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:

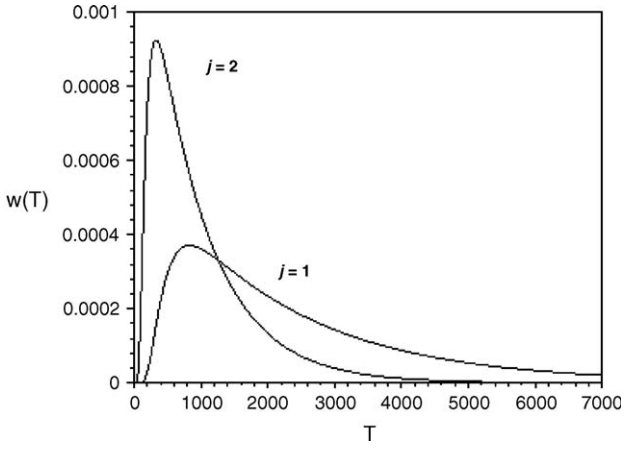

Fig. 1. Variation in the sharpness of the distribution of initiation times of transcription with the jump sizes. Distributions were computed from Eq. (2) with the parameters $a=100$, $k_d=1$ and $D=k_d\frac{(j+1)(2j+1)}{6}$. Here the summation in Eq. (2) was carried out until $n=5000$.

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

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:

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

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