A Dynamic Multiple Input Description of Lactose Operon Expression

[Weijie Wang, Shaoping Wang]

Abstract—**Modeling the regulation of lactose operon expression is significant to understand lactose metabolism at the cell-level, in which the repressor and cAMP-CRP can influence its regulation process. In this paper, a mathematic model of lactose operon expression is established in differential equations with the knowledge of system biology. Considering multiple factors such as lactose, glucose, repressor, permease, β-galactosidase, allolactose and cAMP-CRP, this paper presents dynamic bi-direction regulatory mechanisms in both repressor and cAMP-CRP with the logical "AND" of system biology. The experimental data indicates that the proposed model is more approach to the real condition.**

Keywords—**system biology, mathematical modeling, lactose operon expression**

I. Introduction

Lactose metabolism is an essential process of energy cycle for living organisms [1, 2]. System biology as a bridge to connect molecular-level mechanisms to cell-level behavior is an integrative approach to understand the metabolism[3]. The modeling of gene expression regulation [4-6] is the basis to research system biology, as well as the basis of making theoretical analysis and guidance for the control of gene regulation expression. The regulation of lactose operon of Escherichia coli that can adapt to the changes in different environments has many advantages to be established the model, including small quantity, simple structure and having no complex processing after transcription and translation [7-10].

The model to describe the dynamics of the lactose operon start with G. Yagil and E. Yagil [11] who proposed a relation between effector concentration and the rate of enzyme synthesis. Bliss and Painter[12] initially considered the time of transcription and translation to develop a delay model of the tryptophan operon. And then Maffahy and Savev [13] modeled lactose operon dynamics and included transcriptional and translational delays. Wong and Gladney[14] published a mathematical model which covered inducer exclusion, catabolite repression and much of the relevant biological detail, but failed to treat these elements of the lactose operon. Mccloskey and Palsson [15] reconstruct the metabolic networks at the genome scale that based on genome sequences and enzymatic data. Yildirim and Mackey [16] presented a mathematical model for the regulation of induction in the lac operon in Escherichia coli with five nonlinear differential delay equations due to the transcription and translation process.

Weijie Wang, School of Automation Science and Electrical Engineering , Beihang university
China

While, the more factors to affect the expression of the lactose operon was investigated. Recently, Ozbudak and Thattai [17] suggested that cAMP which can be effected by glucose have influence on lactose utilization network of Escherichia coli. Therefore, the regulation of the expression can be affected by two mechanisms, which are considered in this model by differential equation. In this paper, multiple factors are considered, including lactose, glucose, repressor, permease and cAMP-CRP. The logical "AND" is used to synthesize the two regulatory mechanisms, and the detail about bi-directional dynamic change of production is considered in this model. The proposed model more correspond to reality by comparing with previous experimental data.

The regulation of the lactose operon [18, 19] is described as follow. The diagrammatic drawing of the lactose operon expression of considering multiple factors shown as Fig. 1. There are seven genes involved in lactose regulation, including distant genes (D), regulatory operator (I), promoter (P), operator (O), LacZ, LacY and LacA. Gene D is responsible for producing CRP, which combine with the cAMP to participate in gene regulation. Gene I can produce repressor (R) that inhibit the transcription of lac operon. And promoter is a regular gene to open the transcription when RNAp bind to operator can receive the repressor protein and then the transcription would be suppressed. The main functional genes that involved in metabolic regulation are LacZ and LacY, which are in charge of β-galactosidase (B) and permease (P) respectively, while the function of LacA is not clearly. The β-galactosidase takes the charge of broking down the intracellular lactose (L) into glucose, galactose, and allolactose (A). The permease has a function of transporting outside lactose into the cell.

The lactose operon expression is regulated through two lines. One is about repressor the other is about cAMP. When there is enough glucose, the repressor binds to the operator to prevent the RNAp binds to the promoter region, and the concentrate of cAMP-CRP is low, so that the expression is suppressed. In the absence of glucose available for cellular metabolism, but lactose is transported into the cell by a permease then broken down into glucose, galactose, and allolactose by the β-galactosidase with enough lactose out of the cell. The allolactose binds with the lactose repressor and makes the repressor cannot bind to the operator, meanwhile, cAMP would increase due to lacking of the glucose, and combination of cAMP-CRP can promote the genes to produce mRNA.
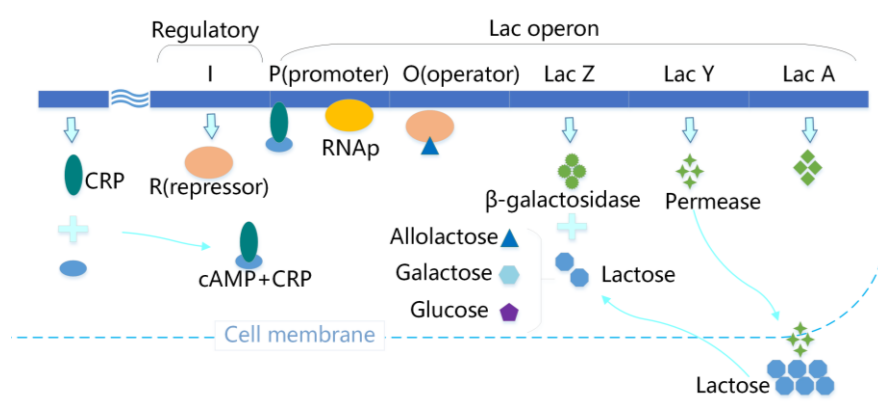
Figure 1 The diagrammatic drawing of the lactose operon expression. Lactose is transported into the cell by a permease then broken down into glucose, galactose, and allolactose by the β-galactosidase. The allolactose binds with the lactose repressor and makes the repressor cannot bind to the operator, meanwhile, cAMP-CRP which would increase due to lacking of the glucose bind on the promoter to facilitate the transcription.

II. Dynamic bi-directional model considering multiple factors

From the description of lactose regulation by existing knowledge, it can be seen that the regulation expression of lactose operon is affected in two aspects. On the one hand, inhibitory effect of repressor, and on the other hand, motivate effect of the combination of CRP and cAMP, which is negatively correlated with glucose. Therefore, the logical "AND" which was proposed in system biology synthesize the two factors, shown as Fig. 2. The transcription rate of lactose operon can be expressed as $f(\text{CAR-cAMP}, \text{Repressor}, t)$, and the mathematic expression shown as follow.

$$\frac{dM}{dt} = \beta_1 \frac{[O]}{[O_T]} \bullet \beta_2 \frac{[D_{CRP\text{-}cAMP}]}{[D_T]} - \gamma_M M = \beta \frac{[O]}{[O_T]} \frac{[D_{CRP\text{-}cAMP}]}{[D_T]} - \gamma_M M \quad (1)$$

Essential to this modeling is the decomposition of the biochemical reaction network into functional units, which can be directly assigned to biological modules.
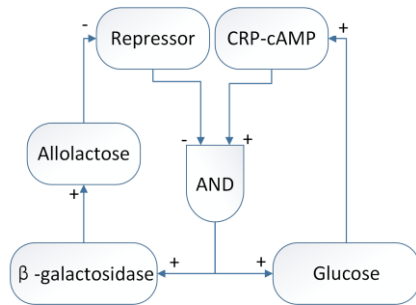

Figure 2 There are two lines to regular the expression of lactose operon. The one is about repressor, the other is about cAMP. The logical "AND" synthesize the two factors.

A. Regulation of repressor

Repressor protein can be produced by a regulatory operon in the preceding of the lac operon. And the repressor can bind to operator region to prevent the RNA Polymerase (RNAp) binding to the promoter region so that the lac operon cannot be fully transcribed. However, allolactose and repressor combined into RA_n, and makes binding of the repressor to the operator region impossible. In that case, the RNAp can easily bind to the promoter initiate transcription of the structural genes to produce mRNA.

1) Binding of repressor to promoter

Repressor, which diffused within the cell, binds to the operator region with the contact at the rate of k_{1on}. And the decomposition rate of OR set to k_{1off}. The reaction of compound can be described as this.
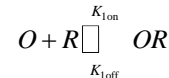
$$O + R \underset{K_{1off}}{\overset{K_{1on}}{\rightleftharpoons}} OR$$

The quantity relationship can be described in a differential equation with the law of mass action

$$\frac{d[OR]}{dt} = k_{1on}[O][R] - k_{1off}[OR] \quad (2)$$

Where [O], [R] and [OR] are the concentrate of operator region, repressor and compound.

Because the reaction reached equilibrium quickly, it is obviously that $\frac{d[OR]}{dt} = 0$. At this point, the concentrate of OR is a stable value, and the dissociation constant can be derived:

$$K_{dOR} = \frac{K_{ORoff}}{K_{ORon}} = \frac{[O][R]}{[OR]} \quad (3)$$

The operator region either contact with repressor or nothing. When operator region is unoccupied, RNAp can be attached to the promoter to transcribe. While, the total of operator equal to unoccupied operator plus occupied operator. Therefore, there is an equation about the operator region as (4).
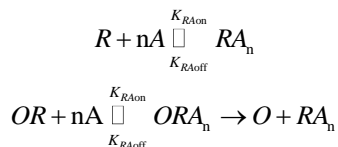
8

Where $[O_T]$ is the total concentration of operator region. Considering (3), the ratio of unoccupied operator region to total operator region can be calculated.

$$\frac{[O]}{[O_T]} = \frac{1}{1 + \dfrac{R}{K_{dOR}}} \quad (4)$$

2) Combination of repressor to allolactose

Allolactose can be conversed from lactose with enzyme β-galactosidase. And allolactose can be combined with repressor no matter whether the repressor bind on the operator region. While, the complex makes the combination between repressor and operator region impossible. The reaction can be described as follow.

Repressor which diffused within the cell binds to the operator region with the contact at the rate of k_{1on}. And the decomposition rate of OR set to k_{1off}. The reaction of compound can be described as follow.

$$R + nA \underset{K_{RAoff}}{\overset{K_{RAon}}{\rightleftharpoons}} RA_n$$

$$OR + nA \underset{K_{RAoff}}{\overset{K_{RAon}}{\rightleftharpoons}} ORA_n \rightarrow O + RA_n$$

Allolactose binds to repressor with the contact at the rate k_{RAon}. The decomposition rate of complex repressor-allolactose (RA_n) is k_{RAoff}. And the allolactose binds to OR at the rate of k_{RAon}. The decomposition rate of RA_n is K_{RAoff}. While, the combination (ORA_n) cannot exist, because RA_n is difficult to attach to operator region.

The dynamical equation of the complex shown as (5).

$$\frac{d[RA_n]}{dt} = K_{RAon}[R][A]^n - K_{RAoff}[RA_n] \quad (5)$$

The same as binding of repressor to promoter, the concentrate of RA_n is a stable value, and the equilibrium equation can be derived.

$$K_{dRA} = \frac{K_{RAoff}}{K_{RAon}} = \frac{[R][A]^n}{[RA_n]} \quad (6)$$

The total amount of repressor is comprised of repressor and RA_n.

$$[R_T] = [R] + [RA_n] \quad (7)$$

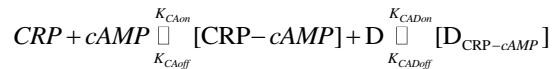Considering the equation above, the ratio of repressor can be deduced with introducing a parameter K_1.

$$K_1 = \frac{[R_T]}{K_{dOR}} \quad (8)$$

$$\frac{[O]}{[O_T]} = \frac{K_{dRE} + A^n}{K_{dRE} + A^n + K_{dRE}K_1} \quad (9)$$

B. Regulation of CRP-cAMP

When there is not enough glucose, the cAMP molecules are synthesized and they bind to cAMP receptor proteins (CRP), forming the complex cAMP-CRP. The cAMP-CRP binds to the region where near the lac promoter to enhance the binding affinity of the promoter for RNAp.

The transcription efficiency that describes the enhancement is equal to the fraction of total binding sites occupied by the CRP-cAMP. The process can be described by the following reaction equation.

$$CRP + cAMP \underset{K_{CAoff}}{\overset{K_{CAon}}{\rightleftharpoons}} [CRP - cAMP] + D \underset{K_{CADoff}}{\overset{K_{CADon}}{\rightleftharpoons}} [D_{CRP-cAMP}]$$

The equations describing their dynamics are consistent with the (2) and (5).

$$\frac{d[CRP - cAMP]}{dt} = K_{CAon}[CRP][cAMP] - K_{CAoff}[CRP - cAMP] \quad (10)$$

$$\frac{d[D_{CRP-cAMP}]}{dt} = K_{CADon}[CRP\text{-}cAMP][D] - K_{CADoff}[D_{CRP-cAMP}] \quad (11)$$

When the reaction is stable, the ratio of total binding sites occupied by the cAMP-CRP complex derived as follow.

$$\frac{[D_{CRP\text{-}cAMP}]}{[D_T]} = \frac{[D_{CRP-cAMP}]}{[D] + [D_{CRP-cAMP}]} = \frac{\dfrac{K_{CAon}}{K_{CAoff}} \cdot \dfrac{K_{CADon}}{K_{CADoff}}[CRP][cAMP]}{1 + \dfrac{K_{CAon}}{K_{CAoff}} \cdot \dfrac{K_{CADon}}{K_{CADoff}}[CRP][cAMP]} \quad (12)$$

The parameters K_{dCAD} is Introduced. The result shown as (14).

$$K_{dCAD} = \frac{K_{CAoff}}{K_{CAon}} \cdot \frac{K_{CADoff}}{K_{CADon}} \quad (13)$$

$$\frac{[D_{CRP\text{-}cAMP}]}{[D_T]} = \frac{[CRP][cAMP]}{[CRP][cAMP] + K_{dCAD}} \quad (14)$$

The concentration of CRP which controlled by another set of genes can be regarded as a definite value. Therefore, the variable that affects the ratio is the concentration of cAMP. The regulation mechanism of cAMP has not been fully mastered. According to the assumption of Wong's, the equation governing the dynamics of the total concentration of cAMP is:

$$\frac{d[cAMP]}{dt} = \frac{k_{cAMP}}{\rho} \frac{K_{b,cAMP}}{[glu] + K_{b,cAMP}} - \gamma_c[cAMP] \quad (15)$$

The dynamical equation of glucose is described as follow.

$$\frac{d[glu]}{dt} = \alpha_A[B]\frac{[A]}{K_A + [A]} + \alpha_L[B]\frac{[L]}{K_L + [L]} - \gamma_g[glu] \quad (16)$$

The first term deals with allolactose conversion to glucose by β-galactosidase. The second term takes into account the conversion of lactose to glucose. The last term accounts for the decrease in internal glucose due to consumption and dilution.

C. Expression of Lactose Operon

A mathematical model for the regulation of expression of lactose operon has been researched by Necmettin Yildirim[16]. His model took into account the dynamics of the β-

9

galactosidase, allolactose, lactose and permease and gave the following equations:

$$\frac{dB}{dt} = \alpha_B \, e^{-\mu \tau_B} \, M - \gamma_B B \qquad (17)$$

$$\frac{dP}{dt} = \alpha_P \, e^{-\mu(\tau_P + \tau_B)} \, M - \gamma_P P \qquad (18)$$

$$\frac{dA}{dt} = \alpha_A \, B \frac{L}{K_L + L} - \beta_A \, B \frac{A}{K_A + A} - \gamma_A A \qquad (19)$$

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{Le} + L_e} - \beta_L P \frac{L}{K_{L1} + L} - \alpha_A \, B \frac{L}{K_L + L} - \gamma_L L \qquad (20)$$

In these equations, the variable of B, P, A, L and L_e represent β-galactosidase, permease, allolactose, lactose and external lactose. μ is the value of the growth rate. The other parameters is constants. B is produced by M, and the rate of change of B is a balance between a production term $\alpha_B f$ and a loss term $\gamma_B B$. Production P is the same as B, the rate of change of P is a balance between a production term $\alpha_P f$ and a loss term $\gamma_P P$. In (19), the first term accounts for production of the lactose. The second term account for breaking down of the allolactose. The final term is a loss term $\gamma_A A$. In (20), the first term accounts for the transport external lactose into cell through the permease. The second term is a converse that lactose transport lactose out to cell. The third term account for breaking down of the lactose. The final term is a loss term $\gamma_L L$.

III. Simulation and Discussion

The model parameters have been set as Tab. 1 by searching published experimental data and existing literature.

Table 1 The model parameters

parameter	value	parameter	value
K_{dRA}	39.68μM2	μ	2.26×10^{-2} / min
K_1	7200	k_{cAMP}	$1 \, \mu M$ / min
$K_{b,cAMP}$	$4 \times 10^{-5} M$	α_A	1.76×10^4 / min
n	2	α_B	1.66×10^{-2} / min
β	9.97×10^{-4} mM/ min	α_P	10 / min
γ_M	0.411/min	α_L	2880 / min
γ_A	0.52 / min	K_A	1.95 mM
γ_B	8.33×10^{-4} / min	K_L	$9.7 \times 10^{-4} M$
γ_L	0	τ_B	2 min
γ_P	0.65 / min	τ_M	0.1 min
γ_g	1.33/min	τ_P	0.83 min

To compare with the actual situation, an experiment by Knorre[20] which is widely used to contrast is introduced. In this experiment, the bacteria were washed and transferred to a lactose-rich and glucose-free medium, so induction of the lactose operon would take place. After the medium change, measured the activity of β-galactosidase.

The contrast result show in Fig. 3. The solid line is the numerical results of activity of β-galactosidase, which based on the model. and the dash dot line is the experimental results of Knorre. It is obvious that the solid line can fit to the experimental results. Especially after the 80min, the fluctuation of the solid line coincided with the fluctuation of experimental data.

Based on the mathematical model established in this paper. The influence of external lactose content on the regulation of lactose operon expression can be quantified, by changing the initial value of the external lactose content. There were two group of lactose concentrate, 3 mM and 1mM, and compare their activity of lactose, allolactose and glucose. The numerical results shown in Fig. 4. The solid line with brown, dotted line with blue and dot dash line with red represented the internal lactose, allolactose and glucose respectively. Obviously, it is content reality that the more external lactose the more activity and longer is expressed. The range of allolactose changes is smaller than internal lactose. Concentrate of the allolactose increase due to the internal lactose. While, when lactose is consumed, the content of allolactose decreases rapidly. The change of glucose activity is consistent with allolactose. When internal lactose is lacking, the activity of glucose begins to decrease significantly.
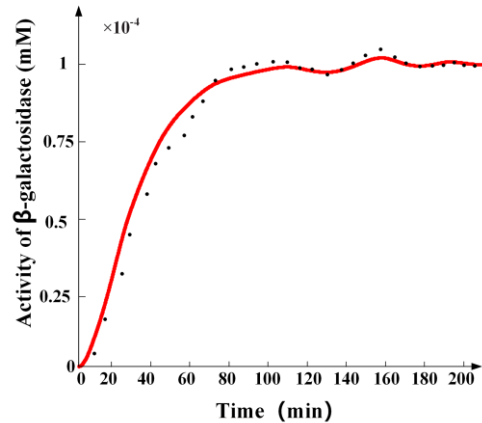

Figure 3 The contrast of the numerical results with experimental results. The solid line is the numerical results of activity of β-galactosidase, and the dash dot line is the experimental results. The solid line can fit to the experimental results
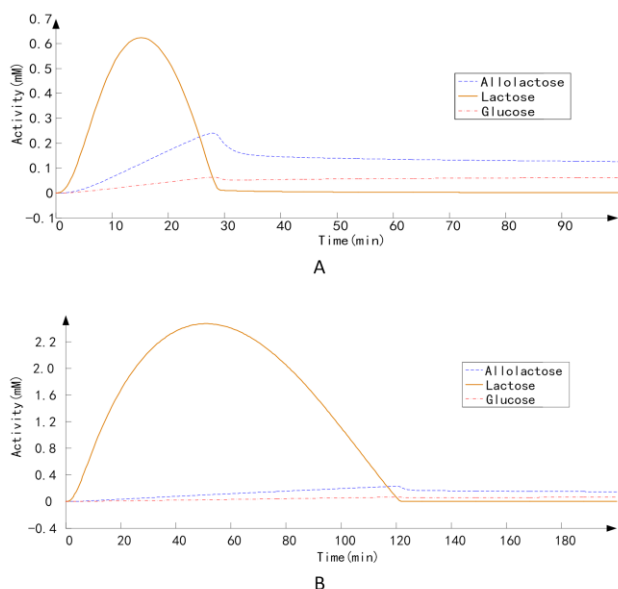
Figure 4 Numerical results with different external concentrate of lactose. The more external lactose, the more activity and longer reaction time. When lactose is lacking, the activity of allolactose and glucose begins to decrease.

IV.   Conclusion

Two regulating circuits, repressor and cAMP respectively effect the regulation of lactose operon expression. A mathematic model of lactose operon expression is established in differential equation with the knowledge of system biology. In the model, the logical "AND" is proposed to synthesize the two regulatory mechanisms. Multiple factors are considered, including lactose, glucose, repressor, permease, β-galactosidase and cAMP-CRP. In addition, the detail about bi-directional dynamic change of production is considered. The modeling more consistent with the reality by comparing with previous experimental data.

Acknowledge

References

[1]   B. Gorke and J. Stulke, "Carbon catabolite repression in bacteria: many ways to make the most out of nutrients," *Nat Rev Microbiol,* vol. 6, pp. 613-24, Aug 2008.

[2]   M. Y. Pavlov and M. Ehrenberg, "Optimal control of gene expression for fast proteome adaptation to environmental change," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 110, p. 20527, 2013.

[3]   D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, and T. Hwa, "A global resource allocation strategy governs growth transition kinetics of Escherichia coli," *Nature,* vol. 551, p. 119, 2017.

[4]   M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa, "Interdependence of cell growth and gene expression: origins and consequences," *Science,* vol. 330, pp. 1099-102, 2010.

[5]   J. J. Lemke, P. Sanchezvazquez, H. L. Burgos, G. Hedberg, W. Ross, and R. L. Gourse, "Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 108, pp. 5712-5717, 2011.

[6]   K. Bettenbrock, S. Fischer, A. Kremling, K. Jahreis, T. Sauter, and E. D. Gilles, "A quantitative approach to catabolite repression in Escherichia coli," *J Biol Chem,* vol. 281, pp. 2578-84, Feb 3 2006.

[7]   J. H. De, "Modeling and simulation of genetic regulatory systems: a literature review," *Journal of Computational Biology A Journal of Computational Molecular Cell Biology,* vol. 9, p. 67, 2002.

[8]   E. Dekel and U. Alon, "Optimality and evolutionary tuning of the expression level of a protein," *Nature,* vol. 436, pp. 588-92, 2005.

[9]   A. Kremling, K. Bettenbrock, B. Laube, K. Jahreis, J. W. Lengeler, and E. D. Gilles, "The Organization of Metabolic Reaction Networks : III. Application for Diauxic Growth on Glucose and Lactose," *Metabolic Engineering,* vol. 3, pp. 362-379, 2001.

[10]   M. Santillán and M. C. Mackey, "Influence of catabolite repression and inducer exclusion on the bistable behavior of the lac operon," *Biophysical Journal,* vol. 86, pp. 1282-92, 2004.

[11]   G. Yagil and E. Yagil, "On the relation between effector concentration and the rate of induced enzyme synthesis," *Biophysical Journal,* vol. 11, pp. 11-27, 1971.

[12]   R. D. Bliss, P. R. Painter, and A. G. Marr, "Role of feedback inhibition in stabilizing the classical operon," *Journal of Theoretical Biology,* vol. 97, pp. 177-93, 1982.

[13]   J. M. Mahaffy and E. S. Savev, "Stability analysis for a ma thematical model of the lac operon," *Quarterly of Applied Mathematics,* vol. 57, pp. 37-53, 1999.

[14]   P. Wong, S. Gladney, and J. D. Keasling, "Mathematical model of the lac operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose," *Biotechnology Progress,* vol. 13, p. 132, 1997.

[15]   D. Mccloskey, B. Ø. Palsson, and A. M. Feist, "Basic and applied uses of genome‐scale metabolic network reconstructions of Escherichia coli," *Molecular Systems Biology,* vol. 9, pp. -, 2013.

[16]   Yildirim, Necmettin, Mackey, and C. Michael, "Feedback Regulation in the lactose operon: A Mathematical Modeling Study and Comparison with Experimental Data," *Biophysical Journal,* vol. 84, pp. 2841-2851, 2003.

[17]   E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and O. A. Van, "Multistability in the lactose utilization network of Escherichia coli," *Nature,* vol. 427, p. 737, 2004.

[18]   M. E. Wall, M. J. Dunlop, and W. S. Hlavacek, "Multiple Functions of a Feed-Forward-Loop Gene Circuit," *Journal of Molecular Biology,* vol. 349, pp. 501-514, 2005.

[19]   S. Semsey, L. Jauffred, Z. Csiszovszki, J. Erdossy, V. Stéger, S. Hansen, *et al.*, "The effect of LacI autoregulation on the performance of the lactose utilization system in Escherichia coli," *Nucleic Acids Research,* vol. 41, pp. 6381-90, 2013.

[20]   W. A. Knorre, "Oscillations of the rate of synthesis of beta-galactosidase in Escherichia coli ML 30 and ML 308," *Biochemical & Biophysical Research Communications,* vol. 31, pp. 812-817, 1968.