Published in final edited form as:

Chem Eng Sci. 2012 March 5; 70: 188–199. doi:10.1016/j.ces.2011.07.062.

MODELING OF GENE REGULATORY PROCESSES BY POPULATION MEDIATED SIGNALING. NEW APPLICATIONS OF POPULATION BALANCES

Che-Chi Shu¹, Anushree Chatterjee², Wei-Shou Hu², and Doraiswami Ramkrishna¹

¹School of Chemical Engineering, Purdue University, West Lafayette, IN 47907

²Department of Chemical Engineering & Materials Science, University of Minnesota, Minneapolis, MN 55455

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.

Keywords

Biological and biomolecular engineering; Population balance; Formulation; Mathematical modelling; Gene regulatory processes; drug resistance transfer

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

© 2011 Elsevier Ltd. All rights reserved.

Correspondence to: Doraiswami Ramkrishna.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

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 determined by Ito stochastic differential equations has been first presented by Ramkrishna (2000). It will be the objective of this paper to build on this population balance framework towards developing applications to gene regulatory processes in which interaction occurs between cellular environment and cells in which intracellular processes undergo stochastic dynamics. The motivation for this objective is to provide the general framework for population-mediated signaling processes on gene regulatory processes that have numerous applications. The importance of this formulation is demonstrated by including a calculation that dispels the notion established in the literature that when bistability exists in gene expression bimodal distributions of protein levels would result.

1. FORMULATION

The population or number density is fundamental to population balance. It is defined in the space of internal and external coordinates. The internal coordinates refer to intracellular species, some of which are stochastic in view of their small numbers and others that are deterministic and are concerned with growth and metabolism. In modeling the effect of populations on a specific gene regulatory process, it would be desirable to uncouple variables associated with growth and metabolism from those involved in the gene regulatory process of specific interest to us. The reason for this is as follows. A population balance with only the stochastic gene regulatory variables would arise on averaging the number density over all variables connected with growth and metabolism. If the gene regulatory process also depended on the variables associated with growth and metabolism, averaging the number density over the latter variables would produce undesirable spatio-temporal dependence of the kinetic constants for gene regulation. We demonstrate this in Appendix B and show how the assumption that the stochastic variables may be considered independent of cell variables involved in growth and metabolism is useful. Thus we will incorporate in the population balance here a division rate μ to account for the rate of increase in the number density by cell division. Further we define the population density $n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)$, comprising only the stochastic intracellular variables, $\mathbf{x}$, 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 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 $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}})$, where $\mathbf{C}_{\mathbf{X}}$ 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 $\mathbf{B}_{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}})d\mathbf{W}$, where $\mathbf{B}_{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}})$ is a matrix (see Table 2) and $d\mathbf{W}$ is a vector of standard Wiener processes (see for example Gardiner, 1997). Cellular motion with respect to fixed coordinates may be described by $\dot{\mathbf{R}}_{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}})$. The population balance equation may then be written as

$$\frac{\partial n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}}) n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}}) n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{X}}) \mathbf{B}_{\mathbf{X}}^T(\mathbf{x}|\mathbf{C}_{\mathbf{X}}) n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t) + \mu(\mathbf{r}, t|\mathbf{C}_{\mathbf{Y}}) n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t) \quad (1)$$

where $\mu(\mathbf{r}, t|\mathbf{C}_{\mathbf{Y}})$ is the Specific growth rate. $\mathbf{C}_{\mathbf{Y}}$ are the extracellular variables relating to cell growth and metabolism and $\mathbf{r}$ are 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

$$\frac{\partial \mathbf{C}_{\mathbf{X}}}{\partial t} + \nabla_{\mathbf{r}} \cdot \mathbf{N}_{\mathbf{X}} = \int \dot{\mathbf{C}}_{\mathbf{X}} n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t) d\mathbf{x} \quad (2)$$

where $\mathbf{N}_{\mathbf{X}}$ are the total flux of $\mathbf{C}_{\mathbf{X}}$ due to convection and diffusion and the vector $\dot{\mathbf{C}}_{\mathbf{X}}$ relates change of extracellular variables to intracellular reactions associated with gene regulation. It is to be noted that Eq. (1) features the variable $\mathbf{C}_{\mathbf{Y}}$ which must be regarded as a *known* time-dependent variable, obtained by simultaneous solution of the population balance equation Eq. (B24) for cell growth and multiplication coupled with Eq. (B21) 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

$$n_{\mathbf{X}}(\mathbf{x}, \mathbf{r}, 0) = N_o(\mathbf{r}) f_{\mathbf{X},o}(\mathbf{x}), \quad \mathbf{C}_{\mathbf{X}}(\mathbf{r}, 0) = \mathbf{C}_{\mathbf{X},o}(\mathbf{r}) \quad (3)$$

where $N_o(\mathbf{r})$ is the initial spatial distribution of the total number density of cells, $f_{\mathbf{X},o}(\mathbf{x})$ the distribution of stochastic variables among the cell population, and $\mathbf{C}_{\mathbf{X},o}(\mathbf{r})$ 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 $\dot{\mathbf{R}}$ depends neither on $\mathbf{x}$ nor $\mathbf{C}_{\mathbf{X}}$ but does only on $\mathbf{r}$ and t . Thus it is possible to define a path of cellular motion by a function $\mathbf{r}' = \mathbf{R}(t|\mathbf{r}_o)$, $0 < t' < t$ satisfying $\mathbf{r}_o = \mathbf{R}(0|\mathbf{r}_o)$ and $\mathbf{r} = \mathbf{R}(t|\mathbf{r}_o)$ which implies that the motion begins at $\mathbf{r}_o$ at time $t' = 0$ and followed until reaching $\mathbf{r}$ 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 $(\mathbf{r}, t)$ and represent this inverse relationship by $\mathbf{r}_o = \mathbf{R}_o(\mathbf{r}, t)$, which defines a time-dependent transformation between coordinates $\mathbf{r}$ and $\mathbf{r}_o$ representing initial locations. The Jacobian of

the forward transformation, $J(\mathbf{r}, t|\mathbf{r}_o) \equiv \left| \frac{\partial \mathbf{R}}{\partial \mathbf{r}_o} \right|$ can be readily shown to satisfy the relations (Aris, 1962).

$$\frac{DJ}{Dt} \equiv \frac{\partial J}{\partial t} + \dot{\mathbf{R}} \cdot \nabla_{\mathbf{r}} J = J \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}$$

The population balance equation (1) becomes on replacing $\dot{\mathbf{R}}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}})$ by $\dot{\mathbf{R}}(\mathbf{r}, t)$

$$\frac{DJ_{n_{\mathbf{x}}}(\mathbf{x}, \mathbf{r}, t)}{Dt} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) J_{n_{\mathbf{x}}}(\mathbf{x}, \mathbf{r}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) J_{n_{\mathbf{x}}}(\mathbf{x}, \mathbf{r}, t) + \mu(\mathbf{r}, t|\mathbf{C}_{\mathbf{y}}) J_{n_{\mathbf{x}}}(\mathbf{x}, \mathbf{r}, t) \quad (4)$$

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

$$J(\mathbf{r}_o, 0|\mathbf{r}_o) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, 0) = N_o(\mathbf{r}_o) f_{\mathbf{x},o}(\mathbf{x})$$

which follows from $J(\mathbf{r}_o, 0|\mathbf{r}_o) = 1$ and (3). We may further define the function

$$P(\mathbf{x}, \mathbf{r}, t) = \frac{1}{N_o(\mathbf{r})} J(\mathbf{r}, t|\mathbf{r}_o) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) \exp \left[- \int_0^t J(\mathbf{r}', t'|\mathbf{r}_o) \mu(\mathbf{r}', t'|\mathbf{C}_{\mathbf{y}}(t')) dt' \right] \quad (5)$$

where $\mathbf{r}_o = \mathbf{R}_o(\mathbf{r}, t)$ is obtained by backtracking along the cell path from position $\mathbf{r}$ at time t . The function $P(\mathbf{x}, \mathbf{r}, t)$ satisfies the partial differential equation

$$\frac{DP(\mathbf{x}, \mathbf{r}, t)}{Dt} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, \mathbf{r}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, \mathbf{r}, t) \quad (6)$$

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 $P(\mathbf{x}, \mathbf{r}, t)$ is given by

$$P(\mathbf{x}, \mathbf{r}, 0) = \frac{1}{N_o(\mathbf{r})} n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, 0) = f_{\mathbf{x},o}(\mathbf{x})$$

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

$$d\mathbf{x} = \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) dt + \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) d\mathbf{W} \quad (7)$$

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).

1.1 Sessile Cells. Growth in Biofilms

The case of sessile cells is encountered in biofilm growth situations. Thus we set $\dot{\mathbf{R}} \equiv \mathbf{0}$ for which $J \equiv 1$ at all locations and times and Eq.(6) becomes

$$\frac{\partial P(\mathbf{x}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, \mathbf{r}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, \mathbf{r}, t) \quad (8)$$

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).

1.2 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,

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C}_{\mathbf{x}}) P(\mathbf{x}, t) \quad (9)$$

which again translates to the Ito equation (7) applicable to the entirely different situation considered earlier. However, the function $\mathbf{C}_{\mathbf{x}}$ must satisfy the differential equation applicable to this well-mixed case, viz.,

$$\frac{d\mathbf{C}_{\mathbf{x}}}{dt} = N(t) \int \dot{\mathbf{C}}_{\mathbf{x}} P(\mathbf{x}, t) d\mathbf{x} = N(t) E[\dot{\mathbf{C}}_{\mathbf{x}}] \quad (10)$$

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 Eq. (5) adapted to the well-mixed case given by.

$$P(\mathbf{x}, t) = \frac{1}{N_o} n_{\mathbf{x}}(\mathbf{x}, t) \exp \left[- \int_0^t \mu(t') dt' \right] \quad (11)$$

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

2. 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 $E[\dot{\mathbf{C}}_{\mathbf{x}}]$. 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.

3. APPLICATION TO TRANSFER OF DRUG RESISTANCE

3.1 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 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 Figure 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_{pre} ; Q_{pre} gives rise to two kinds of m-RNAs known as Q_L mRNA and Q_s mRNA. The Q_L mRNA stimulates downstream genes to produce PrgB protein which, in the real system, triggers conjugation by allowing a donor cell attaching to a recipient cell (Trotter and Dunny, 1990; Waters *et al.*, 2004). In addition, Q_L mRNA also produces a molecule referred to as iCF10-precursor as it eventually matures into iCF10, an inhibitor to cCF10. Q_s mRNA also produces this iCF10-precursor. Thereafter the iCF10-precursor is released into the environment in the process of which it becomes iCF10 preventing the cCF10 from altering the DNA configuration. Another gene referred to as prgX gene encodes an RNA called Anti-Q which binds to Q_{pre} transforming it to Q_s instead of Q_L , thus reducing the production of PrgB protein. Consequently, high concentration of Q_{pre} is required to overcome Anti-Q (Bae *et al.*, 2004; Johnson *et al.*, 2010) and make a cell produce Q_L 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_{pre} is overwhelmed by Anti-Q and almost all Q_{pre} becomes Q_s resulting in nearly no prgB expression. In a nut shell, a subtle balance between concentration of cCF10 and iCF10 control the occurrence of conjugation. Through gene regulation, donor cells change the rate of producing iCF10 nonlinearly responding to add-in cCF10. Conjugative genes, including prgB, are expressed at high ratios of intracellular cCF10 to iCF10 and are not expressed at low 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.

3.2 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, $\mathbf{x}$ 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 $\mathbf{C}_X$ has two components, one the cCF10 (C_1) concentration in the environment and the other that of the iCF10 (C_2). The vector $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_X)$ 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), (ii) application of the Chemical Master Equation to obtain the stochastic version identifying the matrix $\mathbf{B}_X(\mathbf{x}|\mathbf{C}_X)$. Following Haseltine EL, Rawlings JB (2002) and Rao CV, Arkin AP (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.

3.3 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 cell can be described by (12)

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_X)$$

$$\frac{dC_2}{dt} = N\nu_d [K_{1,6}(x_1+x_2) - K_{2,6}C_2] - \left(K_{4,5} + \frac{d\ln V}{dt}\right)C_2 \quad (12)$$

which is a modified version of (10) to account for the additional assumption that the biomass concentration (or the number density N) is maintained constant in the culture by allowing the total volume to expand with growth. In other words, $d\ln V(t)/dt = \mu_d$, the Specific growth rate of a donor cell. The foregoing assumption is essential to provide for a steady

state to be attained. A switch like behavior, bistability of PrgB, is featured in (12) with parameter values shown in Table 3. 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 the cCF10 within a suitable range. A cell can stay at either “on” state or “off”. Unlike ramp behavior, there is no middle value for a cell with bistability because it is unstable. The bistability of pCF10 conjugation system is shown in Figure 2.

3.4 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 Table 4 and Table 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 C_X of population balance model is deterministic. The simulation of pCF10 conjugation system from Fokker-Planck equation of single cell stochastic model is shown as Figure 3a. Notice that the bimodal distribution is consistent with steady state values shown in Figure 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.

3.5 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 cell stochastic model by inclusion of accounting for the reciprocal interaction between the cells and their environment. The simulation of this model is done for three cases.

Case 1: We assume that the volume of the culture is increasing at a rate such that the total number density is independent of time. This assumption is made to allow an eventual steady state in the system. The computational procedure in Section 1.2 is used by assuming that the number density is constant. Further, it is assumed that cCF10 is held constant in the cells' environment by a suitable external add-in. Such situations have been dealt with experimentally (Leonard et al., 1996; Kozłowicz et al., 2006; Fixen et al., 2007; Chandler and Dunny 2008). The Ito version (7) of Eq. (9) is solved sample path wise simultaneously with the equations for the extracellular variable identified for this case in Table 2 which is reproduced below after substituting for the vector $\hat{C}_X$, recognizing that C_1 remains constant.

$$\frac{dC_2}{dt} = v_d N \int [K_{1,6}(x_1 + x_2) - K_{2,6}C_2] P(\mathbf{x}, t) d\mathbf{x} - \left(K_{4,5} + \frac{d \ln V}{dt}\right) C_2 \quad (13)$$

Here again we have $d \ln V / dt = \mu_d$ as in Eq. (12) as the rate of expansion for the extracellular environment for the single cell and the entire well-stirred culture (for the population) is the same. Eq. (13) maybe rewritten as

$$\frac{dC_2}{dt} = v_d N [K_{1,6}E(x_1 + x_2) - K_{2,6}C_2] - \left(K_{4,5} + \frac{d \ln V}{dt}\right) C_2 \quad (14)$$

The simulation is shown in Figure 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 Figure 3a (cell in isolated circumstance), cells in Figure 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 Cook (2011) show unimodal distributions for cells in planktonic growth and bimodal distribution for biofilm growth. The simulation of this process is possible with the population balance model for sessile cells in Section 3.1. Computational effort required for this simulation is in progress and the results will be reported in a future publication as part of an extended investigation relating population balance models of drug resistance transfer to experiments. 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 also consider a growing population of recipient cells which contribute cCF10 with no external add-in of this signaling molecule. The required equations are shown in Table 2. The extracellular variable equations for this case becomes

$$\begin{aligned}\frac{dC_1}{dt} &= N_{r_0} e^{\mu_r t} K_{1,8} - v_d C_1 K_{2,8} N(t) - K_{4,8} C_1 \\ \frac{dC_2}{dt} &= v_d N(t) [K_{1,6} E(x_1 + x_2) - K_{2,6} C_2] - K_{4,5} C_2\end{aligned}\quad (15)$$

where N_{r_0} is the initial number density of recipients and μ_r is its Specific growth rate.

The solution of stochastic Eq. (7) simultaneously with (15) proceeds as discussed in Section 2. The results, presented in Figure 4, include both the protein distribution and the number density in terms of protein levels. That there can be no strict steady state is reflected in the drifting protein distribution to higher protein levels with time shown in Figure 4a. This is because of recipient cells continually contributing cCF10 for protein expression. Figure 4b shows the number density reflects its obvious increase due to replication.

Case 3: In this case, we alter Case 2 by including the conjugation process with transformation of recipient cells to donor cells assuming that such a transformation occurs immediately following conjugation. The model further assumes that the protein 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.

The governing equations are readily written as follows. The rate of conjugation is assumed to depend on the protein level in the conjugating donor cell.

The Fokker Planck Eq. (9) and the transformation (11) continue to be applicable with $\mu(t)$ to be replaced by

$$\mu(t) = \mu_d + k_{con} N_r(t) \int x_5 n_x(\mathbf{x}, t) d\mathbf{x} = \mu_d + k_{con} N_r(t) N(t) E[x_5] \quad (16)$$

where k_{con} is the rate constant for conjugation of donor cells with recipient cells including the proportionality constant for PrgB. The differential equation for the recipient cells is given by

$$\frac{dN_r}{dt} = \mu_r N_r - k_{con} N_r N E[x_5] \quad (17)$$

The environmental variable equations are given by

$$\frac{dC_1}{dt} = N_r K_{1,8} - v_d C_1 K_{2,8} N - K_{4,8} C_1 \quad (18)$$

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

$$\frac{dN}{dt} = \mu_d N + k_{con} N_r N E[x_5] \quad (19)$$

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 Figure 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.

4. 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.

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. (A1), (A2) and (A3).

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_{all}(\mathbf{x}|\mathbf{C}) P(\mathbf{x}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{x_{all}}(\mathbf{x}|\mathbf{C}) \mathbf{B}_{x_{all}}^T(\mathbf{x}|\mathbf{C}) P(\mathbf{x}, t) \quad (A1)$$

$$\dot{\mathbf{X}}_{all}(\mathbf{x}|\mathbf{C}) = \begin{bmatrix} k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3}\right) - (K_{4,1} + \mu_d)x_1 \\ k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{1}{1+K_{3,5}x_3}\right) - (K_{4,2} + \mu_d)x_2 \\ k[K_{1,3}\alpha + K_{1,4}(1-\alpha)] - k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3}\right) - (K_{4,3} + \mu_d)x_3 \\ K_{2,6}C_2 - (K_{4,6} + \mu_d)x_4 \\ K_{1,5}x_2 - (K_{4,9} + \mu_d)x_5 \\ K_{2,8}C_1 - (K_{4,8} + \mu_d)x_6 \end{bmatrix} \quad (A2)$$

$$\text{where : } \alpha = \frac{x_4^4}{x_4^4 + K_{3,8}x_6^4}$$

$$\mathbf{B}_{\mathbf{x}_{all}}(\mathbf{x}|\mathbf{C})\mathbf{B}_{\mathbf{x}_{all}}^T(\mathbf{x}|\mathbf{C}) \equiv \mathbf{B}_{all} = \begin{bmatrix} B_{11} & 0 & B_{13} & 0 & 0 & 0 \\ 0 & B_{22} & 0 & 0 & 0 & 0 \\ B_{31} & 0 & B_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{66} \end{bmatrix} \quad (\text{A3})$$

where:

$$\begin{aligned} B_{11} &= \frac{1}{v_d} \left\{ k[K_{1,1}\alpha + K_{1,2}(1-\alpha)] \left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3} \right) + K_{4,1}x_1 \right\} \\ B_{22} &= \frac{1}{v_d} \left\{ k[K_{1,1}\alpha + K_{1,2}(1-\alpha)] \left(\frac{1}{1+K_{3,5}x_3} \right) + K_{4,2}x_2 \right\} \\ B_{33} &= \frac{1}{v_d} \left\{ k[K_{1,3}\alpha + K_{1,4}(1-\alpha)] + k[K_{1,1}\alpha + K_{1,2}(1-\alpha)] \left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3} \right) + K_{4,3}x_3 \right\} \\ B_{44} &= \frac{1}{v_d} [K_{2,6}C_2 + K_{4,6}x_4] \\ B_{55} &= \frac{1}{v_d} [K_{1,5}x_2 + K_{4,9}x_5] \\ B_{66} &= \frac{1}{v_d} [K_{2,8}C_1 + K_{4,8}x_6] \\ B_{13} &= B_{31} = \frac{1}{v_d} \left\{ -k[K_{1,1}\alpha + K_{1,2}(1-\alpha)] \left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3} \right) \right\} \end{aligned}$$

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 $\mathbf{x}$ into $\mathbf{x}_s$ and $\mathbf{x}_f$, the slow and the fast reaction species. The probability can be described by $P(\mathbf{x}, t) = P(\mathbf{x}_s, \mathbf{x}_f, t) = P(\mathbf{x}_f | \mathbf{x}_s, t)P(\mathbf{x}_s, t)$ with $dP(\mathbf{x}_f | \mathbf{x}_s, t)/dt \approx 0$, refer to Rao CV and Arkin AP (2003), so the noise term of $\mathbf{x}_f$ is negligible and the approximate master equation solely in terms of $\mathbf{x}_s$. The Fokker-Planck equation of $\mathbf{x}_s$ is Eqs. (A4), (A5) and (A6) where $\mathbf{x}_s^T = [x_4 \ x_5 \ x_6]$.

$$\frac{\partial P(\mathbf{x}_s, t)}{\partial t} + \nabla_{\mathbf{x}_s} \cdot \dot{\mathbf{X}}_s(\mathbf{x}_s|\mathbf{C})P(\mathbf{x}, t) = \frac{1}{2} \nabla_{\mathbf{x}_s} \nabla_{\mathbf{x}_s} : \mathbf{B}_{\mathbf{x}_s}(\mathbf{x}_s|\mathbf{C})\mathbf{B}_{\mathbf{x}_s}^T(\mathbf{x}_s|\mathbf{C})P(\mathbf{x}_s, t) \quad (\text{A4})$$

$$\dot{\mathbf{X}}_s(\mathbf{x}_s|\mathbf{C}) = \begin{bmatrix} K_{2,6}C_2 - (K_{4,6} + \mu_d)x_4 \\ K_{1,5}x_3 - (K_{4,9} + \mu_d)x_5 \\ K_{2,8}C_1 - (K_{4,8} + \mu_d)x_6 \end{bmatrix} \quad (\text{A5})$$

$$\mathbf{B}_{\mathbf{x}_s}(\mathbf{x}_s|\mathbf{C})\mathbf{B}_{\mathbf{x}_s}^T(\mathbf{x}_s|\mathbf{C}) = \mathbf{B}_s = \begin{bmatrix} B_{44} & 0 & 0 \\ 0 & B_{55} & 0 \\ 0 & 0 & B_{66} \end{bmatrix} \quad (\text{A6})$$

For convenience, we further put all variables together as Eq. (A7) which is shown in Table 2.

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C})P(\mathbf{x}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C})\mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C})P(\mathbf{x}, t) \quad (\text{A7})$$

First, we combined the drift term of fast reaction species, x_1 , x_2 and x_3 with $\dot{\mathbf{X}}_s(\mathbf{x}_s | \mathbf{C})$ to make $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{C})$. And then, to formulate $\mathbf{B}_{\mathbf{x}}(\mathbf{x}|\mathbf{C})\mathbf{B}_{\mathbf{x}}^T(\mathbf{x}|\mathbf{C}) \equiv \mathbf{B}$, we put zero for the diffusion terms of x_1 , x_2 and x_3 as there are no random.

APPENDIX B

B.1 Number Density

We consider a cell population with intracellular state variables described by a vector $\mathbf{z} \equiv (\mathbf{x}, \mathbf{y})$, which distinguishes between n stochastic variables (processes) $\mathbf{x}$ and m deterministic variables $\mathbf{y}$. The stochastic variables are associated with the gene regulatory processes. Further we let the spatial coordinates of the cell be represented by $\mathbf{r}$. The number density representing the number of cells per unit volume of both internal and spatial coordinates is denoted $n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t)$. If this number density is integrated over the entire domain of all the m intracellular variables in $\mathbf{y}$, we obtain the number density in terms of only the stochastic variables $\mathbf{x}$. Conversely, integration of $n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t)$ over all $\mathbf{x}$ variables will yield the number density in terms of only $\mathbf{y}$. In other words, we may write

$$\int n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t) d\mathbf{y} = n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t), \quad \int n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t) d\mathbf{x} = n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) \quad (\text{B1})$$

The integration range in Eq. (B1) is understood to be over all positive values of cellular variables. The cellular environment is characterized by a vector $\mathbf{C}$ of concentrations of nutrients, signaling molecules, and products of cellular activity, which may depend on spatial and temporal coordinates. The vector $\mathbf{C}$ is regarded as deterministic although it may contain stochastic components transported out from within the cells because it is assumed that there are a large number of cells of each state. The total number density $N(\mathbf{r}, t)$ at location $\mathbf{r}$ and time t is given by

$$N(\mathbf{r}, t) = \int n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t) d\mathbf{z} = \int n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) d\mathbf{x} = \int n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) d\mathbf{y} \quad (\text{B2})$$

It is also useful to recognize probability density functions $f_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t)$, $f_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)$, $f_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t)$ so that

$$\frac{n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t)}{f_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t)} = \frac{n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)}{f_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)} = \frac{n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t)}{f_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t)} = N(\mathbf{r}, t) \quad (\text{B3})$$

If it is assumed that the stochastic variables $\mathbf{x}$ and the deterministic variables $\mathbf{y}$ are *statistically independent* of each other, we have

$$f_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t) = f_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) f_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t), \quad n_{\mathbf{z}}(\mathbf{z}, \mathbf{r}, t) = n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) f_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) = n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) f_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) \quad (\text{B4})$$

We describe the foregoing situation as separation between the processes of cell growth and metabolism from the gene regulatory process of interest and will refer to it as independence between the gene regulatory process and cell division process.

B.2 Growth, Intracellular Kinetics

The rate of change of the intracellular state vector $\mathbf{z}$ is denoted by $(\mathbf{z}|\mathbf{C})$ which displays dependence on the internal state $\mathbf{z}$, the local cell environment $\mathbf{C}$, and freedom from explicit dependence on cell location although spatial dependence is inherited through that of the environment. Note that $(\mathbf{z}|\mathbf{C})$ can be obtained by the product of stoichiometric matrix $\mathbf{A}_{\mathbf{z}}$ and intracellular reaction vector $\boldsymbol{\gamma}$, $(\mathbf{z}|\mathbf{C}) = \mathbf{A}_{\mathbf{z}}\boldsymbol{\gamma}$. We may further note that $(\mathbf{z}|\mathbf{C})$ comprises $\dot{\mathbf{X}}_{\mathbf{z}}(\mathbf{z}|\mathbf{C})$ referring to the mean rate of change of stochastic intracellular variables $\mathbf{x}$ and $\dot{\mathbf{Y}}_{\mathbf{z}}(\mathbf{z}|\mathbf{C})$ referring to the deterministic intracellular variables $\mathbf{y}$. The change in the stochastic variables $\mathbf{x}$ during a small interval dt is expressed as $\dot{\mathbf{X}}_{\mathbf{z}}(\mathbf{z}|\mathbf{C}) dt + \mathbf{B}_{\mathbf{z}}(\mathbf{z}|\mathbf{C}) d\mathbf{W}$, the first term denoting the mean change and the second a stochastic change represented by the standard Wiener process increment $d\mathbf{W}$. The matrix $\mathbf{B}_{\mathbf{z}}(\mathbf{z}|\mathbf{C})$ is $(n \times n)$ so that the matrix

product $\mathbf{B}_Z(\mathbf{z}|\mathbf{C})\mathbf{B}_Z^T(\mathbf{z}|\mathbf{C})$ is an $(n \times n)$, symmetric matrix which will appear in the treatment to follow. The reader, unfamiliar with the standard Wiener process is referred to Gardiner (1997). We simply note here that $d\mathbf{W}$ consists of n (independent) standard Wiener processes characterizing the stochastic variables in the gene regulatory process of interest. Further, the standard Wiener process has the property that its expectations $EdW = 0$, $EdW^2 = dt$. Note the dependence of all the above kinetic expressions on the entire intracellular vector $\mathbf{z}$ and the local environmental vector $\mathbf{C}$. It is further useful to define the following $\mathbf{y}$ - or $\mathbf{x}$ - averaged kinetics

$$\dot{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t|\mathbf{C}) \equiv \frac{\int \dot{\mathbf{X}}_Z(\mathbf{z}|\mathbf{C})n_Z(\mathbf{z}, \mathbf{r}, t)d\mathbf{y}}{n_X(\mathbf{x}, \mathbf{r}, t)}, \quad \dot{\mathbf{Y}}(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) \equiv \frac{\int \dot{\mathbf{Y}}_Z(\mathbf{z}|\mathbf{C})n_Z(\mathbf{z}, \mathbf{r}, t)d\mathbf{x}}{n_Y(\mathbf{y}, \mathbf{r}, t)} \quad (\text{B5})$$

If $\dot{\mathbf{X}}_Z(\mathbf{z}|\mathbf{C})$ does not depend on $\mathbf{y}$ and $\dot{\mathbf{Y}}_Z(\mathbf{z}|\mathbf{C})$ does not depend on $\mathbf{x}$, then, in view of Eqs. (B1) and (B5), we have

$$\dot{\mathbf{X}}_Z(\mathbf{z}|\mathbf{C}) \Rightarrow \dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}), \quad \dot{\mathbf{Y}}_Z(\mathbf{z}|\mathbf{C}) \Rightarrow \dot{\mathbf{Y}}(\mathbf{y}|\mathbf{C}) \quad (\text{B6})$$

which shows *freedom from the spatio-temporal dependence* of these averaged rates displayed in Eq. (B5). The arrow is used to signify the replacement of the term appearing to the left of it by that which appears to its right. We may also define the following $\mathbf{y}$ -averaged matrix $\mathbf{B}_X(\mathbf{x}, \mathbf{r}, t|\mathbf{C})$ as below.

$$\mathbf{B}_X(\mathbf{x}, \mathbf{r}, t|\mathbf{C})\mathbf{B}_X^T(\mathbf{x}, \mathbf{r}, t|\mathbf{C}) = \frac{\int \mathbf{B}_Z(\mathbf{z}|\mathbf{C})\mathbf{B}_Z^T(\mathbf{z}|\mathbf{C})n_Z(\mathbf{z}, \mathbf{r}, t)d\mathbf{z}}{n_X(\mathbf{x}, \mathbf{r}, t)} \quad (\text{B7})$$

If, however, the $\mathbf{y}$ - variables are not involved in the stochastic gene regulatory processes, $\mathbf{B}_Z$ will only involve the stochastic variables $\mathbf{x}$, in which case Eq. (B7) would imply that the matrix $\mathbf{B}_X$ will be free of spatio-temporal dependence so that the notation $\mathbf{B}_X(\mathbf{x}|\mathbf{C})$ may more appropriately replace $\mathbf{B}_Z(\mathbf{z}|\mathbf{C})$.

It is desirable to recognize among the extracellular concentration variables, those that are associated with the stochastic gene regulatory intracellular variables $\mathbf{x}$, and denote them by $\mathbf{C}_X$. Similarly we let $\mathbf{C}_Y$ denote extracellular concentration variables associated with the deterministic intracellular variables $\mathbf{y}$. Thus the term $\mathbf{B}_X(\mathbf{x}|\mathbf{C})$ may be more appropriately denoted as $\mathbf{B}_X(\mathbf{x}|\mathbf{C}_X)$ if extracellular variables $\mathbf{C}_Y$ do not affect stochastic behavior of gene regulatory variables. Thus in Eq. (B6) we may replace $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{C})$ by $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{C}_X)$ and $\dot{\mathbf{Y}}(\mathbf{y}|\mathbf{C})$ by $\dot{\mathbf{Y}}(\mathbf{y}|\mathbf{C}_Y)$.

B.3 Cellular Displacement

Cells may undergo spatial displacement depending either on external mixing devices or because of other mechanisms such as positive or negative chemotaxis. In biofilm environments, cells may be sessile. Since it is not our goal to address in detail the mechanism of such displacements, we will merely denote the rate of spatial cell displacement by $\dot{\mathbf{R}}_Z(\mathbf{z}|\mathbf{C})$ without reference to possible dependence on concentration gradients. We may also derive a cell displacement rate by averaging over all stochastic variables by

$$\dot{\mathbf{R}}_Y(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) \equiv \frac{\int \dot{\mathbf{R}}_Z(\mathbf{z}|\mathbf{C})n_Z(\mathbf{z}, \mathbf{r}, t)d\mathbf{x}}{n_Y(\mathbf{y}, \mathbf{r}, t)} \quad (\text{B8})$$

When the displacement rate is independent of stochastic variables we have the result

$$\dot{\mathbf{R}}_z(\mathbf{z}|\mathbf{C}) \Rightarrow \dot{\mathbf{R}}_y(\mathbf{y}|\mathbf{C}) \quad (\text{B9})$$

which is explicitly independent of time and location. Similarly we can define a displacement rate averaged over all of the variables in $\mathbf{y}$

$$\dot{\mathbf{R}}_x(\mathbf{x}, \mathbf{r}, t|\mathbf{C}) \equiv \frac{\int \dot{\mathbf{R}}_z(\mathbf{z}|\mathbf{C}) n_z(\mathbf{z}, \mathbf{r}, t) d\mathbf{y}}{n_x(\mathbf{x}, \mathbf{r}, t)} \quad (\text{B10})$$

which will similarly simplify to $\dot{\mathbf{R}}_x(\mathbf{x}|\mathbf{C})$ 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 $\sigma_z(\mathbf{z}|\mathbf{C})$, the distribution of states $\mathbf{z}$ of daughters from a parent of state $\mathbf{z}'$ as $p_{\mathbf{z}|\mathbf{z}'}(\mathbf{z}|\mathbf{z}', \mathbf{C})$. 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

$$\frac{\partial n_z(\mathbf{z}, \mathbf{r}, t)}{\partial t} + \nabla_z \cdot \dot{\mathbf{Z}}(\mathbf{z}|\mathbf{C}) n_z(\mathbf{z}, \mathbf{r}, t) + \nabla_r \cdot \dot{\mathbf{R}}_z(\mathbf{z}|\mathbf{C}) n_z(\mathbf{z}, \mathbf{r}, t) = \frac{1}{2} \nabla_x \nabla_x \cdot \mathbf{B}_z(\mathbf{z}|\mathbf{C}) \mathbf{B}_z(\mathbf{z}|\mathbf{C})^T n_z(\mathbf{z}, \mathbf{r}, t) - \sigma_z(\mathbf{z}|\mathbf{C}) n_z(\mathbf{z}, \mathbf{r}, t) + 2 \int \sigma_z(\mathbf{z}'|\mathbf{C}) p_{\mathbf{z}|\mathbf{z}'}(\mathbf{z}|\mathbf{z}', \mathbf{C}) n_z(\mathbf{z}', \mathbf{r}, t) d\mathbf{z}' \quad (\text{B11})$$

In Eq. (B11), we have used subscripts in the gradient operator ∇ to denote the set of partial

derivatives included in the gradient. For example, $\nabla_x \equiv \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right]$. 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 $p_{\mathbf{z}|\mathbf{z}'}(\mathbf{z}|\mathbf{z}', \mathbf{C})$ that will allow non-zero values only when the states of the parent and the off-springs are compatible. Similar to the averaged rates in Eq. (B5), we may also define the following

$$\sigma_y(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) \equiv \frac{\int \sigma_z(\mathbf{z}|\mathbf{C}) n_z(\mathbf{z}, \mathbf{r}, t) d\mathbf{x}}{n_y(\mathbf{y}, \mathbf{r}, t)} \quad (\text{B12})$$

$$p_{y|y'}(\mathbf{y}, \mathbf{r}, t|\mathbf{y}', \mathbf{C}) \equiv \frac{\int d\mathbf{x} \int \sigma_z(\mathbf{z}'|\mathbf{C}) p_{\mathbf{z}|\mathbf{z}'}(\mathbf{z}|\mathbf{z}', \mathbf{C}) n_z(\mathbf{z}', \mathbf{r}, t) d\mathbf{x}'}{\sigma_y(\mathbf{y}', \mathbf{r}, t|\mathbf{C}) n_y(\mathbf{y}', \mathbf{r}, t)}$$

Again we note, as before, that if *cell division and growth are independent of all stochastic intracellular variables* involved in the gene regulatory process of interest here, we have

$$\sigma_y(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) \Rightarrow \sigma_y(\mathbf{y}|\mathbf{C}_y), p_{y|y'}(\mathbf{y}, \mathbf{r}, t|\mathbf{y}', \mathbf{C}) \Rightarrow p_{y|y'}(\mathbf{y}|\mathbf{y}', \mathbf{C}) = \int p_{\mathbf{z}|\mathbf{z}'}(\mathbf{z}|\mathbf{z}', \mathbf{C}) d\mathbf{x} \quad (\text{B13})$$

Eq.(B13) shows that the averaged cell division rate and the daughter state distribution are free of spatio-temporal dependence when the stochastic variables of gene regulation do not affect the process of cell multiplication. It must be apparent that similar division rate and probability of daughter distributions averaged over $\mathbf{y}$ – coordinates may be defined by

$$\sigma_X(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) \equiv \frac{\int \sigma_Z(\mathbf{z} | \mathbf{C}) n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{y}}{n_X(\mathbf{x}, \mathbf{r}, t)}$$

$$p_{X|X'}(\mathbf{x}, \mathbf{r}, t | \mathbf{x}', \mathbf{C}) \equiv \frac{\int d\mathbf{y} \int \sigma_Z(\mathbf{z}' | \mathbf{C}) p_{ZZ'}(\mathbf{z} | \mathbf{z}', \mathbf{C}) n_Z(\mathbf{z}', \mathbf{r}, t) d\mathbf{y}'}{\sigma_X(\mathbf{x}', \mathbf{r}, t | \mathbf{C}) n_X(\mathbf{x}', \mathbf{r}, t)} \quad (\text{B14})$$

If the cell division and growth processes do not involve the gene regulatory process variables, the right hand side of the first of the above expressions becomes on dropping the dependence of σ_Z on $\mathbf{x}$ in the first of Eq. (B14)

$$\frac{\int \sigma_Y(\mathbf{y} | \mathbf{C}_Y) n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{y}}{n_X(\mathbf{x}, \mathbf{r}, t)} = \int \sigma_Y(\mathbf{y} | \mathbf{C}_Y) f_Y(\mathbf{y}, \mathbf{r}, t) d\mathbf{y} \equiv \mu(\mathbf{r}, t | \mathbf{C}_Y) \quad (\text{B15})$$

where we have used Eq. (B4), and $\mu(\mathbf{r}, t | \mathbf{C}_Y)$ is the cell division rate regardless of the cell state at location $\mathbf{r}$ at time t . Thus we have in this case $\sigma_X(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) \Rightarrow \mu(\mathbf{r}, t | \mathbf{C}_Y)$. If we further assume that the partitioning probability for $\mathbf{x}$ – variables and $\mathbf{y}$ – occur independently of each other, then we may represent this as

$$p_{ZZ'}(\mathbf{z} | \mathbf{z}', \mathbf{C}) = p_{X|X'}(\mathbf{x} | \mathbf{x}', \mathbf{C}_X) p_{Y|Y'}(\mathbf{y} | \mathbf{y}', \mathbf{C}_Y) \quad (\text{B16})$$

The right hand side of the second of the expressions in (B14) becomes

$$\frac{\int d\mathbf{y} \int \sigma_Z(\mathbf{z}' | \mathbf{C}) p_{ZZ'}(\mathbf{z} | \mathbf{z}', \mathbf{C}) n_Z(\mathbf{z}', \mathbf{r}, t) d\mathbf{y}'}{\sigma_X(\mathbf{x}', \mathbf{r}, t | \mathbf{C}) n_X(\mathbf{x}', \mathbf{r}, t)} = \frac{p_{X|X'}(\mathbf{x} | \mathbf{x}', \mathbf{C}_X) \int \sigma_Z(\mathbf{y}' | \mathbf{C}_Y) f_Y(\mathbf{y}', \mathbf{r}, t) d\mathbf{y}'}{\mu(\mathbf{r}, t | \mathbf{C}_Y)} = \frac{p_{X|X'}(\mathbf{x} | \mathbf{x}', \mathbf{C}_X) n_X(\mathbf{x}', \mathbf{r}, t)}{\mu(\mathbf{r}, t | \mathbf{C}_Y)} \quad (\text{B17})$$

where we have made use of the second of the relationships in Eq. (B4) and Eq. (B15) and the fact that $\int p_{Y|Y'}(\mathbf{y} | \mathbf{y}', \mathbf{C}) d\mathbf{y} = 1$. Also to be noted is the insertion of $\mathbf{C}_X$ and $\mathbf{C}_Y$ in place of $\mathbf{C}$ to account for the assumption of growth and cell division being separated from the specific gene regulatory process.

B.6 Balance Equation for Cellular Environment

Eq. (B11) must be coupled to an equation for mass conservation of environmental components that are exchanged between the biotic and abiotic phases. This is readily identified as

$$\frac{\partial \mathbf{C}}{\partial t} + \nabla_{\mathbf{r}} \cdot \mathbf{N}_Z = \int \dot{\mathbf{C}}(\mathbf{z} | \mathbf{C}) n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{z} \quad (\text{B18})$$

where $\dot{\mathbf{C}}(\mathbf{z} | \mathbf{C})$ is the vector of rates of change of environmental variables due to intracellular reactions. Note that $\dot{\mathbf{C}}(\mathbf{z} | \mathbf{C})$ can be obtained by the product of stoichiometric matrix $\mathbf{A}_C$ and intracellular reaction vector $\boldsymbol{\gamma}$, $\dot{\mathbf{C}}(\mathbf{z} | \mathbf{C}) = \mathbf{A}_C \boldsymbol{\gamma}$. Recalling $(\mathbf{z} | \mathbf{C}) = \mathbf{A}_Z \boldsymbol{\gamma}$, we know that both $\dot{\mathbf{C}}(\mathbf{z} | \mathbf{C})$ and $(\mathbf{z} | \mathbf{C})$ come from the same intracellular reaction vector $\boldsymbol{\gamma}$. $\mathbf{N}_Z$ is the total flux of extracellular variables due to convection and diffusion. The coupled equations (B11) and (B18) are subject to initial conditions such as

$$n_Z(\mathbf{z}, \mathbf{r}, 0) = N_o(\mathbf{r}) f_{Z,o}(\mathbf{z}), \quad \mathbf{C}(\mathbf{r}, 0) = \mathbf{C}_o(\mathbf{r}) \quad (\text{B19})$$

which describe the initial distribution of cells and environmental variables, the former in terms of internal and external coordinates and the latter in external coordinates.

When no coupling is assumed between the processes of cell growth and division with the gene regulatory process of interest, we may write more specific versions of Eq.(B18) using fluxes $\mathbf{N}_X$ and $\mathbf{N}_Y$ as below.

$$\frac{\partial \mathbf{C}_X}{\partial t} + \nabla_{\mathbf{r}} \cdot \mathbf{N}_X = \int \dot{\mathbf{C}}_X(\mathbf{x}|\mathbf{C}_X) n_X(\mathbf{x}, \mathbf{r}, t) d\mathbf{x} \quad (\text{B20})$$

$$\frac{\partial \mathbf{C}_Y}{\partial t} + \nabla_{\mathbf{r}} \cdot \mathbf{N}_Y = \int \dot{\mathbf{C}}_Y(\mathbf{y}|\mathbf{C}_Y) n_Y(\mathbf{y}, \mathbf{r}, t) d\mathbf{y} \quad (\text{B21})$$

where $\dot{\mathbf{X}}$ and $\dot{\mathbf{Y}}$ are depicted to involve only $\mathbf{C}_X$ and $\mathbf{C}_Y$ respectively instead of $\mathbf{C}$ as implied before. The vector $\dot{\mathbf{C}}_X(\mathbf{x}|\mathbf{C}_X)$ relates change of extracellular variables to intracellular reactions associated with gene regulation only and $\dot{\mathbf{C}}_Y(\mathbf{y}|\mathbf{C}_Y)$ is the vector relates the change of extracellular variables (such as nutrients, fermentation products etc.) to growth and associated metabolic processes. The initial conditions Eq. (B19) must be further embellished in this case as below.

$$\mathbf{C}_X(\mathbf{r}, 0) = \mathbf{C}_{X,o}(\mathbf{r}), \quad \mathbf{C}_Y(\mathbf{r}, 0) = \mathbf{C}_{Y,o}(\mathbf{r}) \quad (\text{B22})$$

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.(B11) 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.

$$\frac{\partial n_Y(\mathbf{y}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{y}} \cdot \dot{\mathbf{Y}}(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) n_Y(\mathbf{y}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_Y(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) n_Y(\mathbf{y}, \mathbf{r}, t) = -\sigma_Y(\mathbf{y}, \mathbf{r}, t|\mathbf{C}) n_Y(\mathbf{y}, \mathbf{r}, t) + 2 \int \sigma_Y(\mathbf{y}', \mathbf{r}, t|\mathbf{C}) p_{Y|Y'}(\mathbf{y}, \mathbf{r}, t|\mathbf{y}', \mathbf{C}) n_Y(\mathbf{y}', \mathbf{r}, t) d\mathbf{y}' \quad (\text{B23})$$

Eq.(B23) resembles the population balance equation derived by Fredrickson (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. (B6), (B9) and (B13) yield the population balance equation

$$\frac{\partial n_Y(\mathbf{y}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{y}} \cdot \dot{\mathbf{Y}}(\mathbf{y}|\mathbf{C}_Y) n_Y(\mathbf{y}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_Y(\mathbf{y}|\mathbf{C}_Y) n_Y(\mathbf{y}, \mathbf{r}, t) = -\sigma_Y(\mathbf{y}|\mathbf{C}_Y) n_Y(\mathbf{y}, \mathbf{r}, t) + 2 \int \sigma_Y(\mathbf{y}'|\mathbf{C}_Y) p_{Y|Y'}(\mathbf{y}|\mathbf{y}', \mathbf{C}_Y) n_Y(\mathbf{y}', \mathbf{r}, t) d\mathbf{y}' \quad (\text{B24})$$

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. (B11) over all $\mathbf{y}$ to obtain

$$\begin{aligned} \frac{\partial n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) = \\ \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) - \sigma_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t | \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) \\ + 2 \int \sigma_{\mathbf{x}}(\mathbf{x}', \mathbf{r}, t | \mathbf{C}) p_{\mathbf{x}|\mathbf{x}'}(\mathbf{x}, \mathbf{r}, t | \mathbf{x}', \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}', \mathbf{r}, t) d\mathbf{x}' \end{aligned} \quad (\text{B25})$$

The use of Eq.(B25) 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 Eq. (B26).

$$\begin{aligned} \frac{\partial n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x} | \mathbf{C}_{\mathbf{x}}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_{\mathbf{x}}(\mathbf{x} | \mathbf{C}_{\mathbf{x}}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) = \\ \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x} | \mathbf{C}_{\mathbf{x}}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x} | \mathbf{C}_{\mathbf{x}}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) - \mu(\mathbf{r}, t | \mathbf{C}_{\mathbf{y}}) n_{\mathbf{x}}(\mathbf{x}, \mathbf{r}, t) + \\ + 2\mu(\mathbf{r}, t | \mathbf{C}_{\mathbf{y}}) \int p_{\mathbf{x}|\mathbf{x}'}(\mathbf{x} | \mathbf{x}', \mathbf{C}) n_{\mathbf{x}}(\mathbf{x}', \mathbf{r}, t) d\mathbf{x}' \end{aligned} \quad (\text{B26})$$

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

$$p_{\mathbf{x}|\mathbf{x}'}(\mathbf{x} | \mathbf{x}', \mathbf{C}) = \delta(\mathbf{x} - \mathbf{x}') \quad (\text{B27})$$

Combining Eqs. (B26) and (B27), 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. (B24) became Eq. (B28) which should be coupled with Eq. (B21).

$$\begin{aligned} \frac{\partial n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t)}{\partial t} + \nabla_{\mathbf{y}} \cdot \dot{\mathbf{Y}}(\mathbf{y} | \mathbf{C}_{\mathbf{y}}) n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) = \\ -\sigma_{\mathbf{y}}(\mathbf{y} | \mathbf{C}_{\mathbf{y}}) n_{\mathbf{y}}(\mathbf{y}, \mathbf{r}, t) + 2 \int \sigma_{\mathbf{y}}(\mathbf{y}' | \mathbf{C}_{\mathbf{y}}) p_{\mathbf{y}|\mathbf{y}'}(\mathbf{y} | \mathbf{y}', \mathbf{C}_{\mathbf{y}}) n_{\mathbf{y}}(\mathbf{y}', \mathbf{r}, t) d\mathbf{y}' \end{aligned} \quad (\text{B28})$$

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 $\mathbf{y}$. We have thus the means to compute $\mu(\mathbf{r}, t | \mathbf{C}_{\mathbf{y}})$, $\mathbf{C}_{\mathbf{y}}(\mathbf{r}, t)$.

Table B1

Nomenclature

$\mathbf{A}_{\mathbf{C}}$	Stoichiometric matrix relating the rates of change of environmental variables to those of the intracellular reactions
$\mathbf{A}_{\mathbf{Z}}$	Stoichiometric matrix relating change of extracellular variables to intracellular stochastic reactions
$\dot{\mathbf{C}}(\mathbf{z} \mathbf{C})$	The vector of rates of change of environmental variables due to intracellular reactions
$\dot{\mathbf{C}}_{\mathbf{x}}(\mathbf{x} \mathbf{C}_{\mathbf{x}})$	The vector of rates of change of environmental variables due to intracellular reactions associated with gene regulation only
$\dot{\mathbf{C}}_{\mathbf{y}}(\mathbf{y} \mathbf{C}_{\mathbf{y}})$	The vector of rates of change of environmental variables due to intracellular reactions associated with growth and duplication
$\mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C})$	If $\mathbf{B}_{\mathbf{Z}}(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{y}$, $\mathbf{B}_{\mathbf{Z}}(\mathbf{z} \mathbf{C})$ can be written as $\mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C})$
$\mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C}_{\mathbf{x}})$	The same as $\mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C})$ but further specified extracellular variables

$\mathbf{B}_Z(\mathbf{z} \mathbf{C})$	The matrix describing stochastic change of Ito stochastic differential equation
$\mathbf{C}$	Concentration of all extracellular variables
$\mathbf{C}_X$	Concentration of extracellular variables relating to stochastic intracellular variables
$\mathbf{C}_{X,o}(\mathbf{r})$	The initial condition of $\mathbf{C}_X$ at $t = 0$
$\mathbf{C}_Y$	Concentration of extracellular variables relating to deterministic intracellular variables
$\mathbf{C}_{Y,o}(\mathbf{r})$	The initial condition of $\mathbf{C}_Y$ at $t = 0$
$f_X(\mathbf{x}, \mathbf{r}, t)$	Probability density function, $f_X(\mathbf{z}, \mathbf{r}, t) = n_X(\mathbf{z}, \mathbf{r}, t)/N(\mathbf{r}, t)$
$f_{X,o}(\mathbf{x})$	$N_o(\mathbf{r}) f_{X,o}(\mathbf{x})$ give the initial condition of n_X at $t = 0$
$f_Y(\mathbf{y}, \mathbf{r}, t)$	Probability density function, $f_Y(\mathbf{z}, \mathbf{r}, t) = n_Y(\mathbf{z}, \mathbf{r}, t)/N(\mathbf{r}, t)$
$f_Z(\mathbf{z}, \mathbf{r}, t)$	Probability density function, $f_Z(\mathbf{z}, \mathbf{r}, t) = n_Z(\mathbf{z}, \mathbf{r}, t)/N(\mathbf{r}, t)$
$f_{Z,o}(\mathbf{z})$	$N_o(\mathbf{r}) f_{Z,o}(\mathbf{z})$ give the initial condition of n_Z at $t = 0$
J	Jacobian of the forward transformation, $J(\mathbf{r}, t \mathbf{r}_o) \equiv \left \frac{\partial \mathbf{R}}{\partial \mathbf{r}_o} \right $
m	Number of stochastic variables
n	Number of deterministic variables
$N(\mathbf{r}, t)$	Total number density at location $\mathbf{r}$ and time t , $N(\mathbf{r}, t) = \int n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{z}$
$N_o(\mathbf{r})$	The initial condition of $N(\mathbf{r}, t)$ at $t = 0$
$\mathbf{N}_X$	The total flux, including convection and diffusion, of extracellular variables relating to stochastic intracellular variables
$\mathbf{N}_Y$	The total flux, including convection and diffusion, of extracellular variables relating to deterministic intracellular variables
$\mathbf{N}_Z$	The total flux of extracellular variables due to convection and diffusion
$n_X(\mathbf{x}, \mathbf{r}, t)$	Number density, $\int n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{y} = n_X(\mathbf{x}, \mathbf{r}, t)$
$n_Y(\mathbf{y}, \mathbf{r}, t)$	Number density, $\int n_Z(\mathbf{z}, \mathbf{r}, t) d\mathbf{x} = n_Y(\mathbf{y}, \mathbf{r}, t)$
$n_Z(\mathbf{z}, \mathbf{r}, t)$	Number density, number of cells per unit volume of both internal and spatial coordinates
$p_{Y Y'}(\mathbf{y} \mathbf{y}', \mathbf{C})$	The distribution of states $\mathbf{y}$ of daughters from a parent of state $\mathbf{y}'$
$p_{Z Z'}(\mathbf{z} \mathbf{z}', \mathbf{C})$	The distribution of states $\mathbf{z}$ of daughters from a parent of state $\mathbf{z}'$
$\mathbf{r}$	Spatial coordinates
$\mathbf{r}_o$	The initial condition of $\mathbf{r}$ at $t = 0$
$\mathbf{R}(d\mathbf{r}_o)$	A function describing the path of cellular motion, $\mathbf{r} = \mathbf{R}(d\mathbf{r}_o)$
$\mathbf{R}_o(\mathbf{r}, t)$	If the inverse relationship of $\mathbf{r}$ and $\mathbf{r}_o$ is unique, $\mathbf{r}_o = \mathbf{R}_o(\mathbf{r}, t)$
$\dot{\mathbf{R}}(\mathbf{r}, t)$	The rate of spatial cell displacement depend on $\mathbf{r}$ and t
$\dot{\mathbf{R}}_X(\mathbf{x} \mathbf{C})$	If $\dot{\mathbf{R}}_Z(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{y}$, $\dot{\mathbf{R}}_Z(\mathbf{z} \mathbf{C})$ can be written as $\dot{\mathbf{R}}_X(\mathbf{x} \mathbf{C})$
$\dot{\mathbf{R}}_Y(\mathbf{y} \mathbf{C})$	If $\dot{\mathbf{R}}_Z(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{y}$, $\dot{\mathbf{R}}_Z(\mathbf{z} \mathbf{C})$ can be written as $\dot{\mathbf{R}}_Y(\mathbf{y} \mathbf{C})$
$\dot{\mathbf{R}}_Z(\mathbf{z} \mathbf{C})$	The rate of spatial cell displacement depend on $\mathbf{z}$ and $\mathbf{C}$
$\mu(\mathbf{r}, t \mathbf{C}_Y)$	the cell division rate regardless of the cell state at location $\mathbf{r}$ at time t
$d\mathbf{W}$	The standard Wiener process increment
$\mathbf{x}$	Stochastic variables
$\dot{\mathbf{X}}(\mathbf{x} \mathbf{C})$	If $\dot{\mathbf{X}}_Z(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{y}$, $\dot{\mathbf{X}}_Z(\mathbf{z} \mathbf{C})$ can be written as $\dot{\mathbf{X}}(\mathbf{x}, \mathbf{C})$
$\dot{\mathbf{X}}(\mathbf{x} \mathbf{C}_X)$	The same as $\dot{\mathbf{X}}(\mathbf{x}, \mathbf{C})$ but further specified extracellular variables
$\dot{\mathbf{X}}_Z(\mathbf{z} \mathbf{C})$	The mean rate of change of stochastic intracellular variables $\mathbf{x}$

$\mathbf{y}$	Deterministic variables
$\dot{\mathbf{Y}}(\mathbf{y} \mathbf{C})$	If $\dot{\mathbf{Y}}_Z(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{x}$, $\dot{\mathbf{Y}}_Z(\mathbf{z} \mathbf{C})$ can be written as $\dot{\mathbf{Y}}(\mathbf{y} \mathbf{C})$
$\dot{\mathbf{Y}}(\mathbf{y} \mathbf{C}_Y)$	The same as $\dot{\mathbf{Y}}(\mathbf{y} \mathbf{C})$ but further specified extracellular variables
$\dot{\mathbf{Y}}_Z(\mathbf{z} \mathbf{C})$	The rate of change of the deterministic intracellular variables $\mathbf{y}$
$\mathbf{z} \equiv (\mathbf{x}, \mathbf{y})$	All intracellular state variables
$\dot{(\mathbf{z} \mathbf{C})}$	The rate of change of the intracellular state vector $\mathbf{z}$
$\boldsymbol{\gamma}$	The vectors describing intracellular reactions
$\sigma_Y(\mathbf{y} \mathbf{C}_Y)$	If $\sigma_Z(\mathbf{z} \mathbf{C})$ does not depend on $\mathbf{x}$, $\sigma_Z(\mathbf{z} \mathbf{C})$ can be written as $\sigma_Y(\mathbf{y} \mathbf{C}_Y)$
$\sigma_Z(\mathbf{z} \mathbf{C})$	The cell division rate

References

- Aris, R. Vectors Tensors, and the Basic Equations of Fluid Mechanics. Prentice-Hall Englewood Cliffs; New Jersey: 1962.
- Bae T, Kozlowicz BK, Dunny GM. Characterization of cis-acting prgQ mutants: evidence for two distinct repression mechanisms by Qa RNA and PrgX protein in pheromone-inducible enterococcal plasmid pCF10. *Molecular Microbiology*. 2004; 51:271–281. [PubMed: 14651627]
- Bensing BA, Dunny GM. Pheromone-inducible expression of an aggregation protein in *Enterococcus faecalis* requires interaction of a plasmid-encoded RNA with components of the ribosome. *Molecular Microbiology*. 1997; 24:295–308. [PubMed: 9159517]
- Buttaro BA, Antiporta MH, Dunny GM. Cell-associated pheromone peptide (cCF10) production and pheromone inhibition in *Enterococcus faecalis*. *Journal of Bacteriology*. 2000; 182:4926–4933. [PubMed: 10940037]
- Chandler JR, Dunny GM. Characterization of the sequence specificity determinants required for processing and control of sex pheromone by the intramembrane protease Eep and the plasmid-encoded protein PrgY. *Journal of Bacteriology*. 2008; 190:1172–1183. [PubMed: 18083822]
- Dunny G, Funk C, Adsit J. Direct stimulation of the transfer of antibiotic-resistance by sex-pheromones in *Streptococcus-faecalis*. *Plasmid*. 1981; 6:270–278. [PubMed: 6796985]
- Ferrell JE. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology*. 2002; 14:140–148. [PubMed: 11891111]
- Fixen KR, Chandler JR, Le T, Kozlowicz BK, Manias DA, Dunny GM. Analysis of the amino acid sequence specificity determinants of the enterococcal cCF10 sex pheromone in interactions with the pheromone-sensing machinery. *Journal of Bacteriology*. 2007; 189:1399–1406. [PubMed: 17098891]
- Fredrickson A, Ramkrishna D, Tsuchiya H. Statistics and dynamics of procaryotic cell populations. *Mathematical Biosciences*. 1967; 1:327–374.
- Fredrickson AG, Mantzaris NV. A new set of population balance equations for microbial and cell cultures. *Chemical Engineering Science*. 2002; 57:2265–2278.
- Gardiner, CW. Handbook of stochastic methods for physics chemistry and natural science. Springer-Verlag; Berlin Heidelberg New: 1997.
- Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*. 2000; 403:339–342. [PubMed: 10659857]
- Johnson CM, Manias DA, Haemig HAH, Shokeen S, Weaver KE, Henkin TM, Dunny GM. Direct Evidence for Control of the Pheromone-Inducible prgQ Operon of *Enterococcus faecalis* Plasmid pCF10 by a Countertranscript-Driven Attenuation Mechanism. *Journal of Bacteriology*. 2010; 192:1634–1642. [PubMed: 20097859]
- Haseltine EL, Rawlings JB. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *Journal of Chemical Physics*. 2002; 117:6959–6969.
- Hjortsø, MA. Population Balances in Biomedical Engineering: Segregation Through the Distribution of Cell States. McGraw-Hill; New York: 2005.

- Redfield R. Is quorum sensing a side effect of diffusion sensing? *TRENDS in Microbiology*. 2002; 10:365–370. [PubMed: 12160634]
- Mantzaris NV. Stochastic and deterministic simulations of heterogeneous cell population dynamics. *Journal of Theoretical Biology*. 2006; 241:690–706. [PubMed: 16487980]
- Kepler TB, Elston TC. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophysical Journal*. 2001; 81:3116–3136. [PubMed: 11720979]
- Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ. Programmable cells: Interfacing natural and engineered gene networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:8414–8419. [PubMed: 15159530]
- Kozlowski BK, Shi K, Gu ZY, Ohlendorf DH, Earhart CA, Dunny GM. Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. *Molecular Microbiology*. 2006; 62:958–969. [PubMed: 17038121]
- Kristich CJ, Li YH, Cvitkovitch DG, Dunny GM. Esp-independent biofilm formation by *Enterococcus faecalis*. *Journal of Bacteriology*. 2004; 186:154–163. [PubMed: 14679235]
- Leonard B, Podbielski A, Hedberg P, Dunny G. *Enterococcus faecalis* pheromone binding protein, PrgZ, recruits a chromosomal oligopeptide permease system to import sex pheromone cCF10 for induction of conjugation. *Proceedings of the National Academy of Sciences of the United States of America*. 1996; 93:260. [PubMed: 8552617]
- Nakayama J, Ruhfel RE, Dunny GM, Isogai A, Suzuki A. The prgQ gene of the *Enterococcus-faecalis* tetracycline resistance plasmid pCF10 encodes a peptide inhibitor, iCF10. *Journal of Bacteriology*. 1994; 176:7405–7408. [PubMed: 7545961]
- Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A. Multistability in the lactose utilization network of *Escherichia coli*. *Nature*. 2004; 427:737–740. [PubMed: 14973486]
- Ramkrishna D. Statistical Models of Cell Populations, *Advances in Biochemical Engineering*. 1979; 11:1–47.
- Ramkrishna, D. Population balances: Theory and applications to particulate systems in engineering. Academic Press; San Diego: 2000.
- Rao CV, Arkin AP. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *Journal of Chemical Physics*. 2003; 118:4999–5010.
- Rao NJ, Borwanker JD, Ramkrish D. Numerical-solution of Ito integral-equations. *Siam Journal on Control*. 1974; 12:124–139.
- Shi K, Brown CK, Gu ZY, Kozlowski B, Danny GM, Ohlendorf DH, Earhart CA. Crystal structure of PrgX and PrgX/pheromone: The role of tetramerization of PrgX in controlling the pheromone induction of pCF10 transfer. Abstracts of the General Meeting of the American Society for Microbiology. 2005; 105:307.
- Subramanian G, Ramkrishna D. On the solution of statistical models of cell populations. *Mathematical Biosciences*. 1971; 10:1–23.
- Subramanian G, Ramkrishna D, Fredrickson A, Tsuchiya H. On the mass distribution model for microbial cell populations. *Bulletin of Mathematical Biology*. 1970; 32:521–537.
- Talay D. Simulation of stochastic differential systems. *Probabilistic methods in applied physics*. 1995:54–96.
- Tian TH, Burrage K. Stochastic models for regulatory networks of the genetic toggle switch. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:8372–8377. [PubMed: 16714385]
- Tomshine J, Kaznessis YN. Optimization of a stochastically simulated gene network model via simulated annealing. *Biophysical Journal*. 2006; 91:3196–3205. [PubMed: 16920827]
- Trotter K, Dunny G. Mutants of *Enterococcus faecalis* deficient as recipients in mating with donors carrying pheromone-inducible plasmids. *Plasmid*. 1990; 24:57–67. [PubMed: 2125350]
- Waters CM, Hirt H, McCormick JK, Schlievert PM, Wells CL, Dunny GM. An amino-terminal domain of *Enterococcus faecalis* aggregation substance is required for aggregation, bacterial internalization by epithelial cells and binding to lipoteichoic acid. *Molecular Microbiology*. 2004; 52:1159–1171. [PubMed: 15130132]
- Wehner M, Wolfer W. Numerical evaluation of path-integral solutions to Fokker-Planck equations. *Physical Review A*. 1983; 27:2663–2670.

Figure 1. The gene regulatory network of pCF10 conjugation system

The *prgQ*-*prgX* gene pair regulates conjugation through sense-antisense interaction, the reaction between Anti-Q and Q_{pre} . The ratio of cCF10 to inhibitor controls the configuration of plasmid DNA which decides the transcription rate of *prgQ*. While pheromone, cCF10, is released by recipient cells in the extracellular environment, the inhibitor iCF10 is encoded from both Q_s and Q_L RNA, products of the *prgQ* gene. Both cCF10 and iCF10 compete for binding to regulatory protein which alters the DNA configuration through binding site Bl and Bs. A cell favors the expression of *prgQ* when the ratio of cCF10 to iCF10 is high. High concentration of Q_{pre} make a cell produces Q_L mRNA which stimulates the expression of *prgB*, an indicator for the onset of conjugation. In the opposite, when the ratio of cCF10 to iCF10 is low, due to sense-antisense interaction almost all Q_{pre} become Q_s which has no effect on stimulating *prgB* expression.

Figure 2. The bistability of 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 3. 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.

Figure 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 from bimodal distribution to unimodal distribution.

Figure 4. Protein distribution of donors in a well mixed system, including donors and recipients, without conjugation

- a) PrgB protein distribution of donors. Pheromone, 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.

Figure 5. The PrgB protein distribution in a well mixed system with conjugation

- a) PrgB protein distribution in donor cells. Donor population changes due to replication and conjugation. Following transfer of plasmid from a donor to a recipient, both cells become donors with drug resistance. This process significantly affects the distribution as seen from a “turn- back” of protein level to lower concentration ranges.
- b) Number density for PrgB protein showing increase due to replication and conjugation with eventual decrease in the range of PrgB distribution, also seen in a).

Table 1

Variables

Notation	Name
x_1	Intracellular concentration of Q_s mRNA
x_2	Intracellular concentration of Q_L mRNA
x_3	Intracellular concentration of Anti-Q RNA
x_4	Intracellular concentration of inhibitor, iCF10
x_5	Concentration of PrgB membrane protein
x_6	Intracellular concentration of pheromone, cCF10
C_1	Extracellular concentration of pheromone, cCF10
C_2	Extracellular concentration of pheromone, iCF10

Table 2

Equations (refer to section 1.2 Cell in Well-mixed Environment)

$n_{\mathbf{x}}(\mathbf{x}, t) = P(\mathbf{x}, t) N(t)$
$\frac{\partial P(\mathbf{x}, t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x} \mathbf{C}) P(\mathbf{x}, t) = \frac{1}{2} \nabla_{\mathbf{x}} \nabla_{\mathbf{x}} : \mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x} \mathbf{C}) P(\mathbf{x}, t)$
$\dot{\mathbf{X}}(\mathbf{x} \mathbf{C}) = \begin{bmatrix} k[K_{1,1}\alpha + K_{1,2}(1-\alpha)](\frac{K_{3,5}x_3}{1+K_{3,5}x_3}) - (K_{4,1} + \mu_d)x_1 \\ k[K_{1,1}\alpha + K_{1,2}(1-\alpha)](\frac{1}{1+K_{3,5}x_3}) - (K_{4,2} + \mu_d)x_2 \\ k[K_{1,3}\alpha + K_{1,4}(1-\alpha)] - k[K_{1,1}\alpha + K_{1,2}(1-\alpha)](\frac{K_{3,5}x_3}{1+K_{3,5}x_3}) - (K_{4,3} + \mu_d)x_3 \\ K_{2,6}C_2 - (K_{4,6} + \mu_d)x_4 \\ K_{1,5}x_2 - (K_{4,9} + \mu_d)x_5 \\ K_{2,8}C_1 - (K_{4,8} + \mu_d)x_6 \end{bmatrix} I$
$\alpha = \frac{x_4^4}{x_4^4 + K_{3,8}x_6^4}$ <p>where:</p>
$\mathbf{B}_{\mathbf{x}}(\mathbf{x} \mathbf{C}) \mathbf{B}_{\mathbf{x}}^T(\mathbf{x} \mathbf{C}) \equiv \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{66} \end{bmatrix}^2$
<p>where:</p> $B_{44} = \frac{1}{v_d} [K_{2,6}C_2 + K_{4,6}x_4]$ $B_{55} = \frac{1}{v_d} [K_{1,5}x_2 + K_{4,9}x_5]$ $B_{66} = \frac{1}{v_d} [K_{2,8}C_1 + K_{4,8}x_6]$
<p>Case 1</p> $N(t) = N_o$ $\frac{d\mathbf{C}_{\mathbf{x}}}{dt} = N_o E[\dot{\mathbf{C}}_{\mathbf{x}}] - \Phi \mathbf{C}_{\mathbf{x}} - \Omega \mathbf{C}_{\mathbf{x}}$ $\mathbf{C}_{\mathbf{x}} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}; \dot{\mathbf{C}}_{\mathbf{x}} = \begin{bmatrix} 0 \\ v_d [K_{1,6}(x_1 + x_2) - K_{2,6}C_2] \end{bmatrix}; \Phi = \begin{bmatrix} 0 & 0 \\ 0 & K_{4,5} \end{bmatrix}; \Omega = \begin{bmatrix} 0 & 0 \\ 0 & (d \ln V / dt) \end{bmatrix};$
<p>Case 2</p> $N(t) = N_o e^{\mu_r t}$ $\frac{d\mathbf{C}_{\mathbf{x}}}{dt} = N(t) E[\dot{\mathbf{C}}_{\mathbf{x}}] - \Phi \mathbf{C}_{\mathbf{x}} + N_r(t) \Psi \quad \text{where: } N_r(t) = N_o e^{\mu_r t}$ $\mathbf{C}_{\mathbf{x}} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}; \dot{\mathbf{C}}_{\mathbf{x}} = \begin{bmatrix} v_d C_1 K_{2,8} \\ v_d [K_{1,6}(x_1 + x_2) - K_{2,6}C_2] \end{bmatrix}; \Phi = \begin{bmatrix} K_{4,8} & 0 \\ 0 & K_{4,5} \end{bmatrix}; \Psi = \begin{bmatrix} K_{1,8} \\ 0 \end{bmatrix};$
<p>Case 3</p>

$$\begin{aligned}
 \frac{dN}{dt} &= \mu_d N + k_{con} N_r N E[x_5] \quad \text{with } N(0) = N_o \\
 \frac{dN_r}{dt} &= \mu_r N_r - k_{con} N_r N E[x_5] \quad \text{with } N_r(0) = N_{r_o} \\
 \frac{d\mathbf{C}_x}{dt} &= N(t)E[\dot{\mathbf{C}}_x] - \Phi \mathbf{C}_x + N_r(t)\Psi \\
 \mathbf{C}_x &= \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}; \quad \dot{\mathbf{C}}_x = \begin{bmatrix} v_d C_1 K_{2,8} \\ v_d [K_{1,6}(x_1 + x_2) - K_{2,6} C_2] \end{bmatrix}; \quad \Phi = \begin{bmatrix} K_{4,8} & 0 \\ 0 & K_{4,5} \end{bmatrix}; \quad \Psi = \begin{bmatrix} K_{1,8} \\ 0 \end{bmatrix};
 \end{aligned}$$

¹Row 1 and 2, $k[K_{1,1}\alpha + K_{1,2}(1 - \alpha)]$ is the total transcription rate of prgQ gene and the ratio after it indicate the partition of Q_{pre} become Q_s or Q_L. The sink term is from degradation and dilution due to cell growth. The first term of row 3 indicates the transcription rate of Anti-Q and the second term is the loss due to reaction with Q_{pre}. Row 4 and 6 assume importation of inhibitor and pheromone linear to extracellular concentration. Row 5 assumes the expression of prgB proportional to Q_L.

²Referring to Appendix A for detail derivatives.

Table 3

Parameters

Constant	Name
k	plasmid copy number, equal to 5
α	DNA of plasmid in loop form, repressed state of <i>prgQ</i>
v_d	volume per donor cell
V	extracellular volume, assuming $d \ln V(t)/dt = \mu$ with initial condition $V(t=0) = N_{d0} v_d$ for case 1 and constant as $V = N_{d0} v_d$ for case 2 and 3, where N_{d0} is the donor cell number at $t=0$
N_o	the initial number density of donor cells

Constant	Name	Value ⁴	
K _{1,1}	transcription rate of <i>prgQ</i> , DNA in loop form	0.0084	(nM/s)
K _{1,2}	transcription rate of <i>prgQ</i> , DNA in un-loop form	0.0876	(nM/s)
K _{1,3}	transcription rate of Anti-Q, DNA in loop form	0.0125	(nM/s)
K _{1,4}	transcription rate of Anti-Q, DNA in un-loop form	0.0014	(nM/s)
K _{1,5}	generation rate of PrgB	0.01	(1/s)
K _{1,6}	generation rate of extracellular inhibitor, iCF10	0.002	(1/s)
K _{1,8}	generation rate of pheromone, cCF10	0.001	(1/s)
K _{2,6}	importation rate of inhibitor, iCF10	0.001	(1/s)
K _{2,8}	importation rate of pheromone, cCF10	0.000258	(1/s)
K _{3,5}	equilibrium constant of Q _{pre} and Anti-Q interaction	0.0443	(1/nM)
K _{3,8}	equilibrium constant of DNA binding reaction	1000	-
K _{4,1}	degradation rate of Q _s mRNA	0.001	(1/s)
K _{4,2}	degradation rate of Q _L mRNA	0.100	(1/s)
K _{4,3}	degradation rate of Anti-Q RNA	0.000136	(1/s)
K _{4,5}	degradation rate of extracellular inhibitor, iCF10	1.00E-06	(1/s)
K _{4,6}	degradation rate of intracellular inhibitor, iCF10	1.00E-06	(1/s)
K _{4,8}	degradation rate of intracellular pheromone, cCF10	1.00E-06	(1/s)
K _{4,9}	degradation rate of PrgB protein	1.00E-06	(1/s)
μ_d	Specific growth rate of donor cells	0.0002567	(1/s)
μ_r	Specific growth rate of recipient cells	0.0002888	(1/s)
k_{con}	conjugation constant	1.00E-09	(1/s nM ³)

⁴The values are adopted from those used for a similar reaction system (Tomshine and Kaznessis, 2006).

Table 4

Fokker-Planck Equation for Single Cell

$$\frac{\partial P(\mathbf{x}_c, t)}{\partial t} + \nabla_{\mathbf{x}_c} \cdot \dot{\mathbf{x}}_c(\mathbf{x}_c|\mathbf{C})P(\mathbf{x}_c, t) = \frac{1}{2} \nabla_{\mathbf{x}_c} \nabla_{\mathbf{x}_c} : \mathbf{B}_{\mathbf{x}_c}(\mathbf{x}_c|\mathbf{C}) \mathbf{B}_{\mathbf{x}_c}^T(\mathbf{x}_c|\mathbf{C}) P(\mathbf{x}_c, t)$$

$$\mathbf{x}_c^T = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad C_2]$$

$$\dot{\mathbf{x}}_c(\mathbf{x}_c|\mathbf{C}) = \begin{bmatrix} k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3}\right) - (K_{4,1} + \mu_d)x_1 \\ k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{1}{1+K_{3,5}x_3}\right) - (K_{4,2} + \mu_d)x_2 \\ k[K_{1,3}\alpha + K_{1,4}(1-\alpha)] - k[K_{1,1}\alpha + K_{1,2}(1-\alpha)]\left(\frac{K_{3,5}x_3}{1+K_{3,5}x_3}\right) - (K_{4,3} + \mu_d)x_3 \\ K_{2,6}C_2 - (K_{4,6} + \mu_d)x_4 \\ K_{1,5}x_2 - (K_{4,9} + \mu_d)x_5 \\ K_{2,8}C_1 - (K_{4,8} + \mu_d)x_6 \\ \frac{v_d}{V}[K_{1,6}(x_1 + x_2) - K_{2,6}C_2] - (K_{4,5} + \mu_d)C_2 \end{bmatrix}$$

$$\mathbf{B}_{\mathbf{x}_c}(\mathbf{x}_c|\mathbf{C}) \mathbf{B}_{\mathbf{x}_c}^T(\mathbf{x}_c|\mathbf{C}) \equiv \mathbf{B}_c = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{44} & 0 & 0 & B_{47} \\ 0 & 0 & 0 & 0 & B_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{66} & 0 \\ 0 & 0 & 0 & B_{74} & 0 & 0 & B_{77} \end{bmatrix}$$

$$B_{77} = \frac{v_d}{V^2} [K_{1,6}(x_1 + x_2) + K_{2,6}C_2] + \frac{1}{V} K_{4,5}C_2$$

$$B_{47} = B_{74} = \frac{-1}{V} K_{2,6}C_2$$