Mathematical Biosciences 283 (2017) 84–90

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Mathematical aspects of the regulation of gene transcription by promoters

Vladimir P. Zhdanov[a,b,*]

[a] Section of Biological Physics, Department of Physics, Chalmers University of Technology, S-41296 Göteborg, Sweden
[b] Boreskov Institute of Catalysis, Russian Academy of Sciences, Novosibirsk 630090, Russia

ARTICLE INFO

Article history:
Received 11 June 2016
Revised 1 September 2016
Accepted 5 November 2016
Available online 8 November 2016

Keywords:
Transcriptional regulation
Transcriptional factors
Lateral interactions
Theory of rate processes
Stochastic kinetics
Bursts

ABSTRACT

Although the transcriptional regulation of gene expression has been a subject of intense experimental and theoretical studies over the past several decades, the understanding of this process is still incomplete. In particular, the models focused on the function of transcription factors usually take into account only the lateral interactions between them in the ground bound state. The rates of attachment and detachment of transcription factors on the promoter sites depend, however, also on the lateral interactions in the activated state. I present general equations describing the effect of both these interactions on the rates of attachment and detachment and illustrate their role in the kinetics of gene expression by using a generic model focused on the function of a gene regulated via two promoter sites. The corresponding analytical treatment and Monte Carlo simulations show that the lateral interaction in the activated state is significant if the genes are expressed in the regime of stochastic bursts of high and low transcriptional activity and RNA and protein populations. In particular, the duration and shape of bursts depend on this interaction.

1. Introduction

Gene expression and the corresponding kinetics are obviously of high interest from very different perspectives. Gene transcription to RNA, performed by polymerase (or, more specifically, by RNA polymerase II), is here one of the central steps. This step is often highly regulated. After several decades of experimental studies, the general principles of the regulation are considered to be well established [1,2]. In particular, the transcription is believed to be controlled by *cis*-regulatory elements, such as promoters and enhancers. Promoters can be located upstream of a transcription start site as well as be spread over a considerable distance [1]. Typically, promoters include common sequence elements, such as a TATA box and an initiator (INR), and other binding sites for transcription factors (chemically, these factors are proteins). Enhancers, stimulating the activity of promoters, also contain binding sites for transcription factors. Their role is in the recruitment of histone-modifying enzymes that create a more favourable chromatin configuration or of a kinase that induces a bound initiation complex to begin transcription.

Experimental investigations of gene expression in general and transcriptional regulation in particular were accompanied by numerous theoretical studies focused on construction and analysis of the corresponding kinetic models. Concerning gene expression in general, we refer to the reviews focused on stochastic effects [3,4], oscillations [5,6], non-coding RNAs [7], and complex genetic networks [8–11]. The bulk of the models related to transcriptional regulation is focused on the function of transcription factors on the promoter sites [12–16]. The conventional approach, used in this area since the pioneering study by Ackers et al. [17], is based on two related assumptions [16]. The first one is that the transcription rate is limited by the initiation, and that the whole rate of the process can be represented as

$$w = \sum_i p_i w_i, \tag{1}$$

where p_i are the probabilities of specific arrangements of bound transcription factors or transcription factors and RNA polymerase, and w_i are the initiation rates corresponding to these arrangements. The second assumption is that the attachment and detachment of these species to DNA are rapid on the time scale of transcription but slow on the time scale of their diffusion so that this subsystem is close to equilibrium, and accordingly the probabilities p_i are determined by the grand canonical distribution (alternatively, the process may depend on transcription-factor diffusion,

* Correspondence to: Boreskov Institute of Catalysis, Russian Academy of Sciences, Novosibirsk 630090, Russia.
E-mail address: zhdanov@catalysis.ru

http://dx.doi.org/10.1016/j.mbs.2016.11.007
0025-5564/

binding, and sliding on DNA [18]). In this case, the explicit expressions for p_i depend on the concentrations of transcription factors inside the cell and lateral (along the DNA) interactions between these factors in the bound state.

If the attachment and detachment of transcription factors occur rapidly, the fluctuations of the RNA or mRNA and protein populations are typically not dramatic and the whole kinetics is close to or at least not far from that predicted by the mean-field kinetic equations. In reality, the ratio between the rates of attachment, detachment, and transcription initiation may be arbitrary. If the former processes are slow, the whole kinetics may look as a series of stochastic bursts of high and low transcriptional activity and RNA and protein populations. Stochastic bursts may also be related to feedback-induced bistability of the kinetics. Such features of gene expression have important consequences for cellular function, being beneficial in some contexts and harmful in others [19–21]. The corresponding theoretical studies are numerous (see e.g. already mentioned reviews [3–16], recent original studies [22–28], and references therein). In fact, expression (1) often remains to be applicable in this case provided p_i are treated as stochastic variables. The kinetic models focused on stochastic effects are typically coarse-grained and based on the use of the "on" and "off" rate constants without specification of the details of the attachment and detachment processes or on the use of the simplest expressions (of the Langmurian or Hill type) for the regulation of the transcription.

Despite the long-standing interest in the mechanisms of transcriptional regulation by promoters, the corresponding theory is still incomplete. Herein, we focus on the way of describing the attachment and detachment rates. As already noted, the available models take into account that these rates are influenced by lateral interactions between the transcription factors in the bound state (to be specific, we refer to transcription factors although our arguments are applicable to RNA polymerase as well). In reality, attachment and detachment are activated processes, and referring to the transition-state theory [29] or Brønsted–Evans–Polanyi relations [30], one can argue that the attachment and detachment rates depend on the energy of the activated state, and in analogy with the bound state this energy is in turn also dependent on the lateral interaction due to steric constrains during attachment. The theory of rate processes complicated by lateral interactions in the bound (ground) and activated states was well developed and is widely used in the context of heterogeneous catalysis [31]. Following the prescriptions of this theory, one can introduce the attachment and detachment rate constants corresponding to given arrangement of bound transcription factors and represent the corresponding rate constants, respectively, as

$$k_i = k_\circ \exp(-\epsilon_i^*/k_B T), \tag{2}$$

$$\kappa_i = \kappa_\circ \exp[-(\epsilon_i^* - \epsilon_i)/k_B T], \tag{3}$$

where ϵ_i and ϵ_i^* are the lateral interactions in the bound (ground) and activated states, and $k_\circ$ and $\kappa_\circ$ are the corresponding rate constants in the absence of these interactions. The physics behind Eqs. (2) and (3) is simple. The repulsive or attractive lateral interaction in the activated state, $\epsilon_i^* > 0$ or $\epsilon_i^* < 0$, shifts the energy of this state upwards or downwards, the activation energy for attachment and detachment increases or decreases, and accordingly the attachment and detachment rate constants decrease or increases, respectively. The repulsive or attractive lateral interaction in the ground state, $\epsilon_i > 0$ or $\epsilon_i < 0$, does not influence the activated state but shifts upwards or downwards the energy of the bound species, the activation energy for detachment decreases or increases, and accordingly the detachment rate constant increases or decreases while the attachment rate remains the same.

In the available models of transcriptional regulation [12–16], the lateral interaction in the ground state is usually taken into account while the interaction in the activated state is ignored. Coulon et al. [32] extended this formalism. They operate with a Gibbs free energy of each promoter state and the energies that must be overcome for transitions. Physically, the corresponding expressions are similar to (2) and (3). In their work (and also in a recent related review [33]), the role of the lateral interaction in the activated state has, however, not been scrutinized and discussed in detail.

To illustrate the effect of the lateral interactions ϵ_i^* and ϵ_i on the kinetics of gene expression, we first of all articulate that the activated state for attachment is the same as for detachment. Physically, this means that the lateral interactions in the activated state, ϵ_i^*, do not influence the attachment-detachment equilibrium. If the attachment and detachment processes are rapid, one can operate with the averaged transcription rate [Eq. (1)] corresponding to equilibrium. In this limit, the kinetics will be close to those predicted by the conventional mean-field equations, and the results will be independent of ϵ_i^*.

In fact, the kinetic equations can be formulated at the level of probabilities of different states of the genes and the mRNA and protein populations (Section 3). In the absence of feedbacks, such equations allow one to calculate *exact* average values of these probabilities and populations irrespective of the level of stochasticity of the kinetics (Section 3). With this reservation, we note that the regime of stochastic bursts is more sensitive to the details of the transcriptional regulation in general and to the lateral interaction in the activated state in particular (Section 4). In our present study, we illustrate all these points in detail by employing a minimal generic model (Section 2) focused on the function of a gene regulated via attachment and detachment of transcription factors on two regulatory sites.

2. Model

Our model includes synthesis of mRNA on a gene existing in a single copy, translation of mRNA to protein (P), and degradation of these species,

$$\text{Gene} \rightarrow \text{Gene} + \text{mRNA}, \tag{4}$$

$$\text{mRNA} \rightarrow \text{mRNA} + \text{P}, \tag{5}$$

$$\text{mRNA} \rightarrow \oslash, \quad \text{and} \quad \text{P} \rightarrow \oslash. \tag{6}$$

Step (4) is considered to be regulated by two transcription factors or, more specifically, by proteins P_1 and P_2 (Fig. 1), which are produced via expression of two other genes, $Gene_1$ and $Gene_2$ (these genes may exist in a few copies), or, more specifically, via translation of $mRNA_1$ and $mRNA_2$ in analogy with steps (4)–(6). Thus, the scheme under consideration include three genes, Gene, $Gene_1$, and $Gene_2$, and six related species, mRNA, $mRNA_1$, $mRNA_2$, P, P_1, and P_2. The populations of these species are designated as N_m, N_{m1}, N_{m2}, N_p, N_{p1}, and N_{p2}, respectively. Focusing on the regulation of step (4) via attachment and detachment of P_1 and P_2 on two proximal promoter sites, we do not describe explicitly the regulation of the other steps. With this specification, P_1 and P_2 can be in the unbound and bound states while the other species are in the unbound state. The introduced populations, N_{p1} and N_{p2}, are considered to correspond to the unbound forms of P_1 and P_2.

The two promoter sites can be in four states (Fig. 1) with the P_1 and P_2 occupations 0,0, 1,0, 0,1, and 1,1, respectively. The probabilities to be in these states are designated as P_{00}, P_{10}, P_{01}, and P_{11}. In this case, there is only one lateral interaction in the bound (ground) state, 1,1, with the both sites occupied. This interaction is designated as ϵ. In addition, there are two lateral interactions in the activated states for the P_1 or P_2 attachment to the sites in states 0,1 and 1,0, respectively, or for the P_1 or P_2 detachment from