Simulation and analysis of some simple biochemical networks

Computer exercise, Tek292, 2013

Contact: Behruz Bozorg (behruz@thep.lu.se, 046-222 9079)

1 Introduction

In this assignment we will start modeling simple reactions and continue by defining more elaborated molecular networks to investigate more complex dynamics. Finally we will also add diffusion and investigate spatial organization in such systems.

Exercises are divided into analytical exercises (**A1-A4**) (which are optimally solved before the scheduled computer exercise) and computer exercises (**C1-C8**), which will be solved using MATLAB software (Mathworks, Natick, MA). In addition there are two extra computer exercises (**C9-C10**), which can be addressed depending on time. The solutions for the exercises must be reported in a report (see guidelines for details).

2 Background

When biological networks have been investigated, it has been found that the same type of dynamical structures, so called *network motifs*, can be found in many systems [1]. Although these network motifs are built with different genes and molecules in different organisms, the interactions are typically the same and hence the dynamical behavior is the same. In this assignment, the goal is to investigate simple systems representing two of the most common dynamical network motifs found in nature.

2.1 Bistable systems

Cells often need to be able to make decisions. For example it can be to determine differentiation paths or respond to environmental signaling. These decisions are often of binary nature, i.e. either on or off. It has been found that cellular decisions often are regulated by a small number of genes that then leads to large downstream changes within the cell. A common motif for a binary decision is to have two genes repressing each other, hence creating a winner take all situation, where either one of the genes, but not both are turned on. Depending on some initial con-

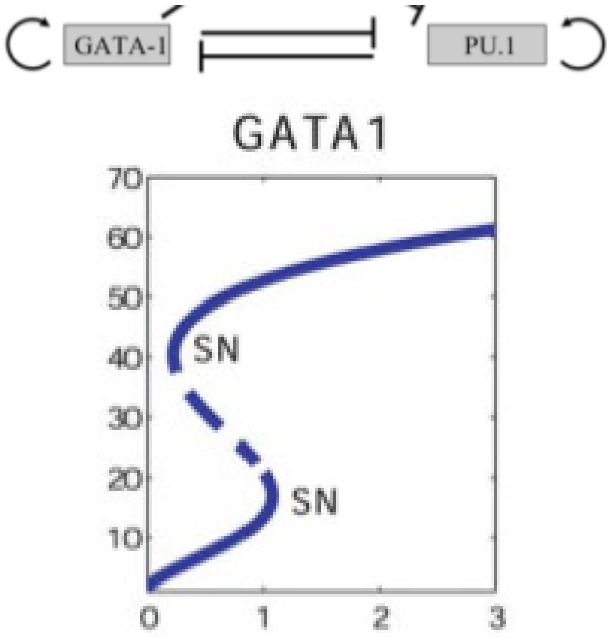


Figure 1: Bistable network determining the differentiation decision between red and white blood cells. The plot shows tha possibility of bistable behavior, where depending on initial concentrations, the GATA-1 gene can either be on or off.

ditions, that may be determined by an external signal, the system can end up in two different states. Hence the network has a bistable dynamics with two different stable fixed points. An example of such a doubly repressing pair of genes is found in the differentiation decision in the hematopoietic erythroid-myeloid switch, where the GATA-1 and PU.1 genes are repressing each other (Fig. 1, [2]). If GATA-1 is expressed the cell will start a downstream transcription program and differentiate into a red blood cell and if PU.1 is expressed it will differentiate into a white blood cell.

2.2 Oscillatory systems

Several biological behaviors are determined from oscillatory systems. Prime examples are the cell cycle and the circadian clock. A network motif that leads to oscillations is to have activation steps that is completed with a negative connection creating a negative feedback loop. An example is a model of the cell cycle where the protein cyclin activate cdc kinases and where the end product is activating degradation of the cyclin molecule, completing the negative feedback loop (Fig. 2, [3]). The dynamics of negative feedback loops will depend on parameter values, where either stable states are possible or oscillations may be the stable

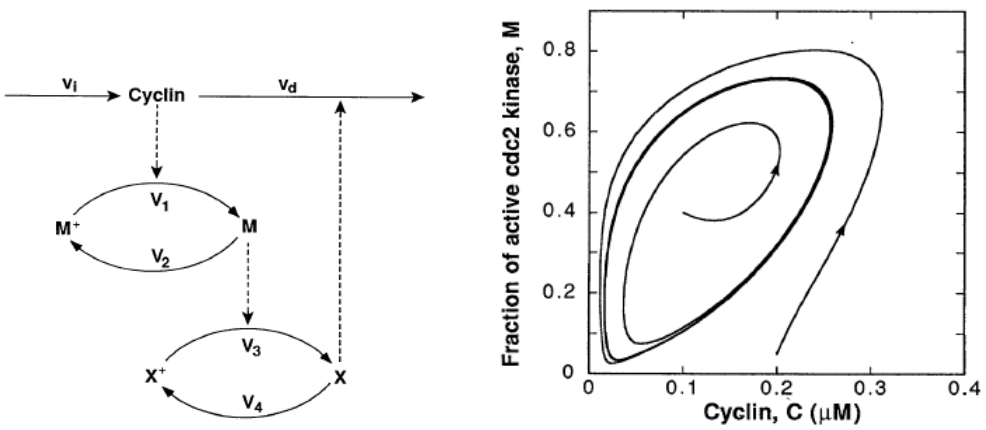


Figure 2: Model network for a cell cycle. Several activation steps are combined with an induced degradation. The model creates oscillations that are displayed in the phase diagram.

behavior. One important feature to get oscillations is to have delays in the interactions.

2.3 Small gene regulatory networks

Construction of small regulatory networks directly within cells, has been made possible in recent years. These systems are typically built within *E. coli* and a main benefit is that analysis of biological parameters, as well as dynamics, is more tractable.

The goal of this exercise is to investigate mathematical models of such systems where bistable and oscillatory dynamics have been engineered into the networks. The full network is a three component repressing network, termed the repressilator (Fig. 3, [4]). It is constructed with three genes that are asymmetrically repressing each other, where gene x_1 represses gene x_2, x_2 represses x_3, and x_3 represses x_1, forming a three node ring. In the process of building the model for this network, we will also analyze the smaller versions of repressor networks, with an one node auto-repressor (gene x_1 represses itself), and a two node repressor network, where x_1 represses x_2, and x_2 represses x_1 (Fig. 4, [5]).

2.4 Spatial pattern formation

A main question in developmental biology is how a group of homogeneous cells can differentiate into spatially heterogeneous cell types. This relates to early differentiation steps forming the base for different tissues as well as for 'global' patterns such as symmetric initiation of organs and visual patterns (Fig. 5). The ideas of reaction-diffusion models where introduced by A. Turing in the 1950s

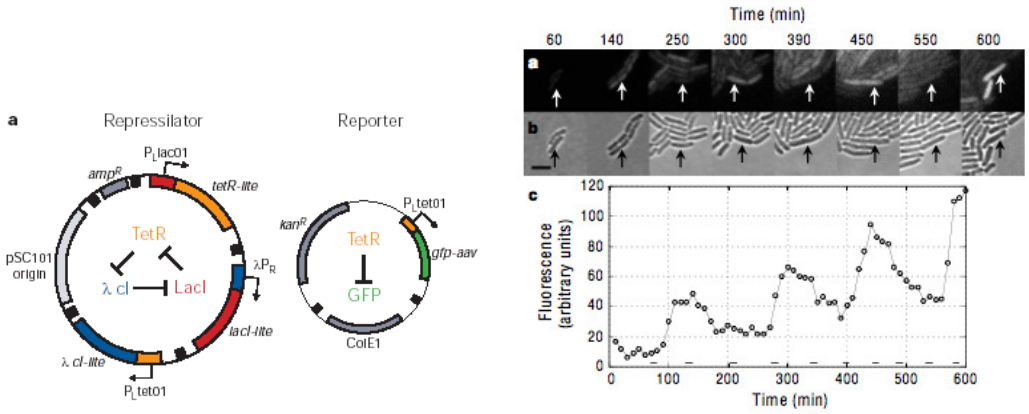


Figure 3: Network built into the *E coli* where three genes are repressing each other. The GFP readout in from the experiment shows oscillations not correlating with the bacteria cell cycle.

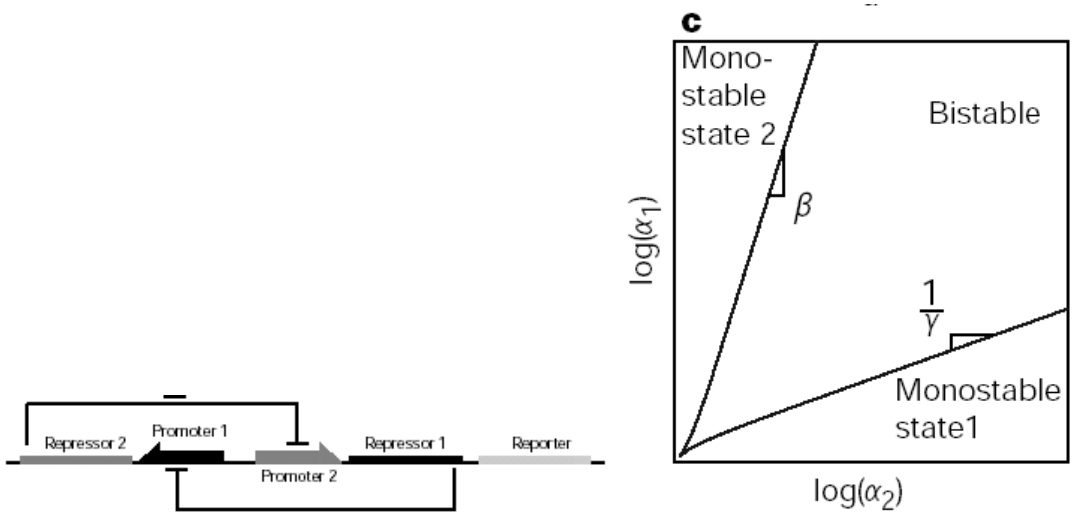
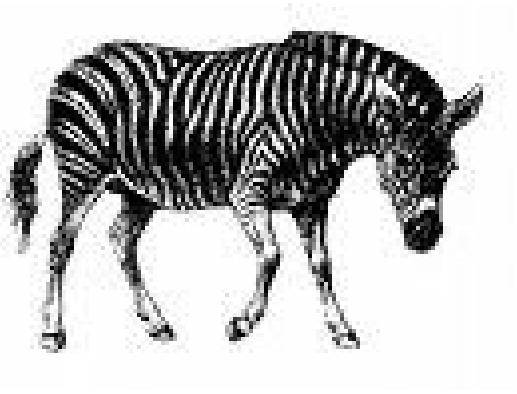


Figure 4: Network built into an *E coli* bacteria. The phase diagram shows how different parameter values in a model of the system leads to different kinds of dynamics with either a single fixed point or bistability.

A)   B)

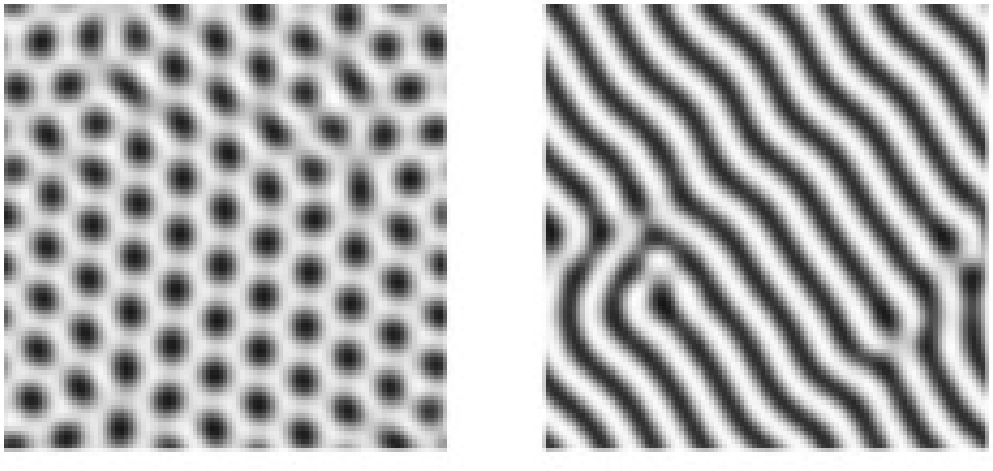


Figure 5: A) Examples of patterns found in nature. B) Simulations of a reaction-diffusion model with different parameters.

[6]. He showed how intracellular reactions combined with intercellular diffusion could lead to spontaneous spatial pattern formation in molecular (morphogen) concentrations. Meinhardt and co-workers has analysed these models in detail and showed that a local activation together with a long-range inhibition is a common mechanism in this type of models, and that depending on parameter values different types of patterns can be found (Fig. 5). In this exercise we will use the *brusselator* model, since its equations can be derived from mass action mechanisms.

3 Simulation of a biochemical network with MATLAB

Simulation of a biochemical network with MATLAB software (Mathworks, Natick, MA) involves numerical solving of dynamical equations.

3.1 Ordinary differential equations

MATLAB allows to solve initial value problems for ordinary differential equations by the family of *ode* solvers. For details see:

`http://www.mathworks.se/help/techdoc/ref/ode23.html`

In particular, the solver *ode*45 allows to solve nonstiff differential equations by Runge-Kutta method. Its syntax is:

```
[T Y] = ode45(odefun,tspan,y0,optionsODE,struct);
```

Where:

- `odefun` is the name of the function where the differential equations are declared.
- `tspan` is a vector specifying the interval of integration.
- `y0` is a vector specifying the initial conditions.
- `optionsODE` contains the optional parameters that change the default integration properties.
- `struct` is a structure containg additional parameters.
- `T` is a column vector of time points.
- `Y` is the solution array. Each row in `Y` corresponds to the solution at a time returned in the corresponding row of `T`.

3.2 Example

As an example, let us implement the bistable erythroid-myeloid switch presented in the beginning [2]. The dynamical equations of GATA-1 ([G]) and PU.1 ([P]) are:

$$\frac{d[G]}{dt} = \frac{\alpha_1 + \alpha_2 \cdot [G]}{1 + \beta_1 + \beta_2 \cdot [G] + \beta_3 \cdot [G] \cdot [P]} - \gamma_1 \cdot [G] \tag{1}$$

$$\frac{d[P]}{dt} = \frac{\delta_1 + \delta_2 \cdot [P]}{1 + \epsilon_1 + \epsilon_2 \cdot [P] + \epsilon_3 \cdot [G] \cdot [P]} - \gamma_2 \cdot [P] \tag{2}$$

The parameter values are: $\alpha_1 = 0.001$, $\alpha_2 = 0.25$, $\beta_1 = 0.001$, $\beta_2 = 0.25$, $\beta_3 = 1$, $\gamma_1 = 0.01$, $\delta_1 = 5$, $\delta_2 = 0.25$, $\epsilon_1 = 5$, $\epsilon_2 = 0.25$, $\epsilon_3 = 1$, $\gamma_2 = 0.01$. These parameters will be defined in a structure called `struct`. The maximum value for `tspan` is set to 2000. The vector of initial conditions `y0` is [0 0]. In the `optionsODE` we decleare the absolute error tolerance `AbsTol` that apply to the individual components of the

solution vector as 10^{-8} for both genes, while the relative error tolerance `RelTol` that applies to all components of the solution vector `Y` as 10^{-8}.

First of all, we create a function called `bistableSwitch` (file `bistableSwitch.m`) where the differential equations are decleared:

```
function dY = bistableSwitch(time,Y,FLAG,struct)

%space allocation for the derivative
dY=zeros(2,1);

%Y(1) is GATA-1
dY(1)=(struct.alpha1+struct.alpha2*Y(1))/...
    (1+struct.beta1+struct.beta2*Y(1)+struct.beta3*Y(1)*Y(2))-...
    struct.gamma1*Y(1);

%Y(2) is PU.1
dY(2)=(struct.delta1+struct.delta2*Y(2))/...
    (1+struct.epsilon1+struct.epsilon2*Y(2)+struct.epsilon3*Y(1)*Y(2))-...
    struct.gamma2*Y(2);
```

This function is called in the main text (file `simulationGATA1PU1.m`) where the system is solved. That is:

```
%Program to simulate the erythroid-myeloid switch

%parameters in a structure
struct.alpha1=0.001;
struct.alpha2=0.25;
struct.beta1=0.001;
struct.beta2=0.25;
struct.beta3=1;
struct.gamma1=0.01;
struct.delta1=5;
struct.delta2=0.25;
struct.epsilon1=5;
struct.epsilon2=0.25;
struct.epsilon3=1;
struct.gamma2=0.01;

%set maximum tspan
tmax=2000;

%call ode45 to solve the system of differential equations
```

```
%first gene is GATA-1 (initial condition = 0)
%second gene is PU.1 (initial condition = 0)
%set the options for ode45
optionsODE=odeset('AbsTol',[1e-8 1e-8],'RelTol',1e-8);
[T Y]=ode45('bistableSwitch',[0 tmax],[0 0],optionsODE,struct);

GATA1=Y(:,1);
PU1=Y(:,2);

%plot the gene profiles versus time
figure(1)
plot(T,GATA1,'-b',T,PU1,'-r');
xlabel('time');
ylabel('expression level');
legend('GATA-1','PU.1');
```

3.3 3D plotting

To produce 3D plot with MATLAB the functions `mesh(X,Y,Z)` or `imagesc(X,Y,Z)` can be used. The first draws a wireframe mesh with color determined by `Z`, so color is proportional to surface height. The second draws `Z` as a heat map. Please note that if `X` and `Y` are vectors of length `n` and `m` respectively, the size of `Z` must be `(m,n)`.

3.4 Literature

The lecture notes may be of help for solving some of the problems. Scientific papers describing the models used in this exercise can also be found at:
`http://www.thep.lu.se/~henrik/tek292/`.

4 Models and Exercises

4.1 Michaelis-Menten

Assume a description of an enzyme reaction as

$$A + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AE \overset{k_+}{\rightarrow} B + E$$

(A1) Write down the differential equations for all molecules. Derive the Michaelis-Menten equation of this process and describe which assumptions you are using. How does the production of [B] depend on the concentrations [A] and [E]?

4.2 Small gene regulation networks

We will use a Hill-type of equation to describe the regulatory repression. In addition to this there is a simple degradation term, resulting in a general system of equation(s) for gene(s) x_i

$$\frac{dx_i}{dt} = \frac{V_i K_i^{n_i}}{K_i^{n_i} + x_{i\oplus 1}^{n_i}} - d_i x_i \tag{3}$$

where the parameter V_i is the transcription rate for unbound DNA, K_i is the Hill constant, related to the repressor concentration leading to half of the maximal production rate. The Hill coefficient, n_i, sets the steepness of the function. Finally, d_i is the degradation rate. The $\oplus$ is a normal plus but modulated at the largest i value to be 1 (i.e. $(i, i \oplus 1)$ equals (1,1) for a single gene system, (1,2) and (2,1) for a two-gene system, and (1,2), (2,3), and (3,1) for the three-gene case).

4.2.1 One gene auto-repression

First, we will investigate the single gene auto-repressor ($x_i = x$). For simplicity, we set some of the parameters to some specific values ($d = n = 1$), leading to a Michaelis-Menten formalism, similar to what has been discussed in the lectures

$$\frac{dx}{dt} = \frac{VK}{K + x} - x. \tag{4}$$

(A2) Describe in words how the production term in Eq. 4 relates to the derivation of Michaelis-Menten done in **(A1)**. Find possible equilibrium states for Eq. 4 and analyse the dynamics.

(C1) Implement the one gene auto-repression model with MATLAB considering parameters from Eq. 4 and simulate the network for different initial concentrations of x. How does the system behave? How does changing the paramters V and K alter the dynamics?

4.2.2 A two gene repressor network, the bistable switch

Now it is time to add a gene and look at the two-gene repressor network we discussed in the lectures [5]. Again, some of the parameters will be fixed during the simulations ($K = d = 1$, for both proteins), resulting in the equations

$$\frac{dx_1}{dt} = \frac{V_1}{1 + x_2^{n_1}} - x_1 \tag{5}$$

$$\frac{dx_2}{dt} = \frac{V_2}{1 + x_1^{n_2}} - x_2 \tag{6}$$

(A3) Find the null-clines for both proteins ($dx_i/dt = 0$). Plot these functions and analyze the behavior in different regions and stability of fixed-points for the two cases $n_1 = n_2 = 1$ and $n_1 = n_2 = 2$ (you may set V_1 and V_2 to values close to 2). Find possible fixed-points for the case $n_1 = n_2 = 1$ analytically.

(C3) Implement the bistable switch with MATLAB for the cases investigated in **(A3)** considering different initial configurations (x_1, x_2) (e.g. (0,0),(0,2),(2,0),(2,2)). Plot the output in the x_2 *versus* x_1 plot (including the null-clines). Describe the dynamics. What happens with a completely symmetric model and initial configuration? How can this network be used to make a decision?

4.2.3 A three gene repressor network, the repressilator

We are now ready to add another protein, resulting in a three gene repressor network termed the repressilator. The model was presented in a paper by Elowitz *et al.* [4], where they studied the system both *in vivo*, by introducing a small genetical network in E.coli, and *in silico*, by performing simulations simular to what you will do in the remaining part of this exercise. The original model included both mRNA and proteins in the dynamics, but here we will simplify matters by only considering mRNA. This results in the following three equations.

$$\frac{dx_1}{dt} = \frac{V_1}{K_1 + x_2^{n_1}} - x_1 \tag{7}$$

$$\frac{dx_2}{dt} = \frac{V_2}{K_2 + x_3^{n_2}} - x_2 \tag{8}$$

$$\frac{dx_3}{dt} = \frac{V_3}{K_3 + x_1^{n_3}} - x_3 \tag{9}$$

The function `repressilator` implements Eqs. 7- 9, while the program `simulationRepressilator` allows the simulation of the network.

(C4) Simulate the model for a number of different initial conditions and analyze the dynamics by plotting $x_i(t)$ and phase planes (x_i versus x_j).

(C5) Try to find an oscillatory behavior by adjusting the parameters. Plot the results. What is typically needed to get oscillations? Can you tune the amplitude or period of the oscillations? How does this network relate to the cell cycle network described in Fig. 2.

4.3 Reaction-diffusion

We will introduce a diffusion-like transport between cells lying on a line, i.e cell i is neighbor with cell $i-1$ and $i+1$, where i is the cell index. The transport is defined in the model by

$$\frac{dc_i}{dt} = D(c_{i-1} - 2c_i + c_{i+1}) \tag{10}$$

where c_i is the concentration in cell i.

4.4 The Brusselator

The Brusselator reaction network consistes of the reactions:

$$A \xrightarrow{k_1} X$$
$$2X + Y \xrightarrow{k_2} 3X$$
$$B + X \xrightarrow{k_3} Y + C$$
$$X \xrightarrow{k_4} D.$$

(A4) Write down the differential equations describing the time evolution of [X] and [Y].

The program `simulationBrusselator` allows the simulation of the network.

(C6) Run the `simulationBrusselator` program for one cell (set `N` to 1) and different parameter values (set `minX` and `maxX` to the same desiderated value; similarly set `minY` and `maxY` to the same desiderated values). Investigate the behavior. Can you get the system to oscillate?

(C7) Increase the number of cells (e.g. set `N` to 100) and use random initial concentrations for X and Y, and parameters leading to oscillations from **(C6)**. Describe the difference between different cells. Introduce diffusion for X (`Dx`). What happens with different cells?

(C8) Use small random deviations in initial concentrations of X and Y. Vary the parameters including the two diffusion parameters `Dx` and `Dy` and describe different behaviors. Can you relate the behavior to spatial differentiation patterns found in biological organisms? How does your model parameters relate to Meinhardts description of the behavior as a local activation and a long-range inhibition?

5 Additional exercises

5.1 Repressilator for populations of cells

(C9) Simulate a repressilator model with two cells (modify the program `simulationRepressilator` and the function `repressilator`). Use parameters leading to oscillations and somewhat different initial concentrations in the two cells. Explain what happens.

(C10) Add diffusion for x_1. What happens?

6 Giudelines for the report

The report should be formulated as a scientific paper including: introduction (description of the problem in general terms providing background information), theory (description of the theory related to the problem including answers to analytical exercises), implementation (description of the code used to implement a network—where applicable), results (description of the results obtained from different simulations), discussion and conclusions (explanation of the results, comparisons, "take home" message, $\cdots$).

The reports must be in English and submitted electronically as a pdf file to the supervisor e-mail address.

It is allowed to collaborate with colleagues during the exercises, but the report must be individual. In the report you are allowed to use information from any source you prefer, but a reference should be given for each cited statement. Any material that is not original and not corrected referenced will be recognized and report to the university disciplinary council.

References

[1] Alon U (2007) An introduction to system biology: design principles of biological circuits. Chapman and Hall.

[2] Chickarmane V, Enver T, Peterson C (2009) Computational modeling of the hematopoietic erythroid-myeloid switch reveals insights into co-operativity, priming and irreversibility. PLoS Comput Biol 5: e1000268.

[3] Goldbeter A (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. Proc Natl Acad Sci USA 88: 9107-9111.

[4] Elowitz M, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335-338.

[5] Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in escherichia coli. Nature 403: 339-342.

[6] Turing AM (1952) The chemical basis for morphogenesis. Philos Trans R Soc London B 237: 3772.