A generic 3D kinetic model of gene expression

Citation for the original published paper (version of record):
Zhdanov, V. (2012). A generic 3D kinetic model of gene expression. Central European Journal of Physics, 10(2): 533-537. http://dx.doi.org/10.2478/s11534-012-0011-5

(article starts on next page)

Cent. Eur. J. Phys. • 10(2) • 2012 • 533-537
DOI: 10.2478/s11534-012-0011-5

VERSITA

Central European Journal of **Physics**

A generic 3D kinetic model of gene expression

Short Communication

Vladimir P. Zhdanov[1,2]*

1 *Department of Applied Physics, Chalmers University of Technology, S-41296 Göteborg, Sweden*

2 *Boreskov Institute of Catalysis, Russian Academy of Sciences, Novosibirsk 630090, Russia*

Received 21 December 2011; accepted 21 January 2012

Abstract: Recent experiments show that mRNAs and proteins can be localized both in prokaryotic and eukaryotic cells. To describe such situations, I present a 3D mean-field kinetic model aimed primarily at gene expression in prokaryotic cells, including the formation of mRNA, its translation into protein, and slow diffusion of these species. Under steady-state conditions, the mRNA and protein spatial distribution is described by simple exponential functions. The protein concentration near the gene transcribed into mRNA is shown to depend on the protein and mRNA diffusion coefficients and degradation rate constants.

PACS (2008): 87.10.-e, 87.16.-b

Keywords: subcellular processes • gene transcription • mRNA synthesis • mRNA translation • mRNA and protein diffusion and degradation

1. Introduction

Gene expression in cells includes polymerase-mediated gene transcription into mRNAs and non-coding RNAs and mRNA translation by ribosomes into proteins. Numerous kinetic models of the interplay of these processes are temporal (see, e.g., a general review by Kaern et al. [1] and reviews focused on stochastic effects [2–4], oscillations [5, 6], non-coding RNAs [7], and complex genetic networks [8–11]). For example, the simplest equations for the mRNA and protein populations, N and n, in a cell are as follows

$$dN/dt = w - kN, \tag{1}$$

$$dn/dt = \upsilon N - \kappa n, \tag{2}$$

where w is the transcription rate, υ is the translation rate constant, and k and κ are the mRNA and protein degradation rate constants. Under steady-state conditions, these equations yield

$$N = w/k, \tag{3}$$

$$n = \upsilon w/(\kappa k). \tag{4}$$

The mRNA and protein formation and degradation are described here on the level of rates or rate constants. In reality, these kinetic parameters depend on the concentrations of polymerase, transcription factors, ribosome, and the enzymes processing mRNAs and proteins.

The temporal equations for describing the kinetics of gene expression do not specify the location of RNAs and proteins. To justify these equations, RNAs and proteins

*E-mail: zhdanov@catalysis.ru

are often assumed to be uniformly distributed in a cell. The conventional theoretical validation of this assumption is based on the estimation of the diffusion length, $l = (D\tau)^{1/2}$, corresponding to the RNA and protein life spans, $\tau = 1/k$ or $1/\kappa$. The RNA and protein gradients are considered to be negligible inside a cell provided that

$$l \gg R, \tag{5}$$

where R is the cell size (e.g., radius if a cell is spherical). According to hydrodynamics, one has $D = k_B T/(6\pi\eta\varrho)$, where ϱ is the RNA or protein radius, and η is the water viscosity. With the corresponding values of the diffusion coefficients, condition (5) is usually fulfilled.

Despite the arguments above, the experimental studies performed during the past decade indicate that RNAs and proteins can be localized both in prokaryotic and eukaryotic cells (see, e.g., recent studies of RNA patterns in bacteria [12–15] and reviews focused on localization of proteins [16, 17] and RNAs [18, 19]). Although the understanding of the biophysics behind such observations is now limited, the results obtained are indicative of at least two general scenarios of localization (reviewed in Ref. [18]). The first scenario implying slow RNA and/or protein diffusion can be validated taking into account that due to macromolecular crowding in cells [20] the RNA and protein diffusion coefficients are often much lower compared to those predicted by hydrodynamics. For this reason, condition (5) may fail, and RNAs and proteins can be localized near the transcription sites, i.e., near chromosomes and/or plasmids. The second general scenario of RNA and protein localization is based on the ability of some of RNAs and/or proteins to aggregate and form complexes (such complexes are often attached to membranes). The latter ("diffusion and capture") scenario can be realized even if the RNA and protein diffusion is fast.

The available theoretical studies of RNA and protein localization are focused on the situations with slow RNA and/or protein diffusion. In particular, we may mention 2D Monte Carlo simulations with mRNA and protein diffusion [21], 3D Monte Carlo simulations with mRNA diffusion [22, 23], mean-field analysis focused on the non-coding RNA diffusion [24], and general mean-field analysis of the stability of solutions of reaction-diffusion equations [25].

Complementing the available experimental and theoretical studies of RNA and protein localization, we present in this Communication a generic 3D spatiotemporal kinetic model of gene expression including mRNA and protein diffusion. Our analysis, aimed primarily at prokaryotic cells, is focused on the case of slow diffusion implying that

$$l \ll R. \tag{6}$$

In this limit, we obtain analytical expressions for the mRNA and protein spatial distributions under steady-state conditions. Despite their simplicity, the results presented are novel (compared to those obtained in Refs. [21–25]) and instructive.

2. General equations

We analyze the transcription of a single gene into mRNA and mRNA translation into protein. The mRNa and protein concentrations, $C(\mathbf{r}, t)$ and $c(\mathbf{r}, t)$, are described inside a cell by the reaction-diffusion equations,

$$\frac{\partial C(\mathbf{r}, t)}{\partial t} = D\nabla^2 C(\mathbf{r}, t) - kC(\mathbf{r}, t) + w(t)\delta(\mathbf{r}), \tag{7}$$

$$\frac{\partial c(\mathbf{r}, t)}{\partial t} = \mathcal{D}\nabla^2 c(\mathbf{r}, t) - \kappa c(\mathbf{r}, t) + \upsilon C(\mathbf{r}, t), \tag{8}$$

where D and $\mathcal{D}$ are the corresponding diffusion coefficients, $\delta(\mathbf{r})$ is the delta function implying that the gene is located at $\mathbf{r} = 0$ and accordingly the mRNA synthesis occurring via gene transcription takes place at $\mathbf{r} = 0$ (the other designations are as in the Introduction). Specifically, the gene is considered to be located in the centre or at one of the poles of a cell. The mRNA translation into protein is assumed to be possible everywhere, and the corresponding rate is proportional to the local mRNA concentration. This assumption, implying that the ribosomes are homogeneously distributed in a cell, is reasonable for prokaryotic cells. In eukaryotic cells, the ribosomes diffuse in the cytoplasm or are attached to the cytosolic side of the membrane of the compartments of the endoplasmic reticulum, and therefore Eq. (8) does not hold. Thus, our model defined by Eqs. (7) and (8), is applicable to prokaryotes. Eq. (7) and the corresponding results may, however, be applicable to eukaryotes as well.

In our treatment, condition (6) is assumed to hold for mRNA and for protein as well. In this case, these species are localized primarily near the gene, and the boundary conditions near the cell membrane become insignificant. Practically, the latter means that we can consider that there are no spatial limitations on diffusion. Adopting this approximation allows us to simplify the integration of Eqs. (7) and (8). In particular, the mRNA concentration can be represented as

$$C(r, t) = \int_{-\infty}^{t} w(t')G(r, t - t')dt', \tag{9}$$

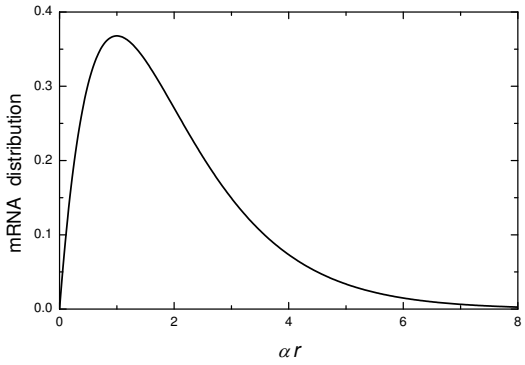

Figure 1. Radial distribution of mRNA according to Eq. (19).

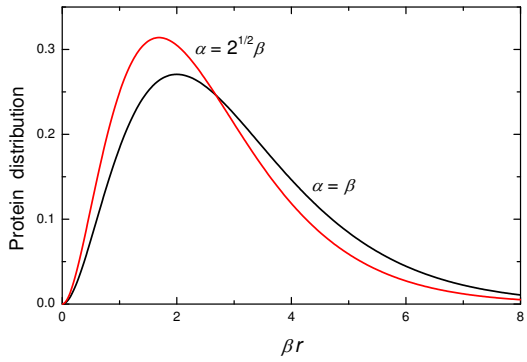

Figure 2. Radial distribution of protein according to Eq. (27) for $\alpha = \beta$ and $\alpha = 2^{1/2}\beta$.

where

$$G(r,t) = (4\pi Dt)^{-3/2} \exp\left(-\frac{r^2}{4Dt} - kt\right) \tag{10}$$

is the Green function for Eq. (7). Substituting (10) into (9) yields

$$C(r,t) = \int_{-\infty}^{t} \frac{w(t')}{[4\pi D(t-t')]^{3/2}} \exp\left(-\frac{r^2}{4D(t-t')} - k(t-t')\right) dt'. \tag{11}$$

For the protein concentration, we have

$$c(\mathbf{r},t) = \upsilon \int_{-\infty}^{t} \int_{V} C(\mathbf{r},t) g(|\mathbf{r}-\mathbf{r}'|, t-t') d^3r' dt', \tag{12}$$

where

$$g(r,t) = (4\pi \mathcal{D}t)^{-3/2} \exp\left(-\frac{r^2}{4\mathcal{D}t} - kt\right) \tag{13}$$

is the Green function for Eq. (8).
If w depends on t and/or the conditions are transient, the mRNA and protein concentrations can be easily calculated numerically by using expressions (11) and (12). Under steady-state conditions (with $w = const$), the mRNA and protein concentrations can be calculated analytically also by using expressions (11) and (12) or by direct integration of Eqs. (7) and (8). In particular, Eq. (7) can be rewritten under steady-state conditions as

$$\frac{D}{r^2}\frac{d}{dr} r^2 \frac{dC(r)}{dr} - kC(r) = 0, \tag{14}$$

with

$$-4\pi r^2 D \left.\frac{dC(r)}{dr}\right|_{r\to 0} = w. \tag{15}$$

Employing expression (11) or Eq. (14), we obtain

$$C(r) = \frac{w\alpha^2}{4\pi k r} \exp(-\alpha r), \tag{16}$$

where $\alpha = (k/D)^{1/2}$. Taking (3) into account, we can rewrite expression (16) as

$$C(r) = \frac{N\alpha^2}{4\pi r} \exp(-\alpha r). \tag{17}$$

The latter expression shows as expected that

$$\int_0^\infty C(r) 4\pi r^2 dr = N. \tag{18}$$

The corresponding radial distribution of mRNA (Fig. 1) is given by

$$F(r) \equiv C(r) 4\pi r^2 / N = \alpha^2 r \exp(-\alpha r). \tag{19}$$

Under steady-state conditions, Eq. (8) is reduced to

$$\mathcal{D}\nabla^2 c(\mathbf{r}) - \kappa c(\mathbf{r}) + \upsilon C(\mathbf{r}) = 0. \tag{20}$$

If $C(\mathbf{r})$ is replaced in this equation by the delta function at $\mathbf{r} = \mathbf{r}'$, i.e.,

$$\mathcal{D}\nabla^2 c(\mathbf{r}) - \kappa c(\mathbf{r}) + \upsilon \delta(\mathbf{r}-\mathbf{r}') = 0, \tag{21}$$

we get

$$c(r) = \frac{\upsilon\beta^2}{4\pi\kappa|\mathbf{r}-\mathbf{r}'|}\exp(-\beta|\mathbf{r}-\mathbf{r}'|), \tag{22}$$

where $\beta = (\kappa/\mathcal{D})^{1/2}$. The solution to Eq. (20) is given by convolution of $C(\mathbf{r})$ and expression (22), i.e.

$$c(\mathbf{r}) = \frac{\upsilon\beta^2}{4\pi\kappa}\int C(\mathbf{r}')\frac{\exp(\beta|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}d^3r'. \tag{23}$$

Substituting (16) into (23) [or (12)], we obtain after integration

$$c(r) = \frac{\upsilon w\alpha^2\beta^2[\exp(-\beta r)-\exp(-\alpha r)]}{4\pi\kappa k r(\alpha^2-\beta^2)}. \tag{24}$$

Taking (4) into account, we rewrite the latter expression as

$$c(r) = \frac{n\alpha^2\beta^2[\exp(-\beta r)-\exp(-\alpha r)]}{4\pi r(\alpha^2-\beta^2)}. \tag{25}$$

Integrally, as expected, we have

$$\int_0^\infty c(r)4\pi r^2 dr = n. \tag{26}$$

The radial distribution of protein (Fig. 2) is defined by

$$f(r) \equiv c(r)4\pi r^2/n = \frac{\alpha^2\beta^2 r[\exp(-\beta r)-\exp(-\alpha r)]}{\alpha^2-\beta^2}. \tag{27}$$

If $\alpha = \beta$, expression (25) is reduced to

$$c(r) = \frac{n\alpha^3\exp(-\alpha r)}{8\pi}. \tag{28}$$

If $\alpha \gg \beta$, expression (25) is simplified as

$$c(r) \simeq \frac{n\beta^2\exp(-\beta r)}{4\pi r}. \tag{29}$$

The latter expression is similar to (17). At $r \to 0$, expression (25) yields

$$c(0) = \frac{n\alpha^2\beta^2}{4\pi(\alpha+\beta)} \equiv \frac{nk\kappa}{4\pi D\mathcal{D}[(k/D)^{1/2}+(\kappa/\mathcal{D})^{1/2}]}. \tag{30}$$

3. Conclusion

The model presented predicts that under steady-state conditions the localized distribution of mRNAs and proteins is described by simple exponential functions [(16) and (24); Figs. 1 and 2]. The shape of these functions is instructive from the tutorial point of view. In addition, the functions can be used to fit measured mRNAs and protein distributions. A good example of such measurements is the recent observation that several mRNAs, chromosomally expressed in *Caulobacter crescentus* and *Escherichia coli*, display limited dispersion from their site of transcription during their lifetime [14].

The protein concentration near the gene is shown to depend on the protein and mRNA diffusion coefficients and degradation rate constants. The corresponding expression (30) may be useful to describe the situations when the transcription of a gene into mRNA is regulated by the protein synthesized by translation of this mRNA or other mRNA transcribed in proximity. The former situations representing self-regulation are abundant (this type of regulation is inherent, e.g., to half of the repressors in *Escherichia coli* [8]). The latter situations are also general and accordingly expected to be potentially important. In this context, it is of interest to notice that the localization of RNAs and proteins is now believed to be a rational explanation why the genes encoding interacting proteins frequently cluster [14] (reviewed in Ref. [18]). This way of transcription and translation is energetically efficient, because functionally related proteins are synthesized close in space and time and accordingly may operate even if their populations are relatively small.

References

[1] M. Kaern, T.C. Elston, W.J. Blake, J.J. Collins, Nature Rev. Genet. 6, 451 (2005)
[2] J. Paulsson, Phys. Life Rev. 2, 157 (2005)
[3] J. Rausenberger, C. Fleck, J. Timmer, M. Kollmann, Progr. Biophys. Molec. Biol. 100, 57 (2009)
[4] A.S. Ribeiro, Math. Biosci. 223, 1 (2010)
[5] G. Tiana, S. Krishna, S. Pigolotti, M.H. Jensen, K. Sneppen, Phys. Biol. 4, R1 (2007)
[6] O. Purcell, N.J. Savery, C.S. Grierson, M. di Bernardo, J. Roy. Soc. Interface 7, 1503 (2010)
[7] V.P. Zhdanov, Phys. Rep. 500, 1 (2011)
[8] U. Alon, Nature Rev. Genet. 8, 450 (2007)
[9] G. Karlebach, R. Shamir, Nature Rev. Molec. Cell Biol. 9, 771 (2008)
[10] M.C. Lagomarsino, B. Bassetti, G. Castellani, D. Remondini, Molec. Biosyst. 5, 335 (2009)

[11] D.F.T. Veiga, B. Dutta, G. Balazsi, Molec. Biosyst. 6, 469 (2010)
[12] M. Valencia-Burtona, et al., Proc. Natl. Acad. Sci. USA 106, 16399 (2009)
[13] J.H. Russell, K.C. Keiler, Proc. Natl. Acad. Sci. USA 106, 16405 (2009)
[14] P. Montero Llopis, et al., Nature 466, 77 (2010)
[15] K. Nevo-Dinur, A. Nussbaum-Shochat, S. Ben-Yehuda, O. Amster-Choder, Science 331, 1081 (2011)
[16] D.Z. Rudner, R. Losick, Cold Spring Harb. Perspect. Biol. 2, a000307 (2010)
[17] E. Evguenieva-Hackenberg, V. Roppelt, C. Lassek, G. Klug, RNA Biol. 8, 49 (2011)
[18] N.E. Broude, Molec. Microbiol. 80, 1137 (2011)
[19] K.C. Keiler, Curr. Opin. Microbiol. 14, 155 (2011)
[20] J.T. Mika, B. Poolman, Curr. Opin. Biotechnol. 22, 117 (2011)
[21] S.A. Isaacson, C.S. Peskin, SIAM J. Sci. Comp. 28, 47 (2006)
[22] V.P. Zhdanov, JETP Lett. 85, 302 (2007)
[23] V.P. Zhdanov, Chem. Phys. Lett. 458, 247 (2008)
[24] V.P. Zhdanov, Biophys. Rev. Lett. 4, 267 (2009)
[25] Q. Ma, G. Shi, S. Xu, Y. Zou, Neural Comput. Applic. 20, 507 (2011)