ELSEVIER

Biophysical Chemistry 121 (2006) 51 – 56

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

Stochastic transcription initiation: Time dependent transcription rates

R. Murugan *

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India

Received 4 December 2005; received in revised form 21 December 2005; accepted 21 December 2005
Available online 27 January 2006

Abstract

The noise in the central process such as transcription, replication and translation of the genomic DNA is very important since it can directly affect the phenotypic and behavioral aspects of an organism as well as the entire cellular function. Here we develop a model on the transcription process based on the assumption that the initiation of the transcription is a stochastic event and the transcription rates may be time dependent random quantities. We derive the central measure properties i.e. mean and the variance, of the distribution of the transcription rates. Our results show that the Fano factor which is a measure of deviation from the Poisson distribution associated with the fluctuations in the number of mRNA molecules deviates from unity due to the randomness in the transcription rates. However when the RNA polymerase molecule searches for the promoter sequences on the DNA lattice by random jumps, the Fano factor approaches the Poisson limit as the jump size associated with the RNA polymerase increases. Since the jump size associated with dynamics of RNAP molecule is positively correlated with the degree of super coiling of DNA, we argue that the super coiled or close-packed structure of DNA might have evolved to keep the noises at the transcriptional level in a minimum.

Keywords: Stochastic gene expression; Transcriptional noise; Transcription initiation

1. Introduction

Site specific interaction of a protein molecule with the DNA lattice is the fundamental process in molecular biology especially in the replication, transcription and the translation of the genomic DNA. Here replication, transcription and translation of the genomic DNA are the central processes of life (called as "central dogma of life") that finally results in the formation of structural and functional building-blocks of the living cell including various biochemical pathways and regulation networks [1]. Due to the microscopic nature of the living cell, the biochemical process and pathways that are happening in vivo are prone to fluctuations which play an important role in modulating various phenotypic (e.g. organisms possessing similar genetic constitution still exhibit diverse phenotypic characteristics) and behavioral aspects (e.g. circadian cycle) of an organism. Among the effects of noises on various levels of cellular organization, the effect of noise on the central process such as transcription, translation and replication of the genomic DNA is very important since it can significantly affect the phenotypic appearance, genetic constitution and genetic inheritance of the organism. Here one should note that the noise at the transcriptional level is critical since even a small fluctuation in the number of mRNA molecules will be amplified to a large fluctuation in the number of protein molecules at the post-translational level.

The strength of noise associated with the number of particular molecule will be measured in terms of the Fano factor [2–6] which is the ratio between the variance (either time averaged or ensemble averaged) and the mean associated with the number fluctuations. Detailed experimental studies show that the Fano factor associated with the fluctuations in the number of mRNA molecules of a single gene is linearly correlated with the transcription rate as well as the mean number of mRNA molecules. Similarly the Fano factor associated with the corresponding protein molecules is linearly correlated with the translation rate as well as the mean number of protein molecules. Almost all the earlier works were mainly focused to explain these facts [Refs. 7–23].

Though noises strongly influence different processes at various cellular organizations of an organism, the living cells are capable of developing the mechanisms to reduce such

* Tel.: +91 22 2278 2737; fax: +91 22 2280 4610, 2280 4611.
E-mail addresses: muruga@tifr.res.in, RMurugan@gmail.com.

0301-4622/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2005.12.010

noises. Since the site specific interaction of protein molecule with the DNA lattice is basic to all the cellular processes including transcription and translation, the fluctuations in the site-specific interaction of protein with the DNA lattice will strongly influence almost all the cellular process, biochemical pathways and networks. Therefore the noise reduction mechanisms those that are developed by the living organisms must be implemented at the DNA–protein interaction levels itself. For example, in case of transcription of the genomic DNA, interaction of RNA polymerase (RNAP, a protein) with the promoter sequences (specific site on the DNA lattice) is the basic/initiating process that is prone to higher fluctuations since the time taken by the RNAP molecule to locate the promoter sequence is a random quantity. Therefore the transcription initiation itself is a stochastic event which in turn is strongly dependent on the initial position of the RNAP molecule on the DNA lattice. In other words the transcription rate must be a time dependent random quantity. However, recently we have shown [17] that when the RNAP molecule searches for the promoter sequence by random jumps with certain critical jump sizes (j, in base-pairs) $j=k_c \geq 2N^{2/3}$ where N is the size of the genome under consideration, the target finding rate is almost independent on the initial position of the RNAP molecule on the DNA lattice. Here one should note that almost all the earlier works in stochastic gene expression strictly assumed a constant transcription rate which is somewhat an over simplification of the underlying process. Moreover it was assumed [19] that under steady state conditions, the Fano factor associated with the fluctuations in the number of mRNA molecules approaches unity or approaches the Poisson limit. However when the transcription rate is a time dependent random quantity, the Fano factor associated with the number of mRNA molecules must be different from unity. In this article we investigate the effects of randomness in the transcription rates on the fluctuations in the number of mRNA molecules and the corresponding Fano factor. Our studies show that the fluctuations in the transcription rate as well as the deviation of the Fano factor associated with the number of mRNA molecules from unity can be significantly reduced when the RNAP molecule searches for the promoter sequences on the DNA lattice by random jumps with higher jump sizes.

2. Transcription initiation is a random event

Let us start with considering a stretch of linear DNA lattice of N base-pairs in length containing a promoter, coding sequence and a terminator of a single gene (this is the minimal configuration of a simple house-keeping gene). We assume that RNAP interacts with the promoter in two-steps viz. RNAP first non-specifically binds to DNA which then performs one dimensional random search for the promoter sequence under non-specifically bound condition [24–26]. Let us assume that the promoter is situated at the lattice position a such that $0<a<N$ where the set of lattice points $\{0,N\}$ constitutes the reflecting boundaries and the lattice point $x=a$ is the only absorbing boundary i.e. whenever RNAP hits the promoter site, the transcription process initiates. Now let us assume that the RNA-polymerase (RNAP) molecule was situated at the lattice position $x=0$ at time $t=0$ and currently performing a one dimensional unbiased random jumps with a jump size of j base-pairs (bps) along the DNA lattice to locate the promoter sequence in order to initiate the transcription event which is then followed by the synthesis of mRNA transcript from the DNA template. When $j=1$ the probability of observing the RNA polymerase at an arbitrary position x on the DNA lattice at time $t>0$ is given as follows:

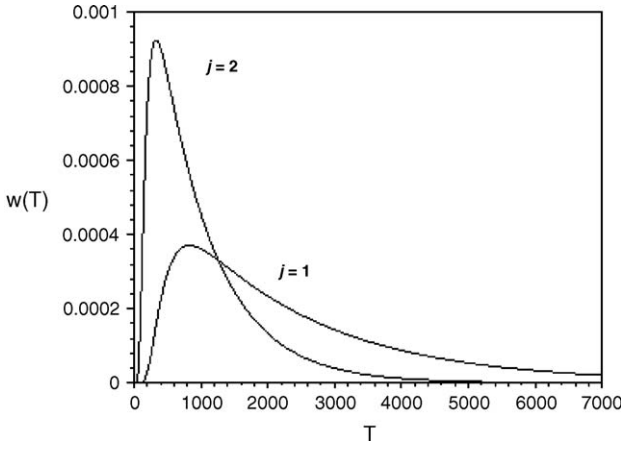

Fig. 1. Variation in the sharpness of the distribution of initiation times of transcription with the jump sizes. Distributions were computed from Eq. (2) with the parameters $a=100$, $k_d=1$ and $D=k_d\frac{(j+1)(2j+1)}{6}$. Here the summation in Eq. (2) was carried out until $n=5000$.

$$P(x,t|0,0)=\frac{2}{a}\sum_{n=1}^{\infty} e^{-\frac{(2n-1)^2\pi^2 D}{8a^2}t}\cos\left((2n-1)\frac{\pi}{2}\frac{x}{a}\right). \quad (1)$$

Eq. (1) is obtained by solving the one dimensional Fokker Plank equation $\frac{D}{2}\partial_x^2 P=\partial_t P$ with the initial condition $P(x,t_0|x_0, t_0)=P(x,0|0,0)=\delta(x)$, using Eigen function expansion method in the interval $0<x<N$ where the set of lattice points $\{0,N\}$ constitutes the reflecting boundary and the lattice point $x=a$ is the probability sink or the absorbing boundary and $D=k_d\frac{(j+1)(2j+1)}{6}$ is the phenomenological one dimensional diffusion coefficient [18] associated with the RNAP molecule where j is the jump size which is assumed to be unity in Eq. (1) and k_d is the maximum three dimensional diffusion controlled collision rate. The jump size j can be positively correlated with the degree of super coiling of the template DNA [Ref. 18] under consideration. The time T taken by the RNA polymerase molecule to find the promoter for the first time (the first passage time) can be shown to be distributed as follows:

$$\varpi(T)=\frac{\pi D}{2a^2}\sum_{n=1}^{\infty}(-1)^{n+1}(2n-1)e^{-\frac{(2n-1)^2\pi^2 D}{8a^2}T}. \quad (2)$$

The mean and the variance associated with the distribution $\varpi(T)$ can be easily computed as follows:

$$m_T=\int_0^{\infty} T\varpi(T)\mathrm{d}T=\langle T\rangle=\frac{a^2}{D} \quad (3)$$

$$v_T=\int_0^{T} T^2\varpi(T)\mathrm{d}T-m_T^2=\langle T^2\rangle-\langle T\rangle^2=\frac{2a^4}{3D^2}. \quad (4)$$

From Eqs. (2) and (4) we can conclude that the sharpness of the distribution of the initiation times of the transcription event can be significantly increased (see Fig. 1) by increasing the jump size associated with the dynamics of RNAP on the DNA lattice in the process of searching for the promoter sequences.

3. Time independent transcription rates

Suppose if we denote the time taken by the transcription machinery to generate a full mRNA transcript is τ_T then the average transcription rate μ_k can be easily given as $\mu_k=[m_T+\tau_T]^{-1}$ where m_T is defined in Eq. (3). Now the entire transcriptional event an be described by the following scheme.

$$\text{RNAP} + \text{Promoter} \xrightarrow{m_T^{-1}} \text{Transcription} - \text{Initiation} \xrightarrow{\tau_T^{-1}} \text{mRNA} \xrightarrow{\gamma_m} \text{Decay}$$

Here we should note that the reaction scheme depicted above is also called as a doubly stochastic Poisson point process. Such doubly stochastic Poisson point process presented in this reaction scheme have been shown to lead to Fano factors greater than unity [6]. Now the birth–death master equation associated with the probability $P_{m,t}$ of observing m number of mRNA molecules at a given time t can be described by the following differential difference equation with the initial condition $P_{m_0,0}=1$.

$$\partial_t P_{m,t} = \mu_k P_{m-1,t} + \gamma_m(m+1)P_{m+1,t} - (\mu_k + \gamma_m m)P_{m,t} \tag{5}$$

Here μ_k is the average transcription rate and γ_m^{-1} is the lifetime associated with the mRNA molecule inside the cell.

Eq. (5) can be solved by the method of generating functions. Now defining the generating function as $G(s,t)=\sum_m s^m P_{m,t}$, one can rewrite Eq. (5) as follows:

$$\partial_t G = \mu_k(s-1)G + \gamma_m(1-s)\partial_s G. \tag{6}$$

Eq. (6) is a first order partial equation which can be solved by the methods of characteristics.

$$G(s,t) = F[(1-s)e^{-\gamma_m t}]e^{\frac{\mu_k s}{\gamma_m}} \tag{7}$$

Now using the initial condition for the generating function $G(s,0)=s^{m_0}$ we obtain the solution to Eq. (6) as follows:

$$G(s,t) = [1-(1-s)e^{-\gamma_m t}]^{m_0} e^{-\frac{\mu_k}{\gamma_m}(1-s)[1-e^{-\gamma_m t}]}. \tag{8}$$

From the definition of generating function $G(s,t)=\sum_m s^m P_{m,t}$ we obtain the probability distribution function $P_{m,t}$ as follows:

$$P_{m,t} = \frac{e^{-\frac{\mu_k}{\gamma_m}[1-e^{-\gamma_m t}]}}{m!} \sum_{n=0}^{m} {}^{m_0}C_n e^{-\gamma_m n t}[1-e^{-\gamma_m t}]^{m_0-n}\left[\frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t})\right]^{m-n}. \tag{9}$$

Noting the initial condition $P_{m_0,0}=1$, at the limit $t\to\infty$ we recover the Poisson distribution of number of mRNA molecules.

$$P_{m,\infty} = \frac{1}{m!} e^{-\frac{\mu_k}{\gamma_m}}\left[\frac{\mu_k}{\gamma_m}\right]^m \tag{10}$$

Now the mean $\langle m_t\rangle$ and the variance $\text{var}\{m_t\}$ associated with the number of mRNA molecules at any time t can be derived directly from the generating function.

$$\langle m_t\rangle = \lim_{s\to 1}\partial_s G = m_0 e^{-\gamma_m t} + \frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t}) \tag{11}$$

$$\text{var}\{m_t\} = \langle m_t^2\rangle - \langle m_t\rangle^2 = m_0 e^{-\gamma_m t}(1-e^{-\gamma_m t}) + \frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t}) \tag{12}$$

Noting the limits $\lim_{t\to\infty}\text{var}\{m_t\}=\frac{\mu_k}{\gamma_m}$ and $\lim_{t\to\infty}\langle m_t\rangle=\frac{\mu_k}{\gamma_m}$ we obtain the steady state limit for the Fano factor $F_{m,t}^{\gamma_m}=\frac{\text{var}\{m_t\}}{\langle m_t\rangle}$ associated with the number of mRNA molecules as follows:

$$F_{m,\infty} = \lim_{t\to\infty}\frac{\text{var}\{m_t\}}{\langle m_t\rangle} = 1. \tag{13}$$

Here the Fano factor is the measure of noises in the number of mRNA molecules at any time. In our case, if the Fano factor is close to unity we can say that the extent of noise/fluctuations in the number of mRNA molecules is a minimum. Whereas deviations of Fano factor from unity can be positively correlated with the increasing noise strength.

4. Time dependent transcription rates

So far we have assumed that the transcription rate is a constant quantity. However as we have seen in the earlier part of this section that the initiation of transcription is in turn a random quantity. Since the mean first passage time associated with the initiation of transcription is distributed as Eq. (2), one can write the transcription rate at any moment as $\mu_k=[T+\tau_T]^{-1}$. In other words one can write the transcription rate as $k_t=\mu_k+\xi_k(t)$ where k_t is the time dependent transcription rate and $\xi_k(t)$ is the delta correlated Gaussian white noise with $\langle\xi_k(t)\rangle=0$ and $\langle\xi_k(t)\xi_k(t')\rangle=D_k\delta(t-t')$. Since the variance associated with distribution of the initiation times T is $v_T=\frac{2a^4}{3D^2}$, one can easily conclude that $D_k=\alpha\frac{2a^4}{3D^2}$ where α is the proportionality constant. Now we can rewrite the master Eq. (5) as follows:

$$\partial_t P_{m,t} = k_t P_{m-1,t} + \gamma_m(m+1)P_{m+1,t} - (k_t + \gamma_m m)P_{m,t}. \tag{14}$$

Proceeding as in the previous case, we obtain the expression for the generating function as follows:

$$\partial_t G = k_t(s-1)G + \gamma_m(1-s)\partial_s G. \tag{15}$$

Solving Eq. (15) we obtain,

$$G(s,t) = F[(1-s)e^{-\gamma_m t}]e^{\frac{\mu_k s}{\gamma_m} + \int_0^s \xi_k(\gamma_m^{-1}\ln[\tau-1]+t-\gamma_m^{-1}\ln(s-1))d\tau} \tag{16}$$

Using the initial condition as $G(s,0)=s^{m_0}$, we obtain the expression for the arbitrary function as $F[(1-s)] = s^{m_0} e^{-\frac{\mu_k s}{\gamma_m} - \int_0^s \xi_k(\gamma_m^{-1}\ln[\tau-1]-\gamma_m^{-1}\ln(s-1))\mathrm{d}\tau}$.

Now defining the new variables as $u = \frac{1}{\gamma_m}\ln\left[\frac{\tau-1}{s-1}\right]+t$, $\tau-1 = (s-1)e^{\gamma_m u}e^{-\gamma_m t}$ and $v = u-t$, $s'=1-(1-s)e^{-\gamma_m t}$ and noting that $\gamma_m e^{-\gamma_m t}e^{\gamma_m u}\mathrm{d}u=\mathrm{d}\tau$, we obtain the solution for the generating function as follows:

$$G(s,t)=[1-(1-s)e^{-\gamma_m t}]^{m_0} e^{-\frac{\mu_k}{\gamma_m}(1-s)[1-e^{-\gamma_m t}]+\gamma_m(1-s)[e^{-\gamma_m t}\int_0^{s'}\xi_k(v)e^{\gamma_m v}\mathrm{d}v-\int_0^s \xi_k(u)e^{\gamma_m u}\mathrm{d}u]} \tag{17}$$

Now defining $I(s,t) = (1-s)e^{-\gamma_m t}\int_0^{s'}\xi_k(v)e^{\gamma_m v}\mathrm{d}v-\int_0^s \xi_k(u)e^{\gamma_m u}\mathrm{d}u$, Eq. (17) can be rewritten as,

$$G(s,t) = [1-(1-s)e^{-\gamma_m t}]^{m_0} e^{-\frac{\mu_k}{\gamma_m}(1-s)[1-e^{-\gamma_m t}]+\gamma_m I(s,t)}. \tag{18}$$

Using the following relation,

$$\partial_s I(s,t) = (1-s)e^{-\gamma_m t}\partial_s\left[\int_0^{s'}\xi_k(v)e^{\gamma_m v}\mathrm{d}v-\int_0^s \xi_k(u)e^{\gamma_m u}\mathrm{d}u\right] -e^{-\gamma_m t}\int_0^{s'}\xi_k(v)e^{\gamma_m v}\mathrm{d}v+\int_0^s \xi_k(u)e^{\gamma_m u}\mathrm{d}u.$$

The probability distribution $P_{m,t}$ associated with the number of mRNA molecules at any time t can be given as,

$$P_{m,t} = \frac{e^{-\frac{\mu_k}{\gamma_m}[1-e^{-\gamma_m t}]}}{m!}\sum_{n=0}^{m} {}^{m_0}C_n e^{-\gamma_m n t}[1-e^{-\gamma_m t}]^{m_0-n} \times\left\{\left[\frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t})\right]^{m-n}+\gamma_m|\partial_s^{m-n}I(s,t)|_{s=1}\right\}. \tag{19}$$

However it is very difficult to proceed further since there are still random terms in Eq. (19). Therefore, we can try to solve the Fokker Plank analogue of Eq. (15). Before proceeding further, we first derive the conditions of validity of Fokker Plank approximation to the master Eq. (5). The Fokker Plank equation (FPE) associated with master Eq. (5) can be written as follows:

$$\partial_t P_{m,t} = -\partial_m[(k-\gamma_m m)P_{m,t}]+\partial_m^2\left[\frac{k+\gamma_m m}{2}P_{m,t}\right] \tag{20}$$

The approximate stochastic differential associated with FPE Eq. (20) can be calculated to be as,

$$\frac{\mathrm{d}m}{\mathrm{d}t} = (k-\gamma_m m)+\sqrt{k+\gamma_m m}\eta(t). \tag{21}$$

Unfortunately Eqs. (20) and (21) are somewhat complicated differential equations whose analytical solutions cannot be obtained in a closed form. However, we can try the following approximation for the non-linear differential Eq. (21).

$$\frac{\mathrm{d}m}{\mathrm{d}t} = (k-\gamma_m m)+\sqrt{D_m}\eta(t) \tag{22}$$

Where the constant D_m is defined such that $\langle\eta(t)\eta(t')\rangle = D_m\delta(t-t')$ i.e. $\langle\eta(t)\eta(t')\rangle=\delta(t-t')$, where $\eta(t)$ is the delta correlated Gaussian white noise with $\langle\eta(t)\rangle=0$. Now the FPE corresponding to Eq. (22) can be written as follows:

$$\partial_t P_{m,t} = -\partial_m(k-\gamma_m m)P_{m,t}+D_m\partial_m^2 P_{m,t}. \tag{23}$$

Eq. (23) can be solved by Fourier transform methods. Now defining the Fourier transform of $P_{m,t}$ as $\tilde{P}_{w,t} = \int_{-\infty}^{+\infty} P_{m,t}e^{-iwm}\mathrm{d}m$, Eq. (23) can be written as,

$$\partial_t\tilde{P}_{w,t} = -D_m(w^2+i\mu_k w)\tilde{P}_{w,t}-\gamma_m w\partial_w\tilde{P}_{w,t}. \tag{24}$$

Solving Eq. (24) with the initial condition as $\tilde{P}_{w,0}=e^{-iwm_0}$ which is simply $\tilde{P}_{w,0} = \int_{-\infty}^{+\infty} P_{m,0}e^{-iwm}\mathrm{d}m = \int_{-\infty}^{+\infty}\delta(m-m_0)e^{-iwm}\mathrm{d}m$ and by performing the inverse Fourier transform one can obtain the probability distribution function $P_{m,t}$ as follows:

$$P_{m,t} = \frac{\gamma_m^{1/2}}{\sqrt{\pi D_m(1-e^{-2\gamma_m t})}}e^{-\frac{\left[m-m_0e^{-\gamma_m t}-\frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t})\right]^2}{2D_m(2\gamma_m)^{-1}(1-e^{-2\gamma_m t})}}. \tag{25}$$

Here one should note that Eq. (25) is simply the Gaussian approximation of the Poisson type distribution given by Eq. (9) and therefore Eq. (22) can be used as an approximation for the original master Eq. (5). Fig. 2 shows the time evolution of distribution function given by Eq. (25) and Fig. 3 shows the equivalence of the distribution functions given by Eqs. (9) and (25) at $t\to\infty$. Noting that $P_{m,0}=\delta(m-m_0)$, the limit at $t\to\infty$ can be obtained as,

$$P_{m,\infty} = \frac{\gamma_m^{1/2}}{\sqrt{\pi D_m}}e^{-\frac{\left[m-\frac{\mu_k}{\gamma_m}\right]^2}{2D_m(2\gamma_m)^{-1}}}. \tag{26}$$

Now the mean and the variance associated with the number of mRNA molecules at any time t can be directly

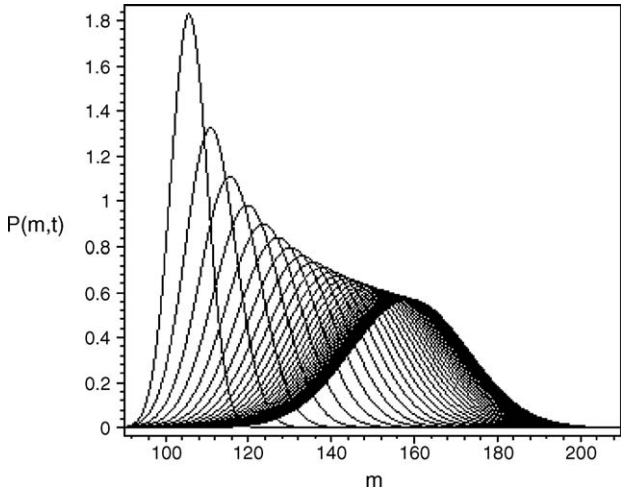

Fig. 2. Time evolution of the probability distribution $P_{m,t}=P(m,t)$ associated with the number of mRNA molecules with time where Eq. (25) was used to compute distributions at different times. Here the parameters are $P_{m_0,0}=\delta(m-m_0)$, $m_0=100$, $\gamma_m=0.01$, $\mu_k=1.6$ and $D_m=1$. The distributions were computed with initial time $t=0.05$ and the final time $t=2$ with the increments of $\Delta t=0.05$.

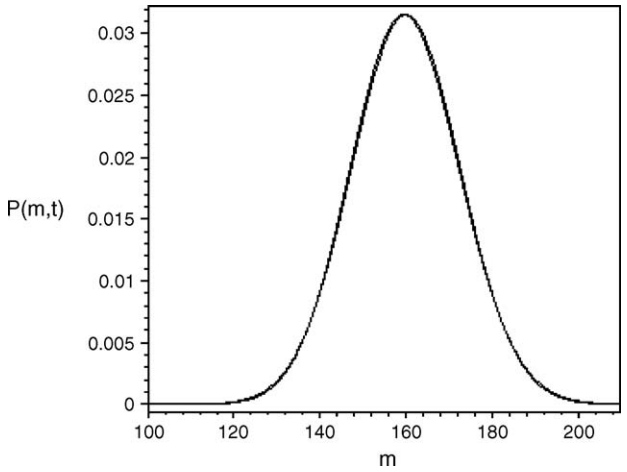

Fig. 3. Equivalence of the distributions given by Eqs. (9) and (25) at $t\to\infty$. Here the parameters are $m_0=100$, $\gamma_m=0.01$, $\mu_k=1.6$ and $D_m=1$. The distributions were computed at the time $t=2$.

obtained by integrating the stochastic differential Eq. (22) as follows:

$$m_t = m_0 e^{-\gamma_m t} + e^{-\gamma_m t}\int_0^t [\mu_k + \eta(\tau)]e^{\gamma_m \tau}\mathrm{d}\tau \tag{27}$$

And subsequently by averaging m_t over time we obtain,

$$\langle m_t\rangle = m_0 e^{-\gamma_m t} + \frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t}) \tag{28}$$

$$\mathrm{var}\{m_t\} = \langle m_t^2\rangle - \langle m_t\rangle^2 = \frac{D_m}{2\gamma_m}\left(1-e^{-2\gamma_m t}\right) \tag{29}$$

By comparing Eqs. (12) and (29) we obtain the conditions of the validity of the stochastic differential Eq. (22) at $t\to\infty$ as follows:

$$\frac{D_m}{2\gamma_m} = \frac{\mu_k}{\gamma_m}. \tag{30}$$

This implies that $D_m=2\mu_k$. With this background, now the stochastic differential equation corresponding to the master Eq. (14) can be written as,

$$\frac{\mathrm{d}m}{\mathrm{d}t} + \gamma_m m = k_t + \eta(t) \tag{31}$$

where the time dependent transcription rate is defined as $k_t=\mu_k+\xi_k(t)$ with $\langle\xi_k(t)\rangle=0$ and $\langle\xi_k(t)\xi_k(t')\rangle=D_k\delta(t-t')$. Here we should note that $D_k=\alpha\frac{2a^4}{3D^2}$, where α is the proportionality constant. Using Kramers–Moyal expansion, the FPE corresponding to Eq. (31) can be written as,

$$\partial_t P_{m,t} = -\partial_m(\mu_k-\gamma_m m)P_{m,t} + (D_m+D_k)\partial_m^2 P_{m,t}. \tag{32}$$

By solving Eq. (32) using Fourier transform method, we obtain the expression for the probability distribution function $P_{m,t}$ in the presence of fluctuating transcription rate as follows:

$$P_{m,t} = \frac{\gamma_m^{1/2}}{\sqrt{\pi(D_m+D_k)(1-e^{-2\gamma_m t})}}\, e^{-\frac{\left[m-m_0e^{-\gamma_m t}-\frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t})\right]^2}{2(D_m+D_k)(2\gamma_m)^{-1}(1-e^{-2\gamma_m t})}}. \tag{33}$$

From Eq. (33) we obtain the limit at $t\to\infty$ as,

$$P_{m,\infty} = \frac{\gamma_m^{1/2}}{\sqrt{\pi(D_m+D_k)}}\, e^{-\frac{\left[m-\frac{\mu_k}{\gamma_m}\right]^2}{2(D_m+D_k)(2\gamma_m)^{-1}}}. \tag{34}$$

Now the mean and the variance associated with the number of mRNA molecules can be given as follows:

$$\langle m_t\rangle = m_0 e^{-\gamma_m t} + \frac{\mu_k}{\gamma_m}(1-e^{-\gamma_m t}) \tag{35}$$

$$\mathrm{var}\{m_t\} = \langle m_t^2\rangle - \langle m_t\rangle^2 = \frac{D_m+D_k}{2\gamma_m}(1-e^{-2\gamma_m t}). \tag{36}$$

Here one should note that though the mean value associated with the number of mRNA molecules at any time t i.e. m_t is not much affected by the fluctuations in the transcription rate μ_k, the variance associated with m_t increases with the strength of fluctuations. Noting the definition of $D_k=\frac{24\alpha a^4}{k_d[(j+1)(2j+1)]^2}$ where j is the jump size associated with the dynamics of RNAP molecule on the DNA lattice in the process of searching for the promoter sequences, and noting the condition $D_m=2\mu_k$, the Fano factor associated with the fluctuations in the number of mRNA molecules under steady state condition as $t\to\infty$ can be written as follows:

$$F_{m,\infty} = \frac{D_m+D_k}{2\gamma_m}\frac{\gamma_m}{\mu_k} = 1+\frac{D_k}{2\mu_k} = 1+\frac{12\alpha a^4}{k_d\mu_k[(j+1)(2j+1)]^2}. \tag{37}$$

Eq. (37) clearly states that the fluctuations in the transcription rate indeed strongly influence the Fano factor associated with the fluctuations in the number of mRNA molecules. However, when RNAP molecule search for the promoter sequences by random jumps with higher jump sizes, we can easily conclude from Eq. (37) that,

$$\lim_{j\to\infty} F_{m,\infty} = \lim_{j\to\infty}\left[1+\frac{12\alpha a^4}{k_d^2\mu_k[(j+1)(2j+1)]^2}\right] = 1. \tag{38}$$

That is to say, we recover the Poisson limit given by Eq. (13) at $j\to\infty$. However the limit $j\to\infty$ is meaningless in real situations since RNAP cannot make such infinite jumps and also it is violating our two-step assumption i.e. the RNAP that is non-specifically bound to the DNA lattice cannot make large jumps with jump size $j>N$ without getting dissociated from the DNA lattice. Moreover, recently we have shown [17] that when the jump size j is close to or beyond certain critical values $j>k_c\approx 2N^{2/3}$ where N is the size of the genome under consideration, the mean and the variance associated with the distribution of the transcription initiation times i.e. $\varpi(T)$ modifies as $\lim_{j\geq k_c}\langle T\rangle=\frac{N}{k_d}$ and $\lim_{j\geq k_c}\mathrm{var}\{T\}=\frac{2N^2}{3k_d^2}$. Therefore under the condition that $j>k_c\approx 2N^{2/3}$, the expression for D_k modifies to $D_k=\frac{24\alpha N^2}{k_d^2}$ from which we obtain the expression for the modified Fano factor under the critical jump limit $j>k_c\approx 2N^{2/3}$ as follows:

$$\lim_{j>k_c} F_{m,\infty} = 1+\frac{12\alpha N^2}{\mu_k k_d^2}. \tag{39}$$

5. Discussion

Under the condition that $j > k_c \approx 2N^{2/3}$ Eq. (39) clearly states that the Fano factor associated with the fluctuations in the number of mRNA molecules is inversely proportional to the transcription rate and directly proportional to the mean value of transcription initiation time m_T. Since the jump size associated with dynamics of RNAP molecule on the DNA lattice is positively correlated with the degree of super coiling of the DNA under consideration, our results indirectly imply that the super coiling of DNA is the main factor that controls the noise level especially at the transcriptional level and we argue that the super coiled or close-packed structure of DNA might have evolved to keep the noises at the transcriptional level minimum apart from the advantage of getting packed into the cellular compartments. Moreover Eq. (39) clearly states that under the critical jump limit $j > k_c \approx 2N^{2/3}$ the deviation of the Fano factor associated with the fluctuations in the number of mRNA molecules from the Poisson limit will be significant only when the genome size under consideration is much larger. For example, in case of *E. coli* the size of the genome is in the order of $N \sim 10^6$ bps and the diffusion controlled collision rate is in the order of $k_d \sim 10^8$. Therefore we can conclude that $\lim_{j>k_c} F_{m,\infty} = 1 + O(10^{-4}) \approx 1$. However when the genome size is much larger as in case of higher organisms we can conclude that the Fano factor associated with the fluctuation in the number of mRNA molecules must be significantly deviating from the Poisson limit.

6. Conclusions

Among the effects of noises at various levels of cellular organization, the noise in the central process such as transcription, replication and translation of the genomic DNA is very important since it can directly affect the phenotypic and behavioral aspects of an organism as well as the entire cellular function. Here we show that the initiation of the transcription is a stochastic event and the transcription rates are time dependent random quantities. We derive the central measure properties, i.e. mean and the variance, of the distribution of transcription rates. Our results show that the Fano factor which is a measure of deviation from the Poisson distribution associated with the fluctuations in the number of mRNA molecules deviates from unity due to the randomness in the transcription rates. However when the RNA polymerase molecule searches for the promoter sequences on the DNA lattice by random jumps, the Fano factor approaches the Poisson limit as the jump size associated with the RNA polymerase increases. Since the jump size associated with dynamics of RNAP molecule is positively correlated with the degree of super coiling of DNA we argue that the super coiled or close-packed structure of DNA might have evolved to keep the noises at the transcriptional level in a minimum. Finally we also show that the Fano factors associated with the fluctuations in the number of mRNA molecules directly proportional to the genome size of the organism under consideration.

Acknowledgements

This work is supported by TIFR, Mumbai. The author is the recipient of Kanwal-Rekhi fellowship for career development.

References

[1] B. Lewin, Genes VIII, Oxford University Press, 2004.
[2] M.C. Teich, IEEE Trans. Biomed. Eng. 36 (1989) 150.
[3] R.G. Turcott, S.B. Lowen, E. Li, D.H. Johnson, C. Tsuchitani, M.C. Teich, Biol. Cybern. 70 (1994) 209.
[4] S.B. Lowen, M.C. Teich, Phys. Rev., A 43 (1991) 4192.
[5] S.B. Lowen, M.C. Teich, J. Acoust. Soc. Am. 99 (1996) 3585.
[6] S.B. Lowen, M.C. Teich, Fractal-Based Point Processes, Wiley-Interscience, 2005.
[7] O.G. Berg, J. Paulsson, M. Ehrenberg, Biophys. J. 79 (2000) 2944.
[8] O.G. Berg, J. Theor. Biol. 71 (1978) 587.
[9] H.H. McAdams, A. Arkin, Trends Genet. 15 (1999) 65.
[10] A. Arkin, J. Ross, H.H. McAdams, Genetics 149 (1998) 1633.
[11] P. van de Putte, N. Goosen, Trends Genet. 8 (1992) 457.
[12] M.S.H. Ko, J. Theor. Biol. 153 (1991) 181.
[13] J. Hasty, J. Pradines, M. Dolnik, J.J. Collins, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 2075.
[14] H. McAdams, A. Arkin, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 814.
[15] T.S. Gardner, C.R. Cantor, J.J. Collins, Nature (London) 403 (2000) 339.
[16] M.B. Elowitz, S. Leibler, Nature (London) 403 (2000) 335.
[17] J. Paulsson, O.G. Berg, M. Ehrenberg, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 7148.
[18] J. Paulsson, M. Ehrenberg, Phys. Rev. Lett. 84 (2000) 5447.
[19] M. Thattai, A. van Oudenaarden, Proc. Natl. Acad. Soc. U. S. A. 98 (2001) 8614.
[20] E.M. Ozbudak, M. Thattai, I. Kurtser, A.D. Grossman, A. van Oudenaarden, Nat. Genet. 31 (2002) 69.
[21] J.M.G. Vilar, H.Y. Kueh, N. Barkai, S. Leibler, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 5988.
[22] N. Barkai, S. Leibler, Nature (London) 403 (2000) 267.
[23] R. Murugan, Biophys. Chemist. 120 (2006) 143.
[24] R. Murugan, Phys. Rev., E 69 (2004) 011911.
[25] O.G. Berg, R.B. Winter, P.H. Von. Hippel, Biochemistry 20 (1981) 6929.
[26] O.G. Berg, R.B. Winter, P.H. Von. Hippel, Biochemistry 20 (1981) 6961.