

RESEARCH ARTICLE

Effects of spatial heterogeneity on bacterial genetic circuits

Carlos Barajas , Domitilla Del Vecchio*

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA

* ddv@mit.edu

Abstract

Intracellular spatial heterogeneity is frequently observed in bacteria, where the chromosome occupies part of the cell's volume and a circuit's DNA often localizes within the cell. How this heterogeneity affects core processes and genetic circuits is still poorly understood. In fact, commonly used ordinary differential equation (ODE) models of genetic circuits assume a well-mixed ensemble of molecules and, as such, do not capture spatial aspects. Reaction-diffusion partial differential equation (PDE) models have been only occasionally used since they are difficult to integrate and do not provide mechanistic understanding of the effects of spatial heterogeneity. In this paper, we derive a reduced ODE model that captures spatial effects, yet has the same dimension as commonly used well-mixed models. In particular, the only difference with respect to a well-mixed ODE model is that the association rate constant of binding reactions is multiplied by a coefficient, which we refer to as the binding correction factor (BCF). The BCF depends on the size of interacting molecules and on their location when fixed in space and it is equal to unity in a well-mixed ODE model. The BCF can be used to investigate how spatial heterogeneity affects the behavior of core processes and genetic circuits. Specifically, our reduced model indicates that transcription and its regulation are more effective for genes located at the cell poles than for genes located on the chromosome. The extent of these effects depends on the value of the BCF, which we found to be close to unity. For translation, the value of the BCF is always greater than unity, it increases with mRNA size, and, with biologically relevant parameters, is substantially larger than unity. Our model has broad validity, has the same dimension as a well-mixed model, yet it incorporates spatial heterogeneity. This simple-to-use model can be used to both analyze and design genetic circuits while accounting for spatial intracellular effects.

 OPEN ACCESS

Citation: Barajas C, Del Vecchio D (2020) Effects of spatial heterogeneity on bacterial genetic circuits. *PLoS Comput Biol* 16(9): e1008159. <https://doi.org/10.1371/journal.pcbi.1008159>

Editor: Alexandre V. Morozov, Rutgers University, UNITED STATES

Received: February 5, 2020

Accepted: July 17, 2020

Published: September 14, 2020

Copyright: © 2020 Barajas, Del Vecchio. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported in part by NSF Expeditions, Grant Number 1521925, AFOSR grant FA9550-14-1-0060, the NSF Graduate Research Fellowships Program, and the Ford Foundation Predoctoral Fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

A general and simple modeling framework to determine how spatial heterogeneity modulates the dynamics of gene networks is currently lacking. To this end, we provide a simple-to-use ordinary differential equation (ODE) model that can be used to both analyze and design genetic circuits while accounting for spatial intracellular effects. We apply our model to several core biological processes and determine that transcription and its

regulation are more effective for genes located at the cell poles than for genes located on the chromosome and this difference increases with regulator size. For translation, we predict the effective binding between ribosomes and mRNA is higher than that predicted by a well-mixed model, and it increases with mRNA size. We provide examples where spatial effects are significant and should be considered but also where a traditional well-mixed model suffices despite severe spatial heterogeneity. Finally, we illustrate how the operation of well-known genetic circuits is impacted by spatial effects.

Introduction

Deterministic models of gene circuits typically assume a well-mixed ensemble of species inside the cell [1, 2]. This assumption allows one to describe genetic circuit dynamics through a set of ODEs, for which a number of established analysis tools are available [1]. However, it is well known that spatial heterogeneity is prevalent inside bacterial cells [3–8]. Depending on the origin of replication, plasmids tend to localize within bacterial cells [9–11]. Furthermore, chromosome genes (endogenous and synthetically integrated ones [12]) are distributed in the cell according to the chromosome complex spatial structure. In bacterial cells, any molecule freely diffusing through the chromosome (e.g., mRNA, ribosome, and protease) experiences what are known as *excluded volume effects*, which capture the tendency of species to be ejected from the nucleoid due to the space occupied by the dense DNA mesh [13]. These excluded volume effects for ribosomes and RNAP in bacteria have been observed experimentally [14].

Despite the strong evidence in support of spatial heterogeneity within bacterial cells, a convenient modeling framework that captures the spatio-temporal organization of molecules inside the cell is largely lacking. As a consequence, how spatial effects modulate genetic circuit dynamics remains also poorly understood. Partial differential equation (PDE) models have been employed on an *ad hoc* basis to numerically capture intracellular spatial dynamics for specific case studies [15–17]. Although a general PDE model of a gene regulatory network (GRN) can be constructed, it is difficult to analyze and impractical for design [18]. Recently, the method of matched asymptotic expansions was used to simplify the PDEs to a set of ODEs to analyze ribosome-mRNA interactions [19]. Similarly, [20] used a compartmentalized model to capture spatial heterogeneity in sRNA-mRNA interactions. However, these results have not been generalized, relied on simulation, and specific parameter values.

In this paper, we provide a general framework to model spatial heterogeneity through an ODE that has the same structure and hence dimensionality as a well-mixed ODE model. To this end, we first introduce a PDE model that captures spatial dynamics. Next, we exploit the time scale separation between molecule diffusion and biochemical reactions to derive a reduced order ODE model of the space averaged dynamics. This model accounts for spatial heterogeneity by multiplying the association rate constant of binding reactions by a factor that depends on the size of freely diffusing species and on the location of spatially fixed species. We call this factor the *binding correction factor* (BCF). Thus, this reduced model has the same dimensionality as traditional well-mixed models, yet it captures spatial effects.

We demonstrate the effects of spatial heterogeneity in genetic circuit behavior by modeling and analyzing several core biological processes. We show that the transcription rate of a gene and the affinity at which transcription factors bind to it, is lower (higher) when the gene is located near mid-cell (cell poles) with respect to the well-mixed model. We show that compared to a well-mixed model, translation rate is always higher and increases with mRNA size. Finally, we consider a genetic clock, a circuit that produces sustained oscillations. We show

that for a parameter range where a well-mixed model predicts sustained oscillations, a model that accounts for spatial heterogeneity of DNA may not show oscillations. All of these phenomena can be recapitulated by our reduced ODE model.

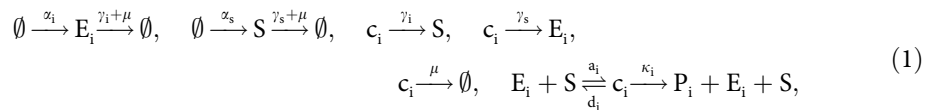
Materials and methods

We use mathematical models to investigate the effects of spatial heterogeneity, specifically DNA localization and excluded volume effects, on genetic circuit behavior. The first part of this section introduces the mathematical model used, a set of nonlinear PDEs. Model reduction is performed on the resulting PDEs to obtain the reduced ODE model that we use to predict how molecule size and location affect genetic circuit's behavior. The numerical method used to simulate the PDEs in this study is discussed in [S1 Text](#): Section 6.

Reaction-diffusion model

A reaction-diffusion model describes the concentration of a species at a given time and location in the cell. We focus on enzymatic-like reactions since they can be used to capture most core processes in the cell. We specialize the model to the cases where the reacting species both freely diffuse or where one freely diffuses while the other one is fixed. For example, mRNA and ribosomes are both freely diffusing, while for RNA polymerase and DNA, one is freely diffusing and the other one is fixed.

Enzymatic-like reactions that model core biological processes. Let S be a substrate being shared by n enzymes E_i , to form product P_i where $i = 1, \dots, n$. The rate at which E_i and S are produced is given by α_i and α_s , respectively. The decay rates of E_i and S are given by γ_i and γ_s , respectively. Here, we assume that E_i and S can be degraded even in complex form, that is, the complex is not protecting them from degradation. Finally, all species are diluted as the cell divides at a rate μ . The biochemical reactions corresponding to this process are given by:



where c_i is the complex formed when E_i binds to S , a_i is the association rate constant, d_i is the dissociation rate constant, and κ_i is the catalytic rate constant of product formation. These enzymatic-like reactions capture many core biological processes such as genes transcribed by RNA polymerase, mRNA translated by ribosomes, or proteins degraded by a common protease [1]. Notice that they differ from the classical enzymatic reactions since the substrate is not converted into product [1].

E. coli actively regulates its geometry to achieve a near-perfect cylindrical shape [21]. Thus, we model the cell as a cylinder of length $2L$ and radius R_c . This geometry is shown in [Fig 1A](#). We assume angular and radial homogeneity ($(R_c/L)^2 \ll 1$) such that the concentration of a species varies only axially (the spatial x direction). Symmetry relative to mid-cell is assumed and hence only half of the cell is considered, that is, $x \in [0, L]$, where $x = 0$ is at mid-cell and $x = L$ is at the cell poles. Furthermore, we assume a constant cross-sectional area along the axial direction.

In [14] it was shown that polysomes were excluded from the dense chromosomal DNA mesh onto the cell poles. These phenomena is generalized for any species that freely diffuses within the DNA mesh and is referred to as “excluded volume effects”. Leveraging the diffusion modeling framework from [13], we now specify the model to capture excluded volume effects. Let $v(x) \in (0, 1]$ be the volume fraction (dimensionless) available to a species to diffuse within the chromosome ([Fig 1B](#)). As derived in [13] and discussed in [S1 Text](#): Section 2, the available

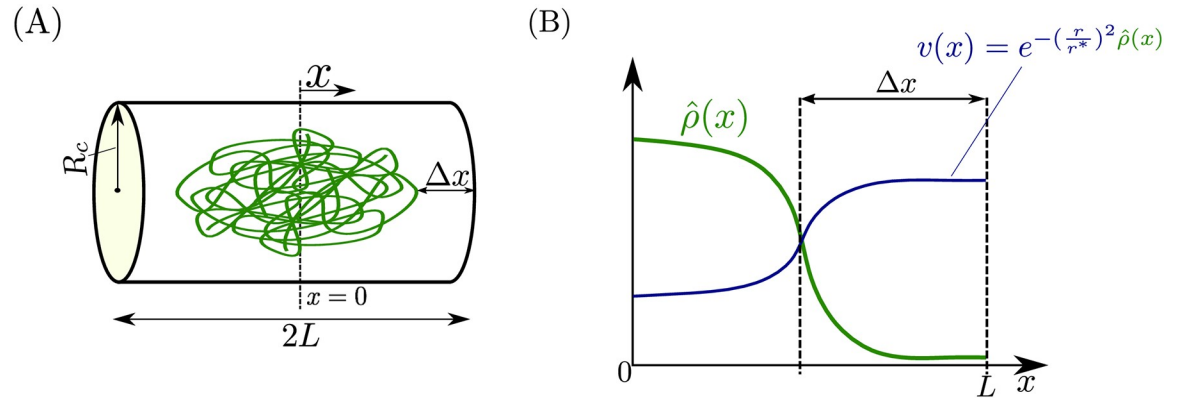


Fig 1. Intracellular spatial geometry. (A) We model the cell as a cylinder of radius R_c and length $2L$. The distance between the end of the chromosome and the cell poles is Δx . (B) The spatial profiles for the normalized local density of DNA length $\hat{\rho}(x)$ and the fraction of available volume $v(x)$ of a freely diffusing species with radius of gyration r within the chromosomal mesh. These two quantities are related by $v(x) = e^{-(r/r^*)^2 \hat{\rho}(x)}$, where r^* is a length scale dependent on the averaged chromosome density in the cell given by (2). The chromosome density is assumed to be monotonically decreasing from mid-cell to the cell poles (as in [13]), thus the available volume profile are monotonically increasing.

<https://doi.org/10.1371/journal.pcbi.1008159.g001>

volume profile $v(x)$ of a species with a radius of gyration r , takes the form

$$v(x) = e^{-(r/r^*)^2 \hat{\rho}(x)}, \quad (r^*)^2 = \frac{V_p}{2\kappa\pi L_p}, \quad (2)$$

where $\hat{\rho}(x)$ is the normalized local density of chromosome DNA length such that $\frac{2}{L} \int_0^L \hat{\rho}(x) dx = 1$, L_p is the total length of chromosome DNA, V_p the volume where the DNA polymer is confined, such that L_p/V_p is the total DNA length per volume, and κ is an empirically determined correction factor (see [13] and S1 Text: Section 2). The quantity $(r^*)^2$ is inversely proportional to the total DNA length per volume. For all simulations in the main text of this study, we model the normalized chromosome density as

$$\hat{\rho}(x) = \frac{1}{1 + e^{20(x/L-1/2)}},$$

as experimentally determined in [13]. This model for $\hat{\rho}(x)$ is monotonically decreasing (i.e., the chromosome is more dense near mid-cell than at the cell poles as shown in Fig 1B). Therefore by (2), the available volume profile is higher near mid-cell than at the cell poles (i.e., $v(0) < v(L)$) as shown in Fig 1B and furthermore, the discrepancy between $v(0)$ and $v(L)$ increases with r/r^* . We note that the specific expressions of $\hat{\rho}(x)$ and r/r^* do not affect the model reduction result of this paper. The main results in this paper are presented for a constant cell length L and chromosome DNA density $\hat{\rho}(x)$, however in S1 Text: Section 9 we relax these assumptions and allow these quantities to vary in time as the cell divides.

For any given species with concentration per unit length given by $y(t, x)$, free to diffuse, with available volume $v(x)$, an expression for the flux term, derived in [13] is given by:

$$J(x, y) = D \left(\underbrace{-\frac{\partial y(t, x)}{\partial x} v(x)}_{\text{towards low concentration}} + \underbrace{y(t, x) \frac{\partial v(x)}{\partial x}}_{\text{towards high available volume}} \right) = -v^2(x) \frac{d}{dx} \left[\frac{y(t, x)}{v(x)} \right], \quad (3)$$

where D is the diffusion coefficient. The flux is driven by two mechanisms: the first is concentration gradient, which pushes molecules from high to low concentrations and the second

drives molecules to regions with a higher volume fraction. This second term is referred to as the excluded volume effect [13]. From (3), if $\left| \frac{\partial y(t,x)}{\partial x} v(x) \right| < \left| y(t,x) \frac{\partial v(x)}{\partial x} \right|$ and $\frac{\partial y(t,x)}{\partial x} \frac{\partial v(x)}{\partial x} > 0$, then the net flux is from low to high concentration, which is the case when species are repelled from the chromosome to high concentration areas in the cell poles. As we will show, this mechanism dictates intracellular heterogeneity in the limit of fast diffusion.

For species S , we denote by $S(t, x)$ its concentration per unit length at time t at location x (similarly for E_i and c_i). Assuming sufficiently high molecular counts, the reaction-diffusion dynamics corresponding to (1) describing the rate of change of the species concentrations at position x , are given by [22]:

$$\begin{aligned} \frac{\partial E_i(t, x)}{\partial t} &= -\frac{d}{dx} [J(x, E_i)] - a_i E_i(t, x) S(t, x) + (\gamma_s + d_i + \kappa_i) c_i(t, x) \\ &\quad + \alpha_i(t, x) - (\gamma_i + \mu) E_i(t, x), \\ \frac{\partial c_i(t, x)}{\partial t} &= -\frac{d}{dx} [J(x, c_i)] + a_i E_i(t, x) S(t, x) - (\gamma_i + \gamma_s + d_i + \kappa_i + \mu) c_i(t, x), \\ \frac{\partial S(t, x)}{\partial t} &= -\frac{d}{dx} [J(x, S)] + \sum_{j=1}^n \left[-a_j E_j(t, x) S(t, x) + (\gamma_j + d_j + \kappa_j) c_j(t, x) \right] \\ &\quad + \alpha_s(t, x) - (\gamma_s + \mu) S(t, x), \end{aligned} \tag{4}$$

where $J(x, \cdot)$ is the flux per unit area per unit time, within the cell. If the species is freely diffusing $J(x, \cdot)$ is given by (3), otherwise if the species is spatially fixed, then $J(x, \cdot) = 0$ for all $x \in [0, L]$. The boundary conditions associated with freely diffusing species of (4) are zero flux at the cell poles and cell center due to the assumed left-right symmetry, which corresponds to:

$$J(0, \cdot) = J(L, \cdot) = 0. \tag{5}$$

Notice that none of the parameters in (4) appearing in (1) depend explicitly on time and space except for the production terms $\alpha_i(t, x)$ and $\alpha_s(t, x)$. The explicit time dependence of the production terms allows us to capture how genes can be activated or repressed externally with a time varying signal [23]. The explicit dependence of the production terms on x allows us to capture where the species is produced within the cell (e.g., DNA in the chromosome or DNA in pole localized plasmid genes).

Dimensionless model: Depending on the parameter regimes, the dynamics of (4) can display time scale separation. For example, diffusion occurs in the order of mili-seconds compared to minutes for dilution due to cell-growth and mRNA degradation [2]. Therefore, we are interested in determining the behavior of (4) in the limit of fast diffusion. We thus rewrite (4) in dimensionless form to make time scale separation explicit. We nondimensionalize the system variables using dilution ($1/\mu$) as the characteristic time scale, the length of the cell (L) as the characteristic length, and μ/a_1 as the characteristic concentration per length scale: $t^* = t\mu$, $y^* = y \frac{a_1}{\mu}$, $x^* = \frac{x}{L}$, where y denotes concentration per unit length and the superscript “*” is used on the dimensionless variable. Concentrations are nondimensionalized through a_1 because this parameter contains a concentration scale, it is fixed in time, and it is assumed to

be nonzero. The dimensionless form of (4) is given by

$$\begin{aligned} \frac{\partial E_i^*(t^*, x^*)}{\partial t^*} &= -\frac{d}{dx^*} [J^*(x^*, E_i^*)] + \frac{1}{\eta_i} \left[-E_i^*(t^*, x^*) S^*(t^*, x^*) \frac{a_i^*}{\tilde{d}_i} + c_i^*(t^*, x^*) \right] \\ &\quad + \alpha_i^*(t^*, x^*) - (\gamma_i^* + 1)(E_i^*(t^*, x^*) + c_i^*(t^*, x^*)), \\ \frac{\partial c_i^*(t^*, x^*)}{\partial t^*} &= -\frac{d}{dx^*} [J^*(x^*, c_i^*)] + \frac{1}{\eta_i} \left[E_i^*(t^*, x^*) S^*(t^*, x^*) \frac{a_i^*}{\tilde{d}_i} - c_i^*(t^*, x^*) \right], \\ \frac{\partial S^*(t^*, x^*)}{\partial t^*} &= -\frac{d}{dx^*} [J^*(x^*, S^*)] + \sum_{j=1}^n \frac{1}{\eta_j} \left[-E_j^*(t^*, x^*) S^*(t^*, x^*) \frac{a_j^*}{\tilde{d}_j} + c_j^*(t^*, x^*) \right] \\ &\quad + \alpha_s^*(t^*, x^*) - (\gamma_s^* + 1)(S^*(t^*, x^*) + \sum_{j=1}^n c_j^*(t^*, x^*)), \end{aligned} \tag{6}$$

where $a_i^* = a_i/a_1$, $\gamma_s^* = \gamma_s/\mu$, $\gamma_i^* = \gamma_i/\mu$, $\tilde{d}_i^* = d_i/\mu$, $\kappa_i^* = \kappa_i/\mu$, $\alpha_i^* = \alpha_i a_1/\mu^2$, $\alpha_s^* = \alpha_s a_1/\mu^2$, $\tilde{d}_i = \gamma_i^* + \gamma_s^* + \tilde{d}_i^* + \kappa_i^* + 1$, $\eta_i = 1/\tilde{d}_i$, and $J^* = J a_1/(\mu^2 L)$. For a freely diffusing species with diffusion coefficient D , the dimensionless parameter that determines the relative speed of diffusion is denoted by $\epsilon = \mu L^2/D$ and fast diffusion corresponds to $\epsilon \ll 1$. Likewise, η_i in (4) determines the relative speed of the binding dynamics, where $\eta_i \ll 1$ implies these reactions are fast. From hereon, unless otherwise specified, we work with variables in their dimensionless form and drop the star superscript for simplifying notation.

Space averaged concentrations: Concentrations per cell are usually the quantities measured experimentally [24] and are the primary quantities of interest. We now derive the space averaged dynamics corresponding to (6), which describe the dynamics of concentrations per half of the cell. We define $\bar{E}_i(t)$, $\bar{S}(t)$, and $\bar{c}_i(t)$ to be the *space averaged enzyme, substrate, and complex concentrations*, respectively, and are given by

$$\bar{E}_i(t) = \int_0^1 E_i(t, x) dx, \quad \bar{c}_i(t) = \int_0^1 c_i(t, x) dx, \quad \bar{S}(t) = \int_0^1 S(t, x) dx,$$

also giving the concentrations per half of the cell. The dynamics governing these space averaged variables are derived by integrating (6) in space and applying the boundary conditions (5) and are given by:

$$\begin{aligned} \frac{d\bar{E}_i(t)}{dt} &= \bar{\alpha}_i(t) - \frac{1}{\eta_i} \left[\bar{E}_i(t) \bar{S}(t) \frac{a_i \theta_i(t)}{\tilde{d}_i} - \bar{c}_i(t) \right] - (\gamma_i + 1)(\bar{E}_i(t) + \bar{c}_i(t)), \\ \frac{d\bar{c}_i(t)}{dt} &= \frac{1}{\eta_i} \left[\bar{E}_i(t) \bar{S}(t) \frac{a_i \theta_i(t)}{\tilde{d}_i} - \bar{c}_i(t) \right], \\ \frac{d\bar{S}(t)}{dt} &= \bar{\alpha}_s(t) - \sum_{j=1}^n \frac{1}{\eta_j} \left[\bar{E}_j(t) \bar{S}(t) \frac{a_j \theta_j(t)}{\tilde{d}_j} - \bar{c}_j(t) \right] - (\gamma_s + 1)(\bar{S}(t) + \sum_{j=1}^n \bar{c}_j(t)), \end{aligned} \tag{7}$$

where overbars denote spatially averaged variables and

$$\theta_i(t) = \frac{\int_0^1 E_i(t, x) S(t, x) dx}{[\int_0^1 E_i(t, x) dx][\int_0^1 S(t, x) dx]}. \tag{8}$$

Therefore, to calculate the space averaged concentrations, one could integrate the outputs of the full PDE (6) directly or use (7) along with (8), as illustrated in Fig 2. Notice that calculating

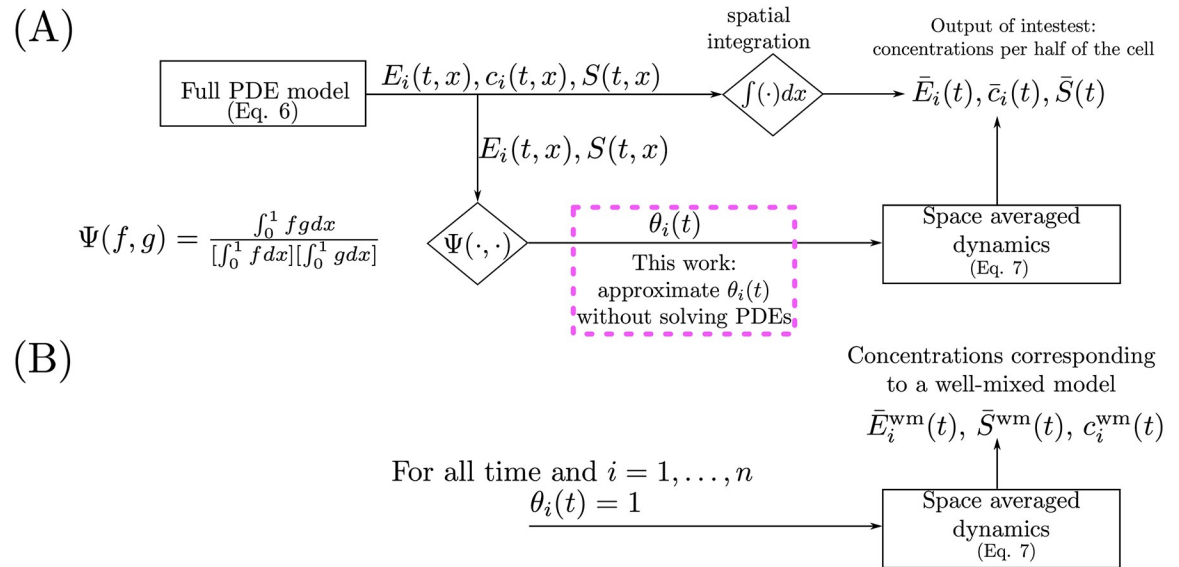


Fig 2. Methods to calculate space averaged concentrations. (A) The time and space dependent solutions of the full PDE (6) are integrated spatially to yield concentrations per half of the cell. Alternatively, space averaged concentrations can be calculated using the space average dynamics (7) with the BCF $\theta_i(t)$ (8) as a time varying parameter. (B) The dynamics of the space averaged concentration is given by the well mixed model (9) when $\theta_i(t) = 1$ for all time and for all $i = 1, \dots, n$.

<https://doi.org/10.1371/journal.pcbi.1008159.g002>

θ_i in (8) requires solving the full PDE system (6) because of its dependence on the product $E_i(t, x)S(t, x)$. Therefore, in general, there is no obvious benefit in working with (7). In this paper, we provide a method to compute a guaranteed approximation of θ_i without solving the PDEs (6).

Well-mixed model: Next, we define what we have been informally referring to as the “well-mixed” model [25]. A standard well-mixed model is derived starting from (1) assuming mass action kinetics, that molecular counts are sufficiently large, and that the intracellular environment is spatially homogeneous (well-mixed) [1]. We let $\bar{E}_i^{wm}(t), \bar{S}^{wm}(t)$ and $\bar{c}_i^{wm}(t)$, denote the well-mixed concentrations of E_i, S , and c_i , respectively, and their dynamics are given by

$$\begin{aligned} \frac{d\bar{E}_i^{wm}(t)}{dt} &= \bar{\alpha}_i(t) - \frac{1}{\eta_i} \left[\bar{E}_i^{wm}(t)\bar{S}^{wm}(t)\frac{a_i}{d_i} - \bar{c}_i^{wm}(t) \right] - (\gamma_i + 1)(\bar{E}_i^{wm}(t) + \bar{c}_i^{wm}(t)), \\ \frac{d\bar{c}_i^{wm}(t)}{dt} &= \frac{1}{\eta_i} \left[\bar{E}_i^{wm}(t)\bar{S}^{wm}(t)\frac{a_i}{d_i} - \bar{c}_i^{wm}(t) \right], \\ \frac{d\bar{S}^{wm}(t)}{dt} &= \bar{\alpha}_s(t) - \sum_{j=1}^n \frac{1}{\eta_j} \left[\bar{E}_j^{wm}(t)\bar{S}^{wm}(t)\frac{a_j}{d_j} - \bar{c}_j^{wm}(t) \right] - (\gamma_s + 1)(\bar{S}^{wm}(t) + \sum_{j=1}^n \bar{c}_j^{wm}(t)). \end{aligned} \tag{9}$$

Comparing (7) and (9), motivates us to define $a'_i(t) = a_i\theta_i(t)$, which can be regarded as the effective association rate constant between E_i and S in (7). We refer to $\theta_i(t)$ as the *binding correction factor* (BCF). The dynamics of the space averaged concentrations (7) coincide with those of the well-mixed model (9) when $\theta_i(t) = 1$ (thus $a'_i(t) = a_i$) for all time and for all $i = 1, \dots, n$. From (8) notice that $E_i(t, x)$ and $S(t, x)$ being spatially constant for all time is not necessary for $\theta_i(t) = 1$ for all time. For example, if $S(t, x)$ is spatially constant while $E_i(t, x)$ has an arbitrary spatial profile (or *vice-versa*), then $\theta_i(t) = 1$. Thus, the space averaged concentrations can coincide with those of a well-mixed model despite severe spatial heterogeneity. In this

Table 1. The flux dynamics and the boundary conditions corresponding to (6) for each case of interest along with a core process example. Here $v_{E_i}(x)$, $v_S(x)$, and $v_{c_i}(x)$, are the available volume profiles of E_i , S, and c_i , respectively. The parameters D_{E_i} , D_{c_i} , and D_S , are the enzyme, complex, and substrate diffusion coefficients, respectively, ϵ is a dimensionless parameter that captures the speed of diffusion (with respect to dilution). A species being spatially fixed translates to the flux being zero throughout the whole spatial domain. In Case II, for $i = 1, \dots, n$, $x_i^* \in (0, 1)$ denotes the location of the fixed species E_i . In Case III, $x_s^* \in (0, 1)$ denotes the location of the fixed species S.

	Case I	Case II	Case III
	All species diffuse	Substrate diffuse and enzymes fixed	Enzymes diffuse and substrate fixed
Dimensionless Flux	$J(x, E_i) = -\frac{1}{\epsilon} \chi_{E_i} v_{E_i}^2 \frac{d}{dx} \left[\frac{E_i}{v_{E_i}} \right]$ $J(x, S) = -\frac{1}{\epsilon} v_S^2 \frac{d}{dx} \left[\frac{S}{v_S} \right]$ $J(x, c_i) = -\frac{1}{\epsilon} \chi_{c_i} v_{c_i}^2 \frac{d}{dx} \left[\frac{c_i}{v_{c_i}} \right]$	$J(x, E_i) = 0$ $J(x, S) = -\frac{1}{\epsilon} v_S^2 \frac{d}{dx} \left[\frac{S}{v_S} \right]$ $J(x, c_i) = 0$	$J(x, E_i) = -\frac{1}{\epsilon} v_{E_i}^2 \frac{d}{dx} \left[\frac{E_i}{v_{E_i}} \right]$ $J(x, S) = 0$ $J(x, c_i) = 0$
Boundary conditions	$J(0, E_i) = J(1, E_i) = 0$ $J(0, c_i) = J(1, c_i) = 0$ $J(0, S) = J(1, S) = 0$	$J(0, S) = J(1, S) = 0$	$J(0, E_i) = J(1, E_i) = 0$
ϵ	$(\mu L^2)/D_S$	$(\mu L^2)/D_S$	$(\mu L^2)/D_{E_i}$
Dimensionless diffusion	$\chi_{E_i} = D_{E_i}/D_S, \chi_{c_i} = D_{c_i}/D_S$	N/A	N/A
Core process	mRNAs binding ribosomes	RNAP binding several genes	Transcription factors binding promoter
Location of fixed species	N/A	x_i^*	x_s^*

<https://doi.org/10.1371/journal.pcbi.1008159.t001>

work, we provide a constant approximation of $\theta_i(t)$ denoted by θ_i^* , which depends on spatial variables such as molecule size and gene location. Under the fast diffusion approximation, we show that θ_i^* is close to $\theta_i(t)$. The space averaged dynamics (7) with $\theta_i(t)$ replaced by θ_i^* , thus provides a reduced ODE model that captures spatial information without having to solve (6). We will compare how solutions to (7) with (8) calculated from the full PDE (6) or with θ_i^* compare to each other and to the solutions of the well-mixed model (9).

Three diffusion cases to capture core biological processes. To use model (6) to describe key biological processes, we consider three cases. In Case I, E_i for all $i = 1, \dots, n$ and S are all freely diffusing within the cell. In Case II, E_i is spatially fixed for all $i = 1, \dots, n$ ($J(x, E_i) = 0$ for all $x \in [0, 1]$ and for all $i = 1, \dots, n$) and S is freely diffusing. In Case III, E_i is freely diffusing for all $i = 1, \dots, n$ and S is spatially fixed ($J(x, S) = 0$ for all $x \in [0, 1]$). Case I may represent mRNA molecules (E_i) competing for ribosomes (S), all freely diffusing in the cell. Case II captures genes (E_i), which are spatially fixed and are transcribed by RNA polymerase (S), which freely diffuses. Case III models transcription factors (E_i), which freely diffuse regulating the same spatially fixed gene (S).

The flux dynamics, the boundary conditions, and a core biological process example for each case are summarized in Table 1. When a species is spatially fixed, the flux is zero through the whole domain, that is, $J(x, \cdot) = 0$ for all $x \in [0, 1]$. The available volume profiles for the enzyme, complex, and substrate are denoted by $v_{E_i}(x)$, $v_{c_i}(x)$, and $v_S(x)$, respectively. The available volume profile for the complex $v_{c_i}(x)$, represents the probability that the complex has enough free volume to hop into the DNA mesh at position x and it equals the product of the probability of the two independent events of the enzyme and the substrate hopping into the DNA mesh [13], thus

$$v_{c_i}(x) = v_{E_i}(x)v_S(x). \tag{10}$$

Furthermore, we define the normalized available volume profiles as

$$\hat{v}_{E_i}(x) = \frac{v_{E_i}(x)}{\int_0^1 v_{E_i}(x) dx}, \quad \hat{v}_{c_i}(x) = \frac{v_{c_i}(x)}{\int_0^1 v_{c_i}(x) dx}, \quad \hat{v}_s(x) = \frac{v_s(x)}{\int_0^1 v_s(x) dx}. \quad (11)$$

Results

Time scale separation

In this section, we provide a time independent approximation of the BCF (8) in the limit of fast diffusion, which depends solely on the size of diffusing species, chromosome density profile ($\hat{\rho}(x)$ and r^*), and the spatial localization of non-diffusing species. With this approximation, we can compute space averaged solutions in (7) without solving the PDEs in (6).

Reduced space averaged dynamics when diffusion is fast and fixed species are localized. For Case II and Case III of Table 1, in which one of the reacting species is fixed, we assume that the *fixed species is spatially localized* to a small space, that is, we have the situation depicted in Fig 3 (see S1 Text: Assumption 3 for the mathematical definition). Practically, for Case II, spatial localization at x_i^* requires that the production rate $\alpha_i(t, x)$ of the fixed species is smaller than some small threshold δ when x is outside the interval $[x_i^* - \delta, x_i^* + \delta]$ for all time and that the space averaged production rate is $\bar{\alpha}_i(t)$ independent of δ (similarly for Case III, x_s^* , and $\alpha_s(t, x)$). From a biological perspective, having the space averaged production rate independent of δ is consistent with the fact that the total amount of DNA in the cell is independent of where the DNA is concentrated. Note that δ is a parameter that controls the amount of

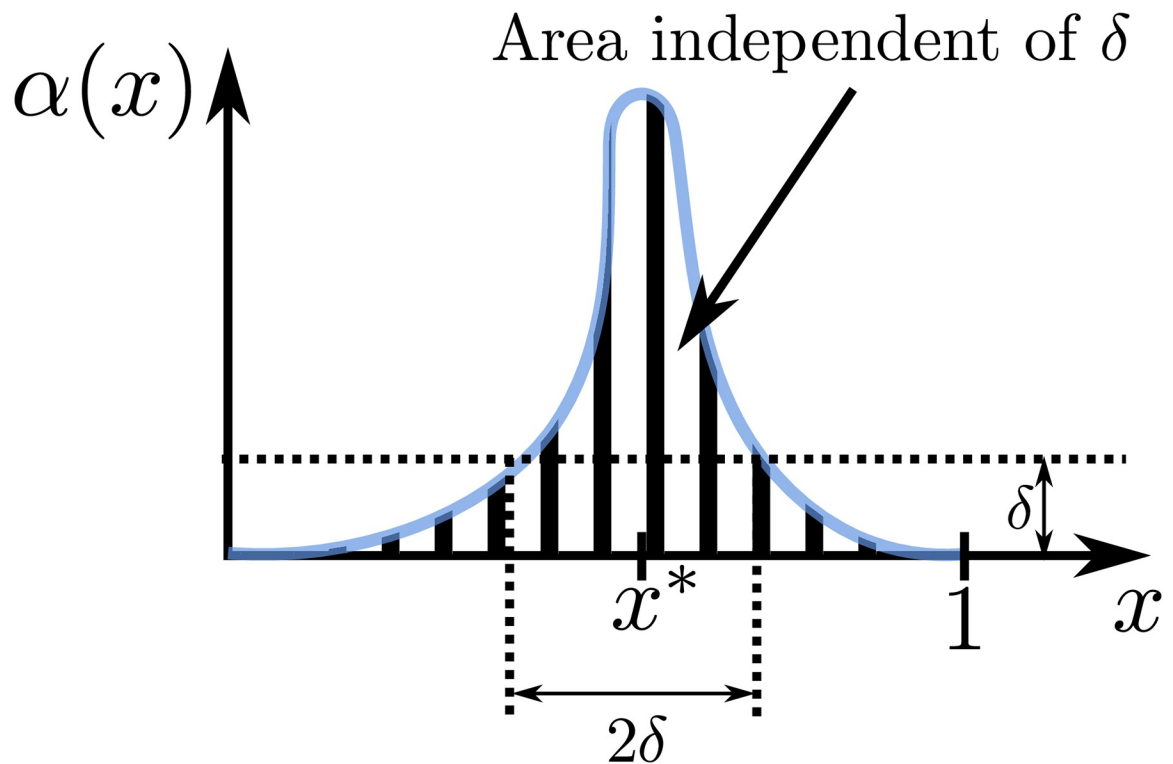


Fig 3. Graphical representation of localization of fixed species. The production rate $\alpha_i(t, x)$ is assumed to be localized at x_i^* if $\alpha_i(t, x) \leq \delta, \forall x \notin [x_i^* - \delta, x_i^* + \delta]$. We assume that the space averaged production $\bar{\alpha}_i(t) = \int_0^1 \alpha_i(t, x) dx$ is independent of δ .

<https://doi.org/10.1371/journal.pcbi.1008159.g003>

localization, such that $\delta \ll 1$ implies the production of spatially fixed species being localized to a small region. Let ϵ be as in Table 1 that appears in (6), the following definition will provide the candidate reduced model that approximates (7) well when $\epsilon \ll 1$ and $\delta \ll 1$. Recall that ϵ is a dimensionless parameter that captures the speed of diffusion (with respect to dilution).

Let x_i^* and x_s^* be the location of the fixed species for Case II and Case III, respectively (see Table 1). For $i = 1, \dots, n$, we define the *reduced space-averaged dynamics* as

$$\begin{aligned} \frac{d\hat{E}_i(t)}{dt} &= \bar{\alpha}_i(t) - \frac{1}{\eta_i} \left[\hat{E}_i(t)\hat{S}(t) \frac{\tilde{a}_i\theta_i^*}{\tilde{d}_i} - \hat{c}_i(t) \right] - (\gamma_i + 1)(\hat{E}_i(t) + \hat{c}_i(t)), \\ \frac{d\hat{c}_i(t)}{dt} &= \frac{1}{\eta_i} \left[\hat{E}_i(t)\hat{S}(t) \frac{\tilde{a}_i\theta_i^*}{\tilde{d}_i} - \hat{c}_i(t) \right], \end{aligned} \tag{12}$$

$$\frac{d\hat{S}(t)}{dt} = \bar{\alpha}_s(t) - \sum_{j=1}^n \frac{1}{\eta_j} \left[\hat{E}_j(t)\hat{S}(t) \frac{\tilde{a}_j\theta_j^*}{\tilde{d}_j} - \hat{c}_j(t) \right] - (\gamma_s + 1)(\hat{S}(t) + \sum_{j=1}^n \hat{c}_j(t)),$$

$$\theta_i^* = \begin{cases} \int_0^1 \hat{v}_{E_i}(x)\hat{v}_s(x)dx & \text{for Case I} \\ \hat{v}_s(x_i^*) & \text{for Case II} \\ \hat{v}_{E_i}(x_s^*) & \text{for Case III} \end{cases}, \tag{13}$$

where $\hat{E}_i(0) = \bar{E}_i(0)$, $\hat{S}(0) = \bar{S}(0)$, $\hat{c}_i(0) = \bar{c}_i(0)$, as given by (7), and $\hat{v}_{E_i}(x)$, $\hat{v}_{c_i}(x)$, and $\hat{v}_s(x)$ are given by (11). Then, we have the following main result of this paper (see S1 Text: Theorem 3 for a formal statement with the proof).

Result 1. Consider system (6) and let $\mathbf{z}(t, x) = [E_1(t, x), \dots, E_n(t, x), c_1(t, x), \dots, c_n(t, x), S(t, x)]^T$ with $\bar{\mathbf{z}}(t) = \int_0^1 \mathbf{z}(t, x)dx$. Consider system (12) and let $\hat{\mathbf{z}}(t) = [\hat{E}_1(t), \dots, \hat{E}_n(t), \hat{c}_1(t), \dots, \hat{c}_n(t), \hat{S}(t)]^T$. Then, for all $t \geq 0$ and ϵ, δ sufficiently small

$$\|\bar{\mathbf{z}}(t) - \hat{\mathbf{z}}(t)\| = \begin{cases} \mathcal{O}(\epsilon) & \text{for Case I} \\ \mathcal{O}(\epsilon) + \mathcal{O}(\delta) & \text{for Case II, III} \end{cases}. \tag{14}$$

By virtue of this result, we can use the simple and convenient ODE model in Eq (12) to describe the space-averaged dynamics of the PDE system (6). In particular, from (12) it appears that spatial effects are lumped into the BCF approximation θ_i^* . Therefore, in order to determine how spatial heterogeneity affects system dynamics, it is sufficient to analyze how dynamics is affected by parameter θ_i^* and how the expression of θ_i^* is, in turn, affected by spatial localization and molecule size (see (13) and (2)).

Remark 1. As discussed in S1 Text: Section 1, as $\epsilon \rightarrow 0^+$, the spatial profile of diffusing molecules approaches that of their available volume profile after a fast transient, that is,

$$\underbrace{E_i(t, x) \approx \bar{E}_i(t)\hat{v}_{E_i}(x)}_{\text{Case I and Case III}}, \quad \underbrace{c_i(t, x) \approx \bar{c}_i(t)\hat{v}_{c_i}(x)}_{\text{Case I}}, \quad \underbrace{S(t, x) \approx \bar{S}(t)\hat{v}_s(x)}_{\text{Case I and Case II}},$$

for the other spatially fixed species we have that their concentrations are localized in a manner as their production terms.

The consequence of Remark 1 is that knowledge of the space averaged dynamics from system (12) also leads to knowledge of the spacial profiles of the species within the cell. This

information is used to propose a method to estimate the BCF from experimental data (See [S1 Text](#): Section 7).

Remark 2. The approximation result holds for $\epsilon \ll 1$, that is, diffusion is much faster than any other time scales in (6). However, in [S1 Text](#): Section 1.4, we motivate why the approximation should still hold (for which the relationship (10) is key) if $\eta_i/\epsilon = \mathcal{O}(1)$ (binding and unbinding between E_i and S occurs at a similar timescale as diffusion), and confirmed via numerical simulations for the upcoming biological examples.

The BCF θ_i^* in (13) is temporally constant and thus the reduced model has the same dimensionality as the well-mixed model (9), yet captures the role of spatial heterogeneity in the interactions between cellular species. Therefore, θ_i^* is a practical and accurate approximation of the BCF when $\epsilon \ll 1$ (sufficient for Case I) and $\delta \ll 1$ (needed for Cases II-III).

Dependence of the BCF on species size and localization. When diffusion is fast and the expression of spatially fixed species is localized, the BCF is well approximated by θ_i^* given in (13). Substituting (2) into (13) and denoting the radius of gyration of E_i and S by $r_{e,i}$ and r_s , respectively, we can rewrite θ_i^* as

$$\theta_i^* = \begin{cases} \frac{\int_0^1 e^{-\frac{r_{e,i}^2+r_s^2}{(r^*)^2}\hat{\rho}(x)} dx}{[\int_0^1 e^{-(r_{e,i}/r^*)^2\hat{\rho}(x)} dx][\int_0^1 e^{-(r_s/r^*)^2\hat{\rho}(x)} dx]} & \text{for Case I} \\ \frac{e^{-(r_s/r^*)^2\hat{\rho}(x_s^*)}}{\int_0^1 e^{-(r_s/r^*)^2\hat{\rho}(x)} dx} & \text{for Case II} \\ \frac{e^{-(r_{e,i}/r^*)^2\hat{\rho}(x_s^*)}}{\int_0^1 e^{-(r_{e,i}/r^*)^2\hat{\rho}(x)} dx} & \text{for Case III} \end{cases} \quad (15)$$

From (15), we observe that θ_i^* depends on the spatial localization of spatially fixed species (i.e., x_i^* and x_s^*), the radius of gyration of diffusing species, r^* (2), and the nominalized local density of DNA length $\hat{\rho}(x)$.

Using (15), we graphically illustrate the dependence of θ_i^* on $r_{E,i}$, r_s , r^* , x_i^* and x_s^* in [Fig 4](#). By analyzing [Fig 4](#), we observe the following:

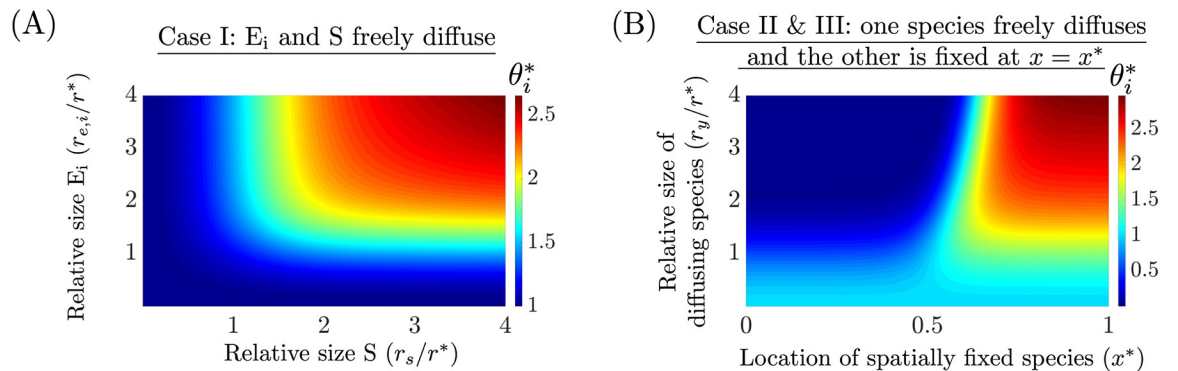


Fig 4. The BCF in the limit of fast diffusion and localization of spatially fixed species. Approximation of the BCF denoted by θ_i^* (15) is provided for Case I and for Case II/Case III. (A) For Case I, where E_i and S both freely diffuse, $\theta_i \geq 1$ and increases when both the size of E_i ($r_{E,i}$) and S (r_s) are sufficiently large (with respect to r^*). (B) For Case II and III, where one of the species diffuses (size r_y) and the other is fixed at $x = x^*$, θ_i^* is different from unity when the radii of the diffusion species is sufficiently large. We observe that $\theta_i^* < 1$ for $x^* \leq 0.4$ and appears to approach zero near $x^* = 0$ for large r_y/r^* . Similarly, for $x^* \geq 0.65$, $\theta_i^* > 1$. Between $0.4 \leq x^* \leq 0.65$ there exists a region that $\theta_i^* = 1$ for all r_y/r^* .

<https://doi.org/10.1371/journal.pcbi.1008159.g004>

Case I: the BCF is always greater than or equal to that of the well-mixed model (9) (where $\theta_i^* = 1$ for all i) and this discrepancy increases with the size of E_i and S . Intuitively, as the size of E_i and S increases, they are pushed out of the chromosome and co-localize near the cell poles, thus they are confined to a smaller volume to interact and hence their effective binding strength increases. If only one of the species is large (with respect to r^*), while the other one is small, then the large species will be ejected from the chromosome and thus will not be homogeneously distributed throughout the cell, however $\theta_i^* \approx 1$, and thus a well-mixed model is valid despite this spatial heterogeneity.

Case II and III: where one of the species diffuses (size r_y) and the other is fixed at $x = x^*$, the BCF is different from unity when r_y is sufficiently large. We observe that $\theta_i^* < 1$ for $x^* \leq 0.4$ and appears to approach zero near $x^* = 0$ for large r_y/r^* . Similarly, for $x^* \geq 0.65$, $\theta_i^* > 1$. This occurs because as the size of the diffusing species increases, the species is ejected from the chromosome onto the cell poles and therefore it is more likely to interact with species fixed at the cell-poles than those near mid-cell. Between $0.4 \leq x^* \leq 0.65$ there exists a region where $\theta_i^* = 1$ for all r_y/r^* . This provides additional evidence that a well-mixed model may be appropriate despite severe intracellular heterogeneity.

When $\hat{\rho}(x)$ is assumed to be a step function, the upper bound for θ_i^* is $1/\Delta x$ for Case I-III, as derived in [S1 Text](#): Section 2, where Δx is the distance between the end of the chromosome and the cell poles as shown in [Fig 1](#). Furthermore, the lower bound for Case I was unity and for Case II-III it was zero.

The value of the BCF provides a measure to determine the extent to which spatial effects modulate the biomolecular dynamics. Therefore, an experimental method to estimate the BCF is desirable. In [S1 Text](#): Section 7, we propose such a method that only requires knowledge of Δx and of the value of concentration of freely diffusing species inside and outside the nucleoid.

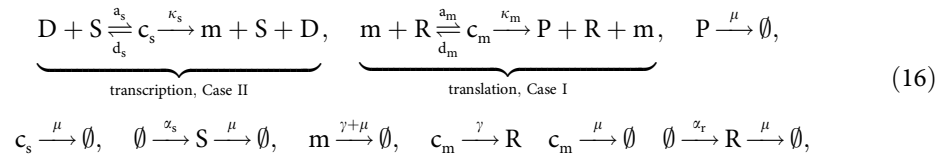
In [S1 Text](#): Section 9, we consider how the BCF can vary temporally as the cell divides and the chromosome density shifts from being concentrated primarily near mid-cell to quarter-cell. We demonstrate that the BCF can vary by over 50% in time for the case where one species is stationary and localized near mid-cell. Furthermore, in [S1 Text](#): Section 10, we show how the BCF is affected when we consider exclusion effects from the DNA of a pole localized high copy plasmid. We show that for the case where both reactant freely diffuse, the BCF decreases as the amount of plasmid DNA increases. For the case where one reactant is spatially fixed and the other freely diffuses, we show that the BCF decreases for a species localized at the cell poles and increases for a species localized near quarter-cell, as the amount of plasmid DNA increases.

Application to core processes and genetic circuits

In this section we apply the results of the time scale separation analysis to determine the effects of intracellular heterogeneity on core processes, such as transcription and translation, and on genetic circuit behavior.

Application of the reduced ODE model to transcription and translation. In this section, we investigate how and the extent to which intracellular heterogeneity affects the core biological processes of transcription and translation, which are responsible for protein production. We model a gene (D) being transcribed by RNAP (S) to form a DNA-RNAP complex (c_s) to produce mRNA (m). The mRNA is then translated by ribosomes (R) to form mRNA-

ribosome complex (c_m) which produces protein P. The chemical reactions are given by



where a_s and d_s are the association and dissociation rate constants, respectively, between RNAP and the gene D, κ_s is the catalytic rate constant of formation of mRNA m , a_m and d_m are the association and dissociation rate constants, respectively, between ribosomes and mRNA, κ_m is the catalytic rate constant of formation of protein P, α_s is the production rate of RNAP, α_r is the ribosome production rate, μ is the cell growth rate constant (set to unity in our nondimensionalization), and γ is the mRNA degradation rate constant. The transcription reaction is in the form of Case II (Table 1) since the gene does not freely diffuse and the RNAP freely diffuses. The translation process falls under Case I, since both mRNA and ribosomes freely diffuse. We assume that the total concentration of D is conserved, so that $D_T(x) = D(t, x) + c_s(t, x)$ and that $D_T(x)$ is localized at $x = x^*$. From (12), the dimensionless reduced space averaged dynamics corresponding to (16) are given by

$$\begin{aligned}
 \frac{d\bar{c}_s(t)}{dt} &= a_s \theta_s^* \bar{D}(t) \bar{S}(t) - (d_s + \kappa_s + 1) \bar{c}_s(t), \\
 \frac{d\bar{S}(t)}{dt} &= \bar{\alpha}_s(t) - a_s \theta_s^* \bar{D}(t) \bar{S}(t) + (\kappa_s + d_s) \bar{c}(t) - \bar{S}(t), \\
 \frac{d\bar{m}(t)}{dt} &= \kappa_s \bar{c}_s(t) - a_m \theta_R^* \bar{R}(t) \bar{m}(t) + (d_m + \kappa_m) \bar{c}(t) - (\gamma + 1) \bar{m}(t) \\
 \frac{d\bar{c}_m(t)}{dt} &= a_m \theta_R^* \bar{R}(t) \bar{m}(t) - (d_m + \kappa_m + 1 + \gamma) \bar{c}(t), \\
 \frac{d\bar{R}(t)}{dt} &= \bar{\alpha}_r(t) - a_m \theta_R^* \bar{R}(t) \bar{m}(t) + (\gamma + d_m + \kappa_m) \bar{c}_m(t) - \bar{R}(t), \\
 \frac{d\bar{P}(t)}{dt} &= \kappa_m \bar{c}_m(t) - \bar{P}(t).
 \end{aligned} \tag{17}$$

Concentration variables are nondimensionalized with respect to the total steady state space averaged RNAP ($\bar{S}_T(\infty) = \bar{\alpha}_s(\infty)/\mu$), since this quantity is a readily available in the literature. Letting r_s , r_m , and r_R , be the radius of gyration of RNAP, mRNA, and ribosomes, respectively, we compute the BCF's via (13) and (2),

$$\begin{aligned}
 \theta_s^* &= \hat{v}_s(x^*) = \frac{e^{-(r_s/r^*)^2 \hat{\rho}(x^*)}}{\int_0^1 e^{-(r_s/r^*)^2 \hat{\rho}(x)} dx}, \\
 \theta_R^* &= \int_0^1 \hat{v}_m(x) \hat{v}_R(x) dx = \frac{\int_0^1 e^{-\frac{r_m^2+r_R^2}{(r^*)^2} \hat{\rho}(x)} dx}{[\int_0^1 e^{-(r_m/r^*)^2 \hat{\rho}(x)} dx][\int_0^1 e^{-(r_R/r^*)^2 \hat{\rho}(x)} dx]},
 \end{aligned} \tag{18}$$

where $\hat{v}_s(x) = \frac{v_s(x)}{\int_0^1 v_s(x) dx}$, $\hat{v}_m(x) = \frac{v_m(x)}{\int_0^1 v_m(x) dx}$, $\hat{v}_R(x) = \frac{v_R(x)}{\int_0^1 v_R(x) dx}$, and $\hat{v}_c(x) = \frac{v_c(x)}{\int_0^1 v_c(x) dx}$ are the normalized available volume profiles of RNAP, mRNA, ribosomes, and of the mRNA-ribosome complex, respectively. Recall that the quantity $(r^*)^2$ is inversely proportional to the total DNA length per volume. We now consider the steady state behavior of system (17) by equating the time derivatives to zero. Specifically, we are interested in how the steady state levels of

produced mRNA and protein are affected by θ_s^* and θ_r^* and, hence, how they depend on spatial quantities such as r_s/r^* , r_m/r^* , r_R/r^* , and x^* .

Total mRNA steady state level: We are interested in investigating the role of spatial effects on the binding between RNAP and the DNA and thus on mRNA production. Here we analyze the steady state total mRNA levels ($\bar{m}_T = \bar{m} + \bar{c}_m$) of (17) rather than the free amount of mRNA (m), since \bar{m}_T is independent of θ_r^* as shown by

$$\bar{m}_T = \kappa_s \bar{c}_s / (\gamma + 1) \text{ with } \bar{c}_s = \bar{D}_T \frac{\bar{S} \theta_s^* / K_s}{1 + \bar{S} \theta_s^* / K_s}, \tag{19}$$

where $\bar{D}_T = \int_0^1 D_T(x) dx$, $K_s = d_s/a_s$, $K_R = d_m/a_m$. If $\theta_s^* = 1$ in (19), then the predicted total mRNA steady state level will be identical to that of a well-mixed model (as in (9)). From (18) and Fig 4B, if the RNAP radius of gyration is sufficiently large with respect to r^* then θ_s^* may be different from unity (depending on x^*), in which case spatial effects arise. If the DNA is localized near mid-cell ($x^* \approx 0$), then it implies that $\theta_s^* < 1$ from Fig 4B and, as a consequence, a decreased steady state total mRNA level will result. Furthermore, for very large values of r_s/r^* and $x^* \approx 0$ we have that $\theta_s^* \rightarrow 0$ and the total mRNA steady state levels will approach zero. Similarly, if the DNA is localized near the cell-poles ($x^* \approx 1$), then it implies that $\theta_s^* > 1$ from Fig 4B and, as a consequence, an increased steady state total mRNA level. This phenomenon occurs because as the excluded volume effects of RNAP are amplified (large r_s/r^*), RNAP will localize primarily in the cell poles and hence transcribe pole-localized DNA more efficiently than DNA near mid-cell (or any region where the local chromosome density is high). When designing genetic circuits, a plasmid backbone is chosen to provide a certain DNA copy number, however the backbone also determines where in the cell the plasmid localizes [9–11]. Therefore, based on our results, localization also affects steady state total mRNA level. If instead of introducing the DNA via a plasmid, the DNA is integrated directly into the chromosome, then the location of integration site should be a parameter to consider.

Fig 5A shows the behavior of the steady state total mRNA level as a function of r_s/r^* and of the location of the transcribed gene, when compared to the level predicted by the well mixed model. Simulations confirm that total mRNA levels are higher for pole localized genes than those near mid-cell and that the discrepancy increases with the size of RNAP relative to r^* . The agreement between the full PDE model (S1 Text: Equation 34) and the reduced ODE model (17) provides numerical validation of the model reduction results (explicitly shown in S1 Text: Fig D). In S1 Text: Fig B, the transient response corresponding to Fig 5, for which the full-PDE and reduced models agree. Furthermore, in S1 Text: Fig B, we also verify that as the size of RNAP increases, it is indeed ejected from the chromosome and adopts its available volume profile (Remark 1). Furthermore, in S1 Text: Fig C, we demonstrate that these results hold independent of the binding and unbinding speed between RNAP and DNA (Remark 2). In S1 Text: Section 8, we propose an experimental method to test the hypothesis that mid-cell genes are transcribed less effectively than pole localized genes.

In [26], it was estimated that $r_s = 6.5 \pm 0.1$ nm, which implies that $r_s/r^* \approx 0.3$. From Fig 4B, this implies that $\theta_s^* \approx 1.06$ when the DNA is at the cell poles and $\theta_s^* \approx 0.94$ when the DNA is near mid-cell, thus we expect the binding strength between RNAP and the DNA to deviate by 6% from that of a well-mixed model. From (19), if $S \theta_s^* / K_s \ll 1$, then $\bar{m}_T = \frac{\theta_s^* \kappa_s \bar{D}_T \bar{S}}{K_s (\gamma + 1)}$; thus in this regime the mRNA concentration is proportional to θ_s^* . So we expect at most a 6% difference in steady state mRNA concentration with respect to what is predicted by a well-mixed model.

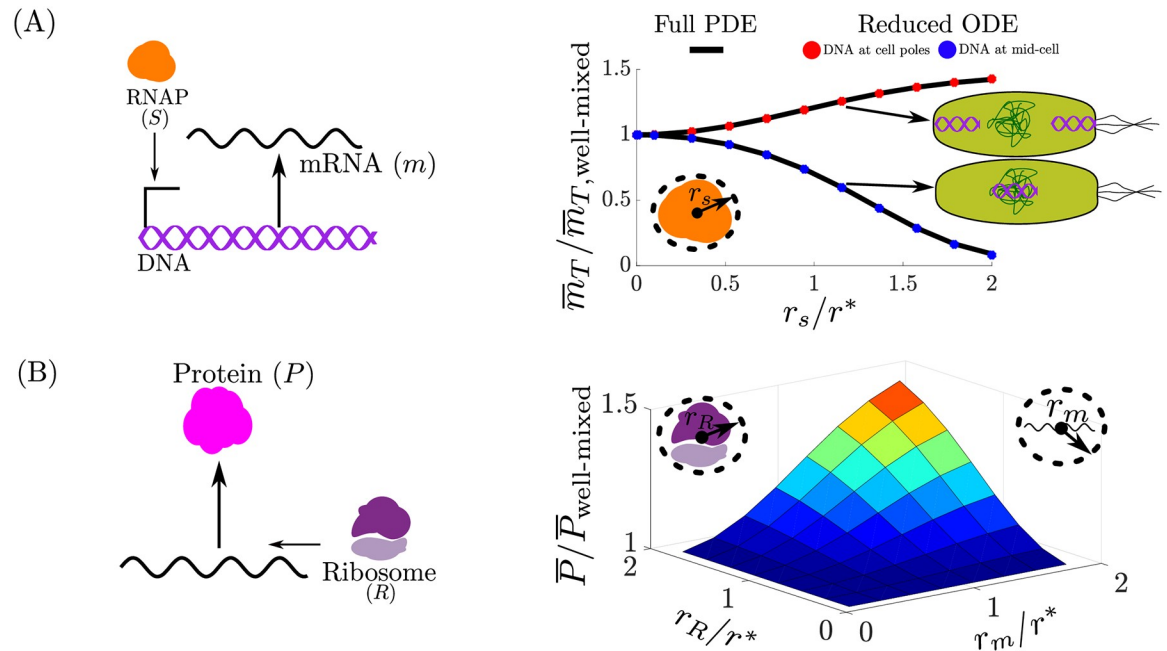


Fig 5. Spatial heterogeneity effects on steady state total mRNA and protein levels. (A) The space averaged total mRNA (\bar{m}_T) concentration predicted by the full PDE model (S1 Text: Equation 34) and the reduced ODE model (17) normalized by that of the well-mixed model ($\bar{m}_{T, \text{well-mixed}}$) as the size of the RNAP (r_s) varies with respect to r^* . With respect to the well-mixed model, the amount of mRNA decreases (increases) when the DNA is localized near mid-cell (cell poles). (B) The space averaged protein concentration (\bar{P}) predicted by the full PDE model (S1 Text: Equation 34) normalized by that of the well-mixed model ($\bar{P}_{\text{well-mixed}}$) as the size of the mRNA (r_m) and ribosome (r_R) varies with respect to r^* . The amount of protein increases when both the mRNA and ribosome size increases. We set $r_s/r^* = 1 \times 10^{-3}$, such that $\theta_s \approx 1$ and thus the result is independent of the spatial location where the gene is expressed. We refer to the well-mixed model as (17) with (18) given by $\theta_s^* = 1$ and $\theta_R^* = 1$. The parameter values and full simulation details are provided in S1 Text: Section 3.

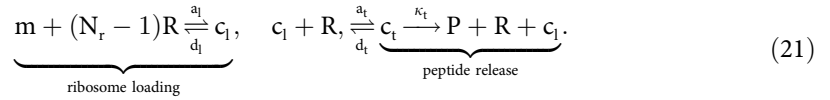
<https://doi.org/10.1371/journal.pcbi.1008159.g005>

Protein steady state level: The steady state protein levels of (17) are given by

$$\bar{P} = \kappa_m \bar{c}_m, \quad \text{with} \quad \bar{c}_m(t) = \bar{m}_T \frac{\bar{R}/K_R \theta_R^*}{1 + \bar{R}/K_R \theta_R^*}. \quad (20)$$

From (19) and (20), if $\theta_s^* = 1$ and $\theta_R^* = 1$, then protein steady state level will be identical to that of a well-mixed model. From (18) and Fig 4A, we conclude that $\theta_R^* \geq 1$ and increases with r_s/r^* and r_m/r^* . Increasing r_s/r^* and r_m/r^* implies that the ribosomes and mRNA are further excluded from the chromosome onto the cell poles. Once localized at the cell-poles, the mRNA and ribosomes are more likely to bind. Fig 5B shows the behavior of the steady state protein levels as a function of r_m and r_R , when compared to the level predicted by the well mixed-model for the full PDE model (S1 Text: Equation 34). Simulations confirm that protein levels with respect to a well-mixed model increases when both the mRNA and ribosome size are sufficiently large. In S1 Text: Fig F, we show that the reduced ODE model (17) is within 2% of the full PDE model (S1 Text: Equation 34) for the result in Fig 5B. In S1 Text: Fig E, we show the transient response corresponding to Fig 5, for which the full-PDE and reduced model agree. Furthermore, in S1 Text: Fig E, we verify that as the size of ribosome and mRNA increase, they are ejected from the chromosome and become distributed according to their available volume profile (Remark 1). Finally, in S1 Text: Figs G and F, we demonstrate that these results hold independent of the binding and unbinding speed between ribosomes and mRNA (Remark 2).

It is well known that most mRNA-ribosome complexes exists in configurations with multiple ribosomes bounded (polysomes) [27, 28]. To capture the prevalence of these polysomes, we model the translation process accounting for the fact that one mRNA can be bound to multiple ribosomes. We first model the mRNA binding simultaneously to $N_r - 1$ ribosomes to form the c_1 complex, to which another ribosome binds to to form the fully loaded c_t complex. The leading ribosome with a complete peptide is released from c_t at a rate κ_t to yield protein P. This is described by the following set of biochemical reactions:



While the ribosome loading reaction in (21) is not in the form of the chemical reactions (1), which assume bimolecular reactions, we can nevertheless apply our results as follows (see S1 Text: Section 3.1 for details). Specifically, ribosome and mRNA profiles will still approach their available volume profiles (Remark 1), that is, $R(t, x) \approx \bar{R}(t)\hat{v}_r(x)$, $m(t, x) \approx \bar{m}(t)\hat{v}_m(x)$, and $c_i(t, x) \approx \bar{c}_i(t)\hat{v}_{c_i}(x)$ where $v_{c_i}(x) = v_r^{N_r-1}(x)v_m(x)$ (recall (10)) and $\hat{v}_{c_i}(x) = v_{c_i}(x)/[\int_0^1 v_{c_i} dx]$. This is verified through simulations in S1 Text: Fig I. By virtue of the reactants in (21) mirroring their available volume profiles and (8), we can approximate the BCF of the loading θ_l^* and translation θ_t^* reactions in (21), given as

$$\theta_l^* = \frac{\int_0^1 v_m(x)v_r^{N_r-1}(x)dx}{[\int_0^1 v_m(x)dx][\int_0^1 v_r(x)dx]^{N_r-1}} = \frac{\int_0^1 e^{-\frac{2}{(r^*)^2}(\frac{r_m}{r^*} + \sqrt{N_r-1}r_R)^2\hat{\rho}(x)} dx}{[\int_0^1 e^{-(r_m/r^*)^2\hat{\rho}(x)} dx][\int_0^1 e^{-(r_R/r^*)^2\hat{\rho}(x)} dx]^{N_r-1}}.$$

$$\theta_t^* = \frac{\int_0^1 v_{c_t}(x)v_r(x)dx}{[\int_0^1 v_{c_t}(x)dx][\int_0^1 v_r(x)dx]} = \frac{\int_0^1 e^{-\frac{2}{(r^*)^2}(\frac{r_m}{r^*} + \sqrt{N_r}r_R)^2\hat{\rho}(x)} dx}{[\int_0^1 e^{-\frac{2}{(r^*)^2}(\frac{r_m}{r^*} + \sqrt{N_r-1}r_R)^2\hat{\rho}(x)} dx][\int_0^1 e^{-(r_R/r^*)^2\hat{\rho}(x)} dx]}.$$

In S1 Text: Fig I, we show computationally that θ_l^* and θ_t^* are good approximations to the BCF. At this point, we can write the ODE corresponding to this system of reactions and just modify the association rate constants by θ_l^* and θ_t^* , as shown in S1 Text: Equation 38. Let $K_d = (d_l/a_l)^{1/(N_r-1)}$, $K_t = (d_t + \kappa_t)/a_t$, $\beta_l = (\gamma + 1)/d_b$, and $\beta_t = (\gamma + 1)/(\kappa_t + d_t)$, if $\beta_l, \beta_t, \bar{R}/K_t \ll 1$ (dilution and mRNA degradation is much slower than the rate of ribosome unbinding and K_t is sufficiently large compared to \bar{R} [29]), then a simple expression for the steady state protein concentration is given by

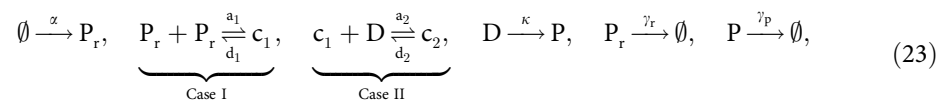
$$\bar{P} = \kappa_t \bar{m}_T \theta_t^* \bar{R} / K_t \underbrace{\frac{\theta_l^* (\bar{R}/K_d)^{N_r-1}}{1 + \theta_l^* (\bar{R}/K_d)^{N_r-1}}}_{\text{ribosome loading}}, \tag{22}$$

where \bar{m}_T is given by (19).

In [13] it was estimated that $r_m = 20$ nm and $r_R = 10$ nm, which implies that $r_m/r^* \approx 0.88$ and $r_R/r^* \approx 0.44$. Assuming the average distance between ribosomes on an mRNA to be 70 nucleotides [30], for a 700 nucleotide mRNA (e.g., GFP or RFP), then we have $N_r = 10$. Thus, for these values, $\theta_l^* \approx 1.56$ and $\theta_t^* \approx 1.07$. This implies that the forward rate in the reaction of 9 ribosomes binding to an mRNA (a_l) is amplified by 56% and the rate at which an additional ribosome binds to this complex (a_t) increases by 7% with respect to a well-mixed model. From (22), this would imply up to 67% increase in protein production with respect to a well-mixed model.

Taken together, these results suggest that while a well-mixed ODE model may be sufficient to describe transcription, it is not sufficiently descriptive to capture spatial effects on translation, particularly ribosome loading. In this case, the BCF should be incorporated in the ODE. Additionally, these results are indicative that for other processes in the cell where complexes of similar size as polysomes are formed, then spatial effects will likely be substantial.

Gene expression regulation by transcription factors. Regulation of gene expression is often performed by transcription factors (TFs) [1]. A transcription factor can either enhance (for activators) or repress (for repressors) transcription. Spatial effects play an identical role in gene regulation via activators as they do in gene regulation via RNAP (Fig 5), thus we focus on transcriptional repressors. In this section, we model transcription regulation where a repressor P_r dimerizes to form dimer c_1 (e.g., TetR dimerizes before binding to a gene [31]) and then blocks transcription of gene D that produces protein P. The biochemical reactions corresponding to this process are:



where α is the production rate of P_r , a_1 (d_1) is the association (dissociation) rate constant to form the c_1 complex, a_2 (d_2) is the association (dissociation) rate constant to form the c_2 complex, κ is the catalytic rate constant to produce protein P, and γ_r and γ_p are the degradation rate constant of P_r and P, respectively. Notice that we have lumped the transcription and translation process to produce P_r into one production reaction and similarly for P. From the results of the previous section, we know that $\bar{\alpha}(t)$ depends on the location where P_r is expressed (higher if its coding DNA is near the cell poles than mid-cell) and the size of its mRNA (higher for longer mRNAs). Similarly, κ depends on the location where P is expressed and its mRNA size. Using the results from the previous section, we can explicitly model these dependences, however, we opt not to do so to solely investigate the role of spatial effects on transcriptional repression. Since the repressor P_r , freely diffuses, the dimerization reaction belongs to Case I. The gene D is spatially fixed and it is repressed by the freely diffusing c_1 , thus this interaction falls under Case II. We assume that the total concentration of D is conserved, so that $D_T(x) = D(t, x) + c_2(t, x)$ and that $D_T(x)$ is localized at $x = x^*$. The reduced ODE model corresponding to (23) obeys

$$\begin{aligned} \frac{d\bar{P}_r(t)}{dt} &= \bar{\alpha}(t) - \gamma_r \bar{P}_r(t), & \frac{d\bar{c}_1(t)}{dt} &= a_1 \theta_1^* \bar{P}_r^2(t) - d_1 \bar{c}_1(t) - a_2 \theta_2^* \bar{c}_1(t) \bar{D}(t) + d_2 \bar{c}_2(t), \\ \frac{d\bar{c}_2(t)}{dt} &= a_2 \theta_2^* \bar{D}(t) \bar{c}_1(t) - d_2 \bar{c}_2(t), & \bar{D}(t) &= 1 - \bar{c}_2(t), & \frac{d\bar{P}(t)}{dt} &= \kappa \bar{D}(t) - \gamma_p \bar{P}(t). \end{aligned} \quad (24)$$

Concentration variables were nondimensionalized with respect to the space averaged total DNA $\bar{D}_T = \int_0^1 D_T(x) dx$. From our main result, the BCF's are given by

$$\theta_1^* = \int_0^1 \hat{v}_{P_r}^2(x) dx = \frac{\int_0^1 v_{P_r}^2(x) dx}{[\int_0^1 v_{P_r}(x) dx]^2}, \quad \theta_2^* = \hat{v}_{c_1}(x^*) = \frac{v_{c_1}(x^*)}{\int_0^1 v_{c_1}(x) dx}, \quad (25)$$

where $v_{P_r}(x) = e^{-(r/r^*)^2 \hat{\rho}(x)}$ and $v_{c_1}(x) = v_{P_r}^2(x)$ (recall (10)) are the available volume profiles of P_r and c_1 , respectively, and r is the radius of gyration of P_r .

We now consider the steady state behavior of system (24) by equating the time derivatives to zero. Specifically, we are interested in how the steady state levels of P is affected by the

spatial quantities r/r^* and x^* . From setting (24) to steady state, we obtain

$$\bar{P} = \frac{\kappa}{\gamma_p} \bar{D}, \quad \text{where} \quad \bar{D} = \frac{1}{1 + (\bar{P}_r/K)^2 \theta^*}, \quad \theta^* = \theta_1^* \theta_2^* = \hat{v}_{P_r}^2(x^*), \quad (26)$$

where $K = \sqrt{K_{d,1} K_{d,2}}$ and $K_{d,i} = d_i/a_i$ for $i = 1, 2$. From (26) and (25), we observe that θ^* contains all the spatial information, which includes the size of P_r and the location of the target gene D. If $\theta^* = 1$, then the protein concentration would be the same as the well-mixed model. The ratio $K/\sqrt{\theta^*}$ can be thought of as an effective disassociation rate constant of the repressor. If D is located near mid-cell ($x^* \approx 0$ in (25)), then for $r/r^* \ll 1$ we have $\theta^* \approx 1$ (see Fig 4), but as r/r^* increases, we have that $\theta^* < 1$ and asymptotically approaches zero as $r/r^* \rightarrow \infty$. Similarly, if D is located near the cell poles ($x^* \approx 1$ in (25)), then for $r/r^* \ll 1$ we have $\theta^* \approx 1$ (see Fig 4), but as r/r^* increases, we have that $\theta^* > 1$. Thus, the efficacy of a transcriptional repressor regulating genes in the chromosome (cell-poles) decreases (increases) with TF size. Intuitively, this occurs because as the TF size increases, excluded volume effects will push it out of the chromosome onto the cell-poles (see Remark 1), thus interacting with DNA near the cell-poles more frequently than with DNA near mid-cell. Numerical simulations validate our predictions as shown in Fig 6, where increasing the transcription factor size leads to higher (lower) repression when the target DNA is localized at the cell poles (mid-cell) with respect to a well-mixed model. The simulation results also show agreement between the predictions of the full PDE (S1 Text: Equation 42) and reduced ODE model (24) (as shown explicitly in S1 Text: Fig L). Furthermore, we demonstrate in S1 Text: Fig J, that this agreement extends to the temporal dynamics. Finally, all our results hold independent of the binding and unbinding speeds of the transcription factor dimerizing and of the dimer binding to the DNA (S1 Text: Fig K). In S1 Text: Section 8, we propose an experimental method to test the hypothesis that mid-cell genes are regulated less effectively than pole localized genes.

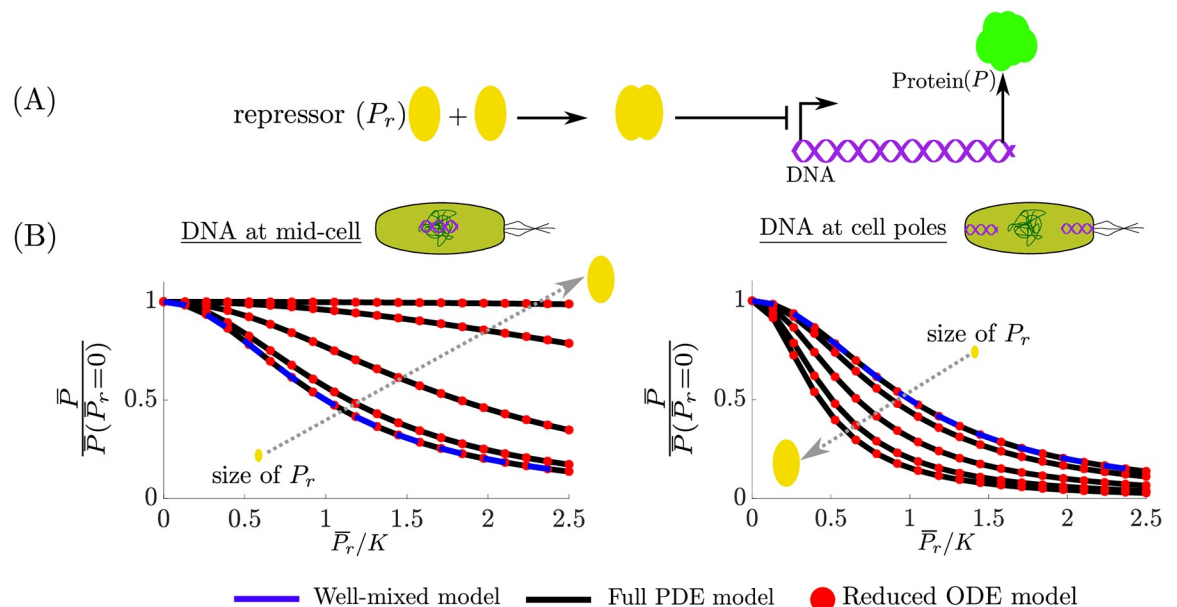


Fig 6. Spatial heterogeneity effects in transcriptional regulation. (A) The repressor P_r dimerizes and regulates the production of protein P. (B) The steady state space-averaged concentration per-cell of \bar{P} normalized by its value when $\bar{P}_r = 0$ (26) for the PDE model (S1 Text: Equation 42), the well-mixed model ((24) with $\theta_1^* = \theta_2^* = 1$), and the reduced ODE model (24) when the DNA is located near mid-cell ($x^* \approx 0$ in (26)) and when the DNA is located at the cell-poles ($x^* \approx 1$) for several sizes of P_r . The parameter values and full simulation details are provided in S1 Text: Section 4.

<https://doi.org/10.1371/journal.pcbi.1008159.g006>

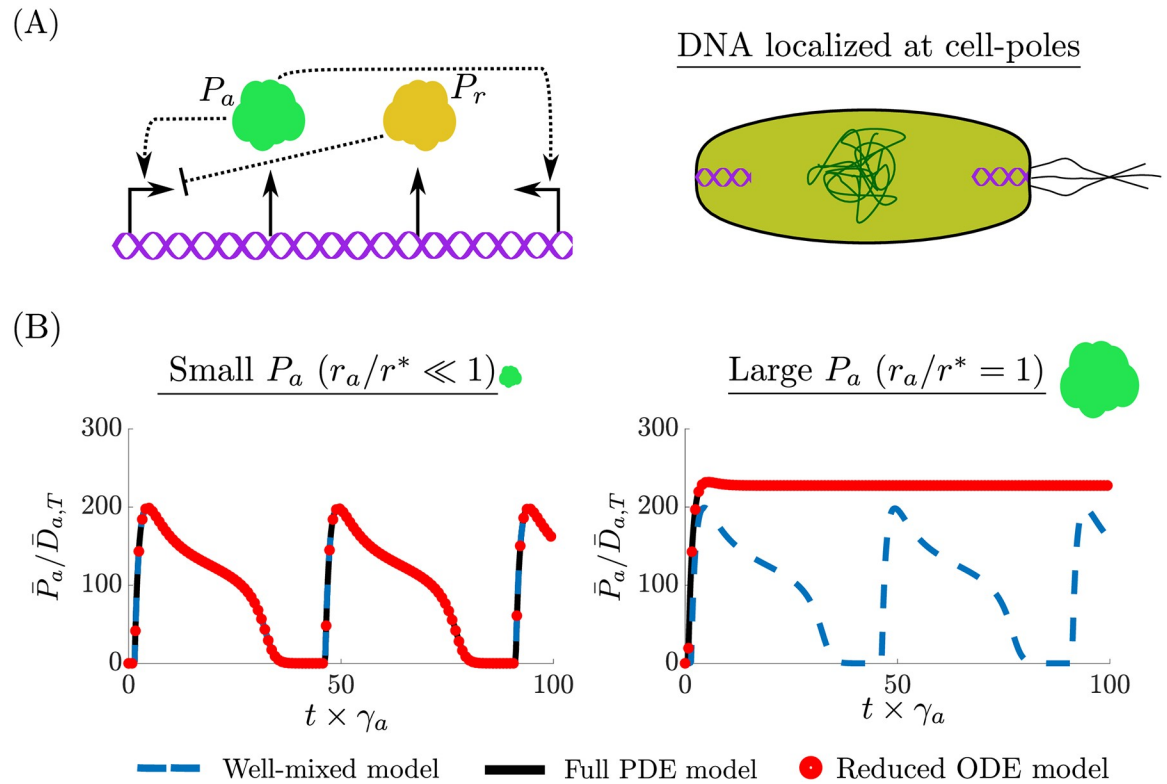


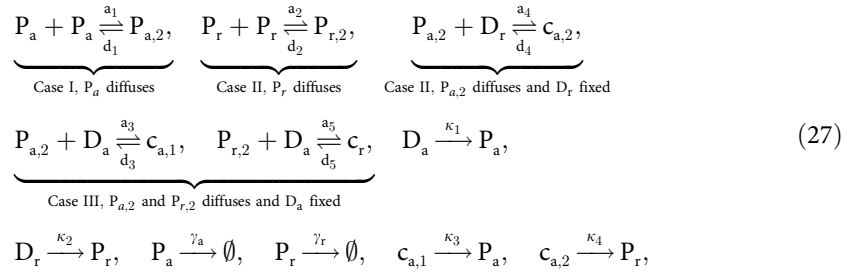
Fig 7. Spatial effects on the dynamics of genetic circuits. (A) The activator-repressor clock where P_r represses P_a and P_a activates itself and P_r . Both proteins are expressed from the same cell-pole localized plasmid. (B) The temporal evolution of P_a is given for the full-PDE model (S1 Text: Equation 44), the reduced ODE model (28), and the well-mixed model (same as (28) with $\theta_{a,1}^* = \theta_{a,2}^* = \theta_r^* = 1$). When P_a is small ($r_a/r^* \ll 1$), all three models predict sustained oscillations. When P_a is large ($r_a/r^* = 1$), the full-PDE model and the reduced ODE model predict the oscillations will cease. For both simulations $r_r/r^* \ll 1 \Rightarrow \theta_r^* \approx 1$. The full simulation details and parameter values are given in S1 Text: Section 5.

<https://doi.org/10.1371/journal.pcbi.1008159.g007>

The reactions in (23) can be easily extended to CRISPRi/dCas9 repression systems [34], where instead of two identical species dimerizing, we have two distinct freely diffusing species bind (dCas9 and guide RNA) to form the complex gRNA-dCas9, which targets a desired DNA sequence. Exploiting the insight gained from analyzing (23), we expect that due to the large size of dCas9 [35] (which is further augmented as it forms a complex with the gRNA), it will regulate pole localized DNA (e.g. ColE1 plasmid DNA [9]) more efficiently than genes in mid-cell (e.g. chromosomally integrated) and thus spatial effects are expected to be more significant when using CRISPRi/dCas9 in genetic circuit design. Specifically, based on approximate values found in the literature, we estimated that $\theta^* \approx 1$ for a transcription factor, while θ^* can range between 0.9 and 1.1 for dCas9-enabled repression. This indicates that a well-mixed model is appropriate for modeling transcription factor-enabled repression of gene expression but may not be sufficient to capture effects of spatial heterogeneity arising with larger repressing complexes such as with dCas9/gRNA (see S1 Text: Section 4 for details).

Genetic oscillator. As a final example, we consider the repressor-activator clock genetic circuit designed in [32] and shown in Fig 7A. This circuit produces sustained oscillations if tuned within an appropriate parameter range [1, 33]. The circuit consists of two proteins P_a and P_r . Protein P_a , is an activator which dimerizes to form $P_{a,2}$ and then binds to its own gene D_a to form complex $c_{a,1}$ to initiate transcription. The dimer $P_{a,2}$ also binds to the gene D_r , which transcribes P_r to form complex $c_{a,2}$ and initiates transcription. Protein P_r , dimerizes

to form $P_{r,2}$ and then represses P_a by binding to D_a to form complex c_r . The biochemical reactions corresponding to this circuit are:



where a_i (d_i) for $i = 1, \dots, 5$ are association (dissociation) rate constants, γ_a (γ_r) is the degradation rate constant of P_a (P_r) κ_1 (κ_2) is the basal rate at which gene D_a (D_r) is transcribed, and κ_3 (κ_4) is the rate at which the DNA-transcription-factor complexes are transcribed for D_a (D_r). We assume that the total concentration of D_a is conserved, so that $D_{a,T}(x) = D_a(t, x) + c_{a,1}(t, x) + c_r(t, x)$ and that $D_{a,T}$ is localized at $x = x_a^*$. Similarly, we assume that the total concentration of D_r is conserved, so that $D_{r,T}(x) = D_r(t, x) + c_{a,2}(t, x)$ and that $D_{r,T}$ is localized at $x = x_r^*$. The reduced ODE model corresponding to (27) is given by:

$$\begin{aligned}
 \frac{d\bar{P}_a(t)}{dt} &= \kappa_1 \bar{D}_a(t) + \kappa_3 \bar{c}_{a,1}(t) - \gamma_a \bar{P}_a(t), & \frac{d\bar{P}_r(t)}{dt} &= \kappa_2 \bar{D}_r(t) + \kappa_4 \bar{c}_{a,2}(t) - \gamma_r \bar{P}_r(t), \\
 \frac{d\bar{P}_{a,2}(t)}{dt} &= a_1 \theta_1^* \bar{P}_a^2(t) - d_1 \bar{P}_{a,2}(t) \\
 &\quad - a_3 \theta_3^* \bar{P}_{a,2}(t) \bar{D}_a(t) + d_3 \bar{c}_{a,1}(t) - a_4 \theta_4^* \bar{P}_{a,2}(t) \bar{D}_r(t) + d_4 \bar{c}_{a,2}(t), \\
 \frac{d\bar{P}_{r,2}(t)}{dt} &= a_2 \theta_2^* \bar{P}_r^2(t) - d_2 \bar{P}_{r,2}(t) - a_5 \theta_5^* \bar{P}_{r,2}(t) \bar{D}_a(t) + d_5 \bar{c}_r(t), \\
 \frac{d\bar{c}_{a,1}(t)}{dt} &= a_3 \theta_3^* \bar{P}_{a,2}(t) \bar{D}_a(t) - d_3 \bar{c}_{a,1}(t), \\
 \frac{d\bar{c}_{a,2}(t)}{dt} &= a_4 \theta_4^* \bar{P}_{a,2}(t) \bar{D}_r(t) - d_4 \bar{c}_{a,2}(t), & \frac{d\bar{c}_r(t)}{dt} &= a_5 \theta_5^* \bar{P}_{r,2}(t) \bar{D}_a(t) - d_5 \bar{c}_r(t), \\
 \bar{D}_a(t) &= 1 - \bar{c}_{a,1}(t) - \bar{c}_r(t), & \bar{D}_r(t) &= \bar{D}_{r,T} - \bar{c}_{a,2}(t).
 \end{aligned} \tag{28}$$

Concentration variables were nondimensionalized with respect to the space averaged total DNA $\bar{D}_{a,T} = \int_0^1 D_{a,T}(x) dx$. Applying our main result, the BCF's are given by

$$\begin{aligned}
 \theta_1^* &= \int_0^1 \hat{v}_{P_a}^2(x) dx, & \theta_2^* &= \int_0^1 \hat{v}_{P_r}^2(x) dx, \\
 \theta_3^* &= \hat{v}_{P_{a,2}}(x_a^*), & \theta_4^* &= \hat{v}_{P_{a,2}}(x_r^*), & \theta_5^* &= \hat{v}_{P_{r,2}}(x_a^*),
 \end{aligned}$$

where $\hat{v}_{P_a}(x)$, $\hat{v}_{P_r}(x)$, $\hat{v}_{P_{a,2}}(x)$, and $\hat{v}_{P_{r,2}}(x)$ are the normalized available volume profiles (i.e., $\hat{v}_{P_a}(x) = v_{P_a}(x) / \int_0^1 v_{P_a}(x) dx$) of P_a , P_r , $P_{a,2}$, and $P_{r,2}$, respectively. From (10), notice that $v_{P_{a,2}}(x) = v_{P_a}^2(x)$ and $v_{P_{r,2}}(x) = v_{P_r}^2(x)$. The available volume profiles are given by

$$v_{P_a}(x) = e^{-(r_a/r^*)^2 \hat{\rho}(x)}, \quad v_{P_r}(x) = e^{-(r_r/r^*)^2 \hat{\rho}(x)}, \tag{29}$$

where r_a and r_r are the radius of gyration of P_a and P_r , respectively. Approximating $\bar{P}_{a,2}(t)$, $\bar{P}_{r,2}(t)$, $\bar{c}_{a,1}(t)$, $\bar{c}_{a,2}(t)$ and $\bar{c}_r(t)$ at their quasi-steady state (since $d_i \gg \gamma_a, \gamma_r$ for $i = 1, \dots, 5$, [1]),

we obtain

$$\begin{aligned} \bar{P}_{a,2}(t) &= \frac{a_1\theta_1}{d_1}\bar{P}_a^2(t), \quad \bar{P}_{r,2}(t) = \frac{a_2\theta_2}{d_2}\bar{P}_r^2(t), \\ \bar{c}_{a,1}(t) &= \frac{a_3\theta_3}{d_3}\bar{P}_{a,2}(t)\bar{D}_a(t), \quad \bar{c}_{a,2}(t) = \frac{a_4\theta_4}{d_4}\bar{P}_{a,2}(t)\bar{D}_r(t), \quad \bar{c}_r(t) = \frac{a_5\theta_5}{d_5}\bar{P}_{r,2}(t)\bar{D}_a(t), \end{aligned}$$

and, therefore, we can further reduce (28) to

$$\frac{d\bar{P}_a}{dt} = \frac{\alpha_{0,A} + \alpha_A \left(\frac{\bar{P}_a}{K_{d,1}}\right)^2}{1 + \left(\frac{\bar{P}_a}{K_{d,1}}\right)^2 + \left(\frac{\bar{P}_r}{K_{d,2}}\right)^2} - \gamma_A \bar{P}_a, \quad \frac{d\bar{P}_r}{dt} = \frac{\alpha_{0,R} + \alpha_R \left(\frac{\bar{P}_a}{K_{d,3}}\right)^2}{1 + \left(\frac{\bar{P}_a}{K_{d,3}}\right)^2} - \gamma_R \bar{P}_r \quad (30)$$

where $\alpha_{0,A} = \kappa_1 \bar{D}_{a,T}$ ($\alpha_{0,R} = \kappa_2 \bar{D}_{r,T}$) is the basal production rate of P_a (P_r), $\alpha_A = \kappa_3$ ($\alpha_R = \kappa_4 \bar{D}_{r,T}$) is the additional production rate of P_a (P_r) due to activation from P_a , and

$$K_{d,1} = \frac{d_1 d_3}{\theta_{A,1}^* a_1 a_3}, \quad K_{d,2} = \frac{d_2 d_5}{\theta_R^* a_2 a_5}, \quad K_{d,3} = \frac{d_1 d_4}{\theta_{A,2}^* a_1 a_4}, \quad (31)$$

$$\theta_{A,1}^* = \theta_1^* \theta_3^* = \hat{v}_{P_a}(x_a^*), \quad \theta_R^* = \theta_2^* \theta_5^* = \hat{v}_{P_r}(x_r^*), \quad \theta_{A,2}^* = \theta_1^* \theta_4^* = \hat{v}_{P_a}(x_r^*). \quad (32)$$

The form of the dynamics given by (30) was theoretically analyzed in [1, 33], and it was shown that the values of $K_{d,i}$ for $i = 1, 2, 3$, were critical in determining whether sustained oscillations occur. From (31), these parameters depend on (32) and thus on the size of P_a and P_r through the available volume profiles (29) and the location of D_a and D_r (x_a^* and x_r^*). Numerical simulations demonstrate how these spatial parameters affect circuit behavior. In our simulation setup, the parameters are chosen such that the well-mixed model ((28) with $\theta_R^* = \theta_{A,1}^* = \theta_{A,2}^* = 1$) oscillates, the DNA of P_a and P_r are localized at the cell poles and have the same copy number (i.e., $x_a^* = x_r^*$ and $\bar{D}_{r,T} = \bar{D}_{a,T} = 1$), the size of P_r is chosen to be small $r_r/r^* \ll 1$ (thus $\theta_R^* \approx 1$), and the size of P_a is varied (thus varying $\theta_{A,1}^*$ and $\theta_{A,2}^*$). Since D_a is localized at the cell poles, it implies $x_a^* \approx 1$ and from (32), we observe that if $r_a/r^* \ll 1 \Rightarrow \theta_{A,1}^* \approx \theta_{A,2}^* \approx 1$ and $\theta_{A,1}^*, \theta_{A,2}^*$ increase as r_a/r^* increases. The results of these simulations are shown in Fig 7. When $r_a/r^* \ll 1$, the full PDE model (S1 Text: Equation 44), the reduced ODE model (28), and the well-mixed model are all in agreement and sustained oscillations are observed. By contrast, when $r_a/r^* = 1$, the PDE and reduced model (which are in agreement with each other as explicitly shown in S1 Text: Fig N) predict that sustained oscillations will no longer occur. Furthermore, in S1 Text: Figure M, we demonstrate that indeed as the size of P_a increases it is excluded from the chromosome onto the cell poles while the spatial profile of P_r is homogeneously distributed throughout the cell since $r_r/r^* \ll 1$ (Remark 1).

Discussion

We derived a reduced order ODE model of genetic circuits with the same dimension as traditional ODE well-mixed models; yet, it captures effects of spatial heterogeneity within bacterial cells (12). In particular, our reduced model is the same as a well-mixed model where all the association rate constants are multiplied by the binding correction factor (BCF). This factor depends on the size and location (if fixed in space) of the reacting species, according to an analytical formula that we derived from first principles (13) and its value can be estimated experimentally through simple procedures (S1 Text: Section 7). We have mathematically

demonstrated that this reduced order model is a good approximation of the space-averaged dynamics resulting from a reaction-diffusion PDE model under the assumption of fast diffusion. It can therefore be used in place of PDE models, providing substantial advantages for both simulation and mathematical analysis.

We applied this model to analyze the effects of spatial heterogeneity on core biological processes and genetic circuits. Specifically, motivated by the fact that DNA, ribosomes, and mRNA have been shown to localize within the cell [4, 5, 14], we analyzed the transcription and translation processes. We determined that mRNA levels are lower (higher) when the gene is localized near the mid-cell (cell poles). We also showed that when the target gene of a transcriptional repressor is near mid-cell (cell poles) the effective repression is lower (higher) with respect to that of the well-mixed model. This discrepancy is amplified as the size of the transcription factor increases. The extent of these spatial effects depends on how different the value of the BCF is from unity. Based on parameters found in the literature, we determined that for the processes of transcription and its regulation the BCF should be close to unity and hence a well-mixed ODE model should be sufficient. However, in situations where the nucleoid is highly compacted (from overexpressing mRNA [13] or translational inhibition [36]), we expect that the available volume profile (2) approaches small values and, as a consequence, the value of the BCF can substantially deviate from unity (13).

Our results provide additional interpretations of well-known biological phenomena. For example, it has been shown that the expression rate of chromosomal genes depends on the locus where the gene is inserted [37]; that the nucleoid dynamically changes shape to control gene expression and transcription regulation [4, 38] (e.g., see *S1 Text*: Section 9, for how a time varying chromosome density modulates the BCF); and that coregulation and coexpression among genes depends on their spatial distance [39]. For a fixed amount of mRNA, we showed that spatial heterogeneity leads to higher translation rates since both mRNA and ribosome are pushed out of the chromosome into a smaller region near the cell poles, which results in larger effective binding affinity. How larger, it depends on the value of the BCF. For a polyosome with 10 translating ribosomes, the value of the BCF can deviate from unity by 56% in the ribosome loading step and by 7% in the peptide release step. These estimates are believed to be conservative since we did not account for the exclusion effects from the peptide chains attached to the translating ribosome, which will result in even more pronounced spatial effects. Therefore, a well-mixed model may not be sufficient to capture the effects of spatial heterogeneity on translation.

Our modeling framework can be easily extended to other aspects of gene expression. For example, we may consider co-transcriptional translation [40]. In this case, as a result of translation being localized at the gene location, the effective ribosome binding site strength will also depend on gene location through the BCF. We may also consider the role of spatial heterogeneity on orthogonal translational machinery [41]. From our models, we predict that one can tune the rate at which orthogonal ribosomes are formed by creating larger synthetic 16S rRNA. Furthermore, once the production of orthogonal ribosomes is placed in a feedback form to decouple genetic circuits [41], our framework suggests that the feedback efficiency may depend on the spatial location of the synthetic 16S rRNA gene. The value of the parameter r^* , whose squared value is inversely proportional to the average chromosome density (2), is critical in determining the extent of spatial effects. In this study we indirectly estimated a value of r^* based on [13]. However, a more comprehensive study should be conducted to estimate r^* for several contexts (*S1 Text*: Section 7), or equivalently to estimate extent of excluded volume effects, which may easily be performed via superresolution imaging [14].

In summary, this paper provides a general and convenient modeling framework to account for DNA localization and excluded volume effects on intracellular species dynamics. While

other phenomena contributing to intracellular spatial heterogeneity, such as crowding [42], sliding, hopping, and dimensionality [17], exist, this is a first step towards creating a general framework to modify current models to capture spatial information. Our model can be used both as an analysis and a design tool for genetic circuits, in which variables such as gene location and regulator size may be considered as additional design parameters.

Supporting information

S1 Text. Supporting information file with mathematical proofs, detailed analysis of examples, generalization and extension of the results and additional simulations.

(PDF)

Acknowledgments

We thank Jean-Jacques Slotine for the technical discussions on contraction theory. We thank Theodore Grunberg and Yili Qian for reviewing the article and helpful discussions.

Author Contributions

Conceptualization: Carlos Barajas, Domitilla Del Vecchio.

Data curation: Carlos Barajas.

Formal analysis: Carlos Barajas.

Funding acquisition: Carlos Barajas, Domitilla Del Vecchio.

Investigation: Carlos Barajas.

Methodology: Carlos Barajas, Domitilla Del Vecchio.

Project administration: Domitilla Del Vecchio.

Resources: Domitilla Del Vecchio.

Software: Carlos Barajas.

Supervision: Domitilla Del Vecchio.

Validation: Carlos Barajas.

Visualization: Carlos Barajas.

Writing – original draft: Carlos Barajas, Domitilla Del Vecchio.

Writing – review & editing: Carlos Barajas, Domitilla Del Vecchio.

References

1. Del Vecchio D, Murray RM. Biomolecular Feedback Systems; 2014.
2. Alon U. An Introduction to Systems Biology: Design Principles of Biological Circuits; 2007.
3. Wingreen NS, Huang KC. Physics of Intracellular Organization in Bacteria. *Annual Review of Microbiology*. 2015; 69(1):361–379. <https://doi.org/10.1146/annurev-micro-091014-104313> PMID: 26488278
4. Weng X, Xiao J. Spatial organization of transcription in bacterial cells. *Trends in Genetics*. 2014; 30(7):287–297. <https://doi.org/10.1016/j.tig.2014.04.008> PMID: 24862529
5. van Gijtenbeek LA, Robinson A, van Oijen AM, Poolman B, Kok J. On the Spatial Organization of mRNA, Plasmids, and Ribosomes in a Bacterial Host Overexpressing Membrane Proteins. *PLoS Genetics*. 2016; 12(12):1–28. <https://doi.org/10.1371/journal.pgen.1006523> PMID: 27977669
6. Saberi S, Emberly E. Non-Equilibrium Polar Localization of Proteins in Bacterial Cells. *PLoS ONE*. 2013; 8(5):1–9. <https://doi.org/10.1371/journal.pone.0064075> PMID: 23700458

7. Brackley CA, Taylor S, Papantonis A, Cook PR, Marenduzzo D. Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. *Proceedings of the National Academy of Sciences*. 2013; 110(38):E3605–E3611. <https://doi.org/10.1073/pnas.1302950110> PMID: 24003126
8. Macnamara CK, Chaplain MAJ. Spatio-temporal models of synthetic genetic oscillators. *Mathematical Biosciences and Engineering*. 2016; 14(1):249–262. <https://doi.org/10.3934/mbe.2017016> PMID: 27879131
9. Yao S, Helinski DR, Toukdarian A. Localization of the naturally occurring plasmid CoIE1 at the cell pole. *Journal of Bacteriology*. 2007; 189(5):1946–1953. <https://doi.org/10.1128/JB.01451-06> PMID: 17158664
10. Ebersbach G, Gerdes K. Plasmid Segregation Mechanisms. *Annual Review of Genetics*. 2005; 39(1):453–479. <https://doi.org/10.1146/annurev.genet.38.072902.091252> PMID: 16285868
11. Wang Y, Penkul P, Milstein JN. Quantitative Localization Microscopy Reveals a Novel Organization of a High-Copy Number Plasmid. *Biophysical Journal*. 2016; 111(3):467–479. <https://doi.org/10.1016/j.bpj.2016.06.033> PMID: 27508432
12. Englaender JA, Jones JA, Cress BF, Kuhlman TE, Linhardt RJ, Koffas MAG. Effect of Genomic Integration Location on Heterologous Protein Expression and Metabolic Engineering in *E. coli*. *ACS Synthetic Biology*. 2017; 6(4):710–720. <https://doi.org/10.1021/acssynbio.6b00350> PMID: 28055177
13. Castellana M, Hsin-Jung Li S, Wingreen NS. Spatial organization of bacterial transcription and translation. *Proceedings of the National Academy of Sciences*. 2016; 113(33):9286–9291. <https://doi.org/10.1073/pnas.1604995113> PMID: 27486246
14. Bakshi S, Siryaporn A, Goulian M, Weisshaar JC. Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Molecular Microbiology*. 2012; 85(1):21–38. <https://doi.org/10.1111/j.1365-2958.2012.08081.x> PMID: 22624875
15. Stoof R, Wood A, Goni-Moreno A. A model for the spatio-temporal design of gene regulatory circuits. *bioRxiv*. 2019; p. 522946. <https://doi.org/10.1101/522946>
16. Halatek J, Frey E. Rethinking pattern formation in reaction-diffusion systems. *Nature Physics*. 2018; 14(5):507–514. <https://doi.org/10.1038/s41567-017-0040-5>
17. Soh S, Byrska M, Kandere-Grzybowska K, Grzybowski BA. Reaction-diffusion systems in intracellular molecular transport and control. *Angewandte Chemie—International Edition*. 2010; 49(25):4170–4198. <https://doi.org/10.1002/anie.200905513> PMID: 20518023
18. de Jong H. Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *Journal of Computational Biology*. 2002; 9(1):67–103. <https://doi.org/10.1089/10665270252833208> PMID: 11911796
19. Barajas C, Vecchio DD. Genetic Circuit-Host Ribosome Transactions: Diffusion-Reaction Model. *Acc*. 2019; p. 1533–1540.
20. Teimouri H, Korkmazhan E, Stavans J, Levine E. Sub-cellular mRNA localization modulates the regulation of gene expression by small RNAs in bacteria. *Physical Biology*. 2017; 14(5). <https://doi.org/10.1088/1478-3975/aa69ac> PMID: 28350301
21. Amir A, van Teeffelen S. Getting into shape: How do rod-like bacteria control their geometry? *Systems and Synthetic Biology*. 2014; 8(3):227–235. <https://doi.org/10.1007/s11693-014-9143-9> PMID: 25136385
22. Leah Edelstein-Keshet. *Mathematical Models in Biology*. Siam. 2012; XXXIII(2):81–87.
23. Mishra D, Rivera PM, Lin A, Del Vecchio D, Weiss R. A load driver device for engineering modularity in biological networks. *Nature Biotechnology*. 2014; 32(12):1268–1275. <https://doi.org/10.1038/nbt.3044> PMID: 25419739
24. Cho SH, Godin JM, Chen CH, Qiao W, Lee H, Lo YH. Review Article: Recent advancements in optofluidic flow cytometer. *Biomicrofluidics*. 2010; 4(4):043001. <https://doi.org/10.1063/1.3511706> PMID: 21267434
25. Chen W, Niepel M, Sorger P. Classic and contemporary approaches to modeling biochemical reactions. *Genes & Development*. 2010; 24:1861–1875. <https://doi.org/10.1101/gad.1945410> PMID: 20810646
26. Meisenberger O, Pilz I, Heumann H. Small-angle x-ray study of DNA-dependent RNA polymerase subunit $\alpha 2$ from *Escherichia coli*. *FEBS Letters*. 1980; 120(1):57–60. [https://doi.org/10.1016/0014-5793\(80\)81045-3](https://doi.org/10.1016/0014-5793(80)81045-3) PMID: 7002611
27. Andreeva I, Belardinelli R, Rodnina MV. Translation initiation in bacterial polysomes through ribosome loading on a standby site on a highly translated mRNA. *Proceedings of the National Academy of Sciences of the United States of America*. 2018; 115(17):4411–4416. <https://doi.org/10.1073/pnas.1718029115> PMID: 29632209

28. Ruscetti FW, Jacobson LA. Accumulation of 70S monoribosomes in *Escherichia coli* after energy source shift-down. *Journal of Bacteriology*. 1972; 111(1):142–151. <https://doi.org/10.1128/JB.111.1.142-151.1972> PMID: 4591472
29. Gyorgy A, Jiménez JI, Yazbek J, Huang HH, Chung H, Weiss R, et al. Isocost Lines Describe the Cellular Economy of Genetic Circuits. *Biophysical Journal*. 2015; 109(3):639–646. <https://doi.org/10.1016/j.bpj.2015.06.034> PMID: 26244745
30. Dennis PP, Bremer H. Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*. 2008; 3(1). <https://doi.org/10.1128/ecosal.5.2.3> PMID: 26443740
31. Ramos JL, Marti M, Molina-henares AJ, Tera W, Brennan R, Tobes R. TetR family of transcriptional. 2005; 69(2):1–31.
32. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*. 2003; 113(5):597–607. [https://doi.org/10.1016/S0092-8674\(03\)00346-5](https://doi.org/10.1016/S0092-8674(03)00346-5) PMID: 12787501
33. Del Vecchio D. Design and analysis of an activator-repressor clock in *E. Coli*. In: *Proceedings of the American Control Conference*; 2007. p. 1589–1594.
34. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*. 2013; 8(11):2180–2196. <https://doi.org/10.1038/nprot.2013.132> PMID: 24136345
35. Gibson GJ, Yang M. What rheumatologists need to know about CRISPR/Cas9. *Nature Reviews Rheumatology*. 2017; 13(4):205–216. <https://doi.org/10.1038/nrrheum.2017.6> PMID: 28202911
36. Sanamrad A, Persson F, Lundius EG, Fange D, Gynna AH, Elf J. Single-particle tracking reveals that free ribosomal subunits are not excluded from the *Escherichia coli* nucleoid. *Proceedings of the National Academy of Sciences*. 2014; 111(31):11413–11418. <https://doi.org/10.1073/pnas.1411558111> PMID: 25056965
37. Sánchez-Romero MA, Lee D, Sánchez-Morán E, Busby SJW. Location and dynamics of an active promoter in *Escherichia coli* K-12. *Biochemical Journal*. 2012; 441(1):481–485. <https://doi.org/10.1042/BJ20111258> PMID: 21936772
38. Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, Kleckner N. Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell*. 2013; 153(4):882–895. <https://doi.org/10.1016/j.cell.2013.04.006> PMID: 23623305
39. Pannier L, Merino E, Marchal K, Collado-Vides J. Effect of genomic distance on coexpression of coregulated genes in *E. coli*. *PLoS ONE*. 2017; 12(4):1–20. <https://doi.org/10.1371/journal.pone.0174887> PMID: 28419102
40. Bakshi S, Choi H, Weisshaar JC. The spatial biology of transcription and translation in rapidly growing *Escherichia coli*. *Frontiers in Microbiology*. 2015; 6(JUL):1–15.
41. Darlington APS, Kim J, Jiménez JI, Bates DG. Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nature Communications*. 2018; 9(1). <https://doi.org/10.1038/s41467-018-02898-6> PMID: 29449554
42. Tabaka M, Kalwarczyk T, Szymanski J, Hou S, Holyst R. The effect of macromolecular crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics*. 2014; 2(September):1–14.