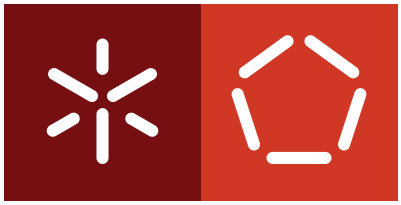
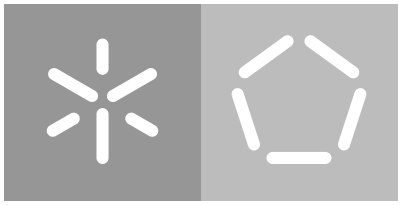


Universidade do Minho
Escola de Engenharia
Engenharia Biológica

Alexandre Rafael Machado Oliveira

A kinetic model of the central carbon metabolism for acrylic acid production in *Escherichia coli*

December 2019



Universidade do Minho
Escola de Engenharia
Engenharia Biológica

Alexandre Rafael Machado Oliveira

A kinetic model of the central carbon metabolism for acrylic acid production in *Escherichia coli*

Master dissertation
Master Degree in Bioinformatics

Dissertation supervised by
Oscar Manuel Lima Dias
Joana Lúcia de Lima Correia Rodrigues

December 2019

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

ACKNOWLEDGMENTS

First of all, I would like to formally acknowledge my supervisors, Doctor Oscar Dias and Doctor Joana Rodrigues, for all the support they provided. To both of them, a very special word of appreciation, not only for their dedication and knowledge transmitted but especially for guiding me through this entire year and for helping me overcome all the obstacles that have arisen from this work.

I also would like to thank all the colleagues that worked by my side at the Biosystems group. Especially to Fernando Cruz, Marta Sampaio, and José Sousa for helping get accustomed to new subjects that I had to learn through the year.

A particular word to all my friends for all the good moments they presented me over the last years. Also, for always giving me motivation and advises, and for making me strive to accomplish more than I ever thought I could. To all of you, I am very thankful and hope you know that you can count on me for the years to come, as I can count on you. In particular, I would also like to thank my dorm-mates, Miguel, Freitas, Nuno and Milk, for all the remarkable moments we had and for making my stay in Braga a living hell, and to Celso for never being there.

Last but not least, a very special word to my family, without whom nothing of this would be possible. Thank you for all the support you provided, and for dealing and helping with my stress and anxiety daily. For always believing in what I was capable, even when I didn't, and for always being there for me no matter how difficult things got. To my parents, I know how much of an effort this last five years required from you, both financially and emotionally. But no matter what, you guys make sure that nothing was missing so that I could pursue a better education. I can undoubtedly say that this is all thanks to you, thus I reserved my most significant praise to both of you.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

MODELO CINÉTICO DO METABOLISMO DO CARBONO CENTRAL PARA PRODUÇÃO DE ÁCIDO ACRÍLICO EM *ESCHERICHIA COLI*

RESUMO

O Ácido Acrílico (AA) é um composto químico que atualmente possui um valor de mercado considerável. Contudo, a maioria do AA comercializado ainda é proveniente da oxidação de derivados do petróleo, como o propileno e o propeno, método que se revela dispendioso e que contribui para o deterioramento do planeta. Assim sendo, existe a necessidade de encontrar novos métodos, mais sustentáveis, para a produção deste composto.

Recentemente a literatura tem demonstrado que o uso de uma via fermentativa, que permita a produção de AA em *Escherichia coli* usando o ácido 3-hidroxipropanoico (3-HP) como intermediário, pode vir a ter aplicações à escala industrial. Atualmente conhecem-se três vias que permitem a produção de AA através do glicerol, malonil-CoA, ou β-alanina, distinguidas apenas na via para produção de 3-HP.

Os objetivos deste trabalho foram a implementação dessas vias num modelo cinético do metabolismo do carbono central de *E. coli* para comparar as vias, e, subsequentemente, encontrar possíveis estratégias de optmização que permitam aumentar a produção de AA. Para isso, este trabalho gerou 12 modelos capazes de simular a produção de 3-HP e AA a partir de glucose our glicerol.

Uma vez que o método para o cálculo dos V_{max} na via heteróloga foi usado com o intuito de evitar a acumulação de intermediários, estes modelos foram mais eficazes na previsão da produção de 3-HP do que de AA. Não obstante, foi concluído que a escolha da fonte de carbono deve depender da via utilizada. Sendo que, para a via do glicerol parece ser mais eficiente usar glicerol, enquanto para as restantes duas vias, o uso de glucose pode ser benéfico. Além disso, este trabalho também sugere que a utilização da via do malonil-CoA pode trazer vantagens para uma produção à escala industrial, uma vez que, ao contrário da via do glicerol, esta não necessita de suplementação de vitaminas, permitindo ainda assim obter uma boa produção de AA. Para finalizar, este trabalho também propõe algumas enzimas que podem ser sobre-expressadas para melhorar a produção de AA. Para a via do glicerol as enzimas são a glicerol-3-fosfato desidrogenase e a glicerol-3-fosfato fosfatase. No caso da via do malonil-CoA é a acetyl-CoA carboxilase. Finalmente, para a via da β-alanina, os alvos sugeridos são a aspartato aminotransferase e a aspartato carboxilase.

Palavras chave: Ácido 3-Hidroxipropanoico, Ácido Acrílico, *E. coli*, Modelos Cinéticos.

A KINETIC MODEL OF THE CENTRAL CARBON METABOLISM FOR ACRYLIC ACID PRODUCTION IN *ESCHERICHIA COLI*

ABSTRACT

Acrylic Acid (AA) is a chemical compound that nowadays has a considerable market value. However, the majority of AA used worldwide still comes from the oxidation of petroleum derivatives, like propylene and propane, which revealed to be expensive and to contribute towards the planet deterioration. Hence, there is currently a need for new and sustainable ways to synthesize this compound.

Recently, the literature has been demonstrating the use of a direct fermentation route that allows *Escherichia coli* to produce AA using 3-hydroxypropionate (3-HP) as an intermediary, and its subsequent application to industrial-scale production. There are three distinct pathways to produce AA, via glycerol, malonyl-CoA, or β-alanine, which mainly differ in the route towards 3-HP production.

The goals of this work were the implementation of these three distinct pathways in a dynamic model of the central carbon metabolism of *E. coli*, in order to compare the pathways, and, subsequently, find possible optimization targets to increase AA production. With that in mind, this work generated 12 models that can simulate 3-HP and AA production from either glucose or glycerol.

Because the method for V_{max} calculation in the heterologous pathway was used to prevented intermediary accumulation, these models were more effective in predicting the production of 3-HP than AA. Despite that, it was concluded that the best carbon source to produce AA depends on the pathway used. With the glycerol route, it seems more efficient to use glycerol, while with the malonyl-CoA and β-alanine routes, glucose appears to be more beneficial. Besides, the work also suggested that using the malonyl-CoA route might be beneficial for industrial-scale production, as it does not need any supplementation of vitamins, contrary to the glycerol route, while also presenting good AA yields. Furthermore, this work also suggests some enzymes that were considered as targets for over-expression in each pathway. For the glycerol route, these enzymes were the glycerol-3-phosphate dehydrogenase and the glycerol-3-phosphate phosphatase. In the malonyl-CoA route, it was the acetyl-CoA carboxylase. Finally, for the β-alanine route, the suggested targets were the aspartate aminotransferase and the aspartate carboxylase.

Keywords: 3-Hydroxypropionate, Acrylic Acid, *E. coli*, Kinetic Models.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

GLSSYMBOLS

3-HP 3-hydroxypropionate.

3-HP-COA 3-hydroxypropionyl-CoA.

3-HPA 3-hydroxypropionaldehyde.

A

AA Acrylic acid.

AA-COA acrylyl-CoA.

ATP adenosine triphosphate.

B

BRENDA Braunschweig Enzyme Database.

C

CCM central carbon metabolism.

COA coenzyme A.

COPASI COmplex PAthway SImulator.

D

DAP dihydroxyacetone phosphate.

F

FAD flavin adenine dinucleotide.

FCC flux control coefficients.

G

GTP guanosine triphosphate.

K

KEGG Kyoto Encyclopedia of Genes and Genomes.

L

LA lactic acid.

LSODE Livermore Solver of Ordinary Differential Equations.

M

MCA metabolic control analysis.

MSA malonic semialdehyde.

N

NAD nicotinamide adenine dinucleotide.

NADP nicotinamide adenine dinucleotide phosphate.

O

ODE ordinary differential equations.

P

PP penthose-phosphate.

S

SA specific activity.

SBML Systems Biology Markup Language.

SED-ML Simulation Experiment Description Markup Language.

T

TCA tricarboxylic acid.

X

XML Extensible Markup Language.

1

INTRODUCTION

1.1 CONTEXT AND MOTIVATION

Acrylic acid (AA) ($C_3H_4O_2$) is an important chemical that is widely used as an industrial feedstock to the production of a wide variety of products daily used such as polymeric flocculants, paints, adhesives, coatings, dispersants, and binders for leather, paper, and textile [1]. Besides, AA is also one of the key components of several polymeric products, such as superabsorbent polymers, which further increases its value [2]. According to the *Allied Market Research*, in 2015 the global market for AA was valued at 12,500 million US dollars. Furthermore, according to their prediction, between 2016 and 2022, the value is expected to increase 6.6% per year, reaching a total of 19,500 million US dollars, which confirms the economic importance of this product [3].

Despite its commercial value, the vast majority of AA is still produced by the oxidation of propylene or propane. In this process, those metabolites are oxidized to acrolein, and then directly converted into AA in a purely chemical process with high energy demand (Figure 1) [4, 2, 5, 6]. Ergo, the principal method for AA production is highly dependent on the world's petroleum reserves, which are not renewable and are in rapid decline, greatly enhancing the production cost. In addition, the oxidization to AA contributes to the carbon dioxide accumulation in the atmosphere, leading to greenhouse effect aggravation, which is one of the most significant problems humankind is currently facing [4, 1]. With this in mind, it is possible to conclude that this chemical process is expensive, has a high energy demand, and it contributes to the planet's deterioration. Hence, the development of an innovative, clean, and sustainable biological method for its production has attracted considerable attention from the scientific community in recent years [4, 7, 1].

In the last decade, several semi-biological methods have arisen and were subsequently optimized. Those methods usually consist of a first step where an adequate organism produces an intermediate through fermentation, and a final step where AA is produced separately by a chemical process [4, 2, 5]. From such semi-biological methods, two stand out because of the promising results obtained in several different works, namely, the *lactic acid (LA)* route and the *3-hydroxypropionate (3-HP)* route (Figure 1).

The first route consists in the bio-based production of LA and a subsequent chemical step to convert it into AA. Even though LA production from bacterial pathway has presented significantly high yields, its dehydration to AA still presents selectivity issues, which affect the resulting yield (around 78%). Alternatives to this catalytic step have been found and have proven to be efficient, as the conversion of LA to 2-acetoxypropionic acid and the subsequent pyrolysis to AA reached a yield over 90%. However, this process still lacks further studying to be applied for industrial-scale production [2, 8, 9]. Finally, the second route consists in the fermentation of simple sugars to produce 3-HP, which is then purified and converted to AA by catalytic dehydration. Compared to the dehydration of LA, this method does not present selectivity issues in the catalytic step and has presented very high yields when converting 3-HP to AA (around 97%). However, despite the current efforts, the fermentation step is still associated with low 3-HP production [4, 2, 8].

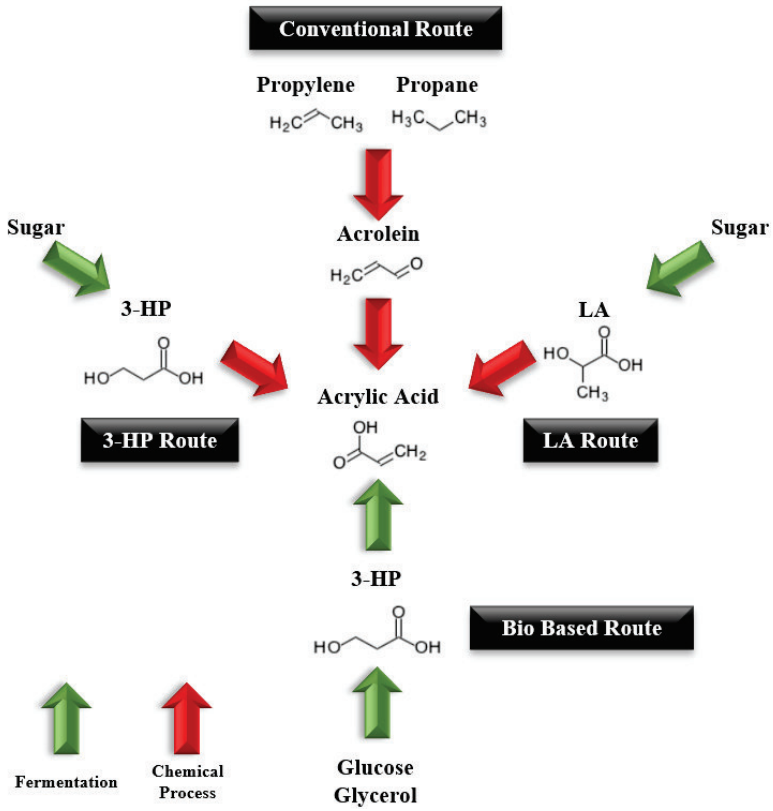


Figure 1.: Representation of the known processes for acrylic acid (AA) production. There are three main ways to produce it: through purely chemical methods, from propylene and propane; through semi-biological methods, where selected organisms produce some intermediates, such as 3-hydroxypropionic acid (3-HP) and lactic acid (LA), that are then transformed into AA through chemical processes; and finally, from a bio-based direct route, where living cells express the entire AA pathway using carbon sources like glucose and glycerol. Adapted from Chu *et al.* (2015) [4].

Although substantial improvements were obtained when compared to the synthesis from propylene and propane, those semi bio-based methods always include a chemical step, such as purification of the intermediate through chemical extraction or the use of chemical catalysts. Generally, those steps are associated with high energy demand, which increases production costs, and consequently, diminishes the economic return that could be achieved if those steps were not required [4, 1]. With this in mind, the ideal method should be a bio-based direct AA route that does not require any chemical catalysts [4].

Fortunately, in recent years, it has been proven that it is possible to use engineered *Escherichia coli* to perform the complete conversion of glucose and glycerol into AA. That can be achieved by inserting the heterologous pathways for AA biosynthesis that were already identified into the bacteria through synthetic biology approaches. In theory, as microbial feedstocks would be cheaper to maintain and the energy demands would be much lower, this method would allow higher profit margins than the other approaches. Unfortunately, despite being a very promising pathway, the AA yields reported by Tong *et al.* (2016) (37.7 mg/L in 48 hours), Chu *et al.* (2015) (0.12 g/L in 15 hours), and Liu and Liu (2016) (13 mg/L in 24 hours) demonstrated that this process still needs to be optimized to compete at an industrial scale with the currently used methods [4, 7, 1, 2].

1.2 GOALS

The goal of this project was to use kinetic models of *E. coli central carbon metabolism (CCM)* to determine which pathway used in the literature to produce AA had a higher potential, and how to optimize the biosynthetic pathways to further improve the production.

In more detail, this dissertation aimed to:

- Identify the distinct pathways for AA production that used 3-HP as an intermediary and their respective kinetic information;
- Insert the heterologous pathways into the model, and perform *in silico* simulations to determine which pathway had a better AA yield;
- Test the use of glucose or glycerol as carbon sources to assess which is associated with an higher AA production;
- Identify possible optimization strategies at a genetic level that may increase AA production.

1.3 THESIS OUTLINE

The document was organized as follows:

- **Chapter 2**: State-of-the-Art
  - Overview of systems biology;
  - Brief introduction to the databases used to obtain information;
  - Theoretical introduction to kinetic models;
  - Brief overview of the available kinetic models for the central carbon metabolism of *E. coli*;
  - Theoretical overview of the bioinformatics tools to be used in this work;
  - Overview of the *E. coli* importance and metabolism;
  - Mapping of the pathways that are relevant for the development of a bio-based direct AA route.
  - Literature review concerning 3-HP and AA production *in vivo*.
- **Chapter 3**: Materials and Methods
  - Selected model and its characteristics;
  - Description of the tasks associated to kinetic data annotation, and kinetic modeling;
  - Kinetic description of all the reactions to be inserted into the model;
  - Detailed information regarding how the time course simulations here performed once the model was assembled;
  - Description of the optimization strategies used to improve AA production.
- **Chapter 4**: Results and Discussion
  - Results of the V_{max} calculation;
  - Results of the *in silico* simulations of AA production, and comparison with published results for *in vivo* production;
  - Results of the *in silico* optimization of AA yields using the methods that were described in chapter three.
- **Chapter 5**: Conclusion and Further Work

2

STATE-OF-THE-ART

This chapter aims to reflect on the importance of systems biology and bioinformatics in metabolic engineering. More specifically, the relevance of kinetic models to predict the production of a product of interest and to identify optimization strategies to improve the yields obtained. Moreover, the state-of-the-art methods will be presented, as well as an overview of *E. coli* importance for this kind of works, and the central metabolism of this organism. This chapter ends with a detailed description of the three main routes to produce AA using 3-HP as an intermediary.

2.1 SYSTEMS BIOLOGY

Biology has entered a new era in the last decades since the first genome was fully sequenced in 1995 [10]. That and a significant improvement in technology have led this field further by combining computational tools, theoretical approaches, and experimental data. Furthermore, the widespread use of this computational biology has aroused the arrival of new study fields such as bioinformatics and systems biology [10, 11].

Systems biology is an interdisciplinary field that relies on computational tools and the theoretical study of biology to identify, deduce, and model relationships between genes, proteins, metabolites, reaction networks, and the cell itself. Thus, the aim of systems biology is not to concentrate on the genetic information and metabolic components themselves, but instead on the nature of the interactions between them [10, 12].

The process of studying those relationships comprises four critical steps. The first step involves the listing of the components of the process that will be studied and the interactions between them. Secondly, those individual components are conjugated into a network and then described as a mathematical model. Thirdly, the resulting models are used to perform *in silico* phenotype predictions that, in the final step, are used to evaluate, describe, and predict the biological functions of that system. Hence, this effort should culminate in a detailed description of the chemical reactions present in the system, and their underlying functions [10].

Since it is an area that generates a considerable amount of data that needs to be interpreted, curated, and then stored, systems biology is heavily dependent on the available computational tools. With this in mind, in the next sections, essential tools used in systems biology will be briefly introduced. Namely, online databases, that are used to store genetic, metabolic and enzymatic information, and metabolic models, on which *in silico* phenotype predictions are performed.

2.1.1 *Online Databases*

Online databases are one of the most relevant tools in systems biology. They allow researchers to access a massive amount of data on metabolites, reactions, enzymes, and also to store a considerable set of metabolic models, which are also very important in this field. With this in mind, all the online databases used to complete this work are briefly described in this section (Table 1).

Table 1.: Databases used in this work, with their respective access link and literature references.

Database	Link	References
BioModels	https://www.ebi.ac.uk/biomodels-main/	[13, 14, 15]
Kyoto Encyclopedia of Genes and Genomes (KEGG)	https://www.genome.jp/kegg/	[16, 17, 18]
MetaCyc	https://metacyc.org/	[19, 20]
Braunschweig Enzyme Database (BRENDA)	https://www.brenda-enzymes.org/	[21, 22]
SABIO-RK	http://sabio.h-its.org/	[23, 24, 25]
eQuilibrator	https://equilibrator.weizmann.ac.il/	[26, 27]

2.1.1.1 *BioModels*

BioModels (Table 1) is an online database that hosts an extensive collection of mathematical models of biological processes. This repository comprises two main branches, one with models retrieved from the publications, and another where the models were generated automatically from pathway data resources [15]. A report from 2015 estimated that this database comprised 1,200 models withdrawn from publication, and over 140,000 models from automated methods. Moreover, BioModels also possesses a comprehensive set of curated models whose reproducibility and accuracy for the replicated biological process was already verified [13, 14, 15].

2.1.1.2 *KEGG*

Kyoto Encyclopedia of Genes and Genomes (KEGG) (Table 1) is a well-known bioinformatics tool that provides a high-level and genomic understanding of the functions and utilities of several organisms. KEGG can be a powerful tool to search for data related to genomics and metabolism, such as enzymes, genes, reactions, and metabolites, which are usually retrieved from molecular-level data [16, 17, 18].

2.1.1.3 *MetaCyc*

MetaCyc (Table 1) integrates the BioCyc Database Collection [19] and is currently a reference database for systems biology. Its 2570 pathways derived from over 54,000 publications crowns this database as the most extensive collection of curated data for metabolic pathways. Moreover, this database contains detailed information on pathways and enzymes that are retrieved from experimental data, and thus providing meaningful knowledge on metabolic pathways, reactions, enzymes, and chemical compounds [19, 20].

2.1.1.4 *BRENDA*

Braunschweig Enzyme Database (BRENDA) (Table 1) is a free online database that comprises data on enzymes and enzymes-ligand systems. All the information that BRENDA contains is retrieved from four sources: literature, text mining procedures, prediction algorithms, and integration of external data. There are over 4.3 million data entries for approximately 84,000 enzymes manually curated directly from over 140,000 literature references. Furthermore, text mining algorithms provide information on kinetic data, occurrence, and enzyme-disease relationships. Furthermore, prediction algorithms provide data for several fields, like genome and locations, and external connections with other databases complete the enzyme information with functional and structural data. Hence, this online resource is a key database in enzyme and enzyme-ligand information in over 30 years [21, 22].

2.1.1.5 *SABIO-RK*

SABIO-RK (Table 1) is an online resource that stores a large amount of information on biochemical reactions and their kinetic properties. All information comprised in SABIO-RK is either manually extracted from the literature or directly learned from laboratory experiments. Additionally, data is consistently checked by automated processes and manually curated [23, 24, 25].

2.1.1.6 *eQuilibrator*

eQuilibrator is an online tool that analysis biochemical reactions and estimates its thermodynamic parameters, like Gibbs free energy and equilibrium constant (K_{eq}), at any given ionic strength, pH and reactant concentration [26, 27]. In order to do that, this tool relies on a broad and comprehensive database that gathers meticulous data on thermodynamic properties of compounds and reactions. That enables eQuilibrator to estimate the necessary energy to produce a compound by an approximation called group contribution, which in turn allows for the thermodynamic analysis of a given system [28, 29, 27].

2.1.2 *Metabolic Models*

Metabolic models are an attempt to simplify and replicate the cellular metabolism [30, 31]. Cellular metabolism comprises a highly complex network of hundreds of metabolites and reactions whose regulatory mechanisms are not yet fully comprehended. Hence, at least to date, the representations that are used do not comprise the full mechanics of cellular behavior and are merely a tool that simplifies the process, thus helping in phenotype prediction [31]. Furthermore, metabolic models have also become important in biotechnology because without them, identifying and optimizing a particular pathway would require researchers to test a high number of combinations to obtain genetic changes that suited their propose, which would be very difficult [32].

There are two types of metabolic models that are widely used in metabolic engineering, namely, stoichiometric and kinetic models. Stoichiometric models describe a set of biochemical pathways as stoichiometric equations that represent the system and are usually used when there is no kinetic information available for the reactions. Whereas kinetic models also combine reaction kinetics with stoichiometry, thus better representing the dynamics that come into play in a metabolic network, but on the other hand, some reactions lack kinetic data so those models cannot always be used [31, 33].

Both types of models are mainly available in a *Systems Biology Markup Language (SBML)* format [34]. This language is a subtype of the *Extensible Markup Language (XML)* language, and therefore, can be broken down into different elements, such as species, reactions, stoichiometries, rate laws, and parameters that are then used to assemble the model. This specific type of format appeared because the diversity of computational tools led to the existence of different formats that could not be interchanged between tools. SBML solved that problem because it allowed different computational tools to translate their specific files into this universal language, and vice-versa, so that it can be used in other tools. Nowadays it has reached a widespread acceptance and has become the standard data language in systems biology [34, 35].

2.1.2.1 *Kinetic Models*

Kinetic models, also known as dynamic models, are widely used in biotechnology. In addition to predicting which reactions take place, kinetic models also predict to what extent those reactions occur, providing a more detailed description of the metabolic network than stoichiometric models do. That is particularly important to biotechnology because they can be used to predict the effects of genetic engineering, and so, to design new pathways for compound production and to optimize their yields [36]. Despite their importance, there is a significant disadvantage associated with dynamic models. Kinetic data is hard to obtain, and some parameters are specific to the assay conditions, resulting in a lack of information on many reactions, which, in turn, has been responsible for the slower progress of this type of models [33].

The best way to describe these dynamic systems is to build mathematical expressions of the reaction kinetics. Then, those expressions are incorporated in the mass balance equations to describe how the biochemical species fluctuate with time [36]. These equations, also called *ordinary differential equations (ODE)*, should be defined simultaneously with the initial conditions:

$$\frac{dX}{dt} = S \cdot v\,(X;P);\; X(0) = X_0 \tag{1}$$

where X stands for the vector of metabolite concentration, S represents the stoichiometric matrix, $v\ (X;\ P)$ the vector of the reaction rate in function of the metabolite concentration (X) and the kinetic data (P), and X_0 the initial metabolite concentration [36].

As mentioned before, these models use kinetic data, so it is crucial to understand kinetic rate expressions and their parameters. Considering this, the main kinetic laws, namely mass action kinetics, Michaelis-Menten kinetics, and two substrate mechanisms, as well as the effect of an inhibitor in their related equations, will be discussed next.

Mass Action Kinetics

The Law of Mass Action [37] that is usually applied to one-step reactions, states that the reaction rate is proportional to the reagents concentrations. So, considering the following reaction (Equation 2):

$$S_1 \xrightarrow{K_1} S_2 \tag{2}$$

assuming that S_1 and S_2 represent two chemical species and K_1 the reaction rate on which S_1 is converted into S_2. For this example the resulting equations would be:

$$x_1(t) = -K_1 \cdot S_1(t) \tag{3}$$

$$x_2(t) = K_1 \cdot S_1(t) \tag{4}$$

where x_i(t) (i=1,2) represents the variation of the species concentration S_i.

Furthermore, this law can be applied to a network of reactions by simply combining the kinetics of their elementary reactions. This results in a model that predicts the behavior of reactions if a dynamic equilibrium state is met [36, 37, 38, 39].

Michaelis-Menten Kinetics

The Michaelis-Menten expression is a kinetic description of the activity of the enzyme. This activity is mainly tracked by the rate of catalysis (V_o), which translates as the number of moles of product formed per second, and depends on the substrate concentration. As stated by this hypothesis, the rate of catalysis increases with the increase of substrate concentration until it begins to asymptotically approach the maximal rate (V_{max}) [40].

Consider the conversion of a substrate (S) to a product (P) catalyzed by an enzyme (E) demonstrated in the following reaction:

$$E + S \underset{\leftarrow}{\overset{K_S}{\rightarrow}} ES \overset{K_{cat}}{\rightarrow} E + P \tag{5}$$

To summarize the reaction, the enzyme firstly binds to the substrate, forming an enzyme-substrate complex (ES). Then, it can either catalyze the conversion to the product or reverse to the original state [40]. Hence, the kinetic equation of this reaction can be described as:

$$V_0 = V_{max} \cdot \frac{[S]}{K_m + [S]} \tag{6}$$

The Michaelis constant (K_m) has two meanings. Firstly, it represents a measure for the strength of the enzyme-substrate complex, in which a low K_m represents a strong binding of the enzyme to the substrate. Secondly, this constant also represents a measure for the concentration of substrate needed for the catalysis to occur, which is the concentration at which half of the V_{max} is reached. The K_m is given by:

$$K_m = \frac{[E] \cdot [S]}{[ES]} \tag{7}$$

The maximal rate (V_{max}) represents the substrate concentration at which all the catalytic sites of the enzyme are saturated with substrate. This rate is given by the product of the catalytic rate constant K_{cat} and the concentration of active sites ($[E]_T$) (Equation 8).

$$V_{max} = K_{cat} \cdot [E]_T \tag{8}$$

Reversible Michaelis-Menten

Although the Michaelis-Menten kinetic was an essential hallmark in the development of kinetic models, in practice, reactions are for the most part reversible [40, 41]. Considering the previous example of the conversion of a substrate (S) to a product (P) (Equation 5), generally, once the product is obtained, it can re-bind to the enzyme (E), regenerating the substrate as demonstrated in the following example:

$$E + S \underset{K_{-1}}{\overset{K_1}{\rightleftarrows}} ES \underset{K_{-2}}{\overset{K_2}{\rightleftarrows}} E + P \tag{9}$$

where K_i represent the reaction rates towards product formation (i = 1, 2) and towards substrate regeneration (i = -1, -2). Because of this, the original Michaelis-Menten was updated to account for the reversibility of the reactions, resulting in the following equation:

$$V_0 = V_{max} \cdot \frac{([S] - \frac{[P]}{K_{eq}})}{K_{m,S} \cdot (1 + \frac{[P]}{K_{m,P}}) + [S]} \tag{10}$$

where $K_{m,S}$ and the $K_{m,P}$ represent the Michaelis-Menten constants of the substrate and product, respectively and the K_{eq} stands for the equilibrium constant [40, 41].

Two Substrate Mechanisms

Despite the prevalence of reactions involving multiple substrates and products in biological systems, the Michaelis-Menten equation was designed considering only single substrate reactions, and, sometimes, it fails to describe certain reactions. That happens because more complex reactions have more sophisticated mechanisms to describe the enzymatic activity involved [42].

So, considering the two substrates (A and B) and two products (P and Q) case:

$$A + B \underset{\leftarrow}{\overset{E}{\rightarrow}} P + Q \tag{11}$$

the reaction can be driven by distinct mechanisms, the sequential mechanism, and the non-sequential mechanism.

In the sequential mechanism, all the substrates must be bound to the enzyme for the reaction to take place, which can occur in an order or in a random sequence. In an ordered mechanism (Figure 2 A), there is a specific order for the substrates to bind to the enzyme. First A must bind, then B, and only then P and Q are produced. On the other hand, in a random fashion (Figure 2 B), B can bind to the enzyme after A or the other way around [43, 42, 44].

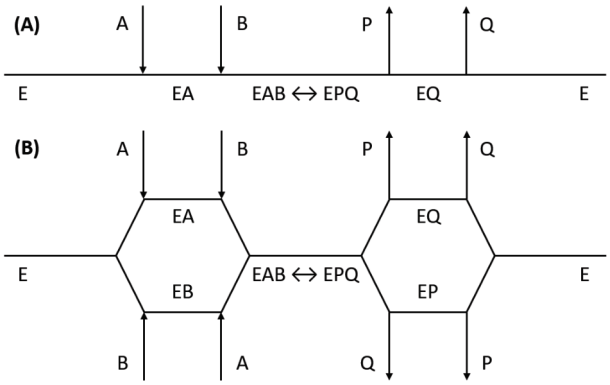


Figure 2.: Representation of the sequential mechanism for multi-substrate reactions. A and B represent the substrates, P and Q the products, and E the enzyme that catalyzes the reaction. (A) In an order sequential mechanism, both substrates must bind to the enzyme for the reaction to occur. In addition, in this case, the order of binding is crucial for the reaction, meaning that B will only bind to the enzyme if A is already bound. (B) On the other hand, the random sequential mechanism does not need a specific order to occur. Adapted from *Michaelis-Menten Kinetics* by Roskoski (2011) [43].

Furthermore, those sequential mechanisms can be described by the following kinetic equation:

$$V_0 = V_{max} \cdot \frac{[A] \cdot [B]}{K_{d,A} \cdot K_{m,b} + K_{m,A} \cdot [B] + K_{m,B} \cdot [A] + [A] \cdot [B]} \tag{12}$$

where $K_{m,A}$ and $K_{m,B}$ are, respectively, the Michaelis-Menten constants for the two substrates of the reaction, A and B, and $K_{d,A}$ is the dissociation constant for A.

Lastly, in the non-sequential mechanism (Figure 3), also known as ping-pong mechanism, there is no need for all the substrates to be bound to the enzyme for the catalysis to take place. That means that A connects to the enzyme and immediately produces P, and only afterward, B will connect and produce Q [43, 42, 44].

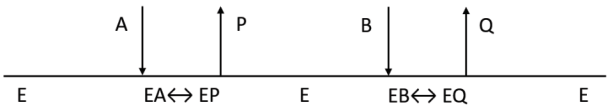


Figure 3.: Representation of a ping-pong mechanism. In this case, a substrate (A) connects to a free enzyme, forming an enzyme-substrate complex (EA). Then A is converted to P, which is then released from the enzyme-product complex (EP). The free enzyme can then bind to the other substrate (B), and form the second product (Q). Adapted from *Michaelis-Menten Kinetics* by Roskoski (2011) [43].

With this in mind, the following kinetic equation can be derived from this specific mechanism:

$$V_0 = V_{max} \cdot \frac{[A] \cdot [B]}{K_{m,A} \cdot [B] + K_{m,B} \cdot [A] + [A] \cdot [B]} \tag{13}$$

where $K_{m,A}$ and $K_{m,B}$ are, respectively, the Michaelis-Menten constants for the two substrates of the reaction, A and B.

Inhibitors

Inhibitors are molecules that block or impair the activity of enzymes by changing their structure or binding to the active center, preventing catalysis to take place. There are four types of reversible inhibition mechanisms, the competitive inhibition, the uncompetitive inhibition, the non-competitive inhibition, and finally, the mixed inhibition [45].

Starting with the competitive mechanism, it describes a situation where the substrate (S) and the inhibitor (I) both bind to the same active site. This way, if the inhibitor binds the enzyme, the substrate cannot, preventing the reaction from taking place, which means that the two molecules are competing for the same active site. Considering this, the parameter that is affected by this type of inhibition is the K_m, and therefore, the kinetic equation for a one substrate one product irreversible reaction resembles the following [45]:

$$V_0 = \frac{V_{max} \cdot [S]}{K_m \left(1 + \frac{I}{K_i}\right) + [S]} \tag{14}$$

In the uncompetitive inhibition, the inhibitor binds the enzyme-substrate complex in the proximity of the binding site, blocking the formation of the product. In this case, in the presence of an inhibitor, the V_{max} and the K_m are both negatively affected, so the equation should be modified to [45]:

$$V_0 = \frac{\frac{V_{max}}{1+\frac{I}{K_i}} \cdot [S]}{\frac{K_m}{1+\frac{I}{K_i}} + [S]} \tag{15}$$

Non-competitive and mixed mechanisms are two similar cases of reversible inhibition. In both cases, the inhibitor binds to the active site of the enzyme or the enzyme-substrate complex. However, the differences between these mechanisms lay in the K_i [45].

In the non-competitive inhibition, the K_i for the enzyme is the same as for the enzyme-substrate complex. This results in a negative impact on the V_{max} value, that is translated to the following modified Michaelis-Menten equation [45]:

$$V_0 = \frac{\frac{V_{max}}{1+\frac{I}{K_i}} \cdot [S]}{K_m + [S]} \tag{16}$$

Finally, in a mixed inhibition, the dissociation constant for the enzyme (K_i) is different from the constant for enzyme-substrate complex (K_i'). This results in the following equation [45]:

$$V_0 = \frac{V_{max} \cdot [S]}{K_m \left(1 + \frac{I}{K_i}\right) + [S] \left(1 + \frac{I}{K_i'}\right)} \tag{17}$$

2.1.2.2 *Escherichia coli Kinetic Models*

E. coli is the main prokaryote used in metabolic engineering. Hence, in this subsection, the kinetic models of *E. coli* assessed for this work will be described.

Chassagnole et al. (2002)

In 2002, Chassagnole and his partners were able to design and validate a kinetic model, comprising the glycolysis and the *penthose-phosphate (PP)* pathway of *E. coli*. This effort resulted in the first dynamic model to connect the transport of sugar to the central carbon metabolism of *E. coli*. This model was also able to replicate experimentally observed dynamics of metabolites, and it has been successfully used in recent works involving dynamic models [46].

Peskov et al. (2012)

Peskov and co-workers aimed to build a dynamic model of *E. coli*'s CCM. To accomplish that, they used assumptions based mainly on experimental data and metabolic and regulatory reconstructions. Furthermore, they also included both *in vitro* and *in vivo* experimental data in the development and verification of the model in order to provide a more detailed description of the metabolism. As a result, a kinetic model was developed to describe *E. coli* aerobic growth in continuous cultures when the concentration of the carbon source is limited [47].

Khodayari and Maranas (2016)

In 2016, Khodayari and Maranas developed a genome-scale kinetic model of *E. coli*. The model, named K-ecoli457, covers 337 metabolites, 457 reactions, and 295 regulatory interactions. Furthermore, this model adequately reproduces fluxomic data for wild-type and 25 mutant strains under distinct growth conditions [48].

Millard et al. (2016)

Millard and colleagues aimed to understand the role of metabolic regulation in *E. coli*. With this goal in mind, they developed a kinetic model of the bacterium CCM. Furthermore, to validate this model, they used 778 independent flux data originated from 266 experiments. Hence, they were able to link cell proliferation and environment to metabolism for the first time. Besides, this work showed that *E. coli* metabolism had far more significant self-regulation capacities, since, without invoking gene expression regulation, kinetic considerations alone were able to explain data obtained in hundreds of studies [49].

Jahan et al. (2016)

Another essential work with kinetic models of *E. coli* was performed by Jahan and coworkers in 2016 [50]. In this study, a model for the central carbon metabolism that included: the glycolytic pathway, *tricarboxylic acid (TCA)* cycle, PP pathway, Entner-Doudoroff pathway, anaplerotic pathways, glyoxylate shunt, oxidative phosphorylation, and the phosphotransferase system was developed. As a result, this model was able to reproduce the dynamics of wild-type *E. coli* and multiple genetic mutants accurately [50].

Matsuoka and Kurata (2017)

In 2017, Matsouka and Kurata developed a model of *E. coli*'s CCM that can simulate the redox regulation of the metabolism under several oxygen concentrations. They used experimental data of a wild-type strain to validate the model, and they came up with a dynamic model that consistently predicts the dynamics of fermentation in *E. coli* [51].

2.2 BIOINFORMATICS TOOLS

In this section, the state-of-the-art of the bioinformatics tools used for model manipulation, simulation, and optimization are presented.

2.2.1 *COPASI*

COmplex PAthway SImulator (COPASI) is an open-source software that has become an essential tool for computational modeling. Despite being able to perform simple analysis with stoichiometric models, COPASI is specialized in the creation, modification, simulation, and optimization of kinetic models [52]. This software is a widely accepted tool, mainly because it has a very user-friendly interface, but it is also controllable via scripting languages in the command line for faster results [52, 53].

2.2.1.1 *Files formats*

COPASI is able to import and export models in SBML [34] format, which enables it to use models from a wide range of online databases, like, for example, BioModels. This tool also stores information in its own XML format (.copasi), and a *Simulation Experiment Description Markup Language (SED-ML)* [54] format to export time course simulations and parameter scans [52, 53].

2.2.1.2 *Creating and modifying a model*

COPASI has a user-friendly interface that allows to create or change a model. Models can be created by adding chemical species that are present in a specific compartment and that enter in a specific reaction. Furthermore, each reaction requires a kinetic law to define the rate at which it occurs, and that law can be chosen from a list of standard functions or manually entered. When entering a new kinetic law, COPASI will automatically change the units based on the expression, making it much more accessible and practical for non-mathematicians. Lastly, COPASI is also able to automatically update the model, meaning that, when a new reaction is added that involves a new species, the species is also added to the model [52].

2.2.1.3 *Simulations*

COPASI can use two different methods to simulate the dynamics of any given model, the deterministic and the stochastic approaches. The deterministic approach, uses the Livermore Solver for Ordinary Differential equations, with an Automatic method switching for stiff and non-stiff problems, and with Root-finding, also known as LSODAR [55]. This integrator is a modified version of the *Livermore Solver of Ordinary Differential Equations (LSODE)* that automatically decides whether a problem can be solved more efficiently. That is achieved through a combination of stiff and non-stiff methods, where the program uses resulting information at the end of each step to automatically decide which one to apply for a more efficient result [55]. The second method uses stochastic formalisms to determine the system solution [56]. To do that, COPASI uses different algorithms, like Gillespie's Direct Method [57], Gibson–Bruck [58], τ-Leap, or adaptive SSA/τ-leap [59]. Moreover, it can also separate reversible reactions into their forward and backward directions to aid its user in adapting deterministic rate equations into their stochastic equivalent, and perform stochastic corrections to rate equations [52, 53].

Furthermore, COPASI also incorporates algorithms that use a combination of both deterministic and stochastic approaches in a more time-efficient manner. They are called hybrid approaches, and COPASI possesses three of them, the Hybrid RungeKutta, the Hybrid LSODA, and the Hybrid RK-45. The first two determine the particle number for each

reaction, and if it is below a threshold, it employs a deterministic approach, and above a stochastic. In turn, the Hybrid RK-45 allows the user to decide which approach to use for each reaction [52].

2.2.1.4 *Metabolic Control Analysis*

A significant advantage of kinetic models lies in predicting the outcome of genetic alterations, such as gene under and over-expression, and gene deletion. One way to perform that is to implement a local parameter sensitivity analysis, to determine to what extent an alteration in a property of the model, like a concentration or a flux, can influence the outcome of the simulation. Although this analysis only considers a local point in the much complex system that is a microorganism, it is undeniable that this tool has some predictive capacities that can be useful to identify targets for metabolic engineering that are suitable for the desired purpose [36].

One important type of sensitivity analysis is the *metabolic control analysis (MCA)* [60, 61, 62]. This tool evaluates how the control of the steady-state fluxes is scattered among the reactions of the system. To do so, the MCA calculates three different metrics, the elasticity coefficients, the *flux control coefficients (FCC)*, and the concentration control coefficients. Of those three, the most important metric for this work is the FCC. This metric evaluates a chosen reaction and returns a value, for each remaining reactions of the system, that reflects how they are impacting its flux. Furthermore, this value can be positive, if one reaction is limiting the flux that arrives a downstream reaction, thus creating a bottleneck, or negative, if it deflects the flux towards other pathways. Either way, even if those results cannot be taken as indisputable facts, this value provides a guideline to identify potential targets for genetic manipulation [36].

2.2.1.5 *Optimization*

Lastly, COPASI is also able to perform optimization tasks to minimize or maximize an objective function defined by the user. Furthermore, any parameter of a reaction, or even from multiple reactions, can be minimized or maximized so that the solution meets the objective function, as well as to what extent they should be modified. To do so, COPASI is equipped with a wide range of local (Hooke & Jeeves [63]; Levenberg–Marquardt [64]; Nelder–Mead [65]; praxis [66]) and global optimization methods (differential evolution [67]; evolutionary strategy [68]; evolutionary programming [69]; genetic algorithm [70]; particle swarm [71], scatter search [72]; random search, simulated annealing [73]) to find the best values toward the set goal.

2.3 *Escherichia coli*

Prokaryotic cells are excellent experimental models for the study of essential aspects of biochemistry and molecular biology because of their simplicity. Additionally, from a vast range of bacterial species that have been studied, *E. coli* is the most comprehended one. Hence, since the beginning of molecular biology, there is a deep understanding of this bacterium's genome and metabolism [74].

E. coli is a gram-negative, fast-growing, and non-sporulating bacillus that was first discovered in 1884 by a German microbiologist named Theodor Escherich. Each cell is about 1 μm long and 0.35 μm wide and, depending on the strain, it may or may not possess flagella to move around the environment, or a pilli to attach to other cells [74, 75, 76]. Furthermore, *E. coli* is a facultative aerobe, meaning that it can grow either in the presence or absence of oxygen, and sensible to extreme temperatures and pH, with an optimal temperature of 37 °C and optimal pH of 7 [77, 78].

In addition, *E. coli* is commonly found in the gastrointestinal tract of most mammals, more specifically in a thin layer of mucus that limits the gut, and is mostly a commensal inhabitant, although some strains are described to cause bloodstream or urinary tract infections. This occasional pathogenicity is due to the presence of, among others, adhesins, toxins, polysaccharide coats, and invasins in the virulent strains [79, 80]. Finally, despite being frequently found inside a host, these bacteria are also able to adapt to the severe conditions of life outside the host and survive in external environments [81].

Nowadays, *E. coli* is the most important model organism that is being used in molecular biology, as it has been used to study even the most basic aspects of life and will most likely continue to be [82]. It has also become essential in the pharmaceutical industry and biotechnology. Historically this happened because this bacteria is very easy to find, as they can be found in every human, easy to grow in a culture medium, and easy to manipulate [75]. Hence, the entire genome of *E. coli* has already been sequenced, consisting of approximately 4.6 million base pairs encoding 4000 different proteins. This bacterium genome is a thousand times smaller than the human genome (3,000 million base pairs), thus bein more straightforward to study [74].

2.3.1 *Importance in Metabolic Engineering*

E. coli is by no means perfect to work with, as it possesses some disadvantages that can represent a threat to culture growth at an industrial scale [83]. Firstly, they do not have the ability to produce every product, such as glycosylated proteins, or proteins with a high number of disulfide bonds [84]. Furthermore, despite being able to grow in a wide range

of conditions, *E. coli* can not survive extreme conditions of pH and temperature. Hence, cultures will be more exposed to microbial contamination and phage attacks [85, 86, 87, 88].

Despite this, *E. coli* is still the preferred prokaryote for both scientific and industrial applications. The increase of knowledge regarding its physiology and genetics allowed researchers to rapidly overcome the organism's limitations and to adapt different strains to surpass the wild type ones [83].

In addition to being extremely easy to grow and maintain in a culture medium, several techniques of genetic manipulation have been developed and perfected for developing new mutants, overcoming those disadvantages and making *E. coli* more profitable to use at the industrial scale [84]. Hence, this organism is still among the top choices for metabolic engineering, has there are currently 484 fully sequenced *E. coli* strains and some of them are the hosts for the industrial production of several chemicals, such as tryptophan, phenylalanine, and lysine [83, 89].

2.3.2 *Central Carbon Metabolism*

The CCM in *E. coli* comprises three main metabolic pathways, namely Glycolysis, TCA cycle, and the PP pathway. These pathways allow the production of the energy required to survive and reproduce. The CCM will be discussed and its importance highlighted in the next section.

2.3.2.1 *Glycolysis*

Glycolysis (Figure 4) is the central pathway for glucose degradation, and in the case of several organisms, the main energy-producing pathway. In this metabolic route one molecule of glucose, a six-carbon molecule, is converted into pyruvate (three-carbon molecule), with an energetic balance of two *adenosine triphosphate (ATP)* and two reduced nicotinamide adenine dinucleotide (NADH) molecules [90].

Glycolysis can be divided into two phases, the preparatory phase and the pay-off phase. The first one consists of the five first reactions of the pathway (Figure 4). In steps one to three, glucose is converted in fructose- 1,6-bisphosphate (F-1,6-P), by phosphorylation and isomerization. Then, this molecule will be cleaved in two different molecules. Two molecules of glyceraldehyde-3-phosphate (Ga3P) will be generated and two molecules of ATP will be invested (Figure 4) [90].

The pay-off phase, consists of the five final steps in the pathway (Figure 4) and culminates with the production of two pyruvate molecules, as well as four ATP and two *nicotinamide adenine dinucleotide (NAD)* molecules, thus providing a final balance of two ATP and two NADH molecules [90].

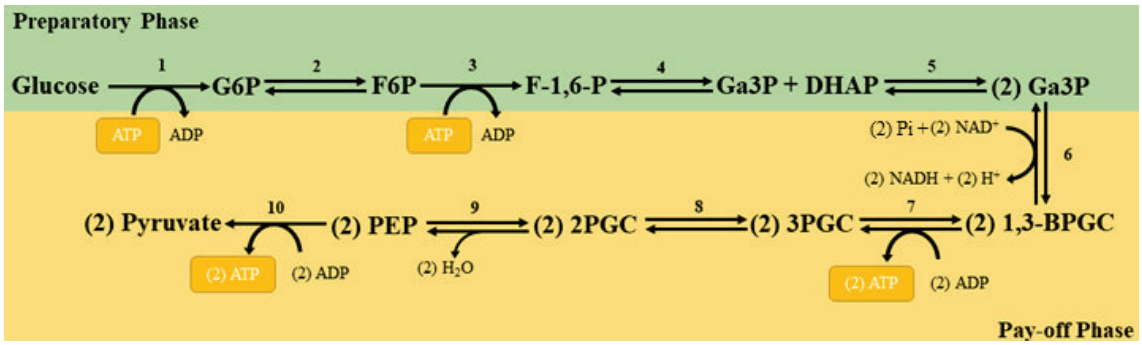


Figure 4.: Schematic representation of the glycolysis pathway. In the preparatory phase, glucose is initially phosphorylated to form glucose-6-phosphate (G6P), converted into fructose-6-phosphate (F6P), and then into fructose-1,6-bisphosphate (F-1,6-P) by a second phosphorylation (Reaction 1-3). Subsequently, the six-carbon molecule is cleaved into glyceraldehyde-3-phosphate (Ga3P) and dihydroxyacetone phosphate (DHAP) (three-carbon molecules) (Reaction 4). The preparatory phase is then completed when the DHAP is converted into another molecule of Ga3P (Reaction 5). The pay-off phase consists of the final five reactions of the pathway. Firstly, Ga3P is put through three consecutive reactions forming sequentially, 1,3-bisphosphoglycerate (1,3-BPGC), 3-phosphoglycerate (3PGC), and 2-phosphoglycerate (2PGC) (Reactions 6-8). Then 2PGC is converted into phosphoenolpyruvate (PEP), which is then converted to pyruvate in the final reaction (Reaction 9 and 10). In this metabolic route, each molecule of glucose is turned into two molecules of pyruvate, two adenosine triphosphate (ATP) and two reduced nicotinamide adenine dinucleotide (NADH). Adapted from *Lehninger Principles of Biochemistry* by Nelson and Cox (2005) [90].

2.3.2.2 *Tricarboxylic Acid cycle*

For most aerobic organisms, glycolysis is not the only step in the breakdown of glucose. The subsequent step is called TCA cycle and is responsible for harnessing most energy these organisms can get from simple sugars. This pathway is divided into two phases. The first phase includes the glycolysis, with an additional reaction were pyruvate is coupled to a *coenzyme A (CoA)* molecule to form acetyl-Coenzyme A (acetyl-CoA) (Figure 5), which also generates one NADH and one CO_2 molecule [90].

The second is the TCA cycle per se, and it includes eight reactions (Figure 5) starting with the fusion of acetyl-CoA and oxaloacetate to produce citric acid. After seven reactions, the pathway culminates in the regeneration of oxaloacetate, thus restarting the cycle. Furthermore, from the two pyruvate molecules generated from the glycolysis, the TCA cycle can generate eighth molecules of NADH, two of reduced flavin adenine dinucleotide (FADH2), and two *guanosine triphosphate (GTP)* [90].

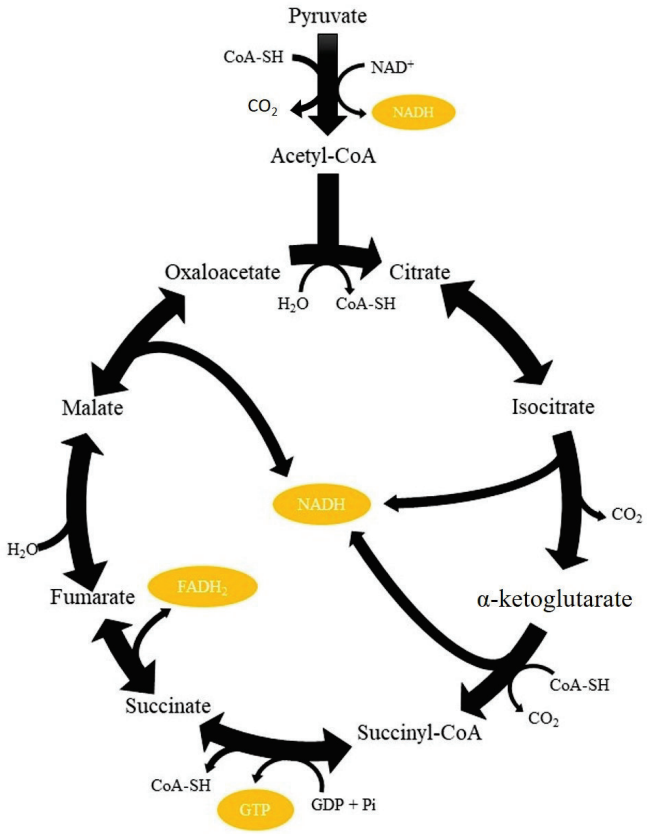


Figure 5.: Schematic representation of the Tricarboxylic Acid Cycle (TCA). In this metabolic route, pyruvate is converted into acetyl coenzyme A (acetyl-CoA) and reduced nicotinamide adenine dinucleotide (NADH). The acetyl-CoA will then be assimilated in the cycle. Through a series of reactions, the pathway will produce three more molecules of NADH, one of reduced flavin adenine dinucleotide (FADH2) and one guanosine triphosphate (GTP). Adapted from *Lehninger Principles of Biochemistry* by Nelson and Cox (2005) [90].

2.3.2.3 *Penthose-phosphate pathway*

The PP pathway is an alternative route for glucose metabolism that is divided into oxidative and non-oxidative phases (Figure 6). The oxidative phase comprises three reactions that deflect glucose-6-phosphate (G6P) from glycolysis and ultimately culminate in the production of ribulose-5-phosphate (R5P) and two molecules of reduced nicotinamide adenine dinucleotide phosphate (NADPH) (Figure 6). This intermediate is vital as, besides being a precursor to the non-oxidative phase, it is also an intermediate for the synthesis of nucleotides, coenzymes, DNA, and RNA. Finally, the non-oxidative phase consists of a couple of steps that ensure the regeneration of glucose-6-phosphate (Figure 6) [90].

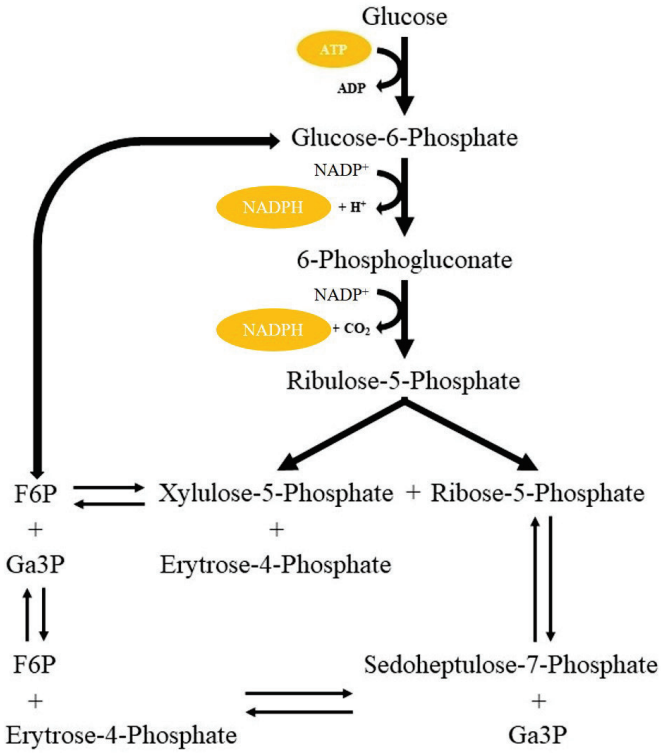


Figure 6.: Schematic representation of the Penthose-phosphate pathway. In the oxidative phase, glucose-6-phosphate is converted into ribulose-5-phosphate, in a process that releases two reduced nicotinamide adenine dinucleotide phosphate (NADPH). On the other hand, in the non-oxidative phase, ribulose-5-phosphate is converted into fructose-6-phosphate (F6P) and glyceraldehyde-3-phosphate (Ga3P), which in turn are responsible for the regeneration of glucose-6-phosphate. Adapted from *Lehninger Principles of Biochemistry* by Nelson and Cox (2005) [90].

2.4 PATHWAYS FOR ACRYLIC ACID PRODUCTION

As in semi-biological methods, the bio-based direct AA route can be divided into two major steps. The first one consists in the conversion of the carbon source into the primary intermediate, 3-HP, which may include different intermediary products. The second concerns the final transformation to AA, for which there is only one known biological pathway able to perform this conversion [4, 1, 91].

When considering the first step to obtain 3-HP, the three main studied pathways are the glycerol, the malonyl-CoA, and β-alanine routes. Since glucose can be converted into glycerol, and glycerol can be converted to an intermediate of the glycolysis, all of them can be used to convert both carbon sources into 3-HP [91, 92, 93].

The distinct bio-based direct AA routes that are suited for this work are described next. This section will be divided into two subsections, where the production of 3-HP will be firstly discussed, and then the AA production.

2.4.1 *3-Hydroxypropionic Acid Production*

The 3-HP production can be carried out using three different intermediates, glycerol, malonyl-CoA and β-alanine (Figure 7). As *E. coli* is able to produce these three metabolites naturally, the focus of this subsection will be the required steps to produce 3-HP.

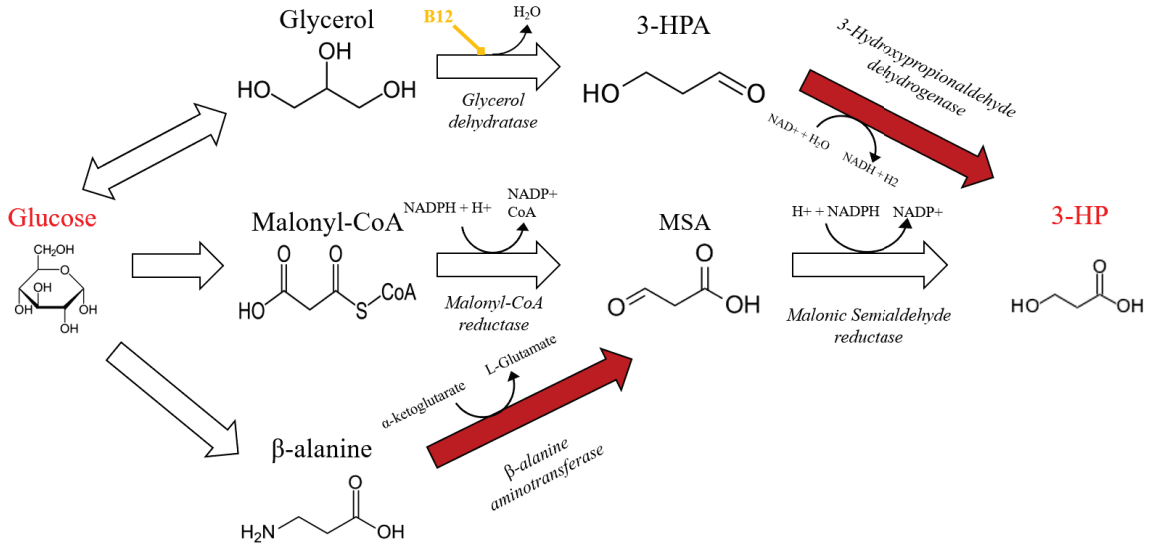


Figure 7.: Pathways to convert glucose to 3-hydroxypropionate (3-HP). 3-HP can be produced from glucose through three distinct intermediates that are naturally produced in *Escherichia coli*: glycerol, malonyl-CoA, and β-Alanine. **Glycerol route:** in the first reaction, vitamin B12 activates the conversion of glycerol to 3-hydroxypropionaldehyde (3-HPA) by the glycerol dehydratase, which is then converted to 3-HP by the aldehyde-dehydrogenase enzyme. **Malonyl-CoA route:** in the first reaction, catalyzed by the malonyl-CoA reductase, the Coenzyme A (CoA) molecule is removed from malonyl-CoA, and at the same time, reduced nicotinamide adenine dinucleotide phosphate (NADPH) is oxidized to nicotinamide adenine dinucleotide phosphate (NADP+), which in turn results in the formation of malonic semialdehyde (MSA). In the final reaction, catalyzed by the malonic semialdehyde reductase, MSA is converted to 3-HP alongside with the oxidation of NADPH. **β-Alanine route:** firstly, β-Alanine and α-ketoglutarate are converted to malonic semialdehyde and L-glutamate by the β-alanine aminotransferase. Then the malonic semialdehyde reductase converts MSA to 3-HP, and also NADPH to NADP+. Furthermore, the arrows highlighted in red represent the reactions of the pathways that are native to *E.coli*. Adapted from Chu *et al.* (2015) [4], Liu *et al.* (2016) [94] and Borodina *et al.* (2015) [95].

2.4.1.1 *Glycerol pathway*

Besides being one of the cheapest carbon sources in the market, glycerol can also be produced from glucose. For this conversion to take place, glucose needs to be directed towards the glycolysis pathway, forming dihydroxyacetone phosphate, an intermediary that is then converted to glycerol [4, 1, 91, 46].

Regarding this route, only two reactions need to take place to obtain 3-HP. The first one consists in the dehydration of glycerol into *3-hydroxypropionaldehyde (3-HPA)* by an enzyme called glycerol-dehydratase. The last enzyme is the 3-hydroxypropionaldehyde dehydrogenase, an it uses 3-HPA, NAD, and H_2O as substrates to produce 3-HP, NADH and H_2 (Figure 7) [4, 1].

Despite being very promising and having good yields associated, this pathway features a significant setback [96]. The reaction responsible for the 3-HPA formation requires vitamin B12 (Figure 7) to take place. This molecule does not interact directly in the reaction, but its presence is required to activate the catalytic activity of the enzyme [93, 97]. That represents a considerable disadvantage to this route since *E. coli* cannot naturally biosynthesize B12, and therefore, the vitamin needs to be supplied to the culture medium, which is an expensive practice that is not at all desired when the main goal is to maximize the profits from AA production [98, 99].

2.4.1.2 *Malonyl-Coenzyme A pathway*

Malonyl-CoA is directly obtained from acetyl-CoA, an intermediate of the TCA cycle that is produced using both glucose and glycerol as the carbon source [91, 97].

Two consecutive reactions are required to produce 3-HP from this intermediary. In the first one, the malonyl-CoA reductase performs the reduction of malonyl-CoA to *malonic semialdehyde (MSA)* coupled with the oxidation of NADPH to *nicotinamide adenine dinucleotide phosphate (NADP)*, while at the same time removing the CoA cofactor from malonyl-CoA. Similarly, in the second reaction, performed by the malonic semialdehyde reductase, MSA is reduced to 3-HP, and at the same time, NADHP is oxidized to NADP (Figure 7) [91, 100, 97].

2.4.1.3 *β-Alanine pathway*

β-alanine can be obtained from oxaloacetate, which is an intermediary of the TCA cycle. Oxaloacetate is converted into L-aspartate, which can then be turned into β-alanine. Hence, as in the remaining routes, it is possible to use this pathway when using the two carbon sources [95, 91, 101, 102].

When β-alanine is used as substrate, two reactions are required to produce 3-HP (Figure 7). The first reaction is catalyzed by β-alanine aminotransferase, in which β-alanine and α-ketoglutarate are used to obtain MSA and L-glutamate. The second one consists in the reduction of MSA to 3-HP which is catalyzed by malonic semialdehyde redutase, a reaction that also belongs to the malonyl-CoA route (Figure 7) [95, 91, 101, 102].

2.4.2 *Acrylic Acid Production*

According to Tong *et al.* (2016) and Chu *et al.* (2015), once 3-HP is obtained, three steps are missing to complete the pathway for AA production. These steps consists in three consecutive reactions that involve three distinct enzymes (Figure 8) [4, 1].

First, 3-hydroxypropionyl-CoA synthase uses ATP to bind a molecule of CoA to 3-HP, resulting into *3-hydroxypropionyl-CoA (3-HP-CoA)*. Then, the 3-hydroxypropionyl-CoA dehydratase turns 3-HP-CoA into *acrylyl-CoA (AA-CoA)*, releasing a water molecule (H_2O) in the process (Figure 8) [4, 91]. In the final step of this pathway, the acrylyl-CoA thioesterase uses H_2O to remove the CoA molecule from AA-CoA, releasing it along with a proton (H^+), producing AA (Figure 8) [4, 91].

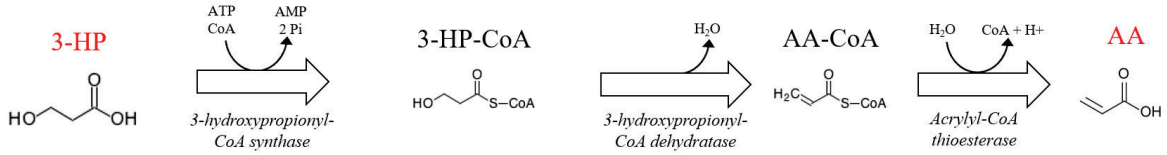


Figure 8.: Pathway for the conversion of 3-hydroxypropionate (3-HP) to acrylic acid (AA). First, the 3-hydroxypropionyl-CoA synthase joins 3-HP with a Coenzyme A (CoA) molecule forming 3-hydroxypropionyl-CoA (3-HP-CoA), while, at the same time, Adenosine Triphosphate (ATP) is dephosphorylated to Adenosine Monophosphate (AMP) and two molecules of P_i. Then, by the action of a 3-hydroxypropionyl-CoA dehydratase, the 3-HP-CoA is converted to acrylyl-CoA (AA-CoA) and H_2O. Finally, AA-CoA is then transformed into AA, with the release of a Coenzyme A molecule (CoA) and a proton (H+) by the acrylyl-CoA thioesterase. Adapted from Chu *et al.* (2015) [4] and Zhijie Liu *et al.* (2016) [7].

2.5 Literature Review of 3-HP and AA Producing Strains

Several studies conducted in recombinant *E. coli* to test and improve 3-HP and AA production from batch cultures, using either glucose or glycerol, were found. As shown in Table 2, 3-HP production through the glycerol pathway is associated with higher yields, reaching up to 8.10 g/L of 3-HP when using glycerol as carbon source [98]. Furthermore, the malonyl-CoA route has also been tested as a viable alternative, and in the work of Liu *et al.* (2016) it reached a production of 3.60 g/L of 3-HP [94] (Table 2). The β-alanine pathway is by far the less studied route in *E. coli* and is associated with the lower reported yields (0.09 g/L of 3-HP) (Table 2).

Concerning AA production, it was not possible to gather as much information as in 3-HP producing routes. However, as shown in Table 3, the reports that tested the full

fermentation pathway from the glycerol and malonyl-CoA routes presented low yields [4, 7, 1].

Table 2.: Literature review on 3-hydroxypropionate (3-HP) production in metabolical engineered *Escherichia coli*.

Paper	Pathway	Carbon Source	Intial Carbon Conc. (g/L)	Titer (g/L)	Reference
Raj *et al.* (2009)	Glycerol	Glycerol	9.20	2.80	[103]
Rathnasingh *et al.* (2009)	Glycerol	Glycerol	18.40	4.40	[104]
Chu *et al.* (2015)	Glycerol	Glycerol	40.00	8.10	[4]
		Glucose	21.50	3.90	[4]
Cheng *et al.* (2015)	Malonyl-CoA	Glucose	10.00	1.80	[105]
Liu *et al.* (2016)	Malonyl-CoA	Glucose	20.00	3.60	[94]
Song *et al.* (2016)	β-alanine	Glucose	15.00	0.09	[93]

Table 3.: Literature review on acrylic acid (AA) production in metabolical engineered *Escherichia coli*.

Paper	Pathway	Carbon Source	Intial Carbon Conc. (g/L)	Titer (g/L)	Reference
Tong *et al.* (2016)	Glycerol	Glycerol	20.00	0.0377	[1]
Chu *et al.* (2015)	Glycerol	Glucose	a	0.12	[4]
Liu and Liu (2016)	Malonyl-CoA	Glucose	20.00	0.013	[7]

a - Information not available.

With this in mind, the studied pathways for AA production (via β-alanine, malonyl-CoA and glycerol) were compared through *in silico* modelling using COPASI. To do so, the seven models described here, namely, the Chassagnole [46], the Peskov [47], the Khodayari [48], the Millard [49], the Jaham [50], and the Matsuoka [51] models, were evaluated to select the one that better complied with the requirements of this work. Then, the three pathways were inserted separately in the chosen model. Furthermore, both glycerol and glucose were tested as the carbon source to assess which of them is associated with higher AA yields. Finally, once AA production was obtained, optimization strategies were identified to enhance the obtained yields. This work allowed the identification of the AA route that has the most potential for higher yields so that it can be implemented *in vivo* for industrial-scale production.

3

MATERIALS AND METHODS

This chapter will focus on the detailed description of the materials and methods used for the assembly, simulation, and optimization of the developed dynamic models.

3.1 MODEL SELECTION

From the seven previously described models, the Millard *et al.* (2016) [49] metabolic model was selected, as it provides a thorough description of the CCM. This model includes the main pathways, namely the glucose phosphotransferase system, glycolysis, gluconeogenesis, pentose phosphate pathway, tricarboxylic acid cycle, glyoxylate shunt, acetate metabolism, anaplerotic reactions, nucleotide interconversion reactions and finally oxidative phosphorylation (Figure 9). In total, this model encompasses 62 metabolites, 68 reactions, and three compartments (extracellular, periplasm, and cytoplasm) [49].

Besides providing a more detailed description of *E. coli* metabolism than previous models, Millard and colleagues were also able to couple the degradation of glucose to oxidative phosphorylation. This fact allowed the inclusion of cofactors like ATP, NAD, NADP and *flavin adenine dinucleotide (FAD)* as part of the metabolism instead of using pseudo-reactions like in previous models [49], which also influenced the decision.

Moreover, this kinetic model is capable of simulating *E. coli* growth in a chemostat under glucose-limited conditions, namely a reactor with 100 liters of culture medium, 0.25 liters of periplasmatic volume, and 1 liter of cytosolic volume, with a glucose feed of 0.23 mM/s. Unfortunately, despite being a detailed model, it does not include the production of glycerol, malonyl-CoA, and β-alanine, which are usually produced by *E.coli*. This means that, the first step to achieve *in silico* production of AA is to extend the CCM to produce those three intermediaries.

The SBML version of the metabolic model is available for download in the BIOMODELS database [15] with the identifier MODEL1505110000.

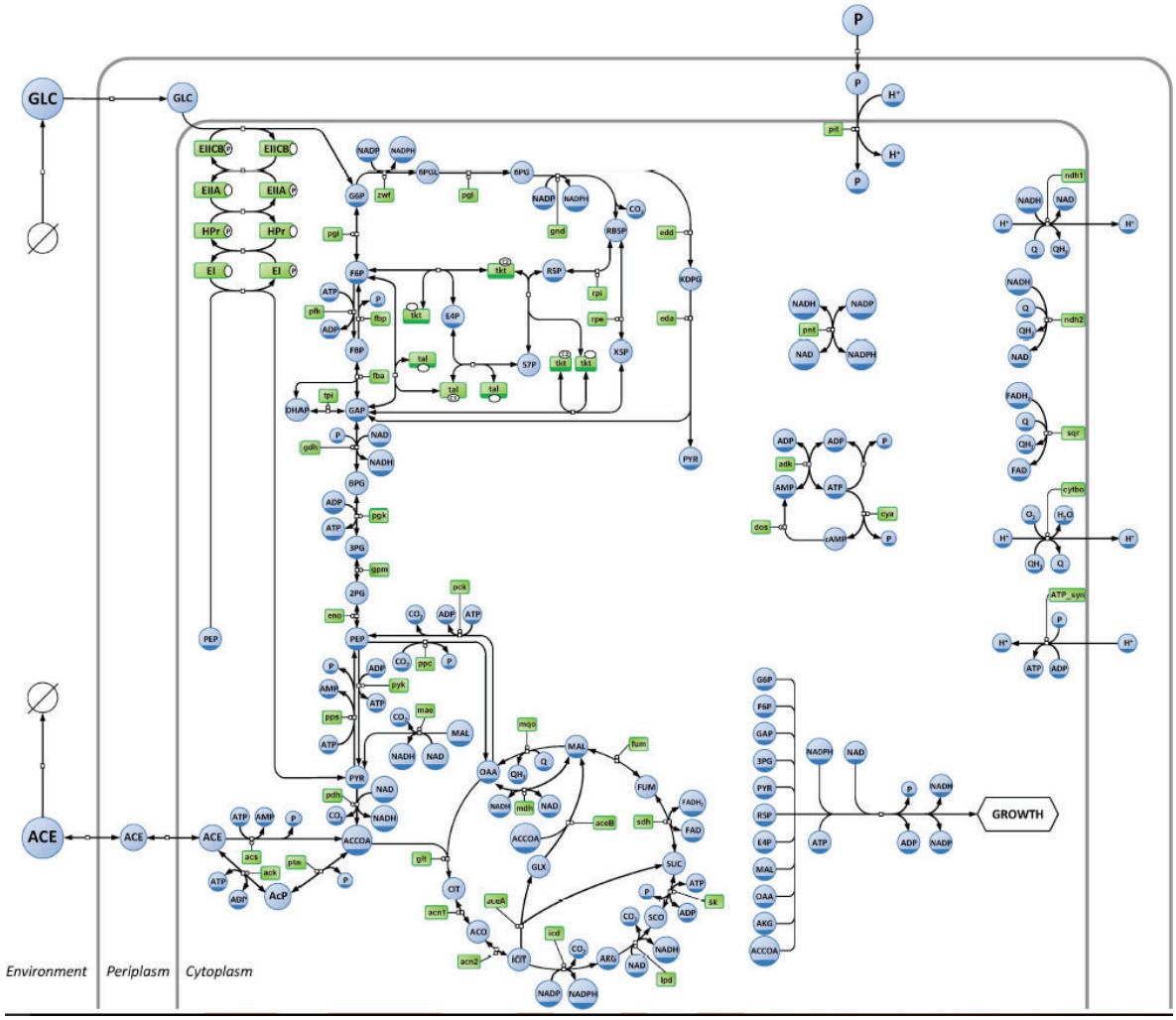


Figure 9.: Representation of the dynamic model of *Escherichia coli* central carbon metabolism developed by Millard *et al.* (2016). This model comprises the glucose phosphotransferase system, glycolysis, gluconeogenesis, pentose phosphate pathway, tricarboxylic acid cycle, glyoxylate shunt, acetate metabolism, and oxidative phosphorylation, with a total of 62 metabolites and 68 reactions. Retrieved from Millard *et al.* (2016) [49].

3.2 KINETIC MODELING

The development of the model was divided into three steps, the extension of the CCM to include the production of the three intermediaries (glycerol, malonyl-CoA and β-alanine), production of 3-HP from those metabolites, and finally, production of AA.

3.2.1 *Parameter Selection*

Kinetic equations and their respective parameters were retrieved from the available literature. Databases like BioCyc [20], BRENDA [22] and Sabio-RK [25] were used to obtain K_m, K_i, K_d, activation constants (K_a), *specific activity (SA)*, and K_{cat} values. Furthermore, the eQuilibrator [26] database was also used to determine equilibrium constants (K_e) required to characterize the reaction reversibility. The last kinetic parameter required to describe each reaction is the V_{max}. Unfortunately, this parameter is highly specific for the concentration of enzyme available, which will in turn depend on the specific conditions of the assay, and is seldom reported in the literature. Therefore, two distinct methods were used to estimate the V_{max}, according to the origin of the reaction (the CCM extension or the heterologous pathways).

Method 1 was employed for reactions that belong to the native metabolism of *E.coli*, the V_{max} estimation was adapted from a method used by Chassagnole and colleagues while developing the kinetic model in 2002 [46]. Initially, a steady-state flux distribution is determined for the original kinetic model. Then, a genome-scale model of *E.coli* K-12 MG1655 (iML1515) [106] is used to predict the flux of the new reactions. For this, the common reactions between the kinetic model and the stoichiometric model are constrained to the previously determined flux distribution ($\pm$ 0.01 mM/s). Then, a flux variability analysis is performed to determine the estimated flux (v). By equalizing v to the rate law of the reaction, the following equation is obtained:

$$v = V_{max} \cdot F(X, K) \Leftrightarrow V_{max} = \frac{v}{F(X, K)} \tag{18}$$

in which X is a vector of parameters and K a vector of steady-state concentrations for the metabolites involved. Furthermore, it is worth noting that for newly added metabolites, the steady-state concentration was assumed to be 1 mM.

Method 2 was used for reactions of the heterologous pathways, the V_{max} was estimated assuming that the total concentration of enzyme was in surplus (100 mM), which allowed calculating this parameter as shown in equation 8.

3.2.2 *Model Extension*

The selected model represents a chemostat system, thus the first step to assemble the AA production model was converting the original model into a batch system. Therefore, the glucose feed and the drain for acetate were deleted. The next step was to extend the model of the CCM to include all three intermediaries and allow the use of glycerol as carbon source. In this subsection, all the V_{max} values were calculated using Method 1, and the remaining parameters were retrieved from articles presenting a more complete characterization of the enzyme's kinetics.

3.2.2.1 *Glycerol, Malonyl-CoA and β-alanine Production*

Regarding the glycerol pathway, two reactions are required to obtain glycerol from *dihydroxyacetone phosphate (DAP)*, namely reactions catalyzed by the glycerol-3-phosphate dehydrogenase (G3pD) and the glycerol-3-phosphate phosphatase (G3pP). Moreover, considering that the latter is not reversible and that this route requires to be reversible to use glycerol as carbon source, the reaction catalyzed by the glycerol kinase (GlyK) was also included. This reaction allows converting glycerol into glycerol-3-phosphate (Figure 10). Only one reaction is required to obtain malonyl-CoA, namely the reaction catalyzed by the acetyl-CoA carboxylase (AccC) (Figure 10). Finally, three reactions are necessary to include the β-alanine route. Two to produce β-alanine, and one to produce L-glutamate, an essential intermediary for β-alanine production also absent from the original model. These reactions are promoted by the aspartate aminotransferase (AspAT), the aspartate carboxylase (AspC), and the L-glutamate dehydrogenase (GluD) (Figure 10) [18, 20].

The reactions and their respective stoichiometry are shown below:

G3pD : DAP + NADPH + H^+ = Glycerol-3-Phosphate + $NADP^+$ (19)

G3pP : Glycerol-3-Phosphate + $H_2O \rightarrow$ Glycerol + P_i (20)

GlyK : Glycerol + ATP $\rightarrow$ Glycerol-3-Phosphate + ADP + H^+ (21)

AccC : Acetyl-CoA + ATP + $HCO_3 \rightarrow$ Malonyl-CoA + ADP + P_i (22)

GluD : α-Ketoglutarate + NADP + H_2O = L-Glutamate + NADPH + NH_3 + H^+ (23)

AspAT : Oxaloacetate + L-Glutamate = L-Aspartate + α-Ketoglutarate (24)

AspC : Aspartate $\rightarrow$ β-alanine + CO_2 (25)

Furthermore, an additional set of pseudo-reactions were included in the model, the Synth reactions. These reactions were inspired by the work of Chassagnole *et al.* (2012) [46] and Machado *et al.* (2014) [107], and are used to represent the pathways involved in the breakdown of the newly added metabolites.

All reactions described next were added to the original model from Millard *et al.* (2016), resulting in the CCMext_GLC model, that depicts the production of all the three intermediaries when glucose is used as the carbon source.

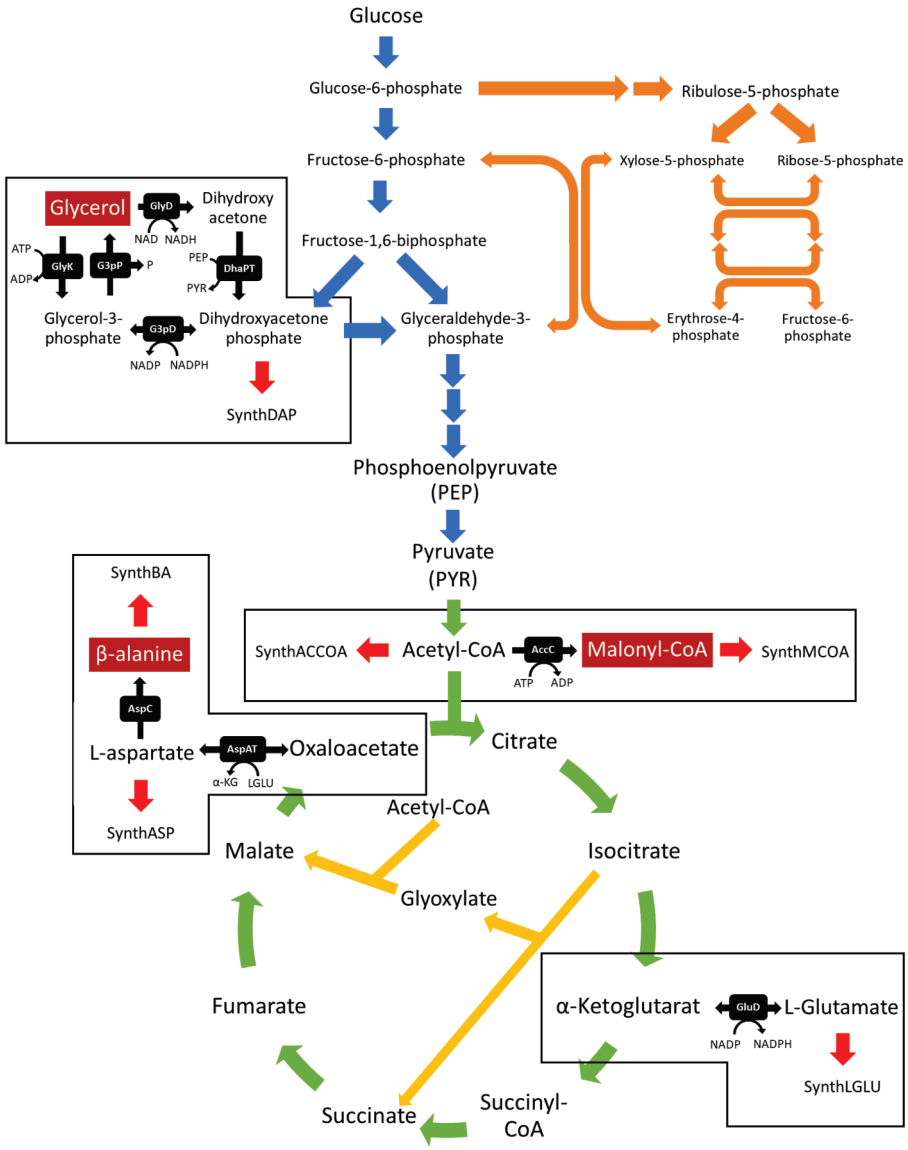


Figure 10.: Representation of the central carbon metabolism of *Escherichia coli* and the reactions added to the kinetic model. The reactions depicted by the blue, orange, green and yellow arrows represent, respectively, the glycolysis, pentose-phosphate pathway, tricarboxylic acid cycle and the glyoxylate shunt, which are all present in the original model. The black arrows represent the seven reactions that were added to obtain glycerol, malonyl-CoA and β-alanine. Finally, the red arrows depict the Synth reactions that were added to account for the deviation of the newly added metabolites to other pathways. Adapted from KeGG [16, 17, 18] and MetaCyc [19, 20].

Glycerol-3-Phosphate Dehydrogenase (G3pD)

This enzyme catalyzes the conversion of DAP to glycerol-3-phosphate (Equation 19). The kinetic properties of G3pD were fully characterized in the work of Edgar and Bell (1978) [108]. As shown in table 4, they reported a K_m of 0.18 mM for DAP ($K_{m,a}$), 0.0034 mM for NADPH ($K_{m,b}$), 0.03 mM for glycerol-3-phosphate ($K_{m,p}$), and finally 0.165 mM for NADP ($K_{m,q}$) and a K_{eq} of 900 for this reaction. They did not accurately identify the kinetic mechanism of the enzyme, but noted that the most probable one was the Rapid Equilibrium Random Bi Bi [108], which was also corroborated by other works for this enzyme [109].

Glycerol-3-Phosphate Phosphatase (G3pP)

The G3pP catalyzes the irreversible transformation of glycerol-3-phosphate to glycerol (Equation 20). This reaction follows a single substrate Michaelis-Menten kinetic. According to literature [110], the respective K_m for glycerol-3-phosphate is equal to 2.9 mM [110], as shown in Table 4.

Glycerol Kinase (GlyK)

This enzyme has been described by Pettigrew *et al.* (1990) [111]. They reported a K_m for glycerol ($K_{m,b}$) of 0.0049 mM and for ATP ($K_{m,a}$) of 0.0084 mM, and a dissociation constant for ATP ($K_{d,a}$) of 0.086 mM. They also reported that the enzyme seemed to have a random bi kinetic mechanism [111].

Acetyl-CoA Carboxylase (AccC)

The kinetic parameters of this enzyme were described by Soriano *et al.* (2005), in which they reported a K_m for acetyl-CoA ($K_{m,a}$) of 0.16 mM and for ATP ($K_{m,b}$) of 0.06 mM [112]. Moreover, a study by Freiberg *et al.* (2004) showed that malonyl-CoA, a product of this reaction, presented competitive inhibition towards acetyl-CoA [113]. The inhibition constant ($K_{i,p}$) they determined for this product was 0.1 mM. Hence, a two substrate order Bi kinetics with substrate inhibition towards acetyl-CoA was adopted.

Glutamate Dehydrogensase (GluD)

This enzyme is responsible for the oxidation of α-ketoglutarate to L-glutamate (Equation 23). This enzyme's kinetic parameters were characterized by Sharkey and Engel (2008). They reported K_m values for α-ketoglutarate ($K_{m,a}$) and for NADP ($K_{m,b}$) of 0.68 mM and 0.018 mM, respectively, and a hill coeficient for NADP of 0.92. Furthermore, they reported that this enzyme exhibited Michaelis-Menten kinetics for α-ketoglutarate and a hill cooperativity kinetics for NADP [114].

Aspartate Aminotransferase (AspAT)

Yagi *et al.* (1985) presented a kinetic description on the AspAT in which K_m values of 15 mM, 0.01 mM, 0.24 mM, and 1.3 mM, were reported for L-glutamate ($K_{m,a}$), oxaloacetate ($K_{m,b}$), aspartate ($K_{m,p}$), and α-ketoglutarate ($K_{m,q}$), respectively [115]. Unfortunately, the underlying mechanism of this enzyme was not specified in this paper, but according to BioCyc, this enzyme is known to follow a reversible ping-pong Bi Bi mechanism [116, 117]. Moreover, to finish the characterization of this enzyme, the eQuilibrator database was accessed to retrieve the K_{eq}, whose value was 3.2.

Aspartate Carboxylase (AspC)

Finally, the last reaction consisted in the conversion of aspartate to β-alanine (Equation 25). Like the G3pP, this enzyme is characterized by a single substrate Michaelis-Menten kinetic, with a K_m of 0.151 mM [118].

Synth Reactions

These reactions were added to account for fluxes of metabolites produced by the new reactions, though not metabolized in the heterologous pathway. The sum of all fluxes from the reactions that metabolize the intermediary metabolites in the stoichiometric model simulation was determined and, assuming mass action kinetics, the value of k was calculated using the same principle as Method 1.

3.2.2.2 *Alternative Glycerol Assimilation Route*

The reaction catalyzed by G3pD has a K_{eq} value of 900, thus requiring excessively high concentrations of glycerol-3-phosphate and NADP to flow in the reverse direction. Hence, glycerol was not assimilated into the CCM, meaning that ATP and other cofactors were not being produced. Since GlyK requires ATP to produce glycerol-3-phosphate, this cofactor was immediately depleted, and the metabolite was not synthesized in a sufficient amount. G3pD did not catalyze the reaction in the reverse direction, thus not allowing to test AA production using glycerol as carbon source.

Therefore, the stoichiometric model was used to identify alternative routes for glycerol intake. Two additional reactions catalyzed by the enzymes glycerol dehydrogenase (GlyD) and the dihydroxyacetone phosphate transferase (DhaPT) were included in model CCMext_GLC (Figure 10), thus creating model CCMext_GLY.

These reactions and their respective stoichiometry are the following:

$$\textbf{GlyD} : \text{Glycerol} + NAD^+ \rightarrow \text{Dihydroxyacetone} + NADH + H^+ \tag{26}$$

$$\textbf{DhaPT} : \text{Dihydroxyacetone} + \text{Phosphoenolpyruvate} \rightarrow \text{DAP} + \text{Pyruvate} \tag{27}$$

Table 4.: Rate Law (RL), equations, and respective parameters of the reactions that belong to the native metabolism of *E.coli*. The following abbreviations were used: RERBB: Rapid Equilibrium Random Bi Bi; MM - Michaelis-Menten; RBB - Random Bi Bi; OBB - Order Bi Bi; PPBB - Ping-Pong Bi Bi.

Reaction	RL	Equation	Parameters
G3pD	RERBB	$V_{max} \cdot \frac{\left(\frac{A \cdot B - \left(\frac{P \cdot Q}{K_{eq}}\right)}{K_{m,a} \cdot K_{m,b}}\right)}{\left(1 + \frac{A}{K_{m,a}} \cdot \frac{B}{K_{m,b}}\right) + \left(1 + \frac{P}{K_{m,p}} \cdot \frac{Q}{K_{m,q}}\right) - 1}$	$K_{m,a}$ = 0.18 mM $K_{m,b}$ = 0.0034 mM $K_{m,p}$ = 0.03 mM $K_{m,q}$ = 0.165 mM K_{eq} = 900
G3pP	MM	$V_{max} \cdot \frac{A}{K_m + A}$	K_m = 2.9 mM
GlyK	RBB	$V_{max} \cdot \frac{A \cdot B}{K_{i,a} \cdot K_{m,b} + K_{m,b} \cdot A + K_{m,a} \cdot B + A \cdot B}$	$K_{m,a}$= 0.0084 mM $K_{m,b}$ = 0.0049 mM $K_{d,a}$ = 0.086 mM
AccC	OBB	$V_{max} \cdot \frac{A}{K_{m,A} \cdot \left(1 + \frac{P}{K_{i,P}}\right) + A} \cdot \frac{B}{K_{m,B} + B}$	$K_{m,a}$ = 0.16 mM $K_{m,b}$ = 0.06 mM $K_{i,p}$ = 0.1 mM
GluD	MM	$V_{max} \cdot \frac{A}{K_{m,a} + A} \cdot \frac{B^n}{(K_{m,b})^n + B^n}$	$K_{m,a}$ = 0.68 mM $K_{m,b}$ = 0.018 mM n = 0.92
AspAT	PPBB	$V_{max} \cdot \frac{\left(\frac{A \cdot B - \left(\frac{P \cdot Q}{K_{eq}}\right)}{K_{m,a} \cdot K_{m,b}}\right)}{\left(1 + \frac{A}{K_{m,a}} \cdot \frac{Q}{K_{m,q}}\right) \cdot \left(1 + \frac{B}{K_{m,b}} \cdot \frac{P}{K_{m,p}}\right)}$	$K_{m,a}$ = 15 mM $K_{m,b}$ = 0.01 mM $K_{m,p}$ = 0.24 mM $K_{m,q}$ = 1.3 mM K_{eq} =3.2
AspC	MM	$V_{max} \cdot \frac{A}{K_m + A}$	K_m = 0.151 mM

Glycerol Dehydrogenase

This reaction was characterized in *E. coli* by Piattoni *et al.* (2013) [119], with K_m values of 76 mM and 0.81 mM, for glycerol ($K_{m,a}$) and NAD ($K_{m,b}$), respectively. Moreover, according to the authors, this enzyme exhibited a behavior that seemed to fit hill cooperativity kinetics, with a hill coefficient (n) of 0.9 (Table 5) [119].

Dihydroxyacetone Phosphate Transferase

This enzyme was not fully characterized for both substrates, as the K_m value for phosphoenolpyruvate was not found in any of the databases. Therefore mass action kinetics were used to represent the dynamics of this reaction, and the k value was calculated using method 1.

Table 5.: Rate Laws, equations, and respective parameters for the alternative glycerol assimilation route. The following abbreviations were used: HC - Hill Cooperativity; MA - Mass Action.

Reaction ID	Rate Law	Equation	Parameters
GlyD	HC	$V_{max} \cdot \frac{A^n}{(K_{m,a})^n + B^n} \cdot \frac{B^n}{(K_{m,b})^n + B^n}$	$K_{m,a}$ = 76 mM $K_{m,b}$ = 0.81 mM n =0.9
DhaPT	MA	$k \cdot A \cdot B$	

3.2.3 *Pathways for acrylic acid production*

Once the model was able to produce glycerol, malonyl-CoA, and β-alanine, the following step was to add the remaining reactions that lead to AA production. With that in mind, this step was divided into two different sub-steps. First, each of the three distinct pathways to produce 3-HP was inserted in the model separately, thus creating distinct models for each carbon source. Then the inclusion of the subsequent reactions for the production of AA. In this subsection, the main factor influencing the strain and parameter choice was the existence of a complete characterization of the underling kinetic mechanism of the respective enzyme. Furthermore, the V_{max} values for the heterologous enzymes were calculated using Method 2. In the end, twelve models were created, six for the synthesis of 3-HP and six for AA (Table A.1).

3.2.3.1 *Glycerol pathway*

Two 3-HP producing models were generated by adding two reactions to the CCMext_-GLC and CCMext_GLY models, creating the 3HP_GlyPath_GLC and the 3HP_GlyPath_-GLY models. As shown in Figure 7, the first reaction is catalyzed by the glycerol dehydratase (GlyDH), that converts glycerol into 3-HPA, and the second reaction is promoted by the 3-hydroxypropionaldehyde dehydrogenase (3hpaD):

$$\textbf{GlyDH} : \text{Glycerol} \rightarrow \text{3-HPA} + H_2O \tag{28}$$

$$\textbf{3hpaD} : \text{3-HPA+NADP}^+ + H_2O \rightarrow \text{3-HP+NADPH+2H}^+ \tag{29}$$

Glycerol Dehydratase (GlyDH)

The kinetic parameters of the enzyme in *Lactobacillus collinoides* were determined by Sauvageot *et al.* (2002). In their work they reported a K_m of 8.3 mM for glycerol, and an activation constant (K_a) of 0.008 mM for vitamin B12. Furthermore, they also reported a specific activity of 0.018 $\mu mol \cdot min^{-1} \cdot mg^{-1}$ when glycerol was the substrate, and a molecular weight of 207 kDa, which allowed the calculation of a K_{cat} value of 0.0621 s^{-1} [120]. Considering the available kinetic parameters, and the fact that B12 is an activator of the reaction and not an intermediary, the specific activation mechanism rate law was assumed (Table 6).

3-hydroxypropionaldehyde Dehydrogenase (3hpaD)

The kinetic properties of the 3hpaD from *E. coli* were studied by Jo *et al.* (2008). This study reported a K_{cat} of 28.54 s^{-1} when 3-HPA was used as a substrate, as well as K_m values of 0.49 mM and 0.06 mM for 3-HPA ($K_{m,a}$) and NADP ($K_{m,b}$), respectively [121]. Considering that no kinetic mechanism was associated with this enzyme, a simple two substrate Michalis-Menten equation was assumed (Table 6). Furthermore, since this reaction had no flux in the stoichiometric model when using the flux constraints of the original model, the V_{max} was calculated using the method for the heterologous reactions (Method 2).

3.2.3.2 *Malonyl-CoA pathway*

The malonyl-CoA route was also included in each carbon source's model, creating models 3-HP_McoaPath_GLC and 3-HP_McoaPath_GLY. These models have two additional reactions, the first catalyzed by the malonyl-CoA reductase (McoaR) and the second catalyzed by malonic semialdehyde reductase (MsaR) (Figure 7):

$$\textbf{McoaR} : \text{Malonyl-CoA + NADPH + H}^+ \rightarrow \text{MSA + CoA + NADP}^+ \tag{30}$$

$$\textbf{MsaR} : \text{MSA + NADPH +H}^+ \rightarrow \text{3-HP + NADP}^+ \tag{31}$$

Table 6.: Rate Laws, equations, and respective parameters for the reactions of the glycerol pathway that culminate with the productions of 3-hydroxypropionate.

Reaction ID	Rate Law	Equation	Parameters
GlyDH	Specific Activation	$\frac{E \cdot K_{cat} \cdot A \cdot Activator}{K_{m,a} \cdot K_a + (K_{m,a} + A) \cdot Activator}$	E = 100 mM K_{cat} = 0.0621 s^{-1} K_m = 8.3 mM K_a = 0.008 mM
3hpaD	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A \cdot B}{K_{m,a} \cdot K_{m,b} + K_{m,b} \cdot A + K_{m,a} \cdot B + A \cdot B}$	E = 100 mM K_{cat} = 28.54 s^{-1} $K_{m,a}$ = 0.49 mM $K_{m,b}$ = 0.06 mM

Malonyl-CoA Reductase (McoaR)

The McoaR catalyzes the conversion of malonyl-CoA to MSA (Equation 30). This enzyme is present in *Chloroflexus aurantiacus* and was described by H*ü*gler *et al.* (2002), which reported a K_{cat} of 50 s^{-1} when malonyl-CoA was used as a substrate, and K_m values of 0.3mM and 0.025 mM for malonyl-CoA ($K_{m,a}$) and NADPH ($K_{m,b}$), respectively. The authors also stated that the enzyme exhibited a two substrate Michaelis-Menten behavior [100] (Table 7).

Malonic Semialdehyde Reductase (MsaR)

The MsaR enzyme, which catalyzes the final reaction of the malonyl-CoA pathway was described by Kockelkorn and Fuchs (2009). The authors characterized this enzyme from *Metallosphaera sedula* and were able to identify the kinetic parameters of the enzyme. The K_{cat} value reported when malonyl-CoA is used as the substrate is 115 s^{-1}, and the K_m values are 0.07 mM for both MSA ($K_{m,a}$) and NADPH ($K_{m,b}$) [122]. Finally, a two substrate Michaelis-Menten mechanism was assumed for this model (Table 7).

Table 7.: Rate Laws, equations, and respective parameters for the reactions of the malonyl-CoA pathway that culminate with the productions of 3-hydroxypropionate.

Reaction ID	Rate Law	Equation	Parameters
McoaR	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A \cdot B}{K_{m,a} \cdot K_{m,b} + K_{m,b} \cdot A + K_{m,a} \cdot B + A \cdot B}$	E = 100 mM K_{cat} = 50 s^{-1} $K_{m,a}$ = 0.3 mM $K_{m,b}$ = 0.025 mM
MsaR	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A \cdot B}{K_{m,a} \cdot K_{m,b} + K_{m,b} \cdot A + K_{m,a} \cdot B + A \cdot B}$	E = 100 mM K_{cat} = 115 s^{-1} $K_{m,a}$ = 0.07 mM $K_{m,b}$ = 0.07 mM

3.2.3.3 *β-Alanine pathway*

Finally, the β-alanine pathway also required adding two additional reactions for the model to produce 3-HP. The first one is the β-alanine aminotransferase (BaAT), and then as the final reaction, the MsaR, a common reaction to the malonyl-CoA pathway (Figure 7):

$$\textbf{BaAT} : \beta\text{-alanine} + \alpha\text{-Ketoglutarate} \rightarrow \text{MSA} + \text{L-Glutamate} \tag{32}$$

$$\textbf{MsaR} : \text{MSA} + \text{NADPH} + \text{H}^{+} \rightarrow \text{3-HP} + \text{NADP}^{+} \tag{33}$$

Two models (3HP_BalaPath_GLC and 3HP_BalaPath_GLY) were created, according to the carbon source.

β-Alanine Aminotransferase (BaAT)

E. coli's BaAT was described by Liu *et al.* (2005). The authors believed that the enzyme followed ping-pong bi bi with competitive substrate inhibition by α-ketoglutarate, and proceeded to estimate the respective parameters. They reported a K_{cat} of 47.4 s^{-1}, and a K_m for β-alanine ($K_{m,a}$) and α-ketoglutarate ($K_{m,b}$), of 5.8 mM and 1.07 mM, respectively. Furthermore, the $K_{i,b}$ for α-ketoglutarate was also estimated, and was equal to 10.2 mM [123] (Table 8). Even though this enzyme was characterized for *E. coli*, it was not present in the stoichiometric model, thus the V_{max} was calculated using the Method 2.

Malonic Semialdehyde Reductase (MsaR)

This enzyme was already characterized for the malonyl-CoA pathway. Thus, the kinetic description used before was adopted for the β-alanine pathway (Table 8).

Table 8.: Rate Laws, equations, and respective parameters for the reactions of the β-alanine pathway that culminate with the productions of 3-hydroxypropionate.

Reaction ID	Rate Law	Equation	Parameters
BaTA	Ping-Pong Bi Bi	$\frac{E \cdot K_{cat} \cdot A \cdot B}{K_{m,b} \cdot A + K_{m,a} \cdot B \cdot \left(1 + \frac{B}{K_{i,B}}\right) + A \cdot B}$	E = 100 mM K_{cat} = 47.4 s^{-1} $K_{m,a}$ = 5.8 mM $K_{m,b}$ = 1.07 mM $K_{i,b}$ = 10.2 mM
MsaR	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A \cdot B}{K_{m,a} \cdot K_{m,b} + K_{m,b} \cdot A + K_{m,a} \cdot B + A \cdot B}$	E = 100 mM K_{cat} = 115 s^{-1} $K_{m,a}$ = 0.07 mM $K_{m,b}$ = 0.07 mM

3.2.3.4 *Acrylic Acid pathway*

When 3-HP is obtained, only three more reactions are required for producing AA, namely:

- **3-Hydroxypropionyl-CoA Synthase (3hpcoaS)**: Alber and Fuchs (2002) studied a variant of the 3hpcoaS from *Chloroflexus aurantiacus*. In their work they stated that the enzyme probably follows a Michaelis-Menten mechanism, and the parameters for that mechanism are a K_{cat} of 36 s^{-1}, and the K_m values for 3-HP ($K_{m,a}$) of 0.015 mM, CoA ($K_{m,b}$) of 0.01 mM, and ATP ($K_{m,c}$) of 0.05 mM [124] (Table 9). The reaction catalyzed by this enzyme is the following:

$$\text{3-HP} + \text{CoA} + \text{ATP} \rightarrow \text{3-HP-CoA} + \text{2P} + \text{AMP} \tag{34}$$

- **3-Hydroxypropionyl-CoA Dehydratase (3hpcoaDH)**: This enzyme is a variant of the 3hpcoaDH from *M. sedula* characterized by Teufel *et al.* (2009), whose work helped to understand the underlying mechanism that controls the enzyme's activity. According to this study, the enzyme follows a Michaelis-Menten kinetics, with a K_{cat} of 96 s^{-1}, and K_m value of 0.06 mM for 3-HP-CoA [125] (Table 9). This enzyme catalyzes the following reaction:

$$\text{3-HP-CoA} = \text{AA-CoA} + H_2O \tag{35}$$

- **Acrylyl-CoA Thioesterase (AcoaTioE)**: No studies characterizing this enzyme's kinetic parameters were found in literature. Hence, an enzyme catalyzing a similar reaction was sought. Ultimately, *E. coli*'s acyl-CoA thioesterase was selected as surrogate. This enzyme was the one used in the work of Chu *et al.* (2015), which using a direct bio-based route [4] achieved AA production for the first time. As opposed to the original reaction, this one was characterized in a report by Zhuang *et al.* (2008), where it was reported that this enzyme kinetics follow the Michaelis-Menten model, with a K_{cat} of 96 s^{-1}, and K_m for AA-CoA of 0.167 mM [126] (Table 9).

$$\text{AA-CoA} + H_2O \rightarrow \text{AA} + \text{CoA} + H^+ \tag{36}$$

The three reactions were added to the six 3-HP producing models, thus creating the six final models able to produce AA, through the different routes, using either glucose or glycerol as carbon source, namely:

- AA_GlyPath_GLC
- AA_GlyPath_GLY
- AA_McoaPath_GLC
- AA_McoaPath_GLY

- AA_BalaPath_GLC
- AA_BalaPath_GLY

Table 9.: Rate Laws, equations, and respective parameters for the reactions required to convert 3-hydroxypropionate into acrylic acid.

Reaction ID	Kinetic Law	Equation	Parameters
3hpcoaS	Michaelis-Menten	$E \cdot K_{cat} \cdot \frac{A}{K_{m,a}+A} \cdot \frac{B}{K_{m,b}+B} \cdot \frac{C}{K_{m,c}+C}$	E = 100 mM K_{cat} = 36 s^{-1} $K_{m,a}$ = 0.015 mM $K_{m,b}$ = 0.01 mM $K_{m,c}$ = 0.05 mM
3hpcoaDH	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A}{K_m+A}$	E = 100 mM K_{cat} = 96 s^{-1} K_m = 0.06 mM
AcoaTioE	Michaelis-Menten	$\frac{E \cdot K_{cat} \cdot A}{K_m+A}$	E = 100 mM K_{cat} = 0.55 s^{-1} K_m = 0.167 mM

3.2.4 *Exchange reaction*

The exchange reactions allow the transport of AA (XCH_AA), glycerol (XCH_GLY), malonyl-CoA (XCH_MCOA), and β-alanine (XCH_BA) from the cytoplasm to the extracellular compartment. Since no kinetic data concerning transport reactions for these metabolites was found, these reactions were included akin to the ones in the original model [49], and they represent the diffusion through the outer membrane [127]. Furthermore, this entire set of reactions follows the reversible Michaelis-Menten kinetics, with a V_{max} of 100 mM/s and a K_m of 10 mM, as it was arbitrarily chosen in the original model [49].

3.3 TIME COURSE SIMULATION

The time course simulations were performed to assess not only AA production but also the yields of the main intermediaries, glycerol, malonyl-CoA, β-alanine and 3-HP. Therefore, these simulations were conducted using a deterministic method (LSODA) from COPASI [53], with a duration of three or six hours, to allow the consumption of all the carbon source.

As one of the goals of this project was also to compare the AA yields associated with glucose and glycerol, these simulations were conducted using both carbon sources. The initial concentrations for glucose and glycerol were, respectively, 55.5 mM (10 g/L) and

217.2 mM (20 g/L), which allowed comparing both carbon sources, and the three pathways for each carbon source. The glycerol concentration is two times higher to maintain the same amount of carbon. Furthermore, to compare the results obtained in these models to the results found in the literature for these routes, the initial concentration of the carbon source for each simulation was set to replicate the initial conditions of the published result.

3.4 OPTIMIZATION STRATEGIES

For each AA producing model that will be used in this section, two new models were generated. One is the exact copy of the respective model, to perform the optimization and simulate AA production, and the other were they are converted to a chemostat system, which was used to perform the MCA. These models represent the first mutant strain (Mutant 0) of each pathway, from which the optimization was performed. The names of the resulting models are presented in Table 10.

Table 10.: Mutant 0 models of each pathway, generated from the previously assembled models to find optimization strategies to improve acrylic acid production.

Pathway	**Model**
Glycerol Route	GlyPath_SIM_Mutanto
	GlyPath_MCA_Mutanto
Malonyl-CoA Route	McoaPath_SIM_Mutanto
	McoaPath_MCA_Mutanto
β-alanine Route	BalaPath_SIM_Mutanto
	BalaPath_MCA_Mutanto

In genetic engineering, the three main techniques to optimize a pathway are genes over-expressions, under-expressions, and knock-outs. In kinetic models, these modifications can be simulated by adjusting the V_{max} value accordingly. For instance, if a gene is over-expressed ten times, the enzyme concentration should present a tenfold increase, which according to equation 8, involves having a V_{max} tenfold higher. Likewise, for under-expressions, a tenfold under-expression is represented by a decrease in the V_{max}. Whereas, knock-outs are simulated *in silico* by setting the V_{max} to zero. Furthermore, in the latter case, before changing the reaction, it is essential to verify if the reaction belongs to the set of critical reactions, as knocking-out those reactions *in vivo* leads to the death of the organism, diminishing AA production.

Therefore, the first step was to obtain the FCC through a MCA for each pathway, using the chemostat models. These coefficients reflect the control each reaction has over AA formation. Thus, the reaction with most influence was selected and modified accordingly to improve the pathway yields, creating new mutant strains.

The following step was to use the batch models from the simulations and to use COPASI's optimization tasks to optimize AA production, by creating *in silico* mutant strains (labeled Mutant 1). The random search method was selected as optimization strategy. This method modifies random combinations of selected parameters to determine the combination that better fits the objective function and exclude the ones that do not fulfill the selected constraints. Furthermore, this method requires a large number of iterations to bestow correct results, as the chance of finding a global solution for the objective function is directly proportional to the number of iterations [128]. Hence, in this work the limit was set to a thousand iterations.

Moreover, it is worth mentioning that the goal of this method is not to meticulously predict the final concentration of AA, but rather finding reactions that are promising targets for optimization. Therefore, the under and over-expressions were limited to 50 times the original V_{max} value. This limit allows overcoming the negative influence of that reaction *in silico*, whilst not impairing its implementation *in vivo*.

Finally, after creating new mutants, this process was repeated to optimize mutant strains. The optimization eventually stopped once either the glucose feed was limiting the AA production, the limiting reaction was already optimized, or the system could no longer reach a stable steady-state point.

4

RESULTS AND DISCUSSION

This chapter will focus on the presentation and discussion of the results obtained with the simulations of 3-HP and AA models, using either glucose or glycerol as the carbon source, and their respective optimization strategies.

4.1 V_{max} CALCULATION

This section is divided into the calculation of the V_{max} values for the CCM extension and the k values for the synth reactions.

4.1.1 *Central Carbon Metabolism Extension*

Using a stoichiometric model, the V_{max} of the enzymes that belong to the native metabolism of *E. coli* were calculated according to Method 1. The resulting values for all the nine reactions required for the extension of the CCM are presented in Table 11:

Table 11.: V_{max} values calculated for the reactions required for the extension of the central carbon metabolism.

Reaction ID	V_{max}
G3pD	0.03095 mM/s
G3pP	0.38996 mM/s
GlyK	0.10557 mM/s
AccC	0.23910 mM/s
GluD	0.52125 mM/s
AspAT	1.96832 mM/s
AspC	1.15×10^{-05} mM/s
GlyD	13.1052 mM/s
DhaPT	0.0586 mM/s

4.1.2 *Synth Reactions*

Using the same principle of Method 1, the k values for the synth reactions were calculated using the fluxes determined with the stoichiometric model. The calculated values are presented in Table 12.

Table 12.: Synth reactions added to the model and respective parameters. These reactions were created for dihydroxyacetone phosphate (DAP), acetyl-CoA (ACCOA), malonyl-CoA (MCOA), L-glutamate (LGLU), aspartate (ASP) and β-alanine (BA).

Reaction ID	Parameter (1/s)
SynthDAP	k = 0.02006
SynthACCOA	k = 0.69799
SynthMCOA	k = 0.09250
SynthLGLU	k = 0.38260
SynthASP	k = 0.05180
SynthBA	k = 0,00001

4.2 ACRYLIC ACID PRODUCTION

The results of the simulations performed, to assess 3-HP and AA production in all models, are presented in this section. Furthermore, the end goal is to understand which is the best carbon source for AA production and which pathway has more potential for AA industrial-scale production, comparing the results with published data from *in vivo* experiments.

4.2.1 *Central Carbon Metabolism Extension*

First it is necessary to analyze the results of the CCM extension to assess the production of the three intermediates (glycerol, malonyl-CoA, and β-alanine), and to compare the results for the two carbon sources used. Concerning glucose as carbon source, the CCMext_GLC was able to consume the 10 g/L of glucose provided in the simulation time frame (Figure 11A), producing 0.17 g/L of glycerol (Figure 11B), a maximum value of 1.20 g/L of malonyl-CoA (Figure 11c), and 0.0009 g/L of β-alanine (Figure 11D). It is worth mentioning that when all glucose is consumed, malonyl-CoA suddenly drops. That is due to the fact that when the production of malonyl-CoA stops, the Synth reaction for malonyl-CoA (SynthMCOA) (Figure 10) is still consuming this metabolite, causing a drop in its concentration.

Even though the simulation of the production of glycerol and malonyl-CoA was successful, the β-alanine model exhibited a major setback. The extremely low V_{max} of the AspC enzyme (Table 4) was limiting the amount of β-alanine produced. The V_{max} was calculated with the flux obtained from the stoichiometric model simulation (Method 1), which resulted in a relevant bottleneck. Thus, this parameter was recalculated using the method developed for the heterologous pathway reactions' parameters (Method 2), shifting the β-alanine flux production limits towards L-aspartate formation. According to Ramjee *et al.* (1997), the K_{cat} value for this enzyme is 0.57 s^{-1} [118]. Therefore, the V_{max} value was set to 57 mM/s in all the models, which drastically altered the production of β-alanine two orders of magnitude (0.037 g/L) (Figure 11E).

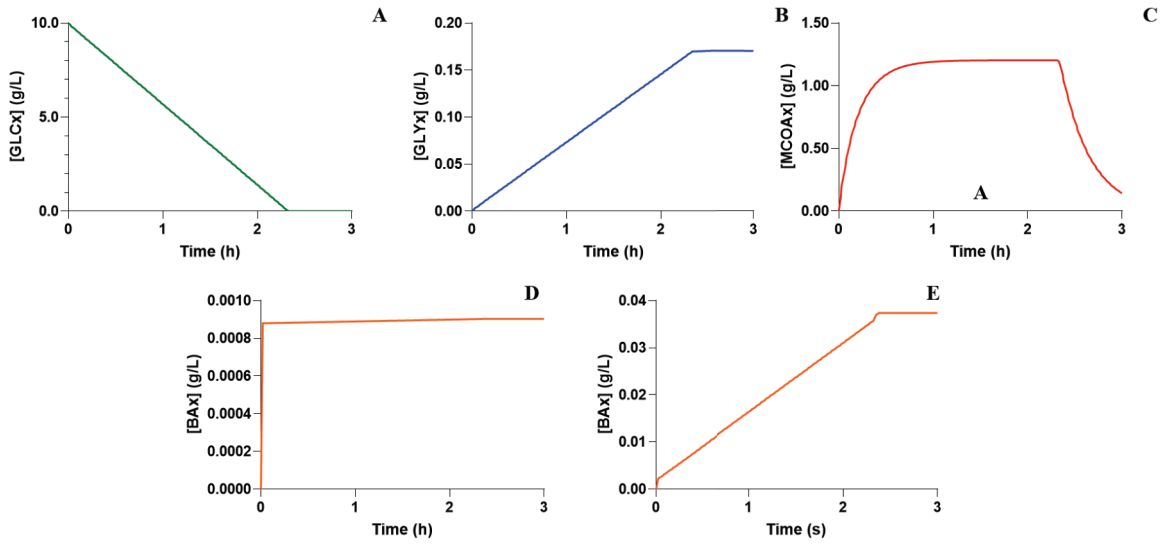


Figure 11.: Results of the CCMext_GLC model concerning glucose (GLC) consumption (A), and the production of glycerol (GLY) (B), malonyl-CoA (MCOA) (C), and β-alanine (BA) before (D) and after (E) increasing V_{max} of the AspC reaction.

Considering glycerol as the only source of carbon, the CCMext_GLY model was able to produce a maximum of 1.33 g/L of malonyl-CoA and 0.0026 g/L of β-alanine, within the simulation time frame, while consuming 20 g/L of glycerol (Figure 12A). When comparing these results with the ones of the CCMext_GLC, it is possible to observe that the malonyl-CoA concentration peaked at a higher value (Figure 12B), but its concentration started to decay much sooner because there was an issue with the assimilation of glycerol, causing the accumulation of intermediaries (Figure 12D). Furthermore, it can also be observed that this carbon source produced a tenfold smaller concentration of β-alanine (Figure 12C).

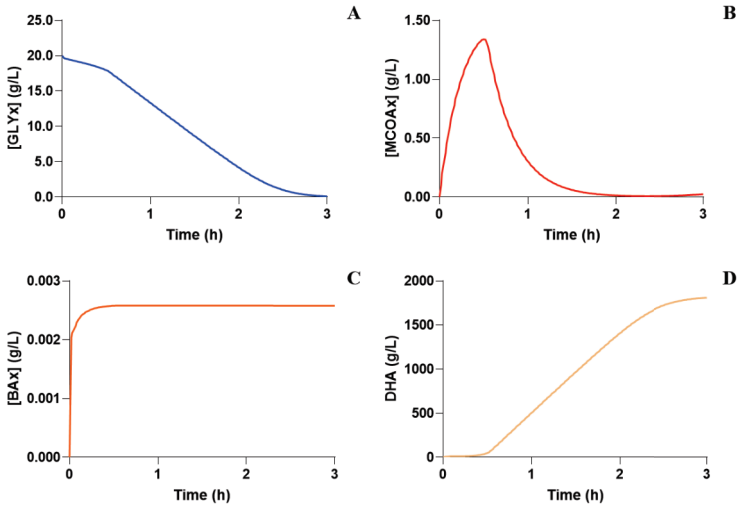


Figure 12.: Results of the CCMext_GLY model concerning glycerol (GLY) consumption (A), and the production of malonyl-CoA (MCOA) (B), and β-alanine (BA) (C).

4.2.2 *Glycerol Models*

Issues were found in the models created for modeling 3-HP and AA production, when glycerol was used as carbon source. The flow of carbon toward the CCM stopped seconds after the simulation started, leading to the accumulation of 3-HP (Figure A.1). After analyzing the system, it was determined that two factors were responsible for this behavior.

The first factor is associated with the heterologous pathway, as the V_{max} of all reactions in this pathway was set as not to limit flux through these reactions. Therefore, these reactions would uptake most of the available glycerol, thus limiting the amount of carbon towards the CCM, which eventually led to the depletion of crucial cofactors for AA syntheses, such as NAD and ATP.

The second factor is associated with the NAD affinity to enzymes GlyD and 3hpD. The K_m value in the GlyD is over tenfold higher than in 3hpD, thus the low concentration of NAD available was mainly used by the 3hpD, which will block the flux towards the CCM, exacerbating the energy production problem even further.

A couple of hypotheses were devised to circumvent these problems. Because the NAD affinity issue was blocking the flux of carbon towards the CCM, the first hypothesis was increasing the affinity of the GlyD enzyme towards NAD. When comparing the affinity with other enzymes present in the model, this enzyme had a significantly higher value. Therefore, other works that characterized this enzyme were searched, to find an alternative for this K_m value. In the work of Zang *et al.* (2010) [129] this enzyme, using another substrate, was described with an affinity towards NAD of 0.0165 mM. Thus, the value was updated and simulated in a period of three hours. However, the problem persisted

as most glycerol was still going towards the formation of 3-HP (Figure A.2). The next strategy was increasing the maximal rate of the GlyD to increase the flux in the CCM, allowing the continuous production of AA. Hence, the V_{max} was recalculated according to Method 2, resulting in a V_{max} of 4298.4 mM/s, which ultimately led to an accumulation of dihydroxyacetone as the activity of the DhaPT enzyme was not enough to quickly consume all the dihydroxyacetone produced by the GlyD (Figure A.3).

The third hypothesis involved decreasing the affinity of the 3hpD towards NAD, to improve the activity of the GlyD, and the maximum rate of the GlyDH to control the flux that is directed towards to heterologous route. Since no further K_m value for the 3hpD was found in the literature, it was assumed that this enzyme had a similar affinity as the GlyD (K_m = 0.8 mM). Furthermore, the V_{max} of the GlyDH was set to 0.621 mM/s, representing an enzyme concentration of 10 mM instead of the 100 mM used in Method 2. These changes allowed directing flux towards the CCM, which resulted in the production of AA without the excessive accumulation of any intermediaries.

After solving the previous issues, simulations were performed and the results are presented next. Regarding the production of 3-HP, COPASI predicted that the 3HP_GlyPath_GLC model is able to convert 10 g/L of glucose into 0.17 g/L of 3-HP in three hours (Figure 13). In contrast, when using 20 g/L of glycerol, the 3HP_GlyPath_GLY model produced 8.3 g/L of 3-HP over the course of six hours (Figure 13).

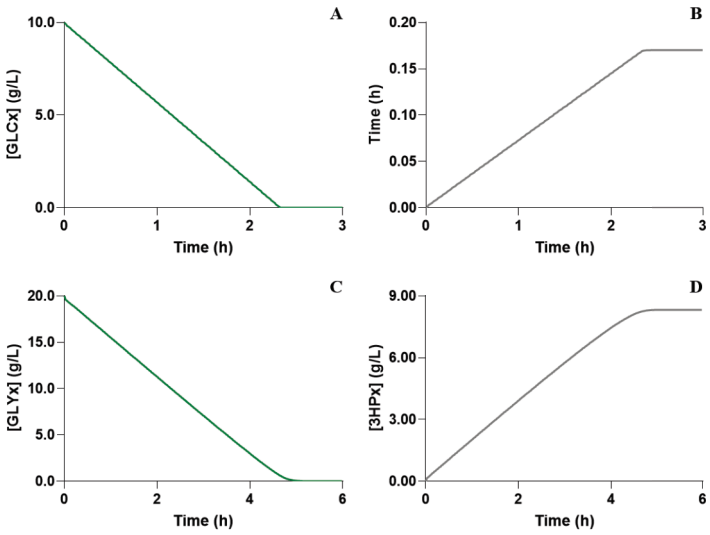


Figure 13.: Simulation results concerning glucose consumption (GLC) (A) and 3-hydroxypropionate (3HP) production (B) for the 3HP_GlyPath_GLC model, and glycerol (GLY) consumption (C) and production of 3HP (D) for the 3HP_GlyPath_GLY model.

The results were compared to the yields retrieved from available literature (Table 2), to assess the predictive capabilities of the final 3-HP models. The 3HP_GlyPath_GLY model was used to simulate the synthesis of 3-HP for different concentrations of glycerol

that replicate the work of different publications (9.2 g/L, 18.4 g/L, and 40 g/L). When using a concentration of 9.2 g/L of glycerol, the model predicted a production of 3.65 g/L of 3-HP, which, as shown in Table 2, is close to what was reported by Raj *et al.* (2009) [103]. However, in the latter cases, there were considerable differences between what was reported and what was predicted. Using 18.4 g/L, the model predicted a 3-HP concentration of 7.62 g/L, which is almost two times higher than what was obtained by Rathnasingh *et al.* (2009) [104] (Table 2). Finally, using 40 g/L of glycerol, 17.11 g/L of 3-HP were synthesized *in silico*, a twofold higher concentration than the yield of Chu *et al.* (2015) [4] (Table 2).

Furthermore, Chu and colleagues also tested 3-HP synthesis using glucose as carbon source [4]. When replicating the initial conditions of this study (Table 2) *in silico*, the 3HP_-GlyPath_GLC model predicted the synthesis of 0.36 g/L of 3-HP. The difference might be associated with the production of glycerol, more specifically, in the flux through G3pD and G3pP (Figure 10), as according to Chu *et al.* (2015) this strain is able to accumulate more glycerol (2.5 g/L) than this model is able to produce, for this amount of glucose (0.34 g/L).

Finally, when considering the expression of the complete pathway for AA production, the results showed a significantly higher amount of AA when glycerol was used as the carbon source. As shown in Figure 14, whereas glycerol allows producing 6.65 g/L of AA, glucose only produces 0.13 g/L, which is consistent with the results obtained for the 3-HP production. Moreover, since the reactions of the heterologous pathway had high V_{max} values, there is no 3-HP accumulation in both cases (Figure 14), which is not in line with what was found in the literature as 3-HP accumulation was always reported [4, 1].

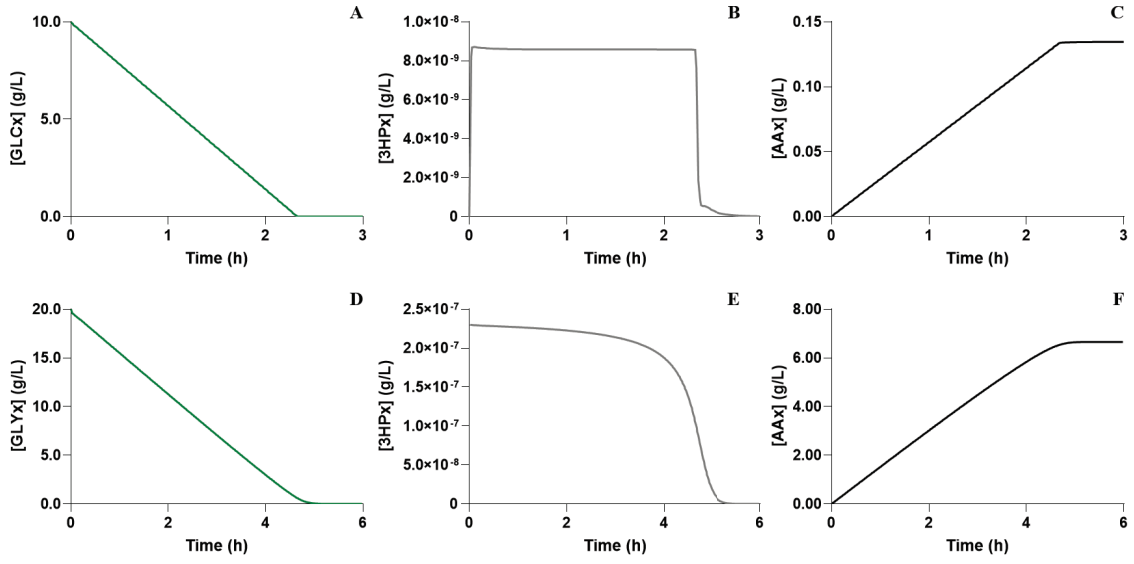


Figure 14.: Simulation results concerning glucose consumption (GLC) (A) and acrylic acid (AA) production (B) for the AA_GlyPath_GLC model, and glycerol (GLY) consumption (C) and production of AA (D) for the AA_GlyPath_GLY model.

Concerning the full pathway to AA production, two works were found in the literature. Using the initial concentration of glycerol used in the work of Tong *et al.* (2016) [1], the AA synthesis predicted by the model (6.65 g/L) was considerably higher than what the authors reported (Table 3). This difference might be related, as previously mentioned, with the high flux in the heterologous pathway preventing the accumulation of intermediaries. Moreover, Chu *et al.* (2015) also tested AA production from glucose (Table 3) in a genetically engineered *E. coli* strain. However, the initial glucose conditions were not stated in the article, thus a comparison with their results was not possible.

When considering a relative similar amount of carbon, the glycerol pathway is associated with higher AA yields when using glycerol, as this carbon source has a straightforward path towards the heterologous pathway, thus reducing the flux in upstream reactions. This result is in agreement with most published works, concerning the use of the glycerol *in vivo*, instead of glucose, as the main carbon source for the glycerol pathway. For instance, in the work of Chu *et al.* (2015), when using relatively similar carbon quantities, the bacteria produced more 3-HP when using glycerol as carbon source.

4.2.3 *Malonyl-CoA Models*

The 3HP_McoaPath_GLC model predicted the production of 1.64 g/L of 3-HP (Figure 15), which is a 9.6-fold increase when comparing to the results of the glycerol pathway, when using 10 g/L glucose. Furthermore, when 20 g/L of glycerol were used, the model 3HP McoaPath GLY only produced 0.19 g/L (Figure 15), which is significantly less than what was produced by glycerol pathway (43-fold decrease). However, the use of a different carbon source did not directly decreased the production of AA. The low availability of phosphoenolpyruvate limits the flux through the DhaPT (Figure 10), inducing the accumulation of dihydroxyacetone, thus restricting the amount of carbon reaching the CCM (Figure A.4). Therefore, this model seems to be impaired when considering glycerol as carbon source, and will probably return erroneous results regarding AA synthesis.

Several works were published using this pathway reporting different yields, when using glucose as carbon source [7]. However, regarding shake flask experiments, two stand out. The first one was performed by Cheng *et al.* (2015), which reported the 3-HP production from 10 g/L of glucose [105]. From such initial condition the 3HP_McoaPath_GLC model produced 1.64 g/L of 3-HP, which is very similar to the final concentration reported (Table 2). Using 20 g/L of glucose the simulations predicted the production of 3.26 g/L of 3-HP, which is in line with the value achieved by Liu *et al.* (2016) (Table 2) [7]. Therefore, regarding the predicted production of 3-HP from glucose, the malonyl-CoA model seemed to match the results found in the literature.

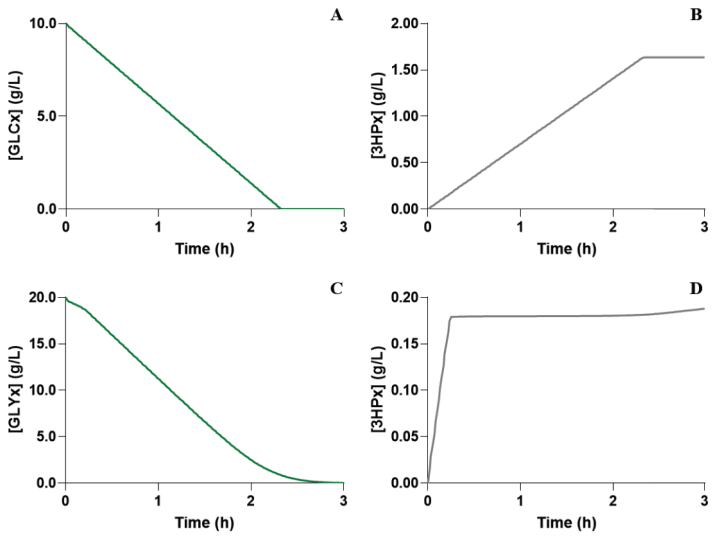


Figure 15.: Simulation results concerning glucose consumption (GLC) (A) and 3-hydroxypropionate (3HP) production (B) for the 3HP_McoaPath_GLC model, and glycerol (GLY) consumption (C) and production of 3HP (D) for the 3HP_McoaPath_GLY model.

Finally, in the AA models, the 3-HP was converted to AA with considerable high yields, which resulted in virtually no intermediary accumulation (Figure 16). Furthermore, when using glucose, the model predicted the production of 1.33 g/L of AA, thus 81% of the 3-HP was converted into AA. When glycerol was used as carbon source, 0.15 g/L of AA were produced, which means that 79% of the produced 3-HP was converted (Figure 16).

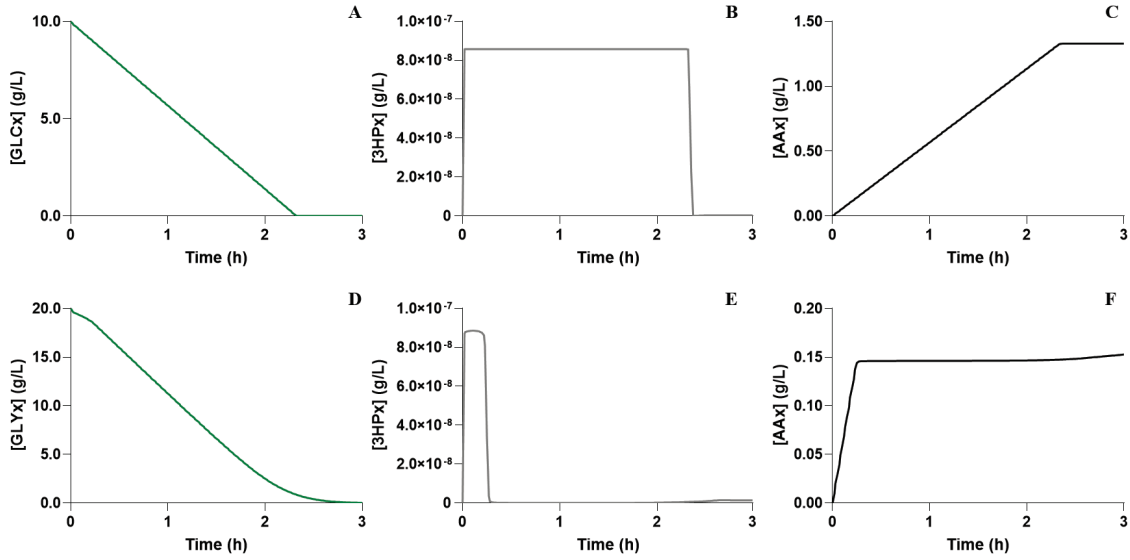


Figure 16.: Simulation results concerning glucose consumption (GLC) (A) and acrylic acid (AA) production (B) for the AA_McoaPath_GLC model, and glycerol (GLY) consumption (C) and production of AA (D) for the AA_McoaPath_GLY model.

Liu and Liu (2016) developed a genetically engineered strain of *E. coli* to produce AA using malonyl-CoA as an intermediate. The resulting strain produced 0.013 g/L of AA from 20 g/L of glucose [7] (Table 3), whereas the model predicted the production of 2.65 g/L of AA for the same environmental conditions. This difference can be associated with several factors. However, as the 3-HP results seemed to be in line with some published results, the most likely reason is that the heterologous pathway reactions V_{max} was calculated assuming excess availability of enzymes. As mentioned before, these conditions prevent intermediary accumulation, which does not reflect what has been reported in the literature [4, 7, 1].

Model predictions for both the 3HP_MCOAPath_GLY and AA_MCOAPath_GLY are skewed due to the accumulation of dihydroxyacetone, not allowing comparison with simulations performed with glucose as carbon source. Regardless, all studies that tested AA or 3-HP through the malonyl-CoA pathway, used glucose, which can be seen as an indicator that this metabolite is, in fact, the best carbon source for this route. This premise should be corroborated as no actual experimental work comparing both carbon sources was found.

4.2.4 *β-alanine Models*

The *β*-alanine pathway is by far the less studied route for 3-HP and AA production in *E. coli*. In fact, to the best of our knowledge, only a few articles have been published concerning this intermediate for 3-HP synthesis in *E. coli* [93]. Works in which AA is produced via the *β*-alanine route are yet to be published, which might be related with the fact that publications using this route reported significantly lower yields when compared with the glycerol and malonyl-CoA pathways [96]. In fact, this study models successfully replicated these results, as this route predicted the production of 0.03 g/L of 3-HP from glucose, representing a 5.6-fold decrease from the glycerol model and 54.6-fold from the malonyl-CoA model (Figure 17). Furthermore, the simulation results when glycerol was used as carbon source also showed a significant decrease in production, with 8.8×10^{-04} g/L of 3-HP (Figure 17). However, as in the malonyl-CoA model, these are misleading results due to issues with the fluxes of the glycerol assimilation *via* (Figure A.5). Hence, a comparison between both carbon sources is once more impaired.

Even though the production of 3-HP via *β*-alanine has not been thoroughly studied, in 2016 Song and colleagues [93] metabolically engineered a strain of *E. coli* to express the heterologous pathway. In their experiment, they grew the cells in a batch system with of glucose as the main carbon source, at a concentration of 15 g/L. From this concentration, a total of 0.09 g/L of 3-HP were synthesized (Table 2), which is a very similar result to the *β*-alanine model predictions (0.05 g/L).

Despite the lack of *in vivo* experiments to compare the results for AA production, the complete bio-based route was still tested to compare with the other two pathways. Once

again, the percentage of 3-HP converted to AA is still high, with approximately 0.03 g/L produced from glucose and 7.2×10^{-04} g/L from glycerol (Figure 18).

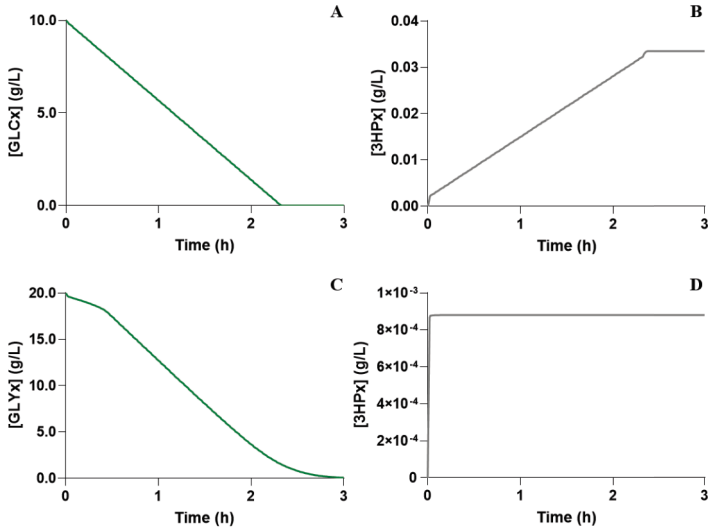


Figure 17.: Simulation results concerning glucose consumption (GLC) (A) and 3-hydroxypropionate (3HP) production (B) for the 3HP_BalaPath_GLC model, and glycerol (GLY) consumption (C) and production of 3HP (D) for the 3HP_BalaPath_GLY model.

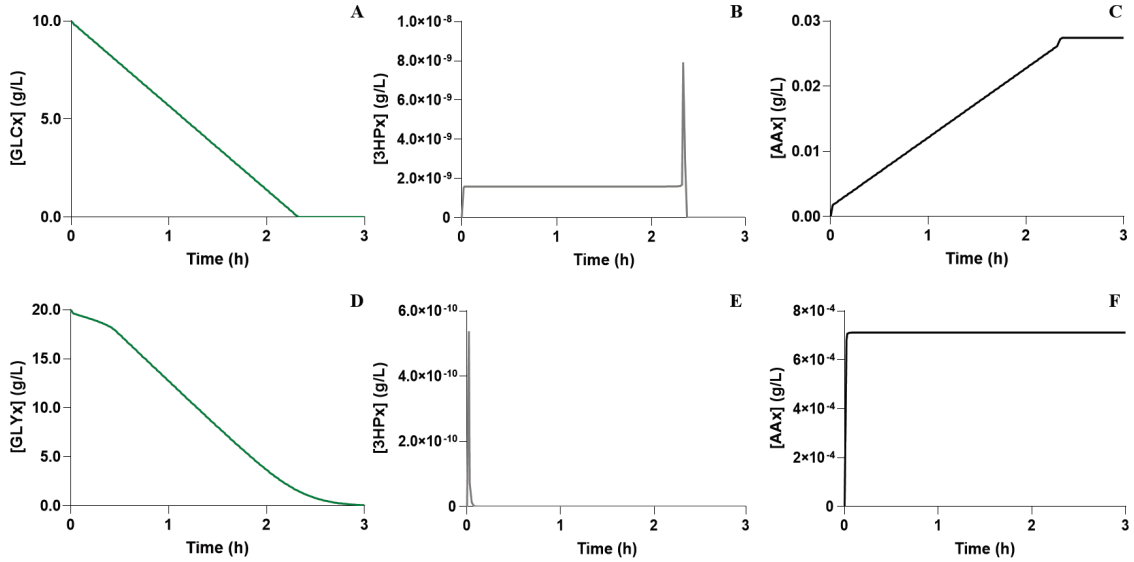


Figure 18.: Simulation results concerning glucose consumption (GLC) (A) and acrylic acid (AA) production (B) for the AA_BalaPath_GLC model, and glycerol (GLY) consumption (C) and production of AA (D) for the AA_BalaPath_GLY model.

From this work it was possible to assess which of the three bio-based routes for AA production was associated with higher yields. First, as shown in Table 14, the results of these models suggest that the β-alanine route is by far the route that produced lower yields. This is in agreement with the literature review for this pathway [93, 96]. On the other hand, the glycerol pathway exhibited the highest yields, when combined with the use of glycerol as the carbon source (Table 13 and 14). Although this pathway is well established [7], a relevant caveat most be recalled. As mentioned before, this pathway has a reaction that is dependent on the presence of vitamin B12, which represents a significant economic disadvantage in an industrial-scale production [93, 96, 97, 98, 99]. Hence, to make this route viable, either the yield must be significantly improved to overcome the cost of the supplementation, or a cheaper way to produce B12 must be found. Therefore, in this context, the malonyl-CoA route gains significance as it does not require the supplementation of vitamins while providing acceptable AA yields (Table 13 and 14) [95, 96, 105].

Table 13.: Summarized results of the 3-hydroxypropionate producing models for the three distinct pathways using either glucose or glycerol as carbon source.

Model	**Carbon Source**	**Yield (g/L)**
Glycerol	Glucose	0.17
Model	Glycerol	8.30
Malonyl-CoA	Glucose	1.64
Model	Glycerol	0.19
β-alanine	Glucose	0.03
Model	Glycerol	8.8×10^{-04}

Table 14.: Summarized results of the acrylic acid producing models for the three distinct pathways using either glucose or glycerol as carbon source.

Model	**Carbon Source**	**Yield (g/L)**
Glycerol	Glucose	0.13
Model	Glycerol	6.65
Malonyl-CoA	Glucose	1.33
Model	Glycerol	0.15
β-alanine	Glucose	0.03
Model	Glycerol	7.2×10^{-04}

4.3 OPTIMIZATION STRATEGIES

Ideally, the optimization should be performed for all the developed model. However, the AA_GlyPath_GLY model did not provide a stable steady-state, thus performing a MCA was not possible. Furthermore, the malonyl-CoA and β-alanine models presented issues when using glycerol as carbon source. Therefore, only models where glucose was used as the carbon source were optimized.

Furthermore, it is worth noting that the optimizations were not performed with the intent of accurately predicting indisputable improvements in AA production, but are rather a guideline in the search for strategies for strain optimization [36]. Therefore, the results presented here should not be used to perform quantitative comparisons between different pathway designs, but rather to provide a general idea of what are the best targets to improve the AA yields *in vivo*.

The optimization strategies identified for each of these three models (AA_GlyPath_GLC, AA_McoaPath_GLC, and AA_BalaPath_GLC) will now be presented.

4.3.1 *Glycerol Model*

The first MCA was performed in the GlyPath_MCA_Mutant_0 model. The FCC that resulted from this analysis showed that the flux towards AA formation is mainly controlled by the G3pD reaction (Equation 19), which exhibited the highest positive FCC (Figure 19). This reaction is a potential bottleneck for AA production, thus a target for overexpression. The GlyPath_SIM_Mutant_0 model was optimized using COPASI's optimization task, which suggested an optimum V_{max} of 1.5475 mM/s, a nearly 50-fold increase in the activity of the G3pD. Hence, the parameter was modified for both models, creating Mutant 1 (GlyPath_MCA_Mutant_1 and GlyPath_SIM_Mutant_1). This model was able to produce 1.60 g/L of a AA (Figure 20), a value eleven fold higher than the initial production.

A second iteration using the Mutant1 models was performed, in which reaction G3pP (Equation 20) was determined to have the highest FCC in the AA yields (Figure 19). The positive coefficient led to the over-expression of the G3pP by increasing its V_{max} 4.7 fold (1.8098 mM/s) for maximizing AA synthesis. Mutant 2 was created (GlyPath_MCA_Mutant_2 and GlyPath_SIM_Mutant_2), producing 3.77 g/L of AA (Figure 20), thus increasing its production over twenty-eight times than the original value.

It is important to mention that enzymes such as the glyceraldehyde-3-phosphate dehydrogenase (GDH) and the GlyK were considered as targets for under-expression or knock-outs, because of their negative coefficients (Figure 19). Since both were included in the critical reactions of *E. coli*, they were only tested as under-expression targets. Unfortunately, reducing the activity of those enzymes did not significantly improve AA production.

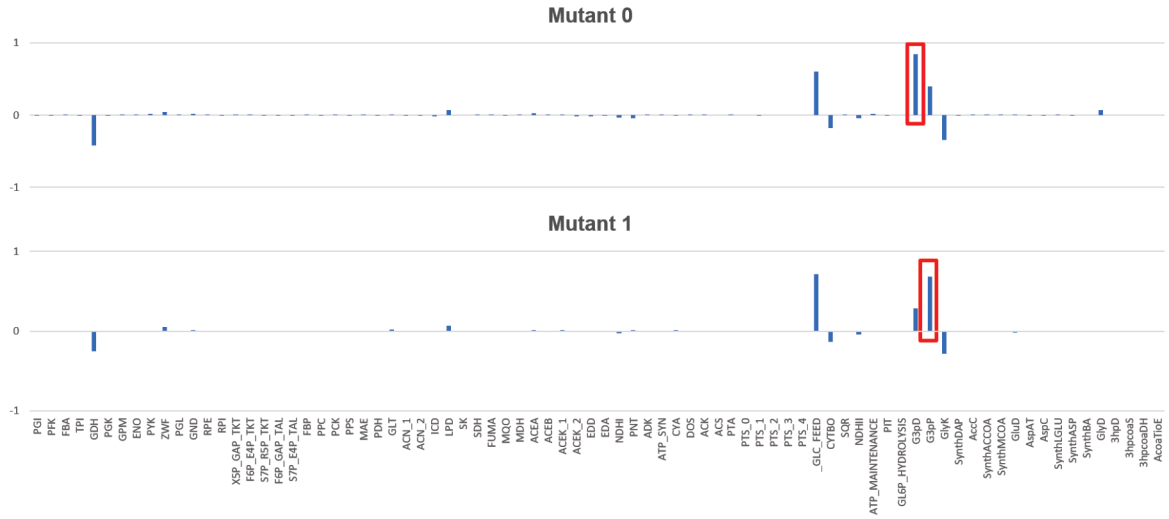


Figure 19.: Flux Control Coefficients for the mutants of the glycerol pathway, where the reaction with the most impact in the production of acrylic acid is highlighted in red.

Another MCA was performed in the Mutant 2 model, but the steady-state analysis could not reach a stable steady-state. Hence, the FCC were not available and the optimization of this pathway was terminated. Nevertheless, this analysis was useful to identify two potential targets for optimization in this pathway, the G3pD, and the G3pP (Figure A.6).

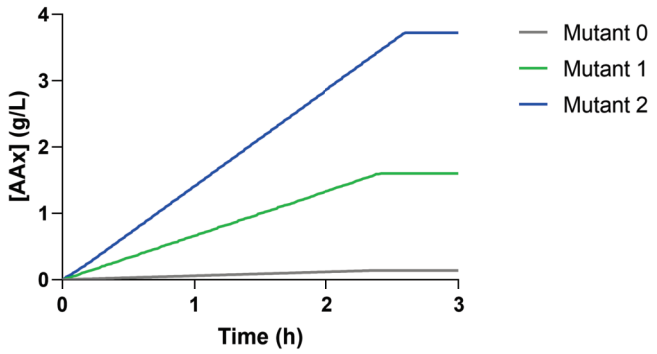


Figure 20.: Acrylic acid production from the three mutants developed for the glycerol model. The simulations were conducted with 10 g/L of glucose for all the mutants. Mutant 0 - model with the heterologous pathway; Mutant 1 - 50-fold increase in the V_{max} of the G3pD; Mutant 2 - 4.7-fold increase in the V_{max} of the G3pP.

4.3.2 *Malonyl-CoA Model*

In the malonyl-CoA chemostat model (McoaPath_MCA_Mutant_0), the FCC values (Figure 21) show that the AccC reaction (Equation 22) has a major influence in AA pro-

duction (Figure A.6). Moreover, the positive FCC indicates that this reaction acts as a bottleneck resulting in the accumulation of acetyl-CoA. Therefore, the best strategy is an over-expression to increase the flux of this reaction.

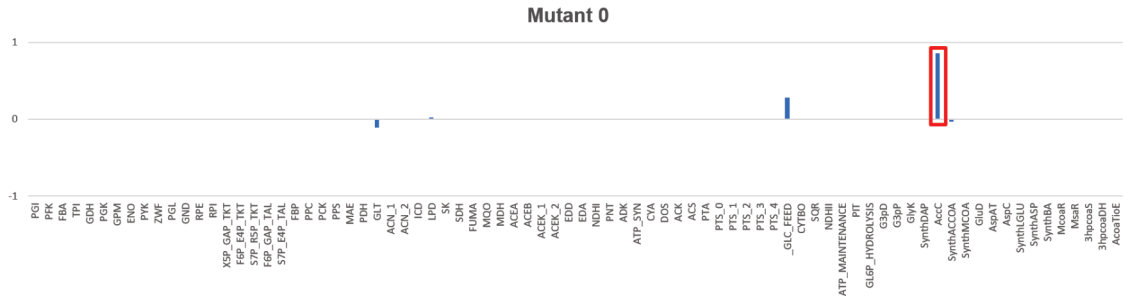


Figure 21.: Flux Control Coefficients for the mutant of the malonyl-CoA pathway, where the reaction with the most impact in the production of acrylic acid is highlighted in red.

Using the batch model (McoaPath_SIM_Mutant_0) it was determined that the optimum V_{max} is 0.5873 mM/s, which corresponds to a 2.4-fold over-expression. Mutant 1 was then created (McoaPath_MCA_Mutant_1 and McoaPath_SIM_Mutant_1), producing 3.21 g/L of AA, which is more than two times higher than the value obtained in Mutant 0 (Figure 22). It was not possible to recalculate the FCC as a new stable steady-state was not available for the new model.

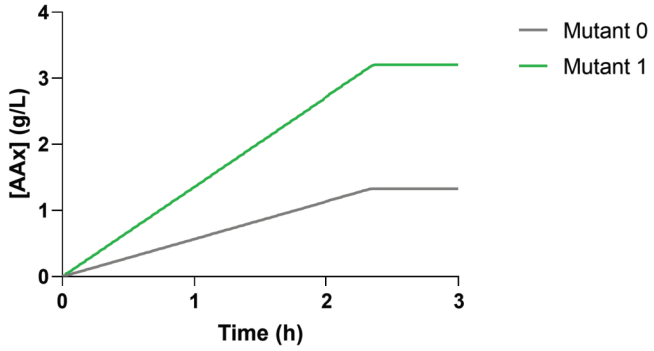


Figure 22.: Acrylic acid production from the two mutants developed for the malonyl-CoA model. The simulations were conducted with 10 g/L of glucose for all the mutants. Mutant 0 - model with the heterologous pathway; Mutant 1 - 2.4-fold increase in the V_{max} of the AccC.

Even though other optimization strategies are recommended in the literature, it seems to be well established that increasing the flux through the AccC reaction is the primary strategy to improve AA production [96], which supports the result obtained. This enzyme is maintained at low concentrations in wild type cells. Hence, it makes sense that the first strategy should be over-expressing it [130]. Nevertheless, other modifications could also be performed to overcome this bottleneck, like redirecting the carbon flux towards the

formation of acetyl-CoA, increasing the availability of ATP, or increasing the specificity of the enzyme towards its substrates [96].

4.3.3 *β-alanine Model*

Performing the MCA with the chemostat model (BalaPath_MCA_Mutant_0), revealed that the β-alanine model has several reactions affecting the AA yields, but the one with the most significant impact is the AspAT (Figure A.6), which is responsible for the conversion of oxaloacetate into L-aspartate (Equation 24). This reaction has a positive coefficient indicating that it is a bottleneck that impairs the downstream flux towards β-alanine. Therefore, the strategy applied was, once again, an over-expression.

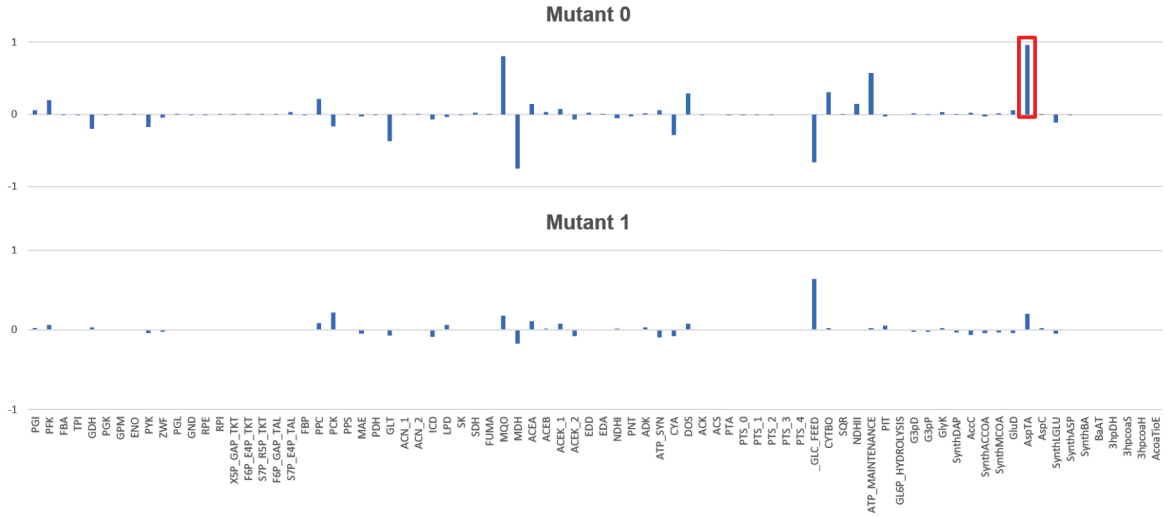


Figure 23.: Flux Control Coefficients for the mutants of the β-alanine pathway, where the reaction with the most impact in the production of acrylic acid is highlighted in red.

Using the BalaPath_SIM_Mutant_0 model, COPASI estimated that a 50-fold over-expression, resulting in a V_{max} of 97.9716 mM/s, would maximize AA production. This approach allowed creating Mutant 1 (BalaPath_MCA_Mutant_1 and BalaPath_SIM_Mutant_1), a strain in which the predicted AA concentration after three hours is 0.87 g/L, which corresponds to a twenty-nine times increase.

After that, another iteration of the method was performed in the models for Mutant 1, but the MCA result showed that the most significant influence on the AA yield was the glucose feed, which is not a viable target for optimization. Other reactions, like the phosphoenolpyruvate carboxykinase (PCK) and the malate quinone oxidoreductase (MQO), with coefficients around 0.2, also seem to have an impact over this pathway. Nevertheless,

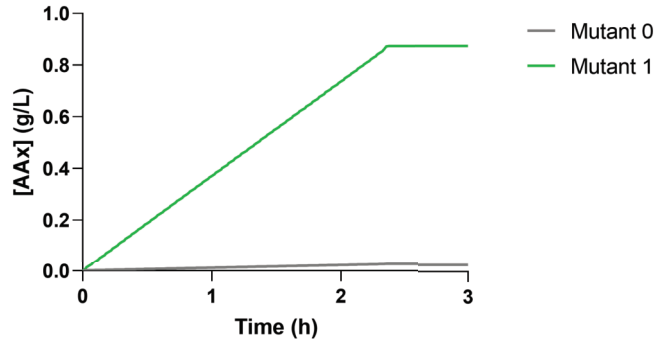


Figure 24.: Acrylic acid production from the two mutants developed for the β-alanine model. The simulations were conducted with 10 g/L of glucose for all the mutants. Mutant 0 - model with the heterologous pathway; Mutant 1 - 50-fold increase in the V_{max} of the AspAT.

over-expressing those reactions did not produce significant increases in AA. Therefore, only the AspAT was identified as a target, which is in agreement with the results of Borodina *et al.* (2015). In that work, it was suggested that the increase of the flux towards L-aspartate was a viable strategy to increase AA production [95]. However, the AspC (Equation 24), which is responsible for the production of β-alanine from L-aspartate, can also be considered as a limiting factor for pathway flux, and a target for optimization, as the $Vmax$ had to be increased for the β-alanine model to work correctly.

As mentioned before, assessing AA production was not the goal of this method. However, when it comes to the pathway with the highest AA yield from glucose, a paradigm shift was observed. As shown in Table 15, the highest AA production was achieved by the glycerol route with 3.77 g/L. Moreover, the malonyl-CoA and β-alanine routes produced 3.21 and 0.87 g/L, respectively.

Table 15.: Summarized results of acrylic acid production for the mutant strains developed for the glycerol, malonyl-CoA, and β-alanine models, using 10 g/L of glucose as substrate.

Model	Strain	Yield (g/L)
Glycerol Route	Mutant 0	0.13
	Mutant 1	1.60
	Mutant 2	3.77
Malonyl-CoA Route	Mutant 0	1.33
	Mutant 1	3.21
β-alanine Route	Mutant 0	0.03
	Mutant 1	0.87

5

CONCLUSION AND FURTHER WORK

5.1 CONCLUSION

In conclusion, this work resulted in the development of 12 distinct models that allow the prediction of 3-HP or AA production using either glucose or glycerol as carbon sources (Table A.1). From the comparison of the results provided from these models with published *in vivo* results, it seems that the models were more effective predicting 3-HP production than the production of AA. That happened because the V_{max} values for the heterologous pathway were calculated with an excessive enzyme concentration, thus the main limiting factor in the synthesis of AA were the fluxes in the CCM. Unfortunately, that does not happen *in vivo* since the reports that tested the full bio-based route showed that some amounts of 3-HP and other intermediates are accumulated in the process.

Regarding AA production, an overall best carbon source was not found. Instead, that answer is specific to each pathway. If the goal is to implement the glycerol pathway, it may be advantageous using glycerol as it allows for a more straightforward path. This route reduces the chance of other reactions diverting the carbon from the heterologous pathway. On the other hand, for the malonyl-CoA and β-alanine routes, results were not robust enough to suggest that one source would have better yields, as models that used glycerol did not work acceptably. Nevertheless, the literature review showed that glucose was the only tested carbon source for these pathways, which might mean that it is beneficial to use this carbon source. However, it is essential to test AA production *in vivo* using both carbon sources, to confirm if it is the case.

Besides, this work also provided more insights on which pathway is capable of providing higher profit margins. Even though the glycerol pathway appears to be correlated with higher AA yields, it is a route that is dependent on vitamin B12 supplementation, which can prove to be expensive on an industrial scale. Therefore, despite producing less AA, it seems to be beneficial to use the malonyl-CoA route, since it still has the second higher yield, it does not need the supplementation of any vitamin, and still was room for optimization.

Finally, since the yields from theses routes remain too low to implement them for industrial-scale production, there is a need for optimization strategies that allow an increase in AA synthesis. From this work, several targets for over-expression were suggested that, theoretically, should increase the amount of AA formed. In the glycerol route, the optimization targets are the G3pD and G3pP. The main goal of these over-expressions was to increase the flux towards glycerol formation, thus increasing its availability for the heterologous pathway. For the malonyl-CoA route, the goal was also to increase the production of malonyl-CoA, which can be achieved by over-expressing the AccC reaction as it is the main limiting factor due to the low bioavailability of the corresponding enzyme in *wild type* cells. Finally, when using the β-alanine route, two targets were suggested, the AspAT and the AspC. The over-expression of these reactions should increase β-alanine formation, which is the main restricting factor in the dynamic model. A final aspect to consider for the optimization of the AA yields is that this model only comprises the CCM. This means that other pathways are almost certainly diverting flux from the AA route. Therefore, as it can be confirmed from the literature, other strategies, like gene knockouts or under-expressions, should be used to force the flux towards the heterologous pathway.

5.2 FURTHER WORK

In the future, further work is required to improve the developed kinetic models. Furthermore, that work should focus on improving the characterization of some key enzymes that presented some difficulties in the development of these models. First, it seems to be crucial to determine the flux of the AspC reaction. The flux determined by the stoichiometric model resulted in a significant limitation for AA formation and had to be modified to allow the β-alanine route to be compared with the other two pathways. Hence, it is essential to obtain experimental data for that reaction, to assess if the modification reflected reality. Moreover, the flux of the G3pD and G3pP reactions, from the glycerol model, seemed to be underestimated due to the low yields when glucose was used as carbon source. Thus, obtaining experimental data on these reactions can also help to improve the models. Another step towards the improvement of the models should be the kinetic characterization of the underlying mechanisms of the DhaPT reaction, and the glycerol transport reactions. This might fix the problems in the assimilation of glycerol, preventing the accumulation of dihydroxyacetone.

After the improvement of the model, the final step should be the validation of the results. Therefore, these heterologous pathways, and their respective optimization strategies, should be implemented *in vivo*, using both carbon sources, to assess the validity of the results provided. Furthermore, these metabolically engineered strains could also be used to determine the flux of the reactions in the heterologous pathway. That would avoid the use of high enzyme concentrations and thus account for the accumulation of intermediates, which would improve the predicting capabilities towards AA formation. Finally, no mechanism for 3-HP and AA transport was described or even identified *in vivo*. Therefore, it becomes important to study the existence of such a mechanism, and if it does exist, determine the kinetic properties that regulate that transport.

BIBLIOGRAPHY

[1] Wenhua Tong, Ying Xu, Mo Xian, Wei Niu, Jiantao Guo, Huizhou Liu, and Guang Zhao. Biosynthetic pathway for acrylic acid from glycerol in recombinant Escherichia coli. *Applied Microbiology and Biotechnology*, 100(11):4901–4907, 2016.

[2] Rolf Beerthuis, Gadi Rothenberg, and N. Raveendran Shiju. Catalytic routes towards acrylic acid, adipic acid and ϵ-caprolactam starting from biorenewables. *Green Chemistry*, 17(1):1341–1361, 2015.

[3] *Allied Market Research*. Acrylic acid market report. Available at https://www.alliedmarketresearch.com/acrylic-acid-market (2019/10/20), 2016.

[4] Hun Su Chu, Jin Ho Ahn, Jiae Yun, In Suk Choi, Tae Wook Nam, and Kwang Myung Cho. Direct fermentation route for the production of acrylic acid. *Metabolic Engineering*, 32(1):23–29, 2015.

[5] Avelino Corma Canos, Sara Iborra, and Alexandra Velty. Chemical routes for the transformation of biomass into chemicals. *Chemical Reviews*, 107(1):2411–2502, 2007.

[6] Signe Unverricht, Heiko Arnold, Andreas Tenten, Ulrich Hammon, Neumann Hans-Peter, and Harth Klaus. Patent No.: US 6,998,504.: Method for the catalytic gas phase oxidation of propene into acrylic acid, 2006.

[7] Zhijie Liu and Tiangang Liu. Production of acrylic acid and propionic acid by constructing a portion of the 3-hydroxypropionate/4-hydroxybutyrate cycle from *Metallosphaera sedula* in *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology*, 43(12):1659–1670, 2016.

[8] Betânia Lunelli, Elis Duarte, Eduardo Toledo, Maria Wolf Maciel, and Rubens Maciel Filho. A new process for acrylic acid synthesis by fermentative process. *Applied Biochemistry and Biotechnology*, 137-140(12):487–499, 2007.

[9] Olan Fruchey, Thomas Malisezewski, and John Sawyer. Wipo patent num.: Wo 2013/036389 a1.: Acrylic acid from lactibe, 2013.

[10] Bernhard Palsson. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, https://www.cambridge.org/pt/academic/subjects/life-sciences/genomics-bioinformatics-and-systems-biology/

systems-biology-properties-reconstructed-networks?format=HB&isbn=9780521859035, 2006.

[11] Jianfeng Feng, Wenjiang Fu, and Fengzhu Sun. *Systems biology: Properties of reconstructed networks*, volume 15 of *Computational Biology*. Springer, 2010.

[12] Hiroaki Kitano. Systems biology: A brief overview. *Science*, 295(5560)):1662–1664, 2002.

[13] Nicolas Le Novere. Biomodels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(1):689–691, 2006.

[14] Chen Li, Marco Donizelli, Nicolas Rodriguez, Harish Dharuri, Lukas Endler, Vijayalakshmi Chelliah, Lu Li, Enuo He, Arnaud Henry, Melanie I. Stefan, Jacky L. Snoep, Michael Hucka, Nicolas Le Novère, and Camille Laibe. Biomodels database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4(92):1–14, 2010.

[15] Vijayalakshmi Chelliah, Nick Juty, Ishan Ajmera, Raza Ali, Marine Dumousseau, Mihai Glont, Michael Hucka, Gaël Jalowicki, Sarah Keating, Vincent Knight-Schrijver, Audald Lloret-Villas, Kedar Nath Natarajan, Jean Baptiste Pettit, Nicolas Rodriguez, Michael Schubert, Sarala M. Wimalaratne, Yangyang Zhao, Henning Hermjakob, Nicolas Le Novère, and Camille Laibe. Biomodels: Ten-year anniversary. *Nucleic Acids Research*, 43(1):542–548, 2015.

[16] Minoru Kanehisa and Susumu Goto. Kegg : Kyoto encyclopedia of genes and genomes. *Nucleic Acid Research*, 28(1):27–30, 2000.

[17] Mao Tanabe and Minoru Kanehisa. Using the kegg database resource. *Current Protocols in Bioinformatics*, 38(1):1.12.1–1.12.43, 2012.

[18] Minoru Kanehisa, Yoko Sato, Miho Furumichi, Kanae Morishima, and Mao Tanabe. New approach for understanding genome variations in kegg. *Nucleic acids research*, 47(1):590–595, 2018.

[19] Ron Caspi, Richard Billington, Luciana Ferrer, Hartmut Foerster, Carol Fulcher, Ingrid Keseler, Anamika Kothari, Markus Krummenacker, Mario Latendresse, Lukas Mueller, Quang Ong, Suzanne Paley, Pallavi Subhraveti, Daniel Weaver, and Peter Karp. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Research*, 44(1):471–480, 2016.

[20] Ron Caspi, Richard Billington, Carol Fulcher, Ingrid Keseler, Anamika Kothari, Markus Krummenacker, Mario Latendresse, Peter Midford, Quang Ong, Wai Kit Ong,

Suzanne Paley, Pallavi Subhraveti, and Peter Karp. The metacyc database of metabolic pathways and enzymes. *Nucleic Acids Research*, 46(1):633–639, 2018.

[21] Sandra Placzek, Ida Schomburg, Antje Chang, Lisa Jeske, Marcus Ulbrich, Jana Tillack, and Dietmar Schomburg. Brenda in 2017: New perspectives and new tools in brenda. *Nucleic Acids Research*, 45(1):380–388, 2017.

[22] Lisa Jeske, Sandra Placzek, Ida Schomburg, Antje Chang, and Dietmar Schomburg. Brenda in 2019: a european elixir core data resource. *Nucleic acids research*, 47(1):542–549, 2019.

[23] Ulrike Wittig, Renate Kania, Martin Golebiewski, Maja Rey, Lei Shi, Lenneke Jong, Enkhjargal Algaa, Andreas Weidemann, Heidrun Sauer-Danzwith, Saqib Mir, Olga Krebs, Meik Bittkowski, Elina Wetsch, Isabel Rojas, and Wolfgang Müller. Sabio-rk - database for biochemical reaction kinetics. *Nucleic Acids Research*, 40(1):790–796, 2012.

[24] Ulrike Wittig, Renate Kania, Meik Bittkowski, Elina Wetsch, Lei Shi, Lenneke Jong, Martin Golebiewski, Maja Rey, Andreas Weidemann, Isabel Rojas, and Wolfgang Müller. Data extraction for the reaction kinetics database sabio-rk. *Perspectives in Science*, 1(1-6):33–40, 2014.

[25] Ulrike Wittig, Maja Rey, Andreas Weidemann, Renate Kania, and Wolfgang Müller. Sabio-rk: An updated resource for manually curated biochemical reaction kinetics. *Nucleic Acids Research*, 46(1):656–660, 2018.

[26] Avi Flamholz, Elad Noor, Arren Bar-Even, and Ron Milo. equilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Research*, 40(1):D770–D775, 2012.

[27] Elad Noor, Arren Bar-Even, Avi Flamholz, Yaniv Lubling, Dan Davidi, and Ron Milo. An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. *Bioinformatics*, 28(15):2037–2044, 2012.

[28] Michael Mavrovouniotis. Estimation of standard gibbs energy changes of biotransformation. *The Journal of Biological Chemistry*, 266(22):14440–14445, 1991.

[29] Matthew Jankowski, Christopher Henry, Linda Broadbelt, and Vassily Hatzimanikatis. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophysical Journal*, 95(1):1487–1499, 2008.

[30] JamesK Bailey. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology progress*, 14(1):8–20, 1998.

[31] Andreas Gombert and Jens Nielsen. Mathematical modelling of metabolism. *Current Opinion in Biotechnology*, 11(1):180–186, 2000.

[32] Satoshi Ohno, Hiroshi Shimizu, and Chikara Furusawa. Fastpros: screening of reaction knockout strategies for metabolic engineering. *Bioinformatics*, 30(7):981–987, 2013.

[33] Kiran Raosaheb Patil, Mats Åkesson, and Jens Nielsen. Use of genome-scale microbial models for metabolic engineering. *Current Opinion in Biotechnology*, 15(1):64–69, 2004.

[34] Michael Hucka, Andrew Finney, Herbert Sauro, Hamid Bolouri, John Doyle, Hiroaki Kitano, Adam Arkin, Benjamin Bornstein, Dennis Bray, Athel Cornish-Bowden, et al. The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.

[35] Frédéric Achard, Guy Vaysseix, and Emmanuel Barillot. Xml, bioinformatics and data integration. *Bioinformatics*, 17(2):115–125, 2001.

[36] Joachim Almquist, Marija Cvijovic, Vassily Hatzimanikatis, Jens Nielsen, and Mats Jirstrand. Kinetic models in industrial biotechnology - improving cell factory performance. *Metabolic Engineering*, 24(1):38–60, 2014.

[37] Peter Waage and Cato Guldberg. Studies concerning affinity. *Journal of Chemical Education*, 63(12):1044–1047, 1986.

[38] Vijaysekhar Chellaboina, Sanjay Bhat, Wassim Haddad, and Dennis Bernstein. Modeling and analysis of mass-action kinetics. *IEEE Control Systems*, 29(4):60–78, 2009.

[39] Tamas Turanyi and Alison Tomlin. *Analysis of Kinetic Reaction Mechanisms*. Springer, 1st edition, 2014.

[40] Jeremy Berg, John Tymoczko, and Lubert Stryer. *Biochemistry*. W.H. Freeman, 5th edition, 2002.

[41] Tamas Keleti. Two rules of enzyme kinetics for reversible michaelis-menten mechanisms. *FEBS Letters*, 208(1):109–112, 1986.

[42] Nuriye Ulusu. Evolution of enzyme kinetic mechanisms. *Journal of Molecular Evolution*, 80(1):251–257, 2015.

[43] Robert Roskoski. Michaelis-menten kinetics. In *xPharm: The Comprehensive Pharmacology Reference*, chapter 5143. Elsevier, 2008.

[44] Irwin Segel. *Enzyme kinetics: behavior and analysis of rapid equilibrium and steady state enzyme systems*. Wiley classics library. Wiley, 1975.

[45] Henry Jakubowski. Biochemistry online: an approach based on chemical logic. Available at https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Online_(Jakubowski) (2019/8/13), 2019.

[46] Christophe Chassagnole, Naruemol Noisommit-Rizzi, Joachim Schmid, Klaus Mauch, and Matthias Reuss. Dynamic modeling of the central carbon metabolism of escherichia coli. *Biotechnology and Bioengineering*, 79(1):53–73, 2002.

[47] Kirill Peskov, Ekaterina Mogilevskaya, and Oleg Demin. Kinetic modelling of central carbon metabolism in escherichia coli. *FEBS Journal*, 279(1):3374–3385, 2012.

[48] Ali Khodayari and Costas Maranas. A genome-scale escherichia coli kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nature communications*, 7(13806), 2016.

[49] Pierre Millard, Kieran Smallbone, and Pedro Mendes. Metabolic regulation is sufficient for global and robust coordination of glucose uptake, catabolism, energy production and growth in *Escherichia coli*. *PLoS computational biology*, 13(2):1–24, 2017.

[50] Nusrat Jahan, Kazuhiro Maeda, Yu Matsuoka, Yurie Sugimoto, and Hiroyuki Kurata. Development of an accurate kinetic model for the central carbon metabolism of escherichia coli. *Microbial Cell Factories*, 15(1):112, 2016.

[51] Yu Matsuoka and Hiroyuki Kurata. Modeling and simulation of the redox regulation of the metabolism in escherichia coli at different oxygen concentrations. *Biotechnology for biofuels*, 10(1):183, 2017.

[52] Frank Bergmann, Stefan Hoops, Brian Klahn, Ursula Kummer, Pedro Mendes, Jürgen Pahle, and Sven Sahle. Copasi and its applications in biotechnology. *Journal of Biotechnology*, 261(1):215–220, 2017.

[53] Stefan Hoops, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. Copasi - a complex pathway simulator. *Bioinformatics*, 22(24):3067–3074, 2006.

[54] Dagmar Waltemath, Richard Adams, Frank Bergmann, Michael Hucka, Fedor Kolpakov, Andrew Miller, Ion Moraru, David Nickerson, Sven Sahle, Jacky Snoep, and Nicolas Le Novère. Reproducible computational biology experiments with sed-ml - the simulation experiment description markup language. *BMC Systems Biology*, 5(1):198, 2011.

[55] Linda Petzold. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM Journal on Scientific and Statistical Computing*, 4(1):136–148, 1983.

[56] Jürgen Pahle. Biochemical simulations: Stochastic, approximate stochastic and hybrid approaches. *Briefings in Bioinformatics*, 10(1):53–64, 2009.

[57] Daniel Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(1):403–434, 1976.

[58] Michael Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A*, 104(1):1876–1889, 2000.

[59] Yang Cao, Daniel Gillespie, and Linda Petzold. Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*, 124(1):044109, 2006.

[60] David Fell. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical Journal*, 286(2):313–330, 1992.

[61] Jens Nielsen. Metabolic engineering: techniques for analysis of targets for genetic manipulations. *Biotechnology and bioengineering*, 58(2-3):125–132, 1998.

[62] Diana Visser and Joseph Heijnen. The mathematics of metabolic control analysis revisited. *Metabolic Engineering*, 4(2):114–123, 2002.

[63] Robert Hooke and T. A. Jeeves. " direct search" solution of numerical and statistical problems. *Journal of the ACM*, 8(2):212–229, 1961.

[64] Kenneth Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.

[65] John Nelder and Roger Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.

[66] Richard Brent. *Algorithms for minimization without derivatives*. Prentice Hall Englewood Cliffs, NJ, 3rd edition, 2013.

[67] Rainer Storn and Kenneth Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous space. *Journal of Global Optimization*, 11(1):341–359, 1997.

[68] Thomas Runarsson and Xin Yao. Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation*, 4(3):284–294, 2000.

[69] David Fogel, Lawrence Fogel, and Wirt Atmar. Meta-evolutionary programming. *Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems & Computers*, 1(1):540–545, 1991.

[70] Zbigniew Michalewicz. *Genetic algorithms + data structures = evolution programs*. Springer Verlag, 3rd edition, 1996.

[71] James Kennedy and Russell Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948, Nov 1995.

[72] Jose Egea, Eva Balsa-Canto, María García, and Julio Banga. Dynamic optimization of nonlinear processes with an enhanced scatter search method. *Industrial and Engineering Chemistry Research*, 48(1):4388–4401, 2009.

[73] Angelo Corana, Michele Marchesi, Claudio Martini, and Sandro Ridella. Corrigenda: "minimizing multimodal functions of continuous variables with the 'simulated annealing' algorithm". *ACM Transactions on Mathematical Software*, 13(3):262–280, 1989.

[74] Geoffrey Cooper and Robert Hausman. *The Cell: A Molecular Approach*. W. H. Freeman, 6th edition edition, 2013.

[75] Zachary Blount. The natural history of model organisms: The unexhausted potential of *E. coli*. *eLife*, 4(1):e05826, 2015.

[76] Heloisa Nascimento, Lucas Silva, Renata Souza, Neusa Silva, and Isabel Scaletsky. Phenotypic and genotypic characteristics associated with biofilm formation in clinical isolates of atypical enteropathogenic *Escherichia coli* (aepec) strains. *BMC Microbiology*, 14(1):184, 2014.

[77] Stanford Shulman, Herbert Friedmann, and Ronald Sims. Theodor escherich: The first pediatric infectious diseases physician? *Clinical Infectious Diseases*, 45(8):1025–1029, 2007.

[78] Upinder Fotadar, Philip Zaveloff, and Louis Terracio. Growth of *Escherichia coli* at elevated temperatures. *Journal of Basic Microbiology*, 45(5):403– 404, 2005.

[79] Mark Sannes, Michael Kuskowski, Krista Owens, Abby Gajewski, and James Johnson. Virulence factor profiles and phylogenetic background of *Escherichia coli* isolates from veterans with bacteremia and uninfected control subjects. *The Journal of Infectious Diseases*, 190(12):2121–2128, 2004.

[80] Jordi Vila, Emma Sáez-López, J. Johnson, Ute Römling, Ulrich Dobrindt, Rafael Cantón, Christian Giske, Thierry Naas, Alessandra Carattoli, Margarita Martínez-Medina, Jordi Bosch, Pilar Retamar, Jesùs Rodríguez-Banõ, Fernando Baquero, and Sara Soto. *Escherichia coli*: An old friend with new tidings. *FEMS Microbiology Reviews*, 40(1):437–463, 2016.

[81] Michael Savageau. *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *The American Naturalist*, 122(6):732–744, 1983.

[82] Carl Zimmer. *Microcosm: E. Coli and the New Science of Life*. Vintage Series. Vintage Books, 2009.

[83] Sammy Pontrelli, Tsan Chiu, Ethan Lan, Frederic Chen, Peiching Chang, and James Liao. *Escherichia coli* as a host for metabolic engineering. *Metabolic Engineering*, 50(1):16–46, 2018.

[84] Hans-Peter Meyer and Diego Schmidhalter. Microbial expression systems and manufacturing from a market and economic perspective. In Eddy C. Agbo, editor, *Innovations in Biotechnology*, chapter 10. IntechOpen, Rijeka, 2012.

[85] Fang Tao, Ji Miao, Gui Shi, and Kechang Zhang. Ethanol fermentation by an acid-tolerant *Zymomonas mobilis* under non-sterilized condition. *Process Biochemistry*, 40(1):183–187, 2005.

[86] David Wernick, Sammy Pontrelli, Alexander Pollock, and James Liao. Sustainable biorefining in wastewater by engineered extreme alkaliphile *Bacillus marmarensis*. *Scientific Reports*, 6(1):20224, 2016.

[87] Aditya Bhalla, Namita Bansal, Sudhir Kumar, Kenneth Bischoff, and Rajesh Sani. Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresource Technology*, 128(1):751–759, 2013.

[88] Alessandra de Melo, Sébastien Levesque, and Sylvain Moineau. Phages as friends and enemies in food processing. *Current Opinion in Biotechnology*, 49(1):185–190, 2018.

[89] NCBI. Genome assembly and annotation report. Available at https://www.ncbi.nlm.nih.gov/genome/genomes/167 (2018/11/04)".

[90] David Nelson and Michael Cox. *Lehninger Principles of Biochemistry*. W.H.Freeman and Company, 6th edition edition, 2005.

[91] Michael Lynch, Ryan Gill, and Tanya Lipscomb. Us 2017/0114377 a1: Method for producing 3-hydroxypropionic acid and other products, 2017.

[92] Akiko Suyama, Yujiro Higuchi, Masahiro Urushihara, Yuka Maeda, and Kaoru Takegawa. Production of 3-hydroxypropionic acid via the malonyl-coa pathway using recombinant fission yeast strains. *Journal of Bioscience and Bioengineering*, 124(4):392–399, 2017.

[93] Chan Woo Song, Je Woong Kim, In Jin Cho, and Sang Yup Lee. Metabolic engineering of *Escherichia coli* for the production of 3-hydroxypropionic acid and malonic acid through β-alanine route. *ACS Synthetic Biology*, 5(11):1256–1263, 2016.

[94] Changshui Liu, Yamei Ding, Rubing Zhang, Huizhou Liu, Mo Xian, and Guang Zhao. Functional balance between enzymes in malonyl-coa pathway for 3-hydroxypropionate biosynthesis. *Metabolic engineering*, 34(1):104–111, 2016.

[95] Irina Borodina, Kanchana Kildegaard, Niels Jensen, Thomas Blicher, JérÔme Maury, Svetlana Sherstyk, Konstantin Schneider, Pedro Lamosa, Markus Herrgård, Inger Rosenstand, Fredrik Öberg, Jochen Forster, and Jens Nielsen. Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via β-alanine. *Metabolic Engineering*, 27(1):57–64, 2015.

[96] Changshui Liu, Yamei Ding, Mo Xian, Min Liu, Huizhou Liu, Qingjun Ma, and Guang Zhao. Malonyl-coa pathway: a promising route for 3-hydroxypropionate biosynthesis. *Critical reviews in biotechnology*, 37(7):933–941, 2017.

[97] Chelladurai Rathnasingh, Subramanian Mohan Raj, Youjin Lee, Christy Catherine, Somasundar Ashok, and Sunghoon Park. Production of 3-hydroxypropionic acid via malonyl-coa pathway using recombinant *Escherichia coli* strains. *Journal of biotechnology*, 157(1):633–640, 2011.

[98] Subramanian Raj, Chelladurai Rathnasingh, Ji-Eun Jo, and Sunghoon Park. Production of 3-hydroxypropionic acid from glycerol by a novel recombinant *Escherichia coli* bl21 strain. *Process Biochemistry*, 43(12):1440–1446, 2008.

[99] Somasundar Ashok, Mugesh Sankaranarayanan, Yeounjoo Ko, Kyeung Eun Jae, Satish Kumar Ainala, Vinod Kumar, and Sunghoon Park. Production of 3-hydroxypropionic acid from glycerol by recombinant *Klebsiella pneumoniae* δdhatδyqhd which can produce vitamin b12 naturally. *Biotechnology and Bioengineering*, 110(2):511–524, 2013.

[100] Michael Hügler, Castor Menendez, Hermann Schägger, and Georg Fuchs. Malonyl-coenzyme a reductase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic co_2 fixation. *Journal of bacteriology*, 184(9):2404–2410, 2002.

[101] Vinod Kumar, Somasundar Ashok, and Sunghoon Park. Recent advances in biological production of 3-hydroxypropionic acid. *Biotechnology Advances*, 31(1):945–961, 2013.

[102] Chan Woo Song, Joungmin Lee, Yoo-Sung Ko, and Sang Yup Lee. Metabolic engineering of *Escherichia coli* for the production of 3-aminopropionic acid. *Metabolic engineering*, 30(1):121–129, 2015.

[103] Subramanian Mohan Raj, Chelladurai Rathnasingh, Woo-Chel Jung, and Sunghoon Park. Effect of process parameters on 3-hydroxypropionic acid production from

glycerol using a recombinant *Escherichia coli*. *Applied microbiology and biotechnology*, 84(4):649–657, 2009.

[104] Chelladurai Rathnasingh, Subramanian Mohan Raj, Ji-Eun Jo, and Sunghoon Park. Development and evaluation of efficient recombinant *Escherichia coli* strains for the production of 3-hydroxypropionic acid from glycerol. *Biotechnology and bioengineering*, 104(4):729–739, 2009.

[105] Zhuan Cheng, Jiaqi Jiang, Hui Wu, Zhimin Li, and Qin Ye. Enhanced production of 3-hydroxypropionic acid from glucose via malonyl-coa pathway by engineered *Escherichia coli*. *Bioresource technology*, 200(1):897–904, 2016.

[106] Jonathan Monk, Colton Lloyd, Elizabeth Brunk, Nathan Mih, Anand Sastry, Zachary King, Rikiya Takeuchi, Wataru Nomura, Zhen Zhang, Hirotada Mori, Adam Feist, and Bernhard Palsson. iml1515, a knowledgebase that computes *Escherichia coli* traits. *Nature Biotechnology*, 35(10):904–908, 2017.

[107] Daniel Machado, Lígia Rodrigues, and Isabel Rocha. A kinetic model for curcumin production in *Escherichia coli*. *BioSystems*, 125(1):16–21, 2014.

[108] Ralph Edgar and Robert Bell. Biosynthesis in escherichia coli of m-glycerol 3-phosphate, a precursor of phospholipid. *Journal of Biological Chemistry*, 253(18):6354–6363, 1978.

[109] Gerd Kloeck and Karlheinz Kreuzberg. Kinetic properties of asn-glycerol-3-phosphate dehydrogenase purified from the unicellular alga *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta*, 991(2):347–352, 1989.

[110] Isabelle Salles, Nynne Forchhammer, Christian Croux, Laurence Girbal, and Philippe Soucaille. Evolution of a *Saccharomyces cerevisiae* metabolic pathway in *Escherichia coli*. *Metabolic Engineering*, 9(2):152–159, 2007.

[111] Donald Pettigrew, Gui Yu, and Youguo Liu. Nucleotide regulation of *Escherichia coli* glycerol kinase: Initial-velocity and substrate binding studies. *Biochemistry*, 29(1):8620–8627, 1990.

[112] Aileen Soriano, Anthony Radice, Amy Herbitter, Erik Langsdorf, Jill Stafford, Sarah Chan, Shihong Wang, Yan-Hui Liu, and Todd Black. *Escherichia coli* acetyl-coenzyme a carboxylase: characterization and development of a high-throughput assay. *Analytical biochemistry*, 349(2):268–276, 2005.

[113] Christoph Freiberg, Nina Brunner, Guido Schiffer, Thomas Lampe, Jens Pohlmann, Michael Brands, Martin Raabe, Dieter Häbich, and Karl Ziegelbauer. Identification

and characterization of the first class of potent bacterial acetyl-coa carboxylase inhibitors with antibacterial activity. *Journal of Biological Chemistry*, 279(25):26066–26073, 2004.

[114] Michael Sharkey and Paul Engel. Apparent negative co-operativity and substrate inhibition in overexpressed glutamate dehydrogenase from *Escherichia coli*. *FEMS microbiology letters*, 281(2):132–139, 2008.

[115] Toshiharu Yagi, Hiroyuki Kagamiyama, Mitsuhiro Nozaki, and Kenji Soda. Glutamate-aspartate transaminase from microorganisms. *Methods in enzymology*, 113(1):83–89, 1985.

[116] Fulcher CA and SRI International. Aspartate aminotransferase. Available at https://biocyc.org/gene?orgid=META&id=YLR027C-MONOMER#tab=RXNS (2019/05/29), 2007.

[117] David Hames and Nigel Hooper. *Biochemistry*. BIOS instant notes. Taylor & Francis, third edition, 2005.

[118] Manoj Ramjee, Ulrich Genschel, Chris Abell, and Alison Smith. *Escherichia coli* l-aspartate-α-decarboxylase: preprotein processing and observation of reaction intermediates by electrospray mass spectrometry. *Biochemical Journal*, 323(1):661–669, 1997.

[119] Claudia Piattoni, Carlos Figueroa, Matías Diez, Ivana Parcerisa, Sebastián Antuña, Raúl Comelli, Sergio Guerrero, Alejandro Beccaria, and Alberto Iglesias. Production and characterization of *Escherichia coli* glycerol dehydrogenase as a tool for glycerol recycling. *Process Biochemistry*, 48(3):406–412, 2013.

[120] Nicolas Sauvageot, Vianney Pichereau, Loïc Louarme, Axel Hartke, Yanick Auffray, and Jean-Marie Laplace. Purification, characterization and subunits identification of the diol dehydratase of *Lactobacillus collinoides*. *European journal of biochemistry*, 269(22):5731–5737, 2002.

[121] Ji-Eun Jo, Subramanian Raj, Chelladurai Rathnasingh, Edwardraja Selvakumar, Woo-Chel Jung, and Sunghoon Park. Cloning, expression, and characterization of an aldehyde dehydrogenase from *Escherichia coli* k-12 that utilizes 3-hydroxypropionaldehyde as a substrate. *Applied microbiology and biotechnology*, 81(1):51, 2008.

[122] Daniel Kockelkorn and Georg Fuchs. Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme a reductase from *Metallosphaera sedula*: enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in *Sulfolobales*. *Journal of bacteriology*, 191(20):6352–6362, 2009.

[123] Wenshe Liu, Peter Peterson, James Langston, Xueguang Jin, Xianzhi Zhou, Andrew Fisher, and Michael Toney. Kinetic and crystallographic analysis of active site mutants of *Escherichia coli* γ-aminobutyrate aminotransferase. *Biochemistry*, 44(8):2982–2992, 2005.

[124] Birgit Alber and Georg Fuchs. Propionyl-coenzyme a synthase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic co2 fixation. *Journal of Biological Chemistry*, 277(14):12137–12143, 2002.

[125] Robin Teufel, Johannes Kung, Daniel Kockelkorn, Birgit Alber, and Georg Fuchs. 3-hydroxypropionyl-coenzyme a dehydratase and acryloyl-coenzyme a reductase, enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in the *Sulfolobales*. *Journal of bacteriology*, 191(14):4572–4581, 2009.

[126] Zhihao Zhuang, Feng Song, Hong Zhao, Ling Li, Jian Cao, Edward Eisenstein, Osnat Herzberg, and Debra Dunaway-Mariano. Divergence of function in the hot dog fold enzyme superfamily: the bacterial thioesterase ycia. *Biochemistry*, 47(9):2789–2796, 2008.

[127] Pedro Mendes, Stephen Oliver, and Douglas Kell. Fitting transporter activities to cellular drug concentrations and fluxes: why the bumblebee can fly. *Trends in pharmacological sciences*, 36(11):710–723, 2015.

[128] COPASI. Random search. Available at https://http://copasi.org/Support/User_Manual/Methods/Optimization_Methods/Random_Search/ (2019/08/21).

[129] Hongfang Zhang, George Lountos, Chi Bun Ching, and Rongrong Jiang. Engineering of glycerol dehydrogenase for improved activity towards 1, 3-butanediol. *Applied microbiology and biotechnology*, 88(1):117–124, 2010.

[130] Wenjuan Zha, Sheryl Rubin-Pitel, Zengyi Shao, and Huimin Zhao. Improving cellular malonyl-coa level in escherichia coli via metabolic engineering. *Metabolic engineering*, 11(3):192–198, 2009.

A

SUPPORTING MATERIAL

A.1 DEVELOPED MODELS

Table A.1.: Kinetic models developed to achieve 3-hydroxypropionate (3HP) and acrylic acid (AA) from either glucose or glycerol. All the models presented in this table are available at https://nextcloud.bio.di.uminho.pt/s/g6y8PjszaYaQr4j.

Model ID	Pathway	Carbon Source	End Product	Metabolites	Reactions
3HP_GlyPath_GLC	Glycerol	Glucose	3-HP	91	89
3HP_GlyPath_GLY	Glycerol	Glucose	AA	94	93
AA_GlyPath_GLC	Glycerol	Glycerol	3-HP	94	92
AA_GlyPath_GLY	Glycerol	Glycerol	AA	97	96
3HP_McoaPath_GLC	Malonyl-CoA	Glucose	3-HP	90	89
3HP_McoaPath_GLY	Malonyl-CoA	Glucose	AA	91	91
AA_McoaPath_GLC	Malonyl-CoA	Glycerol	3-HP	93	92
AA_McoaPath_GLY	Malonyl-CoA	Glycerol	AA	94	94
3HP_BalaPath_GLC	β-alanine	Glucose	3-HP	90	89
3HP_BalaPath_GLY	β-alanine	Glucose	AA	91	91
AA_BalaPath_GLC	β-alanine	Glycerol	3-HP	93	92
AA_BalaPath_GLY	β-alanine	Glycerol	AA	94	94

A.2 GLYCEROL MODEL

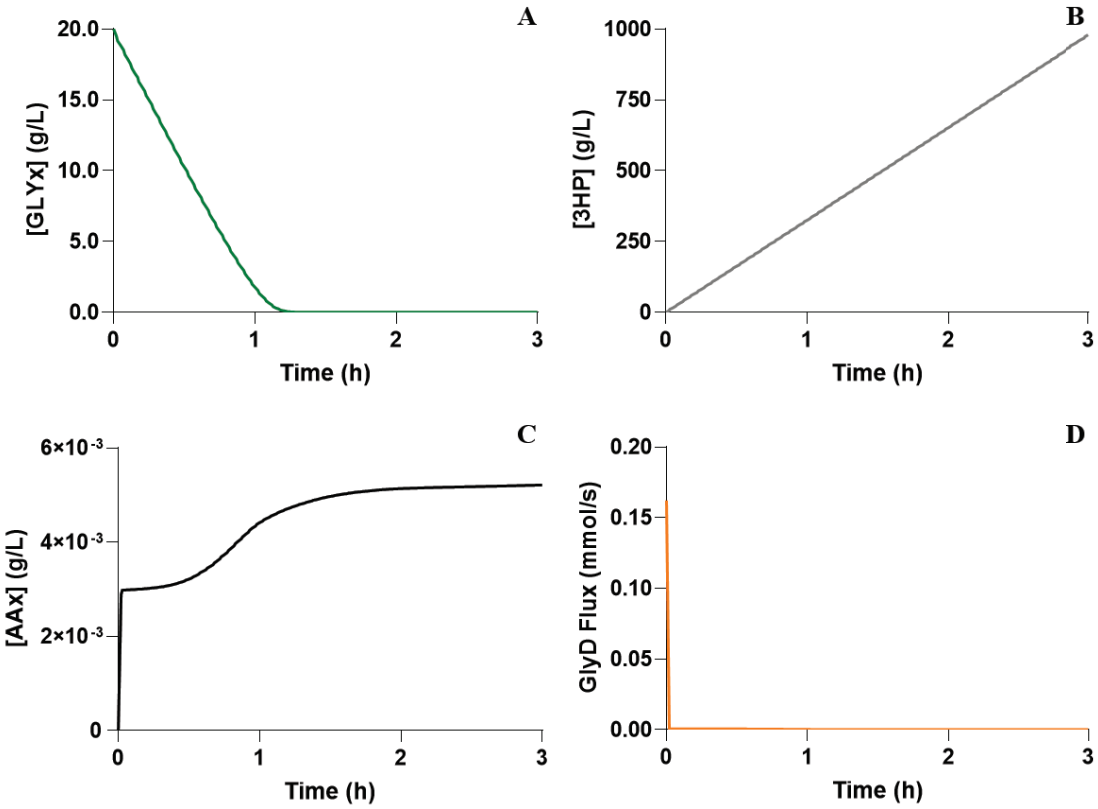


Figure A.1.: Results of the time course simulations concerning glycerol (GLY) consumption (A), and the production of 3-hydroxypropionate (3HP) (B) and acrylic acid (AA) (C), and flux of the glycerol dehydrogenase (GlyD) (D), in the original AA_GlyPath_GLY model.

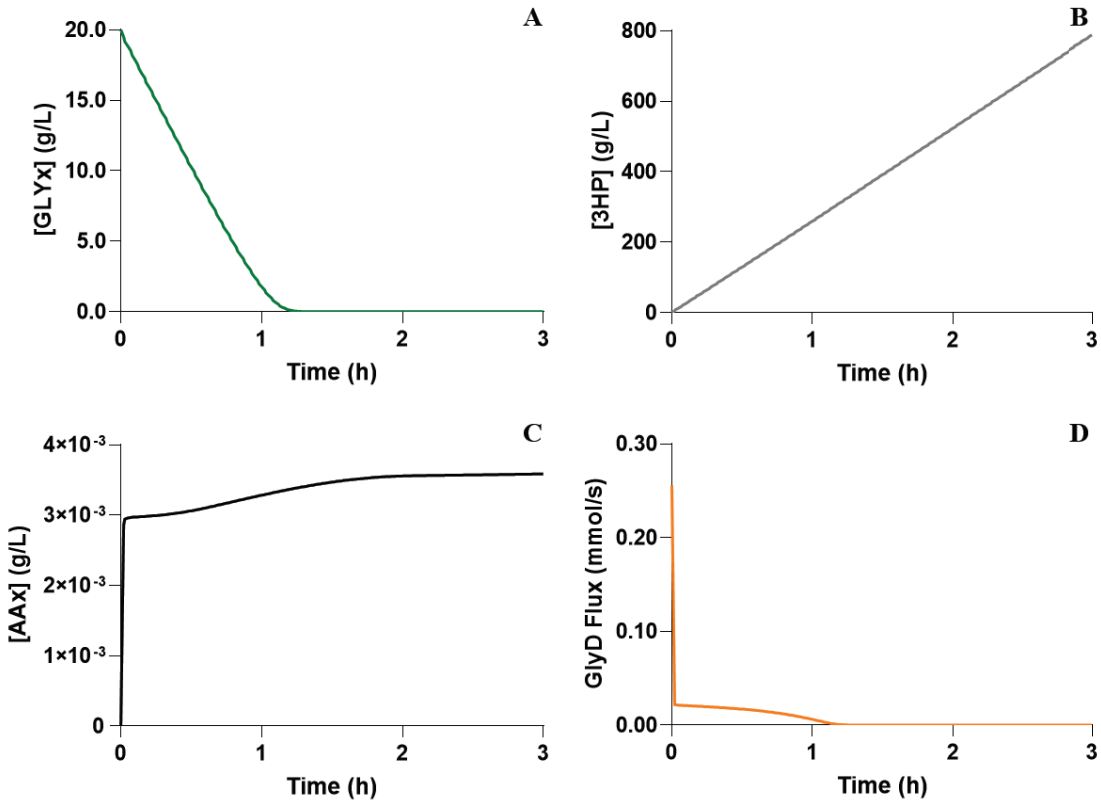


Figure A.2.: Results of the time course simulations concerning glycerol (GLY) consumption (A), and the production of 3-hydroxypropionate (3HP) (B) and acrylic acid (AA) (C), and flux of the glycerol dehydrogenase (GlyD) (D), considering the AA_GlyPath_GLY model after the affinity towards NAD of the GlyD was changed to 0.0165 mM.

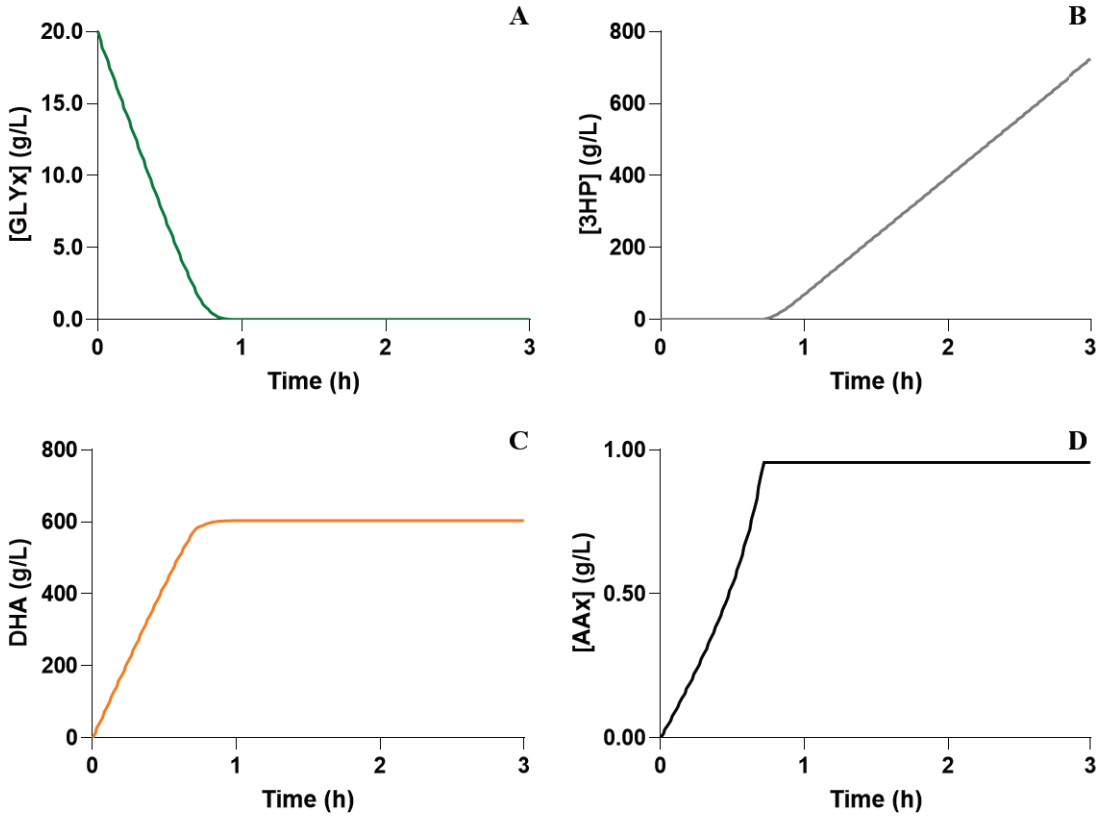


Figure A.3.: Results of the time course simulations concerning glycerol (GLY) consumption (A), and the production of 3-hydroxypropionate (3HP) (B), dihydroxyacetone (DHA) (C), and acrylic acid (AA) (D), considering the AA_GlyPath_GLY model after the V_{max} of there-action glycerol dehydrogenase was changed to 4298.4 mM/s.

A.3 MALONYL-COA MODEL

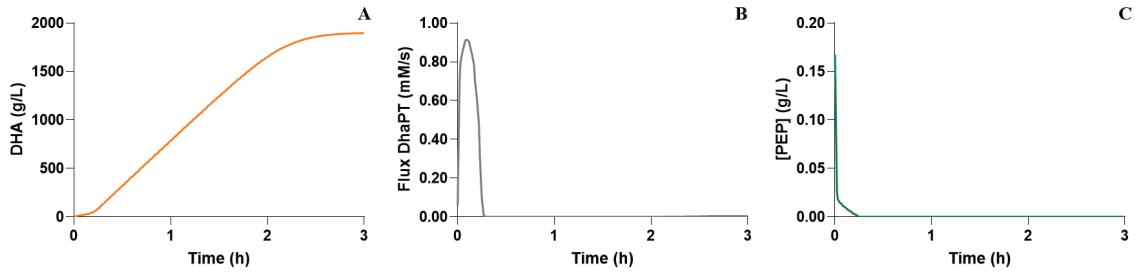


Figure A.4.: Results of the time course simulations obtained with the AA_McoaPath_GLY model. This model presented issues in the assimilation of glycerol, evidenced by the accumulation of dihydroxyacetone (DHA) (A). This accumulation is caused by the low activity of the dihydroxyacetone phosphate transferase (DhaPT) (B), which in turn is caused by the low availability of phosphoenolpyruvate (PEP) in the system (C).

A.4 β-ALANINE MODEL

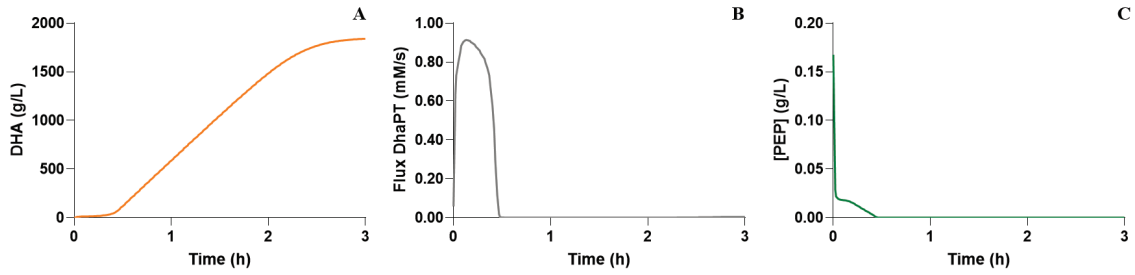


Figure A.5.: Results of the time course simulations obtained with the AA_BalaPath_GLY model. This model presented issues in the assimilation of glycerol, evidenced by the accumulation of dihydroxyacetone (DHA) (A). This accumulation is caused by the low activity of the dihydroxyacetone phosphate transferase (DhaPT) (B), which in turn is caused by the low availability of phosphoenolpyruvate (PEP) in the system (C).

A.5 OPTIMIZATION STRATEGIES

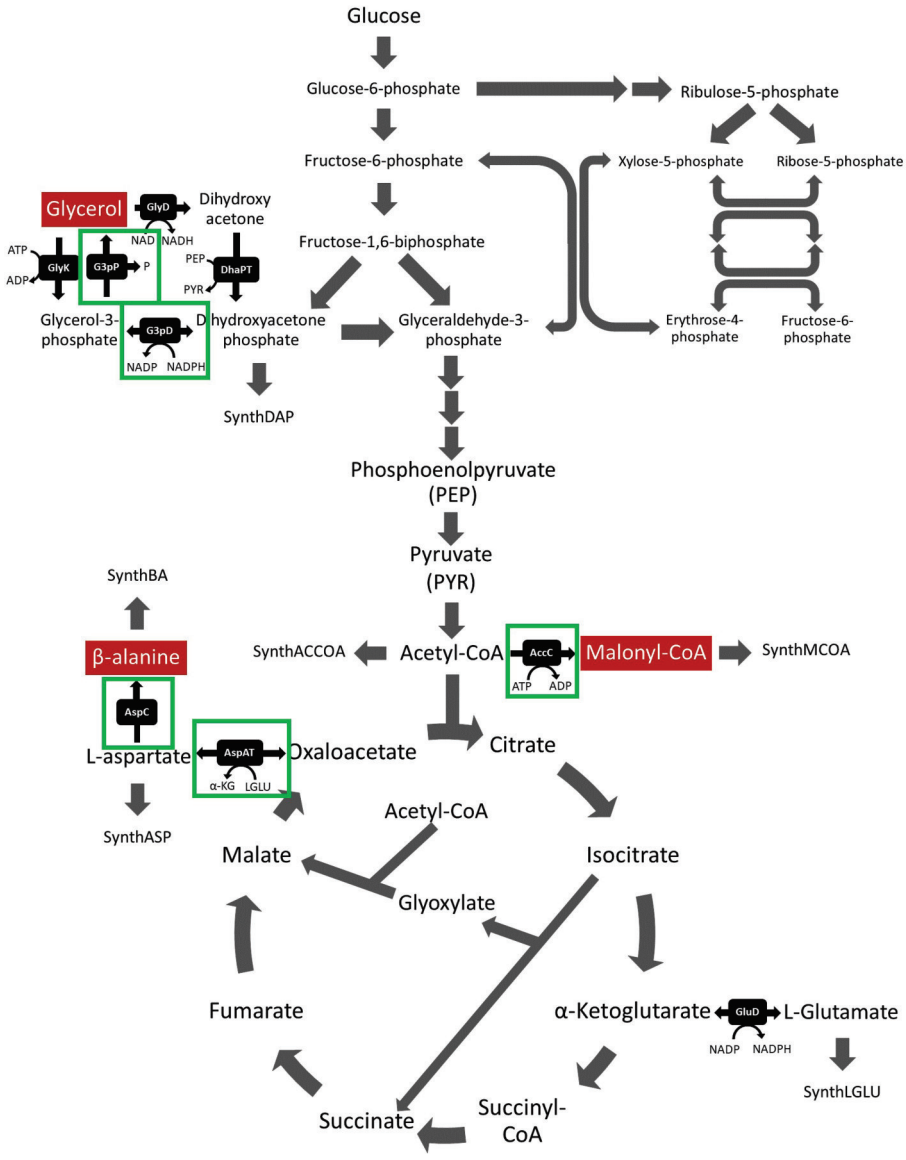


Figure A.6.: Schematic representation of the central carbon metabolism of *Escherichia coli*, where the reactions that were identified as targets for optimization are highlighted in green. For the glycerol pathway, the production of acrylic acid from glucose is limited by the glycerol-3-phosphate dehydrogenase (G3pD) and glycerol-3-phosphate phosphatase (G3pP), and the results suggested an over-expression for both reactions. Moreover, for the malonyl-CoA pathway, the acetyl-CoA carboxylase (AccC) is a bottleneck that limits the flux of the subsequent reaction, and, therefore, they should be over-expressed. Finally, in the β-alanine route, two more reactions were suggested for over-expression, the aspartate aminotransferase (AspAT) and the aspartate carboxylase (AspC).