Chemical Engineering Science 64 (2009) 4247 – 4261

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

A novel free boundary algorithm for the solution of cell population balance models

M.E. Kavousanakis[a,*], N.V. Mantzaris[b], A.G. Boudouvis[a]

[a] National Technical University of Athens, School of Chemical Engineering, Athens 15780, Greece
[b] Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77005, USA

ARTICLE INFO

Article history:
Received 1 December 2008
Received in revised form 3 June 2009
Accepted 17 June 2009
Available online 24 June 2009

Keywords:
Cell heterogeneity
lac operon
Nonlinear dynamics
Numerical analysis
Mathematical modelling

ABSTRACT

There exists an abundance of experimental evidence in a variety of systems, showing that cell populations are heterogeneous systems in the sense that properties such as size, shape, DNA and RNA content are unevenly distributed amongst the cells of the population. The quantitative understanding of heterogeneity is of great significance, since neglecting its effect can lead to false predictions. Cell population balance models are used to address the implications of heterogeneity and can accurately capture the dynamics of heterogeneous cell populations. They are first-order partial-integral differential equations and due to the complexity of formulation, analytical solutions are hard to obtain in the majority of cases. Despite the recent progress, the efficient solution of cell population balance models remains a challenging task. One of the main challenges stems from the fact that the boundaries of the intracellular state space are typically not known *a priori* and using fixed-boundary algorithms leads to inaccuracies and increased computational time demands. Motivated by this challenge, we formulated a free boundary finite element algorithm, capable of solving cell population balance equations more efficiently than the traditional fixed-boundary algorithms. The implementation of the algorithm is accommodated, in the finite element based software package COMSOL Multiphysics. We demonstrate the efficiency of this algorithm using the *lac operon* gene regulatory network as our model system and perform transient and asymptotic behavior analysis. In the latter case, the pseudo-arc-length continuation algorithm is incorporated, in order to investigate the existence of a bistability region, also observed at the single-cell level. Our analysis, revealed the existence of a region of bistability when cell heterogeneity is taken into account; however, its extend shrinks comparing to homogeneous cell populations. The free boundary algorithm can be easily extended for problems of higher dimensionality and we present results for a two-dimensional cell population balance model, which can exhibit an oscillatory behavior. It is shown that oscillations do not persist when the intracellular content is unevenly distributed amongst the daughter cells.

1. Introduction

The recent development in molecular biology, genomics, transcriptomics and proteomics provide powerful tools for the efficient study of processes which occur at the cell level. Systems biology techniques have been applied in order to illuminate and understand the immense complexity at the cell level. However, the biological behavior strongly depends on the complex interactions, which occur amongst the cells of an isogenic population and can lead to major variableness of the phenotypic behavior. This phenomenon—known as heterogeneity of cell populations—has been observed in a plethora of systems; Delbrück (1945) showed significant variations in phage burst sizes. Chung and Stephanopoulos (1995) showed heterogeneity among transcriptional states in sporulating cultures of *Bacillus subtilis*. Heterogeneity has also been observed in the tumbling and swimming states of flagellated bacteria (Spudich and Koshland, 1976), in endothelial cell surface markers (Oh et al., 2004) and in various isogenic *Escherichia coli* systems (Elowitz et al., 2002).

The observed heterogeneity in cell populations is a very important factor which can encumber the efficient production of biotechnological products. On the other hand, the phenotypic variability permits the adaptation of cell populations to abrupt changes of their environment. Furthermore, the main constraining factor in cancer treatment is the observed heterogeneity of the neoplasmatic and metastatic cells. Heterogeneity is evident among histologically similar cancers from different patients (*intertumor* heterogeneity) and among different cells of the same cancer at a single time (progression). Thus, it is crucial to fully understand its impact on the dynamic relation between processes which occur at the single-cell level and at the cell population level.

* Corresponding author. Department of Chemical Engineering, Princeton University, Princeton, NJ 08544, USA. Tel.: +1 609 2582977; fax: +1 609 2580211.
E-mail addresses: mkavousa@princeton.edu (M.E. Kavousanakis), nman@rice.edu (N.V. Mantzaris), boudouvi@chemeng.ntua.gr (A.G. Boudouvis).

0009-2509/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ces.2009.06.054

Cell population heterogeneity has been ascertained to have a strong dependence on the operational properties of gene regulatory networks. Gene regulatory networks are composed of DNA segments in a cell, which interact with each other through their RNA and protein expression products and with other substances in the cell. Thereby, they govern the rates at which genes in the network are transcribed into mRNA. In general, each mRNA molecule is a chemical "blueprint" for a protein product and carries coding information to the sites of protein synthesis, the ribosomes. Some proteins, in turn, serve only to activate other genes and these are the transcription factors, which are the key parts in regulatory networks or cascades. The transcription factors bind to the promoter region at the start of other genes and can either turn them on, initiating the production of another protein (and so on), or inhibit their functionality. The phenotype of each cell is strongly depending on the type and number of genes expressed at each moment and on the subsequent intracellular reactions (metabolism). The intracellular reactions are specifically influenced by a special category of intracellular regulatory molecules. The regulatory molecules exist in small concentrations (Alberts et al., 1994), thus random fluctuations characterize the reaction rates, which they regulate. Recent experimental results (Elowitz et al., 2002; Blake et al., 2003) have demonstrated that gene expression is a stochastic process and works on small synthetic genetic networks, such as the genetic toggle switch of Gardner et al. (2000) and provided additional experimental data on the phenotypic variability originating from stochastic intracellular events. This type of heterogeneity is the so-called *intrinsic* heterogeneity.

The second source of heterogeneity in an isogenic cell population originates from the uneven distribution of the amounts of most intracellular components of mother cells (with the exception of DNA) between the daughter cells. The variability in the daughter cell content and especially in the number of regulatory molecules leads to different phenotypes. This phenomenon repeats itself due to the process of cell cycle and thus leads to further phenotypic variability. The type of heterogeneity originating from the uneven partition of mother cells' intracellular content among the daughter cells is called *extrinsic*. It has been shown experimentally in Elowitz et al. (2002) that extrinsic rather than intrinsic heterogeneity is quantitatively more significant for a wider range of induction levels in various *E. coli* strains. In this paper we will focus on the extrinsic type of heterogeneity.

The basic fundamental question arisen is to illuminate the mechanism which for a given gene regulatory network operating at the single-cell level affects the phenotypic distribution at the cell population level. The quantitative understanding of cell population heterogeneity and its implications in the behavior of the entire cell population is of great significance and the used mathematical model has to be capable of describing the dynamics of a cell population and satisfy two basic features: (a) take into consideration the intrinsic heterogeneity amongst the cells of the isogenic population and (b) include the mathematical formulation of the intracellular processes characterizing the gene regulatory network at each cell.

This can be accomplished with the use and application of a special class of models, known as cell population balance (CPB) models first formulated by Fredrickson and his co-workers in the mid 1960s (Eakman et al., 1966; Tsuchiya et al., 1966; Fredrickson et al., 1967). These are number balances where the main unknown of the problem is the number of cells which at time t have intracellular content between x and $x + \mathrm{d}x$. CPBs are nonlinear partial integro-differential equations and they are characterized by considerable mathematical complexity. In the general case, the CPB problem cannot be solved analytically, although some very interesting approaches exist for simplified versions of the problem (Liou et al., 1997). There exists a body of work focusing on the development of efficient numerical algorithms for the solution of this complex problem and spans over several decades. Subramanian and Ramkrishna (1971) solved the cell population balance with the method of weighted residuals using global basis functions. Orthogonal collocation on finite elements to discretize a simplified version of a yeast cell population model has been applied in Zhu et al. (2000); the same numerical scheme has been applied for bifurcation analysis (Zhang et al., 2002) and model reduction (Zhang et al., 2003) of yeast population models. Spectral algorithms have also been developed for the solution of population balance models in Mantzaris et al. (2001b, 2002). The common feature of these algorithms is that they all assume that the boundaries of the physiological state space (or intracellular content x) are fixed and known *a priori*. However, although the minimum intracellular content might be known from the measurement of the initial distribution, this does not apply to the maximum. Thus, in reality the distribution evolves in an unknown domain. Furthermore, it has been recently shown that CPB models are able to predict multiple solutions at balanced growth (steady state behavior) (Mantzaris, 2006). These steady states on average can vary, in principle, over several orders of magnitude. Hence, a classical fixed boundary algorithm will either fail completely or will produce very poor and inaccurate results for this kind of situations. Clearly, moving boundary algorithms are required to treat these challenges.

The CPB problem formulation is given in Section 2, where we also indicate the inefficiency of up to present numerical algorithms to deal with the aspect of unknown boundaries of the physiological state space. The proposed free boundary algorithm is described in Section 3. The domain of intracellular state space is normalized with respect to its average expression which is treated as an additional degree of freedom for the problem. We illustrate the efficiency of the algorithm for isogenic populations carrying plasmids of positive feedback architectures. The genetic network of interest in this paper, the *lac operon*, is presented in Section 4. The nonlinear nature of the reaction expression, in which the intracellular content participates, gives rise to bistable behavior over a significant region of extracellular inducer concentrations. In order to validate whether the bistable behavior observed at the single-cell level is also present at the population level we numerically solve the CPB problem, with the finite element method using the commercial software package COMSOL Multiphysics.

In Section 5 we present results of the free boundary algorithm application; we choose among different order basis functions the optimal, which compromises numerical stability and accuracy for a given number of degrees of freedom. Furthermore, we study the transient and asymptotic behavior of cell populations, carrying the *lac operon* network. The steady state version of the CPB model is also incorporated in a pseudo-arc-length continuation algorithm and enables the efficient study of the asymptotic behavior of a cell population as a function of the extracellular inducer concentration. The extension of the proposed algorithm for systems of higher dimensionality is presented in Section 6; we apply the algorithm to a cell population with two intracellular species and linear reaction expressions for each of the independent variables. When the intracellular contents are evenly distributed amongst the daughter cells, an oscillatory behavior is expected, as shown in Mantzaris et al. (2001a). The proposed algorithm not only captures the dynamic behavior of such a system; it also predicts that the oscillations do not persist for asymmetric distribution of the intracellular contents during cell division. Finally, in Section 7 we briefly summarize the main results of this study.

2. Cell population balance modelling

Consider a cell population, where each individual is distinguished by its intracellular content x (in molecules or moles) of a given substance. The model which can respect the intrinsic heterogeneity