Optimization of metabolic states

Andreas Kremling[1], Wolfram Liebermeister[2], Elad Noor[3] and Meike T. Wortel[4]

[1] Systems Biotechnology, School of Engineering and Design, Technical University of Munich, Germany; [2] Université Paris-Saclay, INRAE, MaIAGE, 78350 Jouy-en-Josas, France; [3] Department of Plant and Environmental Sciences, Weizmann Institute of Science, 76100 Rehovot, Israel; [4] Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands

Abstract

Cells in a well-mixed and nutrient-rich environment can be expected to experience selection on their growth rate. Such cells optimize their metabolic state to achieve a high growth rate. Metabolic states that lead to a high growth rate are states that realize a maximal biomass production rate at a minimal enzyme cost. Metabolic states that optimize a specific flux at minimal enzyme cost are called enzyme-efficient metabolic states and in this chapter we refer to them as optimal metabolic states. The calculation of optimal metabolic states is facilitated by the result that, in models without further constraints, the flux distributions in enzyme-efficient states are Elementary Flux Modes (EFMs). This result allows for an algorithm to find enzyme-efficient states by: 1) Enumerating the EFMs, 2) calculating the minimal enzyme cost per EFM, and 3) choosing the one with the lowest enzyme investment. This algorithm finds optimal metabolic states for larger models which cannot be optimized by 'brute force', but are still small enough to enumerate the EFMs. Finding optimal metabolic states uncovered the effect of changing external nutrient conditions: As growth conditions are changing, the optimal flux profile either changes continuously (and metabolite and enzyme concentrations change continuously as well) when the same EFM remains optimal, or fluxes change discontinuously together with metabolite and enzyme concentrations when a different EFM becomes optimal.

Contributions: This chapter was drafted and written by Andreas Kremling, Woflram Liebermeister, Elad Noor, and Meike Wortel. It was initially discussed with Jürgen Zanghellini and later reviewed by David S. Tourigny and Hugo Dourado and discussed with Stefan Müller.

Keywords: elementary flux mode (EFM) - enzyme cost minimization - enzyme-efficient state

To cite this chapter: A. Kremling, W. Liebermeister, E. Noor, M.T. Wortel. Optimization of metabolic states (Version July 2025). doi: 10.5281/zenodo.8164468. Chapter from: The Economic Cell Collective (2025). Economic Principles in Cell Biology. No commercial publisher | Online open access book | doi: 10.5281/zenodo.8156386

The authors are listed in alphabetical order.

This is a chapter from the open textbook "Economic Principles in Cell Biology".
Free download from principlescellphysiology.org/book-economic-principles/.
Lecture slides for this chapter are available on the website.

doi: 10.5281/zenodo.8156386

Chapter overview

- Optimal metabolic states in this chapter refer to enzyme-efficient states, which are metabolic states that realize a given objective flux at a minimal enzyme cost.
- In models without further flux constraints, flux distributions of enzyme-efficient states are Elementary Flux Modes (EFMs).
- Elementary Flux Modes can be used to find enzyme-efficient states in networks that would be too large to optimize metabolic states "by brute force".
- Biomass per enzyme efficiency can be converted to cell growth rate by simple approximate formulae.
- The Elementary Flux Mode that is realized in an enzyme-efficient state depends on the external conditions.
- As growth conditions change, either the flux profile changes continuously (together with metabolite and enzyme concentrations), or fluxes change discontinuously, implying jumps also in metabolite and enzyme concentrations.

7.1. Introduction

In a simple economic picture of cells, we assume that cells adjust their metabolic state in each environment to obtain a maximal fitness advantage. This may be impossible in reality, but it remains an interesting question what this best metabolic state would look like, according to our knowledge of cells. So what is the best metabolic state overall (comprising metabolic fluxes, metabolite concentrations and enzyme levels)? What pathways should a cell use, which enzymes should be induced or repressed, and how should this change in a new environment? To answer these questions, we need to remember that all metabolic variables (fluxes, metabolite levels, enzyme levels, and enzyme efficiencies) depend on each other. Physically, fluxes depend on metabolite concentrations through kinetics and enzyme regulation (e.g. competitive inhibition) and metabolites are produced and consumed by the fluxes until a steady state is reached. Hence, if we think in terms of cellular economics (treating enzymes as control variables), then all metabolic variables must be optimized together.

In the previous chapters we saw some ways to predict optimal metabolic fluxes, metabolite concentrations and enzyme levels separately: in Flux Balance Analysis (FBA, Chapter 5 in [1]), we optimized fluxes by maximizing an objective function (typically biomass) while in Enzyme Cost Minimization [2, 3] (Chapter 6 in [1]) metabolite concentrations were optimized by minimizing cost (or, equivalently, maximizing the enzyme efficiencies). Each of these methods is based on a strong assumption: FBA requires measured flux ranges and/or apparent catalytic rates and assumes enzyme saturation effects can be neglected, while enzyme cost minimization requires a given flux distribution. But what if we don't know any of the variables in advance? How can we predict all of them from first principles?

Before thinking about this, let us briefly step back: what do we actually mean by an "optimal state"? What quantity should be maximized in metabolism? There could be very different aims (e.g. production in biotechnology, versus number of offspring and survival in a wild-type cell). However, in both cases an important aim is cell growth – or at least, avoiding strong growth deficits. Below we will see that cell growth depends, to a good approximation, on biomass/enzyme efficiency, that is, biomass production per total enzyme invested. Hence, whenever fast growth is important, cells should maximize this efficiency.

In conclusion, we will consider the following optimality problem: maximize biomass/enzyme efficiency, defined as the production flux per invested enzyme with respect to all metabolic variables (metabolites, enzymes and fluxes) and under all constraints (steady state, enzyme kinetics, etc.). Solutions to this problem are considered optimal states.

7.2. Enzyme-efficient metabolic states use elementary flux modes

The optimization problem in this chapter is to reach maximal objective flux with minimal enzyme investment. The biological interpretation is that this would lead to the highest growth rate, because it optimizes the ratio between gains (fluxes) and costs (enzymes). When we solve this optimization problem with mathematical tools, it is convenient to either find the minimal enzyme investment for a certain flux, or the maximum flux for a fixed enzyme investment. Although one could think of different biological explanations for those two ways to state the optimization problem, mathematically they are equivalent. For the outline of the proof that optimal states are elementary flux modes, it is convenient to fix the objective flux to an arbitrary value (we choose 1) and then minimize the enzyme investment. This leads to the following optimization problem over the fluxes ($\mathbf{v}$), enzymes levels ($\mathbf{e}$) and internal metabolite concentrations ($\mathbf{s}$):

$$
\begin{aligned}
&\underset{\mathbf{v},\mathbf{e},\mathbf{s}}{\text{minimize}} && \sum_{i=1}^{r} h_i \; e_i && && (7.1)\\
&\text{subject to:} && \mathbf{N}\cdot\mathbf{v} = \mathbf{0} && \text{steady state}\\
& && \forall i : v_i = e_i \; \kappa_i(\mathbf{s}) && \text{enzyme kinetics}\\
& && \mathbf{e},\mathbf{s} \geq 0 && \text{positive concentrations}\\
& && v_r = 1 && \text{fixed objective flux}\\
& && \mathbf{s} \leq \mathbf{s}_{\max} && \text{metabolite bounds}
\end{aligned}
$$

where h_i are the weights, $\mathbf{N}$ is the stoichiometry matrix, i is the index of the reactions (ranging from 1 to r), with the last reaction (with index r) representing the objective. This optimization problem states that by adjusting the fluxes ($\mathbf{v}$), metabolite concentrations ($\mathbf{s}$) and enzyme concentrations ($\mathbf{e}$), the total cost (sum of costs – $h_i e_i$ – for every reaction) is minimized, while keeping the objective flux constant (any arbitrary constant can be chosen, here we chose 1). The weights (h_i) can be thought of as the size or production costs of the enzymes (measured, for example, in molecular weight or gene length) We require certain constraints: (i) the metabolic network needs to be in steady state to avoid built-up of intermediates, (ii) enzyme kinetics – the flux of each reaction (v_i) has to be equal to the enzyme concentration (e_i) times a metabolite dependent (e.g., saturation) term ($\kappa_i(\mathbf{s})$), (iii) all enzyme and metabolite concentrations have to be positive, (iv) the objective flux is equal to 1, and (v) the metabolite concentrations are within their given bounds. The latter constraint is optional and is mostly necessary when dealing with irreversible kinetics. Reversible kinetics usually lead to bounded metabolite levels because very high concentrations of products inhibit the reaction that forms the products.

In this section, we will explain why the optimal state is reached at an Elementary Flux Mode (EFM). One important starting point is that, as we have seen before in Chapter 4 in [1], convex optimization problems with only positivity or equality constraints (no other inequalities) lead to an optimal solution at an extreme point of the feasible solution space, and those extreme points are Elementary Flux Modes. However, the optimization problem (7.1) is not convex, mainly due to the hyperbolic dependence of reaction rates on the concentrations of metabolites ($\kappa_i(\mathbf{s})$ is usually not linear in the internal metabolite concentrations).

There are several ways to prove that the solution of this optimization problem is an EFM, of which some are outlined in the papers by Wortel et al. [4] and Müller et al. [5]. Here we will outline a proof by contradiction: assuming a solution to the optimization problem that is not an EFM and showing that this leads to a contradiction.

Theorem 1. *The flux distribution that maximizes an objective flux over the total enzyme cost in a metabolic network without additional constraints is an Elementary Flux Mode.*

Proof. Assume we have some optimal state where the flux distribution is not an EFM. Any optimal solution is

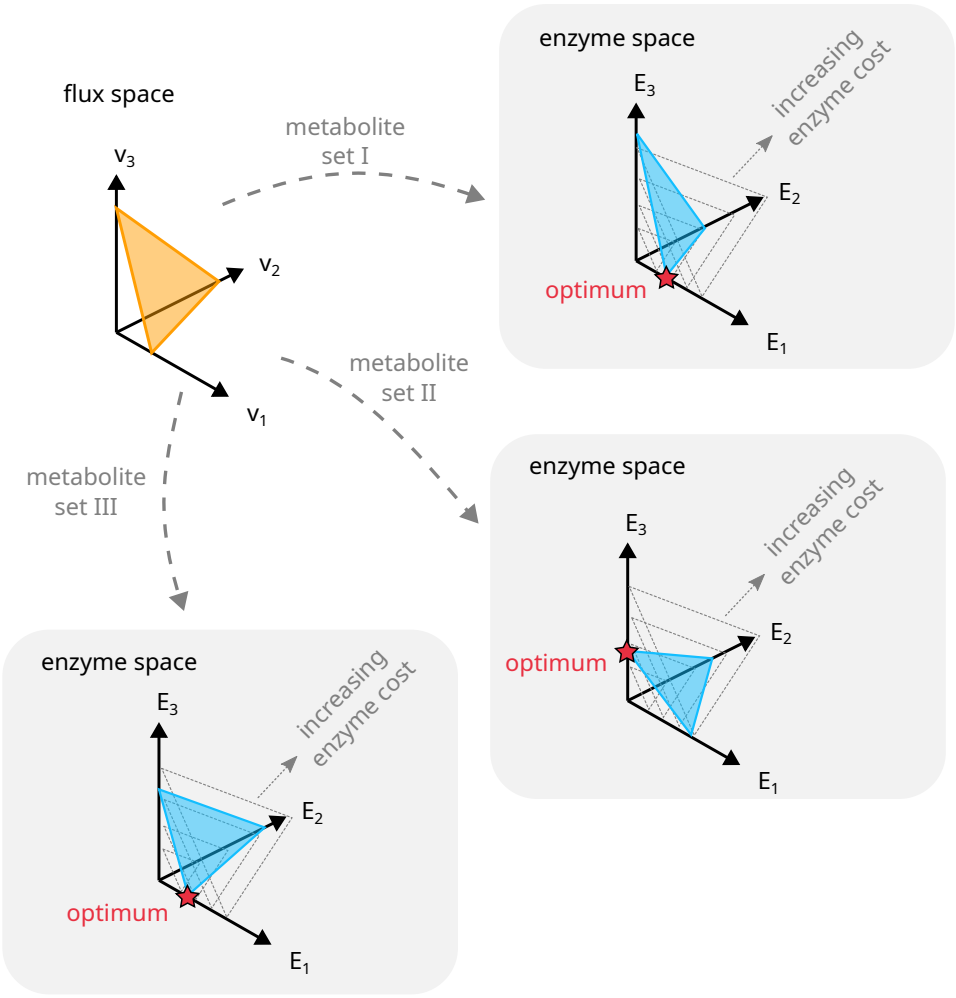


Figure 7.1: Translation from flux to enzyme space retains EFMs as extreme rays – The top left panel shows the feasible flux space with the steady state constraints, all fluxes positive (using splitting of fluxes, as explained in the text, if necessary) and a fixed objective flux. The extreme points here are points where one flux becomes 0 and are elementary flux modes (see Chapter 5 in [1]). Here we show that when we have assumed metabolite concentrations, such as when we keep them at an optimal solution, we get a linear transformation and the extreme rays are maintained. Different metabolite levels, for example solutions to different environmental conditions, can lead to different transformations and therefore different optima (minimal total enzyme), but those are always located at an EFM.

associated with a set of fluxes, enzyme concentrations and metabolite concentrations. Now we set the metabolite concentrations to the concentrations of the assumed optimal state. Then all metabolite-dependent terms ($\kappa_i(\mathbf{s})$) become constants, and we return to a convex problem. As explained in Chapter 10 in [1] and Figure 7.1, the optimum of this problem (now in terms of enzyme concentrations and fluxes) is a flux distribution that is an EFM. But this contradicts our initial assumption that the optimal state from which we took the set of metabolite concentrations was not an EFM. The proof by contradiction shows that the optimal state must be an EFM. □

7.3. Enzyme-efficient states in an example network

To illustrate the proof, we study a simple network representing growth on glucose and pyruvate that we have seen previously in Chapter 5 in [1] (Figure 7.2). We use G and P for glucose and pyruvate in the equations, we use the subscript ex when a metabolite is extracellular and square brackets to denote a concentration. For the use in this chapter, we add enzyme kinetics to this network. We will use the factorized rate law as in Chapter 6 in [1], but then

Box 7.A Kinetics of the example network

The detailed kinetic equations for the example model (Figure 7.2) using the factorized rate law (see Equation (7.2) and Chapters 3 in [1] and 6 in [1]) are:

$$
\begin{aligned}
v_0 &= e_0 \cdot k^+_{\text{cat},0} \cdot \frac{[\text{G}_\text{ex}]/K_{\text{G}_\text{ex}}}{1 + [\text{G}]/K_\text{G} + [\text{G}_\text{ex}]/K_{\text{G}_\text{ex}}} \cdot \left(1 - e^{\Delta_\text{r} G'_0/RT}\right) \\
v_1 &= e_1 \cdot k^+_{\text{cat},1} \cdot \frac{([\text{G}]/K_\text{G})([\text{ADP}]/K_\text{ADP})}{1 + ([\text{P}]/K_\text{P})([\text{P}]/K_\text{P})([\text{ATP}]/K_\text{ATP}) + ([\text{G}]/K_\text{G})([\text{ADP}]/K_\text{ADP})} \cdot \left(1 - e^{\Delta_\text{r} G'_1/RT}\right) \\
v_2 &= e_2 \cdot k^+_{\text{cat},2} \cdot \frac{[\text{P}]/K_\text{P}}{1 + [\text{P}_\text{ex}]/K_{\text{P}_\text{ex}} + [\text{P}]/K_p} \cdot \left(1 - e^{\Delta_\text{r} G'_2/RT}\right) \\
v_3 &= e_3 \cdot k^+_{\text{cat},3} \cdot \frac{([\text{P}]/K_\text{P})([\text{ADP}]/K_\text{ADP})([\text{O}_2]/K_{\text{O}_2})}{1 + ([\text{CO}_2]/K_{\text{CO}_2})([\text{ATP}]/K_\text{ATP}) + ([\text{P}]/K_\text{P})([\text{ADP}]/K_\text{ADP})([\text{O}_2]/K_{\text{O}_2})} \cdot \left(1 - e^{\Delta_\text{r} G'_3/RT}\right) \\
v_4 &= e_4 \cdot k^+_{\text{cat},4} \cdot \frac{[\text{P}_\text{ex}]/K_{\text{P}_\text{ex}}}{1 + [\text{P}_\text{ex}]/K_{\text{P}_\text{ex}} + [\text{P}]/K_\text{P}} \cdot \left(1 - e^{\Delta_\text{r} G'_4/RT}\right) \\
v_\text{BM} &= e_\text{BM} \cdot k^+_{\text{cat,BM}} \cdot \frac{([\text{P}]/K_\text{P})([\text{ATP}]/K_\text{ATP})}{1 + ([\text{BM}]/K_\text{BM})([\text{ADP}]/K_\text{ADP}) + ([\text{P}]/K_\text{P})(\text{ATP}/K_\text{ATP})} \cdot \left(1 - e^{\Delta_\text{r} G'_5/RT}\right)
\end{aligned}
\tag{7.3}
$$

Note that P is a product twice in v_1, as v_1 produces 2P. Note that v_2 and v_4 are the same reaction, but defined in the opposite direction. The standard set of parameters we used for the toy model is all $k^+_{\text{cat},i} = 10\ s^{-1}$ except $k^+_{\text{cat},3} = 0.1\ s^{-1}$, all $\Delta_\text{r} G'^\circ_i/RT = -440$ and all $K_\text{M} = 1$ mM. For the external metabolites $[\text{P}_\text{ex}] = 1$ mM, $[\text{G}_\text{ex}] = 0.05$ mM, $[\text{O}_2] = 0.1$ mM, $[\text{BM}] = 1$ mM and $[\text{CO}_2] = 10$ mM unless mentioned otherwise.

generalized for n_s substrates and n_p products (also compare Eq. (3.10 in [1]) in Chapter 3 in [1]):

$$
v = e \cdot k^+_\text{cat} \cdot \frac{\prod_{n_s}^{j=1} s_j/K_{\text{S},j}}{1 + \prod_{n_p}^{k=1} p_k/K_{\text{P},k} + \prod_{n_s}^{j=1} s_j/K_{\text{S},j}} \cdot \left(1 - e^{\Delta_\text{r} G'/RT}\right) \tag{7.2}
$$

See Box 7.A for all detailed rate laws of the example network. We can simplify this equation by combining the forward catalytic constant, the thermodynamic efficiency factor, the saturation efficiency factor, and the regulation efficiency factor (if that exists) in a function $\kappa(\mathbf{s})$, which only depends on the metabolites, and not on the enzyme concentrations. We will below write κ for $\kappa(\mathbf{s})$.

$$
v_i = e_i \cdot \kappa_i \tag{7.4}
$$

Now we take $v_\text{BM} = 1$ and optimize all fluxes, enzyme concentrations and metabolite concentrations to minimize the enzyme costs ($e_\text{tot} = \sum_i e_i$), while satisfying the constraints posed in Equations (7.1), for different levels of external glucose and standard levels of the other external metabolites. We see that for different concentrations of external glucose, lead to different optimal fluxes, enzyme levels and metabolite levels (Table 7.1).

The table shows that the total enzyme needed for a biomass flux of one decreases with increasing glucose levels, as we expect. In addition, the optimal level of internal glucose increases with increasing external glucose. This is

$[\text{G}_\text{ex}]$	e_tot	v_0	v_1	v_2	v_3	v_4	v_BM	e_0	e_1	e_2	e_3	e_4	e_BM	[G]	[P]	[ATP]	[ADP]
0.01	156.2	5	5	0	9	0	1	54.4	4.4	0	94.4	0	2.9	0.08	15.14	0.05	20.09
0.1	91.3	50	50	99	0	0	1	61.3	11.3	14.2	0	0	4.4	0.13	4.55	0.11	20.09
1	36.2	50	50	99	0	0	1	13.0	8.0	12.5	0	0	2.7	0.60	7.65	0.11	20.09

Table 7.1: Outcomes of the optimization of the example network with standard kinetics, parameter values and external concentrations (see Box 7.A) for varying levels of $[\text{G}_\text{ex}]$.

because a higher external glucose allows for a higher internal glucose while still maintaining a steady glucose influx, and a higher internal glucose allows fewer enzymes to drive further metabolism. Moreover, the fluxes of the solutions follow an EFM (see Figure 7.2b).

We can now reformulate the problem for only the flux and enzyme levels while keeping the metabolite levels as they are in the table. With the metabolite levels in the first row of the table, we can linearly relate the enzyme and flux levels (with the factors κ_i that have become constants now we have set the internal metabolite concentrations), and thus the extreme rays of the enzyme and flux space will be equal and EFMs, as pointed out above (see also Chapter 5 in [1] and Figure 7.1). Optimization in this space will lead to the optimal flux distributions following an EFM (see Box 7.B for the detailed calculations). As fixing part of the optimal solution should lead to the same optimal solution, this required the flux distribution of the optimization over all variables to follow an EFM, as was indeed the case.

We point out two important aspects, using the network (Figure 7.2) as an example. First, it is convenient to split reversible reactions such that fluxes are always positive. In this case, that means that the reversible reaction from P to $\mathrm{P_{ex}}$ is split into the forward reaction v_2 and the reverse reaction v_4, both of which can have only positive flux. This splitting makes sure that EFMs are the extreme rays of the flux space (see Chapter 5 in [1]). This splitting is purely a mathematical convenience; we still assume this to be one reaction in the biological sense, and therefore the kinetic equations of both the forward and the backward reactions will be exactly the same. Depending on in which direction the flux goes, either one of the reactions will be positive and the other zero. Any solution with both reactions positive is infeasible, but minimizing enzyme levels will never lead to such a solution; therefore we do not need to set additional constrains to avoid this. Second, the feasibility of EFMs can depend on external concentrations. In this network, the biomass reaction (v_{BM}) is the objective flux and there are three EFMs leading to the production of biomass: EFM1 consisting of v_0, v_1, v_2 and v_{BM}, EFM2 consisting of v_0, v_1, v_3 and v_{BM} and EFM3 consisting of v_4, v_3 and v_{BM}. However, if $\mathrm{P_{ex}}$ is absent in the environment, the uptake flux v_4 will always be 0 and therefore EFM3 will not be feasible.

7.4. Calculation of optimal states

We can now use the result that states of maximal enzyme efficiency are reached at an elementary flux mode to calculate optimal states in a metabolic network using the following steps:

1. Enumerate the elementary flux modes that include the objective flux
2. Calculate the minimal enzyme for each EFM scaled to an objective flux of 1
3. Compare the EFMs and select the one with minimal enzyme demands

Step 1 is possible for relatively large networks, although usually not for genome scale metabolic networks. Step 2 is a convex optimization problem as we have seen in Chapter 6 in [1] and Step 3 is straightforward. These three steps together are called Enzyme Flux Cost Minimization, because it is similar to Enzyme Cost Minimization, but while that is focused on fixed fluxes, Enzyme Flux Cost Minimization simultaneously finds the optimal fluxes, enzyme and metabolite levels. In this section we will show the method on the example network of Figure 7.2.

First, we describe the network with the stoichiometric matrix ($\mathbf{N}$) and the concentration vector ($\mathbf{s}$):

$$\mathbf{N} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & -1 \\ 0 & 2 & 0 & 10 & 0 & -100 \\ 0 & -2 & 0 & -10 & 0 & 100 \end{pmatrix}, \qquad \mathbf{s} \equiv \begin{pmatrix} [\mathrm{G}] \\ [\mathrm{P}] \\ [\mathrm{ATP}] \\ [\mathrm{ADP}] \end{pmatrix} \tag{7.6}$$

And with the stoichiometric matrix we can describe the steady state constraints:

Box 7.B Optimal metabolic states in the example network

We minimize the enzyme investment for $v_{\mathrm{BM}} = 1$ with $[\mathrm{P_{ex}}] = 0$ (and therefore $v_4 = 0$ and EFM3 is not feasible) for the network in Figure 7.2 (the optimization problem in Equation (7.1)). Assuming all $h_i = 1$, the objective function $\sum_{i=1}^{r} h_i\ e_i = e_0 + e_1 + e_2 + e_3 + e_{\mathrm{BM}}$. The constraints $v_{\mathrm{BM}} = 1$ and $\mathbf{e}, \mathbf{s} \geq 0$ in Eq. (7.1) are straightforward. The steady state of all internal metabolites (G, P, ADP and ATP) leads to the following equalities (the steady states of ADP and ATP lead to the same equality):

$$\begin{aligned} \text{Steady state ATP} &\Longrightarrow 100\ v_{\mathrm{BM}} = 2\ v_1 + 10\ v_3 \\ \text{Steady state P} &\Longrightarrow 2\ v_1 + v_4 = v_2 + v_3 + v_{\mathrm{BM}} \\ \text{Steady state G} &\Longrightarrow v_0 = v_1 \end{aligned}$$

Substituting $v_{BM} = 1$ and $v_4 = 0$ and solving this set of linear equations, we can write all fluxes as functions of v_2: $v_0 = v_1 = 5 + \frac{5}{11} v_2$ and $v_3 = 9 - \frac{1}{11} v_2$ (there is only one independent flux in this system). This means we can draw the feasible flux space on the v_2 line and we can express the objective function in terms of v_2:

$$\begin{aligned} \sum_{i=1}^{r} h_i e_i &= e_0 + e_1 + e_2 + e_3 + e_{\mathrm{BM}} \\ &= v_0/\kappa_0 + v_1/\kappa_1 + v_2/\kappa_2 + v_3/\kappa_3 + v_{\mathrm{BM}}/\kappa_{\mathrm{BM}} \\ &= (5 + 5/11 v_2)/\kappa_0 + (5 + 5/11 v_2)/\kappa_1 + v_2/\kappa_2 + (9 - 1/11 v_2)/\kappa_3 + 1/\kappa_{\mathrm{BM}} \\ &= \underbrace{(5/\kappa_0 + 5/\kappa_1 + 9/\kappa_3 + 1/\kappa_{\mathrm{BM}})}_{\alpha} + \underbrace{(5/(11\kappa_0) + 5/(11\kappa_1) + 1/\kappa_2 - 1/(11\kappa_3))}_{\beta} v_2 \\ &= \alpha + \beta v_2 \end{aligned} \tag{7.5}$$

The kinetic functions (κ_i) depend on several parameters (external metabolite levels $[\mathrm{G_{ex}}]$, $\mathrm{O_2}$, $[\mathrm{CO_2}]$ and $[\mathrm{P_{ex}}]$, catalytic constants, Michaelis constants and Gibbs free energies) and the variables [G], [P], [ATP] and [ADP]. That means that once we have a set of internal metabolite concentrations $\mathbf{s}$, the enzyme levels in the objective function can be written as a constant times the flux: $e_i = v_i/\kappa_i$, with κ_i a constant. For a set of parameters, α and β are positive or negative depending on the choice of $\mathbf{s}$. It is clear that when we minimize this objective function by adjusting v_2, we will always have an optimum at $v_2 = 0$ (when β is positive) or $v_2 = 99$ (when β is negative). $v_2 = 99$ is the maximum of v_2 because then $v_3 = 9 - \frac{1}{11} v_2 = 0$, and higher values of v_2 would lead to negative values for v_3.
In conclusion, the optimum cannot be at a value of $0 < v_2 < 99$. If there would be an optimum with $0 < v_2 < 99$, we can determine $\mathbf{s}$ and calculate whether $\beta > 0$ to find a lower objective value at $v_2 = 0$ or $v_2 = 99$, contradicting that we started with an optimum. Only if $\beta = 0$ there is a range of optima, but this requires very precise parameter values. $v_2 = 0$ and $v_2 = 99$ correspond to EFMs of this network (Figure 7.2).

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{s} = \mathbf{N}\ \mathbf{v} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & -1 & 1 & -1 \\ 0 & 2 & 0 & 10 & 0 & -100 \\ 0 & -2 & 0 & -10 & 0 & 100 \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_{\mathrm{BM}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{7.7}$$

Now we find the EFMs (for example with EFMtool [6]). It can easily be checked that the following EFMs (denoted

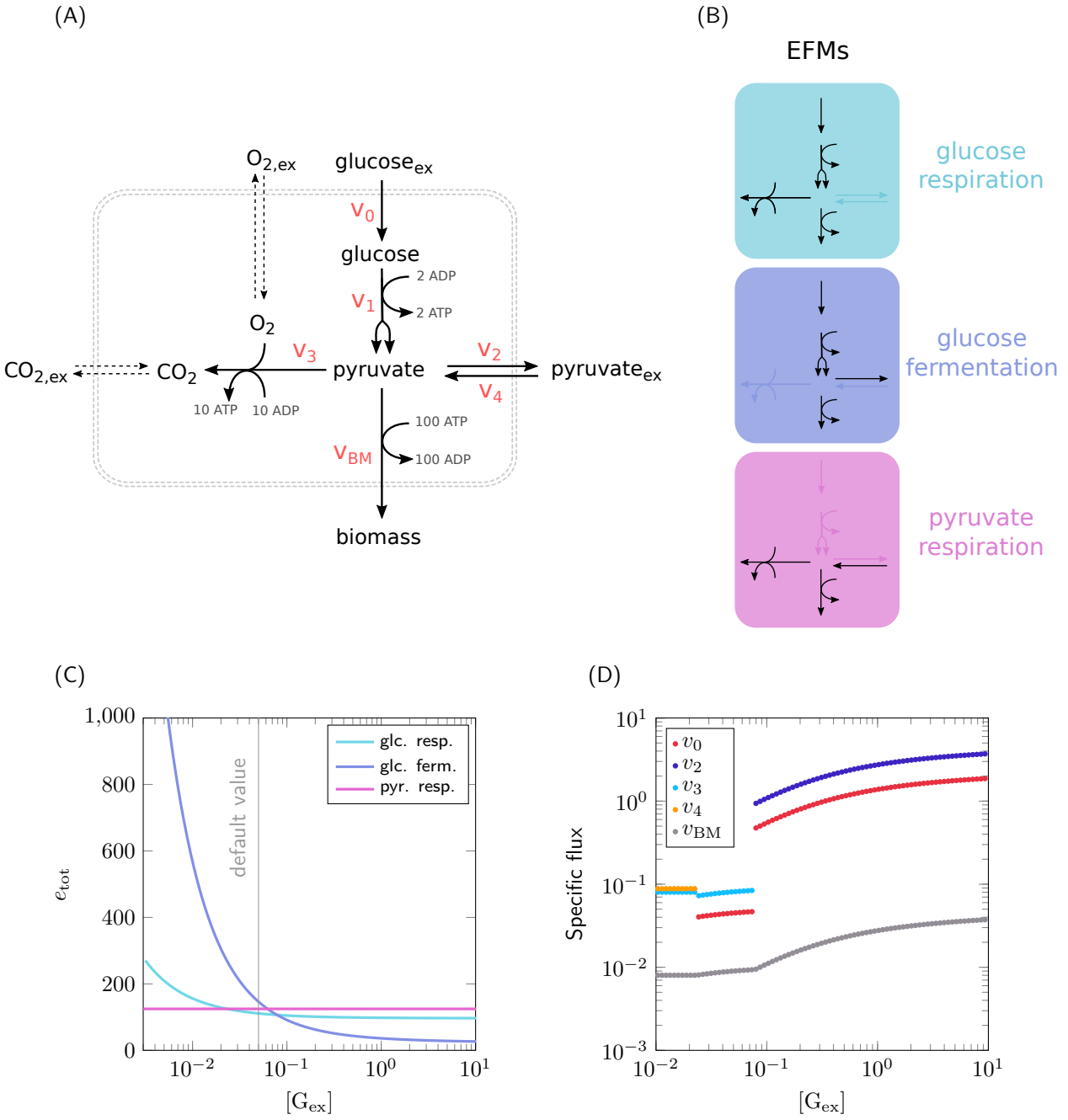


Figure 7.2: States of maximal efficiency in an example model – (A) Example network from Chapter 5 in [1] with added stoichiometry. (B) Three elementary flux modes of this network. (C) Calculated enzyme investment needed for a biomass flux of 1. At a very low concentration of extracellular glucose ($[\mathrm{G_{ex}}]$), EFM3 has the lowest cost. But as we move along the x-axis, at around $[\mathrm{G_{ex}}] = 0.02$ there is a switch to EFM1 and later, at around $[\mathrm{G_{ex}}] = 0.07$, EFM2 becomes the one with the lowest cost. (D) Specific fluxes (flux divided by total enzyme) associated with the optimal EFM for different levels of $\mathrm{G_{ex}}$. Note that v_1 is not shown as it is always equal to v_0. The rates show a discontinuity when there is a switch from one optimal EFM to another.

by vectors $\boldsymbol{f}^{(i)}$) are in the nullspace of the stoichiometric matrix:

$$\boldsymbol{f}^{(1)} = \begin{pmatrix} 5 \\ 5 \\ 0 \\ 9 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{f}^{(2)} = \begin{pmatrix} 50 \\ 50 \\ 99 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{f}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10 \\ 11 \\ 1 \end{pmatrix} \tag{7.8}$$

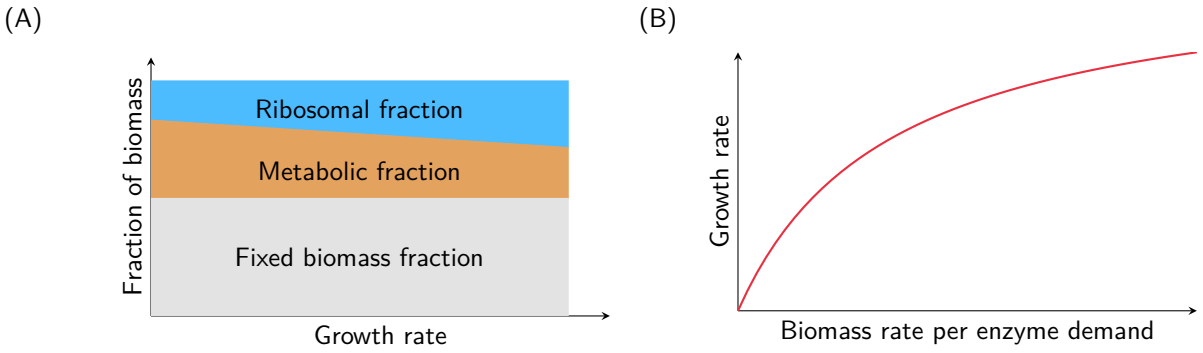


Figure 7.3: Translation of enzyme-specific biomass rate to growth rate – (A) Both from experimental data and a cell-optimization point of view, the ribosomal fraction of the proteome increases with the growth rate, while the metabolic fraction decreases. (B) This leads to a hyperbolic dependency of the growth rate on the biomass production rate per amount of enzymes.

The next step is to perform the convex optimization over the metabolite levels for each one of the three EFMs. Therefore, we express the enzyme levels as a ratio of the flux and the function $f(\mathbf{s})$, using Equation 7.4. Summing over all enzymes, we get a function for the total enzyme cost (level) as a function of fluxes, metabolite concentrations and parameters:

$$e_{\text{tot}} = \sum_i e_i = \sum_i \frac{v_i}{\kappa_i(\mathbf{s})} \,. \tag{7.9}$$

We use the standard parameters (Box 7.A) and replace v_i by the values given by each EFM. We are then left with a convex optimization over the metabolite levels, an Enzyme Cost Minimization problem as in Chapter 6 in [1]. For $[\mathrm{G_{ex}}] = 0.05$ we obtain a total enzyme of 111.1 for EFM1, of 146.3 for EFM2 and 136.5 for EFM3. That means that for these conditions we will conclude that EFM1 is optimal. From the optimization we obtain the metabolite concentrations: $[\mathrm{G}] = 0.08$, $[\mathrm{P}] = 3.93$, $[\mathrm{ATP}] = 0.11$ and $[\mathrm{ADP}] = 20.09$ (that the internal glucose concentration is higher than the external is because we described the transport with regular enzyme kinetics instead of transporter enzyme kinetics, which would have been more realistic). We can next use the rate equations to calculate the enzyme levels from the fluxes and metabolite levels, using the values for the parameters and external concentrations.

We can repeat this procedure for different levels of external concentrations and see that the optimal EFM can change depending on the external concentration (Figure 7.2c). When the optimum shifts to using a different EFM, there is a discontinuity in the fluxes at the external metabolite concentration (Figure 7.2d). Many cells show shifts in metabolic strategies depending on the external conditions and Enzyme Flux Cost Minimization is one way of explaining those shifts.

Above, Enzyme Cost Flux Minimization was used to find the metabolic state with the maximum enzyme efficiency. Although in our calculation we obtain the enzyme concentrations last, it is by enzyme concentrations that cells actually control metabolism. If cells produce enzymes at the concentrations we calculated and reach a steady state, this state will realize the fluxes and metabolite levels that lead to our optimal state.

7.5. Translating enzyme efficiency into cell growth rate

In the section above, we learned how to optimize metabolic states for a maximal overall enzyme efficiency. Why is this quantity relevant? One reason is that overall enzyme efficiency, according to some simple reasoning, determines the cell's growth rate. If microbes compete by growing fast, their fitness is largely determined by their momentary growth rate in their respective environment. In such environments, the biomass/enzyme efficiency will be under selection, which makes it one of the important objective functions in this book. If higher enzyme efficiency means higher growth rate, and if we have a conversion formula for this, we can plot the growth rate of the different EFMs instead of overall enzyme efficiency.

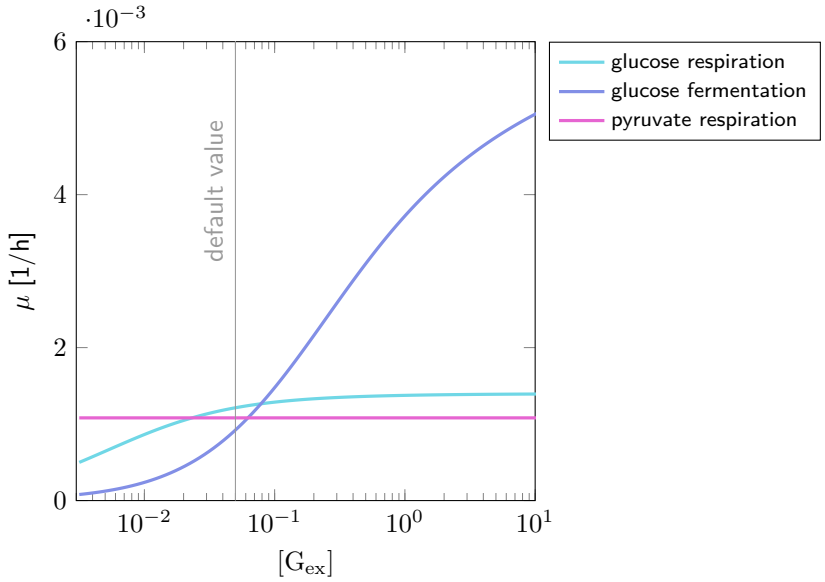


Figure 7.4: Optimal growth rates of the two EFMs for different levels of the external metabolite $\mathrm{G_{ex}}$, computed using Equation (7.10) from the enzyme demands (at a unit biomass production rate) shown in Figure 7.2 (C).

Enzyme-efficient metabolic states allow us to compute specific biomass production rates, i.e. the rate of biomass production per metabolic enzyme invested. If biomass consisted only of enzymes, the ratio "enzyme production rate per total enzyme demand" would give us directly the growth rate. However, biomass does not only consist of metabolic enzymes, but includes ribosomal enzymes, RNA, DNA, lipids, and other compounds. Therefore we need a formula for converting biomass/enzyme efficiency into cellular growth rate.

Mathematically, a cell's growth rate is given by $\mu = v_{\mathrm{BM}}/s_{\mathrm{BM}}$, where v_{BM} is the biomass production rate (biomass produced per cell volume and time) and s_{BM} is the biomass amount per cell volume. If a cell contained nothing but metabolic enzymes (more precisely, the enzymes described in our model), the biomass/enzyme efficiency $\kappa_{\mathrm{BM}} = v_{\mathrm{BM}}/h_{\mathrm{enz}}$ would directly describe the cellular growth rate. Since that is not the case, we need to convert h_{enz} to s_{BM}. The metabolic protein fraction decreases with the growth rate, leading to a hyperbolic dependency of the growth rate on the biomass production rate (Figure 7.3). We may use the empirical approximation $h_{\mathrm{enz}}/s_{\mathrm{BM}} = \kappa_{\mathrm{prot}}(a - b\ \mu)$, where $\kappa_{\mathrm{prot}} = 0.5$ is the fraction of protein mass within the cell dry mass and the parameters $a = 0.27$ and $b = 0.2\,\mathrm{h}$ were fitted to describe the metabolic enzyme fraction in proteomics data, assuming a linear dependence on growth rate [7]. This yields the conversion formula (see also [8]):

$$\mu = \frac{a\ \kappa_{\mathrm{prot}}\ v_{\mathrm{BM}}}{h_{\mathrm{enz}} + b\ \kappa_{\mathrm{prot}}\ v_{\mathrm{BM}}}\,. \tag{7.10}$$

This formula has been used to convert the minimal enzyme cost per biomass flux for different external concentrations in the toy model (Figure 7.2c) to the maximal growth for each EFM (Figure 7.4).

7.6. Application to central metabolism in *E. coli* bacteria

In the previous sections, we saw that finding enzyme-efficient metabolic states can be done by iterating through all possible EFMs and performing the enzyme cost minimization on each one. We demonstrated it on a toy model comprising only 3 EFMs. In Wortel et al. [8], this method was scaled up and applied to a more realistic model covering the central metabolic network, as shown in Figure 7.5A. For this larger network, there are 1566 biomass-generating EFMs. Each reaction is assigned to a single enzyme along with its molecular weight, k_{cat}^{+}, K_{M}, and $\Delta_{\mathrm{r}}G'^{\circ}$, and follows the generalized factorized rate law as in Equation (7.2). These parameters are listed in Appendix section 7.8, and the full procedure for obtaining them is described in Wortel et al. [8], along with other model parameters.

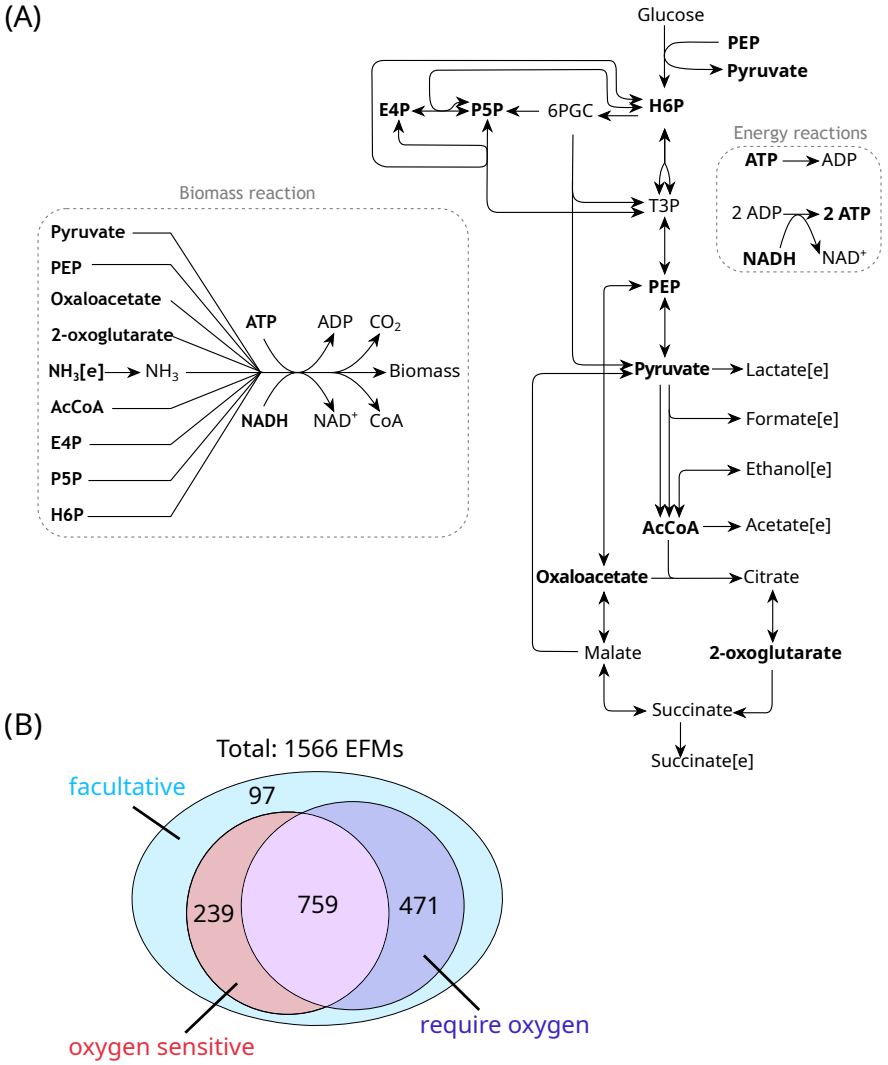


Figure 7.5: Model of central metabolism in *E. coli* bacteria. (A) The metabolic network of the *E. coli* model used by Wortel et al. [8]. Note that only for the purpose of visualization, the network shown here has been condensed by lumping consecutive reactions that are fully coupled (e.g., the reactions between DHAP and PEP are now represented by a single arrow). Furthermore, some groups of metabolites have been merged to a single node: H6P – representing the hexose phosphates G6P, F6P, and FBP; T3P – representing the triose phosphates G3P and DHAP; P5P – representing the pentose phosphates R5P, X5P, and Ru5P. The metabolites that are direct substrates of the biomass reaction are marked in bold. (B) A Venn diagram showing statistics of biomass-producing EFMs in the model and their reliance on oxygen.

First, Wortel et al. [8] wanted to study the effect of environmental conditions on the growth rate of *E. coli*, and see whether the model would be able to recapitulate empirical phenomena. The external glucose concentration was set to 100 mM and oxygen levels were varied between 1 μ and 10 mM. They selected 4 flux modes as representatives (the EFMs *max-gr*, *ana-lac*, and *aero-ace*, and *exp*, which is based on experimentally measured fluxes; the flux distributions are shown in Figure 7.6), and calculated their predicted growth rates in each condition, using Equation (7.10). The results are shown in Figure 7.7. When focusing on a single flux mode, one can see that as the oxgyen level increases so does the growth rate. The increase saturates at some point, which depends on the flux modes and on the kinetic parameters in the model. Indeed, it has long been known that growth rate dependence on a limiting substrate concentration has this specific shape – a relationship generally called the *Monod curve*.

In this specific example, it is interesting to see the Monod curves of the different EFMs, and try to understand the

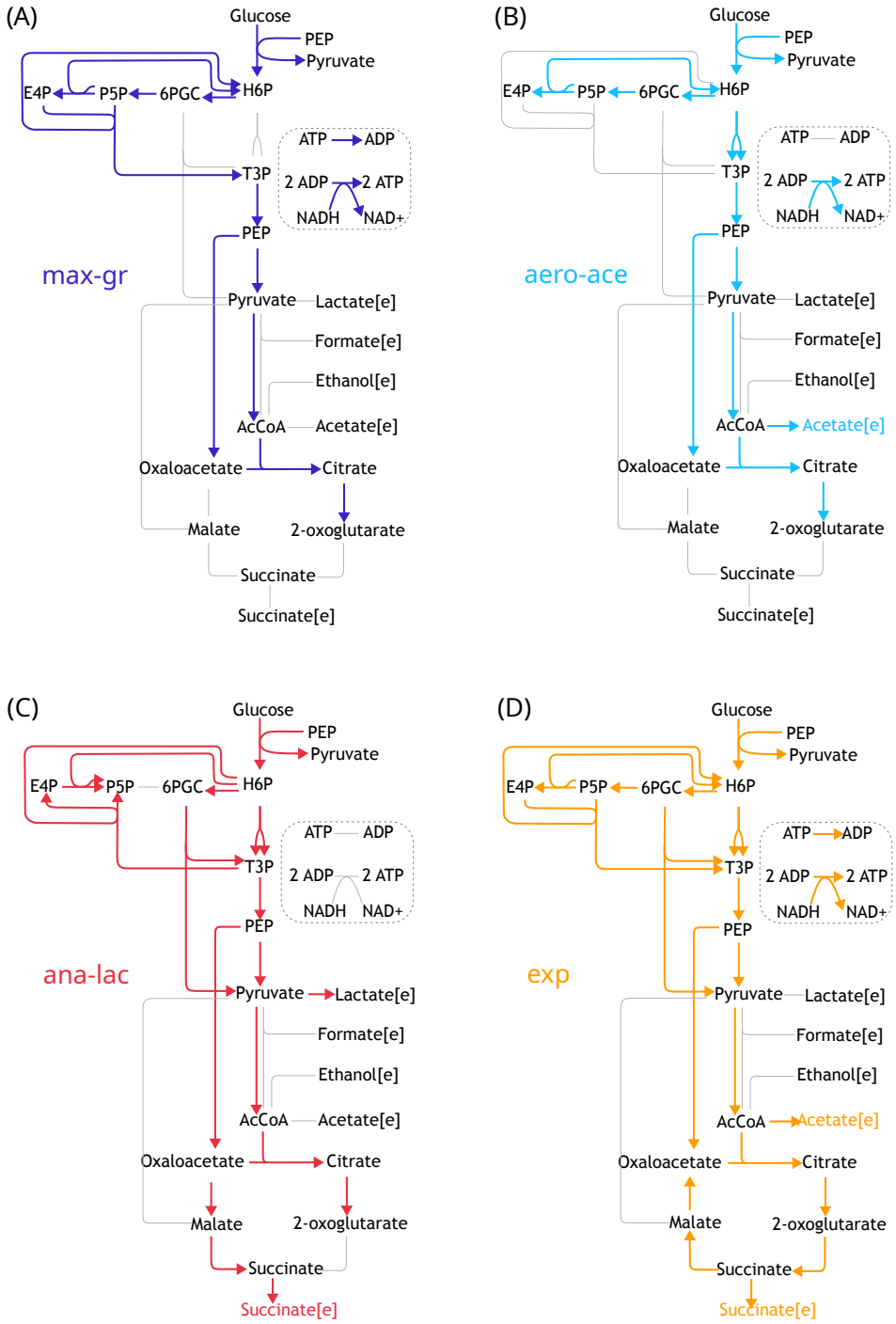


Figure 7.6: The four flux modes chosen for drawing the Monod curves in Figure 7.7: (A) *max-gr* – the EFM with the highest growth rate under the standard conditions chosen in the study, (B) *aero-ace* – an EFM which mixes between respiration and acetate fermentation, (C) *ana-lac* – an EFM that does not require oxygen (i.e. anaerobic) and uses lactate fermentation, and (D) *exp* – which is not an elementary flux mode, but rather one based on the measured flux distribution for *E. coli* growing on minimal media and glucose. The active reactions are highlighted in color (with the flux direction indicated by the arrowhead). The magnitude of each flux is not shown here but can be found in the Supplementary section of Wortel et al. [8]. The biomass reaction is not shown here due to space limitations, but is always active.

differences. First, the EFM called *ana-lac* (red curve), is a flat line. This makes sense because cells that use this EFM do not utilize the oxidative phosphorylation system and therefore do not require oxygen at all for growth. *max-gr*,

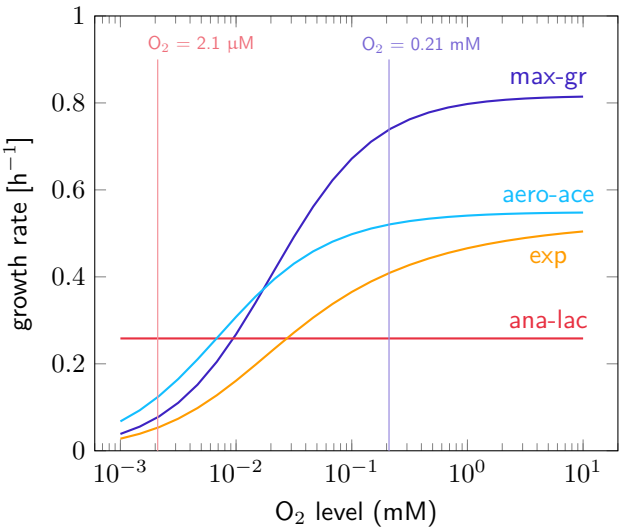


Figure 7.7: Monod curves (cell growth rate as a function of oxygen level) computed using the model shown in Figure 7.5 – Each curve was computed using one of the EFMs and the associated (oxygen-dependent) enzyme demands. The ana-lac strategy (anaerobic growth with lactate secretion) does not use oxygen, therefore it curve is flat.

on the other hand, is very sensitive to the level of oxygen mainly because of the high flux going through oxidative phosphorylation. It is also the EFM with the highest growth rate in standard oxygen levels (0.21 mM), even when taking all the other ∼1500 EFMs into account (not shown here).

Instead of screening only external oxygen levels, we can also screen several model parameters and compute "winning EFMs", their enzyme demands, and the resulting growth rates for our parameter combination. By screening glucose and oxygen concentrations, we obtain the Monod landscape shown in Figure 7.8 (A). Just like in Figure 7.7, there are distinct parameter regions in which optimal growth is reached by specific EFMs. While the *max-gr* EFM remains best when glucose and oxygen levels are high, at low oxygen levels we see a large number of different EFMs, one of them *ana-lac* (see the EFM phase diagram in Figure 7.8 (B)). More results for this model (fluxes plotted in the EFM phase diagram and in flux space, as well as ideal and real enzyme costs for all EFMs), are shown in Appendix Section 7.9.

7.7. Concluding remarks

In this chapter we considered the metabolic network of a cell - and enzyme levels, metabolite concentrations, and fluxes as the state variables - and studied its maximally efficient states. Finding such states can be difficult because fluxes, metabolite concentrations, and enzyme levels are tightly coupled: metabolite concentrations determine enzyme efficiencies, enzyme efficiencies determine optimal enzyme levels, and enzyme levels determine fluxes and metabolite concentrations, which in turn determine enzyme efficiencies. To find an optimal state, all variables need to be optimized at the same time, which is a non-linear optimality problem with (possibly) many local optima. In small toy models, solutions can be found numerically, but for large detailed models, the computational effort becomes enormous. Instead of simplifying the problem (as in the previous chapters) we here used the insight that (in models without extra flux bounds) the optimal solutions must be EFMs.

Recommended readings

- M.T. Wortel, H. Peters, J. Hulshof, B. Teusink, and F.J. Bruggeman. Metabolic states with maximal specific rate carry flux through an elementary flux mode. FEBS Journal, 281(6):15471555, 2014.

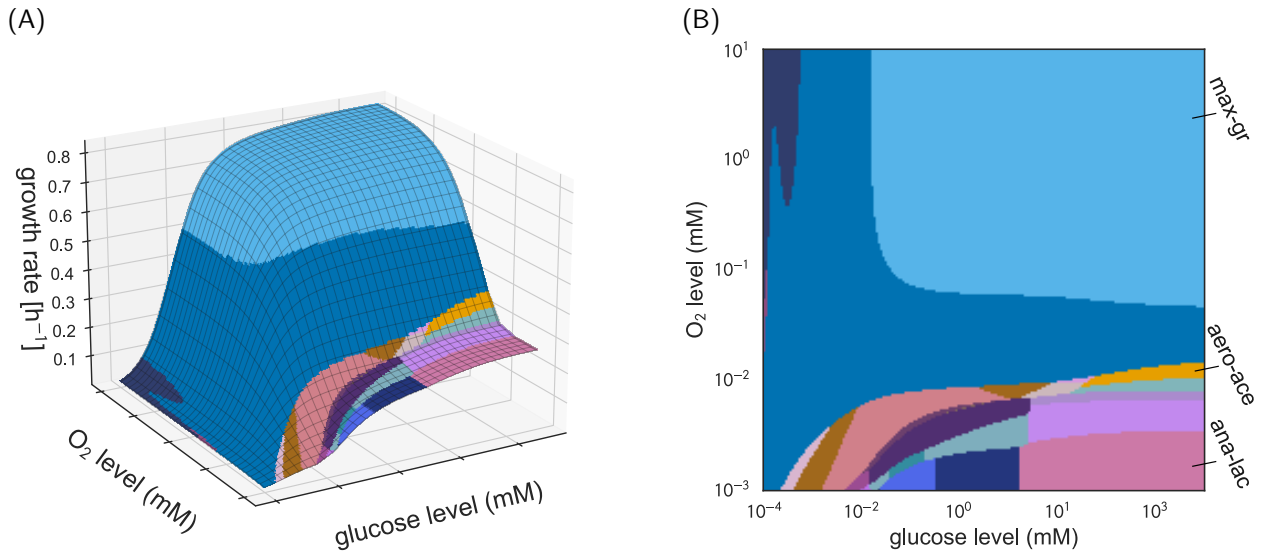


Figure 7.8: Monod landscape – (A) Similar to the 1-dimensional Monod curve (Figure 7.7), the graphics shows the cell growth rate as a function of external glucose and oxygen concentrations, predicted from the *E. coli* model in Figure 7.5. The growth rate of the "winning" EFM – i.e. the one with the highest growth rate under the glucose and oxygen levels matching the x and y values – determines the height of each point. Each color represents the region in which a certain EFM is the "winning" one. (B) EFM phase diagram. The same plot as in (A), seen from above. The "winning EFMs" form a sort of phase diagram. At the boundary between every two regions, the two EFMs lead to the same growth rate (similar to the intersections between curves in Figure 7.7). The EFMs from Figure 7.6 are marked by their names. Note that the colors in this figure do not match the previous colors marking these select EFMs.

- S. Müller, G. Regensburger, and R. Steuer. Enzyme allocation problems in kinetic metabolic networks: Optimal solutions are elementary flux modes. Journal of Theoretical Biology, 347:182190, 2014.
- M.T. Wortel, E. Noor, M. Ferris, F.J. Bruggeman, and W. Liebermeister. Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield. PLoS Computational Biology, 14 (2):e1006010, 2018

Problems

Computer exercises for this chapter can be found on the book website.

Problem 7.1 Effect of oxygen concentration

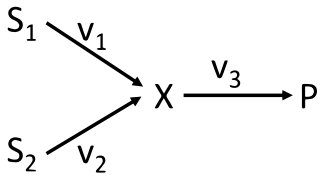
Consider the model in Figure 7.2. What would be the qualitative effect of a change in oxygen concentration on the enzyme cost of the three EFMs and on the choice of the optimal strategy?

Problem 7.2 Effect of external metabolites

Consider the model in Figure 7.2 under standard conditions (Box 7.A and $[\mathrm{G_{ex}}] = 1$, such that EFM2 is optimal, and EFM1 second best (remember that the higher the enzyme cost, the less optimal the EFM). What might happen when we gradually increase the concentration $[\mathrm{P_{ex}}]$? What is the qualitative effect on the enzyme cost of the three EFMs?

Problem 7.3 States of maximal growth rate

Consider the following small toy network:



We want to optimize the specific pathway flux for the production of P (which is v_3/e_{tot}, where we assume all enzymes to have equal costs: i.e. $e_{\mathrm{tot}} = e_1 + e_2 + e_3$) at steady state. We assume mass-action kinetics, meaning the rate is the enzyme concentration times the forward rate constant times the substrate minus the backward rate constant times the product: $v = e(k^+ s - k^- p)$. Unless mentioned otherwise, we use the values $s_2 = 10$, $k_1^+ = 2$, $k_1^- = 1$, $k_2^+ = 3$, $k_2^- = 1$, $k_3^+ = 1$, $k_3^- = 0.1$, $p = 0$ (concentrations are denoted by lower case letters).

(a) Write out the rate equations for all three rates in terms of the parameters and the concentrations.

(b) Give an expression of the total enzyme concentration in terms of fluxes and the metabolite concentrations s_1 and x.

(c) Find the concentration of X for which the specific flux v_3/e_{tot} is maximal for $e_1 = 0$ and $s_2 = 10$, and also give the corresponding value of v_3. HINT: Is is easiest to set $e_{\mathrm{tot}} = 1$ and maximize v_3, replacing e_3 using the equation for the total enzyme cost and the steady state assumption.

(d) Find the concentration of X for which the specific flux v_3/e_T is maximal for $e_2 = 0$ and $s_1 = 10$, and also give the corresponding value of v_3/e_{tot}.

(e) Find the concentration of X for which the enzyme cost is minimal for $e_1 = e_2$ and $s_1 = s_2 = 10$, and also give the corresponding value of v_3/e_{tot}.

(f) What was the best distribution of enzymes from the three options above for $s_1 = 10$?

(g) Find the concentration of X for which the enzyme cost is minimal for $e_1 = 0$ and $s_1 = 50$, and also give the corresponding value of v_3/e_{tot}.

(h) Find the concentration of X for which the enzyme cost is minimal for $e_2 = 0$ and $s_1 = 50$, and also give the corresponding value of v_3/e_{tot}.

(i) Find the concentration of X for which the enzyme cost is minimal for $e_1 = e_2$ and $s_1 = 50$, and also give the corresponding value of v_3/e_{tot}.

(j) What was the best distribution of enzymes from the three options above for $s_1 = 50$?

(k) Interpret the results from this problem in light of the proof shown in this chapter about the optimal specific flux being attained at an EFM.

Appendix sections

7.8. A model of central metabolism in *Escherichia coli*

Metabolite name	Biomass stoichiometric coefficient
AcCoA	-41
ADP	547
2-oxoglutarate	-14
ATP	-547
H_2O	-547
P_i	547
CO_2	2
CoA	41
DHAP	-5
G6P	-4
NAD^+	178
NADH	-178
NH_3	-139
2-oxoglutarateAcetate	-24
PEP	-32
Pyruvate	-38
E4P	-5
R5P	-13

Table 7.2: Stoichiometry of biomass reaction – R70

Reaction ID	EC number	Reaction name	Formula
R1	2.7.1.69	pts	Glucose + PEP $\rightleftharpoons$ G6P + Pyruvate
R2r	5.3.1.9	pgi	G6P $\rightleftharpoons$ F6P
R3	2.7.1.11	pfk	F6P + ATP $\rightleftharpoons$ FBP + ADP
R4	3.1.3.11	fbp	FBP + H_2O $\rightleftharpoons$ F6P + P_i
R5r	4.1.2.13	ald	FBP $\rightleftharpoons$ DHAP + G3P
R6r	5.3.1.1	tim	G3P $\rightleftharpoons$ DHAP
R7ra	1.2.1.12	gap	G3P + NAD^+ + P_i $\rightleftharpoons$ BPG + NADH
R7rb	2.7.2.3	pgk	BPG + ADP $\rightleftharpoons$ 3PG + ATP
R7rc	5.4.2.11 / 5.4.2.12	pgm	3PG $\rightleftharpoons$ 2PG
R8r	4.2.1.11	pgh	2PG $\rightleftharpoons$ PEP
R9	2.7.1.40	pyk	PEP + ADP $\rightleftharpoons$ Pyruvate + ATP
RR9	2.7.9.2	pps	Pyruvate + 2 ATP $\rightleftharpoons$ PEP + 2 ADP + P_i

Table 7.3: Glycolysis

Reaction ID	EC number	Reaction name	Formula
R10a	1.1.1.49	zwf	G6P + NAD^+ $\rightleftharpoons$ 6PGL + NADH
R10b	3.1.1.31	glh	6PGL $\rightleftharpoons$ 6PGC
R10c	1.1.1.44	pgd	6PGC + NAD^+ $\rightleftharpoons$ NADH + CO_2 + Ru5P
R11r	5.1.3.1	rpe	Ru5P $\rightleftharpoons$ X5P
R12r	5.3.1.6	rpi	Ru5P $\rightleftharpoons$ R5P
R13r	2.2.1.1	txt1	R5P + X5P $\rightleftharpoons$ S7P + G3P
R14r	2.2.1.2	tal	G3P + S7P $\rightleftharpoons$ E4P + F6P
R15r	2.2.1.1	txt2	E4P + X5P $\rightleftharpoons$ G3P + F6P
R60	4.2.1.12	edd	6PGC $\rightleftharpoons$ KDPG
R61r	4.1.2.14	eda	KDPG $\rightleftharpoons$ G3P + Pyruvate

Table 7.4: Pentose Phosphate Pathway

Reaction ID	EC number	Reaction name	Formula
R20	2.3.1.54	pfl	Pyruvate + CoA $\rightleftharpoons$ AcCoA + Formate
R21	1.2.4.1 / 2.3.1.12	pdh	Pyruvate + NAD^+ + CoA $\rightleftharpoons$ AcCoA + CO_2 + NADH
R22	2.3.3.1	csn	2-oxoglutarateacetate + AcCoA $\rightleftharpoons$ Citrate + CoA
R23r	4.2.1.3	acn	Citrate $\rightleftharpoons$ iso-Citrate
R24	1.1.1.41	icd	iso-Citrate + NAD^+ $\rightleftharpoons$ 2-oxoglutarate + NADH + CO_2
R25	1.2.4.2	kgd	2-oxoglutarate + NAD^+ + CoA $\rightleftharpoons$ NADH + Succinateinyl-CoA + CO_2
R26r	6.2.1.5	scs	Succinateinyl-CoA + ADP + P_i $\rightleftharpoons$ Succinateinate + ATP + CoA
R27	1.3.5.1	sdh	Succinateinate + ADP + O_2[e] + P_i $\rightleftharpoons$ Fumarate + ATP
R27b	1.3.5.4	frd	Fumarate + NADH $\rightleftharpoons$ Succinateinate + NAD^+
R28r	4.2.1.2	fum	Fumarate $\rightleftharpoons$ Malate
R29r	1.1.1.37	mdh	Malate + NAD^+ $\rightleftharpoons$ 2-oxoglutarateacetate + NADH

Table 7.5: TCA Cycle

Reaction ID	EC number	Reaction name	Formula
R40	4.1.1.31	ppc	PEP + CO_2 ⇌ 2-oxoglutarateacetate + P_i
R41	1.1.1.38	me	Malate + NAD^+ ⇌ Pyruvate + NADH + CO_2
R42	4.1.1.49	ppck	2-oxoglutarateacetate + ATP ⇌ PEP + ADP + CO_2

Table 7.6: Anaplerotic Reactions

Reaction ID	EC number	Reaction name	Formula
R53r	1.1.1.27	ldh	Pyruvate + NADH ⇌ Lactate + NAD^+
R54ra	1.2.1.10	ada	AcCoA + NADH ⇌ Acetaldehyde + NAD^+ + CoA
R54rb	1.1.1.1	adh	Acetaldehyde + NADH ⇌ ETOH + NAD^+
R55a	2.3.1.8	pta	AcCoA + P_i ⇌ Acetyl-P + CoA
R55b	2.7.2.1	ack	Acetyl-P + ADP ⇌ Acetate + ATP

Table 7.7: Redox-associated reactions

Reaction ID	Reaction name	Formula
R80	oxphos	NADH + 2 ADP + 0.5 O_2[e] + 2 P_i ⇌ NAD^+ + 2 ATP + 3 H_2O
R82	atpmain	ATP + H_2O ⇌ ADP + P_i + ATP_{main}

Table 7.8: Oxidative phosphorylation

Reaction ID	Reaction name	Formula
R90	ex_{etoh}	ETOH ⇌ ETOH[e]
R91	ex_{ace}	Acetate ⇌ Acetate[e]
R93	ex_{NH_3}	NH_3[e] ⇌ NH_3
R94	ex_{lac}	Lactate ⇌ Lactate[e]
R95	ex_{suc}	Succinateinate ⇌ Succinateinate[e]
R96	ex_{for}	Formate ⇌ Formate[e]
R97r	ex_{CO_2}	CO_2 ⇌ CO_2[e]

Table 7.9: Membrane Transport Reactions

Reaction ID	k_{cat}^{+} [1/s]	K_{eq} [unitless]	Enzyme molecular weight [Da]
R1	100	N/A	$2.6{\cdot}10^{5}$
R10a	240	N/A	$5.6{\cdot}10^{4}$
R10b	410	N/A	$3.6{\cdot}10^{4}$
R10c	110	N/A	$1.0{\cdot}10^{5}$
R11r	130	2.3	$2.5{\cdot}10^{4}$
R12r	1400	2.3	$1.9{\cdot}10^{4}$
R13r	46	3.7	$7.3{\cdot}10^{4}$
R14r	17	0.9	$3.5{\cdot}10^{4}$
R15r	75	38	$7.3{\cdot}10^{4}$
R20	4800	N/A	$8.5{\cdot}10^{4}$
R21	38	N/A	$2.8{\cdot}10^{5}$
R22	360	N/A	$9.6{\cdot}10^{4}$
R23r	33	0.074	$9.6{\cdot}10^{4}$
R24	110	N/A	$4.6{\cdot}10^{4}$
R25	150	N/A	$1.2{\cdot}10^{6}$
R26r	89	0.52	$7.1{\cdot}10^{4}$
R27	78	N/A	$7.9{\cdot}10^{5}$
R27b	180	N/A	$1.8{\cdot}10^{5}$
R28r	280	4.7	$6.0{\cdot}10^{4}$
R29r	210	$6.1{\cdot}10^{-5}$	$3.2{\cdot}10^{4}$
R2r	320	0.51	$6.2{\cdot}10^{4}$
R3	110	N/A	$1.4{\cdot}10^{5}$
R4	25	N/A	$3.7{\cdot}10^{4}$
R40	120	N/A	$2.0{\cdot}10^{5}$
R41	76	N/A	$6.3{\cdot}10^{4}$
R42	51	N/A	$6.0{\cdot}10^{4}$
R53r	140	$2.1{\cdot}10^{4}$	$3.7{\cdot}10^{4}$
R54ra	0.35	$2.3{\cdot}10^{-3}$	$9.6{\cdot}10^{4}$
R54rb	320	$2.8{\cdot}10^{3}$	$9.6{\cdot}10^{4}$
R55a	91	N/A	$7.7{\cdot}10^{4}$
R55b	59	N/A	$4.3{\cdot}10^{4}$
R5r	8.0	$3.0{\cdot}10^{-4}$	$3.9{\cdot}10^{4}$
R60	250	N/A	$6.5{\cdot}10^{4}$
R61r	80	$9.6{\cdot}10^{-3}$	$2.2{\cdot}10^{4}$
R6r	7800	11	$5.4{\cdot}10^{4}$
R70	99	N/A	$6.0{\cdot}10^{4}$
R7ra	230	0.088	$3.6{\cdot}10^{4}$
R7rb	390	730	$4.1{\cdot}10^{4}$
R7rc	53	0.16	$2.9{\cdot}10^{4}$
R80	$4.0{\cdot}10^{6}$	N/A	$9.1{\cdot}10^{5}$
R82	180	N/A	$6.0{\cdot}10^{4}$
R8r	210	3.5	$4.6{\cdot}10^{4}$
R9	510	N/A	$5.0{\cdot}10^{4}$
R90	100	N/A	N/A
R91	100	N/A	$5.9{\cdot}10^{4}$
R93	100	N/A	$4.5{\cdot}10^{4}$
R94	100	N/A	$5.9{\cdot}10^{4}$
R95	100	N/A	$4.5{\cdot}10^{4}$
R96	100	N/A	$3.1{\cdot}10^{4}$
R97r	100	N/A	N/A
RR9	13	N/A	$8.7{\cdot}10^{4}$

Table 7.10: Kinetic parameters associated with reactions

Reaction ID	Metabolite name	K_{M} [mM]
R1	G6P	0.102
R1	Glucose	0.116
R1	PEP	0.0983
R1	Pyruvate	0.102
R10a	G6P	0.314
R10a	6PGL	0.129
R10a	NAD^+	0.863
R10a	NADH	0.129
R10b	6PGL	0.168
R10b	6PGC	0.0594
R10c	CO_2	0.0626
R10c	6PGC	0.101
R10c	Ru5P	0.0626
R10c	NAD^+	0.0591
R10c	NADH	0.0626
R11r	Ru5P	0.0878
R11r	X5P	0.114
R12r	R5P	1.25
R12r	Ru5P	0.558
R13r	G3P	1.23
R13r	R5P	0.972
R13r	S7P	2.11
R13r	X5P	0.157
R14r	E4P	0.175
R14r	F6P	0.888
R14r	G3P	0.578
R14r	S7P	0.206
R15r	E4P	0.0934
R15r	F6P	0.737
R15r	G3P	1.27
R15r	X5P	0.152
R20	AcCoA	0.0352
R20	CoA	0.0168
R20	Formate	6.35
R20	Pyruvate	2.18
R21	AcCoA	0.159
R21	CO_2	0.159
R21	CoA	0.0629
R21	Pyruvate	0.291
R21	NAD^+	0.0629
R21	NADH	0.159
R22	AcCoA	0.0867
R22	Citrate	0.0756
R22	CoA	0.0756
R22	2-oxoglutarate	0.0287
R23r	Citrate	3.49
R23r	iso-Citrate	2.42
R24	2-oxoglutarate	0.483
R24	CO_2	2.02
R24	iso-Citrate	0.0227
R24	NAD^+	1.06
R24	NADH	0.0119
R25	2-oxoglutarate	0.0670
R25	CO_2	0.108
R25	CoA	0.0927
R25	Succinyl-CoA	0.108
R25	NAD^+	0.0927
R25	NADH	0.108
R26r	CoA	0.00731
R26r	Succinate	0.237
R26r	Succinyl-CoA	0.0105
R26r	ADP	0.0560
R26r	ATP	0.0812
R27	Fumarate	0.0812
R27	O_2[e]	0.371
R27	Succinate	0.0756
R27	ADP	0.371
R27	ATP	0.0270
R27b	Fumarate	0.0201
R27b	Succinate	0.205
R27b	NAD^+	0.0431
R27b	NADH	0.232
R28r	Fumarate	0.314
R28r	Malate	0.615
R29r	Malate	3.19
R29r	2-oxoglutarate	0.0283
R29r	NAD^+	0.460
R29r	NADH	0.0321
R2r	F6P	0.162
R2r	G6P	0.273
R3	F6P	0.116
R3	FBP	0.113
R3	ADP	0.113
R3	ATP	0.141
R4	F6P	0.171
R4	FBP	0.0161
R40	CO_2	0.115
R40	2-oxoglutarate	0.0426
R40	PEP	0.364
R41	CO_2	0.0885
R41	Malate	0.361
R41	Pyruvate	0.0885
R41	NAD^+	0.0691
R41	NADH	0.0885
R42	CO_2	5.21
R42	2-oxoglutarate	0.571
R42	PEP	0.0643
R42	ADP	0.0484
R42	ATP	0.0750
R53r	LACTATE	0.517

Michaelis constants – part I

Reaction ID	Metabolite name	K_{M} [mM]
R53r	Pyruvate	0.0193
R53r	NAD^+	0.517
R53r	NADH	0.0193
R54ra	AcCoA	0.0242
R54ra	Acetaldehyde	1.80
R54ra	CoA	0.00786
R54ra	NAD^+	0.0415
R54ra	NADH	0.113
R54rb	Acetaldehyde	0.0593
R54rb	ETOH	5.49
R54rb	NAD^+	0.169
R54rb	NADH	0.0593
R55a	AcCoA	0.0424
R55a	Acetyl-P	0.313
R55a	CoA	0.0860
R55b	Acetate	3.44
R55b	Acetyl-P	0.154
R55b	ADP	0.402
R55b	ATP	0.0714
R5r	DHAP	0.0782
R5r	FBP	0.204
R5r	G3P	0.0782
R60	6PGC	0.0434
R60	KDPG	0.150
R61r	G3P	0.00146
R61r	KDPG	0.561
R61r	Pyruvate	0.00146
R6r	DHAP	0.0750
R6r	G3P	0.745
R70	AcCoA	0.462
R70	2-oxoglutarate	0.352
R70	BIOMASS	0.0998
R70	CO_2	0.0996
R70	CoA	0.891
R70	E4P	0.0144
R70	G6P	4.31
R70	NH_3	0.0151
R70	2-oxoglutarate	0.00672
R70	PEP	0.169
R70	Pyruvate	0.319
R70	R5P	0.881
R70	ADP	0.0293
R70	ATP	0.342
R70	NAD^+	1.43
R70	NADH	0.0913
R7ra	DPG	0.0576
R7ra	G3P	0.687
R7ra	NAD^+	0.0558
R7ra	NADH	0.0576
R7rb	DPG	0.0426
R7rb	3PG	0.235
R7rb	ADP	0.0426
R7rb	ATP	0.235
R7rc	3PG	0.132
R7rc	2PG	0.0755
R80	O_2[e]	0.116
R80	ADP	0.136
R80	ATP	0.0737
R80	NAD^+	0.0859
R80	NADH	0.116
R82	ATP_{main}	0.130
R82	ADP	0.130
R82	ATP	0.0769
R8r	PEP	0.131
R8r	2PG	0.108
R9	PEP	0.291
R9	Pyruvate	0.0476
R9	ADP	0.218
R9	ATP	8.45
R90	ETOH	0.100
R90	ETOH[e]	0.100
R91	Acetate	0.100
R91	Acetate[e]	0.100
R93	NH_3	0.0999
R93	NH_3[e]	0.100
R94	Lactate	0.100
R94	Lactate[e]	0.100
R95	Succinate	0.100
R95	Succinate[e]	0.100
R96	Formate	0.0999
R96	Formate[e]	0.100
R97r	CO_2	0.0999
R97r	CO_2[e]	0.100
RR9	PEP	0.0934
RR9	Pyruvate	0.0864
RR9	ADP	0.0873
RR9	ATP	0.0350

Michaelis constants – part II

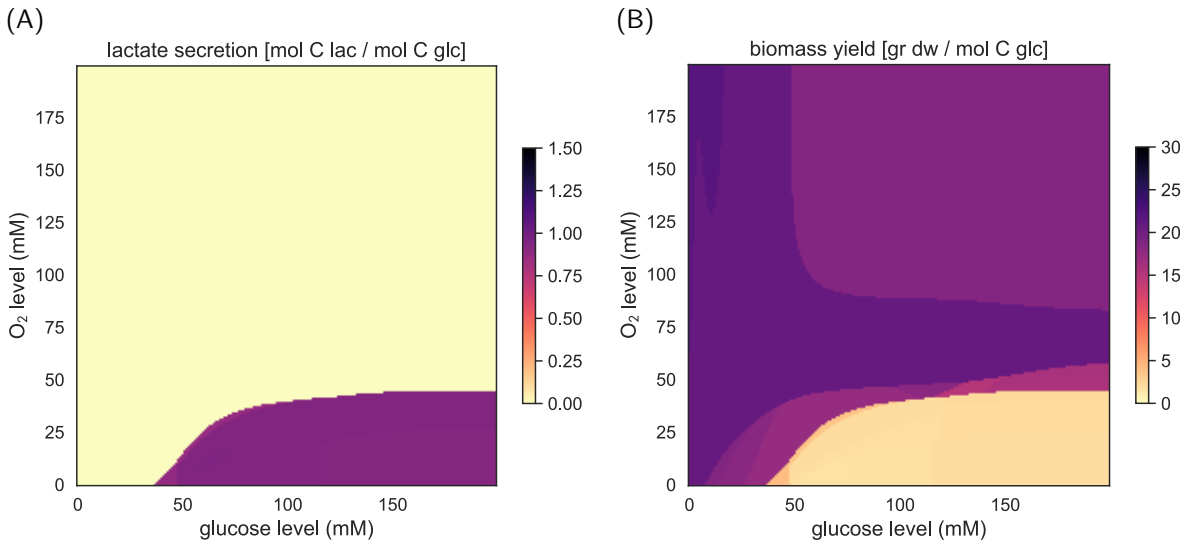


Figure 7.9: Metabolic strategies in the *E. coli* model, depending on external glucose and oxygen concentrations, In the EFM pahse diagram. Each region represents the winning EFM as explained in Figure 7.8. Here, the colors represent the flux in one specific reaction based on the winning EFM in that region. (A) The lactate secretion flux is strikingly equal to 0 in most regions. The only conditions where lactate is secreted is at low oxygen and medium/high glucose concentrations. (B) The biomass yield is, in general, high if and only if lactate is not secreted. This makes sense because the carbon coming from the glucose is often the limiting nutrient for growth, and there is a trade-off between using it for biomass versus fermentation products such as lactate. Interestingly, the region with high glucose and high oxygen levels (upper right quadrant) is occupied by an EFM that doesn't achieve the highest possible yield (i.e. *max-gr*). In low glucose and high oxygen, or in medium oxygen levels, the winning EFMs are the ones with relatively higher biomass yields.

7.9. More results for the *E. coli* central metabolism model

This section contains additional results for the *E. coli* central metabolism model from Chapter 7 in [1], in particular, fluxes plotted in the EFM phase diagram and in flux space, as well as ideal and real enzyme costs for all EFMs.

Each point in the Monod landscape in Figure 7.8 (A) corresponds to a state of the model, and the calculations that lead to the growth rate and the "winning EFM" shown yield a full description of this state, including all fluxes, metabolite concentrations and enzyme levels. These data can be explored and visualized in many ways. For illustration, Figure 7.10 shows a variant of Figure 7.8 (A) in the horizontal axes do not describe the external concentrations of glucose and oxygen, but their uptake rates, and the vertical axes shows the biomass production rate. Since uptake rates are a function of external concentrations, and the growth rate directly depends on the biomass production rate, we might have expected that this yields the same picture, just a bit stretched along each of the axes. However, the picture looks very different: instead of forming a continuous surface, the points now fall on disconnected rays, apparently with one ray for each colored region of the surface. In fact, when looking at the picture closely, we can see that each of the winning EFMs gives rise to exactly one ray. But this, after all, is logical. In the new plot, all axes refer to reaction rates, and for each EFM all rates come in fixed ratios, giving rise to a ray. So, if our solutions are EFMs, this picture cannot be continuous – in line with the fact that, in the original Monod landscape 7.8 (A), when moving from one region to the other one, one would notice a discrete jump of the reaction rates. But why is the new picture not a continuos surface, if uptake rates depends smoothly on external metabolite concentrations? In fact, they do not only depend on these concentrations, but also on resource allocation to the transporter. If this allocation shows a discrete jump (again, when moving from one region to another one), then also the rate shows a jump. The comparison between the two plots shows us what we gain by considering enzyme kinetics as compared to a pure stoichiometric model. With the biomass rate as a proxy for cell growth, each EFM defines fixed ratios between this growth rate and each of the metabolic fluxes, including the uptake rates. When continuously scaling an EFM, the glucose uptake,

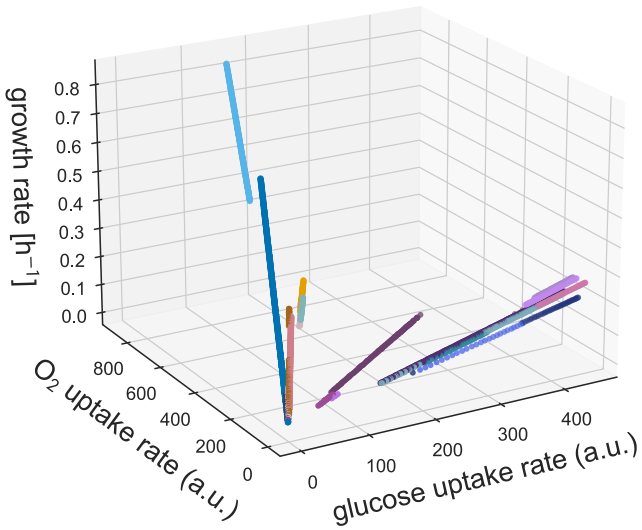


Figure 7.10: Proportional scaling of fluxes within each EFM – In the diagram with glucose uptake, oxygen uptake, and biomass production rate on the axes, each colored line corresponds to one EFM, and shows the possible combinations of fluxes obtained from the model behind Figure 7.8 (which also shares the EFM colors with this figure). Importantly, here the x and y axes represent uptake rates and not substrate concentrations. Therefore, as expected, each EFM yields a straight line (because of the proportional scaling of different fluxes for each EFM). Since – according to our reasoning – optimal flux distributions must be EFMs, only these combinations of fluxes are actually possible. When glucose and oxygen concentrations are varied smoothly in Figure 7.8, the corresponding movement in this plot would be along the lines and sometimes, jumps between different lines (when the system moves from one region to another one in Figure 7.8).

oxygen uptake, and biomass production will scale proportionally. So the rays in Figure 7.10 reflect what we can know about possible metabolic fluxes based on network structure alone; but to get to the Monod landscape, as a function of concentrations, we had to use kinetic information and a extra principle of economical enzyme usage.

The phase diagram of "winning EFMs" can also be used to visualize other (optimized) quantities as functions of glucose and oxygen concentrations. Figure 7.9 shows as example the (biomass-specific) lactate secretion (showing that also a number of other winning EFMs, apart from *ana-lac*, secrete lactate) and the biomass yield on glucose.

Finally, a statistics over all EFMs shows that the range of possible enzyme demands per biomass production rate is quite large: as shown in Figure 7.11, they vary over more than two orders of magnitude, making some EFMs a hundred-fold more enzyme-expensive than others. The same plot also shows how enzyme costs depend on the fact that enzymes do not operate at their full capacity (reaching their k_{cat} value), but at best at the enzyme efficiencies predicted by enzyme cost minimization. For our *E. coli* model and the aerobic glucose conditions studied, if all enzymes could operate at their k_{cat} values, this would decrease to overall enzyme demand by a factor of at least 1.4, or maximally 4.7, depending on the EFM in question. But still, in this case, for determining enzyme costs the choice of the right EFM (even assuming "ideal" enzymes) is much more important than considering the actual, "non-ideal" way in which enzymes operate. But this may not always hold: under low-oxygen conditions, the enzyme demands of some EFMs may increase much more drastically.

Solutions to problems

Problem 7.1 (Effect of oxygen concentration)

The oxygen concentration affects only the rate v_3, and an increase in oxygen increases this rate for the same enzyme concentration. Since EFM2 does not contain v_3, the enzyme cost of this EFM will not change. EFM1 and EFM3

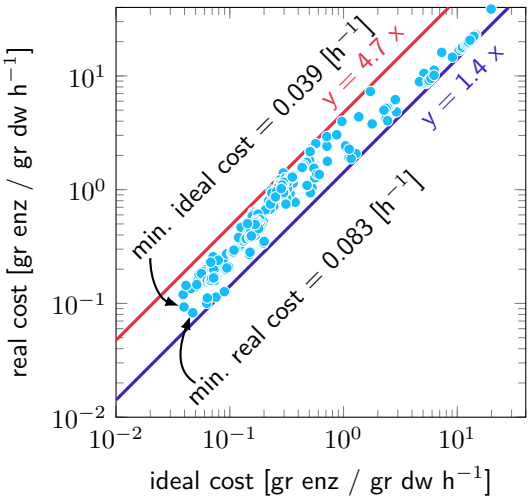


Figure 7.11: Ideal and real enzyme costs of elementary flux modes – For each EFM (shown as a cyan dot), the ideal enzyme cost per biomass production rate (i.e. assuming that all the enzymes are saturated) is compared to the actual cost (calculated using Enzyme Cost Minimization, assuming standard aerobic glucose conditions). The costs span a wide range from the most enzyme-efficient EFMs on the lower left to the least enzyme-efficient ones on the upper right. For different EFMs, the ratio of actual and ideal costs varies between 1.4 and 4.7. Here, the EFM with the minimal actual cost is among the top 5 in terms of ideal cost.

benefit from an increase of the oxygen concentration, because they will have to invest less enzyme in v_3 to obtain the same rate. Therefore, those EFMs can become more beneficial and the optimal EFM could shift to one of those EFMs.

Problem 7.2 (Effect of external metabolites)

Increasing $[\mathrm{P_{ex}}]$ will benefit EFM3 and decrease the benefit of EFM2 (because for EFM2 $[\mathrm{P_{ex}}]$ will inhibit reaction v_2 and therefore more enzyme is needed for reaction v_2 and the enzyme cost of EFM2 will increase. Qualitatively, with increasing $[\mathrm{P_{ex}}]$, EFM2 might become more expensive, and either EFM1 or EFM3 will become beneficial, or both at different $\mathrm{P_{ex}}$ concentrations, depending on the kinetics of the reactions.

Problem 7.3 (States of maximal growth rate)

(a) $v_1 = e_1(k_1^+ s_1 - k_1^- X)$, $v_2 = e_2(k_2^+ s_2 - k_2^- x)$ and $v_3 = e_3(k_3^+ x - k_3^- p)$

(b) $e_{\mathrm{tot}} = \frac{v_1}{(2s_1 - x)} + \frac{v_2}{(30-x)} + \frac{v_3}{x}$

(c) When $e_1 = 0$, v_1 is also 0 and to achieve steady state $v_2 = v_3$ and using the rate equations and filling in the parameters we get $e_{\mathrm{tot}} = \frac{v_3}{(30-x)} + \frac{v_3}{x}$. We now set $e_{\mathrm{tot}} = 1$ to obtain $v_3 = (1 - \frac{v_3}{30-x})x$ which we can rewrite to $v_3 = \frac{x}{1+\frac{x}{30-x}}$. We can take the derivative to x and set it equal to 0 to find the optimum, which leads to $x = 15$ and $v_3/e_{\mathrm{tot}} = \frac{15}{2}$. Note that in this specific case we did not need to set $e_{\mathrm{tot}} = 1$ and could have maximized v_3/e_{tot} directly, but in general this does not always work. In the rest of the answers we assume we set $e_{\mathrm{tot}} = 1$ and therefore $v_3 = v_3/e_{\mathrm{tot}}$.

(d) $e_2 = 0$ implies $v_2 = 0$ and filling in $s_1 = 10$ leads to $v_3 = (1 - \frac{v_3}{20-x})x$, which can be rewritten to $v_3 = \frac{x}{1+\frac{x}{20-x}}$. This is optimal when $x = 10$ and $v_3/e_{\mathrm{tot}} = 5$.

(e) $e_1 = e_2$ implies $\frac{v1}{20-x} = \frac{v2}{30-x}$. From the steady state we know that $v_2 = v_3 - v_1$. Filling this in and solving for v_1 leads to $v_1 = v_3 \frac{x-20}{x-50}$. From the total enzyme and by replacing e_1 by e_2 we get $e_3 = 1 - 2\frac{v_1}{20-x}$. Putting this in the equation for v_3 and using the previous equality to replace v_1 leads to: $v_3 = (1 + \frac{v_3}{x-25})x$. Solving this for v_3 gives $v_3 = x - \frac{x^2}{25}$, which is optimal for $x = 12.5$ with $v_3/e_{\mathrm{tot}} = \frac{25}{4}$.

(f) It was optimal to invest all enzyme in e_2 and none in e_1, because that lead to the highest specific flux v_3 (namely $v_3 = 7.5$).

(g) Since $v_1 = 0$, S_1 is not involved in any reaction and the solution is the same as above, $x = 15$ and $v_3/e_{\mathrm{tot}} = \frac{15}{2}$.

(h) Similar calculations as above but now with $s_1 = 50$ lead to $v_3 = \frac{x}{1+\frac{x}{100-x}}$, which is optimal when $x = 50$ and gives $v_3/e_{\text{tot}} = 25$.

(i) Similar calculations as above but now with $s_1 = 50$ lead to $v_3 = x - \frac{x^2}{65}$. v_3/e_{tot} is maximal at $x = 32.5$ and takes the value 16.25.

(j) Now s_1 increased the optimal strategy would be to invest all enzymes in e_1, and have $e_2 = 0$, because that leads to the highest specific flux of v_3, namely $v_3/e_{\text{tot}} = 25$.

(k) The two EFMs in this pathway that produce P are $v_1 = v_3$ with $v_2 = 0$ and $v_2 = v_3$ with $v_1 = 0$. In the problem we saw when we optimize the specific flux, we always obtained one of those EFMs as the best solution, from the options that we tested. Therefore, these results are in agreement with the proof outlined in this chapter.

Bibliography

[1] The Economic Cell Collective, editor. *Economic Principles in Cell Biology*. Free online book, 2023. doi: 10.5281/zenodo.8156386.

[2] Avi Flamholz, Elad Noor, Arren Bar-Even, Wolfram Liebermeister, and Ron Milo. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. U. S. A.*, 110(24):10039–10044, June 2013. doi: 10.1073/pnas.1215283110.

[3] Elad Noor, Avi Flamholz, Arren Bar-Even, Dan Davidi, Ron Milo, and Wolfram Liebermeister. The protein cost of metabolic fluxes: Prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.*, 12(11): e1005167, November 2016. doi: 10.1371/journal.pcbi.1005167.

[4] M.T. Wortel, H. Peters, J. Hulshof, B. Teusink, and F.J. Bruggeman. Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS Journal*, 281(6):1547–1555, 2014.

[5] Stefan Müller, Georg Regensburger, and Ralf Steuer. Enzyme allocation problems in kinetic metabolic networks: Optimal solutions are elementary flux modes. *Journal of Theoretical Biology*, 347:182–190, 2014.

[6] Marco Terzer and Jörg Stelling. Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics*, 24(19):2229–2235, 2008. ISSN 1367-4803. doi: 10.1093/bioinformatics/btn401.

[7] Matthew Scott, Carl W Gunderson, Eduard M Mateescu, Zhongge Zhang, and Terence Hwa. Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007):1099–1102, 2010. doi: 10.1126/science.1192588.

[8] Meike T Wortel, Elad Noor, Michael Ferris, Frank J Bruggeman, and Wolfram Liebermeister. Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield. *PLoS Comput. Biol.*, 14(2):e1006010, February 2018. doi: 10.1371/journal.pcbi.1006010.