Model of population heterogeneity and metabolic regulation in bioreactors

Mark A. Pinto and Charles D. Immanuel

Dept. of Chemical Engineering and Chemical Technology, Imperial College, London SW7 2AZ, UK
c.immanuel@imperial.ac.uk; Phone: +44 (0)20 7594 5594; Fax: +44 (0)20 7594 6606

Abstract— In this paper, a population balance model of a bioreactor is discussed. The importance of accounting for the heterogeneity of the cell population with respect to key metabolites is elicited. Further, a framework is presented for incorporating key regulatory metabolisms into the model. The preliminary results provide qualitative consistency with what is expected based on process understanding, and also quantitative consistency among the different cases studied so far. These serve as a first step to a more rigorous model validation that will be pursued in future studies.

Introduction

There is a growing interest in exploiting biological processes for the production of commodity chemicals, say for example biopolymers [1]. Thus, there is a large need to better understand and model bioprocesses and bioreactors. The traditional approaches to bioreactor modelling are based on the so-called lumped approach, which assume uniformity and homogeneity in the cell population and their behaviour. This approach results in simple models such as the Monod models, which are easy to develop but are highly inadequate in most cases [2]. In any population, there is a wide variation in the activities of the cells [3]. While a few are in the viable and productive stage, others are at rest while even others are in a senescent or death stage. There is a complex internal metabolic scheme that determines the particular state of the cell. Further, the cells participate in a repeated cycle of changes. Thus, there is a need to explicitly account for the distributed character of the cell populations.

In addition to this internal population heterogeneity, there are complex schemes of external regulations that drive the cells through the cell cycles. For instance, in multicellular organisms, the growth of individual cells is carefully regulated through the mitogen activated protein kinase (MAPK) signalling pathway. In this pathway, the binding of an extra-cellular growth factor to a ligand on the cells surface triggers a signalling cascade which results in the conversion of the protein Ras from its inactive GDP (guanosine diphosphate) bound state to its active GTP (guanosine triphosphate) bound state. The activation of Ras then triggers a further cascade of enzyme catalysed reactions that culminate in the entry of the cell into the growth phase of its cell cycle from its dormant state. Thus, unlike unicellular organisms where cellular growth is dependent only on available substrate, growth of cells in multicellular organisms is dependent not only on available substrate but on extracellular regulation as well. It is therefore important to consider these regulatory aspects when modelling multicellular organisms.

In this paper, one of the first attempts to develop a combined population balance model and a model for regulatory mechanisms is discussed. Moreover, as a significant departure from most previous population balance studies, a multi-dimensional characterisation of the population heterogeneity is adopted in this study. A reliable solution technique is employed for these highly computation-intensive models.

Cell cycle and regulation

A realistic model of the cellular biological processes has to account for the cell cycles in the heterogeneous medium. The cells in the population undergo a cycle of changes through their life time. The individual cells synthesise metabolites including proteins, RNA and DNA, and thereby grow in size. Upon the duplication of the genetic carriers – the DNAs, the cells undergo cell division into two usually identical daughter cells, contributing new and younger members to the population. The cell cycle is divided into two regimes – the interphase and the mitosis phase (M-phase). The interphase is that regime of the cell cycle that characterises the synthesis of metabolites, cell growth and DNA replication. The M-phase characterises the cell division. In the case of the Eukaryotes, the interphase is further divided into the G1 phase, the S phase, and the G2 phase. The G1 phase accounts for the synthesis of the metabolites required for DNA replication, the S phase accounts for DNA synthesis, the G2 phase accounts for the gap between DNA synthesis and cell division that serves to ensure proper replication of the DNA. There could also be a G0 phase, which represents an additional gap during which the cells are at rest. The cell division (the M phase) is itself

divided into several phases (Prophase, Metaphase, Anaphase and Telophase) that represent the processes of separation of the DNA, cytokinesis and formation of the cell membrane dividing the daughter cells. The cells usually cease to divide (senescent state) after a certain number of generations of reproduction. Cell senescence usually sets in after 60-80 generations. The cells also undergo apoptosis and necrosis (cell death). The duration of the cell cycle depends upon the organism, and in the human body, the cell cycle duration also depends upon the organs that the cells constitute.

The G1 and S phases are characterised by the complex metabolic reaction scheme that occurs within the cells. The amount of information available on the various pathways (glycolysis pathway, TCA cycle pathway *etc.*) far exceeds the amount that can be modelled reasonably with the available computational resources. Thus, in developing models for bioreactors and for biomedical applications, the metabolic pathways that are the most relevant will need to be adopted. Techniques such as the cybernetic modelling concepts [4] can be exploited to account for both the simplification that is introduced into the metabolic pathways, and more importantly to account for the selectivity and regulations that are inherent within the cells.

The passage from one phase to the other is dictated by several regulatory mechanisms (check points), some of which are triggered by external impulses. The most important regulation is the RAS signalling, which regulates growth and proliferation as well as survival of the cell populations. The RAS signalling operates through several pathways such as the RAF pathway and the Mitogen-Activated Protein Kinase (MAPK) pathway, which determine the transition of the cells into the synthesis phase of the cell cycle [5]. The protein p53, a type of cyclin activated kinase inhibitor, is crucial to maintain normal cell cycles by supressing cell proliferation and regulating cell growth. Mutants of p53 and the perturbations that appear in these regulatory mechanisms are the major reasons behind the occurence of cancer and malignancy, as has been borne out by several studies [6], [7]. In addition, the cell senescence is contingent on telomere shortening (telomeres being precursors to replicated DNAs), which is perturbed by the expression of the enzyme telomerase that leads to 'immortality' of cells (return from the senescent state back to a proliferative state) and hence to cancer [6]. Similarly, cell apoptosis is regulated by the transforming growth factor TGF-β [8]. Population balance provides a natural framework to model these various mechanisms.

Population Balance Model for Bioreactors

In this study, a model is developed with specific regard to a bioreactor. The general population balance model is given in Equation (1). In this equation, $F(\mathbf{x}, \mathbf{t})$ accounts for the population density of the cells at generation n (after n divisions), $\mathbf{x}$ accounts for the internal coordinates (the concentration of the metabolites within the cells), and N is the number of internal coordinates considered in the problem. $\frac{dx_i}{dt}$ accounts for the rate of change of the metabolites due to kinetic pathways within the cells. $\Re_{div}(\mathbf{x}, \mathbf{t})$ accounts for the formation of cells due to cell division undergone by the mother cells. Likewise, $\Re'_{div}(\mathbf{x}, \mathbf{t})$ accounts for the depletion of mother cells due to cell division. $\Re_{div}(\mathbf{x}, \mathbf{t}) = \mathbf{0}$ beyond the threshold generation at which cell senescence occurs. In bioreactor models, cell senescence is usually not of significance as the residence time of the cells is much shorter. But this assumes importance for cancer modelling. $\Re_{death}(\mathbf{x}, \mathbf{t})$ accounts for cell death and $\Re_{cell-cell}(\mathbf{x}, \mathbf{t})$ can be employed to account for cell-cell interactions such as the recombination aspects in bacterial populations.

$$\frac{\partial}{\partial t}F(\mathbf{x}, \mathbf{t}) + \Sigma_{\mathbf{i=1}}^{\mathbf{N}} \frac{\partial}{\partial \mathbf{x_i}}(F(\mathbf{x}, \mathbf{t})\frac{\mathbf{dx_i}}{\mathbf{dt}}) = \Re_{\mathbf{div}}(\mathbf{x}, \mathbf{t})$$
$$-\Re'_{div}(\mathbf{x}, \mathbf{t}) - \Re_{\mathbf{death}}(\mathbf{x}, \mathbf{t}) \quad (1)$$

The cell division terms are modelled as in Equations (2) and (3). The extracellular substrate and product metabolites are modelled as in Equations (4) and (5), respectively, accounting for the diffusion out of the cells and any deactivation in the aqueous medium. D is the dilution rate (reciprocal of the residence time) in the bioreactor.

$$\Re'_{div}(\mathbf{x}, \mathbf{t}) = \mathbf{\Gamma}(\mathbf{x})F(\mathbf{x}, \mathbf{t}) \quad (2)$$

$$\Re_{div}(\mathbf{x}, \mathbf{t}) = \int_{\mathbf{x_1}}^{\mathbf{x_1,max}} \int_{\mathbf{x_2}}^{\mathbf{x_2,max}} \int_{\mathbf{x_3}}^{\mathbf{x_3,max}} \int_{\mathbf{x_4}}^{\mathbf{x_4,max}} \int_{\mathbf{x_5}}^{\mathbf{x_5,max}}$$
$$\int_{x_6}^{x_6,max} \Gamma(x)F(x,t)dx_1 dx_2 dx_3 dx_4 dx_5 dx_6 \quad (3)$$

$$\frac{dx_{ex}^s}{dt} = Dx_0 - \int_{\mathbf{x}} \kappa(x_{ex}^s - x^s)F(x,t)d\mathbf{x} - Dx_{ex}^s \quad (4)$$

$$\frac{dx_{ex}^p}{dt} = \int_{\mathbf{x}} \kappa(x^p - x_{ex}^p)F(x,t)d\mathbf{x} - k_w x_{ex}^p - Dx_{ex}^p \quad (5)$$

The cell death events are modelled as below:

$$\Re_{death}(\mathbf{x}, \mathbf{t}) = \mathbf{\Gamma_d}(\mathbf{x})F(\mathbf{x}, \mathbf{t}) \quad (6)$$

In this particular study, the glycolytic pathway is being considered. The key internal metobolites chosen are the same as the ones chosen in a previous study [9], [10]. These are glucose (x_1); glyceraldehyde 6-phosphate (x_2); 1,3 biphospho glycerate (x_3); pyruvate (x_4); NADH (x_5); and ATP (x_6). The growth models for these internal metabolites are given in Equations (7)-(12): The nutrient glucose is a continuous input to the cells (J_0 in Equation (7)), while

the pyruvate (x_4) is a metabolite produced by the cells that is excreted to the aqueous medium (with mass transfer coefficient κ). x_{5T} accounts for the combined NAD and NADH, which is assumed to be constant although the two quantities might change individually. Likewise, x_{6T} accounts for the combined ATP and ADP, which is also assumed to remain constant.

$$\frac{dx_1}{dt} = J_0 - \frac{k_1 x_1 x_6}{1+(\frac{x_6}{K_I})^q} \tag{7}$$

$$\frac{dx_2}{dt} = \frac{2k_1 x_1 x_6}{1+(\frac{x_6}{K_I})^q} - k_2 x_2 (x_{5T} - x_5) - k_6 x_2 x_5 \tag{8}$$

$$\frac{dx_3}{dt} = k_2 x_2 (x_{5T} - x_5) - k_3 x_3 (x_{6T} - x_6) \tag{9}$$

$$\frac{dx_4}{dt} = k_3 x_3 (x_{6T} - x_6) - k_4 x_4 x_5 - \kappa(x_4 - x_{4,ex}) \tag{10}$$

$$\frac{dx_5}{dt} = k_2 x_2 (x_{5T} - x_5) - k_4 x_4 x_5 - k_6 x_2 x_5 \tag{11}$$

$$\frac{dx_6}{dt} = -\frac{2k_1 x_1 x_6}{1+(\frac{x_6}{K_I})^q} + 2k_3 x_3 (x_{6T} - x_6) - k_5 x_6 \tag{12}$$
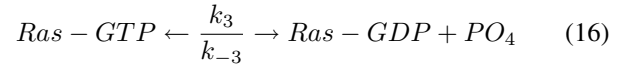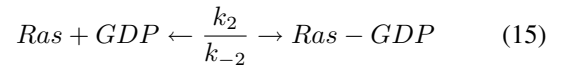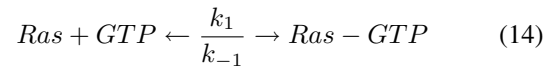
Model for regulatory metabolism

As indicated previously, the RAS signalling and the MAPK pathways are the major regulators of cell growth. This pathway consists of a series of steps in an enzyme cascade that culminates in the upregulation of genes responsible for cell cycle progression and cell division. The signalling cascade begins when insulin binds to the insulin receptor which then undergoes autophosphorylation. The insulin receptor then phosphorylates the protein insulin receptor substrate-1 (IRS-1). The protein Grb2 then binds to the phosphorylated IRS-1. This is followed by the binding of the protein Sos to Grb2 and, subsequently, the binding of the protein Ras to Sos. Before Ras is bound to Sos, it is in its inactive state where it is bound to GDP. Binding of the inactive form of Ras to Sos, causes GDP release from Ras and the binding of Ras to GTP resulting in the active GTP bound form of Ras. Ras then activates Raf-1, the first of three protein kinases Raf-1, ERK (extracellular regulated kinase) and MEK (mitogen-activated ERK-activating kinase) that form a cascade in which each protein activates the next. The activated phosphorylated form of ERK then moves into the nucleus where it activates nuclear transcription factors such as Elk1 by phosphorylation. The activation of these nuclear transcription factors upregulates the transcription of genes needed for cell cycle progression and division which results in an increase in cell cycle proteins such as cyclins and, ultimately, cell cycle progression and cell division.

The MAPK signalling cascade as described above has been modelled in detail [11]. However, in the model presented here, a simplified reaction scheme has been adopted. This is necessary as the objective of the current work is to develop a multi-dimensional cell population balance model. Thus, individual aspects might need to be simplified to keep the multi-scale model computationally-tractable. In this model of Ras signalling, Ras is free to bind with either GDP or GTP. Only the Ras bound to GTP is active and can influence cell cycle progression. The growth rates of the state variables x are influenced by the concentration of GTP bound Ras, that is [Ras-GTP] as shown in Equation (13). Note that the unregulated rates are the ones shown in Equations (7)-(12). Thus, in effect, the intermediate steps in the regulation mechanism are lumped together, which approach is also justified by the lack of detailed kinetic parameters for a more elaborate scheme. Ths simplified scheme is shown in Equations (14) - (16). The concentration of the Ras bound GTP is in turn modelled employing Equations (17)-(21) based on the adopted simplified scheme. In this model, it is assumed that the concentration of phosphate, PO_4, is sufficiently high to be considered constant.

$$\frac{dx_1^{regulated}}{dt} = \frac{dx_1^{unregulated}}{dt} \times [Ras - GTP] \tag{13}$$

$$Ras + GTP \leftarrow \frac{k_1}{k_{-1}} \rightarrow Ras - GTP \tag{14}$$

$$Ras + GDP \leftarrow \frac{k_2}{k_{-2}} \rightarrow Ras - GDP \tag{15}$$

$$Ras - GTP \leftarrow \frac{k_3}{k_{-3}} \rightarrow Ras - GDP + PO_4 \tag{16}$$

$$\frac{[Ras]}{dt} = -k_1[Ras][GTP] + k_{-1}[Ras - GTP] - k_2[Ras][GDP] + k_{-2}[Ras - GDP] \tag{17}$$

$$\frac{[Ras-GTP]}{dt} = k_1[Ras][GTP] - k_{-1}[Ras - GTP] - k_3[Ras - GTP] + k_{-3}[Ras - GDP][PO_4] \tag{18}$$

$$\frac{[Ras-GDP]}{dt} = k_2[Ras][GDP] - k_{-2}[Ras - GDP] + k_3[Ras - GTP] - k_{-3}[Ras - GDP][PO_4] \tag{19}$$

$$\frac{[GTP]}{dt} = -k_1[Ras][GTP] + k_{-1}[Ras - GTP] \tag{20}$$

$$\frac{[GDP]}{dt} = -k_2[Ras][GDP] + k_{-2}[Ras - GDP] \tag{21}$$

RESULTS AND DISCUSSION

Numerical simulation results under different cases are presented next. For the six-dimensional cell population balance model formulated above, the rate equations are computationally intensive and stiff, and using traditional solution techniques such as discretisation methods or the methods of weighted residuals would result in a large

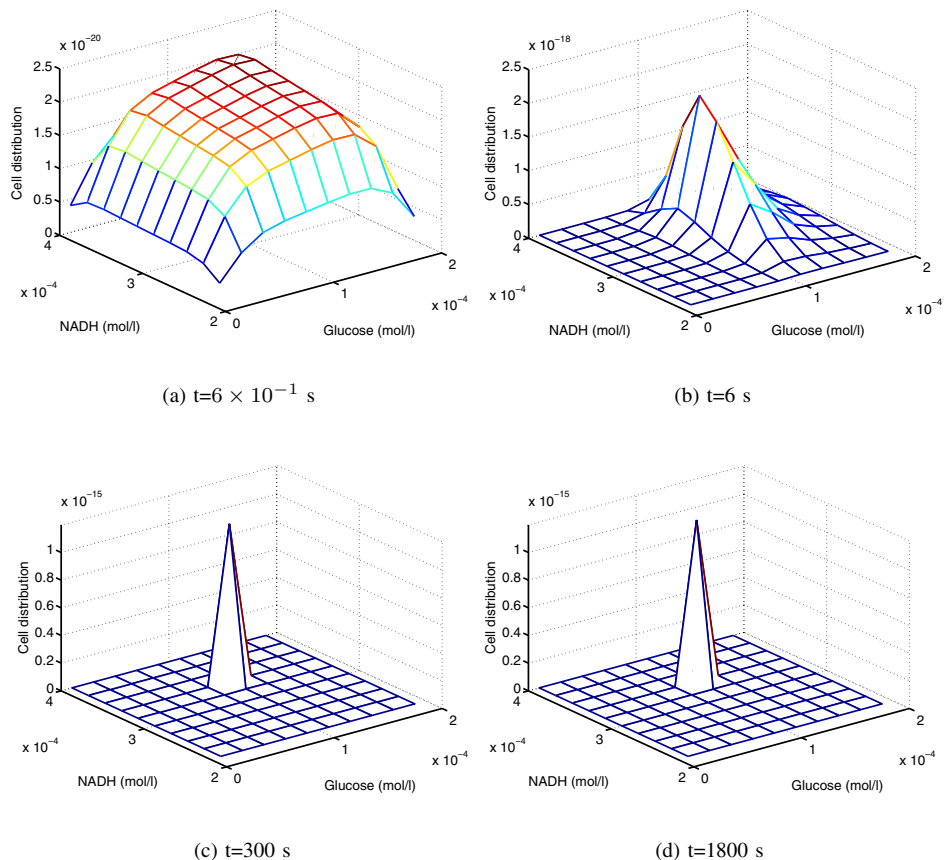(a) t=6 × 10^{−1} s

(b) t=6 s

(c) t=300 s

(d) t=1800 s

Fig. 1.   Two-dimensional plots of the evolution of the cell populations that depicts a synchronisation of the population.

system of equations and, consequently, very large computational requirements, even if the discretisation parameters were kept small. The largest cell population balance model solved using traditional methods was three-dimensional [12]. Simplified kinetics were used and, further, cell division, growth, differentiation and death as well as their regulation were neglected. A Monte Carlo type method involving the extrapolation of data obtained from a few cells (akin to statistical techniques) has subsequently been used to simulate a six-dimensional population distribution [9]. More recently, a finite volume technique proposed to solve multi-dimensional population balance models [13] was found to be significantly more efficient and accurate than the Lax-Wendroff finite difference method which, in turn, however, was found to be the least accurate of several finite difference methods evaluated to solve single-dimensional cell population balance models [12]. Further studies are therefore necessary to compare the finite volume technique [13] with the Hierarchical Two-Tier solution strategy employed here [14], [10]. In this strategy, the rates of the contributing sub-processes are calculated in the first tier and the population balance distribution is updated in the second tier. Semi-

analytical solutions have been derived for rate processes such as cell division [15] which are in terms of complex and computation-intensive integrals, thereby reducing the integrals to simpler terms major portions of which are computed once and used throughout. These measures result in a substantial reduction of the computational load rendering the technique at least comparable in efficiency to the statistically-based techniques used before.

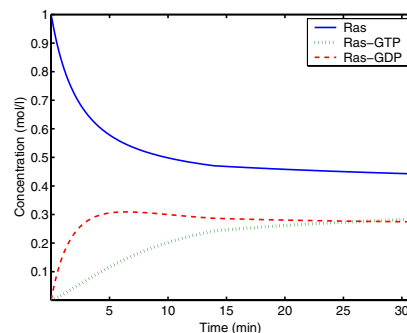The six-dimensional population balance model as formu-

Fig. 2.   The variation of key metobolites that participate in the Ras signalling and regulation of cell growth.

(a) t=6 s

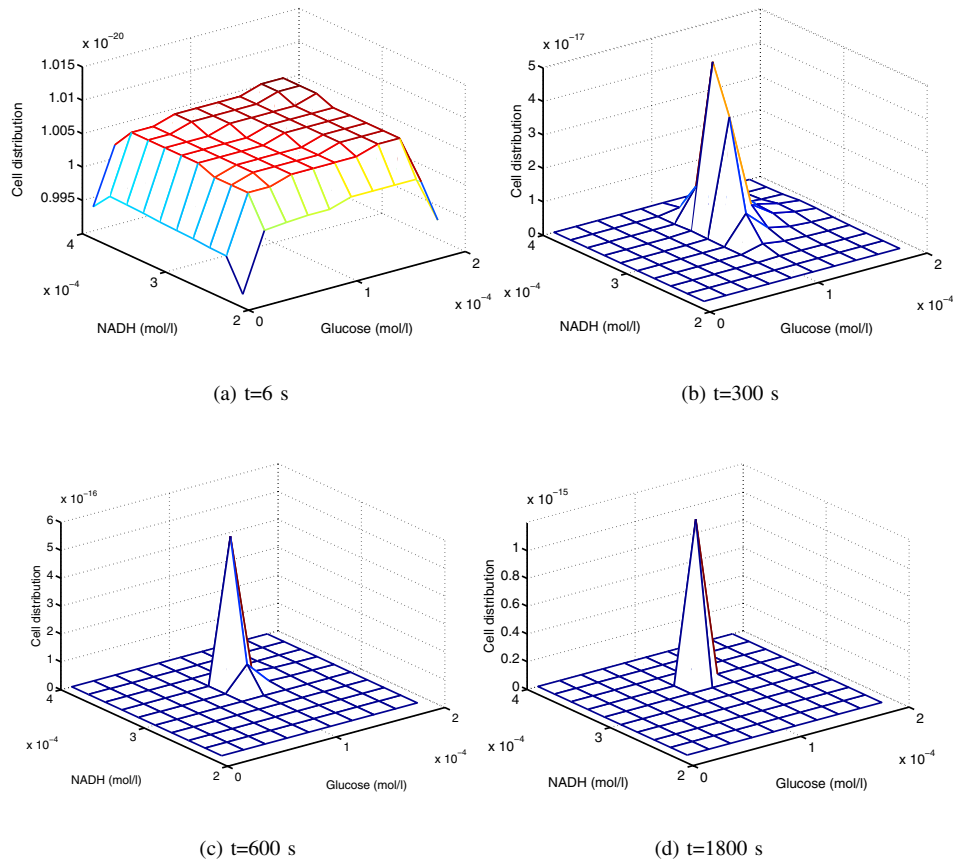(b) t=300 s

(c) t=600 s

(d) t=1800 s

Fig. 3. The effect of Ras signalling and cell growth regulation of the evolution of the cell populations. The figures indicated a reduced rate of synchronisation compared to Figure 1.

lated above was initially simulated without considering the death and division phenomena. In simulating the model, the cell population is initially assumed to be distributed uniformly. The results from this simulation, as shown in Figure 1, indicate the synchronisation of the cell population, an effect which has been observed by other workers [9]. Most of the synchronisation is complete in five minutes (Figure 1c) in this case. The co-ordinates of G3P/DHP, 1,3 BPG, pyruvate and ATP at which this synchronization takes place are 0.95 mM, 0.10 mM, 0.10 mM and 0.60 mM respectively.

The results presented next discuss the effect of incorporation of the Ras signalling aspects as formulated previously, into the population balance model. The cell division and death phenomena are neglected as in the previous case. The variations in concentrations of Ras, GTP bound Ras and GDP bound Ras are shown in Figure 2. It can be seen that after approximately thirty minutes, the system reaches a steady-state. The simulation of the population balance model incorporating this Ras signalling results in a decrease in the rate at which the cell population becomes synchro-

nized (Figures 3). This is particularly evident initially when the concentration of GTP bound Ras is very low. However, most of the cell population does get synchronized in thirty minutes. The incorporation of Ras signalling into the cell population balance model also influences the dynamics of extracellular pyruvate as can be seen from Figure 4. Without Ras signalling incorporated, the extracellular concentration of pyruvate first drops before rising again to reach a steady state. However, when Ras signalling has been incorporated, the drop in the extracellular concentration of pyruvate is much more pronounced though the steady state concentration is the same as in the case without Ras signalling. The overall diminishment of the metabolic rates results in this non-intuitive result when viewed in the face of the entire population. Thus, it is evident that the introduction of Ras signalling into the cell population balance model influences the dynamics of the model.

In the results presented next, the population balance with both cell growth and cell division is discussed. In order for the model to be realistic, a value for the division kernel must be chosen that results in a realistic cell doubling time.
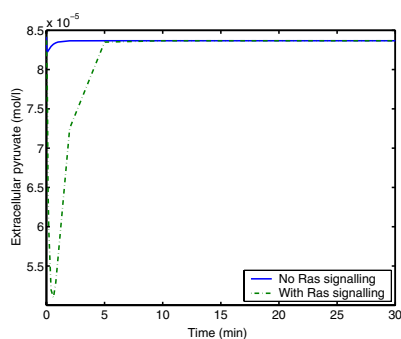
Fig. 4. The effect of Ras growth signalling on the extracellular concentration of a key metabolite.

Fig. 5. Variation of the doubling time with the value of the division kernel.

Therefore, the cell population balance model was simulated for several values of the kernel to see the resulting effect on cell doubling time. These simulations were carried out for both cases: with and without the incorporation of Ras signalling. As can be seen from Figure 5, cell doubling time varies exponentially with the division constant. Also, as can be seen, the incorporation of Ras signalling does not affect the doubling time of the cells. This is to be expected, as the Ras signalling only affects cell growth, with effects on cell division being indirect and less strong. The doubling time of mammalian cells being approximately 24 hours [16], a corresponding cell division constant value of $k_{div} = 5.00 \times 10^{-4}$ is chosen for this study. It is seen that with this choice of the cell division constant, the effect of the cell division phenomenon on the cell population distribution is minimal and less obvious, although there is a profound effect on lumped variables such as the total number of cells in the population and extracellular product concentrations (figure not shown for these here). Similar studies were performed to incorporate cell death and apoptosis into the model. The immediate next step in these studies is to incorporate the regulatory mechanisms that influence the cell division and cell death processes to further enhance the multi-scale character of the model. The next step will be to carry out experimental studies for model validation and future control studies.

CONCLUSION

The population balance modelling of biological systems, and bioreactors in particular, was discussed. The study is based on the precept of the inadequacy of a lumped model to capture the complex mechanisms underlying biological processes [2]. Thus, a model that accounts for the heterogeneity in the cell population is presented. As a significant departure from previous population balance models of bioreactors that mainly account for distribution in the cell mass, the heterogeneity in the population with respect to several key cell
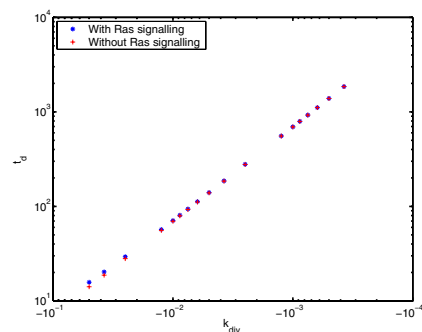
metabolites is accounted for in the model [9]. Further, this is one of the first attempts to simultaneously account for the complex regulatory metabolism into the population balance models. The first results presented here are qualitatively and quantitatively consistent with expectation, encouraging more rigorous validation efforts. An efficient numerical scheme was employed to solve these computation-intensive and multi-scale models [17], [10].

REFERENCES

[1] G. Stephanopoulos, "Invited comment: Chemical and Biological Engineering," *Chem. Eng. Sci.*, vol. 58, pp. 4931–4933, 2003.
[2] P. Daoutidis and M. A. Henson, "Dynamics and Control of Cell Populations in Continuous Bioreactors," in *AIChE Symposium Series: Chemical Process Control-VI*, ser. 326, J. B. Rawlings, B. A. Ogunnaike, and J. W. Eaton, Eds., vol. 98, 2001, pp. 274–289.
[3] D. L. Nelson and M. M. Cox, *Legninger Principles of Biochemistry*. W. H. Freeman and Co., New York, 2004.
[4] D. S. Kompala, D. Ramkrishna, and G. T. Tsao, "Cybernetic Modeling of Microbial Growth on Multiple Substrates," *Biotech. Bioeng.*, vol. 26, pp. 1272–1281, 1984.
[5] J. Downward, "Targetting RAS Signalling Pathways in Cancer Therapy," *Nature Reviews: Cancer*, vol. 3, pp. 11–22, 2003.
[6] N. F. Mathon and A. C. Lloyd, "Cell Senescence and Cancer," *Nature Reviews: Cancer*, vol. 3, pp. 203–213, 2001.
[7] D. Ruggero and P. P. Pandolfi, "Does the Ribosome Translate Cancer?" *Nature Reviews: Cancer*, vol. 3, pp. 179–192, 2003.
[8] P. M. Siegel and J. Massague, "Cytostatic and Apoptotic Actions of TGF-beta in Homeostasis and Cancer," *Nature Reviews: Cancer*, vol. 3, pp. 807–820, 2003.
[9] M. A. Henson, D. Muller, and M. Reuss, "Cell population modelling of yeast glycolytic oscillations," *Biochem. J.*, vol. 368, pp. 433–446, 2002.
[10] C. D. Immanuel, "Population Balance Model for Cellular Processes in Biological Systems: Biochemical and Biomedical Applications," *Proceedings of IFAC Dynamics and Control of Process Systems, Cambridge, MA*, 2004.
[11] A. R. Asthagiri and D. A. Lauffenburger, "A Computational Study of Feedback Effects on Signal Dynamics in a Mitogen-Activated Protein Kinase (MAPK) Pathway Model," *Biotechnol. Prog.*, vol. 17, pp. 227–239, 2001.
[12] N. V. Mantzaris, P. Daoutidis, and F. Srienc, "Numerical solution of multi-variable cell population balance models. i. finite difference methods," *Computational Chemical Engineering*, vol. 25, pp. 1411–1440, 2001.
[13] R. Gunawan, I. Fusman, and R. D. Braatz, "High resolution algorithms for multidimensional population balance models," *AIChE Journal*, vol. 50(11), pp. 2738–2749, 2004.
[14] C. D. Immanuel and F. J. Doyle III, "Efficient Solution Technique for Multi-dimensional Population Balance Model Describing Granulation Processes," *AIChE Annual Meeting, Particle Technology Forum, San Francisco*, November 2003.
[15] M. A. Pinto, *Population balance modelling of eukaryotic cells*. M.Sc. thesis, Imperial College London, University of London, 2004.
[16] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. Garland Science, New York, 2002.
[17] C. D. Immanuel and F. J. Doyle III, "Computationally-Efficient Solution of Population Balance Models Incorporating Nucleation, Growth and Coagulation," *Chem. Eng. Sci.*, vol. 58, no. 16, pp. 3681–3698, 2003.