20th European Symposium on Computer Aided Process Engineering – ESCAPE20
S. Pierucci and G. Buzzi Ferraris (Editors)

Combining Genetic Circuit and Microbial Growth Kinetic Models: A Challenge for Biological Modelling

Michalis Koutinas,[a] Alexandros Kiparissides,[a] Ming-Chi Lam,[a,b] Rafael Silva-Rocha,[c] Victor de Lorenzo,[c] Vitor A.P. Martins dos Santos,[b] Efstratios N. Pistikopoulos,[a] Athanasios Mantalaris[a,*]

[a] *Department of Chemical Engineering, Imperial College, London SW7 2AZ, UK*
[b]*Systems and Synthetic Biology Group, Helmholtz Center for Infection Research, Inhoffenstrasse 7, D-38124, Braunschweig, Germany*
[c]*Centro Nacional de Biotecnologia, Consejo Superior de Investigaciones Cientificas, Darwin 3, Cantoblanco, 28049 Madrid, Spain*
[*]*Author to whom correspondence should be addressed: a.mantalaris@imperial.ac.uk*

Abstract

A modelling framework that consists of model building, validation and analysis, leading to model-based design of experiments and to the application of optimisation-based model-predictive control strategies for the development of optimised bioprocesses is presented. An example of this framework is given with the construction and experimental validation of a dynamic mathematical model of the *Ps/Pr* promoters system of the TOL plasmid, which is used for the metabolism of *m*-xylene by *Pseudomonas putida* mt-2. Furthermore, the genetic circuit model is combined with the growth kinetics of the strain in batch cultures, demonstrating how the description of key genetic circuits can facilitate the improvement of existing growth kinetic models that fail to predict unusual growth patterns. Consequently, the dynamic model is combined with global sensitivity analysis, which is used to identify the presence of significant model parameters, constituting a model-based methodology for the formulation of genetic circuit optimization methods.

Keywords: Dynamic modelling, Sensitivity analysis, Genetic circuit, pWW0 (TOL) plasmid, *Pseudomonas putida*.

1. Introduction

Genetic circuits are groups of elements that interact producing certain behaviour [1]. Based on our capability to engineer genetic circuits, fundamental biological processes can be studied systematically and targets can be identified for genetic modification, producing the desired properties. However, the extensive experimentation required to understand the function of genetic circuits, is often limited by the time and cost required. Although the experimental techniques required for the study of genetic circuits are very sophisticated, reliable mathematical models are equally important in reducing substantially the trial-and-error experimentation. In line with this, dynamic modelling can be used for characterisation of the cellular function integrating biological information into predictive models [2]. Furthermore, the molecular and genetic events responsible for the growth kinetics of a microorganism can be extensively influenced by the presence of mixtures of substrates leading to unusual growth patterns, which cannot be accurately predicted from existing models [3]. Thus, a novel approach combining

genetic circuit and growth kinetic models constitutes an improved version of the currently used models for the prediction of microbial growth kinetics.
Complex biological models may include a large number of parameters, which can be difficult to estimate and may incorporate expensive and time consuming experiments. Global sensitivity analysis (GSA) is a tool used to quantify the importance of model parameters and their interactions with regards to the model output [4]. Analyzing the properties of parameters included in genetic circuit models provides the identification of the most significant ones with respect to the output of interest. Thus, the experiments required for circuit optimisation can be aimed at genetic modifications altering these parameters alone, consequently reducing the cost and the number of experiments required. Application of GSA methods to biological systems has been limited to a few examples [5]. In this work, we present a modeling framework that consists of model building, validation and analysis providing a solid basis for genetic circuit optimization. A combined model has been constructed describing the function of the *Ps/Pr* promoters system and the growth kinetics of *P. putida* mt-2. The model's prediction has been compared to that of models accounting for enzymatic interactions. Finally, preliminary model analysis has been performed with the application of Sobol' GSA method [6].

2. Results and discussion

2.1. Genetic circuit model

P. putida mt-2 is equipped with the TOL plasmid (pWW0), specifying a pathway for the catabolism of *m*-xylene. The enzymes required for these reactions are produced by the two gene operons of TOL (upper- and *meta*- operon), while two genes (*xylS* and *xylR*) control the regulation of transcription of the operons. These four transcriptional units are driven by four different promoters (*Pu*, *Pm*, *Ps* and *Pr*). The *Ps/Pr* system has been reconstructed into its various interacting molecular components and has been described as a combination of logic gates (Fig. 1) producing a representation in an analogy to electronic circuitry. Based on the logic model of the *Ps/Pr* system, the Hill functions were used as input functions to the genes and a dynamic mathematical model of the system was generated, as described below.

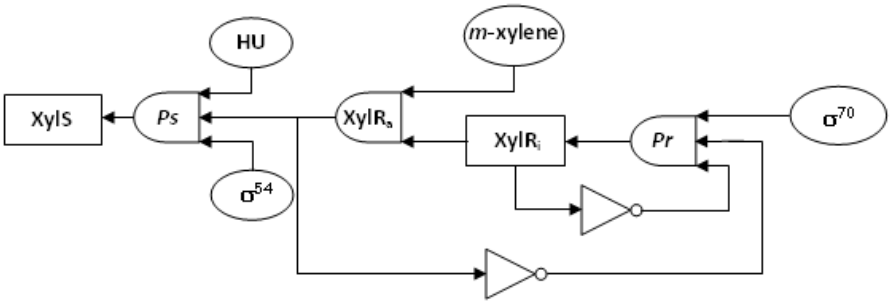


Figure 1. Logic diagram of the *Ps/Pr* system of TOL plasmid pWW0. ⬭ : input; ▭ : output; ⊓ : AND; ▷○ : NOT.

P. putida mt-2 degrades aromatic substrates through a series of events leading to coordinated expression from the upper- and *meta*-cleavage pathways of TOL. The master regulator of the two pathways, the XylR protein, is transcribed by the *xylR* gene from two σ^{70} tandem promoters (*Pr1* and *Pr2*). After binding with *m*-xylene, the inactive dimer form of the XylR protein ($XylR_i$) binds ATP and oligomerizes to form a hexamer. This leads to the formation of the active form of XylR ($XylR_a$), which induces transcription of the *Pu* promoter synthesizing the upper-pathway enzymes. The

synthesis of $XylR_i$, as well as the forward and reverse reactions for XylR activation/deactivation are expressed by Eqs. (1-2) (all symbols are defined in Table 1):

$$\frac{dXylR_i}{dt} = \frac{\beta_{XylR_i} \Pr_{TC}}{K_{\Pr,XylR_i} + \Pr_{TC}} - r_{XylR} XylR_i Xyl + 3r_{R,XylR} XylR_a (Xyl_{INI} - Xyl) - \alpha_{XylR_i} XylR_i \quad (1)$$

$$\frac{dXylR_a}{dt} = \frac{1}{3} r_{XylR} XylR_i Xyl - r_{R,XylR} XylR_a (Xyl_{INI} - Xyl) - \alpha_{XylR_a} XylR_a \quad (2)$$

Table 1. List of symbols.

Symbols	Definition
$XylR_i$ / $XylR_a$	concentrations of the inactive and active forms of XylR protein
r_{XylR} / $r_{R,XylR}$	$XylR_i$ oligomerization and $XylR_a$ dissociation constants
Xyl / Xyl_{INI} / Suc	total *m*-xylene, total *m*-xylene initial and total succinate concentrations
Pr_{TC} / Ps_{TC}	relative activities of *Pr* and *Ps*
t	time
$K_{Pr,XylRi}$ / β_{XylRi}	$XylR_i$ and maximal $XylR_i$ translation rates
α_{XylRi} / α_{XylRa}	degradation/dilution rates due to cellular volume increase for $XylR_i$ and $XylR_a$
β_0 / β_{Ps} / β_{Pr}	basal and maximal expression levels of *Ps* and *Pr*
$K_{XylRa,Ps}$ / K_{XylRi} / K_{XylRa}	activation and repression coefficients of *Ps* and *Pr* due to $XylR_i$ and/or $XylR_a$
$n_{Pr,i}$ / $n_{Pr,a}$ / $n_{Ps,a}$	Hill coefficients of *Pr* and *Ps* due to $XylR_i$ and/or $XylR_a$ binding
$K_{SUC,Pr}$ / $K_{SUC,Ps}$	inhibition constant of succinate on *Pr* and *Ps*
α_{Pr} / α_{Ps}	deactivation rates of *Pr* and *Ps*
μ_1 / μ_2 / $\mu_{max,1}$ / $\mu_{max,2}$	specific and maximum specific growth rates of biomass on *m*-xylene and succinate
S_1 / S_2	*m*-xylene and succinate concentrations
$K_{S,1}$ / $K_{I,1}$ / $K_{S,2}$	*m*-xylene and succinate saturation and/or inhibition constants
X	biomass concentration
MWt_1 / MWt_2	*m*-xylene and succinate molecular weight
Y_1 / Y_2	yield coefficient for biomass on *m*-xylene and succinate
$K_{I,1,2}$ / $K_{I,1\text{-}P,2}$	*m*-xylene inhibition and by-product inhibition on succinate constant

For simplification of the model developed we express both *xylR* promoters as a single *Pr* promoter. The XylR protein activates *Ps* and represses its own synthesis. During the experiments, the *Pr* promoter was slightly repressed in the presence of both substrates as compared to the case when only *m*-xylene was present. Therefore, we presumed that succinate is repressive for *Pr* in the presence of *m*-xylene and that the concentration of σ^{70} is at a constant level. The function of *Pr* promoter activity is expressed by Eq. (3).

$$\frac{d\Pr_{TC}}{dt} = \frac{\beta_{\Pr}}{1 + \left(\frac{XylR_i}{K_{XylR_i}}\right)^{n_{Pr,i}} + \left(\frac{XylR_a}{K_{XylR_a}}\right)^{n_{Pr,a}}} \frac{1}{1 + K_{SUC,\Pr} Suc^2} - \alpha_{\Pr} \Pr_{TC} \quad (3)$$

The *xylS* gene is expressed constitutively at a basal expression level β_0 (Eq. 4) but boosted in the presence of *m*-xylene synthesizing the XylS protein and stimulating the induction of the *meta*-pathway. Activation of *Ps* is assisted by the HU protein, which stabilizes the correct architecture of the promoter. The experiments of this study confirmed in that *Ps* is negatively affected in the presence of succinate. Thus, we consider that succinate is repressive for *Ps* promoter and we assume that the concentration of HU and σ^{54} is constant at housekeeping level.

$$\frac{dPs_{TC}}{dt} = \beta_0 + \beta_{Ps} \frac{XylR_a^{\,n_{Ps,a}}}{K_{XylR_a,Ps}^{\,n_{Ps,a}} + XylR_a^{\,n_{Ps,a}}} \frac{1}{1 + \left(\frac{Suc}{K_{SUC,Ps}}\right)^2} - \alpha_{Ps} Ps_{TC} \quad (4)$$

2.2. Growth kinetic model

P. putida mt-2 was first cultivated in the presence of succinate and *m*-xylene as single substrates. The biodegradation of 0.9 mM *m*-xylene fed in a batch experiment was modelled assuming substrate inhibition [7] (Eqs. 5-6).

$$\mu_1 = \frac{\mu_{\max,1} S_1}{K_{S,1} + S_1 + \frac{S_1^3}{K_{I,1}^2}}, \quad (5); \quad \frac{dS_1}{dt} = -\frac{1}{MWt_1}\frac{\mu_1}{Y_1} X, \quad (6)$$

A 1 h lag-phase occurred following the introduction of *m*-xylene (data not shown). Since the culture was pre-grown in succinate, the lag might be due to the change in substrate requiring the induction of new enzymes for *m*-xylene biodegradation. Thus, we assume that the transition from the lag to the accelerating phase takes place when the activity of *Ps* increased from its basal level by 65-fold, an amount which corresponds to the activation of the TOL pathway and to the induction of its enzymes. The genetic circuit model was used to calculate *Ps* promoter's activity over time, estimating the duration of the lag-phase, and the growth kinetic model was used after the lag-phase.

The growth kinetics of mt-2 was studied in the presence of succinate. The consumption of 13.6 mM succinate fed in a batch experiment was modelled using Eqs. (7-8).

$$\frac{dS_2}{dt} = -\frac{1}{MWt_2}\frac{\mu_2}{Y_2} X, \quad (7); \quad \mu_2 = \frac{\mu_{\max,2} S_2}{K_{S,2} + S_2}, \quad (8)$$

The strain was cultured in a batch experiment in the presence of 14 mM succinate and 1.04 mM *m*-xylene. Unlike the cases of simultaneous or diauxic growth often observed when a mixture of substrates is available, mt-2 displayed a different growth pattern. Following the initial lag-phase, *m*-xylene degradation started first followed by succinate degradation resulting in two phases were both substrates were utilised individually and one phase were both substrates were utilised simultaneously. *m*-xylene is sensed by *P. putida* mainly as a stressor to be extruded rather than as a nutrient to be metabolised. Consequently, the lag-phase in succinate degradation can be attributed to the presence of the stressor and the duration of the lag on succinate might depend on the time required to inactivate *m*-xylene. In order to consider the inhibitory effects of *m*-xylene and its intermediates on succinate degradation, a new succinate degradation model is suggested (Eq. 9). We assume that a major intermediate in *m*-xylene degradation accumulates over time proportionally to the removal of *m*-xylene. Furthermore, inhibition of growth on succinate due to the presence of *m*-xylene is also considered.

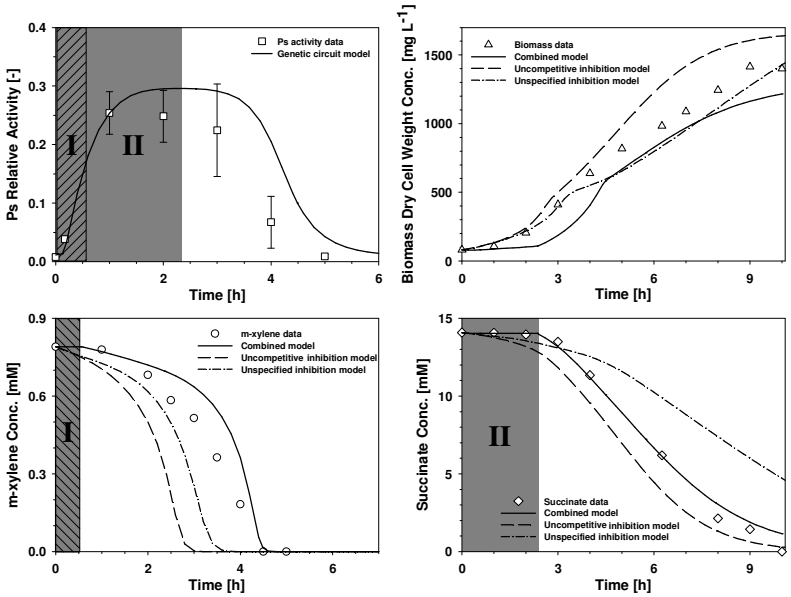
$$\mu_2 = \frac{\mu_{\max,2} S_2}{K_{S,2} + S_2} \frac{K_{I,1,2}}{K_{I,1,2} + S_1} \frac{K_{I,1-P,2}}{K_{I,1-P,2} + (1 - S_1)} \quad (9)$$

Due to the overall repression of *Ps*, we assumed that in the presence of both substrates, *m*-xylene degradation started when the activity of *Ps* increased from its basal level by 14-fold. Also, as mentioned above, the lag-phase in succinate degradation is attributed to the stress effect caused by *m*-xylene. Thus, we assume that growth on succinate starts when the cellular metabolic resources are redistributed towards succinate assimilation indicating the onset of the TOL pathway deinduction, which is expressed by the time point where *Ps* activity starts decreasing from its maximum value. The model parameter values were obtained from the experiments presented above.

The mixed-substrate experiment was also modelled with the SKIP model, which is used when the type of substrate interactions cannot be directly specified, and with cell growth

models accounting for competitive, noncompetitive and uncompetitive inhibition. The uncompetitive inhibition and SKIP models satisfactorily described the experimental data, while the competitive and non-competitive inhibition models failed to follow the experimental results (data not shown).

The predictive capability of the model was tested with an independent experiment. The initial succinate concentration was maintained at 14.1 mM, while *m*-xylene concentration was reduced to 0.8 mM. The duration of the lag-phase for each substrate was calculated from the genetic circuit model (Fig. 2) as described above. The combined mathematical model underpredicted the biomass concentration (Fig. 3) and overpredicted to minor extent the *m*-xylene concentration over time (Fig. 4). However, the model closely tracked succinate concentration (Fig. 5) and overall produced a satisfactory description of the experimental data. In contrast, the competitive inhibition and the SKIP model failed to describe the experimental results confirming that only the combined model can be predictive under different experimental conditions.



Figures 2-5. Comparison of the combined model and the substrate interaction models prediction. I) lag-phase on *m*-xylene; II) lag-phase on succinate.

2.3. Model analysis

The ability of the Sobol' method to distinguish between individual and total sensitivity index (SI) enables us to identify interacting factors within the system gaining valuable insight into its dynamics. The dimensionality of the sensitivity analysis problem is defined by the number of model parameters; therefore a feasibility constraint regarding the maximum possible number of individually scanned parameters is imposed implicitly in terms of computational time. This constraint is unavoidable due to the - increasing with dimension - number of model evaluations required for the Monte Carlo integrals to converge. Researchers in the field of GSA often resolve to parameter grouping in order to reduce the dimensionality of the problem, thus solving a more tractable version of the original problem. Therefore the parameters of the model have been divided in 16 groups according to their biological function, to make the computation of GSA feasible.

The sensitivities of the parameter groups have been calculated at different time points for $XylR_a$ concentration as the output variable and the results are shown in Fig. 6. The

most significant parameter groups are: i) Gp1 ($\mu_{max,1}$), ii) Gp10 (r_{XylR}, $r_{R,XylR}$), iii) Gp11 (β_{XylRi}, α_{XylRi}, $K_{Pr,XylRi}$) and Gp12 (α_{XylRa}). Therefore, parameters related to $XylR_i$ and $XylR_a$ oligomerization and dissociation respectively, $XylR_i$ translation and degradation, $XylR_a$ degradation and the maximum specific growth rate on *m*-xylene are the most significant. The identification of significant parameters not only enriches our knowledge on the intricate mechanisms that govern cellular behaviour, but may be an indication of which are the most significant parameters to genetically modify towards the production of improved cellular behaviour.

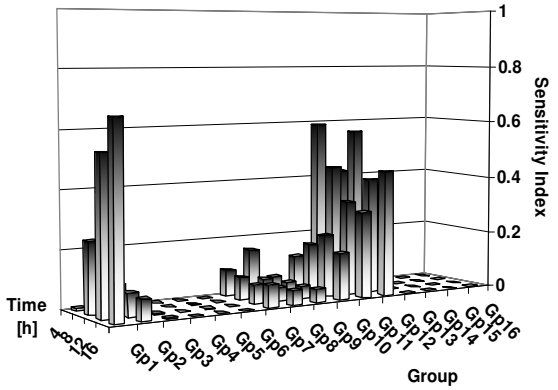


Figure 6. Preliminary calculation of SI for various parameter groups. Indexes were calculated at different time points for $XylR_a$ concentration as the output variable.

3. Conclusions

The mathematical model successfully combines the prediction of a key genetic circuit to the growth kinetics of the microorganism, producing a reliable description of the system. The combination of the constructed model with GSA constitutes a model-based methodology identifying the driving mechanisms of the system, which can be used for hypothesis testing and network optimisation. Thus, the modeling framework presented in this study enables the formulation of genetic circuit optimisation methods, opening a window into the direct re-programming of cellular behaviour and, subsequently, the development of optimised and novel, high-added value biocatalysts.

4. Acknowledgements

This work was supported by the following projects: a) PROBACTYS (EU – FP6), b) PSYSMO (BBSRC – ERA-NET program) and c) TARPOL (EU – FP7).

References

[1] R. Weiss, S. Basu, S. Hooshangi, A. Kalmbach, D. Karig, R. Mehreja and I. Netravali, Nat. Comput., 2 (2003) 47
[2] F.R. Sidoli, A. Mantalaris and S.P. Asprey, Cytotechnology, 44 (2004) 27
[3] K.F. Reardon, D.C. Mosteller and J.D.B. Rogers, Biotechnol. Bioeng., 69 (2000) 385
[4] A. Kiparissides, S.S. Kucherenko, A. Mantalaris and E.N. Pistikopoulos, Ind. Eng. Chem. Res., 48 (2009) 7168
[5] C. Kontoravdi, S.P. Asprey, E.N. Pistikopoulos and A. Mantalaris, Biotechnol. Prog., 21 (2005) 1128
[6] I.M. Sobol', Math. Comput. Simul., 55 (2001) 271
[7] T. Yano and S. Koga, Biotechnol. Bioeng., 11 (1969) 139