June 2019

Maximizing and Modeling Malonyl-CoA Production in Escherichia coli

Tatiana Thompson Silveira Mello
Louisiana State University and Agricultural and Mechanical College

MAXIMIZING AND MODELING MALONYL-COA PRODUCTION IN *ESCHERICHIA COLI*

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

in

The Department of Mechanical Engineering

by
Tatiana Thompson Silveira Mello
Mech. Eng., University of Campinas, 2016
Biol. Scien., University of Campinas, 2011
August 2019

Acknowledgments

I would like to thank Dr. Marcio de Queiroz and Dr. Grover Waldrop for advising and guiding me during these past years. Many thoughts, ideas, and advices were shared during this pleasant experience helping me to accomplish not only another professional step in my life, but also to become a better person.

Also I would like to thank my family, that besides being far away from here, always believed on me and made themselves present at all stages of this degree. Finally, I would like to thank my best partner, Daniel, who shared with me all this convoluted and adventurous experience of living abroad and discovering a new world out of our comfort zone.

Table of Contents

List of Tables

List of Figures

Abstract

In E. *coli*, fatty acid synthesis is catalyzed by the enzyme acetyl-CoA carboxylase (ACC), which converts acetyl-CoA into malonyl-CoA. Malonyl-CoA is a major building block for numerous of bioproducts. Multiple parameters regulate the homeostatic cellular concentration of malonyl-CoA, keeping it at a very low level. Understanding how these parameters affect the bacterial production of malonyl-CoA is fundamental to maximizing it and its bioproducts. To this end, competing pathways consuming malonyl-CoA can be eliminated, and optimal nutritional and environmental conditions can be provided to the fermentation broth. Most previous studies utilized genetic modifications, expensive consumables, and high-cost quantification methods, making unfeasible the development of an economically-attractive process with high product yield.

In this work, we propose a low-cost, simple, and effective method to maximize and quantify malonyl-CoA production in E. *coli*. The enzyme, 1,3,6,8-tetrahydroxynapthalenesynthase (THNS) catalyzes the condensation of five molecules of malonyl-CoA forming 1,3,6,8-tetrahydroxynaphthalene (THN), which is auto-oxidized into flaviolin. Flaviolin can be measured spectrophotometrically, meaning that direct quantification of malonyl-CoA is accomplished by measuring the absorbance of flaviolin. Results showed that the main parameters associated with malonyl-CoA maximization are the use of rich medium supplemented with glucose and metals, the incubation temperature of 37°C, the inoculum incubation time, and the use of lactose as an inducer. Moreover, a kinetic model of the bacterium metabolism was built in order to guide the maximization of malonyl-CoA. By performing a stability analysis and a metabolic control analysis, the system was determined to be stable and flux controlled by phosphoglucose isomerase. This suggests that this enzyme should be modified to optimize the production of malonyl-CoA.

Chapter 1
Introduction

1.1. Motivation

As the world′s population steadily increases, there is a concomitant increase in the demand for energy. The current petrochemical industry controls most of the production of transportation energy and other petrochemicals. The instability of petroleum prices, the limited availability, and its impact in the environment make it necessary to look for alternative feedstocks that can sustain the profitability of energy companies [3]. Although there are renewable, environmental-friendly sources for producing biochemicals and biofuels, the cost of using these alternative sources is not yet competitive with the traditional petrochemical industry. Therefore, there is an imperative necessity to continue exploring renewable sources to produce fuels and chemicals, as well as to optimize these processes.

Industrial biotechnology uses microorganisms and enzymes to produce a wide range of chemical compounds. One example is the production of bulk chemicals by using biological pathways instead of the standard petrochemical processes. This biological approach has several potential advantages over established petrochemical processes:

- Fossil-based products release carbon, increasing atmospheric greenhouse gases. On the other hand, bioproducts hold the potential of a much shorter carbon cycle and, consequently, a strongly decreased negative impact on the environment;
- World reserve of fossil feedstock is limited and unequally distributed;
- Mild conditions under which biotechnological processes are generally performed can be performed in any place, irrespective of the specific environmental conditions;
- Microbial and enzyme-based catalysis offers access to an enormous range of known and yet to be discovered molecules with potential applications in the pharmaceutical, food, chemical, and fuels industries.

The advances of biotechnology over the past 50 years directly have impacted the number and quality of bioproducts in the marketplace. By modifying wild strains, cells can tolerate better external conditions; by adding or removing reactions from the original pathway when overexpressing or underexpressing native genes, and modifying proteins characteristics, desired products have their yield increased. Therefore, the combination of a powerful microorganism with several tools from biotechnology has the potential to produce highly efficient bioprocesses.

With this in mind, the selection of an optimum platform capable of transforming the feedstock into the desired product is a key stage to synthesize bioproducts. Even though different microorganisms are able to produce bioproducts as either an intermediate or end product through a range of metabolic pathways, this study focuses on one in particular: *Escherichia coli*.

E. *coli* is one of the most utilized microbial strains at the bioindustrial level due to its simplicity, low experimental cost, ease of genetic modification, and intensive academic studies in the past decades [4]. This microorganism is a prokaryote and therefore has a simpler structure and no cellular compartmentalization. Several industrially relevant compounds are produced by E. *coli* such as bulk chemicals [5], biofuels [6], and organic acids [7] such as succinic acid. Moreover, the development of new sciences such as genomics, transcriptomics, and proteomics increased the understanding of this bacterium. This large amount of information on the metabolism of this bacteria has resulted in the development of mathematical models of E. *coli* metabolism, such as the central carbon metabolism for glucose utilization [8] and the genome-scale modeling [9].

In this scenario, malonyl-CoA has been intensively investigated in the past decades by the biotechnology industry. Malonyl-CoA is a three-carbon product found in most organisms, and is a major building block for dozens of bioproducts [6]. Through a biosynthetic pathway, malonyl-CoA can be converted into flavonoids or polyketides. Recently, flavonoid- derived compounds were identified as potential pharmaceutical candidates act-

ing as bactericide, fungicide, chemo preventive, and immunosuppressants [10]. Polyketides were found as precursors for the production of semi-synthetic antibiotics [11, 2]. When catalyzed by malonyl-CoA reductase, malonyl-CoA is converted to 3-hydroxypropionic acid, a bulk chemical used in the production of acrylics. Moreover, malonyl-CoA is the substrate for fatty acids biosynthesis, which are the precursor for biodiesel production (Fig 1.1).

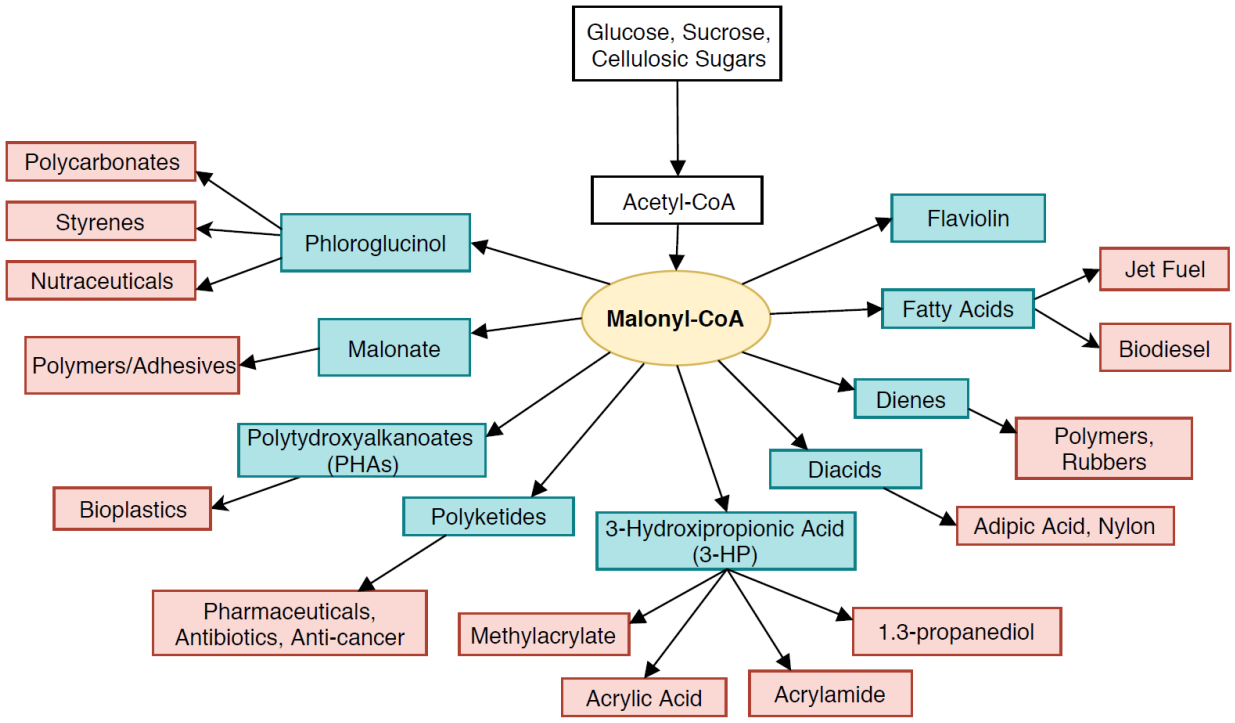


Figure 1.1. Malonyl-CoA and its derived bioproducts

The production of malonyl-CoA derived bioproducts is still difficult due to its intracelullar concentration being very low. Therefore, studies attempted to overcome this constraint by genetically modifying E. *coli*, and by utilizing chemical inhibitors of fatty acids synthesis. Unfortunately, these strategies are costly and cannot compete with the traditional fossil fuel-based methods. Therefore, we propose a low-cost method to increase malonyl-CoA in E. *coli* by adjusting the fermentation conditions to 'optimum' ones.

The method will be guided by extensive experimentation and by a kinetic model of the process. The mathematical model will assist in the creation of bacterial strains with high

productivity rates by determining rate limiting steps and the conditions that maximize production. In other words, the kinetic model of the malonyl-CoA metabolic pathway in E. *coli* will quantitatively predict how changes in one pathway can affect the synthesis rate of the bioproduct derived from malonyl-CoA.

1.2. Background

1.2.1. Metabolism of E. *coli*

Cellular metabolism consists of all reactions that uptake nutrients from the environment, and then utilize them for cell function and reproduction. Pathways in the central metabolism are especially relevant in this study, as they allow the utilization of different carbon sources, such as glucose, glycerol, sucrose, and acetate. Glucose and other sugars have been widely studied, as they represent an efficient carbon source for cellular growth [12].

Firstly, the phosphotransferase system (PTS) transports the glucose present in the media and phosphorylates it into glucose-6-phosphate (G6P). Glycolysis is for the most part a linear pathway that starts from the metabolite glucose-6-phosphate (G6P) and culminates in the production of pyruvate. The first reaction in glycolysis is the conversion of G6P into fructose-6-phosphate (F6P) mediated by phosphoglucose isomerase (PGI). F6P is an essential metabolite that either goes to the pentose phosphate pathway or continues in the glycolytic pathway by adding a second phosphorylation that results in fructose-bi-phosphate (FBP). FBP is reversibly dissociated into two metabolites, glyceraldehyde-3-phosphate (GAP) and di-hydroxyacetone-phosphate (DHAP). DHAP can be transformed into GAP by the action of a triosephosphate isomerase (TPI) and continue with the glycolytic pathway. GAP is interconverted through four successive reactions into phosphoenolpyruvate (PEP), resulting in the production of one molecule of ATP and one molecule of NADH. The last reaction in glycolysis is the conversion of PEP into pyruvate (PYR), producing one ATP. This reaction is catalyzed by the action of two pyruvate kinases, both subject to genetic and allosteric regulation [13]. Thereafter, pyruvate dehydrogenase catalyzes the

formation of acetyl-CoA (AcCoA) from pyruvate [14]. Most of the formed acetyl-CoA is directed to the tricarboxylic acid (TCA) cycle, whereas a smaller part is catalyzed by acetyl-CoA carboxylase (ACC) to form a malonyl-CoA [15]. Therefore, a portion of the acetyl-CoA flux from the central metabolism pathway is diverted to synthesize malonyl-CoA for fatty acid and lipid membrane synthesis. Malonyl-CoA is ultimately also a substrate for the production of fatty acids. The above process is depicted in Fig. 1.2.

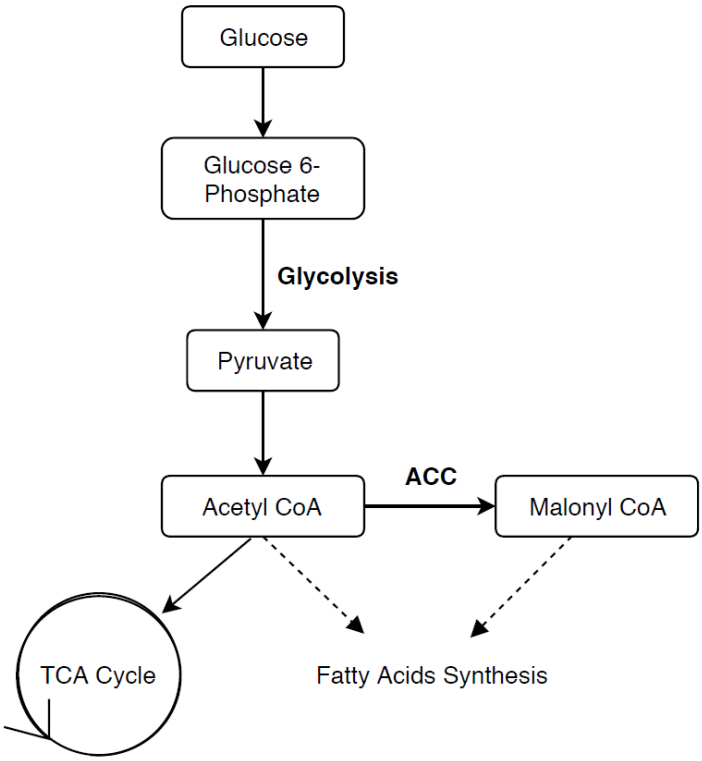


Figure 1.2. E. *coli* metabolism

Acetyl-CoA carboxylase uses bicarbonate (HCO3-), which is the hydrated form of CO_2, as a substrate. As bicarbonate is consumed by ACC, more gaseous CO_2 dissolves in the water and reacts to form bicarbonate. Bacterial ACC is composed of three proteins: biotin carboxylase (BC), carboxyl transferase (CT), and biotin carboxyl carrier protein (BCCP). These three proteins are produced by four different genes. BC and BCCP are each encoded by a single gene, whereas CT requires two genes that code for an α and β subunit. The biotin moiety is attached to a specific lysine residue (lysine 122) in BCCP via a post-

translational modification reaction catalyzed by biotin ligase [16].

In the first half reaction, BC carboxylates biotin attached to BCCP. In the second half reaction, CT transfers the carboxyl group from biotin to acetyl-CoA to produce malonyl-CoA. Extensive studies have shown that enzymatic activity requires all three of these proteins to form a macromolecular complex, hereafter referred to as holo ACC:

Reaction 1:

$$BCCP - biotin + MgATP + HCO_3^- \underset{Mg^{++}}{\overset{BC}{\rightleftharpoons}} BCCP - biotin - CO_2^- + MgADP + Pi$$

Reaction 2:

$$BCCP - biotin - CO_2^- + Acetyl - CoA \overset{CT}{\rightleftharpoons} BCCP - biotin + Malonyl - CoA$$

Sum:

$$MgATP + HCO_3^- + Acetyl - CoA \overset{BC-BCCP-CT}{\rightleftharpoons} MgADP + Pi + Malonyl - CoA$$

The last reaction is the rate-limiting step to produce fatty acids and lipid membrane synthesis, and therefore vital for the cellular metabolism and cell membrane maintenance. Intracellular malonyl-CoA concentration is tightly regulated to be very low [17], so as to coordinate the rate of fatty acid biosynthesis with phospholipid production, macromolecule synthesis, and cell growth. This tight regulation of malonyl-CoA thereby limits the yield of the mentioned bioproducts of Fig. 1.1. Maximizing the enzymatic activity of ACC *in vivo* will enhance the production of malonyl-CoA, and therefore increase the yield of these industrially valuable products.

1.2.2. Enzyme Kinetics

Metabolic pathways are composed of individual enzyme-catalyzed steps. Enzymes catalyze reactions by binding the reactants (substrates) and facilitating their conversion to the reaction products. The details of this molecular mechanism varies widely and has correspondingly different effects at the macroscopic level of enzyme kinetics. There are enzymes that have multiple substrates, in which the affinity of binding can depend on the order in which the different substrates bind. Besides substrate and product binding, effectors can

bind non-competitively to sites different from the active site (allosteric regulation) [18]. Moreover, the physiological environment (pH, ion concentrations, temperature) influences binding and catalysis and hence the kinetics of product formation.

Enzymatic reactions follow the mass action law which states that the rate of a chemical reaction is proportional to the product of the concentrations of the reactants [19]. Mathematically, this law will take the form of ordinary differential equations (ODEs).

The simplest and most widely used expression is the rate law Michaelis-Menten kinetics, which describes an enzymatic kinetic reaction that transforms irreversibly one substrate (S) into one product (P), through one enzyme (E) that forms an intermediate complex (ES) by binding with the substrate (Eq 1.1). The reaction rate k_2 is called the enzyme's catalytic constant.

$$E + S \underset{\mathrm{k_{-1}}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} ES \xrightarrow{\mathrm{k_2}} E + P \tag{1.1}$$

The rate v, of an enzyme is the speed at which it generates its products. It can be regarded as a function of the form $v(P_1, \ldots, P_m)$, where P are the concentrations of the metabolites. The rate in the following equation is the net rate of product formation:

$$v(S, P) = \frac{dP}{dt} = k_2 \, [ES] \tag{1.2}$$

The second reaction is irreversible, meaning that the product concentration does not affect v. Assuming that the substrate is in excess relative to the enzyme concentration ($S \gg E$), the enzyme-complex ES is expected to rise extremely fast after binding of the enzyme to the substrate, and thereafter remain constant. Since the time scale is reasonably small (from 0.1 to 1000 seconds), we can assume steady-state conditions, i.e., $\frac{d[ES]}{dt} = 0$. Applying the law of mass action, and denoting concentrations by $[S]$ (substrate), $[E]$ (free enzyme), $[ES]$ (enzyme-substrate complex), and $[P]$ (product), the following differential equation

model is obtained:

$$\frac{d[S]}{dt} = +k_{-1}[ES] - k_1[S][E] \tag{1.3}$$

$$\frac{d[E]}{dt} = +k_{-1}[ES] + k_2[ES] - k_1[S][E] \tag{1.4}$$

$$\frac{d[ES]}{dt} = +k_1[S][E] - k_2[ES] - k_{-1}[ES] \tag{1.5}$$

$$\frac{d[P]}{dt} = k_2[ES] \tag{1.6}$$

Under steady state conditions, we have:

$$(k_2 + k_{-_1})[ES] = k_1[E][S] \tag{1.7}$$

considering that the concentration of enzyme inside the cell remains constant when it is not consumed, we can write $E = E_{tot} + ES$, where E_{tot} is composed partly of free enzyme and partly of enzyme bound to the substrate. Combining this relationship with Eq. 1.7, and solving for ES:

$$(k_2 + k_{-_1})[ES] = k_1([E_{tot}] + [ES])[S]$$

$$[ES] = \frac{k_1[E_{tot}][S]}{(k_2 + k_{-_1}) + k_1[S]} = \frac{[E_{tot}][S]}{K_m + [S]}$$

where $K_m = \frac{k_{-1}+k_2}{k_1}$ is called the Michaelis-Menten constant.

Now, using Eq 1.2 and assuming that the amount of substrate is very large approaching the value k_2E_{tot}, the maximum reaction rate is $v_{max} = k_2E_{tot}$, the basic equation of enzyme kinetics becomes:

$$v(S,P) = \frac{v_{max}\ [S]}{[S] + K_m} \tag{1.8}$$

Since v_{max} is directly proportional to the total amount of enzyme present in the cell, we can see that the rate of the enzymatic reaction Eq. 1.8 can be manipulated by adjusting the amount of enzyme.

The more general form for metabolic reactions is written in the form of v as a function of S and P, i.e., $v = v(S, P)$. For example, the reaction, where A is substrate, E is enzyme, X is enzyme-substrate complex and P is product, can be written as:

$$A + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X \underset{k_4}{\overset{k_3}{\rightleftharpoons}} P + E$$

$$\frac{d[A]}{dt} = v_2 - v_1$$

$$\frac{d[E]}{dt} = v_2 + v_3 - v_1 - v_4$$

$$\frac{d[X]}{dt} = v_1 + v_4 - v_2 - v_3$$

$$\frac{d[P]}{dt} = v_3 - v_4$$

where rate equations are given by:

$$v_1 = k_1 [A] [E]$$

$$v_2 = k_2 [X]$$

$$v_3 = k_3 [X]$$

$$v_4 = k_4 [P] [E]$$

1.2.3. Pathways

In order to visualize the cell's complete metabolic network, the metabolic pathways need to be studied. Pathways can be linear, either forward or backwards, or contain cycles or other sophisticated types of interaction between individual enzymatic reactions [19].

Assume a pathway consisting of reactions with m metabolites where each metabolite has concentration C_i for $1 \leq i \leq m$, and each reaction has rate v_j for $1 \leq i \leq r$. Each v_j is a function of all metabolites concentration: $v_j = v_j(C_1, \ldots, C_m)$. We can arrange all rates into a $r \times 1$ vector and all concentrations into a $m \times 1$ matrix. Often, there are reactions that produce or consume more than one molecule of a given metabolite. Hence, the stoichiometric matrix (N) needs to be introduced, where N is an $m \times r$ matrix whose

elements are integers. For example, if k molecules of the metabolite i appear as a product in reaction j, then $N_{ij} = k$, while if k molecules of metabolite i appear as substrate in reaction j, then $N_{ij} = -k$ [20]. Therefore, the reaction $2C_1 + C_2 \longrightarrow 4C_3 + 2C_4$ has:

$$N = \begin{bmatrix} N_{11} \\ N_{21} \\ N_{31} \\ N_{41} \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 4 \\ 2 \end{bmatrix} \tag{1.9}$$

as the stoichiometric matrix, thus we can write a set of equations in the matrix form $\frac{d[C]}{dt} = Nv(C)$ that govern the time dynamics of a specific pathway.

1.2.4. Stability Analysis

The concept of stability is central to the analysis of dynamical systems. Systems are said to be stable if small perturbations, whether in the initial condition or due to external changes, do not give rise to large sustained modifications in the behavior of the system [21].

Given a set of dynamic equations in the form:

$$\dot{x} = \frac{dx}{dt} = f(t, x, u) \tag{1.10}$$

where $\dot{x}$ denotes the derivative of x with respect to the time variable t. The vector x is called the state of the system and u is the input variable. When the input u is identically zero, the equation takes the form:

$$\dot{x} = f(x, t, 0) = f(x, t)$$

When $f(x, t)$ is not a function of time, we can write $\dot{x} = f(x)$, in which case the system is said to be autonomous. Autonomous systems are invariant to shifts in the initial time, meaning that changing the time variable from t to $\tau = t - \alpha$ does not change the right-hand

side of the state equation [22].

It is important to understand the concept of equilibrium points concept when dealing with the state equation. "A point $x = x_e$ in the state space is said to be an equilibrium point of the autonomous system $\dot{x} = f(x)$ if it has the property that whenever the state of the system starts at x_e it remains at x_e for all future time" [22]. Thus, the equilibrium points of an autonomous system are the real roots of the equation $f(x_e) = 0$. Clearly, if

$$\dot{x} = \frac{dx}{dt} = f(x_e) = 0$$

it follows that x_e is constant and, by definition, it is an equilibrium point.

Nonlinear systems can be linearized in order to facilitate their stability analysis. Therefore, the qualitative behavior of a nonlinear system near an equilibrium point can be determined through linearization with respect to that point [23]. Using the Taylor series, and assuming a small neighborhood around the equilibrium point, the system can be written as:

$$\dot{x} = Ax \tag{1.11}$$

where A is the Jacobian matrix evaluated at the equilibrium point x_e ($A = \frac{\partial f}{\partial x}_{x=x_e}$). If x is an $n \times 1$ vector, then A is an $n \times n$ matrix, where the i, jth element is the partial derivative of the ith component of f with respect to the jth element of x:

$$a_{i,j} = \frac{\partial f_i(x)}{\partial x_j}_{x=x_e}$$

The behavior of the linearized system (Eq. 1.11) can provide information about the local behavior of the original nonlinear model (Eq. 1.10). Specifically, this information is obtained from the eigenvalues of A which are determined from the characteristic equation $det(\lambda_i I - A) = 0$. The stability of a linear system depends on the eigenvalues λ_i, according to the following criteria:

- If *all* $Re(\lambda_i) < 0$, the equilibrium point is *stable*;
- If *any* $Re(\lambda_i) > 0$, the equilibrium point is *unstable*;
- If $Re(\lambda_i) = 0$, it is indeterminate, and nothing can be concluded about the stability of the equilibrium point.

1.2.5. Metabolic Control Analysis

The advantage of Metabolic Control Analysis (MCA) is that it requires a smaller number of experiments and parameter values in comparison to other analysis methods. By studying control coefficients, we get responses to perturbation of metabolic parameters. Therefore, kinetic modeling can be combined with MCA theory in order to elucidate the control structure of a pathway, and study the effects of perturbations on a system in steady state [24]. MCA is also referred as the analysis of the sensitivity of metabolites [25].

It is known that in a cell, if the production and degradation of metabolite rates are very fast and approximately at the same level, then the concentration of a metabolite is at steady state [25]. Steady state levels show more similar properties to metabolic levels, making possible the understanding of the enzyme kinetics in the reaction just by using parameters. Moreover, it is possible to predict how an enzyme works under diverse conditions by changing substrate concentrations.

Since metabolic networks are complex systems, it is not intuitive which reactions are primarily responsible for control of the system. Consequently, by calculating the so-called flux control coefficients or sensitivity coefficients, we can determine which metabolites have more effect on the metabolic pathway. Modelling efforts often focus on the steady-state

behavior since experimental observations of metabolism are often carried out in steady-state [19]. The variables of MCA are the steady state metabolite concentrations, and reaction fluxes, while the parameters are the enzyme concentrations.

MCA describes how changing an enzyme activity affects metabolite concentration and reaction fluxes, and can be used to predict the response of a system to perturbations. A control coefficient is a relative measure of how much a perturbation affects a system variable (e.g. fluxes or concentrations). The relative sensitivities of metabolite concentrations are called concentration control coefficients, and are defined by:

$$C_S = \frac{e_j}{C_{ss}} \frac{\partial C_{ss}}{\partial e_j} \tag{1.12}$$

where C_{ss} is the steady-state concentration of specie C and e is the abundance of enzyme j. Likewise, the flux control coefficients are given by

$$C_J = \frac{e_j}{J_k} \frac{\partial J_k}{\partial e_j} \tag{1.13}$$

where J_k is the flux through reaction k.

The rate of a metabolic reaction is proportional to the abundance of the enzyme (e_j). Perturbation is known to change in the concentration e_j of active enzyme. Thus, a relative change in enzyme concentration results in an equivalent relative change in the corresponding reaction rate (e.g. 2% increase in e_j leads to a 2% increase in rate v_k) [19].

The flux control summation theorem states that all C_J of a specific flux J_k should add to one, regardless of the form of the kinetics of the reaction. When a C_J is close to one, it means that the corresponding enzyme is the rate-limiting of the system. However usually the control of the flux is distributed among more than one enzyme. On the other hand, the concentration control summation theorem states that the sum of all C_S of one metabolite is equal to zero for any metabolite of the pathway. Therefore, for a system with m metabolites and k reactions:

$$\sum_{i=1}^{k} C_{J_k} = 1$$

$$\sum_{i=1}^{m} C_{S_m} = 0$$

While the control coefficients provide information about how flux J_k or concentration C_S depends on a reaction rate v_i, a new term called elasticity is included in the MCA to show the relationship between the reaction rate v_i and the concentration of metabolite C_S [19]:

$$\epsilon_{C_{S_m}} = \frac{C_{S_m}}{v_i} \frac{\partial v_i}{\partial C_{S_m}}$$

Since elasticities are partial derivatives of the rate laws, they are properties of individual reactions. In contrast, control coefficients will provide an insight on how that reaction behaves within the whole network. Another theorem, called the Connectivity Theorem, relates control coefficients to the elasticities [26]. The connectivity theorem for flux control is as follows:

$$\sum_{i=1}^{k} C_{J_k} \cdot \epsilon_{C_{S_m}} = 1$$

Applying MCA to the kinetic model, the elastciities can be calculated for each reaction. By using the summation and connectivity theorem, the control coefficients can be found, providing information about the control steps in the metabolic pathway and facilitating enzymatic manipulation.

1.3. Literature Review

A good strategy to maximize malonyl-CoA in E. *coli* is to inhibit the fatty acid synthesis since it would not only eliminate a competing pathway consuming malonyl-CoA, but also alleviate the inherent negative regulation (Fig. 1.2). Several studies have attempted to use metabolic engineering to decrease the fatty acids synthesis in recombinant E. *coli* strains, and enhance the carbon flux towards acetyl and malonyl-CoA. Lynch et al achieved the highest titer reported thus far of 3-Hydroxypropionic acid, a bioproduct derived from malonyl-CoA, by knocking down out genes of the fatty acids synthesis pathway [27]. Lu et al introduced four distinct genetic changes into the E. *coli* genome, in which one of them overexpresses ACC to overproduce more efficiently fatty acids which are used to synthesize microbial biodiesel [6]. Furthermore, many approaches treated the cells with expensive antibiotics, such as cerulenin or thiolactomycin, to inhibit β-keto-acyl ACP synthase and acetyl-CoA ACP transacylase of fatty acid synthesis, thereby enhancing malonyl-CoA production [15, 28, 29, 30].

Acetyl-CoA carboxylase is well known to be the first enzyme of the biosynthetic sequence of fatty acid synthesis, but it is possible that this enzyme is not the crucial pacemaker of fatty acids synthesis. Therefore, the goal is to understand how nutrients and physiological conditions play a role into the optimization of ACC, and maximization of malonyl-CoA in E. *coli*, without modifying any gene of the microorganism and/or treating the cells with fatty acids synthesis inhibitors or any other expensive resource.

So far, no research groups have been able to enhance the production of malonyl-CoA at a low enough cost that is competitive with current petrochemical-based processes.(Table 1.1). The highest titer achieved for 3-HP by Lynch et al [27] depends on the utilization of cerulenin which costs approximately $\$20\text{mg}^{-1}$ and IPTG which costs $\$65.90\text{g}^{-1}$, making the process impractical for industrial scale production. Moreover, most studies utilize complex and expensive methods to quantify malonyl-CoA, such as Gas Chromatography–Mass Spectrometry (GC-MS) or High Performance Liquid Chromatography (HPLC).

Table 1.1. Different studies carried out for the production of malonyl-CoA

End product	Fermentation condition	Strategy to maximize malonyl-CoA	Inducer	Malonyl-CoA (mM/ gDCW)	Reference
-	Shake flask	Genetic modifications	IPTG	0.05	[10]
-	Shake flask	Genetic modifications	IPTG	0.53	[17]
Polyketides	Fed-batch; 5L bioreactor	Genetic modifications; Cerulenin	IPTG	1.50	[30]
Naringenin (Flavonoid)	Shake flask	Genetic modifications	IPTG	1.74	[31]
3-hydroxypropionic acid	Shake flask	Genetic modifications	IPTG	3.41	[32]
3-hydroxypropionic acid	Shake flask	Genetic modifications	IPTG	7.01	[33]
3-hydroxypropionic acid	Fed-batch; 1L bioreactor	Genetic modifications; Cerulenin	IPTG	110.17	[27]

Modeling the metabolism of a microorganism is an important and essential tool to rationally exploit the different metabolic pathways and properties of this complex biological system. In particular, kinetic models are capable of representing the complex biochemistry of cells in a more complete way compared to most other types of models [34]. Genome-scale, metabolic-networking, and constraint-based modeling have been carried in order to simulate the best strategies to enhance intracellular availability of malonyl-CoA [10, 30, 31, 35]. However, a kinetic model that includes all the enzymatic reactions, from the glucose uptake up to the bioproduct formation from malonyl-CoA, has not been done yet. Therefore, the second outcome of this project is to model the metabolism of E. *coli* for the production of malonyl-CoA and its bioproduct.

Chapter 2
Materials and Methods

2.1. Experimental Work

2.1.1. Strains

The E. *coli* strain BL21(DE3) and the expression vector pCDFDuet-1 were from Novagen. Restriction enzymes, dNTPs, and T4 DNA ligase were from New England Biolabs. Primers were purchased from MWG Biotech. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) was from Gold Biotechnology. In addition, the plasmid pLB0056, which contains the genes for holo ACC and biotin ligase, was a gift from Dr. Park Sunghoon of Pusan National University, Busan, Republic of Korea [36]. All other reagents were from Sigma.

The gene for 1,3,6,8-tetrahydroxynapthalene synthase (THNS) was a gift from Hans Liao of OPX Biotechnologies in Boulder, CO. The THNS gene was amplified using the forward primer 5'-CTTCTTGGATCCGATGACCACTCTGTGCCGC-3' and backward primer 5'-CTTCTTAAGCTTTCATTAATCGGCGGTCTG-3'. The PCR product was cut with BamHI and HindIII then inserted into pAEP9, which was cut with the same two restriction enzymes. This generated the plasmid pSEB1, containing not only the gene for THNS but also the genes for the α and β subunits of CT, which were cloned into a mini operon on pCDFDuet-1. The E. *coli* strain BL21(DE3) was transformed with pSEB1 and pAEP7, which contained the genes for BCCP and BC in a mini operon cloned into pET-28. Co-transformation of pSEB1 and pAEP7 was possible because pET28 and pCDFDuet-1 have different origins of replication. E. *coli* strain BL21(DE3) was transformed with the plasmid pAER1 which contains the amplified THNS gene. In order to evaluate the different subunits of ACC individually, the THNS gene from pSEB1 was subcloned into the BamHI and HindIII sites of pCDFDuet-1 to generate pAER1 (Table 2.1).

Table 2.1. List of all plasmids used in this work

Plasmid	Description
pAER1	THNS
pAEP3	Holo BCCP
pAEP7	BCCP and BC
pAEP9	α and β of CT
pSEB1	pAEP9 and THNS in pCDFDuet-1
pLBOO56	Holo ACC and Biotin Ligase

2.1.2. Fermentation Conditions

Unless noted otherwise, the following phases for the *in vivo* assay were the standard procedure used for all experiments (Fig 2.1).

- Plate: Luria Bertani (LB) agar plates were streaked with a colony from the permanent, and then incubated for 20 hours at 37°C.

- Inoculum: 10 ml of medium was added to a 125 ml flask and supplemented with carbon sources. A single colony from the LB agar plate was used to inoculate the flasks, and 30 μl of antibiotics (50 mg.ml^{-1}) were added. Flasks were covered with aluminum foil, with holes to increase aeration. Incubation was done overnight in a shaking water bath at 37°C and 250 rpm.

- Cultures: 5 ml of autoclaved medium was added to a 125 ml flask, 1% (v/v) of the inoculum was transferred and gene overexpression was induced with lactose or IPTG. The flask was covered with aluminum foil, with holes to increase aeration. Incubation was done overnight in a shaking water bath at 37°C and 250 rpm.

- Reading: 1 ml sample of the culture was centrifuged for 150 seconds at 13,500 rcf (relative centrifugal force), and 250 μl of the supernatant was added to 750 μl of water (1:4 dilution). The absorbance (A) was measured at 340nm (OD_{340}, L= 1 cm path length) using a Cary 60 UV-Vis spectrophotometer from Agilent Technologies. The blank standard was a 1:4 dilution of the medium. The concentration of flaviolin (c) was determined using the extinction coefficient ϵ= 3,068 M^{-1}.cm^{-1} [37] and Beer's

Law ($A = \epsilon Lc$). All experiments were done in triplicate. Results were reported as the concentration of flaviolin (μM) per gram wet-weight (gWW) of bacterial cells in a 1ml sample.

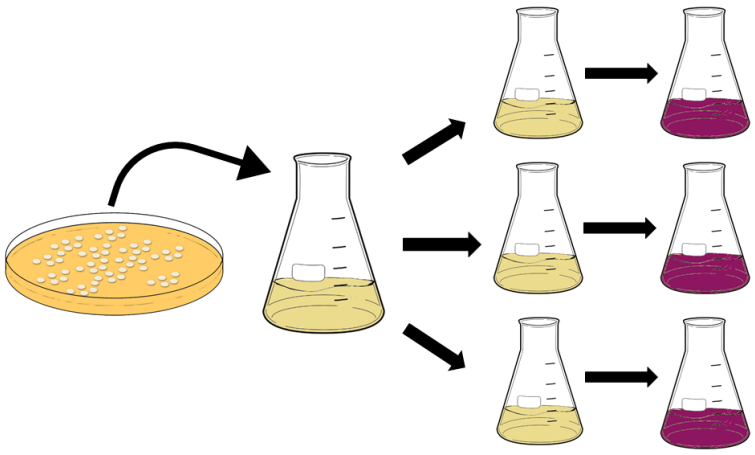


Figure 2.1. Fermentation steps utilized to maximize intracellular malonyl-CoA

2.1.3. Growth Media

E. *coli* cultures were grown in different media in order to understand the role of each macro and micro nutrient. The preparation method for media and stock solutions are described below.

Media

- LBT rich medium contained (per liter) used in the inoculum and cultures: 10 g tryptone, 5 g yeast extract, 5 g NaCl, and adjust the volume with tap water to 1000 ml. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- LBM rich medium contained (per liter) used in the inoculum and cultures: 10 g tryptone, 5 g yeast extract, 5 g NaCl, glucose and micronutrients solution as indicated, and adjust the volume with distilled deionized water to 1000 ml. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- 2XYT rich medium (per liter): 16 g tryptone, 10 g yeast extract, 5 g NaCl, and adjust the volume with tap water to 1000 ml. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- TB rich medium (per liter): 12 g tryptone, 24 g yeast extract, 4 ml glycerol, and adjust the volume with tap water to 1000 ml. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- M9 minimal media (per liter): 780 ml sterile water, 200 ml 5xM9 salts, 2 mL of 1M $MgSO_4$, 0.1 mL 1M $CaCl_2$, 5 ml of thiamine hydrochloride (5mg/ml), 1 mL 20% glucose, 3 ml of antibiotics, and adjust the volume with distilled deionized water to 1000 ml.

Stock Solutions

- To make 1M $MgSO_4$: 120.37 g dissolved in 1000 ml water. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- To make 1M CaCl2: 147.01g g dissolved in 1000 ml water. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

- To make 1M glucose stock solution: 900 mL DI water, 180.16 g glucose, and adjust the volume to 1000 ml. Filter sterilized by passing it thought a 0.22 μm filter and stored at 4 °C.

- To make 200g/L (20%) glucose stock solution: 900 ml DI water, 200 g glucose, and adjust the volume to 1000 ml. Filter sterilized by passing it through a 0.22 μm filter, and stored at 4°C.

- To make 1M $CaCl_2$-$2H_2O$: 15 ml DI water, 2.94 g $CaCl_2$-$2H_2O$, and adjust the volume to 20 ml. Filter sterilized by passing it through a 0.22 μm filter, and stored at room temperature.

- To make 0.1M $MnCl_2$-$4H_2O$: 7 ml DI water, 0.1979 g $MnCl_2$-$4H_2O$, and adjust the volume to 10 ml. Filter sterilized by passing it through a 0.22 μm filter, and stored at room temperature.

- To make 0.1 M $CuCl_2$-$2H_2O$: 7 ml DI water, 0.1705 g $CuCl_2$-$2H_2O$, and adjust the volume to 10 ml. Filter sterilized by passing it through a 0.22 μm filter, and stored at room temperature.

- To make 0.1 M $MgSO_4$: 17 ml DI water, 0.2407 g $MgSO_4$, and adjust the volume to 20 ml. Filter sterilized by passing it through a 0.22 μm filter, and stored at room temperature.

- To make 5xM9 salts (per liter): 900 ml DI water, 64g $Na_2HPO_4.7H_2O$, 15g. KH_2PO_4, 2.5g NaCl, 5.0 g NH_4Cl, and adjust the volume to 1000 ml. Sterilized by autoclaving for 25 minutes at 15 psi on the liquid cycle.

2.2. Mathematical Model

The mathematical model was developed following the workflow diagram illustrated in Fig. 2.2. The development of the model started with the collection of appropriate data from similar mathematical models existent in the literature [8, 13, 38, 39, 40]. After an iterative process where experimental data, computational simulations, and predictions were compared, the model was validated for use in metabolic engineering studies. The strategy followed to construct the kinetic model was:

1. Set simulation software: Matlab (R2015b, The Mathworks, Inc.);

2. Define metabolic network listing metabolites and reactions to be included in the model;

3. Define kinetics of each reaction based on literature review [8, 13, 38, 39, 40];

4. Build system of ODEs by applying a mass balance to each substrate, metabolite, product, and biomass;

5. Define initial conditions of substrates, internal metabolites, products, and biomass. Substrates, products, and the biomass initial concentrations were obtained from experiments, while internal metabolites were estimated based on literature review [41]. The mathematical model is not sensitive to these estimated values because the system promptly adjusts values of the internal metabolites after the first few seconds of simulation.

6. Obtain *in vitro* kinetic parameters from the mentioned literature, and if necessary calibrate them comparing with experimental data.

7. Simulate model and solve the ODEs system, providing quantitative data.

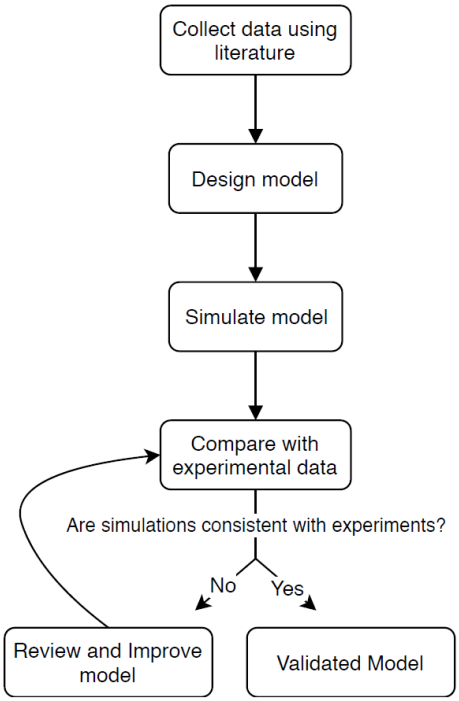


Figure 2.2. Workflow diagram for mathematical model construction

Chapter 3
Experimental Work

In order to maximize intracellular concentration of malonyl-CoA, the first step is to optimize ACC activity in the cell. Therefore, it is essential to measure its activity *in vivo*. Most previous studies utilize sophisticated and complex methods to quantify intermediate compounds, such as, High Performance Liquid Chromatography (HPLC), isotopic labeling, mass spectrometry, and enzymatic-specific assays [10, 15, 29, 36, 42]. All these high cost methods go against the requirement of developing an economically viable process. Therefore, the intracellular malonyl-CoA was quantified using the enzyme 1,3,6,8-tetrahydroxynapthalene synthase (THNS). THNS catalyzes condensation of five molecules of malonyl coenzyme A (CoA) to form 1,3,6,8-tetrahydroxynaphthalene (THN) (Fig 3.1). THN is readily converted into 2,5,7-trihydroxy-1,4-naphthoquinone (flaviolin) by auto-oxidation and secreted out of the cell [17]. Flaviolin is randomly polymerized to form a red-brown compound which shows a strong protection of the hosts against ultraviolet (UV) radiation [43, 44]. Therefore, the *in vivo* quantification of malonyl-CoA can be directly done measuring the absorbance of flaviolin with a spectrophotometer, making the quantitative data acquisition fast and economically feasible. The doubling time of a strain with THNS gene was seen as around 52 minutes, which is close to the doubling time of a wild strain (55 minutes), showing that the cell did not have its cellular growth compromised when the gene was overexpressed.

When E. *coli* cultures are grown to produce large quantities of bioproducts, the cost of production is important. It is therefore advantageous to grow cells in the medium that

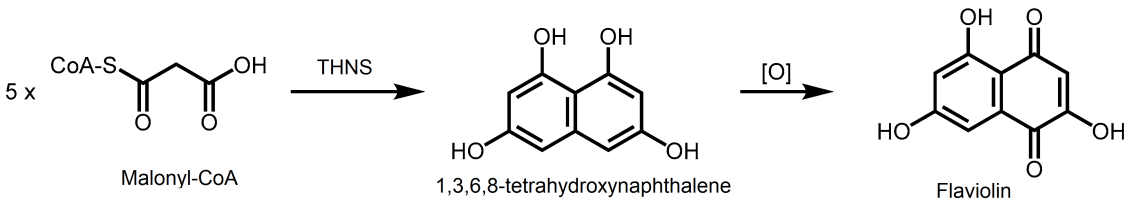


Figure 3.1. Reaction catalyzed by 1,3,6,8-tetrahydroxynaphtalene synthase (THNS) [2]

achieves the highest culture yield for a minimum amount of carbon source and nutrients. Sezonov et al. saw that growth in LB broth is carbon limited, indicating that it contains $<$ 100 μM fermentable sugar equivalents utilizable by E. *coli* (free sugars, sugar phosphates, oligosaccharides, nucleotides, etc.) [45]. Since LB broth lacks recoverable sugars and has high concentrations of catabolizable amino acids, these are likely depleted sequentially during the stationary phase of growth. This phenomenon causes a constantly variation in the physiological state of the cells. Furthermore, it was seen that the composition of LB is not constant, since it is composed of yeast extract, tryptone and sodium chloride these nutritional components vary from batch to batch when bought ready from manufacturer (i.e. batch A may have more yeast extract than batch B).

The reproducibility was difficult to reach when the experiment was repeated many times under the same conditions, indicating that the system is very sensitive, and that an accurate and careful control is required to guarantee an optimized system. In order to overcome this sensitivity issue, the effects of several parameters on the system behavior were carefully studied. Among the studied parameters were: time of incubation of the inoculum, incubation temperature, trace metals presence in the medium, carbon source availability in the inoculum, type of carbon source, minimal and rich medium, pH of the inoculum, aeration, amount and type of inducer for gene expression, time of induction and temperature shift (Fig 3.2). The results of these investigations allowed us to develop a tight method in which all conditions were well known and established .

3.1. Medium

3.1.1. Rich Medium

E. *coli* cells were grown in a LBT or 2XYT 10 ml inoculum with antibiotic where appropriate. The inoculums were incubated aerobically for 24 hours at 37°C. Cultures were grown in triplicate in 5 mL LB and a 1% (v/v) inoculum was introduced. No carbon supplementation was done during any step. Induction was done with the addition of 250 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was

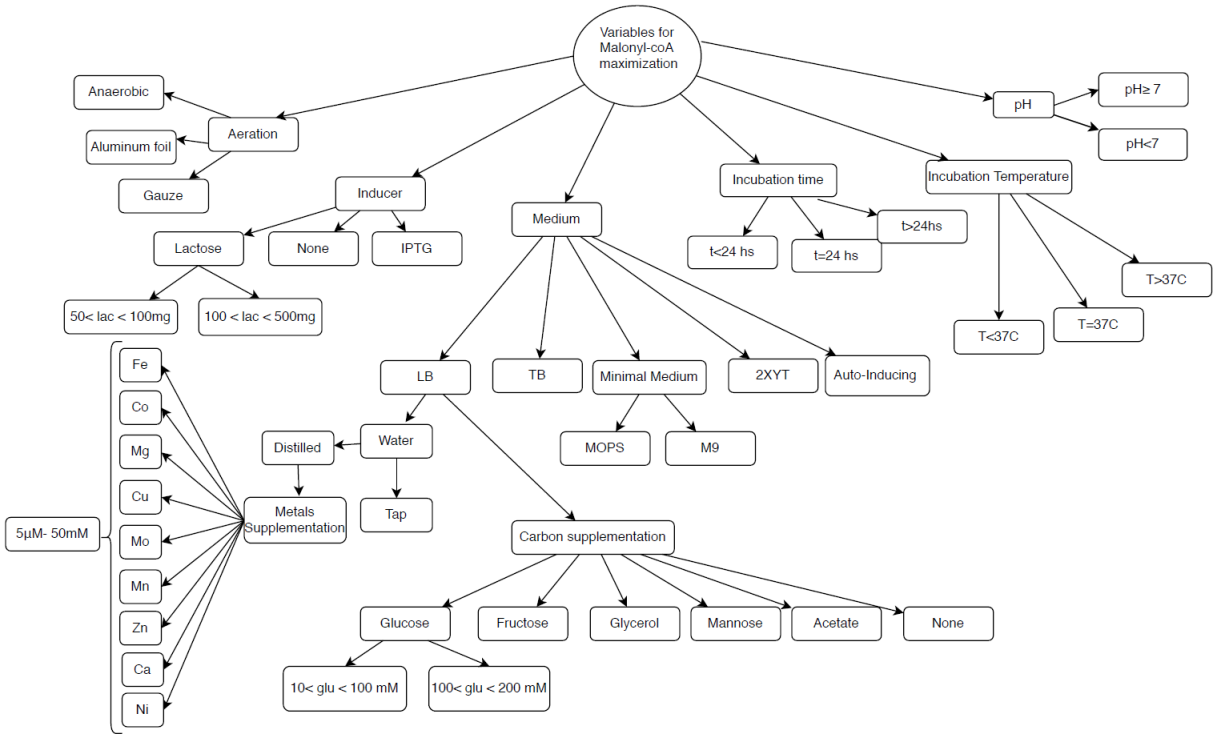


Figure 3.2. Experimental parameters studied to maximize malonyl-CoA production in E. *coli*

measured (Fig 3.3).

It was seen that a richer medium such as 2XYT was more effective in the production of flaviolin (Fig 3.3). The basic components of LB and 2XYT medium are tryptone, yeast extract, and sodium chloride. However, 2XYT has 50% more yeast extract and 62.5% more tryptone than the LB, which leads to the hypothesis that the lack of a carbon source in the inoculum is responsible for the difference in flaviolin production. Even without the addition of an inducer (lactose), the flaviolin production was higher in the inoculums made with 2XYT. Accordingly to [45], the main carbon sources in this medium are not sugars but catabolizable aminoacids. However, it was seen in [46] that rich media, such as LB and 2XYT, may cause an unintended induction. This hypothesis is reasonable since tryptone is an enzymatic digest of casein- a milk protein that contains lactose.

3.1.2. Carbon Supplementation

It is unclear how extra carbon sources such as glucose affect THN biosynthesis [47]. Sun et al suggested that different nutrients may enable E. *coli* to generate different metabolites

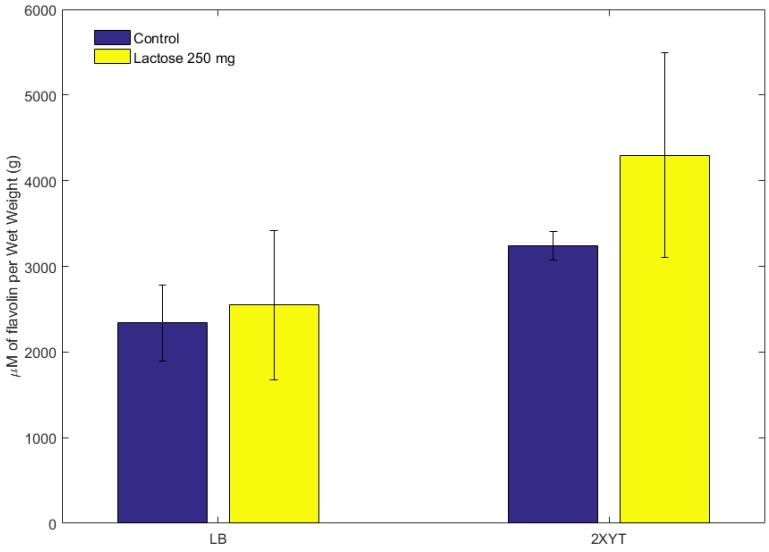


Figure 3.3. Effects of type of media in the production of flaviolin. The control group was not induced with lactose, while the second one indicates use of 250 mg

[47]. The addition of glucose yielded the production of four different compounds, while supplementation of sodium pyruvate into the induced broth of E. *coli* resulted in another two different bioproducts [47]. Therefore, to better understand the role of carbon sources on the production of malonyl-coA and ultimately, flaviolin, seven different carbon sources (glucose ($C_6H_12O_6$), glycerol ($C_3H_8O_3$), fructose ($C_6H_12O_6$), mannose ($C_6H_12O_6$), sodium acetate ($C_2H_3O_2Na$), sorbitol ($C_6H_14O_6$), and arabinose ($C_5H_10O_5$)) were used to supplement the 10 ml of LBT inoculum at different concentrations (0.1, 1, 10, 100 and 1000 mM). A control in which no carbon source was supplemented in the inoculum was also studied. The inoculums were incubated aerobically for 24 hours at 37 °C. Cultures were grown in triplicate in 5ml of LBT medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 100 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.4).

The addition of carbon sources to the medium (Fig 3.4) overcame the difference of product formation when using LB instead of 2XYT. During the exponential phase, the depletion of a carbon source in rich medium and the formation of waste products, inhibits

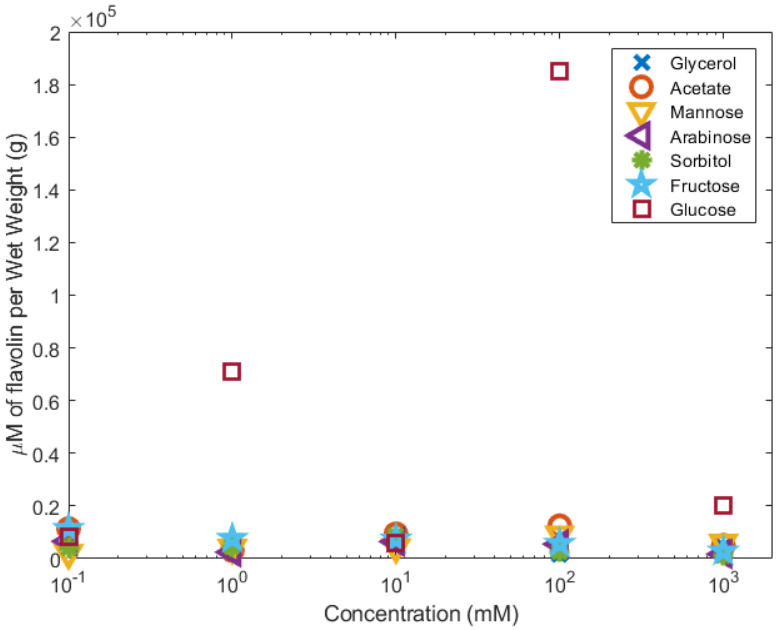


Figure 3.4. Carbon source effect (0, 0.1mM, 1mM, 10mM, 100mM and 1M) on the flaviolin production

cell growth and can be toxic to cells (Fig 3.5). This type of unstable growth is typically found in batch flask systems at the end of the growth curve when the substrate is nearly all consumed. Thus, supplementing glucose to the inoculum stabilizes the growth, providing carbon source to culture, and preventing an unintended induction [46].

3.1.3. Glucose Supplementation

From the results gathered in Fig. 3.4, glucose was chosen as the best carbon source to supplement the inoculum, stabilizing the growth while not compromising the induction of the gene of interest during the induction phase. Different concentrations of glucose (0-180 mM) were applied to inoculums made with LBT. Cultures were grown in triplicate in 5ml of LBT medium without carbon supplementation and with 1% (v/v) inoculum. Induction was done with the addition of 100 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.6).

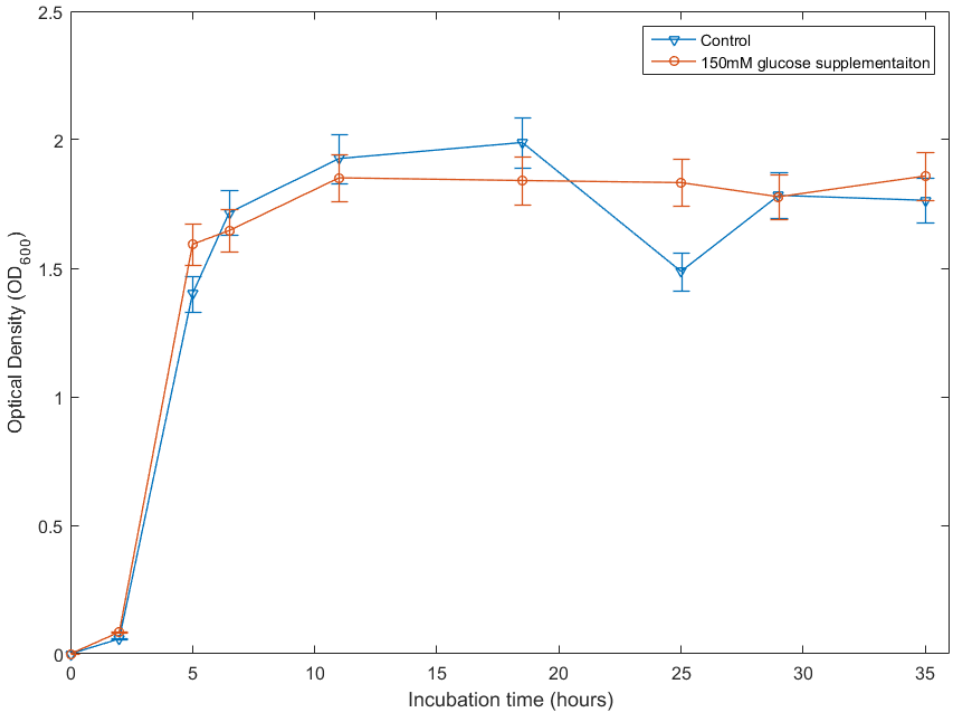


Figure 3.5. Bacterial growth curve comparing the influence of glucose supplementation in the inoculum

The glucose dependence experiment showed that there is an optimal glucose concentration for the inoculum which can result in a high production of flaviolin in the cultures. The range 60-80 mM of glucose supplementation showed the best results and stability for the malonyl-CoA maximization. Likely, metabolic pathways other than the one of our interest are being activated by certain amounts of glucose in the inoculum. Finding the right balance between the amount of glucose and the triggering of malonyl-coA to produce THN requires a tight method to guarantee reproducibility.

3.1.4. Minimal Medium

Samples 1 and 2 used LBT medium in the cultures (Table 3.1). 10 ml of LBT medium was used to prepare the inoculums. Sample 2 was inoculated with a 150 mM glucose supplemented inoculum. The inoculums were incubated aerobically for 24 hours at 37°C. Cultures were grown in triplicate in 5ml of LBT medium. Induction was done with the addition of 100 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin

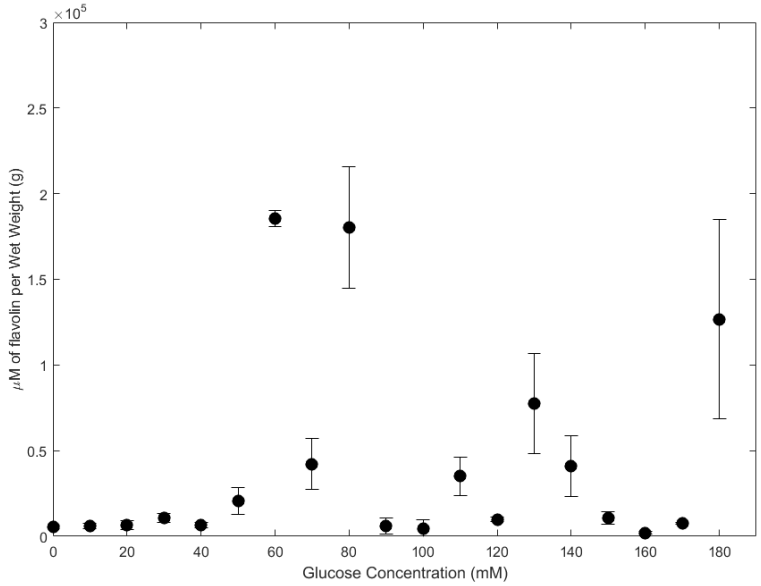


Figure 3.6. Production of flaviolin as a function of glucose concentration in the inoculums absorbance was measured. Samples 3 and 4 used M9 minimal medium in the cultures (Table 3.1). 20ml of fresh LBT medium was used to prepare inoculums. Sample 4 was inoculated with a 150 mM glucose supplemented inoculum. The inoculums were incubated aerobically for 24 hours at 37°C. Using an adapted procedure adopted by [48], cells were harvested by centrifugation (9500rpm/10minutes), washed and resuspended in triplicate in 5 ml M9 fresh medium without carbon sources (glucose), and a 1% (v/v) inoculum was introduced. Once OD_{600} reached 0.2 (approximately 1 hour), 0.1 ml of glucose 20% was added to the cultures. Protein expression was induced after all glucose was depleted (another 45 minutes of incubation) with 100 mg of lactose. Induced cultures were grown for another 22 hours aerobically in a water bath shaker (250 rpm) before flaviolin absorbance was measured (Fig 3.7).

Table 3.1. Medium utilized to prepare inoculum and cultures

Sample	Inoculum	Cultures
1	LBT	LBT
2	LBT 150mM Glu	LBT
3	LBT	M9
4	LBT 150mM Glu	M9

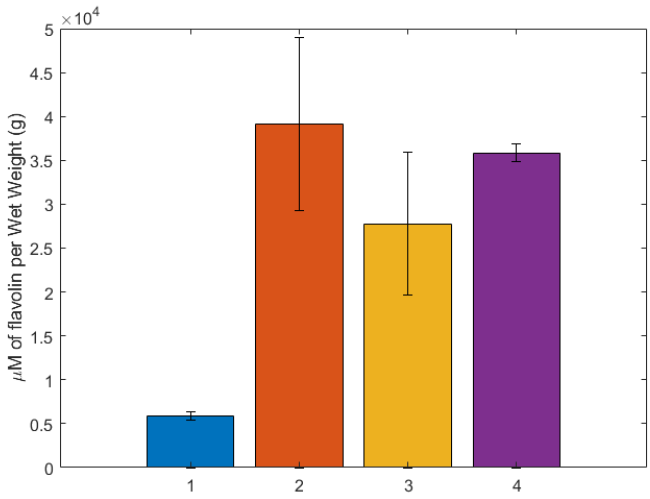


Figure 3.7. Minimal medium effect on the flaviolin production

As expected, cells grown in M9 Minimal Media revealed to contain all basic components for gene expression since they were previously grown in a richer medium supplemented with glucose, and then centrifuged and transferred to M9. Assuming that cells re-suspended in M9 Minimal Medium (without carbon and nitrogen sources) no longer have nutritive molecules in the medium, the residual LBT will be consumed at the beginning of the incubation, whereas approximately 1 hour of incubation ($OD_{600} \sim 0.2$) was enough to fully eliminate LBT. Consequently, since the cells were first grown in a richer medium, they contained all the main components of the molecular machinery, and increased the protein expression in comparison to full growth in M9 medium [48]. Although, the yield of flaviolin was not the maximum when comparing with LBT media, the primary advantages of a defined media, such as M9 Minimal Media, is that the results are more reproducible because the experiment can be better controlled, and it is cheaper than richer medium.

3.1.5. Richer Medium Supplemented with Glucose

10 ml of LBT and Terrific Broth (TB) media were used to prepare the inoculums (Table 3.2). 80mM glucose supplemented inoculum was used to inoculate samples A1, A2, C1 and C2. The inoculums were incubated aerobically for 22 hours at 37°C. Cultures were grown in triplicate in 5ml of medium and a 1% (v/v) inoculum was introduced. Induction was

done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.8).

Table 3.2. Medium utilized to prepare inoculum and cultures

Sample	Inoculum	Cultures
A1	LBT 80mM Glu	LBT
A2	LBT 80mM Glu	TB
B1	TB	LBT
B2	TB	TB
C1	TB 80mM Glu	LBT
C2	TB 80mM Glu	TB

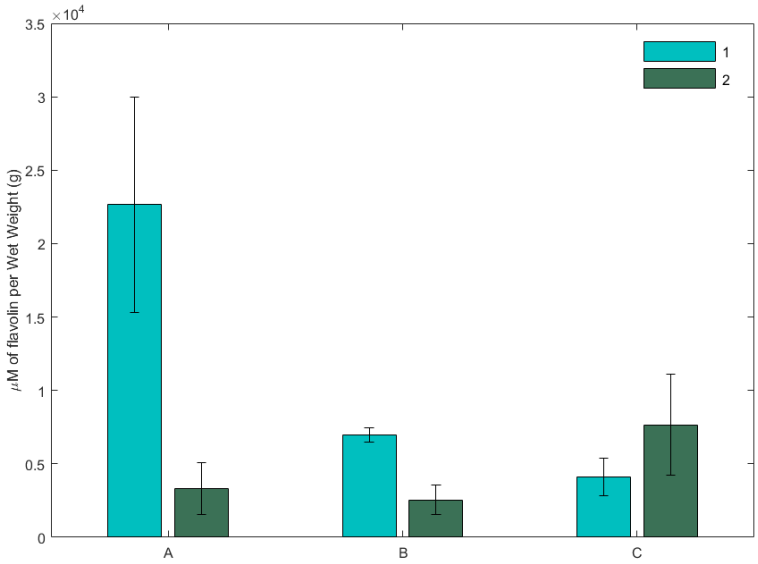


Figure 3.8. Richer medium effect on the flaviolin production

Fig. 3.8 compared the effect that richer medium supplemented with glucose has in the maximization of malonyl-CoA. Sample A1, which had the inoculum and culture made with less rich medium (LBT) had the best result when compared with the other samples. The fact of possessing more nutrients delays the establishment of a stationary phase, but it also increases the yield of viable cells [46]. It was seen that a high cellular growth is inversely proportional to the flaviolin production, probably due to the fact that the fatty acids synthesis pathway is activated, declining the concentration of free malonyl-CoA.

3.1.6. Metals Supplementation - Saturating the Medium

The activity of some enzymes depend on the binding of metals to its active site. Since it is unknown which metal and what best concentration maximizes malonyl-CoA, we can provide a high amount of different metals in order to saturate the production of target enzymes with little effect on growth. Therefore, individual metals and a mixture of trace metals were supplied for production of target proteins known and unknown to bind to specific metals. All flasks contained LBT medium while distilled deionized water was used in place of tap water. Each flask had different initial concentrations of metals in the inoculum and cultures. The different metals and their concentrations are summarized in Table 3.3. 10 ml of LBT medium supplemented with 80mM of glucose and the metals were used to prepare the inoculums. Inoculums were incubated aerobically for 23 hours at 37°C. Cultures were grown in triplicate in 5ml of LBT medium supplemented with metals and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.9).

Table 3.3. Metals concentration in the inoculum and medium

Sample	Metal Concentration in the Medium
1	-
2	0.2x Trace Metals, 50mM Fe, 2 mM Mg
3	0.4x Trace Metals, 2mM Mg
4	1x Trace Metals, 2mM Mg
5	5x Trace Metals, 2mM Mg
6	50 mM Fe
7	5 μM Mo
8	40 μM Co
9	100 μM Zn
10	200 μM Ca
11	20 μM Ni
12	2000 μM Mg
13	100 μM Mn
14	700 μM Cu

It is known that trace metals are required for maximal growth and protein expression in defined media. Once distilled water was used, it was assumed that it did not have

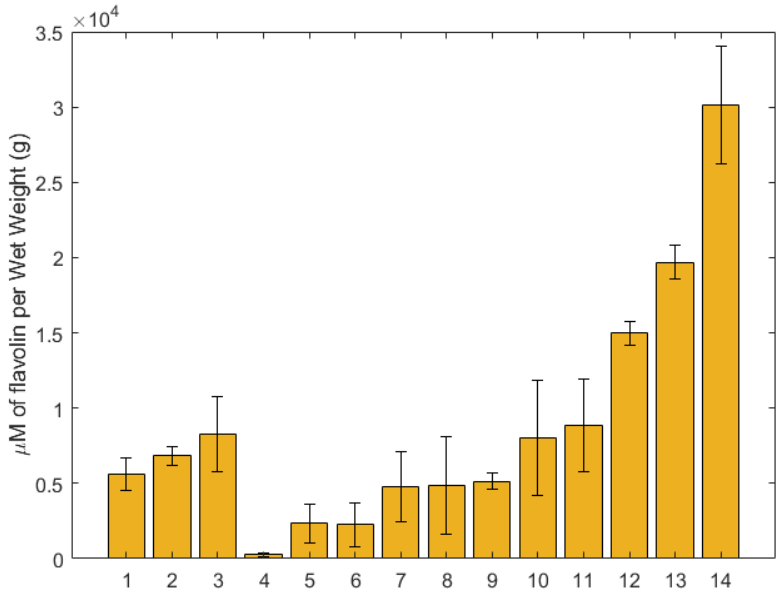


Figure 3.9. Saturation of target proteins with supplementation of different metals in the medium

substantial amounts of trace metals. Therefore individual metal ions and a mix of trace metals were tested at high concentrations in order to saturate target proteins and check for possible toxicity. The 1× concentration of metal mix supplies 20 μM calcium, 10 μM manganese, and zinc, and 2μM cobalt, copper, nickel, molybdate, and borate. Maybe these amounts that are not toxic to growth, saturated potential binding sites in target proteins once the results were minimal. Concentration of 5× trace metals was slightly inhibitory for the malonyl-CoA maximization but the cultures showed high density and growth. The presence of 50 mM of iron clearly inhibited the product formation in the samples 2 and 6 when comparing with samples 1 and 3. In the same way, molybdenium, cobalt, and zinc (Samples 7-9) when alone did not have a positive effect in the flaviolin production. On the other hand, calcium, nickel, magnesium, manganese and copper (Samples 10-14) showed a noticeable improvement in the malonyl-CoA maximization when present in the medium at high concentrations. The role of each metal is still unknown, but magnesium, for example, could be due to the fact that the malonyl-CoA synthesis requires ATP and magnesium ion for biotin carboxylation from acetyl-CoA. The color obtained in the copper

sample was considerably different from the color observed in samples which had a high amount of flaviolin. Therefore, the brown color observed in sample 14, may indicate that copper was oxidized and changed the solution color instead of representing a key metal for malonyl-CoA maximization.

3.1.7. Metals Supplementation - Mimicking Metals Concentrations in Water Samples

During experimentation, LBT medium was observed to be an important parameter that affects the production of malonyl-CoA. Since LBT was prepared using tap water, the concentration of metals could vary in each experiment. The amount of metals not only varies from day-to-day for a given site but also across sites due to the type and age of the piping system. Therefore, we used 5 different samples of tap water collected from different sites to investigate which one lead to the maximization of malonyl-CoA. Samples 4 and 5 showed the best results in the maximization of malonyl-CoA as shown in Figure 3.10. An elemental analysis of the water samples (Table 3.4) was done in order to determine metals composition of each one. Metals stock solutions of $CaCl_2$-$2H_2O$, $MnCl_2$-$4H_2O$, $CuCl_2$-$2H_2O$ and $MgSO_4$ were prepared in order to mimic the metal composition of the tap water from samples 4 and 5. Medium was prepared utilizing the procedure described for LBT, except for the use of distilled deionized water instead of tap water and the supplementation of the metals in the medium (Table 3.4). 10 ml of the aforementioned medium supplemented with 60 mM of glucose was used to prepare the inoculums. Inoculums were incubated aerobically for 22 hours at 37°C. Cultures were grown in triplicate in 5ml of this medium supplemented with metals and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.10).

After running the water elemental analysis, it was seen that Sample 5 from Table 3.4 resulted in the maximal production of flaviolin. As it was seen previously, calcium, magnesium, and manganese are important metals to the malonyl-CoA maximization. However,

its concentration cannot saturate target proteins in order to keep the cellular homeostasis. Consequently, medium LBM was defined with the addition of these three metals.

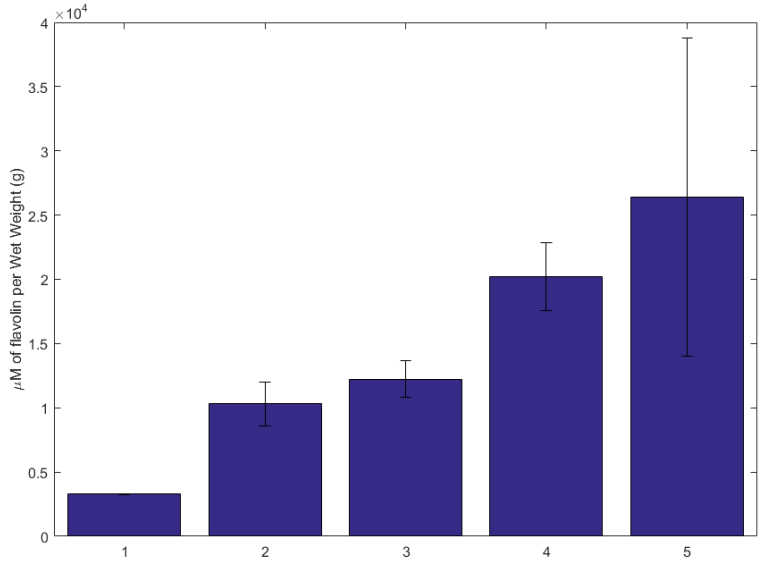


Figure 3.10. Effect of using different water samples in the flaviolin production

Table 3.4. Elemental analysis of water samples

Sample	Metal Concentration in the Water Sample
1	-
2	0.003 mg/L Al, 3.50 mg/L Ca, 0.088 mg/L Cu, 0.308 mg/L Mg, 0.016 mg/L Mn
3	0.048 mg/L Al, 1.42 mg/L Ca, 0.003 mg/L Fe, 0.185 mg/L Mg, 0.011 mg/L Mn
4	0.008 mg/L Al, 4.27 mg/L Ca, 0.063 mg/L Cu, 0.309 mg/L Mg, 0.032 mg/L Mn
5	0.035 mg/L Al, 1.58 mg/L Ca, 0.199 mg/L Mg, 0.012 mg/L Mn

3.2. Induction

3.2.1. Inducer Type

For industrial purposes, large quantities of protein are often required; thus the ability to maximize the protein yield from a specific culture volume is important. Proper protein folding is an important factor in the biological activity of overexpressed enzymes, while the inducer concentration and temperature of induction are the most important factors affecting protein folding.

The *in vivo* assay for ACC employs the plasmid pAER1 which contains the gene coding for THNS. This gene is controlled by lac operon and its expression can be induced with either lactose or Isopropyl β-D-1-thiogalactopyranoside (IPTG). To determine which inducer is more effective, E. *coli* cultures containing the plasmid pAER1 with the THNS gene were induced with various amounts of lactose or IPTG. The inoculum was prepared using LBM supplemented with 60 mM of glucose and incubated for 22 hours at 37°C. Cultures were grown in triplicate in LBM, and a 1% (v/v) inoculum was inoculated. The inducer was added at the time of the inoculation. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.11).

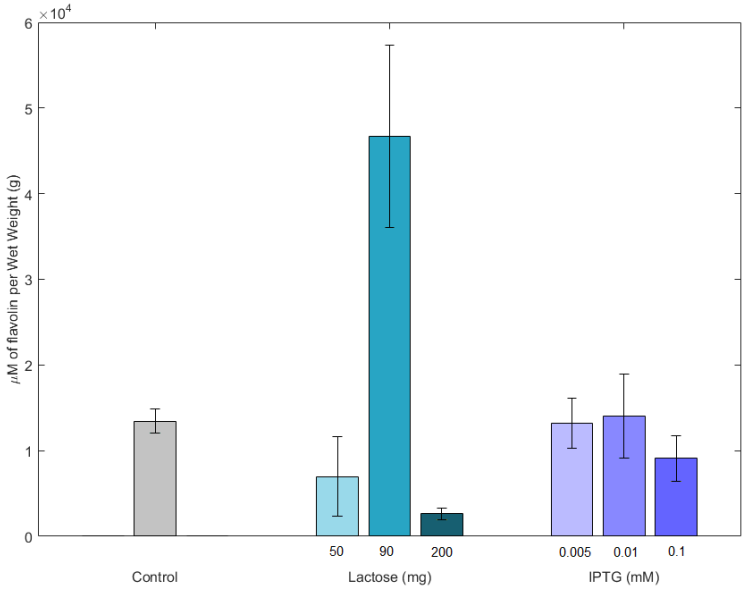


Figure 3.11. Inducer effect (Lactose vs IPTG) on the production of flaviolin

An ideal concentration of lactose is that one that not only induce a high gene expression, but also supplies maintenance in the bacterial metabolism. Lactose produced significantly more flaviolin per gram wet weight (gWW) than IPTG, and 90 mg of lactose produced the most flaviolin per gWW. Natural sugar inducers, such as lactose, have been shown to induce less stress and toxicity than IPTG in E. coli BL21(DE3) strains [49].

3.2.2. Temperature Shift and Time of Induction

It is known that heterologous proteins expressed at high levels in E. *coli* often fail to reach their native conformation and have a tendency to form inclusion bodies. However,this can be minimized by culturing cells at a reduced growth temperature [50]. Also, [51] saw that plasmid stability can be improved when the induction phase is carried out at low temperatures. Furthermore, since *lac* operon can be inhibited by glucose, the ideal time to induct the culture is when glucose is nearly exhausted. Usually, this occurs when OD_{600} reaches 0.3, which, in our case, corresponds to roughly 2 hours of incubation.

Further experimentation was done to investigate how temperature and time of induction affect the overexpression of our target protein. With an 80 mM glucose supplemented LBT inoculum incubated for 24 hours at 37°C, four cases of temperature shift were performed. Cultures were grown in triplicate in 5ml of LBT medium without carbon supplementation, and a 1% (v/v) inoculum was introduced. The time of induction was also evaluated, by inducing half of the samples with 100 mg of lactose just after inoculating the inoculum ($t_{induction}$=0 hour). The other half was induced when OD_{600} = 0.3 (2 hours of growth) was reached ($t_{induction}$=2 hours). After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.12). Blue bars had the temperature kept at 37°C during the induction (T_{ind}) and the growth (T_{gwth}). Orange bars had the inducer added at T_{ind}=25°C , after that the temperature was shifted to T_{gwth}=37°C. Yellow bars had the inducer added at T_{ind}=37°C , after that the temperature was shifted to T_{gwth}=25°C. Purple bars had the temperature kept at 25°C during induction and growth.

There was no statistical difference between adding lactose at inoculation or after inoculation (Fig 3.12). Therefore, for convenience, genes were induced at inoculation for all of the subsequent experiments.

3.3. Inoculum

It is known that some metabolic pathways are activated in specific ranges of temperature. For this reason, we tested the impact of the temperature in the inoculum phase by

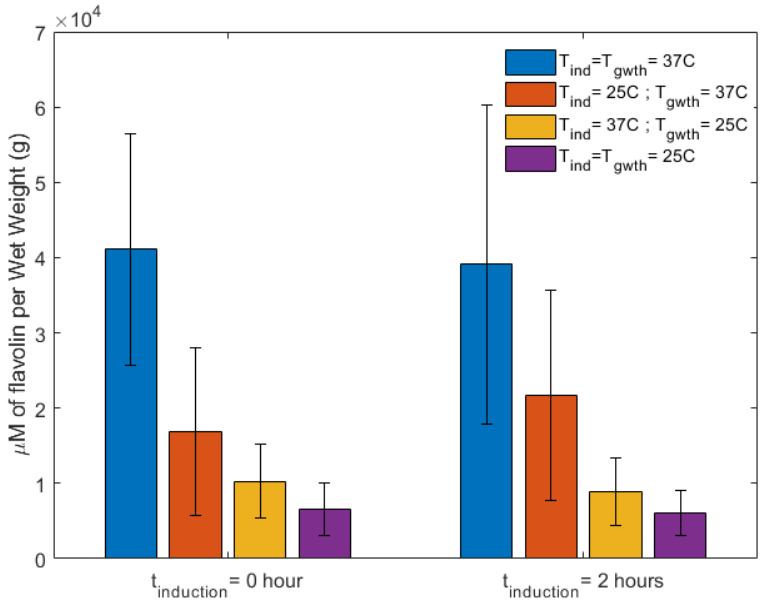


Figure 3.12. Flaviolin production when varying the time and temperature of induction

cultivating samples at different temperatures. E. *coli* cells were grown for 22 hours at 30, 37, and 39°C in inoculums made with LBM medium supplemented with 60 mM of glucose. Cultures were grown in triplicate in 5ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation, the flaviolin absorbance was measured (Fig 3.13).

Usually, standard microbiology procedure indicates "overnight" as the period to incubate the inoculum. However, usually the OD_{600} is not specified, and neither the number of hours needed to optimize genes overexpression. In order to determine the number of incubation hours that maximizes THNS during induction phase, an experiment varying the time for incubation of the inoculum was done. E. *coli* cells were grown at 37°C in inoculums made with LBM medium supplemented with 60 mM of glucose. Inoculums were incubated for 19.5, 20, 20.5, 21, 21.5, 22 and 22.5 hours, and the respective optical density at 600 nm (OD_{600}) was measured. Cultures were grown in triplicate in 5ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures,

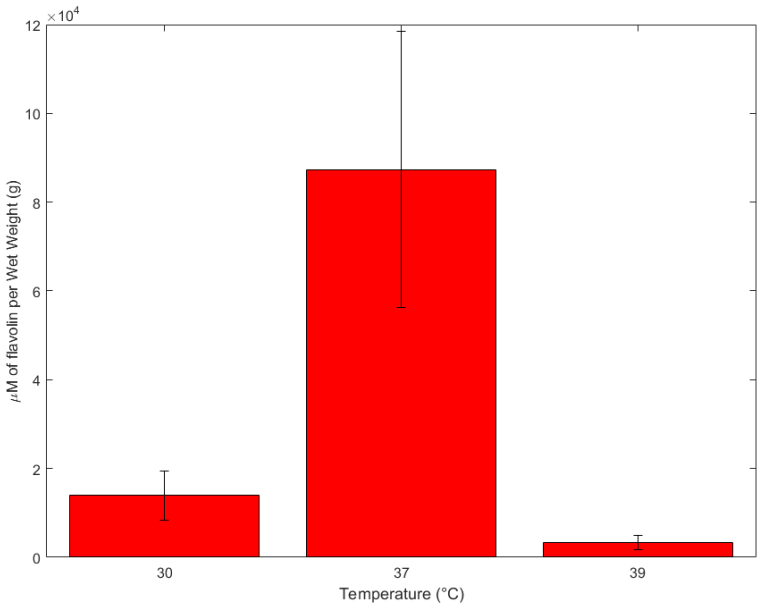


Figure 3.13. Temperature influence in the production of flaviolin

the flaviolin absorbance was measured (Fig 3.14). Blue circles of Fig 3.14 represent the production of flaviolin (left axis) in cultures made from inoculums that were incubated for different periods, and red squares represent the OD_{600} (right axis) of these inoculums.

The utilization of an inoculum was seen as a key feature in the proposed method to maximize malonyl-CoA. The bacterial growth begins after lag phase, which is defined as the transition to the exponential phase after the population has doubled. Lag phase is when the cell physiologically adapts to the culture conditions [52]. Therefore, the protein synthesis and gene induction may take some time lasting from minutes to hours. However this period can be adjusted since it depends on the type of medium and inoculum size. As a result, since the cultures were made from the stationary phase inoculum, there will be a small lag phase as the stationary phase cells adjust to the new conditions and shift physiologically from stationary phase cells to exponential phase cells. Moreover, once the cultures grew in a medium, which does not have glucose, and the inoculum cells have consumed all the glucose that were grown, there is no physiologically adaptation to adjust these cells to the new medium, resulting in a insignificant lag phase length, which increases the rate of product formation.

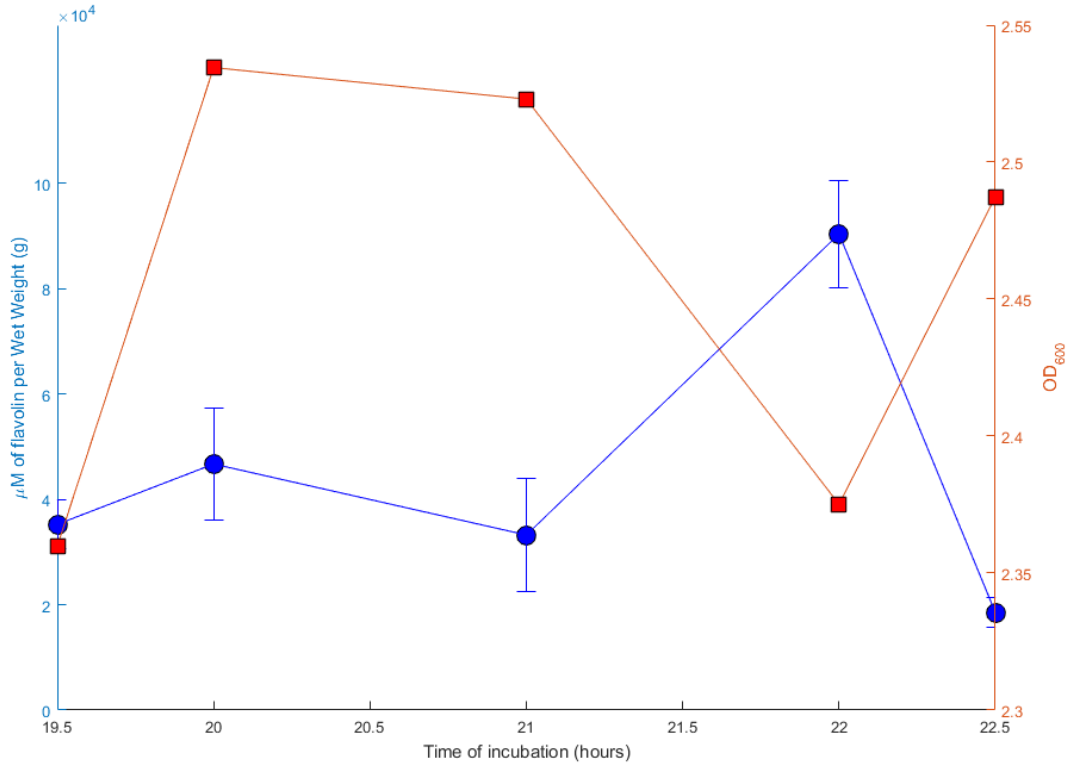


Figure 3.14. Effect of time of incubation of the inoculum in the production of flaviolin

We observed an acidic pH (4-5) in the cultures grown at 30°C (Fig 3.13), which was likely due to a different metabolic pathway being activated, resulting in the low production of flaviolin in these samples or the lag phase length being increased since the cells had to adapt to the new environment. The same behavior was seen when the inoculum was kept at 39°C, indicating that a tight control of the incubation temperature at 37°C is essential to maximize malonyl-CoA production.

The OD_{600} of the inoculum varied from 2.35 to 2.54 within a period of 3 hours (Fig 3.14). The production of flaviolin was maximum when the optical density decreased abruptly, showing that there is an optimum metabolic condition at which the cells need to be harvested and used to inoculate the cultures. Therefore, it was defined that the incubation time of the inoculum would be when the OD_{600} is suddenly lowered at the stationary phase, which corresponds to approximately 22 hours.

3.4. Aeration

It is known that keeping a reasonably good aeration is essential to maintaining a neutral pH and obtaining a good cellular growth [46]. Furthermore, the biosynthesis of microbial products in shake flasks may be limited by inadequate supply of oxygen to the cultures [53]. Therefore, we tested the influence of three types of closure in order to understand the role of oxygen supply in the production of malonyl-CoA. E. *coli* cells were grown for 22 hours at 37°C in inoculums made with LBM medium supplemented with 60 mM of glucose. Cultures were grown in triplicate in 5ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. Both flasks utilized to make the inoculum and the cultures were covered with aluminum foil with holes, cotton gauze, or parafilm. Three layers of parafilm sealed completely the exchange of air, leading to an anaerobic environment inside the flask. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.15).

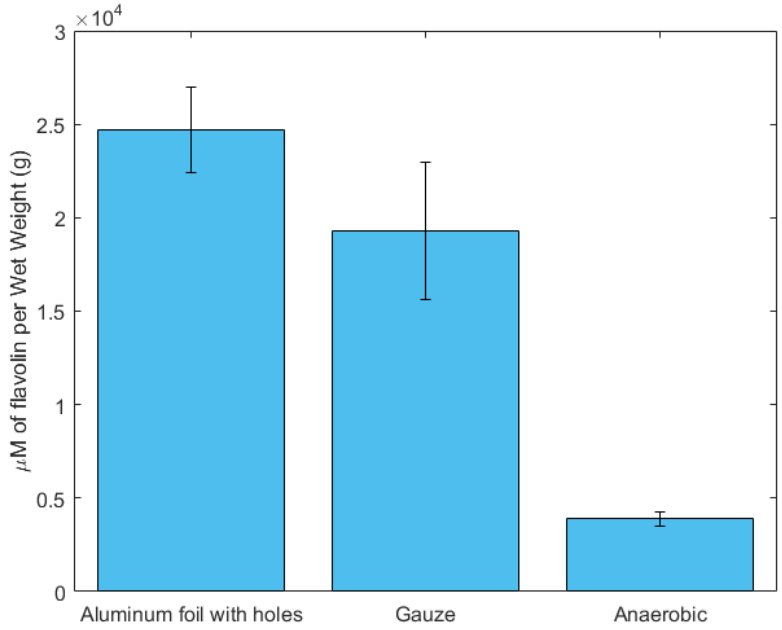


Figure 3.15. Effect of the types of closure in the flaviolin production

The rate of aeration is not only important for the bacterial respiration and auto-oxidation of THN to flaviolin, but it also has an important effect on saturation density,

acidity, and induction [46]. Figure 3.15 showed that covering the flasks with aluminum foil with holes on top to increase the aeration increases the production of flaviolin when compared with gauze covering, or when constraining the transfer of air to the cultures. Also, Studier et al [46] concluded that a higher rate of aeration demands more lactose to induce a high-level of protein production. Therefore an optmized aeration is desired in order to use the minimum amount of lactose to maximize malonyl-CoA.

3.5. Overexpressing Acetyl-CoA Carboxylase

Since the initial objective of optimizing acc activity *in vivo* is to increase production of malonyl-CoA, the original hypothesis was that this could be accomplished by overproduction of holo ACC. Therefore, the effect of overexpressing holo ACC genes on the amount of flaviolin produced was investigated. The strain consisted of pAEP7 and pSEB1, which together contained the genes for THNS and all three ACC subunits (one gene coding for BCCP and BC; two genes coding for CT) (Table 2.1). Additionally, in order to provide enough biotin ligase which is responsible for the biotinylation of BCCP, another strain containing pAER1 and pLB0056, which coded for the genes of all three ACC subunits as well as for biotin ligase, was tested. Three strains of E. *coli* (pAER1, pAEP7+pSEB1, pAER1 + pLB0056) were grown for 19.5 hours at 37°C in inoculums made with LBM medium supplemented with 60 mM of glucose. Cultures were grown in triplicate in 5 ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.16).

The effect of overproducing all three ACC subunits did not result in an increase in the level of malonyl-CoA (Fig. 3.16) as many other studies have shown [10, 17, 27, 30, 31, 36]. Neither the strain containing paEP7 and pSEB1, or the strain containing pAER1 and pLB0056 result in an increase in flaviolin biosynthesis. Therefore, overexpressing the three subunits of ACC, as well as biotin ligase, were not sufficient to maximize malonyl-CoA as compared with the control (pAER1).

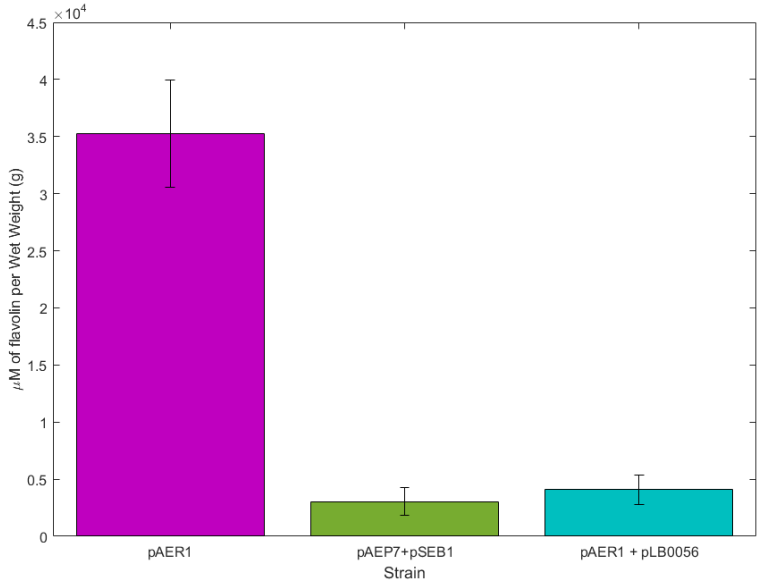


Figure 3.16. Effect of overexpression of holo ACC (pAEP7+pSEB1) and holo ACC with biotin ligase (pAER1+pLB0056) comparing with only pAER1

3.6. pH

In order to define the pH level that maximizes malonyl-CoA, the pH of LBM medium was adjusted from 6.7 to 7, 7.4 and 8.4 with addition of NaOH 5N. E. *coli* cells were grown for 20 hours at 37°C in inoculums made with LBM medium supplemented with 60 mM of glucose. Cultures were grown in triplicate in 5ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. After 24 hours of incubation of the cultures, the flaviolin absorbance was measured (Fig 3.17)

Previous research has revealed that environmental pH partly influences the cytoplasmic or intracellular pH, which affects enzyme activity and reaction rates, protein stability, structure of nucleic acids, and many other biological molecules [54]. Adjusting the medium pH from 6.7 to 7 increased substantially the flaviolin production. This might be a consequence of, when glucose is abundant in the inoculum, cells grow rapidly and the acetate assimilation pathway is activated ('acetate overflow') [10]. High cellular concentration of acetate decreases the pH, resulting in a toxic environment for the cells which inhibits growth [55].

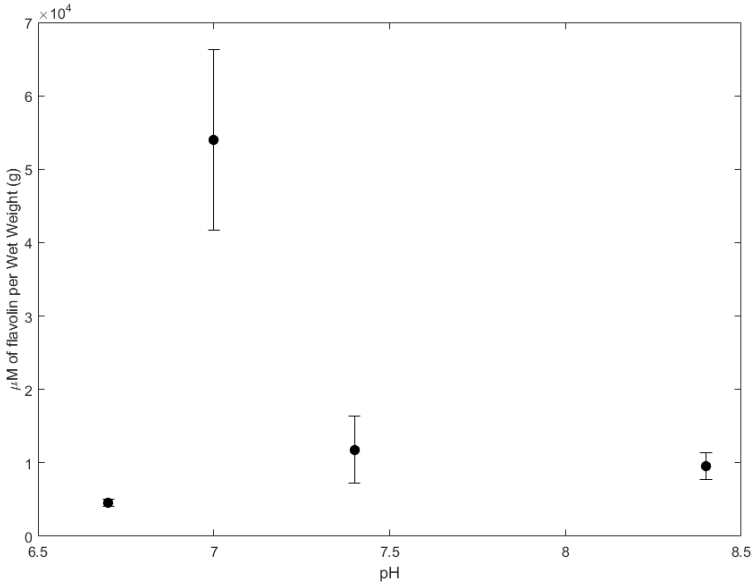


Figure 3.17. Effect of adjusting the medium pH to 7 in the flaviolin production

Consequently, the acetate pathway is activated, and enzymes such as phosphotransacetylase and acetate kinase convert acetyl-CoA to acetate, turning the environment toxic and inhibiting the pathway of interest. Decreasing the undesirable conversion from acetyl-CoA to acetate by adding more alkaline to the pH may be beneficial for malonyl-CoA-derived bioproducts. Therefore, the adjustment of the pH from 6.7 to 7 can prevent the undesired activation of others metabolic pathways and guarantees the control of flaviolin production.

3.7. Time Incubation

The number of hours for the biomass to convert the substrate into the product is important for industry in order to maximize the production of flaviolin per gram of cells. Thus, E. *coli* cells were grown for 20 hours at 37°C in inoculums made with LBM medium with pH adjusted to 7 and supplemented with 60 mM of glucose. Cultures were grown in triplicate in 5ml of LBM medium without carbon supplementation and a 1% (v/v) inoculum was introduced. Induction was done with the addition of 90 mg of lactose. Absorbance and wet weight of the culture were measured after the following hours of incubation: 4, 8, 12, 20, 24, and 48 (Fig. 3.18).

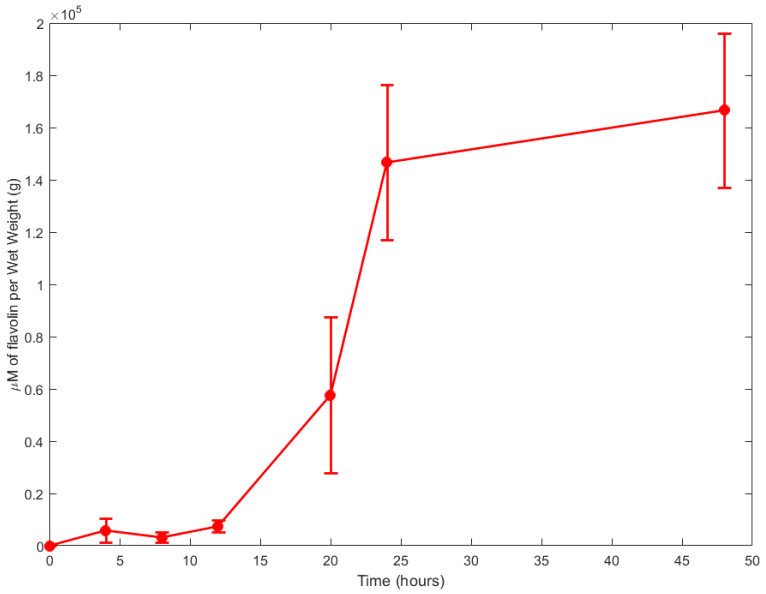


Figure 3.18. Production of flaviolin along the time of incubation

The time course experiment showed that incubating the cultures for 24 hours guarantees the maximum production per amount of cell with less error associated. Specifically, the highest production of flaviolin was 150 mM/gWW. We observed an inverse proportion relationship between the amount of biomass and the flaviolin production. This is expected since the malonyl-CoA produced during the first hours of induction is promptly consumed by THNS, resulting in a low availability of this substrate for the fatty acids synthesis. Consequently, smaller biomass production is observed not only because of the time course, but also due to the malonyl-CoA availability.

3.8. Process Economics

The operational cost to run the proposed method is the sum of raw materials, and a shaking bath incubator. Raw materials account for the cost of all fermentation media used in the inoculum, culture, and agar plate. Materials cost was estimated based on online pricing information (Table 3.5 and A.1).

Table 3.5. Materials used in the fermentation process

Consumables	Price ($/gram)	Quantity (g/5ml)	Quantity (g/ L)	Cost ($/5 ml)	Cost ($/L)
Antibiotic	0.785	0.0015	0.0015	0.0012	0.0012
Yeast Extract	0.1762	0.1	4.8	0.0176	0.8458
Triptone	0.3112	0.2	2.4	0.0622	0.7469
Sodium Chloride	0.063	0.1	2.4	0.0063	0.1512
Agar	0.29	0.5	0.5	0.145	0.145
Glucose	0.031	0.1802	0.1802	0.0055	0.0056
Calcium Chloride	0.176	0.0006	0.1411	0.0001	0.0248
Magnesium Sulfate	0.147	9.8567e-05	0.0237	$1.44893e^{-5}$	0.0035
Manganese Chloride	0.242	$1.978e^{-6}$	0.0005	$4.78676e^{-7}$	0.0001
Lactose	0.0526	0.09	21.6	0.0047	1.1362
Distilled Water	$2.5e^{-5}$	50	1200	0.0013	0.03
			Total	0.2440	3.0902

It was calculated that the cost to produce 1 liter of fermentation broth is approximately $3.09. Therefore, considering an average of 68 mM/gWW of flaviolin with 24 hours of incubation, and 2.554 mM of flaviolin production in 5 ml of fermentation broth, the product yield is 0.03 gram of flaviolin per gram of biomass, while the titer is 1.053 grams of flaviolin per liter of broth. Meaning that it would cost $3.09 to produce approximately 1 gram of flaviolin.

However, the ultimate goal is not to produce flaviolin but to maximize malonyl-CoA. The reaction of Figure 3.1 assists on the estimation of malonyl-CoA production with the proposed method. Once five molecules of malonyl-CoA are converted into flaviolin, the malonyl-CoA concentration would be five times the flaviolin concentration, which was measured spectophotometrically. Therefore, using an average of 2.554 mM of flaviolin produced after 24 hours of incubation, the intracellular malonyl-CoA concentration would be 12.77 mM resulting in a titer of 10.9 g/L with $0.29 per gram of malonyl-CoA. Table 3.6 summarizes these results.

Table 3.6. Average indicators for flaviolin and malonyl-CoA production

		Flaviolin	Malonyl-CoA
Production	(mM/gWW)	68	340
Concentration	(mM)	2.55	12.77
Titer	(g/L)	1.053	10.9
Cost	($/g)	3.09	0.29

The experimental method can be scaled up to large volume bioreactors using the current industrial infrastructure of biochemical plants. The operating costs associated with the scale up can be estimated based on Table 3.7. Raw materials represent the biggest expenditure; therefore, utilizing a low cost method as the one described in this work would result in an attractive investment. Assuming that the investor has most of the equipment, the cost to implement the method would be minimum.

Table 3.7. Operating cost items and ranges [1]

Cost Item	Type of Cost	Range of Values (% of Total)
Raw Materials	Direct	10-80
Labor	Direct	20-50
Consumables	Direct	1-50
Lab/QC/QA	Direct	2-50
Waste Disposal	Direct	1-20
Utilities	Direct	1-30
Equipment-Dependent	Indirect	10-70
Miscellaneous	Indirect	0-20

Chapter 4
Mathematical Model

One of the most challenging goals is the design of biological systems based on quantitative predictions with the aid of mathematical models. Several obstacles are encountered when addressing quantitative design. This is because of the well-known difficulties in measuring enzyme kinetics under *in vivo* conditions since this is a precondition for a reliable quantitative analysis of the metabolic pathway.

There is a need for decreasing the complexity of metabolic networks since the many biochemical details of metabolic networks can be overwhelming. When the purpose of the model is to amplify the flux of specific products in the industrial manufacturing of metabolites scenario, the main task is reduced to an amplification of the central metabolic pathways.

Another important factor for generating the model is the range of validity. It is known that dynamic models are based on measurement of intracellular responses to external disturbances, therefore many of the effects observed can be attributed to cellular functions beyond the intended functions [8]. Consequently, to avoid the superposition of side effects, the time scale of the experimental observation used for model identification should be as brief as possible. Other challenges in engineering of biological systems are the environmental conditions, genetic mutations, crosstalk, cell death, and incomplete models.

4.1. Kinetic Modeling

This chapter presents the development and results of a kinetic model for the fermentative metabolism of glucose in E. *coli*, and the use of the model for the clarification of the control structure of the associated metabolic pathway. Despite efforts to establish pathways and mechanisms of glucose fermentation in E. *coli* and malonyl-CoA maximization, the quantitative analysis of this metabolic process had not been performed before this thesis. The dynamics and steady state of metabolites, substrates, products, and fluxes

(also called metabolic network) can be quantitatively predicted by kinetic models. Kinetic models are composed of a set of ordinary differential equations (ODEs) depicting enzymatic events that occur inside the cell and transport reactions through the cytoplasmic membrane.

With the aim of describing the effects of changes in enzyme activity on metabolite concentrations and reaction fluxes, a stability analysis and a sensitivity approach called Metabolic Control Analysis are conducted.

4.1.1. Model of Central Carbon Metabolism of E. *coli*

The system is a microorganism (E. *coli*), in which many of the reactions are catalyzed by enzymes. The kinetics of one reaction is defined by the rate law of the reaction plus the value of all involved kinetic parameters. Michaelis-Menten equation, for instance, is most widely used relationship to derive enzyme kinetics relationships. This equation describes an enzymatic kinetic reactions that transform irreversibly one substrate into one product. A mass balance for each metabolite is defined as an ODE system using the kinetics of each reaction.

The dynamic model of the glycolysis pathway, and the conversion of pyruvate to malonyl-CoA involves mass balance equations for extracellular glucose and for the intracellular metabolites. The mass balance equations of the metabolites take on the following form [8]:

$$\frac{dC_i}{dt} = \sum_j a_{ij} v_j - \mu C_i$$

where C_{ij} is the concentration of metabolite i, μ is the specific growth rate, and a_{ij} is the stoichiometric coefficient for this metabolite in reaction j, and v_j is the reaction rate.

The key feature of the kinetic model is that it predicts the reaction rate at given values of the concentrations of substrate(s) and product(s). Therefore, considering a generic form for each enzymatic reaction v_i:

$$v_i = v_{imax} f_i(C, k)$$

where v_{imax} is the maximum rate of the reaction v_i and f_i is a function dependent on metabolite concentrations C and a set of parameters k [8].

Based on [56], which analyzed and compared the most used and complete kinetic models, the model in [39] was chosen as reference for the developed model. That is, Kadir et al model [39] can be modified and simplified in order to quantitatively predict the physiological behavior of E. *coli* and its production of malonyl-CoA.

The experimental conditions consisted of the aerobic growth of E. *coli* cultivated in LBM, derived from 60 mM glucose inoculum, and induced with 90 mg of lactose (50 mM). E. *coli* and other bacteria can feed on both lactose and glucose, but when both sugars are present glucose is consumed before lactose. This phenomenon is known as diauxic growth [57], and it represents an optimal thermodynamic solution given that glucose is a cheaper energy source than lactose. That is to take advantage of lactose, the bacteria needs to expend energy in producing the enzymes needed to transport and metabolize this sugar [58]. These enzymes are the permeases, which are localized on the outer membrane and are responsible for uptaking the external lactose and transport it inside the cell. A regulatory phenomenon known as inducer exclusion happens when glucose inhibits the functionality of the permeases, decreasing the rate of which external lactose is uptaken. This mechanism was included in the kinetic model, assuming that the transport of lactose followed saturation kinetics and was inhibited by the concentration of extracellular glucose [57].

The kinetic model has the phosphotransferase system (PTS) to uptake external glucose and convert it into glucose-6-phosphate (G6P), and permeases to uptake lactose which will be converted to G6P by β-galactosidase. G6P enters into the glycolysis pathway, which has pyruvate as the ultimate product. Pyruvate dehydrogenase converts pyruvate into acetyl-CoA, which is then converted into malonyl-CoA by ACC catalysis. Finally, THNS converts malonyl-CoA into THN which is auto-oxidized into flaviolin. This last reaction

does not depend upon cell metabolism, and therefore was not included in the model (Fig 4.1). The metabolites were modeled as variables in the model, and its nomenclature and initial concentrations are listed in Table 4.1.

Based on mass balance, the kinetic rates are summarized in Table 4.2, where the kinetic parameters abbreviations and values are in Table 4.3. Most of the kinetic rates and types of glycolysis were taken from analytical experiments reported in the literature (Table 4.4).

Table 4.1. Initial concentration of metabolites

Metabolite	Abbreviation	Concentration (mM)
Biomass	X	4.8
External Glucose	GLCEX	30
External Lactose	LACEX	50
Internal Lactose	LACIN	0
Phosphoenolpyruvate	PEP	0.01
Pyruvate	PYR	0.01
Glucose-6-Phosphate	G6P	0.01
Fructose-6-Phosphate	F6P	0.01
Fructose-1,6-biphosphate	F16BP	0.01
Glyceraldehyde-3-Phosphate	GAP	0.01
Acetil-coenzyme A	AcCoA	0.01
Malonyl-coenzyme A	McoA	0.01
1,3,6,8tetrahydroxynapthalene	THN	0

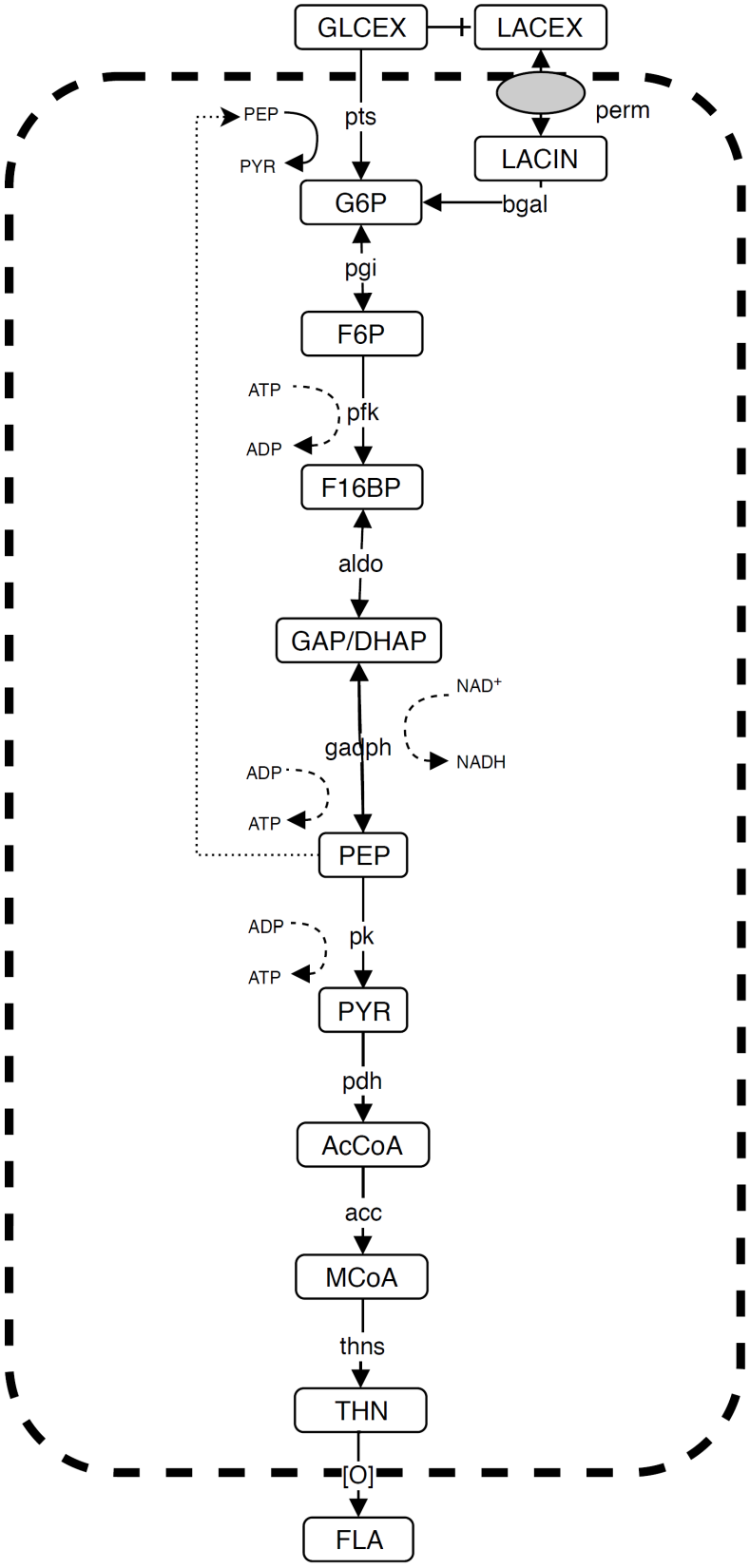


Figure 4.1. Malonyl-CoA metabolic pathway

Table 4.2. Kinetic Rates

Reaction	Equation	Reference
Cell Growth	$\mu = \mu_m \left(1 - \frac{[X]}{X_m}\right) \left(\frac{[G6P]}{K_s + [G6P]}\right)$	[39]
Phosphotransferase System	$\mathrm{v}_{pts} = \frac{v_{max,pts}[GLCEX]\frac{[PEP]}{[PYR]}}{\left(K_{pts,a1} + K_{pts,a2}\frac{[PEP]}{[PYR]} + K_{pts,a3}[GLCEX] + [GLCEX]\frac{[PEP]}{[PYR]}\right)\left(1 + \frac{[G6P]^{nG6P}}{K_{pts,G6P}}\right)}$	[39]
Lactose Uptake	$\mathrm{v}_{perm} = k_{LACIN}[PERM]\left(\left(\frac{[LACEX]}{[LACEX]+K_{t,lac}}\right)\left(\frac{K_{i,glu}}{K_{i,glu+[GLCEX]}}\right) - \frac{[LACIN]}{[LACIN]+K_{t,Lacin}}\right)$	[57]
β-Galactosidase	$\mathrm{v}_{betagal} = k_{cat,lac}[\beta GAL]\left(\frac{LACIN}{[LACIN]+K_{m,lac}}\right)$	[57]
Phosphoglucoseisomerase	$\mathrm{v}_{pgi} = \frac{v_{max,\ pgi}\left([G6P] - \frac{[F6P]}{K_{eq,pgi}}\right)}{K_{m,G6P}\left(1 + \frac{[F6P]}{K_{m,F6P}}\right) + [G6P]}$	[39]
Phosphofructokinase	$\mathrm{v}_{pfk} = \frac{v_{max,\ pfk}K_{pk,ATP}[F6P]}{K_{pk,ATP,ADP}\left([F6P] + K_{s,F6P}K_{b(ADP,AMP)} + \frac{[PEP]}{K_{pfk,PEP}}\right)\left(1 + \frac{L_{pfk}}{\left(1+[F6P]\left(\frac{K_{a,(ADP,AMP)}}{K_{s,F6P}\left(K_{b(ADP,AMP)} + \frac{[PEP]}{K_{pfk,\ PEP}}\right)}\right)^{n_{pfk}}\right)}\right)}$	[39]
Aldolase	$\mathrm{v}_{aldo} = \frac{v_{max,aldo}\left([F16BP] - \frac{[GAP]^2}{K_{aldo,eq}}\right)}{(K_{aldo,\quad F16BP} + [F16BP] + \frac{K_{aldo,GAP}[GAP]}{\left[K_{aldo,eq}V_{blf}\right]} + \frac{K_{aldo,DHAP}[GAP]}{\left[K_{aldo,eq}V_{blf}\right]} + \frac{[F16BP][GAP]}{K_{aldo,gap\ \ inh}} + \frac{[GAP]^2}{\left[K_{aldo,eq}V_{blf}\right]}}$	[39]
	$\mathrm{v}_{gapdh} = \frac{v_{max,gapdh}\left([GAP] - \frac{[PEP][NADH]}{K_{gapdh,eq}}\right)}{\left(K_{gapdh,GAP} + \left(1 + \frac{[PEP]}{K_{gapdh,pgp}}\right) + [GAP]\right)\left(\frac{K_{gapdh,NAD}}{[NAD]}\left(1 + \frac{[NADH]}{K_{gapdh,NADH}}\right) + 1\right)}$	[39]
Pyruvate Kinase	$\mathrm{v}_{pk} = \frac{v_{max,pk}[PEP]\left(\frac{[PEP]}{K_{pk,PEP}} + 1\right)^{n_{pk}-1}[ADP]}{K_{pk,PEP}\left(L_{pk}\left(\frac{1 + \frac{[ATP]}{K_{pk,ATP}}}{\frac{[F16BP]}{K_{pk,F16BP}} + \frac{[AMP]}{K_{pk,AMP}} + 1}\right)^{n_{pk}} + \left(\frac{[PEP]}{K_{pk,PEP}} + 1\right)^{n_{pk}}\right)\left([ADP] + K_{pk,ADP}\right)\}}$	[39]
Pyruvate Kinase	$\mathrm{v}_{pdh} = \frac{\frac{v_{max,\ pdh}}{[NAD]}\left(\frac{1}{1 + \frac{K_{pdh,i}[NADH]}{[NAD]}}\right)\left(\frac{[PYR]}{K_{m,PYR}}\right)\left(\frac{1}{K_{m,NAD}}\right)\left(\frac{[CoA]}{K_{m,CoA}}\right)}{\left(1 + \frac{[PYR]}{K_{m,PYR}}\right)\left(\frac{1}{[NAD]} + \frac{1}{K_{m,NAD}} + \frac{[NADH]}{K_{m,NADH}[NAD]}\right)\left(1 + \frac{[CoA]}{K_{m,CoA}} + \frac{[AcCoA]}{K_{m,AcCoA}}\right)}$	[39]
Acetyl-CoA Carboxilase	$\mathrm{v}_{acc} = \frac{v_{max}[ATP][AcCoA]}{K_{i,ATP}K_{AcCoA} + K_{ATP}[AcCoA] + K_{AcCoA}[ATP] + [ATP][AcCoA]}$	[59]
1,3,6,8-tetrahydroxynaphthalene synthase	$\mathrm{v}_{thns} = \frac{v_{max,thns}[McoA]^5}{K_{m,thns}^5[McoA]^5}$	[60]

Table 4.3. Kinetic Parameters

Reaction	Parameters Values	Reference
Cell Growth	$\mu_m = 0.9782h^{-1}$	Measured
	$X_m = 62.23$	Measured
	$K_s = 0.2$	Measured
Phosphotransferase system (pts)	$v_{max,pts} = 25.739mmol/gDCW.h$	[39]
	$K_{pts,a1} = 1mM$	[39]
	$K_{pts,a2} = 0.1mM$	Optimized
	$K_{pts,a3} = 1.0mM$	[39]
	$n_{G6P} = 4$	[39]
	$K_{pts,G6P} = 0.5mM$	[39]
Lactose Uptake (perm)	$k_{cat,in} = 1288800h^{-1}$	Optimized
	$k_{t,lac} = 0.26mM$	[57]
	$K_{i,glu} = 0.276mM$	[57]
	$K_{t,lac,in} = 14.62mM$	[57]
β-galactosidase (betagal)	$K_{m,lac} = 0.014mM$	Optimized
	$k_{cat,lac} = 572400h^{-1}$	[57]
Phosphoglucoisomerase (pgi)	$v_{max,pgi} = 26.371mmol/gDCW.h$	[39]
	$K_{eq,pgi} = 0.43mM$	[39]
	$K_{m,G6P} = 2.46mM$	[39]
	$K_{m,F6P} = 0.2mM$	[39]

Table 4.3 cont'd.

Reaction	Parameters Values	Reference
Phosphofructokinase (pfk)	$v_{max,pfk} = 24.613 mmol/gDCW.h$	[39]
	$K_{pfk,ATP} = 4.27 mM$	[39]
	$K_{pfk,ATP,ADP} = 4.6944 mM$	[39]
	$K_{s,F6P} = 0.14 mM$	[39]
	$K_{b_{ADP,AMP}} = 3.88 mM$	[8]
	$n_{pfk} = 4$	[39]
	$L_{pfk} = 10$	Optimized
	$K_{a,ADP,AMP} = 1.1118 mM$	[39]
	$K_{pfk,PEP} = 3.26 mM$	[39]
Aldolase (aldo)	$v_{max,aldo} = 2.8337 mmol/gDCW.h$	[39]
	$K_{aldo,eq} = 1.4 mM$	Optimized
	$K_{aldo,F16BP} = 0.133 mM$	[39]
	$K_{aldo,GAP} = 0.088 mM$	[39]
	$V_{blf} = 2$	[39]
	$K_{aldo,DHAP} = 0.088 mM$	[39]
	$K_{aldo,gadph} = 0.6 mM$	[39]
Glyceraldehyde 3-phosphate dehydrogenase (gapdh)	$v_{max,gapdh} = 121.29 mmol/gDCW.h$	[39]
	$K_{gapdh,eq} = 0.63$	[39]
	$K_{ggapdh,GAP} = 0.15 mM$	[39]
	$K_{gapdh,pgp} = 0.1 mM$	[39]
	$K_{gapdh,NAD} = 0.45 mM$	[39]
	$K_{gapdh,NADH} = 0.02 mM$	[39]

Table 4.3 cont'd.

Reaction	Parameters Values	Reference
Pyruvate kinase (pk)	$v_{max,pk} = 1.0849 mmol/gDCW.h$	[39]
	$K_{pk,PEP} = 0.31 mM$	[39]
	$n_{pk} = 4$	[39]
	$L_{pk} = 1$	Optimized
	$K_{pk,F16BP} = 0.19 mM$	[39]
	$K_{pk,AMP} = 0.2 mM$	[39]
	$K_{pk,ADP} = 0.26 mM$	[39]
	$K_{pk,ATP} = 22.5 mM$	[39]
Pyruvate dehydrogenase (pdh)	$v_{max,pdh} = 27171 mmol/gDCW.h$	[39]
	$K_{m,PYR} = 6.8 mM$	[56]
	$K_{m,NAD} = 0.4 mM$	[39]
	$K_{m,CoA} = 0.014 mM$	[39]
	$K_{m,NADH} = 0.1 mM$	[39]
	$K_{m,AcCoA} = 0.008 mM$	[39]
	$K_{pdh,i} = 46.4 mM$	[39]
Acetyl-CoA carboxilase (acc)	$v_{max,acc} = 1176 mmol/gDCW.h$	[61]
	$K_{i,ATP} = 0.1 mM$	[61]
	$K_{AcCoA} = 6.3 mM$	[61]
	$K_ATP = 0.01 mM$	[61]
1,3,6,8-tetrahydroxynaphthalene synthase (thns)	$v_{max,thns} = 0.30816 mmol/gDCW.h$	[60]
	$K_{m,thns} = 0.00358 mM$	[60]

Table 4.4. Kinetic type of the different enzymatic reactions

Reaction	Kinetic type description	Reference
pts	Nonspecific kinetics	[39]
perm	Reversible Michaelis-Menten with competitive inhibition	[57]
betagal	Michaelis-Menten	[57]
pgi	Reversible Michaelis-Menten	[39]
pfk	Allosteric Regulation	[39]
aldo	Ordered uni-bi mechanism	[39]
gadph	Two-substrate Reversible Michaelis-Menten	[39]
pk	Allosteric Regulation	[39]
pdh	Michaelis-Menten with non-competitive inhibition	[39]
acc	Sequential	[59]
thns	Michaelis-Menten	[60]

Once all the kinetic equations are defined, these rates are integrated into a system of ODEs where $\frac{d[\cdot]}{dt}$ is the metabolite consumption/production, $[\cdot]$ denotes the metabolite concentration, μ is the specific growth rate, and v_i is the kinetic rate described in Table 4.2. The term μ represents the dilution effect due to the increase in cellular volume followed by growth. The ODEs for a batch system are listed below:

$$\frac{d\,[X]}{dt} = \mu\,[X]$$

$$\frac{d\,[GLCEX]}{dt} = \ -v_{pts}\,[X]$$

$$\frac{d\,[LACEX]}{dt} = -v_{perm}\,[X]$$

$$\frac{d\,[LACIN]}{dt} = v_{perm} - v_{betagal} - \mu\,[LACIN]$$

$$\frac{d\,[PEP]}{dt} = v_{gapdh} - v_{pts} - v_{pk} - \mu\,[PEP]$$

$$\frac{d\,[PYR]}{dt} = v_{pk} + v_{pts} - v_{pdh} - \mu\,[PYR]$$

$$\frac{d\,[G6P\,]}{dt} = v_{pts} + v_{betagal} - v_{pgi} - \mu\,[G6P]$$

$$\frac{d\,[F6P\,]}{dt} = v_{pgi} - v_{pfk} - \mu\,[F6P]$$

$$\frac{d\,[F16BP]}{dt} = v_{pfk} - v_{aldo} - \mu\,[F16BP]$$

$$\frac{d\,[GAP\,]}{dt} = 2v_{aldo} - v_{gadph} - \mu\,[GAP]$$

$$\frac{d\,[AcCoA]}{dt} = v_{pdh} - v_{acc} - \mu\,[AcCoA]$$

$$\frac{d\,[MCoA]}{dt} = v_{acc} - v_{thns} - \mu\,[MCoA]$$

$$\frac{d\,[THN]}{dt} = v_{thns} - \mu\,[THN]$$

This system of ODEs was solved using Matlab R2015b and the function *ode45* (Appendix A.2.2) with a total time of simulation of 48 hours. The concentrations of cofactors such as CoA, ATP, ADP, AMP, NADH, NAD+ and of proteins permease and β-galactosidase were assumed to be constant according with [39] and [57], respectively (Table 4.5).

According to [62] the element composition of biomass is $C_1H_{1.666}O_{0.27}N_{0.2}$, therefore the molecular weight of E. *coli* can be assumed to be 20.786 g/mol. The average of experimental biomass formed after 48 hours of the inoculum incubation is of 0.01 g per ml of fermentation broth (481.1mM). 50 μL of the inoculum is inoculated in each culture, which means that the initial molarity of biomass in the 5 ml culture is 4.8mM.

The utilized medium (LBM) is composed of tryptophan and yeast extract. According to [45], E. *coli* can catabolize approximately 30 mM of aminoacids and sugars from Luria Bertani medium. Therefore, the initial condition of the substrate was considered as 30 mM (Table 4.1). Kinetic parameters were primarily from [39], but some were adjusted using experimental data and computational optimization (Table 4.3).

Table 4.5. Cofactors concentrations

Metabolite	Abbreviation	Concentration (mM)
Adenosine diphosphate	ADP	0.595
Adenosine monophosphate	AMP	0.955
Adenosine triphosphate	ATP	4.27
Coenzyme A	CoA	0.001
Diphosphopyridine nucleotide, oxidized	NAD+	1.47
Diphosphopyridine nucleotide, reduced	NADH	0.1
Permease	PERM	0.03
β-galactosidase	BETAGAL	0.012

4.1.2. Parameters Calibration

Finding reliable parameters for the model is a non-trivial task. While *in vitro* parameters and optimization attempt to solve this problem, differences are frequently found, and optimized parameters can differ by several orders of magnitude with respect to *in vitro* parameters [8, 13, 38, 39, 40]. Values from literature were used as an initial approach, followed by optimization of relevant parameters.

The growth parameters were calculated using experimental measurements of the initial and final amounts of biomass and substrate. X is the cell concentration and X_m is the

final value of X in the batch culture. X_m is set as the product between the initial substrate concentration and $Y_{s/x}$, where $Y_{s/x}$ is the yield coefficient and represents the ratio between the biomass produced and the substrate consumed.

In order to calculate the maximum specific growth rate (μ_m), we constructed a growth curve with measurements of OD_{600} along 24 hours (Fig 4.2). To find μ_m, the exponential phase must be identified in the growth curve since the slope describes the maximum specific growth rate [52]. The Dry Cell Weight (DCW) was calculated as 0.33 times the measured OD_{600} value of the slope points [29]. Taking the logarithm value of DCW, μ_m is calculated as:

$$\mu_m = \frac{log(DCW_{t_2}) - log(DCW_{t_1})}{t_2 - t_1} = 0.9782$$

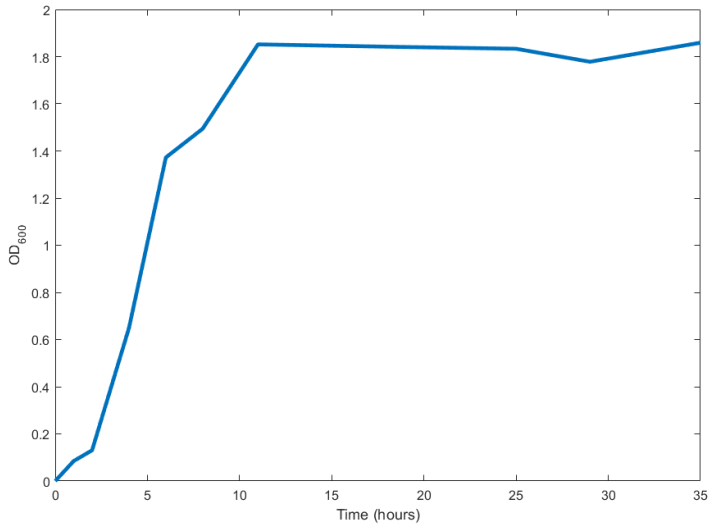


Figure 4.2. Growth Curve utilized to calculate the specific growth rate (μ_m) from the exponential phase

The parameters $K_{pts,a2}$, $k_{cat,in}$, $K_{m,lac}$, L_{pfk}, $K_{aldo,eq}$ and L_{pk} from Table 4.3 were calibrated by defining the kinetic model and then finding the parameter values that reproduce the experimental data with the best fit. First, the values from the literature were used as an initial approach [39], and then they were adjusted until finding values which minimize the difference between simulated and experimental concentrations of the metabolites.

4.1.3. Results

Figures 4.3 and 4.4 show the solution to the system of ODEs. The model was able to correctly describe the substrate metabolism (glucose and lactose), THN production, and biomass. Simulated bacterial growth curve (biomass) shows a curve in agreement with experimental data (Fig 4.2). Unfortunately, the glucose consumption cannot be measured experimentally in a low cost manner. Blood glucose meters detect maximum 550 mg/dL of glucose, which is approximately 29 mM of glucose. Since the system has more glucose than the instrument can detect, it was impossible to gather this data using a low cost measurement. Finally, the simulation of THN product formation (2.6 mM) is slightly above the experimental data (2.5 mM). This can be attributed to the exclusion of by-products and other metabolic pathways in this simplified kinetic model which would consume malonyl-CoA and acetyl-CoA, as the TCA cycle for example.

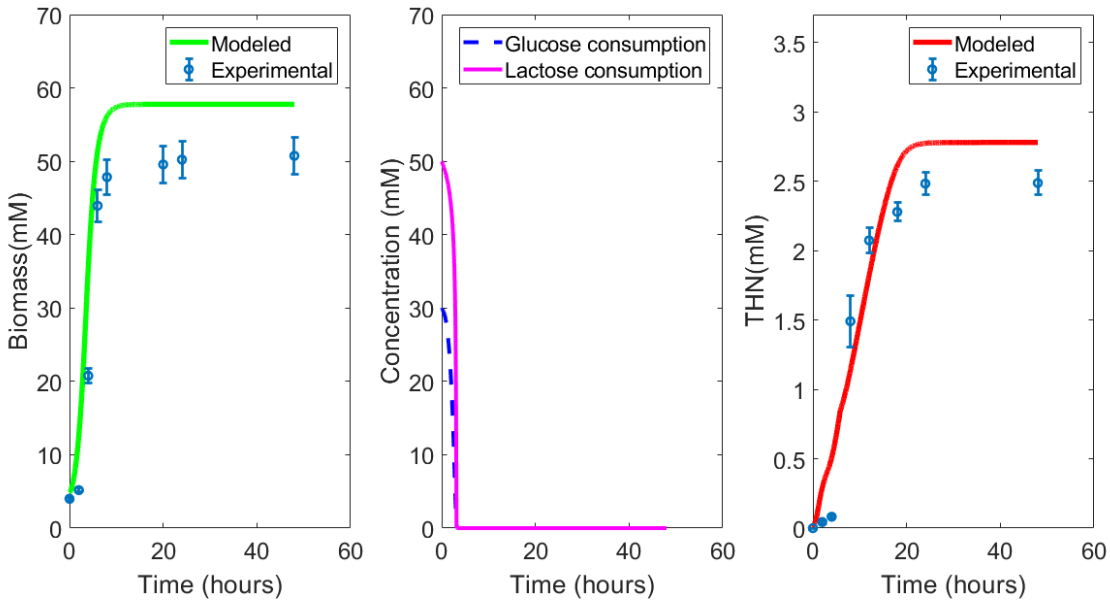


Figure 4.3. Simulated growth curve (green), glucose (blue dashed) and lactose consumption (magenta), and THN production (red). Experimental data are represented by the $'o'$.

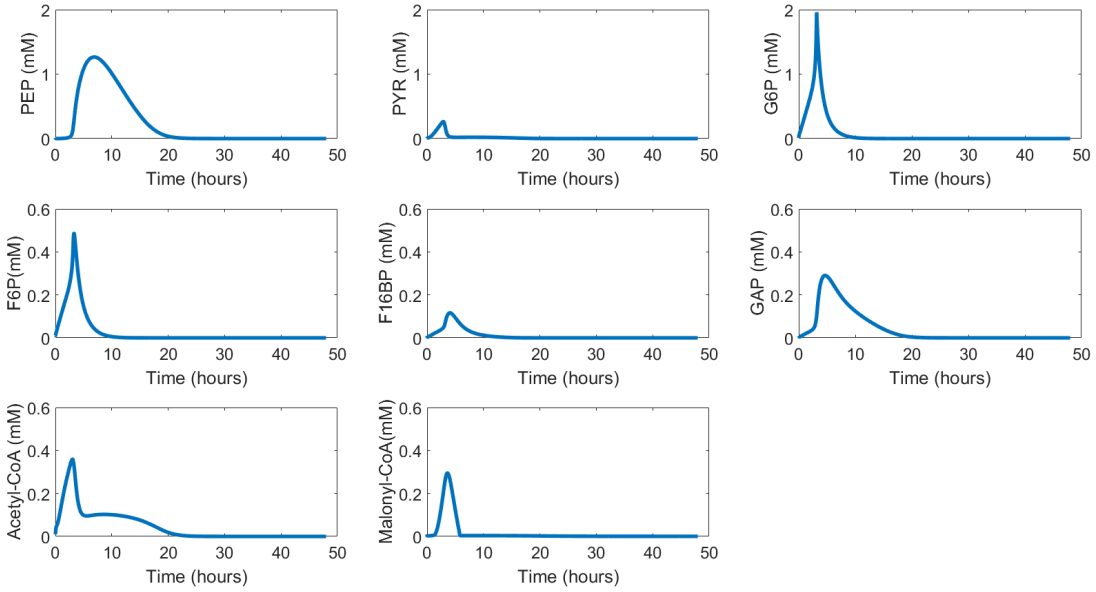


Figure 4.4. Prediction of changes in the concentration of intracellular metabolites

4.2. Stability Analysis

A stability analysis was performed because it is assumed that under physiological conditions and with the environment maintained constant, the metabolite concentration may be stable. However, the pathway may exhibits oscillations at some specific points.

If an experiment on a specific pathway is repeated many times, it is very likely that the initial concentrations of the various metabolites will be different at each time. However, it is expected that the pathway settles to the same fixed point (equilibrium point). This means that if the pathway is at a stable equilibrium point and a small change is made to some of the concentrations, then after some transient behaviors the pathway settles back to the same point. On the other hand, if the equilibrium point is unstable, a small perturbation would cause the system to change drastically and move away from its original position.

To analyze a system, the plant model needs to be provided. Plants are usually dynamic systems. In this case, the E. *coli* metabolism is the plant, and the set of ODEs of kinetic modeling is the computed model. First, the equilibrium points of the original nonlinear system need to be found. Using the function *fsolve* of Matlab, the roots of the nonlinear

systems were found by providing an initial point x_0.

The roots of the system are by definition the equilibrium points, and because the system is closed, meaning that none of the variables can assume negative values, none of the equilibrium points can be negative. Therefore, once all the equilibrium points found by the $fsolve$ function contained one or more negative values, it was decided that the equilibrium points of the system would be those at steady state. The stability of the kinetic model is checked by substituting the equilibrium point of each metabolite as the initial condition of the system. Once these values remain the same during all simulation time, it is possible to conclude that these are the model's equilibrium points. Therefore, by analyzing the graphs of the kinetic model, the equilibrium points were chosen as the value that each variable assumes at the end of the simulation (Table 4.6).

Table 4.6. Equilibrium points of the kinetic model

Variable	Metabolite	Equilibrium point value
x1	X	57.764
x2	GLCEX	0
x3	PEP	2.24e-10
x4	PYR	2.803e-12
x5	G6P	8.708e-19
x6	F6P	3.379e-19
x7	F16BP	6.713e-20
x8	GAP	2.469e-11
x9	AcCoA	1.515e-10
x10	MCoA	0.0004025
x10	THN	2.7698
x12	LACEX	0
x13	LACIN	0

The set of ODEs built in the previous section have a nonlinear behavior, therefore the system needs to be linearized by finding the Jacobian matrix A. Utilizing the function *jacobian*, the first-order partial derivatives of the kinetic functions with respect of the variables were found, and the variables ($x1$-$x13$) were substituted by its respective equilibrium points found previously (Appendix A.3.1):

$$A = \frac{\partial f}{\partial x}_{x=x_e}$$

Hence, the nonlinear system was linearized,

$$\frac{d[x]}{dt} = Ax$$

where A is the 13×13 Jacobian matrix evaluated at the equilibrium points, and x is the state vector.

In order to analyze the stability of the system, the eigenvalues of A need to be computed. Using the Matlab *eig* function, we calculated the eigenvalues of A:

$$\lambda_{13} = \begin{bmatrix} -11808.77 \\ -0.06835774 \\ -3.056301e^{-19} \\ -490631.3 \\ -8589.963 \\ -29.87355 \\ -1.823951 \\ -43.18644 \\ -38.34954 \\ -21.30602 \\ -1.047352 \\ -0.6496196 \\ -3.647823e^{-18} \end{bmatrix}$$

Note that the eigenvalues of the matrix A are negative real numbers, although two of them are approximately zero (-3.056301e^{-19} and -3.647823e^{-18}). This leads to the conclusion that the system is stable at the equilibrium point. Hence, to check this condition,

the initial conditions of the kinetic model were substituted by the equilibrium points (Fig 4.5 and 4.6). Analyzing these results, it is possible to see that all substrates, intermediate metabolites, and products remain the same during all simulation time. Therefore, the kinetic behavior of the system when analyzed at its equilibrium point shows that the system is stable.

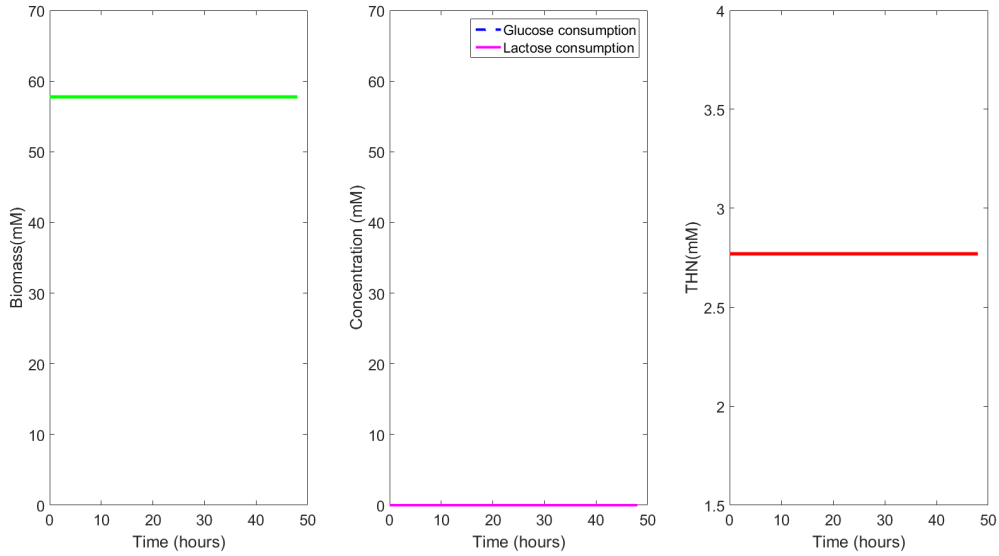


Figure 4.5. Biomass, Glucose, Lactose and THN response at the equilibrium point

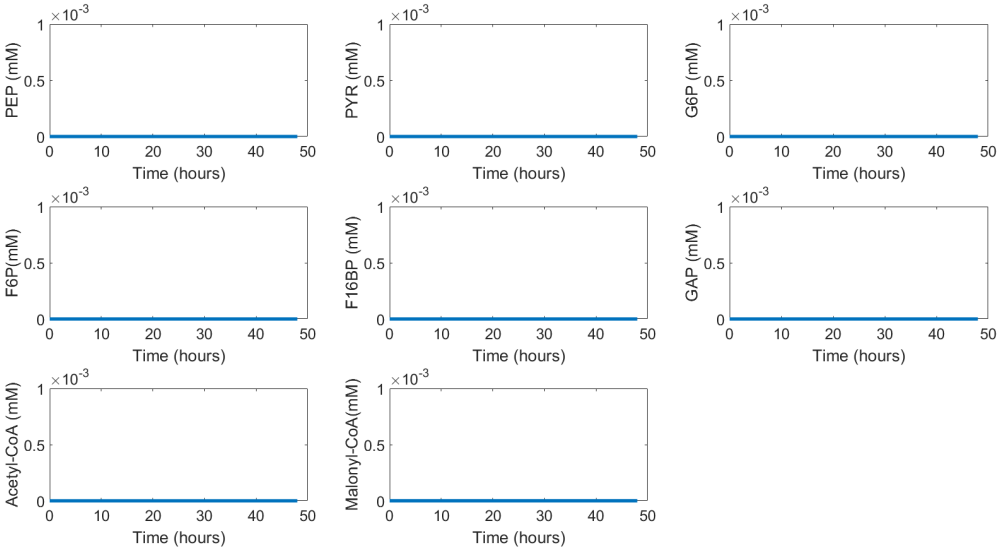


Figure 4.6. Metabolites response at the equilibrium point

4.3. Metabolic Control Analysis

The mass balance equations for all metabolites from Section 4.1 can be arranged in a 1×13 vector $\frac{dx}{dt}$. Using the concept of stoichiometric matrix introduced in the Section 1.2.3, we can rewrite the ODEs:

$$\frac{dx}{dt} = N \cdot v$$

where N is the stoichiometric matrix and v is the vector with all kinetic rates (Appendix A.4).

The Jacobian matrix (A) introduced in Section 4.2, can be rewritten:

$$A = \frac{\partial f}{\partial x}_{x=x_e} = \frac{\partial(\frac{dx}{dt})}{\partial x}_{x=x_e} = \frac{\partial(Nv)}{\partial x}_{x=x_e} = N\frac{\partial v}{\partial x}_{x=x_e} \tag{4.1}$$

Using the relationships introduced by [20] and [63] and Eq. 4.1, the control matrices (Eq. 1.12 and 1.13) can be rewritten as follows:

$$C_S = \frac{e_j}{C_{ss}}\frac{\partial C_{ss}}{\partial e_j} = -\left(N\frac{\partial v}{\partial x}\right)^{-1} N = -A^{-1}N$$

$$C_J = \frac{e_j}{J_k}\frac{\partial J_k}{\partial e_j} = I - \frac{\partial v}{\partial x}\left(N\frac{\partial v}{\partial x}\right)^{-1} N = I - \frac{\partial v}{\partial x}_{x=x_e} A^{-1}N$$

The 13×12 matrix C_S and 12×12 matrix C_J are unscaled, meaning that the coefficients do not reflect rates of change in proportion to the concentration of x and reaction rates v at the equilibrium point.

Denoting the diagonal matrix ($diag$) whose diagonal entries are given by x, where x is any vector:

$$diag(x)_{ij} = \begin{cases} 0 & if \quad i \neq j \\ x_i & if \quad i = j, \end{cases}$$

we can scale the control matrices [20]:

$$\bar{C_S} = (diag(x_e))^{-1} C_S diag(v(x_e)) \tag{4.2}$$

$$\bar{C_J} = (diag(v(x_e)))^{-1} C_J diag(v(x_e)) \tag{4.3}$$

Applying the summation theorem to Eq. 4.2 and 4.3:

$$\sum_{i=1}^{k} (\bar{C_J}\mathbf{1})_i = 1 \tag{4.4}$$

$$\sum_{i=1}^{m} (\bar{C_S}\mathbf{1})_i = 0 \tag{4.5}$$

where k is the number of rates and m is the number of variables. Therefore, implementing the equations 4.2, 4.3, 4.4 and 4.5 in Matlab, and using the kinetic rates, parameters and ODEs described previously, we can find the control matrices C_J and C_S:

$$\bar{C_J} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \bar{C_S} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Since the summation of all values of $\bar{C_J}$ should be one, this result showed at that all the flux control in the glucose fermentation pathway is done by one single enzyme. Specifically, the enzyme phosphoglucose isomerase (pgi) is responsible for the control of the pathway flux since it is the only coefficient that does not have a null value. All other flux control coefficients have null values, meaning that no control is performed by the corresponding enzymes.

Phosphoglucose isomerase (pgi) is the enzyme responsible for the catalysis of $G6P$ into $F6P$. Analyzing the accumulation and consumption of these two metabolites over time (Fig. 4.7), it is notable that the enzymatic type of reaction is the responsible for the slow and low formation of $F6P$. The conversion of $G6P$ into $F6P$ follows reversible Michaelis-Menten kinetics, meaning that three reversible stages take place: (i) pgi binds to $G6P$ forming the enzyme-substrate complex, (ii) this complex is converted into enzyme-product complex, and (iii) desorption of the product $F6P$. Analyzing the concentration of $G6P$ and $F6P$ over time, it is clear that the former is accumulated in a faster rate than the latter, meaning that the reaction is favored towards $G6P$ formation.

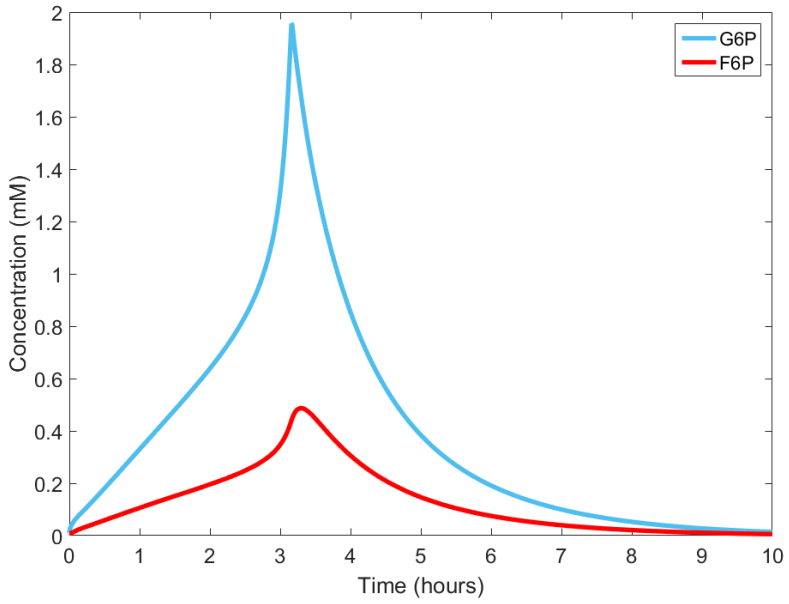


Figure 4.7. $G6P$ and $F6P$ concentration varying with the time

We can test this hypothesis by altering the kinetic parameters of *pgi* in the model. After changing all the parameters by one order of magnitude up or down, it was possible to see that the main ones associated with the production of THN were $v_{max,pgi}$ and $K_{eq,pgi}$. Increasing $v_{max,pgi}$ and $K_{eq,pgi}$ by 10 times (263.71 mmol/gDCW.h and 4.3 mM, respectively), we can see the production the biomass formation decreasing by half, the lactose and glucose consumption did not change, and the production of THN doubled (black curves of Fig 4.8).

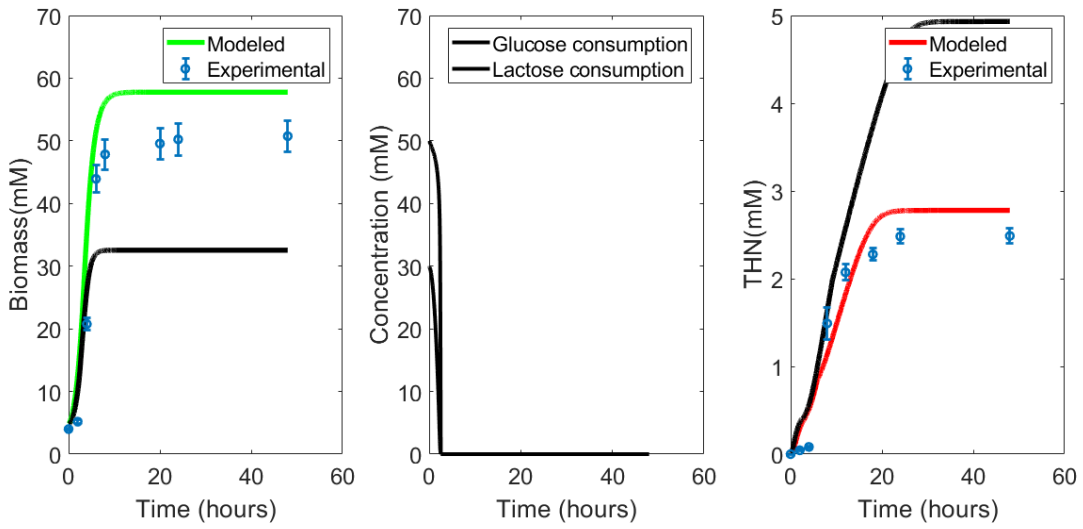


Figure 4.8. Effect of optimizing phosphoglucose isomerase parameters in the kinetic model

4.4. Discussion

Besides being a simplified version of the bacterial metabolism, the proposed model was able to correctly represent the utilization of glucose and lactose and predict the synthesis of THN and biomass. Moreover, the stability analysis showed that at the calculated equilibrium point the system has a stable behavior, meaning that small perturbations in the fermentation conditions will not result in major changes in the bacterial metabolism.

MCA showed that the enzyme phosphoglucose isomerase controls by itself the flux towards malonyl-CoA production. The conversion of $G6P$ in $F6P$ was seen as favored in the reverse way, which is probably achieved by the stronger binding with the product than with the reactant, and because β-galactosidase may be faster than *pgi*, accumulating more

$G6P$ than $F6P$. Hence, the enzyme kinetics of *pgi* can be ameliorated to reduce the time of glucose fermentation toward the production of malonyl-CoA in E. *coli*, while changes in other enzymes involved in the pathway will have no effect in this metabolic process.

In order to represent the whole bacterial metabolism, and therefore generate more reliable results, other metabolic pathways that consume intermediate metabolites, such as the TCA cycle and the fatty acids synthesis, should be included in the studied kinetic model. As a result, it is possible to virtually knock-in or knock-out genes that may play a role in the malonyl-CoA production, and predict which genetic modification should be done in order to maximize the final product. Therefore, the implementation of a complete kinetic model can be even more useful and reliable in saving time and resources in experimental work.

Chapter 5
Conclusions and Future Directions

In this study, we described an integrated experimental and computational approach for maximizing production of malonyl-CoA in E. *coli*, a key precursor for numerous bio-products, such as biodiesel, flavonoids, bioplastics, and polyketides. Therefore, a low-cost method that improves cellular malonyl-CoA level highlights the potential of this work for biosynthesis of these commercially-important compounds.

The proposed experimental method produced an average of 68 mM of flaviolin/gWW and 340 mM of malonyl-CoA/gWW with a cost of $3.09 per liter of fermentation broth. These results were greater than the ones reported by previous studies (Table 1.1). Such achievement was done without any genetic modification, but only with the study of how each fermentation parameter affected the flaviolin production, and thereby maximized malonyl-CoA. Moreover, all the fermentation conditions as well as the quantification method did not require complex and expensive materials or equipment. Hence, the lab scale method can be easily scaled up with a competitive cost when compared with current industrial process.

The kinetic model constructed allows the further prediction of minimal set of genetic modifications that guarantee an optimum genotype, which are not obvious to predict by experimental inspection.

The first next step of this study would be to scale up the proposed method using a bioreactor and a defined medium. This would provide a better insight into production rates, yield, and other process variables.

The second next step would be to incorporate in the kinetic model other pathways, such as the TCA cycle, the aminoacids and the fatty acids synthesis. The inclusion of other by-products, such as acetic acid, would also improve the model accuracy.

References

[1] D. Petrides and R. G. Harrison, "Bioprocess Design and Economics." Oxford University Press, 2003. [Online]. Available: http://www.intelligen.comhttp//www.oup-usa.org/isbn/0195123409.html

[2] Y. P. Lim, M. K. Go, and W. S. Yew, "Exploiting the biosynthetic potential of type III polyketide synthases," *Molecules*, vol. 21, no. 6, pp. 1–38, 2016.

[3] R. A. Voloshin, M. V. Rodionova, S. K. Zharmukhamedov, T. N. Veziroglu, and S. I. Allakhverdiev, "Review : Biofuel production from plant and algal biomass," *International Journal of Hydrogen Energy*, vol. 41, no. 39, pp. 17 257–17 273, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijhydene.2016.07.084

[4] J. M. Clomburg and R. Gonzalez, "Biofuel production in Escherichia coli: the role of metabolic engineering and synthetic biology," *Applied Microbiology and Biotechnology*, vol. 86, no. 2, pp. 419–434, 2010. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/20143230http://link.springer.com/10.1007/s00253-010-2446-1

[5] I. Borodina, K. R. Kildegaard, N. B. Jensen, T. H. Blicher, J. Maury, S. Sherstyk, K. Schneider, P. Lamosa, M. J. Herrgard, I. Rosenstand, F. Oberg, J. Forster, and J. Nielsen, "Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine," *Metabolic Engineering*, vol. 27, pp. 57–64, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.ymben.2014.10.003

[6] X. Lu, H. Vora, and C. Khosla, "Overproduction of free fatty acids in E. coli: Implications for biodiesel production," *Metabolic Engineering*, vol. 10, no. 6, pp. 333–339, 2008.

[7] C. S. Henry, L. J. Broadbelt, and V. Hatzimanikatis, "Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate," *Biotechnology and Bioengineering*, vol. 106, no. 3, pp. 462–473, 2010.

[8] C. Chassagnole, N. Noisommit-Rizzi, J. W. Schmid, K. Mauch, and M. Reuss, "Dynamic modeling of the central carbon metabolism of Escherichia coli," *Biotechnology and Bioengineering*, vol. 79, no. 1, pp. 53–73, 2002.

[9] J. D. Orth, T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist, and B. O. Palsson, "A comprehensive genome-scale reconstruction of Escherichia coli metabolism–2011," *Molecular Systems Biology*, vol. 7, no. 1, pp. 535–535, apr 2014.

[10] Z. L. Fowler, W. W. Gikandi, and M. A. G. Koffas, "Increased malonyl coenzyme A biosynthesis by tuning the Escherichia coli metabolic network and its application to flavanone production," *Applied and Environmental Microbiology*, vol. 75, no. 18, pp. 5831–5839, 2009.

[11] J. Cortés, J. Velasco, G. Foster, A. P. Blackaby, B. A. Rudd, and B. Wilkinson, "Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in Saccharopolyspora erythraea," *Molecular Microbiology*, vol. 44, no. 5, pp. 1213–1224, 2002.

[12] M. K. Choudhary, "Metabolic engineering of Escherichia coli for the efficient utilization of plant sugar mixture," Ph.D. dissertation, Iowa State University, 2008. [Online]. Available: https://lib.dr.iastate.edu/rtd/15312

[13] K. Peskov, E. Mogilevskaya, and O. Demin, "Kinetic modelling of central carbon metabolism in Escherichia coli," *FEBS Journal*, vol. 279, no. 18, pp. 3374–3385, 2012.

[14] T. Romeo and J. L. Snoep, "Glycolysis and Flux Control," *EcoSal Plus*, vol. 1, no. 2, jul 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26443505http://www.asmscience.org/content/journal/ecosalplus/10.1128/ecosalplus.3.5.1

[15] W. Zha, S. B. Rubin-Pitel, Z. Shao, and H. Zhao, "Improving cellular malonyl-CoA level in Escherichia coli via metabolic engineering," *Metabolic Engineering*, vol. 11, no. 3, pp. 192–198, 2009.

[16] J. E. Cronan and G. L. Waldrop, "Multi-subunit acetyl-CoA carboxylases." *Progress in lipid research*, vol. 41, no. 5, pp. 407–35, sep 2002. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12121720

[17] B. Zhao, F. P. Guengerich, A. Bellamine, D. C. Lamb, M. Izumikawa, L. Lei, L. M. Podust, M. Sundaramoorthy, J. A. Kalaitzis, L. M. Reddy, S. L. Kelly, B. S. Moore, D. Stec, M. Voehler, J. R. Falck, T. Shimada, and M. R. Waterman, "Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of Streptomyces coelicolor A3(2) cytochrome P450 158A2," *Journal of Biological Chemistry*, vol. 280, no. 12, pp. 11 599–11 607, 2005.

[18] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 4th ed. Wiley, 2012.

[19] B. P. Ingalls, "Mathematical Modelling in Systems Biology : An Introduction," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, pp. 1–396, 2014.

[20] J. Gunawardena, "Notes on Metabolic Control Analysis," Tech. Rep., 2002. [Online]. Available: http://bip.cnrs-mrs.fr/bip10/rls.htm

[21] P. A. Iglesias and B. P. Ingalls, *Control theory and systems biology.* MIT Press, 2010.

[22] H. J. Marquez, *Nonlinear control systems: Analysis and Design.* John Wiley, 2003.

[23] H. K. Khalil, *Nonlinear systems.* Prentice Hall, 2002.

[24] A. Dixon, "Designing Predictive Mathematical Models for the Metabolic Pathways Associated with Polyhydroxybutyrate Synthesis in Escherichia coli," Ph.D. dissertation, Utah State University, 2011.

[25] A. K. Gombert and J. Nielsen, "Mathematical modelling of metabolism." *Current opinion in biotechnology*, vol. 11, no. 2, pp. 180–6, apr 2000. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10753761

[26] D. A. Fell, "Metabolic Control Analysis: a survey of its theoretical and experimental development," *Biochem. J*, vol. 286, pp. 313–330, 1992. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1132899/pdf/biochemj00128-0011.pdf

[27] M. D. Lynch, R. T. Gill, and T. E. W. Lipscomb, "Methods for producing 3-hydroxypropionic acid and other products," 2014.

[28] Y. M. Zhang, S. W. White, and C. O. Rock, "Inhibiting bacterial fatty acid synthesis," *Journal of Biological Chemistry*, vol. 281, no. 26, pp. 17 541–17 544, 2006.

[29] M. D. Lynch, T. W. Lipscomb, A. D. Trahan, A. Singh, and T. Wolter, "Microbial Production of Chemical Products and Related Compositions, Methods and Systems," 2014.

[30] D. Yang, W. J. Kim, S. M. Yoo, J. H. Choi, S. H. Ha, M. H. Lee, and S. Y. Lee, "Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria," *Proceedings of the National Academy of Sciences*, vol. 115, no. 40, pp. 9835–9844, oct 2018. [Online]. Available: http://www.pnas.org/lookup/doi/10.1073/pnas.1808567115

[31] P. Xu, S. Ranganathan, Z. L. Fowler, C. D. Maranas, and M. A. Koffas, "Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA," *Metabolic Engineering*, vol. 13, no. 5, pp. 578–587, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.ymben.2011.06.008

[32] S. Mohan Raj, C. Rathnasingh, W.-C. Jung, and S. Park, "Effect of process parameters on 3-hydroxypropionic acid production from glycerol using a recombinant Escherichia coli," *Applied Microbiology and Biotechnology*, vol. 84, no. 4, pp. 649–657, sep 2009. [Online]. Available: http://link.springer.com/10.1007/s00253-009-1986-8

[33] F. Matsuda, B. F. Cress, H. Shimizu, K. Tokuyama, M. A. Koffas, and Y. Toya, "Magnesium starvation improves production of malonyl-CoA-derived metabolites in Escherichia coli," *Metabolic Engineering*, vol. 52, no. August 2018, pp. 215–223, 2018.

[34] R. S. Costa, A. Hartmann, and S. Vinga, "Kinetic modeling of cell metabolism for microbial production," *Journal of Biotechnology*, vol. 219, pp. 126–141, feb 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168165615302248

[35] A. O. Johnson, M. Gonzalez-Villanueva, L. Wong, A. Steinbüchel, K. L. Tee, P. Xu, and T. S. Wong, "Design and application of genetically-encoded malonyl-CoA biosensors for metabolic engineering of microbial cell factories," *Metabolic Engineering*, vol. 44, no. July, pp. 253–264, 2017. [Online]. Available: https://doi.org/10.1016/j.ymben.2017.10.011

[36] C. Rathnasingh, S. M. Raj, Y. Lee, C. Catherine, S. Ashok, and S. Park, "Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant Escherichia coli strains," *Journal of Biotechnology*, vol. 157, no. 4, pp. 633–640, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.jbiotec.2011.06.008

[37] S. Krauser, P. Kiefer, and E. Heinzle, "Multienzyme Whole-Cell InSitu Biocatalysis for the Production of Flaviolin in Permeabilized Cells of Escherichia coli," *ChemCatChem*, vol. 4, no. 6, pp. 786–788, 2012.

[38] N. Jahan, K. Maeda, Y. Matsuoka, Y. Sugimoto, and H. Kurata, "Development of an accurate kinetic model for the central carbon metabolism of Escherichia coli," *Microbial Cell Factories*, vol. 15, no. 1, pp. 1–20, 2016.

[39] T. A. Kadir, A. A. Mannan, A. M. Kierzek, J. McFadden, and K. Shimizu, "Modeling and simulation of the main metabolism in Escherichia coli and its several single-gene knockout mutants with experimental verification," *Microbial Cell Factories*, vol. 9, pp. 1–21, 2010.

[40] W. Won, C. Park, S. Y. Lee, K. S. Lee, and J. Lee, "Parameter estimation and dynamic control analysis of central carbon metabolism in Escherichia coli," *Biotechnology and Bioprocess Engineering*, vol. 16, no. 2, pp. 216–228, 2011.

[41] Y. Toya, N. Ishii, K. Nakahigashi, T. Hirasawa, T. Soga, M. Tomita, and K. Shimizu, "13C-Metabolic flux analysis for batch culture of Escherichia coli and its pyk and pgi gene knockout mutants based on mass isotopomer distribution of intracellular metabolites," *Biotechnology Progress*, vol. 26, no. 4, pp. 975–992, 2010.

[42] S. Atsumi, A. F. Cann, M. R. Connor, C. R. Shen, K. M. Smith, M. P. Brynildsen, K. J. Chou, T. Hanai, and J. C. Liao, "Metabolic engineering of Escherichia coli for 1-butanol production," *Metabolic Engineering*, vol. 10, no. 6, pp. 305–311, 2008.

[43] I. Miyahisa, M. Kaneko, N. Funa, H. Kawasaki, H. Kojima, Y. Ohnishi, and S. Horinouchi, "Efficient production of (2S)-flavanones by Escherichia coli containing an artificial biosynthetic gene cluster," *Applied Microbiology and Biotechnology*, vol. 68, no. 4, pp. 498–504, 2005.

[44] J. Zeng, R. Decker, and J. Zhan, "Biochemical characterization of a type III polyketide biosynthetic gene cluster from Streptomyces toxytricini," *Applied Biochemistry and Biotechnology*, vol. 166, no. 4, pp. 1020–1033, 2012.

[45] G. Sezonov, D. Joseleau-Petit, and R. D'Ari, "Escherichia coli physiology in Luria-Bertani broth," *Journal of Bacteriology*, vol. 189, no. 23, pp. 8746–8749, 2007.

[46] F. W. Studier, "Protein production by auto-induction in high density shaking cultures." *Protein expression and purification*, vol. 41, no. 1, pp. 207–234, 2005.

[47] L. Sun, J. Zeng, S. Zhang, T. Gladwin, and J. Zhan, "Effects of exogenous nutrients on polyketide biosynthesis in Escherichia coli," *Applied Microbiology and Biotechnology*, vol. 99, no. 2, pp. 693–701, 2014.

[48] C. Bracken, J. Marley, and M. Lu, "A method for efficient isotopic labeling of recombinant proteins," *J Biomol NMR*, vol. 20, pp. 71–75, 2001. [Online]. Available: http://www.ncbi.nih.gov/entrez/query.fcgi?cmd=retrieve{&}db=pubmed{&}dopt=abstract{&}list{_}uids=11430757

[49] P. Dvorak, L. Chrast, P. I. Nikel, R. Fedr, K. Soucek, M. Sedlackova, R. Chaloupkova, V. de Lorenzo, Z. Prokop, and J. Damborsky, "Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21(DE3) carrying a synthetic metabolic pathway," *Microbial Cell Factories*, vol. 14, no. 1, p. 201, 2015. [Online]. Available: http://www.microbialcellfactories.com/content/14/1/201

[50] M. Gadgil, V. Kapur, and W.-S. Hu, "Transcriptional response of Escherichia coli to temperature shift." *Biotechnology progress*, vol. 21, no. 3, pp. 689–699, 2005.

[51] Y. Zhang, L. Taiming, and J. Liu, "Low temperature and glucose enhanced T7 RNA polymerase-based plasmid stability for increasing expression of glucagon-like peptide-2 in Escherichia coli." *Protein expression and purification*, vol. 29, no. 1, pp. 132–9, may 2003. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12729734

[52] R. M. Maier, *Bacterial Growth*. Elsevier, jan 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B978012394626300003X

[53] L. McDaniel and E. Bailey, "Effect of Shaking Speed and Type of Closure," *Applied Microbiology*, vol. 17, no. 2, pp. 286–290, 1969.

[54] K. A. Presser, D. A. Ratkowsky, and T. Ross, "Modelling the growth rate of Escherichia coli as a function of pH and lactic acid concentration," *Applied and Environmental Microbiology*, vol. 63, no. 6, pp. 2355–2360, 1997.

[55] G. W. Luli and W. R. Strohl, "Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations," *Applied and Environmental Microbiology*, vol. 56, no. 4, pp. 1004–1011, 1990.

[56] A. P. Lima, V. Baixinho, D. Machado, and I. Rocha, "A Comparative Analysis of Dynamic Models of the Central Carbon Metabolism of Escherichia coli," *IFAC-PapersOnLine*, vol. 49, no. 26, pp. 270–276, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ifacol.2016.12.137

[57] P. Wong, S. Gladney, and J. D. Keasling, "Mathematical model of the lac operon: Inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose," *Biotechnology Progress*, vol. 13, no. 2, pp. 132–143, 1997.

[58] M. Santillán, "Bistable behavior in a model of the lac operon in Escherichia coli with variable growth rate," *Biophysical Journal*, vol. 94, no. 6, pp. 2065–2081, 2008.

[59] T. C. Broussard, A. E. Price, S. M. Laborde, and G. L. Waldrop, "Complex formation and regulation of escherichia coli acetyl-CoA carboxylase," *Biochemistry*, vol. 52, no. 19, pp. 3346–3357, 2013.

[60] M. Izumikawa, P. R. Shipley, J. N. Hopke, T. O'Hare, L. Xiang, J. P. Noel, and B. S. Moore, "Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from Streptomyces coelicolor A3(2)," *Journal of Industrial Microbiology and Biotechnology*, vol. 30, no. 8, pp. 510–515, 2003.

[61] C. Z. Blanchard and G. L. Waldrop, "Overexpression and kinetic characterization of the carboxyltransferase component of acetyl-CoA carboxylase." *The Journal of biological chemistry*, vol. 273, no. 30, pp. 19 140–5, jul 1998. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/9668099

[62] P. Mongiovi, "Empirical elemental formula for biomass - Bacteria Escherichia coli - BNID 101800." [Online]. Available: https://bionumbers.hms.harvard.edu/bionumber.aspx?id=101800{&}ver=14

[63] J.-H. S. Hofmeyr, "Metabolic control analysis in a nutshell," Tech. Rep. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.324.922{&}rep=rep1{&}type=pdf

Appendix A
Model and Experiment Details

A.1. Cost of Consumables

Table A.1. Cost of consumables utilized to maximize malonyl-CoA production

Material	Cost ($)	Package Size (g)	Price ($/gram)	Reference
Streptomycin	78.5	100	0.785	https://bit.ly/2EQpN86
Yeast Extract	88.1	500	0.1762	https://bit.ly/2UuEMv4
Triptone	77.8	250	0.3112	https://bit.ly/2EQ7Y9i
Bacteriological Agar	145	500	0.29	https://bit.ly/2NRyXFG
Sodium Chloride	63	1000	0.063	https://bit.ly/2TqXV0G
Dextrose	31	1000	0.031	https://bit.ly/2ESFxb4
Calcium Chloride Dihydrate	88	500	0.176	https://bit.ly/2XPwTmm
Magnesium Sulfate Anydrous	147	1000	0.147	https://bit.ly/2TsNLAv
Manganese Chloride	121	500	0.242	https://bit.ly/2SXpm0Q
D-Lactose Monohydrate	52.6	1000	0.0526	https://bit.ly/2UoFr1m

A.2. Kinetic Model

A.2.1. Enzymatic Reactions

List of reaction stoichiometries of the kinetic model [39]:

- Phosphotransferase system:

$$GLCEXT + PEP \rightarrow G6P + PYR$$

- Lactose Uptake

$$LACEX \rightarrow LACIN$$

$$LACIN \rightarrow G6P$$

- Glycolysis pathway:

$$G6P \rightarrow F6P$$

$$F16BP \rightarrow 2GAP/DHAP$$

$$GAP + NAD + ADP \rightarrow PEP + NADH + ATP$$

$$PEP + ADP \rightarrow PYR + ATP$$

- THN synthesis:

$$PYR + CoA + NAD^{+} \rightarrow AcCoA + CO_2 + NADH$$

$$AcCoA + ATP + HCO_3^{-} \rightarrow MCoA + ADP$$

$$5MCoA \rightarrow THN + 5CO_2$$

A.2.2. Matlab Code

```
function model_qss_lactose

tspan=0:0.01:48;

%Define Initial conditions (mM)

x0=[4.8 30 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0 50 0];

% Variables
% X          x(1)
% GLCEX      x(2)
% PEP        x(3)
% PYR        x(4)
% G6P        x(5)
% F6P        x(6)
% F16BP      x(7)
% GAP        x(8)
% AcCoA      x(9)
% MCoA       x(10)
% THN        x(11)
%LACEX   x(12)
%LACIN   x(13)

```

```
24
25 %Parameters [mM and mmol/g.h]
26
27 k(1)=0.9782;          %u_m
28 k(2)=0.2;             %K_s
29 k(3)=62.23;           %X_m
30
31 k(7)=25.739;         %v_(max,pts)
32 k(8)=1;              %K_(pts,a1)
33 k(9)=0.1;            %K_(pts,a2)
34 k(10)=1;             %K_(pts,a3)
35 k(11)=4;             %n_G6P
36 k(12)=0.5;           %K_(pts,G6P)
37
38 k(13)=26.3711;      %v_(max,pgi)
39 k(14)=0.43;          %K_(eq,pgi)
40 k(15)=2.46;          %K_(m,G6P)
41 k(16)=0.2;           %K_(m,F6P)
42
43 k(17)=24.613;       %v_(max,pfk)
44 k(18)=4.27;          %K_(pk,ATP)
45 k(19)=4.6944;        %K_(pk,ATP,ADP)
46 k(20)=0.14;          %K_(s,F6P)
47 k(21)=3.88;          %K_b(ADP,AMP)
48 k(22)=10;            %L_pfk
49 k(23)=1.1118;        %K_(a,(ADP,AMP) )
50 k(24)=4;             %n_pfk
51 k(25)=3.26;          %K_PEP
52
53 k(26)=2.8337;       %v_(max,aldo)
54 k(27)=1.4;           %K_(aldo,eq)
55 k(28)=0.133;         %K_(aldo,F16BP)
56 k(29)=0.088;         %K_(aldo,GAP)
```

```
57 k(30)=2;              %V_blf
58 k(31)=0.088;          % K_(aldo,DHAP)
59 k(32)=0.6;            %K_(aldo,gap inh)
60
61 k(33)=121.29;         %v_(max,gapdh)
62 k(34)=0.63;           %K_(gapdh,eq)
63 k(35)=0.15;           %K_(ggapdh,GAP)
64 k(36)=0.1;            %K_(gapdh,pgp )
65 k(37)=0.45;           %K_(gapdh,NAD)
66 k(38)=0.02;           %K_(gapdh,NADH)
67
68 k(39)=1.0849;         %v_(max,pk)
69 k(40)=0.31;           %K_(pk,PEP)
70 k(41)=4;              %n_pk
71 k(42)=1;              %L_pk
72 k(43)=0.19;           %K_(pk,F16BP)
73 k(44)=0.2;            %K_(pk,AMP)
74 k(45)=0.26;           %K_(pk,ADP)
75 k(46)=22.5;           %K_(pk,ATP)
76
77 k(47)=27171;          %v_(max,pdh)
78 k(48)=6.8;              %K_(m,PYR)
79 k(49)=0.4;            %K_(m,NAD)
80 k(50)=0.014;          %K_(m,CoA)
81 k(51)=0.1;            %K_(m,NADH)
82 k(52)=0.008;          %K_(m,AcCoA)
83 k(53)=46.4;          %K_(pdh,i)
84
85 k(54)=11.76;           %v_(max,acc)
86 k(55)=0.1;            %K_(i,ATP)
87 k(56)=0.1;           %K_ATP
88 k(57)=6.3;            %K_AcCoA
89 k(58)=0.30816;          %v_max_thns
```

```
k(59)=0.00358;     %K_(m,thns)

k(60)=1288800; %k_lac
k(61)= 0.26; % kt,lac
k(62)= 0.271 ;   %Ki_glu
k(63)=14.62; %Kt,LAcin
k(64)= 572400; %kcal,lac
k(65)= 0.014; %Km,lac

%Cofactors concentrations
ADP =    0.595;
AMP =    0.955;
ATP =    4.27;
CoA =    0.001;
NAD =    1.47;
NADH=    0.1;
PERM= 3*10^-5;
BETAGAL=0.012;

%Figures
options = odeset('RelTol',1e-8,'AbsTol',1e-10);
[t,x]=ode15s(@model_lactose, tspan, x0,options, k);

figure(1)
subplot(131)
t_exp=[0 2 4 6 8 20 24 48];
exp_X1=[4     5.1780    20.7778    43.9311    47.8401    49.5657    50.2316
      50.7498];

plot(t, x(:,1),'green','LineWidth',4);
hold on
plot(t_exp,exp_X1, 'o',  'LineWidth', 4)
ylim([0 70])
```

```
122 xlabel('Time (hours)');
123 ylabel('Biomass(mM)');
124 legend('Modeled', 'Experimental')
125 set(gca,'fontsize', 18)
126
127 subplot(132)
128
129 plot(t,x(:,2),'b—', t, x(:,12), 'm–', 'LineWidth',3);
130 legend('Glucose consumption', 'Lactose consumption');
131 xlabel('Time (hours)');
132 ylabel('Concentration (mM)');
133 ylim([0 70])
134 set(gca,'fontsize', 18)
135
136
137 subplot(133)
138 t_exp=[0 2 4 8 12 18 24 48];
139 exp_X11=[0 0.04380704 0.081981747 1.490352021 2.073011734 2.278226858
       2.482069 2.487075];
140
141 plot(t,x(:,11), 'red','LineWidth',4);
142 hold on
143 plot(t_exp,exp_X11, 'o', 'LineWidth', 4)
144
145 legend('Modeled', 'Experimental');
146 xlabel('Time (hours)');
147 ylabel('THN(mM)');
148 set(gca,'fontsize', 18)
149 ylim([0 3.7])
150
151 figure(5)
152
153 subplot(331)
```

```
154 plot(t, x(:,3),'LineWidth',4);
155 xlabel('Time (hours)');
156 ylabel('PEP (mM)');
157 set(gca,'fontsize', 18)
158 ylim([0 2])
159
160 subplot(332)
161  plot(t, x(:,4),'LineWidth',4);
162 xlabel('Time (hours)');
163 ylabel('PYR (mM)');
164 set(gca,'fontsize', 18)
165 ylim([0 2])
166
167 subplot(333)
168 plot(t, x(:,5), 'LineWidth',4);
169 xlabel('Time (hours)');
170 ylabel('G6P (mM)');
171 set(gca,'fontsize', 18)
172 ylim([0 2])
173
174  subplot(334)
175 plot(t, x(:,6),'LineWidth',4);
176 xlabel('Time (hours)');
177 ylabel('F6P(mM)');
178 set(gca,'fontsize', 18)
179 ylim([0 0.6])
180
181 subplot(335)
182  plot(t, x(:,7),'LineWidth',4);
183 xlabel('Time (hours)');
184 ylabel('F16BP (mM)');
185 set(gca,'fontsize', 18)
186 ylim([0 0.6])
```

```
187
188 subplot(336)
189 plot(t, x(:,8),'LineWidth',4);
190 xlabel('Time (hours)');
191 ylabel('GAP (mM)');
192 set(gca,'fontsize', 18)
193 ylim([0 0.6])
194
195 subplot(337)
196 plot(t, x(:,9),'LineWidth',4);
197 xlabel('Time (hours)');
198 ylabel('Acetyl-CoA (mM)');
199 set(gca,'fontsize', 18)
200 ylim([0 0.6])
201
202 subplot(338)
203 plot(t, x(:,10),'LineWidth',4);
204 xlabel('Time (hours)');
205 ylabel('Malonyl-CoA(mM)');
206 set(gca,'fontsize', 18)
207 ylim([0 0.6])
208
209 figure(6)
210 plot(t,x(:,13))
211
212
213 function [dxdt]= model_lactose(t,x,k)
214
215
216 %Cofactors
217 ADP =   0.595;
218 AMP =   0.955;
219 ATP =   4.27;
```

```
220 CoA =    0.001;
221 NAD =    1.47;
222 NADH=    0.1;
223 PERM= 3*10^-5; %M
224 BETAGAL=0.012; %mM
225
226
227 %rates
228
229
230 mu= k(1)*(1-x(1)/k(3))*(x(5)/(k(2)+x(5)));
231 v_pts= (k(7)*x(2)*x(3))/((k(8)*x(4)+k(9)*x(3)+k(10)*x(2)*x(4)+x(2)*x(3))*(1+(
        x(5)^k(11))/k(12)));
232
233 v_pgi= (k(13)*((x(5)-x(6)/k(14))))/(k(15)*(1+x(6)/k(16)+x(5)));
234 v_pk=((k(39)*x(3)*(x(3)/k(40)+1)^(k(41)-1))*ADP)/((k(40)*(k(42)*((1+ATP/k(46)
        )/(x(7)/k(43)+AMP/k(44)+1)))^k(41)+(x(3)/k(40)+1)^k(41))*(ADP+k(45)));
235 v_pdh=(k(47)/NAD*(1/(1+k(53)*NADH/NAD))*x(4)/k(48)*(1/k(49))*(CoA/k(50)))
        /((1+x(4)/k(48))*(1/NAD+1/k(49)+NADH/(k(51)*NAD)*(1+CoA/k(50)+x(9)/k(52))
        ));
236 v_gapdh= (k(33)*(x(8)-x(3)*NADH/(k(34)*NAD)))/((k(35)+(1+x(3)/k(36))+x(8))*((
        k(37)/NAD)*(1+NADH/k(38))+1));
237 v_pfk= (k(17)*k(18)*x(6))/(k(19)*(x(6)+k(20)*k(21)+x(3)/k(25))*(1+(k(22))/(1+
        x(6)*(k(23)/(k(20)*(k(21)+x(3)/k(25))^k(24))))));
238 v_aldo= (k(26)*(x(7)-(x(8))^2/k(27)))/(k(28)+x(7)+k(29)*x(8)/(k(27)*k(30))+k
        (31)*x(8)/(k(27)*k(30))+x(7)*x(8)/k(32)+(x(8))^2/(k(27)*k(30)));
239 v_acc=(k(54)*ATP*x(9))/(k(55)*k(57)+k(56)*x(9)+k(57)*ATP+ATP*x(9));
240 v_thns=(k(58)*(x(10))^5)/(k(59)^5+x(10)^5);
241 v_tlac=k(60)*PERM*((x(12)/(x(12)+k(61))*(k(62)/(k(62)+x(2))))-(x(13)/(x(13)+k
        (63))));
242 v_catlac=k(64)*BETAGAL*(x(13)/(x(13)+k(65)));
243
244 %Differential equations based on mass balance
```

```
245
246 dx1dt=mu*x(1);
247 dx2dt=-v_pts*x(1);
248 dx3dt=v_gapdh-v_pts-v_pk-mu*x(3);
249 dx4dt= v_pk+v_pts-v_pdh-mu*x(4);
250 dx5dt= v_pts+v_catlac-v_pgi-mu*x(5);
251 dx6dt=v_pgi-v_pfk-mu*x(6);
252 dx7dt=v_pfk-v_aldo-mu*x(7);
253 dx8dt=2*v_aldo-v_gapdh-mu*x(8);
254 dx9dt=v_pdh-v_acc-mu*x(9);
255 dx10dt=v_acc-v_thns-mu*x(10);
256 dx11dt= v_thns-mu*x(11);
257 dx12dt=-v_tlac*x(1);
258 dx13dt=v_tlac-v_catlac-mu*x(13);
259
260
261 dxdt=[dx1dt; dx2dt; dx3dt; dx4dt; dx5dt; dx6dt; dx7dt; dx8dt; dx9dt; dx10dt;
       dx11dt; dx12dt; dx13dt];
262
263 return
```

A.3. Stability Analysis

A.3.1. Jacobian Matrix

$$A = \begin{bmatrix} -3.6478e^{-18} & 0 & 0 & 0 & 20.274 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -11809.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1.5331e^{-29} & -204.43 & -6.6561 & 0 & -7.8619e^{-11} & 0 & -3.4393e^{-13} & 37.18 & 0 & 0 & 0 & 0 & 0 \\ 3.9716e^{-31} & 204.43 & 0.75457 & -29.874 & -2.0367e^{-12} & 0 & 3.4393e^{-13} & 0 & 3.771e^{-9} & 0 & 0 & 0 & 0 \\ 5.9598e^{-38} & 204.43 & 0 & 0 & -10.72 & 24.93 & 0 & 0 & 0 & 0 & 0 & 0 & 490633.0 \\ 2.3126e^{-38} & 0 & 7.1494e^{-19} & 0 & 10.72 & -28.677 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4.5944e^{-39} & 0 & -7.1494e^{-19} & 0 & -2.3561e^{-20} & 3.7468 & -21.306 & 7.5149e^{-10} & 0 & 0 & 0 & 0 & 0 \\ 1.6898e^{-30} & 0 & 5.9016 & 0 & -8.6656e^{-12} & 0 & 42.612 & -37.18 & 0 & 0 & 0 & 0 & 0 \\ 1.0369e^{-29} & 0 & 0 & 29.874 & -5.3173e^{-11} & 0 & 0 & 0 & -1.824 & 0 & 0 & 0 & 0 \\ 2.7506e^{-23} & 0 & 0 & 0 & -0.00014106 & 0 & 0 & 0 & 1.824 & -0.068358 & 0 & 0 & 0 \\ 1.9017e^{-19} & 0 & 0 & 0 & -0.97522 & 0 & 0 & 0 & 0 & 0.068358 & -3.0563e^{-19} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8590.0 & 152.76 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 148.71 & -490633.0 \end{bmatrix}$$

A.3.2. Matlab Code

```
clc
close all
clear all

syms x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13


   X=[x1;x2;x3;x4;x5;x6;x7;x8;x9;x10;x11;x12;x13];
   A=jacobian(@model_Stab_lacglu,X);


A1=subs(A,[{x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13}],[57.7644,0,2.24e
      -10,5.803e-12,8.708e-19, 3.379e-19, 6.713e-20, 2.469e-11,1.515e
      -10,0.0004019,2.7786,0,0]);
EB=vpa(eig(A1),5)   %find A eigenvalues

StaB= vpa(EB,7);  % simplify to 7 significant digits
if all(real(StaB)<0)
     StaB
end

function dxdt = model_Stab_lacglu

syms x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 X

k(1)=0.9782;          %u_m
k(2)=0.2;             %K_s
k(3)=62.23;           %X_m

k(7)=25.739;          %v_(max,pts)
k(8)=1;               %K_(pts,a1)
```

```
30 k(9)=0.1;          %K_(pts,a2)
31 k(10)=1;           %K_(pts,a3)
32 k(11)=4;           %n_G6P
33 k(12)=0.5;         %K_(pts,G6P)
34
35 k(13)=26.3711;     %v_(max,pgi)
36 k(14)=0.43;        %K_(eq,pgi)
37 k(15)=2.46;        %K_(m,G6P)
38 k(16)=0.2;         %K_(m,F6P)
39
40 k(17)=24.613;      %v_(max,pfk)
41 k(18)=4.27;        %K_(pk,ATP)
42 k(19)=4.6944;      %K_(pk,ATP,ADP)
43 k(20)=0.14;        %K_(s,F6P)
44 k(21)=3.88;        %K_b(ADP,AMP)
45 k(22)=10;          %L_pfk
46 k(23)=1.1118;      %K_(a,(ADP,AMP) )
47 k(24)=4;           %n_pfk
48 k(25)=3.26;        %K_PEP
49
50 k(26)=2.8337;      %v_(max,aldo)
51 k(27)=1.4;         %K_(aldo,eq)
52 k(28)=0.133;       %K_(aldo,F16BP)
53 k(29)=0.088;       %K_(aldo,GAP)
54 k(30)=2;           %V_blf
55 k(31)=0.088;       % K_(aldo,DHAP)
56 k(32)=0.6;         %K_(aldo,gap inh)
57
58 k(33)=121.29;      %v_(max,gapdh)
59 k(34)=0.63;        %K_(gapdh,eq)
60 k(35)=0.15;        %K_(ggapdh,GAP)
61 k(36)=0.1;         %K_(gapdh,pgp )
62 k(37)=0.45;        %K_(gapdh,NAD)
```

```
63 k(38)=0.02;          %K_(gapdh,NADH)
64
65 k(39)=1.0849;        %v_(max,pk)
66 k(40)=0.31;          %K_(pk,PEP)
67 k(41)=4;             %n_pk
68 k(42)=1;             %L_pk
69 k(43)=0.19;          %K_(pk,F16BP)
70 k(44)=0.2;           %K_(pk,AMP)
71 k(45)=0.26;          %K_(pk,ADP)
72 k(46)=22.5;          %K_(pk,ATP)
73
74 k(47)=27171;         %v_(max,pdh)
75 k(48)=6.8;           %K_(m,PYR)
76 k(49)=0.4;           %K_(m,NAD)
77 k(50)=0.014;         %K_(m,CoA)
78 k(51)=0.1;           %K_(m,NADH)
79 k(52)=0.008;         %K_(m,AcCoA)
80 k(53)=46.4;          %K_(pdh,i)
81
82 k(54)=11.76;          %v_(max,acc)
83 k(55)=0.1;            %K_(i,ATP)
84 k(56)=0.1;            %K_ATP
85 k(57)=6.3;            %K_AcCoA
86 k(58)=0.30816;        %v_max_thns
87 k(59)=0.00358;        %K_(m,thns)
88
89 k(60)=1288800;          %k_lac,in
90 k(61)= 0.26;            %kt,lac
91 k(62)= 0.271 ;          %Ki_glu
92 k(63)=14.62;            %Kt,LAcin
93 k(64)= 572400;          %kcal,lac
94 k(65)= 0.014;           %Km,lac
95
```

```
96 %Cofactors concentrations
97 ADP = 0.595;
98 AMP =    0.955;
99 ATP =    4.27;
100 CoA =    0.001;
101 NAD =    1.47;
102 NADH= 0.1;
103 PERM= 3*10^-5;
104 BETAGAL=0.012;
105
106
107
108 mu= k(1)*(1-x1/k(3))*(x5/(k(2)+x5));
109 v_pts= (k(7)*x2*x3)/((k(8)*x4+k(9)*x3+k(10)*x2*x4+x2*x3)*(1+(x5^k(11))/k(12))
        );
110 v_pgi= (k(13)*((x5-x6/k(14))))/(k(15)*(1+x6/k(16)+x5));
111 v_pk=(k(39)*x3*(x3/k(40)+1)^(k(41)-1)*ADP)/((k(40)*(k(42)*((1+ATP/k(46))/(x7/
        k(43)+AMP/k(44)+1)))^k(41)+(x3/k(40)+1)^k(41))*(ADP+k(45)));
112 v_pdh=(k(47)/NAD*(1/(1+k(53)*NADH/NAD))*x4/k(48)*(1/k(49))*(CoA/k(50)))/((1+
        x4/k(48))*(1/NAD+1/k(49)+NADH/(k(51)*NAD)*(1+CoA/k(50)+x9/k(52))));
113 v_gapdh= (k(33)*(x8-x3*NADH/k(34)))/((k(35)+(1+x3/k(36))+x8)*((k(37)/NAD)*(1+
        NADH/k(38))+1));
114 v_pfk= (k(17)*k(18)*x6)/(k(19)*(x6+k(20)*k(21)+x3/k(25))*(1+(k(22))/(1+x6*(k
        (23)/(k(20)*(k(21)+x3/k(25))^k(24))))));
115 v_aldo= (k(26)*(x7-(x8)^2/k(27)))/(k(28)+x7+k(29)*x8/(k(27)*k(30))+k(31)*x8/(
        k(27)*k(30))+x7*x8/k(32)+(x8)^2/(k(27)*k(30)));
116 v_acc=(k(54)*ATP*x9)/(k(55)*k(57)+k(56)*x9+k(57)*ATP+ATP*x9);
117 v_thns=(k(58)*(x10)^5)/(k(59)^5+(x10)^5);
118 v_tlac=k(60)*PERM*((x12/(x12+k(61))*(k(62)/(k(62)+x2)))-(x13/(x13+k(63))));
119 v_catlac=k(64)*BETAGAL*(x13/(x13+k(65)));
120
121
122 dx1dt=mu*x1;
```

```
123 dx2dt=-v_pts*x1;
124 dx3dt=v_gapdh-v_pts-v_pk-mu*x3;
125 dx4dt= v_pk+v_pts-v_pdh-mu*x4;
126 dx5dt= v_pts+v_catlac-v_pgi-mu*x5;
127 dx6dt=v_pgi-v_pfk-mu*x6;
128 dx7dt=v_pfk-v_aldo-mu*x7;
129 dx8dt=2*v_aldo-v_gapdh-mu*x8;
130 dx9dt=v_pdh-v_acc-mu*x9;
131 dx10dt=v_acc-v_thns-mu*x10;
132 dx11dt= v_thns-mu*x11;
133 dx12dt=-v_tlac*x1;
134 dx13dt=v_tlac-v_catlac-mu*x13;
135
136
137 v=[mu v_pts v_pgi v_pk v_pdh v_gapdh v_pfk v_aldo v_acc v_thns v_tlac
       v_catlac];
138 X=[x1;x2;x3;x4;x5;x6;x7;x8;x9;x10;x11;x12;x13];
139
140 dxdt=[dx1dt; dx2dt; dx3dt; dx4dt; dx5dt; dx6dt; dx7dt; dx8dt; dx9dt; dx10dt;
       dx11dt; dx12dt; dx13dt];
141
142 end
```

A.4. MCA

- Stoichiometric matrix (N)

$$N = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\
-1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
-1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
\end{bmatrix}$$

- Kinetic rate vector (v)

$$v = \begin{bmatrix} \mu \\ \mathrm{v}_{pts} \\ \mathrm{v}_{pgi} \\ \mathrm{v}_{pk} \\ \mathrm{v}_{pdh} \\ \mathrm{v}_{gapdh} \\ \mathrm{v}_{pfk} \\ \mathrm{v}_{also} \\ \mathrm{v}_{acc} \\ \mathrm{v}_{thns} \\ \mathrm{v}_{lac} \\ \mathrm{v}_{catlac} \end{bmatrix}$$

- Matlab Code

```
clc
close all
clear all

syms x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13

options=optimset('Display', 'iter');
 X=[x1;x2;x3;x4;x5;x6;x7;x8;x9;x10;x11;x12;x13];
 A=jacobian(@model_Stab_lacglu,X);
 A1=subs(A,[{x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13}],[57.7644,3.59e
    -61,2.24e-10,5.803e-12,8.708e-19, 3.379e-19, 6.713e-20, 2.469e-11,1.515e
    -10,0.0004019,2.7786,1.27e-130,1.092e-136]);

A_i=inv(A1); %inverse of jacobian  with equilibrium points

```

```
14 dv_dx=jacobian(@rates,X);
15
16 xe=[57.7644 3.59e-61 2.24e-10 5.803e-12 8.708e-19 3.379e-19  6.713e-20 2.469e
       -11 1.515e-10 0.0004019 2.7786 1.27e-130 1.092e-136];                %%
       equilibrium point
17
18 dvdx_xe=subs(dv_dx,[{x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13
       }],[57.7644,3.59e-61,2.24e-10,5.803e-12,8.708e-19, 3.379e-19, 6.713e-20,
       2.469e-11,1.515e-10,0.0004019,2.7786,1.27e-130,1.092e-136]);
19 v_xe=subs(@rates,[{x1,x2,x3,x4,x5,x6,x7,x8,x9,x10,x11,x12,x13}],[57.7644,3.59
       e-61,2.24e-10,5.803e-12,8.708e-19, 3.379e-19, 6.713e-20, 2.469e-11,1.515e
       -10,0.0004019,2.7786,1.27e-130,1.092e-136]);
20
21 N= [1  0  0  0  0  0  0  0  0  0  0  0;
22       0 -1  0  0  0  0  0  0  0  0  0  0;
23       -1 -1  1 -1  0  0  0  0  0  0  0  0;
24       -1  1  0  1 -1  0  0  0  0  0  0  0;
25       -1  1  0  0  0 -1  0  0  0  0  0  1;
26       -1  0  0  0  0  1 -1  0  0  0  0  0;
27       -1  0  0  0  0  0  1 -1  0  0  0  0;
28       -1  0 -1  0  0  0  0  2  0  0  0  0;
29       -1  0  0  0  1  0  0  0 -1  0  0  0;
30       -1  0  0  0  0  0  0  0  1 -1  0  0;
31       -1  0  0  0  0  0  0  0  0  1  0  0 ;
32       0  0  0  0  0  0  0  0  0  0 -1  0;
33       -1  0  0  0  0  0  0  0  0  0  1  -1];
34
35
36  Cs= vpa(-A_i*N, 5)    %%%% CONCENTRATION CONTRL MATRIX %%%%%% unscaled
37
38  Cj=vpa(eye(12)-dvdx_xe*A_i*N,5); %%%% FLUX CONTROL MATRIX %%%%%%%unscaled
39
40
```

```
%%%      SCALING      %%%%%

xe_new= diag(xe);
v_xe_new=diag(v_xe);

Cs_new=vpa(inv(xe_new)*Cs*v_xe_new,5)
Cj_new=vpa(inv(v_xe_new)*Cj*v_xe_new,5)

Sum_Cj=vpa(sum(Cj_new,2),5)



function v=rates(~,~)
syms x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 X

k1=0.9782;
k2=0.2;
k3=62.23;
k4=0;
k5=0;
k6=0;
k7=25.739;
k8=1;
k9=0.1;
k10=1;
k11=4;
k12=0.5;
k13=26.3711;
k14=0.43;
k15=2.46;
k16=0.2;
k17=24.613;
k18=4.27;
```

```
 74 k19=4.6944;
 75 k20=0.14;
 76 k21=3.88;
 77 k22=10;
 78 k23=1.1118;
 79 k24=4;
 80 k25=3.26;
 81 k26=2.8337;
 82 k27=1.4;
 83 k28=0.133;
 84 k29=0.088;
 85 k30=2;
 86 k31=0.088;
 87 k32=0.6;
 88 k33=121.29;
 89 k34=0.63;
 90 k35=0.15;
 91 k36=0.1;
 92 k37=0.45;
 93 k38=0.02;
 94 k39=1.0849;
 95 k40=0.31;
 96 k41=4;
 97 k42=1;
 98 k43=0.19;
 99 k44=0.2;
100 k45=0.26;
101 k46=22.5;
102 k47=27171;
103 k48=6.8;
104 k49=0.4;
105 k50=0.014;
106 k51=0.1;
```

```
107 k52=0.008;
108 k53=46.4;
109 k54=11.76;
110 k55=0.1;
111 k56=0.1;
112 k57=6.3;
113 k58=0.30816;
114 k59=0.00358;
115 k60=1288800;
116 k61=0.26;
117 k62=0.271;
118 k63=14.62;
119 k64=572400;
120 k65=0.014;
121
122 %Cofactors concentrations
123 ADP = 0.595;
124 AMP =  0.955;
125 ATP =  4.27;
126 CoA =  0.001;
127 NAD =  1.47;
128 NADH= 0.1;
129 PERM= 3*10^-5; %M
130 BETAGAL=0.012; %mM
131
132 mu= k1*(1-x1/k3)*(x5/(k2+x5));
133 v_pts= (k7*x2*x3)/((k8*x4+k9*x3+k10*x2*x4+x2*x3)*(1+(x5^k11)/k12));
134 v_pgi= (k13*((x5-x6/k14)))/(k15*(1+x6/k16+x5));
135 v_pk=(k39*x3*(x3/k40+1)^(k41-1)*ADP)/((k40*(k42*((1+ATP/k46)/(x7/k43+AMP/k44
      +1)))^k41+(x3/k40+1)^k41)*(ADP+k45));
136 v_pdh=(k47/NAD*(1/(1+k53*NADH/NAD))*x4/k48*(1/k49)*(CoA/k50))/((1+x4/k48)*(1/
      NAD+1/k49+NADH/(k51*NAD)*(1+CoA/k50+x9/k52)));
137 v_gapdh= (k33*(x8-x3*NADH/k34))/((k35+(1+x3/k36)+x8)*((k37/NAD)*(1+NADH/k38)
```

```
    +1));
138 v_pfk= (k17*k18*x6)/(k19*(x6+k20*k21+x3/k25)*(1+k22/(1+x6*(k23/(k20*(k21+x3/
    k25)^k24)))));
139 v_aldo= (k26*(x7-(x8)^2/k27))/(k28+x7+k29*x8/(k27*k30)+k31*x8/(k27*k30)+x7*x8
    /k32+(x8)^2/(k27*k30));
140 v_acc=(k54*ATP*x9)/(k55*k57+k56*x9+k57*ATP+ATP*x9);
141 v_thns=(k58*(x10)^5)/(k59^5+(x10)^5);
142 v_tlac=k60*PERM*((x12/(x12+k61)*(k62/(k62+x2)))-(x13/(x13+k63)));
143 v_catlac=k64*BETAGAL*(x13/(x13+k65));
144
145
146 v=[mu; v_pts; v_pgi; v_pk; v_pdh; v_gapdh; v_pfk; v_aldo ;v_acc ;v_thns;
    v_tlac; v_catlac];
147 X=[x1;x2;x3;x4;x5;x6;x7;x8;x9;x10;x11;x12;x13];
148
149
150 end
```

A.5. Pictures of Experiments

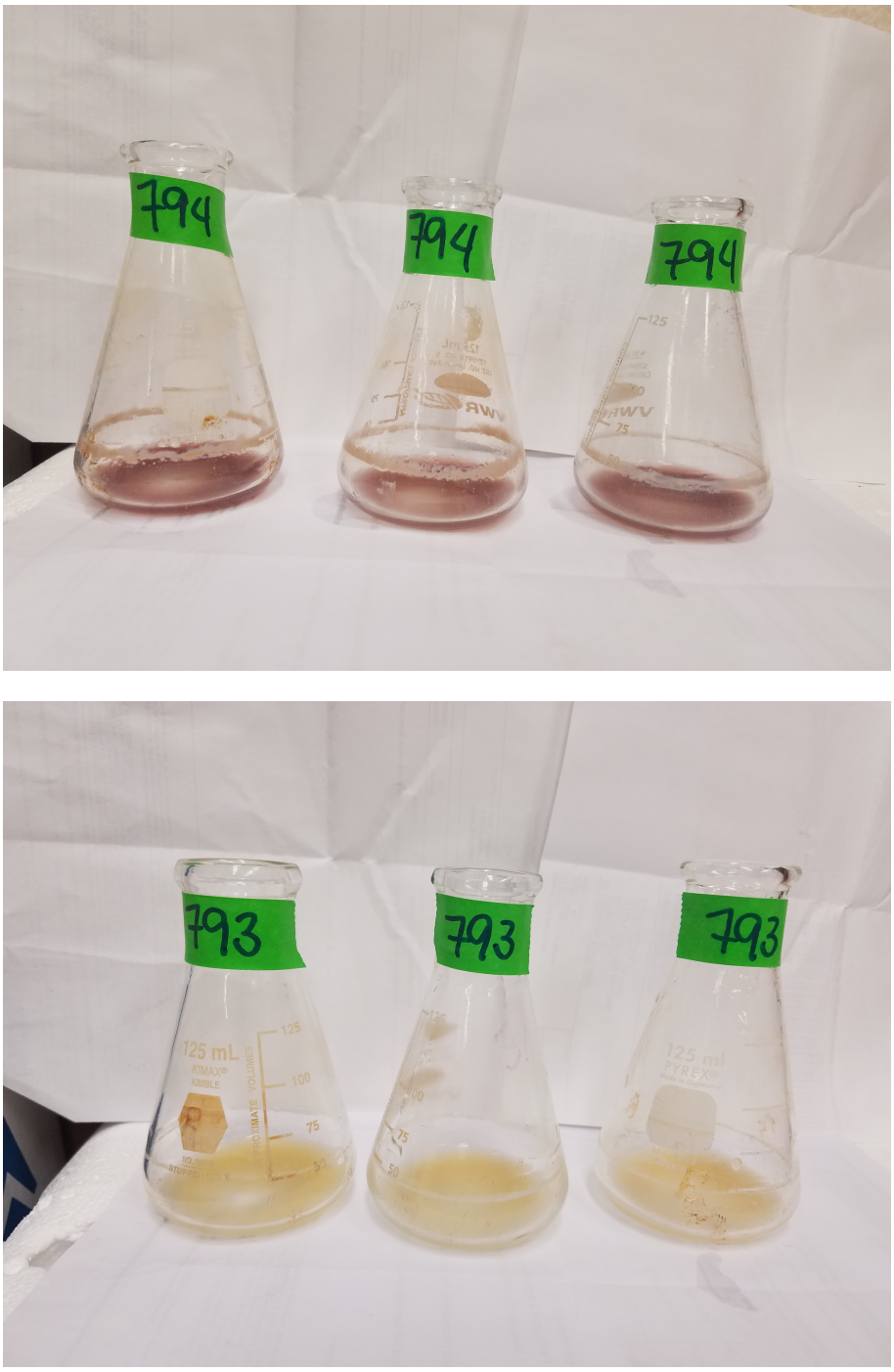
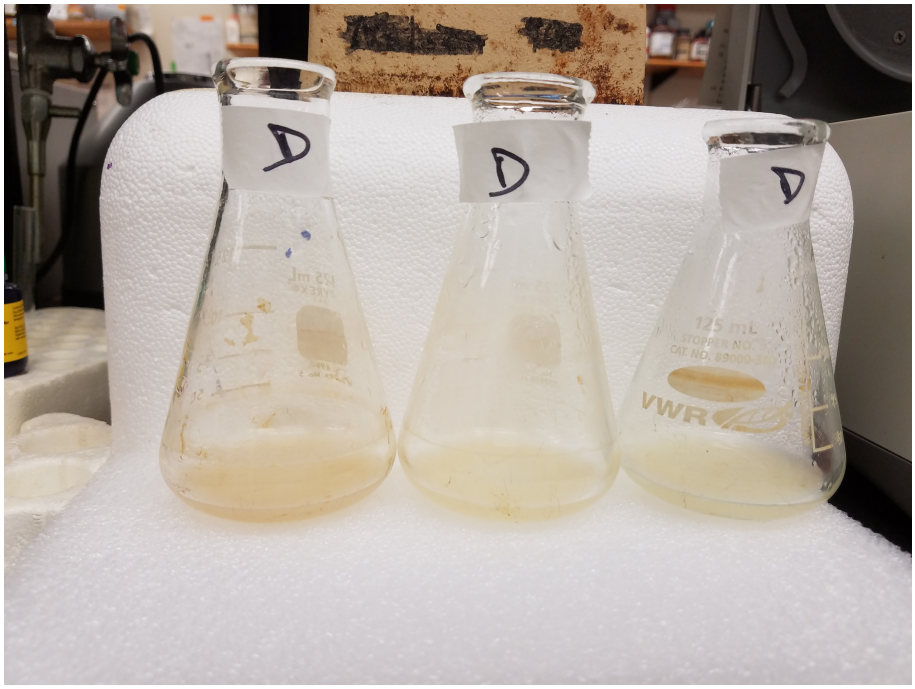


Figure A.1. Flaviolin production without genetic modification (top) and with holo-ACC being overexpressed (bottom)



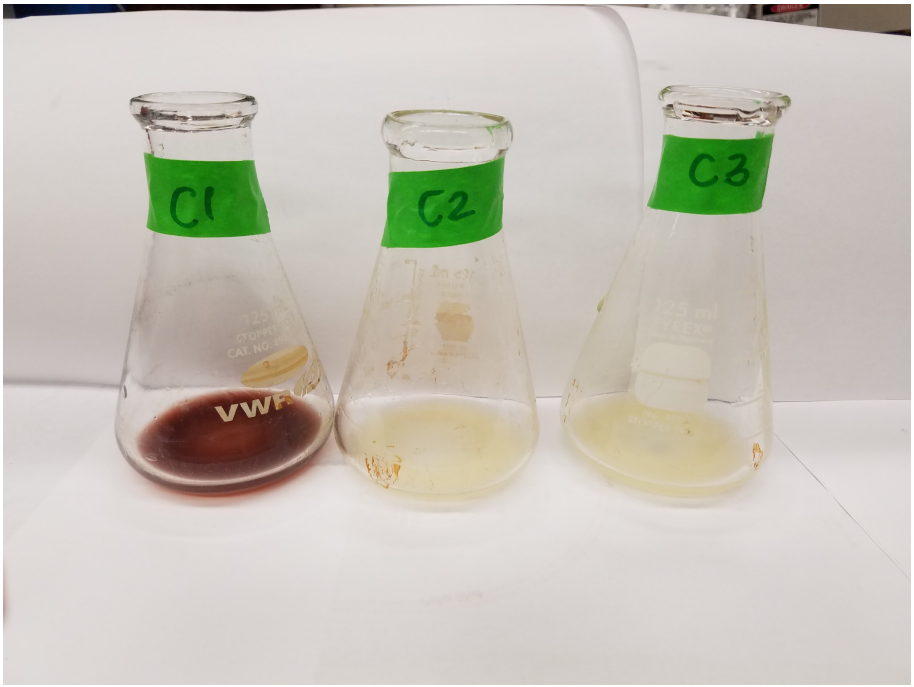


Figure A.2. Lack of reproducibility observed at the beginning of this study

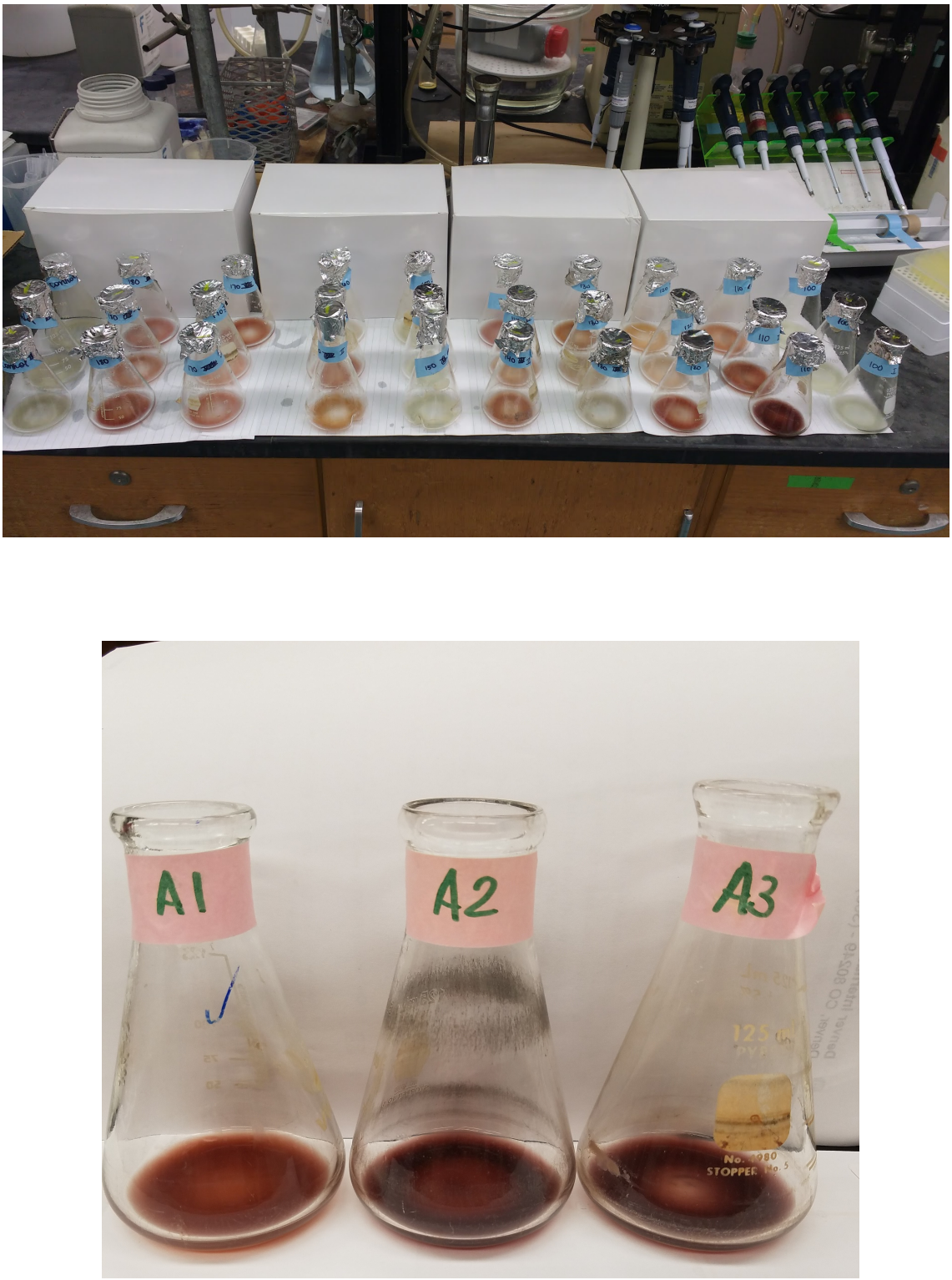


Figure A.3. Glucose experiment (top) and samples with the highest flaviolin production (bottom)

Vita

Tatiana T. S. Mello was born in Piracicaba, Brazil. Before coming to Louisiana State University, she completed two bachelor degrees at University of Campinas (Unicamp), Brazil: Biological Sciences (2011) and Mechanical Engineering (2016). Her research interest is in Bioprocess Engineering, which deals with the design of the process and equipment for the manufacturing of bioproducts.