Journal of Physics: Conference Series

Dynamic model of gene regulation for the lac operon

To cite this article: Maia Angelova and Asma Ben-Halim 2011 *J. Phys.: Conf. Ser.* **286** 012007

View the article online for updates and enhancements.

Related content

- DNA looping increases the range of bistability in a stochastic model of the lac genetic switch
  Tyler M Earnest, Elijah Roberts, Michael Assaf et al.
- Cost--benefit theory and optimal design of gene regulation functions
  Tomer Kalisky, Erez Dekel and Uri Alon
- Stochastic switching in biology: from genotype to phenotype
  Paul C Bressloff

This content was downloaded from IP address 128.184.188.40 on 18/10/2017 at 04:34

Dynamic model of gene regulation for the lac operon

Maia Angelova and Asma Ben-Halim

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.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 network and is one of the best-studied 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 delays. We investigate an analytical and numerical solution and analyse the range of values for the parameters corresponding to a stable solution.

1. Introduction

Bacterial 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, 3, 4]. 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 the immediate DNA products can also be regulated in the cell. The proteins involved in the regulatory functions are produced by other genes. This gives rise to genetic regulatory network consisting of regulatory interactions between DNA, RNA, proteins and small molecules. 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 target genes.

The *lac* operon [1] 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, lactose is transported into the cell by a permease. 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 [5, 6].

A number of mathematical models of the *lac* operon have been developed (see for example [3, 4, 5, 7, 8, 11, 12]). 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 have investigated the analytical and numerical solutions and analysed the range of values for the parameters leading to stable solutions. The effects of the noise term on the concentrations levels are also presented.

Published under licence by IOP Publishing Ltd

2. Deterministic model

Deterministic models have been widely used to analyse genetic regulatory systems [2, 3, 4]. In many cases they are given 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]. However, the deterministic model based on ODEs, 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 systems with small number of gene responses. 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 ODEs. Time delays are common and substantial in biochemical processes. They are essential for the system and 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 a deterministic model (without noise) following the works of Yildirim and Mackey [8, 9, 10, 11]. 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 full model [8] consists of five DDEs describing the lactose system in *e.coli* of positive feedback. This model is simplified [9], assuming a constant quantity of lactose L inside the cell (equilibrium of internal and external cellular lactose) and ignoring the dynamics of permease (assumed as a constant permease),

$$\frac{dM}{dt} = \alpha_M f(A_{\tau_M}) - \widetilde{\gamma}_M M, \tag{1}$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \widetilde{\gamma}_B B, \tag{2}$$

$$\frac{dA}{dt} = \alpha_A B h(L) - \beta_A B g(A) - \widetilde{\gamma}_A A. \tag{3}$$

Here M is the mRNA concentration, n is the number of molecules of allolactose required to inactivate the repressor R. The model takes into account the delays in the response of mRNA, β-galactosidase and allolactose. The factor $e^{-\mu\tau_M}$ accounts for the dilution of A through growth during the transcriptional period, where μ is the bacterial growth rate, $A_{\tau_M} = A(t-\tau_M)$. The description of the same system without time delays is achieved with $\tau_Y = 0, Y = A, B, M$. The rate of change of M is a balance between a production term $\alpha_M f$,

$$f(A_{\tau_M}) = \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}. \tag{4}$$

The functions $h(L)$ and $g(A)$ do not depend on time delays,

$$h(L) = \frac{L}{K_L + L}, \quad g(A) = \frac{A}{K_A + A}. \tag{5}$$

The parameters of the system, $K_i, i = 1, 2, L$, $\alpha_Y, Y = A, B, M$ and β_A, are deterministic mass-action kinetic rate constants. K_1 is the equilibrium constant for the repressor-allolactose reaction, $K = 1 + K_2 R_{tot}$, and R_{tot} is the total amount of repressor R. The loss of rate of B is given by $\widetilde{\gamma}_B B$. Similarly, $\widetilde{\gamma}_M$ and $\widetilde{\gamma}_A$ represent the loss of rate of M and A respectively. $\widetilde{\gamma}_Y Y = (\gamma_Y + \mu)Y$ is made up of degradation term $\gamma_Y Y$ and effective loss due to dilution μY for $Y = M, B, A$. The values of these parameters are estimated experimentally in [12].

The model was tested using the estimates for the parameters in [12]. Details of the numerical and analytical solution are given in [9]. The solutions indicate that the steady states depend on the intracellular lactose concentration L and growth rate μ. 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 using Routh-Hurwitz criterium [4]. The numerical solutions are obtained with Matlab (using Euler and Rounge-Kutta methods, dde23 routine and symbolic toolbox).

The steady state point (M_*, B_*, A_*), is depicted when the derivatives of M, B and A are equal to zero, $\frac{dM}{dt} = 0$, $\frac{dB}{dt} = 0$, $\frac{dA}{dt} = 0$. Applied to equations (1)-(3), this gives a condition for the steady state,

$$f(A_*) = \theta \frac{A_*}{\left[h(L) - \frac{\beta_A}{\alpha_A} g(A_*)\right]}, \quad \theta = \frac{\widetilde{\gamma}_A \widetilde{\gamma}_B \widetilde{\gamma}_M e^{\mu \tau_B}}{\alpha_A \alpha_B \alpha_M}. \tag{6}$$

A graphical solution for $n = 2$ is illustrated on Figure 1 with initial conditions $A = M = 0, B = 1$. The dashed curve represents the left hand side, $f(A_*)$, of (6), while the right hand side, which is a monotonically increasing function of L, is plotted for three values of L. Steady state exists for intracellular lactose in the range $40\mu\text{M} \leq L \leq 55.4\mu\text{M}$. The location of the steady state is given by the intersection of the curves representing each side of (6). Depending on the value of L, one, two or three steady states are possible for $n = 2$. For example, three steady states exist for $L = 40\mu\text{M}$. The range of allolactose concentration is always positive, thus $h(L) - \frac{\beta_A}{\alpha_A} g(A_*) > 0$, which bounds the region where no steady states exist.

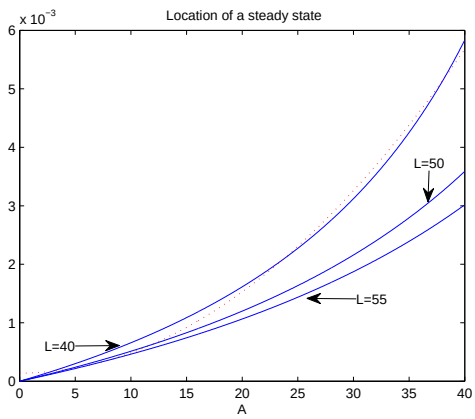

Figure 1. Steady states obtained by graphical solution of (6), A (μM).

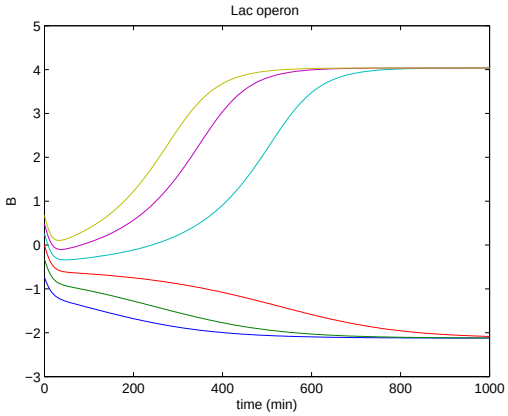

Figure 2. Bistability of β-galactosidase (μM) as a function of time (min) for various initial values of B.

β-galactosidase regulatory pathway is the most essential of the regulatory mechanisms in the lactose operon [1, 5, 6]. The concentration of β-galactosidase as a function of time was simulated and shows (Figure 2) that the steady states display bistability, in agreement with [8, 9].

The concentrations of mRNA and β-galactosidase are plotted on Figures 3 and 4 without time delay and with time delay respectively with initial values $M = 0, B = 1$. mRNA levels reach saturation faster then β-galactosidase levels. B is more sensitive to time delays, which also reduce the levels of concentration.

3. Stochastic correction

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

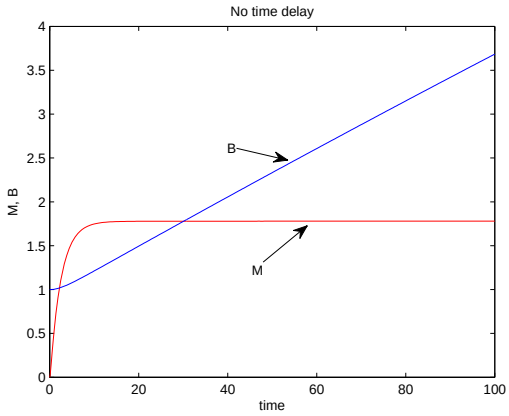

Figure 3. M (μM) and B (μM) versus time (min) without time delay.

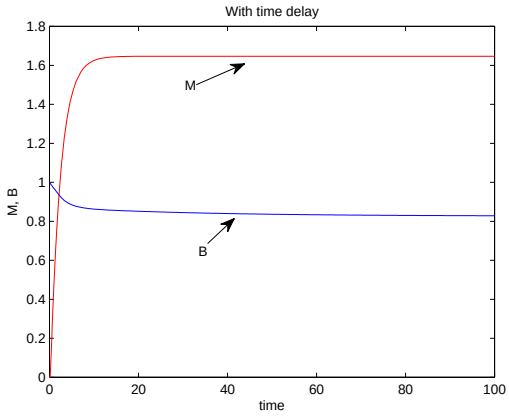

Figure 4. M (μM) and B (μM) versus time (min) with time delay.

modification of proteins, gene expression is also part of the regulatory process. Gene regulation is a stochastic process as it happens in living organisms and there are many fluctuations due to delay of response and other processes taking place at the same time. For small systems such as the *lac*, a stochastic model is more appropriate, as the gene responses are few and their nature is probabilistic. We have adjusted the deterministic model using the Langevin approach [3, 7] by adding a noise term to equations (1)-(3) to account for the stochastic behavior of the system.

Let $Y = M, B, A$ and F is the function representing the right hand side of the equations (1)-(3),

$$\frac{dY}{dt} = F(Y). \tag{7}$$

We add a noise term,

$$\frac{dY}{dt} = F(Y) + \sigma\frac{dW}{dt} \tag{8}$$

where W is independent random variable of Wiener process, which satisfies the conditions of this process, namely $W_t - W_0 \sim \mathcal{N}(\mu\sigma^2)$ [14]. Here $\mathcal{N}(\mu\sigma^2)$ is a normal distribution with expected value (mean) μ and variation σ^2. μ and σ are also known as drift and volatility respectively. Thus, the stochastic differential equation (SDE) is

$$dY = F(Y)dt + \sigma dW \tag{9}$$

We have used Euler-Maruyama method, which is a technique for approximate numerical solution of SDE, and is a simple generalisation of the Euler method [15]. Let $h = t_{i+1} - t_i$, where $i = 0, ..., N$,is the iterations step and N is the number of iterations. Then

$$Y_{i+t} = Y_i + F(Y_i)h + \sigma Z_t\sqrt{h} \tag{10}$$

where Z_t are normally distributed random numbers (derivatives of a Wiener process) with mean 0 and variance 1. Applying to the model,

$$\frac{dM}{dt} = \alpha_M f(A_{\tau_M}) - \widetilde{\gamma}_M M + \lambda, \tag{11}$$

$$\frac{dB}{dt} = \alpha_B M - \widetilde{\gamma}_B B + \lambda, \tag{12}$$

$$\frac{dA}{dt} = \alpha_A B h(L) - \beta_A B g(A) - \widetilde{\gamma}_A A + \lambda. \tag{13}$$

where $\lambda = \sigma Z_t \sqrt{h}$. These equations were solved with and without time delays. The solutions for M and B, with initial conditions $M = A = 0$ and $B = 1$, and different noise terms are illustrated in Figures 5 and 6 respectively. The fluctuations of M and B are clearly present, as expected from the stochastic nature of the process. The levels of concentration are sensitive to the noise. The fluctuations of mRNA concentration are above and below mRNA levels without noise, while β-gallactosidase concentration mainly decreases with time in the presence of the noise term.

The effect of the stochasticity was investigated in [13], where different noise terms were investigated experimentally and by simulation and the behaviour of the deterministic and stochastic models compared. The conclusion of this work is that it is still unclear whether the stochastic gene expression in the *lac* operon is detrimental or beneficial for the cells.

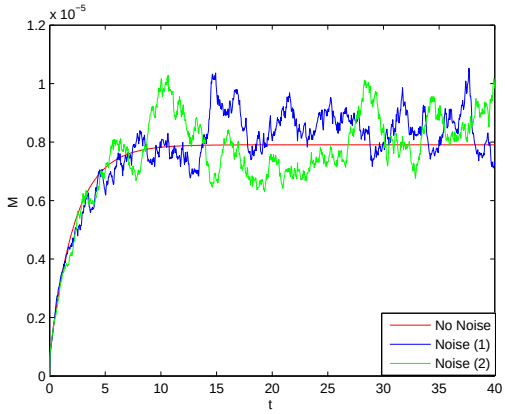

Figure 5. mRNA with different noise, M (μM).

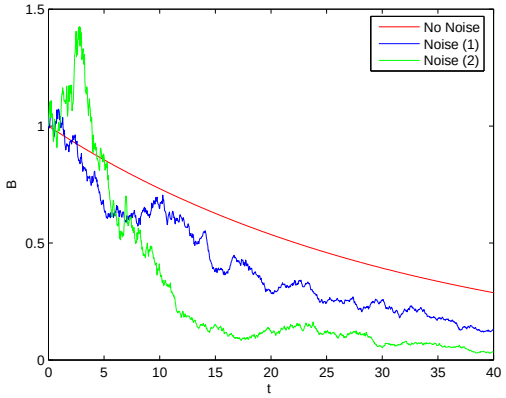

Figure 6. β-gallactosidase with different noise, B (μM).

References

[1] Ptashne M 2004 *A genetic switch* Cold Spring Harbor Laboratory Press: Cold Spring Harbor New York.
[2] Bower J and Bolouri H Editors 2001 *Computational Modelling of Genetic and Biochemical Networks* MIT Press:Boston Massachusets.
[3] De Jong H 2002 J Comp Biology 2002, 9, 67-103.
[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] Yagil G and Yagil E 1971 *Biophys J* **11** 11–27.
[7] Mettetal J. T, D. Muzzey J. M., Pedraza E. M., Ozbudak and van Oudenaarden A. 2006. PNAS 103: 7304-7309.
[8] Yildirim N and Mackey M C(2003) *Biophys J* **84** 2841–2851.
[9] Yildirim N, Santillan M and Mackey M C 2004 /it Chaos **14** 279–291.
[10] Mackey M C,Santillan M and Yildirim N 2004 *CR Biologies***327** 211–224.
[11] Santillan M, Mackey M.C. and Zeron E.S. 2007 . Biophysical J. 92: 3830-3842.
[12] Ackers G, Johnson A D and Shea M A 1982 *PNAS* **79** 1129–1133.
[13] van Hoek M and Hogeweg P 2007 *PLoS Comput Biol* **3**(6) doi:10.1371/journal.pcbi.0030111.
[14] Stark H and Woods J W 2002 *Probability and Random Processes with Applications to Signal Processing* Prentice Hall: New Jersey.
[15] Kloeden P E and Platen E 1999 Numerical Solutions of Stochastic Differential Equations Springer:Berlin,