Available online at www.sciencedirect.com
ScienceDirect
Journal of Biotechnology 128 (2007) 362–375
Journal of BIOTECHNOLOGY
www.elsevier.com/locate/jbiotec

Cell population heterogeneity in expression of a gene-switching network with fluorescent markers of different half-lives

Stephanie Portle [a,b], Thomas B. Causey [c], Kim Wolf [a], George N. Bennett [c], Ka-Yiu San [b,a], Nikos Mantzaris [a,b,*]

[a] *Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77251-1892, United States*
[b] *Department of Bioengineering, Rice University, Houston, TX 77251-1892, United States*
[c] *Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77251-1892, United States*

Received 15 August 2006; received in revised form 20 September 2006; accepted 26 September 2006

Abstract

We studied the distribution of expression levels amongst the cells of an *Escherichia coli* population carrying a gene-switching network, known as the genetic toggle. We employed two green fluorescent protein (GFP) reporter proteins with different half-lives and characterized the effect of isopropyl-β-D-thiogalactopyranoside (IPTG) inducer concentration on fluorescence distribution characteristics. Our flow cytometric measurements indicated that there is a spread of fluorescence phenotypes of one to three orders of magnitude, due to the highly heterogeneous nature of the cell populations under investigation. Moreover, the shape of the distribution at a specific quasi-time-invariant reference state, defined for comparison purposes, strongly depended on inducer concentration. For very low and very high inducer concentrations, the distributions at the reference state are unimodal. On the contrary, for intermediate IPTG concentrations, two distinct subpopulations were formed below and above a single-cell threshold, resulting in distributions with a bimodal shape. The region of inducer concentrations where bimodality is observed is the same and independent of GFP half-life. Bimodal number density functions are not only obtained at the reference state. Transient studies revealed that even in cases where the distribution at the reference state is unimodal, the distribution becomes bimodal for a period of time required for the population to pass through the single-cell induction threshold. However, this feature was only captured by the system with the reduced half-life GFP. A simple single-cell model was used to shed light into the effect of inducer concentration and GFP half-life on the shape of the experimentally measured number density functions. The wide range of fluorescent phenotypes and the inability of the average population properties to fully characterize network behavior, indicate the importance of taking into account cell population heterogeneity when designing such a gene-switching network for biotechnological and biomedical applications.

Keywords: Heterogeneity; Genetic toggle; GFP; Flow cytometry

* Corresponding author at: Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77251-1892, United States. Tel.: +1 713 348 2955; fax: +1 713 348 2955.
E-mail address: nman@rice.edu (N. Mantzaris).

0168-1656/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2006.09.026

1. Introduction

It is a well-known fact that cells of isogenic populations exhibit different phenotypes. This phenomenon, known as cell-to-cell variability or cell population heterogeneity, has been observed and studied in many systems from the mid 1940s until today. Some of the most representative experimental efforts include Delbrück's studies of phage burst size variability (Delbruck, 1945), Powell's observations on the variation in cell division times for various strains of bacteria and different conditions (Powell, 1956), studies of *lac* operon expression levels by Novick and Weiner (1957) and Maloney and Rotman (1973), Spudich and Koshland's work on the tumbling and smooth-swimming states of flagellated bacteria (Spudich and Koshland, 1976), studies on spore formation in *Bacillus subtilis* (Chung and Stephanopoulos, 1995), and work to differentiate and find inducer-dependence of extrinsic and intrinsic sources of heterogeneity by Elowitz et al. (2002) using artificial genetic networks in *Escherichia coli*.

Cell population heterogeneity stems from several sources. First, there are differences in the microenvironments surrounding the cells. If cells "feel" different input signals, such as a toxin or food source, they will act differently one from the other. However, even in uniform environments cell populations still behave heterogeneously. Each cell of a population undergoes the cell cycle during which it grows and then at a certain point it divides to yield two daughter cells. It has been experimentally established that most molecules of the mother cell excluding chromosomal DNA partition unequally between the two daughter cells (Sweeney et al., 1994; Tran et al., 2000; Kelleher et al., 2000; Bellaiche et al., 2001; Orgogozo et al., 2002). Thus, the daughter cells initially have different intracellular states, and as a result, they exhibit different phenotypes. Since each cell undergoes the cell cycle many times before it dies, this phenomenon repeats itself and leads to a distribution of phenotypes amongst the cells of the population. In addition to unequal partitioning of cellular material at division, cell population heterogeneity may result from stochastic phenomena at the single-cell level which are unrelated to division. Specifically, the phenotype of each cell is largely determined by the function of a well-known class of molecules, known as regulatory molecules, which typically exist in very low concentrations. Therefore, their action can be purely stochastic. Hence, even cells which have identical content at a given point in time may exhibit differences in state and phenotype at the next instant. As a result, cell population heterogeneity is tightly related to regulation of gene expression.

Gene expression is regulated on a variety of levels, from before transcription to after translation (Alberts et al., 2002). In bacterial systems, it is thought that most genes are mainly regulated at the transcriptional level (Thieffry et al., 1998; Weinzierl, 1999). Transcriptional regulation is performed by regulatory molecules that can have positive or negative effects with respect to expression of the gene. However, naturally occurring regulatory networks are usually highly interconnected to each other and to various metabolic pathways, while in many cases several of their functional aspects are simply unknown. Thus, understanding the relationship between complex, natural regulatory networks and the distribution of phenotypes at the cell population level constitutes a major challenge.

On the other hand, artificial gene regulatory networks are typically simpler and consist of elements with well-defined, pre-determined functions, which do not interfere with the rest of cellular function. Thus, they offer the great opportunity to study in isolation the complex interplay between cell population heterogeneity and specific regulatory architectures. In addition, since they are constructed *de novo*, they offer the unique capability to engineer the cellular phenotype by using combinations of such networks (Guet et al., 2002). Thus, their potential use for biotechnological and biomedical applications is theoretically enormous. However, in order to achieve such ambitious goals, it is of fundamental importance to first investigate in detail the relationship between simple artificial genetic networks and cell population heterogeneity.

Several groups have designed artificial genetic networks with distinct dynamical features, such as an oscillator (Elowitz and Liebler, 2000; Fung et al., 2005) or a bistable genetic switch (Chen et al., 1993). Besckai et al. have studied negative and positive feedback genetic architectures and how various components and conditions contribute to the level of gene expression (Becskei and Serrano, 2000; Becskei et al., 2001). In this work we will focus on a particular artificial gene-switching network, known as the genetic toggle, which was constructed by Gardner et al. (2000). The network is composed of two promoter–operator–repressor sets,

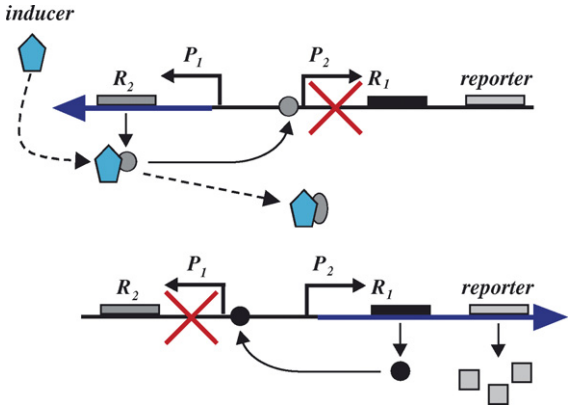

Fig. 1. Cartoon of the genetic toggle network in pTAK117, modified from Judd et al. (2000): $P_1 = P_L s1con$, $R_2 = lac$ repressor (*lacI*) (light circles), $P_2 = Ptrc\text{-}2$, $R_1 = lambda$ repressor (*cIts*) (dark circles), reporter = GFP (*gfpmut3*) (squares), inducer = IPTG (pentagons).

connected in a mutually repressive way: expression of one repressor inhibits expression of the other repressor (see Fig. 1). This leads to each repressor competing for dominance in the cell. In the corresponding plasmid called pTAK117 that we used in this work, the two repressors are the *lac* repressor and the temperature sensitive *cI* (*lambda*) repressor from phage lambda. The *lac* repressor inhibits the function of the Ptrc-2 promoter controlling expression of the *lambda* repressor, which in turn inhibits the function of the P_Ls1con promoter controlling expression of the *lac* repressor. The stable GFP variant *gfpmut3* (Cormack et al., 1996) is co-expressed with the *lambda* repressor, thus functioning as a reporter protein for *lambda* repressor expression levels. This two-gene network has two possible states. In the absence of inducers at low temperature, *lac* repressors dominate (GFP expression is low). To switch to domination of *lambda* repressors, the extracellular inducer isopropyl-β-D-thiogalactopyranoside (IPTG, a lactose analog) can be added, which will bind to the *lac* repressors and derepress *cI* expression. A similar genetic switch had also been constructed earlier by Chen et al. (1991, 1993) using the same repressor proteins but different promoters (λP_L and tac) fused with different reporter genes (vibrio hemoglobin and chroramphenicol acetyl transferase gene).

In their original work where the pTAK117 construct was presented, Gardner et al. (2000) mainly focused on illustrating the presence of the two aforementioned states of the genetic toggle network. Using an experimental characterization assay consisting of three successive dilution steps and applying changes in IPTG inducer concentration and temperature at different points in time, they also showed the stability of the induced and uninduced states in the pTAK117 system. Moreover, their flow cytometric measurements of the entire fluorescent distributions at three IPTG concentrations revealed some interesting patterns at the cell population level and clearly illustrated high extents of heterogeneous behavior. However, the half-life of *gfpmut3* reporter protein is known to be more than 24 h (Andersen et al., 1998), while the half-life of the *lambda* repressor is of the order of 1 h (Keiler et al., 1996). Thus, for a more complete understanding of the relationship between the specific genetic toggle architecture and cell population heterogeneity, a reporter protein with a half-life close to that of the *lambda* repressor must also be used.

Multiple GFP variants have been employed by several researchers for reporting promoter activity or protein expression levels. Bi et al. (2002) studied the dynamics of cell death using two GFP variants with different half-lives. Their results showed that the destabilized variant exhibited better dynamic characteristics. Furthermore, Sternberg et al. (1999) compared the use of a stable and unstable GFP variant in studying cell growth in biofilms. They found that the stable GFP was good for estimating growth rates during exponential growth, but for less favorable or stable conditions than a chemostat, the unstable variant had distinct advantages.

Motivated by the above considerations, we first constructed a variant of plasmid pTAK117 containing a reduced half-life GFP gene (*gfpaav*) in place of *gfpmut3*. This new gfp is a variant of *gfpmut3* itself, mutated by the addition of a degradation tag at the C-terminus (Andersen et al., 1998). The excitation and emission spectra of these proteins are very similar, assuring us that the only difference between these proteins is in degradation rate. The approximate half-life of *gfpaav* is 60 min. We then performed detailed characterization experiments with both the long and reduced half-life GFP reporter proteins. The distribution measurements were carried out with flow cytometry, a high throughput method that quantifies light scattering and fluorescence properties of individual cells. The parameters relevant to this study are the two light scattering properties, forward and side scatter (measures of cel-

lular size and internal complexity, respectively), and green and red fluorescence (FSC, SSC, FL1 and FL3, respectively). Green fluorescence represents intracellular GFP content, while red fluorescence represents propidium iodide (PI) content, a measure of viability. Our experimental and modeling results aim at elucidating the level of control that IPTG and protein half-life can exert on cell distribution characteristics.

2. Materials and methods

2.1. Strain

E. coli strain JM2.300 (*lambda-*, *lacI22 rpsL135* (StrR), *thi-1*, CGSC strain 5002) was used for all experiments. This strain has a mutant *lac* repressor that is non-functional. It was transformed with either plasmid pTAK117 or plasmid pGFPaav (see below). The cells were made chemically competent and transformed with plasmid as described in *RbCl Transformation Procedure for Improved Efficiency* (1994).

2.2. Plasmid construction

Two plasmids were used in this work. The first one, pTAK117 with the long half-life GFP was a gift by Professor Collins. Standard methods were used for construction of the plasmid with the reduced half-life (pGFPaav) and polymerase chain reaction (Miller, 1992; Sambrook and Russell, 2001). The *gfpmut3* was excised from pTAK117 by digestion with *Kpn*I–*Nhe*I. The resulting 5366 bp fragment was isolated from a 0.8% NuSieve GTG agarose (ISC BioExpress) tris-borate–EDTA gel using a Qiagen, Qiaex II Gel Extraction Kit. The following primers (obtained from Integrated DNA Technologies, INC.) were used to amplify gfp_{aav} from plasmid pZE21-GFPAAV (Elowitz and Liebler, 2000): N-terminus, 5′GGAGAAA*GGTACC***ATG**CGTAAAGGAGAAG-AACT3′ (the sequence in italics is the *Kpn*I cloning site and the bold sequence the gfp_{aav} start codon) and C-terminus, 5′AAGCTT*GCTAGC*-**TTA**AACTGCTGCAGCGTAGT3′ (the sequence in italics is the *Nhe*I cloning site and the bold sequence the gfp_{aav} stop codon). Platinum *Pfx* DNA polymerase (Invitrogen) was used for DNA amplification with a RoboCycler Gradient 96 Temperature Cycler (Stratagene). The gfp_{aav} PCR product was cleaned using a Wizard PCR Preps (Promega). The DNA was then digested with *Kpn*I–*Nhe*I and *DpnI* to remove any template plasmid. After digestion the restriction endonucleases were removed using the Wizard PCR Preps kit and the gfp_{aav} DNA was suspended in de-ionized water. Two hundred and forty nanograms of the gfp_{aav} DNA was ligated to 80 ng of the 5366 bp DNA fragment using 400 units of T4 DNA ligase (New England Biolabs) in a total volume of 10 μL. The ligation was performed at 16 °C for 18 h. Three microliters of the ligation reaction were transformed into *E. coli* JM2.300 chemically competent cells using the RbCl procedure (1994). Luria-Bertani (LB) agar plates and broth were used for growth and selection of bacteria (Sambrook and Russell, 2001). All LB agar plates and broth cultures contained 50 mg L^{-1} ampicillin. Transformants were selected on LB agar plates containing ampicillin. After selection, plasmid was isolated from LB broth cultures for screening with *Nde*I. One plasmid was chosen that had the correct banding pattern after digestion with *Nde*I for conformation of the gfp_{aav} insert by DNA sequencing (LS Labs, Inc.). The primers used for sequencing flanked the gfp_{aav} gene: N-terminus, 5′-GGGCATCAAATTAAACCACACC-3′, C-terminus, 5′-CCAAAACAGCCAAGCTAGCG-3′. After the gfp_{aav} sequence was confirmed the plasmid was named pGNB10030. It is referred to in this paper by its gfp variant, pGFPaav.

2.3. Shake flask experiments

Cells were grown to exponential phase in a single shake flask. First, cells were grown overnight in 5 mL of LB medium (10 g/L NaCl, 10 g/L BactoTryptone (BD Biosciences), 5 g/L yeast extract (Fisher BioTech), 100 mg/L ampicillin (Fisher BioTech)) for 12 h and then subcultured at a low cell density (∼1–2000 cells/mL) in 400 mL of prewarmed and aerated LSRB medium (4 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, 100 mg/L ampicillin, the appropriate concentration of IPTG (Fisher BioTech)) in the range of 10–2000 μM at 32 °C, shaking at 250 rpm in an orbital shaker (New Brunswick Scientific), covered from light. The 2 L flasks were capped with foam to allow oxygen transfer. Samples were withdrawn approximately once per generation (40 min) until the cultures started to

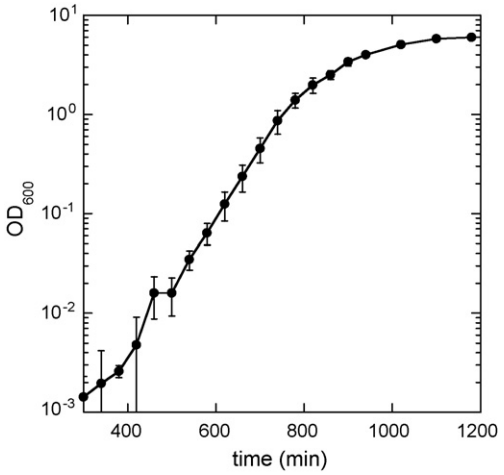

Fig. 2. Optical density (OD_{600}) as a function of time for [IPTG] = 30 μM. The cell population remains in exponential phase until about 700 min, after which point it starts to transition to stationary phase.

transition to stationary phase (see Fig. 2 for a representative optical density (OD_{600}) curve. Cell density was measured using a spectrophotometer (Bausch & Lomb, Spectronic 1001) at 600 nm. Samples were not taken until cells reached an OD_{600} of approximately 0.0001 due to the low initial cell concentration. Cells were washed twice in cold PBS (58 mM Na_2HPO_4, 17 mM NaH_2PO_4, 68 mM NaCl, in Milli-Q H_2O, pH adjusted between 7.3 and 7.4, filter sterilized). The samples were kept on ice and shielded from light before being measured with the flow cytometer. Triplicate experiments were performed for each IPTG concentration.

2.4. PI staining

Propidium iodide (PI) is used to identify dead or injured cells (Alsharif and Godfrey, 2002; Lewis et al., 2004), as PI can traverse the membrane of cells with compromised membranes and bind to the DNA (fluorescing red). The cells were stained with PI (Sigma) at a concentration of 2 mg/L and left at room temperature for 5 min.

2.5. Flow cytometry

The flow cytometer used was a FACScalibur (BD Biosciences), with a 15 mW, 488 nm, air-cooled argon-ion laser. All parameters were logarithmically amplified, with the following settings: FSC E01, SSC 381 V, FL1 601 V, FL2 500 V, FL3 575 V. Compensation of FL1–0.9% FL2 was applied to remove overlap of PI emission spectra into FL1. A side scatter threshold was applied to gate out much of the noise (at channel 130). Twenty thousand to 40,000 events were collected for each sample. The flow cytometer was calibrated with EGFP calibration beads (Clontech).

2.6. Gating data

Gates were applied in the post-processing of the flow cytometer files according to red fluorescence (PI) and light scattering. Cells that stained strongly with PI had lower green fluorescence than the rest, possibly due to GFP leakage out of the cells (Lehtinen et al., 2004). Since these cells did not represent the network dynamics, all cells staining strongly with PI were excluded. Moreover, since *E. coli* cells are so small and close to the size of the debris being sensed by the cytometer, gates were applied to better identify the cells. After observing the change in light scattering values while the culture was in exponential phase (between 300 and 700 min), identical FSC and SSC gates were applied to include events above a linear value of 30 (channel value of 378). Additionally, an FL3 (red fluorescence) gate was applied to exclude events above a relative red fluorescence level of 40 (channel value of 410), excluding dead or injured cells.

2.7. Software

First, the listmode files from the flow cytometer were converted from fcs format to ASCII format with MFI software (Martz, 1992), after confirming that FACScalibur data can be successfully converted using MFI. Data were then processed with a FORTRAN code we developed to apply the aforementioned gates and calculate average fluorescence per cell and number density functions, excluding events in the first or last channel (Leary, 1998). Graphs were constructed with Kaleidagraph (Synergy Software, Version 3.52).

3. Experimental results

In order to assess the effects of inducer concentration and GFP half-life on cell distribution charac-

teristics it is important to first choose a point where the distributions will be compared to each other. Typically, comparisons are made at the same point in time (e.g. Gardner et al., 2000; Khlebnikov and Keasling, 2002). However, cellular physiology changes not only as a function of time, but also as a result of different extracellular conditions. Thus, it is more meaningful to compare distributions when cell cultures grown under different conditions exhibit similar qualitative behavior.

Fig. 3a and b shows a representative example of the time evolution for the average FSC (a measure of cell size), and average green fluorescence of the entire cell population for both strains and for [IPTG] = 100 μM. Notice that independent of GFP half-life, the average population properties initially increase with time. They subsequently reach a plateau lasting for approximately three generations (460–540 min), while afterwards, the average size and average fluorescence start dropping with time as the cells prepare to enter into the stationary phase. Moreover, during the period of time where the average population properties reach a plateau, the fluorescence number density functions become practically indistinguishable from each other (Fig. 3c). Very similar qualitative behavior was observed for all other IPTG conditions studied.

Thus, based on these observations, green fluorescence distributions were compared at points in time where the average properties reach this characteristic plateau and the number density functions overlap. This quasi-steady-state of the cell population will henceforth be referred to as the "reference state". We note that the aforementioned definition of the reference state is in the spirit of that of balanced growth, according to which a batch cell culture reaches the state of balanced growth when the number density function with respect to all cellular properties becomes time-invariant (see Fredrickson et al., 1967 for a thorough discussion). However, it is impossible to measure all cellular characteristics. More importantly, due to the existence of stationary phase, it is not realistic to expect that the number density function will stay forever in a time-invariant state.

It was found that for low (below 30 μM) and high (above 50 μM) IPTG concentrations the reference state was reached at either 500 or 540 min starting from an initial cell concentration of 2000 cells/mL, well before cells exit logarithmic phase (they start changing to sta-

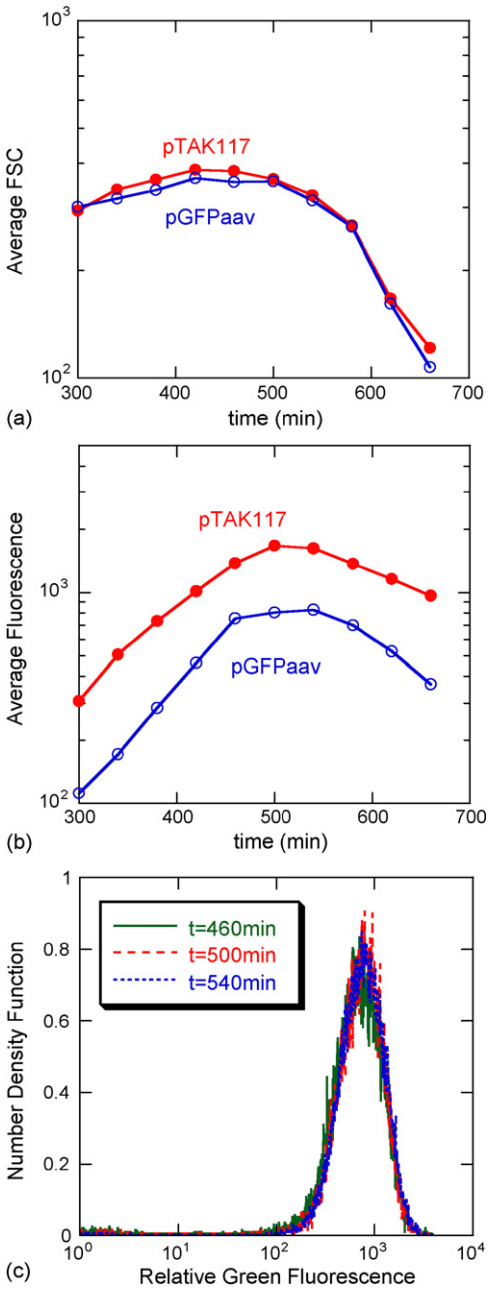

Fig. 3. Defining the reference state. All results at [IPTG] = 100 μM. Filled circles: pTAK117; open circles pGFPaav. (a) Average forward scatter (FSC) as a function of time. (b) Average green fluorescence as a function of time. (c) Number density function for pGFPaav at t = 460 min (solid line), 500 min (dashed line) and 540 min (dotted line). Number density functions practically overlap for three generations, while the average FSC and green fluorescence remain almost the same (holds true regardless of GFP half-life and [IPTG]).

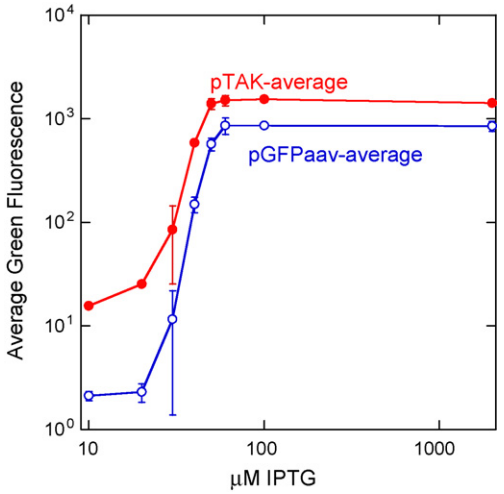

Fig. 4. Average green fluorescence at the reference state as a function of inducer concentration. Filled circles: pTAK117. Open circles: pGFPaav. Error bars represent the standard deviation based on triplicate experiments. Some error bars are smaller than the markers.

tionary phase in FSC and SSC as stated above just after those values plateau). However, for intermediate IPTG concentrations (30–50 μM) more time is required for the reference state to be reached. It was found that IPTG concentration does not affect the growth rate of the culture. Moreover, no differences were found in the time to reach the reference state between the GFP reporter proteins. Therefore, this qualitative difference for intermediate IPTG concentrations was attributed to the balance existing between the *lambda* and *lac* repressor concentrations, which in turn leads to slower evolution of the genetic toggle dynamics at intermediate [IPTG] (see also modeling section). Thus, in order to avoid reaching the stationary phase before the aforementioned criteria for the reference state are met at intermediate [IPTG] lower initial cell concentrations were used (1–10 cells/mL).

After establishing a reference state where comparisons are meaningful, we focus on the experimental characterization of the state of the entire cell population. Fig. 4 shows the average relative green fluorescence at the reference state as a function of [IPTG] for both markers. Due to the structure of the genetic toggle network and the biological function of the inducer, the average fluorescence exhibits a monotonically increasing, sigmoidal dependence on [IPTG]. A relatively sharp rise in average fluorescence is observed for both

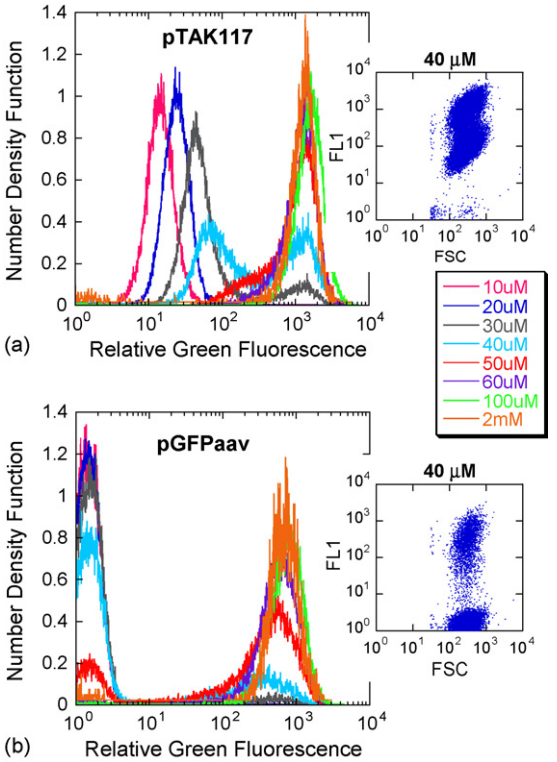

Fig. 5. Number density functions at the reference state for different inducer concentrations. (a) pTAK117 and (b) pGFPaav as determined by flow cytometry as described in methods. Included are scatter plots of FL1 vs. FSC for each variant, showing that each subpopulation of the bimodal distributions have the same FSC (size) range, and therefore differences in fluorescence are not due to age of the cells.

pTAK117 and pGFPaav between 20 and 60 μM above which the average fluorescence remains at maximal levels. Since the green fluorescent protein encoded by *gfpaav* has a shorter half-life than that expressed by the *gfpmut3* gene present in pTAK117, the average fluorescence levels of the cell cultures with the former gene are lower than those carrying the latter.

One would obtain only limited understanding of network behavior if the experimental characterization were based solely on average population properties. A much more in-depth insight is obtained by studying the fluorescence distributions of the entire cell population. Fig. 5a and b shows the number density functions for different inducer concentrations at the reference state for both pTAK117 (Fig. 5a) and pGFPaav (Fig. 5b). First, notice that expression levels vary by one to even

three orders of magnitude for some [IPTG]. Thus, the cell populations appear very heterogeneous while the extent of heterogeneity strongly depends on [IPTG].

Moreover, for low (below 30 μM) and high (above 50 μM) [IPTG], all number density functions are unimodal. However, in the intermediate range of [IPTG], where the transition from low to high average expression level occurs (see Fig. 4), the number density functions are bimodal. There appears to be a clear threshold at the single-cell level for these intermediate IPTG concentrations. The cells below and above this threshold form well-defined subpopulations, which lead to the observed bimodality. We note that Gardner et al. (2000) also reported bimodal distributions for pTAK117 at [IPTG] = 40 μM using a different experimental protocol, while our results show bimodal number density functions for [IPTG] between 30 and 50 μM for both reporter proteins. Therefore, these results indicate that the specific genetic architecture at the single-cell level leads to this characteristic bimodality feature at the cell population level for intermediate inducer concentrations.

Despite the fact that the range of [IPTG] for which the number density function is bimodal is the same for both reporter proteins, there are significant differences between pTAK117 and pGFPaav in the value of the single-cell threshold separating the two subpopulations forming the bimodal distributions. Specifically, as estimated by eye in Fig. 5a and b, the switch between low and high expression levels occurs at relative fluorescence values around 300 and 4 for pTAK117 and pGFPaav, respectively.

Notice also that for both reporter proteins, at high IPTG concentrations, the unimodal number density functions practically overlap one with the other. Moreover, the high fluorescence parts of the bimodal number density functions obtained at intermediate induction levels appear to be centered around the same mean fluorescence value. Thus, it appears that for [IPTG] above 30 μM and for a range of [IPTG] nearly two orders of magnitude (up to 2 mM), all cells crossing the single-cell threshold become maximally induced. On the contrary, for low [IPTG], the population shifts toward higher expression levels as [IPTG] increases, indicating that the balance between the *lac* and *lambda* repressor concentrations shifts to allow slightly higher expression levels of GFP. This behavior is more visible in the case of pTAK117 (Fig. 5a) where the single-cell fluorescence threshold for induction is approximately 75 times higher than in the case of pGFPaav. Due to the low GFP level at low [IPTG] in pGFPaav, the fluorescence level for some subthreshold cells falls below the quantitative sensitivity of the flow cytometer, and some are below the four decade range of the cytometer, making it difficult to compare number density functions at low expression levels for the two systems. However, since the half-life of the reporter expressed by pGFPaav is much closer to that of the *lambda* repressor, the sharper separation between the two subpopulations forming the bimodal distributions in the case of pGFPaav and the significantly larger separation between the mean low and high fluorescence values of unimodal distributions is believed to be a more accurate representation of network behavior.

A more complete view of gene switching dynamics can be obtained through transient studies. Figs. 6 and 7 show the time evolution of the green fluorescent number density function for both pTAK117 and pGFPaav and for [IPTG] = 40 and 60 μM, respectively. In both cases, the initial number density function (t = 300 min) is unimodal. In the first case (Fig. 6), for both reporter proteins, a second subpopulation grows out of the main population in the middle of the exponential growth phase. Notice also that the subpopulation consisting of cells at high induction levels becomes gradually more significant until the reference state is reached, where the overall number density function obtains a stable bimodal shape. As also shown in Fig. 5, the separation between the low and high expressing peaks in the bimodal number density function is more pronounced in the case of the reduced half-life GFP both transiently and at the reference state. This qualitative route towards bimodality is representative of all intermediate inducer concentrations where bimodal number density functions are obtained at the reference state ([IPTG] = 30–50 μM).

However, bimodality can also be observed transiently even in cases where the number density function is unimodal at the reference state (see Fig. 7b). In the case of the reduced half-life GFP, as the population passes through the single-cell switching point, a separate subpopulation at high fluorescence is formed and continues to co-exist with the low fluorescence subpopulation, thus resulting in a bimodal number density function. However, in contrast to the case of [IPTG] = 40 μM (Fig. 6b) the percentage of cells

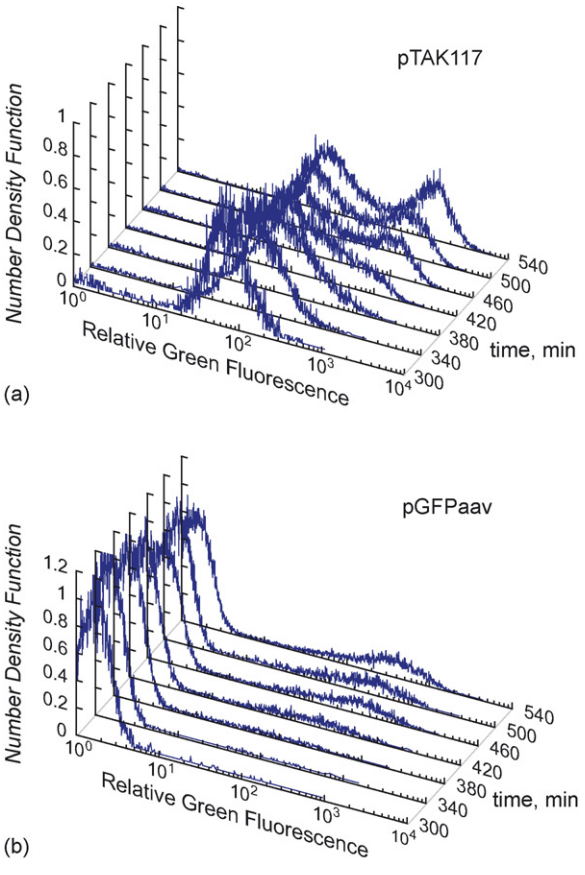

Fig. 6. Time evolution of the number density function at [IPTG] = 40 μM for (a) pTAK117 and (b) pGFPaav as determined by flow cytometry as described in methods.

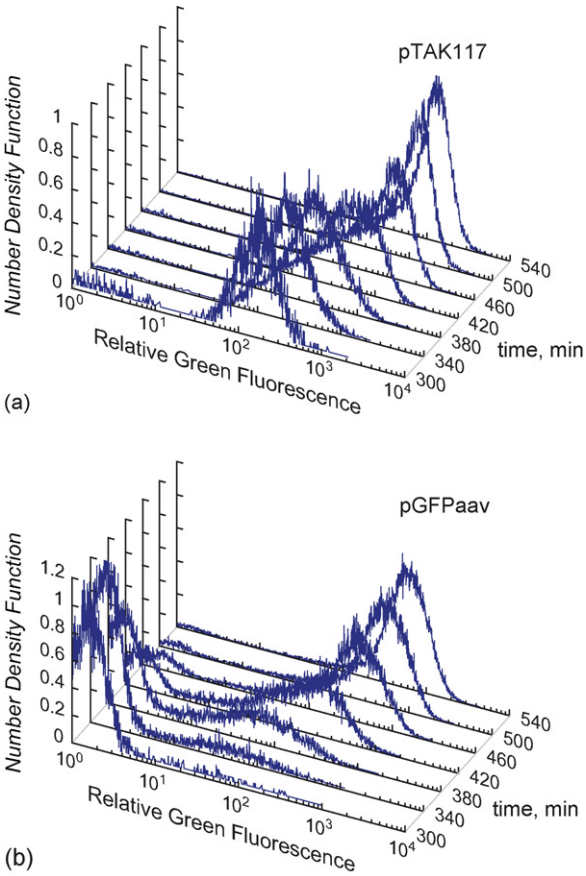

Fig. 7. Time evolution of the number density function at [IPTG] = 60 μM for (a) pTAK117 and (b) pGFPaav as determined by flow cytometry as described in methods.

below threshold continually decreases with time until it completely vanishes. Thus, the number density function eventually becomes unimodal at the reference state due to the higher concentration of IPTG, which enables even the low expressing cells to become fully induced after some period of time. Contrary to pGFPaav, transient bimodality is not observed in the case of pTAK117 at [IPTG] = 60 μM (Fig. 7a). The number density function becomes distorted as it passes through the single-cell switching point of 300 but it never becomes bimodal. This is attributed to the much higher stability of the protein expressed by *gfpmut3* compared to that of the *lambda* repressor, which does not allow enough time for visualizing the transition of cells from low to high *lambda* repressor expression levels that actually occurs at the single-cell level.

To obtain an understanding of transient bimodality we note that there exists a threshold in intracellular content due to the architecture of the genetic toggle network. As can be seen in Fig. 5 this threshold is around 300 for pTAK117 and 4 for pGFPaav in units of relative fluorescence as measured by the flow cytometer. Cells which cross this threshold become induced at a much higher rate than cells below this threshold. Due to this difference in induction rates above and below the single-cell threshold and the fact that the cell population is heterogeneous, as cells cross this threshold transiently, two distinct subpopulations are formed, thus, resulting in the observed bimodal shapes. For high enough extracellular IPTG concentrations however, bimodality is only transient since the high IPTG

concentration eventually allows all cells of the population to cross the single-cell threshold.

4. Mathematical modeling

In order to obtain a more in-depth understanding of some key features that were experimentally observed we utilized the simple, deterministic single-cell model that was presented by Gardner et al. (2000). According to this model, the expression dynamics of the *lac* and *lambda* repressors are given by the following dimensionless equations:

$$\frac{du}{dt} = \frac{a_1}{1+v^{\beta}} - u \quad (1)$$

$$\frac{dv}{dt} = \frac{a_2}{1+(u/(1+([\mathrm{IPTG}]/K))^{\eta})^{\gamma}} - v \quad (2)$$

where u and v are the dimensionless intracellular amounts of the *lac* and *lambda* repressors, respectively; a_1 and a_2 the dimensionless effective synthesis rates of the P_Ls1con and Ptrc-2 promoters, respectively; β the cooperativity of repression of the P_Ls1con promoter; γ the cooperativity of repression of the Ptrc-2 promoter; K the dissociation constant of IPTG from the *lac* repressor; η is the cooperativity of IPTG binding. The *E. coli* chromosome of the strain used in the experiments has a mutant *lac* repressor, which was assumed to not interact with the inducer, IPTG, or with other functional repressors encoded on the plasmid to form a multimer, or with the operator sites on the plasmids. Any effect from the extra *lacO* site in the chromosome was assumed to be negligible because the plasmid is maintained at about 30 copies per cell (Amann et al., 1988). For a more detailed description of the model see Gardner et al. (2000), while for its derivation and underlying assumptions see Edelstein-Keshet (1988).

We are also interested in predicting the dynamics of GFP since this is the marker that is experimentally measured. Since GFP is co-expressed with the *lambda* repressor, its synthesis rate is assumed to be proportional to that of the *lambda* repressor, with a proportionality constant ξ. Moreover, let δ denotes the ratio of the GFP degradation rate over that of the *lambda* repressor. Then, the dynamics of the dimensionless intracellular GFP concentration is given by the equation:

$$\frac{d\mathrm{GFP}}{dt} = \frac{\xi a_2}{1+(u/(1+([\mathrm{IPTG}]/K))^{\eta})^{\gamma}} - \delta\,\mathrm{GFP} \quad (3)$$

Table 1
Dimensionless parameter values used in simulations

Parameter	Value
a_1	156.25
a_2	15.625
β	2.5
γ	1.0
K	2.9618×10^{-5}
η	2.0015
ξ	1.0
δ	1.0 (pGFPaav); 0.03 (pTAK117)

All parameters used in our simulations are the same as the ones presented in Gardner et al. (2000) and are also given in Table 1 for completeness. For simplicity we assumed $\xi = 1$. Moreover, since the more stable GFP of pTAK117 has a half-life more than 24 h, while that of pGFPaav is approximately the same as that of the *lambda* repressor, we assumed $\delta = 0.03$ when simulating cells with pTAK117 and $\delta = 1$ when simulating cells carrying pGFPaav. Using these parameter values we simulated this simple model aiming at understanding the fate of single cells for different inducer concentrations. As the initial condition we used an uninduced state, since in all the experiments presented all cells were initially at low expression levels as well.

Fig. 8a shows transient simulation results for three different inducer concentrations. Notice that at low [IPTG] (e.g. [IPTG] = 20 μM), uninduced cells remain at low expression levels until steady state is reached since the available IPTG molecules are not enough to relieve the strong inhibition that the *lac* repressor molecules exert on the Ptrc-2 promoter. On the contrary, for high [IPTG] (e.g. [IPTG] = 100 μM), even cells which start from very low expression levels quickly reach the induced steady state. The high concentration of IPTG in this case rapidly decreases the number of free *lacI* molecules and shifts the intracellular state in favor of the *lambda* repressors in a very rapid fashion. However, for intermediate [IPTG] (e.g. [IPTG] = 40 μM), there exists a much more even competition between the number of free *lac* and *lambda* repressor molecules. Thus, in order for a cell starting with high *lac* and low *lambda* repressor concen-

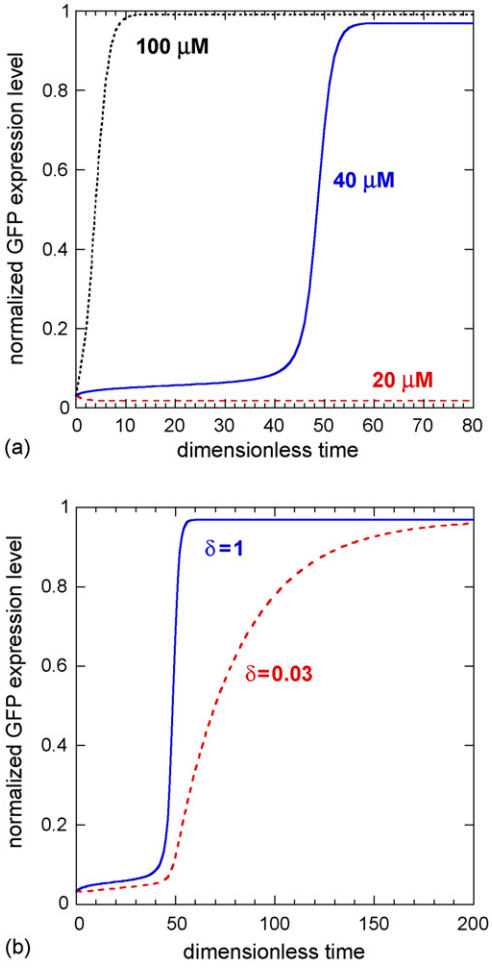

Fig. 8. Dynamics of the normalized GFP expression level using models (1–3) for: (a) different IPTG concentrations; [IPTG] = 20 μM (dashed line), [IPTG] = 40 μM (solid line) and [IPTG] = 100 μM (dotted line). In all simulations $\delta = 1$, while all other parameter values are given in Table 1 and (b) different GFP half-lives. Solid line: $\delta = 1$ (pGFPaav). Dashed line: $\delta = 0.03$ (pTAK117). In both simulations [IPTG] = 40 μM while all other parameter values are given in Table 1. Normalization was performed with the maximum possible GFP expression level (a_2/δ).

trations to overcome the presence of many free *lac* repressor molecules and eventually reach a state with high *lambda* repressor content, a significant amount of time is required, thus leading to the distinct sigmoidal dynamics shown in Fig. 8a. These results are consistent with the experimental finding that at intermediate [IPTG], more time is required in order for the system to reach the reference state.

These results also shed light into the effect that IPTG concentration has on the shape of the number density functions. Since uninduced cells at low [IPTG] cells remain at low expression levels, while at high [IPTG], they very quickly become induced, it is logical to expect the unimodal shape of the number density functions presented in Fig. 5. However, at intermediate IPTG concentrations there exists a very pronounced lag phase that an uninduced cell is required to pass before switching to high expression levels. The duration of this lag phase is a function of the initial condition; it decreases for higher initial expression levels. Thus, some of the cells present in the initial population manage to overcome the hurdle and eventually become induced, while other cells divide before having the opportunity to pass the single-cell threshold and become induced. Since, at each division event, cellular material partitions between two daughter cells, the expression levels of the two daughter cells are always smaller than those of the mother. Hence, we hypothesize that the combined action of division and the delay in surpassing the single-cell threshold shown in Fig. 8a, causes some cells to become "trapped" to the uninduced state, while others have enough time to become induced, thus resulting in the observed bimodal shape of the number density function at intermediate [IPTG].

Using this rationale, it is possible to obtain additional insight into the effect of GFP half-life on the bimodal shape of the number density functions presented in Fig. 5. Fig. 8b shows the single-cell dynamics for an intermediate [IPTG] (40 μM) for the pTAK117 ($\delta = 0.03$) and pGFPaav ($\delta = 1$) systems. Although longer GFP half-life promotes faster protein build-up, it also makes degradation the slowest step of the process compared to transcription and translation. Hence, in the case of the reduced GFP half-life, and once the single-cell threshold is passed, the transition from low to high expression levels is much faster than in the case of the long half-life GFP. As a result, more cells will exist for longer periods of time at intermediate expression levels (between the uninduced and induced states) for the system with slower GFP degradation. This in turn offers a possible explanation of the less sharp separation between the two subpopulations forming the bimodal number density function in the case of the pTAK117 system at intermediate IPTG concentrations.

Clearly, the model used for these simulations is fully deterministic and does not take into account stochas-

tic effects at the single-cell level (see e.g. Vilar et al., 2003). Moreover, it simulates only non-dividing single cells and therefore cannot predict entire cell property distributions such as cell population balance models (e.g. Mantzaris, 2006). In order to obtain a deeper quantitative understanding of the systems under consideration more detailed modeling is required. However, the hypotheses formed using this simple model, constitute a basis for further thinking that can contribute towards elucidating the complex relationship between single-cell genetic architecture and cell population heterogeneity.

5. Summary and conclusions

The interplay between a specific, prototype genetic architecture at the single-cell level and the distribution of phenotypes at the cell population level was studied with the use of flow cytometry. The network under investigation was the genetic toggle (Gardner et al., 2000) consisting of two promoter–repressor pairs. Two green fluorescent proteins of different half-lives were placed after the *lambda* repressor as reporters in order to study the effect of network structure on cell population dynamics as well as assess the potential of using an extracellular inducer (IPTG) to control the distribution characteristics. Both reporter proteins exhibited the same sigmoidal induction patterns for the average expression levels. However, experimental characterization with flow cytometry offered a more in-depth view of network behavior. Specifically, for very low and very high inducer concentrations, the distribution of fluorescent phenotypes was found to be unimodal at a well-defined, quasi-time-invariant reference state. On the contrary, at intermediate [IPTG], where the average fluorescence switches from low to high expression levels, the distributions become bimodal consisting of two subpopulations below and above a specific single-cell threshold. The region of inducer concentrations where bimodality was observed at the reference state was the same for both reporter proteins. Moreover, bimodality was not only observed at the reference state. Even at inducer concentrations where the distribution eventually becomes unimodal, transient studies revealed that as the cell population passes through the single-cell induction threshold, two subpopulations are formed for some period of time. Thus, the existence of a single-cell threshold and the bimodal shape of the fluorescence distribution is a robust pattern of this particular genetic network.

The single-cell induction threshold of fluorescence was found to be significantly lower in the case of the reduced half-life GFP protein. Thus, the corresponding bimodal distributions were visibly more asymmetrical for the system with the higher GFP degradation rate. Moreover, unlike the case of the reduced half-life GFP, it was not possible to visualize the bimodal transition of the population from low to high expression levels for high inducer concentrations using the more stable GFP. Since the half-life of the *lambda* repressor is much closer to that of the less stable GFP, the results obtained with the *gfpaav* gene downstream the *lambda* repressor at the reference state, as well as transiently, offer a more accurate representation of the genetic toggle behavior. However, the results obtained with the long half-life GFP are also valuable in understanding how the expression levels will be distributed amongst the cells of the population when a high-stability protein is placed under the control of the genetic toggle network downstream of the *lambda* repressor. Finally, a simple single-cell model was used to shed light into the effect of inducer concentration and GFP half-life on the shape of the experimentally measured number density functions.

The wide spread of fluorescent phenotypes from one to three orders of magnitude depending on the inducer concentration and the inability of the average population properties to fully characterize network behavior indicate the importance of taking into account cell population heterogeneity when designing such a gene-switching network for biotechnological applications.

Acknowledgements

The authors would like to thank J. Collins for donation of the plasmid pTAK117, Michael Elowitz for the plasmid pZE21-GFPAAV, Ronald J Parry for the use of shakers in his lab and Kyriacos Zygourakis for useful suggestions throughout the course of this work. Partial support by grants NSF/BES-QSB 0331324**,** NIH/NIGMS R01GM071888, equipment grant NSF/BES 0420840 and NSF IGERT Training Grant in Cellular Engineering (DGE0114264) is also gratefully acknowledged.

References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Molecular Biology of the Cell, fourth ed. Garland Science.

Alsharif, R., Godfrey, W., 2002. Bacterial Detection and Live/Dead Discrimination by Flow Cytometry. Becton, Dickinson and Company, San Jose, CA. Application Note, Microbial Cytometry. http://www.bdbiosciences.com/pdfs/whitePapers/23-6287-01.pdf.

Amann, E., Ochs, B., Abel, K.-J., 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. Gene 69, 301–315.

Andersen, J.B., Sternberg, C., Poulsen, L.K., Bjorn, S.P., Givskov, M., Molin, S., 1998. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl. Environ. Microbiol. 64 (6), 2240–2246.

Becskei, A., Serrano, L., 2000. Engineering stability in gene networks by autoregulation. Nature 405 (6786), 590–593.

Becskei, A., Seraphin, B., Serrano, L., 2001. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. EMBO J. 20 (10), 2528–2535.

Bellaiche, Y., Gho, M., Kaltschmidt, J.A., Brand, A.H., Schweisguth, F., 2001. Frizzled regulates localization of cell-fate determinants and mitotic spindle rotation during asymmetric cell division. Nat. Cell Biol. 3 (1), 50–57.

Bi, X.J., Wirth, M., Beer, C., Kim, E.J., Gu, M.B., Zeng, A.P., 2002. Dynamic characterization of recombinant Chinese hamster ovary cells containing an inducible c-fos promoter GFP expression system as a biomarker. J. Biotechnol. 93 (3), 231–242.

Chen, W., Bailey, J.E., Lee, S.B., 1991. Molecular design of expression systems: comparison of different repressor control configurations using molecular mechanism models. Biotechnol. Bioeng. 38 (7), 679–687.

Chen, W., Kallio, P.T., Bailey, J.E., 1993. Construction and characterization of a novel cross-regulation system for regulating cloned gene-expression in *Escherichia coli*. Gene 130 (1), 15–22.

Chung, J.D., Stephanopoulos, G., 1995. Studies of transcriptional state heterogeneity in sporulating cultures of *Bacillus subtilis*. Biotechnol. Bioeng. 47 (2), 234–242.

Cormack, B.P., Valdivia, R.H., Falkow, S., 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173 (1), 33–38.

Delbruck, M., 1945. The burst size distribution in the growth of bacterial viruses (Bacteriophages). J. Bacteriol. 50 (2), 131–135.

Edelstein-Keshet, L., 1988. Mathematical Models in Biology. McGraw-Hill, New York.

Elowitz, M.B., Liebler, S., 2000. A synthetic oscillatory network of transcriptional regulators. Nature 403, 335–338.

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. Science 297 (5584), 1183–1186.

Fredrickson, A.G., Ramkrishna, D., Tsuchiya, H.M., 1967. Statistics and dynamics of procaryotic cell populations. Math. Biosci. 1, 327–374.

Fung, E., Wong, W.W., Suen, J.K., Bulter, T., Lee, S.G., Liao, J.C., 2005. A synthetic gene-metabolic oscillator. Nature 435 (7038), 118–122.

Gardner, T.S., Cantor, C.R., Collins, J.J., 2000. Construction of a genetic toggle switch in *Escherichia coli*. Nature 403, 339–342.

Guet, C.C., Elowitz, M.B., Hsing, W., Leibler, S., 2002. Combinatorial synthesis of genetic networks. Science 296, 1466–1469.

Judd, E.M., Laub, M.T., McAdams, H.H., 2000. Toggles and oscillators: new genetic circuit designs. BioEssays 22, 507–509.

Keiler, K.C., Waller, P.R.H., Sauer, R.T., 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science 271 (5251), 990–993.

Kelleher, J.F., Mandell, M.A., Moulder, G., Hill, K.L., L'Hernault, S.W., Barstead, R., Titus, M.A., 2000. Myosin VI is required for asymmetric segregation of cellular components during C-elegans spermatogenesis. Curr. Biol. 10 (23), 1489–1496.

Khlebnikov, A., Keasling, J.D., 2002. Effect of lacY expression on homogeneity of induction from the Ptac and Ptrc promoters by natural and synthetic inducers. Biotechnol. Progr. 18, 672–674.

Leary, J.F., 1998. Flow Cytometry Data Display & Analysis – The Basics, in PATH 6311 – Molecular Cytometry.

Lehtinen, J., Nuutila, J., Lilius, E.M., 2004. Green fluorescent protein-propidium iodide (GFP-PI) based assay for flow cytometric measurement of bacterial viability. Cytometry Part A 60A (2), 165–172.

Lewis, G., Taylor, I.W., Nienow, A.W., Hewitt, C.J., 2004. The application of multi-parameter flow cytometry to the study of recombinant *Escherichia coli* batch fermentation processes. J. Ind. Microbiol. Biotechnol. 31 (7), 311–322.

Maloney, P.C., Rotman, B., 1973. Distribution of suboptimally induced Beta-D-galactosidase in *Escherichia coli*—enzyme content of individual cells. J. Mol. Biol. 73 (1), 77–91.

Mantzaris, N.V., Stochastic and Deterministic Simulations of Cell Population Dynamics. J. Theor. Biol., published online February 20, 2006.

Martz, E., 1992. MFI: A Flow Cytometry List Mode Data Analysis Program Optimized for Batch Processing under MS-DOS. http://www.umass.edu/microbio/mfi.

Miller, J.H., 1992. A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria. Cold Spring Harbor Press, Plainview, NY.

Novick, A., Weiner, M., 1957. Enzyme induction as an all-or-none phenomenon. Proc. Natl. Acad. Sci. U.S.A. 43 (7), 553–566.

Orgogozo, V., Schweisguth, F., Bellaiche, Y., 2002. Binary cell death decision regulated by unequal partitioning of Numb at mitosis. Development 129 (20), 4677–4684.

Powell, E.O., 1956. Growth rate and generation time of bacteria, with special reference to continuous culture. J. Gen. Microbiol. 15 (3), 492–511.

RTPIE, 1994. RbCl Transformation Procedure For Improved Efficiency. The NEB Transcript. 6(1), p. 7.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Lab. Press, Plainview, NY.

Spudich, J.L., Koshland, D.E., 1976. Non-genetic individuality—chance in single cell. Nature 262 (5568), 467–471.

Sternberg, C., Christensen, B.B., Johansen, T., Nielsen, A.T., Andersen, J.B., Givskov, M., Molin, S., 1999. Distribution of bacterial growth activity in flow-chamber biofilms. Appl. Environ. Microbiol. 65 (9), 4108–4117.

Sweeney, P.J., Srienc, F., Fredrickson, A.G., 1994. Measurement of unequal DNA partitioning in tetrahymena-pyriformis using slit-scanning flow-cytometry. Biotechnol. Progr. 10 (1), 19–25.

Thieffry, D., Huerta, A.M., Perez-Rueda, E., Collado-Vides, J., 1998. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. Bioessays 20 (5), 433–440.

Tran, J., Brenner, T.J., DiNardo, S., 2000. Somatic control over the germline stem cell lineage during Drosophila spermatogenesis. Nature 407 (6805), 754–757.

Vilar, J.M.G., Guet, C.C., Leibler, S., 2003. Modeling network dynamics: the lac operon, a case study. J. Cell Biol. 161 (3), 471–476.

Weinzierl, R.O.J., 1999. Mechanisms of Gene Expression. Imperial College Press, London, UK.