **activator**. An activator will bind to the DNA so that it promotes the initiation of transcription. Typically, we expect the transcription rate to increase as the concentration of activator increases. Once again, the Hill type function will be useful to describe the interaction effect. It is slightly different from the previous one:

$f(R) = \frac{\beta . A^n}{{K_m}^n + A^n}$. The Hill function can be derived from considering the transcription factor binding/unbinding on the promoter region to be at equilibrium (similar to the enzyme-substrate assumption in the Michaelis-Menten formula). This function has 3 parameters: β,*n*,K_m:

- β is the maximal expression rate when there is a lot of activators, i.e. $f(A \gg K_m) = \beta$.
- K_m is the activation coefficient (units of concentration), it is equal to the concentration of activator needed to activate by 50% the overall expression, i.e $f(K_m) = \frac{\beta}{2}$
- *n* is the Hill Coefficient. It controls the steepness of the switch between no-repression to full-repression.

$$\begin{array}{c} Activator \\ \downarrow \\ Gene \rightarrow mRNA \rightarrow Protein \end{array}$$



$$\frac{d[mRNA]}{dt} = \frac{k_1 A^n}{K_m^n + A^n} - d_1[mRNA]$$

$$\frac{d[\mathrm{Pr}otein]}{dt} = k_2[mRNA] - d_2[\mathrm{Pr}otein]$$

Consider K_m=100, n=2. k_1, k_2, d_1, d_2 as found previously

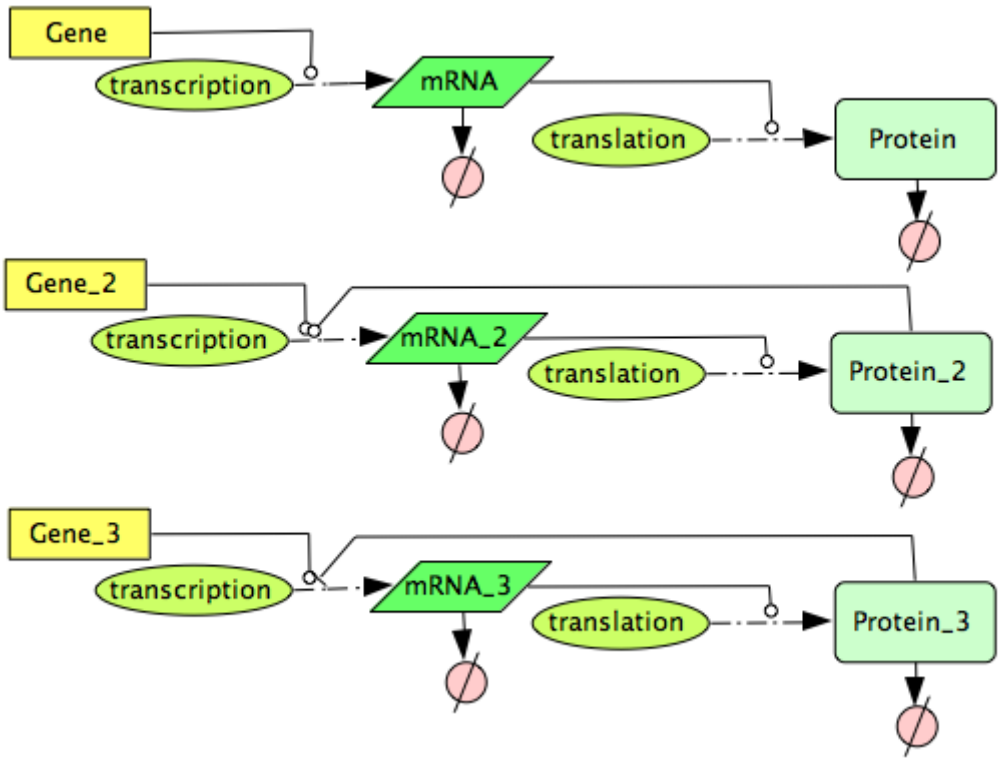
6. We want to establish the transfer function between the [Activator] concentration and the protein steady-state level. To do so plot $[Protein]_{steady-state} = F([Activator])$ using the results of the simulations where the Activator concentration varies from [0 to 1000] by steps of 10.
7. Suggest an application where this genetic circuit might be useful.

4 Possitive and negative feedback.

We have just seen that some genes can be controlled by DNA-binding proteins. This type of interactions will enable the construction of genetic networks, called transcription networks. One gene produces a protein, which then binds to one or more other genes, and control them (positively or negatively), and so on.

At the same time, it has been observed that some genes can also regulate themselves (positively or negatively). When a gene is regulated by the very same protein it produces, the motif is called a feedback. It can be a positive feedback, or

a negative, depending on the type of interaction there is between the protein and the promoter region. In this section, we will explore those 2 options.



Questions:

1. Provide the 3 ODE systems you need to describe the 3 different motifs:constitutive, negative feedback, and positive feedback.
2. Create all the necessary kinetic laws considering the parameters as follow:

    - **Constitutive Expression**: K_1=.346; d_1=.138 ; K_2=6.931 ; d_2=0.017 .
    - **Negative Feedback Expression**: K_1=33.847 ; n=1; Km=10 ; d_1=.138 ; K_2=6.931 ; d_2=0.017
    - **Positive Feedback Expression**: K_1=.346 ; n=1; Km=10 ; d_1=.138 ; K_2=6.931 ; d_2= 0.017

3. Compare the different models with regards to their protein steady-state level, and also with regards to how fast each model is reaching its steady-state.
4. Comment on the benefits and drawbacks of using a negative, or positive, feedback.

Practical 3

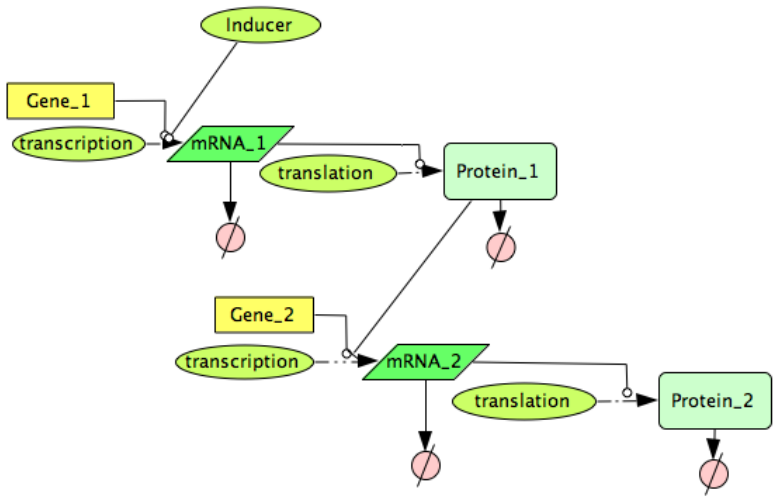
Objectives

- To explore computationally some simple genetic circuits:
  - Genetic cascade to be tunned
  - A genetic oscillator: the Repressilator.
  - A mystery device.

1. Genetic cascade to be tunned

In Practical 2, we have explored how gene expression can be controlled (positively or negatively) by other proteins. Here, we start to assemble genetic circuit by connecting genes. The model studied here is a simple genetic cascade, where a Gene_1 produces a protein which repressed the expression of Gene_2. Additionally, you are able to induce the expression of Gene_1 thanks to an inducer.

Build the model network:



Questions:

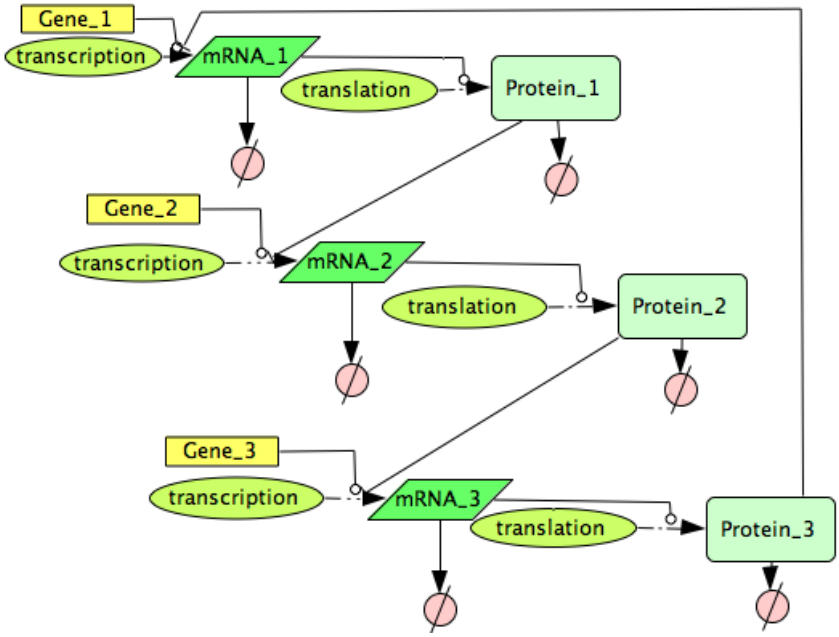
Using the default kinetics parameters, build the following 3 transfer functions:

- Between the [inducer] concentration, and the Protein_1 steady-state expression level. Take the inducer level to vary between [0, 10000]. Comment on the behaviour observed (Illustrations)
- Between the Protein_1 concentration (to be kept constant during the simulations), and the Protein_2 steady-state expression level. Make

sure that the range covered by the concentration of Protein_1 helps to explore the full dynamic range of system behaviour.

- Between the [inducer] concentration, and the Protein_2 steady-state expression level. Take the inducer level to vary between [0, 10000]. Comment on the behaviour observed.
- We want to tune this cascade so that when using our maximal [inducer] concentration we are able to completely shut-down the expression of Protein_2. Suggest 2 independent strategies to achieve that result.

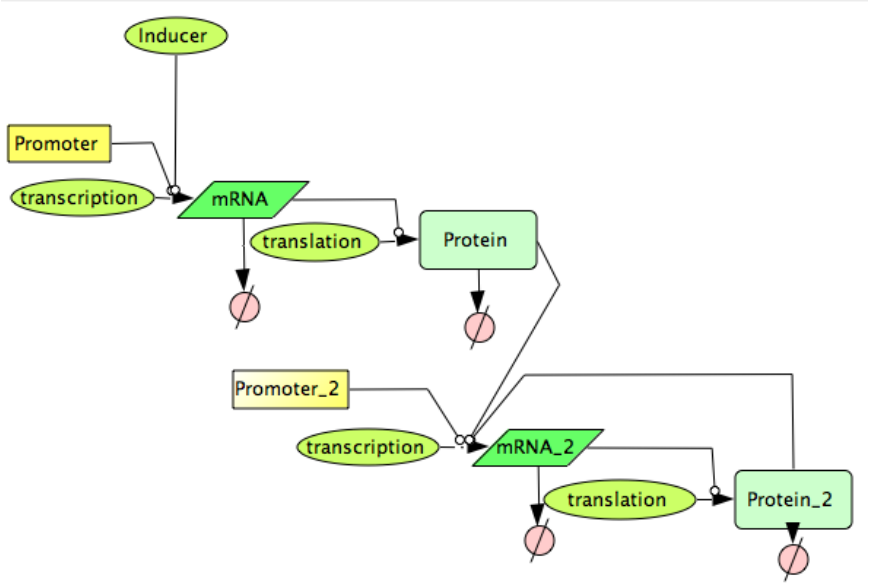
2. The repressilator



Questions:

- Focus on the modelling section of the published paper, and build your own 'Repressilator' model in CellDesigner.
- Run a simple sensitivity analysis on each parameter, and suggest some strategies to control the amplitude of the oscillations.
- Do the same to suggest ways to control the frequency of the oscillations.
- **Bonus**: Could you explain the drift observed on the experimental data ? Provide a simulation supporting your hypothesis.

4. Another device



Questions:

- Suggest a strategy to characterise this model (Establish clear simulation scenario to illustrate the properties of this genetic circuit. Tip: The cell designer function related to 'Change Amount' might be handy)
- From your characterisation, how would you call this device?
- Imagine an application where this genetic device would be useful.