Dynamic model of gene regulation for the lac operon

Maia Angelova

Intelligent Modelling Lab, School of Computing, Engineering and Information Sciences, Northumbria University, Newcastle upon Tyne NE2 1XE, UK

E-mail: maia.angelova@northumbria.ac.uk

Asma Ben-Halim

Intelligent Modelling Lab, School of Computing, Engineering and Information Sciences, Northumbria University, Newcastle upon Tyne NE2 1XE, UK

E-mail: asma.benhalim@northumbria.ac.uk

Abstract. Gene regulatory network is a collection of DNA which interact with each other and with other matter in the cell. The *lac* operon is an example of a relatively simple genetic networks and is one of the well-studied and best-understood structures in the Escherichia coli bacteria. In this work we consider a deterministic model of the *lac* operon with a noise term (representing the stochastic nature of the regulation). The model is written in terms of a system of simultaneous first order differential equations with delay. We investigate an analytical and numerical solution and analyse the range of values for the parameters corresponding to a stable solution. We compare our results with other works on the *lac* system.

1. Introduction

In living organisms the genome plays a central role in controlling the cellular processes, such as the response of a cell to environmental signals, the differentiation of cells and groups of cells in the unfolding of developmental programmes, the replication of the DNA preceding the cell division and many others [1, 2]. Proteins synthesized from genes may function as transcription factors binding to regulatory sites of other genes, as enzymes catalyzing metabolic reactions, or as components of signal transduction pathways. The degradation of proteins and immediate DNA products can also be regulated in the cell. The proteins involved in the above regulatory functions are produced by other genes. This gives rise to genetic regulatory systems, structured by networks of regulatory interactions between DNA, RNA, proteins and small molecules. A gene regulatory network is a collection of DNA which interact with each other and with other matter in the cell. A simple network consists of one or more input gene, metabolic, and signaling pathways, regulatory proteins that integrate the input signals, several target genes, and the RNA and proteins produced from those target genes.

The *lac* operon [3] is an example of a relatively simple genetic network and is one of the well-studied and best-understood structures in the Escherichia coli (*E.coli*). It consists of a promoter, and operator region and three larger structural genes, *lacZ*, *lacY* and *lacA*, with a preceding regulatory operon responsible for producing a repressor (R) protein. In the absence of glucose available for cellular metabolism, but in the presence of external lactose (L_e), lactose is transported into the cell by a permease (P). Intracellular lactose (L) is then broken down

into allolactose (A) first and then glucose and galactose by the enzyme β-galactosidase (B). The allolactose feeds back to bind with the lactose repressor and enable the transcription process to proceed.

A number of mathematical models of the *lac* operon have been developed (see for example [2, 4, 5, 6, 7, 10]. In this work we consider a deterministic model of the lac operon with a noise term representing the stochastic nature of the regulation. The model is written in terms of a system of simultaneous first order differential equations with delay. We investigate an analytical and numerical solution and analyse the range of values for the parameters corresponding to a stable solution. We compare our results with other works on the lac system.

2. The deterministic model

Deterministic models have been widely used to analyse genetic regulatory systems [1, 2, 4]. On many occasions they are described by a system of ordinary differential equations (ODE). The ODEs formalism models the concentration of RNAs, proteins and other molecules by time dependent variables. The interactions in the network have a form of functional and differential relations between concentrations. The network dynamics can be described by Michaelis-Menten enzyme kinetics [4]. Note that the deterministic model, based on differential equations, describes an average response of the system. It assumes that the concentration varies continuously and deterministically, both of which assumptions may be questionable in case of gene regulation particularly of small systems, as the small number of molecules in such networks compromises such assumption. For large systems there is a large number of gene responses and the gene regulatory network can be realistically described with a set of deterministic differential equations.

Time delays are common and substantial in biochemical processes.They are essential for the system as they can protect it against transient loss of input signal, improve the accuracy of reading the information and filter non-beneficial pulses. However, time delays are not always beneficial as they may play a negative role in the stability of the gene network.

First, we will consider the deterministic model (without noise) following the works by Mackey *et al* [7, 8, 9, 10]. The model is written in terms of a system of simultaneous first order ODEs with time delays (DDEs) and is based on Michaelis-Menten enzyme kinetics. The model [7] consists of five DDEs describing the lactose system in *e.coli* of positive feedback,

$$\begin{aligned}
\frac{dM}{dt} &= \alpha_M f(A_{\tau_M}) + \Gamma_0 - \widetilde{\gamma}_M M, \\
\frac{dB}{dt} &= \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \widetilde{\gamma}_B B, \\
\frac{dA}{dt} &= \alpha_A B h(L) - \beta_A B g(A) - \widetilde{\gamma}_A A, \\
\frac{dL}{dt} &= \alpha_L P h'(L_e) - \beta_{L_1} P h(L_1) - \beta_{L_2} B h(L_2) - \widetilde{\gamma}_L L, \\
\frac{dP}{dt} &= \alpha_P e^{-\mu(\tau_P+\tau_B)} M_{\tau_P+\tau_B} - \widetilde{\gamma}_P P.
\end{aligned}$$

Here M is the mRNA concentration, n is the number of molecules of allolactose required to inactivate the repressor. Two time delays are included in this system, τ_M and τ_P. The factor $e^{-\mu\tau_M}$ accounts for the dilution of A through growth during the transcriptional period, where μ is the rate of degradation. To consider the system without time delays, one can choose these parameters to be 0. The parameters of the system, $K_i, i = 1, 2, L, L_e$, $\alpha_j, j = A, B, L, M, P$ and $\beta_k, k = A, L, L_1, L_2$ are deterministic mass-action kinetic rate constants. K_1 is the equilibrium constant for the repressor-allolactose reaction, K_2 is the equilibrium constant for the operator-repressor reaction, $K = 1+K_2R_{tot}$, and R_{tot} is the total amount of repressor. The rate of change

of M is a balance between a production term $\alpha_M f$, where

$$f(A_{\tau_M}) = \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} \tag{1}$$

$$h(L) = \frac{L}{K_L + L}, \quad g(A) = \frac{A}{K_A + A}. \tag{2}$$

The loss of rate of B is given by $\widetilde{\gamma}_B B$. Similarly, the loss of rate of M, A, L and P is represented by $\widetilde{\gamma}_M$,$\widetilde{\gamma}_A$, $\widetilde{\gamma}_L$ and $\widetilde{\gamma}_P$ respectively. The values of these parameters as well as parameters α_M , α_B , α_A, α_L and α_P are estimated experimentally in [11].

This model is difficult to solve analytically and a simplified model was proposed in [8] with a constant quantity of lactose inside the cell, $L = const$ (equilibrium of internal and external cellular lactose). Ignoring the dynamics of permease, $P = const$ (assumed as a constant permease concentration), leads to a system of three delayed DDEs with two delays,

$$\frac{dM}{dt} = \alpha_M f(A_{\tau_M}) - \widetilde{\gamma}_M M, \tag{3}$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \widetilde{\gamma}_B B, \tag{4}$$

$$\frac{dA}{dt} = \alpha_A B h(L) - \beta_A g(A) - \widetilde{\gamma}_A A. \tag{5}$$

The model has been tested the estimates for parameters of Ackers *et al* [11]. Details of the numerical and analytical solution is given in [8].

At the steady state point (M_*, B_*, A_*), the delay differential are determined for $\tau = 0$, and

$$\frac{dM}{dt} = 0, \quad \frac{dB}{dt} = 0, \quad \frac{dA}{dt} = 0.$$

Thus, a definition of the steady state is given by

$$f(A_*) = \theta \frac{A_*}{\left[h(L) - \frac{\beta_A}{\alpha_A} g(A_*)\right]}, \tag{6}$$

where $\theta = \frac{\widetilde{\gamma}_M \widetilde{\gamma}_B \widetilde{\gamma}_A}{\alpha_M \alpha_B e^{-\mu\tau_B} \widetilde{\gamma}_M \widetilde{\gamma}_B}$ The solution can be found graphically by plotting the left and right hand sides of (1) and given on Figure 1. The dashed curve represent the left hand side (LHS), and other three curves represents the right hand side (RHS) of 1. The intersection between the two curves (RHS and LHS) gives the location of the steady state. The range $40\mu M \leq L \leq 55.4\mu M$ has been used as in this range three steady state exist. The range of allolactose concentration depends on $h(L) - \frac{\beta_A}{\alpha_A} g(A_*) \succ 0$.

β-galactosidase regulatory pathway is the most essential of the regulatory mechanisms in the lactose operon [3]. The behaviour of the β-galactosidase levels as a function of time are simulated. The solution showed that the steady states display bistability. The steady states are depending on the extracellular lactose concentration and growth rate. We have further investigated the analytical and numerical solution of this simplified model and analysed the range of values for the parameters corresponding to a stable solution. The stability is using Routh-Huirwitz conditions [4].

Figure 9 presents the stable and unstable regions of the $L - A$ space. The S -shaped curve gives three steady states obtained from 1, the range for L is [40, 55.4]μM , the straight line is obtained from (**??**), and the third curve is from equation $\eta_0 + \vartheta < 0$ and separates the region where states are unstable. For the internal lactose concentration, when $L = 45\mu M$ we have

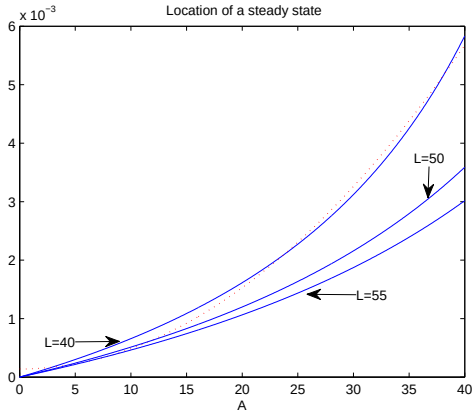

Figure 1. Steady state solutions of $f(A)$ for three values of the lactose.

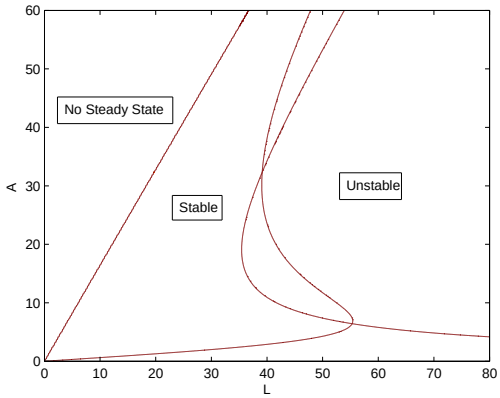

Figure 2. Numerical solution showing stable and unstable regions in $L - A$ space.

found three steady states: $A_* = 54.8 \quad 5.56 \quad 11.0256$ $B_* = 0.0119 \quad 0.00027940 \quad 0.00063667$ $M_* = 0.0237 \quad 0.00055673 \quad 0.0013$ MATLAB dde23 routine and symbolic toolbox were used for the numerical solutions. The bistability of the β-galactosidase versus time is shown on Figure 3.

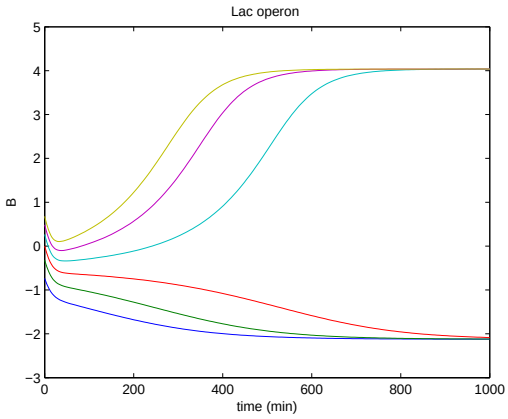

Figure 3. Bistability of β-galactosidase.

3. Stochastic model

Gene expression is a complex process regulated at several stages in the synthesis of proteins. Apart from the regulation of DNA transcription, RNA translation and post-translational modification of proteins, gene expression is also part of the regulatory process. Gene regulation is a stochastic process as first it happens in a living organism and there are many fluctuations due to delay of response and other processes taking place in the same time.

For small systems, with a small number of genes (such as the *lac*), a stochastic model is more appropriate, as the responses are few and their nature is probabilistic. However, a compromise can be reached by using a deterministic model and adding a noise term to account for the stochastic behaviors of the system. In this section we'll apply the Langevin approach by introducing a noise term [1, 2, 6, 4].

We have add noise term for the continuous time differential equations, if

$$\frac{dy}{dt} = f(y)$$

To add noise

$$\frac{dy}{dt} = f(y) + \sigma\frac{dW}{dt}$$

where W is the vector of Winer processes, σ is the standerd deviation. Multiply both sides by dt to get stochastic differential equation (SDE).

$$dy = f(y)dt + \sigma dW$$

The simplest stochastic numerical approximation is the Euler Maruyama method by consider Itô calculus SDE.

Times $h = t_{i+1} - t_i$, where $i = 0, ..., n$

$$y_{i+t} = y_i + f(y_i)h + \sigma Z_i\sqrt{h}$$

where Z_t are distributed random numbers (derivative of a Winer process) with mean 0 and variance 1. This is Eular Maruyama method, and it is possible for higher order methods. We have used this method for the *lac* operon model with out time delay.

$$\begin{aligned}
\frac{dM}{dt} &= \alpha_M\frac{1 + K_1(A)^n}{K + K_1(A)^n} - \widetilde{\gamma}_M M + \lambda, \\
\frac{dB}{dt} &= \alpha_B M - \widetilde{\gamma}_B B + \lambda, \\
\frac{dA}{dt} &= \alpha_A B\frac{L}{K_L + L} - \beta_A B\frac{A}{K_A + A} - \widetilde{\gamma}_A A + \lambda.
\end{aligned}$$

where $\lambda = \sigma Z\sqrt{h}$, apply this equations to the Euler method by using Matlab and with different numbers of σ.

3.1. Steady state analysis

The model of the lac operon is given by three nonlinear ODEs with time delay and parameter gives

4. Discussion and further work

mRNAwithNoiseAndWithNoNoise.eps

References

[1] Bower J and Bolouri H Editors 2001 *Computational Modelling of Genetic and Biochemical Networks* MIT Press:Boston Massachusets.
[2] De Jong H 2002 J Comp Biology 2002, 9, 67-103.
[3] Ptashne M 2004 *A genetic switch* Cold Spring Harbor Laboratory Press: Cold Spring Harbor New York.
[4] Murray J D 2002 *Mathematical Biology* 3rd Edition Volume 1, IAM **17** Springer:NY Berlin.
[5] Goodwin BC 1969 *Eur J Biochem* **10** 515–522.
[6] Mettetal J. T, D. Muzzey J. M., Pedraza E. M., Ozbudak and van Oudenaarden A. 2006. PNAS 103: 7304-7309.
[7] Yildirim N and Mackey M C(2003) *Biophys J* **84** 2841–2851.
[8] Yildirim N, Santillan M and Mackey M C 2004 2004 /it CHAOS **14** 279– .
[9] Mackey M C,Santillan M and Yildirim N 2004 *CR Biologies***327** 211–224.
[10] Santillan M, Mackey M.C. and Zeron E.S. 2007 . Biophysical J. 92: 3830-3842.
[11] Ackers G, Johnson A D and Shea M A 1982 *PNAS* **79** 1129–1133.

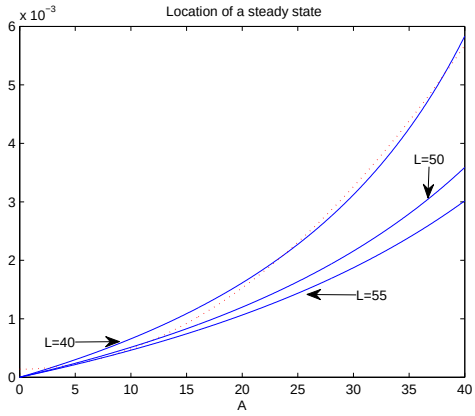

Figure 4.

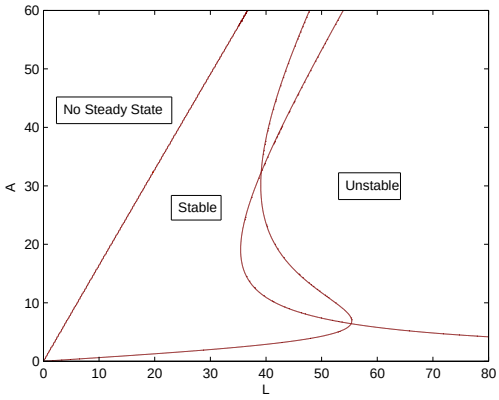

Figure 5. Stable solution.

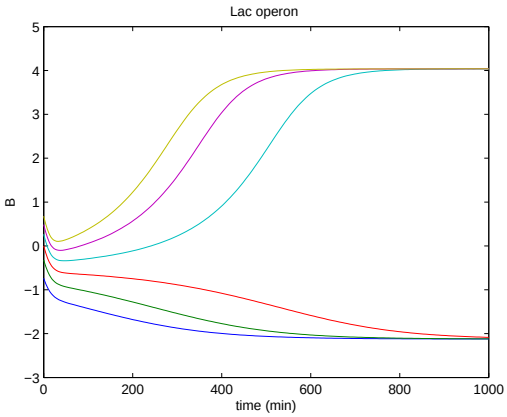

Figure 6. lac

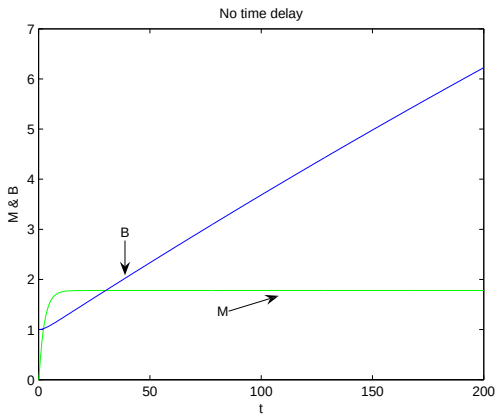

Figure 7. No delay

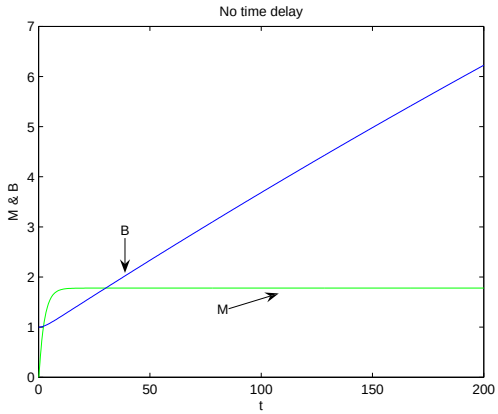

Figure 8. No time delay

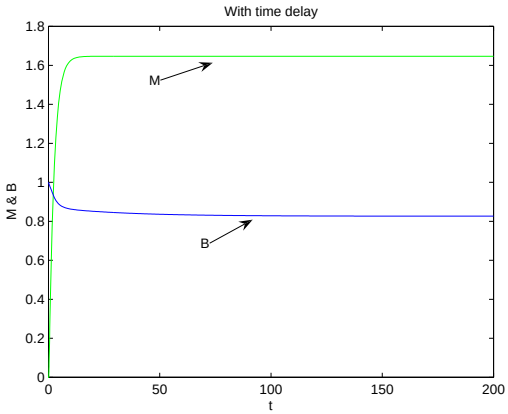

Figure 9. Time delay.

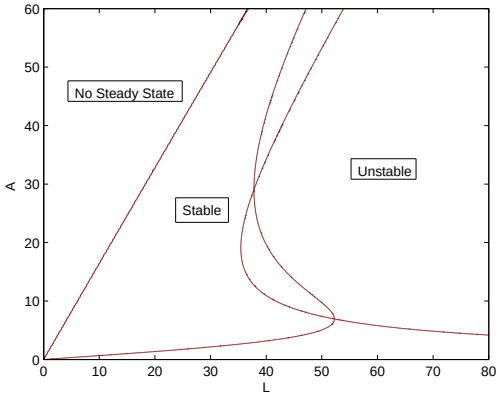

Figure 10. No delay

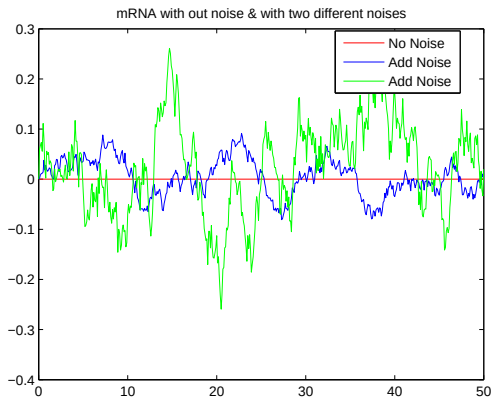

Figure 11. Three

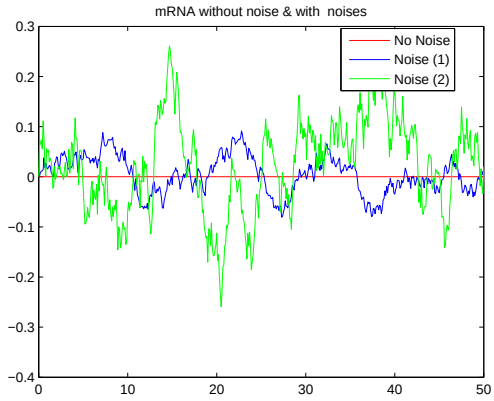

Figure 12. mRNA with noise and with no Noise