



Regular article

Experimentally integrated dynamic modelling for intuitive optimisation of cell based processes and manufacture

Adrian J. Stacey^{a,b}, Elizabeth A. Cheeseman^a, Katie E. Glen^a, Rebecca L.L. Moore^a, Robert J. Thomas^{a,*}

^a Centre for Biological Engineering, Wolfson School of Mechanical and Manufacturing Engineering, Loughborough University, Loughborough, Leicestershire, LE11 3GR, UK

^b Advanced Bioprocess Design Ltd, 90 Main Street, East Leake, Loughborough, LE12 6 PG, UK

ARTICLE INFO

Article history:

Received 27 July 2017

Received in revised form 4 December 2017

Accepted 10 January 2018

Available online 10 January 2018

Keywords:

Cell therapy

Manufacture

Model

Bioprocess

Dynamic

Mechanistic

ABSTRACT

Dynamic mechanistic modelling of cell culture is a key tool in bioprocess development to support optimisation and risk assessment. However, the approach is underutilised in the bioprocess industry due to challenges including lack of accessible tools to support a structured approach, the difficulty of realising computationally tractable (low parameter) yet realistic models, and the specialised skill sets required. We have proposed that these issues could be partly addressed by developing a parsimonious framework comprising a set of model building blocks, based on the ordinary differential equation modelling paradigm, representing common cell culture dynamics and modulation thereof. The framework is designed to avoid obvious pathological behaviours. Further, specific model instances within the framework can be simply visualised as a directed graph with vertices representing system species, dynamics and modulations, and arcs representing the interactions between them. The directed graph can be extended to describe the timing and nature of experimental interventions. A visual and intuitive route to describing models with an associated mathematical framework enables realisation in a software interface and integration with standard mathematical tools such as those for sensitivity analysis and parameter estimation. Such a framework is sufficient to represent some of the simple mechanisms underpinning bioprocesses that nonetheless lead to highly non-linear and counterintuitive outcomes. It also has a relatively low learning burden for users without formal mathematical training. The concept could be extended to include, for example, partial differential equation-based approaches to stochastic or spatially complex systems built up from a small number of parametrically parsimonious and well-behaved building blocks.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cell culture processes are a major element of manufacture in many biologic and emerging cell-based therapies. The dynamics of these processes, such as cell growth, consumption and production rates, have long been recognised as important determinants of process outcomes. For instance, in protein producing cell line cultures, reduction of major nutrients and accumulation of metabolic by-products such as ammonia and lactate can result in growth inhibition, reduced culture viability, and altered product titre [1]. Evidence is accumulating that newer cell-based products proposed

in the regenerative medicine field add further complexity. Cell-released factors (identified or unidentified), metabolic substrate availability, and metabolic by-products can have tissue specific feedback relationships with lineage trajectory and growth rate, and as such are likely to be highly product specific [2–4]. Each candidate process and product will require appropriate understanding, description, and control of these dynamic relationships to achieve product optimisation and robust process control [5,6].

Best practice of process development and optimisation, as articulated in structured approaches such as Quality by Design or Six Sigma, has quantitative modelling at its heart [7]. In cell culture, current models fall roughly into two camps, namely those that empirically map between process parameters and target outcomes and those that consider the dynamics and mechanisms by which process parameters affect process outcomes. Empirical mapping treats the biological and experimental system as a “black box” and thus provides limited process insight for risk assessment, extrapo-

* Corresponding author.

E-mail addresses: adrian@conceptiverse.com (A.J. Stacey),

e.a.cheeseman@lboro.ac.uk (E.A. Cheeseman), k.e.glen@lboro.ac.uk (K.E. Glen),

r.l.l.moore@lboro.ac.uk (R.L.L. Moore), R.J.Thomas@lboro.ac.uk (R.J. Thomas).

lation or hypothesis-based iterative development [8]. Further, due to the complex and dynamic nature of the cell culture environment, experimental design or manufacturing conditions optimised for single time points or across time intervals often fail when considered over longer relevant time-courses. Conversely, a mechanistic approach entails the formulation and evaluation of hypotheses concerning the dynamics of the culture in terms of their consequences for culture trajectory. Mathematical and mechanistic formalisation of hypotheses concerning these dynamics represents the starting point effectively to link short time-scale dynamics with their consequences over a longer time-scale. There is increasing awareness, particularly within cell therapy, that such process insight is critical in diminishing risk in transfer to manufacture and in delivering consistent product quality [9,10].

Low parameter-count approaches of the empirical paradigm (exemplified by Design of Experiments) have been used extensively to improve cell based manufacture processes, including for cell therapies, facilitated by low-barrier and generic software tools for experimental design and analysis [11–13]. The mechanistic approach is employed by specialist groups within academia to elucidate biological systems and process knowledge, but application within commercial process development remains sporadic [14–16]. Consequently the hypothesis-testing power, and precision of hypothesis expression that mechanistic modelling enables, is significantly under-utilised, particularly early in product and process development when teams or companies are small and resources relatively limited. Reasons for restricted application include lack of low barrier turnkey tools and standardised workflows necessitating a diverse skill-set covering both biological hypothesis development and specialist modelling techniques. The tendency therefore is for the majority of biological experimentation and hypothesis testing, or of mathematical modelling of biological systems, to proceed in isolation, or in a poorly integrated manner. The requirement to address such deficiencies has been recognised in several technology road-mapping exercises such as that conducted by the National Cell Manufacturing Consortium [17].

We aim to facilitate wider application of dynamic mechanistic modelling by addressing the challenges that prevent the development of a broadly accessible turnkey software package for mechanistic cell process modelling. Such a package must (i) minimise mathematical knowledge required for model development (ii) enable precise articulation and interdisciplinary communication of population-dynamic hypotheses (iii) support development of relevant and robust mathematical models and (iv) allow linking of models to data from complex time-course experimentation for verification. To achieve this we aimed to develop a conceptual framework and mathematical formulation that could facilitate the expression of a broad range of biological phenomena in a consistent form and that could be conveniently expressed within a visual (software) interface to provide an intuitive bridge between biological description of a dynamic system and precise mathematical expression thereof.

2. Methods

2.1. Model framework development

Many complex time-courses can be efficiently described within an ordinary differential equation (ODE) modelling paradigm, in which the evolution of the system is described in terms of the dependency of the rates of change of the system variables on other system variables or additional temporal factors. Within this paradigm, mathematical expression of biological dynamics conventionally uses established functions, such as logistic and Monod (for macroscopic kinetics) and flux equations (microscopic kinet-

ics) [16,18]. However, such formulations often conflate multiple mechanisms, for example, in terms of growth and saturation, which prevents straightforward reconfiguration to express a full range of dynamic hypotheses. To address this we developed a modelling approach in which - by restricting the repertoire of mathematical forms to a set of carefully chosen building blocks - the constituents of a system, and the relationships between them can be expressed intuitively, the former in terms of natural language, and the latter in terms of directed graphs (digraphs). A digraph constitutes a system of asymmetric relationships in which directional arrows (arcs) define the relationships between points (vertices) that represent the elements in an organisational structure. A secondary benefit of a limited ODE approach is that it can be designed to ensure that ill-formed or badly behaved model formulations are naturally avoided (for example, those that would, under certain conditions, predict a negative quantity of a necessarily positive system constituent), in contrast to a naïve deployment of, for instance, flux equations.

From the perspective of a population or sub-population of cells the bulk of these dynamics can be characterised as:

- population growth due to cell division
- population decline due to cell death, and
- concomitant decline of one population and growth of a second due to interconversion of cells from one to the other, for example, due to a phenotypic change.

On an instantaneous time-frame, these dynamics can be approximated as being essentially additive. In terms of the primary modulators of population dynamics e.g. non-cell species within the media or direct cell feedback, the number of general forms is similarly small:

- production of a species by cells (such as a by-product of metabolic activity or a signalling molecule such as a cytokine)
- consumption (or destruction) of a species by the cells
- species decay, and
- conversion from one species to another, for example due to cell activity.

In the absence of further data during model formulation, each of these dynamics can be assumed to take the simplest possible form that avoids pathologies in the model, namely:

- cell growth:

$$\frac{dX}{dt} = rX$$

where X is the cell density and r the (positive) specific growth rate

- decline of cell population or species:

$$\frac{dX}{dt} = -rX$$

where X is cell density or species concentration and r the (positive) specific decay rate

- interconversion of population or species:

$$\frac{dX}{dt} = -rX$$

And

$$\frac{dY}{dt} = rX$$

where X and Y are the source and sink densities or concentrations respectively, and r the positive specific interconversion rate

- consumption or destruction of species:

$$\frac{dX}{dt} = \frac{-rk}{(k+X)}$$

where X is species density or concentration, r is the maximum specific rate of consumption or destruction, and k determines the value of X for which the consumption or destruction is half the value of r .

These dynamics may themselves be modulated by secondary factors. For example, cell growth may be inhibited in the absence of source nutrients or the presence of a growth inhibitor, cell death promoted by the presence of a toxin, species consumption promoted by the presence or activity of the cells. There is broad mechanistic support for these modulations in the metabolic and cell signalling literature; we propose the following main patterns of modulation:

- promotion when promoting factor has a threshold:

$$M = (1 - \exp(a(X - b)))^{-1}$$

where X is the density of the promoting factor (e.g. nutrients) and a and b denote the sensitivity and threshold for promotion respectively ($a, b > 0$);

- promotion without the requirement for a threshold:

$$M = X(a + X)^{-1}$$

where X is the density of the promoting factor (e.g. nutrients) and a ($a > 0$) is the concentration to reach half maximum rate (that is, Monod kinetics);

- inhibition:

$$M = 1 - (1 - \exp(a(X - b)))^{-1}$$

where X is the density of the inhibitory factor (e.g. metabolic by-products) and a and b denote the sensitivity and threshold for inhibition respectively;

- optimality:

$$M = \exp(-a(X - b)^2)$$

where X is the density of the critical factor for optimal activity (e.g. pH) and a and b denote the sensitivity and optimum respectively.

By limiting the options for both dynamics and modulation, model pathologies are avoided. Alternative formulations would enable the development of singularities; for example, if inhibition were expressed as $M = X^{-a}$ ($a > 0$); similarly, alternatives would support non-biological behaviour, such as driving nutrient concentration negative if consumption were expressed simply as $\frac{dX}{dt} = -aY$ where Y is the density of the consuming cells).

As a first approximation we consider a multiplicative relationship between modulatory factors and associated dynamics. Noting that modulation may depend either on the absolute system state

or the rate of change thereof, this results in a general model of the form:

$$\frac{dX_i}{dt} = \sum_j \left(D_{i,j}(\mathbf{X}) \prod_k M_{i,j,k} \left(\mathbf{X}, \frac{d\mathbf{X}}{dt} \right) \right)$$

where the vector \mathbf{X} represents the current state of the entire system, D the dynamics and M the modulator effects on these dynamics as described above.

2.2. Model expression and manipulation

To reduce the mathematical skill barrier for model utilisation, and hence broaden accessibility, it was necessary to develop an intuitive and common language method for model description. Due to the decoupling of the various components, the general structure of the mathematical model can be expressed as a directed graph (digraph) to visualise the relationships between the individual constituents (Fig. 1A). The vertices on the digraph represent intuitively meaningful phenomena: species, dynamics and modulation respectively (species can be used to denote a quantitative component of a system, such as the cell density, the concentration of some analyte or metabolite, or an abstraction of a bulk property of the cells, such as average cell maturity). The arcs from vertex to vertex correspond to interactions between these vertices. Predominantly, these represent the association of a dynamic with a given species (for example, a growth dynamic associated with cell density) and the association of a modulator with a given dynamic and a given species on which it depends (for example, inhibition of cell growth due to the presence of an inhibitor). Additional vertices are required to indicate modulation that depends on a dynamic (for example, the production of inhibitor at a rate related to the total growth rate of the producing cells) and the species that results from a conversion dynamic (for example, to indicate that the death of cells results in the production of dead cells from live ones). The digraph can naturally be extended to include experimental interventions, such as step changes in species, defined by a schedule and the nature of the operation (Fig. 1A).

3. Results

3.1. Model development and evaluation

The general model can be used to express a wide array of specific mechanistic hypotheses which can be visualised as digraphs. A very simple example is a system in which cells constitutively produced an autocrine growth inhibitor; this model may be sufficient where the precursor of a metabolic inhibitor is saturating with respect to production of the inhibitor (such as glucose with respect to lactate in high glucose medium [19]), or an inhibitory cytokine is stably produced (Fig. 1B) [20]. An alternative configuration exemplifies a more complex system where a cell-produced species is involved in an autocrine survival loop, and a further extrinsic cytotoxic species is also present. This is similar to TNF α mediated survival and cytotoxic resistance in certain cell lines [21]. In this example we have further proposed that production of the survival species is also dependent on cell population growth rate, and the non-viable cells have a constant decay rate to remove from the system (Fig. 1C). Multiple further reconfigurations could be envisaged to introduce further hypothetical mechanisms such as growth inhibition in addition to cell death, further modulating species, or to describe nuances of mechanism such as promotion of growth as opposed to promotion of survival.

The digraph lends itself to straightforward presentation on a user interface (Fig. 2A). In this example, we show a further hypothetical model which incorporates two subpopulations of cells,

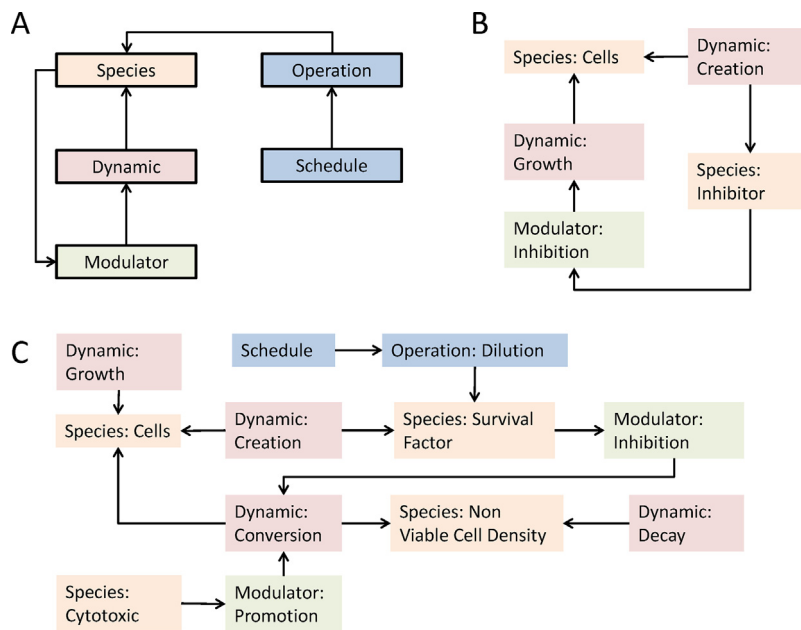


Fig. 1. Expression of general model structure in visual form. A) The model can be considered a directed graph (digraph), with the vertices corresponding to Species (orange), Dynamics (pink) or Modulators (green), and the arrows (arcs) to interactions between the vertices. Arrows denote the asymmetry between linked vertices: the arrow points to the affected and from the driving vertex. The digraph can naturally be extended to include experimental interventions (blue) B) An illustration of a specific instance of the general model corresponding to a system where cells constitutively produce an inhibitor which suppresses their own growth (e.g. cytokine-mediated feedback inhibition). C) An exemplar description of a more complex model (akin to autocrine mediated cell survival with an extrinsic cytotoxic factor) including an extension of the digraph to describe a scheduled point-dilution of the survival factor (e.g. media change) representing a step change in the media properties. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

whereby progenitor cells divide asymmetrically to produce a new version of the parent and a more mature lineage committed cell, which in turn produces an inhibitory cytokine which restricts the growth of the progenitors. This has parallels to negative feedback loops observed in multiple relevant systems e.g. haematopoietic stem cell expansion or caspase mediated inhibition of immature erythroblasts by more mature erythroblasts [2,22]. This results in an initial increase in progenitor population (growth of progenitors exceeds conversion to mature cells) followed by a decline (conversion to mature cells exceeds growth of progenitors due to the production of a cytokine that inhibits progenitor growth) (Fig. 2B). The vertices can be manipulated by the user to enable parameterisation of the model, and arcs created or removed to provide the appropriate interactions between these vertices, or to specify experimental or operational interventions (Fig. 2A). The close relationship between system conceptualisation and user interface representation expedites model development and exploration, making it natural for the user to ask “what if” questions and consider alternative hypotheses around system behaviour. This representation therefore constitutes a graphical but precise language by means of which key system dynamic concepts can be articulated.

3.2. Experimental intervention

Intervention in a biological culture system, via operations such as media exchange, dilution or cell sorting, is commonly used to test hypotheses of underlying mechanisms. Such interventions are also common to many production processes. The modelling framework proposed enables easy visualisation of a broad range of such interventions by a small extension to the digraph, in which interventions are associated with a schedule (to indicate when they occur) and the affected species (Figs. 1C, 2A). By allowing interventions of this kind to be integrated, the modelling framework enables the user the flexibility to describe and investigate *in silico*

possible experimental protocols for testing diverse hypotheses or simulating specific process operation (Fig. 2C). From a mathematical perspective, the formulation of such interventions is selected to avoid model pathologies such as those that would potentially drive variables negative for poorly chosen parameters:

- dilution:

$$X \rightarrow aX + (1 - a)b$$

where a ($0 \leq a \leq 1$) is the proportion of the original media or cell population that remains, and b the concentration or density in the replacement ($b \geq 0$);

- reset:

$$X \rightarrow b$$

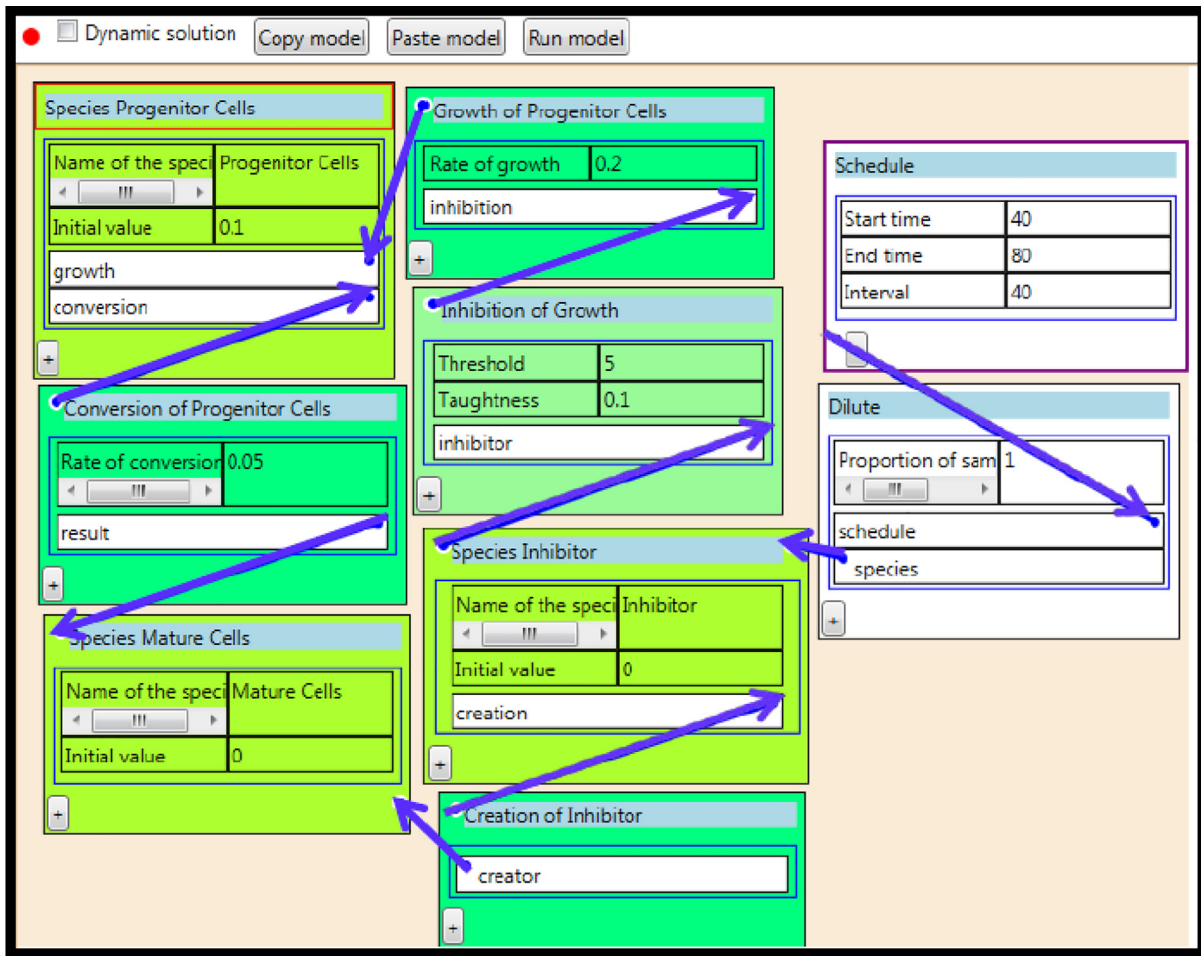
where b is the new cell density or species concentration ($b \geq 0$).

3.3. Model application for experimental design and process optimisation

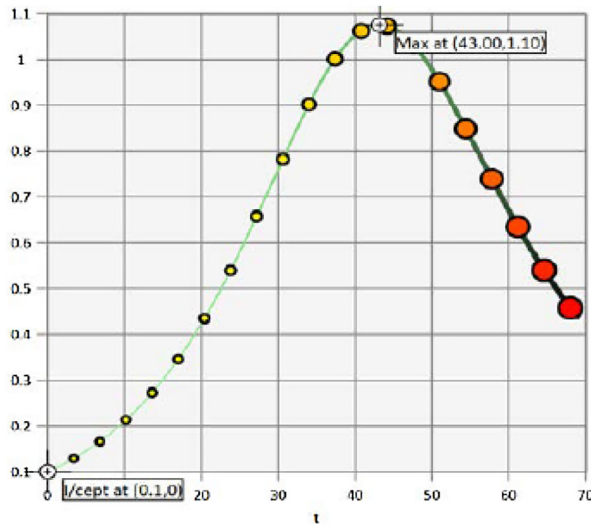
Due to the ease with which digraphs can be presented to, and manipulated by, users in the context of a piece of software, it is straightforward to realise this framework as a software tool. In conjunction with well-established numerical techniques for optimisation and sensitivity analysis, this provides a rapid conduit from model articulation to optimisation with respect to a production goal. Further, experimental design aimed at accurately estimating model parameters is facilitated by, for example, sensitivity analysis of process values with respect to both model parameters and experimental interventions.

Using the mature cell-progenitor negative feedback model as an exemplar we can show how important process decisions pertaining

A



B



C

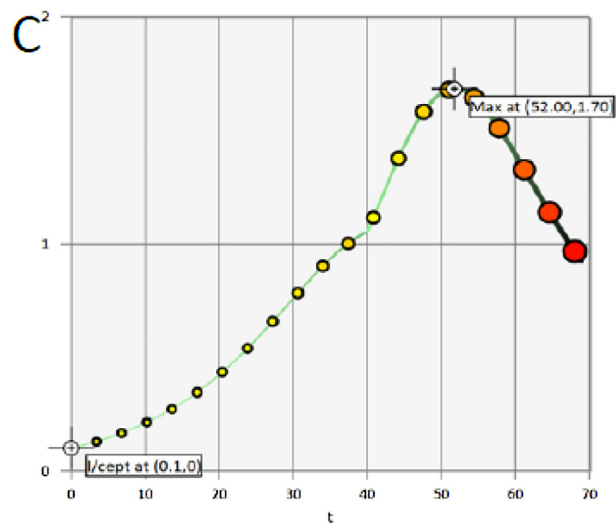


Fig. 2. The digraph method for describing mechanistic hypotheses can simply be incorporated into an interface A) Illustrates the model structure described in a graphical user interface to manipulate the model system; in this instance a progenitor species has a growth dynamic and also a conversion to a more mature phenotype. The mature cells produce an inhibitor which suppresses the growth rate of the progenitors leading to maturation. An experimental intervention is also specified with removal of all inhibitor at 40 h B) A typical model trajectory, without an experimental intervention, with the y axis denoting the density of progenitor cells, the line darkness the concentration of the inhibitor, and the size of the points the density of product C) The trajectory of the same model with the specified experimental intervention of complete inhibitor dilution at 40 h.

to optimisation can be identified. In this system, cytokine-mediated negative feedback suppresses the bulk up of progenitors in the presence of mature cells, and the productivity of a volume of medium is therefore determined by the accumulation of inhibitory

cytokines. Mathematically the specific instance of the general model described in 2.1 is:

$$\frac{dS}{dt} = \frac{r_g}{(1 + \exp(a(I - b)))} - r_c S$$

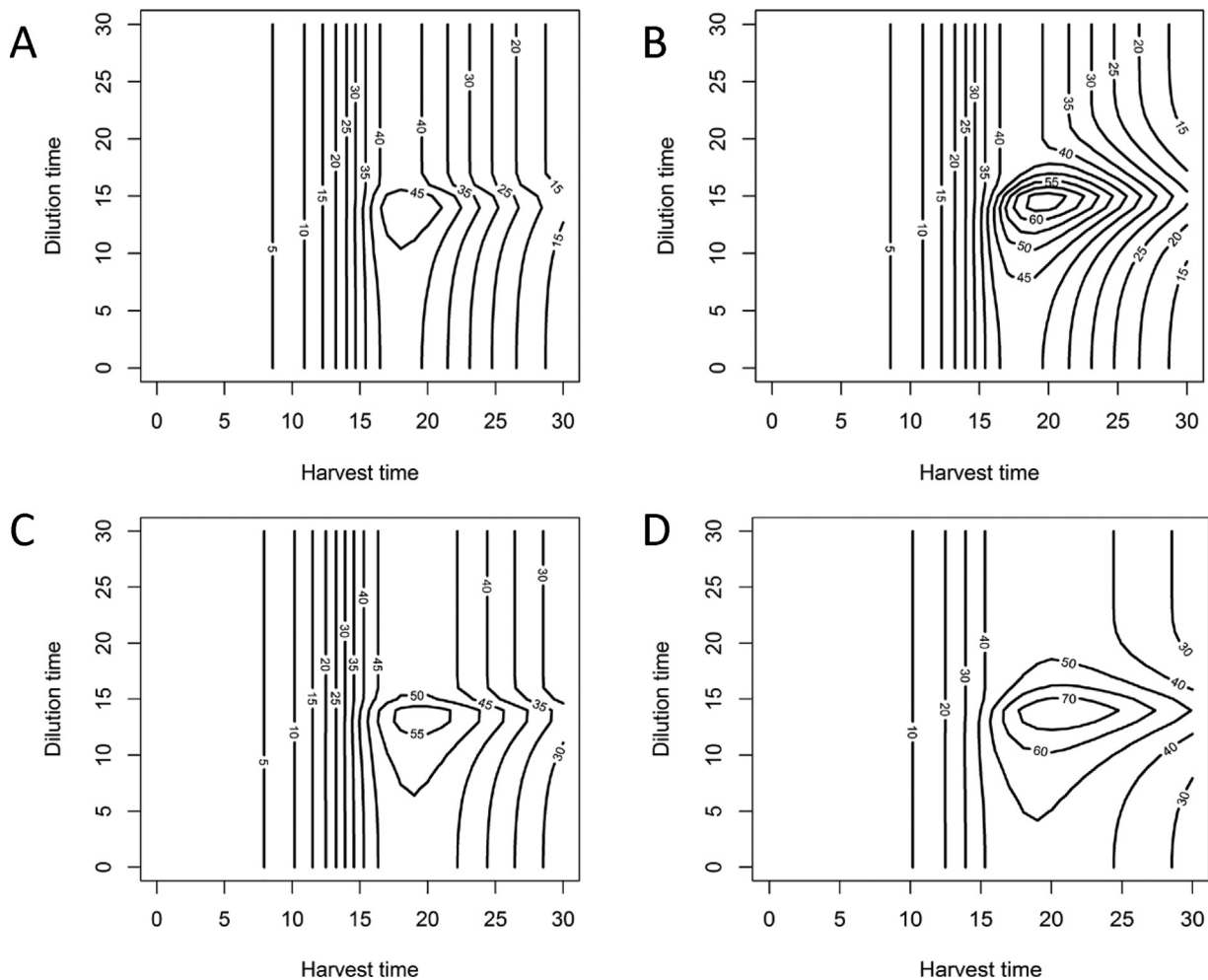


Fig. 3. Models can be used to identify optimal process operation. Using the inhibitory feedback model (described in Fig. 2) with the addition of a mature cell product death dynamic as an exemplar, optimal harvest and dilution time for maximum mature product yield is shown given a 25% (A, B) and 75% (C, D) medium exchange and death rates of mature product of 0.1 (A, C) and 0.2 (B, D). Units are arbitrary. The dynamics of the system lead to a large degree of non-linearity; insensitivity to dilution time for much of the harvest time space would lead to very poor estimations of optima from an uninformed Design of Experiment type approach.

$$\frac{dP}{dt} = r_c S - r_d P$$

$$\frac{dI}{dt} = r_p P$$

describing the evolution of Progenitor Density (S), Product Density (P) and Inhibitor Concentration (I) over time, where r_g is growth rate, r_c conversion rate from progenitor to product, r_d product death rate, r_p production rate of inhibitor, a is inhibitor sensitivity and b inhibitor threshold.

By reducing the concentration of such cytokines by a media dilution, there is the potential to increase volumetric productivity. Two key parameters, namely, the timing of media dilution and harvest of mature cells, determine the efficacy of this strategy. The modelling approach described here allows the user to explore the sensitivity of the volumetric productivity to these parameters, and in particular to identify targeted experiments to optimise the process. As an example of the former, the optimal combination of dilution time and mature product harvest time for maximum mature cell yield can be visualised for different proportions of medium exchange given different underlying death rates of mature cells (Fig. 3). Due to the highly non-linear nature of the system, a Design of Experiments approach would fail to find the optimum without

a considerable amount of experimentation; the class of mechanistic models described here enables a more directed and informed exploration of parameter space.

System sensitivity to driving mechanisms is critical in assessing process risk and opportunities for process improvement. A Monte Carlo simulation of the mature cell-progenitor negative feedback model described can be used to assess the sensitivity of its constituent and output species (inhibitor, progenitor cell, and mature cell product concentrations) over a culture time course to variability in model driving parameters (rates of growth, death, production, conversion and the modulating effect of the cytokine inhibitor). For the purpose of example, model parameters were selected using a growth rate reported in an erythroblast culture system, with other parameters selected to generate a base case model with a duration to terminal maturation of approximately 2 weeks [23] (initial progenitor density $1e5 \text{ cells mL}^{-1}$, growth rate $r_g = 0.05 \text{ h}^{-1}$, conversion rate $r_c = 0.035 \text{ h}^{-1}$, product death rate $r_d = 0.05 \text{ h}^{-1}$, production rate of inhibitor $r_p = 1e-5 \text{ cells}^{-1}$, inhibitor sensitivity $a = 0.1$, threshold inhibitor concentration $b = 50$). Sensitivity was assessed by determining the standard deviation in the variable of interest (cells, inhibitor) that results from a normal distribution in the driving parameters with a standard deviation 10% of the base case parameter value (Fig. 4). The sensitivities clearly differ from variable to variable and in a nonlinear fashion both temporally and in terms

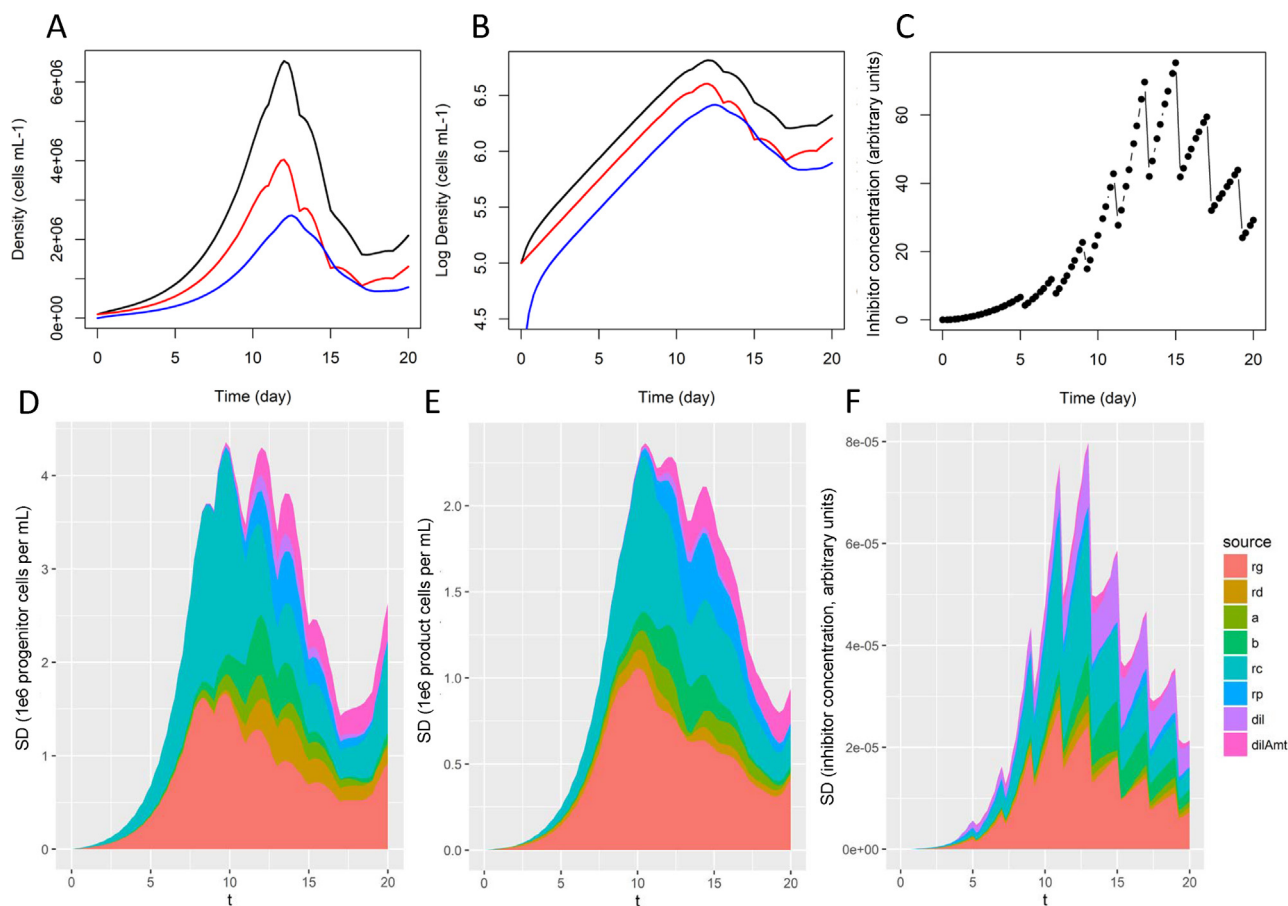


Fig. 4. Models can be used to identify process sensitivities. A and B show, for an illustrative base case of the hypothetical model parameterised to reflect the time course of an erythroblast proliferation and differentiation process reported in the literature, the time-courses of progenitor density (red), mature product density (blue) and total cell density (black) (t = time(days), SD = Standard Deviation); C show the corresponding inhibitor concentration time-course. Media exchange (without cell density reduction) provides dilution of inhibitor every second day (dilution factor $dilAmt = 0.5$) starting at day 5 ($dil = 120$ h). The areas in D, E and F show the sensitivity of, respectively, progenitor density, product density and inhibitor concentration to the driving parameters over time, measured as the standard deviation in the variable of interest that results from a normal distribution with standard deviation 0.1 in the driving parameter relative to the base case. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the driving variable in question. Some illustrative aspects of these trajectories are as follows: (a) Progenitor and product sensitivity to growth rate align with the underlying trajectories as the growth dominates early during the culture and the ratio of growth rate to inhibitory effects becomes increasingly important as culture progresses. There are peaks in progenitor sensitivity to growth rate coinciding with dilution events, corresponding to release from inhibition during these periods, but these effects are temporally delayed and dispersed for product. (b) Death of product has an effect on progenitors only by means of modulating product density and hence inhibitor production; its effect is transitory as the system regulates itself due to the inhibitor-mediated feedback effect of product on progenitors. Consequently effect on product is minimal. (c) The sensitivity to proportion of dilution develops after peak progenitor density for both, but its effect after this point remains relatively constant due to the self-regulating nature of the system. These profiles are highly non-linear, and not straightforward to predict without a formal mechanistic modelling approach. Such an approach is therefore powerful in identifying candidates for process optimisation or aspects of the system where control is critical.

4. Discussion

We have developed a modelling framework focussed on cell based processes for rapid model development and application.

The framework comprises a mathematical formalisation representing common cell culture dynamics that decouples aspects of the system thereby enabling compatibility with an intuitive interface. The restricted model building blocks aim to achieve the necessary trade-off between computationally tractable yet realistic models [14]. Digraphs provide a visual framework that simply supports expression of mechanistic hypotheses without requiring detailed knowledge of the mathematics involved. These elements are amenable to implementation in a software tool and integration with established methods for model solving and exploration. This approach bridges the gap between biology and modelling, which, for many practical applications such as cell processing (outside specialist fields such as systems biology), exist in two distinct skill sets without a widely shared language. The approach has promise for addressing several previously identified barriers to wider implementation of mechanistic modelling: support for systematic development of models following a best practice stepwise path of building, parameterisation and analysis [24,25]. Specifically it could support a structured and formalised approach from proposal of a biological system (hypothesis), through experimental design, expression as a digraph, conversion to ODE form, solution and analysis.

There are a number of evident benefits of such an approach to a biological or interdisciplinary team working with cell based processes. Mechanistic modelling, in contrast to a purely statistical (or

descriptive) approach such as DoE, provides greater process insight and therefore potential for rational determination of areas of process risk or to target for process optimisation [14]. It also enables more freedom in experimental design and source data allowing a user to explore (and develop) alternate models and parameterisation in a less restricted manner based on existing system knowledge. Further, this freedom enables data from experiments designed and conducted in tandem with a modelling approach to be scavenged and used to identify or refine critical process dynamic parameters. This latter point is of particular value as it offers the potential for maximising the process knowledge returned from investment in experimentation during early process development which will often target a spectrum of process issues (e.g. media formulation, unit operation optimisation and macro-scale process development) but may not conform to a consistent design. Although these benefits are common to an ODE based mechanistic modelling approach, the framework we describe greatly reduces the barrier to realisation. The digraph model description is designed to enhance interdisciplinary team communication and encourage exploration of alternative hypotheses. The intuitive structure of the digraphs supports rapid formulation, mathematical formalisation and testing of multiple hypotheses. By limiting the repertoire of functional forms included in the framework, the user is to some degree protected from pathological model formulations; the framework itself embodies an amount of modelling expertise. Experimental simulation *in silico* can quickly contribute to improved experimental design, particularly where partial system knowledge exists (eg approximate growth rates) [26,27]. The net effect is to streamline the process of formal articulation of biological phenomena, and thereby provide the potential to capitalise better on the power of modelling including robust exploration of alternative hypotheses, system optimisation and sensitivity analysis. The business and clinical impact of this will be improved design space understanding, associated risk-analyses and science-based regulatory submissions [9].

The approach we have developed builds on a rich heritage of mathematical modelling in earlier biological and bio-industrial systems. Modelling of biological growth developed from the recognition that external environmental forces impose limits on growth and that this can be represented by the logistic model. Multiple subsequent modifications of the logistic model were developed to account for e.g. specific environmental/organism growth limitations, specific product harvesting scenarios, or metabolic modulation i.e. Chapman-Richards. Other modifications such as the Gompertz or Gomp-ex model have created various asymmetries in the asymptotes at either end of the curve (to account, for example, for lag phase). Other specific case models have been developed i.e. the Monod equation to represent substrate dependent growth. In each of these a specific modification has been introduced based on a hypothetical or known insight into growth dynamics to increase the ability to represent real world growth behaviour; comparisons of these variants are frequently conducted to choose growth models for specific systems, and software has been developed to support selection from alternatives [28–30]. However, the various logistic model modifications are limited to crude approximation of multiple mechanisms to represent growth, with no simple route to unpack mechanisms if required to represent more complex system behaviour. Recently there has been significant work generating more elaborate ODE models describing metabolic processes and their relationships to cell growth or protein productivity. These approaches are partially mechanistic and biologically structured models, often incorporating the concept of compartments where multiple similar reaction rates may be combined to simplify the model. These models offer more mechanistic insight into the drivers of system behaviour and, once structured, can be parameterised with relatively limited experi-

mental data therefore lending themselves to semi-automated serial application across different bio-production systems [31,32]. However, the relatively elaborate ODE frameworks remain specific to the described metabolic processes and do not offer flexibility to simply incorporate other growth influences or population complexity. Our framework recognises the value in industrial bioprocess of a compartmental approach for cost effective model development, and the advantage of a biologically informed model structure, however retains the freedom to create a wide array of different model structures within constraints that protect from obvious modelling pathologies. A researcher may, for instance, start with something close to an existing logistic model, but can simply combine elements such as additional growth dynamics, key metabolic influences and population transitions, developing the level of compartmentalisation as required for the process at hand. This flexibility is valuable for cell based therapy processes where culture knowledge is limited relative to earlier bio-industrial applications and a framework is required for exploring alternative mechanisms and determining the best structure (and compartmentalisation) to deal with these types of processes. This represents a research and development stepping stone to generic model forms tailored for the cell therapy field offering standardised experimental design for rapid parameterisation of models describing proliferation, differentiation, and product generation.

The twin goals of achieving both wide user accessibility and the capability to model sufficient phenomena for broad application across biological systems create a tension between an extended model library underlying the interface and a minimised learning burden for users. The ODE-system approach to expressing dynamics, with non-linear model components which can be composited in a straightforward manner, can describe a range of important biological dynamics related to cell culture (e.g. growth, death, and interconversion of both non-cell and cell species) and thereby support the modelling of a broad range of processes (e.g. maturation, cytokine-mediated inhibition, lineage commitment) without incurring model pathologies. Our relatively parsimonious ODE framework constitutes one amongst a number of broadly mechanistic approaches; it is, however, particularly appropriate to well-mixed systems that are relatively challenging to characterise analytically. Partial differential equation extensions to this approach could handle spatial heterogeneity efficiently (e.g. reaction-advection-diffusion equations [33]), evolution of probability distributions describing system state [34], and migration of populations through lineage space [35]. In all these cases, the degrees of model freedom is much greater and as such a more extensive data set would be required for fitting or testing. The same applies for more complex systems models, such as highly-parameterised network models or those addressable by biologically-targeted modelling languages (e.g. SBML) [36]. A broader modelling framework will provide a greater expressiveness [37]; however, the disadvantage of this is increased time investment in appropriating such frameworks, and the far larger modelling knowledge required to avoid pathological model behaviour and promote selection of appropriate model structures. In practice the approach we have taken here could be extended into these domains, by developing parallel approaches whereby stochastic or spatially complex systems can be built up from a small number of parametrically parsimonious and well-behaved building blocks.

Funding

This work was supported by an Engineering and Physical Sciences Research Council Fellowship grant (EP/K00705X/1)

References

- [1] S.S. Ozturk, M.R. Riley, B.O. Palsson, Effects of ammonia and lactate on hybridoma growth, metabolism, and antibody production, *Biotechnol. Bioeng.* 39 (1992) 418–431, <http://dx.doi.org/10.1002/bit.260390408>.
- [2] E. Csaszar, D.C. Kirouac, M. Yu, W. Wang, W. Qiao, M.P. Cooke, A.E. Boitano, C. Ito, P.W. Zandstra, Rapid expansion of human hematopoietic stem cells by automated control of inhibitory feedback signaling, *Cell Stem Cell* 10 (2012) 218–229, <http://dx.doi.org/10.1016/j.stem.2012.01.003>.
- [3] D. Baksh, J.E. Davies, P.W. Zandstra, Soluble factor cross-talk between human bone marrow-derived hematopoietic and mesenchymal cells enhances in vitro CFU-F and CFU-O growth and reveals heterogeneity in the mesenchymal progenitor cell compartment, *Blood* 106 (2005) 3012–3019, <http://dx.doi.org/10.1182/blood-2005-01-0433>.
- [4] D. Busse, M. de la Rosa, K. Hobiger, K. Thurley, M. Flossdorf, A. Scheffold, T. Höfer, Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 3058–3063, <http://dx.doi.org/10.1073/pnas.0812851107>.
- [5] S. Kyriakopoulos, C. Kontoravdi, A framework for the systematic design of fed-batch strategies in mammalian cell culture, *Biotechnol. Bioeng.* 111 (2014) 2466–2476, <http://dx.doi.org/10.1002/bit.25319>.
- [6] J.C. Ryall, T. Cliff, S. Dalton, V. Sartorelli, Metabolic reprogramming of stem cell epigenetics, *Cell Stem Cell* 17 (2015) 651–662, <http://dx.doi.org/10.1016/j.stem.2015.11.012>.
- [7] U.S. Department of Health and Human Services Food and Drug Administration, ICH Q8(R2) Pharmaceutical Development, 2009.
- [8] T. Marique, M. Cherlet, V. Hendrick, F. Godia, G. Kretzmer, J. Wérenne, A general artificial neural network for the modelization of culture kinetics of different CHO strains, in: *Cytotechnology*, 2001, pp. 55–60, <http://dx.doi.org/10.1023/A:1014084802708>.
- [9] Y.Y. Lipsitz, N.E. Timmins, P.W. Zandstra, Quality cell therapy manufacturing by design, *Nat. Biotechnol.* 34 (2016) 393–400, <http://dx.doi.org/10.1038/nbt.3525>.
- [10] B. Ben Yahia, L. Malphettes, E. Heinzle, Macroscopic modeling of mammalian cell growth and metabolism, *Appl. Microbiol. Biotechnol.* 99 (2015) 7009–7024, <http://dx.doi.org/10.1007/s00253-015-6743-6>.
- [11] C. Chun, K. Heineken, D. Szeto, T. Ryll, S. Chamow, J.D. Chung, Application of factorial design to accelerate identification of CHO growth factor requirements, *Biotechnol. Prog.* 19 (2003) 52–57, <http://dx.doi.org/10.1021/bp025575>.
- [12] M.M. Hunt, G. Meng, D.E. Rancourt, I.D. Gates, M.S. Kallos, Factorial experimental design for the culture of human embryonic stem cells as aggregates in stirred suspension bioreactors reveals the potential for interaction effects between bioprocess parameters, *Tissue Eng. Part C Methods* 20 (2014) 76–89, <http://dx.doi.org/10.1089/ten.tec.2013.0040>.
- [13] P.D. Mitchell, E. Ratcliffe, P. Hourd, D.J. Williams, R.J. Thomas, A Quality by Design approach to risk reduction & optimisation for human embryonic stem cell cryopreservation processes, *Tissue Eng. Part C Methods* 20 (2014) 1–32, <http://dx.doi.org/10.1089/ten.TEC.2013.0595>.
- [14] M. Koutinas, A. Kiparissides, E.N. Pistikopoulos, A. Mantalaris, Bioprocess systems engineering: transferring traditional process engineering principles to industrial biotechnology, *Comput. Struct. Biotechnol. J.* 3 (2012) e201210022, <http://dx.doi.org/10.5936/csbt.201210022>.
- [15] H.T.B. Pham, G. Larsson, S.O. Enfors, Growth and energy metabolism in aerobic fed-batch cultures of *Saccharomyces cerevisiae*: simulation and model verification, *Biotechnol. Bioeng.* 60 (1998) 474–482, [http://dx.doi.org/10.1002/\(SICI\)1097-0290\(19981120\)60:4<474::AID-BIT9>3.0.CO;2-J](http://dx.doi.org/10.1002/(SICI)1097-0290(19981120)60:4<474::AID-BIT9>3.0.CO;2-J).
- [16] P. Farzan, B. Mistry, M.G. Ierapetritou, Review of the important challenges and opportunities related to modeling of mammalian cell bioreactors, *AIChE J.* 63 (2017) 398–408, <http://dx.doi.org/10.1002/aic.15442>.
- [17] R. Roy, G. Milton, S. Dane, R. Lichtner, L. Brindle, Achieving large-scale, cost-effective, reproducible manufacturing of high-quality cells, in: *A Technology Roadmap to 2025*, 2016.
- [18] F.R. Sidoli, A. Mantalaris, S.P. Asprey, Modelling of mammalian cells and cell culture processes, *Cytotechnology* 44 (2004) 27–46, <http://dx.doi.org/10.1023/B:CYTO.0000043397.94527.84>.
- [19] M. Del Mar Romero, D. Sabater, J.A. Fernández-López, X. Remesar, M. Alemany, Glycerol production from glucose and fructose by 3T3-L1 Cells: a mechanism of Adipocyte defense from excess substrate, *PLoS One* 10 (2015), <http://dx.doi.org/10.1371/journal.pone.0139502>.
- [20] A. Wimmel, B. Wiedenmann, S. Rosewicz, Autocrine growth inhibition by transforming growth factor beta-1 (TGFbeta-1) in human neuroendocrine tumour cells, *Gut* 52 (2003) 1308–1316.
- [21] M. Tsubaki, M. Komai, T. Itoh, M. Imano, K. Sakamoto, H. Shimaoka, N. Ogawa, K. Mashimo, D. Fujiwara, T. Takeda, J. Mukai, K. Sakaguchi, T. Satou, S. Nishida, Inhibition of the tumour necrosis factor-alpha autocrine loop enhances the sensitivity of multiple myeloma cells to anticancer drugs, *Eur. J. Cancer* 49 (2013) 3708–3717, <http://dx.doi.org/10.1016/j.ejca.2013.07.010>.
- [22] R. De Maria, A. Zeuner, A. Eramo, C. Domenichelli, D. Bonci, F. Grignani, S.M. Srinivasula, E.S. Alnemri, U. Testa, C. Peschle, Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1, *Nature* 401 (1999) 489–493, [http://dx.doi.org/10.1016/S0887-7963\(00\)80020-6](http://dx.doi.org/10.1016/S0887-7963(00)80020-6).
- [23] R. Bayley, F. Ahmed, K. Glen, M. McCall, A. Stacey, R. Thomas, The productivity limit of manufacturing blood cell therapy in scalable stirred bioreactors, *J. Tissue Eng. Regen. Med.* (2017), <http://dx.doi.org/10.1002/term.2337>, ePub ahead of print.
- [24] A. Kiparissides, M. Koutinas, C. Kontoravdi, A. Mantalaris, E.N. Pistikopoulos, Closing the loop in biological systems modeling—From the in silico to the in vitro, *Automatica* 47 (2011) 1147–1155, <http://dx.doi.org/10.1016/j.automatica.2011.01.013>.
- [25] C. Kontoravdi, E.N. Pistikopoulos, A. Mantalaris, Systematic development of predictive mathematical models for animal cell cultures, *Comput. Chem. Eng.* 34 (2010) 1192–1198, <http://dx.doi.org/10.1016/j.compchemeng.2010.03.012>.
- [26] A. Kiparissides, S.S. Kucherenko, A. Mantalaris, E.N. Pistikopoulos, Global sensitivity analysis challenges in biological systems modeling, *Ind. Eng. Chem. Res.* 28 (2009) 7168–7180, <http://dx.doi.org/10.1021/ie900139x>.
- [27] M. Baltes, R. Schneider, C. Sturm, M. Reuss, Optimal experimental design for parameter estimation in unstructured growth models, *Biotechnol. Prog.* 10 (1994) 480–488, <http://dx.doi.org/10.1021/bp00029a005>.
- [28] M. Kahm, G. Hasenbrink, H. Lichtenberg-frate, J. Ludwig, M. Kschischo, Grofit: fitting biological growth curves, *J. Stat. Software* 33 (2010) 1–21, <http://dx.doi.org/10.1038/npre.2010.4508.1>.
- [29] N. Shirsat, A. Mohd, J. Whelan, N.J. English, B. Glennon, M. Al-Rubeai, Revisiting Verhulst and Monod models: analysis of batch and fed-batch cultures, *Cytotechnology* 67 (2015) 515–530, <http://dx.doi.org/10.1007/s10616-014-9712-5>.
- [30] J. López-Meza, D. Araíz-Hernández, L.M. Carrillo-Cocom, F. López-Pacheco, M. del R. Rocha-Pizaña, M.M. Alvarez, Using simple models to describe the kinetics of growth, glucose consumption, and monoclonal antibody formation in naïve and infliximab producer CHO cells, *Cytotechnology* 68 (2016) 1287–1300, <http://dx.doi.org/10.1007/s10616-015-9889-2>.
- [31] E. Violet, T. Fischer, A software supported compartmental modeling approach for *Paenibacillus polymyxa*, *IFAC Proc Vol.* 2010 (2010) 377–382, <http://dx.doi.org/10.3182/20100707-3-BE-2012.0048>.
- [32] C. Kontoravdi, D. Wong, C. Lam, Y.Y. Lee, M.G.S. Yap, E.N. Pistikopoulos, A. Mantalaris, Modeling amino acid metabolism in mammalian cells—toward the development of a model library, *Biotechnol. Prog.* 23 (2007) 1261–1269, <http://dx.doi.org/10.1021/bp070106z>.
- [33] L. Zhang, A.D. Lander, Q. Nie, A reaction–diffusion mechanism influences cell lineage progression as a basis for formation, regeneration, and stability of intestinal crypts, *BMC Syst. Biol.* 6 (2012) 93, <http://dx.doi.org/10.1186/1752-0509-6-93>.
- [34] M.J. Grote, V. Palumberi, B. Wagner, A. Barbero, I. Martin, Dynamic formation of oriented patches in chondrocyte cell cultures, *J. Math. Biol.* 63 (2011) 757–777, <http://dx.doi.org/10.1007/s00285-010-0390-4>.
- [35] J. Wu, M.R. Rostami, E.S. Tzanakakis, Stem cell modeling: from gene networks to cell populations, *Curr. Opin. Chem. Eng.* 2 (2013) 17–25, <http://dx.doi.org/10.1016/j.coche.2013.01.001>.
- [36] A. Shapiro, M. Finney, B. Hucka, A. Bornstein, A. Funahashi, S.M. Jouraku, N. Keating, J. re, M.J. Matthews, SBML models and MathSBML, in: *Introduction to Systems Biology*, 2007, pp. 395–421, http://dx.doi.org/10.1007/978-1-59745-531-2_20.
- [37] M. Read, P.S. Andrews, J. Timmis, V. Kumar, Modelling biological behaviours with the unified modelling language: an immunological case study and critique, *J. R. Soc. Interface* 11 (2014), <http://dx.doi.org/10.1098/rsif.2014.0704>, 20140704.