Applying asymptotic methods to synthetic biology: modelling the reaction kinetics of the mevalonate pathway

Mohit P. Dalwadi[a,*], Marco Garavaglia[a], Joseph P. Webb[b], John R. King[a,c], Nigel P. Minton[a]

[a] *Synthetic Biology Research Centre, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*
[b] *Department of Molecular Biology and Biotechnology, University of Sheffield, Western Bank, Sheffield S10 2TN, UK*
[c] *School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*

Abstract

The mevalonate pathway is normally found in eukaryotes, and allows for the production of isoprenoids, a useful class of organic compounds. This pathway has been successfully introduced to *Escherichia coli*, enabling a biosynthetic production route for many isoprenoids. In this paper, we develop and solve a mathematical model for the concentration of metabolites in the mevalonate pathway over time, accounting for the loss of acetyl-CoA to other metabolic pathways. Additionally, we successfully test our theoretical predictions experimentally by introducing part of the pathway into *Cupriavidus necator*. In our model, we exploit the natural separation of time scales as well as of metabolite concentrations to make significant asymptotic progress in understanding the system. We confirm that our asymptotic results agree well with numerical simulations, the former enabling us to predict the most important reactions to increase isopentenyl diphosphate production whilst minimizing the levels of HMG-CoA, which inhibits cell growth. Thus, our mathematical model allows us to recommend the upregulation of certain combinations of enzymes to improve production through the mevalonate pathway.

Keywords: asymptotic analysis, metabolic pathways, isoprenoid production

*Corresponding author. *E-mail address:* mohit.dalwadi@nottingham.ac.uk

1. Introduction

Isoprenoids are a diverse class of naturally occurring organic chemicals found in all organisms. In plants, isoprenoids are the cause of many aromas and, in animals, isoprenoids form steroids and sterols. The wide range of isoprenoid products is one reason why the successful introduction of a viable isoprenoid pathway to *Escherichia coli* by [1] (in this case to produce amorpha-4,11-diene, a precursor to the antimalarial compound artemisinin) was a major breakthrough in synthetic biology. Since then, a significant amount of experimental work has been carried out to improve the yield from this pathway (see, for example, [2, 3, 4, 5, 6]).

There are two main pathways from pyruvate to isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP), and these two products can react to make isoprenoid compounds. IDP and DMADP are essentially interchangeable due to the enzyme isopentenyl diphosphate isomerase that allows conversion between the two. The first pathway is known as the mevalonate pathway, and starts from acetyl coenzyme A (acetyl-CoA), mainly derived from pyruvate, which is converted to IDP via the key pathway intermediate mevalonate. The second pathway is known as the non-mevalonate or, alternatively, the 2-C-methyl-*D*-erythritol 4-phosphate/1-deoxy-*D*-xylulose 5-phosphate (MEP/DOXP) pathway, and also converts pyruvate to IDP. The mevalonate pathway was the first to be discovered, and occurs naturally in eukaryotes. The non-mevalonate pathway mainly occurs in bacteria (with some exceptions), and some plants.

The reason for introducing the mevalonate pathway to *Escherichia coli*, a bacterium that naturally expresses only the non-mevalonate pathway, is to bypass the natural negative feedback mechanisms in place that would ordinarily prevent the overproduction of isoprenoids. We are interested in mathematically modelling this mevalonate pathway, with the goal of understanding how to further modify the pathway by, for example, upregulating genes that control certain enzymes, in order to produce more IDP. The mevalonate pathway we will model comprises the following:

$$\text{Pyruvate} \xrightarrow{k_1} \text{Acetyl-CoA}, \tag{1a}$$

$$\text{Acetyl-CoA} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \text{Acetoacetyl-CoA}, \tag{1b}$$

$$\text{Acetyl-CoA} \xrightarrow{A} \phi, \tag{1c}$$

$$\text{Acetoacetyl-CoA} + \text{Acetyl-CoA} \xrightarrow{k_3} \text{HMG-CoA}, \tag{1d}$$

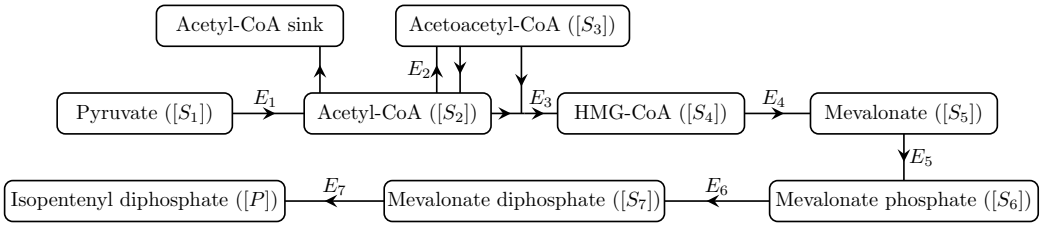

Figure 1: A schematic network diagram for the pathway we consider in this paper, where arrows denote the direction of the reactions. We only track the metabolites included in this Figure and, specifically, not any involved in the acetyl-CoA sink. Where we write E_i (for $i = 1,\dots,7$) next to a reaction arrow, this denotes a specific enzyme that controls the reaction. Hence, E_1 corresponds to the pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12, and EC 1.8.1.4), E_2 corresponds to acetyl-CoA acetyltransferase (EC 2.3.1.9), E_3 corresponds to HMG-CoA synthase (EC 2.3.3.10), E_4 corresponds to HMG-CoA reductase (EC 1.1.1.34), E_5 corresponds to mevalonate kinase (EC 2.7.1.36), E_6 corresponds to phosphomevalonate kinase (EC 2.7.4.2), and E_7 corresponds to mevalonate diphosphate decarboxylase (EC 4.1.1.33).

$$\text{HMG-CoA} \xrightarrow{k_4} \text{Mevalonate}, \tag{1e}$$

$$\text{Mevalonate} \xrightarrow{k_5} \text{Mevalonate phosphate}, \tag{1f}$$

$$\text{Mevalonate phosphate} \xrightarrow{k_6} \text{Mevalonate diphosphate}, \tag{1g}$$

$$\text{Mevalonate diphosphate} \xrightarrow{k_7} \text{Isopentenyl diphosphate}, \tag{1h}$$

where (1c) represents the loss of acetyl-CoA to other metabolic pathways, such as the citric acid cycle or any pathways directly involved with fatty acid biosynthesis. Moreover, we will not keep track of any metabolites associated with this acetyl-CoA sink. We show a schematic of the pathway in Figure 1.

In general, the aim of our model is to determine which reactions are the most important for IDP production without resorting to expensive and time-consuming experiments. We are also interested in determining which reactions have the most significant control over the levels of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA), linked to the inhibition of cell growth due to its inhibition of fatty acid biosynthesis [5]. Specifically, we will be interested in maintaining low levels of HMG-CoA whilst increasing IDP production.

We make several key modelling assumptions to facilitate analysis of our system. Firstly, we assume that the formation rate of enzyme complex is much quicker than the rate of substrate consumption, and thus the reaction rates are governed by Michaelis–Menten-type laws, the specific form of which we obtain from the literature. We also consider a system that is well mixed and thus spa-

tially independent. Additionally, we consider the case where pyruvate is instantaneously introduced to a system containing all of the relevant enzymes, but none of the intermediate metabolites, allowing for a cleaner mathematical analysis. We first investigate the case where pyruvate is continuously replenished and held at a constant concentration, then the case where pyruvate is never replenished. We show that the second case shares many similarities with the first until the pyruvate is depleted to a certain critical level, which we determine. Understanding these extreme cases allows us to determine the key reactions in this pathway, and to suggest targets for upregulation.

As is the case with many biological systems, there are many parameters in the system. Thus, a comprehensive understanding of the system using a purely experimental approach would be very time consuming. This reasoning also applies to investigating our mathematical model using a fully numerical approach, although the time taken to investigate the system would be shorter than the purely experimental approach. To get around this issue, we supplement and guide our numerical simulations by determining asymptotic approximations (see, for example, [7, 8, 9]) of the metabolite concentrations. This will enhance our physical insight into the underlying system and allow us to determine how the concentrations vary as functions of the experimental parameters. Moreover, this approach allows us to bypass the issue we have with the uncertainty in the parameters, as this method only requires an idea of the order of magnitude of each parameter.

Finally, to test our theoretical predictions, we introduce part of the mevalonate pathway (from acetyl-CoA to mevalonate) into *Cupriavidus necator* by transforming it with a plasmid harbouring the *mvaE* and *mvaS* genes from *Enterococcus faecalis* under the control of the P_{BAD} *L*-arabinose inducible promoter. The *mvaE* and *mvaS* genes code for the enzymes responsible for the conversion of acetyl-CoA to mevalonate (in the first half of the mevalonate pathway). We show that mevalonate can be produced by our bacterial chassis, and confirm that our experimental results are successfully predicted by our model.

The outline of this paper is as follows. We introduce a mathematical model to describe the nonlinear reaction kinetics in §2. We solve this system in §3, where we give both numerical and asymptotic solutions to describe the system behaviour. In that section, we consider the continuous replenishment of pyruvate case first, then we consider the no replenishment of pyruvate case. We discuss the experimental procedure and results in §4, and compare these results to our model predictions. We finish by discussing our results and comparing the two regimes in §5.

2. Model description

The dimensional system we consider is derived from Michaelis–Menten-type laws found in the literature, where each variable is defined in Table 1. The forms of each reaction rate are obtained from the corresponding reference in Table 2. The dimensional system is as follows

$$\frac{\mathrm{d}[S_1]}{\mathrm{d}\tau} = -\frac{k_1 E_1 K_1^i [S_1]}{K_1^i \left([S_1] + K_1^M\right) + [S_1][S_2]}, \tag{2a}$$

$$\frac{\mathrm{d}[S_2]}{\mathrm{d}\tau} = \frac{k_1 E_1 K_1^i [S_1]}{K_1^i \left([S_1] + K_1^M\right) + [S_1][S_2]} - \frac{k_2 E_2 [S_2]}{[S_2] + K_2^M} + \frac{k_{-2} E_2 [S_3]}{[S_3] + K_{-2}^M} - A[S_2], \tag{2b}$$

$$\frac{\mathrm{d}[S_3]}{\mathrm{d}\tau} = \frac{k_2 E_2 [S_2]}{[S_2] + K_2^M} - \frac{k_{-2} E_2 [S_3]}{[S_3] + K_{-2}^M} - \frac{k_3 E_3 K_3^i [S_2][S_3]}{K_3^i [S_2][S_3] + K_{3,a}^M [S_3]\left([S_3] + K_3^i\right) + K_3^i K_{3,b}^M [S_2]}, \tag{2c}$$

$$\frac{\mathrm{d}[S_4]}{\mathrm{d}\tau} = \frac{k_3 E_3 K_3^i [S_2][S_3]}{K_3^i [S_2][S_3] + K_{3,a}^M [S_3]\left([S_3] + K_3^i\right) + K_3^i K_{3,b}^M [S_2]} - \frac{k_4 E_4 [S_4]}{[S_4] + K_4^M}, \tag{2d}$$

$$\frac{\mathrm{d}[S_5]}{\mathrm{d}\tau} = \frac{k_4 E_4 [S_4]}{[S_4] + K_4^M} - \frac{k_5 E_5 [S_5]}{[S_5] + K_5^M}, \tag{2e}$$

$$\frac{\mathrm{d}[S_6]}{\mathrm{d}\tau} = \frac{k_5 E_5 [S_5]}{[S_5] + K_5^M} - \frac{k_6 E_6 [S_6]}{[S_6] + K_6^M}, \tag{2f}$$

$$\frac{\mathrm{d}[S_7]}{\mathrm{d}\tau} = \frac{k_6 [S_6]}{[S_6] + K_6^M} - \frac{k_7 E_7 [S_7]}{[S_7] + K_7^M}, \tag{2g}$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}\tau} = \frac{k_7 E_7 [S_7]}{[S_7] + K_7^M}. \tag{2h}$$

Most of the terms in the governing equations (2) are standard Michaelis–Menten reaction velocities, for which we obtain the relevant kinetic parameters from the references indicated in Table 2. The two modified Michaelis–Menten terms we include are for reactions (1a) and (1d), which describe different types of inhibition. In Kresze and Ronft [10], it is shown that acetyl-CoA has an inhibitory effect on (1a) which is uncompetitive with pyruvate. Therefore, for this reaction velocity we use the standard form for uncompetitive inhibition [11]. In Middleton [12], it is shown that acetoacetyl-CoA has an inhibitory effect on (1d) which is competitive with acetyl-CoA, and we therefore take this reaction velocity to have the standard form for competitive inhibition, as stated in Middleton [12] and Yung-Chi and Prusoff [11]. The parameter A represents the total loss of acetyl-CoA to other

Original variable	Description	Nondimensionalisation
$[S_1]$	Pyruvate	$[S_1] = S_0 S_1$
$[S_2]$	Acetyl CoA	$[S_2] = S_0 S_2$
$[S_3]$	Acetoacetyl-CoA	$[S_3] = S_0 S_3$
$[S_4]$	HMG-CoA	$[S_4] = S_0 S_4$
$[S_5]$	Mevalonate	$[S_5] = S_0 S_5$
$[S_6]$	Mevalonate phosphate	$[S_6] = S_0 S_6$
$[S_7]$	Mevalonate diphosphate	$[S_7] = S_0 S_7$
$[P]$	Isopentenyl diphosphate	$[P] = S_0 P$
τ	Time	$\tau = (S_0/k_1 E_1)t$

Table 1: Dimensional and dimensionless variable definitions.

metabolic pathways, for example, the citric acid cycle or fatty acid biosynthesis, and we assume that this occurs with first-order kinetics. While this parameter is difficult to measure experimentally, we will bypass this issue by considering a distinguished limit in the dimensionless system when we perform an asymptotic analysis. The parameters E_i, where $i = 1, \ldots, 7$, denote the enzyme concentrations for the reactions they control.

We use initial conditions that correspond to the scenario where pyruvate is instantaneously introduced to a system containing all of the relevant enzymes, but none of the intermediate metabolites. That is, we use $[S_1](0) = S_0$, $[S_2](0) = [S_3](0) = [S_4](0) = [S_5](0) = [S_6](0) = [S_7](0) = [P](0) = 0$. Here, S_0 represents the initial or typical level of pyruvate present in the system. As any given metabolite may already be present in the real-world system, our approach to the initial conditions is a modelling choice. That is, we choose to reduce the number of uncertain parameters in the system in order to facilitate a more simplified analysis of the system.

To nondimensionalize the system variables, we scale each dimensional metabolite concentration with S_0, the initial concentration of pyruvate. Additionally, we scale time with $S_0/(k_1 E_1)$, the characteristic time of the first reaction, which occurs between pyruvate and acetyl-CoA. We summarise these scalings in Table 1. To form dimensionless parameters, we first note that estimates of the kinetic parameters can vary significantly in different environments (Table 2). Given the uncertainty in the parameters, we seek to understand how the system behaves for different values of these parameters; we seek asymptotic solutions in terms of the system parameters, allowing us to explicitly determine how a variation in parameter values affects the system. We can explore how the system behaves as these

Dimensional	Organism	Range
$k_1 = 10\,\mathrm{s}^{-1}$	*Saccharomyces cerevisiae* [10, 13]	4 – $30\,\mathrm{s}^{-1}$ [10, 13, 14]
$k_2 = 200\,\mathrm{s}^{-1}$	*Enterococcus faecalis* [15]	10 – $260\,\mathrm{s}^{-1}$ [15, 16, 17]
$k_{-2} = 3000\,\mathrm{s}^{-1}$	*Enterococcus faecalis* [15]	80 – $3600\,\mathrm{s}^{-1}$ [15, 18]
$k_3 = 6\,\mathrm{s}^{-1}$	*Saccharomyces cerevisiae* [12]	0.5 – $14\,\mathrm{s}^{-1}$ [12, 19, 20]
$k_4 = 10\,\mathrm{s}^{-1}$	*Enterococcus faecalis* [15]	1 – $20\,\mathrm{s}^{-1}$ [15, 21, 22]
$k_5 = 20\,\mathrm{s}^{-1}$	*Methanosarcina mazei* [23]	4 – $40\,\mathrm{s}^{-1}$ [23, 24]
$k_6 = 4\,\mathrm{s}^{-1}$	*Saccharomyces cerevisiae* [25]	2 – $6\,\mathrm{s}^{-1}$ [25, 26]
$k_7 = 1\,\mathrm{s}^{-1}$	*Saccharomyces cerevisiae* [27]	0.1 – $5.5\,\mathrm{s}^{-1}$ [27, 28]
$K_1^M = 0.65\,\mathrm{mM}$	*Saccharomyces cerevisiae* [10]	0.13 – $1\,\mathrm{mM}$ [10, 29, 14]
$K_1^i = 0.014\,\mathrm{mM}$	*Saccharomyces cerevisiae* [10]	0.014 – $0.018\,\mathrm{mM}$ [10, 29]
$K_2^M = 1\,\mathrm{mM}$	*Enterococcus faecalis* [15]	0.06 – $1.2\,\mathrm{mM}$ [15, 16, 17]
$K_{-2}^M = 0.01\,\mathrm{mM}$	*Enterococcus faecalis* [15]	0.01 – $0.09\,\mathrm{mM}$ [15, 18]
$K_{3,a}^M = 0.015\,\mathrm{mM}$	*Saccharomyces cerevisiae* [12]	0.01 – $0.04\,\mathrm{mM}$ [12, 19, 20]
$K_{3,b}^M = 0.003\,\mathrm{mM}$	*Saccharomyces cerevisiae* [12]	0.0001 – $0.01\,\mathrm{mM}$ [12, 20]
$K_3^i = 0.01\,\mathrm{mM}$	*Saccharomyces cerevisiae* [12]	0.008 – $0.02\,\mathrm{mM}$ [12, 20]
$K_4^M = 0.02\,\mathrm{mM}$	*Enterococcus faecalis* [15]	0.015 – $0.065\,\mathrm{mM}$ [15, 21, 22]
$K_5^M = 0.1\,\mathrm{mM}$	*Methanosarcina mazei* [23]	0.06 – $0.24\,\mathrm{mM}$ [23, 24]
$K_6^M = 0.9\,\mathrm{mM}$	*Saccharomyces cerevisiae* [25]	0.004 – $0.9\,\mathrm{mM}$ [25, 26]
$K_7^M = 0.2\,\mathrm{mM}$	*Saccharomyces cerevisiae* [27]	0.03 – $0.9\,\mathrm{mM}$ [27, 28]
$A \quad [\mathrm{s}^{-1}]$		
Dimensionless parameters		
$\bar{k}_2 = \varepsilon k_2 E_2/k_1 E_1 = 0.2$	$\bar{K}_1^M = K_1^M/S_0 = 0.65$	$\bar{K}_4^M = K_4^M/\varepsilon S_0 = 2$
$\bar{k}_{-2} = \varepsilon k_{-2} E_2/k_1 E_1 = 3$	$\bar{K}_1^i = K_1^i/\varepsilon S_0 = 1.4$	$\bar{K}_5^M = K_5^M/S_0 = 0.1$
$\bar{k}_3 = k_3 E_3/k_1 E_1 = 0.6$	$\bar{K}_2^M = K_2^M/S_0 = 1$	$\bar{K}_6^M = K_6^M/S_0 = 0.9$
$\bar{k}_4 = k_4 E_4/k_1 E_1 = 1$	$\bar{K}_{-2}^M = K_{-2}^M/\varepsilon S_0 = 1$	$\bar{K}_7^M = K_7^M/S_0 = 0.2$
$\bar{k}_5 = k_5 E_5/k_1 E_1 = 2$	$\bar{K}_{3,a}^M = K_{3,a}^M/\varepsilon S_0 = 1.5$	$\bar{A} = A S_0/k_1 E_1 = 1$
$\bar{k}_6 = k_6 E_6/k_1 E_1 = 0.4$	$\bar{K}_{3,b}^M = K_{3,b}^M/\varepsilon S_0 = 0.3$	
$\bar{k}_7 = k_7 E_7/k_1 E_1 = 0.1$	$\bar{K}_3^i = K_3^i/\varepsilon S_0 = 1$	

Table 2: Parameters. We use the value $S_0 = 1\,\mathrm{mM}$, and assume that $E_i = E_j$ for $i, j = 1, \ldots, 7$. Different values of E_j can be considered by varying the appropriate dimensionless parameter. We introduce the small dimensionless parameter $\varepsilon = 0.01$, to formally account for the large difference in magnitude between parameters, and choose $\bar{A} = 1$ in the simulations, as there is a distinguished asymptotic limit when $\bar{A} = O(1)$.

parameters vary within an order of magnitude by first scaling each rate constant with the rate constant of the first reaction, each Michaelis constant with the initial pyruvate concentration, and each enzyme concentration with concentration of the first enzyme. Then, we use the typical dimensional values in Table 2 to introduce an artificial small dimensionless parameter $\varepsilon = 0.01$ into the system, and write each dimensionless parameter as $c\varepsilon^j$, where c is an $O(1)$ parameter (between 0.1 and 10), and j is an integer. The resultant dimensionless parameters in our system are given in Table 2, and this approach allows us to interrogate the system using an asymptotic analysis (see, for example, [7, 8, 9]). Although, as is always the case with an asymptotic analysis, there may theoretically be an issue in equating terms with the same powers of ε when extreme $O(1)$ parameters are multiplied together, we will show that our asymptotic and numerical results show excellent agreement, and thus the approach is reliable for this system.

Therefore, we nondimensionalize using the dimensionless variables defined in Table 1 and, using the dimensionless kinetic parameters defined in Table 2, we obtain the dimensionless system

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = -\frac{\varepsilon\bar{K}_1^i S_1}{\varepsilon\bar{K}_1^i\left(S_1+\bar{K}_1^M\right)+S_1S_2}, \tag{3a}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = \frac{\varepsilon\bar{K}_1^i S_1}{\varepsilon\bar{K}_1^i\left(S_1+\bar{K}_1^M\right)+S_1S_2} - \frac{\bar{k}_2 S_2}{\varepsilon\left(S_2+\bar{K}_2^M\right)} + \frac{\bar{k}_{-2}S_3}{\varepsilon\left(S_3+\varepsilon\bar{K}_{-2}^M\right)} - \bar{A}S_2, \tag{3b}$$

$$\begin{aligned}\frac{\mathrm{d}S_3}{\mathrm{d}t} = {} & \frac{\bar{k}_2 S_2}{\varepsilon\left(S_2+\bar{K}_2^M\right)} - \frac{\bar{k}_{-2}S_3}{\varepsilon\left(S_3+\varepsilon\bar{K}_{-2}^M\right)} \\ & - \frac{\bar{k}_3\bar{K}_3^i S_2S_3}{\bar{K}_3^i S_2S_3 + \bar{K}_{3,a}^M S_3\left(S_3+\varepsilon\bar{K}_3^i\right)+\varepsilon\bar{K}_3^i\bar{K}_{3,b}^M S_2},\end{aligned} \tag{3c}$$

$$\frac{\mathrm{d}S_4}{\mathrm{d}t} = \frac{\bar{k}_3\bar{K}_3^i S_2S_3}{\bar{K}_3^i S_2S_3 + \bar{K}_{3,a}^M S_3\left(S_3+\varepsilon\bar{K}_3^i\right)+\varepsilon\bar{K}_3^i\bar{K}_{3,b}^M S_2} - \frac{\bar{k}_4 S_4}{S_4+\varepsilon\bar{K}_4^M}, \tag{3d}$$

$$\frac{\mathrm{d}S_5}{\mathrm{d}t} = \frac{\bar{k}_4 S_4}{S_4+\varepsilon\bar{K}_4^M} - \frac{\bar{k}_5 S_5}{S_5+\bar{K}_5^M}, \tag{3e}$$

$$\frac{\mathrm{d}S_6}{\mathrm{d}t} = \frac{\bar{k}_5 S_5}{S_5+\bar{K}_5^M} - \frac{\bar{k}_6 S_6}{S_6+\bar{K}_6^M}, \tag{3f}$$

$$\frac{\mathrm{d}S_7}{\mathrm{d}t} = \frac{\bar{k}_6 S_6}{S_6+\bar{K}_6^M} - \frac{\bar{k}_7 S_7}{S_7+\bar{K}_7^M}, \tag{3g}$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{\bar{k}_7 S_7}{S_7 + \bar{K}_7^M}. \tag{3h}$$

In forming the dimensionless variables, we have used the physically plausible value of $S_0 = 1\,\mathrm{mM}$. Finally, the initial conditions become $S_1(0) = 1$, $S_2(0) = S_3(0) = S_4(0) = S_5(0) = S_6(0) = S_7(0) = P(0) = 0$.

We solve this system for two cases. The first case is where (3a) does not hold and we instead impose $S_1(t) \equiv 1$. This corresponds to the scenario where pyruvate is continuously replenished and held at a constant value. The second case is where (3a) does hold, and we will show that the two systems are equivalent for $t = O(1)$. In both cases, we are interested in determining how to maximise the production of IDP whilst minimizing the levels of HMG-CoA, linked to the inhibition of cell growth.

3. Solutions

3.1. Numerical results

We solve the system presented in §2 numerically, using ode15s in MATLAB with a relative tolerance of 10^{-14}. We use the parameter values given in Table 2 for the continuous replenishment of pyruvate (Figure 2) and the no replenishment of pyruvate (Figure 3) cases. In each case, we are also able to model the over-expression of an enzyme by increasing the dimensionless turnover numbers (the parameters denoted by a lower-case k with a subscript) given in Table 2.

Although the dynamics for each case are different, we can observe general trends. Most notably, we see that increasing E_3 increases both the maximum levels of HMG-CoA and of IDP production, whereas increasing E_4 decreases the maximum levels of HMG-CoA but has a negligible affect on IDP production. Comparing the continuously and never replenished pyruvate cases, we see that the initial dynamics appear to be similar between cases for the same parameter values (until $t \approx 2$ for HMG-CoA and $t \approx 12$ for IDP), but the dynamics diverge after a longer time. In the continuous replenishment case, the concentration of HMG-CoA and the production rate of IDP tends to a constant value whereas, in the no replenishment case, the concentration of HMG-CoA increases to a maximum level before decreasing, and the concentration of IDP tends to a constant value. To understand these phenomena in more detail, and to determine how we can increase IDP production whilst minimizing the maximum levels of (cell growth inhibiting) HMG-CoA in terms of the parameter values, we now perform an asymptotic analysis.

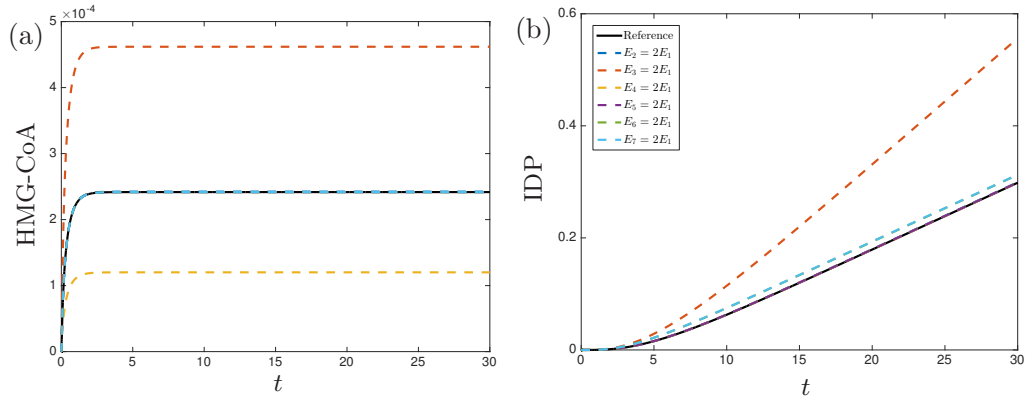

Figure 2: The numerically determined concentrations of (a) HMG-CoA and (b) IDP in the continuous replenishment of pyruvate case. The solid black lines denote the solutions using the reference parameter values given in Table 2 with $E_i = E_j$ for $i \neq j$, and the dashed lines denote the solutions when a particular enzyme is over-expressed. The solutions when E_2, E_5, E_6, or E_7 are doubled are near identical to the reference concentration in (a). The solutions when E_2, E_4, or E_5 are doubled are near identical to the reference concentration in (b), and the solutions when E_6 or E_7 are doubled are near identical to each other.

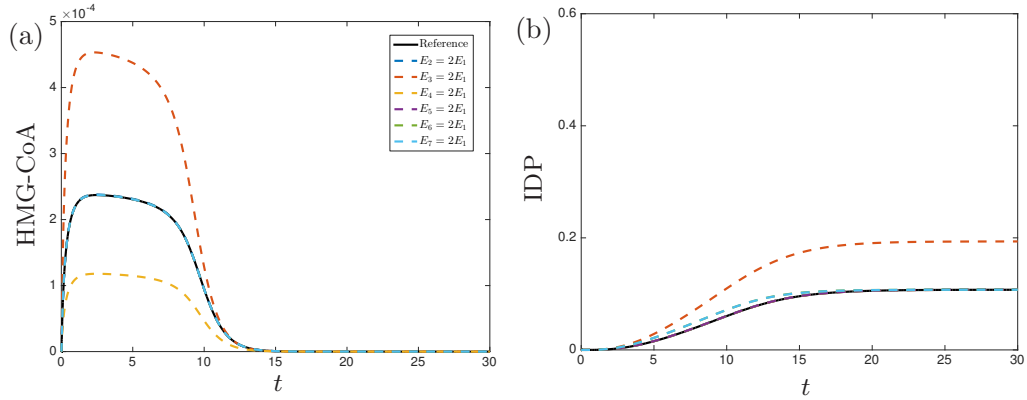

Figure 3: The numerically determined concentrations of (a) HMG-CoA and (b) IDP in the no replenishment of pyruvate case. The solid black lines denote the solutions using the parameter values given in Table 2 with $E_i = E_j$ for $i \neq j$, and the dashed lines denote the solutions when a particular enzyme is over-expressed. The solutions when E_2, E_5, E_6, or E_7 are doubled are near identical to the reference concentration in (a). The solutions when E_2, E_4, or E_5 are doubled are near identical to the reference concentration in (b), and the solutions when E_6 or E_7 are doubled are near identical to each other. We see that the solutions in this no replenishment of pyruvate case are very similar to the solutions in the continuous replenishment of pyruvate case until $t \approx 2$ for HMG-CoA and until $t \approx 12$ for IDP.

3.2. Asymptotic results

3.2.1. Continuous replenishment of pyruvate

Here, we consider the system (3b–h), and impose $\bar{S}_1(t) \equiv 1$. As the sink reaction (1c) is an amalgamation of all sinks of acetyl-CoA (S_2), it is difficult to obtain accurate estimates of $\bar{A}$. We proceed by assuming the distinguished limit $\bar{A} = O(1)$, and we discuss the further limit $\bar{A} = O(1/\varepsilon)$ in Appendix A. In pursuing an asymptotic analysis (see, for example, [7, 8, 9]) for small ε, we make the following scalings: $(S_2, S_5, S_6, S_7, P) = \varepsilon^{1/2}(\bar{S}_2, \bar{S}_5, \bar{S}_6, \bar{S}_7, \bar{P})$, $(S_3, S_4) = \varepsilon^{3/2}(\bar{S}_3, \bar{S}_4)$ and we obtain the $t = O(1)$ governing equations

$$\varepsilon \frac{\mathrm{d}\bar{S}_2}{\mathrm{d}t} = \frac{\varepsilon \bar{K}_1^i \bar{S}_1}{\varepsilon^{1/2}\bar{K}_1^i\left(\bar{S}_1 + \bar{K}_1^M\right) + \bar{S}_1\bar{S}_2} - \frac{\bar{k}_2\bar{S}_2}{\varepsilon^{1/2}\bar{S}_2 + \bar{K}_2^M} + \frac{\bar{k}_{-2}\bar{S}_3}{\varepsilon^{1/2}\bar{S}_3 + \bar{K}_{-2}^M} - \varepsilon\bar{A}\bar{S}_2, \tag{4a}$$

$$\varepsilon^2 \frac{\mathrm{d}\bar{S}_3}{\mathrm{d}t} = \frac{\bar{k}_2\bar{S}_2}{\varepsilon^{1/2}\bar{S}_2 + \bar{K}_2^M} - \frac{\bar{k}_{-2}\bar{S}_3}{\varepsilon^{1/2}\bar{S}_3 + \bar{K}_{-2}^M} - \frac{\varepsilon\bar{k}_3\bar{K}_3^i\bar{S}_2\bar{S}_3}{\varepsilon^{1/2}\bar{K}_3^i\bar{S}_2\bar{S}_3 + \varepsilon\bar{K}_{3,a}^M\bar{S}_3\left(\varepsilon^{1/2}\bar{S}_3 + \bar{K}_3^i\right) + \bar{K}_3^i\bar{K}_{3,b}^M\bar{S}_2}, \tag{4b}$$

$$\varepsilon \frac{\mathrm{d}\bar{S}_4}{\mathrm{d}t} = \frac{\bar{k}_3\bar{K}_3^i\bar{S}_2\bar{S}_3}{\varepsilon^{1/2}\bar{K}_3^i\bar{S}_2\bar{S}_3 + \varepsilon\bar{K}_{3,a}^M\bar{S}_3\left(\varepsilon^{1/2}\bar{S}_3 + \bar{K}_3^i\right) + \bar{K}_3^i\bar{K}_{3,b}^M\bar{S}_2} - \frac{\bar{k}_4\bar{S}_4}{\varepsilon^{1/2}\bar{S}_4 + \bar{K}_4^M}, \tag{4c}$$

$$\frac{\mathrm{d}\bar{S}_5}{\mathrm{d}t} = \frac{\bar{k}_4\bar{S}_4}{\varepsilon^{1/2}\bar{S}_4 + \bar{K}_4^M} - \frac{\bar{k}_5\bar{S}_5}{\varepsilon^{1/2}\bar{S}_5 + \bar{K}_5^M}, \tag{4d}$$

$$\frac{\mathrm{d}\bar{S}_6}{\mathrm{d}t} = \frac{\bar{k}_5\bar{S}_5}{\varepsilon^{1/2}\bar{S}_5 + \bar{K}_5^M} - \frac{\bar{k}_6\bar{S}_6}{\varepsilon^{1/2}\bar{S}_6 + \bar{K}_6^M}, \tag{4e}$$

$$\frac{\mathrm{d}\bar{S}_7}{\mathrm{d}t} = \frac{\bar{k}_6\bar{S}_6}{\varepsilon^{1/2}\bar{S}_6 + \bar{K}_6^M} - \frac{\bar{k}_7\bar{S}_7}{\varepsilon^{1/2}\bar{S}_7 + \bar{K}_7^M}, \tag{4f}$$

$$\frac{\mathrm{d}\bar{P}}{\mathrm{d}t} = \frac{\bar{k}_7\bar{S}_7}{\varepsilon^{1/2}\bar{S}_7 + \bar{K}_7^M}. \tag{4g}$$

The leading-order version of (4) is given by

$$\frac{\mathrm{d}\bar{S}_2}{\mathrm{d}t} = \frac{\bar{K}_1^i}{\bar{S}_2} - v_3\bar{S}_3 - \bar{A}\bar{S}_2, \quad 0 = v_2\bar{S}_2 - v_{-2}\bar{S}_3, \quad 0 = v_3\bar{S}_3 - v_4\bar{S}_4,$$
$$\frac{\mathrm{d}\bar{S}_5}{\mathrm{d}t} = v_4\bar{S}_4 - v_5\bar{S}_5, \quad \frac{\mathrm{d}\bar{S}_6}{\mathrm{d}t} = v_5\bar{S}_5 - v_6\bar{S}_6, \quad \frac{\mathrm{d}\bar{S}_7}{\mathrm{d}t} = v_6\bar{S}_6 - v_7\bar{S}_7, \quad \frac{\mathrm{d}\bar{P}}{\mathrm{d}t} = v_7\bar{S}_7, \tag{5}$$

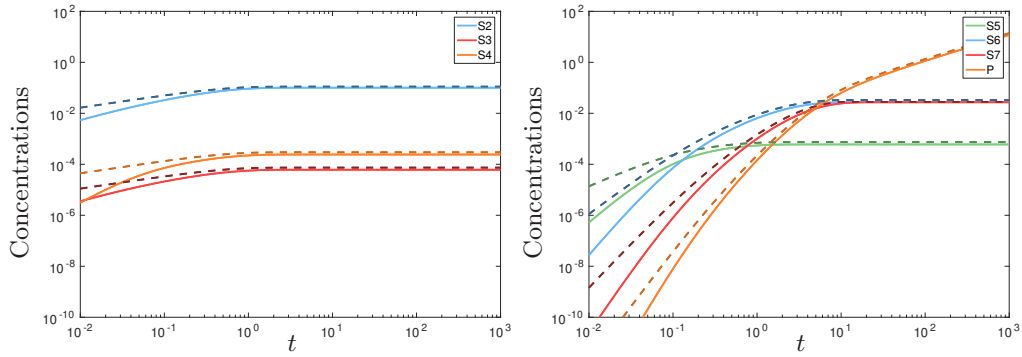

Figure 4: The numerical and asymptotic solutions for the metabolite concentrations in the continuous replenishment of pyruvate case. The solid light lines denote the numerical solutions, and the dashed darker lines denote the asymptotic solutions given in (6). We see good agreement between the numerical and asymptotic solutions for $t = O(1)$, and the system attains its steady state solution in this region.

where $v_j = k_j/K_j^M$ for $j \in \{2,-2,4,5,6,7\}$, and $v_3 = \bar{k}_3/\bar{K}_{3,b}^M$, and (5) is solved by

$$\bar{S}_2 = \left(\frac{\bar{K}_1^i\left(1-e^{-2\alpha t}\right)}{\alpha}\right)^{1/2}, \quad \bar{S}_3 = \frac{v_2}{v_{-2}}\bar{S}_2, \quad \bar{S}_4 = \frac{v_2 v_3}{v_{-2}v_4}\bar{S}_2,$$

$$\bar{S}_5 = \frac{v_2 v_3}{v_{-2}}\int_0^t e^{v_5(s-t)}\bar{S}_2(s)\,\mathrm{d}s, \quad \bar{S}_6 = \frac{v_2 v_3 v_5}{v_{-2}}\int_0^t \frac{e^{v_5(s-t)}-e^{v_6(s-t)}}{v_6-v_5}\bar{S}_2(s)\,\mathrm{d}s,$$

$$\bar{S}_7 = \frac{v_2 v_3 v_5 v_6}{v_{-2}}\int_0^t \frac{(v_7-v_6)\,e^{v_5(s-t)}-(v_7-v_5)\,e^{v_6(s-t)}+(v_6-v_5)\,e^{v_7(s-t)}}{(v_7-v_6)\,(v_7-v_5)\,(v_6-v_5)}\bar{S}_2(s)\,\mathrm{d}s,$$

$$\bar{P} = \frac{v_2 v_3 v_5 v_6 v_7}{v_{-2}}\int_0^t \frac{\frac{v_7-v_6}{v_5}\left(1-e^{v_5(s-t)}\right)-\frac{v_7-v_5}{v_6}\left(1-e^{v_6(s-t)}\right)+\frac{v_6-v_5}{v_7}\left(1-e^{v_7(s-t)}\right)}{(v_7-v_6)\,(v_7-v_5)\,(v_6-v_5)}\bar{S}_2(s)\,\mathrm{d}s. \tag{6}$$

where $\alpha = \bar{A} + v_2 v_3/v_{-2}$. We see that these asymptotic solutions agree well with the numerical ones, except at early time where there is an additional asymptotic region that we do not consider (Figure 4).

Although the leading-order solution for $\bar{P}$, given in (6), depends on parameters from almost all of the reactions, the dependence on v_5, v_6, and v_7 is significantly reduced as $t \to \infty$, and we may determine that the leading-order long-time be-

haviour in this regime is given by

$$\bar{P} \sim \frac{v_2 v_3}{v_{-2}} \left(\frac{\bar{K}_1^i}{\alpha} \right)^{1/2} t \quad \text{as } t \to \infty. \tag{7}$$

Hence, the reactions involving acetyl-CoA are the most important for the production of IDP. We can see that the numerical results shown in Figure 4b agree with the analytic result that $\bar{P}$ tends to a linear function of time. We may also deduce the leading-order long-time behaviour of HMG-CoA in the form

$$\bar{S}_4 \to \frac{v_2 v_3}{v_{-2} v_4} \left(\frac{\bar{K}_1^i}{\alpha} \right)^{1/2} \quad \text{as } t \to \infty. \tag{8}$$

In dimensional terms, the long-time behaviour of the isopentenyl diphosphate (IDP) concentration is

$$[P] \sim \omega E_3 \left(\frac{E_1 k_1 K_1^i}{A + \omega E_3} \right)^{1/2} \tau \quad \text{as } \tau \to \infty, \tag{9a}$$

and the maximum level of HMG-CoA present in the system is

$$[S_4] \sim \frac{\omega E_3 K_4^M}{k_4 E_4} \left(\frac{E_1 k_1 K_1^i}{A + \omega E_3} \right)^{1/2}, \tag{9b}$$

where

$$\omega = \frac{k_2 k_3 K_{-2}^M}{k_{-2} K_2^M K_{3,b}^M}. \tag{9c}$$

From the explicit results (9), we see that the long-time production of IDP depends on the enzyme concentrations E_1 and E_3, whereas the maximum concentration of HMG-CoA depends on the enzyme concentrations E_1, E_3, and E_4. Recalling that our goal is to maximize IDP whilst minimizing HMG-CoA, and noting that the dependence on E_1 and E_3 is the same for both metabolites of interest, the only way our goal can be achieved by varying enzyme concentration is to significantly increase E_4, so that it compensates for any increase in E_1 or E_3. That is, our model suggests that we should overexpress HMG-CoA reductase, the enzyme that catalyses the reaction from HMG-CoA to mevalonate. Importantly, we are able to deduce that this is the most significant leading-order effect, and

this explicit result is possible due to our asymptotic analysis. The significance of this reaction is in agreement with [4], where it was shown that over-expression of HMG-CoA reductase alleviated the inhibition of cell growth, benefiting IDP production. Increasing E_3, the enzyme that catalyses the reaction from acetoacetyl-CoA and acetyl-CoA to HMG-CoA would increase the levels of IDP produced by a single cell, but would also produce more HMG-CoA, which would reduce the number of cells in the system (though this is not formally taken into account by our model). The same is true of increasing E_1, the enzyme that catalyses the reaction from pyruvate to acetyl-CoA, and decreasing A, any enzyme that catalyses reactions from acetyl-CoA to sinks of acetyl-CoA. Finally, we note that these results all agree with the numerical results in Figure 2.

3.2.2. No replenishment of pyruvate

We now consider the case where pyruvate can be depleted. Note that this case, unlike the previous, is singularly perturbed on the long timescale. From (3a), we must also consider the governing equation

$$\frac{\mathrm{d}\bar{S}_1}{\mathrm{d}t} = -\frac{\varepsilon^{1/2}\bar{K}_1^i\bar{S}_1}{\varepsilon^{1/2}\bar{K}_1^i\left(\bar{S}_1+\bar{K}_1^M\right)+\bar{S}_1\bar{S}_2}, \tag{10}$$

instead of imposing $\bar{S}_1(t)\equiv 1$. By seeking a power series representation for $\bar{S}_1$ in terms of the small parameter $\varepsilon^{1/2}$, and using the solution (6a) for $\bar{S}_2$, we obtain the following asymptotic solution to (10) for $t=O(1)$:

$$\bar{S}_1(t) \sim 1-\left(\frac{\varepsilon\bar{K}_1^i}{\alpha}\right)^{1/2}\left(\alpha t+\log\left\{1+\left[1-e^{-2\alpha t}\right]^{1/2}\right\}\right)+O(\varepsilon). \tag{11}$$

Importantly, we find that $\bar{S}_1(t)\sim 1+O(\varepsilon^{1/2})$ for $t=O(1)$, and thus the no replenishment case is equivalent to the continuous replenishment case at leading order for $t=O(1)$.

As suggested by (11), the $O(1)$ decay of pyruvate occurs over the timescale $t=O(\varepsilon^{-1/2})$. Further, as the dominant balances over this timescale do not change, the solution (11) allows us to deduce that the depletion of pyruvate occurs when

$$t^* = 1/(\varepsilon\alpha\bar{K}_1^i)^{1/2}+O(1). \tag{12}$$

For completion, we show the dynamics of this depletion, where the levels of pyruvate become exponentially small, in Appendix B. Using the depletion time (12) in the long-time IDP solution (7) and reverting back to the original unscaled

version of IDP, we determine that the total amount of IDP produced in the no replenishment of pyruvate case is

$$\lim_{t\to\infty} P = \frac{v_2 v_3}{\alpha v_{-2}} + O(\varepsilon^{1/2}). \tag{13}$$

Further, we may deduce that the maximum level of HMG-CoA in the system is again given by (8). In dimensional terms, the long-time behaviour of the isopentenyl diphosphate (IDP) concentration is

$$[P] \sim \frac{\omega S_0 E_3}{A + \omega E_3} \quad \text{as } \tau \to \infty, \tag{14}$$

and the maximum level of HMG-CoA present in the system is

$$[S_4] \sim \frac{\omega E_3 K_4^M}{k_4 E_4} \left(\frac{E_1 k_1 K_1^i}{A + \omega E_3} \right)^{1/2}, \tag{15}$$

where ω is defined in (9c).

Hence, our main conclusions from the continuous replenishment case are still valid for this case. For this no-replenishment-of-pyruvate case, we are also able to further determine that a lower value of K_1^i, perhaps achievable by introducing heterologous enzymes for the reaction (1a), would decrease the maximum amount of HMG-CoA present whilst having no significant effect on the total IDP produced.

4. Experimental validation

To validate the predictions made in this model, we carried out *in vivo* experiments in *Cupriavidus necator* H16, a gram-negative bacterium previously known as *Ralstonia eutropha*. *C. necator* is a facultative chemolithoautotrophic microorganism of relevant biotechnological interest since it can be exploited as a bacterial chassis for the production of chemicals. As proof of concept, we sought to introduce the upper part of the mevalonate pathway, leading to the production of isoprenoid precursors, into *C. necator*. To this end, *C. necator* H16 was transformed with a plasmid (pBBR1JW3) carrying the *mvaE* and *mvaS* genes from *E. faecalis* of the *L*-arabinose inducible promoter P_{BAD}. These two genes code for the enzymes Acetyl-CoA acetyltransferase/HMG-CoA reductase (MvaE) and Hydroxymethylglutaryl-CoA synthase (MvaS), respectively. MvaE is a bifunctional enzyme that catalyses two reactions, from acetyl-CoA to acetoacetyl-CoA and from HMG-CoA to mevalonate, while MvaS converts acetoacetyl-CoA and

acetyl-CoA to HMG-CoA. In our model, increasing the expression of MvaE corresponds to increasing E_2 and E_4, whereas increasing the expression of MvaS corresponds to increasing E_3 (a schematic of the pathway is shown in Figure 1). This plasmid provides a path from pyruvate to mevalonate (S_1 to S_5) within the modified *C. necator*.

4.1. Setting up of bacterial cultures for mevalonate production

Single colonies of *C. necator* H16/pBBR1-USERcassette1 and *C. necator* H16/pBBR1JW3 (pBBR1::*araC*/P_{BAD}-*mvaES*) were used to inoculate 5ml of LB solution with 300μg/ml kanamycin (in 50ml tubes) and grown overnight at 30 °C with shaking (200 rpm). The following morning, the optical density (OD_{600}) of the cultures was measured and normalised to $OD_{600} = 0.2$ in 100ml of LB solution with 300μg/ml kanamycin (in 500ml flasks). The bacterial cultures were then incubated at 30 °C with shaking (200 rpm) until they reached an OD_{600} of around 0.6 - 0.7. At this point, 1ml samples were collected from each culture, centrifuged at 14000 rpm for one minute and the cell pellets were stored at -20 °C to be used as pre-induction protein samples for SDS-PAGE. In addition, 1ml of 1% (w/v); 2% (w/v); or 20% (w/v) *L*-arabinose solutions were added to the corresponding cultures, thus obtaining *L*-arabinose final concentrations of 0.01%; 0.02%; and 0.2%, respectively. Following 18 hours of incubation at 30 °C with shaking (200 rpm), the OD_{600} of the cultures was measured. The bacterial cultures were then centrifuged (8000 rpm for five minutes), concentrated to $OD_{600} = 15$ in 1% fructose minimal medium (FMM) –NH_4Cl with 300μg/ml kanamycin and transferred to 250ml flasks. After 24 hours of incubation at 30 °C with shaking (200 rpm), 500μl samples were taken from the cultures and centrifuged (14000 rpm for one minute). Supernatants were then analysed by HPLC to quantify the mevalonate titres produced by *C. necator* H16/pBBR1JW3 in response to the different *L*-arabinose concentrations taken into account.

4.2. Model comparison

We can use our model to predict the behaviour of the shorter pathway from pyruvate to mevalonate, by taking the limit $\bar{k}_5, \bar{k}_6, \bar{k}_7 \to 0$. In this case, the results for $[S_4]$ remain the same and, in dimensional terms, the long-time results for $[S_5]$ are

$$[S_5] \sim \omega E_3 \left(\frac{E_1 k_1 K_1^i}{A + \omega E_3} \right)^{1/2} \tau \quad \text{as } \tau \to \infty, \tag{16}$$

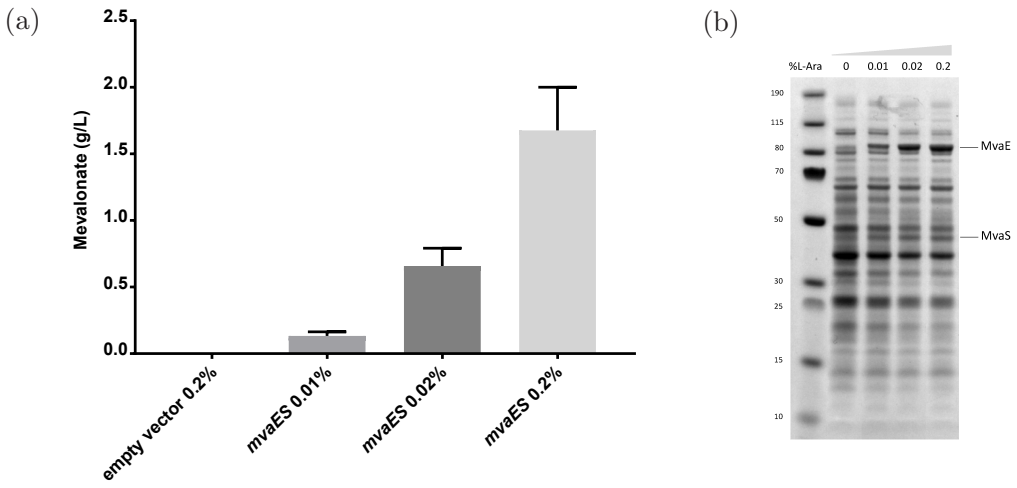

Figure 5: Experimental results for (a) mevalonate production by *C. necator* H16/pBBR1JW3 (pBBR1::*araC*/P_{BAD}-*mvaES*) after 24 hours of growth in 1% (w/v) fructose minimal medium (FMM) in response to the following *L*-arabinose concentrations: 0.01% (w/v - grey bar); 0.02% (w/v - dark grey bar); and 0.2% (w/v - light grey bar). As a negative control, mevalonate production was assessed also with the *C. necator* H16/pBBR1-USERcassette1 strain (empty vector), that does not harbour the *mvaE* and *mvaS* genes. (b) SDS-PAGE of pre- (lane 1) and post-induction protein extracts of *C. necator* H16/pBBR1JW3. Expression of MvaE (86 KDa) and MvaS (42 KDa) was assessed following induction with 0.01% (lane 2); 0.02% (lane 3); and 0.2% (lane 4) *L*-arabinose. While MvaE expression levels increased with increasing *L*-arabinose concentrations, production of MvaS appeared to remain constant.

for continuous replenishment of pyruvate and

$$[S_5] \to \frac{\omega S_0 E_3}{A + \omega E_3} \quad \text{as } \tau \to \infty, \tag{17}$$

for no replenishment of pyruvate, where $\omega = k_2 k_3 K_{-2}^M / (k_{-2} K_2^M K_{3,b}^M)$, as previously defined in (9c). All variables are defined in §2. These are the same results our model predicts for $[P]$, so our previous predictions for maximizing IDP while minimizing HMG-CoA will also be our current predictions for maximizing mevalonate while minimizing HMG-CoA in this experiment.

HPLC data show that increasing the concentration of the inducer *L*-arabinose leads to significant increments in the amount of mevalonate produced after 24 hours of growth in the presence of fructose as a carbon source (Figure 5a). From the three levels of *L*-arabinose we considered, there appear to be diminishing re-

turns on the effectiveness of increasing *L*-arabinose. From an SDS-PAGE, we see an increase in the expression of MvaE as the levels of *L*-arabinose are increased (Figure 5b). However, production of MvaS appears to be constant as the levels of *L*-arabinose are increased. In our model, this corresponds to increasing E_2 and E_4 while keeping E_3 (and all other enzyme levels) constant. As discussed above, our model predicts that this will decrease the maximum levels of HMG-CoA in the system, and thus produce more IDP. Hence, our model is verified by these experimental results.

5. Conclusions

We developed and solved a mathematical model for the kinetics of the mevalonate pathway, derived using generalised Michaelis–Menten kinetics and including the effect of a sink from acetyl-CoA into other metabolic pathways. We considered two extreme cases, namely where the pyruvate was continuously and never replenished. We used asymptotic analysis to gain physical insight into the system behaviour, allowing us to evaluate the effect of upregulating different reactions without resorting to an expensive parameter sweep. The system we considered here has eight dependent variables with 20 kinetic parameters. Our asymptotic analysis enabled us to give analytic expressions for each dependent variable in the continuous replenishment case, and to reduce the entire system to numerically solving a coupled nonlinear system of two dependent variables with one parameter in the never replenished case. We then validated our model by performing experiments that agreed with our predictions.

The main experimental goal is to maximise IDP production, whilst minimizing the maximum levels of HMG-CoA, a metabolite that is linked to the inhibition of cell growth due to its inhibition of fatty acid biosynthesis [5]. In terms of over-expressing enzymes, we see that over-expressing E_3, the enzyme that catalyses the reaction from acetoacetyl-CoA and acetyl-CoA to HMG-CoA will have a positive effect on both IDP production and maximum levels of HMG-CoA. Thus, over-expressing this enzyme by itself will not have significant effects. However, we additionally note that over-expressing E_4, the enzyme that catalyses the reaction from HMG-CoA to mevalonate, will decrease the levels of HMG-CoA without having a significant effect on IDP production. Therefore, over-expressing both E_3 and E_4 should have a much larger effect on IDP production than just over-expressing E_3. The importance of the reaction from HMG-CoA to mevalonate has been previously noted by experiments [4]. We also note that increasing E_1 has a positive effect on both IDP production for continuously replenished pyruvate,

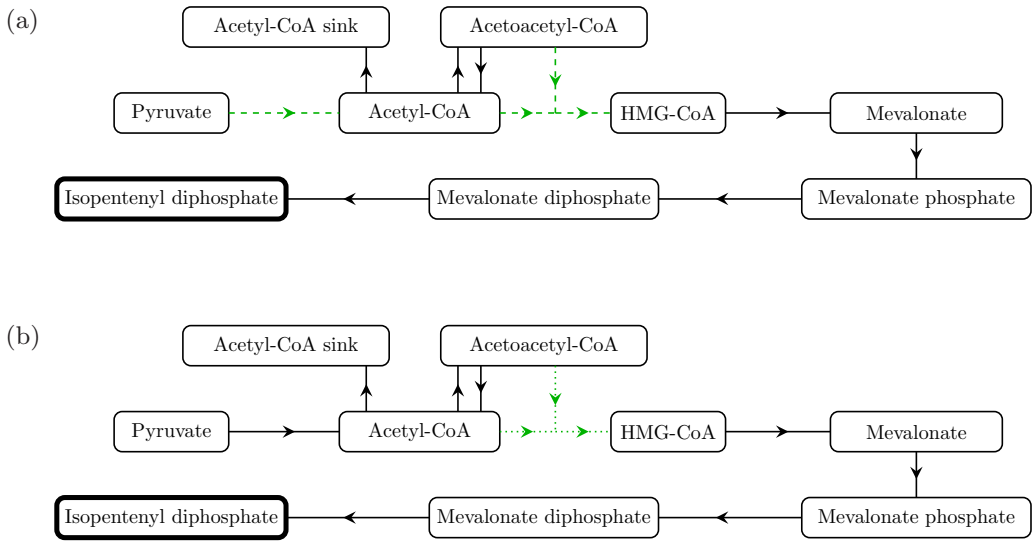

Figure 6: Schematic showing how overexpression of enzymes can affect IDP production in the cases of pyruvate being (a) continuously replenished and (b) never replenished. A green arrow denotes that overexpressing the enzyme corresponding to that reaction results in greater production of IDP, but comes with diminishing returns. A dashed or dotted arrow denotes whether these diminishing returns are unbounded or bounded, respectively.

and on the maximum levels of HMG-CoA. Thus, as for E_3, this effect can be amplified by also over-expressing E_4 at the same time. We illustrate these results in Figures 6 and 7.

Our results also suggest that the reactions between acetyl-CoA and acetoacetyl-CoA can have a significant effect on both IDP production and maximum levels of HMG-CoA. However, this effect cannot be achieved by over- or under-expressing E_2, the enzyme that catalyses this reaction, as the enzyme catalyses the reaction in both directions. We see that this positive effect can be attained by using an enzyme that strongly favours the forward reaction. That is, an enzyme with a large ratio of $k_2K_{-2}^M/k_{-2}K_2^M$. The final reaction that plays an important role at leading order is the reaction from acetyl-CoA to a sink, which has a negative leading-order effect on both IDP production and the maximum levels of HMG-CoA. Although we consider the dimensionless sink coefficient to be of $O(1)$ in our main analysis, we also show in Appendix A that a large sink coefficient does not significantly affect the system behaviour. However, we note that the importance of acetyl-CoA to other metabolic pathways represented by this sink, such as the citric acid cycle, means that there are likely to be negative effects to the cell if this reaction is

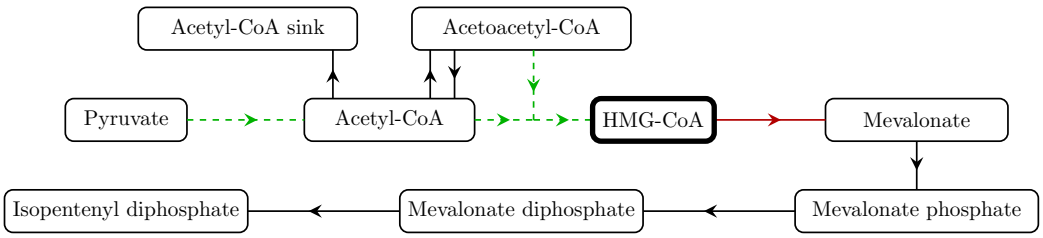

Figure 7: Schematic showing how overexpression of enzymes can affect levels of HMG-CoA. A green/red arrow denotes that overexpressing the enzyme corresponding to that reaction results in greater/lesser amounts of HMG-CoA, and a dashed arrow denotes that over-expression results in unbounded but diminishing returns. The results are the same for the cases of pyruvate being continuously and never replenished.

significantly altered. Our results suggest that the reactions and enzymes we have mentioned are the only significant reactions at leading order, and we predict that the over-expression of other enzymes in the pathway will not have a significant effect on either IDP production or on the maximum levels of HMG-CoA.

In the case where pyruvate was never replenished, we found that the depletion of pyruvate occurred slowly enough that, for intermediate time, our system reproduced the steady-state behaviour we would eventually expect from the continuously replenished pyruvate case. However, whilst the system was initially robust to this slow depletion, we additionally calculated the point at which the low levels of pyruvate affected the entire system, then calculated the dynamics of this depletion. This required the use of the method of matched asymptotic expansions with logarithmic matching terms involving the small variable to track metabolite concentrations up to the first correction order over two timescales, thus allowing us to obtain the leading-order depletion at the third and final timescale.

In our experiments, we were able to introduce a pathway from pyruvate to mevalonate into the bacterium *C. necator*. This was achieved by transforming *C. necator* with a plasmid harbouring the *E. faecalis mvaE* and *mvaS* genes under the control of a promoter induced by the presence of *L*-arabinose (P_{BAD}). The proteins MvaE and MvaS, respectively encoded by these two genes, are responsible for the conversion of acetyl-CoA to mevalonate. A protein expression analysis showed that production of MvaE increased as the levels of *L*-arabinose present in the cultivation medium were increased. On the other hand, expression of MvaS appeared to be independent of *L*-arabinose concentration. In any case, the effect of increasing MvaE expression was to produce more mevalonate, an experimental result that agreed with the theoretical predictions from our model. We note that

our model could be verified further by directly measuring the levels of HMG-CoA. However, as HMG-CoA is not secreted by the cell and is likely to be unstable, direct measurement is significantly more technical; as carried out in Pitera et al. [4], the experiments would have to be repeated while rapidly quenching the cellular metabolism, before examining the endo-metabolites using mass spectrometry.

In the derivation of this model, we chose initial conditions that modelled the instantaneous addition of pyruvate to a well-mixed solution of enzymes. These conditions were chosen for mathematical convenience and are likely to differ significantly from the conditions within a cell producing isopentenyl-diphosphate. However, as the systems we have considered both tend to stable steady states, this difference is unlikely to be a practical issue for the long time results, although it would affect the initial transients.

Finally, we note that this work highlights how a simple mathematical model can be used to predict the biologically relevant behaviour of a system. Moreover, this work shows how asymptotic analysis can play an important role in reducing the computational complexity of a derived system and be used to overcome uncertainty issues with parameter values. We hope that the predictions from our model can help to build a more efficient path to biological discovery.

Acknowledgements

The authors would like to thank Prof. Eriko Takano (director of the Manchester Synthetic Biology Research Centre, SYNBIOCHEM) and Dr. Swathi Alagesan for donating some of the plasmids used in this research paper. This work was supported by the Biotechnology and Biological Sciences Research Council [grant number BB/L013940/1]; and the Engineering and Physical Sciences Research Council, jointly funding the first grant number.

Appendix A. Large $\bar{A}$ limit

There is another distinguished limit in the system (3) when $\bar{A} = O(1/\varepsilon)$. We briefly summarise the system behaviour in this limit for $t = O(1)$. The leading-order system (5) is mainly unchanged, apart from the first equation for $\bar{S}_2$, which becomes

$$\frac{\varepsilon \bar{K}_1^i S_1}{\varepsilon^{1/2} \bar{K}_1^i \left(S_1 + \bar{K}_1^M\right) + S_1 S_2} = \varepsilon \bar{A} S_2, \tag{A.1}$$

where we now have $S_2 = O(\varepsilon^{1/2})$, providing an asymptotic balance between the two terms. Rearranging (A.1) for S_2 in terms of S_1, we obtain

$$S_2 = \frac{\varepsilon^{1/2}}{2}\left[\left\{\left(\frac{\bar{K}_1^i(S_1+\bar{K}_1^M)}{S_1}\right)^2 + \frac{4\bar{K}_1^i}{\varepsilon\bar{A}}\right\}^{1/2} - \frac{\bar{K}_1^i\left(S_1+\bar{K}_1^M\right)}{S_1}\right]. \tag{A.2}$$

For the continuous-replenishment case we have $S_1(t) \equiv 1$, and for the no-replenishment case the governing equation for S_1 is given by (10). Finally, focusing on the continuous-replenishment case, we note that we can re-write (7), the long-time production of IDP, as

$$\bar{P} \sim \left(\frac{\nu_2\nu_3}{\nu_{-2}}t\right)\lim_{t\to\infty} S_2(t) \quad \text{as } t \to \infty. \tag{A.3}$$

Thus, from (A.2) we see that the long-time production of IDP gains some further dependence on the kinetic properties of E_1 in the limit of large $\bar{A}$, but retains the same qualitative behaviour.

Appendix B. Dynamics of pyruvate depletion

In this appendix, we further investigate the dynamics of pyruvate depletion. We only investigate $\bar{S}_1$ and $\bar{S}_2$ here, as the leading-order versions of the remaining metabolites can be obtained by suitably rescaling the solutions (6b–g) for the timescales we present in this appendix.

To obtain the leading-order solutions during the important depletion timescale, we require the long-time behaviour of the $O(\varepsilon^{1/2})$ correction to $\bar{S}_2$ at $t = O(1)$. To this end, we may use the leading-order solution (6) to write the asymptotic expansion

$$\bar{S}_2 \sim \left(\frac{\bar{K}_1^i\left(1-e^{-2\alpha t}\right)}{\alpha}\right)^{1/2} + \varepsilon^{1/2}\bar{K}_1^i s_2(t) \quad \text{as } \varepsilon \to 0, \tag{B.1}$$

where $s_2(t) = O(1)$ is the first correction term for $\bar{S}_2$. We combine (4a,b), (11), and (B.1) to obtain

$$\frac{1}{\alpha}\frac{\mathrm{d}s_2}{\mathrm{d}t} + \left(\frac{2-e^{-2\alpha t}}{1-e^{-2\alpha t}}\right)s_2 = \frac{2\beta}{1-e^{-2\alpha t}} + 2\gamma\left(1-e^{-2\alpha t}\right), \tag{B.2a}$$

where

$$\beta = -\frac{\bar{K}_1^M + 1}{2}, \tag{B.2b}$$

$$\gamma = \frac{v_2 v_3}{2\alpha^2 v_{-2}} \left(\frac{1}{\bar{K}_2^M} - \frac{v_2}{\bar{k}_{-2}} + \frac{v_2 v_3}{v_{-2}\bar{k}_3} \right). \tag{B.2c}$$

Solving (B.2a) in the long-time limit, we deduce that

$$s_2 \sim \beta + \gamma + \text{exponentially small terms} \quad \text{as } t \to \infty, \tag{B.3}$$

and thus we are able to determine the long-time behaviour

$$\bar{S}_2 \sim \left(\frac{\bar{K}_1^i}{\alpha} \right)^{1/2} + \varepsilon^{1/2} \bar{K}_1^i (\beta + \gamma) \quad \text{as } t \to \infty. \tag{B.4}$$

The $O(1)$ depletion of pyruvate occurs over the timescale $t = O(\varepsilon^{-1/2})$, and we investigate this using the scaling $t = T/\varepsilon^{1/2}$, where $T = O(1)$. Writing $\bar{S}_1$ and $\bar{S}_2$ as the asymptotic expansions

$$\bar{S}_1 \sim \bar{S}_1^{(0)} + \varepsilon^{1/2} \bar{S}_1^{(1)}, \quad \bar{S}_2 \sim \bar{S}_2^{(0)} + \varepsilon^{1/2} \bar{S}_2^{(1)} \quad \text{as } \varepsilon \to 0, \tag{B.5}$$

the appropriate governing equations, obtained from (10) and (4a,b), are

$$\frac{\mathrm{d}\bar{S}_1^{(0)}}{\mathrm{d}T} = -\frac{\bar{K}_1^i}{\bar{S}_2^{(0)}}, \tag{B.6a}$$

$$\frac{\mathrm{d}\bar{S}_1^{(1)}}{\mathrm{d}T} = \frac{\bar{K}_1^i}{\left(\bar{S}_2^{(0)}\right)^2} \left(\bar{S}_2^{(1)} + \bar{K}_1^i \left(1 + \frac{\bar{K}_1^M}{\bar{S}_1^{(0)}} \right) \right), \tag{B.6b}$$

$$0 = \frac{\bar{K}_1^i}{\bar{S}_2^{(0)}} - \alpha \bar{S}_2^{(0)}, \tag{B.6c}$$

$$\frac{\mathrm{d}\bar{S}_2^{(0)}}{\mathrm{d}T} = -2\alpha \left(\bar{S}_2^{(1)} + \bar{K}_1^i \left(\frac{1}{2} \left(1 + \frac{\bar{K}_1^M}{\bar{S}_1^{(0)}} \right) - \gamma \right) \right), \tag{B.6d}$$

with matching conditions, valid when $T \to 0^+$,

$$\bar{S}_1^{(0)} \sim 1 - \left(\alpha \bar{K}_1^i \right)^{1/2} T, \quad \bar{S}_1^{(1)} \sim -\left(\frac{\bar{K}_1^i}{\alpha} \right)^{1/2} \log 2,$$

$$\bar{S}_2^{(0)} \sim \left(\frac{\bar{K}_1^i}{\alpha}\right)^{1/2}, \quad \bar{S}_2^{(1)} \sim \bar{K}_1^i(\beta+\gamma), \tag{B.6e}$$

obtained using Van Dyke's matching rule [30] on the $t = O(1)$ results (11) and (B.4).

The system (B.6) is thus solved by

$$\bar{S}_1^{(0)} = 1 - \left(\alpha\bar{K}_1^i\right)^{1/2} T, \tag{B.7a}$$

$$\bar{S}_1^{(1)} = \alpha\bar{K}_1^i\left(\gamma+\frac{1}{2}\right)T - \frac{\bar{K}_1^M\left(\alpha\bar{K}_1^i\right)^{1/2}}{2}\log\left(1-\left(\alpha\bar{K}_1^i\right)^{1/2}T\right) - \left(\frac{\bar{K}_1^i}{\alpha}\right)^{1/2}\log 2, \tag{B.7b}$$

$$\bar{S}_2^{(0)} = \left(\frac{\bar{K}_1^i}{\alpha}\right)^{1/2}, \tag{B.7c}$$

$$\bar{S}_2^{(1)} = \bar{K}_1^i\left(\beta+\gamma-\frac{\left(\alpha\bar{K}_1^i\right)^{1/2}T}{2\left(1-\left(\alpha\bar{K}_1^i\right)^{1/2}T\right)}\right). \tag{B.7d}$$

The most significant asymptotic regime of interest for the depletion dynamics occurs when $\bar{S}_1 = O(\varepsilon^{1/2})$, which results in the depletion of several metabolites in the system. This region occurs when $t = 1/(\varepsilon\alpha\bar{K}_1^i)^{1/2} + \hat{t}/\alpha$ (equivalently, when $T = 1/(\alpha\bar{K}_1^i)^{1/2} + \varepsilon^{1/2}(\hat{t}/\alpha)$), where $\hat{t} = O(1)$. We additionally scale $(\bar{S}_1, \bar{S}_2) = (\bar{K}_1^i/\alpha)^{1/2}(\varepsilon^{1/2}\hat{S}_1(\hat{t}), \hat{S}_2(\hat{t}))$. Substituting these scalings into (10) and (4a,b) results in the leading-order long-time equations

$$\frac{\mathrm{d}\hat{S}_1}{\mathrm{d}\hat{t}} = -\frac{\hat{S}_1}{\alpha\bar{K}_1^M + \hat{S}_1\hat{S}_2}, \tag{B.8a}$$

$$\frac{\mathrm{d}\hat{S}_2}{\mathrm{d}\hat{t}} = \frac{\hat{S}_1}{\alpha\bar{K}_1^M + \hat{S}_1\hat{S}_2} - \hat{S}_2. \tag{B.8b}$$

The leading-order matching conditions as $\hat{t} \to -\infty$ are $\hat{S}_1 \sim -\hat{t}$ and $\hat{S}_2 \to 1$, and the remaining variables in the system are still given by (6c-h). In practice, we find that this leading-order system is sensitive to the $O(\varepsilon^{1/2})$ matching condition for $\hat{S}_1$ as $\hat{t} \to -\infty$. We can deduce this directly by determining the behaviour of $\hat{S}_1$ as $\hat{t} \to -\infty$ from the system (B.8). That is, we deduce that

$$\hat{S}_1 \sim -\hat{t} - (\alpha\bar{K}_1^M/2)\log(-\hat{t}) + C \quad \text{as } \hat{t} \to -\infty. \tag{B.9a}$$

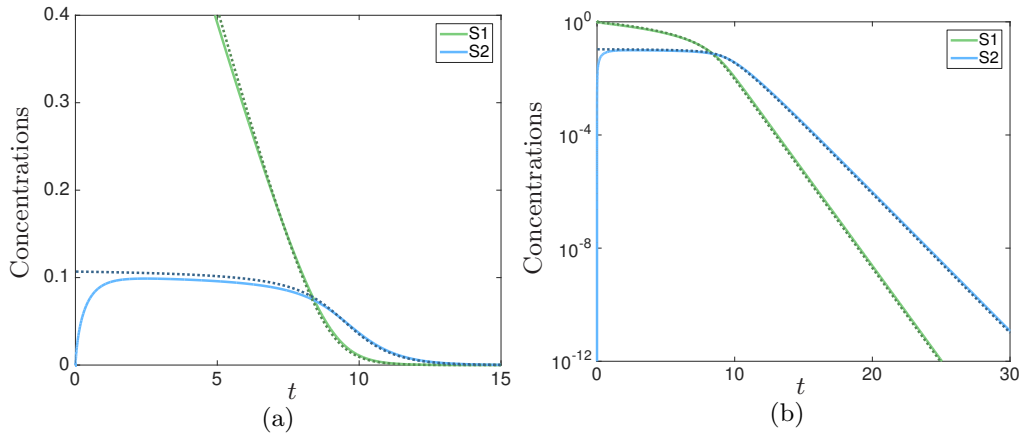

Figure B.8: The solutions to the full and reduced systems for the concentrations of S_1 and S_2 in the no replenishment of pyruvate case. We show (a) A linear plot (b) A log-lin plot. The solid light lines denote the numerical solutions to the full problem, and the dashed darker lines denote the solutions to the reduced problem given in (B.8). We see good agreement between the numerical and asymptotic solutions apart from at early time, as we expect.

We may determine C by matching the solution to $\bar{S}_1$ up to $O(\varepsilon^{1/2})$ when $t = O(\varepsilon^{-1/2})$, given in (B.7a,b). From this matching, we deduce that

$$C = -\log 2 + \frac{1}{2}\left(\frac{\nu_2\nu_3}{\alpha\nu_{-2}}\left(\frac{1}{\bar{K}_2^M} - \frac{\nu_2}{k_{-2}} + \frac{\nu_2\nu_3}{k_3\nu_{-2}}\right) + \alpha\right) - \frac{\alpha\bar{K}_1^M}{4}\log\left(\frac{\varepsilon\bar{K}_1^i}{\alpha}\right) + O(\varepsilon^{1/2}). \tag{B.9b}$$

Using (B.9) as the matching condition for $\hat{S}_1$, we may numerically solve (B.8) to obtain the depletion dynamics for $\hat{S}_1$ and $\hat{S}_2$, then use the latter solution in (6b–g) to deduce the depletion dynamics for the remaining metabolites. We see excellent agreement between these asymptotically determined solutions and the full numerical solutions (Figure B.8).

As the system (B.8) is governed by two coupled autonomous nonlinear equations, we further note that we are able to reduce these to a single nonlinear ordinary differential equation by dividing one equation by the other and seeking a solution to $\bar{S}_2(\bar{S}_1)$. Although this first-order differential equation can be transformed into a linear second-order differential equation with solutions involving modified Bessel functions, reintroducing t into this system is nontrivial, and the solutions do not provide any insight into the underlying system behaviour. Therefore, we do not consider this route further.

[1] V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, Engineering a mevalonate pathway in Escherichia coli for production of terpenoids, Nature Biotechnol 21 (7) (2003) 796–802.

[2] K. Tabata, S.-I. Hashimoto, Production of mevalonate by a metabolically-engineered Escherichia coli, Biotech Lett 26 (19) (2004) 1487–1491.

[3] J. D. Newman, J. Marshall, M. Chang, F. Nowroozi, E. Paradise, D. Pitera, K. L. Newman, J. D. Keasling, High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered Escherichia coli, Biotech Bioeng 95 (4) (2006) 684–691.

[4] D. J. Pitera, C. J. Paddon, J. D. Newman, J. D. Keasling, Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli, Metab Eng 9 (2) (2007) 193–207.

[5] L. Kizer, D. J. Pitera, B. F. Pfleger, J. D. Keasling, Application of functional genomics to pathway optimization for increased isoprenoid production, Appl Environ Microb 74 (10) (2008) 3229–3241.

[6] J. R. Anthony, L. C. Anthony, F. Nowroozi, G. Kwon, J. D. Newman, J. D. Keasling, Optimization of the mevalonate-based isoprenoid biosynthetic pathway in Escherichia coli for production of the anti-malarial drug precursor amorpha-4,11-diene, Metab Eng 11 (1) (2009) 13–19.

[7] E. J. Hinch, Perturbation methods, Cambridge University Press, 1991.

[8] J. Kevorkian, J. D. Cole, Perturbation methods in applied mathematics, vol. 34, Springer Science & Business Media, 2013.

[9] R. O'Malley, Jr., Singular perturbation methods for ordinary differential equations, vol. 89, Springer Science & Business Media, 2012.

[10] G.-B. Kresze, H. Ronft, Pyruvate dehydrogenase complex from baker's yeast. 1. Purification and some kinetic and regulatory properties, Eur J Biochem 119 (3) (1981) 573–579.

[11] C. Yung-Chi, W. H. Prusoff, Relationship between the inhibition constant (K_I) and the concentration of inhibitor which causes 50 per cent inhibition (I_50) of an enzymatic reaction, Biochem Pharmacol 22 (23) (1973) 3099–3108.

[12] B. Middleton, The kinetic mechanism of 3-hydroxy-3-methylglutaryl-coenzyme A synthase from baker's yeast, Biochem J 126 (1) (1972) 35–47.

[13] G.-B. Kresze, H. Ronft, Pyruvate dehydrogenase complex from baker's yeast. 2. Molecular structure, dissociation, and implications for the origin of mitochondria, Eur J Biochem 119 (3) (1981) 581–587.

[14] J. L. Snoep, M. T. De Mattos, M. Starrenburg, J. Hugenholtz, Isolation, characterization, and physiological role of the pyruvate dehydrogenase complex and alpha-acetolactate synthase of Lactococcus lactis subsp. lactis bv. diacetylactis, J Bacteriol 174 (14) (1992) 4838–4841.

[15] M. Hedl, A. Sutherlin, E. I. Wilding, M. Mazzulla, D. McDevitt, P. Lane, J. W. Burgner, K. R. Lehnbeuter, C. V. Stauffacher, M. N. Gwynn, V. W. Rodwell, Enterococcus faecalis acetoacetyl-coenzyme A thiolase/3-hydroxy-3-methylglutaryl-coenzyme A reductase, a dual-function protein of isopentenyl diphosphate biosynthesis, J Bacteriol 184 (8) (2002) 2116–2122.

[16] E. Okamura, T. Tomita, R. Sawa, M. Nishiyama, T. Kuzuyama, Unprecedented acetoacetyl-coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway, Proc Natl Acad Sci USA 107 (25) (2010) 11265–11270.

[17] K. Matsumoto, Y. Tanaka, T. Watanabe, R. Motohashi, K. Ikeda, K. Tobitani, M. Yao, I. Tanaka, S. Taguchi, Directed evolution and structural analysis of NADPH-dependent Acetoacetyl Coenzyme A (Acetoacetyl-CoA) reductase from Ralstonia eutropha reveals two mutations responsible for enhanced kinetics, Appl Environ Microbiol 79 (19) (2013) 6134–6139.

[18] J. J. Reddick, J. K. Williams, The mmgA gene from Bacillus subtilis encodes a degradative acetoacetyl-CoA thiolase, Biotechnol Lett 30 (6) (2008) 1045–1050.

[19] A. Sutherlin, M. Hedl, B. Sanchez-Neri, J. W. Burgner, C. V. Stauffacher, V. W. Rodwell, Enterococcus faecalis 3-hydroxy-3-methylglutaryl coenzyme A synthase, an enzyme of isopentenyl diphosphate biosynthesis, J Bacteriol 184 (15) (2002) 4065–4070.

[20] D. A. Nagegowda, C. Mee-Len, et al., Brassica juncea 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase 1: expression and characterization of recombinant wild-type and mutant enzymes, Biochemical Journal 383 (3) (2004) 517–527.

[21] S. M. Ma, D. E. Garcia, A. M. Redding-Johanson, G. D. Friedland, R. Chan, T. S. Batth, J. R. Haliburton, D. Chivian, J. D. Keasling, C. J. Petzold, T. S. Lee, C. S. R, Optimization of a heterologous mevalonate pathway through the use of variant HMG-CoA reductases, Metab Eng 13 (5) (2011) 588–597.

[22] Y. P. Ching, S. P. Davies, D. G. Hardie, Analysis of the Specificity of the AMP-Activated Protein Kinase by Site-Directed Mutagenesis of Bacterially Expressed 3-hydroxy 3-methylglutaryl-CoA Reductase, Using a Single Primer Variant of the Unique-site-elimination Method, Eur J Biochem 237 (3) (1996) 800–808.

[23] Y. A. Primak, M. Du, M. C. Miller, D. H. Wells, A. T. Nielsen, W. Weyler, Z. Q. Beck, Characterization of a feedback-resistant mevalonate kinase from the archaeon Methanosarcina mazei, Appl Environ Microbiol 77 (21) (2011) 7772–7778.

[24] K.-X. Huang, A. Scott, G. N. Bennett, Overexpression, purification, and characterization of the thermostable mevalonate kinase from Methanococcus jannaschii, Prot Express Purif 17 (1) (1999) 33–40.

[25] D. E. Garcia, J. D. Keasling, Kinetics of phosphomevalonate kinase from Saccharomyces cerevisiae, PLoS One 9 (1) (2014) e87112.

[26] N. E. Voynova, S. E. Rios, H. M. Miziorko, Staphylococcus aureus mevalonate ki-

nase: isolation and characterization of an enzyme of the isoprenoid biosynthetic pathway, J Bacteriol 186 (1) (2004) 61–67.

[27] D. Krepkiy, H. M. Miziorko, Identification of active site residues in mevalonate diphosphate decarboxylase: implications for a family of phosphotransferases, Prot sci 13 (7) (2004) 1875–1881.

[28] Y. Qiu, J. Gao, F. Guo, Y. Qiao, D. Li, Mutation and inhibition studies of mevalonate 5-diphosphate decarboxylase, Bioorg Medicinal Chem Lett 17 (22) (2007) 6164–6168.

[29] J. T. Pronk, H. Yde Steensma, J. P. van Dijken, Pyruvate metabolism in Saccharomyces cerevisiae, Yeast 12 (16) (1996) 1607–1633.

[30] M. van Dyke, Perturbation methods in fluid dynamics, Parabolic Press, 1975.