Bachelor Wiskunde

Bachelor thesis

A bacterial growth law modelled by EGMs

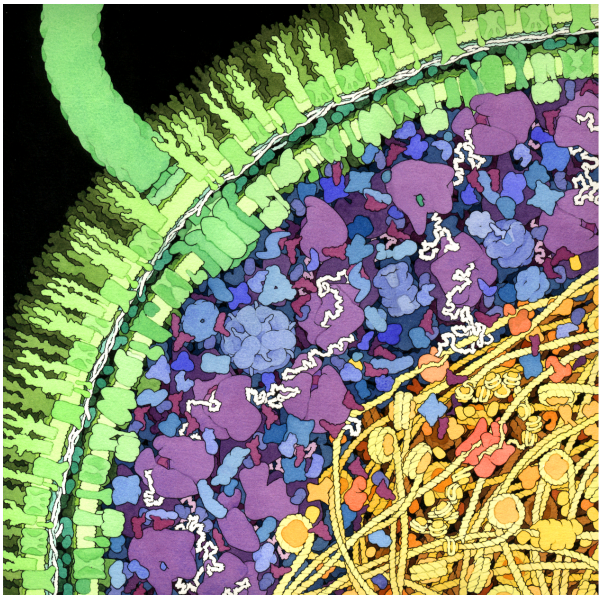
by

Wander van der Meer

15 July, 2017

Supervisor: dr. Bob Planqué

Second examiner: dr. Frank Bruggeman



Department of Mathematics
Faculty of Sciences

VU UNIVERSITY AMSTERDAM

Abstract

In biology, it is widely known that the ribosomal fraction of all the enzymes, including the ribosomes, in a cell is affinely connected to the cell's growth rate. It is however poorly understood why this is the case. The theory of EGMs, minimal networks connecting metabolites, enzymes and ribosomes, does give plenty insight in this so-called bacterial growth law. A redundant network can, through the use of this theory, be reduced to a minimal network in which the optimal growth rate is obtained. In the relatively simple case of five metabolites and seven enzymes, it is shown what the minimal network, consisting of equal metabolites and enzymes, is and that the ribosomal fraction of all the enzymes is affinely connected to the growth rate.

Title: A bacterial growth law modelled by EGMs
Author: Wander van der Meer, wanderv.dmeer@gmail.com, 2537538
Supervisor: dr. Bob Planqué
Second examiner: dr. Frank Bruggeman
Date: 15 July, 2017

Cover image: Artist impression of the interior of an E. coli bacteria by David S. Goodsell, 1999

Department of Mathematics
VU University Amsterdam
De Boelelaan 1081, 1081 HV Amsterdam
http://www.math.vu.nl/

Contents

1 Introduction

For around sixty years, biologists have known about the relation between the growth rate of a bacterial cell and one of the so-called bacterial growth laws. It has been observed that the fraction of ribosomes in a cell is an affine function of the growth rate [6] [8]. This observation will from now on be called the *bacterial growth law*. There are other growth laws, but those are not taken into consideration.

This bacterial growth law is not fully understood. Various models have been proposed to explain it [7] [2], but these are incorrect as they all try to describe kinetics whereas their first assumption is that there are no kinetics in the cell. The kinetics in the cell occur due to dilution in the cell. We will see later on that this dilution is essentially the growth rate. Therefore, neglecting kinetics will be followed by a skewed definition of the growth rate, which is meaningless.

The metabolism of (bacterial) cells is most easily modeled by grouping their contituents into three subgroups: the metabolites, the enzymes and the ribosomes [1]. The metabolites are converted into other metabolites by the enzymes through a chemical reaction. This is done until the metabolites become amino acids, which are a subgroup of the metabolites. These amino acids are converted by the ribosomes to constitute the enzymes, including the ribosome. This way, the interior of the cell is capable of making everything in its interior. A detailed map showing how the various metabolites and enzymes work together can be found at [4].

The theory of Elementary Growth Modes, in short EGMs, tries to remedy aforementioned shortcoming. An EGM is a minimal type of network, consisting of metabolites and an equal amount of enzymes in which the maximal growth rate is achieved. In this framework, some interesting features arise. One of them is that knowing the growth rate in steady state means that the composition of the entire cell is known.

The goal of this thesis is to understand the theory of EGMs and use this to show the bacterial growth law in a relatively simple, artificial pathway.

Most of the theory discussed in this thesis is based on *Growth and Metabolism* by dr. B. Planqué and dr. F. Bruggeman, which is not published.

2 Theory

First, we will first discuss in a basic way some of the principles of growth rate and stoiciometry. Then, we will introduce degrees of freedom because this simple view doesn't allow us to optimise the growth rate. We can tweak these degrees of freedom to optimize the growth rate, from which a framework will follow in which we can calculate metabolite and enzyme concentrations in metabolic pathways.

2.1 Basic

Let us denote the copy number of a compound k as n_k. The concentration of compound k is then defined as $c_k = n_k/V$, where $V = V(\boldsymbol{n})$ is the volume of the cell, as a function of the copy number of the various compounds $\boldsymbol{n}$ in the cell. We assume that the cell is well-stirred and that all changes copy number of the compounds are due to chemical reactions which take place in the cell, so that

$$\dot{n}_k = V \sum_j N_{kj} v_j(\boldsymbol{c}) \quad \text{for all } k. \tag{2.1}$$

Here N denotes the stoichiometry matrix, which tells us the amounts of molecules that are needed for and produced by every chemical reaction, and v_j denotes the j-th reaction rate, as a function of the concentrations $\boldsymbol{c}$. Since $n_k = Vc_k$ we can rewrite this as

$$\dot{V} c_k + V \dot{c}_k = V \sum_j N_{kj} v_j(\boldsymbol{c}) \quad \text{for all } k. \tag{2.2}$$

At balanced growth, all concentrations are constant in time, so

$$\dot{V} c_k = V \sum_j N_{kj} v_j(\boldsymbol{c}) \quad \text{for all } k. \tag{2.3}$$

We can use the chain rule to rewrite $\dot{V} = \sum_l \frac{\partial V}{\partial n_l} \dot{n}_l$ such that

$$\left(\sum_l \frac{\partial V}{\partial n_l} V \sum_j N_{lj} v_j(\boldsymbol{c}) \right) c_k = V \sum_j N_{kj} v_j(\boldsymbol{c}) \quad \text{for all } k. \tag{2.4}$$

The right-hand side of eq. (2.4) is constant in time, and as such, the left-hand side as to be as well. In particular, $\frac{\partial V}{\partial n_k}$ must be independent of time for all k, such that they cannot dependent explicitly on $\boldsymbol{n}$, so we can write

$$\frac{\partial V}{\partial n_k} = \rho_k \quad \text{for all } k, \tag{2.5}$$

for some constants ρ_k. Therefore, we must set $V = \sum_k \rho_k n_k$, in balanced growth. A self-evident interpretation of this definition of V is that ρ_k is the molar volume of compound k and that volume increases to keep osmotic pressure constant as copy numbers of the various compounds in the cell grow through inflow of water.

With this definition of V, we also have

$$\sum_k \rho_k c_k = \frac{1}{V} \sum_k \rho_k n_k = 1, \tag{2.6}$$

at balanced growth. We define the growth rate as the relative increase of volume,

$$\mu(t) = \frac{\dot{V}}{V}. \tag{2.7}$$

Since $\boldsymbol{\rho} \cdot \boldsymbol{c} = 1$ we write

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\rho} \cdot \boldsymbol{c} &= \sum_k \rho_k \dot{c}_k \\ &= \sum_k \rho_k \left(\sum_j N_{kj} v_j(\boldsymbol{c}) - \frac{\dot{V}}{V} c_k \right) \\ &= \boldsymbol{\rho} \cdot N\boldsymbol{v}(\boldsymbol{c}) - \boldsymbol{\rho} \cdot (\mu(t)\boldsymbol{c}) \\ &= \boldsymbol{\rho} \cdot N\boldsymbol{v}(\boldsymbol{c}) - \mu(t) \end{aligned} \tag{2.8}$$

from which we can conclude that

$$\mu(t) = \boldsymbol{\rho} \cdot N\boldsymbol{v}(\boldsymbol{c}). \tag{2.9}$$

However, we are now unable to optimize the growth rate through this equation. Simply put, this is because there are no degrees of freedom in the model. The model is fully described by all the concentrations, the molar volume of every compound, the stoichiometry matrix and the various reaction rates, making the growth rate a property of the system in which everything is fixed when considering steady state.

2.2 Advanced

2.2.1 Whole-cell model

To add degrees of freedom, we seperate the enzymes and the ribosomes from the metabolites, so $\boldsymbol{c} = (\boldsymbol{x}, \boldsymbol{e}, r)$, see also [5]. There are m different metabolites with concentrations $x_1, x_2, \ldots, x_m$, n different enzymes with concentrations $e_1, e_2, \ldots, e_n$ and one type of ribosome with concentration r. The dynamical system for all concentrations is still given by eq. (2.2), rewritten as

$$\dot{\boldsymbol{c}} = N\boldsymbol{v}(\boldsymbol{c}) - \mu\boldsymbol{c}. \tag{2.10}$$

We divide N into four different matrices corresponding to metabolism and the synthesis of enzymes and ribosomes through

$$N = \begin{pmatrix} P & -M \\ O & I \end{pmatrix}, \tag{2.11}$$

where P is an $m \times n$ matrix, $m < n$, M is an $m \times (n+1)$ matrix with non-negative entries, O is an $(n+1) \times n$ matrix with entries all equal to zero and I is the $(n+1) \times (n+1)$ identity matrix. The matrix P corresponds to the stoichiometric matrix in a metabolic pathway, i.e. in column j the metabolites corresponding to the rows with the negative entries are converted into metabolites corresponding to the rows with the positive entries, catalysed by enzyme j. The matrix M has on the jth column the quantities of metabolites, corresponding to their row numbers, needed to construct enzyme j.

The vector of reaction rates $\boldsymbol{v} = (v_1, \ldots, v_{2n+1})$ is also split up into the metabolic reaction rates $v_1, \ldots, v_n$, enzyme synthesis rates $(w_1, \ldots, w_n)$ and the ribosome synthesis rate $w_r = w_{n+1}$, which are all functions of the concentrations $\boldsymbol{c}$. Each reaction has an enzyme associated to it, which not only makes the reaction possible but also catalyses it. We will therefore write for the metabolic reaction rates

$$v_j(\boldsymbol{c}) = e_j f_j(\boldsymbol{x}) \quad \text{for } j = 1, \ldots, n. \tag{2.12}$$

For enzyme synthesis, we know that they are synthesised by the ribosome, which can only synthesise one enzyme or ribosome at a time. We therefore write α_j as the fraction of the ribosomes that synthesise enzyme j and α_{n+1} as the fraction of the ribosomes that synthesise ribosomes. Evidently, this means that

$$\alpha_1 + \cdots + \alpha_{n+1} = 1, \tag{2.13}$$

assuming that the ribosomes always synthesise something. These α_j are the aforementioned degrees of freedom. For the enzyme synthesis rates we write

$$w_j(\boldsymbol{c}) = r\alpha_j g_j(\boldsymbol{x}) \quad \text{for } j = 1, \ldots, n+1. \tag{2.14}$$

One more step is needed, and that is subdividing the molar volumes $\boldsymbol{\rho}$ into $\rho_1, \ldots, \rho_m$ for the m metabolites, $\sigma_1, \ldots, \sigma_n$ for the n enzymes and $\sigma_r = \sigma_{n+1}$ for the ribosome. The full model, using eqs. (2.2) and (2.9), can now be written as follows:

$$\begin{cases} \dot{x}_i &= \sum_{j=1}^{n} P_{ij} e_j f_j(\boldsymbol{x}) \;-\; \sum_{j=1}^{n+1} M_{ij} r\alpha_j g_j(\boldsymbol{x}) \;-\; \mu x_i, \quad \text{for } i = 1, \ldots, m, \\ \dot{e}_j &= \qquad\qquad r\alpha_j g_j(\boldsymbol{x}) \;-\; \mu e_j, \quad \text{for } j = 1, \ldots, n, \\ \dot{r} &= \qquad\qquad r\alpha_{n+1} g_{n+1}(\boldsymbol{x}) \;-\; \mu r, \\ \mu &= \sum_{i=1}^{m} \rho_i \sum_{j=1}^{n} P_{ij} e_j f_j(\boldsymbol{x}) + \sum_{j=1}^{n+1} \left(\sigma_j - \sum_{i=1}^{m} \rho_i M_{ij}\right) r\alpha_j g_j(\boldsymbol{x}). \end{cases} \tag{2.15}$$

In the next sections, some characterisations and properties of balanced growth are treated.

2.2.2 At balanced growth all copy numbers increase at the same rate

From the steady state equations

$$0 = \dot{c}_k = \frac{\mathrm{d}}{\mathrm{d}t}\frac{n_k}{V} = \frac{\dot{n}_k}{V} - \frac{\dot{V} n_k}{V^2} = \frac{\dot{n}_k}{V} - \mu\frac{n_k}{V} \quad \text{for all } k, \tag{2.16}$$

it follows that

$$\mu = \frac{\dot{n}_k}{n_k} \quad \text{for all } k, \tag{2.17}$$

in other words, the copy number of all compounds increases at the same rate, namely the growth rate, in steady state.

2.2.3 Relation between the growth rate and the enzyme and ribosome concentrations

In steady state, it follows from eq. (2.15) that

$$r\alpha_{n+1}g_{n+1}(\boldsymbol{x}) = \mu r, \quad r\alpha_j g_j(\boldsymbol{x}) = \mu e_j, \quad \text{for } j = 1, \ldots, n, \tag{2.18}$$

from which it follows that, first, $\mu = \alpha_{n+1}g_{n+1}(\boldsymbol{x})$, and by adding all $n+1$ of these equations that

$$r\left(\mu + \sum_{j=1}^{n} \alpha_j g_j(\boldsymbol{x})\right) = \mu\left(r + \sum_{j=1}^{n} e_j\right), \tag{2.19}$$

from which it follows that

$$\frac{\mu}{\mu + \sum_{j=1}^{n} \alpha_j g_j(\boldsymbol{x})} = \frac{r}{r + \sum_{j=1}^{n} e_j}. \tag{2.20}$$

Recall the definition of the growth rate in terms of metabolism in eq. (2.15). Using $a_j = \sum_{i=1}^{m} \rho_i P_{ij}$ and $b_j = \sum_{i=1}^{m} \rho_i M_{ij}$ to make notation more concise, we rewrite this as

$$\begin{aligned}
\mu &= \sum_{i=1}^{m} \rho_i \sum_{j=1}^{n} P_{ij} e_j f_j(\boldsymbol{x}) + \sum_{j=1}^{n+1} \left(\sigma_j - \sum_{i=1}^{m} \rho_i M_{ij}\right) r\alpha_j g_j(\boldsymbol{x}) \\
&= \sum_{i=1}^{m} \rho_i \sum_{j=1}^{n} P_{ij} r \frac{\alpha_j g_j(\boldsymbol{x}) f_j(\boldsymbol{x})}{\mu} + \sum_{j=1}^{n+1} \left(\sigma_j - \sum_{i=1}^{m} \rho_i M_{ij}\right) r\alpha_j g_j(\boldsymbol{x}) \\
&= r\left(\left(\sigma_{n+1} - b_{n+1}\right)\mu + \sum_{j=1}^{n} \left(a_j \frac{f_j(\boldsymbol{x})}{\mu} + \left(\sigma_j - b_j\right)\right) \alpha_j g_j(\boldsymbol{x})\right),
\end{aligned} \tag{2.21}$$

such that the steady state ribosome concentration is given by

$$r = \frac{\mu}{\left(\sigma_{n+1} - b_{n+1}\right)\mu + \sum_{j=1}^{n} \left(a_j \frac{f_j(\boldsymbol{x})}{\mu} + \left(\sigma_j - b_j\right)\right) \alpha_j g_j(\boldsymbol{x})} \tag{2.22}$$

and the enzyme concentrations are given by

$$e_j = \frac{r\alpha_j g_j(\boldsymbol{x})}{\mu} = \frac{\alpha_j g_j(\boldsymbol{x})}{(\sigma_{n+1} - b_{n+1})\,\mu + \sum_{j=1}^{n}\left(a_j \frac{f_j(\boldsymbol{x})}{\mu} + (\sigma_j - b_j)\right)\alpha_j g_j(\boldsymbol{x})}. \tag{2.23}$$

2.2.4 Growth rate versus ribosomal fraction of enzymes

Recall equation (2.20). In our model, we assume that $g_j(\boldsymbol{x}) = g(\boldsymbol{x})$ for all $j = 1, \ldots, n{+}1$, and we know that $\mu = \alpha_{n+1} g_{n+1}(\boldsymbol{x})$, so this fraction reduces to

$$\frac{r}{r + \sum_j e_j} = \frac{\alpha_{n+1} g(\boldsymbol{x})}{\alpha_{n+1} g(\boldsymbol{x}) + g(\boldsymbol{x}) \sum_j \alpha_j} = \frac{\alpha_{n+1} g(\boldsymbol{x})}{g(\boldsymbol{x}) \sum_{j=1}^{n+1} \alpha_j} = \alpha_{n+1}. \tag{2.24}$$

Likewise, since $e_j = \frac{r\alpha_j g(\boldsymbol{x})}{\mu}$, we find

$$\frac{e_j}{r + \sum_j e_j} = \frac{\alpha_j g(\boldsymbol{x})}{\mu} \frac{r}{r + \sum_j e_j} = \alpha_j \frac{g(\boldsymbol{x})}{\mu} \alpha_{n+1} = \alpha_j, \tag{2.25}$$

since $\mu = \alpha_{n+1} g(\boldsymbol{x})$. Therefore, the fractions of enzymes and ribosomes is independent of the growth rate. A non-optimized system (a system with fixed $\boldsymbol{\alpha}$) can therefore not explain the bacterial growth law.

This is different in optimized growth, since in that case, we find that the α_j are dependent on how much the cell if 'fed' with a substrate with concentration x_0 outside of the cell and we will then find different fractions of enzymes and a different growth rate at different x_0.

Note that it doesn't matter if we speak of the fraction of ribosomal concentration of the total concentration of enzymes or of the copy number, since from

$$\frac{r}{r + \sum_j e_j} = \frac{Vr}{V\left(r + \sum_j e_j\right)} = \frac{n_r}{n_r + \sum_{j=m+1}^{n} n_j}, \tag{2.26}$$

it follows that they are the same quantity. This quantity is therefore concisely called 'the ribosomal fraction of enzymes'.

2.2.5 Turning the model into a Linear Program

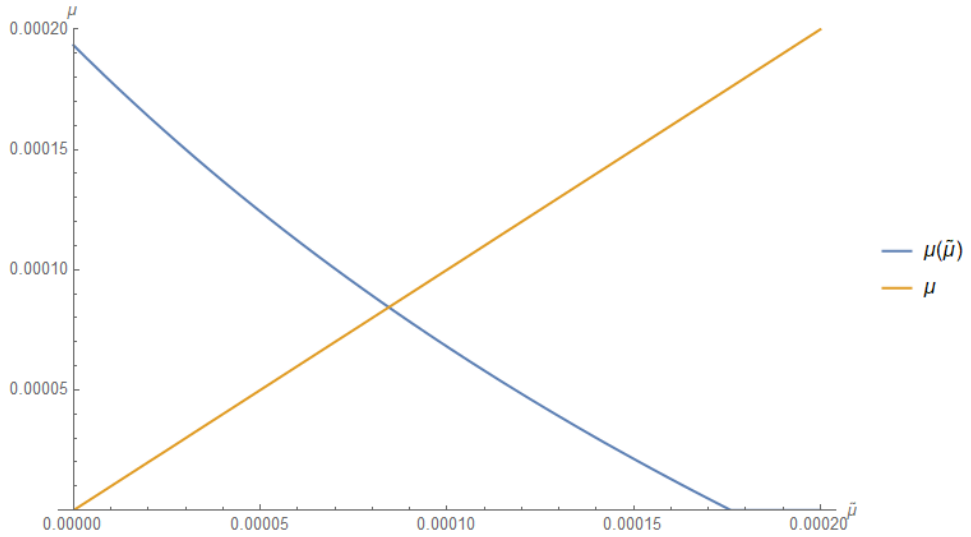


Figure 2.1: In blue the curve $\mu(\tilde{\mu})$ for non particular, fixed $\boldsymbol{x}$ governed by eq. (2.31). In orange, the identity curve. At the intersection, we find $\tilde{\mu} = \mu$.

Recall eq. (2.22). We rewrite this into the quadratic equation for μ

$$\frac{\mu}{r} = (\sigma_{n+1} - b_{n+1})\,\mu + \sum_{j=1}^{n} \left(a_j \frac{f_j(\boldsymbol{x})}{\mu} + (\sigma_j - b_j) \right) \alpha_j g_j(\boldsymbol{x}). \tag{2.27}$$

Following eq. (2.15), in steady state, i.e. $\dot{\boldsymbol{c}} = 0$ and $e_j = \frac{r\alpha_j g_j(\boldsymbol{x})}{\mu}$, the metabolic concentration of x_i for $i = 1, \ldots, m$ is described by

$$\begin{aligned} \frac{\mu x_i}{r} &= \sum_{j=1}^{n} P_{ij} \frac{e_j}{r} f_j(\boldsymbol{x}) - \sum_{j=1}^{n+1} M_{ij} \alpha_j g_j(\boldsymbol{x}) \\ &= \sum_{j=1}^{n} P_{ij} \frac{\alpha_j g_j(\boldsymbol{x})}{\mu} f_j(\boldsymbol{x}) - \sum_{j=1}^{n} M_{ij} \alpha_j g_j(\boldsymbol{x}) - M_{ir} \alpha_r g_r(\boldsymbol{x}) \\ &= \sum_{j=1}^{n} \left(\frac{P_{ij} f_j(\boldsymbol{x})}{\mu} - M_{ij} \right) \alpha_j g_j(\boldsymbol{x}) - M_{ir}\mu. \end{aligned} \tag{2.28}$$

Plugging in eq. (2.27) gives

$$\begin{aligned} x_i \left(\sum_{j=1}^{n} \left(\frac{a_j f_j(\boldsymbol{x})}{\mu} + \sigma_j - b_j \right) \alpha_j g_j(\boldsymbol{x}) + (\sigma_r - b_r)\,\mu \right) \\ = \sum_{j=1}^{n} \left(\frac{P_{ij} f_j(\boldsymbol{x})}{\mu} - M_{ij} \right) \alpha_j g_j(\boldsymbol{x}) - M_{ir}\mu, \end{aligned} \tag{2.29}$$

which are m equations in $\boldsymbol{x}$ and μ only. These equations are generally nonlinear in $\boldsymbol{x}$ and quadratic in μ. Solving these and eq. (2.13) for $\boldsymbol{x}$ and μ is no straightforward exercise.

Now, denote the μ in the denominators by $\tilde{\mu}$. We can construct, for fixed $\tilde{\mu}$, a linear problem in μ. Let us write for $i = 1, \ldots, m$:

$$\begin{aligned} A_{ij}(\boldsymbol{x}, \tilde{\mu}) &= \left(\frac{a_j f_j(\boldsymbol{x})}{\tilde{\mu}} + \sigma_j - b_j \right) x_i - \frac{P_{ij} f_j(\boldsymbol{x})}{\tilde{\mu}} + M_{ij} \quad \text{for } j = 1, \ldots, n, \\ A_{i,n+1}(\boldsymbol{x}, \tilde{\mu}) &= (\sigma_r - b_r) x_i + M_{ir}. \end{aligned} \tag{2.30}$$

The linear problem is then, for fixed $\tilde{\mu}$ and for $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n, \mu / g_{n+1}(\boldsymbol{x}))$,

$$\max_{\boldsymbol{\alpha}} \left\{ \mu \;\middle|\; A(\boldsymbol{x}, \tilde{\mu}) \boldsymbol{\alpha} = 0, \sum_{j=1}^{n} \alpha_j + \frac{\mu}{g_{n+1}(\boldsymbol{x})} = 1, \alpha_j, \mu \geq 0 \right\}, \tag{2.31}$$

which generally is either easily solvable, or has no feasible solutions. The maximiser, denoted by $\boldsymbol{\alpha}^{\text{opt}}$, is an extreme ray of the cone spanned by the constraints. It will therefore be on top of some of the inequality constraints, which means as many entries of $\boldsymbol{\alpha}^{\text{opt}}$ as possible are zero. This means that the maximal growth rate will be in an EGM, since that is the most minimal pathway. We will generally not immediately find $\tilde{\mu} = \mu^{\text{opt}} = \alpha_{n+1}^{\text{opt}} g_{n+1}(\boldsymbol{x})$. We need to vary both $\boldsymbol{x}$ and $\tilde{\mu}$ until we find $\tilde{\mu} = \mu^{\text{opt}}$.

For fixed $\boldsymbol{x}$, eq. (2.31) gives a vector $\left(\alpha_1, \ldots, \alpha_n, \frac{\mu}{g(\boldsymbol{x})}\right)$, which will hold different μ for different $\tilde{\mu}$, which we can see as a function $\mu(\tilde{\mu})$. See Figure 2.1 for an example of this function for fixed $\boldsymbol{x}$. We need to find the intersection of this curve with the identity to find $\tilde{\mu} = \mu$. We then need to do this for different $\boldsymbol{x}$ to find the optimal growth rate.

3 Results

In this section, we will look at an example pathway and use the theory of EGMs to find the bacterial growth law.

3.1 Pathways

We consider five metabolites and seven enzymes, which are (redundantly) connected through four distinct EGMs given by the pathway given in Figure 3.1. A similar, real-life example of this pathway can be found at the pathway starting at *prephenate* in [4]. The four EGMs that can be distilled from this pathway are given by the following sequences of reactions:

EGM_1	v_1, v_2, v_3, v_4, v_5
EGM_2	v_1, v_2, v_3, v_6, v_7
EGM_3	v_1, v_2, v_4, v_6
EGM_4	v_1, v_3, v_5, v_7

This is of course under the assumption that the stoichiometry allows for these non-redundant pathways to be possible. Note that in each of these four EGMs, the amount of metabolites m is equal to the amount of enzymes n. A priori, there is no reason to expect that optimal growth rate is obtained in the same EGM for different values of extracellular substrate concentration x_0.

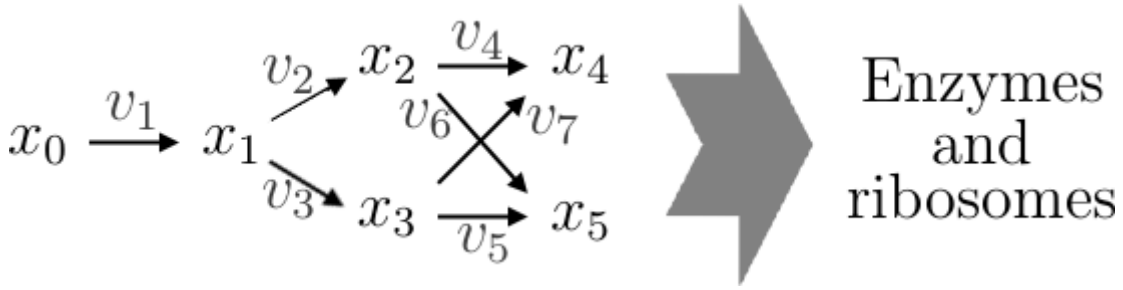


Figure 3.1: Five metabolites, x_1 through x_5, connected by seven different enzyms that together with the ribosome construct the whole cell. In this case, x_4 and x_5 can be seen as the amino acids. The substrate outside of the cell from which it grows is called x_0.

We find optimal growth rate through the following procedure: First, we define stoichiometry matrices P and M, flux rates $f_j(\boldsymbol{x})$, $g_j(\boldsymbol{x})$ and the molar volumes $\boldsymbol{\rho}$, $\boldsymbol{\sigma}$ and σ_r. Since these are properties of the various compounds in consideration, they are fixed.

Next, we describe the linear problem from section 2.2.5 and solve it for various $\boldsymbol{x}$ and $\tilde{\mu}$ until we find the optimal μ and the accompanying concentrations.

3.2 Assumptions

In this section, various assumptions made will be highlighted and explained, if not yet already done.

3.2.1 Definitions

We use

$$P = \begin{pmatrix} 1 & -1 & -2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -2 & 0 & -3 \\ 0 & 0 & 0 & 3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 3 & 2 & 0 \end{pmatrix}, \; M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 30 & 10 & 30 & 0 & 30 & 10 & 0 & 30 \\ 40 & 40 & 0 & 50 & 20 & 40 & 10 & 30 \end{pmatrix} \tag{3.1}$$

for the metabolic and enzymatic stoichiometry.

We define the sign of x as

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0; \\ 0 & x = 0; \\ 1 & x > 0, \end{cases} \tag{3.2}$$

then, for the fluxes, we define

$$f_1(\boldsymbol{x}) = k_{cat,1} \frac{x_0 - x_1/k_1}{\sum_l k_{cat,l} + x_0 + x_1/k_1}, \tag{3.3}$$

where x_0 is the concentration of the source metabolite *outside* the cell. For $j = 2, \ldots, n$ we define

$$f_j(\boldsymbol{x}) = k_{cat,j} \frac{\sum_i \operatorname{sgn}(x_i) x_i/k_i}{\sum_l k_{cat,l} + \sum_i |\operatorname{sgn}(x_i)| x_i/k_i} \tag{3.4}$$

and

$$g(\boldsymbol{x}) = g_j(\boldsymbol{x}) = k_{cat,r} \frac{x_4 x_5}{\sum_l k_{cat,l} + x_4 + x_5 + x_4 x_5} \quad \text{for all } j. \tag{3.5}$$

The molar volumes $\boldsymbol{\rho}$ for the metabolites and $\boldsymbol{\sigma}$ for the enzymes and the ribosome are chosen in such a way that $\boldsymbol{\rho} = \boldsymbol{k}$ and $\boldsymbol{\sigma} = 10 \times \boldsymbol{k}_{\text{cat}}$. This factor 10 means that the enzymes take up way more space than the metabolites do.

3.2.2 Fluxes

The fluxes $f_j(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are defined in the way above to resemble the Michaelis Menten kinetics formula [3]. The fluxes used have to have the following properties:

1. Larger source metabolite concentration results in a higher flux;
2. Larger target metabolite concentration results in a smaller flux;
3. Maximal flux is obtained when source metabolite concentration is infinite and target metabolite concentration is zero.

The fluxes have to remain non-negative. We exclude, e.g., the possibility that x_3 creates x_4, which creates x_2 (requiring negative $f_4(\boldsymbol{x})$), which creates x_5. Starting conditions which give solutions with negative $f_j(\boldsymbol{x})$ or $g(\boldsymbol{x})$ must be discarded.

The constants $\boldsymbol{k}_{\text{cat}}$ and $\boldsymbol{k}$ typically vary wildly and since it's not the goal to investigate the effect of these constants on the ribosomal fraction of enzymes, they are chosen randomly between 0.3 and 0.8.

3.3 Finding the bacterial growth law

For various concentrations x_0 between 7/40 and 11/16 in steps of 1/80, the optimal growth rate and the fraction of ribosomes of the total enzyme have been calculated. This is shown in Figure 3.2. The growth rates vary between 1.5×10^{-4} and 1.7×10^{-4}.

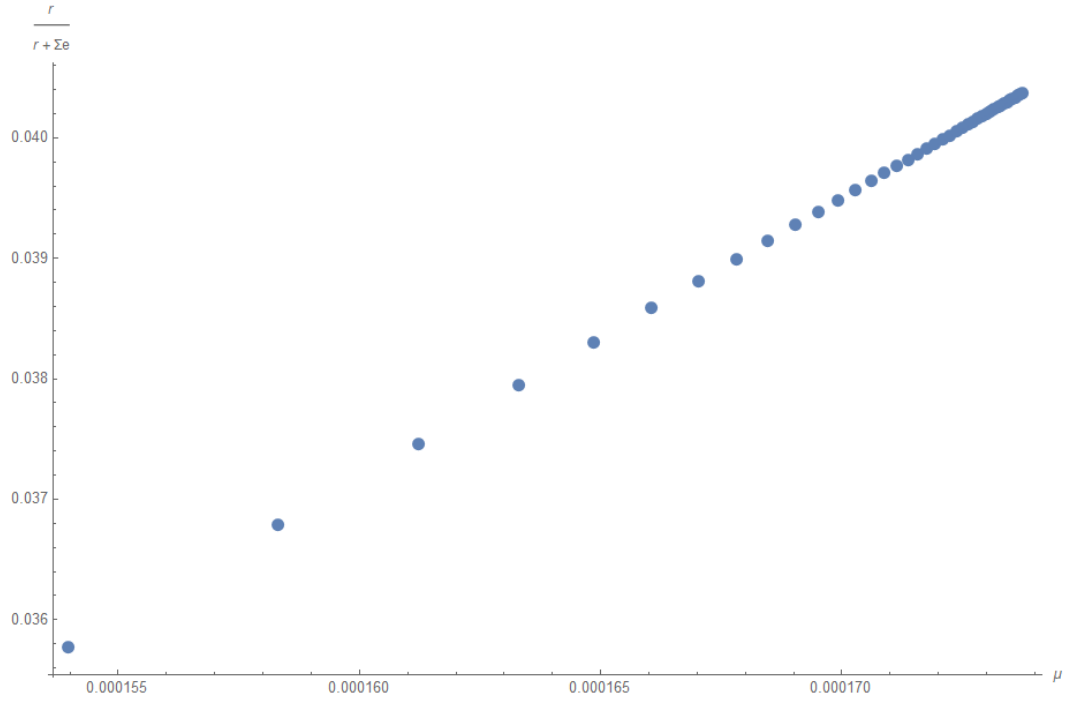


Figure 3.2: The optimal growth rate versus the fraction of ribosomes of the total enzymes for concentrations x_0 between 7/40 and 11/16, in steps of 1/80.

When smaller concentrations of x_0 than 7/40 are chosen, the system becomes unstable and an optimum couldn't be found. When performing linear regression to find a best

straight line fit through these data points, the y-axis intercept is found to be at 9.6×10^{-8}, with an R^2 of one. A priori, there is no reason for this line to go through the origin. A cell that doesn't grow, still needs ribosomes to sustain itself.

For every x_0 considered, we find that $\alpha_6 = \alpha_7 = 0$, so the optimal system will always find itself in EGM_1.

3.4 Conclusion

It can be concluded that it is possible to find an affine connection between the growth rate and the ribosomal fraction of the total enzyme in the theory of EGMs. Although this is only found for a relatively small range in growth rate, it is done in a relative large portion of meaningful extracellular substrate concentrations. It can also be seen that for higher extracellular substrace concentration x_0, the optimal growth rate will barely increase, showing a saturation in the system. Looking at larger x_0 than 11/16 is therefore not meaningful.

4 Discussion

Previously, we have calculated optimal growth rates, but we haven't considered the fact that these are actually *locally optimal*. Eq. (2.31) gives us a vector $\left(\alpha_1, \ldots, \alpha_n, \frac{\mu}{g(\boldsymbol{x})}\right)$ for fixed $\boldsymbol{x}$ and $\tilde{\mu}$. Generally, the function $\mu(\tilde{\mu})$ prescribed by eq. (2.31) is found to be strictly descending down to the y-axis, see Figure 2.1 for an example. To find the growth rate μ, we have to equate said curve to the identity curve. It is unknown whether there is always at most one intersection for positive μ. If there is more than one, we might be looking at growth rate which isn't optimal, without knowing it immediately.

On the other hand, it is also unknown whether there is a different set of values for $\boldsymbol{\alpha}$ that gives the same optimal growth rate μ^{opt}. This different vector $\boldsymbol{\alpha}'$ could even describe a completely different EGM. Yet, both EGMs must abide to the bacterial growth law. For such $\boldsymbol{\alpha}'$ we know that

$$\mu^{\text{opt}} = \alpha_{n+1} g_{n+1}(\boldsymbol{x}) = \alpha'_{n+1} g_{n+1}(\boldsymbol{x}'). \tag{4.1}$$

Since we know that $\alpha_{n+1} = r/(r + \sum_j e_j)$, this could mean that $\boldsymbol{\alpha}'$ gives a different fraction of ribosomes in the cell. If this happens, it would make finding the bacterial growth law using this model a lot harder.

Furthermore, it has been noted earlier that the system malfunctions for extracellular substrate concentrations smaller than 7/40. This could be caused by eq. (2.31) being to close to zero, making it nigh impossible to pinpoint the exact intercept with the identity function. If this is the cause, then this could be remedied by increasing the working precision. This will, however, be more time-consuming as this will lead to slower calculations.

Bibliography

[1] Campbell NA, Reece JB, Urry L, Cain ML, Wasserman SA, Minorsky PV, Jackson RB, *Biology: A Global Approach*, London, Pearson, 2015.

[2] Klumpp S, Hwa T, *Bacterial growth: global effects on gene expression, growth feedback and proteome partition* Current Opinion in Biotechnology, 28, 96-102, 2014.

[3] Michaelis L, Menten ML, *Die Kinetik der Invertinwirkung*, Biochem Z, **49**, 333369, 1913.

[4] Michal G, `http://biochemical-pathways.com/#/map/1`.

[5] Molenaar D, van Berlo R, de Ridder D, Teusink B, *Description of the Growth Models.*

[6] Schaechter M, Maaløe O, Kjeldgaard NO, *Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium*, Journal of General Microbiology, **19**, 592-606, 1958.

[7] Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T, *Interdependence of cell growth and gene expression: origins and consequences*, Science, **330**, 1099-1102 2010.

[8] Scott M, Hwa T, *Bacterial growth laws and their applications*, Current Opinion in Biotechnology, 22, 559-565, 2011.