



Effect of Rate Constants in TF Regulated Gene Expression

Enakshi Guru¹, Sanghamitra Chatterjee²

¹Department of Physics, Camellia Institute of Technology, Kolkata, India

²Department of Electronics, Camellia Institute of Technology, Kolkata, India

Email: enakshiguru@hotmail.com, sangha30@rediffmail.com

Received 27 October 2014; revised 30 November 2014; accepted 23 December 2014

Copyright © 2014 by authors and OALib.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Simple mathematical model of gene expression has been studied thoroughly in this paper. We assume that protein synthesis and protein decay are deterministic in nature whereas transition from active to inactive state is totally stochastic. The probability distribution function of transcript number is calculated and evidence of binary and graded response has been observed for different reaction rates. Slow and fast kinetics have been studied explicitly. Different rate constants play a significant role in calculating distribution function (PDF) and mean value of TF regulated gene expression. It is seen that rate constants control the variance as well. A small change in the value of the rate constant gives a significant change in average protein value. The randomness of gene expression is reduced with increasing value of one of the rate constants. The main focus of this paper is to observe the effect of rate constants in transcription factor regulated gene expression.

Keywords

Transcription Factor Regulation, Binary and Graded Response, Rate of Protein Activation, Fast and Slow Kinetics, Probability Distribution Function

Subject Areas: Biophysics, Computational Biology, Synthetic Biology

1. Introduction

Gene expression and its regulation are studied in details in cell dynamics [1]-[4]. Transcription and translation are two main steps of gene expression and these two processes involve in several biochemical reactions. The sequence along with the strand of the DNA is copied or transcribed in mRNA in the first step which is known as transcription. The sequence of mRNA molecules is translated into protein in the second step of gene expression which is called translation. System regulation is achieved by binding a regulatory molecule *i.e.* activator or repressor to the operator region of DNA in eukaryotic cell. This activator or transcription factor binds at the pro-

moter region and regulates the gene expression [5] [6]. The rate of production of protein is a function of transcription factor. This regulatory function is either a monotonic increasing or decreasing function depending on the nature of the transcription factor (activator or repressor). Hill function describes this regulation function well. TF interacts either in one site or in the multiple sites of the promoter. Cellular inducing signal activates the TFs which then bind to appropriate enhancer sequence of DNA. Cooperative bindings among the TF form the bound TF complexes or oligomers are sometimes essential for the initiation of gene expression in eukaryotes. Gutierrez *et al.* studied the role of cooperativity in gene regulation and they had shown that cooperativity is a new source of bimodality [7]. They investigated that the Hill coefficient and the level of noise increased with the increase in interaction energy between activators [7] [8]. The response to an inducing signal is measured by the amount of protein synthesized by cell in gene expression. The response is found to be either graded or binary [9] [10] depending on the relative stability of the activated and deactivated gene.

In Binary gene expression [11] [12] two distinct subpopulations of genetically identical cells are observed which means that states are either low or high. A population of similar cells with identical genetic structure is expected to give similar gene expression under same environmental condition. But in real case two distinct subpopulations are observed. In one of the subpopulations, the protein level is low whereas in the other subpopulation the protein level is high. A fraction of cells in which the proteins levels are intermediate are considered as low state. Binary gene expression has also been known as the “all-or-none” phenomenon. Novick and Weiner [11] gave the first evidence of this type of gene expression. TF regulates the gene expression by increasing the concentration of activator or repressor and raises the fraction of cells in the high or low subpopulation. The origin of binary gene expression is purely stochastic. In binary gene expression two distinct peaks are observed. In graded response the output varies continuously as the amount of input stimulus varies, until the steady state is achieved. Graded and binary response both can occur in the same system if the rate constants vary accordingly.

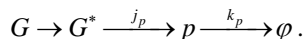
Stochastic gene expression is associated with a series of biochemical events which are probabilistic in nature. The time evolution is not a deterministic process and the biochemical events in gene expression also random in nature [13]-[15]. Binding or unbinding of RNAP at promoter region of DNA and also the regulatory molecules in the operator region are probabilistic process. Probabilistic nature is ignored when the participant molecules are large in number. We have studied the stochastic process in *E. coli* in case of toggle switch model [16]. Proteins that express in a cell exhibit fluctuations around a mean value as a function of time. In a population of similar cells, the protein levels in individual cells at a specific instant of time are not identical but spread around a mean value. Expression of gene also depends on the specific values of rate constant significantly. The rate of reaction plays a significant role in gene expression. As the value of rate of reaction increases randomness increases, system requires more time to reach in equilibrium. The system becomes almost deterministic in nature for very low value of reaction rate. Mean value of protein, around which the protein level fluctuates, also changes with the rate constant. Higher value of rate constant mean value is logically in high state. Fluctuation around mean position creates noise [17]-[21], which is measured extensively in this paper. Systems biology predicts theoretical modelling with which one can explain the noble features of cellular phenomenon and several genetic diseases.

This paper is organized as follows. In Section 2 we describe a simple model of gene expression combining deterministic and stochastic approach. In this section we assume that gene can switches from inactive to active state randomly but protein synthesis and protein decay are deterministic process. In Section 3 we calculate the probability distribution function of stochastically expressed gene. By changing the rate constant we describe the binary and graded gene expression in the same system. In binary gene expression the expression level is either low or high which means that the response is digital in nature. Two distinct subpopulations in a system of genetically identical cell are shown whereas for graded response the graph shows unimodal peaks. In Section 4 we describe the transcription factor induced gene. We observe the change of dose response curve that is probability distribution function of induced gene with increased value of Hill coefficient. We also study the slow and first kinetics of reaction rate equation. In Section 5 we adopt Gillespie model to calculate the protein concentration stochastically.

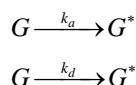
We observe the effect of rate constants in different aspect of stochastic calculations. We study the change in variance, fano factor and mean protein level with rate constants. We have tried to find out the role of different rate constants *i.e.* protein activation and deactivation rate, protein decay rate, protein degradation rate due to cell growth and dilution rate to cell division, in TF regulated gene expression and also incorporated this in mathematical model of gene expression.

2. Mathematical Model of Gene Expression for Single Gene

In a simple stochastic model of gene expression [6] [10] [22] a gene can be either in inactive (G) state or in active (G^*) state. Only in the active state, gene can synthesis a protein (p) at the rate called j_p . The gene degrades with another rate constant called k_p and the degradation product is given by φ . Protein decay occurs in case of both active and inactive states but protein synthesis occurs only in the active states. The schematic diagram of the whole process is given below



j_p signifies degradation rate due to cell growth and k_p signifies dilution rate to cell division. Random transition occurs from active gene to inactive gene. The nature of the transition is stochastic and given by the following equation:



where k_a and k_d are the activation rate constants and the deactivation rate constants respectively. The time evolution of the system of reaction is not a continuous process as the molecular population levels in a reaction changes only by discrete amount. Half time for the gene activation is $T_a = \frac{\log 2}{k_a}$ and the half time for gene deactivation is $T_d = \frac{\log 2}{k_d}$. The half time associated with the protein degradation is $T_p = \frac{\log 2}{k_p}$.

For simplicity we assume protein synthesis and protein decay occurs deterministically according to the following equation

$$\frac{dx}{dt} = \frac{j_p}{X_{\max}} z - k_p x$$

where $x = X/X_{\max}$, X and X_{\max} being the protein concentration at time t and the maximum protein concentration respectively. The value of z flips between “0” and “1” at random time intervals. In the deterministic approach the time evolution of the system is assumed to be continuous but actually it is not deterministic as molecular population level changes randomly with time. Deterministic rate approach is justified when the number of molecules is larger than the fluctuating molecule.

System biologists [6] [10] [22] had studied probabilistic interpretation of the rate constants. The probability density function of transcription number which describes the distribution of protein levels in the steady state is given by the beta distribution

$$p(x) = N(1-x)^{r_2-1} x^{r_1-1}$$

$$\text{with } r_1 = k_a/k_p, \quad r_2 = k_d/k_p \quad \text{and} \quad N = \frac{\Gamma(r_1+r_2)}{\Gamma(r_1)\Gamma(r_2)}.$$

In induced gene expression, activators or transcription factors (TFs) promote transitions to the active state G^* of the gene. In this case, k_a and k_d are functions of the concentration of the TF molecules.

3. Binary and Graded Response for TF Regulated Single Gene

Considering the parameters r_1 and r_2 , binary and graded response takes place in the same model. By changing the rate constant of protein decay k_p , r_1 and r_2 can be changed. Concentration of Transcription factor with the probability distribution are plotted in **Figure 1** and **Figure 2**. When $r_1 < 1$, $r_2 < 1$ bimodal distribution in the protein levels is obtained which is shown in **Figure 1**. In induced gene expression, as the concentration (S) of the TF molecules is increased, a binary response is observed *i.e.* the fraction of cells go up in the “high” sub-population by sharp transition.

In the binary gene expression if the gene is in inactive state, the mean protein level (x) will be zero and if it is in active state the mean protein level x will be one, X will be equal to X_{\max} *i.e.* the maximum protein

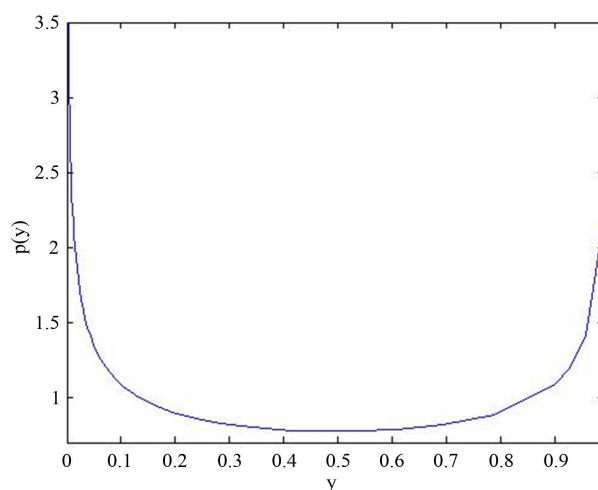


Figure 1. Plot of probability density function with protein concentration gives binary response in protein level for $r_1 < 1$ and $r_2 < 1$.

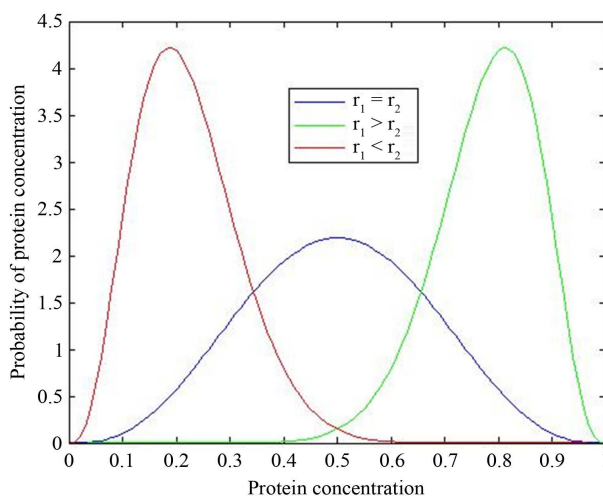


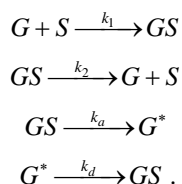
Figure 2. Protein concentration changes with probability distribution function for different rate constants r_1 and r_2 .

synthesis. Slow transitions between the inactive and active promoter states are expected to be relevant in eukaryotic gene expression. Two distinct subpopulations originate due to slow transitions between the inactive and active states of the gene. If transcription factor (TF) influences the transition from G to G^* then k_a and k_d will be the function of concentration (S) of TF molecules. If the concentration of the activator (S) molecule increases, the fraction of cell goes to high state from lower on. Thus we get binary response. When random activation and deactivation processes are taken place if the activation rate (k_a) and deactivation rate (k_d) are faster than rate of protein synthesis (k_p) then the average protein level should be between 0 to 1. We have taken $k_a = 2$ and $k_d = 3$ for our analysis. If the activation and deactivation rate constants are slower than protein decay mean protein level will be either low or at high state, depending on whether the gene is in deactivated or activated state. If the residence time of gene in activated state is sufficient, protein expression will be maximum and high state is obtained. If the gene is in deactivated state for a long time accumulated protein decays fully and it goes to low state. When r_1 and r_2 both are greater than 1 then the three possible cases are achieved which are shown in **Figure 2** for three conditions respectively that is, $r_1 > r_2$, $r_2 > r_1$ and $r_1 = r_2$. A single peak is observed in **Figure 2** which is unimodal distribution with graded response. The peak position shifts with the values of r_1 and r_2 . Peak will be at the middle position when $r_1 = r_2$. When $r_1 > r_2$ the peak is shifted to-

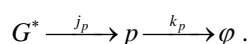
wards right and when $r_2 > r_1$ the peak is shifted towards left.

4. Effect of Hill Coefficient in TF Regulated Gene Expression

Transcription factor are the proteins (S) which bind at the appropriate regions of G and regulate its expression in transcription. The concentration of TF affects the expression of protein. The variation of concentration of TF gives the nonlinear response of amount of protein synthesized by the target gene. If we assume that G goes to active state by the presence of stimulus S then the reaction scheme will be



Other states will be as before



The effective rate constant (K'_a) of the above equations have the form of Hill function where

$$K'_a = k_a \frac{(S/K)^n}{1 + (S/K)^n}$$

where K is the activation coefficient which describes the dissociation constant, n is called Hill coefficient or degree of cooperativity which establishes how close the Hill function is to the step function.

In **Figure 3** it has been shown that effective rate constants vary with concentration of TF. The Dose response curve is sigmoid in nature, gives the maximum steepness at intermediate level. A small change in the concentration of TF about the inflection point, (the point where the tangent to the curve has maximum slope) give wide change of effective rate of protein synthesis. Inducer increases the number of expressed cell but the not the average protein concentration. For $n=1$ the plot is hyperbolic. For $n>1$ it is sigmoidal of S shaped.

Figure 4 shows the variation of probability distribution function with the variation of concentration of TF. It is also evident from the graph that with the increasing value of hill coefficient effective rate constant changes and as a result the peak value of PDF is being increased. For smaller value of n ($n=3$) peak height is small. When $n=6$, the peak of the distribution function attains higher value and shifts towards right. As hill coefficient n increases more ($n=12$) the vertical rise of PDF continues but the shifting of peak horizontally is not so significant.

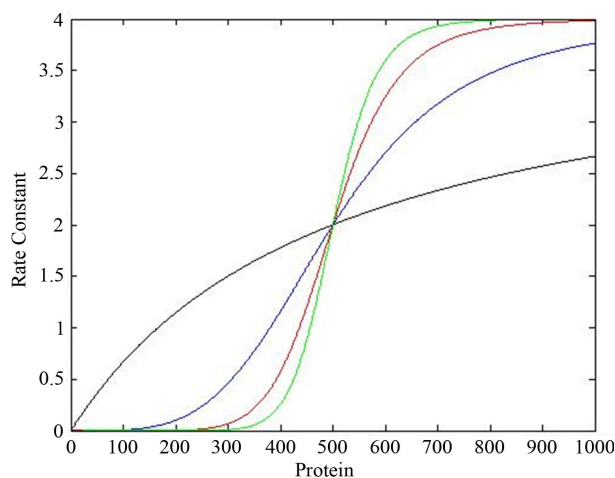


Figure 3. Rate constant varies with concentration of TF, curve becomes stiffer with rising hill coefficient.

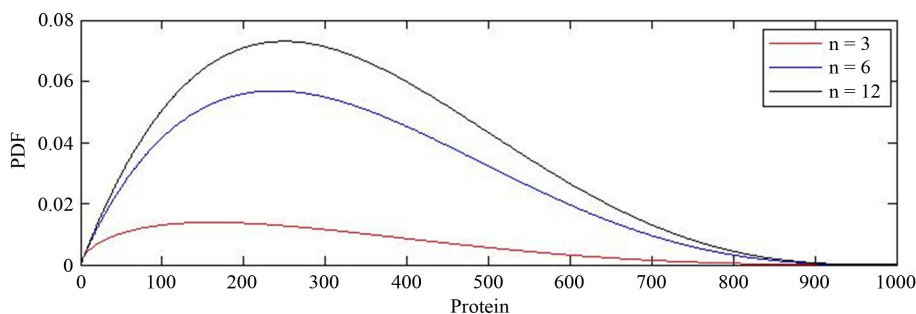


Figure 4. Plot of concentration of protein vs PDF for different hill function, with the increase in hill coefficient the value of PDF increases and position of peak also shift.

Let $T_a = T_d = T_p / \gamma$ so $k_a = \frac{\log 2}{T_a}$ and $k_d = \frac{\log 2}{T_d}$ and $k_p = \frac{\log 2}{T_p} = \frac{\log 2}{\gamma T_a} = \alpha k_a$ where α and γ are

numeric constant and $\alpha > 0$. With the increase in α , we get faster expression of kinetics. In **Figure 5** we plot the same graph as **Figure 4** (*i.e.* protein concentration vs PDF) for slow and fast kinetics. As α increases the gene expression kinetics become faster and variance is also reduced but the mean protein level remains unaltered and we get the peak in the same position of the protein level.

5. Effect of Rate Constant in Stochastic Gene Expression

The role of stochastic gene expression has been studied thoroughly by constructing mathematical model of gene expression using Gillespie model [23]. The time evolution of the number of biomolecules of a particular type is determined by this technique. The average protein vs time and histogram of average protein concentration is plotted in **Figure 6(a)** and **Figure 6(b)**. Histogram depicts the unimodal distribution which have been also observed in deterministic process (**Figure 2**, **Figure 4** and **Figure 5**).

In the stochastic analysis random fluctuation around the mean protein level gives rise to noise which is observed clearly in **Figure 6**. In **Figure 7** the rate constant is increased ten times than what is done in **Figure 6**. It is seen clearly randomness decreases. So the rate constants have a significant role in case of stochasticity in gene expression.

Noise plays an important role in gene expression. Amount of protein expressed randomly is always fluctuate around a mean value. This variation is calculated by fano factor. It is clear from the **Table 1** that with increasing value of k_p (for a fixed value of j_p) variance decreases *i.e.* noise is less. So the fano factor also decreases with k_p . The value of average protein level around which protein level fluctuates decreases with the increase of rate constant of protein. This is simply because as the rate of protein dissociation increases less amount of protein will sustain. It is revealed from **Figure 6** and **Figure 7** that with increasing value of k_p randomness reduces. For a fixed value of k_p variance, fano factor and mean protein level increases. The expression of fano factor is given by $(\varphi) = \frac{\sigma^2}{N}$ where N is average protein abundance and σ is standard deviation. Similarly

the expression of noise or coefficient of variation is given by $(\eta) = \frac{\sigma}{N}$. The following table (**Table 1**) gives the clear idea that how rate constants k_p and j_p are related with variance and noise.

The plot of average protein concentration with time for different rate constants is given in **Figure 7**. It has been observed that with the higher rate constant ($k_p = 5$) system reaches in the equilibrium faster than the medium ($k_p = 0.5$) and low rate constant ($k_p = 0.05$). It is evident from the **Figure 6** and **Figure 7** that as the activation rate constant k_p decreases, the gene get more time to stay in inactive state so time requires to go in the active state is much larger than for smaller k_p value. We also plot the protein concentration with time for different values of rate constants k_p . It is revealed from the **Figure 8** that for rate constant $k_p = 0.05$ more time is required to reach in equilibrium. The system will reach equilibrium faster for high values of rate constant of protein decay (k_p).

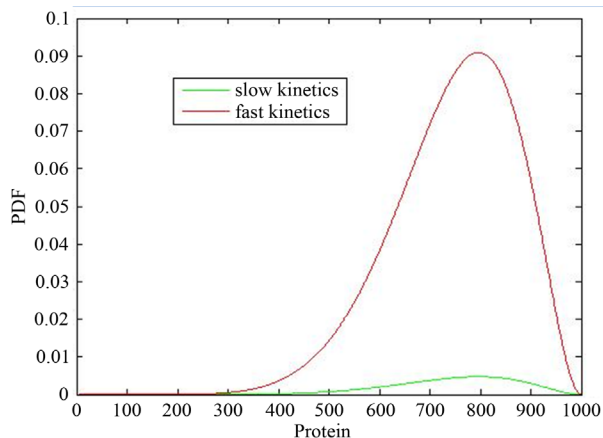


Figure 5. Slow and first kinetics occurs for different rate constants.

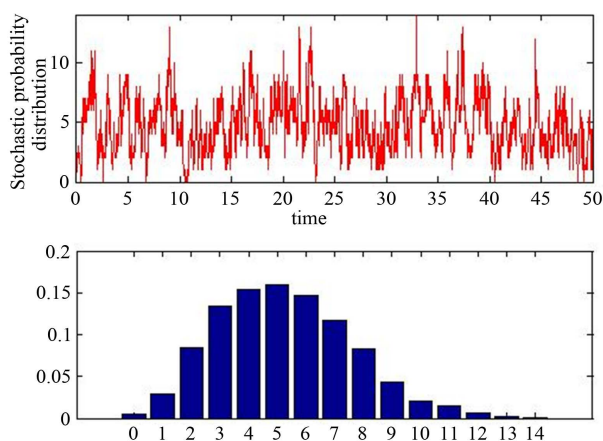


Figure 6. (a) Fluctuation of protein about mean level with time; (b) Unimodal distribution of protein level is observed in the histogram.

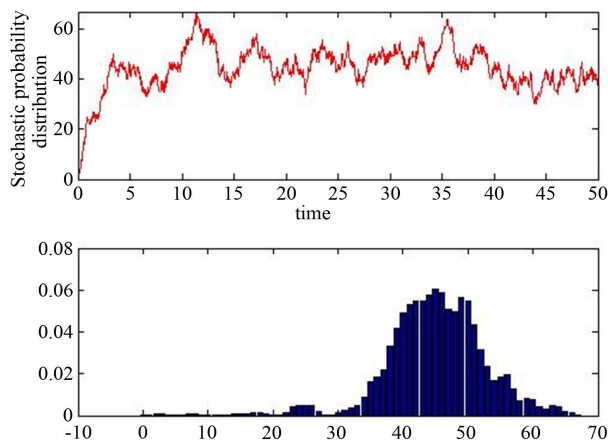


Figure 7. Randomness reduces as we increase the rate constant k_p from 0.05 to 0.5 and unimodal distribution of protein level is also observed in the histogram.

We calculate the variance in protein concentration from the stochastic model of Gillespie for different rate constant. In **Figure 9** we plot the variance for different values of rate constants (k_p). We get the same time

Table 1. Shows the dependence of mean protein level, variance and fano factor with different k_p and j_p values.

j_p	k_p	Variance	Fano factor	Mean protein level
25	0.05	1.00×10^4	24.82	404.92
	0.5	60.33	1.28	47.25
	5	4.4	0.86	5.1
100	0.05	1.79×10^5	110.13	1640
	0.5	511	2.56	200.25
	5	20.37	0.98	20.73
1000	0.05	1.88×10^7	1130	16700
	0.5	1.97×10^5	15.07	2000
	5	229.55	1.15	200.09

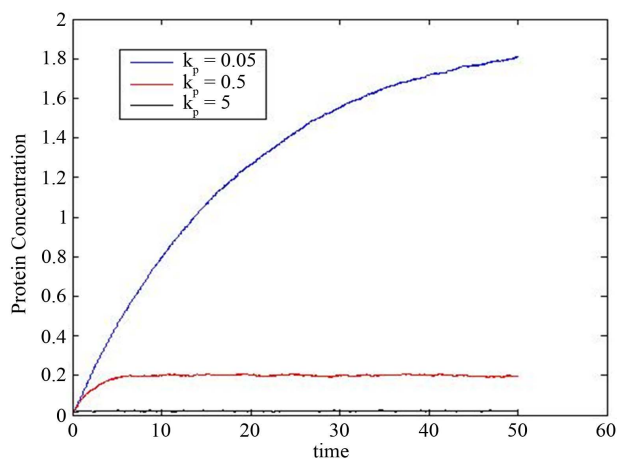


Figure 8. Plot of protein concentration with time for different rate constants (k_p).

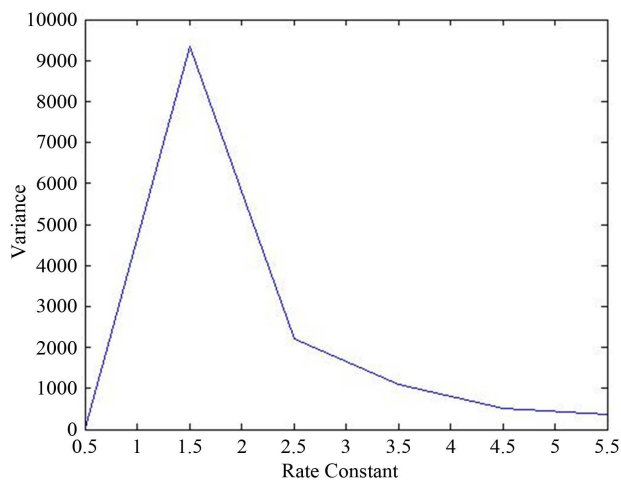


Figure 9. Rate constant k_p vs variance.

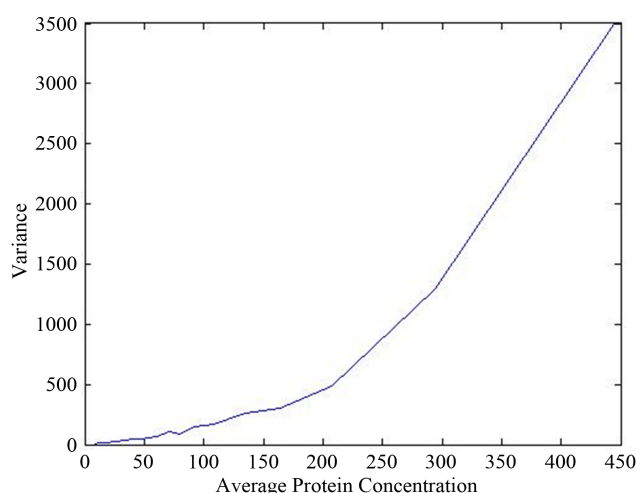


Figure 10. Change in variance with average protein concentration.

dependence with rate constant in deterministic cases also (Figure 5). With the increase of rate constant variance increases and reaches a peak for a particular value of k_p ($k_p = 1.5$), after that with the increase of rate constant variance decreases.

Rate constant k_p has a significant role to control the variance. It is clearly observed from the above figure that variance goes into saturation after $k_p = 4.5$.

The change of variance with average protein concentration is shown in Figure 10. It is clearly observed from this picture that variance increases linearly with average protein concentration when concentration is high.

6. Summary

In this paper we have studied simple mathematical model of gene expression in both deterministic and stochastic process. Activation of gene is considered to be in random manner but protein synthesis and decay are occurred in a deterministic way. We observe binary and graded gene expression in the same model by changing the value of activation and deactivation rate constant and also the rate constant of protein decay. We have seen the effect of rate constant in TF regulated gene expression. We also observe probability distribution function of TF regulated gene expression for different Hill coefficient. By changing the rate constant of protein decay we can control the kinematics of the rate equation. As a result fast and slow curve of gene expression are obtained. We calculate the protein concentration with time in stochastic method using Gillespie algorithm. Stochasticity of gene expression also depends on rate constant of protein decay. We observe that stochasticity decreases with increasing value of rate constants. We have shown that mean protein level increases with the rate of protein synthesis. We have also seen that the equilibrium reaches earlier for higher value of rate constant of protein decay. We plot histogram which shows unimodal distribution of protein concentration which is in agreement with deterministic process. We have calculated the fano factor and variance of the system for different rate constant of protein decay and degradation rate. We have observed that variance changes with decay rate of protein. Variance is maximum for a fixed value of this decay rate. Variance decreases on either side with the variation of rate constant. We also observe that variance increases with average protein concentration. We have studied extensively the nature of stochastic burst with the deterministic rate constants and two techniques give almost same observations in our paper.

7. Conclusion

It has been concluded from this paper that although the process of TF regulated gene expression is stochastic in nature even then it is highly dependent on the specific range of values of different rate constants. Different aspects of gene expressions regarding rate constants have been thoroughly explored so that a clear picture can be viewed. This provides a document to get an overview of different rate constants (decay constant, growth constant etc.) in TF regulated gene expression.

References

- [1] Alberts, B., Bray, D., Lewis, J., Ruff, M. and Roberts, K. (1989) *Molecular Biology of the Cell*. Garland Publishing, New York.
- [2] Andrianantoandro, E., Basu, S., Karig, D.K. and Weiss, R. (2006) Synthetic Biology: New Engineering Rules for an Emerging Discipline. *Molecular Synthetic Biology*, **2**, 1-4.
- [3] Hasty, J., McMillen, D., Isaacs, F. and Collins, J.J. (2001) Computational Studies of Gene Regulatory Network: In Numeromolecular Biology. *Nature Reviews Genetics*, **2**, 268. <http://dx.doi.org/10.1038/35066056>
- [4] Hasty, J., McMillen, D. and Collins, J.J. (2002) Engineered Genetic Circuits. *Nature*, **420**, 224-230. <http://dx.doi.org/10.1038/nature01257>
- [5] Alon, U. (2006) *An Introduction to System Biology: Design Principles of Biological Circuits*. Chapman and Hall, London.
- [6] Karmakar, R. and Bose, I. (2006) Stochastic Model of Transcription Factor-Regulated Gene Expression. *Physical Biology*, **3**, 200. <http://dx.doi.org/10.1088/1478-3975/3/3/005>
- [7] Gutierrez, P.S., Monteoliva, D. and Diambra, L. (2012) Cooperative Binding of Transcription Factors Promotes Bimodal Gene Expression Response. *PLoS One*, **7**, e44812. <http://dx.doi.org/10.1371/journal.pone.0044812>
- [8] Gutierrez, P.S., Monteoliva, D. and Diambra, L. (2009) Role of Cooperative Binding on Noise Expression. *Physical Review E*, **80**, Article ID: 011914. <http://dx.doi.org/10.1103/PhysRevE.80.011914>
- [9] Pirone, J.R. and Elston, T.C. (2004) Fluctuation in Transcription Factor Binding Can Explain the Graded and Binary Response Observed in Inducible Gene Expression. *Journal of Theoretical Biology*, **226**, 111-121. <http://dx.doi.org/10.1016/j.jtbi.2003.08.008>
- [10] Karmakar, R. and Bose, I. (2004) Graded and Binary Response in Stochastic Gene Expression. *Physical Biology*, **1**, 197. <http://dx.doi.org/10.1088/1478-3967/1/4/001>
- [11] Novick, A. and Weiner, M. (1957) Enzyme Induction as an All-or-None Phenomenon. *Proceedings of the National Academy of Sciences of the United States of America*, **43**, 553-566. <http://dx.doi.org/10.1073/pnas.43.7.553>
- [12] Blake, W.J., Kaern, M., Cantor, C.R. and Collins, J.J. (2003) Noise in Eukaryotic Gene Expression. *Nature*, **422**, 633-637. <http://dx.doi.org/10.1038/nature01546>
- [13] Blake, W.J., Kaern, M., Cantor, C.R. and Collins, J.J. (2005) Stochasticity in Gene Expression: From the Theories of Phenotypes. *Nature Reviews Genetics*, **6**, 451-464. <http://dx.doi.org/10.1038/nrg1615>
- [14] McAdams, H.H. and Arkin, A. (1997) Stochastic Mechanism in Gene Expression. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 814-819. <http://dx.doi.org/10.1073/pnas.94.3.814>
- [15] Elowitz, M.B., Liebler, S., Siggia, A.J. and Swain, E.D. (2002) Stochastic Gene Expression in a Single Cell. *Science*, **297**, 1183-1186.
- [16] Guru, E. and Chatterjee, S. (2013) Study of Synthetic Biomolecular Network in Escherichia Coli. *International Journal of Biophysics*, **3**, 38-50.
- [17] Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. and Elowitz, M.B. (2005) Gene Regulation at the Single Cell Level. *Science*, **307**, 1962-1965. <http://dx.doi.org/10.1126/science.1106914>
- [18] Thattai, M. and Van Oudenaarden, A. (2001) Intrinsic Noise in Gene Regulatory Network. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 8614-8619. <http://dx.doi.org/10.1073/pnas.151588598>
- [19] Ozbudak, E.M., Thattai, M., Kurlser, I., Grossman, A.D. and Van Qudenaarden, A. (2002) Regulation of Noise in the Expression of a Single Gene. *Nature Genetics*, **31**, 69-73. <http://dx.doi.org/10.1038/ng869>
- [20] Swain, P.S., Elowitz, M.B. and Siggia, E.D. (2002) Intrinsic and Extrinsic Contribution to Stochasticity in Gene Expression. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 12795-12800. <http://dx.doi.org/10.1073/pnas.162041399>
- [21] Thattai, V. and Qudenaarden, A. (2001) Intrinsic Noise in Gene Regulatory Networks. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 8614-8619. <http://dx.doi.org/10.1073/pnas.151588598>
- [22] Cook, D.L., Gerber, A.N. and Tapscott, S.J. (1998) Stochastic Gene Expression: Implications for Haploinsufficiency. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 15641-15647. <http://dx.doi.org/10.1073/pnas.95.26.15641>
- [23] Gillespie, D.T. (1977) Exact Stochastic Simulation of Coupled Chemical Reaction. *The Journal of Physical Chemistry*, **81**, 2340-2361. <http://dx.doi.org/10.1021/j100540a008>