THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Engineering synthetic pathways for adipic acid biosynthesis

Jae Ho Shin

Department of Biology and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2021

Engineering synthetic pathways for adipic acid biosynthesis
JAE HO SHIN
ISBN 978-91-7905-419-9

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4886
ISSN 0346-718X

Division of Industrial Biotechnology
Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
Various techniques and aspects involved in bio-based production of adipic acid. Codon-optimization, molecular cloning, DNA sequence analysis, and analytical chemistry are depicted.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2021

Engineering synthetic pathways for adipic acid biosynthesis

JAE HO SHIN

Division of Industrial Biotechnology
Department of Biology and Biological Engineering
Chalmers University of Technology

ABSTRACT

Utilization of petroleum in consumer product manufacturing is causing irreversible environmental damage. Its impact on land, sea, and air calls for the development of more sustainable technologies based on the use of renewable materials such as lignocellulosic biomass and its conversion into platform chemicals. Engineering microorganisms to produce chemicals is an important undertaking to address such issues and bio-based production of adipic acid especially has gained recent attention. In the present thesis I assess the *in vivo* and *in silico* action of enzymes involved in microbial production of adipic acid from simple sugar molecules. The aim of this work was to comprehensively map out the metabolic pathways leading to adipic acid biosynthesis and to investigate the enzymatic components of the L-lysine pathway, the reverse β-oxidation pathway, and *cis*,*cis*-muconic acid reduction.

Investigation of theoretical and *in silico* aspects in the deamination step in the L-lysine pathway revealed deamination of L-lysine to be chemically difficult to occur. Removal of the β-amino group from β-D-lysine was deemed more feasible than the α-amino group from L-lysine, and an alternative route *via* β-D-lysine deamination was suggested. Homology modeling and molecular docking studies shed light on the substrate binding mechanisms of enzymes responsible for the reduction of the intermediates in the L-lysine pathway. Potential mechanism and feasibility of α,β-reduction were explained in terms of substrate interaction in the enzyme-binding pockets. *Corynebacterium glutamicum* was chosen as the host chassis for achieving adipic acid synthesis *via* reverse β-oxidation. Stepwise construction of a five-step synthetic pathway demonstrated functionality of each step in *C. glutamicum*. Biosynthesized and secreted 3-hydroxyadipate was detected in the cultivation broth using GC/MS. Weak *trans*-2-hexenedioic acid and adipic acid signals were observed using LC/MS after concentrating the cultivation broth. Dehydration of 3-hydroxyadipyl-CoA was identified as a potential bottleneck in the pathway. While implementing the reverse β-oxidation pathway, a new pathway involving *cis*,*cis*-muconic acid and 3-oxoadipic acid was identified and experimented on. The modified strategy for bio-conversion of benzoic acid to *cis*,*cis*-muconic acid was successful and molecular docking studies were carried out to better understand how oxidoreductases might reduce *cis,cis*-muconic acid.

Taking multiple approaches to generate adipic acid revealed different challenges in each pathway. One approach led to biosynthesis of adipic acid.

Keywords: *Corynebacterium glutamicum*, reverse β-oxidation, α,β-unsaturated bonds, β-D-lysine, enoate reductase, *cis,cis*-muconic acid, non-decarboxylative condensation.

List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. Emma Skoog*, Jae Ho Shin*, Veronica Saez-Jimenez*, Valeria Mapelli, Lisbeth Olsson, "Biobased adipic acid–The challenge of developing the production host" *Biotechnol. Adv.,* **36**, 2248-2263, (2018) *Equal contribution

II. Emma Karlsson*, Jae Ho Shin*, Gunnar Westman, Leif A. Eriksson, Lisbeth Olsson, Valeria Mapelli, "*In silico* and *in vitro* studies of the reduction of unsaturated α,β bonds of *trans*-2-hexenedioic acid and 6-amino-*trans*-2-hexenoic acid – Important steps towards biobased production of adipic acid" *PloS One*, **13**(2), e0193503, (2018) *Equal contribution

III. Veronica Saez-Jimenez, Željka Sanader Maršić, Matteo Lambrughi, Jae Ho Shin, Robin van Havere, Elena Papaleo, Lisbeth Olsson, Valeria Mapelli, "Structure-function investigation of 3-methylaspartate ammonia lyase reveals substrate molecular determinants for the deamination reaction" *Plos One*, **15**(5), e0233467, (2020)

IV. Jae Ho Shin, Aaron John Christian Andersen, Puck Achterberg, Lisbeth Olsson, "Adipic acid biosynthesis by reverse β-oxidation pathway in engineered *Corynebacterium glutamicum*" (*Manuscript in preparation*)

V. Jae Ho Shin, Clarisse Breard, Lisbeth Olsson, "Assessment of bioconversion for benzoic acid to adipic acid by engineered *Corynebacterium glutamicum*" (*Manuscript in preparation*)

Author's contributions

Author's contributions

I. Took part in conceptualization, design, and planning of the review. Wrote about direct production of adipic acid from biological feedstock.

II. Formal analysis, investigation, writing – original draft, writing – review & editing. Conceptualized, designed, planned, and performed the *in silico* work. Conceptualized and designed the theory of the alternative pathway. Wrote the *in silico* portion of the manuscript.

III. Formal analysis. Conceptualized and developed the theory regarding hypothetical α,β-elimination of β-D-lysine. Designed the β-lysine pathway based on the theory of α,β-elimination.

IV. Conceptualized, designed, planned, and performed most of the cloning, sequencing, cultivation, and chemical analysis. Carried out preliminary studies on sample purification and derivatization. Wrote the manuscript.

V. Conceptualized, designed, planned, and performed most of the molecular docking, cloning, sequencing, cultivation, and chemical analysis experiments. Initiated and optimized analytical methods for benzoic acid and *cis,cis*-muconic acid measurements. Wrote the manuscript.

List of Contents

"...And why do we fall, Bruce?

So we can learn to pick ourselves up."

Thomas Wayne

Preface

This doctoral thesis fulfils the requirements for a PhD degree at the Department of Biology and Biological Engineering, Division of Industrial Biotechnology, Chalmers University of Technology, Sweden. The work presented in this thesis was performed between 2016 and 2021 and was funded by the Swedish Research Council (grant no. 2016-03344) and Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (grant no. 2013-78). The aim of this work was to evaluate different metabolic strategies for adipic acid production from renewable materials. The majority of the work was carried out at the Division of Industrial Biotechnology, Chalmers University of Technology, under the supervision of Professor Lisbeth Olsson. Valeria Mapelli and Professor Leif Eriksson co-supervised part of work.

Many collaborative relationships were built during this thesis work. The *in silico* studies on the enzymes NemA and Oye1 were performed under the guidance of Professor Leif Eriksson, Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden. *In vitro* characterization of NemA and Oye1 was carried out by Emma Skoog, a previous Ph.D. student at the Division of Industrial Biotechnology. *In vitro* characterization of the deamination reaction was carried out by Veronica Saez-Jimenez at the Division of Industrial Biotechnology. Metabolite analysis using GC/MS was performed at the Chalmers Mass Spectrometry Infrastructure, Chalmers University of Technology, with input from Otto Savolainen and Jacob Kindbom. *In silico* studies on enoate reductase were performed in collaboration with Clarisse Breard of the Paris Institute of Technology for Life, Food and Environmental Sciences, Paris, France for her master's thesis studies. Metabolite analysis using LC/MS was performed by Aaron John Christian Andersen at the Metabolomics Core, Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark.

Jae Ho Shin

February, 2021

1. Introduction

Since the first genetically modified organism was reported in 1973 (Cohen *et al*. 1973; Cohen and Chang 1973), biotechnology has emerged as one of the most important tools to improve quality of life over the last half century. Owing to the power of recombinant DNA technology, production of the world's first synthetic human insulin (Goeddel *et al*. 1979b) and the first recombinant drug, human growth hormone (Goeddel *et al*. 1979a), followed soon after. Over the next 50 years, the rapid development of biotechnology has not only allowed the establishment of routines to produce recombinant proteins (Baneyx 1999) and antibiotics (Hopwood *et al*. 1985), but has been applied also to vaccines (Slade *et al*. 2009), agriculture (Blancquaert *et al*. 2015), transgenic crops (Paine *et al*. 2005), and genetically modified animals (Jaenisch 1988). Given a growing awareness of air and ocean pollution, as well as man-made climate change, biotechnology has been employed to contribute to solve existing and upcoming environmental challenges (Zechendorf 1999), and achieve a more sustainable future (Griggs *et al*. 2013). In this thesis, the power of recombinant technology was applied to help address one of the greatest issues of my generation: petroleum dependency.

One of the key contribution in addressing climate change and our reliance on fossil fuels is a continuous expansion of the catalogue of chemicals produced by genetically modified organisms (Straathof 2014), an important undertaking of industrial biotechnology (Erickson *et al*. 2012; Singh 2011). Genetic engineering of microorganisms allows facile manipulation of cell physiology and metabolism. In particular, the latter provides the opportunity to intentionally produce a desired metabolite or compound at high yield and productivity (Sauer *et al*. 2008). Depending on its chemical nature, the biologically produced compound may be compatible with existing petrochemical infrastructure and require only minor changes to downstream processes (Bozell 2010). The increased amount of omics data, along with better understanding of the genetic makeup of organisms, has facilitated the generation of more efficient and robust microorganisms (Brunk *et al*. 2016). The major motivation of my thesis work lies in the urgency to develop petroleum-independent adipic acid manufacturing.

Adipic acid, an aliphatic dicarboxylic acid, is used as a building block for fibers, polyesters, resins, and various polyamides including nylon-6,6 (Polen *et al*. 2013). Presently, industrial production of adipic acid relies on a mixture of cyclohexanone and cyclohexanol, also known as KA oil (ketone and alcohol), as feedstock (Van De Vyver and Román-Leshkov 2013). Oxidation of KA oil is achieved by employing nitric acid (HNO_3) as oxidant, which emits nitrous oxide (N_2O), a major contributor to global warming, and atmospheric pollution (Rahman *et al*. 2016; Yan *et al*. 2020). KA oil is obtained from benzene, which is derived from petroleum (Rahman *et al*. 2016). An estimated 2.6–3.3 million tons of adipic acid is produced annually in the abovementioned manner (Kruyer *et al*. 2020; Polen *et al*. 2013). An estimated 65–80% of adipic acid is used in nylon-6,6 manufacturing for textiles, carpets, tires, and automobile parts (Polen *et al*. 2013; Shimizu *et al*. 2000).

The dependence on petroleum during conventional adipic acid production and consequent negative environmental impact have sparked the search for more sustainable, bio-based alternatives. While multiple academic teams have explored and successfully carried out adipic acid biosynthesis (**Figure 1** and **paper I**), technology transfer to the industry has been less effective (Babu *et al*. 2015; Cheong *et al*. 2016; Deng *et al*. 2016; Kallscheuer *et al*. 2017a; Kallscheuer *et al*. 2017b; Li *et al*. 2020; Yu *et al*. 2014). Successful examples of bio-based adipic acid biosynthesis by academic teams include the use of glycerol, which is neither a sugar nor of lignocellulosic origin, as a substrate (Zhao *et al*. 2018). Alternative production hosts, enzymes, and pathways are already available, yet remain underdeveloped. Ideally, successful bio-based adipic acid production involves efficient enzymes, a well-characterized biosynthetic pathway, a robust host, and a readily available production process.

Initially, biological adipic acid production was preceded by the biosynthesis of *cis,cis*-muconic acid (Niu *et al*. 2002) (**Figure 2**). *cis,cis*-Muconic acid biosynthesis had been the focus of research between the 1970s and 1990s (Tsuji and Kuwahara 1977), in part as an effort to understand biodegradation of aromatic compounds (Bang *et al*. 1996; Yoshikawa *et al*. 1990). Later on, the focus shifted toward production of *cis,cis*-muconic acid from glucose rather than aromatic compounds (Niu *et al*. 2002) and various strategies were attempted (Choi *et al*. 2020; Khalil *et al*. 2020; Kruyer and Peralta-Yahya 2017; Xie *et al*. 2014). Still, very little attention has been paid to biological conversion of *cis,cis*-muconic acid to adipic acid. Following the characterization of enoate reductase (Joo *et al*. 2017), the first biological production of adipic acid *via cis,cis*-muconic acid was reported (Raj *et al*. 2018).

Research on biological production of adipic acid was then shifted to ω- and β-oxidation of fatty acid molecules (Beardslee and Picataggio 2012; Picataggio *et al*. 1991). The conversion begins with ω-oxidation of fatty acids by cytochrome P450 monooxygenase from *Candida tropicalis* ATCC 750 (Picataggio *et al*. 1992), aided by NADPH-cytochrome P450 oxidoreductase (CPR) (Sutter *et al*. 1990). Dicarboxylic acids obtained by the ω-oxidation process can be subsequently shortened by β-oxidation (**Figure 3**). With fine-tuned β-oxidation, varying lengths of dicarboxylic acids can be selectively processed down to the six-carbon backbone molecule, adipic acid (Picataggio *et al*. 1991). Selectivity of dicarboxylic acid length can be adjusted by deletion of POX4, as the latter exhibits broad substrate specificity. The

overall process yields adipic acid from various fatty acid molecules and will be discussed in more detail in **section 5.1**.

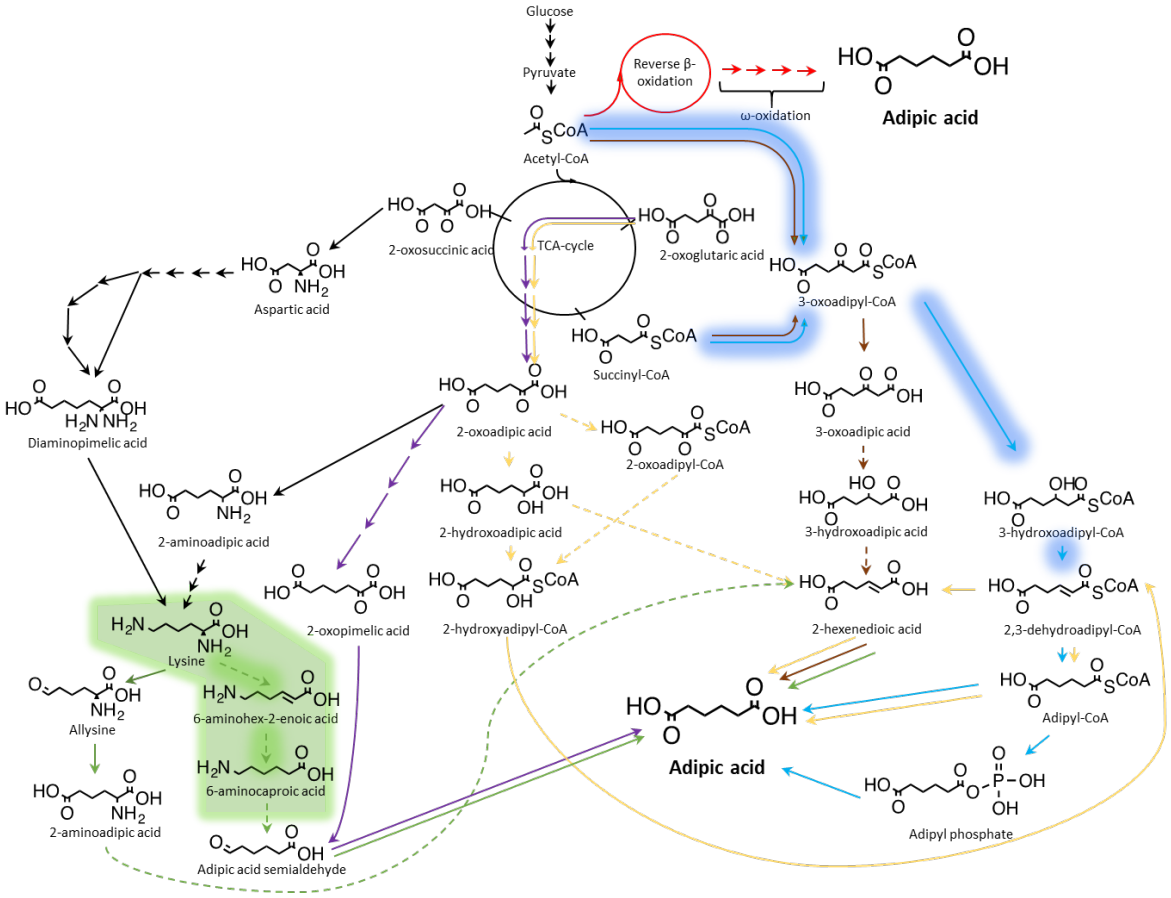


Figure 1. Overview of the metabolic pathways mediating direct production of adipic acid from glucose. Reverse adipate degradation pathway (blue lines), reverse *β*- then *ω*-oxidation pathway (red lines), 2-oxopimelic acid pathway (purple lines), 3-oxoadipic acid pathway (brown lines), 2-oxoadipic acid alternative pathway (yellow lines), lysine alternative pathway (green yellow). Enzymatic activity has been demonstrated for the steps indicated by full lines. Enzymes studied in **papers II** and **III** are highlighted in green. Enzymes studied in **paper IV** are highlighted in blue. Figure adapted from **Figure 3** of **paper I**.

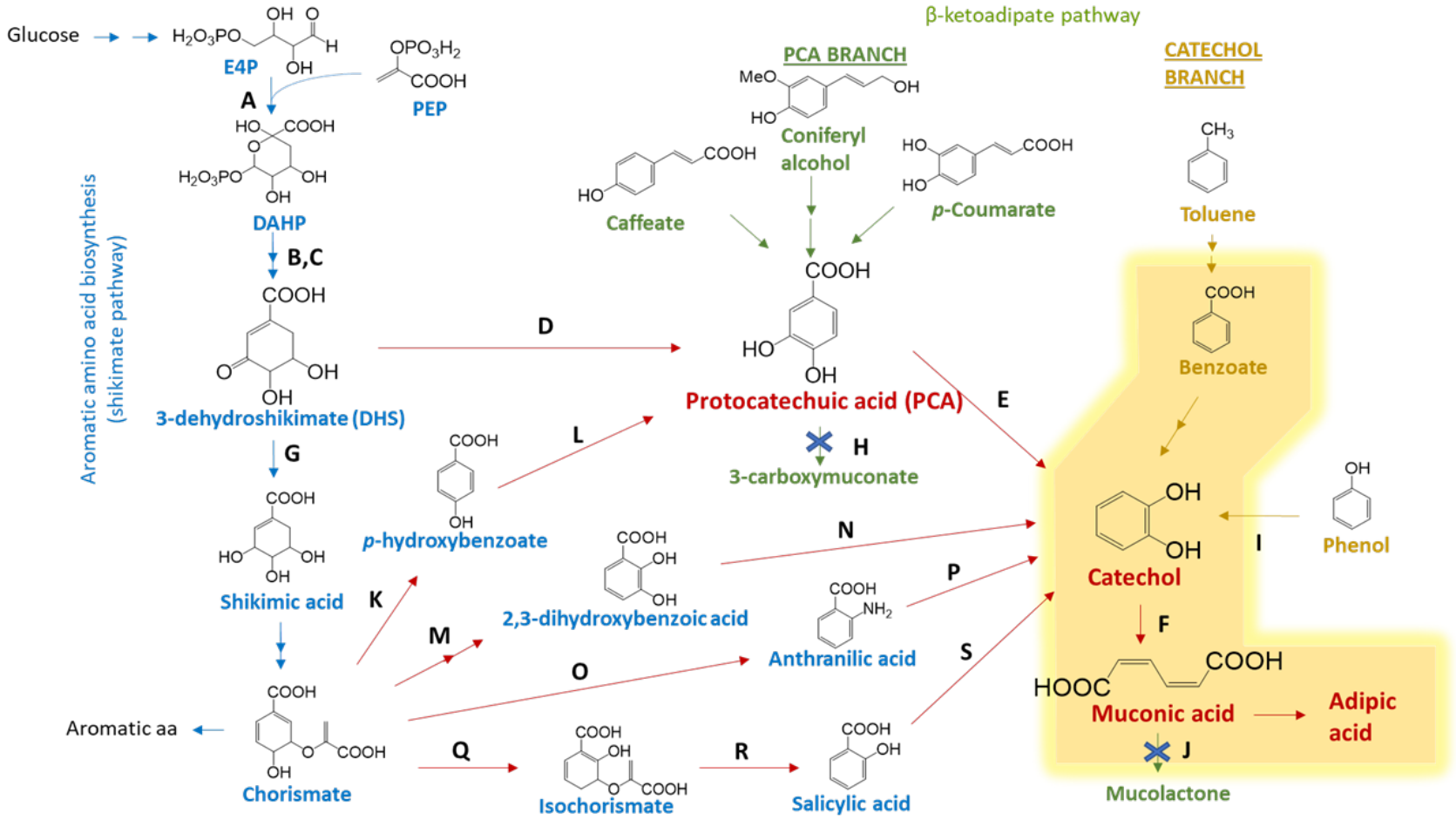


Figure 2. **Metabolic pathways leading to adipic acid *via cis,cis*-muconic acid.** My contribution in **paper V** is highlighted in orange. E4P, erythrose 4-phosphate; PEP, phosphoenolpyruvate; DAHP, 3-deoxy-D-arabinoheptulosonate-7-phosphate; A, DAHP synthase; B, 3-dehydroquinate (DHQ) synthase; C, DHQ dehydratase; D, DHS dehydratase; E, PCA decarboxylase; F, catechol 1,2-dioxygenase; G, 3-dehydroshikimate dehydrogenase; H, protocatechuate 3,4 dioxygenase; I, phenol monooxygenase; J, muconatecycloisomerase; K, chorismate pyruvate-lyase; L, 4-hydroxybenzoic acid hydrolase; M, isochorismate synthase, isochorismatase and 2,3-dihydro-2,3-DHBA dehydrogenase; N, 2,3-dihydroxybenzoate decarboxylase; O, anthranilate synthase; P, anthranilate 1,2-dioxygenase; Q, isochorismate synthase; R, isochorismate pyruvate lyase; S, salicylate 1-monoxygenase. Figure adapted from **Figure 6** of **paper I**.

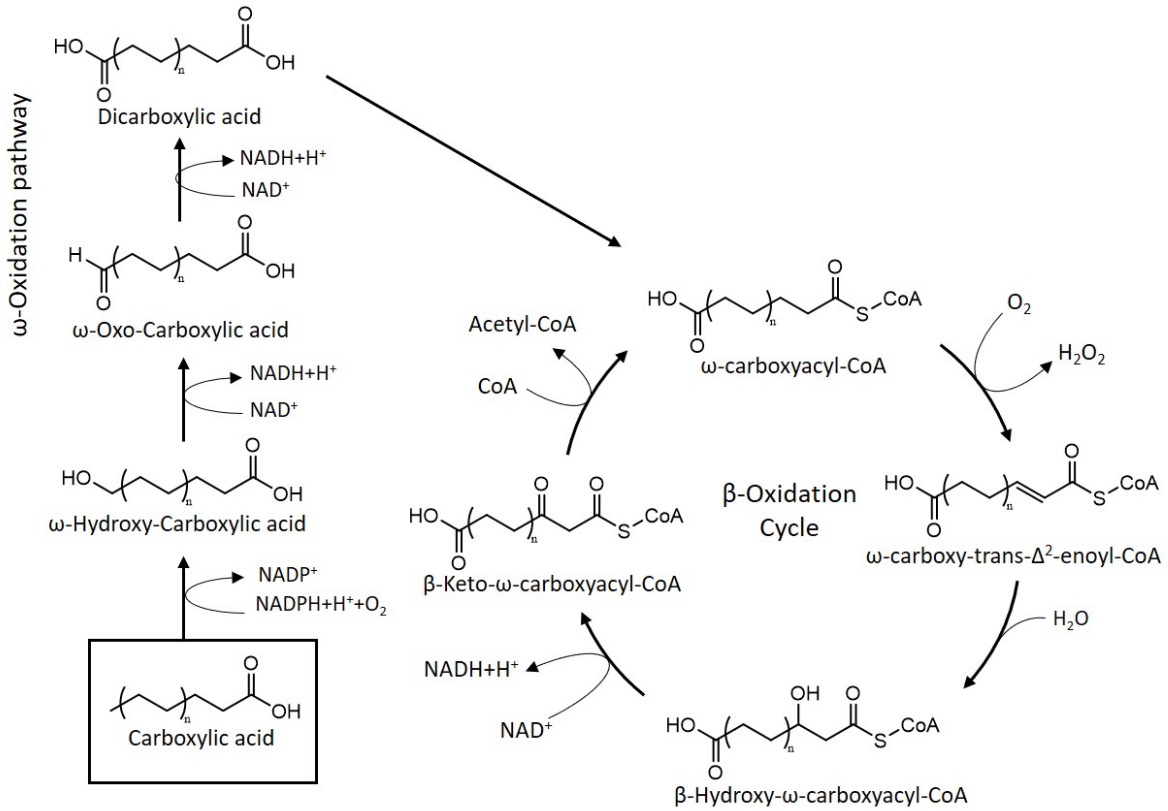


Figure 3. Production of adipic acid via ω- and β-oxidation of fatty acids. Conversion of a fatty acid (carboxylic acid) to dicarboxylic acid by means of the ω-oxidation pathway, followed by its entry in the β-oxidation cycle, whereby two carbons are lost per cycle until adipic acid is formed. Figure adapted from **Figure 4** of **paper I**.

The next wave of studies on biological adipic acid production centered on the reverse β-oxidation pathways (**Figure 1**). Attempts to revert the functionality of β-oxidation were initially motivated by fuel biosynthesis (Dekishima *et al*. 2011; Dellomonaco *et al*. 2011). In subsequent years, the strategy was applied in various microorganisms to biosynthesize adipic acid from renewable feedstock, including glucose and glycerol (**sections 4.2** and **5.5**). Initial attempts to produce adipic acid from glucose using reverse β-oxidation proved to be challenging due to low titers (Babu *et al*. 2015; Cheong *et al*. 2016; Deng *et al*. 2016; Kallscheuer *et al*. 2017a; Kallscheuer *et al*. 2017b; Li *et al*. 2020; Yu *et al*. 2014) and the trend has shifted toward using glycerol owing to a possibly more favorable redox balance (Zhao *et al*. 2018). The enzymatic components for the reverse β-oxidation pathway were obtained from various organisms, as outlined in **section 4.2**. Because multiple enzyme options are available in each step of reverse β-oxidation, they can be introduced in a combinatorial manner (**paper IV**).

The overall goal of this thesis was to assess the individual enzymes involved in adipic acid biosynthesis *via* the L-lysine pathway, the reverse β-oxidation pathway, and *cis,cis*-muconic acid reduction (**Figures 1** and **2**). The L-lysine pathway has been theoretically conceptualized as a patent application (US Patent 7,799,545), but has not been experimentally demonstrated (**section 4.1**). Specifically, enzymes involved in the elimination of the α-amino group of L-lysine and subsequent reduction of the α,β-unsaturated bond of 6-amino-*trans*-2-hexenoic acid have not been well understood. Adipic acid biosynthesis from glucose *via* the reverse β-oxidation pathway is rather inefficient and has not been introduced yet in a microorganism with robust acid-secretion capability (**section 4.2**). Moreover, the key factor determining the ability of oxidoreductases to reduce *cis,cis*-muconic acid has not been fully elucidated (**section 4.3**). In this thesis, I investigated enzymatic reactions involved in different synthetic pathways leading to adipic acid (**paper I**) by harnessing the potential of recombinant technology (**paper IV** and **V**), as well as *in silico* technologies (**papers II**, **III** and **V**) required to address the aforementioned issues. The intention was to improve our understanding of synthetic metabolism and building a foundation to develop cell factories capable of producing adipic acid from renewable resources such as simple sugars. The scientific questions raised during my thesis work and how I addressed them are listed below.

Is adipic acid a naturally occurring compound?

The natural occurrence of the target compound of this thesis, adipic acid, is considered rare (Polen *et al*. 2013); however, as it does occur in nature, its biosynthetic pathway can be identified and overproduction can be implemented in a designated host. Adipic acid has been detected in bacteria (Deng and Mao 2015), plants (Vavruch 1954), and patients with diabetic ketoacidosis (Pettersen *et al*. 1972). Whereas its biosynthetic pathway in plants has not been fully elucidated, the one in bacteria has been recently identified and can be engineered into other organisms. This raises the question of which genes and host organisms should be selected. Given the strong interest in bio-based adipic acid production, multiple pathways have been explored simultaneously by different research groups, further widening the options (**paper I**).

Which adipic acid biosynthetic pathways are known, and which enzymatic steps are involved?

A substantial effort to produce adipic acid from renewable feedstocks has already been made and numerous pathways have been identified (Averesch *et al*. 2018; Deng *et al*. 2016; Polen *et al*. 2013) (**paper I**). The details of the individual enzymatic steps contributing to adipic acid biosynthesis (**Figure 1**) are addressed in **paper I**. Prior to the recent discovery of the bacterial pathway (Deng and Mao 2015), many groups introduced the reverse β-oxidation pathway in *Escherichia coli* (Babu *et al*. 2015; Cheong *et al*. 2016; Kallscheuer *et al*. 2017a; Yu *et al*. 2014). The enzymatic components for the reverse β-oxidation pathway were obtained from various organisms, as outlined in **section 4.2**. In my thesis, I approached the problem by first identifying

the enzymes for the reverse β-oxidation pathway (**paper I**) and then applying them in a combinatorial manner in *Corynebacterium glutamicum* (**paper IV**).

Can a theoretical but unproven pathway be demonstrated?

Although a bacterial adipic acid biosynthetic pathway has been discovered, it remains to be determined whether it is the most efficient and robust option. By exploring additional pathways, it may be possible to develop a superior one capable of using substrates other than glycerol. One such example is provided by the production of adipic acid *via* the L-lysine pathway (**Figure 1**). The L-lysine pathway remains solely theoretical because not all enzymatic steps have been identified. Removal of the α-amino group from L-lysine and subsequent reduction of the α,β-unsaturated bond represent the 'metabolic gap' in the pathway (Shin *et al*. 2013), as all other metabolic steps between glucose and adipic acid are already known (**papers I** and **II**). The incomplete L-lysine pathway can be addressed both in theoretical and experimental terms. One can employ simulations to compute enzymatic steps and assess the thermodynamic feasibility and efficiency of the pathway. Experimentally, one can introduce a series of enzymes to build and then evaluate the pathways *in vivo*. Experimental aspects of the reduction and deamination reactions in the L-lysine pathway are addressed in **papers II** and **III**, respectively. Furthermore, interactions of enzymes and substrates in the L-lysine pathway are assessed *in silico*. As *in silico* studies shed light on enzyme-binding pockets and molecular interactions, this approach can be used to screen for possible enzyme candidates.

Which organism is the most suitable for a given pathway?

When multiple pathways are available, one can turn around the question and ask which pathway is most suitable for a given host organism. While yeast generally tolerates adipic acid better than bacteria (Karlsson *et al*. 2017), implementation of the *cis*,*cis*-muconic acid pathway (Raj *et al*. 2018) or β-oxidation pathway in yeast does not yield more adipic acid than in *E. coli*. Even though *E. coli* has been shown to generate the highest adipic acid production (Zhao *et al*. 2018), the pathway should be further improved to allow for the use of substrates other than glycerol (**paper I**). Additionally, the pathway should be implemented in diverse host organisms to allow more options. The ideal production scheme would involve a robust, acid-tolerant strain and strong flux to adipic acid from glucose. I addressed this challenge in **papers IV** and **V** by investigating alternative combinations of pathways and production hosts. *C. glutamicum* presents interesting characteristics and, to the best of my knowledge, has not been reported previously as a production host for adipic acid. *C. glutamicum* is known to secrete multiple compounds including succinic acid and *cis*,*cis*-muconic acid under different conditions (**section 5.4**). Thus, **paper IV** provides *in vivo* feasibility assessment of reverse β-oxidation in *C. glutamicum*; whereas **paper V** investigates *in vivo* reduction of *cis*,*cis*-muconic acid in *C. glutamicum* (**section 4.3**).

In general, developing a process for the bioproduction of chemicals from renewable feedstock often requires 1) pathway identification, 2) host selection, 3) implementation of individual enzymatic steps, 4) pathway and host organism optimization, and 5) process

optimization and scale-up (Shin *et al*. 2013). In my thesis, I addressed the first three steps. Further work on the last two steps will ultimately lead to improved adipic acid production using biomass.

2. Computational assessment of enzyme-substrate interaction

With advances in biotechnology, genome sequencing becomes cheaper with each year (Wetterstrand 2020) and nearly 9 trillion bases have been sequenced as of the end of 2020 (NCBI 2020a). As a consequence of increasing genomic knowledge, the number of known protein sequences amounts now to nearly 200 million (EBI 2020; NCBI 2020b). For practical reasons, not all proteins can be crystallized; however, 11,000 crystal structures are reported annually and 170,000 of them have been deposited by November of 2020 (RCSB 2020). At present, it is practically and realistically near to impossible to fully understand all the 200 million known protein sequences at the structural level. This is where computational tools prove to be powerful and useful in a complementary manner by broadening our knowledge of enzymatic processes. Homology modeling (Salmaso and Moro 2018) and molecular docking studies (Haddad *et al*. 2020) have been particularly useful in this thesis for **papers II** and **V**. The computational methods ensured the substrate binding to the enzyme pocket during the enzyme candidate screening process.

2.1. Protein structure prediction

For practical reasons, structural studies and crystallography simply cannot keep up with the growing number of known protein sequences. To this end, homology modeling offers an important tool to complement structural studies as it allows to predict the 3D structure of a protein based solely on its sequence. Various tools are available for both academic and commercial purposes (Bordoli *et al*. 2009; Lindorff-Larsen *et al*. 2011; Misura *et al*. 2006; Yang *et al*. 2014). Besides differences between algorithms and software, there are also several methods to build a homology model, including physics-based modeling and knowledge-based modeling (Kuhlman and Bradley 2019). The latter is sometimes referred to as template-based modeling

and relies on existing knowledge about template structures, whereas the former uses conformational sampling and known examples of sequence homology to predict local structures.

The template-based approach depends on databases to identify both sequence and structural similarities. Protein structure databases such as PDB can be queried using the protein sequence of interest to find a suitable structural template for subsequent model construction. Structures can then be predicted based on multiple sequence alignment with the template sequences (**Figure 4**) (Bordoli *et al*. 2009; Eddy 1998). For instance, the alpha helices and beta sheets of an unknown protein can be constructed using already characterized structures with similar amino acid sequences. Multiple sequence alignment with BLAST, or equivalent, is done automatically in the homology modeling process. Because the template-based method relies heavily on sequence alignment, the structural coverage of a database and its accuracy are of crucial importance. As of 2017, nearly 70% of all human proteins and nearly 95% of all human drug targets had a structural homologue in the database (Somody *et al*. 2017). While the same level of coverage might not be true for all bacterial proteins, the proteins studied within the scope of this thesis each had at least one structural homologue as template (**papers II** and **V**). Novel protein folds were not discovered during the studies and thus their validation was not required. Proteins in this thesis exhibited strong homology with each template (87% identity for Oye and 29-32% for enoate reductases).

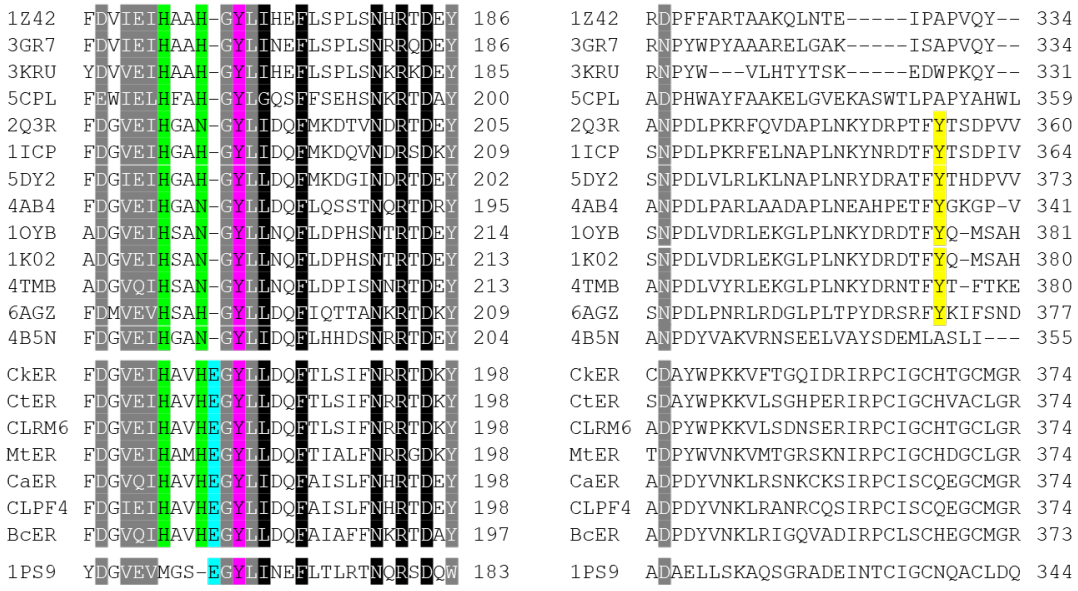


Figure 4. Exemplary multiple alignment of protein sequences. Members of the Old yellow enzyme family (labeled with PDB ID) and enoate reductases are aligned for comparison. Residues involved in substrate binding are highlighted in light green. The catalytic tyrosine residue is shown in magenta. Identical residues are highlighted black. Similar residues are highlighted in grey. E179 or equivalent residue unique to ERs is highlighted in cyan. E164 for *E. coli* 2,4-dienoyl CoA reductase (1PS9) involved in substrate binding is also highlighted in cyan. Conserved tyrosine residues prohibiting *cis,cis*-muconic acid binding is highlighted in yellow. Figure adapted from **Figure 5** of **paper V**.

Physics-based modeling can be useful when the sequence of interest does not present significant homology with other known globular structures. The algorithm looks for the lowest possible energy the protein structure conformation can take in the 3D space. In other words, given the amino acid sequence, the algorithm looks for the most stable 3D structure of the protein, *i.e.*, that with the lowest energy conformation. Because the physics-based method does not rely on a database for homology search, novel folds are likely to be discovered (Kuhlman and Bradley 2019). Thus, it requires heavy computational power and shorter amino acid sequences are advantageous. Proteins with a length of 10–80 amino acids have been predicted with great accuracy using physics-based methods (Lindorff-Larsen *et al*. 2011). Given recent advances in computational power and machine learning capabilities, the prediction of protein structures has become even faster and more accurate (Callaway 2020). Artificial intelligence could successfully predict protein sequences of various lengths at the recent CASP 14 protein folding competition (Service 2020). For proteins studied in **papers II** and **V**, at least one template structure was identified, and the physics-based approach was not necessary.

For all the homology models built during this thesis work, template-based methods were used. In **paper II**, the structure of NemA (E.C. 1.3.1.-) from *E. coli* was required but was not available from the database. Thus, a homology model was built using the Prime module in Maestro software. The template used for NemA model construction was pentaerythritol tetranitrate reductase (PDB: 3P7Y) (Toogood *et al*. 2011). The identities were 87% and the positives 94% between the NemA sequence and the template. In **paper V**, the structure of enoate reductases was required to study binding feasibility of *cis,cis*-muconic acid. As no such structure was available, homology models (**Figure 5**) were built using I-TASSER software (Yang *et al*. 2014).

2.2.Molecular docking

Molecular docking applies computational tools to study the interaction between small molecules (*e.g.*, substrates) and macromolecules (*e.g.*, enzymes) (Morris *et al*. 1998; Trott and Olson 2009), and is widely used in chemistry and biology. Over the last two decades, computational tools including quantum mechanics and molecular docking have been crucial for protein and pathway engineering (Siegel *et al*. 2010; Wu *et al*. 2011; Zhang *et al*. 2008). While molecular docking is mainly used for drug discovery or structure-based drug design (Meng *et al*. 2012a; Morris *et al*. 1998; de Ruyck *et al*. 2016), the technique is common also in metabolic pathway engineering and synthetic biology (Chen *et al*. 2016; Fisher *et al*. 2014; Li *et al*. 2016; Marcheschi *et al*. 2012; Zhang *et al*. 2008). As a technique dealing with small molecule-protein interaction at the atomic level (Meng *et al*. 2012a; Morris *et al*. 1998), molecular docking can greatly assist in better understanding the enzymes to be implemented in pathway engineering. Molecular docking was employed in **papers II** and **V** for substrate-enzyme interaction studies.

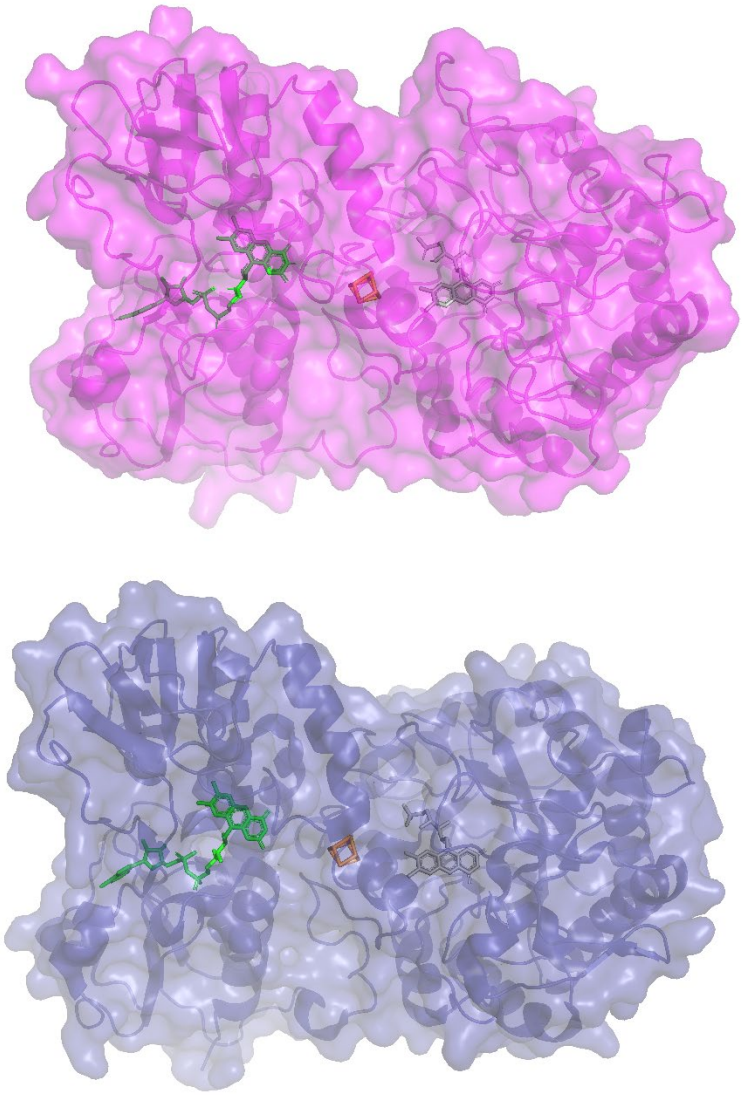


Figure 5. Exemplary homology modeling result. The enoate reductase homology model is shown in magenta. The template used for the model, 2-naphthoyl-CoA reductase (PDB: 6QKG) (Willistein *et al*. 2019), is shown in deep blue. FAD, yellow; [4Fe-4S] cluster, orange; FMN, grey. Figure adapted from **Figure S1** of **paper V**.

In **paper II**, two target compounds, *trans*-2-hexenedioic acid and 6-amino-*trans*-2-hexenoic acid, and two control compounds, *trans*-2-hexenal and *trans*-2-hexenoic acid, were used as substrates for docking to NemA and the Old yellow enzyme (Oye) family member Oye1 (PDB: 1OYB) (Fox and Karplus 1994). The four substrates were constructed in a computational language and loaded in the docking software for preparation. Reduction of *trans*-2-hexenedioic

acid or 6-amino-*trans*-2-hexenoic acid had not been reported previously; here, it was attempted to increase knowledge of the L-lysine pathway (**Figure 1**). The rationale for choosing the L-lysine pathway is discussed in more detail in a later chapter (**section 4.1**) as well as in **paper II**. Reduction of various α,β-unsaturated aldehydes (Stueckler *et al*. 2010), α,β-unsaturated ketoesters (Brenna *et al*. 2015), and α,β-unsaturated ketones (Brenna *et al*. 2015; Brenna *et al*. 2013; Padhi *et al*. 2009) is well documented. The control compound *trans*-2-hexenal was chosen because it is a known substrate for members of the Oye family (Williams *et al*. 2004). Enzymatic reduction of a hexenoic acid is rarely described in the literature. Because the target molecules are bi-functional, the control compound *trans*-2-hexenoic acid was chosen to investigate the monocarboxylate functional group on a molecule with a 6-carbon backbone. The Oye family was specifically chosen due to its wide substrate specificity (Williams *et al*. 2004). Using such promiscuous enzyme was considered an advantage for studying an unknown reaction. Thus, molecular docking was performed prior to *in vitro* studies as part of the enzyme screening process in **paper II**. Had the enzymes not had binding pockets capable of accommodating the target substrates, they would have been omitted from further studies. The results obtained in **paper II** revealed that two of the Oye family members had binding pockets accommodating *trans*-2-hexenedioic acid or 6-amino-*trans*-2-hexenoic acid. Further insight on Oye members was gained in **paper V**, when they were compared to another family of enzymes.

In **paper V**, the docking of two target compounds, *cis*,*cis*-muconic acid and *trans*-2-hexenedioic acid, and one control compound, *trans*-cinnamic acid, (**Figure 6**) to Oye members and enoate reductase class of enzymes was studied. Given that *trans*-cinnamic acid is a known substrate for enoate reductases (Joo *et al*. 2017), it was used as the control. The initial aim of the molecular docking experiment in **paper V** was to identify whether significant differences among the binding residues of enoate reductases exist. Enoate reductases were previously described to be capable of reducing *cis,cis*-muconic acid and *trans*-2-hexenedioic acid (Joo *et al*. 2017); whereas Oye family members had been found incapable of reducing these two substrates (Joo *et al*. 2017). Nonetheless, Oye members were included in **paper V** because some of them bound to carboxylic acids in **paper II**. Moreover, additional members of the Oye members not covered in **paper II** were investigated in **Paper V**. Molecular docking in **paper V** showed that all enoate reductases and some Oye members could bind to the target substrates. However, the results also demonstrated clear structural differences among these two types of enzymes. For instance, whereas the binding pocket structure was similar among the two, the degree of solvent exposure was very different. Enoate reductases exhibit two α,β domains that are not present in Oye family members. More importantly, enoate reductases harbor a FAD and [4Fe-4S] cluster, which allow more reductive power compared to the Oye family. Additional minor differences in binding pocket residues, potentially capable of inhibiting carboxylic acid binding, were observed among Oye members.

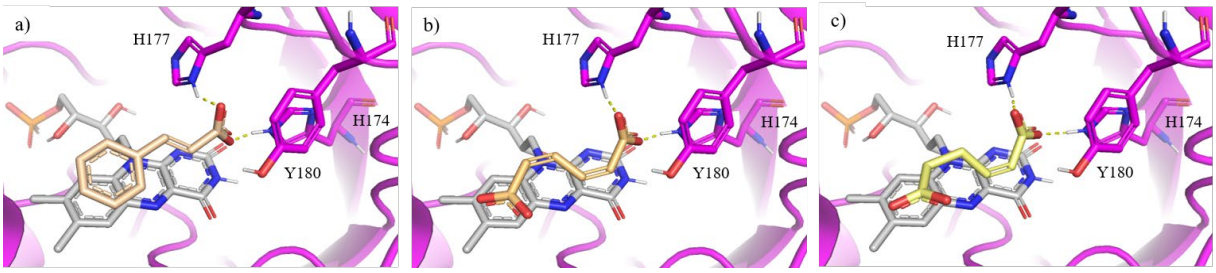


Figure 6. Molecular docking of substrates to an enzyme model. Binding of (a) cinnamic acid, (b) *cis,cis*-muconic acid, and (c) hexenedioic acid to enoate reductase (magenta). Figure adapted from **Figure 3** of **paper V**.

3. Analytical methods for adipic acid pathway intermediates

Chemical analysis is an important aspect in enzyme characterization and *in vivo* assessments of metabolic pathways. Metabolites present in low quantities are challenging to analyze especially when the amount is below the detection limit. Multiple analytical methods and instruments may be utilized in a complementary manner to provide better insight of the metabolic pathways. In **papers VI** and **V**, various analytical methods were used to evaluate metabolites in the cultivation broth. Producing non-native chemicals, which do not occur naturally in an organism, is challenging for numerous reasons. Heterologous gene expression does not always yield high intracellular protein levels and misfolded or insoluble products are common. Furthermore, the overall metabolism of the host organism might change as a result of introducing just a few genes. Sometimes, this goes unnoticed because the change is minor, yet, it shows that a powerful analytical method is needed to accurately detect target metabolites. Analytical methods to detect metabolites were extremely crucial in my studies to monitor the progression of the metabolic pathways.

3.1. High-performance liquid chromatography

High-performance liquid chromatography (HPLC) is commonly employed to measure various types of metabolites, including sugars, acetic acid, succinic acid, fumaric acid, and lactic acid. Here, HPLC was used in **papers IV** and **V** for metabolite analysis (**Table 1**). While compounds can be resolved easily by HPLC, the limit of detection is quite poor. The compounds listed above are often detected at a g/L scale and detection is feasible even at a mg/L scale; however, detection at a µg/L or ng/L scale is more problematic. In my case, the detection limit for adipic acid by HPLC being approximately 50 mg/L and production at the µg/L scale, HPLC was

employed in **paper IV** only to assess other compounds such as glucose. Additionally, HPLC was employed in **paper V** to analyze benzoic acid and *cis,cis*-muconic acid.

Table 1. List of analytical methods used in this thesis to detect adipic acid pathway intermediates, adipic acid, sugars, and central metabolites.

	Metabolites analyzed	Note
HPLC	Sugars, organic acids, central metabolites	•Easy sample preparation •Signals within the detection range •Routine method available in the lab •Sensitivity not effected by cultivation broth
	Benzoic acid and *cis,cis*-muconic acid	•Easy sample preparation •Instrument operated with reverse phase column •Signals within the detection range •Very high reproducibility
	Adipic acid	•Current production level is lower than the detection limit •Derivatization needed for increased signal •BPB derivatization step compromised by cultivation broth •Instrument operated with a reverse phase column
GC/MS	3-Oxoadipic acid, 3-hydroxyadipic acid, *trans*-2-hexenedioic acid, adipic acid	•MSTFA derivatization required •Very high sensitivity •Sensitivity compromised by cultivation broth •Removal of cultivation impurities desired •Requires staff for the instrument maintenance •Cultivation broth damages instrument •Currently only 3-hydroxyadipic acid production level within the detection range •Qualitative result provided
LC/MS	3-Oxoadipic acid and 3-hydroxyadipic acid	•Sensitivity compromised by cultivation broth •Not compatible with the current SPE protocols
	Trans-2-hexenedioic acid and adipic acid	•Sample preparation is easy •Analysis takes place at a different university •Need logistics for sample transport •Compatible with the current SPE protocols

3.2. Gas chromatography/mass spectrometry

Gas chromatography/mass spectrometry (GC/MS) is a widely used technique for identifying and quantifying compounds. The ability to provide qualitative confirmation of compounds present in the cultivation broth was the reason for choosing GC/MS as a means to further characterize the engineered strains in **paper IV** (**Table 1**). GC/MS can be used to identify or quantify compounds

within microorganisms for fluxome analysis (Wittmann *et al*., 2004), as well as those excreted by engineered microorganisms (Kallscheuer *et al*. 2017a). The target compound (adipic acid) and its intermediates were investigated in **paper IV**. The *m/z* fragmentation pattern can be used to identify the compound in the National Institute of Standards and Technology (NIST) database, while the instrument's software provides the top 100 likely hits that match the fragmentation pattern, as well as their probability. Essentially, the probability defines the confidence level for a given compound based on the *m/z* fragmentation pattern, while the latter is then matched to the database. Because the library search is partially dependent on the peaks and fragmentation quality, the probability can be obtained also for each concentration of adipic acid standards. Here, the compounds of interest were functionalized by carboxylic acid and, hence, a derivatization step was needed to make the compounds more volatile and suitable for GC/MS analysis. To this end, the metabolites were derivatized by N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) (Paczia *et al*., 2012). Given the uncertainty regarding the amount of the compound of interest present in the cultivation broth, the limit of detection of the GC/MS instrument was assessed first (**Figure 7**). Crucially, in GC/MS analysis, the signal corresponding to the target compound may be influenced by the presence of other chemicals such as the constituents of the cultivation broth. The purpose of limit of detection assessment was to detect a miniscule amount of adipic acid in the cultivation broth, which likely affected the obtained signals. Thus, the limit of detection was evaluated both in the absence and presence of cultivation broth. If the adipic acid standard matched the data library entry for adipic acid, then, increasing the concentration of standard should augment the chance of a match. This scenario was assessed with or without cultivation medium (**Figure 7**). In the absence of cultivation broth, even 1 μg/L adipic acid standard displayed 72% probability of matching the library entry. In contrast, in the presence of cultivation broth, spiking with 1 μg/L adipic acid led to only 1.17% probability (**Figure 7**). At 1 mg/L adipic acid, the probability of matching was 20% in the presence of cultivation broth and 95% in its absence. These results suggested that the presence of untreated cultivation broth posed a challenge when detecting less than 10 mg/L of adipic acid by GC/MS. This meant that the engineered strains had to be sufficiently efficient to produce 10 mg/L or more of adipic acid if this method was to be employed.

3.3. Carboxylic acid derivatization for reverse-phase HPLC

Given the difficulty of measuring tiny concentrations of compounds by GC/MC in the presence of impurities, another approach was pursued. Because simply concentrating the samples would not be adequate, labeling and amplifying the signal of the target compound was deemed more appropriate. Even if not widely used, derivatization of free fatty acid molecules with chromophores, such as 4′-bromophenacyl bromide (BPB) (Durst *et al*. 1975; Grushka *et al*. 1975; L'emeillat *et al*. 1981), 2-bromo-2′-acetonaphthone (Zhang *et al*. 2014) or 2-bromoacetophenone (Grushka *et al*. 1975), have been documented before. Measurement of adipic acid below the level of detection by GC/MS was attempted using BPB derivatization. KOH served as the alkalizing agent and crown ether (18-crown-6) as catalyst. A successful measurement of adipic acid standard was accomplished (**Figure 8**), with 3.12 mg/L as the lowest reliably detected amount; below this level, peak shape and size became unclear (orange point, **Figure 8**). Construction of the standard curve for adipic acid in the presence of cultivation

medium was subsequently attempted; however, the signal did not correlate to the concentration of adipic acid and any attempts to measure sub-mg/L amounts of adipic acid were abandoned.

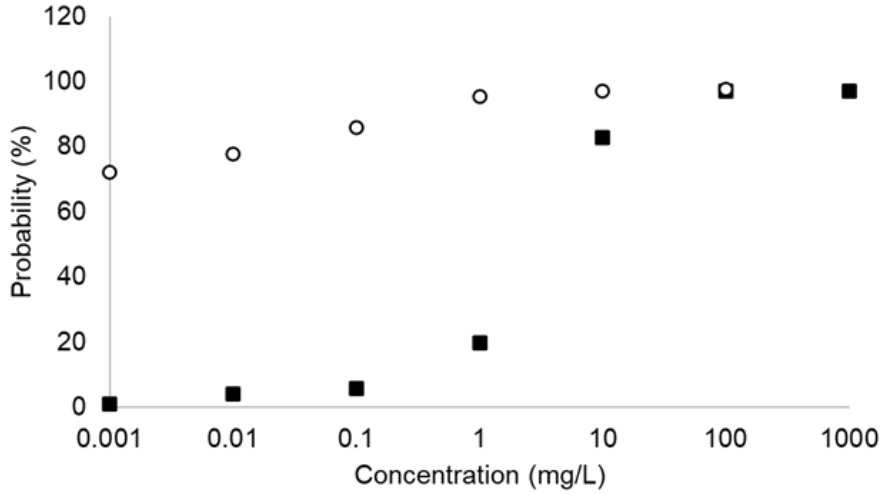


Figure 7. Probability of the *m/z* fragmentation pattern matching the adipic acid entry in the library. The probability as a function of adipic acid standard concentration in the presence (dark square) or absence (empty circle) of cultivation medium is shown.

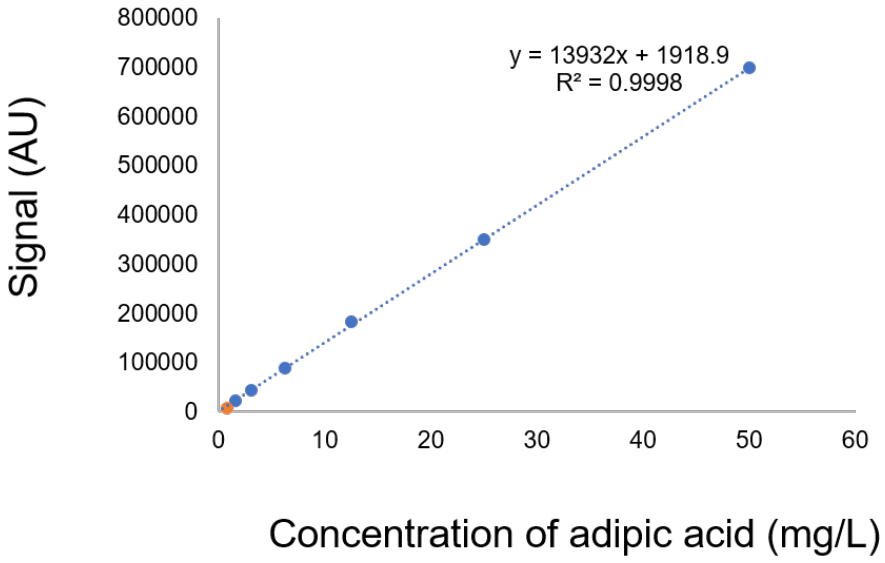


Figure 8. Standard curve for BPB-derivatized adipic acid. The blue plot spans the range between 3.12 and 50 mg/L adipic acid. The orange point corresponds to 1.56 mg/L adipic acid.

3.4. Solid-phase extraction

Considering that the ratio of target compounds to interfering compounds was too low, selective sequestration of target analytes was needed. Thus, solid-phase extraction (SPE) was explored as a means of removing unwanted compounds from the cultivation broth and selectively extracting adipic acid or intermediate compounds. As all target compounds were derivatives of carboxylic acids, it was reasoned that ion exchange through an anionic column could be employed. However, all attempts to extract carboxylic acid from the cultivation broth based on anion exchange chromatography failed. Neither manipulation of SPE protocols according to the manufacturer's guidance and suggestions, nor use of a counter-intuitive approach such as cation-exchange column, improved the results. As the carboxylic acids were by no means retained in the resin columns, we hypothesized that the excessively high ionic strength of the cultivation broth prevented the correct interaction between the resin and carboxylic acids.

Given that the target analytes were negatively charged, using a hydrophobic or non-charged resin was not the most obvious option. However, the first evidence of successful carboxylic acid retention from the cultivation broth was actually observed with hydrophobic SPE resins (**Appendix A**). A hydrophobic column, Strata C8, was used initially to retain adipic acid and to remove impurities from the cultivation broth. Strata C8 is a silica-based resin functionalized by octane to enable hydrophobic interactions. At each step, SPE was modified to improve retention of adipic acid (**Appendix A**) based upon the assumption that the neutralized form of adipic acid was retained by the hydrophobic column in the presence of excessively high ionic strength or a more acidic pH, while impurities would flow through. Acidification of the cultivation broth prior to loading on the column ensured the neutralization of adipic acid. However, even under these conditions, adipic acid was not fully retained by the Strata C8 column and was eluted in the load and wash fractions. For an accurate quantification, all adipic acid must be in the elution fraction. To try and improve retention, a different attempt was made by switching to another column called Strata X. While both columns interact with neutral compounds, the polymer-based Strata X resin retains compounds based on hydrophobic interactions and hydrogen bonding. Retention was substantially improved as no adipic acid was present either in the load or in the wash fractions when using the Strata X resin (**Appendix A**). As much as 5 mL of cultivation broth was loaded onto the SPE column, yielding a final elution volume of 1 mL and corresponding to a five-fold concentration. Not all impurities were removed by Strata X, but identification of flow through compounds was not pursued and the obtained result was deemed satisfactory. Retention was higher compared to the Strata C8 column even when the optimized protocols were applied (**Figure 9**). Testing with 100 µg of adipic acid, all adipic acid was recovered by Strata X while 83 µg by Strata C8. Additionally, retention of *trans*-2-hexenedioic acid was successfully achieved by SPE (**paper IV**); whereas 3-hydroxyadipic acid and 3-oxoadipic acid were not retained by the Strata X resin.

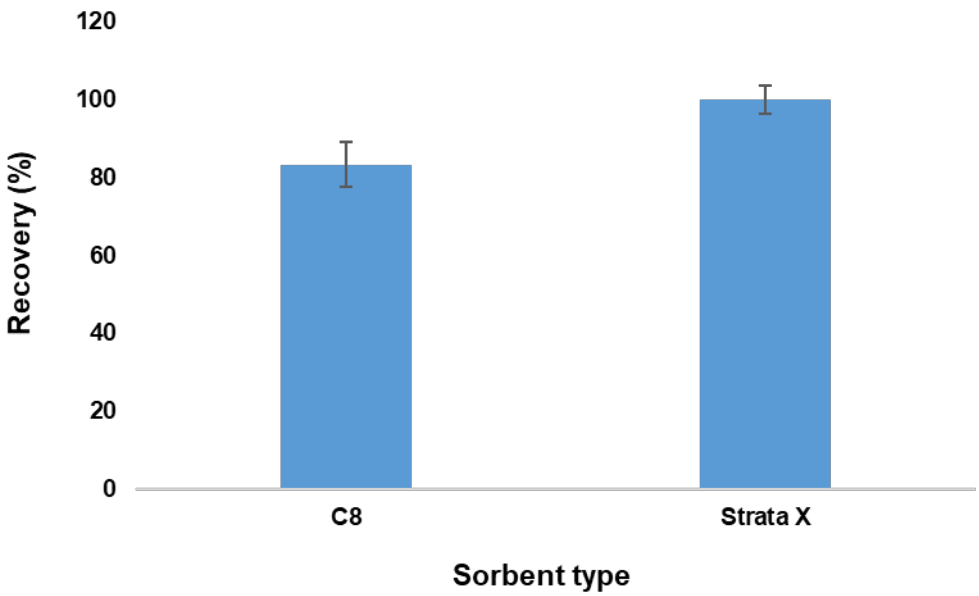


Figure 9. Adipic acid recovery by two different hydrophobic SPE columns. Adipic acid was spiked in the cultivation broth at a concentration of 100 mg/L. The recovery rate was normalized to that obtained by the Strata X resin. (N = 3; error bars = standard deviation).

3.5. Liquid chromatography/mass spectrometry

Based on the progress made with SPE, *trans*-2-hexenedioic acid was measured using liquid chromatography/mass spectrometry (LC/MS). It was assumed that both the target analytes and the impurities from cultivation broth were concentrated during the SPE process. Because of this increased amount of impurities, GC/MS was replaced by LC/MS to avoid damaging the instrument. Unlike GC/MS, the power of ionization is much lower and there is no fragmentation in LC/MS. However, dimers, adducts, and loss of water or carbon dioxide from the molecules may occur, yielding specific *m/z* patterns.

LC/MS allowed to investigate the production of *trans*-2-hexenedioic acid in the engineered strains. Details about the construction of the strain for *trans*-2-hexenedioic acid are provided in **paper IV**. Cultivation broth from the engineered strain was collected and cells were removed by centrifugation and subsequent SPE treatment of supernatant followed. The Strata X column was loaded with 5 mL cultivation broth, which was concentrated to 1 mL by SPE. 5 µL of the sample was then analyzed by LC/MS for the presence of *trans*-2-hexenedioic acid (**Figure 10**). The engineered strain, harboring plasmid pJS143, produced a miniscule amount of *trans*-2-hexenedioic acid; whereas the latter compound was not detected in the control strain harboring plasmid pJS146 (**paper IV**). Plasmid pJS143 expressed four genes (*paaJHF* and *tesB*) and details of the metabolic pathway are described at a later point (**section 4.2**). The control plasmid expressed three genes (*paaJH* and *tesB*) from the same *trans*-2-hexenedioic acid biosynthesis pathway but lacked the one responsible for the last step (*paaF*). Further analysis with GC/MS shall validate the identity of *trans*-2-hexenedioic acid.

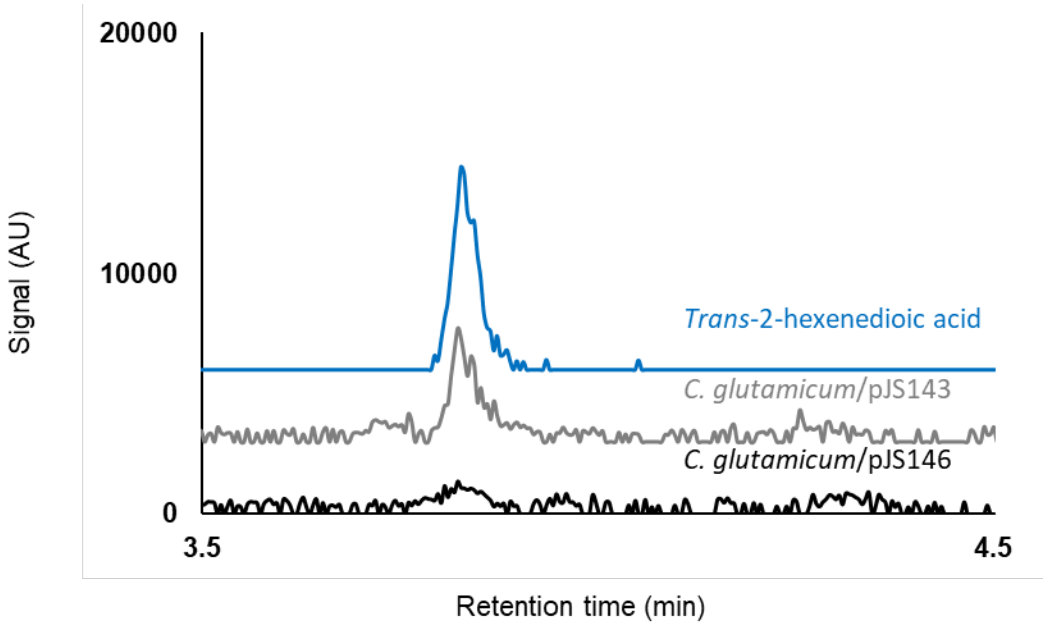


Figure 10. Detection of *trans*-2-hexenedioic acid using LC/MS. Representative chromatograms of SPE-treated samples are shown (extracted ion: 143.03498). The blue line indicates the *trans*-2-hexenedioic acid standard. Grey line indicates *C. glutamicum* harboring the *trans*-2-hexenedioic acid producing vector pJS143. Black line indicates *C. glutamicum* harboring 3-hydroxyadipic acid producing vector pJS146. Time intervals are shown for clarity.

4. Synthetic metabolic pathways for adipic acid production

Multiple synthetic metabolic pathways for the biological production of adipic acid have been reported (Deng *et al*. 2016; Kallscheuer *et al*. 2017a; Kruyer and Peralta-Yahya 2017). While demonstration of a pathway's functional status is a crucial step, further optimization and continuous improvements are necessary in order to augment the titer, yield, and productivity (**paper I**). Additionally, the use of glucose as feedstock for the biosynthesis of adipic acid requires further exploration and implementation of new strategies. While glycerol-based strategies for adipic acid biosynthesis has been successfully demonstrated, sugar-based substrates would permit transition opportunities toward lignocellulosic substrates. The L-lysine pathway (**papers II** and **III**), the reverse β-oxidation pathway (**paper VI**) and reduction of *cis*,*cis*-muconic acid (**paper V**) were investigated in my thesis.

4.1. Exploration of the L-lysine pathway

Nearly a decade ago, US patent number 7,799,545 describing various synthetic pathways leading to adipic acid was published. The theoretical outline of adipic acid biosynthesis *via* L-lysine was first described in this patent. A handful of groups from academia and the industry have attempted to experimentally demonstrate the pathway (J.H. Seo, personal communications), but no success has been reported since then. Considering that L-lysine is a biologically and industrially attainable intermediate, the interest for this pathway remains.

Implementation of synthetic pathways in L-lysine-producing organisms is advantageous, given the industrial relevance of L-lysine and the estimated 2.4 million tons of it being produced annually (Lee and Wendisch 2017) by *C. glutamicum* fermentation. Biosynthesis of 1,5-

diaminopentane (Kind *et al*. 2014; Kind and Wittmann 2011), glutaric acid (Rohles *et al*. 2018), and 5-aminovaleric acid (Rohles *et al*. 2016) has been successfully demonstrated by engineering L-lysine-producing strains. Any attempt to implement adipic acid production in L-lysine-generating organisms is a logical approach as long as the pathways are well established (**paper II**). However, as adipic acid production *via* L-lysine has not been established, the biochemical and physical aspects of the pathway need to be explored prior to addressing current challenges. There are three biochemical steps in converting L-lysine to adipic acid: α,β-elimination, α,β-reduction, and transamination at the ω-position. The transamination at the ω-position has already been addressed in the literature (Fedorchuk *et al*. 2020; Takehara *et al*. 2018; Yagi *et al*. 1991).

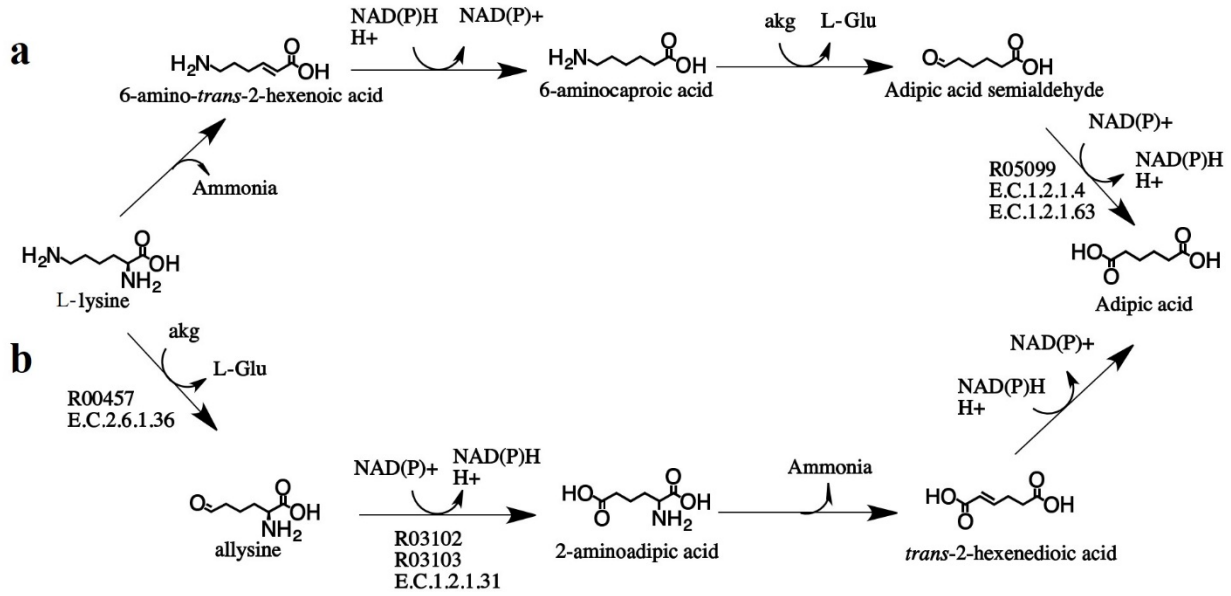


Figure 11. The theoretical L-lysine pathways for adipic acid biosynthesis. (a) The L-lysine pathway to adipic acid first described by patent application U.S. 7,491,520 from Genomatica. (b) The alternative L-lysine pathway described in **paper II**. Figure adapted from **Figures 1** and **2** of **paper II**.

The α,β-reduction step in the L-lysine pathway can occur either on *trans*-2-hexenedioic acid or on 6-amino-2-*trans*-hexenoic acid (**Figure 11**); the difference being whether α,β-reduction happens before or after transamination of the terminal amino group. Reduction of 6-amino-*trans*-2-hexenoic acid has been initially described in patent application U.S. 7,491,520 (Raemakers-Franken *et al*. 2007); whereas reduction of *trans*-2-hexenedioic acid has not been detailed. To better understand the reaction mechanisms, molecular docking studies were carried out (**paper II**). To this end, it was first necessary to choose an enzyme structure. The crystal structure of Oye1 (PDB: 1OYB) from *S. pastorianus* had been often employed in computational studies of various α,β-unsaturated molecules (**section 2.2**). An additional enzyme, NemA from *E. coli*, was chosen for molecular docking studies, as it already showed activity toward 6-amino-*trans*-2-hexenoic acid (Raemakers-Franken *et al*. 2007). Because a crystal structure for NemA was not available, a homology model was used instead.

Molecular docking studies revealed that both NemA and Oye1 possessed binding pockets large enough to accommodate 6-amino-*trans*-2-hexenoic acid or *trans*-2-hexenedioic acid (**paper II**). Binding of substrates to the enzyme occurred at an orientation and angle that were acceptable for catalysis during α,β-reduction. Subsequent purification of the enzymes was intended for experimental demonstration of the reduction reaction; however, *in vitro* experiments could not prove catalysis. One reason why the NemA result differed from a previous report (Raemakers-Franken *et al*. 2007) could be ascribed to different reaction conditions. Activation of the double bond is often required for α,β-reduction to occur (Lonsdale and Reetz 2015), and whether an Oye-type enzyme facilitates such activation on 6-amino-*trans*-2-hexenoic acid or *trans*-2-hexenedioic acid is unknown. While an experimental demonstration could not be provided, I was able to rule out the inability of the NemA or Oye1 enzymes to spatially accommodate *trans*-2-hexenedioic acid or 6-amino-*trans*-2-hexenoic acid (**paper II**). The next challenge would be to screen for more enzymes to identify those capable of performing the catalytic step. Once a possible candidate is selected in the future, mutation strategies can be devised to further improve catalysis. Along the course of this study, an enzyme capable of reducing *trans*-2-hexenedioic acid was reported (Joo *et al*. 2017). Furthermore, subjecting NemA to enzyme engineering could potentially improve its catalytic activity, which currently takes 48 h to detect (Raemakers-Franken *et al*. 2007). The discovery of new enzymes with which to carry out reduction may open new possibilities. Once the mechanisms of transamination and reduction are established, elimination of the α-amino group remains the only piece of the puzzle to be determined. A look at the synthetic pathway reveals that formation of 6-amino-*trans*-2-hexenoic acid or *trans*-2-hexenedioic acid follows removal of the α-amino group and a β-hydrogen to create an unsaturated bond between the α and β positions (**Figure 11**).

Removal of the α-amino group from L-lysine must be thoroughly examined in order to design a logical experimental strategy for adipic acid biosynthesis *via* the L-lysine pathway (**paper III**). One option would be oxidative deamination of the amino group, whereby the resulting molecule is an α-keto acid, which can be then converted to 6-aminocaproic acid *via* carbon elongation and subsequent decarboxylation (Cheng *et al*. 2019). The other option is non-oxidative deamination, which theoretically leads to the formation of an unsaturated bond between α- and β-carbons. Arylalanine ammonia lyases, grouped under the E.C. 4.3.1.- category that includes histidine ammonia lyase and phenylalanine ammonia lyase, utilize 4-methylideneimidazole-5-one (MIO) as a cofactor to catalyze the non-oxidative deamination (Poppe 2001; Schwede *et al*. 1999). Unlike L-phenylalanine or L-histidine, the L-lysine backbone lacks an aromatic ring and a nucleophile, thus making the MIO-catalyzed reaction on L-lysine appear chemically difficult. Additionally, as the mechanism of MIO-catalyzed deamination remains uncertain (Parmeggiani *et al*. 2018), it is unclear whether such reactions are suitable for elimination on L-lysine. Moreover, a hypothetical non-oxidative removal of the α-amino group from L-lysine would proceed in conjunction with a base-catalyzed abstraction of the β-hydrogen (**Figure 12**). The acidity of β-hydrogen seems to be close to the weakest possible value and makes it very difficult to be abstracted (Bordwell 1988). To the best of my knowledge, no enzyme capable of carrying out such reaction on L-lysine has been reported. Indeed, such reaction seems nearly impossible on L-lysine unless the MIO-catalyzed reactions are ultimately proven to function *via* E2 elimination which represents the concerted reaction (**Figure 12**).

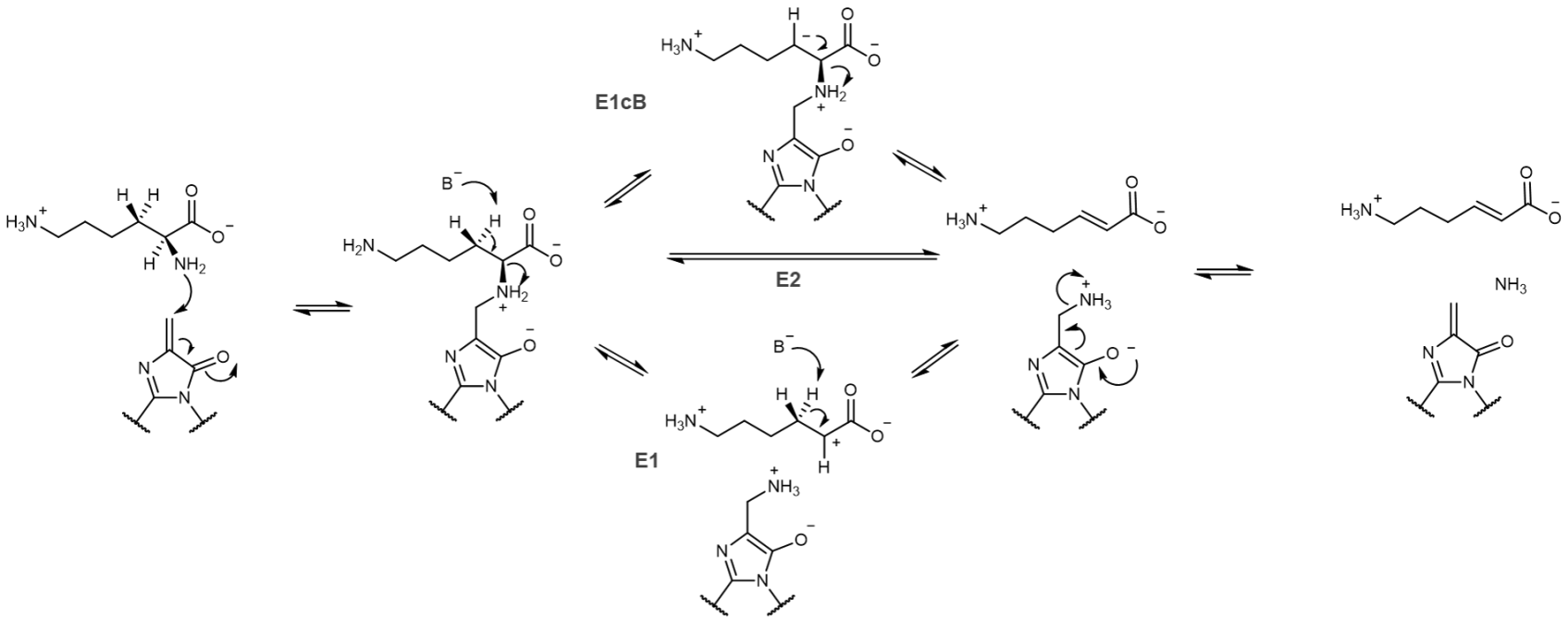


Figure 12. **Hypothetical α,β-deamination of L-lysine *via* a MIO intermediate.** Base-catalyzed abstraction of β-hydrogen after the initial Amino-MIO intermediate formation may proceed hypothetically *via* three possible options. E1cB mechanism leads to carbanion formation, E2 to a concerted reaction, and E1 to a carbocation formation. The exact mechanism by which MIO-enzymes operate has not been fully elucidated.

By E1cB mechanism, an anion can be created on the β-position of L-lysine by base abstraction of the β-hydrogen (Parmeggiani *et al*. 2018). In E1 mechanism, a carbocation can be crated on the α-position of L-lysine by MIO departure. The E1 and E1cB are less likely to occur on L-lysine by MIO catalyzed reaction due to unstable intermediate formation (**Figure 12**). Therefore, the desire to eliminate the α-amino group from L-lysine begs for a different strategy.

While elimination of the α-amino group from L-lysine seems chemically challenging, I hypothesize that replacing the amino group with a hydrogen in the β-position would facilitate the elimination reaction (**paper III**). In the hypothetical α,β-elimination of β-lysine, abstraction of the α-hydrogen yields a lone electron pair on the α-carbon, which can be stabilized by resonance through the α-carboxylate group (Porter and Bright 1980; Puthan Veetil *et al*. 2009). This would potentially allow for the formation of an enediolate intermediate, which would collapse to facilitate the departure of ammonia from the β-carbon. Abstraction of the β-hydrogen, followed by collapse of the enediolate intermediate has been shown to occur during non-oxidative deamination by L-aspartate ammonia lyase (Fibriansah *et al*. 2011; Shi *et al*. 1997). The structure of β-lysine and the hypothetical α,β-elimination of β-lysine resemble those of L-aspartate (**Figure 13**). Structurally, the β-hydrogen in L-aspartate is in the α position to the terminal carboxylate and thus resembles the α-hydrogen in β-lysine. To have the elimination product in the *trans* conformation, β-lysine must be in the D-relative configuration and the side chain of β-D-lysine must be in the anti-position to the α-carboxylate in the reaction (**Figure 14**).

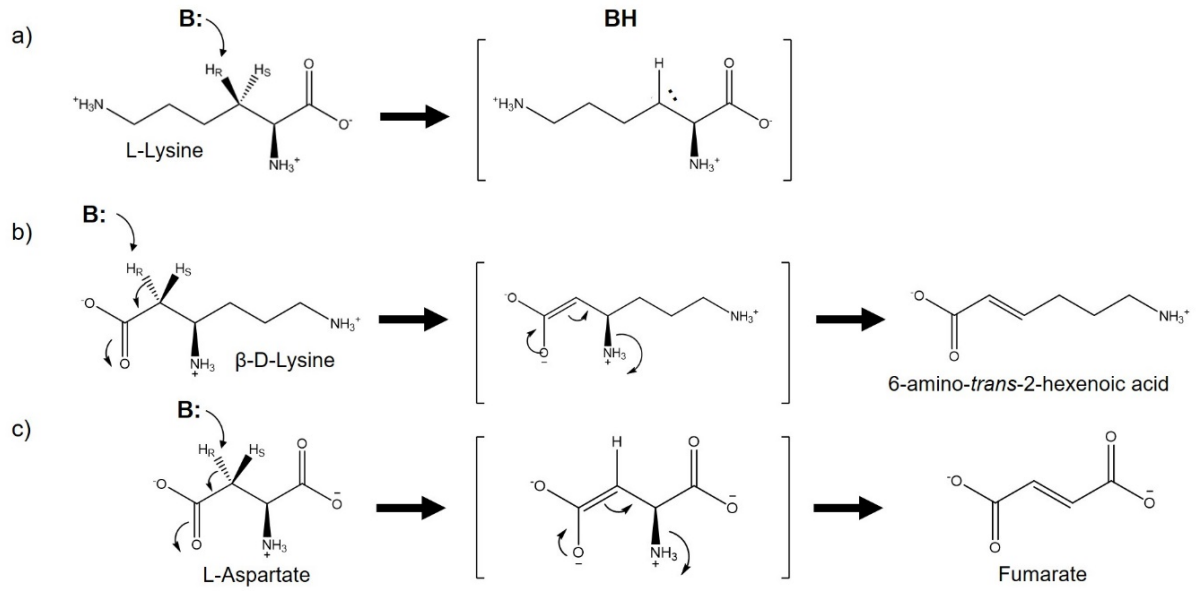


Figure 13. Hypothetical base abstraction of the α-hydrogen in L-lysine and β-hydrogen in β-lysine. (a) An unstable lone-pair election results from base-abstraction of the β-hydrogen in L-lysine. (b) Hypothetical α,β-deamination of β-lysine *via* transient enediolate intermediate formation. (c) α,β-elimination proceeds *via* the enediolate intermediate in L-aspartate ammonia lyase. B:, base; BH, conjugate acid of the base; H_S, Pro-S hydrogen; H_R, pro-R hydrogen

The hypothetical elimination proceeds with initial abstraction of the pro-R proton of β-D-lysine, followed by enediolate intermediate formation (**Figures 13** and **14**). A hypothetical enzyme would have suitable residues for stabilizing the enediolate intermediate. Collapse of the enediolate intermediate would then yield the elimination product of 6-amino-*trans*-2-hexenoic acid in the *trans* conformation. The need for a trans conformation can be rejected if α,β-reduction of 6-amino-hexenoic acid in the *cis* form is experimentally demonstrated. Regardless of the configuration of the product, the hypothetical α,β-elimination of β-lysine requires an engineered enzyme to be created (**Figure 14**). To the best of my knowledge, a naturally occurring enzyme capable of carrying out the elimination reaction on β-lysine has not been characterized nor identified. Aspects of the α,β-elimination theory on β-lysine, as well as trial experiments with aspartase-type enzymes are covered in **paper III**.

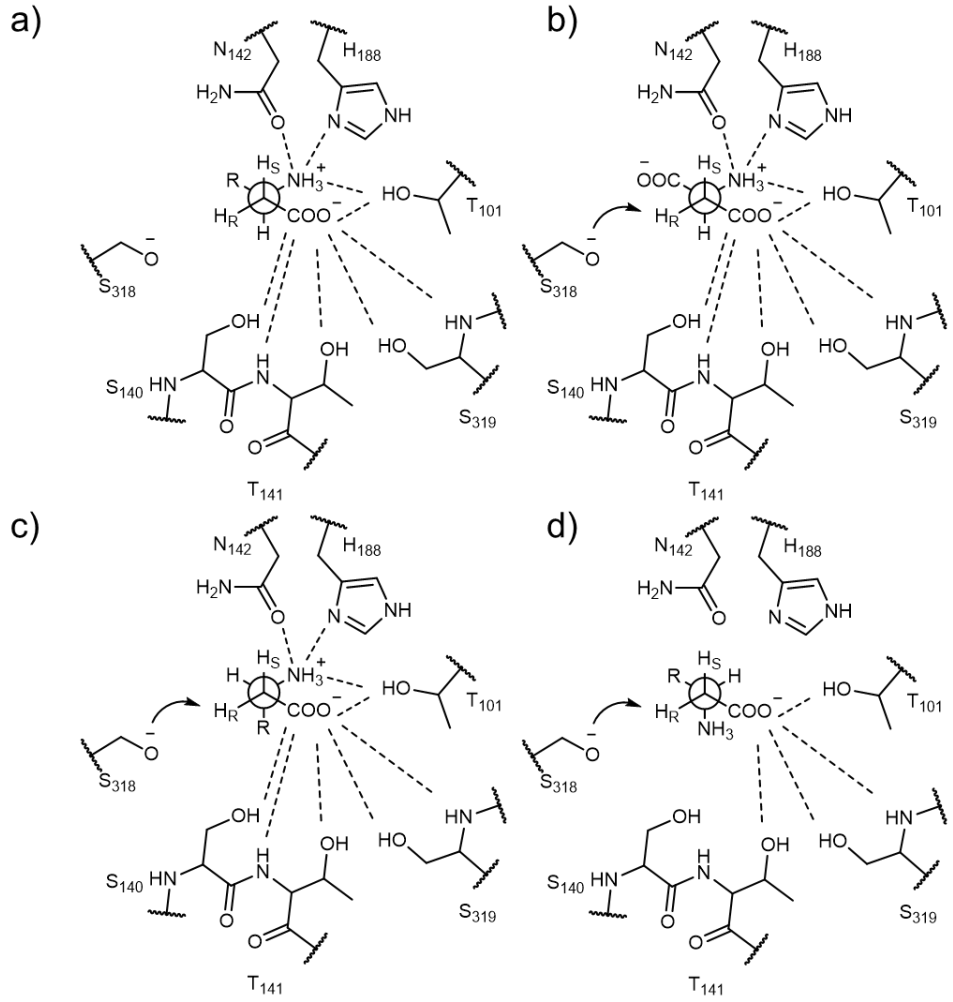


Figure 14. Proposed enzyme binding mechanism to β-lysine by a hypothetical enzyme catalyzing α,β-elimination. Newman projection of substrates prior to abstraction of the β-hydrogen of (a) β-D-lysine, (b) L-aspartate, and (c, d) L-lysine. The theoretical enzyme binding configuration and residues are based on aspartate ammonia lyase (Fibriansah *et al*. 2011).

4.2. Production of adipic acid and analogues *via* the reverse β-oxidation pathway

Functional reverse β-oxidation for adipic acid biosynthesis was first implemented by engineering *E. coli* (Yu *et al*. 2014), with other examples following thereafter (Babu *et al*. 2015; Kallscheuer *et al*. 2017a; Zhao *et al*. 2018). Reverse β-oxidation utilizes intermediates in the central metabolism as building blocks to construct a 6-carbon backbone and to build adipic acid after a number of biotransformation steps. The first three steps of reverse β-oxidation are catalyzed by the same enzymes that are often involved in phenylacetate metabolism (Teufel *et al*. 2010). In the phenylacetate metabolism, metabolites enter the tricarboxylic (TCA) cycle after being converted to succinyl-CoA and acetyl-CoA. Instead, reverse β-oxidation, utilizes succinyl-CoA and acetyl-CoA as starting materials to drive the last three steps of phenylacetate metabolism in the opposite direction. The last step of phenylacetate metabolism is thiolytic cleavage of 3-oxoadipyl-CoA (Teufel *et al*. 2010) and, therefore, the first step of reverse β-oxidation is a non-decarboxylative Claisen condensation to form 3-oxoadipyl-CoA from succinyl-CoA and acetyl-CoA (Cheong *et al*. 2016). A few enzymes are known to accomplish such condensation reaction (**Table 2**). Using *E. coli* as the host chassis, overexpression of 3-oxoadipyl-CoA thiolase (E.C. 2.3.1.174) encoded by *E. coli paaJ* (Babu *et al*. 2015; Cheong *et al*. 2016; Kallscheuer *et al*. 2017a; Yu *et al*. 2014) or its homolog encoded by *Tfu_0875* from *Thermobifida fusca* (Zhao *et al*. 2018) achieved successful 3-oxoadiypl-CoA biosynthesis. *T. fusca* is also capable of carrying out the same reaction *in vivo* owing to the enzymatic activity of the *Tfu_0875* gene product even at basal expression level (Deng and Mao 2015). In **paper IV**, overexpression of *E. coli paaJ* in *C. glutamicum* was proven, albeit indirectly, to be functional (step 1 in **Figure 15**).

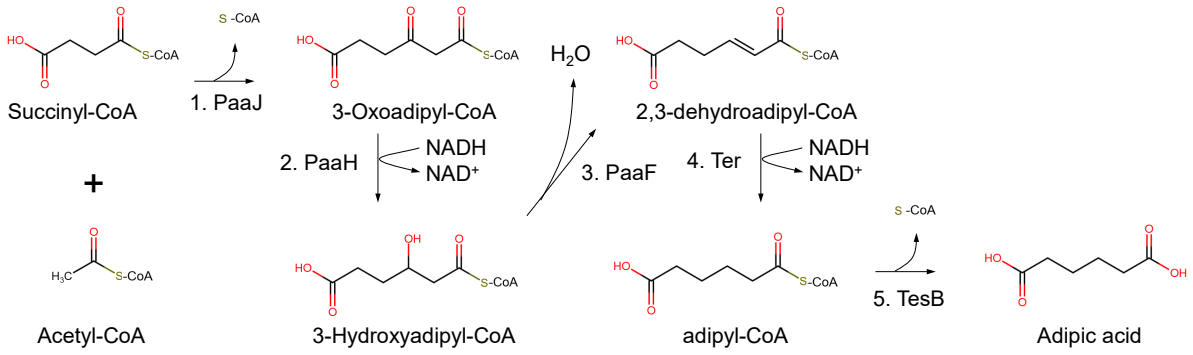


Figure 15. Reverse β-oxidation for adipic acid biosynthesis from central metabolites. The synthetic pathway implemented in *C. glutamicum* is shown. PaaJ, 3-oxoadipyl-CoA thioase; PaaH, 3-hydroxyadipyl-CoA dehydrogenase; PaaF, 2,3-dehydroadipyl-CoA hydratase; Ter, *trans*-enoyl-CoA reductase; TesB, acyl-CoA thioesterase. Figure adapted from **Figure 1** of **paper IV**.

Table 2. The enzymes used in the reverse β-oxidation pathway. Table adapted from **Table 1** from **paper I.**

Reaction	Substrate	Product	Enzyme	Donor	Note	Reference
Condensation	Acetyl-CoA and succinyl-CoA	3-oxoadipyl-CoA	Tfu_0875	*T. fusca* (native)	*T. fusca* host	Deng and Mao 2015
			Tfu_0875	*T. fusca*	*E. coli* expression	Zhao et al. 2018
			PaaJ	*E. coli* (native)	Reversible enzyme	Yu et al. 2014; Kallscheuer et al. 2017a; Babu et al. 2015; Cheong et al. 2016
Reduction	3-oxoadipyl-CoA	3-hydroxyadipyl-CoA	Tfu_2399	*T. fusca* (native)	*T. fusca*	Deng and Mao 2015
			Tfu_2399	*T. fusca*	Expression in *E. coli*	Zhao et al. 2018
			PaaH1 (3-hydroxy acyl-CoA reductase)	*R. eutropha*	Outperforms Hbd	Yu et al. 2014
			PaaH	*E. coli* (native)	pET22b-based expression	Kallscheuer 2017a; Cheong et al. 2016; Babu et al. 2015
			PaaH	*E. coli*	Expression in *P. putida*	Niu et al. 2020
Dehydration	3-hydroxyadipyl-CoA	2,3-dehydroadipyl-CoA	Crt (crotonase)	*Clostridium acetobutylicum*	Weak performance	Dekishima et al. 2011; Yu et al. 2014
			Ech (enoyl CoA hydratase)	*R. eutropha*	Outperformed and replaced Crt	Yu et al. 2014
			Tfu_0067	*T. fusca* (native)	*T. fusca*	Deng and Mao 2015
			Tfu_0067	*T. fusca*	*E. coli*	Zhao et al. 2018
			PaaZ	*E. coli* (native)	Expression in *E. coli*	Babu et al. 2015
			PaaF	*E. coli* (native)	pET22b-based expression	Kallscheuer et al. 2017a
			PaaF	*E. coli*	Expression in *P. putida*	Niu et al. 2020

Continued

Reaction	Substrate	Product	Enzyme	Donor	Note	Reference
Reduction	2,3-dehydroadipyl-CoA	Adipyl-CoA	Ter (trans enoyl CoA reductase)	*Euglena gracilis*	Expression in *E. coli*	Yu et al. 2014
			Ter	*Treponema denticola*	Also used in hexanol production	Shen et al. 2011; Tucci et al. 2007; Cheong et al. 2016
			Tfu_1647	*T. fusca* (native)	*T. fusca*	Deng and Mao 2015
			Tfu_1647	*T. fusca*	*E. coli*	Zhao et al. 2018
			DcaA (acyl-CoA dehydrogenase)	*Acinetobacter baylyi*	Outperformed Acd of *Cupriavidus necator*	Kallscheuer et al. 2016
			Acd (acyl-CoA dehydrogenase)	*Cupriavidus necator*	pET22b-based expression	Kallscheuer et al. 2016
Termination	Adipyl-CoA	Adipic acid	Ptb-buk	*Clostridium acetobutylicum*	Two step reaction *via* replacement of CoA by phosphate followed by removal of phosphate	Yu et al. 2014
			TesB (thioesterase)	*A. baylyi* ADP1	pET22b-based expression	Kallscheuer 2016
			Tfu_2577 and Tfu2576	*T. fusca* (native)	*T. fusca*	Deng and Mao 2015
			Tfu_2577 and Tfu2576	*T. fusca*	*E. coli*	Zhao et al. 2018
			TesB	*E. coli*	pET22b-based expression	Kallscheuer 2016
			Acd (acyl-CoA dehydrogenase)	*Thermococcus gammatolerans*	Replaced by TesB due to low activity	Kallscheuer 2016

The second step of reverse β-oxidation is the reduction of 3-oxoadipyl-CoA to 3-hydroxyadipyl-CoA. Reduction of 3-oxoadipyl-CoA can be carried out by various enzymes listed in **Table 2**. Overexpression of *E. coli* 3-hydroxyadipyl-CoA dehydrogenase (E.C. 1.1.1.157) encoded by *paaH* from *E. coli* (Babu *et al*. 2015; Cheong *et al*. 2016; Kallscheuer *et al*. 2017a) or its *Tfu_2399* homolog from *T. fusca* (Zhao *et al*. 2018), 3-hydroxyacyl-CoA reductase encoded by *paaH1* from *Ralstonia eutropha* (Yu *et al*. 2014), or 3-hydroxybutyryl-CoA reductase encoded by *hbd* from *Clostridium acetobutylicum* (Yu *et al*. 2014) has been reported to lead to reduction of 3-oxoadipyl-CoA in *E. coli*. Similarly, reduction of 3-oxoadipyl-CoA (step 2 in **Figure 15**) has been achieved by implementing *E. coli* 3-hydroxyadipyl-CoA dehydrogenase in *Pseudomonas putida* (Niu *et al*. 2020) or *C. glutamicum* (**paper IV**).

The next step is dehydration of 3-hydroxyadipyl-CoA to yield 2,3-dehydroadipyl-CoA (step 3 in **Figure 15**). Overexpression of the *E. coli* 2,3-dehydroadipyl-CoA hydratase gene in *E. coli* (Kallscheuer *et al*. 2017a) or *P. putida* (Niu *et al*. 2020) mediates this reaction and 2,3-dehydroadipyl-CoA can be obtained (**Table 2**). The latter can be produced in *E. coli* also by overexpressing 3-hydroxyadipyl-CoA dehydrogenase encoded by *Tfu_0067* from *T. fusca*, crotonase encoded by *crt* from *C. acetobutylicum* (Yu *et al*. 2014), putative enoyl-CoA hydratase encoded by *ech* from *R. eutropha* (Yu *et al*. 2014), or the gene product of *paaZ* from *E. coli* (Babu *et al*. 2015). Dehydration of 3-hydroxyadipyl-CoA through overexpression of *E. coli* 2,3-dehydroadipyl-CoA hydratase in *C. glutamicum* could not be fully confirmed (**paper IV**). Release of CoA from 2,3-dehydroadipyl-CoA yields *trans*-2-hexenedioic acid. While a miniscule amount of *trans*-2-hexenedioic acid was observed by LC/MS (**Figure 10**), another method, such as GC/MS or LC with tandem MS (LC-MS/MS), is required to confirm the presence of this byproduct.

Reduction of 2,3-dehydroadipyl-CoA (step 4 in **Figure 15**; **Table 2**) yields adipyl-CoA. This can be accomplished by overexpression of *trans*-enoyl-CoA reductase encoded by *ter* from *Treponema denticola* (Cheong *et al*. 2016), *bcd* from *C. acetobutylicum* (Babu *et al*. 2015), 5-carboxy-2-pentenoyl-CoA reductase encoded by *Tfu_1647* from *T. fusca* (Zhao *et al*. 2018), *trans*-enoyl-CoA reductase from *Euglena gracilis* (Dekishima *et al*. 2011; Yu *et al*. 2014), or *dcaA* from *Acinetobacter baylyi* (Kallscheuer *et al*. 2017a) in *E. coli*. In the final step, release of CoA from adipyl-CoA yields adipic acid (step 5 in **Figure 15**). Overexpression of acyl-CoA thioesterase (Acot8) from *Mus musculus* (Babu *et al*. 2015; Cheong *et al*. 2016), its homolog from *A. baylyi* (Kallscheuer *et al*. 2017a), or succinyl-CoA synthetase (*Tfu_2577*, *Tfu_2576*) from *T. fusca* has been shown to release adipic acid from adipyl-CoA in *E. coli* (**Table 2**).

Overexpression of *paaJ*, *paaH*, and *tesB* in *C. glutamicum* grown in shake-flasks allowed for 3-hydroxyadipic acid biosynthesis and its detection following excretion from the cell (**paper IV**). Derivatizing the samples and subjecting them to GC/MS analysis would provide insight on the formation of pathway intermediates. Detection of chemically derivatized 3-hydroxyadipic acid was an encouraging result as it proved that part of the pathway was functional. Non-decarboxylative Claisen condensation of acetyl-CoA and succinyl-CoA to form 3-oxoadipyl-CoA, reduction of 3-oxoadipyl-CoA, and finally release of CoA to form 3-hydroxyadipic acid have been verified for the first time in *C. glutamicum* (**Figure 16**).

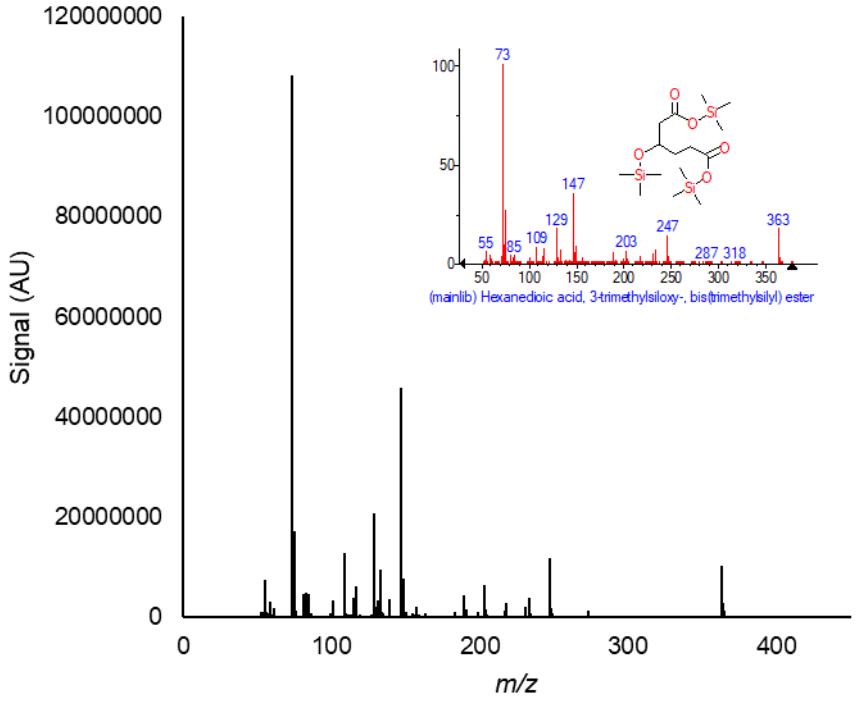


Figure 16. Fragmentation pattern of trimethylsilylated 3-hydroxyadipic acid biosynthesized by engineered *C. glutamicum*. The engineered strain harboring *paaJ, paaH,* and *tesB* was cultivated in shake-flasks. Cell-free cultivation broth was chemically derivatized and analyzed by GC/MS. Inset shows MSTFA-derivatized 3-hydroxyadipic acid NIST database entry. Figure adapted from **Figure 2** of **paper IV**.

The 3-hydroxyadipic acid precursor can be converted to adipic acid either biologically or chemically (Shi *et al*. 2020). Additionally, it can co-polymerize with other monomers to form various non-traditional polyesters and polyamides. 3-Hydroxyadipic acid synthesis requires heavy metals and strong acids (Shi *et al*. 2020). In comparison, the method described in my thesis relied on an engineered microorganism as the biocatalyst and simple sugar as the starting material (**paper IV**). Introduction of the downstream genes *paaF* and *ter* to the parental strain did not yield convincing amounts of *trans*-2-hexenedioic acid (**Figure 10**) nor adipic acid (**Figure 17**). However, detection of a small amount of adipic acid in the cultivation broth was possible after SPE treatment of samples prior to LC/MS analysis. Again, a more sensitive method, such as GC/MS or LC-MS/MS, would further verify the identity of adipic acid in the cultivation broth.

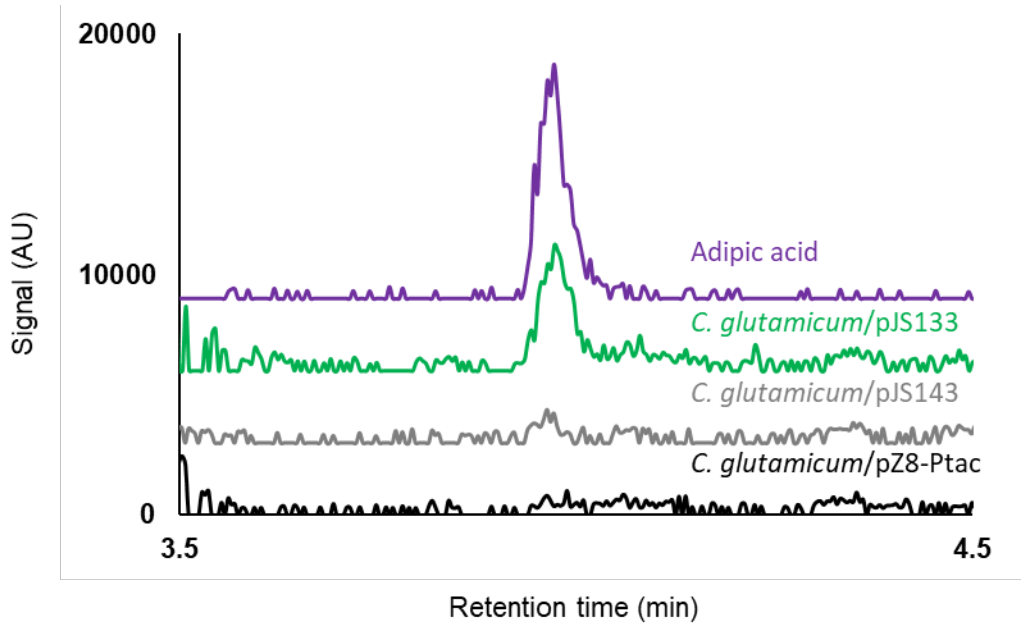


Figure 17. Adipic acid detection by LC/MS. SPE-treated cultivation broth of engineered *C. glutamicum* strains was analyzed by LC/MS (extracted ion: 145.05063). Representative chromatograms from triplicate data are shown. Purple line: adipic acid standard. Green line: *C. glutamicum* harboring the adipic acid-producing vector pJS133. Grey line: *C. glutamicum* harboring the *trans*-2-hexenedioic acid producing vector pJS143. Black line: *C. glutamicum* harboring an empty vector. Figure adapted from **paper IV**.

Prominence of 3-hydroxyadipic acid, together with only small amounts of *trans*-2-hexenedioic acid and adipic acid, in the cultivation broth of the strain harboring the full-length pathway pointed to a possible bottleneck caused by the inefficient dehydration of 3-hydroxyadipyl-CoA to 2,3-dehydroadipyl-CoA (**paper IV**). In the design of the full-length vector, pJS133, *paaJ* and *paaF* genes were organized in an operon with *paaF* downstream of *paaJ*. Downstream genes are known to be expressed more weakly than upstream genes in operon configurations (Lim *et al*. 2011). To ensure stronger flux from 3-hydroxyadipyl-CoA to 2,3-dehydroadipyl-CoA, stronger expression of *paaF* was needed. Optimized codon usage has been demonstrated to improve protein expression as well as the synthetic pathway metabolic flux (Shin *et al*. 2016). Assessment of codon bias of the constructs revealed some discrepancy with the codon usage in *C. glutamicum* (**Figure 18**). To improve pathway flux and overcome the bottleneck, a new plasmid vector with optimized codon usage was constructed, transformed into *C. glutamicum*, and evaluated by flask-cultivation. As the *C. glutamicum* harboring the improved vector failed to increase the amount of adipic acid, there is likely another reason for the inefficient conversion of 3-hydroxyadipyl-CoA to 2,3-dehydroadipyl-CoA in *C. glutamicum*. The effect of codon change in PaaF can be investigated more in detail by either individual protein assessment or re-organizing the operon and placing a promoter directly upstream of *paaF* in the construct.

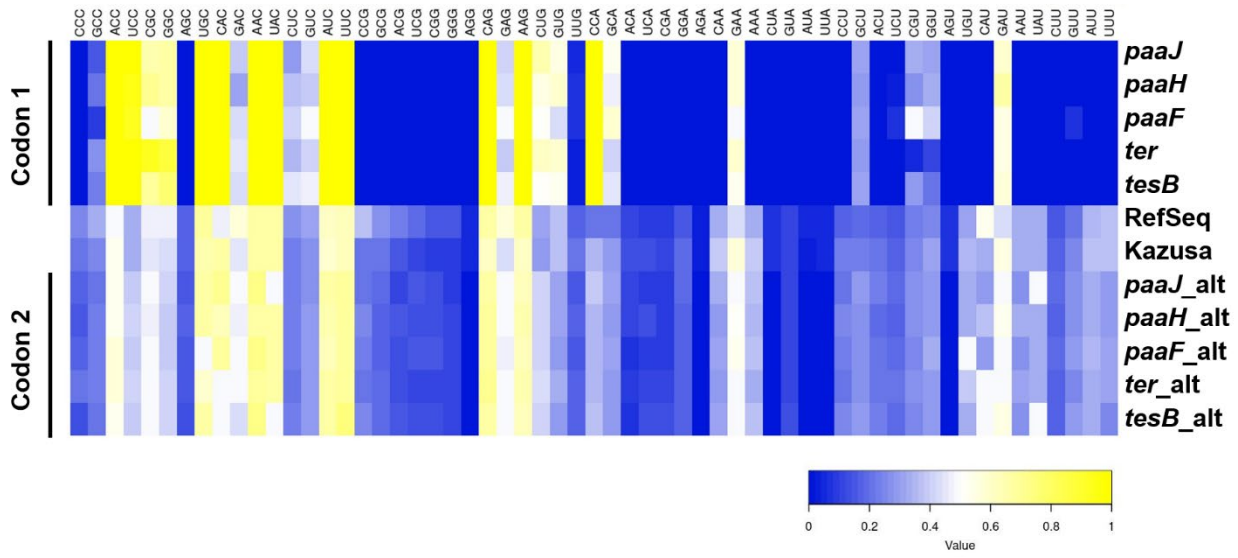


Figure 18. Codon usage analysis of genes in the adipic acid vector. The codon usage pattern of five genes involved in adipic acid production is shown (codon 1). The codon usage pattern in the *C. glutamicum* genome was analyzed from two databases, including RefSeq and Kazusa. A different plasmid vector for producing adipic acid was constructed in agreement with the codon usage pattern from the above databases (codon 2). Codons ending with G and C are organized on the left side of the heatmap. Stop codons, start codons, and the tryptophan codon are omitted for clarity.

4.3. Evaluation of *cis*,*cis*-muconic acid reduction in *C. glutamicum*

One of the most notable examples in bio-based adipic acid production involves the generation of *cis,cis*-muconic acid from glucose by engineered *E. coli* (Draths and Frost 1994). Because this engineering strategy relied on the protocatechuic acid pathway, absence of shikimate dehydrogenase (E.C. 1.1.1.25; encoded by *aroE*) was required to accumulate 3-dehydroshikimic acid. Further introduction of dehydroshikimate dehydratase (E.C. 4.2.1.118; encoded by *aroZ*) and protocatechuate decarboxylase (E.C. 4.1.1.63; encoded by *aroY*), both from *Klebsiella pneumoniae*, allowed carbon flux in the engineered *E. coli* to proceed to catechol. Finally, introduction of *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase (E.C. 1.13.11.1; encoded by *catA*) enabled production of *cis,cis*-muconic acid from glucose. A study encompassing more genetic engineering and process engineering steps was published later on (Niu *et al*. 2002). As mentioned earlier in the introduction of this thesis and in **paper I**, many other studies followed and the biosynthesis of *cis,cis*-muconic acid has been extensively investigated. Most studies have focused on chemical catalysis for the conversion of *cis,cis*-muconic acid to adipic acid.

While it has been long speculated that enzymatic reduction of *cis,cis*-muconic acid would be possible, it was considered a challenge because many enzymes capable of such reduction depended on a Fe-S cluster and required an anaerobic environment. Many bioprocesses affecting *cis,cis*-muconic acid are aerobic as degradation of aromatic compounds requires oxidation

(**papers I** and **V**). While enoate reductases from *Clostridium* species have been considered as feasible candidates, they are oxygen-sensitive. A breakthrough occurred in 2017, when oxygen-tolerant enoate reductases were reported to reduce *cis,cis*-muconic acid (Joo *et al*. 2017). Particularly, enoate reductase from *Bacillus coagulans* or *C. acetobutylicum* not only reduced *cis,cis*-muconic acid, but reduced also *trans*-2-hexenedioic acid to adipic acid; whereas enzymes belonging to the Oye family were not able to catalyze the same reduction steps. Because the reasons for such reaction disparity were not provided in that study, I have investigated the differences between Oye members and enoate reductases (**paper V**). Molecular docking of *cis,cis*-muconic acid or *trans*-2-hexenedioic acid to enoate reductases from *B. coagulans*, *C. acetobutylicum*, *Clostridium kluyveri*, *Clostridium ljungdahlii*, and *Moorella thermoacetica* proceeded in a seamless way. Interestingly, not all members of the Oye family were able to accommodate *cis,cis*-muconic acid or *trans*-2-hexenedioic acid, as some of them did not possess a pocket shape that would correctly fit *cis,cis*-muconic acid for reduction (Supplementary information in **paper V**). Further, it was found that enoate reductases were much larger than Oye members and there was only homology between the enoate reductases and the TIM barrel domain of Oye members. Additionally, the enzyme pocket of Oye members was much more exposed to the solvent; whereas that of enoate reductases was buried deeply inside. It was further speculated that the FAD cofactor-binding domain was displayed only by enoate reductases and not Oye family members.

4.4. Conclusions on synthetic pathways for adipic acid biosynthesis

Multiple pathways exist for adipic acid biosynthesis, and each is met with different challenges. Properly functioning L-lysine pathway would allow adipic acid biosynthesis from glucose when the pathway is introduced into a L-lysine overproducing microorganism. However, a metabolic gap exists in the pathway since not all steps have been well characterized. Examining the MIO catalyzed reactions revealed that mechanism *via* E2 is likely necessary for deamination of L-lysine. If the MIO-catalysis does not proceed *via* E2, then focus should shift to an alternative pathway *via* β-D-lysine.

Since adipic acid production from glucose with reverse β-oxidation pathway has not been efficient in *E. coli*, *C. glutamicum* was explored and adipic acid biosynthesis was demonstrated. While further engineering must be done to improved titer, a weak signal was observed and indirectly indicated that the pathway was functioning.

Molecular docking studies of oxidoreductases revealed that enoate reductases instead of Oye members should be investigated for follow up studies. The enoate reductase have the cofactors and sufficient reducing power for reduction of *cis,cis*-muconic acid and *trans*-2-hexenedioic acid.

5. Host chassis for adipic acid biosynthesis

While *E. coli* and *S. cerevisiae* are often employed for genetic manipulation, they are not always the best choice for producing the desired compound(s). Each microorganism has adapted to its own niche and has developed different characteristics and capabilities. Choosing the correct host strain is as important as choosing the right enzymes and pathways (**paper I**).

5.1. Yeast

Yeasts are some of the model organisms most often selected for genetic manipulation, and production of adipic acid in yeast has been contemplated by both academic and industrial laboratories (Beardslee and Picataggio 2012; Picataggio *et al*. 1992; Raj *et al*. 2018; Sandström *et al*. 2014). Verdezyne, one of the most anticipated adipic acid biotechnology companies of the past decade, attempted to produce adipic acid from palm oil feedstock (Beardslee and Picataggio 2012). The host organism was *Candida tropicalis* and the technology was built upon ω-oxidation of fatty acids (Picataggio *et al*. 1992). In the laboratory, this approach was demonstrated using coconut oil as feedstock (Beardslee and Picataggio 2012); whereas in industrial settings, implementation was based on crude palm oil. Having demonstrated a 50 g/L titer of adipic acid (Beardslee and Picataggio 2012), a new pilot plant was nearly 80–85% complete when the project was unfortunately discontinued (**paper I**). Although this event clearly reflected the challenges of bio-based production, efforts to produce adipic acid continued within academia (Ju *et al*. 2020). Taking a different approach, an engineered *S. cerevisiae* strain was shown to produce adipic acid *via* the *cis*,*cis*-muconic acid pathway using glucose as feedstock (Raj *et al*. 2018). While production of *cis*,*cis*-muconic acid from renewable feedstock had been demonstrated multiple times, its conversion to adipic acid had relied heavily on chemical processes. Eventually, Mahadevan and colleagues discovered an enoate reductase (Joo *et al*.

2017) and implemented it in *S. cerevisiae* for successful production of adipic acid. While the titer remains at the sub-g/L scale, this example demonstrates ongoing efforts to develop petroleum-independent production of adipic acid. Shifting between aerobic and anaerobic phases during cultivation, as was done in that study (Joo *et al*. 2017), could potentially overcome the enoate reductase inactivity observed in **paper V**.

5.2. Thermobifida

Even though adipic acid occurs rarely in nature, a microbial pathway that produced adipic acid was first reported in *T. fusca* in 2015 (Deng and Mao 2015). This metabolic pathway was observed while trying to produce 1-propanol (Deng and Fong 2011) and was identified as reverse β-oxidation. It consisted of five biochemical steps starting from the central metabolism and involved six genes: *Tfu_0875*, *Tfu_2399*, *Tfu_0067*, *Tfu_1647*, *Tfu_2577*, and *Tfu_2576*. Implementing the *Tfu_0067* gene in *C. glutamicum* could address the bottleneck caused by the dehydration reaction in **paper IV**. The above six genes were later implemented in *E. coli* where they further improved adipic acid biosynthesis (Zhao *et al*. 2018).

5.3. Pseudomonads

Pseudomonas putida and other Pseudomonads have been known to carry out soil detoxification and bioremediation through degradation of aromatic compounds. Production of *cis,cis*-muconic acid has been a notable example of the application of *P. putida*. Niu and colleagues, who pioneered the engineering of *E. coli* for *cis,cis*-muconic acid production, have recently succeeded in generating adipic acid from engineered *P. putida* (Niu *et al*. 2020). The starting aromatic compounds, including protocatechuic acid and catechol, were converted to 3-ketoadipic acid *via* *cis,cis*-muconic acid. Specifically, 3-ketoadipic acid was converted *via* the reverse degradation pathway in a manner akin to that described in **paper IV**. A similar approach was attempted during preliminary studies carried out in **paper V**, whereby strains developed in **paper IV** were tested for their capacity to convert benzoic acid to adipic acid.

5.4. *Corynebacterium*

A growing wealth of genomic knowledge has allowed to move beyond traditional model organisms, such as *E. coli* and *S. cerevisiae*, to more diverse targets. As a result, *C. glutamicum* has become one of the most well characterized model organisms for metabolic engineering (Becker and Wittmann 2015; Becker and Wittmann 2012; Lee *et al*. 2016). Originally discovered in Japan, *C. glutamicum* has been used as a workhorse for the production of amino acids for more than half of the previous century. Indeed, global production of L-glutamate and L-lysine for the food industry relies mainly on *C. glutamicum*, which generates 2.4 million tons of L-lysine and 3.1 million tons of L-glutamate annually (Lee and Wendisch 2017). Rational engineering of *C. glutamicum* for dietary supplement production has been mostly avoided,

perhaps due to regulation concerning the use of GRAS organisms. In the industry, strain improvement has been based mostly on random mutagenic methods followed by massive screening. In contrast, in academia, the preferred approach has been that of selective genetic engineering.

Another reason for the establishment of *C. glutamicum* in academic research could be linked to its exceptional capability to overproduce primary metabolites. Genetic engineering or process engineering strategies have been applied to *C. glutamicum* to separately produce enormous amounts of primary metabolites and amino acids, including succinic acid, L-ornithine (Meiswinkel *et al*. 2013), L-arginine, L-glutamate (Yang *et al*. 2015), L-lysine (Becker *et al*. 2011), L-alanine (Yamamoto *et al*. 2012), L-valine (Hasegawa *et al*. 2012), ethanol (Inui *et al*. 2005), L-lactate (Tsuge *et al*. 2019), D-lactate (Okino *et al*. 2008), and ectoine (Becker *et al*. 2013). The production titers for these compounds are impressive: L-lysine·HCl 120 g/L; L-arginine, 92.5 g/L; L-glutamate 116 g/L; L-valine, 150 g/L; L-alanine, 216 g/L; and D-lactate, 120 g/L. The growing number of engineering tools will further facilitate manipulation of the *C. glutamicum* genetic make-up, expanding the construction of synthetic pathways.

Exploiting the powerful metabolism of *C. glutamicum* combined with new synthetic pathways is posed to increase production of various chemicals. Cadaverine, 5-aminovaleric acid, and glutaric acid are notable examples of L-lysine-derived compounds. Implementing a single gene, *cadA*, in *C. glutamicum* has enabled very efficient production of cadaverine (Kind and Wittmann 2011). Similarly, introduction of the genes *davB*, and *davA* has led to successful production of 5-aminovaleric acid (Rohles *et al*. 2016). Finally, implementation of the genes *davT* and *davD* (or equivalent) has resulted in successful production of glutaric acid *via* L-lysine (Rohles *et al*. 2018). As *C. glutamicum* is also known to produce vast amounts of L-glutamate, the *gadB* (or equivalent) gene has been introduced to successfully boost production of gamma-aminobutyrate (Baritugo *et al*. 2018). Moreover, as *C. glutamicum* can be modified to generate L-ornithine, the *speC* gene has been introduced to enable production of putrescine (Schneider and Wendisch 2010). Application of *C. glutamicum* for adipic acid was investigated in **papers VI** and **V**.

5.5. *E. coli*

As *E. coli* is one of the most genetically amenable microorganisms, it has served as a host for producing a myriad of chemicals (Choi *et al*. 2016). Not surprisingly, the first engineered organism for biological production of adipic acid was *E. coli* (Yu *et al*. 2014), in which the introduction of genes responsible for reverse β-oxidation generated sub-mg/L amounts of adipic acid from glucose. Several attempts followed to further increase production of adipic acid through reverse β-oxidation in *E. coli* (Babu *et al*. 2015; Clomburg *et al*. 2015; Dellomonaco *et al*. 2011). Notably, glycerol proved much more effective than glucose as feedstock, affording greater production titers (Zhao *et al*. 2018). Moreover, reverse β-oxidation genes from *Thermobifida* proved to be very effective. While glycerol, a sugar alcohol, cannot be directly obtained from lignocellulosic materials, production of elevated titers of adipic acid in *E. coli*

from glycerol demonstrated that this could be achieved independently of petroleum. Furthermore, production of adipic acid *via cis*,*cis*-muconic acid has been demonstrated in *E. coli* using an enoate reductase from *C. acetobutylicum* (Sun *et al*. 2018). Introducing enoate reductase in *E. coli* generated nearly 10 times more adipic acid (Sun *et al*. 2018) than in yeast (Raj *et al*. 2018). Even though adipic acid production remained at the sub-g/L scale, this experiment demonstrated the possibility of more than one metabolic pathway to produce adipic acid from a simple sugar such as glucose.

6. Current production of chemicals and compounds using microbes

Technologies for generating adipic acid have not reached a point, whereby industrial-level production using renewable feedstock is feasible. The combination of adipic acid production and microbial cell factories is still in its infancy. However, technology for producing other chemicals has already matured to an industrial scale. Some examples are discussed hereafter, with the aim of showing where adipic acid research and production should proceed.

6.1. Successful examples of cell factories producing chemicals and compounds

Whereas 1,4-BDO was initially derived only from petroleum, it is now partially produced by bacteria. Biobased 1,4-BDO is used for the commercial manufacturing of compostable plastic bags, which are hugely popular in Europe (Schilling 2019). From an initial biological production of 18 g/L (Yim *et al*. 2011), continuous strain improvement (Bart and Cavallaro 2015; Barton *et al*. 2014; Beardslee and Picataggio 2012; Burgard *et al*. 2016; Burgard *et al*. 2003; Burk 2010; Challener 2014) has now led to an estimated output of 200,000–600,000 L of 1,4-BDO from glucose (Wehrs *et al*. 2019), which makes up a significant chunk of the existing 1.3 million ton global market (Burk 2010). This example demonstrates that building operational bio-refineries can lead to the production of bio-based compounds with real world applications. Many similar examples will follow in the future and, although there has not been another success story in years, I envision one day commercially viable petroleum-free adipic acid production.

6.2. Production using combined bio- and chemical-conversion

Not all steps in bio-based production must be carried out biologically. In certain instances, processes that combine biological and chemical approaches can effectively generate chemicals from renewable resources. Bioplastics that rely on at least one chemical step in the production scheme have already been commercialized. In this category of chemicals, which includes bioethanol or polylactic acid (PLA), the monomers or precursor molecules are produced biologically, and subsequent conversion to end-products is carried out by chemical means.

Just as fuel production does not take up the entire petroleum utilization, bioethanol serves both as fuel and as a platform chemical for the manufacturing of various consumer products. Production of bio-based ethanol is driven largely by the bio-energy sector, as well as the cooperation between governments and the energy industry. The two largest producers are the U.S.A. and Brazil, and the primary feedstock is sugarcane, switch grass and maize. Ethanol produced during conventional yeast fermentation, can undergo additional processes including dehydration. Chemical dehydration of ethanol yields ethylene, which is the monomer required in the manufacturing of high- or low-density polyethylene (PE). Sometimes, downstream processes can be local, *i.e*., chemical conversion is carried out near the ethanol production site. This is the case for the 'I'm green™' PE products from Braskem, a Brazilian company, which are produced from lignocellulosic ethanol. Localizing ethylene production at the site of yeast fermentation of sugarcane simplifies the logistics and shipping, driving down costs. Lignocellulosic ethanol can be further converted to ethylene glycol through oxidation of ethylene. Ethylene glycol is a constituent of polyethylene terephthalate (PET) bottles. By replacing petroleum-based ethylene glycol with lignocellulosic ethylene glycol, 30% (w/w) of a PET bottle is of biological origin. This is the case of 'Plant bottles' advertised by the Coca-Cola Company. Various co-polymers, which are made of more than one monomer, can partially originate from petroleum-based and bio-based feedstock (see **section 6.3**).

Another example of a product that results from combining biological and chemical processes is biodegradable plastic generated from lactic acid as the monomer. Lactic acid fermentation is well known and production titers above 200 g/L have been reported (Meng *et al*. 2012b). In practice, biologically produced lactic acid can be polymerized to PLA in a chemical process known as ring opening polymerization. A similar strategy is employed by NatureWorks to produce 150,000 tons of PLA annually (Vink and Davies 2015). PLA is currently used in variety of consumer products, including coffee capsules, utensils, tea bags, food packaging, and many others.

6.3. Mixing bio-based and petroleum-based resources

Products of a partially chemical and partially biological origin encompass various compounds, including the polyamides nylon-5,10, nylon-4,10, and nylon-6,10. These polymers require a diamine and a diacid for their production. While theoretically it should be possible, production using only bio-based monomers has not yet been realized. Nylon-5,10 is a polyamide comprised

of 1,5-diaminopentane (also known also as cadaverine) and decanedioic acid (known also as sebacic acid) (Kind *et al*. 2014). It is produced from petroleum-derived sebacic acid and bio-based cadaverine and is employed for example for toothbrush production in China (Kyulavska *et al*. 2019). Cadaverine fermentation has been widely successful in Asia and Europe using engineered *C. glutamicum* strains. Sebacic acid, in theory, can be produced biologically through microbial castor oil fermentation (Jeon *et al*. 2019). Accordingly, nylon-5,10 could be manufactured entirely from biological feedstocks (Kind *et al*. 2014). Instead, biologically produced sebacic acid has been utilized mostly with 1,4-diaminobutane (putrescine) for the synthesis of nylon-4,10. In practice, castor oil is converted to sebacic acid and polymerized with petroleum-derived putrescine. This is commercially sold under the brand name EcoPaXX® by the Dutch company DSM. While the use of *C. glutamicum* for putrescine production has been quite successful (Schneider and Wendisch 2010), the technology has not been taken up by the industry to the best of my knowledge. Successful technology transfer of biological putrescine production would allow the generation of nylon-4,10 entirely from biological feedstock. Another commercialized product consisting of bio-based sebacic acid is nylon-6,10, sold under the name Ultramid® by BASF. In this case, 1,6-diaminohexane is petroleum-derived and sebacic acid is biologically derived.

7. Conclusion

In my thesis work, enzymatic reactions in multiple metabolic pathways leading to adipic acid biosynthesis were investigated. Particularly, I investigated the enzymatic components encompassing the L-lysine pathway (**papers II** and **III**), the reverse β-oxidation pathway (**paper IV**), and *cis*,*cis*-muconic acid reduction (**paper V**). Selecting the pathway, host organism, and enzymes for adipic acid biosynthesis was addressed in **paper I**. Investigation of combining different pathways and host strains was addressed in **papers IV** and **V**. Various tools including the recombinant technology (**papers IV** and **V**), analytical chemistry (**papers IV** and **V**), and computational chemistry (**papers II**, **III**, and **V**) were used in my thesis.

Comprehensively mapping out the enzyme components and metabolic pathways leading up to adipic acid revealed similarities and differences in metabolic engineering strategies (**paper I**). Glycerol and palm oil, rather than glucose, were observed as preferred carbon sources. Multiple pathways including the alternative L-lysine pathway converged at *trans*-2-hexenedioic acid (**papers I** and **II**). In **paper II**, the binding pocket of Oye1 and *E. coli* NemA were determined not to be the limiting factor for possible α,β-reduction of 6-amino-*trans*-2-hexenoic acid and *trans*-2-hexenedioic acid. Both Oye1 and NemA were found capable of forming hydrogen bonds with carboxylic acid substrates and binding to them in the correct orientations. In **paper III**, investigation of the deamination reaction illustrated that L-lysine might not be the best intermediate for α,β-elimination-assisted conversion. Instead, β-D-lysine was found to be a more suitable option, potentially allowing for the generation of 6-amino-*trans*-2-hexenoic acid.

In **paper IV**, *C. glutamicum* was determined as a feasible production host for implementation of the reverse β-oxidation pathway. Enzymes involved in the reverse β-oxidation was identified and compared in **paper I** and selected components were implemented in **paper IV**. Non-decarboxylative condensation of succinyl-CoA and acetyl-CoA was achieved for the first time *in vivo* in *C. glutamicum* to the best of my knowledge. Biosynthesis of 3-hydroxyadipic acid in engineered *C. glutamicum* was observed. 3-hydroxyadipic acid was identified as a potentially useful platform chemical for establishing the bio-based industry concept. Identifying

analytical methods capable of quantifying tiny amounts of *trans*-2-hexenedioic acid and adipic acid was very challenging throughout. Near the end of the thesis work, a weak adipic acid signal was observed by LC/MS, yet it was considered insufficient to provide convincing proof of successful adipic acid biosynthesis. A more sensitive method such as LC-MS/MS would be necessary to confirm the presence of adipic acid with greater certainty.

In **paper V**, the investigation surrounding the reduction of *trans*-2-hexenedioic acid covered in **paper II** was revisited. Benzoic acid conversion to *cis,cis*-muconic acid was confirmed. Computational tools were used to investigate Oye family members and enoate reductases. Major differences between the two families were identified. Putative reasons for the Oye family members lacking the capability to reduce *trans*-2-hexenedioic acid were elucidated and the enoate reductase family was determined to be more feasible for *trans*-2-hexenedioic acid reduction.

Investigation of enzymes in multiple adipic acid pathways revealed new opportunities for biosynthesis of adipic acid. Theoretical consideration for L-lysine pathway is now more developed than when it was first introduced. New alternative routes for utilizing L-lysine as intermediate is now laid out using either β-D-lysine or aminoadipic acid. *C. glutamicum* can be considered as a new host organism to produce adipic acid in addition to yeast, *P. putida*, *T. fusca*, and *E. coli*. Functionality of the reverse β-oxidation pathway in *C. glutamicum* is demonstrated and can be further improved.

8. Outlook

The existing dependence on petroleum-based industrial production, together with its standards and practices, has stimulated numerous technological advances aimed at the renewable production of platform chemicals for consumer product manufacturing. In such a circular economy scenario, plants capture and convert atmospheric carbon dioxide into simple sugars, which are then further transformed to platform chemicals.

The use of engineered organisms as microbial cell factories has already resulted in the production of a handful of chemicals, such as bioethanol and 1,4-BDO. Many technical challenges have already been overcome, allowing technology readiness to reach a higher level. Nevertheless, many challenges still lie ahead, especially with respect to expanding the catalogue of chemicals amenable to renewable and sustainable production by other industrial or academic entities. One of the remaining challenges is to create suitable microorganisms capable of efficiently producing adipic acid. Recent bankruptcy of an adipic acid-producing company, Verdezyne, showcases the intrinsic difficulties of developing such technology and processes, and bringing them to market. Strengthening cooperation between governments and the industry could avoid another setback. Research must be continued not only by identifying better enzymes and metabolic pathways, but also by exploring more types of substrates. Use of fatty acids as substrates in adipic acid production provided a new opportunity for international cooperation as shown also by the location of the Verdezyne plant in Malaysia, a leading palm oil producer. I personally wonder whether a different strategy with cheaper substrate and different investor partnerships would have faced the same destiny. This is not to criticize the management or the technology of said venture, but to envisage a better future for bio-based adipic acid production, where opportunities are available on multiple fronts. Multiple pathways, host strains, and substrates have already been documented but have not been implemented at the industrial level. Seeing multiple options in the commercial market is the ultimate goal. My thesis addresses such opportunities by investigating various pathways and host combinations. Continued development of bio-based adipic acid production would entail further characterization of the L-lysine

pathway, improvement of the reverse β-oxidation pathway, and investigation of enoate reductase mechanisms.

A deeper theoretical understanding of α,β-elimination of β-D-lysine is needed to further assess whether the chance of elimination is indeed better than for L-lysine. This requires better understanding of MIO-mediated enzymatic processes and, ultimately, the determination of whether the reaction proceeds *via* the E2 mechanism. Only if the E2 mechanism is confirmed, can L-lysine be considered as an interesting compound for further studies. If E2 is not identified as the reaction mechanism for MIO-mediated enzymes, then β-D-lysine should be investigated instead. Similarly, the alternative pathway presented in **paper II** involves L-2-aminoadipic acid as one of the intermediates. The amino group on aminoadipic acid might be more feasible for elimination if located in the β position rather than in the α position. Thus, future studies may include but should not be limited to D-3-aminoadipic acid as a potential intermediate. Successful elimination of the amino group from aminoadipic acid yields *trans*-2-hexenedioic acid, which is a recurring intermediate in the reduction of *cis,cis*-muconic acid.

While computational and structural aspects of enoate reductases have been provided, more experimental results are required. Reduction of *trans*-2-hexenedioic acid in *C. glutamicum* should be investigated under additional cultivation conditions. Shifting between aerobic and anaerobic conditions during cultivation should be also investigated. Moreover, enoate reductase expression needs to be confirmed in *C. glutamicum*. It is still unknown whether enoate reductases can functionally mature in *C. glutamicum* under the conditions tested. As analytics were challenging, improving adipic acid titer will benefit detection and verification. SPE, too, would benefit from further optimization. While reduction of *trans*-2-hexenedioic acid should be investigated further, investigation of 6-amino-*trans*-2-hexenoic acid reduction requires starting from anew. The reaction demonstrated in U.S. patent 7,491,520 could not be experimentally confirmed in this work. As demonstrated in **paper V**, the differences in structure and reducing power between Oyes and enoate reductases, point to the use of the latter rather than the former in future studies. Additionally, results provided in **paper V** cast some doubt on the validity of the patent's claims. While some Oye members can bind to the carboxylic group, reducing power is greater in enoate reductases. Additional enzyme candidates should be explored to identify those with greater oxygen tolerance.

Biosynthesis of adipic acid *via* reverse β-oxidation in *C. glutamicum* was technically achieved, but more challenges lie ahead to improve the titer. First, identification of *trans*-2-hexenedioic acid and adipic acid produced by engineered strains using more qualitative analytical methods such as LC-MS/MS is required. Alternatively, the SPE protocol can be optimized to accommodate a larger cultivation volume, thus increasing the adipic acid concentration for analysis and enhancing the removal of impurities. Second, degradation of 3-oxoadipyl-CoA needs to be blocked by deletion of the *pcaF* gene. Overall, genome engineering as well as pathway improvement are needed. Higher titers will improve the analytical steps as a higher concentration of adipic acid augments the chance for correct identification.

If theoretically given an opportunity to continue with the project, I would have taken the following steps to ameliorate adipic acid production by *C. glutamicum*: 1) validate adipic acid

identity, 2) establish intellectual property, 3) bring in industrial interest for expedited development, 4) increase adipic acid titer, yield, and productivity, and 5) establish a demonstration plant operation.

Finally, multiple vendors already exist for commercial bioethanol. The same needs to be true also for bio-based adipic acid to ensure a sustainable future.

9. Acknowledgements

To **Lisbeth**, thanks for your flexibility, honesty, leadership, and positivity. I'm very grateful for your patience in the uncertain moments throughout the Ph.D. journey. You've given me a second chance at Ph.D. and provided support no matter the circumstances. I cannot thank you enough. I hope you get well soon.

To **Calle**, I appreciate your educational philosophy, patience, and fairness. Thank you for keeping me company at the kårhuset during lunch time and for being an awesome examiner.

To **Leif**, thank you for hosting me at your laboratory for the computational work. It was pleasant having chats during lunch time.

To **Emma**, **Vero**, and **Vale**, it was fun working with you all for the L-lysine projects. I wish you all the best in your future projects and positions.

To **Aaron**, it was pleasure working with you albeit virtually because of the new circumstances we all faced. Thanks for your expertise and discussions.

To **Sonali** and **Antonio**, it was amazing during my time at the Lundberg lab thanks to you guys. I am forever in your debt for your help and friendship. Thanks for all the good memories we shared.

To the **Hedberg** family (**Karin, Greger** and **Jonas**), thank you for hosting me during my first and a half year in Sweden. I shared such pleasant moments thanks to your generosity. I will never forget you.

To **Jeroen** and **Fabio**, you two have been simultaneously terrible and amazing. Thanks for all the good memories. Thanks for all the support not only during normal times, but especially during the pandemic.

To **Punchalee** and **Thenia**, thank you for not only running the lab smoothly but also for providing support when needed. Especially, Pun, thank you so much for all the help during the pandemic.

To the 'hey google' group (**Elena**, **Dan**, **Andrea**, and others), thank you for all the pizzas, movies, and laughter we shared. Thanks to all of you, the long winters of Sweden were spent joyfully. Thanks for the good memories in Kiruna as well.

To **Fredrik** and **Vera**, you guys have been the mature and sociable individuals in the group whom I always had the pleasure to talk to. I will miss our conversations very much.

To **Peter**, thank you so much for introducing the DTU core facilities to me. Extremely crucial data in my thesis was obtained thanks to this suggestion. I cannot thank you enough.

To **Puck**, it was very fun working with you in the lab. You should be very proud of the work you have done here. Good luck with rest of your studies.

To **Clarisse**, you did such an amazing job with computational studies. It's a shame that we never got to work together in person due to the special circumstances we found ourselves in. Good luck with rest of your studies.

To **Amanda**, **David**, **Thenia**, **Julia**, **Venky**, **Silfa**, **Suzanna**, and the **rest of members** in the lab, thank you all so much for being great colleagues. Stay strong and healthy.

To my master's thesis advisors, **Bill** and **Chuck**, thank you for encouraging me to pursue a higher degree though it did not work out for me the first time. Your wisdom and scientific rigor continue to inspire me.

To my wife, **Mika**, thank you for always putting up with me and dealing with my tantrums. You have supported me throughout the uncertain moments in the past year and I cannot thank you enough.

To my brother, **Steven**, thanks for being a great brother and always being on my side. Thanks for being a good son to mom and dad while I'm overseas.

To my beloved **parents**, thank you always for supporting me. It has been 10 years since I embarked on this uncertain journey of doctoral studies. Thank you for supporting me regardless of what life decisions I have made.

10. References

Averesch NJH, Martínez VS, Nielsen LK, Krömer JO. Toward Synthetic Biology Strategies for Adipic Acid Production: An in Silico Tool for Combined Thermodynamics and Stoichiometric Analysis of Metabolic Networks. *ACS Synth. Biol.* 2018;7(2):490–509.

Babu T, Yun EJ, Kim S, Kim DH, Liu KH, Kim SR, *et al*. Engineering *Escherichia coli* for the production of adipic acid through the reversed β-oxidation pathway. *Process Biochem.* 2015;50(12):2066–71.

Baneyx F. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 1999;10(5):411–21.

Bang SG, Choi WJ, Choi CY, Cho MH. Production of *cis,cis*-muconic acid from benzoic acid via microbial transformation. *Biotechnol. Bioprocess Eng.* 1996;1(1):36–40.

Baritugo KA, Kim HT, David Y, Khang TU, Hyun SM, Kang KH, *et al*. Enhanced production of gamma-aminobutyrate (GABA) in recombinant *Corynebacterium glutamicum* strains from empty fruit bunch biosugar solution. *Microb. Cell Fact.* 2018;17(1).

Bart JCJ, Cavallaro S. Transiting from adipic acid to bioadipic acid. Part II. Biosynthetic pathways. *Ind. Eng. Chem. Res.* 2015;54(2):567–76.

Barton NR, Burgard AP, Burk MJ, Crater JS, Osterhout RE, Pharkya P, *et al*. An integrated biotechnology platform for developing sustainable chemical processes. *J. Ind. Microbiol. Biotechnol.* 2014;42(3):349–60.

Beardslee T, Picataggio S. Bio-based adipic acid from renewable oils. *Lipid Technol.* 2012;24(10):223–5.

Becker J, Schäfer R, Kohlstedt M, Harder BJ, Borchert NS, Stöveken N, *et al*. Systems metabolic engineering of *Corynebacterium glutamicum* for production of the chemical chaperone ectoine. *Microb. Cell Fact.* 2013;12(1).

Becker J, Wittmann C. Bio-based production of chemicals, materials and fuels - *Corynebacterium glutamicum* as versatile cell factory. *Curr. Opin. Biotechnol.* 2012 Aug;23(4):631–40.

Becker J, Wittmann C. Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angew. Chemie - Int. Ed.* 2015;54(11):3328–50.

Becker J, Zelder O, Häfner S, Schröder H, Wittmann C. From zero to hero-Design-based systems metabolic engineering of *Corynebacterium glutamicum* for l-lysine production. *Metab. Eng.* 2011;13(2):159–68.

Blancquaert D, Van Daele J, Strobbe S, Kiekens F, Storozhenko S, De Steur H, *et al*. Improving folate (Vitamin B 9) stability in biofortified rice through metabolic engineering. *Nat. Biotechnol.* 2015;33(10):1076–8.

Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* 2009;4(1):1–13.

Bordwell FG. Equilibrium acidities in dimethyl sulfoxide solution. *Acc. Chem. Res.* 1988 Dec 1;21(12):456–63.

Bozell JJ. Connecting biomass and petroleum processing with a chemical bridge. *Science.* 2010;329(5991):522–3.

Brenna E, Cosi SL, Ferrandi EE, Gatti FG, Monti D, Parmeggiani F, *et al*. Substrate scope and synthetic applications of the enantioselective reduction of α-alkyl-β-arylenones mediated by Old Yellow Enzymes. *Org. Biomol. Chem.* 2013;11(18):2988.

Brenna E, Gatti FG, Monti D, Parmeggiani F, Sacchetti A, Valoti J. Substrate-engineering approach to the stereoselective chemo-multienzymatic cascade synthesis of *Nicotiana tabacum* lactone. *J. Mol. Catal. B Enzym.* 2015;114:77–85.

Brunk E, George KW, Alonso-Gutierrez J, Thompson M, Baidoo E, Wang G, *et al*. Characterizing Strain Variation in Engineered E. coli Using a Multi-Omics-Based Workflow. *Cell Syst.* 2016;2(5):335–46.

Burgard A, Burk MJ, Osterhout R, Van Dien S, Yim H. Development of a commercial scale process for production of 1,4-butanediol from sugar. *Curr. Opin. Biotechnol.* 2016;42:118–25.

Burgard AP, Pharkya P, Maranas CD. OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol. Bioeng.* 2003;84(6):647–57.

Burk MJ. Sustainable production of industrial chemicals from sugars. *Int. Sugar J.* 2010;112(1333):30–5.

Callaway E. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. *Nature.* 2020; 588: 203-204.

Challener C. Bioproducts Genomatica BDO process recognised. *Chem. Ind.* 2014;78(5):6.

Chen S, Jia N, Ding MZ, Yuan YJ. Comparative analysis of l-sorbose dehydrogenase by docking

strategy for 2-keto-l-gulonic acid production in *Ketogulonicigenium vulgare* and *Bacillus endophyticus* consortium. *J. Ind. Microbiol. Biotechnol.* 2016;43(11):1507–16.

Cheng J, Hu G, Xu Y, Torrens-Spence MP, Zhou X, Wang D, *et al*. Production of nonnatural straight-chain amino acid 6-aminocaproate via an artificial iterative carbon-chain-extension cycle. *Metab. Eng.* 2019;55:23–32.

Cheong S, Clomburg JM, Gonzalez R. Energy-and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. *Nat. Biotechnol.* 2016;34(5):556–61.

Choi KR, Shin JH, Cho JS, Yang D, Lee SY. Systems Metabolic Engineering of *Escherichia coli*. *EcoSal Plus*. 2016;7(1).

Choi S, Lee HN, Park E, Lee SJ, Kim ES. Recent advances in microbial production of *cis,cis*-muconic acid. *Biomolecules*. 2020;10(9):1–14.

Clomburg JM, Blankschien MD, Vick JE, Chou A, Kim S, Gonzalez R. Integrated engineering of β-oxidation reversal and ω-oxidation pathways for the synthesis of medium chain ω-functionalized carboxylic acids. *Metab. Eng.* Elsevier; 2015;28:202–12.

Cohen SN, Chang ACY. Recircularization and autonomous replication of a sheared R factor DNA segment in *Escherichia coli* transformants. *Proc. Natl. Acad. Sci. U. S. A.* 1973;70(5):1293–7.

Cohen SN, Chang ACY, Boyer HW, Helling RB. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 1973;70(11):3240–4.

Dekishima Y, Lan EI, Shen CR, Cho KM, Liao JC. Extending carbon chain length of 1-butanol pathway for 1-hexanol synthesis from glucose by engineered *Escherichia coli*. *J. Am. Chem. Soc.* 2011;133(30):11399–401.

Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature*. 2011;476(7360):355–9.

Deng Y, Fong SS. Metabolic engineering of *Thermobifida fusca* for direct aerobic bioconversion of untreated lignocellulosic biomass to 1-propanol. *Metab. Eng.* 2011;13(5):570–7.

Deng Y, Ma L, Mao Y. Biological production of adipic acid from renewable substrates: Current and future methods. *Biochem. Eng. J.* 2016;105:16–26.

Deng Y, Mao Y. Production of adipic acid by the native-occurring pathway in *Thermobifida fusca* B6. *J. Appl. Microbiol.* 2015;119(4):1057–63.

Draths KM, Frost JW. Environmentally Compatible Synthesis of Adipic Acid from d-Glucose. *J. Am. Chem. Soc.* 1994;116(1):399–400.

Durst HD, Milano M, Kikta EJ, Connelly SA, Grushka E. Phenacyl Esters of Fatty Acids Via Crown Ether Catalysts for Enhanced Ultraviolet Detection in Liquid Chromatography. *Anal. Chem.* 1975;47(11):1797–801.

EBI. TrEMBLstats. 2020 [accessed 2020 Nov 30]. Available from: https://www.ebi.ac.uk/uniprot/TrEMBLstats

Eddy SR. Profile hidden Markov models. *Bioinformatics*. 1998;14(9):755–63.

Erickson B, Nelson JE, Winters P. Perspective on opportunities in industrial biotechnology in renewable chemicals. *Biotechnol. J.* 2012;7(2):176–85.

Fedorchuk TP, Khusnutdinova AN, Evdokimova E, Flick R, Di Leo R, Stogios P, *et al*. One-Pot Biocatalytic Transformation of Adipic Acid to 6-Aminocaproic Acid and 1,6-Hexamethylenediamine Using Carboxylic Acid Reductases and Transaminases. *J. Am. Chem. Soc.* 2020;142(2):1038–48.

Fibriansah G, Veetil VP, Poelarends GJ, Thunnissen AMWH. Structural basis for the catalytic mechanism of aspartate ammonia lyase. *Biochemistry*. 2011;50(27):6053–62.

Fisher AK, Freedman BG, Bevan DR, Senger RS. A review of metabolic and enzymatic engineering strategies for designing and optimizing performance of microbial cell factories. *Comput. Struct. Biotechnol. J.* 2014;11(18):91–9.

Fox KM, Karplus PA. Old yellow enzyme at 2 Å resolution: overall structure, ligand binding, and comparison with related flavoproteins. *Structure*. 1994;2(11):1089–105.

Goeddel D V., Heyneker HL, Hozumi T, Arentzen R, Itakura K, Yansura DG, *et al*. Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. *Nature*. 1979a;281(5732):544–8.

Goeddel D V., Kleid DG, Bolivar F. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci. U. S. A.* 1979b;76(1):106–10.

Griggs D, Stafford-Smith M, Gaffney O, Rockström J, Öhman MC, Shyamsundar P, *et al*. Policy: Sustainable development goals for people and planet. *Nature*. 2013;495(7441):305–7.

Grushka E, Durst HD, Kikta EJ. Liquid chromatographic separation and detection of nanogram quantities of biologically important dicarboxylic acids. *J. Chromatogr. A*. 1975;112(C):673–8.

Haddad Y, Adam V, Heger Z, Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLoS Comput. Biol.* 2020;16(4): e1007449.

Hasegawa S, Uematsu K, Natsuma Y, Suda M, Hiraga K, Jojima T, *et al*. Improvement of the redox balance increases L-valine production by *Corynebacterium glutamicum* under oxygen deprivation conditions. *Appl. Environ. Microbiol.* 2012;78(3):865–75.

Hopwood DA, Malpartida F, Kieser HM, Ikeda H, Duncan J, Fujii I, *et al*. Production of "hybrid" antibiotics by genetic engineering. *Nature*. 1985;314(6012):642–4.

Inui M, Kawaguchi H, Murakami S, Vertès AA, Yukawa H. Metabolic engineering of *Corynebacterium glutamicum* for fuel ethanol production under oxygen-deprivation conditions. *J. Mol. Microbiol. Biotechnol.* 2005;8(4):243–54.

Jaenisch R. Transgenic animals. *Science.* 1988;240(4858):1468–74.

Jeon WY, Jang MJ, Park GY, Lee HJ, Seo SH, Lee HS, *et al*. Microbial production of sebacic acid from a renewable source: Production, purification, and polymerization. *Green Chem.* 2019;21(23):6491–501.

Joo J chan, Khusnutdinova AN, Flick R, Kim T, Bornscheuer UT, Yakunin F, *et al*. Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chem. Sci.* 2017;8:1406–13.

Ju JH, Oh BR, Heo SY, Lee YU, Shon J hoon, Kim CH, *et al*. Production of adipic acid by short- and long-chain fatty acid acyl-CoA oxidase engineered in yeast *Candida tropicalis*. *Bioprocess Biosyst. Eng.* 2020;43(1):33–43.

Kallscheuer N, Gätgens J, Lübcke M, Pietruszka J, Bott M, Polen T. Improved production of adipate with *Escherichia coli* by reversal of β-oxidation. *Appl. Microbiol. Biotechnol.* Springer Berlin Heidelberg; 2017a Dec 8;101:2371–82.

Kallscheuer N, Polen T, Bott M, Marienhagen J. Reversal of β-oxidative pathways for the microbial production of chemicals and polymer building blocks. *Metab. Eng.* 2017b;42:33–42.

Karlsson E, Mapelli V, Olsson L. Adipic acid tolerance screening for potential adipic acid production hosts. *Microb. Cell Fact.* 2017;16(1).

Khalil I, Quintens G, Junkers T, Dusselier M. Muconic acid isomers as platform chemicals and monomers in the biobased economy. *Green Chem.* 2020;22(5):1517–41.

Kind S, Neubauer S, Becker J, Yamamoto M, Völkert M, Abendroth G von, *et al*. From zero to hero - Production of bio-based nylon from renewable resources using engineered *Corynebacterium glutamicum*. *Metab. Eng.* 2014;25:113–23.

Kind S, Wittmann C. Bio-based production of the platform chemical 1,5-diaminopentane. *Appl. Microbiol. Biotechnol.* 2011;91(5):1287–96.

Kruyer NS, Peralta-Yahya P. Metabolic engineering strategies to bio-adipic acid production. *Curr. Opin. Biotechnol.* 2017;45:136–43.

Kruyer NS, Wauldron N, Bommarius AS, Peralta-Yahya P. Fully biological production of adipic acid analogs from branched catechols. *Sci. Rep.* 2020;10(1).

Kuhlman B, Bradley P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* 2019;20(11):681–97.

Kyulavska M, Toncheva-Moncheva N, Rydz J. Biobased polyamide ecomaterials and their susceptibility to biodegradation. *Handb. Ecomater.* 2019;4:2901–34.

L'emeillat Y, Ménez JF, Berthou F, Bardou L. Quantitative gas chromatographic determination of low-molecular-weight straight-chain carboxylic acids as their p-bromophenacyl esters after extractive alkylation in acidic medium. *J. Chromatogr. A*. 1981;206(1):89–100.

Lee JH, Wendisch VF. Production of amino acids – Genetic and metabolic engineering approaches. *Bioresour. Technol.* 2017;245:1575–87.

Lee JY, Na YA, Kim E, Lee HS, Kim P. The actinobacterium *Corynebacterium glutamicum*, an industrial workhorse. *J. Microbiol. Biotechnol.* 2016;26(5):807–22.

Li G, Huang D, Sui X, Li S, Huang B, Zhang X, *et al*. Advances in microbial production of medium-chain dicarboxylic acids for nylon materials. *React. Chem. Eng.* 2020;5(2):221–38.

Li N, Zhang YZ, Li DD, Niu YH, Liu J, Li SX, *et al*. Overexpression, homology modeling and coenzyme docking studies of the cytochrome P450nor2 from *Cylindrocarpon tonkinense*. *Mol. Biol.* 2016;50(2):320–7.

Lim HN, Lee Y, Hussein R. Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 2011;108(26):10626–31.

Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. *Science.* 2011;334(6055):517–20.

Lonsdale R, Reetz MT. Reduction of α,β-Unsaturated Ketones by Old Yellow Enzymes: Mechanistic Insights from Quantum Mechanics/Molecular Mechanics Calculations. *J. Am. Chem. Soc.* 2015;137(46):14733–42.

Marcheschi RJ, Li H, Zhang K, Noey EL, Kim S, Chaubey A, *et al*. A synthetic recursive "+1" pathway for carbon chain elongation. *ACS Chem. Biol.* 2012;7(4):689–97.

Meiswinkel TM, Gopinath V, Lindner SN, Nampoothiri KM, Wendisch VF. Accelerated pentose utilization by *Corynebacterium glutamicum* for accelerated production of lysine, glutamate, ornithine and putrescine. *Microb. Biotechnol.* 2013;6(2):131–40.

Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided-Drug Des.* 2012a;7(2):146–57.

Meng Y, Xue Y, Yu B, Gao C, Ma Y. Efficient production of l-lactic acid with high optical purity by alkaliphilic *Bacillus sp.* WL-S20. *Bioresour. Technol.* 2012b;116:334–9.

Misura KMS, Chivian D, Rohl CA, Kim DE, Baker D. Physically realistic homology models built with ROSETTA can be more accurate than their templates. *Proc. Natl. Acad. Sci. U. S. A.* 2006;103(14):5361–6.

Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, *et al*. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 1998;19(14):1639–62.

NCBI. GenBank and WGS Statistics. 2020a [accessed 2020 Oct 20]. Available from: https://www.ncbi.nlm.nih.gov/genbank/statistics/

NCBI. RefSeq. 2020b [accessed 2021 Jan 12]. Available from: https://www.ncbi.nlm.nih.gov/refseq/

Niu W, Draths KM, Frost JW. Benzene-free synthesis of adipic acid. *Biotechnol. Prog.* 2002;18(2):201–11.

Niu W, Willett H, Mueller J, He X, Kramer L, Ma B, *et al*. Direct biosynthesis of adipic acid from lignin-derived aromatics using engineered *Pseudomonas putida* KT2440. *Metab. Eng.* 2020;59:151–61.

Okino S, Suda M, Fujikura K, Inui M, Yukawa H. Production of D-lactic acid by *Corynebacterium glutamicum* under oxygen deprivation. *Appl. Microbiol. Biotechnol.* 2008;78(3):449–54.

Padhi SK, Bougioukou DJ, Stewart JD. Site-saturation mutagenesis of tryptophan 116 of

Saccharomyces pastorianus old yellow enzyme uncovers stereocomplementary variants. *J. Am. Chem. Soc.* 2009;131(9):3271–80.

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, *et al*. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* 2005;23(4):482–7.

Parmeggiani F, Weise NJ, Ahmed ST, Turner NJ. Synthetic and therapeutic applications of ammonia-lyases and aminomutases. *Chem. Rev.* 2018;118(1):73–118.

Pettersen JE, Jeelum E, Eedjarn L. The occurrence of adipic and suberic acid in urine from ketotic patients. *Clin. Chim. Acta*. 1972;38(1):17–24.

Picataggio S, Deanda K, Mielenz J. Determination of *Candida tropicalis* acyl coenzyme A oxidase isozyme function by sequential gene disruption. *Mol. Cell. Biol.* 1991;11(9):4333–9.

Picataggio S, Rohrer T, Deanda K, Lanning D, Reynolds R, Mielenz J, *et al*. Metabolic engineering of *Candida tropicalis* for the production of long-chain dicarboxylic acids. *Nat. Biotechnol.* 1992;10(8):893–9.

Polen T, Spelberg M, Bott M. Toward biotechnological production of adipic acid and precursors from biorenewables. *J. Biotechnol.* Elsevier B.V.; 2013;167(2):75–84.

Poppe L. Methylidene-imidazolone: A novel electrophile for substrate activation. *Curr. Opin. Chem. Biol.* 2001;5(5):512–24.

Porter DJ, Bright HJ. 3-Carbanionic substrate analogues bind very tightly to fumarase and aspartase. *J. Biol. Chem.* 1980;255(10):4772–80.

Puthan Veetil V, Raj H, Quax WJ, Janssen DB, Poelarends GJ. Site-directed mutagenesis, kinetic and inhibition studies of aspartate ammonia lyase from *Bacillus sp*. YM55-1. *FEBS J.* 2009;276(11):2994–3007.

Raemakers-Franken PC, Nossin PMM, Brandts P, Wubbolts MG, Peeters WPH, Ernste S, *et al*. Biochemical Synthesis of 6-Amino Caproc Acid. US Pat. Appl. 7491520B2. 2007.

Rahman A, Mupa M, Mahamadi C. A Mini Review on New Emerging Trends for the Synthesis of Adipic Acid from Metal-Nano Heterogeneous Catalysts. *Catal. Letters*. 2016;146(4):788–99.

Raj K, Partow S, Correia K, Khusnutdinova AN, Yakunin AF, Mahadevan R. Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* 2018;6:28–32.

RCSB. PDB Statistics: Overall Growth of Released Structures Per Year. 2020 [accessed 2020 Sep 15]. Available from: https://www.rcsb.org/stats/growth/growth-released-structures

Rohles CM, Gießelmann G, Kohlstedt M, Wittmann C, Becker J. Systems metabolic engineering of *Corynebacterium glutamicum* for the production of the carbon-5 platform chemicals 5-aminovalerate and glutarate. *Microb. Cell Fact.* 2016;15(1).

Rohles CM, Gläser L, Kohlstedt M, Gießelmann G, Pearson S, Del Campo A, *et al*. A bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6,5 using metabolically engineered: *Corynebacterium glutamicum*. *Green Chem.* 2018;20(20):4662–74.

de Ruyck J, Brysbaert G, Blossey R, Lensink MF. Molecular docking as a popular tool in drug design, an in silico travel. *Adv. Appl. Bioinforma. Chem.* 2016;9(1).

Salmaso V, Moro S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol*. 2018;9:923

Sandström AG, Almqvist H, Portugal-Nunes D, Neves D, Lidén G, Gorwa-Grauslund MF. *Saccharomyces cerevisiae*: A potential host for carboxylic acid production from lignocellulosic feedstock? *Appl. Microbiol. Biotechnol.* 2014;98(17):7299–318.

Sauer M, Porro D, Mattanovich D, Branduardi P. Microbial production of organic acids: expanding the markets. *Trends Biotechnol.* 2008;26(2):100–8.

Schilling C. Five Ways Biotech Supports the Transition to a More Circular Economy. *Ind. Biotechnol.* 2019;15(4):234–6.

Schneider J, Wendisch VF. Putrescine production by engineered *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 2010;88(4):859–68.

Schwede TF, Rétey J, Schulz GE. Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. *Biochemistry*. 1999;38(17):5355–61.

Service R. 'The game has changed.' AI triumphs at solving protein structures. *Science.* 2020;370(6521):1144–5.

Shi H, Zhang L, Wu Y, Yu R, Peng Y, Wang Y, *et al*. Production of Adipic Acid Derivatives from d-Glucaric Acid by Hydrodeoxygenation Mediated with Hydroiodic Acid. *Catal. Letters*. 2020;

Shi W, Dunbar J, Jayasekera MMK, Viola RE, Farber GK. The structure of L-aspartate ammonia-lyase from *Escherichia coli*. *Biochemistry*. 1997;36(30):9136–44.

Shimizu A, Tanaka K, Fujimori M. Abatement technologies for N2O emissions in the adipic acid industry. *Chemosph. - Glob. Chang. Sci.* 2000;2(3–4):425–34.

Shin JH, Kim HU, Kim DI, Lee SY. Production of bulk chemicals via novel metabolic pathways in microorganisms. *Biotechnol. Adv.* 2013;31(6):925–35.

Shin JH, Park SH, Oh YH, Choi JW, Lee MH, Cho JS, *et al*. Metabolic engineering of *Corynebacterium glutamicum* for enhanced production of 5-aminovaleric acid. *Microb. Cell Fact.* 2016;15(1).

Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, St.Clair JL, *et al*. Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction. *Science.* 2010;329(5989):309–13.

Singh R. Facts, growth, and opportunities in industrial biotechnology. *Org. Process Res. Dev.* 2011;15(1):175–9.

Slade BA, Leidel L, Vellozzi C, Woo EJ, Hua W, Sutherland A, *et al*. Postlicensure safety surveillance for quadrivalent human papillomavirus recombinant vaccine. *JAMA - J. Am. Med. Assoc.* 2009;302(7):750–7.

Somody JC, MacKinnon SS, Windemuth A. Structural coverage of the proteome for pharmaceutical applications. *Drug Discov. Today*. 2017;22(12):1792–9.

Straathof AJJ. Transformation of biomass into commodity chemicals using enzymes or cells. *Chem. Rev.* 2014;114(3):1871–908.

Stueckler C, Mueller NJ, Winkler CK, Glueck SM, Gruber K, Steinkellner G, *et al*. Bioreduction of alpha-methylcinnamaldehyde derivatives: chemo-enzymatic asymmetric synthesis of Lilial and Helional. *Dalton Trans.* 2010;39(36):8472–6.

Sun J, Raza M, Sun X, Yuan Q. Biosynthesis of adipic acid via microaerobic hydrogenation of *cis,cis*-muconic acid by oxygen-sensitive enoate reductase. *J. Biotechnol.* 2018;280:49–54.

Sutter TR, Sangard D, Loper JC. Isolation and characterization of the alkane-inducible NADPH-cytochrome P-450 oxidoreductase gene from *Candida tropicalis*. Identification of invariant residues within similar amino acid sequences of divergent flavoproteins. *J. Biol. Chem.* 1990;265(27):16428–36.

Takehara I, Fujii T, Tanimoto Y, Kato DI, Takeo M, Negoro S. Metabolic pathway of 6-aminohexanoate in the nylon oligomer-degrading bacterium Arthrobacter sp. KI72: identification of the enzymes responsible for the conversion of 6-aminohexanoate to adipate. *Appl. Microbiol. Biotechnol.* 2018;102(2):801–14.

Teufel R, Mascaraque V, Ismail W, Voss M, Perera J, Eisenreich W, *et al*. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proc. Natl. Acad. Sci.* 2010;107(32):14390–5.

Toogood HS, Fryszkowska A, Hulley M, Sakuma M, Mansell D, Stephens GM, *et al*. A Site-Saturated Mutagenesis Study of Pentaerythritol Tetranitrate Reductase Reveals that Residues 181 and 184 Influence Ligand Binding, Stereochemistry and Reactivity. *ChemBioChem*. 2011;12(5):738–49.

Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 2009;31(2):455–61.

Tsuge Y, Kato N, Yamamoto S, Suda M, Jojima T, Inui M. Metabolic engineering of *Corynebacterium glutamicum* for hyperproduction of polymer-grade l- and d-lactic acid. *Appl. Microbiol. Biotechnol.* 2019;103:3381–3391.

Tsuji M, Kuwahara M. Accumulation of *cis, cis*-muconic acid from benzoic acid by a mutant induced from *Corynebacterium glutamicum*. *J. Ferment. Technol.* 1977;55(2):95–7.

Vavruch I. Chromatograficka studie repnych semen a cukrovky ii.* Bezduslkate organicke kyseliny. *Chem. List.* 1954;46(453):442–5.

Vink ETH, Davies S. Life Cycle Inventory and Impact Assessment Data for 2014 Ingeo® Polylactide Production. *Ind. Biotechnol.* 2015;11(3):167–80.

Van De Vyver S, Román-Leshkov Y. Emerging catalytic processes for the production of adipic acid. *Catal. Sci. Technol.* 2013;3(6):1465–79.

Wehrs M, Tanjore D, Eng T, Lievense J, Pray TR, Mukhopadhyay A. Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends Microbiol.* 2019;27(6):524–37.

Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). 2020 [accessed 2020 Nov 17]. Available from: www.genome.gov/sequencingcostsdata

Williams RE, Rathbone DA, Scrutton NS, Bruce NC. Biotransformation of explosives by the old yellow enzyme family of flavoproteins. *Appl. Environ. Microbiol.* 2004;70(6):3566–74.

Willistein M, Bechtel DF, Müller CS, Demmer U, Heimann L, Kayastha K, *et al*. Low potential enzymatic hydride transfer via highly cooperative and inversely functionalized flavin cofactors. *Nat. Commun.* 2019;10(1).

Wu D, Wang Q, Assary RS, Broadbelt LJ, Krilov G. A computational approach to design and evaluate enzymatic reaction pathways: Application to 1-butanol production from pyruvate. *J. Chem. Inf. Model.* 2011;51(7):1634–47.

Xie NZ, Liang H, Huang RB, Xu P. Biotechnological production of muconic acid: Current status and future prospects. Biotechnol. Adv. 2014. p. 615–22.

Yagi T, Misono H, Tanizawa K, Yoshimura T, Soda K. Characterization of the half and overall reactions catalyzed by L-lysine:2-oxoglutarate 6-aminotransferase. *J. Biochem.* 1991;109(1):61–5.

Yamamoto S, Gunji W, Suzuki H, Toda H, Suda M, Jojima T, *et al*. Overexpression of genes encoding glycolytic enzymes in *Corynebacterium glutamicum* enhances glucose metabolism and alanine production under oxygen deprivation conditions. *Appl. Environ. Microbiol.* 2012;78(12):4447–57.

Yan W, Zhang G, Wang J, Liu M, Sun Y, Zhou Z, *et al*. Recent Progress in Adipic Acid Synthesis Over Heterogeneous Catalysts. *Front. Chem.* 2020;8(185).

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. *Nat. Methods*. 2014;12(1):7–8.

Yang T, Rao Z, Kimani BG, Xu M, Zhang X, Yang ST. Two-step production of gamma-aminobutyric acid from cassava powder using *Corynebacterium glutamicum* and *Lactobacillus plantarum*. *J. Ind. Microbiol. Biotechnol.* 2015;42(8):1157–65.

Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, *et al*. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* 2011;7(7):445–52.

Yoshikawa N, Mizuno S, Ohta K, Suzuki M. Microbial production of *cis,cis*-muconic acid. *J. Biotechnol.* 1990;14(2):203–10.

Yu J Le, Xia XX, Zhong JJ, Qian ZG. Direct biosynthesis of adipic acid from a synthetic pathway in recombinant *Escherichia coli*. *Biotechnol. Bioeng.* 2014;111(12):2580–6.

Zechendorf B. Sustainable development: How can biotechnology contribute? *Trends Biotechnol.* 1999;17(6):219–25.

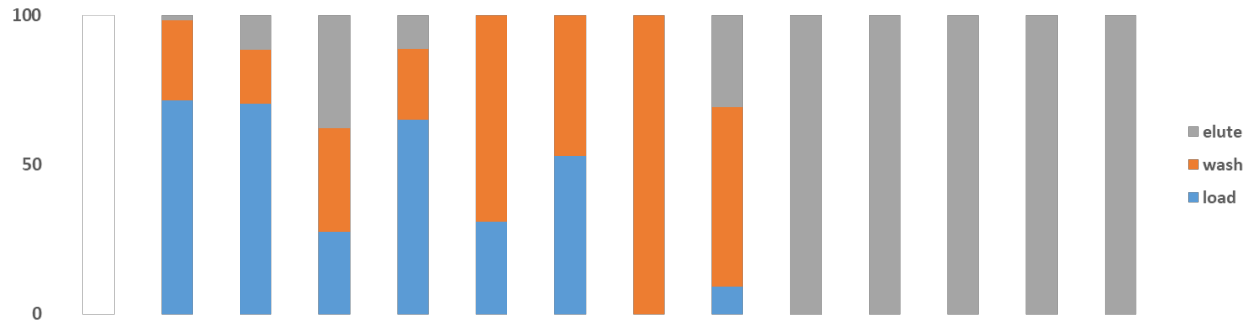
Zhang JF, Zhang ZQ, Dong WC, Jiang Y. A new derivatization method to enhance sensitivity for the determination of low levels of valproic acid in human plasma. *J. Chromatogr. Sci.*

2014;52(10):1173–80.

Zhang K, Sawaya MR, Eisenberg DS, Liao JC. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc. Natl. Acad. Sci. U. S. A.* 2008;105(52):20653–8.

Zhao M, Huang D, Zhang X, Koffas MAG, Zhou J, Deng Y. Metabolic engineering of *Escherichia coli* for producing adipic acid through the reverse adipate-degradation pathway. *Metab. Eng.* 2018;47:254–62.

11. Appendix



Loaded adipic acid (µg)	400	400	400	400	40	40	40	40	40	40	400	120	200
Loaded sample volume (mL)	1	1	1	1	1	1	1	1	1	1	1	3	5
Column type	C8	C8	C8	C8	C8	C8	C8	C8	Strata X	Strata X	Strata X	Strata X	Strata X
Column size (mL)	1	1	3	1	1	3	3	3	3	3	3	3	3
Pretreat	FA	FA	FA	FA	FA	FA	HCl	FA	FA	HCl	FA	FA	FA
Pretreat volume (µL)	150	150	150	150	12	100	120	100	12	120	12	33	55
Pretreat (%)	13	13	13	13	1	8	5	8	1	5	1	1	1
Formic acid in equilibration (%)	0	0	0.1	0.1	.1	8	0	0	0	0	0	0	0
HCl in equilibrition (%)	0	0	0	0	0	0	0.1	0	0	0.1	0	0	0
Wash MeOH (%)	5	0	0	0	0	0	0	0	0	0	0	0	0
Wash formic acid (%)	0	0	0.1	0.1	0.1	8	0	0	0	0	0	0	0
Wash HCl(%)	0	0	0	0	0	0	0.1	0	0	0.1	0	0	0
Elution MeOH (%)	100	100	100	10	100	100	100	100	100	100	100	100	100
Elution formic acid (%)	0	0	0	0.1	0	0	0	0	0	0	0	0	0

Appendix A. Optimization of solid-phase extraction. Various conditions to selectively isolate adipic acid from cultivation broth are shown.C8 and Strata X columns were used with different treatment conditions. Data were not reproduced (n = 1). MeOH, methanol; FA, formic acid; HCl, hydrochloric acid.