Bistability versus Bimodal Distributions in Gene Regulatory Processes from Population Balance

Che-Chi Shu[1], Anushree Chatterjee[2], Gary Dunny[3], Wei-Shou Hu[2], Doraiswami Ramkrishna[1]*

1 School of Chemical Engineering, Purdue University, Indiana, United States of America, 2 Department of Chemical Engineering and Materials Science, University of Minnesota, Minnesota, United States of America, 3 Department of Microbiology, University of Minnesota, Minnesota, United States of America

Abstract

In recent times, stochastic treatments of gene regulatory processes have appeared in the literature in which a cell exposed to a signaling molecule in its environment triggers the synthesis of a specific protein through a network of intracellular reactions. The stochastic nature of this process leads to a distribution of protein levels in a population of cells as determined by a Fokker-Planck equation. Often instability 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 signaling molecule within a suitable range. A consequence of such bistability has been the appearance of bimodal distributions indicating two different populations, one in the "off" state and the other in the "on" state. The bimodal distribution can come about from stochastic analysis of a single cell. However, the concerted action of the population altering the extracellular concentration in the environment of individual cells and hence their behavior can only be accomplished by an appropriate population balance model which accounts for the reciprocal effects of interaction between the population and its environment. In this study, we show how to formulate a population balance model in which stochastic gene expression in individual cells is incorporated. Interestingly, the simulation of the model shows that bistability is neither sufficient nor necessary for bimodal distributions in a population. The original notion of linking bistability with bimodal distribution from single cell stochastic model is therefore only a special consequence of a population balance model.

Introduction

In the study of cell populations, with vastly improved flow cytometry, access to multivariate distribution measures of cell populations has advanced considerably, calling for a concomitant application of theory sensitive to population heterogeneity. In this regard, the population balance framework of Fredrickson et al. [1] has provided the requisite modeling machinery for the same. While this recognition generally exists in the literature, the modeling of gene regulatory processes has been at the single cell level based on it being viewed as an "average" cell. Since gene regulatory processes typically involve a small number of molecules, the reaction network is stochastic in its dynamics, a feature that is included in the single cell analysis. A further issue of importance, that of bistability, occurs when two levels of gene expression, one high and referred to as "on," and the other low and referred to as "off" exist for a given concentration of the signaling molecule. This issue is very much a part of the stochastic modeling of the single cell [2,3]. Several kinds of stochastic models have been developed; two of them that have been broadly used are the Stochastic Simulation Algorithm (SSA) [4,5], and the Fokker-Planck equation or Stochastic Differential Equations (SDE) [6–8]. The Stochastic model certainly cures the drawback of the deterministic model which describes only the averaged behavior on large populations without realizing the fluctuating behaviors in different cells.

Bistability has been studied extensively through experiments, theoretical analysis, and numerical simulations [2,3,9–11]. A bistable system is characterized by the existence of two stable steady states. The modes relating to two stable steady states appear as a bimodal distribution of the population. The coexistence of bistability and bimodal distribution has been shown in many publications [2,3,9,12–14].

However, almost all of the modeling works on stochastic gene regulation relate to processes at the single-cell level. The outcome of numerous simulated trajectories of single cell behavior has been interpreted as population behavior. A cell is assumed to act totally independently of other cells without regard to the fact that the signaling environment is continuously altered by the concerted action of all members of the population. That no interaction between other cells has been taken into consideration in these models could indeed lead to serious bias. The drawback of the single cell model may be overcome by applying the Population Balance approach [15]. A detailed general framework of the application of population balances to microbial populations was developed by Fredrickson et al. [1]. However, the population balance model (PBM) in the cited work and many others that followed in the literature are based on deterministic behavior of the particulate entities. Ramkrishna [15] shows how the PBM can accommodate random particulate behavior described by stochastic differential equations. In this study, we demonstrate formulating a stochastic

Author Summary

Traditionally cells in a population have been assumed to behave identically by using deterministic mathematical equations describing average cell behavior, thus ignoring its inherent randomness. A single cell stochastic model has therefore evolved in the literature to overcome this drawback. However, this single cell perspective does not account for interaction between the cell population and its environment. Since stochastic behavior leads to each cell acting differently, the cumulative impact of individual cells on their environment and consequent influence of the latter on each cell could constitute a behavior at variance. Thus in nature, cells are constantly under the influence of a highly dynamic environment which in turn is influenced by the dynamics of the cell population. A typical single cell stochastic model ignores such an interaction between the population and its environment, and uses probability distribution of a single cell to represent the entire population, which may lead to inappropriate predictions. In this study, we propose a population balance model coupled with stochastic gene regulation to demonstrate the behavior of a population in which its interactive behavior with its environment is considered. Our simulation results show that bistability is neither sufficient nor necessary for bimodal distributions in a population.

Figure 1. Schematic showing single cell bistability and population bimodal distribution. For cCF10 pheromone (signaling molecule) induced pCF10 conjugation system, concentration of PrgB protein indicates the level of conjugation. Our model indicates that (**A**) due to different plasmid copy number in culture of cells, bimodal population distribution can arise from cells without bistability and (**B**) due to interaction with each other, cells with bistability can abandon bimodal population distribution.
doi:10.1371/journal.pcbi.1002140.g001

gene regulation incorporating PBM, which is capable of tracing time evolution of the behavior of the entire cell population. A system of pheromone-induced conjugative plasmid transfer [16] contributing to the dissemination of antibiotic resistance and the virulence of *Enterococcus faecalis* infections [17,18] has been simulated in this study as an example of the critical difference between stochastic gene regulation incorporating PBM and single-cell stochastic model.

It is our objective in this paper to formulate population balance models with stochastic gene expression in single cells. Further, we explore circumstances under which bimodal distributions are observed in protein distributions; in particular we investigate the generally prevailing view in the literature that bistability and bimodality of protein distribution occur concurrently [2,3,9,12–14]. An exception to this view appears in the work of Karmakar and Bose [19,20], who showed that bimodal distributions can arise without bistability when the reaction time of the downstream gene regulation is short relative to the time required for change of DNA conformation. Other similar publications can also be found in literature [21–25]. While these cited works show bimodal distributions without bistability, it must be understood that their conclusions are based on mechanistic differences in the behavior of isolated single cells. In this study, we approach the issue of the relationship between bistability and bimodality from a rational viewpoint; i.e., to examine the nature of protein distribution from cells with and without bistability within the framework of population balances. Thus, circumstances will be investigated for Figure 1**A**, in which bimodal distributions can arise without bistability, and for Figure 1**B**, in which unimodal distributions can arise even when bistability exists.

Methods

The single cell deterministic (average) equations of pCF10 System

The gene regulatory network for pCF10 based conjugation system is shown in Figure 2**A**. Under natural circumstances, pCF10 deficient recipient cells release a pheromone called cCF10 into the extracellular environment, whereas pCF10 carrying donor
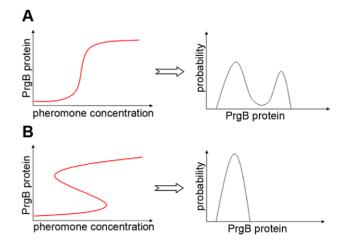
cells release an inhibitor molecule, iCF10 into the environment [26]. Both iCF10 and cCF10 are transported into the donor cells to interact with pCF10 DNA favoring off vs on state respectively. A pair of divergent genes *prgQ* and *prgX* present on pCF10 DNA regulates the genetic switch controlling onset of conjugation. The transcription of prgQ gene results in the formation of Q_{PRE}; Q_{PRE} gives rise to two kinds of RNAs known as Q_L RNA and Q_s RNA. In the opposite direction, the *prgX* encodes the PrgX repressor and a non-coding antisense RNA called Anti-Q which may bind to Q_{PRE} [27,28]. A Q_{PRE} bound with Anti-Q leads to shorter Q_s RNA. On the other hand, the final product of free Q_{PRE} is longer Q_L RNA. Under off conditions iCF10 bound PrgX tetramer represses *prgQ*; small amounts of Q_{PRE} are nearly all bound by overwhelming Anti-Q and result in $Q_{S,}$ which is incapable of inducing conjugation, predominantly expressed. In the on state iCF10 bound PrgX tetramer is replaced by cCF10 bound PrgX dimer which relieves repression of *prgQ*, thus causing expression of a longer Q_L transcript from *prgQ* gene [29]. The Q_L RNA consequently results in expression of PrgB protein, an indicator for the onset of conjugation [30].

The deterministic (average) equations based on mass-action kinetics that represent the gene regulatory network in Figure 2**A** are represented below with its associated nomenclature listed in Table 1.

$$\frac{dq_s}{dt} = k[K_{1,1}o + K_{1,2}(1-o)](\frac{K_{3,5}\,q_a}{1+K_{3,5}\,q_a}) - (K_{4,1}+\mu)\,q_s \quad (1)$$

$$\frac{dq_L}{dt} = k\,[K_{1,1}o + K_{1,2}(1-o)]\,(\frac{1}{1+K_{3,5}\,q_a}) - (K_{4,2}+\mu)\,q_L \quad (2)$$

$$\frac{dq_a}{dt} = k\,[K_{1,3}o + K_{1,4}(1-o)] - k\,[K_{1,1}o + K_{1,2}(1-o)]$$
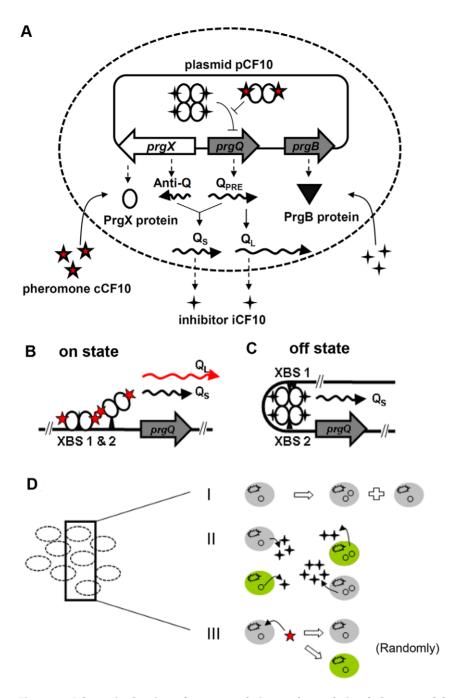$$(\frac{K_{3,5}\,q_a}{1+K_{3,5}\,q_a}) - (K_{4,3}+\mu)\,q_a \quad (3)$$

Figure 2. Schematic drawing of gene regulation and population balance model. (**A**) The gene reaction network of pCF10 based conjugation system. The *prgQ-prgX* gene pair regulates conjugation. 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 iCF10 and cCF10 compete for binding to PrgX protein which is assumed to exist at a constant concentration in this study. In the off state, iCF10-bound PrgX tetramers repress *prgQ* gene expression via formation of a DNA loop. Under these conditions, the nascent *prgQ* transcript Q_{pre} interacts with the non-coding antisense RNA, Anti-Q, to give rise to shorter Q_S RNA. In the on state PrgX-cCF10 dimers relieve *prgQ* repression to give rise to increased production of Q_{pre}, which tends to titrate Anti-Q, allowing production of the longer Q_L RNA and consequently PrgB protein. The concentration of PrgB protein indicates the level of conjugation. (**B**) the DNA configuration of on state. (**C**) the DNA configuration of off state. (**D**) Schematic depicting the Population Balance Model (PBM) with stochastic gene regulation. Grey color indicates a cell in off state and green indicates on state. Properties of the PBM include I) Uneven distribution of plasmids to daughter cells. II) Cells with different plasmid copy number or different states act differently and influence each others. III) A cell acts random according to stochastic intracellular gene regulation.
doi:10.1371/journal.pcbi.1002140.g002

$$\frac{di}{dt} = K_{2,6}I - (K_{4,6} + \mu)\,i \quad (4)$$

$$\frac{db}{dt} = K_{1,7} + \frac{K_{1,9}}{1 + \exp[-K_{3,9}(q_L - K_{1,8})]} - (K_{4,9} + \mu)\,b \quad (5)$$

Eq. (5), representing the mass balance of PrgB, treats the production rate as a sigmoid function of Q_L [31]. A variation considering the production rate of PrgB as a linear function of Q_L is contained in the modified differential equation (6) below.

Table 1. Nomenclature of pCF10 system.

Notation	Name
o	DNA of plasmid in loop form, repressed state of prgQ
k	Plasmid copy number
r	Order of DNA binding reaction, equal to 4
c	Intracellular concentration of pheromone, cCF10
q_s	Intracellular concentration of Q_s mRNA
q_L	Intracellular concentration of Q_L mRNA
q_a	Intracellular concentration of Anti-Q RNA
i	Intracellular concentration of inhibitor, iCF10
b	Concentration of PrgB membrane protein
I	Extracellular concentration of inhibitor, iCF10
I_n	Extracellular number of inhibitor, iCF10
C	Extracellular concentration of pheromone, cCF10
v_d	Volume per donor cell
V	Total volume

$$\frac{db}{dt} = K_{1,5}q_L - (K_{4,9} + \mu)\, b \qquad (6)$$

The differential equation which remains is that expressing the mass balance of intracellular pheromone concentration given by

$$\frac{dc}{dt} = K_{2,8}C - (K_{4,8} + \mu)\, c \qquad (7)$$

In Eqs. (1) to (3), $o \equiv \dfrac{i^r}{i^r + K_{3,8}\, c^r}$, represents DNA in repressed configuration. The above equations are based on mechanisms already published in the literature [28,30,32–34]. The parameter values for the simulations, adopted from those used for a similar reaction system [35], are summarized in Table 2. RNA species Q_S, Q_L and Anti-Q are described in Eqs.(1), (2) and(3), considering rate of production, degradation and dilution due to growth. The transcription rate of RNA species is modeled by a non-linear function of iCF10, cCF10 and Anti-Q to take into consideration effects of Transcriptional Interference and Antisense interaction, and are discussed in greater detail elsewhere [36]. Transport of both iCF10 and cCF10 into the donor cell is modeled as a first order reaction (Eqs. (4) and (7)). Eq. (5) considers the production rate of PrgB as a sigmoid function [31] with respect to Q_L and has been used to simulate Figures 3**B** and 4. Instead of sigmoid function, for reducing computational burden, Eq. (6) assumes that the expression of PrgB is linear in Q_L and has been used to simulate Figures 3**A**, Figure 5, Figure 6 and Text S1. The trends simulated by this deterministic model are consistent with experimental observations [17] (refer to Text S1). While the above notation for concentration serves to remind the reader of the reaction species to which it belongs, it is not convenient for their compact representation in the upcoming equations of population balance. Therefore, the intracellular concentrations are renamed as shown in Eq.(8).

$$\mathbf{x}^T = [\, x_1 \equiv q_s, \quad x_2 \equiv q_L, \quad x_3 \equiv q_a, \quad x_4 \equiv i, \quad x_5 \equiv b, \quad x_6 \equiv c \,] \quad (8)$$

Table 2. Values of Parameters of pCF10 system.

Reaction constant	Name	Value	
$K_{1,1}$	transcription rate of prgQ, DNA in loop form	0.0084	(nM/s)
$K_{1,2}$	transcription rate of prgQ, 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 for first order reaction	0.01	(1/s)
$K_{1,6}$	generation rate of extracellular inhibitor, iCF10	0.005	(1/s)
$K_{1,7}$	basic generating rate of membrane protein PrgB	0.00155	(nM/s)
$K_{1,8}$	threshold concentration of Q_L	12.00	(nM)
$K_{1,9}$	rate constant of generating membrane protein PrgB	0.0031	(nM/s)
$K_{2,6}$	importation rate of inhibitor, iCF10	0.001	(1/s)
$K_{2,8}$	importation rate of pheromone, cCF10	2.57E-04	(1/s)
$K_{3,5}$	equilibrium constant of Q_{PRE} and Anti-Q reaction	0.0443	(1/nM)
$K_{3,8}$	equilibrium constant of DNA binding reaction	1.00E06	-
$K_{3,9}$	constant of sigmoid function for Q_L to PrgB	12.00	(1/nM)
$K_{4,1}$	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	0.0001359	(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	net specific growth rate of donor cells	0.0002567	(1/s)

Note that the extracellular concentration variable C and I are spared from inclusion in the vector $\mathbf{x}$. Eq. (9) provides an explanation of the different symbols in the foregoing differential Eqs. (1)–(7).

$$\dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y}) = \begin{bmatrix} k\,[K_{1,1}o + K_{1,2}(1-o)]\,(\dfrac{K_{3,5}\,q_a}{1+K_{3,5}\,q_a}) - (K_{4,1} + \mu)\,q_s \\[2ex] k\,[K_{1,1}o + K_{1,2}(1-o)]\,(\dfrac{1}{1+K_{3,5}\,q_a}) - (K_{4,2} + \mu)\,q_L \\[2ex] k\,[K_{1,3}o + K_{1,4}(1-o)] - k\,[K_{1,1}o + K_{1,2}(1-o)] \\ (\dfrac{K_{3,5}\,q_a}{1+K_{3,5}\,q_a}) - (K_{4,3} + \mu)\,q_a \\[2ex] K_{2,6}I - (K_{4,6} + \mu)\,i \\[2ex] K_{1,7} + \dfrac{K_{1,9}}{1+\exp[-K_{3,9}\,(q_L - K_{1,8})]} - (K_{4,9} + \mu)\,b \\ or \quad K_{1,5}q_L - (K_{4,9} + \mu)\,b \\[1ex] K_{2,8}C - (K_{4,8} + \mu)\,c \end{bmatrix} \quad (9)$$
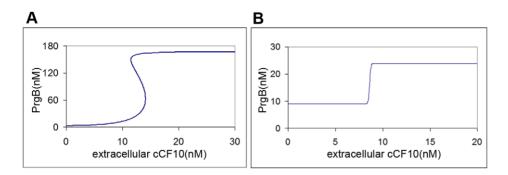
Figure 3. The steady-state behaviors. (**A**) System demonstrates bistability using the parameter values as shown in Table 2. (**B**) The possibility of bistability can be analytically excluded by setting the parameter value of $K_{4,2}=K_{4,1}$ independent of other parameters. For both (**A**) and (**B**), plasmid copy number k is equal to 5.
doi:10.1371/journal.pcbi.1002140.g003

where $o=\dfrac{i^r}{i^r+K_{3,8}\,c^r}$, $\mathbf{Y}=\begin{bmatrix}C\\I\end{bmatrix}$

Further, the differential equation for the mass balance of I remains to be identified (C is modeled as add-in with constant concentration). Towards this end, we define I_n as the *number* of inhibitor molecules in the extracellular space. Assuming that the fraction of extracellular volume to total volume as constant, we identify Eq.(10) for I_n as

$$\frac{dI_n}{dt}=K_{1,6}(q_s+q_L)v_d-K'_{2,6}I-K_{4,5}I_n \quad (10)$$

where v_d is the volume per cell. The first term on the right hand side of Eq. (10) represents the number of inhibitor molecules exiting the cell per unit time, the second their uptake rate by cells, and the third their degradation in the extracellular volume. For each cell, the uptake rate depends on the extracellular inhibitor concentration, so that the total uptake rate is proportional to the product of number of cells and the extracellular inhibitor concentration. Note that $K'_{2,6}$ is not a constant because the uptake of inhibitor occurs by active transport and its rate depends on PrgZ protein [37]. Assuming that the uptake rate is proportional to number of the PrgZ protein and that the latter is proportional to the volume of the cell, we rewrite $K'_{2,6}=K_{2,6}v_d$ which yields $K_{2,6}$ as a reasonable constant. Next, we define φ as the volume fraction of extracellular volume so that $1-\varphi$ would be volume fraction of the cell, from which we have $I_n=\varphi v_d I/(1-\varphi)=I\varphi V$ where V is the total volume. Substituting this into Eq. (10), using renamed variables, we obtain the differential equation for I as

$$\frac{dI}{dt}=K_{1,6}(q_s+q_L)\frac{(1-\varphi)}{\varphi}+K_{2,6}I\frac{(1-\varphi)}{\varphi}-(K_{4,5}+\frac{d\ln(\varphi V)}{dt})I \quad (11)$$

The last term on the right hand side can also be represented as $\dfrac{d\ln(\varphi V)}{dt}=\dfrac{d\ln V}{dt}=\mu$, a dilution term resulting from assuming that the volume fraction of cells remains constant as 0.5 so that an increase of cell volume by growth also results in an increase of extracellular volume to the same extent. Note that the above assumption about the extent of the extracellular volume has no influence on our conclusion. In this regard the reader is referred to the toy example where volume is modeled as constant.
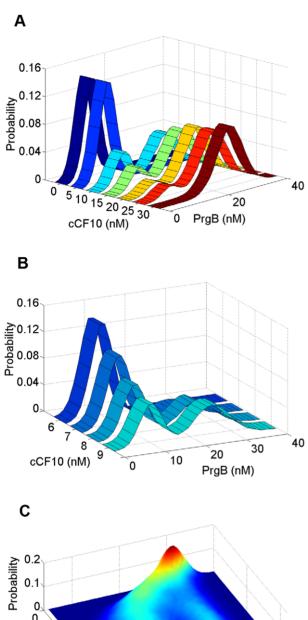
Population Balance Model (PBM) with deterministic intracellular behaviors
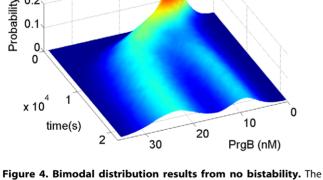
The effort of applying PBM on analyzing cell behaviors can be traced back to mid-twentieth century [1,38]. More recently, the number of publications applying PBM has notably increased on analyzing complex cellular behavior (e.g. Mantzaris [39]). Thus the behavior of an entire culture of microorganisms can be simulated by PBM in the form of a multivariate population distribution. A generic formulation of PBM is presented by Ramkrishna [15]. This formulation distinguishes a vector of internal coordinates $\mathbf{x}^T=(x_1,x_2,...x_d)$ and a vector of external coordinates $\mathbf{r}^T=(r_1,r_2,r_3)$; the former represents d different quantities associated with the cell and the latter denotes the position vector of the cell. Cells with the same coordinates are viewed as indistinguishable. The dynamics associated with the intracellular variables through cellular processes (including gene regulation) can be described by a rate vector $\dot{\mathbf{X}}$ containing the deterministic reaction rates in terms of internal coordinates $\mathbf{x}$. The vector $\mathbf{x}$ includes quantities such as cell mass, various intracellular components associated with gene expression, and so on. The vector $\mathbf{Y}$ is a vector of extracellular variables influencing the intracellular processes; which may include concentrations of nutrient, signaling molecules, inhibitor and so on. The motion of cells with respect to a fixed coordinate frame may be written as Eq. (13), where the vector $\dot{\mathbf{R}}$ describes the velocity of the cell which may be caused by the mixing of cells (for instance, in a planktonic growth situation) or zero when imbedded in a biofilm without motion.

$$\frac{d\mathbf{x}}{dt}=\dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y}) \quad (12)$$

$$\frac{d\mathbf{r}}{dt}=\dot{\mathbf{R}}(\mathbf{x},\mathbf{r},t) \quad (13)$$

where "$|\mathbf{Y}$" means "given $\mathbf{Y}$". The notation is used to recognize the influence of extracellular variables, $\mathbf{Y}$ on intracellular processes. We use $n(\mathbf{x},\mathbf{r},t)$ to denote the actual number density and denote its *expectation* by number density $f(\mathbf{x},\mathbf{r},t)$ [40], for a given set of internal coordinates $\mathbf{x}$, position coordinates $\mathbf{r}$ at a certain time t. The term density here refers to number of cells per unit volume of space of internal coordinates, $\mathbf{x}$ as well as that of external coordinates, $\mathbf{r}$. The existence of this density in physical

A

B

C

Figure 4. Bimodal distribution results from no bistability. The parameters used in this simulation are shown in Table 2 except for $K_{4,2} = K_{4,1} = 0.001$ (1/s). (**A**) The stationary distribution responding to different concentrations of extracellular pheromone, cCF10. As the concentration of pheromone increases, cell population migrates from state of low PrgB, viewed as off state, to state of high PrgB, viewed as on state. When pheromone concentration reaches 30 nM, all cells stay at on state. (**B**) The change of PrgB protein distribution for cCF10 concentration from 6 nM to 9 nM. (**C**) The dynamic behavior responding to constant concentration of pheromone, cCF10. This simulation is done with extracellular pheromone concentration maintained at 10 nM.
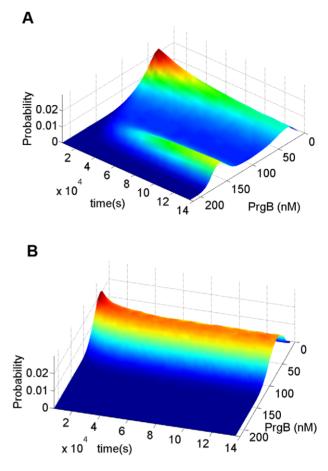doi:10.1371/journal.pcbi.1002140.g004

A

B

Figure 5. Population distribution for a system with bistability. Both (**A**) and (**B**) use the same parameters shown in Table 2. Extracellular pheromone concentration is maintained at 13.5 nM for both simulations. Only noise term of protein and peptides are taken into consideration [48,49]. Thirty thousand cells are used. (**A**) Outcome from single cell stochastic model described in Eq.(30). Due to stochasticity, bimodal distribution comes from initial unimodal distribution. Single cell stochastic model ignores the interaction between cells. Extracellular inhibitor is used only by the donor cell secreting it. (**B**) Outcome from PBM described in Eq.(29). The extracellular inhibitor is utilized by the whole population. The simulation outcome shows that the population effect leads to unimodal distribution.
doi:10.1371/journal.pcbi.1002140.g005

space must be recognized even when there is no explicit dependence of cell numbers with position. The population balance equation is a state-specific balance due to various processes such as by conjugation, cell division, and so on. It may be written as

$$\frac{\partial f(\mathbf{x},\mathbf{r},t)}{\partial t} + \nabla_{\mathbf{x}}\cdot\dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})f(\mathbf{x},\mathbf{r},t) + \nabla_{\mathbf{r}}\cdot\dot{\mathbf{R}}(\mathbf{x},\mathbf{r},t)f(\mathbf{x},\mathbf{r},t) = h(\mathbf{x},\mathbf{r},t) \quad (14)$$

The function $h(\mathbf{x},\mathbf{r},t)$ in Eq. (14) represents the net (number) rate of production of cells of state $\mathbf{x}$ at a particular location $\mathbf{r}$ and time t. Note in particular that h, although represented as a simple function of its arguments, acquires its dependence on them through being a *functional* of the number density $f(\mathbf{x},\mathbf{r},t)$. Eq. (14) is coupled with a conservation equation written for $\mathbf{Y}$, for which we define $\dot{\mathbf{Y}}(\mathbf{x})$ as the rate at which a cell of state $\mathbf{x}$ "consumes" or "contributes" to the extracellular variables in $\mathbf{Y}$. The extracellular reaction rate is described by $\gamma(\mathbf{Y})$. Noting that $\mathbf{Y}(\mathbf{r},t)$ is a function
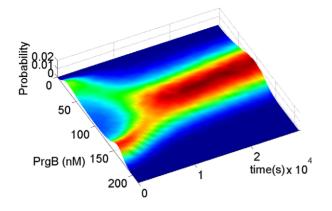
Figure 6. Population distribution, started with bimodal distribution, merging into a unimodal distribution due to population effect. Bimodal distribution generated from single cell stochastic model with extracellular pheromone concentration equal to 13.5 nM served as initial distribution of PBM. Simulation was done with plasmid copy number equal to 5 and parameter values listed on Table 2. Population effect of extracellular inhibitor causes two modes to merge into one.
doi:10.1371/journal.pcbi.1002140.g006

of position and time, the conservation equation for $\mathbf{Y}$ is described in Eq. (15) where $\mathbf{N_Y}$ is the total flux of the various components in $\mathbf{Y}$ including convective as well as diffusive transport.

$$\frac{\partial \mathbf{Y}}{\partial t} + \nabla \cdot \mathbf{N_Y} = -\int \dot{\mathbf{Y}}(\mathbf{x})f(\mathbf{x},\mathbf{r},t)d\mathbf{x} + \gamma(\mathbf{Y}) \qquad (15)$$

The population balance model is defined by Eqs. (14) and (15), properly supplemented by initial and boundary conditions.

PBM with stochastic intracellular behaviors

We next incorporate the intracellular stochastic behavior of gene regulation into the population balance model described in Eqs. (14) and (15). The formulation of population balance equations has been presented by Ramkrishna [15] when the internal state is a stochastic process described by the stochastic differential equation

$$d\mathbf{x} = \dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})dt + \mathbf{B}(\mathbf{x}|\mathbf{Y})d\mathbf{W}(t) \qquad (16)$$

where $\mathbf{B}(\mathbf{x}|\mathbf{Y})$ is the term that determines the magnitude of stochastic fluctuations (refer to Gillespie's Chemical Langevin equation [41]) of signal transduction reactions on the associated intracellular variables; $d\mathbf{W}(t)$, a vector, represents the increment of a standard Wiener process (during the time interval dt). We further note that the SDE are based on Ito formulation. Also, its equivalent Fokker-Planck equation can be written as Eq.(17).

$$\frac{\partial P(\mathbf{x},t)}{\partial t} = -\nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})P(\mathbf{x},t) + \frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}(\mathbf{x}|\mathbf{Y})\mathbf{B}(\mathbf{x}|\mathbf{Y})^T P(\mathbf{x},t) \quad (17)$$

where " $:$ " represent double dot product so $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}(\mathbf{x}|\mathbf{Y})\mathbf{B}(\mathbf{x}|\mathbf{Y})^T$ can also be written as $\sum_i \sum_j \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\mathbf{B}(\mathbf{x}|\mathbf{Y})\mathbf{B}(\mathbf{x}|\mathbf{Y})^T$.

The population balance equation with position coordinate can be written as Eq.(18)

$$\frac{\partial f(\mathbf{x},\mathbf{r},t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})f(\mathbf{x},\mathbf{r},t) + \nabla_{\mathbf{r}} \cdot \dot{\mathbf{R}}(\mathbf{x},\mathbf{r},t)f(\mathbf{x},\mathbf{r},t)$$
$$= \frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}(\mathbf{x}|\mathbf{Y})\mathbf{B}(\mathbf{x}|\mathbf{Y})^T f(\mathbf{x},\mathbf{r},t) + h(\mathbf{x},\mathbf{r},t) \qquad (18)$$

which is the number balance of cells of state $\mathbf{x}$ accounting for stochastic changes in internal coordinates as defined by the Ito SDE, Eq. (16). The derivation of this equation is available in Ramkrishna [15]. Next, we consider the cells to be completely (uniformly) mixed so that spatial coordinates may be eliminated. The resulting population balance equation is given by

$$\frac{\partial f(\mathbf{x},t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})f(\mathbf{x},t)$$
$$= \frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}(\mathbf{x}|\mathbf{Y})\mathbf{B}(\mathbf{x}|\mathbf{Y})^T f(\mathbf{x},t) + h(\mathbf{x},t) \qquad (19)$$

Eq. (19) is coupled to a version of Eq. (15) in environmental variables $\mathbf{Y}$ modified to account for a well-mixed system given by

$$\frac{d\mathbf{Y}}{dt} = -\int E\dot{\mathbf{Y}}(\mathbf{x},t)f(\mathbf{x},t)d\mathbf{x} + \gamma(\mathbf{Y}) \qquad (20)$$

where $E\dot{\mathbf{Y}}(\mathbf{x},t)$ is the expected rate of consumption of extracellular variables by cells of state $\mathbf{x}$. Note that $\dot{\mathbf{Y}}(\mathbf{x},t)$ is stochastic in view of the single cell behavior being stochastic so that the cumulative contribution from a large collection of cells to the environment is deterministic given by $E\dot{\mathbf{Y}}(\mathbf{x},t)$. Eqs. (19) and (20) must be considered with initial and boundary conditions. Note that number density function, $f(\mathbf{x},t)$, satisfies "natural boundary conditions" (i.e., vanishing of the function and its gradient at infinity). With this background, we are in a position to consider the population balance model for the system of interest, viz., conjugation of plasmid pCF10 system.

PBM for conjugation of plasmid pCF10 system

Since the conjugative response can be influenced by the number of copies of pCF10 plasmid, we include plasmid copy number as a discrete internal coordinate for the cell in view of its effect on cell dynamics as the number of plasmids in each cell becomes important. As a result, the number density Eq. (3) is further embellished with a discrete variable representing copy number. We define the expected number density of cells $f_k(\mathbf{x},t)$, with internal state $\mathbf{x}$ and plasmid copy number k, where k varies between 1 and $k_{\max}$. Kinetics of gene expression is denoted as $\dot{\mathbf{X}}_k(\mathbf{x}|\mathbf{Y})$ to account for the effect of plasmid copy number. Cell division rate, denoted by μ, is assumed to be a constant. A more complicated situation of cell growth, which is not taken into consideration in this study, including the dependence of μ on intracellular stochastic state x can be found in Tanase-Nicola's work [42]. However, this cited work suffers from neglecting interaction between cells through population effects on the environment. The random partitioning of plasmid between daughter cells is denoted by $p_{k|j}$ where j is the copy number of the dividing parent and k is that of the daughter cell. Plasmids replication is assumed to occur instantaneously prior to cell division. The population balance equation with stochastic intracellular behavior for this case can be written as

$$\frac{\partial f_k(\mathbf{x},t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_k(\mathbf{x}|\mathbf{Y})f_k(\mathbf{x},t) =$$
$$\frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}_k(\mathbf{x}|\mathbf{Y})\mathbf{B}_k(\mathbf{x}|\mathbf{Y})^T f_k(\mathbf{x},t) + \qquad (21)$$
$$-\mu f_k(\mathbf{x},t) + 2\sum_j f_j(\mathbf{x},t)\mu p_{k|j}, \qquad k = 1,2,\dots,k_{\max}$$

Eq. (21) represents a number balance of cells of state $\mathbf{x}$ with copy

number k. The first term on the left hand side represents the rate at which such cells accumulate, the second denotes the net flux by "convective" transport (in internal coordinate space), while the first term on the right hand side represents the mean fluctuation due to random effects, the second the loss of cells by division and the last term the gain of cells of copy number k by division of other cells. The factor 2 accounts for doubling of the population by binary division. The intracellular variables of parent and daughter cells are considered to be same. The extracellular components include only pheromone and inhibitor. We do not concern ourselves with the resistance transfer process in this paper as our focus is only on the expression of the protein PrgB in the donor cells. We enunciate three further model assumptions: (i) the population density is maintained constant in the growing population by appropriate dilution; in other words, the system volume is allowed to expand suitably. This assumption is only made to provide for a true steady state in the protein level distribution. A similar assumption is also made in the single cell analysis. (ii) constant extracellular pheromone concentration, which implies that we need only consider the component I for the dynamics of the vector $\mathbf{Y}$, and (iii) that inhibitor is produced and secreted directly into the environment through intracellular reaction as described by Eq. (11).

The mass balance of extracellular inhibitor is adopted from Eq. (20) acknowledging any copy number dependence of the rate of consumption of extracellular variables by cells. A macroscopic mass balance for the extracellular variables, based on assumption of perfect mixing, is given by.

$$\frac{dI}{dt} = \sum_k \int E[K_{1,6}(x_1+x_2) - K_{2,6}I] \frac{v_d}{\varphi} f_k(\mathbf{x},t)d\mathbf{x} - (K_{4,5}+\mu)I \quad (22)$$

In Eq. (22), we have used the renamed concentration variables x_1 and x_2 in place of q_s and q_L respectively. It is also worth noting that the dilution term on the second term on the right hand side comes about from the assumption that the system volume is allowed to expand to keep the population density constant. The motivation for this assumption, as pointed out earlier, is the attainment of a true steady state in the population with respect to protein levels as in the single cell analysis. A schematic illustration of PBM described by Eqs. (21) and (22) is shown in Figure 2**B**.

We restrict ourselves in this work to the case where plasmid copy number of the daughter cells is the same as that of the parent cell. In other words, if a cell of copy number k divides, just prior to division, the plasmids double in number and are equally shared by the two daughter cells. Of course more general cases are admissible in the model framework which can account for plasmid replication and uneven distribution among daughter cells. The specific assumptions in this work will, however, suffice for demonstration of population balance modeling. The even partitioning of plasmids among daughter cells in a population of *uniform* copy number distribution is described in Eq. (23).

$$p_{k|j} = \delta_{jk} = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases} \quad (23)$$

Using Eq. (23) converts the population balance Eq. (21) into Eq. (24) shown below with $k = 1,2,...,9$, which yields an average copy number of 5 close to experimental observation.

$$\frac{\partial f_k(\mathbf{x},t)}{\partial t} + \nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_k(\mathbf{x}|\mathbf{Y}) f_k(\mathbf{x},t) =$$
$$\frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}_k(\mathbf{x}|\mathbf{Y})\mathbf{B}_k(\mathbf{x}|\mathbf{Y})^T f_k(\mathbf{x},t) + \mu f_k(\mathbf{x},t) \quad (24)$$

The $\mathbf{B}_k(\mathbf{x}|\mathbf{Y})$ of Eq. (24) is identified below

$$\mathbf{B}_k(\mathbf{x}|\mathbf{Y}) = \begin{bmatrix} [0]_{3\times 3} & [0]_{3\times 6} \\ [0]_{3\times 3} & [\mathbf{B}_p]_{3\times 6} \end{bmatrix}$$

where $[\mathbf{B}_p]_{3\times 6} = \begin{bmatrix} B_{p,11} & B_{p,12} & 0 & 0 & 0 & 0 \\ 0 & 0 & B_{p,23} & B_{p,24} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{p,35} & B_{p,36} \end{bmatrix}$

$$B_{p,11} = \sqrt{\frac{K_{2,6}I}{v_d}},\ B_{p,12} = \sqrt{\frac{K_{4,6}i}{v_d}},\ B_{p,23} =$$

$$\left\{ K_{1,7} + \frac{K_{1,9}}{1 + \exp[-K_{3,9}(q_L - K_{1,8})]} \right\}^{1/2} v_d^{-1/2} or \sqrt{\frac{K_{1,5}q_L}{v_d}}$$

$$B_{p,24} = \sqrt{\frac{K_{4,9}b}{v_d}},\ B_{p,35} = \sqrt{\frac{K_{2,8}C}{v_d}},\ B_{p,36} = \sqrt{\frac{K_{4,8}c}{v_d}}$$

and the $\dot{\mathbf{X}}_k(\mathbf{x}|\mathbf{Y})$ is exactly the same as $\dot{\mathbf{X}}(\mathbf{x}|\mathbf{Y})$ which is identified in Eq.(9). The reason for using index k is to clearly state that PBM account for population heterogeneity of plasmid copy number. Note that Eq. (24) should be coupled with Eq. (22). The overall expected number density, denoted $f_T(\mathbf{x},t)$, can be obtained in Eq. (25)

$$f_T(\mathbf{x},t) = \sum_k f_k(\mathbf{x},t) \quad (25)$$

The solution method for PBM

The difficulty involved in solving PBM with stochastic intracellular behaviors comes from the natural boundary condition (i.e., vanishing of the function and its gradient at infinity). To handle this problem, we transform the population balance equation (24) into a Fokker-Planck equation using the transformation shown in Eq. (13) to obtain Eq. (14). The integral over $\mathbf{x}$ coordinates is unity because N_{ko} represents the initial number density of cells with copy number k.

$$\tilde{f}_k(\mathbf{x},t) = \frac{1}{N_{ko}} f_k(\mathbf{x},t) e^{-\mu t} \quad (26)$$

$$\frac{\partial \tilde{f}_k(\mathbf{x},t)}{\partial t} = -\nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_k(\mathbf{x}|\mathbf{Y})\tilde{f}_k(\mathbf{x},t) +$$
$$\frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}_k(\mathbf{x}|\mathbf{Y})\mathbf{B}_k(\mathbf{x}|\mathbf{Y})^T \tilde{f}_k(\mathbf{x},t) \quad (27)$$

Eq. (27) is a Fokker-Planck type differential equation whose solution represents the probability distribution for the stochastic process defined by the Ito SDE (refer to Eq. (16)). Thus the sample-pathwise solution to Ito SDE will provide an alternative route to calculate expectations of all quantities associated with the stochastic process $\mathbf{x}$, including the quantity $E\dot{\mathbf{Y}}_k(\mathbf{x},t)$, for substitution into Eq. (22). We solved Ito SDE relating to Eq. (27) by using the Euler algorithm [43]. The computation proceeds in a stepwise manner for each discrete interval to keep abreast of environmental variables.

Results

We first compare the single cell model approach with that of population balances, delineating the differences between them. The single cell approach would use Ito SDE with variations including extracellular variables, i.e., the inhibitor in the present context, (as the inducer molecule concentration is assumed to be held constant). The secretion of inhibitor to the exterior and subsequent transport back into the cell will enter the model as a singular experience of the cell in question. Thus the environmental inhibitor concentration will display dynamics as a result of interaction with the single cell. On the other hand, in the population balance model, stochastic behavior of numerous individual cells will provide different secretions of the inhibitor to the environment which, because of mixing effects, would change the environment of all cells in the population. This will undoubtedly produce different dynamics of the system. To further elucidate the difference between the two approaches, we note that the environmental inhibitor I in the population balance model changes as a result of the cumulative (expected) exchange with the population whereas the single cell model will account only for the (stochastic) exchange rate with the cell's environment.

Bimodal distribution from plasmid unevenly distributing to daughter cells

Although bistability is featured by Eqs. (1)–(10) for the range of parameter values shown in Table 2 (Figure 3**A**), bistability can be (analytically) excluded by forcing parameter $K_{4,2}$, the degradation rate of Q_L mRNA, equal to $K_{4,1}$, the degradation rate of Q_s mRNA, regardless of the values of other parameters (Figure 3**B**). The detailed derivation is shown in the Text S1.

The parameters used in this simulation are shown in Table 2, except $K_{4,2}=K_{4,1}=0.001$ (1/s). Stochasticity is restricted to protein alone to reduce computational time. Forty-five thousand cells are used. The outcome is shown in Figure 4. Results of stationary distribution responding to different pheromone concentration are shown in Figure 4**A** and 4**B**. At zero and low pheromone concentration, all cells are at off mode. With increasing pheromone concentration, the cell population gradually migrates from mode of low PrgB, viewed as off state, to mode of high PrgB, viewed as on state. Finally, when the pheromone concentration exceeds 30 nM, all cells stay at on state. The transcription rate of various RNA species is directly proportional to the plasmid copy number. The difference between cells of various plasmid copy numbers causes population heterogeneity resulting in a distribution across the population. The results of dynamic behavior are shown in Figure 4**C**. This simulation is done for extracellular pheromone concentration equal to 10 nM. The initial distribution is obtained by simulating cells with no pheromone added. Population from initial unimodal distribution finally develops into a bimodal distribution. Figure 4 demonstrates that bistability is not necessary for a bimodal distribution and that the bimodal distribution arises directly out of population heterogeneity.

Resolving bistability versus bimodal distribution

If there is no plasmid copy number distribution, the population balance model can then be written as:

$$\frac{\partial f_5(\mathbf{x},t)}{\partial t} = -\nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_5(\mathbf{x}|\mathbf{Y}) f_5(\mathbf{x},t) + \frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}_5(\mathbf{x}|\mathbf{Y})\mathbf{B}_5(\mathbf{x}|\mathbf{Y})^T f_5(\mathbf{x},t) + \mu f_5(\mathbf{x},t)$$

(28)

$\dot{\mathbf{X}}_5(\mathbf{x}|\mathbf{Y})$ and $\mathbf{B}_5(\mathbf{x}|\mathbf{Y})$ indicate Eq. (28) describing system with plasmid copy number equal to 5.

With the solution method described in Methods we convert the population balance Eq. (28) into

$$\frac{\partial \tilde{f}_5(\mathbf{x},t)}{\partial t} = -\nabla_{\mathbf{x}} \cdot \dot{\mathbf{X}}_5(\mathbf{x}|\mathbf{Y}) \tilde{f}_5(\mathbf{x},t) + \frac{1}{2}\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} : \mathbf{B}_5(\mathbf{x}|\mathbf{Y})\mathbf{B}_5(\mathbf{x}|\mathbf{Y})^T \tilde{f}_5(\mathbf{x},t)$$

(29)

Notice that above equation should be coupled with environmental equation of extracellular inhibitor Eq.(22).

The single cell stochastic model may be written as

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

(30)

Where index ie indicates both intracellular and extracellular variables are involved. For reasons that have already been elucidated earlier, Eqs. (29) and (30) are not the same. In Eq. (29), the vector $\mathbf{x}$ is different from the vector $\mathbf{x}_{ie}$ in Eq. (30) because the latter also includes the extracellular inhibitor as a stochastic variable.

The outcome of the simulation is shown in Figures 5**A** and 5**B**. While the outcome of Eq.(30) shows a bimodal distribution that of Eq.(29) shows a unimodal distribution, thus indicating the strong impact of the population on the behavior of individual cells.

In order to exclude the possibility that Figure 5**B** was a result of insufficient simulation time to develop into a bimodal distribution, a simulation was conducted with PBM using initial distribution as a bimodal distribution calculated from the single cell Fokker-Planck equation. The simulation outcome, Figure 6, shows a result consistent with Figure 5**B**. A bimodal distribution under the influence of the population effect finally merges into one mode.

Toy example

The purpose of this example, shown in Figure 7, is to elucidate the key elements of the more complicated model of the pCF10 System. To simplify the discussion, we use symbols to denote not only the molecular species but their concentrations. The precursor of the signal molecule, denoted as X_{so}, needs membrane protein, m_p, to mature into intracellular signal molecule, x_s. Two kinds of gene, xp gene and xi gene, encode product and inhibitor, respectively. As the signal molecule dominates, the transcription rate of xp gene is high and that of xi gene is low. On the other hand, inhibitor favors xi gene instead of xp gene by consuming signal molecule. Further, defining the intracellular concentration of inhibitor as x_{in}, the concentration of product as x_p, the extracellular concentration of inhibitor as X_{ex}, and letting v be the volume per cell, we formulate the mass balance equations for the single cell as

$$\frac{dx_s}{dt} = K_{ss}X_{so}m_p - (K_{ds}+v)x_s$$

(31)

$$\frac{dx_p}{dt} = K_{Gp}G_{pa} - (K_{dp}+v)x_p$$

(32)

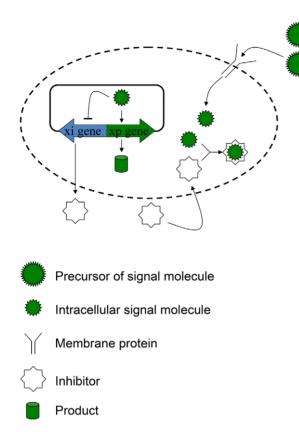$$\frac{dx_{in}}{dt} = K_{iI}X_{ex} - (K_{di}+v)x_{in}$$

(33)

Figure 7. Reaction network for gene regulation in toy example. After uptake by membrane protein, the precursor of signal molecule matures into signal molecule. When signal molecule binds to DNA, it favors the expression of xp gene encoding product and represses the transcription of xi gene which encodes the inhibitor. The binding reaction of inhibitor to signal molecule is fast and irreversible so either signal molecule or inhibitor dominates.
doi:10.1371/journal.pcbi.1002140.g007

Table 3. Parameters of toy example.

Reaction Constant	Name	Value	
K_{ss}	generation rate constant of intracellular signal molecule	1	(1/arb.u. min)
K_{ds}	degradation rate constant of intracellular signal molecule	0.994	(1/min)
K_{Gp}	generation rate constant of product	50	(arb.u./ volume min)
K_{dp}	degradation rate constant of intracellular signal molecule	0.994	(1/min)
K_{iI}	uptake rate of inhibitor	1	(1/min)
K_{di}	degradation rate constant of intracellular signal molecule	0.994	(1/min)
K_{GI}	generation rate constant of extracellular inhibitor	80	(arb.u./ volume min)
m_p	membrane protein	6	(arb.u.)
υ	net specific growth rate of cells	0.006	(1/min)
v	volume per cell	1E-15	(volume)
$K_I^{on/off}$	"configuration factor" of xi gene	0.5	-
$K_p^{on/off}$	"configuration factor" of xp gene	2.4	-

doi:10.1371/journal.pcbi.1002140.t003

signal molecule with m_p which is assumed to be controlled by a certain gene and maintained constant in each cell; $K_{Gp}G_{pa}$ is the generation rate of product where K_{Gp} is a basic rate multiplied by a "configuration factor", G_{pa}; $K_{iI}X_{ex}$ describes the uptake rate of inhibitor where X_{ex} is the extracellular inhibitor concentration; $K_{ds}x_s$, $K_{dp}x_p$ and $K_{di}x_{in}$ are degradation terms and υ is cell growth rate. For extracellular inhibitor, the generation term includes a basic transcription rate, K_{GI}, multiplied by a "configuration factor", G_{Ia}. Because extracellular inhibitor is the accumulated result from all the cells, Eq.(34) further accounts for number density, N. Note that although steady state exists for deterministic equation, there is no true steady state for stochastic model or stochastic gene regulation incorporating PBM. However, the effect from the increment of N is small enough to consider the system as in pseudo steady state (refer to Figure 8 and 9).

The PBM with stochastic intracellular behaviors of this system is shown in Eq. (35) with environmental equation, Eq. (36). Following the method introduced in the section "The solution method for PBM", Eq.(35) can be transferred into Eq. (37).

$$\frac{dX_{ex}}{dt} = NvK_{GI}G_{Ia} - NvK_{iI}X_{ex} \tag{34}$$

where G_{pa} and G_{Ia} describe how DNA configurations change generation rate and are defined as

$$G_{pa} = \begin{cases} K_p^{on/off} & \\ 1 & \end{cases} \quad when \quad \begin{matrix} x_{in} < x_s \\ x_s \le x_{in} \end{matrix},$$

$$G_{Ia} = \begin{cases} K_I^{on/off} & \\ 1 & \end{cases} \quad when \quad \begin{matrix} x_{in} < x_s \\ x_s \le x_{in} \end{matrix}$$

and the values of parameters are identified in Table 3. Note that $K_p^{on/off}$ is greater than one because the signal molecule favors the transcription of xp gene and $K_I^{on/off}$ is less than one as the signal molecule prevents the transcription of xi gene. Eqs. (32) and (34), together with G_{pa} and G_{Ia} as defined above, imply that binding of intracellular inhibitor to signal molecule is an irreversible reaction with suitably large rates so that only the signal molecule or inhibitor dominates the system. The toy example is composed of seven reactions with the system at constant volume and the volume of cells negligible compared to that of the system. For intracellular variables, $K_{ss}X_{so}m_p$ describes the generation rate of intracellular

$$\frac{\partial f_t(\mathbf{x}_t, t)}{\partial t} = -\nabla_{\mathbf{x}_t} \cdot \dot{\mathbf{X}}_t(\mathbf{x}_t | \mathbf{Y}_t) f_t(\mathbf{x}_t, t)$$
$$+ \frac{1}{2}\nabla_{\mathbf{x}_t}\nabla_{\mathbf{x}_t} : \mathbf{B}_t(\mathbf{x}_t | \mathbf{Y}_t)\mathbf{B}_t(\mathbf{x}_t | \mathbf{Y}_t)^T f_t(\mathbf{x}_t, t) + \upsilon f_t(\mathbf{x}_t, t) \tag{35}$$

$$\frac{dX_{ex}}{dt} = \int (K_{GI}G_{Ia} - K_{iI}X_{ex}) v f_t(\mathbf{x}_t, t)d\mathbf{x}_t \tag{36}$$

$$\frac{\partial \tilde{f}_t(\mathbf{x}_t, t)}{\partial t} = -\nabla_{\mathbf{x}_t} \cdot \mathbf{X}_t(\mathbf{x}_t | \mathbf{Y}_t)\tilde{f}_t(\mathbf{x}_t, t)$$
$$+ \frac{1}{2}\nabla_{\mathbf{x}_t}\nabla_{\mathbf{x}_t} : \mathbf{B}_t(\mathbf{x}_t | \mathbf{Y}_t)\mathbf{B}_t(\mathbf{x}_t | \mathbf{Y}_t)^T \tilde{f}_t(\mathbf{x}_t, t) \tag{37}$$
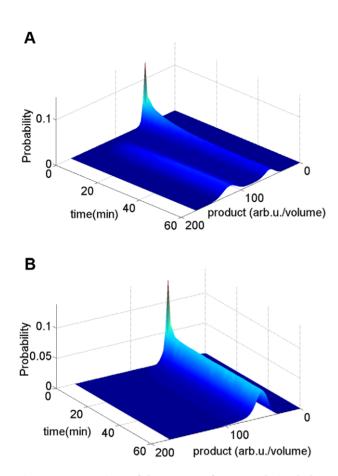
A

B

Figure 8. Comparison of the outcome from population balance model and single cell stochastic model in toy example. Thirty thousand cells are used with constant concentration, 10 arbitrary unit/volume, of add-in signal molecular precursor. (**A**) Bimodal distribution is observed from single cell stochastic model. (**B**) Unimodal distribution is observed from population balance model.
doi:10.1371/journal.pcbi.1002140.g008

where $\tilde{f}_t(\mathbf{x}_t,t)=\frac{1}{N_o}f_t(\mathbf{x}_t,t)e^{-\upsilon t}$ and N_o is the initial number density of cells. Each of the terms in Eq. (37) are identified below

$$\mathbf{x}_t^T=\begin{bmatrix} x_s & x_p & x_{in} \end{bmatrix}$$

$$\dot{\mathbf{X}}_t(\mathbf{x}_t|\mathbf{Y}_t)=\begin{bmatrix} K_{ss}X_{so}m_p-(K_{ds}+\upsilon)x_s \\ K_{Gp}G_{pa}-(K_{dp}+\upsilon)x_p \\ K_{iI}X_{ex}-(K_{di}+\upsilon)x_{in} \end{bmatrix} \text{ where } \mathbf{Y}_t=\begin{bmatrix} X_{so} \\ X_{ex} \end{bmatrix}$$

$$\mathbf{B}_t(\mathbf{x}_t|\mathbf{Y}_t)=\begin{bmatrix} \sqrt{\frac{K_{ss}X_{so}m_p}{\upsilon}} & \sqrt{\frac{K_{ds}x_s}{\upsilon}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{K_{Gp}G_{pa}}{\upsilon}} & \sqrt{\frac{K_{dp}x_p}{\upsilon}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\frac{K_{iI}X_{ex}}{\upsilon}} & \sqrt{\frac{K_{di}x_{in}}{\upsilon}} \end{bmatrix}$$

With the parameter values shown in Table 3, the deterministic Eq. (31)–(34) are featured with bistability and the outcome of single cell stochastic model, Eq.(38), shows corresponding bimodal distribution, Figure 8**A**. However, the prediction of PBM shows unimodal
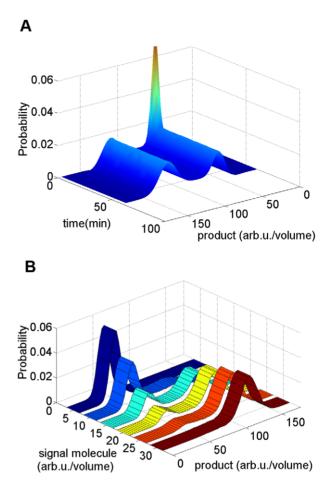
A

B

Figure 9. Bimodal distribution from no bistability. Thirty five thousand cells are used for simulation. (**A**) Although there is no true steady state, the effect from the increment of N on standard deviation of the product protein distribution is small enough to consider the system as pseudo steady state. (**B**) The plot of product protein distribution with different add-in signal molecular precursor.
doi:10.1371/journal.pcbi.1002140.g009

distribution, Figure 8**B**. The same as mentioned in pCF10 system, such qualitative difference is raised from the interaction between cells.

$$\frac{\partial P_t(\mathbf{x}_{ts},t)}{\partial t}=-\nabla_{\mathbf{x}_{ts}}\cdot\dot{\mathbf{X}}_{ts}(\mathbf{x}_{ts})P_{ts}(\mathbf{x}_{ts},t)$$
$$+\frac{1}{2}\nabla_{\mathbf{x}_{ts}}\nabla_{\mathbf{x}_{ts}}:\mathbf{B}_{ts}(\mathbf{x}_{ts})\mathbf{B}_{ts}(\mathbf{x}_{ts})^T P_t(\mathbf{x}_{ts},t) \quad (38)$$

Each of the terms in Eq. (38) are identified below.

$$\mathbf{x}_{ts}^T=\begin{bmatrix} x_s & x_p & x_{in} & X_{ex} \end{bmatrix}$$

$$\dot{\mathbf{X}}_{ts}(\mathbf{x}_{ts})=\begin{bmatrix} K_{ss}X_{so}m_p-(K_{ds}+\upsilon)x_s \\ K_{Gp}G_{pa}-(K_{dp}+\upsilon)x_p \\ K_{iI}X_{ex}-(K_{di}+\upsilon)x_{in} \\ NK_{GI}G_{Ia}-NK_{iI}X_{ex} \end{bmatrix}$$

$$\mathbf{B}_{ts}(\mathbf{x}_{ts}) = \begin{bmatrix} \sqrt{\frac{K_{ss}X_{so}m_p}{v}} & \sqrt{\frac{K_{dc}x_s}{v}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{K_{Gp}G_{pa}}{v}} & \sqrt{\frac{K_{dp}x_p}{v}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\frac{K_{iI}X_{ex}}{v}} & \sqrt{\frac{K_{di}x_{in}}{v}} & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\frac{NK_{iI}X_{ex}v}{V_t}} & 0 & \sqrt{\frac{NK_{GI}G_{Ia}v}{V_t}} \end{bmatrix}$$

where V_t is the system volume.

Next we demonstrate bimodal distribution from no bistability. Note that the bistability can be excluded by forced G_{Ia} equal to one. In other words, when $K_I^{on/off} = 1$, there is no bistability. Instead of $m_p = 6$, we assign $m_p = 3,4,...,9$ for each subpopulation and each of them starts with the same biomass concentration. The simulation outcome is shown in Figure 9. The bimodal distribution comes from population heterogeneity.

Discussion

In this paper, we have investigated how population effect changes the behaviors of a culture of cells and demonstrated that the single cell approach does not account for the effects of population heterogeneity and is therefore at risk of producing erroneous results. For this application, we demonstrated that bistability is neither necessary nor sufficient for bimodal distribution.

In incorporating stochastic effects, we have relied on a continuous description of the intracellular variables by SDE. As the stochasticity arises from the randomness of chemical transformations of a small number of reacting molecules, the variables are essentially discrete. The SSA uses the chemical Master equation which is based on discrete variables, whereas the SDE approach has found various justifications in the literature. For example, van Kampen [8] uses system size expansion to obtain continuous descriptions of the stochastic variables. Although the continuous description is known to be appropriate for relatively larger number of molecules, publications exist in the literature that demonstrate the usefulness of continuous description of discrete variables for as low as even ten particles [43]. Estimates of the expressed protein level in the system of interest here range in the thousands in the on state and roughly in the range 14–35 particles per cell in the off state. Arguments for these estimates are included in the Text S1.

Hence the adoption of SDE may be regarded as appropriate for this application. In addition, the analysis of populations involves several cells of small variations about a given state so that intracellular behavior averaged among them qualifies for the SDE approach even more than in an isolated single cell.

In the section of resolving bistability versus bimodal distribution, if cells act independently from each other, bimodal distribution can be observed. However, cells change distribution from bimodal to unimodal due to population effect. In general, planktonic cells diffuse freely in the culture and such an isolated situation is hardly reached, but for a cell immobilized by extracellular matrix, such as biofilms [44], an isolated situation may be possible [45]. This study simulates the response of *E. faecalis* donor cells, harboring plasmid pCF10, to pheromone concentration. At low concentrations of pheromone as found in the natural situation [46,47], for a perfectly mixed system all cells are predicted to be at off-state as shown in Figure 5**B**. However, for unmixed systems, non-uniformity of inhibitor concentration can lead to some cells being at the off-state, others at the on-state which together make up a bimodal distribution for the integrated population. This provides a possible mechanism for the observation of bimodal behaviors under naturally occurring conditions such as biofilms involved in dissemination of antibiotic resistance as has been shown in recent work (Cook 2010, unpublished work). Our effort in this paper has been to show that a cell in a population can behave in a significantly different manner as its environment is altered by the concerted action of other cells. A natural follow-up to this paper is the modeling of the transfer of drug resistance accounting for the presence of donor and recipient cells in different environments for which population balances will indeed provide the proper framework.

Supporting Information

Text S1 Supporting information includes the proof of how bistability is excluded from pCF10 system, qualitative consistency between model predictions of pCF10 system and experimental observation, and the estimation of protein number per cell. (DOC)

Acknowledgments

Author Contributions

References

1. Fredrickson A, Ramkrishna D, Tsuchiya H (1967) Statistics and dynamics of procaryotic cell populations. Math Biosci 1: 327–374.
2. Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, et al. (2004) Programmable cells: Interfacing natural and engineered gene networks. Proc Natl Acad Sci U S A 101: 8414–8419.
3. Tian TH, Burrage K (2006) Stochastic models for regulatory networks of the genetic toggle switch. Proc Natl Acad Sci U S A 103: 8372–8377.
4. Gillespie DT (1977) Exact Stochastic Simulation of Coupled Chemical-Reactions. J Phys Chem 81: 2340–2361.
5. Shah B, Ramkrishna D, Borwanker J (1977) Simulation of particulate systems using the concept of the interval of quiescence. AIChE J 23: 897–904.
6. Gardiner CW (1997) Handbook of stochastic methods for physics chemistry and natural science; Haken PDDhcmH, ed. Springer-Verlag Berlin Heidelberg New.
7. Risken H (1996) The Fokker-Planck equation: Methods of solution and applications Springer Verlag.
8. Van Kampen NG (2007) Stochastic Processes in Physics and Chemistry Elsevier.
9. Kepler TB, Elston TC (2001) Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. Biophys J 81: 3116–3136.
10. Maamar H, Dubnau D (2005) Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop. Mol Microbiol 56: 615–624.
11. Yildirim N, Mackey MC (2003) Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data. Biophys J 84: 2841–2851.
12. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339–342.
13. Ferrell JE (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Curr Opin Cell Biol 14: 140–148.
14. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A (2004) Multistability in the lactose utilization network of Escherichia coli. Nature 427: 737–740.
15. Ramkrishna D (2000) Population balances: Theory and applications to particulate systems in engineering Academic Press San Diego.
16. Kozlowicz BK, Shi K, Gu ZY, Ohlendorf DH, Earhart CA, et al. (2006) Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. Mol Microbiol 62: 958–969.
17. Hirt H, Manias DA, Bryan EM, Klein JR, Marklund JK, et al. (2005) Characterization of the pheromone response of the Enterococcus faecalis conjugative plasmid pCF10: Complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction. J Bacteriol 187: 1044–1054.

18. Chuang ON, Schlievert PM, Wells CL, Manias DA, Tripp TJ, et al. (2009) Multiple Functional Domains of Enterococcus faecalis Aggregation Substance Asc10 Contribute to Endocarditis Virulence. Infect Immun 77: 539–548.

19. Karmakar R, Bose I (2004) Graded and binary responses in stochastic gene expression. Phys Biol 1: 197–204.

20. Karmakar R, Bose I (2007) Positive feedback, stochasticity and genetic competence. Phys Biol 4: 29–37.

21. Pirone JR, Elston TC (2004) Fluctuations in transcription factor binding can explain the graded and binary responses observed in inducible gene expression. J Theor Biol 226: 111–121.

22. To TL, Maheshri N (2010) Noise Can Induce Bimodality in Positive Transcriptional Feedback Loops Without Bistability. Science 327: 1142–1145.

23. Hornos JEM, Schultz D, Innocentini GCP, Wang J, Walczak AM, et al. (2005) Self-regulating gene: An exact solution. Phys Rev E 72.

24. Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proc Natl Acad Sci U S A 105: 17256–17261.

25. Kærn M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet 6: 451–464.

26. Dunny G (2007) The peptide pheromone-inducible conjugation system of Enterococcus faecalis plasmid pCF10: cell–cell signalling, gene transfer, complexity and evolution. Phil Trans R Soc B 362: 1185–1193.

27. Johnson CM, Manias DA, Haemig HAH, Shokeen S, Weaver KE, et al. (2010) Direct Evidence for Control of the Pheromone-Inducible prgQ Operon of Enterococcus faecalis Plasmid pCF10 by a Countertranscript-Driven Attenuation Mechanism. J Bacteriol 192: 1634–1642.

28. Bae T, Kozlowicz BK, Dunny GM (2004) Characterization of cis-acting prgQ mutants: evidence for two distinct repression mechanisms by Qa RNA and PrgX protein in pheromone-inducible enterococcal plasmid pCF10. Mol Microbiol 51: 271–281.

29. Shi K, Brown CK, Gu ZY, Kozlowicz B, Danny GM, et al. (2005) Crystal structure of PrgX and PrgX/pheromone: The role of tetramerization of PrgX in controlling the pheromone induction of pCF10 transfer. Abstr Gen Meet Am Soc Microbiol 105: 307.

30. Bensing BA, Dunny GM (1997) Pheromone-inducible expression of an aggregation protein in Enterococcus faecalis requires interaction of a plasmid-encoded RNA with components of the ribosome. Mol Microbiol 24: 295–308.

31. Mestl T, Plahte E, Omholt SW (1995) A Mathematical framework for describing and analyzing gene regulatory networks. J Theor Biol 176: 291–300.

32. Buttaro BA, Antiporta MH, Dunny GM (2000) Cell-associated pheromone peptide (cCF10) production and pheromone inhibition in Enterococcus faecalis. J Bacteriol 182: 4926–4933.

33. Fixen KR, Chandler JR, Le T, Kozlowicz BK, Manias DA, et al. (2007) Analysis of the amino acid sequence specificity determinants of the enterococcal cCF10 sex pheromone in interactions with the pheromone-sensing machinery. J Bacteriol 189: 1399–1406.

34. Nakayama J, Ruhfel RE, Dunny GM, Isogai A, Suzuki A (1994) The PrgQ gene of the Enterococcus-faeaclis tetracycline resistance plasmid pCF10 rncodes a peptide inhibitor, iCF10. J Bacteriol 176: 7405–7408.

35. Tomshine J, Kaznessis YN (2006) Optimization of a stochastically simulated gene network model via simulated annealing. Biophys J 91: 3196–3205.

36. Chatterjee A, Johnson CM, Shu CC, Kaznessis YN, Ramkrishna D, et al. (2011) Convergent transcription confers a bistable switch in Enterococcus faecalis conjugation. Proc Natl Acad Sci U S A 108: 9721–9726.

37. Leonard B, Podbielski A, Hedberg P, Dunny G (1996) Enterococcus faecalis pheromone binding protein, PrgZ, recruits a chromosomal oligopeptide permease system to import sex pheromone cCF10 for induction of conjugation. Proc Natl Acad Sci U S A 93: 260–264.

38. Subramanian G, Ramkrishna D (1971) On the solution of statistical models of cell populations. Math Biosci 10: 1–23.

39. Mantzaris NV, Daoutidis P, Srienc F (2001) Numerical solution of multi-variable cell population balance models: I. Finite difference methods. Comput Chem Eng 25: 1411–1440.

40. Ramkrishna D, Borwanker Jd (1973) Puristic analysis of population balance I. Chem Eng Sci 28: 1423–1435.

41. Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58: 35–55.

42. T nase-Nicola S, Ten Wolde PR (2008) Regulatory control and the costs and benefits of biochemical noise. PLoS Comput Biol 4: e1000125.

43. Adalsteinsson D, McMillen D, Elston T (2004) Biochemical Network Stochastic Simulator(BioNetS): software for stochastic modeling of biochemical networks. BMC bioinformatics 5: 24.

44. Kristich CJ, Li YH, Cvitkovitch DG, Dunny GM (2004) Esp-independent biofilm formation by Enterococcus faecalis. J Bacteriol 186: 154–163.

45. Redfield R (2002) Is quorum sensing a side effect of diffusion sensing? Trends Microbiol 10: 365–370.

46. Dunny G, Yuhasz M, Ehrenfeld E (1982) Genetic and physiological analysis of conjugation in Streptococcus-faecalis. J Bacteriol 151: 855–859.

47. Andrup L, Andersen K (1999) A comparison of the kinetics of plasmid transfer in the conjugation systems encoded by the F plasmid from Escherichia coli and plasmid pCF10 from Enterococcus faecalis. Microbiology 145: 2001–2009.

48. Haseltine EL, Rawlings JB (2002) Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. J Chem Phys 117: 6959–6969.

49. Rao CV, Arkin AP (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. J Chem Phys 118: 4999–5010.