Metabolic Engineering 13 (2011) 401–413

Contents lists available at ScienceDirect

Metabolic Engineering

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

Linking genes to microbial growth kinetics—An integrated biochemical systems engineering approach

Michalis Koutinas [a], Alexandros Kiparissides [a], Rafael Silva-Rocha [b], Ming-Chi Lam [c], Vitor A.P. Martins dos Santos [c,1], Victor de Lorenzo [b], Efstratios N. Pistikopoulos [a], Athanasios Mantalaris [a,*]

[a] Centre for Process Systems Engineering, Department of Chemical Engineering, South Kensington Campus, Imperial College London, SW7 2AZ, London, UK
[b] Centro Nacional de Biotecnologia, Consejo Superior de Investigaciones Cientificas, Darwin 3, Cantoblanco, 28049 Madrid, Spain
[c] Systems and Synthetic Biology Group, Helmholtz Center for Infection Research (HZI), Inhoffenstrasse 7, D-38124, Braunschweig, Germany

ARTICLE INFO

Article history:
Received 23 October 2010
Received in revised form
31 January 2011
Accepted 3 February 2011
Available online 16 February 2011

Keywords:
Metabolic engineering
Genetic circuit
pWW0 (TOL) plasmid
Pseudomonas putida
m-xylene
Dynamic model

ABSTRACT

The majority of models describing the kinetic properties of a microorganism for a given substrate are unstructured and empirical. They are formulated in this manner so that the complex mechanism of cell growth is simplified. Herein, a novel approach for modelling microbial growth kinetics is proposed, linking biomass growth and substrate consumption rates to the gene regulatory programmes that control these processes. A dynamic model of the TOL (pWW0) plasmid of *Pseudomonas putida* mt-2 has been developed, describing the molecular interactions that lead to the transcription of the *upper* and *meta* operons, known to produce the enzymes for the oxidative catabolism of *m*-xylene. The genetic circuit model was combined with a growth kinetic model decoupling biomass growth and substrate consumption rates, which are expressed as independent functions of the rate-limiting enzymes produced by the operons. Estimation of model parameters and validation of the model's predictive capability were successfully performed in batch cultures of mt-2 fed with different concentrations of *m*-xylene, as confirmed by relative mRNA concentration measurements of the promoters encoded in TOL. The growth formation and substrate utilisation patterns could not be accurately described by traditional Monod-type models for a wide range of conditions, demonstrating the critical importance of gene regulation for the development of advanced models closely predicting complex bioprocesses. In contrast, the proposed strategy, which utilises quantitative information pertaining to upstream molecular events that control the production of rate-limiting enzymes, predicts the catabolism of a substrate and biomass formation and could be of central importance for the design of optimal bioprocesses.

1. Introduction

Monitoring of bioprocess performance is generally conducted by measuring macroscopic operating parameters, ignoring the molecular interactions controlling the process (Kovarova-Kovar and Egli, 1998). However, in many cases biomass utilisation and substrate consumption patterns cannot be accurately predicted by models developed merely based on bulk measurements, due to regulation at both the enzyme and the genetic level (Rogers and Reardon, 2000). Especially in bioprocesses with mixed microbial populations, multiple substrates and fluctuating substrate concentrations, traditional Michaelis–Menten and Monod models do not capture the description of substrate degradation (Park et al., 2008). Previous studies have demonstrated that enzymatic measurements can be successfully used to construct mechanistic models with improved predictive capabilities (Melchiorsen et al., 2001). Nevertheless, the application of experimentally validated models of key genetic circuits, describing the upstream molecular and genetic events that control the synthesis of enzymes, to improve the description of the kinetic properties of a microorganism has not been demonstrated yet. The current state of the art is rather limited to the recent work of Douma et al. (2010) presenting the development of a simple dynamic gene regulation model to describe biomass and penicillin production in a chemostat. It was shown that although the gene regulation model was only validated using enzyme activity assays instead of measuring mRNA levels, it nevertheless improved significantly the prediction of the bioprocess demonstrating the benefits of accounting for genetic events in biochemical engineering.

* Corresponding author. Fax: +44 20 7594 5638.
E-mail address: a.mantalaris@imperial.ac.uk (A. Mantalaris).
[1] Current address: Chair for Systems and Synthetic Biology, Wageningen University, Dreijenplein 310, 6703 HB Wageningen, The Netherlands.

1096-7176/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2011.02.001

The application of modern molecular tools to link the molecular to meso- and macro-scale events that instrumentally affect the composition, physiological state and activities of a microbial population is becoming a primary research goal for environmental biotechnology. Competitive RT-PCR has been previously applied to improve the monitoring of *in situ* microbial function and activity in methanotrophs (Han and Semrau, 2004). Furthermore, RT-PCR has also been used to understand the link between the photosynthetic capacity of autotrophic plankton and transcription of the gene responsible for primary carbon fixation during photosynthesis (Corredor et al., 2004). Although it is emphasised that molecular biology methods can be useful for optimising bioprocess performance, nonetheless substrate consumption is only correlated to mRNA transcript levels using best-fit lines without considering the regulatory processes controlling the transcription from the relevant genes. Thus, even though an almost linear relationship between mRNA levels and cellular activity might occur, this relationship can be strongly dependent on the growth conditions – thus, emphasising the need for establishing detailed mechanistic models that link transcript numbers and substrate utilisation rates. The limitation of investigations that do not consider the regulatory phenomena affecting gene expression has been previously shown in a gas-phase biofilter (Gunsch et al., 2007). Specifically, although relative gene expression was consistent with biofilter performance, no direct mathematical correlation could be established between the microscopic and macroscopic levels.

Pseudomonas putida is a metabolically versatile soil bacterium, capable of metabolising a large number of industrially important aromatics (Pieper et al., 2004). *P. putida* strains have been ranked among the most solvent-tolerant bacteria known (Nicolaou et al., 2010), suitable for enhancing their biotechnological production of compounds by metabolic engineering (Ewering et al., 2006). Numerous *Pseudomonads* exhibit a wide biotechnological potential, producing a series of bulk and fine chemicals, which has led to a growing interest in studying specific metabolic pathways at the gene expression and regulation levels (Ballerstedt et al., 2007). In line with this, mathematical models of promoter/regulator systems (Van Dien and de Lorenzo, 2003) and genome-scale metabolic models have been developed (Nogales et al., 2008; Puchalka et al., 2008) aiming at exploring the vast biotechnological potential of this bacterium. *P. putida* mt-2 is the best characterised toluene-degrading bacterium. This strain harbours the TOL plasmid (pWW0), which specifies a pathway for the oxidative catabolism of toluene and *m*-xylene (Timmis, 2002). The required genetic information for the metabolism of these compounds is encoded by the *xyl* operons of the plasmid, synthesising the relevant biocatalytic enzymes for conversion of substrates to Krebs cycle intermediates, while *xylS* and *xylR* are involved in transcriptional control (Ramos et al., 1997). The complex interactions between TOL plasmid-encoded regulators, a set of sigma factors and DNA-bending proteins, resulting in expression by the catabolic operons, render the TOL plasmid a paradigm of specific and global regulation (Aranda-Olmero et al., 2005).

This work establishes a quantitative framework that links molecular to macroscale events in microbial systems. To this end, we have recently paved the way developing a dynamic mathematical model of the *Ps/Pr* node of the TOL plasmid, involved in the metabolism of *m*-xylene by *P. putida* mt-2 (Koutinas et al., 2010). Herein, this model is extended to account for the regulatory effects pertinent to the function of *Pu* and *Pm* promoters, driving the transcription from the operons, to construct a complete model of the genetic interactions encountered in TOL. The mRNA transcript levels computed in the genetic circuit model are linked to specific growth and substrate utilisation rates through computation of the rate-limiting enzymes synthesised by the operons. The parameter values of the combined model are estimated in a single experiment and its predictive capability is evaluated through a series of independent experiments. Our results show that the combined dynamic model effectively describes the function and dynamics of the system, at both the microscopic (promoter mRNA concentration) and macroscopic (substrate and biomass concentration) levels. The considerable accuracy of the combined model in predicting the performance of the system for a wider range of conditions, compared to Monod-type models, highlights the importance of this novel modelling approach in capturing essential molecular dynamics and using these effectively for the prediction of bioprocess reaction kinetics. Such a strategy may provide fundamental information for the realisation of bioprocesses pertaining complex regulatory programmes that govern their dynamics.

2. Materials and methods

2.1. Growth conditions

All subcultures of *P. putida* mt-2 were pregrown overnight at 30 °C in M9 minimal medium (Sambrook et al., 1989) supplemented with 15 mM succinate. Triplicate cultures were prepared by diluting the overnight culture in minimal medium to an initial optical density of 0.1 at 600 nm (UV-2101PC, Shimadzu UK Ltd, UK). The minimal medium was supplemented with *m*-xylene at a concentration level in agreement to the requirements of each experiment. The incubation of the cultures was performed using conical Erlenmeyer flasks with 2.35 L total volume (0.4 L culture volume), which were continuously stirred at 1250 rpm via a Heidolph MR 3001 K (Heidolph, UK) magnetic stirrer. The flasks were filled with medium to one-sixth of their volume, to ensure that sufficient oxygen is available, and closed gas-tight with Teflon coated lids to avoid losses of the volatile organic compound. Temperature was maintained constant at 30 °C. All chemicals used were obtained from Sigma-Aldrich Company Ltd (UK) and were of ANALAR grade. *m*-xylene was obtained from VWR International Ltd (UK) 99% pure.

2.2. Analyses

Gas Chromatograph (GC) analysis was employed for determination of the *m*-xylene concentration in the gaseous and aqueous samples using an Agilent 6850 Series II Gas Chromatograph with a FID detector and a 'J&W Scientific' (Agilent Technologies UK Limited, UK) column with HP-1 stationary phase (30 m × 0.32 mm × 0.25 μm). Gaseous samples of 25 μL were injected into the GC and the temperature programme run at 70 °C for 3 min and then increased to 80 °C at a rate of 5 °C min^{-1}. Biomedium *m*-xylene concentration was determined experimentally as previously described (Koutinas et al., 2010). The coefficient of variation for 5 samples was 4.6% at a concentration level of 0.07 mM *m*-xylene.

Biomass concentration was determined by absorbance at 600 nm on a UV-2101PC scanning spectrophotometer (Shimadzu, UK) interpolating from a previously established dry weight calibration curve. The coefficient of variation for 5 samples was 4.2% at a concentration level of 583 $mg_{biomass}$ L^{-1}.

2.3. Isolation of total RNA, cDNA synthesis and quantitative real-time PCR

Quantitative Real-Time PCR (Q-PCR) was performed to determine the expression of *xylR* (*Pr* promoter), *xylS* (*Ps* promoter), *xylU* (*Pu* promoter), *xylX* (*Pm* promoter) and *rpoN* (housekeeping) genes during the course of the experiments. The method was performed as previously described (Koutinas et al., 2010) and the primer sequences used are shown in Table 1.

Table 1
Primers used in quantitative real time-PCR.

Primer	Description	Source
5 *xyl*R 907 RT	5′-AACTGTTTGGTGTCGATAAGG-3′	[29]
3 *xyl*R 1009 RT	3′-ATCACCTCATCAAGAAAGATGG-5′	[29]
5 *xyl*S 210 RT	5′-GGATTAGAGACCTGTTATCATCTG-3′	[29]
3 *xyl*S 318 RT	3′-GATTGAGCAGCAATAGTTCG-5′	[29]
5 *xyl*U 204 RT	5′-GCAGTTATCGGCTTCATCTC-3′	This study
3 *xyl*U 306 RT	3′-CATATAGTCGGTTGAGGTTAGC-5′	This study
5 *xyl*X 047 RT	5′-TGAAGAAGATGAGAACGAGG-3′	This study
3 *xyl*X 157 RT	3′-AGATAAATCCAGTTGCCCTC-5′	This study
5 *rpo*N 1067 RT	5′-TAACGAAACCCTGATGAAGG-3′	[29]
3 *rpo*N 1169 RT	3′-AATGTCATGCAGTACCAACG-5′	[29]

2.4. Statistical analysis

SigmaStat (Systat Software UK Ltd, UK, version 3.5) was used for calculation of 95% confidence intervals between the experimental data of the promoters mRNA concentration and the model's prediction, in order to elucidate the precision of the model in predicting the experimental results.

2.5. Parameter estimation in gPROMS

All parameter estimation experiments and model simulations were carried out on an Intel® Core™2 Duo (E4600–2.4, 2.39) personal computer with 3.24 GB of RAM memory. All model simulations and parameter estimation experiments were implemented in the advanced process modelling environment gPROMS® (Process Systems Enterprise, 2010). gPROMS is an equation-oriented modelling system used for building, validating and executing first-principles models within a flowsheeting framework. Parameter estimation in gPROMS is based on the Maximum Likelihood formulation which provides simultaneous estimation of parameters in both the physical model of the process as well as the variance model of the measuring instruments. gPROMS attempts to determine values for the uncertain physical and variance model parameters, θ, that maximise the probability that the mathematical model will predict the measurement values obtained from the experiments. Assuming independent, normally distributed measurement errors, ε_{ijk}, with zero means and standard deviations, σ_{ijk}, this maximum likelihood goal can be captured through the following objective function:

$$\Phi = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\min_{\theta}\left\{\sum_{i=1}^{NE}\sum_{j=1}^{NVi}\sum_{k=1}^{NMij}\left[\ln(\sigma_{ihk}^2) + \frac{(\bar{z}_{ijk}-z_{ijk})^2}{\sigma_{ihk}^2}\right]\right\}$$

where N stands for total number of measurements taken during all the experiments, θ is the set of model parameters to be estimated, NE is the number of experiments performed, NV_i is the number of variables measured in the ith experiment and NM_{ij} is the number of measurements of the jth variable in the ith experiment. The variance of the kth measurement of variable j in experiment i is denoted as σ_{ijk}^2, while $\bar{z}_{ijk}$ is the kth measured value of variable j in experiment i and z_{ijk} is the kth (model-)predicted value of variable j in experiment i. The above formulation can be reduced to a recursive least squares parameter estimation if no variance model for the sensor is selected.

3. Results

3.1. Mathematical modelling of the TOL genetic circuit

A complete model of the transcriptional regulation of TOL was constructed based on existing biological knowledge of its function

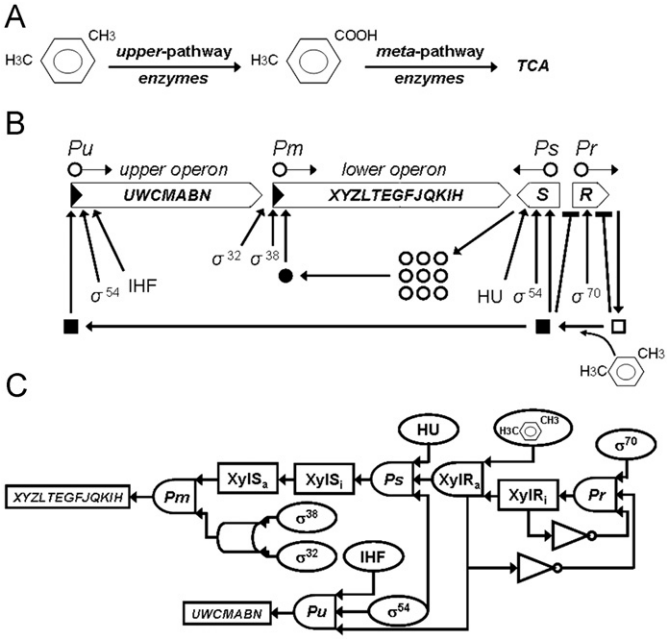

Fig. 1. Transcriptional regulation of the TOL pathway. (A) The enzymes encoded in the *upper* operon sequentially transform *m*-xylene into 3-methylbenzoate. The latter is then transformed into acetate and pyruvate through the action of the enzymes synthesised by the *meta* operon. The *meta* pathway products are channelled into the Krebs cycle yielding the precursor molecules required to support biomass growth. (B) The *upper* operon is transcribed from the σ^{54} promoter *Pu* upon activation by the central regulator of the pathway (*XylR*) bound to specific effectors such as *m*-xylene. The *meta* operon is transcribed from the *Pm* promoter, which is activated by either *XylS* and 3-methylbenzoate as a co-inducer, or by high levels of *XylS* expression alone (considered in this work). *xylS* and *xylR* are transcribed from the divergent and overlapping promoters *Ps* and *Pr*, respectively. The regulation of these promoters is connected, because *Ps* is activated by XylR, which also binds and downregulates its own production from *Pr*. (C) A logic representation of the transcriptional regulation in TOL has been produced with the use of logic gates. The various regulatory modules of TOL have been rationally combined to construct a conceptual representation of the environmental signal (*m*-xylene) activating the existing gene expression programme of the plasmid, which changes the cell metabolism through the control of the production rate of the downstream protein expression machinery. ○: inactive form of *XylS* ($XylS_i$); ●: active form of *XylS* ($XylS_a$); □: inactive form of *XylR* ($XylR_i$); ■: active form of *XylR* ($XylR_a$); ⬭: input; ▭: output; ⋂: AND; ⋃: OR; ▷: NOT.

(Moreno et al., 2010; Ramos et al., 1997) and specified by various molecular components, which interact to guide the catabolism of *m*-xylene (Fig. 1A, B). Consequently, the action of the various molecules, genes and gene products has been conceptually described as a combination of logic gates (Fig. 1C), based on biochemical inverters (Weiss et al., 2003), to produce a simple representation of the plasmid's regulatory logic and its expression. Based on this logic representation, Hill functions were used as input functions to the genes (Alon, 2006) obtaining a dynamic mathematical model of the system controlling the production of the enzymes required for the metabolism of *m*-xylene and similar compounds.

3.2. Pr promoter

Expression of *xylR* is driven by two constitutive σ^{70}-dependent tandem promoters (*Pr1* and *Pr2*) synthesising the central regulator of the pathway, the XylR protein. Following binding with an aromatic effector such as *m*-xylene, the inactive dimer form of the protein ($XylR_i$) binds ATP, triggering the multimerization of the regulator to form a hexamer (Bertoni et al., 1998). The multimerized protein undergoes conformational changes producing a transcriptionally competent form of *XylR* ($XylR_a$) (Devos et al., 2002), which

results in rapid transition between the active and inactive forms of the protein (Shingler, 2003). Consequently, $XylR_a$ induces transcription from *Pu* and *Ps* triggering the response of the pathway towards *m*-xylene biodegradation. The synthesis of $XylR_i$ from *Pr* driven transcription and the mechanism for XylR activation/deactivation are expressed by Eqs. (1)–(3).

$$\frac{dXylR_i}{dt} = \beta_{XylR_i}\mathrm{Pr}_{TC} - r_{XylR}XylR_i + 3r_{R,XylR}XylR_a - \alpha_{XylR_i}XylR_i \quad (1)$$

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

$$r_{XylR} = r_{R,XylR}\frac{Xyl}{Xyl_{INI}} \quad (3)$$

$XylR_i$ and $XylR_a$ refer to the concentrations of the inactive and active forms of *XylR* protein, respectively. β_{XylRi} is the translation rate of *Pr* promoter's mRNA, Pr_{TC} is the relative mRNA concentration of *Pr*, r_{XylR} is the $XylR_i$ oligomerization constant, $r_{R,XylR}$ is the $XylR_a$ dissociation constant, *Xyl* is the concentration of *m*-xylene, Xyl_{INI} is the initial *m*-xylene concentration at the parameter estimation experiment, *t* is the time, and α_{XylRi} accounts for $XylR_i$ degradation and dilution due to cellular volume increase.

In *Ps* the upstream activator sequences (UAS), where *XylR* targets, overlap the *Pr* promoters triggering a transcriptional switch between *Pr* and *Ps* (Bertoni et al., 1998). Thus, both $XylR_i$ and $XylR_a$ are known to repress expression from *Pr* reducing their own synthesis (Eq. (4)). For simplification of the model developed we have assumed that the concentration of σ^{70} is constant at housekeeping level (previously shown in Ishihama, 2000) and we express both *xylR* tandem promoters as a single *Pr* promoter.

$$\frac{dPr_{TC}}{dt} = \beta_{Pr}\left(\frac{K^3_{XylR_i}}{K^3_{XylR_i}+XylR_i^3} + \frac{K_{XylR_i}}{K_{XylR_i}+XylR_a}\right) - \alpha_{Pr}Pr_{TC} \quad (4)$$

β_{Pr} stands for the maximal expression level of *Pr*, K_{XylRi} is the repression coefficient of *Pr* due to $XylR_i$ binding, and α_{Pr} is the mRNA degradation rate of *Pr*. The Hill coefficients of *Pr* due to XylR binding have been set to three and one, respectively, in order to account for the presence of three dimers or one hexamer on the binding site at a dynamic equilibrium.

Comparing the expression used for *Pr* to the one presented in our previous work with the *Pr*/*Ps* system (Koutinas et al., 2010), the following minor differences can be identified. Given that succinate has not been used in the experiments of this study, the catabolite repression effect of this compound on *Pr* has been removed from Eq. (4). The Hill coefficient values for the terms accounting for repression of *Pr* due to $XylR_i$ and $XylR_a$ binding have been set to 3 and 1, respectively, to account for the three $XylR_i$ dimers required to form a single $XylR_a$ hexamer. Furthermore, the activation/deactivation mechanism for *XylR* has been modified to consider that only $XylR_i$ is activated by the presence of *m*-xylene to form $XylR_a$, while degradation of the latter is omitted to account for the fast transition between the two forms of this protein.

3.3. Pu promoter

The *upper* pathway of TOL is controlled by the σ^{54}-dependent promoter *Pu*, which is activated at a distance by $XylR_a$ (Velazquez et al., 2006). The binding of the regulator to upstream activating sequences and the looping out of the complex closely to the σ^{54}-RNAp complex is a process assisted by the integration host factor (IHF). This structural change optimises the promoter's geometry stimulating *Pu* to drive the synthesis of the enzymes coded by the *upper* operon (Carmona et al., 1999). σ^{54} is known to maintain a balanced intracellular concentration through the growth time (Cases et al., 1996; Jishage et al., 1996). Furthermore, when *P. putida* is grown in a rich LB medium the concentration of IHF is highest at early stationary phase (Valls et al., 2002), delaying the increase in the transcriptional output from *Pu*. Based on the fact that such a delay was not monitored in the present study, under M9 medium conditions, and due to the lack of experimental information about σ^{54} and IHF concentrations, it is assumed that σ^{54} and IHF are not limiting. Thus, the relative mRNA concentration of *Pu* is given by

$$\frac{dPu_{TC}}{dt} = \beta_{Pu}\frac{XylR_a}{K_{XylR_a,Pu}+XylR_a} - \alpha_{Pu}Pu_{TC} \quad (5)$$

Pu_{TC} stands for the relative mRNA concentration of *Pu*, β_{Pu} is its maximal expression level, $K_{XylRa,Pu}$ is the activation coefficient of *Pu* due to $XylR_a$ binding, and α_{Pu} is the mRNA degradation rate of the promoter.

Pu controls the genetic information (genes *xylUWCMABN*) for the *upper* pathway, which is located in the *upper* operon. The enzymatic products of *xylC* (benzaldehyde dehydrogenase), *xylM* and *xylA* (xylene monooxygenase), and *xylB* (benzyl alcohol dehydrogenase) are involved in the oxidation of aromatic effectors to the corresponding benzyl alcohol, benzaldehyde, and benzoate, respectively (Van Dien and de Lorenzo, 2003). Furthermore, it is widely accepted that control of the enzyme level on the flux of a pathway is distributed between all participating enzymes (Douma et al., 2010). Nevertheless, these enzymes do not exert the same level of control on the flux of the pathway, which is usually dominated by a single rate-limiting enzyme. The production of this rate-limiting enzyme for the conversion of *m*-xylene to 3-methylbenzoate is expressed as a function of the relative mRNA concentration of *Pu* (Eq. (6)).

$$\frac{dXylU}{dt} = \beta_{XylU}Pu_{TC} - \alpha_{XylU}XylU \quad (6)$$

XylU stands for the concentration of the assumed rate-limiting enzyme of the *upper* pathway, β_{XylU} is the translation rate based on *Pu* mRNA, and α_{XylU} accounts for XylU degradation and dilution due to cellular volume increase.

3.4. Ps promoter

The *xylS* gene is transcribed at low levels from the constitutive σ^{70}-dependent *Ps2* promoter and its expression is boosted in the presence of *m*-xylene, through $XylR_a$ regulated induction of the σ^{54}-dependent *Ps1* promoter (Gonzalez-Perez et al., 2004). We have previously shown that in the absence of TOL pathway effectors activation of *Ps* is very low (Koutinas et al., 2010). Therefore, the basal expression level of *Ps* is considered insignificant and it is not accounted for the mRNA concentration of *Ps*. Thus, we assume that the induction of *Ps2* is negligible under the conditions of our experiments and we account only for the $XylR_a$ induced *Ps1* promoter, which is referred in the rest of the manuscript as the *Ps* promoter. Similarly to the effect of IHF on *Pu*, the binding of HU protein on *Ps* facilitates the correct architecture of the promoter bringing into contact the $XylR_a$ bound to the UAS and the σ^{54}-RNAp complex (Perez-Martin and de Lorenzo, 1995). Based on previous findings showing that HU is abundant in *Escherichia coli* (Ishihama, 1999) and σ^{54} is a constitutive protein (Merrick, 1993), it is assumed that the concentrations of HU and σ^{54} are constant at housekeeping level. Furthermore, although IHF may significantly repress *Ps* activity (Marques et al., 1998), under inducing conditions this effect is not significant due to strong $XylR_a$ binding to *Ps* (Holtel et al., 1992). Thus, it is assumed that *Ps* is IHF-independent and its relative mRNA concentration

is given by

$$\frac{dPs_{TC}}{dt} = \beta_{Ps}\frac{XylR_a}{K_{XylR_a,Ps}+XylR_a} - \alpha_{Ps}Ps_{TC} \quad (7)$$

Ps_{TC} stands for the relative mRNA concentration of *Ps*, β_{Ps} is its maximal expression level, $K_{XylRa,Ps}$ is the activation coefficient of *Ps* by $XylR_a$, and α_{Ps} is the mRNA degradation rate of the promoter.

The mRNA translated from *Ps1* is 10 times more efficient than that from *Ps2* (Gonzalez-Perez et al., 2004). This pattern of translation, together with the pattern of transcription described above, lead to two types of activation of the *XylS* protein. In cells growing with *meta* pathway effectors, such as benzoate, the small amounts of *XylS* produced constitutively from *Ps2* stimulate transcription from *Pm* following binding to the effector. However, in cells growing with *upper* pathway effectors, such as *m*-xylene, the high *XylS* concentration synthesised by *Ps1* suffice to stimulate *Pm* even when *XylS* effectors are not present (Dominguez-Cuevas et al., 2008). In both cases the activation of *XylS* occurs due to conformational changes, driven either by the addition of *meta* pathway effectors or by the presence of high *XylS* concentrations, leading to dimerisation of *XylS* monomers. The synthesis of *XylS* from *Ps* driven transcription and the mechanism for *XylS* activation/deactivation in the presence of *m*-xylene are expressed by

$$\frac{dXylS_i}{dt} = \beta_{XylS_i}Ps_{TC} - r_{XylS}XylS_i + 2r_{R,XylS}XylS_a - \alpha_{XylS_i}XylS_i \quad (8)$$

$$\frac{dXylS_a}{dt} = \frac{1}{2}r_{XylS}XylS_i - r_{R,XylS}XylS_a \quad (9)$$

where $XylS_i$ and $XylS_a$ refer to the concentrations of the inactive and active forms of *XylS* protein, respectively, β_{XylSi} is the translation rate based on *Ps* mRNA, r_{XylS} is the $XylS_i$ oligomerization constant, $r_{R,XylS}$ is the $XylS_a$ dissociation constant, and α_{XylSi} accounts for $XylS_i$ degradation and dilution due to cellular volume increase.

3.5. Pm promoter

The *meta*-cleavage pathway operon is under the control of a single *Pm* promoter and its transcription is stimulated by the presence of the activated form of the *XylS* protein. This process is mediated by RNAp with σ^{32} or σ^{38} depending on the growth phase (Gonzalez-Perez et al., 2002). It was recently shown that the $XylS_i$ monomer also binds to *Pm* (Dominguez-Cuevas et al., 2008). Thus, the relative mRNA concentration of *Pm* is given by

$$\frac{dPm_{TC}}{dt} = \beta_{Pm}\frac{XylS_a}{K_{XylS_i,Pm}+XylS_a}\frac{XylS_i^2}{K^2_{XylS_i,Pm}+XylS_i^2} - \alpha_{Pm}Pm_{TC} \quad (10)$$

Pm_{TC} stands for the relative mRNA concentration of *Pm*, β_{Pm} is its maximal expression level, $K_{XylSi,Pm}$ is the activation coefficient of *Pm* by *XylS*, and α_{Pm} is the mRNA degradation rate of the promoter.

The *meta* operon comprises 13 genes (*xylXYZLTEGFJQKIH*) consisting one of the largest operons in prokaryotes (Ramos et al., 1997). The transcription originating at the *Pm* promoter leads to the formation of nine enzymes involved in the *meta* pathway. Thus, the benzoate yielded through the action of the enzymes synthesised in the *upper* operon is *cis*-dioxygenated in the *meta* pathway to form 3-methylcatechol, which is cleaved in *meta* and consequently channelled into the Krebs cycle (Velazquez et al., 2005). Similarly to the assumption made for the *upper* pathway, the conversion of benzoate to Krebs cycle intermediates is modelled as being exclusively controlled by a single rate-limiting enzyme. Therefore, the synthesis of the rate-limiting enzyme becomes a function of the relative mRNA concentration of *Pm*, as shown below

$$\frac{dXylM}{dt} = \beta_{XylM}Pm_{TC} - \alpha_{XylM}XylM \quad (11)$$

XylM stands for the concentration of the assumed rate-limiting enzyme of the *meta* pathway, β_{XylM} is the translation rate based on *Pm* mRNA, and α_{XylM} accounts for *XylM* degradation and dilution due to cellular volume increase.

3.6. Coupling TOL to the growth kinetics

Predicting degradation of substrates and dynamics of growth with models based on enzyme kinetics is not straightforward (Littlejohns and Daugulis, 2008), because the production of enzymes in a catabolic pathway is often subject to transcriptional regulation controlling their expression (Kovarova-Kovar and Egli, 1998). This mechanism clearly indicates that unless the residual concentrations of growth-controlling substrates are linked to the regulatory loops governing the elements coding the genetic basis for the synthesis of enzymes, the applicability of the developed model will be usually restricted to a narrow range of conditions. In order to account for the regulatory events controlling the metabolism of *m*-xylene in *P. putida* mt-2, a mathematical model coupling the mRNA transcript levels computed in the TOL model to specific growth and substrate utilisation rates has been developed.

Given that the *upper* pathway drives the first reaction for *m*-xylene bioconversion to the various intermediates involved in TOL, *m*-xylene consumption has been modelled as a function of the concentration of the rate-limiting enzyme controlled by *Pu* (Eq. (12)). Furthermore, the *meta* pathway channels the products of the TOL pathway from *m*-xylene biodegradation into the Krebs cycle providing the precursor molecules required for the anabolic processes which result in biomass growth. Thus, the specific growth rate on *m*-xylene is expressed as a function of the concentration of the rate-limiting enzyme controlled by *Pm* (Eq. (13)). Consequently, biomass production is modelled using Eq. (14) and its decay is expressed by Eq. (15).

$$\frac{dXyl}{dt} = -\frac{1}{MW_{m-x}}\frac{\beta_{XylU,m-x}XylU}{K_{XylU,m-x}+XylU}X \quad (12)$$

$$\mu = \frac{\beta_{XylM,b}XylM}{K_{XylM,b}+XylM} \quad (13)$$

$$\frac{dX}{dt} = (\mu - d)X \quad (14)$$

$$d = \frac{d_{max}Xyl}{K_d+Xyl} \quad (15)$$

MW_{m-x} is the molecular weight of *m*-xylene, $\beta_{XylU,m-x}$ stands for the maximum *m*-xylene metabolic quotient based on *XylU*, $K_{XylU,m-x}$ is the saturation constant for *XylU*, *X* refers to biomass concentration, μ is the specific growth rate of biomass on *m*-xylene, $\beta_{XylM,b}$ is the maximum specific growth rate of biomass based on *XylM*, $K_{XylM,b}$ is the saturation constant for *XylM*, *d* accounts for the decay rate, d_{max} is the maximum decay rate, and K_d refers to the decay saturation constant.

In many cases, microbial growth kinetics are modelled by applying the Monod equation to describe the specific growth rate (Eq. (16)). Furthermore, when an inhibitory substrate such as *m*-xylene is fed, a Monod term accounting for substrate inhibition can be used (Eq. (17)) (Morgado et al., 2004; Yano and Koga, 1969). Therefore, the consumption of the substrate is given by

Eq. (18) and biomass production is expressed by Eq. (14)

$$\mu = \frac{\mu_{max,1} Xyl}{K_{S,1} + Xyl} \tag{16}$$

$$\mu = \frac{\mu_{max,2} Xyl}{K_{S,2} + Xyl + Xyl^3/K_I^2} \tag{17}$$

$$\frac{dXyl}{dt} = -\frac{\mu}{Y} X \tag{18}$$

$\mu_{max,1}$ and $\mu_{max,2}$ stand for the maximum specific growth rates of biomass on *m*-xylene, $K_{s,1}$ and $K_{s,2}$ are the *m*-xylene saturation constants, K_I refers to the *m*-xylene inhibition constant, and Y is the yield coefficient for biomass on *m*-xylene.

Comparing the specific growth and substrate consumption rates of the two models based on enzyme kinetics (Monod, and Yano and Koga) and that of the combined model that accounts for transcriptional regulation of the *m*-xylene metabolising pathway, three main structural advantages of the latter can be highlighted:

i. Biomass growth and substrate consumption have been decoupled and they are modelled as two independent processes.
ii. Both rates are independent of the substrate concentration and they are expressed as a function of the relevant rate-limiting enzyme controlling each process. In this way the expression of key genes for the bioprocess has been linked to the growth kinetics, since the genetic circuit model is validated based on measurements of mRNA transcript levels as shown below.
iii. The combined model accounts for non-constant biomass yield, which is regulated by the production of the *upper* pathway rate-limiting enzyme.

The advantages of the combined model summarised above address limitations of models developed based on enzyme kinetics, which are unstructured and empirical. Therefore, while empirical/unstructured models constitute an indispensable tool in all fields of biotechnology, their applicability is often compromised because they do not account for complex cellular mechanisms.

3.7. Parameter estimation experiment

3.7.1. Genetic circuit model prediction

The model presented above was applied in gPROMS (Process Systems Enterprise, 2010) and the parameter values, given in Table 2, were obtained from a batch experiment in the presence of 1 mM *m*-xylene as the sole carbon source. Under induced conditions, $XylR_i$ interacts with the substrate to form the master regulator of the system ($XylR_a$). Thus, *Pr* should be repressed due to binding of its two protein products, while *Pu* and *Ps* were expected to be activated as a result of their interaction with $XylR_a$. Furthermore, the expression of *xylS* due to the presence of $XylR_a$ results in *XylS* hyperproduction driving the increase in *Pm* promoter's relative mRNA concentration. Experimental measurements of the promoters' relative mRNA levels (Fig. 2A–D) confirmed the expected system behaviour described above showing that the genetic circuit model can effectively describe the experimental results.

3.7.2. Lag-phase

When *P. putida* strains equipped with the TOL plasmid are exposed to aromatic compounds three distinct cellular responses are generated. The expression programs produced comprise a nutritional signal triggering the metabolic programme encoded in TOL, a toxic signal activating a solvent extrusion and tolerance response, and interference with the protein folding machinery setting off a heat-shock response (Dominguez-Cuevas et al., 2006). The general stress response, reflected by the latter two expression programs, includes the interaction of the aromatic compound with the phospholipid bilayer of the cell membrane disturbing its structure (Sikkema et al., 1995). This change of the membrane's structure results in the disruption of xylene monooxygenase (XMO), the enzyme initiating *m*-xylene degradation, the activity of which is known to be membrane-bound (Buhler et al., 2006). Therefore, although the genes included in TOL are transcribed without delay (apart from the genes of the *meta* pathway which show a short delay) following introduction of *m*-xylene, due to the activity loss of XMO immediate bioconversion of the aromatic compound to TOL pathway intermediates does not occur. This results in a lag-phase where the substrate is not consumed and biomass is not produced. The mathematical description of the mechanism causing the lag-phase was out of the scope of our work. Thus, from the biomass and *m*-xylene concentration profiles for each experiment we have observed the duration of the lag-phase and we have fitted the three models developed to the average values of the experimental points during this period. However, since the response of the plasmid starts immediately following introduction of *m*-xylene, the TOL model was simulated from the beginning of the experiment.

Table 2
Parameter values used for model simulation.

Parameter	Equation employed	Value
d_{max}	15	0.35 h^{-1}
K_d	15	1.79 mM
$K_{XylM,b}$	14	15.4 mM
$K_{XylRa,Ps}$	7	25.6 mM
$K_{XylRa,Pu}$	5	25.6 mM
K_{XylRi}	4	13.8 mM
$K_{XylSi,Pm}$	10	4.85 mM
$K_{XylU,m-x}$	12	4.83 mM
MW_{m-x}	12	106 g_{m-xyl} mol^{-1}
$r_{R,XylR}$	1	632 h^{-1}
$r_{R,XylS}$	8	47.4 h^{-1}
r_{XylS}	8	13.3 h^{-1}
α_{Pm}	10	3.77 h^{-1}
α_{Pr}	4	6.38 h^{-1}
α_{Ps}	7	3.37 h^{-1}
α_{Pu}	5	3.56 h^{-1}
α_{XylM}	11	5.94 h^{-1}
α_{XylRi}	1	7.45 h^{11}
α_{XylSi}	8	60 h^{-1}
α_{XylU}	6	8.61×10^{-2} h^{-1}
β_{Pm}	10	5.1 h^{-1}
β_{Pr}	4	16.1 h^{-1}
β_{Ps}	7	5.45 h^{-1}
β_{Pu}	5	5.71 h^{-1}
β_{XylM}	11	3.88×10^2 mM h^{-1}
$\beta_{XylM,b}$	13	0.58 h^{-1}
β_{XylRi}	1	7.13×10^2 mM h^{-1}
β_{XylSi}	8	1.81×10^3 mM h^{-1}
β_{XylU}	6	1.48×10^3 mM h^{-1}
$\beta_{XylU,m-x}$	12	0.19 g_{m-xyl} g_{biom}^{-1} h^{-1}

3.7.3. Combined model simulation and parameter estimation

For the description of *m*-xylene depletion and biomass growth during the parameter estimation experiment the combined, Monod, and Yano and Koga models were employed. According to *m*-xylene and biomass measurements, the duration of the lag-phase was 90 min for this experiment. Thus, the genetic circuit model (Eqs. (1)–(11)) was used to compute the concentrations of the rate-limiting enzymes *XylU* and *XylM* controlling *m*-xylene biodegradation and biomass production, respectively, (Fig. 3A, B).

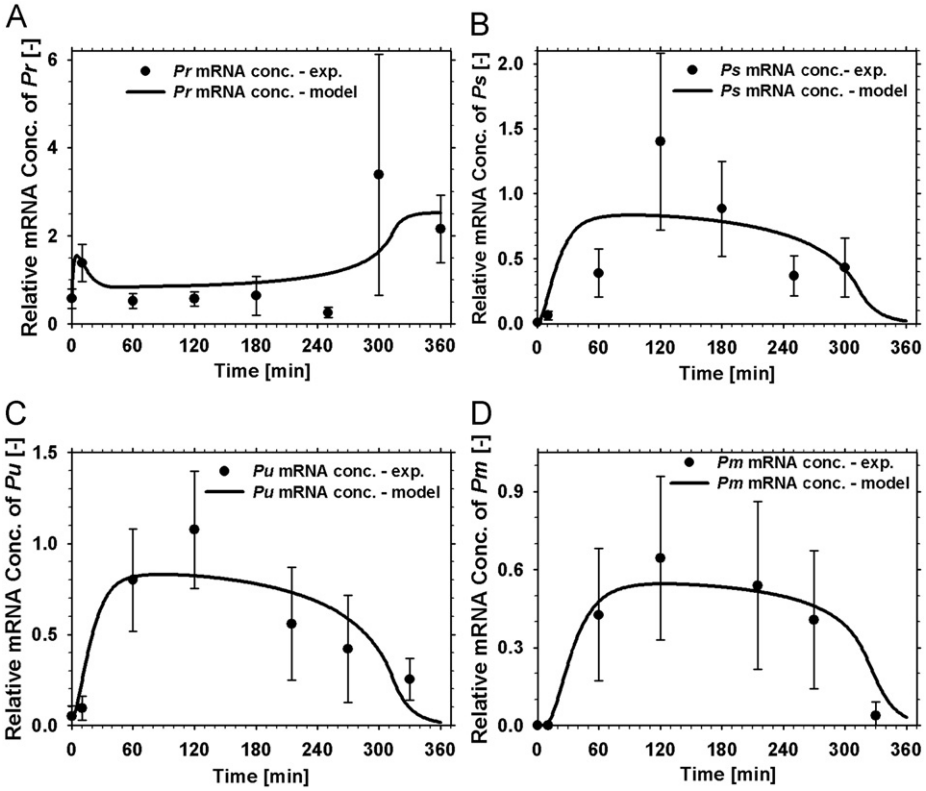

Fig. 2. Comparing the prediction of the genetic circuit model with the expression of the promoters encoded in TOL. Shown are the measured relative mRNA levels and model simulations for (A) *Pr*, (B) *Ps*, (C) *Pu*, and (D) *Pm* under the presence of 1 mM *m*-xylene in the parameter estimation experiment. The results are obtained as an average from six individual measurements at each point and the error bars are calculated for standard deviation.

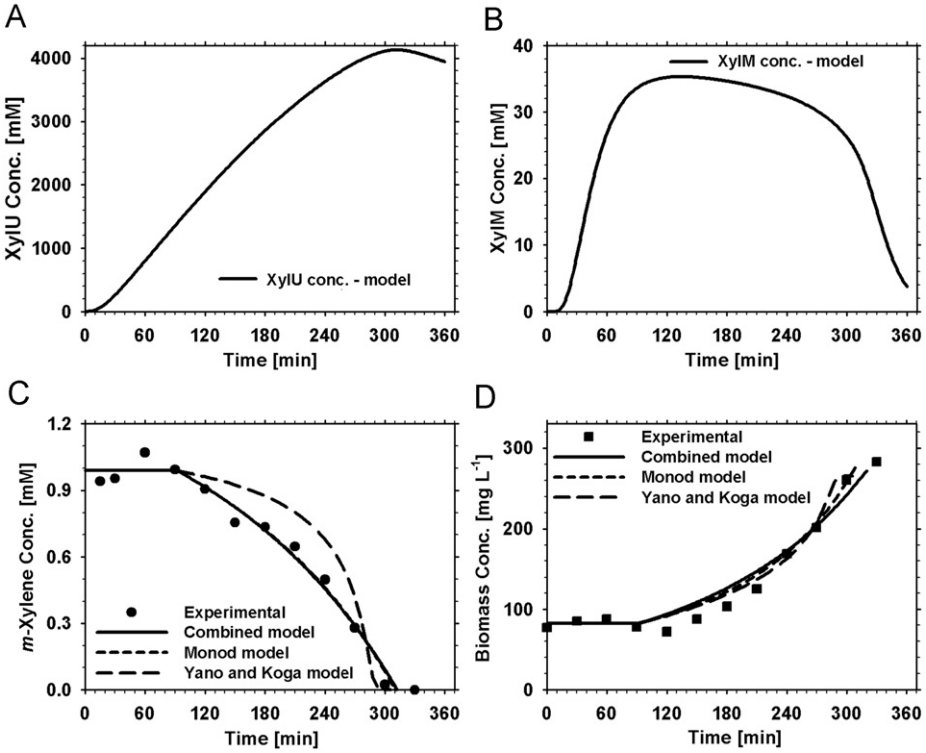

Fig. 3. *Pseudomonas putida* mt-2 growth kinetics in the parameter estimation experiment. The dynamic genetic circuit model was used to calculate the concentration profiles of the rate-limiting enzymes (A) *XylU* and (B) *XylM* synthesised in the TOL operons. Consequently, the concentration of each enzyme was used to model (C) *m*-xylene biodegradation due to the induction of the *upper* pathway and (D) biomass growth resulting from the metabolic products of the *meta* operon channelled into the Krebs cycle.

Table 3
Correlation coefficients between experimental and modelling results.

Experiment	R^2 (biomass)	R^2 (*m*-xylene)
1 (Combined model)	0.961	0.985
1 (Monod model)	0.974	0.986
1(Yano and Koga model)	0.975	0.910
2 (Combined model)	0.928	0.955
2 (Monod model)	0.608	0.817
2 (Yano and Koga model)	0.845	0.294
3 (Combined model)	0.952	1.000
3 (Monod model)	0.950	0.989
3 (Yano and Koga model)	0.872	0.314
4 (Combined model)	0.928	0.955
4 (Monod model)	0.609	0.817
4 (Yano and Koga model)	0.845	0.294

1. 1 mM *m*-xylene (parameter estimation experiment); 2. 0.4 mM *m*-xylene (predictive experiment); 3. 0.7 mM *m*-xylene (predictive experiment); 4. 1.3 mM *m*-xylene (predictive experiment).

However, Eqs. (12)–(15) were employed after the lag-phase to predict the concentrations of substrate and biomass using as initial conditions for the variables their values at 90 min (Fig. 3C, D). Based on the predictions of the three models, the high correlation coefficient values obtained (Table 3) between experimental and modelling results confirm that the experimental results were adequately described by all the models studied. Furthermore, the parameter values for the Monod, and Yano and Koga models were estimated in this experiment. It would be interesting to compare the values of the estimated parameters against values that exist in literature. However, since the biological mechanisms that follow mRNA transcription are presented here in a lumped fashion, relevant parameters could not be found for most of the parameters of the developed model. Nevertheless, parameter values relevant to mRNA transcription and degradation are in accordance with basic biological expectations (Feng et al., 2004; Hooshangi et al., 2005; Stricker et al., 2008).

3.8. Predictive experiments

Having obtained the parameter values for the growth kinetic models their predictive capability was tested with three independent experiments, to facilitate their validation under a wide range of conditions. Consequently, *P. putida* mt-2 was grown in three triplicate batch cultures utilising initial *m*-xylene concentration of 0.4, 0.7 and 1.3 mM, respectively.

3.8.1. Genetic circuit model prediction

The overall trend of the model's prediction for *Pu* and *Pm* promoters is accurate for the experiments with initial *m*-xylene concentration of 0.4 and 0.7 mM, as confirmed by correlating the model's prediction with 95% confidence intervals of *Pu* and *Pm* mRNA levels (Fig. 4A–D). Similarly, when 1.3 mM *m*-xylene was fed the model closely tracked the relative mRNA level of *Pu* for the first 120 min, with some discrepancies for the remaining of the culture when *Pu* relative mRNA concentration was slightly overpredicted (Fig. 4E). However, as confirmed by 95% confidence intervals of the experimental data, the genetic circuit model underpredicted the relative mRNA level of *Pm* for most of the experiment with 1.3 mM *m*-xylene (Fig. 4F). The higher *m*-xylene concentration used in this experiment might have resulted in accumulation of *meta* pathway effectors, such as 3-methylbenzoate, that bind *XylS* and enhance transcription from *Pm* (Dominguez-Cuevas et al., 2008). Nevertheless, although this effect might have caused an increase in the mRNA concentration of *Pm*, even in this case the model follows the general trend of *Pm* promoter's mRNA level. Thus, the results of the predictive experiments confirm that the existing model structure satisfactorily describes the experimental data of *Pu* and *Pm* relative mRNA concentrations based on the level of biological information available for the system.

The profiles of the promoters' relative mRNA concentration are in agreement with previous studies, showing that when the culture is grown in M9 medium under *m*-xylene inducing conditions *Pu* promoter is quickly induced upon exposure to the aromatic compound, while induction from *Pm* occurs approximately 15 min later (Velazquez et al., 2005). Expression of the *upper* operon is controlled by the *m*-xylene responsive regulator *XylR*, while the *meta* operon is under the control of a two-stage cascade established by the $XylR_a$ activation of *Ps* and activation of *Pm* due to $XylS_a$ binding (Fig. 1C). Therefore, since cascades are known to produce temporal programs of gene expression (Hooshangi et al., 2005), the delay in *Pm* induction is an expected behaviour.

The part of the mathematical model describing the function of *Pr* and *Ps* promoters, as well as the production of $XylR_i$ and its subsequent multimerization to $XylR_a$ has been recently experimentally validated (Koutinas et al., 2010). Therefore, measuring the relative mRNA concentration of these promoters in the predictive experiments was not considered necessary and their predicted profiles, as well as the predicted concentration profiles of the protein regulators of the system ($XylR_i$, $XylR_a$, $XylS_i$ and $XylS_a$) are presented in Fig. 5.

3.8.2. Comparison of the combined model with Monod-type models

Based on the prediction of the genetic circuit model, the concentrations of the rate-limiting enzymes controlled by *Pu* and *Pm* were calculated for the predictive experiments (Fig. 5D, H). The duration of the lag-phase for *m*-xylene biodegradation was 10, 150 and 100 min for the experiments with initial substrate concentrations of 0.4, 0.7 and 1.3 mM, respectively (Fig. 6A–F). Based on the correlation coefficient values calculated between experimental and modelling results (Table 3), only the combined model was able to accurately describe the measured biomass and *m*-xylene concentration profiles for all three experiments. The Yano and Koga model accounting for substrate inhibition was only capable to predict the experiment performed in the *upper* *m*-xylene concentration tested with reasonable accuracy, failing to predict the growth kinetics when 0.4 or 0.7 mM *m*-xylene were used. Likewise, although the Monod model closely tracked *m*-xylene and biomass concentrations in the batch cultures fed with 0.7 and 1.3 mM *m*-xylene, it unsuccessfully described the growth kinetics at the lowest substrate concentration tested. Furthermore, it is interesting to note that apart from underpredicting *m*-xylene biodegradation and overpredicting biomass concentration under 0.4 mM *m*-xylene conditions, the Monod model calculated 32% higher biomass production compared to that measured experimentally. On the contrary the combined model accurately described all three experiments confirming that when quantitative information of the regulatory effects controlling a bioprocess is considered the model's prediction can be valid under a wide range of conditions.

Although the developed model has more parameters than the Monod-type models used for comparison purposes, the fact that the combined model has been successfully tested in three independent experiments under a wide range of conditions indicates that the model's predictive capability is satisfactory. Furthermore, although variables accounting for the protein products of the 4 promoters have been introduced, none of these proteins have been experimentally measured. Therefore, the parameters associated with these variables cannot be uniquely identified from the available experimental data. To mitigate this

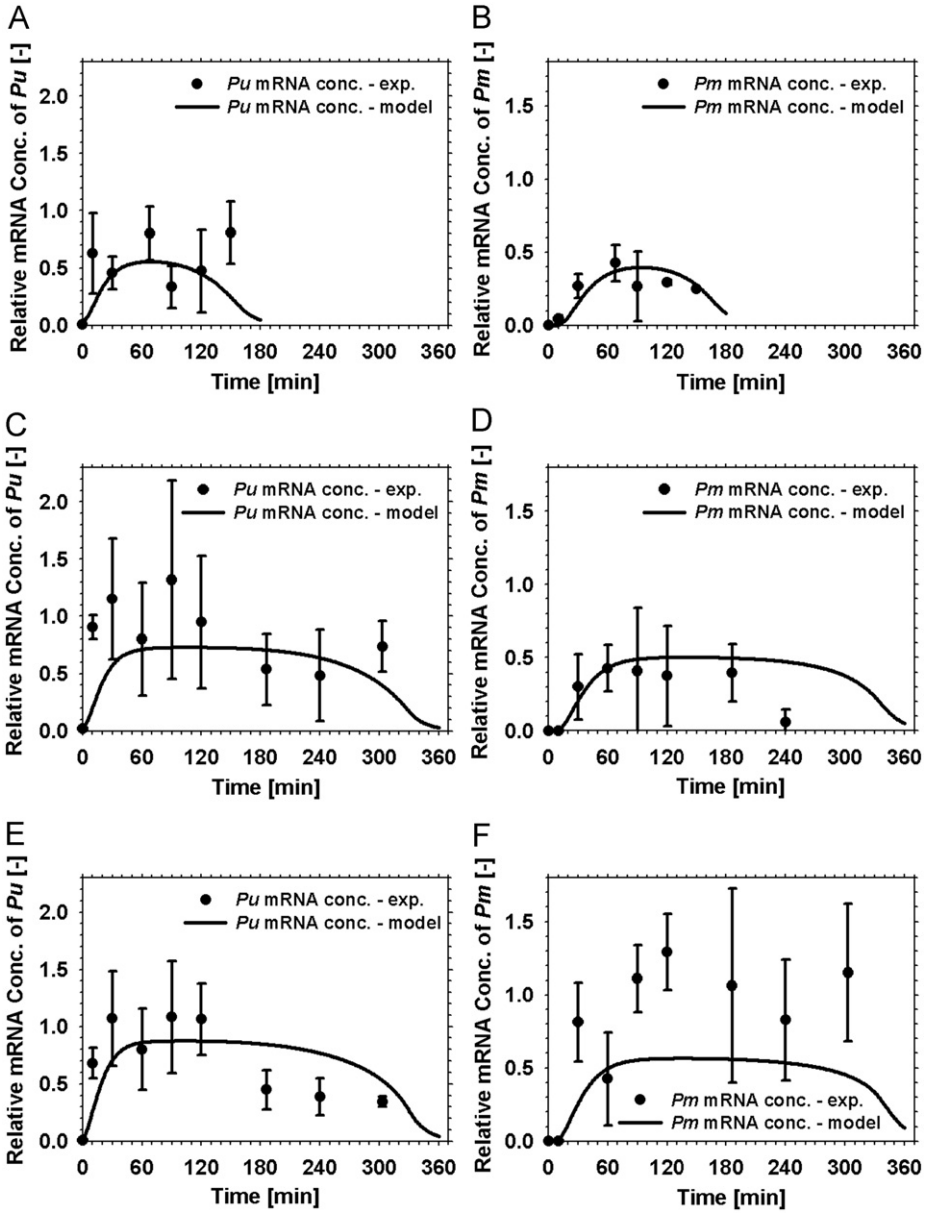

Fig. 4. An overview of the genetic circuit model prediction of *Pu* and *Pm* promoters relative mRNA concentration. Shown are simulation and experimental results for the three predictive experiments. (A, B) 0.4 mM *m*-xylene, (C, D) 0.7 mM *m*-xylene, and (E, F) 1.3 mM *m*-xylene. The results are obtained as an average from six individual measurements at each point and the error bars are calculated for standard deviation.

source of parametric uncertainty Global Sensitivity Analysis (GSA) has been performed scanning all model parameters over a space $\pm 90\%$ from their estimated value. According to the GSA results (Supplemental Material), none of the parameters associated with the genetic circuit has a significant effect on the model outputs of interest (*m*-xylene and biomass concentrations). The GSA results mathematically guarantee that the estimated parameter values are adequate approximations.

4. Discussion

Understanding the functional relationship between gene expression and growth kinetics is a big challenge. The TOL genetic circuit studied for the metabolism of *m*-xylene is abundant among *Pseudomonas* spp. and given its metabolic versatility, the elements driving its response remain important modules for the construction of microbes with new activities. We have built and validated a mathematical model that captures the essential regulatory features of TOL taking into account the observed dynamics of the regulatory circuit in a quantitative manner. The dynamic behaviour of the genetic circuit has been coupled to the growth kinetics of the strain developing a systematic framework that links molecular-level understanding of a system to macroscopic bioprocess behaviour. We demonstrate that substrate consumption and biomass production can be decoupled and described in a mechanistic way based on the molecular phenomena controlling each of the two processes, while the presence of non-constant yields can be also considered reducing the error often generated when constant yield is assumed.

The combination of advanced genetic techniques with mathematical models capturing essential quantitative features of experimental measurements can bring a breakthrough from descriptive and empirical approaches to a modern form of mathematical reasoning enriching our understanding of bioprocesses (Bailey, 1998). Advances in sequencing and genetic

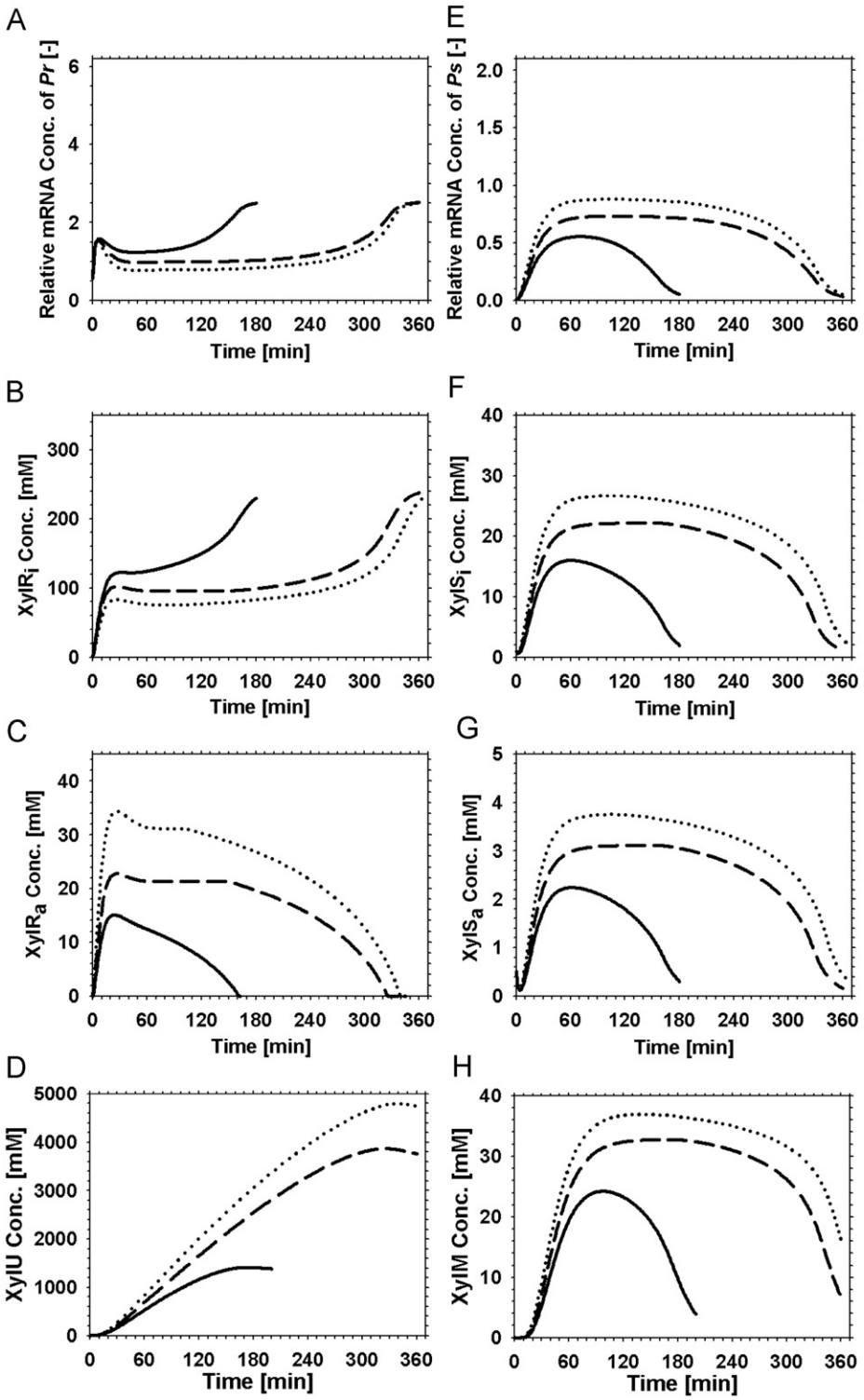

Fig. 5. Predicted dynamic profiles of the TOL regulatory elements controlling the synthesis of catabolic rate-limiting enzymes. Shown are model simulations for the three predictive experiments. The genetic circuit model is applied to calculate the time course of the relative mRNA level of (A) *Pr*. The induction of *Pr* is used to determine the production of (B) $XylR_i$, which in the presence of *m*-xylene is multimerized to form (C) $XylR_a$. The latter stimulates transcription from *Pu* regulating the synthesis of (D) *XylU*. $XylR_a$ also regulates transcription from (E) *Ps* synthesising (F) $XylS_i$. The hyperproduction of $XylS_i$ results in its dimerisation to form (G) $XylS_a$, which subsequently regulates the synthesis of (H) *XylM*. –––: predicted concentration – 0.4 mM *m*-xylene; – –: predicted concentration—0.7 mM *m*-xylene;……: predicted concentration—1.3 mM *m*-xylene.

engineering have made this approach feasible through the study of naturally occurring genetic circuits with mathematical models (Hasty et al., 2002). There has been a variety of previous studies focused on quantitative modelling of genetic circuits, including the broadly studied lactose (*lac*) operon (Lee and Bailey, 1984), bistable switches (Isaacs et al., 2003), cell cycle regulatory systems (Laub et al., 2007), circadian clocks (Hardin, 2005), and oscillating networks (Elowitz and Leibler, 2000) harboured by wild-type microorganisms. Furthermore, artificial genetic circuits embedded within cellular genetic circuits may provide

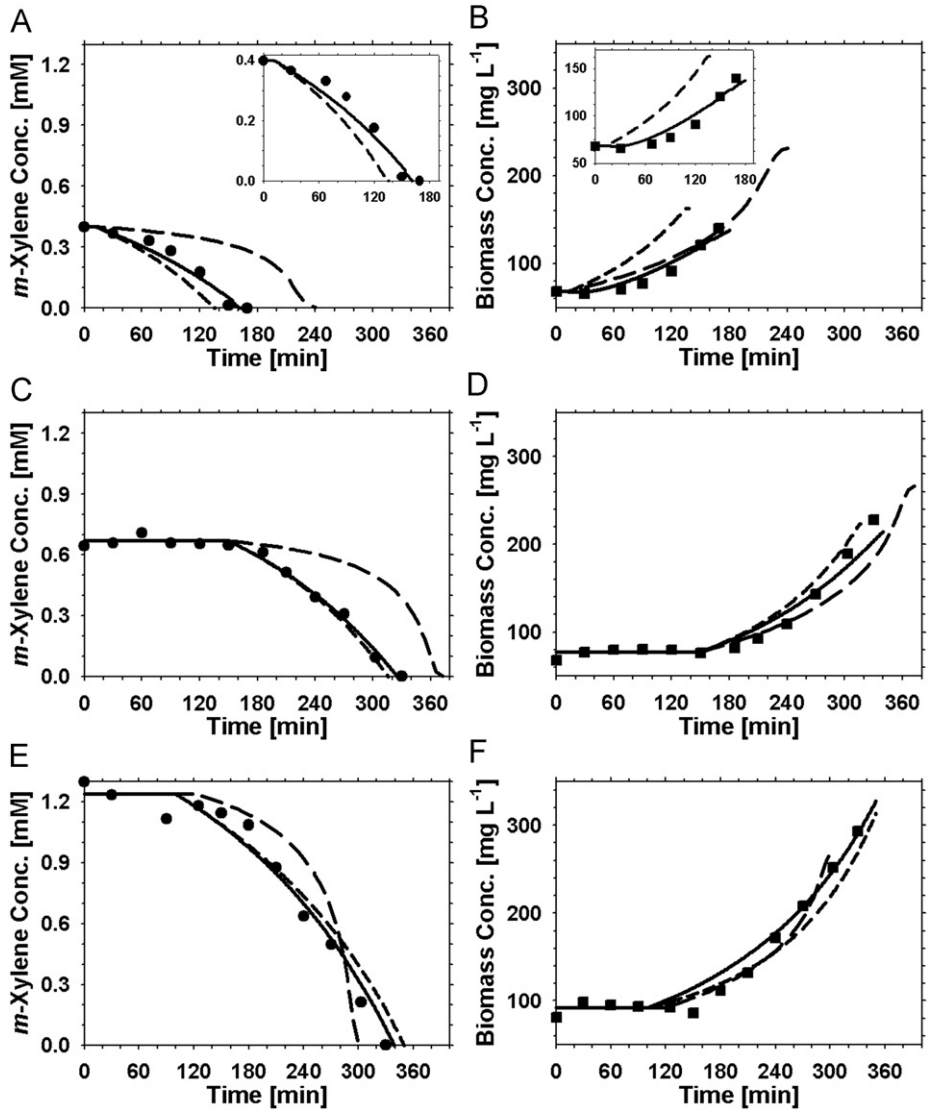

Fig. 6. Comparisons of *Pseudomonas putida* mt-2 growth kinetics predictions for the three models developed. The TOL model was used to calculate the concentration profiles of the rate-limiting enzymes *XylU* and *XylM* regulating *m*-xylene biodegradation and biomass growth, respectively. Therefore, the concentration of each enzyme was used to predict *m*-xylene and biomass dynamic profiles in the combined model, which is compared to the Monod, and Yano and Koga models. Shown are simulation and experimental results for the three predictive experiments. (A, B) 0.4 mM *m*-xylene—the insets show a zoom in of *m*-xylene and biomass concentration profiles, (C, D) 0.7 mM *m*-xylene, and (E, F) 1.3 mM *m*-xylene. ●: *m*-xylene concentration—experimental; ■ : biomass concentration—experimental;----- : combined model; – – –: Monod model; –– –: Yano and Koga model.

opportunities for sophisticated interface engineering between living and non-living systems (Kramer et al., 2005). Therefore, mathematical modelling approaches employed to study the function of key genetic circuits for a bioprocess can reveal important links of the particular circuit to the function of the biological system (Han, 2008).

Mathematical modelling is becoming one of the prominent areas of biological studies and it has the potential to drive bioprocessing towards a more precise engineering discipline. Various modelling approaches have been employed to study the properties of biological systems. Flux balance analysis (FBA) has been useful for analysing the behaviour of large metabolic networks and predicting the phenotypic properties of microorganisms (Alper et al., 2005). Sorting of large amounts of biological data can be also done with the use of Boolean logic, making the assumption that gene expression is discrete (de Jong, 2002). Furthermore, dynamic analysis employing a set of ordinary differential equations (ODEs) can be applied to describe biological systems in a mechanistic way, providing information about the kinetics of molecular interactions (Pecou, 2005). However, ODE models are yet more useful for the study of smaller systems mainly due to (i) partially known parameters, (ii) molecular mechanisms not yet described in detail, and (iii) complexity in the analysis of nonlinear differential equations. Moreover, the stochastic kinetics modelling approach considering the stochastic nature of biochemical reactions has been applied to describe the concentration of molecular species within the cell (Hasty et al., 2001).

In the past few years, the effort to build a whole-cell model has made the development of integrative modelling approaches necessary for the analysis of cellular metabolism. Thus, large-scale metabolic (FBA) and regulatory network (Boolean) models have been used as a scaffold with which ODE-based models can be integrated to study detailed models of sub-cellular networks in the context of their global effects (Covert et al., 2008; Lee et al., 2008). In this article we integrate two dynamic models in the regulatory and metabolic cellular level to describe the *m*-xylene degrading behaviour of *P. putida*. We suggest that focusing on the

gene expression profiles of a specific pathway that is important for a bioprocess may provide accurate description of the metabolic dynamics avoiding the development of laborious large-scale models.

As the structure of an increasing number of genetic circuits is elucidated, shedding light on dynamic aspects of key regulatory systems, it will be possible to advance traditional bioprocess modelling into a modern era where accounting for the interaction between genes will be required for the design of optimal bioprocesses. Our study demonstrates how to link the power of molecular biology tools with knowledge of the functional relationships between genetic circuit components to develop advanced mechanistic models accurately predicting cellular functions. Without question, in order to improve the accuracy of models in the field the existing assumptions of unstructured and empirical models, usually limiting their applicability to a narrow range of conditions, should be replaced by quantitative descriptions of the regulatory effects controlling upstream the production of catabolic enzymes. Future studies can use the modelling framework developed here in conjunction with the current progress in molecular biology to decipher more detailed models of the relationship between the dynamics of key circuit components and how their underlying biochemical interactions affect the function of biological systems.

Acknowledgments

This work was supported with the following projects: (a) PROBACTYS (FP6—NEST-PATHFINDER EU call on Synthetic Biology), (b) PSYSMO (BBSRC—ERA-NET programme on the Systems Biology of Microorganisms) and (c) TARPOL (FP7 EU—KBBE Coordination Action for SynBio in Environmental Sciences).

Appendix A. Supplementary materials

Supplementary materials associated with this article can be found in the online version at doi:10.1016/j.ymben.2011.02.001.

References

Alon, U., 2006. An Introduction to Systems Biology: Design Principles of Biological Circuits. CRC Press, Florida.

Alper, H., Jin, Y.S., Moxley, J.F., Stephanopoulos, G., 2005. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. Metab. Eng. 7, 155–164.

Aranda-Olmero, I., Ramos, J.L., Marques, S., 2005. Integration of signals through Crc and PtsN in catabolite repression of *Pseudomonas putida* TOL plasmid pWW0. Appl. Environ. Microbiol. 71, 4191–4198.

Bailey, J.E., 1998. Mathematical modelling and analysis in biochemical engineering: past accomplishments and future opportunities. Biotechnol. Prog. 14, 8–20.

Ballerstedt, H., Volkers, R.J.M., Mars, A.E., Hallsworth, J.E., dos Santos, V.A.M., Puchalka, J., van Duuren, J., Eggink, G., Timmis, K.N., de Bont, J.A.M., Wery, J., 2007. Genomotyping of *Pseudomonas putida* strains using *P. putida* KT2440-based high-density DNA microarrays: implications for transcriptomics studies. Appl. Microbiol. Biotechnol. (75), 1133–1142.

Bertoni, G., Marques, S., de Lorenzo, V., 1998. Activation of the toluene-responsive regulator XylR causes a transcriptional switch between sigma 54 and sigma 70 promoters at the divergent *Pr*/*Ps* region of the TOL plasmid. Mol. Microbiol. 27, 651–659.

Buhler, B., Straathof, A.J.J., Witholt, B., Schmid, A., 2006. Analysis of two-liquid-phase multistep biooxidation based on a process model: indications for biological energy shortage. Org. Process Res. Dev. 10, 628–643.

Carmona, M., de Lorenzo, V., Bertoni, G., 1999. Recruitment of RNA polymerase is a rate-limiting step for the activation of the σ^{54} promoter *Pu* of *Pseudomonas putida*. J. Biol. Chem. 274, 33790–33794.

Cases, I., de Lorenzo, V., Perez-Martin, J., 1996. Involvement of σ^{54} in exponential silencing of the *Pseudomonas putida* TOL plasmid *Pu* promoter. Mol. Microbiol. 19, 7–17.

Corredor, J.E., Wawrik, B., Paul, J.H., Tran, H., Kerkhof, L., Lopez, A., Dieppa, O., 2004. Cardenas, Geochemical rate-RNA integration study:ribulose-1,5-bisphosphate carboxylase/oxygenase gene transcription and photosynthetic capacity of planktonic photoautotrophs. Appl. Environ. Microbiol. 70, 5459–5468.

Covert, M.W., Xiao, N., Chen, T.J., Karr, J.R., 2008. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics 24, 2044–2050.

De Jong, H., 2002. Modeling and simulation of genetic regulatory systems: a literature review. J. Comput. Biol. 9, 67–103.

Devos, D., Garmendia, J., de Lorenzo, V., Valencia, A., 2002. Deciphering the action of aromatic effectors on the prokaryotic enhancer-binding protein XylR: a structural node of its N-terminal domain. Environ. Microbiol. 4, 29–41.

Dominguez-Cuevas, P., Gonzalez-Pastor, J.-E., Marques, S., Ramqos, J.L., de Lorenzo, V., 2006. Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. J. Biol. Chem. 281, 11981–11991.

Dominguez-Cuevas, P., Marin, P., Busby, S., Ramos, J.L., Marques, S., 2008. Roles of effectors in *XylS*-dependent transcription activation: intramolecular domain derepression and DNA binding. J. Bacteriol. 190, 3118–3128.

Dominguez-Cuevas, P., Ramos, J.L., Marques, S., 2008. Sequential *XylS*-CTD binding to the *Pm* promoter induces DNA bending prior to activation. J. Bacteriol. 192, 2682–2690.

Douma, R.D., Verheijen, P.J.T., de Laat, W.T.A.M., Heijnen, J.J., van Gulik, W.M., 2010. Dynamic gene expression regulation model for growth and penicillin production in *Penicillium chrysogenum*. Biotechnol. Bioeng. 106, 608–618.

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

Ewering, C., Heuser, F., Benolken, J.K., Bramer, C.O., Steinbuchel, A., 2006. Metabolic engineering of strains of *Ralstonia eutropha* and *Pseudomonas putida* for biotechnological production of 2-methylcitric acid. Metab. Eng. 8, 587–602.

Feng, X.-J., Hooshangi, S., Chen, D., Li, G., Weiss, R., Rabitz, H., 2004. Optimizing genetic circuits by global sensitivity analysis. Biophys. J. 87, 2195–2202.

Gonzalez-Perez, M.M., Marques, S., Dominguez-Cuevas, P., Ramos, J.L., 2002. *XylS* activator and RNA polymerase binding sites at the Pm promoter overlap. FEBS Lett. 519, 117–122.

Gonzalez-Perez, M.-M., Ramos, J.L., Marques, S., 2004. Cellular *XylS* levels are a function of transcription of *xylS* from two independent promoters and the differential efficiency of translation of the two mRNAs. J. Bacteriol. 186, 1898–1901.

Gunsch, C.K., Kinney, K.A., Szaniszlo, P.J., Whitman, C.P., 2007. Relative gene expression quantification in a fungal gas-phase biofilter. Biotechnol. Bioeng. 98, 101–111.

Han, J.-D.J., 2008. Understanding biological functions through molecular networks. Cell Res. 18, 224–237.

Han, J.-I., Semrau, J.D., 2004. Quantification of gene expression in methanotrophs by competitive reverse transcription-polymerase chain reaction. Environ. Microbiol. 6, 388–399.

Hardin, P.E., 2005. The circadian timekeeping system of *Drosophila*. Curr. Biol. 15, R714–R722.

Hasty, J., McMillen, D., Collins, J.J., 2002. Engineered gene circuits. Nature 420, 224–230.

Hasty, J., McMillen, D., Isaacs, F., Collins, J.J., 2001. Computational studies of gene regulatory networks: *in numero* molecular biology. Nat. Rev. 2, 268–279.

Holtel, A., Timmis, K.N., Ramos, J.L., 1992. Upstream binding sequences of the XylR activator protein and integration host factor in the *xylS* gene promoter region of the *Pseudomonas* TOL plasmid. Nucl. Acids Res. 20, 1755–1762.

Hooshangi, S., Thiberge, S., Weiss, R., 2005. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc. Natl. Acad. Sci. USA 102, 3581–3586.

Isaacs, F.J., Hasty, J., Cantor, C.R., Collins, J.J., 2003. Prediction and measurement of an autoregulatory genetic module. Proc. Natl. Acad. Sci. USA 100, 7714–7719.

Ishihama, A., 1999. Modulation of the nucleoid, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival. Genes Cells, 135–143.

Ishihama, A., 2000. Functional modulation of *Escherichia coli* RNA polymerase. Annu. Rev. Microbiol. 54, 499–518.

Jishage, M., Iwata, A., Ueda, S., Ishihama, A., 1996. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. J. Bacteriol. 178, 5447–5451.

Koutinas, M., Lam, M.-C., Kiparissides, A., Silva-Rocha, R., Godinho, M., Livingston, A.G., Pistikopoulos, E.N., de Lorenzo, V., dos Santos, V.A.P.M., Mantalaris, A., 2010. The regulatory logic of *m*-xylene biodegradation by *Pseudomonas putida* mt-2 exposed by dynamic modelling of the principal node *Ps*/*Pr* of the TOL plasmid. Environ. Microbiol. (12), 1705–1718.

Kovarova-Kovar, K., Egli, T., 1998. Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. Microbiol. Mol. Biol. Rev. 62, 646–666.

Kramer, B.P., Fischer, M., Fussenegger, M., 2005. Semi-synthetic mammalian gene regulatory networks. Metab. Eng. 7, 241–250.

Laub, M.T., Shapiro, L., McAdams, H.H., 2007. Systems biology of *Caulobacter*. Annu. Rev. Genet. 41, 429–441.

Lee, J.M., Gianchandani, E.P., Eddy, J.A., Papin, J.A., 2008. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. PLoS Comput. Biol. 4, e1000086.

Lee, S.B., Bailey, J.E., 1984. Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* operator function. Biotechnol. Bioeng. 26, 1372–1382.

Littlejohns, J.V., Daugulis, A.J., 2008. Kinetics and interactions of BTEX compounds during degradation by a bacterial consortium. Process Biochem. 43, 1068–1076.

Marques, S., Gallegos, M.T., Manzanera, M., Holtel, A., Timmis, K.N., 1998. Activation and repression of transcription at the double tandem divergent promoters for the *xylR* and *xylS* genes of the TOL plasmid of *Pseudomonas putida*. J. Bacteriol. 180, 2889–2894.

Melchiorsen, C.R., Jensen, N.B.S., Christensen, B., Jokumsen, K.V., Villadsen, J., 2001. Dynamics of pyruvate metabolism in *Lactococcus lactis*. Biotechnol. Bioeng. 74, 271–279.

Merrick, M.J., 1993. In a class of its own—the RNA polymerase sigma factor σ^{54} (σ^{N}). Mol. Microbiol. 10, 903–909.

Moreno, R., Fonseca, P., Rojo, F., 2010. The Crc global regulator inhibits the *Pseudomonas putida* pWW0 toluene/xylene assimilation pathway by repressing the translation of regulatory and structural genes. J. Biol. Chem. 285, 24412–24419.

Morgado, J., Merlin, G., Gonthier, Y., Eyraud, A., 2004. A mechanistic model for m-xylene treatment with a peat-bed biofilter. Environ. Technol. 25, 123–132.

Nicolaou, S.A., Gaida, S.M., Papoutsakis, E.T., 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. Metab. Eng. 12, 307–331.

Nogales, J., Palsson, B.O., Thiele, I., 2008. A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: *i*JN746 as a cell factory. BMC Syst. Biol. 2, 79.

Park, J., Lang, J., Thamaraiselvi, K., Kukor, J.J., Abriola, L.M., 2008. Induction kinetics of aerobic toluene degradation as a function of carbon starvation history. Process Biochem. 43, 1345–1351.

Pecou, E., 2005. Splitting the dynamics of large biochemical interaction networks. J. Theor. Biol. 232, 375–384.

Perez-Martin, J., de Lorenzo, V., 1995. The σ^{54}-dependent promoter *Ps* of the TOL plasmid of *Pseudomonas putida* requires HU for transcriptional activation in vivo by XylR. J. Bacteriol. 177, 3758–3763.

Pieper, D.H., dos Santos, V.A.P.M., Golyshin, P.N., 2004. Genomic and mechanistic insights into the biodegradation of organic pollutants. Curr. Opin. Biotechnol. 15, 215–224.

Process Systems Enterprise, gPROMS, 〈www.psenterprise.com/gproms〉, 1997–2010.

Puchalka, J., Oberhardt, M.A., Godinho, M., Bielecka, A., Regenhardt, D., Timmis, K.N., Papin, J.A., dos Santos, V.A.P.M., 2008. Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. PLoS Comput. Biol. (4), e1000210.

Ramos, J.L., Marques, S., Timmis, K.N., 1997. Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. Annu. Rev. Microbiol. 51, 341–372.

Rogers, J.B., Reardon, K.F., 2000. Modeling substrate interactions during the biodegradation of mixtures of toluene and phenol by *Burkholderia* species JS150. Biotechnol. Bioeng. 70, 428–435.

Sambrook, J., Fritsch, E.F., Maniatis, E., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Press, New York.

Shingler, V., 2003. Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. Environ. Microbiol. 5, 1226–1241.

Sikkema, J., de Bont, J.A.M., Poolman, B., 1995. Mechanisms of membrane toxicity of hydrocarbons. Microbiol. Rev. 59, 201–222.

Stricker, J., Cookson, S., Bennett, M.R., Mather, W.H., Tsimring, L.S., Hasty, J., 2008. A fast, robust and tunable synthetic gene oscillator. Nature 456, 516–519.

Timmis, K.N., 2002. *Pseudomonas putida*: a cosmopolitan opportunist *par* excellence. Environ. Microbiol. 4, 779–781.

Weiss, R., Basu, S., Hooshangi, S., Kalmbach, A., Karig, D., Mehreja, R., Netravali, I., 2003. Genetic circuit building blocks for cellular computation, communications, and signal processing. Nat. Comput. 2, 47–84.

Yano, T., Koga, S., 1969. Dynamic behaviour of chemostat subject to substrate inhibition. Biotechnol. Bioeng. 11, 139–153.

Valls, M., Buckle, M., de Lorenzo, V., 2002. *In vivo* UV laser footprinting of the *Pseudomonas putida* σ^{54} *Pu* promoter reveals that integration host factor couples transcriptional activity to growth phase. J. Biol. Chem. 277, 2169–2175.

Van Dien, S.J., de Lorenzo, V., 2003. Deciphering environmental signal integration in σ^{54}-dependent promoters with a simple mathematical model. J. Theor. Biol. 224, 437–449.

Velazquez, F., Fernandez, S., de Lorenzo, V., 2006. The upstream-activating sequences of the σ^{54} promoter *Pu* of *Pseudomonas putida* filter transcription read through from upstream genes. J. Biol. Chem. 281, 11940–11948.

Velazquez, F., Parro, V., de Lorenzo, V., 2005. Inferring the genetic network of *m*-xylene metabolism through expression profiling of the *xyl* genes of *Pseudomonas putida* mt-2. Mol. Macrobiol. 57, 1557–1569.