Modeling of gene regulatory processes by population-mediated signaling: New applications of population balances

Che-Chi Shua, Anushree Chatterjeeb, Wei-Shou Hub, Doraiswami Ramkrishna

a School of Chemical Engineering, Purdue University, West Lafayette, IN 47907, United States
b Department of Chemical Engineering & Materials Science, University of Minnesota, Minneapolis, MN 55455, United States

Abstract

Population balance modeling is considered for cell populations in gene regulatory processes in which one or more intracellular variables undergo stochastic dynamics as determined by Ito stochastic differential equations. This paper addresses formulation and computational issues with sample applications to the spread of drug resistance among bacterial cells. It is shown that predictions from population balances can display qualitative differences from those made with single cell models which are usually encountered in the literature. Such differences are deemed to be important.

1. Introduction

We address in this paper the modeling of gene regulatory processes in which a signaling molecule in the cellular environment initiates a set of intracellular reactions culminating in the synthesis of a protein through the expression of a specific gene. Because the number of molecules participating in reaction is small, the behavior of the reaction system is characteristically stochastic in nature. Consequently some cells display high protein levels representing “on” situation of the “gene switch”, while other cells with low protein levels represent the “off” state of the switch. The “on” and “off” states often occur with bistability, a scenario that has attracted numerous publications in the literature in which the stochastic analysis of the reaction system provides for bimodal protein level distributions among the cell population, which can also be compared with experiments in a flow cytometer through the use of fluorescent dyes. Such bimodal distributions arise as solutions of a Fokker–Planck equation representing the stochastic behavior of a single cell. A large number of publications exist in the literature that lead to the impression that bistability and bimodal distributions occur together. Since observations on gene expression in single cells cannot be made directly, existence of bistability has been inferred (Gardner et al., 2000; Kepler and Elston, 2001; Ferrell, 2002; Kobayashi et al., 2004; Ozbudak et al., 2004; Tian and Burrage, 2006) by observations of bimodal distributions of protein levels using a cytometer. Our focus in this paper is on situations in which the gene regulatory process is influenced by the behavior of a population of cells. For example, the signaling molecule may arise as secretions of a population of other cells. In some situations, reaction species, transported out of the cell, may subsequently participate again in the intracellular reactions by reentering the cells but not before redistribution by mixing in the extracellular environment. The single cell analysis that appears in the literature for describing gene expression in a population is inadequate to address the foregoing scenarios in which extracellular environment is altered by interaction with the cells and by transport processes. Consequently, it becomes apparent that the problem requires the framework of population balances.

A detailed development of the population balance framework has been presented by Fredrickson et al. (1967) for addressing the dynamics of microbial populations. Subsequent additions have been made to this development (Ramkrishna, 1979; Fredrickson and Mantzaris, 2002; Hjortso, 2005). These developments have, however, dealt with deterministic intracellular processes and therefore cannot account for stochasticity in gene regulatory processes. Although Ramkrishna (1979) indicates how random growth rate can be treated in population balance models of cell populations, the formulation of population balance models in which stochastic changes occur in intracellular variables as...
coordinates may be described by growth. For growth in biofilms the cells may be regarded as sessile, variables connected with growth and metabolism. The spatial gene regulation as the internal coordinates assuming independence only the stochastic intracellular variables, comprising only the stochastic intracellular variables, connected with gene regulation as the internal coordinates assuming independence between the stochastic variables and the deterministic variables connected with growth and metabolism. The spatial coordinates introduced can account for different scenarios of growth. For growth in biofilms the cells may be regarded as sessile, while in the so-called planktonic growth cells may be regarded as well mixed so that we may dispense with spatial coordinates. The kinetics of gene regulation are contained in the function \(X(x|C_x)\), where \(C_x\) is the vector of extracellular variables associated with gene regulation (for example, in the context of our application, this vector would contain the concentrations of the signaling molecule and the inhibitor molecule). The stochasticity of gene expression is represented by \(B_x(x|C_x)dW\), where \(B_x(x|C_x)\) is a matrix (see Table 2) and \(dW\) is a vector of standard Wiener processes (see for example Gardiner, 1997). Cellular motion with respect to fixed coordinates may be described by \(R_x(x|C_x)\). The population balance equation may then be written as

\[
\frac{\partial n(x,r,t)}{\partial t} + \nabla_x \cdot X(x|C_x)n(x,r,t) + \nabla_r \cdot R(x|C_x)n(x,r,t) = \frac{1}{2} \nabla_x \cdot \nabla_r [B_x(x|C_x)B_x^T(x|C_x)n(x,r,t) + \mu(r,t|C_y)h_y(x,r,t)]
\]

where \(\mu(r,t|C_y)\) is the specific growth rate. \(C_y\) are the extracellular variables relating to cell growth and metabolism and \(r\) are the spatial coordinates. Details of the derivation of Eq. (1) are available in Ramkrishna (2000). As pointed out earlier, Appendix B shows how the spatio-temporal dependence of the averaged kinetic parameters has been eliminated. Eq. (1) is of course coupled with mass balances for the environmental variables given by

\[
\frac{\partial C_x}{\partial t} + \nabla_x \cdot N_x = \int C_x n(x,r,t) dx
\]

where \(N_x\) are the total flux of \(C_x\) due to convection and diffusion and the vector \(C_x\) relates change of extracellular variables to intracellular reactions associated with gene regulation. It is to be noted that Eq. (1) features the variable \(C_y\) which must be regarded as a known time-dependent variable, obtained by simultaneous solution of the population balance equation (B.24) for cell growth and multiplication coupled with Eq. (B.21) for associated extracellular variables (see Appendix B).

Although computational methods can be devised for the simultaneous solution of Eqs. (1) and (2), their implementation would be highly time consuming. In what follows, however, we introduce additional considerations that will further simplify this aspect. Let the initial conditions for Eqs. (1) and (2) be given by

\[
n(x,r,0) = N_0(r|x_0,0)|C_x(r,0) = C_{x0}(r)
\]

where \(N_0(r)\) is the initial spatial distribution of the total number density of cells, \(f_0(x)\) the distribution of stochastic variables among the cell population, and \(C_{x0}(r)\) the spatial distribution of extracellular variables connected to the gene regulatory process.

Our goal is to transform Eq. (1) further to something more amenable to solution. Towards this end, we assume that cell motion is effected by an external device rather than by intracellular mechanisms so that \(R\) depends neither on \(x\) nor \(C_x\) but does only on \(r\) and \(t\). Thus it is possible to define a path of cellular motion by a function \(r = R(r(t),0)\), \(0 < t < t\) satisfying \(r_0 = R(0,0)\) and \(r = R(r(t))\), which implies that the motion begins at \(r_0\) at time \(t = 0\) and followed until reaching \(r\) at time \(t = t\). In the absence of diffusion, we presume that it is possible to uniquely solve for the initial position of the cell given its current coordinates \((r,t)\) and represent this inverse relationship by \(r_0 = R_0(r,t)\), which defines a time-dependent transformation between \(r\) and \(x\) representing initial locations. The Jacobian of the forward transformation, \(J(r(t),r_0) = |\partial R/\partial r|\), can be readily shown to satisfy the relations (Aris, 1962)

\[
\frac{D}{Dt} = \frac{\partial}{\partial t} + \nabla_r \cdot R
\]

The population balance equation (1) becomes on replacing \(R_x(x|C_x)\) by \(R(r,t)\)

\[
\frac{Dn(x,r,t)}{Dt} + \nabla_x \cdot X(x|C_x)n(x,r,t) + \nabla_r \cdot R(x|C_x)n(x,r,t) = \frac{1}{2} \nabla_x \cdot \nabla_r [B_x(x|C_x)B_x^T(x|C_x)n(x,r,t) + \mu(r,t|C_y)h_y(x,r,t)]
\]

The initial condition for (4) is readily seen to be

\[
j(r_0,0|x_0)n(x_0,0) = N_0(r_0|x_0)
\]

which follows from \(j(r_0,0|x_0) = 1\) and (3). We may further define the function

\[
P(x,r,t) = \frac{1}{N_0(r(t))}j(r(t),r_0)n(x,r,t)
\]

\[
\times \exp \left[ - \int_0^t \mu(r(t)|C_y)dt \right]
\]

where \(r_0 = R_0(r(t))\) is obtained by backtracking along the cell path from position \(r\) at time \(t\). The function \(P(x,r,t)\) satisfies the partial differential equation

\[
\frac{DP(x,r,t)}{Dt} + \nabla_x \cdot X(x|C_x)P(x,r,t) = \frac{1}{2} \nabla_x \cdot \nabla_r [B_x(x|C_x)B_x^T(x|C_x)P(x,r,t)]
\]
which is a Fokker–Planck equation with time measured along the path of cell motion so that quantities depending on the cell’s spatial coordinates become readily identified functions of time. The initial condition for \( P(x,r,t) \) is given by

\[
P(x,r,0) = \frac{1}{N_0(r)} \delta(x,r,0) = f_{x0}(x)
\]

The advantage of Eq. (6) lies in its possession of an Ito form of stochastic differential equations given by

\[
dx = X(x; C_x) dt + B_x(x; C_x) dW \tag{7}
\]

which can be solved directly by established numerical methods. However, it is important to bear in mind that the solution of Eq. (7) must be considered simultaneously with (2).

2.1. Sessile cells: growth in biofilms

The case of sessile cells is encountered in biofilm growth situations. Thus we set \( R = 0 \) for which \( f = 1 \) at all locations and times and Eq. (6) becomes

\[
\frac{\partial P(x,r,t)}{\partial t} + \nabla_x \cdot X(x; C_x) P(x,r,t) = \frac{1}{2} \nabla_x \nabla_r : B_x(x; C_x) B_{x}^T(x; C_x) P(x,r,t) \tag{8}
\]

implying the same Ito stochastic equations as (7) except that time is followed at the fixed location of the cell. Again, one must account for the coupling of (7) with (2).

2.2. Cells in well-mixed environment

If we regard the cells to be in a closed, well-mixed spatial environment, then Eq. (8) may be integrated over this domain to obtain the following Fokker–Planck equation:

\[
\frac{\partial P(x,t)}{\partial t} + \nabla_x \cdot X(x; C_x) P(x,t) = \frac{1}{2} \nabla_x \nabla_r : B_x(x; C_x) B_{x}^T(x; C_x) P(x,t) \tag{9}
\]

which again translates to the Ito equation (7) applicable to the entirely different situation considered earlier. However, the function \( C_x \) must satisfy the differential equation applicable to this well-mixed case, viz.,

\[
\frac{dC_x}{dt} = N(t) \int C_x P(x,t) dx = N(t) \mathbb{E}[C_x] \tag{10}
\]

Eq. (10) must be modified slightly to account for any volume changes and degradation processes that may exist in the extracellular environment (such situations are considered in examples to follow). The number density can be obtained from the general transformation equation (5) adapted to the well-mixed case given by

\[
P(x,t) = \frac{1}{N_0} f_{x}(x,t) \exp \left[ - \int_0^t \mu(t') dt' \right] \tag{11}
\]

In the following section, we discuss some of the computational details of the problems just posed.

3. Computational issues

We have in the previous section identified a situation of separation of processes associated with growth and multiplication of cells from the gene regulatory process specific to an application. As stated before, the population balance model for the gene regulatory process is given by Eq. (10) and the stochastic differential equations (7) or alternatively the Fokker–Planck equation (8) to be solved simultaneously. Although the Fokker–Planck equation can be solved directly by path-integral methods (Wehner and Wolfer, 1983), the computational burden tremendously increases with increase in the number of variables. Thus, we opt for the solution of Ito equations (7) with stochastic algorithms (Rao et al., 1974; Talay, 1995) to obtain sample-pathwise realizations that can be averaged to obtain expected values such as \( \mathbb{E}[C_x] \). In this study, we choose the Euler method (Talay, 1995) because our focus is to demonstrate phenomena instead of pursuing higher order accuracy. Furthermore, we seek the solution only for an infinitesimal interval at each step. In next section, we then consider an application to the transfer of drug resistance to present some interesting results.

4. Application to transfer of drug resistance

4.1. The biological background

We are concerned with a gene regulatory process connected with the transfer of drug resistance in Enterococcus faecalis (Dunny et al., 1981) induced by a molecule called cCF10, pheromone, released by a population of recipient cells. The actual transfer of drug resistance occurs as a result of conjugation between the so-called donor cells (containing plasmids referred to as pCF10 on which genes encode drug resistance) and recipient cells not resistant to the drug. The gene regulatory network of pCF10 conjugation is shown in Fig. 1.

The transport of cCF10 across the cell membrane occurs by an active transport mechanism involving a membrane protein. On entering the cell, the cCF10 alters the DNA configuration and favors the expression of what is referred to as a prgQ gene, resulting in the formation of a molecule called Q specifying the synthesis of a protein called Q which binds to regulatory protein which alters the DNA configuration through sense-antisense interaction. The ratio of cCF10 to inhibitor controls the configuration of Q which has no effect on stimulating prgB expression. High concentration of cCF10 prevents the expression of prgB protein, thereby reducing the production of PrgB protein. Consequently, high concentration of Q.
Q_{prQ} is required to overcome Anti-Q (Bae et al., 2004; Chatterjee et al., 2011; Johnson et al., 2010) and make a cell produce Q_{prQ} mRNA (Shi et al., 2005) towards stimulating downstream gene expression including that of prgB. On the other hand, when the ratio of cCF10 to iCF10 is low the expression of prgQ gene is insufficient. Small amount of Q_{prQ} becomes Q_{prQ} in nearly no prgB expression. In a nut shell, a subtle change of producing iCF10 nonlinearly responding to additive ratios of the same. The foregoing observations are made because of their implications to the transfer of drug resistance which is the application of eventual interest.

4.2. Population balance equation of pCF10 conjugation system

We consider a well-mixed system of cells in a closed domain. The vector of intracellular stochastic variables, x has 6 components in this application listed in Table 1. The population balance equation transformed to the Fokker–Planck form, is given by Eq. (9). The extracellular variable vector C_{x} has two components, one the cCF10 (C_{1}) concentration in the environment and the other that of the iCF10 (C_{2}). The vector X(x,C_{x}) is identified in Table 2 and results from (i) mechanisms of the concerned reactions published elsewhere (Nakayama et al., 1994; Bensing and Dunny, 1997; Buttaro et al., 2000; Bae et al., 2004; Fixen et al., 2007) and (ii) application of the Chemical Master Equation to obtain the stochastic version identifying the matrix B_{x}(x,C_{x}). Following Haseltine and Rawlings (2002) and Rao and Arkin (2003), the noise terms of RNA variables are set to zero to reduce computational burden. As a result, some of the coefficients of the foregoing matrix become zero. The details of the derivation of the stochastic matrix whose coefficients are identified in Table 2 are provided in Appendix A. The parameter values are shown in Table 3 and the Euler algorithm (Talay, 1995) is applied for solving Ito stochastic differential equation.

4.3. Bistability from deterministic equations of single cell

Following the mass action law and the mechanism of the pCF10 conjugation system, the deterministic behavior of single
shown in Table 4. The main difference between Tables 4 and 2 is in which implies cells are fully isolated from each other. The Fokker–Planck equation of single cell stochastic model has been based on the single cell stochastic model formulation according to mass action law are featured with bistability in pCF10 conjugation system. The deterministic equations formulated according to mass action law are featured with bistability in pCF10 conjugation system with the parameter values shown in Table 4. Bistability occurs as a consequence of two (stable) steady state protein levels, one at the low end representing the “off” state, and the other at the high end representing the “on” state for a given concentration of cCF10 within a suitable range, 25.9–31.7 nM for this case. The middle steady state is unstable. Through bistability a cell behavior becomes switch like, either “off” or “on”. For a constant cCF10 concentration, a jump from “off” state to “on” is not possible for deterministic model but possible for stochastic model.

4.4. Single cell stochastic model

The general approach in the literature for modeling gene regulatory processes has been based on the single cell stochastic model which implies cells are fully isolated from each other. The Fokker–Planck equation of pCF10 conjugation system for single cell model is shown in Table 4. The main difference between Tables 4 and 2 is in the extracellular equation. The single cell stochastic model ignores the extracellular concentration change due to the population. Consequently, the extracellular concentration of the single cell model is a stochastic process included in the probability function. On the other hand, the extracellular variable Cx of population balance model is deterministic. The simulation of pCF10 conjugation system from Fig. 3a. Notice that the bimodal distribution is consistent with steady state values shown in Fig. 2. Although incorporating stochasticity allows the single cell model to analyze cell random behavior, its result cannot be interpreted as a population distribution but only as the probability distribution of a cell.

4.5. Different predictions from population balance model and single cell stochastic model

The population balance model of this paper provides the appropriate tool for curing the foregoing drawback of a single
Fokker–Planck equation for single cell.

\[
\frac{\partial p(x,t)}{\partial t} + \nabla_x \cdot \left( \frac{1}{2} \nabla_x \right) p(x,t) = \frac{1}{2} \nabla_x \cdot \left( \frac{1}{2} \nabla_x \right) p(x,t)
\]

\[
X(x,C) = \begin{bmatrix}
K_{1,1} x + K_{1,2} (1-x) & K_{1,3} x + K_{1,4} (1-x) \\
K_{2,1} x + K_{2,2} (1-x) & K_{2,3} x + K_{2,4} (1-x)
\end{bmatrix}
\]

\[
B_{22} = B_{44} = 0
\]

\[
B_{24} = B_{42} = 0
\]

\[
B_{26} = B_{62} = 0
\]

\[
B_{46} = B_{64} = 0
\]

\[
\frac{dC_2}{dt} = v_2 N \int [K_{1,6}(x_1 + x_2) - K_{2,6} C_2] p(x,t) dx - \left( K_{4,5} + \frac{d\ln V}{dt} \right) C_2
\]

Fig. 3. The difference between single cell stochastic model and population balance model. (a) The result from single cell stochastic model. The extracellular cCF10 concentration is 30 nM. From deterministic model, we know the corresponding stable steady state values are 35 and 163 nM. The bimodal distribution shows consistent with these values. (b) The result from population balance model. Cells are assumed well stirred in the system with extracellular cCF10 concentration equaling to 30 nM. Because cells indirectly influence others through extracellular inhibitor the pattern of cell behavior is altered form bimodal distribution to unimodal distribution.

Fig. 4. Protein distribution of donors in a well mixed system, including donors and recipients, without conjugation. (a) PrgB protein distribution of donors. Phoromone, cCF10, concentration increases due to growth of the recipient population, somewhat faster than donors. Consequently, PrgB levels increase showing higher variance to stochastic fluctuations. (b) Number density for PrgB protein. Compare with protein distribution shown in (a), to note the increase in cell numbers due to replication.
the entire well-stirred culture (for the population) is the same. Eq. (13) maybe rewritten as

$$\frac{dC_1}{dt} = v_d N [K_{1,6}E(x_1 + x_2) - K_{2,6}C_2] - \left( K_{4,5} + \frac{dnV}{dt} \right) C_2$$  \hspace{1cm} (14)

The simulation is shown in Fig. 3b assuming a well-stirred system described by equations in Table 2. The population balance model clearly produces a qualitatively different result from that of the single cell stochastic model as the former produces a unimodal distribution while the latter obtains a bimodal distribution. Compared to Fig. 3a (cell in isolated circumstance), cells in Fig. 3b (well-stirred system) have less ability to spread drug resistance because all cells are at “off” state. In general, a planktonic cell moves in the culture medium so that the well-stirred assumption may be plausible. For a cell immobilized by an extracellular matrix, such as biofilms (Kristich et al., 2004), an isolated situation may be possible (Redfield, 2002). In fact, unpublished results of collaborating group show unimodal distribution in the new donor cells is the same as those in the old donor cells conjugating with the recipient cells. Relaxing this assumption, which is introduced to simplify the model, would considerably increase the complexity of the model.

For the purposes of this paper, we consider two other cases.

**Case 2.** We consider the same environment as in Case 1 with the following difference. We hold the culture volume constant but allow the population of donor cells to increase exponentially. We consider the same environment as in Case 1 with the following difference. We hold the culture volume constant but allow the population of donor cells to increase exponentially.

The environmental variable equations are given by

$$\frac{dC_1}{dt} = v_d N [K_{1,8}E(x_1 + x_2) - K_{2,8}C_2]$$

while the differential equation for $C_2$ remains the same as in (15). Eqs. (17) and (18) must be solved simultaneously with

$$\frac{dN}{dt} = \mu N - k_{con}NNE[x]$$

The computational procedure involves solution of Eq. (7) for computing the right-hand sides of Eqs. (15), (17) and (19). The results of computation are presented in Fig. 5 for the number density (5a) as well as the probability density (5b). Both show protein level distributions increase at the beginning, subsequently decreasing because of an increase in the number of donors from transformed recipient cells. It must be noted, however, that these results were obtained for models with simplifications that are not necessarily true.
5. Concluding remarks

The motivation of this paper has been to expound a methodology of population balances that includes stochastic particulate behavior for application to a general class of important biological applications in which gene regulatory processes play a significant role. This paper shows that population balance models can display substantially different dynamic behavior relevant to the applications from the single cell approach that has been used in the literature for the analysis of gene regulatory processes. A particularly interesting extension of this work is the study of conjugation including recipient and donor cell populations in the spread of infection.

Nomenclature

\( A_C \) stoichiometric matrix relating the rates of change of environmental variables to those of the intracellular reactions

\( A_Z \) stoichiometric matrix relating change of extracellular variables to intracellular stochastic reactions

\( \mathbf{C}(z|C) \) vector of rates of change of environmental variables due to intracellular reactions

\( \mathbf{C}_{X}(\mathbf{x}|C_X) \) vector of rates of change of environmental variables due to intracellular reactions associated with gene regulation only

\( \mathbf{C}_{Y}(y|C_Y) \) vector of rates of change of environmental variables due to intracellular reactions associated with growth and duplication

\( B_{X}(x|C) \) if \( B_{X}(z|C) \) does not depend on \( y \), \( B_{X}(z|C) \) can be written as \( B_{X}(x|C) \)

\( B_{X}(x|C_X) \) same as \( B_{X}(x|C) \) but further specified extracellular variables

\( B_{Z}(z|C) \) matrix describing stochastic change of \( z \) to stochastic differential equation

\( C \) concentration of all extracellular variables

\( C_X \) concentration of extracellular variables relating to stochastic intracellular variables

\( C_{X}(r|C) \) initial condition of \( C_X \) at \( t = 0 \)

\( C_{Y}(r|C) \) initial condition of \( C_Y \) at \( t = 0 \)

\( f_{X}(x,r,t) \) probability density function, \( f_{X}(z,r,t) = \frac{G_{x}(z,r,t)}{N(r,t)} \)

\( f_{X}(x) \) give the initial condition of \( n_{x} \) at \( t = 0 \)

\( f_{Y}(y,r,t) \) probability density function, \( f_{Y}(z,r,t) = \frac{G_{y}(z,r,t)}{N(r,t)} \)

\( f_{Z}(z,r,t) \) probability density function, \( f_{Z}(z,r,t) = \frac{G_{z}(z,r,t)}{N(r,t)} \)

\( f_{Z}(z) \) \( n_{z}(r) \)\( f_{z}(z) \) give the initial condition of \( n_{z} \) at \( t = 0 \)

\( j \) Jacobian of the forward transformation, \( j(r,t,n_{r}) = \frac{\partial R}{\partial r} \)

\( m \) number of stochastic variables

\( n \) number of deterministic variables

\( N(r,t) \) total number density at location \( r \) and time \( t \)

\( N(i,0) \) initial condition of \( N(r,t) \) at \( t = 0 \)

\( N_X \) total flux, including convection and diffusion, of extracellular variables relating to stochastic intracellular variables

\( N_Y \) total flux, including convection and diffusion, of extracellular variables relating to deterministic intracellular variables

\( N_Z \) total flux of extracellular variables due to convection and diffusion

\( n_{X}(x,r,t) \) number density, \( \frac{d}{dt} n_{X}(x,r,t) \)

\( n_{Y}(y,r,t) \) number density, \( \frac{d}{dt} n_{Y}(y,r,t) \)

\( n_{Z}(z,r,t) \) number density, number of cells per unit volume of both internal and spatial coordinates

\( p_{Y}(y|x,c) \) distribution of states \( y \) of daughters from a parent of state \( x \)

\( p_{Z}(z|x,c) \) distribution of states \( z \) of daughters from a parent of state \( x \)

\( z \) spatial coordinates

\( r \) initial condition of \( r \) at \( t = 0 \)

\( R(r|C) \) a function describing the path of cellular motion, \( r = R(r|C) \)

\( R_{d}(r,t) \) if the inverse relationship of \( r \) and \( r_{0} \) is unique, \( r_{0} = R_{d}(r,t) \)

\( R_{0}(r,t) \) rate of spatial cell displacement depend on \( r \) and \( t \)

\( R_{X}(x|C) \) if \( R_{X}(z|C) \) does not depend on \( y \), \( R_{X}(z|C) \) can be written as \( R_{X}(x|C) \)

\( R_{Y}(y|C) \) if \( R_{Y}(z|C) \) does not depend on \( y \), \( R_{Y}(z|C) \) can be written as \( R_{Y}(y|C) \)

\( R_{Z}(z|C) \) rate of spatial cell displacement depend on \( z \) and \( C \)

\( \mu(r,t|C_{Y}) \) cell division rate regardless of the cell state at location \( r \) at time \( t \)

\( dW \) standard Wiener process increment

\( x \) deterministic variables

\( X(x|C) \) if \( X_{X}(z|C) \) does not depend on \( y \), \( X_{X}(z|C) \) can be written as \( X(x|C) \)

\( X(x|C_X) \) same as \( X(x|C) \) but further specified extracellular variables

\( Y(y|C) \) if \( Y_{Y}(z|C) \) does not depend on \( x \), \( Y_{Y}(z|C) \) can be written as \( Y(y|C) \)

\( Y(y|C_Y) \) same as \( Y(y|C) \) but further specified extracellular variables

\( Y_{Z}(z|C) \) rate of change of the deterministic intracellular variables \( y \)

\( z \equiv (x,y) \) all intracellular state variables

\( Z(z|C) \) rate of change of the intracellular state vector \( z \)

\( \gamma \) vectors describing intracellular reactions

\( \sigma_{Y}(y|C_{Y}) \) if \( \sigma_{Z}(y|C_{Y}) \) does not depend on \( x \), \( \sigma_{Z}(y|C_{Y}) \) can be written as \( \sigma_{Y}(y|C_{Y}) \)

\( \sigma_{Z}(z|C) \) cell division rate

Acknowledgments

The work is supported by a grant from NIH (GM081888) to WSH.

Appendix A

The Fokker–Planck equation of pCF10 conjugation system for intracellular variables is obtained by expanding chemical master equation. If the stochasticity of all 6 intracellular variables is taken into consideration we have Eqs. (A.1), (A.2) and (A.3)

\[
\frac{\partial p(x,t)}{\partial t} + \nabla_x X_{x0}(x,C) p(x,t) = \frac{1}{2} \nabla_x \nabla_x' X_{x0}(x,C) B_{x0}(x,C) B_{x0}(x,C)' p(x,t)
\]

\[
\begin{align*}
\begin{bmatrix}
 k[K_{11}z+K_{12}(1-z)](\frac{K_{11}z+\mu_{1}}{1+K_{11}z+\mu_{1}}) & -(K_{11}z+\mu_{1})x_{1} \\
 k[K_{11}z+K_{12}(1-z)](\frac{K_{11}z+\mu_{1}}{1+K_{11}z+\mu_{1}}) & -(K_{11}z+\mu_{1})x_{2}
\end{bmatrix}
\end{align*}
\]

\[
X_{x0}(x,C)=
\begin{align*}
\begin{bmatrix}
 k[K_{11}z+K_{12}(1-z)](\frac{K_{11}z+\mu_{1}}{1+K_{11}z+\mu_{1}}) & -(K_{11}z+\mu_{1})x_{1} \\
 k[K_{11}z+K_{12}(1-z)](\frac{K_{11}z+\mu_{1}}{1+K_{11}z+\mu_{1}}) & -(K_{11}z+\mu_{1})x_{2}
\end{bmatrix}
\end{align*}
\]

\[
\begin{align*}
\begin{bmatrix}
 K_{S1}S_{C1} - (K_{S1} + \mu_{S1})x_{4} \\
 K_{S2}S_{C2} - (K_{S2} + \mu_{S2})x_{5} \\
 K_{S3}S_{C3} - (K_{S3} + \mu_{S3})x_{6}
\end{bmatrix}
\end{align*}
\]
\[ \alpha = \frac{x_4}{x_2 + K_3 S X_0} \]

\[ \mathbf{B}_{\mathbf{x}_d}(\mathbf{x}|C) \mathbf{B}_{\mathbf{x}_d}^T(\mathbf{x}|C) = \mathbf{B}_{\mathbf{all}} = \begin{bmatrix} B_{11} & 0 & B_{13} & 0 & 0 & 0 \\ 0 & B_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & B_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{66} & 0 \end{bmatrix} \]  \hspace{1cm} (A.3)

where

\[ B_{11} = \frac{1}{v_d} \left( k \{ K_{1,1} x_2 + K_{1,2} (1-x_2) \} \left( \frac{K_3 S X_1}{1 + K_3 S X_3} + K_4 x_1 \right) \right) \]

\[ B_{22} = \frac{1}{v_d} \left( k \{ K_{1,1} x_2 + K_{1,2} (1-x_2) \} \left( \frac{1}{1 + K_3 S X_3} + K_4 x_2 \right) \right) \]

\[ B_{33} = \frac{1}{v_d} \left( k \{ K_{1,1} x_2 + K_{1,2} (1-x_2) \} + k \{ K_{1,1} x_2 + K_{1,2} (1-x_2) \} \cdot \frac{K_3 S X_1}{1 + K_3 S X_3} + K_4 x_3 \right) \]

\[ B_{44} = \frac{1}{v_d} \left[ K_{2,6} C_2 + K_4 x_4 \right] \]

\[ B_{55} = \frac{1}{v_d} \left[ K_1 x_2 + K_4 x_5 \right] \]

\[ B_{66} = \frac{1}{v_d} \left[ K_{2,6} C_1 + K_4 x_6 \right] \]

\[ B_{13} = B_{31} = \frac{1}{v_d} \left\{ -k \{ K_{1,1} x_2 + K_{1,2} (1-x_2) \} \left( \frac{K_3 S X_1}{1 + K_3 S X_3} \right) \right\} \]

However, the responding time for RNA variable is much shorter than protein and peptide. A quasi-steady-state assumption can be applied while formulate the chemical master equation by separating variables into x and \( \mathbf{x}_d \), the slow and the fast reaction species. The probability can be described by \( P(\mathbf{x},t) = P(\mathbf{x}_d,\mathbf{x}_t) = P(\mathbf{x}_d) P(\mathbf{x}_t) \) with \( dP(\mathbf{x}_d) = \mu \mathbf{x}_d \mathbf{x}_d \mathbf{P}(\mathbf{x}_d) dt \approx 0 \), refer to Rao and Arkin (2003), so the noise term of \( \mathbf{x}_t \) is negligible and the approximate master equation solely in terms of \( \mathbf{x}_d \). The Fokker–Planck equation of \( \mathbf{x}_d \) is Eqs. (A.4), (A.5) and (A.6) where \( \mathbf{x}_d' = [x_4 \ x_5 \ x_6] \)

\[ \frac{\partial P(\mathbf{x}_d,t)}{\partial t} + \nabla_{\mathbf{x}_d} \cdot \mathbf{X}_d(\mathbf{x}_d|C) \mathbf{P}(\mathbf{x}_d,t) \]

\[ = \frac{1}{2} \nabla_{\mathbf{x}_d} \cdot \nabla_{\mathbf{x}_d} : \mathbf{B}_{\mathbf{x}_d}(\mathbf{x}_d|C) \mathbf{B}_{\mathbf{x}_d}^T(\mathbf{x}_d|C) \mathbf{P}(\mathbf{x}_d,t) \]  \hspace{1cm} (A.4)

\[ \mathbf{X}_d(\mathbf{x}_d|C) = \begin{bmatrix} K_{2,6} C_2 - (K_{4,6} + \mu_d) x_4 \\ K_{1,5} x_2 - (K_{4,5} + \mu_d) x_5 \\ K_{2,6} C_1 - (K_{4,6} + \mu_d) x_6 \end{bmatrix} \]  \hspace{1cm} (A.5)

\[ \mathbf{B}_{\mathbf{x}_d}(\mathbf{x}_d|C) \mathbf{B}_{\mathbf{x}_d}^T(\mathbf{x}_d|C) = \mathbf{B}_d = \begin{bmatrix} B_{44} & 0 & 0 & 0 \\ 0 & B_{55} & 0 & 0 \\ 0 & 0 & B_{66} & 0 \end{bmatrix} \]  \hspace{1cm} (A.6)

For convenience, we further put all variables together as Eq. (A.7) which is shown in Table 2.

\[ \frac{\partial \mathbf{P}(\mathbf{x},t)}{\partial t} + \nabla \cdot \mathbf{X}(\mathbf{x}|C) \mathbf{P}(\mathbf{x},t) \]

\[ = \frac{1}{2} \nabla \cdot \nabla : \mathbf{B}_d(\mathbf{x}|C) \mathbf{B}_d(\mathbf{x}|C) \mathbf{P}(\mathbf{x},t) \]  \hspace{1cm} (A.7)

First, we combined the drift term of fast reaction species, \( x_1 \), \( x_2 \) and \( x_3 \) with \( \mathbf{X}_d(\mathbf{x}_d|C) \) to make \( \mathbf{X}(\mathbf{x}|C) \). And then, to formulate \( \mathbf{B}_d(\mathbf{x}|C) \mathbf{B}_d(\mathbf{x}|C) = \mathbf{B} \), we put zero for the diffusion terms of \( x_1 \), \( x_2 \) and \( x_3 \) as there are no random.
It is further useful to define the following kinetics. Note the dependence of all the above kinetic expressions on the external mixing devices or because of other mechanisms such as product BZ variables by derive a cell displacement rate by averaging over all stochastic variables. Thus in Eq. (B.6) we may replace if, however, the term BX variables do not affect stochastic behavior of gene regulatory variables. Thus in Eq. (B.6) we may replace

\[ R_x(x, t) = \int \mathcal{R}_x(z) n_x(z, t) \, dz \]

(B.8)

cell displacement rate as

\[ \frac{\partial n_x(x, t)}{\partial t} + \nabla_x \cdot Z_x(z) n_x(z, t) = \mathcal{R}_x(z) n_x(z, t) \]

(B.11)

In Eq. (B.11), we have used subscripts in the gradient operator \( \nabla \) to denote the set of partial derivatives included in the gradient. For example,

\[ \nabla_x = \left( \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n} \right) \]

(B.12)

For the method of derivation of this equation connected with the stochastic terms, the reader is referred to Ramkrishna (2000). Specification of the integration range is avoided in favor of implying constraints in the function \( p_{x|z}(z, x, t) \) that will allow non-zero values only when the states of the parent and the offspring are compatible. Similar to the averaged rates in Eq. (B.5), we may also define the following

\[ \sigma_y(y, t) = \int \mathcal{R}_y(z) n_y(z, t) \, dz \]

(B.10)

which will similarly simplify to \( R_y(x) \) if cell division and growth variables are not involved in their motion.

### B.4. Cell division rates and daughter state distribution

Following Fredrickson et al. (1967) we define the cell division rate as \( \sigma_y(z, t) \), the distribution of states \( z \) of daughters from a parent of state \( x \) as \( p_{x|z}(z, x, t) \). These functions are important phenomenological inputs to the population balance equation in the next section to describe the evolution of the population.

### B.5. Population balance equation

Following Ramkrishna (2000), we may now identify the population balance equation as

\[ \frac{\partial n_x(x, t)}{\partial t} + \nabla_x \cdot Z_x(z) n_x(z, t) + \nabla_x \cdot R_x(z) n_x(z, t) = \mathcal{R}_x(z) n_x(z, t) \]

(B.11)

In Eq. (B.11), we have used subscripts in the gradient operator \( \nabla \) to denote the set of partial derivatives included in the gradient. For example,

\[ \nabla_x = \left( \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n} \right) \]

(B.12)

For the method of derivation of this equation connected with the stochastic terms, the reader is referred to Ramkrishna (2000). Specification of the integration range is avoided in favor of implying constraints in the function \( p_{x|z}(z, x, t) \) that will allow non-zero values only when the states of the parent and the offspring are compatible. Similar to the averaged rates in Eq. (B.5), we may also define the following

\[ \sigma_y(y, t) = \int \mathcal{R}_y(z) n_y(z, t) \, dz \]

(B.10)

which will similarly simplify to \( R_y(x) \) if cell division and growth variables are not involved in their motion.
In addition, we must recognize that the flux of cells must vanish at the boundary of the domain to which they are confined in internal as well as external coordinates. For an open system, it is of course possible to entertain cells entering the spatial domain of interest through some boundary by stipulating the flux there.

### B.7. Some preliminary considerations

Our considerations in this section are designed towards generating reasonable circumstances under which mathematical solutions can be obtained to the problem just defined. If Eq. (B.11) is integrated over the entire domain of all stochastic variables we obtain from using the vanishing of the flux at the boundary the following equation:

\[
\frac{\partial n_\gamma(y,r,t)}{\partial t} = \nabla \cdot (\nabla n_{\gamma}(y,r,t)) + \nabla \cdot (n_{\gamma}(y,r,t) \nabla r(y,r,t)) + \nabla \cdot (n_{\gamma}(y,r,t) \nabla r_{\gamma}(y,r,t)) dy
\]

(B.23)

Eq. (B.23) resembles the population balance equation derived by Fredrickson et al. (1967) with the exception of the strongly detracting spatio-temporal dependence of quantities concerned with cell growth and cell division. However, if we disengage the stochastic variables from cell growth and division, Eqs. (B.6), (B.9) and (B.13) yield the population balance equation

\[
\frac{\partial n_\gamma(y,r,t)}{\partial t} + \nabla \cdot (\nabla n_{\gamma}(y,r,t)) + \nabla \cdot (n_{\gamma}(y,r,t) \nabla r_{\gamma}(y,r,t)) dy
\]

(B.24)

which is the more familiar equation of Fredrickson et al. (1967) in view of the cell growth and division functions being free of spatio-temporal coordinates.

Next we integrate Eq. (B.11) over all \( y \) to obtain

\[
\frac{\partial n(x,r,t)}{\partial t} + \nabla \cdot (\nabla n(x,r,t)) + \nabla \cdot (n(x,r,t) \nabla r(x,r,t))
\]

\[
= \frac{1}{2} \nabla \cdot (B(x,r,t) \nabla n(x,r,t)) - \frac{1}{2} \nabla \cdot (B(x,r,t) \nabla n(x,r,t))
\]

(B.25)

The use of Eq. (B.25) is thwarted by the spatio-temporal dependence of all phenomenological quantities, however. If we invoke independence of the stochastic variables associated with the specific gene regulatory process of interest, we have the following:

\[
\frac{\partial n(x,r,t)}{\partial t} + \nabla \cdot (\nabla n(x,r,t)) + \nabla \cdot (n(x,r,t) \nabla r(x,r,t))
\]

\[
= \frac{1}{2} \nabla \cdot (B(x,r,t) \nabla n(x,r,t)) - \frac{1}{2} \nabla \cdot (B(x,r,t) \nabla n(x,r,t))
\]

(B.26)

We make one further specific assumption about the partitioning of the gene regulatory variables between two daughter cells, viz., that they refer to intracellular concentrations and are equally shared by the daughter cells. This implies that

\[
p_{X|X} (x', x) = \delta(x - x')
\]

(B.27)

Combining Eqs. (B.26) and (B.27), we obtain the population balance equation of main interest to this paper referred to as Eq. (1).
B.8. Computational issues of cell growth

In this section, we discuss the computational issues associated with cell growth and multiplication. If the cell motion is only affected by an external device as described in the main text, Eq. (B.24) became Eq. (B.28) which should be coupled with Eq. (B.21)

\[
\frac{\partial n_Y(y,r,t)}{\partial t} + \nabla_y \cdot j_Y(Y,C_Y)n_Y(y,r,t) = \frac{\partial Y(y,C_Y)}{\partial t}n_Y(y,r,t) \\
+ \nabla_r \cdot \mathbf{R}_Y(y,r,t) = -\sigma_Y(y,C_Y)n_Y(y,r,t) \\
+ 2 \int \sigma_Y(y,C_Y)n_Y(y,r,t) j_Y(y,C_Y) dy' (B.28)
\]

The solution to this problem as it has been addressed by numerous investigators in the literature over the last few decades (Subramanian et al., 1970; Subramanian and Ramkrishna, 1971; Mantzaris, 2006) for varying degrees of complexity with respect to the choice of intracellular variables \(y\). We have thus the means to compute \( \mu(r,t|C_Y), \sigma_Y(r,t) \).

References


