METABOLIC ENGINEERING FOR CHEMICAL PRODUCTION VIA EXTENDING AMINO ACID PATHWAY

by

JIAN WANG

(Under the Direction of YAJUN YAN)

ABSTRACT

Microbial-based chemical production from renewable and cheap carbon sources serves as a green replacement to petroleum-based chemical production or plant-based extraction of natural products. The enabling technology, known as metabolic engineering, incorporates enzyme mining, synthetic pathway design and assembly, and systematic host engineering to improve titers, productivities and yields of target products in microbial hosts. To expand the chemical space accessible from microbial-based system, a key principle of metabolic engineering design is to extend existing metabolic pathways and redirect carbon flux en route to these native metabolic pathways. This takes advantage of native metabolic highways and reduce toxic or foreign intermediates or bottleneck enzymes that may impair cell fitness or production performance. Herein, we focused on extending amino acid pathways in *Escherichia coli* and designing synthetic pathways to produce new chemical products. By extending aromatic amino acid pathways, we were able to produce high-value natural products including caffeic acid derived phenethyl esters and amides, and improve the production of tryptophan and its pathway derivatives anthranilate and muconic acid. Then, by extending charged amino acid pathways, we established a novel metabolic platform capable of producing industrially important C3-C5 diols

from amino acids. With the presented platform, seven different diol-convertible amino acids can be converted to diols including 1,3 propanediol (1,3 PDO), 1,4-butanediol (1,4 BDO) and 1,5-pentanediol (1,5 PDO). Finally, by harnessing acetyl-CoA mediated carbon extension pathways, two C5 dicarboxylates, glutarate and 2-methyl-succinate, were successfully produced with applications as bioplastic monomers. Our work presents new metabolic platforms that enriches the pathway repertoire for both natural and non-natural chemical products by extending existing microbial metabolic infrastructures.

INDEX WORDS: Metabolic engineering, bioproduction, amino acid catabolism, pathway engineering, shikimate pathway, caffeic acid, aromatic compounds, diols, dicarboxylates, carbon flux, *Escherichia coli*

# METABOLIC ENGINEERING FOR CHEMICAL PRODUCTION VIA EXTENDING AMINO ACID PATHWAY

by

JIAN WANG

BE, Huazhong Agricultural University, China, 2010

MS, Graduate University Chinese Academy of Sciences, China, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

# METABOLIC ENGINEERING FOR CHEMICAL PRODUCTION VIA EXTENDING AMINO ACID PATHWAY

by

JIAN WANG

| | |
|---|---|
| Major Professor: | Yajun Yan |
| Committee: | Jim Kastner |
| | Michael W. W. Adams |
| | William Kisaalita |

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
December 2019

# DEDICATION

This dissertation is dedicated to my family in China that has been supporting and encouraging me throughout my years at UGA. To my parents, for your support and understanding for me studying abroad for so long. To my brother, for being such a wonderful brother, and being with dad and mom when I am not around. To my wife, Xiaoxuan Ge (Grace), for her loving support and being with me all the time. Finally, to my upcoming baby, for giving me a chance to start a new journey of being a farther.

ACKNOWLEDGEMENTS

I would like to thank my major professor Dr. Yajun Yan for your exceptional support, advising and guidance, which helped me accomplish challenging research projects and develop the scientific mindset for future career. I enjoyed the processes of discussion and coming up with new ideas or solutions to address scientific problems with you. I would like to thank my committee members, Dr. William Kisaalita, Dr. Jim Kastner and Dr. Michael W. W. Adams, for their helpful suggestions and assistance in improving my research and dissertation. I would also like to express my gratitude to all members of Yan lab. It's always joyful journey working and having fun with you. Thank you for all the time we've been together, sharing knowledge, cooperating on research, and talking about science and sports. In particular, I thank Dr. Yaping Yang and Dr. Yuheng Lin for their advice and training that greatly contributed to my accomplishments.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

# INTRODUCTION AND LITERATURE REVIEW[1]

[1]Jian Wang, Sanjay Guleria, Mattheos AG Koffas, Yajun Yan. *Current opinion in biotechnology*, 2016. 37, 97-104.

### 1.1 Metabolic engineering in microbes

The heavy dependence on fossil sources for energy generation and chemical production has encountered increasing challenges, mainly owing to the vast depletion of fossil resources, the growing demand of fossil-derived products, and the emerging environmental issues [1-3]. Due to the ever-increasing concerns about future fossil supplies and climate change, the term "metabolic engineering" was coined in early 1990s, which refers to a science of improving production of metabolite and protein products from renewable biomass or recalcitrant wastes by altering pathway distributions and rates in microbial hosts [4,5]. Since then, numerous chemicals that are used to be chemically produced from petroleum resources or natural products that are used to be extracted from plants tissues can be produced in engineered microbes. The feasibility of manipulating microbes to produce next-generation biofuels (ethanol, butanol and isobutanol, isopentanol, fatty alcohols and biodiesels, etc.), commodity chemicals (succinate and 1,4-butanediol, etc.), nutraceuticals (phytochemicals, prebiotics, polysaccaharides and poly amino acids, etc.) and drug precursors (artemisinin and taxadiene, etc.) has been well-demonstrated with high productivity and yield for industrial applications [6-8]. The roaring development of metabolic engineering was remarkably empowered by the notable advances in all fields adjacent to metabolic engineering, including the accumulation of databases of genes, enzymes and metabolic pathways, the growth of genetic dissection and manipulation of different microorganisms, the expansion of enzyme repertoire that allow achieving desired reactions by rational design or directed evolution, the ease of gene synthesis and multiplexed gene assembling, and the emergence of new genetic tools that enable more precise control over carbon flux and metabolic pathways, etc. [6]. With continuing advance of metabolic engineering, bio-

based economy could serve as a green and sustainable replacement to fossil-based chemical industry.

### 1.1.1 Renewable feedstocks

To achieve cost-efficiency for microbial-based bioproduction, reducing the capital cost of carbon feedstocks is of utmost importance. Fermentative sugars are cheap and the major carbon sources for microbial cell growth and bioproduction in industrial settings. These primarily include C6 sugars like glucose, fructose, and mannose, or C5 sugars like xylose and arabinose derived from lignocellulosic biomass (**Figure 1.1**). As a renewable alternative to fossil reservoirs, lignocellulosic biomass (derived from agricultural, wood, or forest residues) represents a great but underutilized source of fermentable sugars (60-70 wt %) [9]. The recent development of pretreatment methods on lignocellulosic biomass enabled efficient release of these C6 and C5 sugars, providing cheap and non-food feedstocks for renewable chemical production in microbes [10]. Within microbes, C6 and C5 sugars can directly enter the Embden-Meyerhoff-Parnas (EMP, also known as "glycolysis") pathway or the pentose phosphate (PP) pathway, and flow into the tricarboxylic acid (TCA) cycle for generating metabolite building blocks and energy for biomass formation and product synthesis. Due to the abundance of lignocellulosic biomass in nature and its low cost and easy accessibility, tremendous efforts have been directed to bio-based production of fuels and chemicals from lignocellulosic biomass to offset the global fossil fuel consumption [11-13].

Except fermentative sugars, lignin as another major composition in lignocellulosic biomass (15 to 40% weight), has been underutilized due to its structural complexity, heterogeneity and recalcitrance, which remains a major challenge to biorefineries [14,15]. Despite being a rich pool of

more energy dense aromatic compounds, lignin is often combusted to generate low efficient energy [16]. For utilization of lignin, depolymerization mechanisms and approaches have been intensively studied and developed. White-rot and brown-rot fungi are identified lignin degraders in nature that are equipped with peroxidases, laccases or Fenton chemistry, yet difficult for genetic engineering and protein expression [17-19]. Additionally, thermochemical treatments, such as pyrolysis and hydrolysis, generate relatively high content of lignin breakdown products (**Figure 1.1**) [20-24]. Due to the differences in depolymerization approaches and lignocellulosic resources, various aromatic monomers could be generated for direct valorization or be further degraded into microbial central carbon metabolism [15]. Many microorganisms natively contain aromatic-catabolizing pathways, by which a diverse battery of enzymes funnels these aromatic monomers to common intermediates like catechol and protocatechuate [25]. Without aromatic ring cleavage, hydroxylation or methylation of these aromatic monomers could generate value-added aromatic compounds like gallic acid, pyrogallol, and vanillin, etc.; while with meta- or ortho-cleavage, catechol or protocatechuate could be converted to *cis,cis*-muconic acid, succinate, pyruvate or acetyl-CoA (**Figure 1.1**) [26-28]. Although limited efforts have been put towards biological lignin valorization, lignin has shown the potential as a cheap substrate and microbial cell factories have been developed for the funneling of the broad lignin depolymerized monomeric compounds into more specific value-added fuels and chemicals [29].

Glycerol (C3) is a low-cost carbon source that is generated as an inevitable byproduct of biodiesel fuel and oleochemical production from vegetable oils or animal fats [30,31]. The increased supply of crude glycerol and the drop of its price is emerging ascribed to the tremendous growth in the production of biodiesel worldwide. From 2000 to 2010, the glycerol price dramatically decreased from about $2000/ton to under $600/ton in US [32]. In particular, the price for crude

glycerol (80%) ranged between $198-220/ton in US in 2018, making it a competitive feedstock against fermentative sugars [30,33]. The development of processes converting inexpensive glycerol into higher value chemicals could therefore add new values to the production of biodiesel. Except its increasing abundance, an additional advantage of glycerol is its higher reduced state compared to sugars (such as glucose, xylose, etc.), which would theoretically allow higher product yields of reduced chemicals [34,35]. This is especially true when it comes to those highly reduced chemicals, whose production from sugars is limited by the availability of reducing equivalents [36,37]. There are generally two major ways for glycerol utilization. One is to utilize glycerol as carbon source for cell growth and bioproduction, as glycerol can serve as a native substrate for most microbes. The other way is to utilize glycerol as substrate by directly converting it to 3-hydroxypropionic acid (3-HP) and 1,3-propanediol (1,3 PDO) without entering glycerol assimilation pathway [38-40]. This is also the major biological process that leads to commercialization for 1,3 PDO production.

C1 carbon sources, including CO, $CO_2$, methane ($CH_4$) and methanol ($CH_3OH$), are an abundant and low-cost waste gasses produced from a wide range of industrial processes, with total global carbon emission from fossil resources equivalent to approximately 9.7 gigatons in 2014 [41]. Obligately anaerobic acetogenic bacteria like *Clostridium ljungdahlii*, *C. autoethanogenum*, *C. carboxidivorans*, *C. aceticum* and *Sporomusa ovata* could employ the Wood-Ljungdahl (WL) pathway to synthesize acetyl-CoA from either CO or $CO_2$ and $H_2$, which could be used as the substrate for 2,3-butanediol, butyrate and butanol and 2-oxobutyrate (**Figure 1.1**) [41]. Cyanobacteria are a diverse group of photosynthetic microorganisms that are capable of using sunlight as energy and $CO_2$ as carbon source. These photoautotrophs like *Synechocystis* and *Synechococcus* use the carbon-fixing enzyme ribulose-1, 5-bisphosphate

carboxylase/oxygenase (RuBisCO) to capture $CO_2$ and serve as photobiocatalysts for chemical production [42,43]. The model bacterium *Escherichia coli* is not known as a $CO_2$-fixing strain. However, due to its fast growth rate and genetic tractability, increasing interests have been shown to engineer $CO_2$-fixing *E. coli*. The most successful attempt was able to disrupt glycolysis by knocking out the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) and establish a fully functional Calvin-Benson-Bassham (CBB) cycle for sugar synthesis from $CO_2$ in *E. coli*, in the presence of organic acid like pyruvate [44]. Natural gas methane ($CH_4$), as well as its derivative methanol ($CH_3OH$), are considered as next-generation carbon feedstocks due to their easy access and low cost [45]. Currently, aerobic methanotrophs represent the only available catalytic route for methane bioconversion, which involves activating methane to methanol via methane monooxygenase (MMO) followed by converting methanol to formaldehyde en route to fuel and chemical production [46,47]. Metabolic engineering of methanotrophs enables bioconversion of single-carbon methane to multi-carbon and higher-value products including lactic acid (C3), succinate (C4), 1,4 butanediol (1,4 BDO, C4), isobutanol (C4), isoprene (C5) and astaxanthin (C40) [45]. Interestingly, an increasing effort has been taken towards engineering non-methylotrophs like *E. coli* into methylotrophs. Implementation of an NAD-dependent methanol dehydrogenase (Mdh2) and the ribulose monophosphate (RuMP) cycle by expressing hexulose-6-phosphate synthase (Hps) and 6-phospho-3-hexuloisomerase (Phi) from *Bacillus methanolicus* allowed successful establishment and optimization of methanol utilization in *E. coli* [48-50]. Very recently, introduction of the natural serine cycle from *Methylobacterium extorquens AM1* into *E. coli* established an alternative pathway for methanol assimilation, which allowed condensation of methanol (or formate) with bicarbonate ($HCO_3^-$) to produce ethanol [51].

Compared with the above-mentioned carbon sources, protein substrates are less explored as feedstocks for bioproduction. Protein biomass have unique features including structurally diverse components (e.g. 20 different kinds of amino acids) and enrichment of nitrogen (**Figure 1.1**). When protein deaminated and converted to fuel or chemicals, the reduced nitrogen could be recycled as nitrogen fertilizers [52]. The source for protein biomass mainly covers industrial food, agricultural and animal waste [53]. Noteworthy, proteins are also the major constitutes of the fast-growing and $CO_2$-fixing photosynthetic microorganisms, making them an green source for proteins. Indeed, microalgae are ideal biomass sources that provide both carbohydrates and high concentrations of proteins (up to 60-71% in *Spirulina maxima*) [54,55]. A major challenge for protein biomass utilization is its complicated pretreatment that requires a proteolysis process before it can be assimilated and converted by microbes. Although limited efforts being taken on protein biomass utilization, fundamental success have been laid towards the utilization of proteins, including consolidated protein conversion via secreted proteases and nitrogen flux driven conversion of amino acids to biofuels [52,56].

### 1.1.2 Microbial hosts

Microbial hosts are serving as the workhorse for heterologous pathway overexpression and carbon flux redirection to desirable target products. There are several traits that render certain microbes as industrially relevant production strains [57]. The first requisite is genetic competence and tractability that permits host strains to readily accept and stably express foreign DNA in a programable manner. The second requisite is the accessibility of genomic and proteomic information for genome-wide characterization and prediction of cellular responses to different manipulations and environments. The third trait is the availability and accommodation of well-

characterized metabolic engineering toolkits that allow easy manipulation of the genotypes and thus phenotypes. The final requirement is the accumulated knowledge on the physiology and of wildtype or mutant laboratory strains. Within different microbial hosts, even with identical pathways, the production performance may have significant difference. Therefore, when choosing microbial hosts, a full knowledge of both the background of the strains and the target biosynthesis pathways is needed. Conventional model microbial hosts include the bacteria *E. coli* and *Corynebacterium glutamicum*, the yeasts *S. cerevisiae* and *Yarrowia lypolytica*, the filamentous fungus *Aspergillus niger* and the cyanobacterium *Synechocystis* sp. PCC 6803 [57,58].

The Gram-negative bacterium *E. coli* and Gram-positive bacterium *C. glutamicum* have been extensively harnessed as robust industrial microbial hosts, because of their well-characterized genetic context, genetic amenability and the Generally Regarded As Safe (GRAS) status. *E. coli* is the most widely used host for both laboratory and industrial settings, ascribed to its rapid growth, simple culture condition, genetic and metabolic plasticity, wealth of biochemical and physiological knowledge, and the versatile tools for genomic engineering and regulation [59]. It has been harnessed as the host for industrial production of various products, including recombinant proteins, high-value amino acids, nucleotides, dicarboxylates, diols, diamines, polyhydroxybutyate, etc. [60]. Novel genetic tools like promoter and ribosome binding site (RBS) engineering, multiplex automated genome engineering (MAGE), small regulatory RNAs (sRNAs) technique, the CRISPR-(d)Cas9 system have been successfully developed, demonstrated and applied in *E. coli* [61-63]. This will expedite the process of *E. coli*-based host engineering and pathway design for new or improved production capabilities. *C. glutamicum* plays an important role for the high-level production of amino acids including glutamate, lysine, leucine, serine and arginine, etc. [64-67]. These metabolic traits made *C. glutamicum* an ideal

platform for deriving new products from amino acids. For instances, extending glutamate pathway in *C. glutamicum* led to production of 12.37g/L gamma-amino butyric acid (GABA), while extending lysine pathway allowed production of C5 platform chemicals including 5-aminovalerate (28 g/L), cadaverine (103 g/L), and glutarate (90 g/L) [68-71].

The yeast host *S. cerevisiae* is known as its ability to produce ethanol with high titer and productivity due to its tolerance to ethanol. Indeed, *S. cerevisiae* represents the most commonly used eukaryotic organisms in metabolic engineering and almost all pathways that are functional *in E. coli* could be transferred into *S. cerevisiae* [72]. It also possesses several additional advantages. Except being generally more industrially robust than *E. coli*, it is an ideal host for functional expression of eukaryotic cytochrome P450 enzymes [73]. This makes *S. cerevisiae* a widely used production host for complicated natural products with long pathways and P450 enzymes. Expression of multiple-gene plant pathways in *S. cerevisiae* permitted the production of benzylisoquinoline alkaloids including dihydrosanguinarine and its oxidized derivative sanguinarine, reticuline and its derivatives sanguinarine and berberine [74-76]. Aside from *S. cerevisiae*, *Yarrowia lipolytica* has become a preferred non-conventional yeast host for metabolic engineering, which has been used for production of chemicals derived from acetyl-CoA, fatty acids and lipids [77-79]. One unique feature of *Y. lipolytica* is its high flux through the pentose phosphate pathway that could generate NADPH for fatty acid biosynthesis [80]. Currently, *Y. lipolytica* has been engineered as a production host for omega-3 (like eicosapentaenoic acid and docosahexaenoic acid, etc.) and omega-6 fatty acids (like γ-linolenic acid and arachidonic acid, etc.), carotenoids and biodiesel, among which eicosapentaenoic acid (EPA) production has already been commercialized [77,81].

Comparatively, the fungal host *A. niger* and the cyanobacterial host *Synechocystis* sp. PCC 6803 have been less utilized. *A. niger* is used for production of organic acids and heterologous proteins. Its unique compartmentalization feature of separating mitochondrial TCA cycle with cytosol downstream enzymes made it an ideal host for production of TCA cycle derived products. *A. niger* is known for commercial level production of citric acid (109-200 g/L) [82,83]. Its relative host *A. terreus*, has been engineered for industrial production of itaconate (>80 g/L), an unsaturated dicarboxylate with high potential as a building block for biopolymers [84]. *Synechocystis* sp. PCC 6803 is the model cyanobacterial host strain for metabolic engineering explorations. Cyanobacteria are photoautotrophic and can utilize sunlight as the energy supply and capture $CO_2$ for cell growth and bioproduction. This metabolic trait renders cyanobacteria ideal metabolic platform for light-driven and carbon-efficient production. Although in early stage of proof-of-concept demonstration, *Synechocystis* sp. PCC 6803 has been utilized as production host for ethanol, poly-β-hydroxybutyrate (PHB), isoprenoids like limonene and 13R-manoyl oxide (13R-MO), etc. [85-89].

### 1.2 Chemical production via extending central carbon metabolism

Microbes like the *E. coli* can produce almost all essential metabolites and cofactors for bioproduction via a highly evolved metabolic network. Together with its fast growth and GRAS status, *E. coli* has been widely recognized as an ideal chassis for creating designer cell factory with purpose-driven functions. A key principle of metabolic pathway design is to extend native carbon metabolism or metabolic nodes with natural or synthetic biosynthetic pathways. This takes advantage of native metabolic highways and reduce toxic or foreign intermediates or bottleneck enzymes that may impair cell fitness or production performance.

### 1.2.1 Glycolysis

Glycolysis is the entry pathway for sugar-based carbon metabolism that degrades one molecule of C6 sugar to two molecules of pyruvates (C3). During sugar breakdown, all intermediates including sugar phosphates could potentially serve as precursors for chemical production (**Figure 1.2**). Glucose-6-phosphate (G6P) is a metabolite branching point for glycolysis and the oxidative pentose phosphate pathway, which can be re-routed into heterologous production of *myo*-inositol (MI) via *myo*-inositol-1-phosphate synthase (INO1) from *S. cerevisiae* [90]. MI can be further converted into other useful products, such as the biopolymer precursor glucaric acid, and *scyllo*-inositol, a potential therapeutic for Alzheimers [91,92]. Glucaric acid was identified as one of the 12 top value-added chemicals from biomass by US DOE [93]. Glucaric acid is synthesized from MI with the successive catalytic reaction of *myo*-inositol oxygenase (MIOX) from mice and a uronate dehydrogenase (Udh) from *Pseudomonas syringae*. Although its biosynthesis pathway has been uncovered for decades, its titer is still for from commercial merits [92]. Currently, different strategies have been implemented to improve glucaric acid titers, including the use of synthetic scaffolds to co-localize all pathway enzymes, increasing MIOX stability and MI transport, and dynamic knockdown of competing central metabolism [94-97]. The highest titer of glucaric acid could reach 4.58 g/L [94]. Further engineering and optimization is still needed to improve its production to industrial levels.

Pyruvate is a central metabolite in glycolysis and also a common starting point for natural or synthetic pathways (**Figure 1.2**). The most well-known pathway is the 2-*C*-methyl-*D*-erythritol 4-phosphate (MEP) pathway for isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) production. In the *E. coli*, the MEP isoprenoid pathway is initiated by condensation of glyceraldehyde-3 phosphate (G3P) and pyruvate from glycolysis [98].

Overexpression and engineering of MEP pathway in *E. coli* has afforded production of a wide range of chemicals, from simples biofuels like isopentenol and isoprene, to complex drug precursors like taxadiene and lycopene [99-102]. Especially, the high-value product taxadiene, the precursor to chemotherapeutic drug taxol (paclitaxel), reached a high titer of 1.0 g/L via engineering MEP pathway in *E. coli* [101,103]. Condensation of two pyruvate unlocks the gate to generate acetolactate, which can be further directed to 2,3-butanediol (2,3-BDO) or isobutanol via two different decarboxylation pathways [104,105]. This has been serving as a industrially viable pathway for bio-isobutanol production, with titer reaching more than 50 g/L by engineered *E. coli* strains in 1-L bioreactor [106].

Pyruvate decarboxylation generates C2 metabolite acetyl-CoA. As one of the major metabolites in central metabolism, acetyl-CoA could serve as a starting point not only for energy generation via TCA cycle but also for chemical production via heterologous pathways (**Figure 1.2**). Acetyl-CoA can be directly reduced to ethanol by alcohol dehydrogenase E (AdhE) at the expense of two NAD(P)H. This lays the basis for bioethanol fermentation or brewery from biomass [107,108]. Noteworthy, the majority of acetyl-CoA derived chemicals are initiated from condensation of two molecules of acetyl-CoA to acetoacetyl-CoA, which can be later converted to poly-3-hydroxybutyric acid (PHB), 3-hydroxybutyrate (3HB), acetone and butanol [109]. PHB is a biodegradable polyester that is naturally produced by *Cupriavidus necator* (formerly *Ralstonia eutropha*). It is synthesized by condensation of two molecules of acetyl-CoA to acetoacetyl-CoA (with an acetoacetyl-CoA thiolase encoded by *phaA*), subsequent reduction to 3-hydroxybutyryl-CoA (with a reductase encoded by *phaB*), and polymerization to PHB (with a synthase encoded by *phaC*) [110]. Hijacking the PHB pathway by replacing *phaC* with a acyl-CoA thioesterase (*tesB* from *E. coli*) could convert 3-hydroxybutyryl-CoA to 3HB, a versatile chiral building block for

the synthesis of fine chemicals such as vitamins, antibiotics and pheromones [111]. Acetone and butanol pathway naturally exist in clostridia species. Acetoacetyl-CoA can be hydrolyzed to acetoacetate by a CoA-transferase (*ctfAB* from *C. acetobutylicum*) or a thioesterase (*ybgC* from *Haemophilus influenzae*), which was then decarboxylated to acetone by an acetoacetate decarboxylase (*adc* from *C. acetobutylicum*) [112]. When extensively reducing acetoacetyl-CoA with 3-hydroxybutyryl-CoA dehydrogenase (HBD), crotonase (Crt), butyryl-CoA dehydrogenase (Bcd) and aldehyde/alcohol dehydrogenase (AdhE2), butanol will be produced at the expense of four NAD(P)H molecules [113].

The mevalonate pathway, also known as 3-hydroxy-3-methyl-glutaryl-CoA (HMG-CoA) reductase pathway from *S. cerevisiae*, is one of the two major gateways for isoprenoids in eukaryotes, archaea, and bacteria [114]. It initiates from condensation of three molecules of acetyl-CoA and ends with the generation of C5 units IPP and DMAPP. IPP and DMAPP can be directly converted to fuel chemicals like prenol and isoprenol, or further chain-elongated to isoprenoids like monoterpenes and geraniol (C10), sesquiterpenes and farnesol (C15), diterpenes and geranygeraniol (C20), triterpenes (C30), phytoene, lycopene, and carotene (C40) [115-117]. The most successful application of MVA pathway is its capability to achieve high level production of amorpha-4,11-diene, an direct precursor to antimalarial drug artemisinin [114]. Pathway optimization and host engineering finally enabled production of amorpha-4,11-diene with titers of 27.4 g/L in *E. coli* and 40 g/L in *S. cerevisiae*, respectively [118,119]. The pathway for these chemicals has been well-established in model host *E. coli* and could be easily accommodated into any acetyl-CoA generating devices, including $CO_2$-fixing bacteria.

Condensation of acetyl-CoA and pyruvate via the citramalate synthase (CimA) from *Methanoccus jannaschii* leads to generation of 2-ketobutyrate (2-KB) [120]. 2-KB can be directly

decarboxylated to produce 1-propanol, or further extended to 2-ketovalerate or 2-keto-3-methyl-valerate via leucine pathway (LeuABCD) or isoleucine pathway (IlvIHBN), enabling production of 1-butanol and 2-methyl-1-butanol [120,121]. Higher alcohols (C3 or higher) have been considered as an attractive gasoline substitute because of higher energy densities, less hygroscopicity, and are less volatility than ethanol [121].

### 1.2.2 TCA cycle

The TCA cycle is essential for energy generation and is capable of generating multiple metabolite intermediates that can serve as precursors for chemical production (**Figure 1.3**) [122]. The C4 dicarboxylates including succinate, fumarate and malate are native products from TCA cycle and have been recognized as one of the twelve bio-based building block chemicals identified by the US Department of Energy (DOE) [93]. These C4 dicarboxylates have shown wide applications in both biomaterial and food industry and are thus hot targets for metabolic engineering. With the short biosynthetic pathways, they can be easily produced and optimized with high titers from multiple carbon sources including glucose, xylose and glycerol [35,123-125]. Particularly, *Escherichia coli* has been extensively used as a host to produce succinate with high productivity and yield, and could eventually achieve 86.6 g/L succinate with a yield of 1.41 mol/mol glucose (theoretical yield 1.5 mol/mol) [125,126]. 4-Hydroxybutyrate (4-HB), a naturally occurring neurotransmitter and a psychoactive drug and a comonomer for polyesters, can be derived from reduction of succinyl-CoA or decarboxylated from α-oxoglutarate. Systematic engineering of TCA cycle based on aeration levels enabled development of an *E. coli* strain that was capable of producing 103.4 g/L of 4-HB with a yield of 0.419 g/g glycerol in fed-batch fermentation under microaerobic condition [127]. 4-HB can be further reduced to 1,4-

butanediol (1,4-BDO), whose *de novo* production from glucose was systematically improved from 18 g/L to more than 120 g/L with a yield of 0.40 g/g (80% of theoretical) [128,129]. Itaconate is another one of the top 12 value-added chemicals from biomass [93]. *cis*-Aconitate is an intermediate between isocitrate and citrate in TCA cycle, and can be decarboxylated to itaconate by *cis*-aconitate decarboxylase (CadA), without any consumption of energy or reducing equivalent [130]. TCA cycle also serves the biosynthetic machinery for several important amino acids. Oxaloacetic acid and α-ketoglutarate are α-keto acids that can be aminated to generate aspartate and glutamate, respectively. Aspartate then serves as a direct precursor for two additional amino acids, e.g. asparagine and lysine.

### 1.2.3 Fatty acid biosynthesis pathway

Acetyl-CoA carboxylase mediated aerobic carbon-fixation to produce malonyl-CoA, which is the initial step for fatty acid biosynthesis (**Figure 1.4**). When malonyl-CoA enters fatty acid biosynthesis pathway, it will be iteratively incorporated into a "+2" carbon chain elongation cycle to generate fatty acyl-ACP (C8-C18). Fatty acyl-ACP can be directly converted to fatty acids by thioesterases (TES), or reduced to fatty aldehyde by acyl-ACP reductase (AAR) [131]. Fatty aldehyde can be further reduced to fatty alcohols by aldehyde reductase or decarboxylated to fatty alkanes by fatty aldehyde decarbonylase (FAD) [132]. Both fatty alcohols and alkanes are important fuel chemicals to replace petroleum derived fuels. Carboxylic acid reductase (CAR) from *Mycobacterium marinum* is capable of directly reducing fatty acids to fatty aldehydes, which can be further converted to fatty alcohols and alkanes [133]. Very recently, a blue light-dependent fatty acid decarboxylase from microalga *Chlorella variabilis* NC64A enabled direct production of fatty alkane (C13-C17) from fatty acids (C8-C18) [134]. Fatty acids can be also

activated to fatty acyl-CoA by acyl-CoA ligases (ACL), which can be coupled with alcohols by acyltransferase (AT) to produce biodiesel or wax esters [135].

The carboxylation of acetyl-CoA to malonyl-CoA, rendered the feasibility of generating of versatile malonyl-CoA derived chemicals (**Figure 1.4**). Except those long-chain chemicals from derived fatty acids, enzymatic release of CoA from malonyl-CoA or oxidation of malonic semi-aldehyde could potentially enable production of a useful C3 dicarboxylate malonate [136]. 3-Hydroxypropionyl-CoA dehydratase mediated dehydration of 3-hydroxypropionyl-CoA to acryloyl-CoA, provides an opportunity to produce acrylic acid and its derived esters. Acrylic acid and its derived esters are versatile monomers for polymer industry with a huge market. Acryloyl-CoA can be converted to acrylic acid by promiscuous thioesterases or CoA transferases from bacteria like *E. coli*, or coupled with alcohols like ethanol to produce ethyl acrylate by promiscuous acyltransferases. The acrylate pathway from renewable carbon source has been developed but the synthetic pathway for acrylate esters was not established yet [137,138]. The 3-hydroxypropionic acid (3-HP) is a commercially valuable chemical that has been ranked as one of the 12 platform chemicals selected by the US Department of Energy (DOE) [139]. It serves as a building block for polymers and other industrially important chemicals such as acrylic acid and 1,3-propanediol (1,3 PDO). 3-HP is generated from reduction of malonyl-CoA to malonate semialdehyde via malonyl-CoA reductase (MCR) at the expense of two molecules of NADPH [140]. Although other biosynthetic routes of 3-HP or 1,3 PDO from glycerol or glucose have been developed, this sustainable process from fixing $CO_2$ to 3-HP is still attractive.

### 1.3 Chemical production via extending amino acid pathways

Amino acids are the basic constitutional units for almost all life, since proteins are constructed by 20 proteinogenic amino acids. nutritional supplements for human. Amino acids are also important nutritional supplements and feed additives, among which 9 proteinogenic amino acids (e.g. methionine, threonine, leucine, isoleucine, valine, lysine, histidine, phenylalanine, and tryptophan) are essential for humans [141]. Microorganisms harbor all amino acid biosynthetic pathways and can synthesize amino acids from simple carbon sources (**Figure 1.5**). Amino acid biosynthetic network is tightly regulated, rendering limited amino acid synthesis for cellular demands. However, with genome engineering and pathway overexpression, amino acid pathway can be reprogramed and carbon flux can be directed to amino acids. For example, the major amino acids in market include glutamate (Glu) and lysine (Lys), whose titers could reach 100 g/L (with a yield of 0.60 g/g glucose) and 120 g/L (with a yield of 0.55 g/g glucose) in *C. glutamicum*, respectively [142-144]. The threonine (Thr) overproduction was achieved in engineered *E. coli* strain, with the highest titer of 82.4 g/L (with a high yield of 0.393 g/g glucose) by fed-batch culture [145]. The structural diversity and ease of achieving high titers in microbial hosts making amino acids ideal precursors for new value-added natural products or commodity chemicals.

#### 1.3.1 Aromatic amino acids

Phytochemicals are a broad spectrum of plant-derived bioactive secondary metabolites that are commonly found in fruits, vegetables, beans and grains. They are involved in plant defenses against biotic or abiotic stresses, and also exhibit health-protecting or disease-preventing effects on humans [146]. Phytochemicals is a vast and very important repertoire for nutraceuticals, which

include polyphenolic compounds (flavonoids, isoflavonoids, stilbenoids and curcuminoids), alkaloids, terpenoids (monoterpenes, diterpenes, tetraterpenes, polyterpenes, steroid saponins, lycopene and carotenoids) and their derived compounds.

Amongst the phytochemicals, polyphenolic compounds have been used for decades as effective antioxidants and food ingredients in the nutraceutical industry. Polyphenolic scaffolds such as flavonoids and stilbenoids are derived from the phenylpropanoic pathway that extends from aromatic amino acids phenylalanine (Phe) and tyrosine (Tyr) (**Figure 1.6**) [147]. Phenylalanine and tyrosine are converted to pivotal intermediates cinnamic acid and *p*-coumaric acid via phenylalanine ammonia lyase (PAL) and tyrosine ammonia lyase (TAL), which are further converted to cinnamoyl-CoA and *p*-coumaroyl-CoA by 4-coumaroyl-CoA ligase (4CL), respectively. Then, three moles of malonyl-CoA condense with cinnamoyl-CoA or *p*-coumaroyl-CoA to construct the backbone flavanones, stibenes or curcuminoids by type III polyketide synthases, such as chalcone synthase (CHS), curcuminoid synthase (CUS) or stilbene synthase (STS) [148]. With the action of tailoring enzymes such as hydroxylases, *O*-methyltransferases, glucosyltransferase, lipases and prenyltransferase, diverse polyphenolic compounds are generated [149].

*E. coli* and *S. cerevisiae* have generally been developed as platform organisms for *de novo* or semi-*de novo* production of almost all kinds of polyphenolic compounds. Some of the biggest advantages of *E. coli* include its fast growth and ease of genetic manipulation; for *S. cerevisiae*, its Generally Regarded As Safe (GRAS) status and its ability to functionally express plant metabolic enzymes. In all cases, in order to achieve polyphenolic compound production in microbes, plant-originated enzymes are firstly overexpressed via either codon optimization or construction of enzyme chimeras that facilitate bacterial expression [150]. The first strategy is

feeding precursors to the heterologous hosts to produce corresponding polyphenolic compounds. For example, supplementation with phenylalanine to *E. coli* expressing heterologous pathways produced chrysin (9.4 mg/L) or galangin (1.1 mg/L) while supplementation with tyrosine led to the production of apigenin (13 mg/L) or kaempferol (15.1 mg/L) [151]. This feeding approach is widely used when precursors are low or cannot be biosynthesized in microbes. Flavonoid glucosides like luteolin 4'-*O*- and 7'-*O*-glucosides were produced by feeding luteolin to *E. coli* expressing glycosyltransferases UGT71G1 [152]. By expressing glycosyltransferases with different substrate specificities, quercetin 3-*O*-glucoside, quercetin-3-*O*-glucuronide, quercetin 3-*O*-galactoside and quercetin-3-*O*-rhamnoside were respectively produced [152-154]. The second approach is *de novo* biosynthesis of phenolic compounds from simple carbon sources like glucose and glycerol, which is more appealing for industrial applications. *E. coli* hosts can be easily metabolically engineered to overproduce phenylalanine and tyrosine from simple carbon sources. The intracellular concentration of malonyl-CoA, the extender unit for polyphenolic backbones, can be greatly enhanced in *E. coli* by overexpressing acetyl-CoA carboxylase (ACC) or genome-scale metabolic optimization or by repressing malonyl-CoA consumption pathways using antisense RNA [155,156]. *E. coli* has been successfully utilized for *de novo* biosynthesis of certain polyphenolic compounds such as pinocembrin, pinosylvin, sakuranetin and ponciretin [157-159]. Furthermore, substrate promiscuity of authentic or engineered enzymes has provided with new possibilities to produce novel or non-natural polyphenolic compounds [160,161].

Alkaloids are amino acid-derived nitrogenous compounds that have important therapeutic values, including anticancer and antimalarial effects [147]. For a long time, production of alkaloids was limited to plants because of their complex structure and long biosynthetic pathways. The three most well-known alkaloids are benzylisoquinoline alkaloids (BIAs) derived from tyrosine,

monoterpene indole alkaloids (MIAs) derived from tryptophan and glucosinolates. In recent years, elucidation of many BIA biosynthetic pathways enabled the reconstruction of BIA pathways in both *E. coli* and *S. cerevisiae* [162-164]. (*S*)-reticuline, a branch-point intermediate for BIA, was biosynthesized from simple carbon sources with titer of 46.0 mg/L [162]. Bacteria consortia of *E. coli* producing (*S*)-reticuline and *S. cerevisiae* producing either monooxygenase CYP80G2 or the berberine bridge enzyme (BBE) yielded 7.2 mg/L magnoflorine and 8.3 mg/L scoulerine, respectively (**Figure 1.6**) [76]. *S. cerevisiae* was also enabled to produce (*R,S*)-reticuline and was further engineered to generate scoulerine, tetrahydrocolumbamine and tetrahydroberberine from (*S*)-reticuline and salutaridine from (*R*)-reticuline [75]. However, metabolic engineering of MIA alkaloids in microbes is limited. Recently, *de novo* production of the MIA alkaloid strictosidine has been achieved in yeast by introducing 21 new genes and deleting 3 genes in the yeast genome [165]. Glucosinolates are sulfur-rich, amino acid-derived natural compounds. Tryptophan-derived indolylglucosinolate (IG) has recently been produced in *S. cerevisiae* by inserting eight plant genes into the genome, which represents a rare example of glucosinolate production in microbes (**Figure 1.6**) [166]. Although plant platforms are more suitable for scalable alkaloid production, metabolic engineering in microbes has shown potential capability for cost-efficient production of these plant-derived compounds [149,165].

### 1.3.2 Aliphatic amino acids

Aliphatic amino acids are nonpolar and hydrophobic amino acids, including glycine (Gly), alanine (Ala), valine (Val), leucine (Leu) and isoleucine (Ile). The last three are also known as branched-chain amino acids (BCAAs). A major derivation of aliphatic amino acids is to produce branched-chain alcohols via 2-keto acid decarboxylation pathway, also referred as Ehrlich

pathway (**Figure 1.7**) [167]. 2-Keto acids are metabolite intermediates in amino acid biosynthesis pathways, which can be released from counterpart amino acids by transamination reactions. These 2-keto acids can be easily decarboxylated to aldehydes by a broad-substrate-range 2-keto-acid decarboxylase KivD from *Lactococcus lactis* and then to alcohols by promiscuous alcohol dehydrogenases (ADHs). Using this strategy, amino acid biosynthesis pathways could be hijacked to produce various 2-keto acids, like 2-ketobutyrate (2-KB) and 2-keto-3-methyl-valerate (2-KMV) from isoleucine pathway, 2-keto-isovalerate (2-KIV) from valine pathway, 2-keto-4-methyl-pentanoate (2-KMP) from leucine pathway, and 2-ketovalerate (2-KV) from norvaline pathway. Eventually, 1-propanol, 2-methyl-1-butanol, isobutanol, 3-methyl-1-butanol and 1-butanol were released from those branched-chain amino acid pathways [168]. Further combination of Ehrlich pathway and the 'synthetic recursive +1' carbon elongation pathway could produce various long-chain (C5-C8) alcohols, including 1-pentanol, -3-methyl-1-pentanol, 1-hexanol, 1-heptanol and 1-octanol [169,170].

### 1.3.3 Charged amino acids

The charged amino acids include two acidic amino acids aspartate (Asp) and glutamate (Glu) and three basic amino acids lysine (Lys), arginine (Arg), and histidine (His). Additionally, further amination of aspartate and glutamate generate two polar amino acids, asparagine (Asn) and glutamine (Gln). Charged amino acids have been less utilized due to their bifunctional terminal groups. Aspartate has been utilized as a precursor for β-alanine, an important platform chemical for production of acrylamide and acrylonitrile [171,172]. The bio-based production of β-alanine could reach 32.3 g/L by *E. coli* in fed-batch fermentation [171]. Via this β-alanine route, several new products could be produced with corresponding synthetic pathways, including 3-HP,

malonic acid, and poly(3-hydroxypropionate) (P3HP) [136,173]. As mentioned earlier, glutamate and lysine have been generally used as precursors to produce C4 and C5 platform chemicals including GABA, 5-aminovalerate, cadaverine, and glutarate. Recently, glutamate and lysine pathway have been extended to produce C4 (butyrolactam), C5 (valerolactam) and C6 (caprolactam) lactams, with butyrolactam reaching a high titer of 54 g/L by *E. coli* in fed-batch fermentation [174]. With the capability to achieve high titers of charged amino acids, further development of new pathways would accelerate their upgrading.

### 1.4 Research Objectives

The primary goal of this work is to develop new synthetic pathways for natural or non-natural chemical products by extending existing metabolic pathways especially amino acid pathways in *E. coli* (**Figure 1.8**). Chapter 2 describes the extension of tyrosine and phenylalanine pathways for the synthesis of high-value phytochemicals, including a panel of caffeic acid derived phenethyl esters and amides. This enables the establishment of synthetic pathway in microbes to produce plant-derived chemicals, whose natural pathway remains enigmatic. Chapter 3 describes the metabolic design and creation of a pyruvate-driven and growth-coupled production platform for the other aromatic amino acid tryptophan. Based on this platform, we demonstrate the production enhancement of three pathway derivatives including tryptophan, anthranilate and muconic acid. In Chapter 4, by extending charged amino acid pathways, we establish a novel metabolic platform capable of producing industrially important C3-C5 diols from amino acids. With the presented platform, seven different amino acids can be converted to diols including 1,3 PDO, 1,4 BDO and 1,5 PDO. In Chapter 5 and 6, we explore and establish two novel pathways derived from acetyl-CoA mediated carbon extension pathway, which afford production of two

C5 dicarboxylates glutarate and 2-methyl-succinate with applications as bioplastic monomers. Our work highlights the potential of creating diverse new pathways from native metabolic highways. With those established pathways, high-value or high-volume chemicals can be produced from renewable feedstocks, representing potential opportunities for industrial applications.

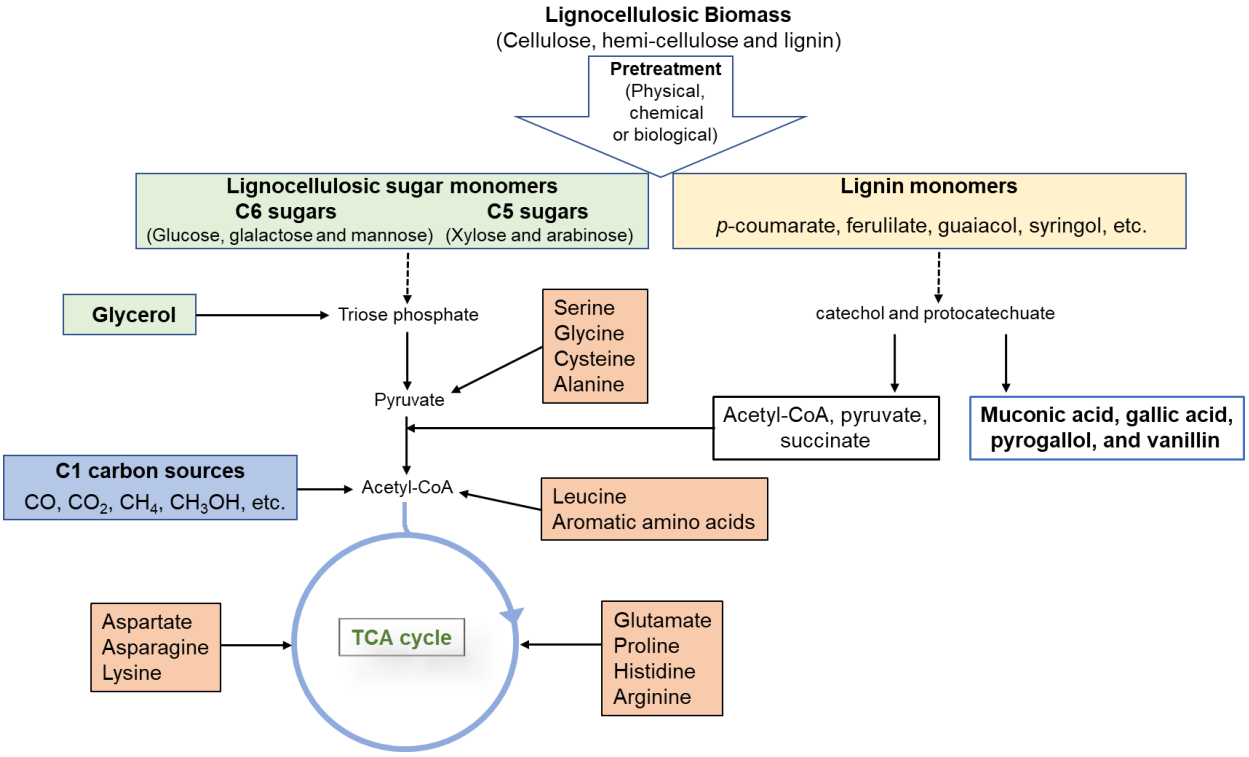



**Figure 1.1 Renewable feedstocks utilized by microbes for bio-based production.** With purpose-driven metabolic engineering approaches, various low-cost carbon sources could be utilized by engineered microbial hosts. These carbon sources including fermentative sugars and lignin monomers from lignocellulosic biomass, glycerol, C1 carbon syngasses and protein derived amino acids.

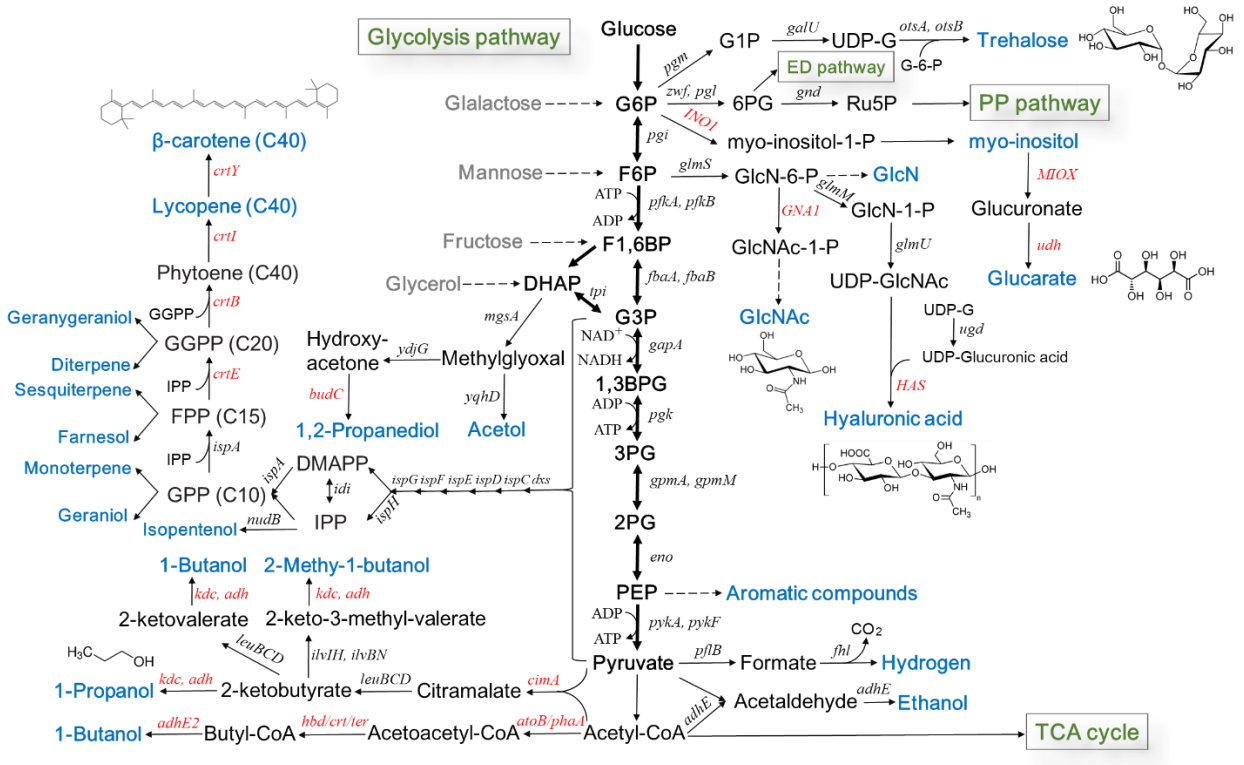



**Figure 1.2 Glycolysis pathway and metabolic pathways for various value-added chemicals.** Arrows indicate the metabolite flux of the enzymatic reactions and the thicker arrows indicate the glycolysis pathway. Endogenous genes are shown in black and heterologous genes are shown in red. Metabolites referred: G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; F1,6BP, fructose 1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde 3-phosphate; 1,3BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; G1P, glucose-1-phosphate; UDP-G, UDP-D-glucose; 6PG, D-gluconate 6-phosphate; Ru5P, ribulose 5-phosphate. GlcN, glucosamine; GlcNAc, N-acetylglucosamine; DMAPP, Dimethylallyl diphosphate; IPP, Isopentenyl diphosphate; GPP, geranylgeranyl phosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl diphosphate.

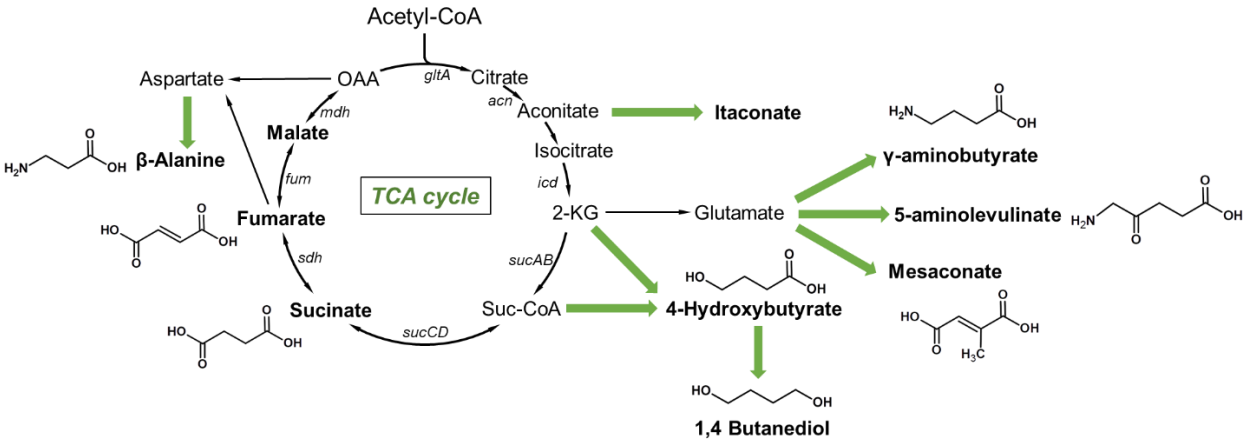



**Figure 1.3 TCA cycle derived metabolic pathways for value-added chemicals.** Thin arrows indicate the metabolite flux of TCA cycle, and the thick green arrows indicate the synthetic pathways for target products. Metabolites abbreviations: OAA, oxaloacetate; 2-KG, 2-ketoglutarate.

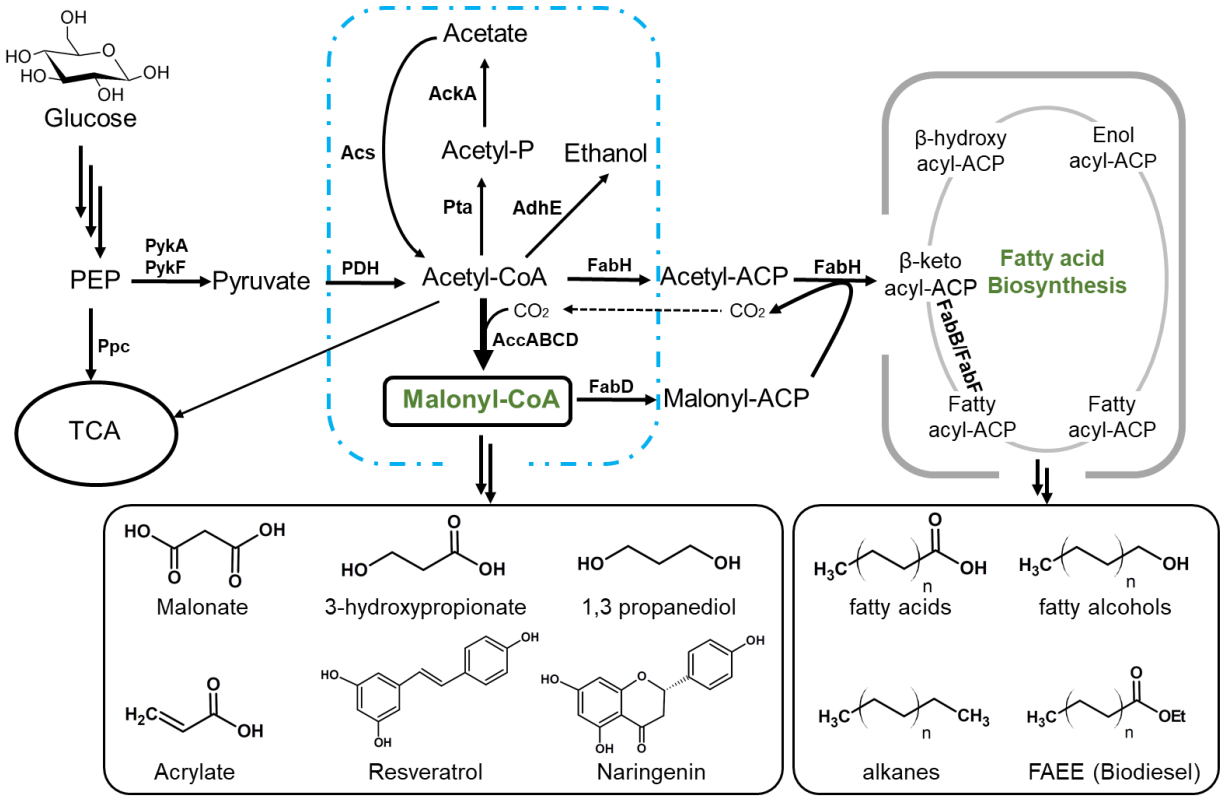



**Figure 1.4 Bioproduction of value-added chemicals by extending fatty acid pathways.** Malonyl-CoA is produced from carboxylation of acetyl-CoA, and further condensation of acetyl-ACP and malonyl-ACP leads to fatty acid biosynthesis pathway. Typical examples of malonyl-CoA and fatty acids derived chemicals are represented. Metabolites abbreviations: PEP, phosphoenolpyruvate; acetyl-P, acetyl phosphate; ACP, acyl carrier protein.

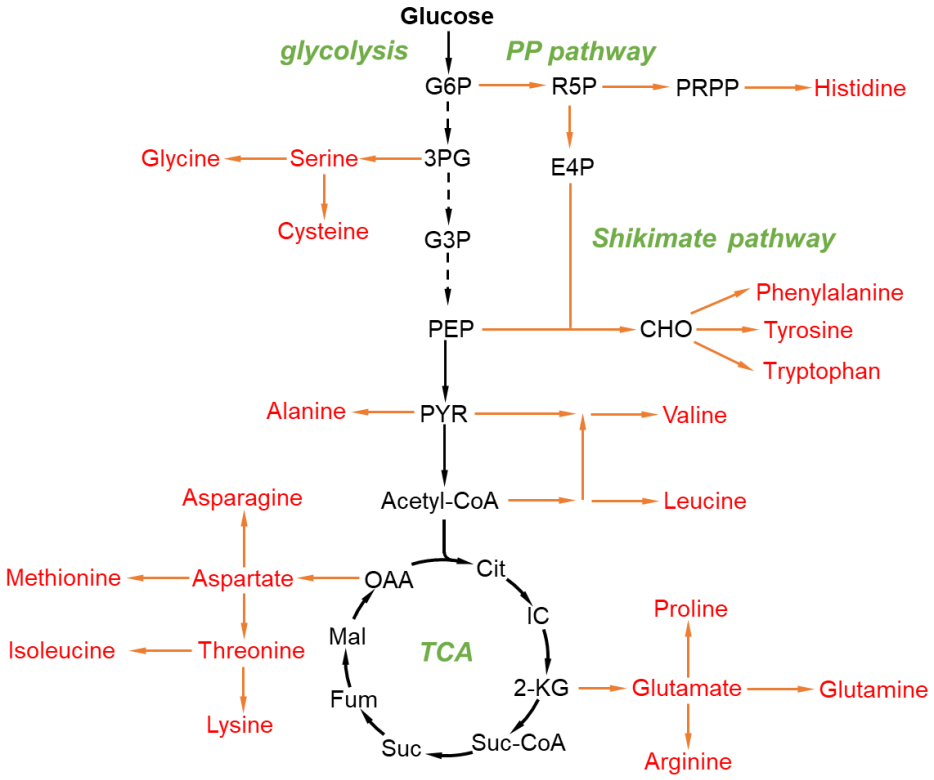



**Figure 1.5 Amino acid biosynthetic pathways in *E. coli*.** The scheme represents the natural pathways for 20 different amino acid biosynthetic pathways that are derived from central carbon metabolism. Metabolites abbreviations: Cit, citrate; IC, isocitrate; 2-KG, 2-ketoglutarate; Suc, succinate; Fum, fumarate; Mal, malate; OAA, oxaloacetate.

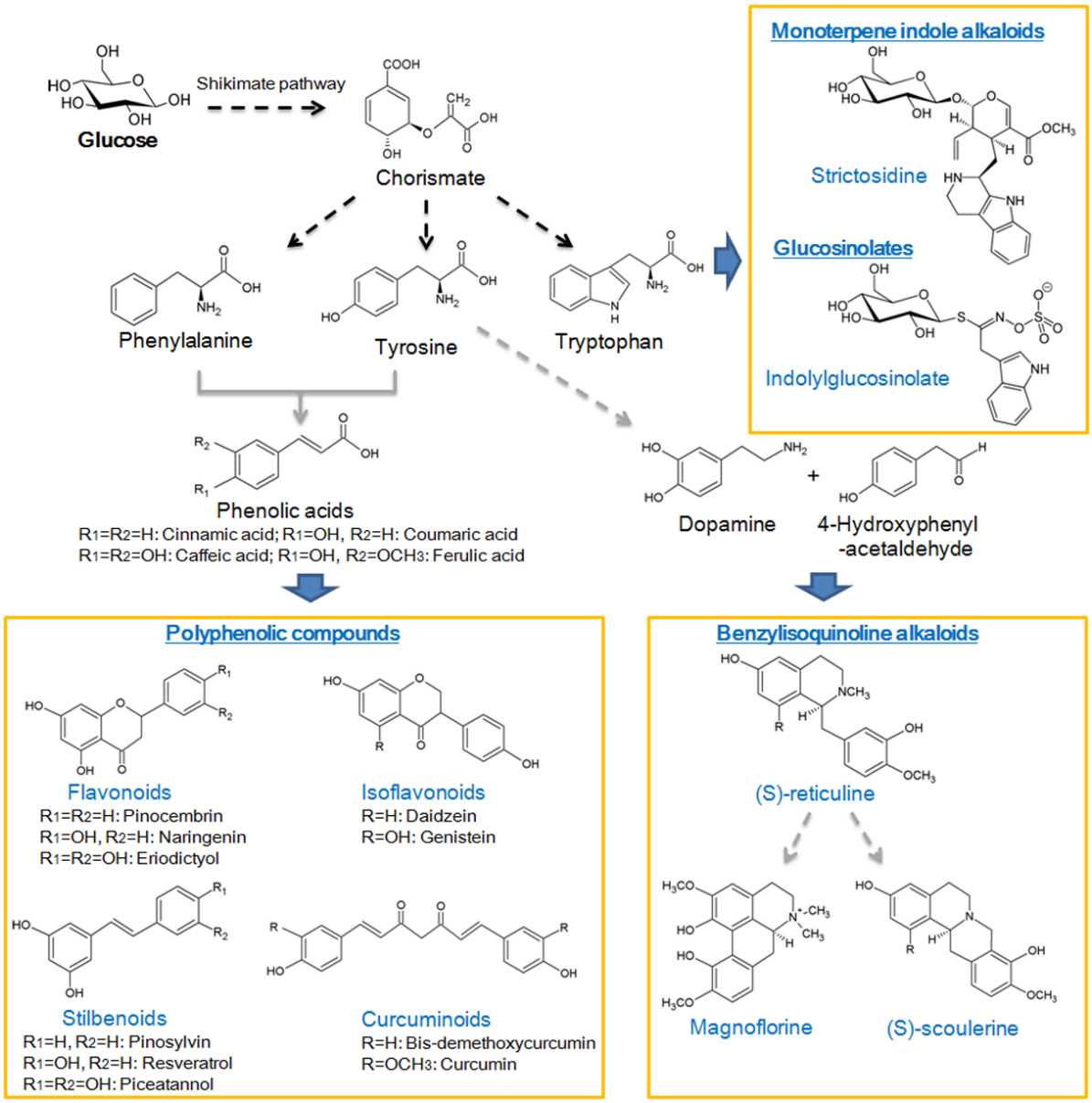



**Figure 1.6 Examples of microbial-based production of phytochemicals from aromatic amino acids.** Arrows represent the metabolic flux from primary shikimate pathway to target phytochemicals. Polyphenolic compounds are derived from phenylalanine or tyrosine, benzylisoquinoline alkaloids (BIA) are derived from tyrosine, monoterpene indole alkaloids (MIA) and indolylglucosinolate (IG) are derived from tryptophan.

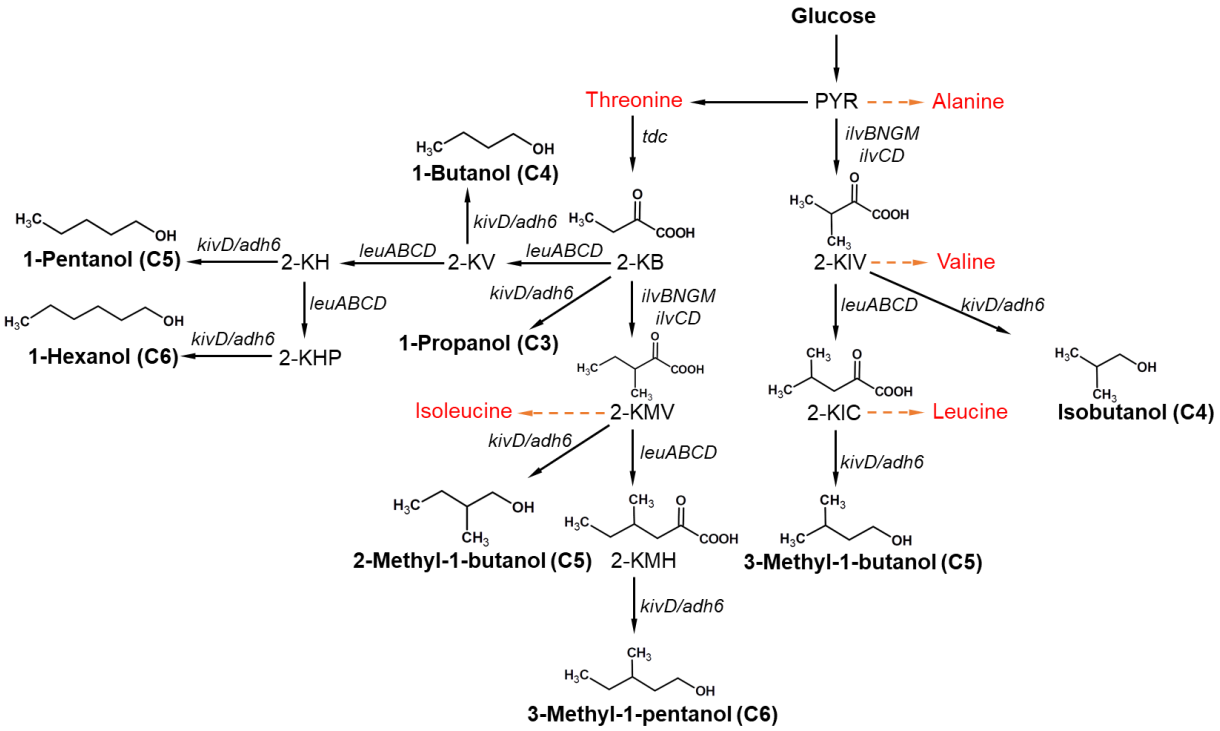



**Figure 1.7 Metabolic engineering of biofuels via extending aliphatic amino acid pathways.** Aliphatic amino acid pathways can generate various 2-keo acids that can be further extended and converted linear or branched-chain alcohols. The scheme represents the typical alcohols produced by from extending amino acid pathways via 2-keto acid decarboxylation. Metabolites abbreviations: 2-KB, 2-ketobutyrate; 2-KIV, 2-ketoisoveralate; 2-KIC, 2-ketoisocaproate; 2-KMV, 2-keto-3-methylvalerate; 2-KMH, 2-keto-4-methylhexanoate; 2-KV, 2-ketovalerate; 2-KH, 2-ketohexanoate; 2-KHP, 2-ketoheptanoate.

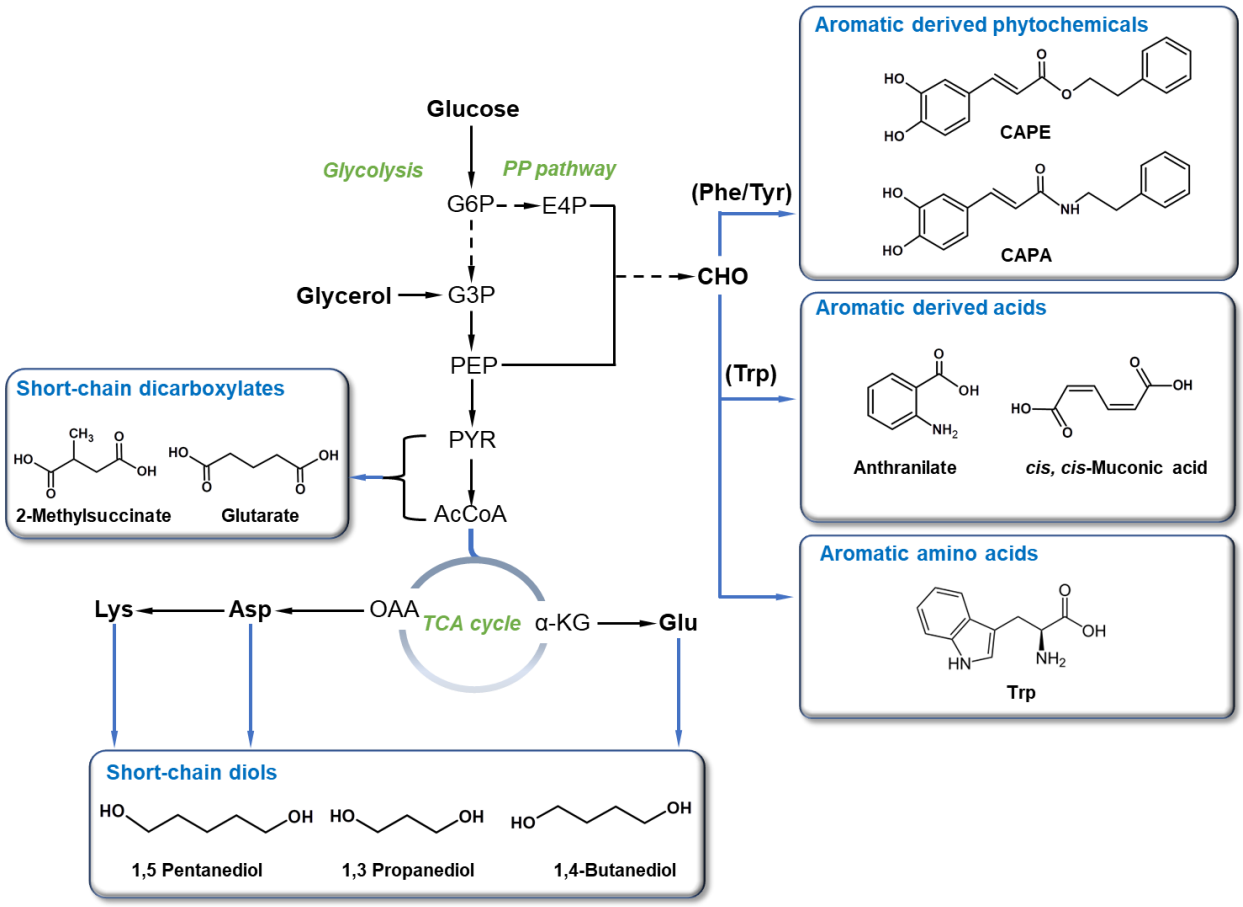



**Figure 1.8 An overview of extending amino acid pathways for chemical production in this dissertation.** Black arrows indicated native pathways and blue arrows indicate synthetic pathways leading to target products. These target products (represented in boxes) include short-chain dicarboxylates and diols, aromatic derived high-value phytochemicals, anthranilate, *cis,cis*-muconic acid, and Trp.

# CHAPTER 2

# ENGINEERING A BACTERIAL PLATFORM FOR TOTAL BIOSYNTHESIS OF CAFFEIC ACID DERIVED PHENETHYL ESTERS AND AMIDES[1]

[1]Jian Wang, Monika Mahajani, Sheneika L Jackson, Yaping Yang, Mengyin Chen, Eric M Ferreira, Yuheng Lin, Yajun Yan. *Metabolic Engineering*, 2017. 44, 89-99.

## 2.1 Abstract

Caffeic acid has been widely recognized as a versatile pharmacophore for synthesis of new chemical entities, among which caffeic acid derived phenethyl esters and amides are the most extensively-investigated bioactive compounds with potential therapeutical applications. However, the natural biosynthetic routes for caffeic acid derived phenethyl esters or amides remain enigmatic, limiting their bio-based production. Herein, product-directed design of biosynthetic schemes allowed the development of thermodynamically favorable pathways for these compounds via acyltransferase (ATF) mediated trans-esterification. Production based screening identified a microbial O-ATF from *Saccharomyces cerevisiae* and a plant N-ATF from *Capsicum annuum* capable of forming caffeic acid derived esters and amides, respectively. Subsequent combinatorial incorporation of caffeic acid with various aromatic alcohol or amine biosynthetic pathways permitted the *de novo* bacterial production of a panel of caffeic acid derived phenethyl esters or amides in *Escherichia coli* for the first time. Particularly, host strain engineering via systematic knocking out endogenous caffeoyl-CoA degrading thioesterase and pathway optimization via titrating co-substrates enabled production enhancement of five caffeic acid derived phenethyl esters and amides, with titers ranging from 9.2 to 369.1 mg/L. This platform expanded the capabilities of bacterial production of high-value natural aromatic esters and amides from renewable carbon source via tailoring non-natural biosynthetic pathways.

## 2.2 Introduction

Nature is the largest arsenal for biologically active compounds and motivates cost-effective manufacturing of valuable natural products. The ever-increasing research and discoveries have brought us a deeper understanding of the biosynthetic machineries of numerous natural products, especially those from plants, on both genetic and enzymatic levels. Instead of direct extraction from plant tissues with low yields or chemical synthesis with high costs and low specificity, knowledge-based microbial production is a sustainable and cost-effective biomanufacturing alternative with high specificity and less complexity. Especially, genetic amenability and the Generally Regarded As Safe (GRAS) status make the well-characterized microbes like *Saccharomyces cerevisiae* and *Escherichia coli* robust industrial microbial platforms for production of non-cognate natural products from simple carbon sources [175]. The last decades have witnessed tremendous advances in achieving microbial-based production of plant-derived chemicals with pharmacological activities including polyphenols, alkaloids and terpenoids [8,149]. In most cases, the entire plant biosynthetic pathways can be directly recruited into microbial workhorses and repurposed to synthesize desired phytochemicals [165,176,177]. However, the limited availability of natural pathways or functional enzymes sometimes presents a significant challenge for biosynthesis through metabolic engineering. In such situations, portions of natural pathways or totally artificial pathways have to be created by capitalizing and assembling enzymes from different organisms or pathways [128,178-181]. These alternative non-natural pathways might afford production of natural compounds via tailoring similar biochemical pathways or integrating promiscuous enzymes, by compensating the absence of natural pathways.

Caffeic acid (CA) is a plant-specific secondary metabolite with potent antioxidant properties [178,182]. The CA scaffold is abundantly found in biologically active natural compounds like caffeic

acid derived flavonoids, esters and amides, thus prompting an upsurge in the development of CA-based therapeutically bioactive compounds [183]. Caffeic acid phenethyl ester (CAPE) is one of the most promising therapeutically bioactive polyphenol components from honey bee propolis, which has been recognized as a potential drug candidate for its potent antioxidant activity [184,185]. Numerous research studies have demonstrated that CAPE possesses versatile pharmacological properties such as antibacterial, anti-inflammatory, antiproliferative, antiviral, anticarcinogenic, antidiabetic, antileukemic, chemopreventive and neuroprotective effects [186,187]. Recent preclinical studies have shown that CAPE was efficient in inhibiting different kinds of cancers such as colorectal cancer, lung cancer, breast cancer, pancreatic cancer, gastric cancer, melanoma, glioma, cholangiocarcinoma and hepatocellular carcinoma in micromolar concentrations [188-195]. CAPE ester analogs including caffeic acid (4-hydroxylphenethyl) ester (CAHPE), caffeic acid (3,4-dihydroxyphenethyl) ester (CADPE) and a series of amide analogs including caffeic acid phenethyl amide (CAPA), caffeoyl tyramine (CATA) and caffeoyl dopamine (CADA) are all potent antioxidants [184,196]. Notably, the amide analogs showed improved stability compared to CAPE [197].

How caffeic acid derived phenethyl esters and amides are biosynthesized is still unclear in nature. As the prototype, CAPE has been previously chemically or enzymatically synthesized. Chemical synthesis of CAPE via acid-catalyzed esterification or base-catalyzed alkylation routinely involves environmentally harmful chemicals and even carcinogenic agents like hexamethylphosphoramide (HMPA) [198]. The most common enzymatic strategy to synthesize CAPE was via a lipase-catalyzed esterification involving caffeic acid and 2-phenylethanol, whose efficiency was hindered by non-specificity [199,200]. In particular, chemical synthesis as well as enzymatic catalysis suffers from a thermodynamic disadvantage, as hydrolysis is favored at

ambient temperatures (ΔG= −5 kcal/mol) [201]. Thus, reconstituting a feasible microbial-based production of CAPE as well as its ester and amide analogs with reliable bio-safety and high specificity will be more attractive, ascribed to its ecological acceptability and economic benefits. Although short and medium-chain esters as well as long-chain fatty esters have been extensively biosynthesized in microbial hosts, aromatic esters and its analog amides like caffeic acid derived ones are rarely produced [135,201-205].

Due to the inaccessibility of natural pathways for synthesis of caffeic acid derived phenethyl esters and amides, we turned to establishing artificial pathways. By examination of the CAPE and CAPA chemical structures, we reasoned that acyltransferase mediated trans-esterification of caffeoyl-CoA with 2-phenylethanol (PE) or 2-phenylethylamine (PA) would facilitate ester or amide formation since the release of free CoA from the acyl groups upon esterification is thermodynamically favored (ΔG= −7.5 kcal/mol) [201]. In the present work, we took advantage of this CoA-dependent trans-esterification pathway and constituted an *E. coli* platform for production of a collection of caffeic acid derived phenethyl esters and amides from glucose. To construct this platform, we first identified and characterized the acyltransferases that can catalyze the trans-esterification of caffeoyl-CoA with aromatic alcohols or amines. Next, assembly of the whole pathway by combinatorial incorporation of an artificial caffeic acid pathway and various aromatic alcohol or amine biosynthetic pathways enabled successful production of a panel of caffeic acid derived phenethyl esters or amides. Finally, systematic investigation of endogenous (thio)esterases on caffeoyl-CoA degradation enhanced production of both caffeic acid derived esters and amides. Specifically, balancing co-substrates CA and PE with an engineered promoter library enhanced production of CAPE. This work affords the first bacterial production of authentic caffeic acid derived phenethyl esters and amides via assembling product-directed non-

natural metabolic pathways, which expands the capabilities of bacterial production of high-value aromatic esters and amides.

## 2.3 Materials and methods

### 2.3.1 Bacterial strains, plasmids and chemicals

Bacterial strains and plasmids used in this study were listed in **Table S2.1**. *E. coli* XL1-Blue was used for plasmid construction and preservation, and *E. coli* BL21 Star (DE3) was used for protein expression and purification. *E. coli* BW25113 (F') and its derived knockout strains were used for production. The derived knockout strains were generated from *E. coli* BW25113 (F') via P1 phage transduction method according to standard protocols [206]. pZE12-luc (high-copy number), pCS27 (medium-copy number) and pSA74 (low-copy number) were used as backbone plasmids for pathway construction. pETDuet-1 was used for cloning genes to be expressed. Standard chemicals were purchased from Sigma-Aldrich unless otherwise specified. All chemicals and solvents were purchased from Alfa Aesar (Ward Hill, Ma), Sigma-Aldrich (St. Louis, MO), Fisher Scientific (Waltham, MA), J.T. Baker (Center Valley, PA), and Matrix Scientific (Columbia, SC).

### 2.3.2 Chemical synthesis of standards

Standard of esters and amides were synthesized to high purity (> 95%) according to standard protocols of carbodiimide couplings. These chemicals included *p*-coumaric acid phenethyl ester, *p*-coumaric acid 4-hydroxyphenethyl ester, caffeic acid phenethyl ester (CAPE), caffeic acid 4-hydroxyphenethyl ester (CAHPE), *p*-coumaroyl phenylethylamine, *p*-coumaroyl tyramine, *p*-coumaroyl dopamine, caffeic acid phenethyl amide (CAPA), caffeoyl tyramine (CATA), and

caffeoyl dopamine (CADA). If phenolic groups were present, protection as acetates was carried out prior to coupling, and deacetylations were performed afterward. See the **Supplementary Note** for full details. Caffeic acid (3,4-dihydroxyphenethyl) ester (CADPE) was not synthesized and CADPE quantification was estimated based on the synthesized CAHPE standard because of their structural similarity.

### 2.3.3 Plasmid constructions

All genes were amplified by polymerase chain reaction (PCR) using Phusion High-Fidelity DNA polymerase (New England BioLabs). Plasmid constructions and DNA manipulations were performed following the standard molecular cloning protocols [207]. pZE-*TH* is a plasmid previously constructed to express *hpaBC* from *E. coli* BL21 Star (DE3) and the codon-optimized *tal* gene from *R. glutinis* [208]. pZE-*Eht1-4CL2* was constructed by inserting *Eht1* from *S. cerevisiae* and *4CL2* from *A. thaliana* into the high copy number plasmid pZE12-luc between *Acc65*I and *Xba*I. Similarly, *Atf1*, *Atf2*, *Eeb1* from *S. cerevisiae*, *CAT* from pSA74 plasmid and the codon-optimized *HCT* from *N. tabacum* were constructed into pZE12-luc with *4CL2*. pZE-*THT-4CL2* was obtained by inserting the codon-optimized *THT* from *C. annuum* and *4CL2* into pZE12-luc between *Acc65*I and *Xba*I. pZE-*KA* was constructed by inserting *kivD* from *L. lactis* and *adh6* from *S. cerevisiae* into pZE12-luc in between *Acc65*I and *Xba*I. pCS-*TPTA* containing *tyrA**, *ppsA*, *tktA*, *aroG** was constructed previously [178]. pCS-*TPTA-EC* and pCS-*TPTA-TC* were constructed by amplifying *Eht1-4CL2* and *THT-4CL2* respectively and inserting into pCS-*TPTA* in between *Spe*I and *Sac*I. *LbPDC* from *L. brevis*, *MjTDC* from *M. jannaschii*, *PpDDC* from *P. putida* were firstly inserted into the medium copy number plasmid pCS27 to obtain pCS-*LbTDC*, pCS-*MiTDC* and pCS-*PpDDC*, and then amplified and inserted into pZE-*TH* in between *Spe*I

and *Sph*I to generate pZE-*TH-LbTDC*, pZE-*TH-MjTDC* and pZE-*TH-PpDDC*. To obtain pZE-*TH-KA, kivD-adh6* was amplified from pZE-*KA* and inserted into pZE-*TH* in between *Spe*I and *Sph*I. To express and purify Eht1, THT and 4CL2, each gene was respectively amplified and constructed into pETDuet-1, generating pETDuet-*Eht1*, pETDuet-*THT*, pETDuet-*4CL2*. To express and purify 15 native putative (thio)esterases from *E. coli* including BioH, FadM, TesB, YbaC, EntH, YbfF, YciA, PaaI, PaaY, YeiG, YpfH, YqiA, YjfP, YbgC and YdiI, the corresponding genes were amplified and inserted into pETDuet-1 as described previously [209]. pCS-P*lpp1.0-egfp* and pCS-P*lpp2.0-egfp* were respectively constructed by inserting P*lpp1.0* and P*lpp2.0* with *egfp* into pCS27. pSA-P*lppx-pheA** plasmids were respectively constructed by inserting P*lppx* and *pheA** into pSA74 between *Xho*I and *Bam*HI. All plasmids constructed were listed in **Table S2.1**.

### 2.3.4 Culture media and conditions

Luria-Bertani (LB) medium was used for inoculation and cell propagation. M9 minimal medium containing 20 g/L glucose and 5 or 3 g/L yeast extract was used for shake flask experiments. When needed, ampicillin, kanamycin, and chloramphenicol were added to the medium at 100, 50, and 34 μg/mL, respectively. Triplicate seed cultures were inoculated at 37 °C overnight in 3 mL liquid LB medium on a gyratory shaker (New Brunswick). Shake flask experiments were performed by inoculating with 2 % inoculum from overnight seed cultures into 20 mL medium and grown in a rotary shaker at 250 rpm. For feeding experiments, cells were inoculated in M9 medium with 5 g/L yeast extract at 37 °C for 3 h and then transferred into a 30 °C shaker. Isopropyl-β-D-thiogalactoside (IPTG) with a final concentration of 0.5 mM and substrates with a final concentration of 500 mg/L were added. For *de novo* production, cells were

inoculated in M9 medium with 3 g/L yeast extract in a 30 °C shaker and 0.5 mM IPTG was added for induction at initial inoculation. Samples (1 mL) were harvested by centrifugation every 12 h until 48 h and the supernatants were subjected to HPLC analysis.

### 2.3.5 Generation of P*lpp* promoter library

To generate a P*lpp* promoter library, the basal plasmid pCS-*Plpp2.0-egfp* was constructed that *egfp* was under control of a synthesized *Plpp* variant promoter *Plpp2.0*. Degenerate primers containing randomized sequences covering the -10 box region of *Plpp2.0* were synthesized and applied to amplify the whole basal plasmid with a reverse primer adjacent to the degenerate primer. Both primers contain an *Eco*RI restriction site (5'-GAATTC-3'). The linearized PCR products were digested by *Eco*RI, ligated using DNA ligase. *E. coli* XL1-Blue were transformed with the ligated products for colony selections based on eGFP fluorescence difference. Approximately 180 colonies were picked and inoculated in LB medium. 1% of overnight inoculum was transferred into 200 μL medium in black 96-well plates and the eGFP fluorescence intensities were monitored by microplate reader (Biotek) using excitation wavelength of 520 nm and emission wavelength of 485 nm. A second round of screening was performed in test tubes by choosing colonies with gradient eGFP fluorescence for confirmation and plasmids were subjected for sequencing analysis.

### 2.3.6 Protein expression and purification

*E. coli* BL21 Star (DE3) were transformed with pETDuet-1 derived expression plasmids. Overnight inoculants were transferred into 50 mL LB medium and induced with 0.5 mM IPTG when the optical density ($OD_{600}$) of culture reached 0.6-0.8 and inoculated at 30 °C on a rotary

shaker for another 9-12 h. The cultures were harvested by centrifugation and the cells were lysed using Mini Bead Beater (Biospec). Protein in the supernatant was purified using His-Spin Protein Miniprep kit (Zymo Research, Irvine, CA). The purified protein was identified by 12 % SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and the protein concentration was measured using a BCA Protein Assay Kit (Pierce).

### 2.3.7 *In vitro* enzyme assays

The purified $His_6$-Eht1 and $His_6$-THT were used for *in vitro* enzyme assays. The 1 mL reaction system contained Tris-HCl (100 mM, pH=7.5), $MgCl_2$ (5 mM), ATP (5 mM), coenzyme A (0.25 mM), caffeic acid (0.2 mM), PE (1 mM) or PA (1 mM) and purified 4CL2 (0.2 μM) with or without 1 μM purified Eht1 or THT. After 30 min reaction, 100 μL 20 % HCl was added to stop the reaction and the products were subjected to HPLC analysis.

For caffeoyl-CoA degradation assay, the 1 mL reaction system contained Tris-HCl (100 mM, pH=7.5), $MgCl_2$ (5 mM), ATP (5 mM), coenzyme A (0.25 mM), caffeic acid (0.2 mM) and purified 4CL2 (0.2 μM) with or without 1 μM purified (thio)esterases or 200 μL crude enzyme extracts. The reactions were terminated in 30 min by 20 % HCl and analyzed by HPLC. To prepare the crude enzyme extracts, *E. coli* single knockout strains of putative (thio)esterase genes were purchased from Keio collection (CGSC) and inoculated in 3 mL LB medium overnight at 37 °C on a gyratory shaker at 250 rpm. 1 mL cell culture was harvested by centrifugation and cell pellets were resuspended in 1 mL of 10 mM Tris-HCl at pH 8.0 and added to 2 mL tubes containing 500 μL of 0.1 mm diameter glass beads for vortexing on a bead beater. The soluble fraction was collected by centrifugation and used as the crude enzyme extracts for the caffeoyl-CoA degradation assay.

### 2.3.8 HPLC Analysis

All samples from the cultures were centrifuged at 13,000 rpm for 10 min and the supernatants were subjected to HPLC (Dionex Ultimate 3000) with a reverse phase ZORBAX SB-C18 column and an Ultimate 3000 Photodiode Array Detector. Mobile phases used for peak detection were 0.1 % trifluoroacetic acid (TFA) in water (solvent A) and 100 % methanol (solvent B). Analytes were separated following a program with 5 to 80 % solvent B for 16 min, 80 to 5 % solvent B for 1min and 5 % solvent B for additional 4 min at a flow rate of 1 mL/min. Caffeic acid and its derived esters and amides were detected and quantified under UV light of wavelength 340 nm, while *p*-coumaric acid and its derived esters and amides were detected and quantified under wavelength 315 nm.

### 2.3.9 MS analysis

The caffeic acid derived esters and amides were extracted from the cultures with equivalent volume of acetyl acetate. Then the extract was dried by a vacuum speed evaporator, dissolved in dimethylsulphoxide (DMSO) and diluted with water. Electrospray ionization (ESI)-mass spectrometry (MS) and MS/MS of extracts were performed on a Thermo Scientific Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA).

## 2.4 Results

### 2.4.1 Designing biosynthetic pathways for caffeic acid derived phenethyl esters and amides

The natural biosynthetic pathways of caffeic acid derived phenethyl esters or amides were not fully elucidated. However, the mechanism of acyltransferase catalyzed formation of their analogs like rosmarinic acid and hydroxycinnamoyl anthranilates indicated the opportunity to establish

artificial biosynthetic routes based on acyltransferase-mediated trans-esterification [210,211]. Product-directed bioprospecting envisioned that, using caffeoyl-CoA as the scaffold architecture, caffeic acid derived phenethyl ester and amide analogs might be produced by acyltransferase-mediated coupling with different aromatic alcohol or amine counterparts. Thus, to render the biosynthesis of caffeic acid derived phenethyl esters and amides, three biosynthetic modules are to be introduced into *E. coli* chassis to assemble the entire biosynthetic pathway: the caffeic acid biosynthetic module, the aromatic alcohol or amine biosynthetic module, and finally the acyltransferase catalyzed trans-esterification module (**Figure 2.1**).

The caffeic acid biosynthetic module and the aromatic alcohol or amine biosynthetic module were both extended from aromatic amino acid pathways via either deamination or decarboxylation. In our previous work, an artificial and high-yielding two-step caffeic acid biosynthetic module has been constructed and optimized via successive deamination of tyrosine to *p*-coumaric acid by a tyrosine ammonia lyase from *Rhodotorula glutinis* (RgTAL) and hydroxylation of *p*-coumaric acid by a 4-hydroxyphenylacetate 3-hydroxylase (HpaBC) from *E. coli* [178,208]. To generate aromatic alcohols including PE and tyrosol, a promiscuous 2-keto-acid decarboxylase (KDC) KivD from *Lactococcus lactis* and an alcohol dehydrogenase (ADH) Adh6 from *S. cerevisiae* were involved, which were respectively responsible for successive decarboxylation and dehydrogenation of phenylpyruvate and 4-hydroxyphenylpyruvate [121] (**Figure 2.1**). 3-Hydroxytyrosol could be produced from 4-hydroxyphenylpyruvate or tyrosol owing to the catalytic promiscuity of KivD or HpaBC [212]. To test the functionality of KivD and Adh6, 500 mg/L of phenylpyruvate, 4-hydroxyphenylpyruvate and 3, 4-hydroxyphenylpyruvate were fed to *E. coli* BW25113 (F') (**Table S2.1**) co-expressing the two enzymes. High performance liquid chromatography (HPLC) analysis confirmed that 292.8, 145.3 and 193.3

mg/L of PE, tyrosol and 3-hydroxytyrosol were respectively produced in the media at 24 h after substrate feeding and induction by 0.5 mM IPTG (**Figure S2.1A**). Notably, *E. coli* also consumed phenylpyruvate, 4-hydroxyphenylpyruvate and 3, 4-hydroxyphenylpyruvate to produce corresponding aromatic amino acids (data not shown).

To generate aromatic amines including PA, tyramine (TA) and dopamine (DA), aromatic amino acid decarboxylases (AADCs) of different bacterial origins were introduced and screened [213-215]. The phenylalanine decarboxylase (LbPDC) from *Lactobacillus brevis*, tyrosine decarboxylase (MjTDC) from *Methanocaldococcus jannaschii* and L-Dopa decarboxylase (PpDDC) from *Pseudomonas putida* were over-expressed on pCS27 plasmids and tested by feeding aromatic amino acids including phenylalanine, tyrosine and L-Dopa (500 mg/L). LbPDC enabled production of 322.8, 560.0 and 432.6 mg/L PA, TA and DA from counterpart aromatic amino acid precursors, while MjTDC decarboxylated tyrosine and L-Dopa, producing 342.9 and 185.2 mg/L TA and DA, respectively (**Figure S2.1B** and **C**). PpDDC almost specifically decarboxylated L-Dopa, producing 477.6 mg/L DA and trace amounts of PA (30.4 mg/L) (**Figure S2.1D**). These results indicated that LbPDC had broad substrate specificity and high activity towards all three aromatic amino acids, while MjTDC had a higher preference towards tyrosine and PpDDC exhibited substrate specificity only towards L-Dopa as demonstrated previously [162]. Thus, all feeding experiments indicated that precursor biosynthetic modules could be readily recruited by capitalizing corresponding heterologous enzymes.

The final steps for biosynthesis of caffeic acid phenethyl esters or amides included (1) activation of caffeic acid into caffeoyl-CoA by an ATP-dependent plant enzyme 4-coumarate-CoA ligase (4CL) and (2) enzymatic trans-esterification of caffeoyl-CoA with alcohol or amine counterparts by functional alcohol O-acyltransferases (O-ATFs) or amine N-acyltransferases (N-

ATFs). 4CL2 from *Arabidopsis thaliana* has been demonstrated with high catalytic turnover and specificity towards caffeic acid [216]. However, native or functional acyltransferases for caffeic acid derived esters or amides were not well-documented.

### 2.4.2 Identification and characterization of functional ATFs

To identify functional O-ATF for CAPE formation, four candidate alcohol O-acyltransferases including Atf1, Atf2, Eeb1 and Eht1 from *S. cerevisiae* were screened. Additionally, chloramphenicol acetyltransferase (CAT) was also tested due to its capability of using PE as an acyl acceptor; codon-optimized hydroxycinnamoyl transferase (HCT) from *Nicotiana tabacum* was examined due to its viability of producing caffeoyl shikimate and chlorogenic acid [201,217]. All six O-ATFs were respectively co-expressed with 4CL2 from *A. thaliana* in *E. coli* BW25113 (F'). When fed with caffeic acid (500 mg/L) and 2-phenylethanol (500 mg/L), only *E. coli* cells co-expressing Eht1 and 4CL2 were capable of producing CAPE with a titer of 52.9 mg/L at 24 h post-feeding (**Table S2.2**). The retention time (RT) of CAPE product (RT=16.3 min) was in good agreement with the CAPE standard in HPLC analysis and LC/MS analysis with a product ion of m/z 283.0 in negative mode (**Figure 2.2A**). The fragmentation pattern of CAPE product in MS/MS analysis was consistent with CAPE standard as previously reported [218]. This indicated that Eht1 catalyzed the trans-esterification of caffeoyl-CoA with PE and released authentic CAPE. Similarly, 35.0 mg/L CAHPE was also produced when fed with CA and tyrosol (**Table S2.2**, **Figure S2.2A** and **D**), demonstrating the viability of Eht1 for formation of CAPE ester analogs. Regarding N-ATF for caffeic acid derived amide formation, we selected an optimal candidate hydroxycinnamoyl-CoA: tyramine N-(hydroxycinnamoyl) transferase (THT) from *Capsicum annuum*, which has been well demonstrated with a broad substrate specificity and

accepts acyl donors such as cinnamoyl-, 4-coumaroyl-, caffeoyl- and feruloyl-CoA and a wide range of acyl acceptors [219]. When fed with corresponding precursors (500 mg/L), *E. coli* cells co-expressing codon-optimized THT and 4CL2 produced 117.5, 132.8 and 79.9 mg/L CAPA, CATA and CADA, respectively (**Table S2.2**). All products were in good agreement with synthesized standards in HPLC analysis and were further confirmed by LC/MS analysis (**Figure 2.2B, Figure S2.3A, E** and **F**).

Interestingly, both Eht1 and THT exhibited catalytic promiscuity towards *p*-coumaric acid that respectively afforded production of *p*-coumaric acid phenethyl ester (49.6 mg/L), *p*-coumaric acid (4-hydroxyphenethyl) ester (6.8 mg/L) (**Figure S2.2A-C**), *p*-coumaroyl phenylethylamine (32.4 mg/L), *p*-coumaroyl tyramine (77.3 mg/L) and *p*-coumaroyl dopamine (126.9 mg/L) when fed with corresponding precursors (**Figure S2.3A-D**). This underlined that both Eht1 and THT could accommodate hydroxycinnamoyl-CoA as acyl donors.

To validate the catalytic function of Eht1 and THT, the two enzymes as well as 4CL2 were expressed with N-terminally $His_6$-tagged proteins and purified to homogeneity from *E. coli* BL21 Star (DE3) for a coupled enzyme assay (**Figure 2.3A**). In the presence of ATP and free CoA, 0.2 mM CA was almost completely converted to caffeoyl-CoA by $His_6$-4CL2 (0.2 μM) within 30 min (**Figure 2.3B**). When $His_6$-Eht1 (1 μM) and PE (1 mM) were added, CAPE was produced albeit in a trace amount (RT=16.3 min), indicating that Eht1 catalyzed the trans-esterification of caffeoyl-CoA with PE for CAPE formation with limited efficiency. In contrast, THT catalyzed the trans-esterification of caffeoyl-CoA with PA but with much higher activity, in which caffeoyl-CoA was completely consumed to form CAPA (RT=13.4 min) in 30 min (**Figure 2.3B**). To test if Eht1 or THT has any thioesterase activity, caffeic acid was first incubated with $His_6$-4CL2 and then with $His_6$-Eht1 or $His_6$-THT for 2 h. $His_6$-Eht1 resulted in substantial

degradation of caffeoyl-CoA to caffeic acid, while $His_6$-THT did not result in any measurable degradation (**Figure 2.3B**). The *in vitro* enzyme assays indicated that both Eht1 and THT exhibited acyltransferase activity towards caffeoyl-CoA, whereas Eht1 was endowed with additional innate thioesterase activity towards caffeoyl-CoA. This was consistent with higher *in vivo* bioconversion ratio of amide analogs than ester analogs when exogenously fed with similar amounts of co-substrates.

### 2.4.3 *De novo* production of CAPE via tuning phenylalanine branch

After identification and validation of functional acyltransferases, we sought to achieve *de novo* production of caffeic acid derived esters and amides by combinatorial incorporation of all pathway genes in *E. coli.* We first attempted to achieve CAPE production in M9 minimal medium containing 20 g/L glucose as a proof-of-concept demonstration. The entire CAPE biosynthetic pathway involves four over-expressed endogenous genes and six exogenous genes, which were assembled into two plasmids: (1) pZE-*TH-KA* containing the codon-optimized *tal* from *R. glutinis* and *hpaBC* from *E. coli* BL21 Star (DE3) in separate operons while *kivD* from *L. lactis* and *adh6* from *S. cerevisiae* in another operon; (2) pCS-*TPTA-EC* containing *tyrA**, *pps*, *tktA* and *aroG** from *E. coli* in one operon while *Eht1* from *S. cerevisiae* and *4CL2* from *A. thaliana* in another operon (**Figure 2.4D**). More specifically, feedback inhibition resistant chorismate mutase-prephenate dehydrogenase (*tyrA**) and 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (*aroG**), as well as PEP synthase (*pps*) and transketolase (*tktA*) were over-expressed to increase carbon flux into shikimate pathway. Additionally, *tyrR* was knocking out in *E. coli* BW25113 (F') to deregulate shikimate pathway genes. Introduction of pZE-*TH-KA* and pCS-*TPTA-EC* into *E. coli* BW25113 (F') Δ*tyrR* enabled successful production of CAPE with a

maximum titer of 6.0 mg/L at 24 h (**Figure 2.4D**). In addition, the most hydroxylated CAPE analog CADPE, was also accumulated with maximum titer of 7.7 mg/L at 48 h (data not shown). This might be ascribed to the catalytic promiscuity of KivD and HpaBC. Interestingly, when *E. coli* BW25113 (F') Δ*tyrR* was transformed with pZE-*TH-KA* and pCS-*TPTA* for shake flask experiments, no production of CAPE or CADPE was observed as expected, whereas CA was produced almost 2-fold than PE during shake flask experiments (**Figure 2.4C**). This indicated an imbalance of co-substrates CA and PE in the parental host.

Since CA and PE are both extended from the shikimate pathway at the branch point chorismate, we attempted to balance the co-substrates by titrating the expression of the *pheA* via an engineered P*lpp* promoter library in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*. P*lpp* is a strong constitutive promoter of lipoprotein Lpp in *E. coli* [220]. One enhanced variant P*lpp2.0* was previously generated from wild type P*lpp* (P*lpp1.0*) via mutating the -35 and -10 boxes [220]. A plasmid-based eGFP reporter system was constructed by placing *egfp* under control of P*lpp2.0* (pCS-P*lpp2.0-egfp*) and -10 box of P*lpp2.0* were subjected to random mutations (**Figure 2.4A**). Fluorescence-based screening enabled generation of a P*lpp* promoter library (P$lpp_x$) with varying promoter strengths from 0.03 to 2.0-fold relative to P*lpp1.0* (**Figure 2.4B, Figure S2.4**). Such promoter based tuning of PheA might permit modulation of endogenous carbon flux between the tyrosine and phenylalanine branches, thereby allowing maximum productivity with an optimal trade-off between co-substrates CA and PE [221,222].

We expressed *pheAev1* (named *pheA** here), coding a truncated feedback inhibition resistant PheA (1-303 aa) with two point mutants Thr259Ser and Tyr230Leu[223], under control of P*lpp* promoter library on the low copy number plasmid (pSA-P$lpp_x$-*pheA**). Introduction of pSA-P$lpp_x$-*pheA** along with pZE-*TH-KA* and pCS-*TPTA* into *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*

enabled shunting chorismate into phenylalanine branch for PE formation. The production profiles in shake flask experiments showed that CA was produced with a titer around 2-fold to PE in *E. coli* BW25113 (F') Δ*tyrR*, whereas the production of CA and PE was well balanced to an almost equivalent titer only when *pheA** was expressed under control of P*lpp0.03* or P*lpp0.2* in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* (**Figure 2.4C**). Subsequent combinatorial introduction of pSA-P*lpp0.2-pheA** with the CAPE pathway plasmids pZE-*TH-KA* and pCS-*TPTA-EC* into *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* afforded CAPE production of 8.6 mg/L, a 43.3% increase compared with the unbalanced control strain (**Figure 2.4D**). The results underlined that titrating of the tyrosine and phenylalanine branches via promoter-mediated tuning of branching enzyme PheA* enabled balance of the co-substrates, which in turn effectively improved the CAPE titer.

### 2.4.4 Host strain engineering with deletion of thioesterase(s)

Considering the prevalence of endogenous (thio)esterases in *E. coli* cells, the thioester intermediate caffeoyl-CoA might be degraded to caffeic acid, as has been reported that *p*-coumaroyl-CoA could be degraded by endogenous (thio)esterase(s) like YbgC [224]. We set out to systematically investigate the caffeoyl-CoA degradation activity of 15 native known or putative (thio)esterases from host *E. coli* cells, including BioH, FadM, TesB, YbaC, EntH, YbfF, YciA, PaaI, PaaY, YeiG, YpfH, YqiA, YjfP, YbgC and YdiI. First, to individually demonstrate their *in vitro* activity, all (thio)esterases were over-expressed as N-terminally $His_6$-tagged proteins and purified to homogeneity from *E. coli* BL21 Star (DE3) as described previously [209]. Caffeoyl-CoA degradation assay was performed by first incubating caffeic acid (0.2 mM) with $His_6$-4CL2 (0.2 μM) and then subjected to each purified (thio)esterase (1 μM). Among all (thio)esterases tested, only TesB, YjfP, YbgC and YdiI led to severe degradation of caffeoyl-CoA and thus formation

of caffeic acid (**Figure 2.5A**). Considering the expression level or *in vivo* conditions that might affect the *in situ* contribution of (thio)esterase(s) on caffeoyl-CoA degradation, we further verified if deletion of these potential caffeoyl-CoA degrading (thio)esterases in native *E. coli* host cells would alleviate caffeoyl-CoA degradation. Cell lysates of *E. coli* BW25113 single gene knockout strains (Δ*tesB*, Δ*yjfP*, Δ*ybgC* and Δ*ydiI*) were used as crude enzymes for the caffeoyl-CoA degradation assay. Among all four knockout strains tested, only cell lysates of the *ydiI* knockout strain showed significantly reduced caffeoyl-CoA degradation (**Figure 2.5B**), indicating that the thioesterase YdiI played a major role in endogenous caffeoyl-CoA degradation in native *E. coli* cells. Thus, it is practical to solely knock out *ydiI*; ideally that would maintain caffeoyl-CoA availability for production of caffeic acid derived compounds.

To examine that hypothesis, CA (500 mg/L) and PE or PA (500 mg/L) were fed into the engineered host strain *E. coli* BW25113 (F') Δ*ydiI* co-expressing Eht1 and 4CL2 or THT and 4CL2, respectively. Compared with BW25113 (F') as the host, 58.4 mg/L CAPE and 173.8 mg/L CAPA were produced in the *E. coli* BW25113 (F') Δ*ydiI* host, representing 10.2 % and 74.7 % increases, respectively (**Figure 2.5C**). This suggested that deletion of *ydiI* in *E. coli* host did facilitate production of CAPE and especially CAPA via enhancing the endogenous caffeoyl-CoA pool. Thus, the final platform strain *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*/Δ*ydiI* were created and transformed with pZE-*TH-KA*, pCS-*TPTA-EC* and pSA-*Plpp0.2-pheA** (**Figure 2.4D**). Shake flask experiments confirmed that the final strain produced 9.2 mg/L CAPE at 48 h, which is just slightly increased compared to that without *ydiI* knockout. However, this is not surprising considering the low activity of Eht1 and the slight increase even in feeding experiments.

### 2.4.5 *De novo* production of CAPE ester and amide analogs

We next sought to produce hydroxylated CAPE analogs including CAHPE and CADPE that are derived from the tyrosine branch in the optimized host strains. Both CAHPE and CADPE were produced by leveraging promiscuous enzymes from CAPE pathway. Simply, the pathway plasmids pZE-*TH-KA* and pCS-*TPTA-EC* were introduced into *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*/Δ*ydiI*. Shake flask experiments of the transformants confirmed that CAHPE and CADPE were simultaneously produced with a maximum titer of 14.2 and 47.9 mg/L at 48 h, respectively, without any detectable accumulation of CAPE (**Figure 2.6A**). This indicated that the catalytic promiscuity of HpaBC enabled higher production of the most hydroxylated CADPE by hydroxylating tyrosol and/or CAHPE. This was underlined by its capability of hydroxylating bulk aromatic compounds [225] and explained the accumulation of CADPE in CAPE production strain.

In contrast to production of aromatic alcohols that utilizes the catalytic promiscuity of KivD, aromatic amino acid decarboxylases LbPDC, MjTDC and PpDDC were utilized to convert phenylalanine, tyrosine and L-Dopa to aromatic amines phenylethylamine, tyramine and dopamine, respectively. Unfortunately, CAPA production was not successful via either introducing pZE-*TH-LbPDC* with pCS-*TPTA-TC* into *E. coli* BW25113 (F') Δ*tyrR* or using P*lpp* library based balancing of CA and PA in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*/Δ*ydiI* (data not shown). This was most likely ascribed to the promiscuity of LbPDC towards all aromatic amino acids and its relative low activity towards phenylalanine (**Figure S2.1B**). To produce CATA and CADA, the entire pathway plasmids pCS-*TPTA-TC* and pZE-*TH-MjTDC* or pZE-*TH-PpDDC* were introduced into *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*. Shake flask experiments of both strains confirmed that 17.0 mg/L CATA and 195.7 mg/L CADA were produced. However, when

BW25113 (F') Δ*tyrR*/Δ*pheA*/Δ*ydiI* was applied as host, 52.9 mg/L CATA and 369.1 mg/L CADA were produced, which were respectively 3.1-fold and 1.9-fold to those produced in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* (**Figure 2.6B**). This further underlined the contribution of *ydiI* knockout to the production enhancement of caffeic acid derived amides. Additionally, higher production of caffeic acid derived amides than esters was consistent with the higher activity of THT as well as the higher stability of caffeic acid derived phenethyl amides than esters. This engineered strain would represent a good bacterial platform for production of caffeic acid derived esters and amides from simple carbon sources.

## 2.5 Discussion

Hydroxycinnamic acid esters (HCEs) and amides (HCAAs) are a widely distributed and phytochemically important group of plant secondary metabolites with interesting pharmaceutical potentials. Among those HCEs and HCAAs, caffeic acid derived esters and amides are of the most interest because of potent anti-oxidative activity and pharmaceutical applications. Development of microbial constructs for production of these valuable compounds has been increasingly achieved in recent years, including successful production of hydroxycinnamoyl shikimates, hydroxycinnamoyl anthranilates, rosmarinic acid, chlorogenic acid, *p*-coumaroyl phenethylamine, and *p*-coumaroyl tyramine [211,226-231]. Generally, expression of specific plant-derived BAHD acyltransferases that were capable of catalyzing O- or N-acylation would afford synthesis of multiple hydroxycinnamate and even benzoate conjugates [232]. However, natural pathways for some of these, like caffeic acid derived phenethyl esters and amides, have remained elusive. Very recently, a benzoyl-CoA:benzyl alcohol/phenylethanol benzoyltransferase (BPBT) from petunia (*Petunia hybrida*) has been demonstrated with promiscuous activity to

hydroxycinnamoyl-CoA and 2-phenylethanol for synthesis of CAPE [232]. In the present study, we identified a functional O-acyltransferase Eht1 from the yeast *S. cerevisiae* and assembled an artificial pathway for CAPE production from glucose in *E. coli*. Based on that, we exploited the biosynthetic platform for production of a panel of CAPE ester and amide analogs. Particularly, formation of esters or amides was favored by the release of CoA from caffeoyl-CoA and the driving force of the irreversible loss of $CO_2$ or $NH_3$ during precursor preparations.

To mimic the natural acyltransferase-mediated synthesis of caffeic acid based esters, six O-ATFs including Atf1, Atf2, Eht1, Eeb1, CAT and HCT with broad substrate specificity were screened. Eht1 has been reported to prefer larger acyl-CoA units especially medium chain octanoyl-CoA (C8) as CoA donor and larger alcohols including 2-phenylethanol as acceptors [201,233]. Herein, we expanded its capability of accommodating *p*-coumaroyl-CoA or caffeoyl-CoA, affording production of *p*-coumaric acid or caffeic acid derived phenethyl esters. However, the low yield of CAPE formation indicated that Eht1 might have low activity towards aromatic acyl-CoA and/or aromatic alcohols, as implied by the reduced activity of Eht1 towards longer chain acyl-CoA (>C8) [233]. Eht1 did not exhibit any esterase activity towards the final ester CAPE (data not shown), but exhibited innate thioesterase activity towards caffeoyl-CoA (**Figure 2.3B**). This was perhaps not unexpected, given that Eht1 has been previously demonstrated to function not only as an acyltransferase but also as a thioesterase, which might contribute to the low production level of CA-derived esters [233]. Indeed, formation of CA-derived amides was more readily achieved, probably because THT had higher catalytic activity with no detectable thioesterase activity and the amide linkage is more rigid and hydrolysis resistant [197,231]. Especially, the high activity of pathway enzymes like HpaBC, DDC and THT concomitantly contributed to the high titer of the most hydroxylated CA-derived amide CADA.

Biosynthetic pathways with competing branching routes will always lead to an imbalance of co-substrates, ascribed to the undesirable carbon flux distribution and catalytic capacities of pathway enzymes. As co-substrates of CAPE, CA and PE are respectively derived from tyrosine and phenylalanine branches from chorismate. We created a P*lpp*-based constitutive promoter library and enabled balance of CA and PE via tuning the expression level of branching enzyme PheA*. Interestingly, low expression of *pheA** in an engineered CA producer afforded almost equivalent co-substrates, which might be due to the high activity of PheA* [223]. The enhancement of CAPE production highlighted the feasibility of promoter mediated re-tuning of carbon flux to maximize titers of desired conjugated products. Unfortunately, attempts to balance CA and PA for CAPA production were unsuccessful, primarily because PDC unspecifically decarboxylates Phe, Tyr, and L-Dopa.

Furthermore, considering the prevalence of thioesterase(s) in *E. coli* cells that might degrade the thioester intermediate caffeoyl-CoA, unbiased deletion of all thioesterases was impractical and untenable. Systematic investigation of caffeoyl-CoA degrading activity of 15 endogenous thioesterases in *E. coli* allowed identification of one thioesterase YdiI that might play a predominant role in endogenous degradation of intracellular caffeoyl-CoA. YdiI is the 1,4-dihydroxy-2-naphthoyl-CoA thioesterase enzyme that belongs to the 4-hydroxybenzoyl-CoA thioesterase subfamily, and has been previously reported to show high *in vitro* catalytic activity towards bulk CoA thioesters including *p*-coumaroyl-CoA [234,235] and salicoyl-CoA [209]. Knocking out *ydiI* in *E. coli* cells significantly reduced degradation of endogenous caffeoyl-CoA. Given the low catalytic activity and innate thioesterase activity of Eht1, caffeic acid derived esters were not significantly improved in *ydiI* knockout strains. However, the titers of caffeic acid derived amides were significantly improved in *ydiI* knockout strains, supporting the contribution of *ydiI*

deletion to enhancing caffeoyl-CoA availability. *E. coli* can be easily manipulated to a tyrosine overproducer and numerous engineering work has been performed to achieve that [236-240]. Herein, by deleting caffeoyl-CoA degrading thioesterases like *ydiI*, the biosynthetic platform could be further expanded to increase the endogenous caffeoyl- or *p*-coumaroyl-CoA pool for synthesis of CoA-dependent compounds like esters and amides achieved in this work, and many other valuable polyphenol compounds.

In conclusion, we established the artificial biosynthetic pathways for caffeic acid phenethyl esters and amides, which illustrate and demonstrate the engineering of a bacterial platform capable of producing a total of five valuable natural compounds from glucose. Further optimizations include: (1) balancing co-substrates via implementation of an engineered promoter library and (2) enhancing caffeoyl-CoA pool via deletion of thioesterase gene *ydiI*. The biosynthetic platform developed here highlights a general strategy of assembly of non-natural pathways for production of natural aromatic amino acid derived products. Identification of efficient O-ATFs or engineering of rate-limiting Eht1 is needed to achieve high titers of caffeic acid derived phenethyl esters. Future improvements like protein engineering, pathway engineering, and fermentation optimization might enable a significant advance in the metabolic engineering of caffeic acid derived phenethyl esters and amides biosynthesis, and the future establishment of a potential industrial process.

## 2.6 Tables and figures

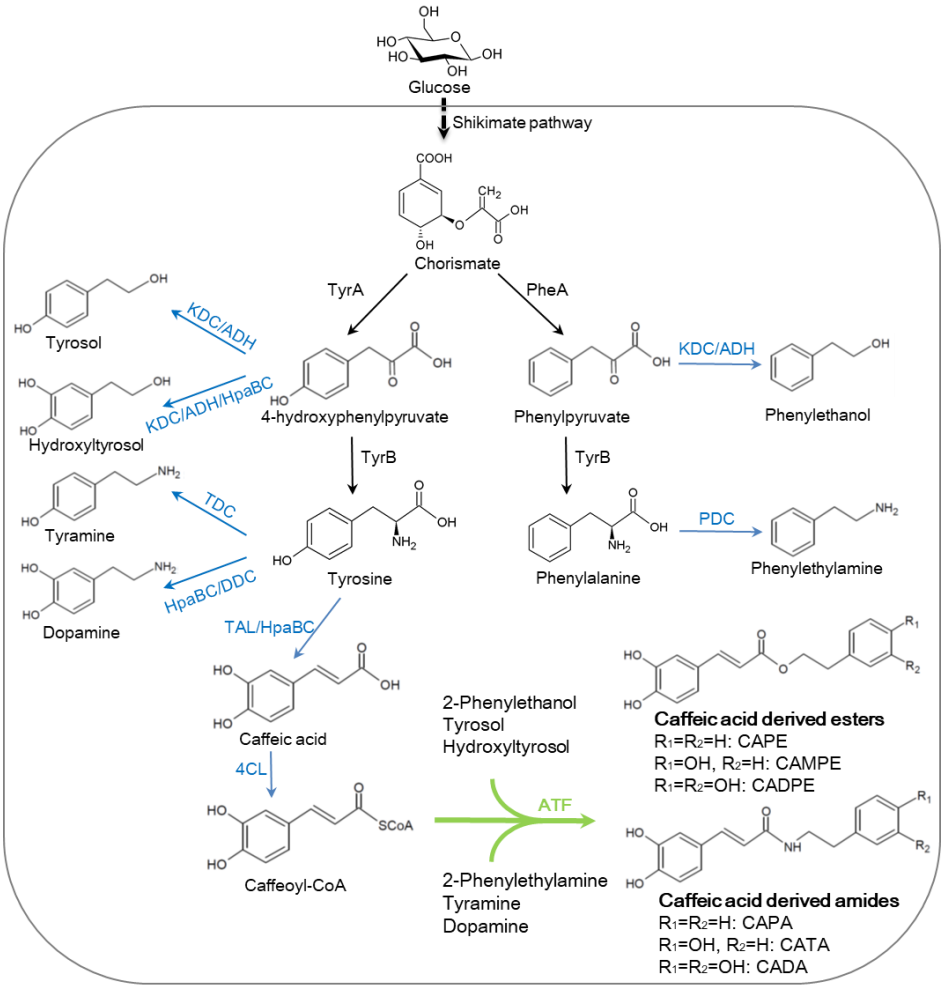



**Figure 2.1 Engineering artificial biosynthetic routes for caffeic acid derived phenethyl esters and amides in *E. coli*.** Endogenous shikimate pathway is responsible for production of aromatic amino acids from glucose in *E. coli*. Heterologous enzymes for substrates are shown in blue and ATF for the final trans-esterification step is shown in green. TyrA, chorismate mutase/prephenate dehydrogenase; PheA, chorismate mutase / prephenate dehydratase; TyrB, tyrosine aminotransferase; KDC, 2-keto acid decarboxylase; ADH, alcohol dehydrogenase; HpaBC, 4-hydroxyphenylacetate 3-hydroxylase; PDC, phenylalanine decarboxylase; TDC, tyrosine decarboxylase; DDC, L-Dopa decarboxylase; ATF, acyltransferase.

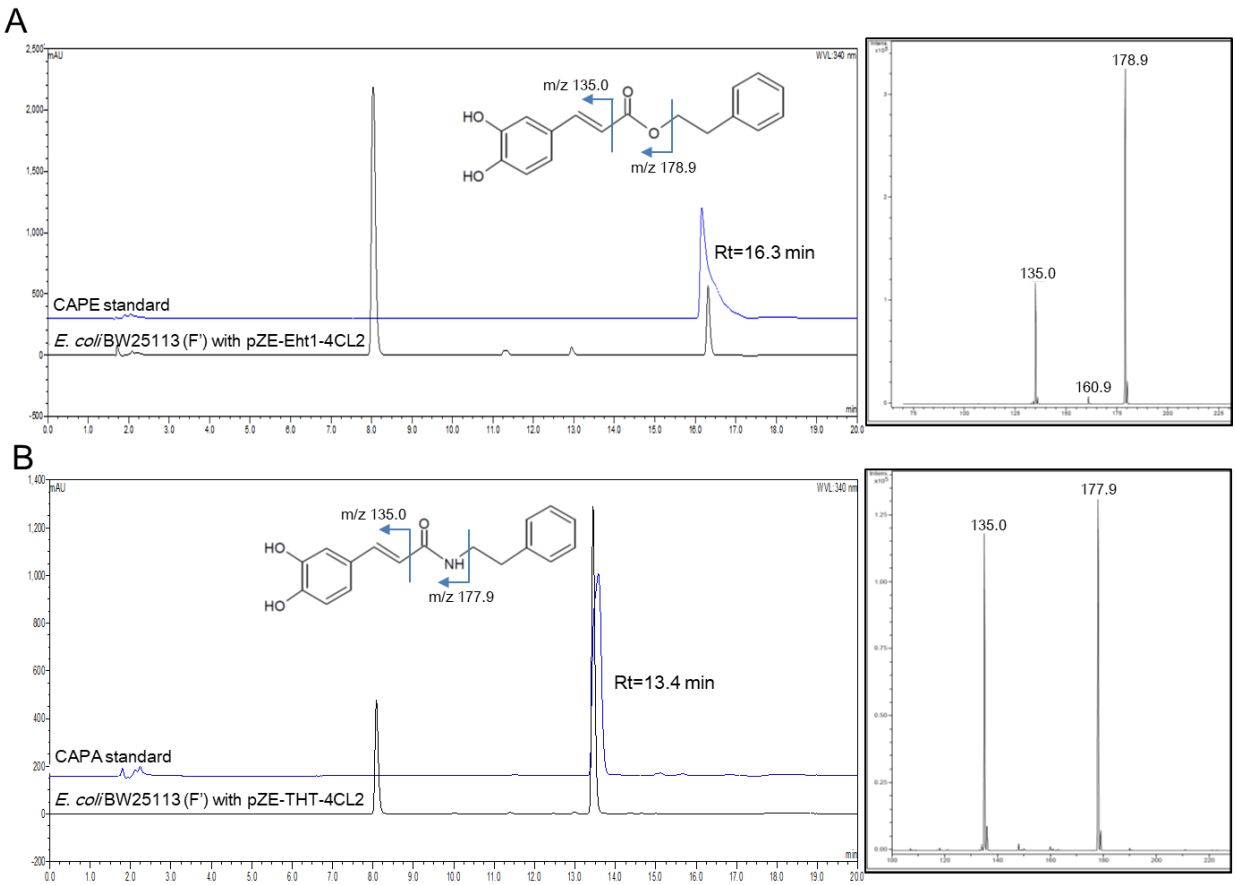



**Figure 2.2 Acyltransferases mediated production of caffeic acid derived phenethyl ester and amide.** (A) HPLC analysis of culture supernatants of *E. coli* BW25113 (F') co-expressing Eht1 and 4CL2 when fed with CA and PE after 24 h. (B) HPLC analysis of culture supernatants of *E. coli* BW25113 (F') co-expressing THT and 4CL2 when fed with CA and PA after 24 h. 200 mg/L CAPE purchased and CAPA synthesized were respectively used as standard. The peak at retention time of 8.1 min corresponds to CA. The right panels in (A) and (B) are tandem mass results for both negative product ions (m/z 283.0 for CAPE and m/z 282.0 for CAPA) in ESI-MS analysis.

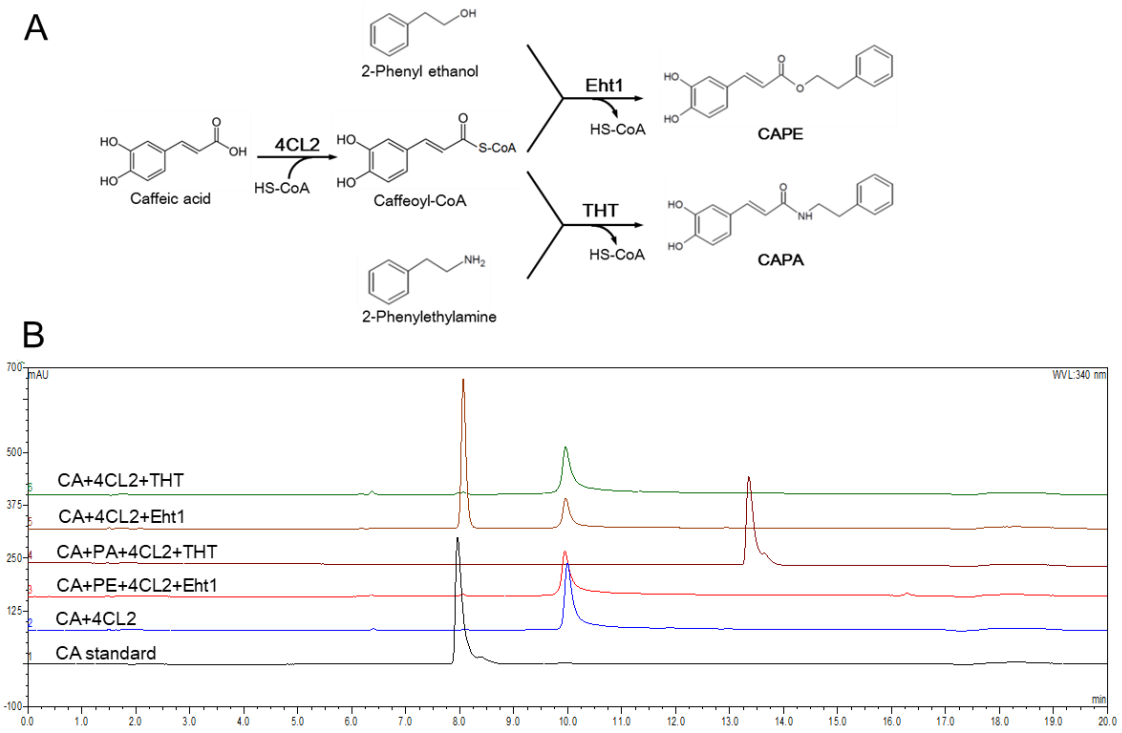



**Figure 2.3 *In vitro* charaterization of Eht1 and THT.** (A) Scheme of coupled *in vitro* enzyme assay. 4CL2 converts caffeic acid into caffeoyl-CoA in the presence of free CoA. Eht1 and THT respectively transfer caffeoyl-CoA to acceptors 2-phenylethanol and 2-phenylethylalamine. (B) *In vitro* enzymatic synthesis of caffeic acid derived phenethyl esters or amides. Each channel denotes: 1, CA standard; 2, CA with $His_6$-4CL2; 3, CA and PE with $His_6$-4CL2 and $His_6$-Eht1; 4, CA and PA with $His_6$-4CL2 and $His_6$-THT; 5, CA with $His_6$-4CL2 and $His_6$-Eht1; 6, CA with $His_6$-4CL2 and $His_6$-THT.

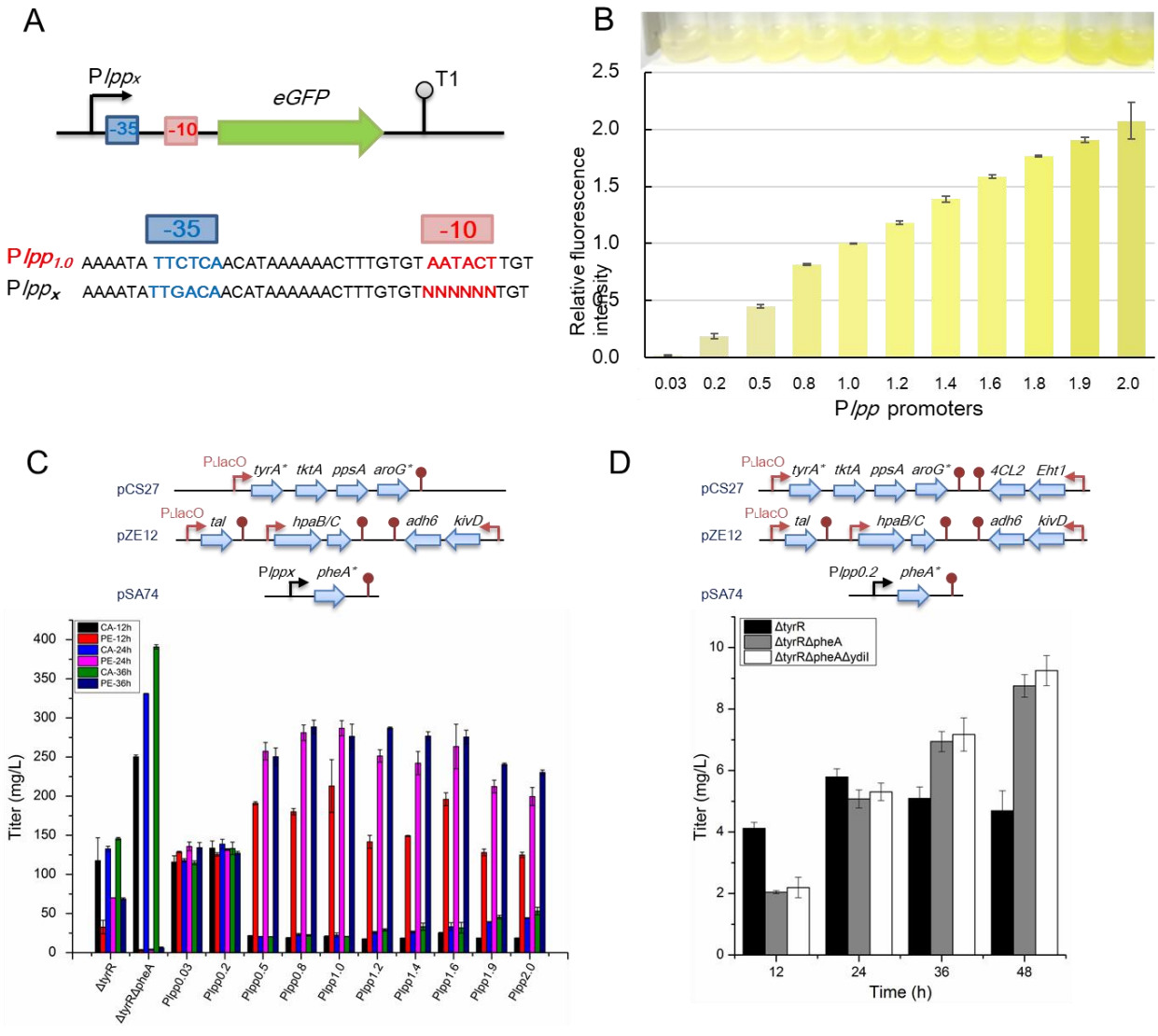

A
P lppx
-35
-10
eGFP
T1
-35
-10
P lpp1.0 AAAATA TTCTCAACATAAAAAACTTTGTGT AATACT TGT
P lppx AAAATATTGACAACATAAAAAACTTTGTGTNNNNNNTGT
B
Relative fluorescence intensity
0.0
0.5
1.0
1.5
2.0
2.5
0.03 0.2 0.5 0.8 1.0 1.2 1.4 1.6 1.8 1.9 2.0
P lpp promoters
C
PLlacO tyrA* tktA ppsA aroG*
pCS27
PLlacO tal hpaB/C adh6 kivD
pZE12
Plppx pheA*
pSA74
CA-12h
PE-12h
CA-24h
PE-24h
CA-36h
PE-36h
Titer (mg/L)
0 50 100 150 200 250 300 350 400
ΔtyrR
ΔtyrRΔpheA
Plpp0.03
Plpp0.2
Plpp0.5
Plpp0.8
Plpp1.0
Plpp1.2
Plpp1.4
Plpp1.6
Plpp1.9
Plpp2.0
D
PLlacO tyrA* tktA ppsA aroG* 4CL2 Eht1
pCS27
PLlacO tal hpaB/C adh6 kivD
pZE12
Plpp0.2 pheA*
pSA74
ΔtyrR
ΔtyrRΔpheA
ΔtyrRΔpheAΔydiI
Titer (mg/L)
0 2 4 6 8 10
12 24 36 48
Time (h)

**Figure 2.4 *De novo* production of CAPE by balancing precursors via promoter engineering.** (A) Promoter engineering of the -10 box in *Plpp* based on eGFP fluorescence. In wild type P*lpp1.0*, the -35 box (blue box) is TTCTCA and the -10 box (red box) is AATACT. In engineered P*lppx* promoters, the -35 box is the same with P*lpp2.0* as TTGACA, while the -10 box is subjected to mutations. (B) Fluorescence based screening enabled generation of a *Plppx* promoter library with promoter strength ranging from 0.03 to 2.0 folds of *Plpp1.0*. The fluorescence intensities of selected clones were normalized based on P*lpp1.0*. (C) Balancing of co-substrates CA and PE by titrating expression of *pheA** in BW25113 (F') Δ*tyrR*/Δ*pheA* using P*lpp* promoter library. CA and PE titers were measured and compared in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* transformed with pCS-*TPTA*, pZE-*TH-KA* and pSA-P$lpp_x$-*pheA**. *E. coli* BW25113 (F') Δ*tyrR* and Δ*tyrR*/Δ*pheA* transformed with pCS-*TPTA* and pZE-*TH-KA* were applied as controls. (D) *De novo* production of CAPE in *E. coli* BW25113 (F') Δ*tyrR* (black), Δ*tyrR*/Δ*pheA* (grey) and Δ*tyrR*/Δ*pheA*/Δ*ydiI* (white). *E. coli* BW25113 (F') Δ*tyrR* was transformed with pZE-*TH-KA* and pCS-*TPTA-EC*, while *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* and Δ*tyrR*/Δ*pheA*/Δ*ydiI* were transformed with pZE-*TH-KA* and pCS-*TPTA-EC* and pSA-*Plpp0.2-pheA**.

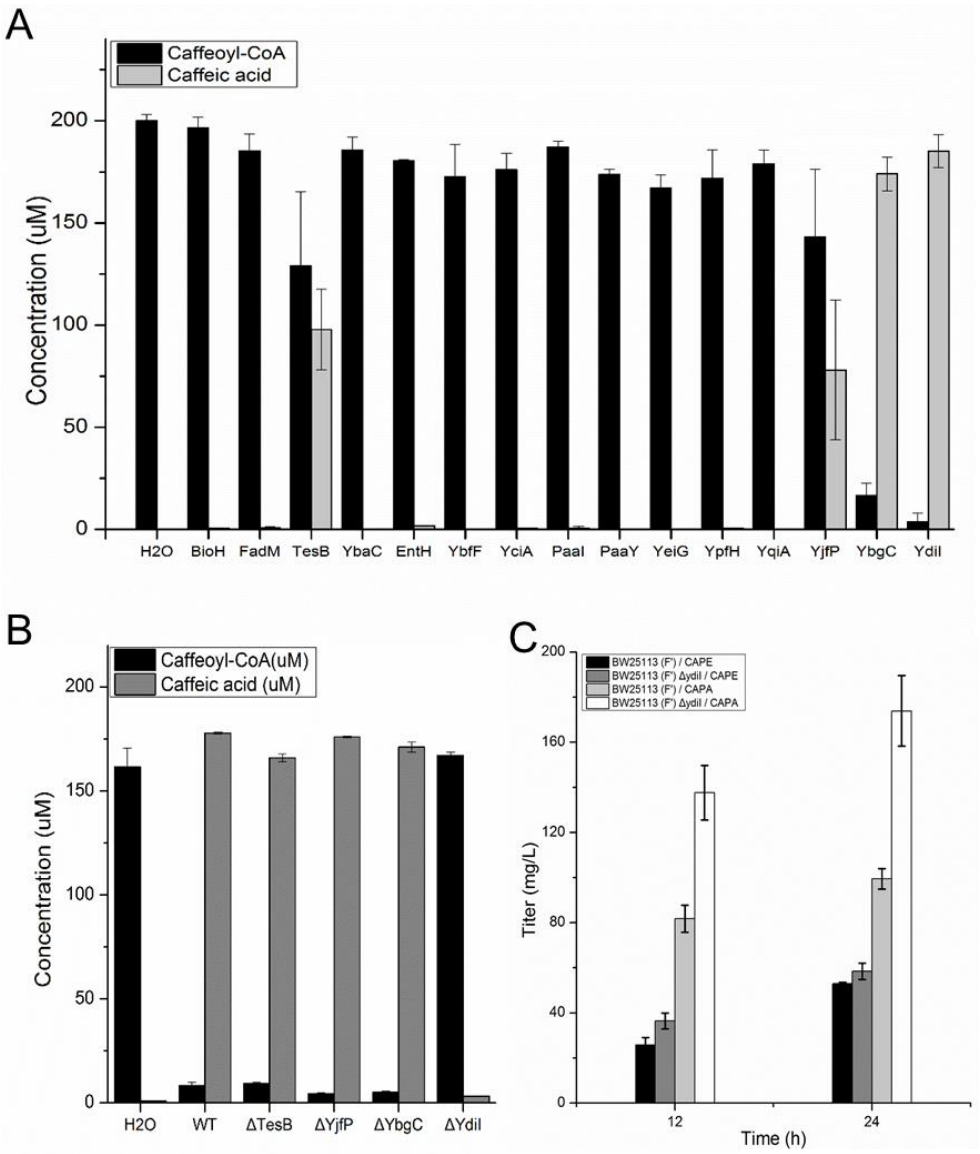



**Figure 2.5 Characterization of (thio)esterases on caffeoyl-CoA degradation.** (A) Degradation of caffeoyl-CoA by 15 purified (thio)esterases from *E. coli*. All (thio)esterases were N-terminally $His_6$-tagged. (B) Degradation of caffeoyl-CoA by crude enzyme of *E. coli* BW25113 (F') and *E. coli* BW25113 single gene knockout strains including Δ*tesB*, Δ*yjfP*, Δ*ybgC* and Δ*ydiI*. (C) Production of CAPE and CAPA in *E. coli* BW25113 (F') and *E. coli* BW25113 (F') Δ*ydiI*.

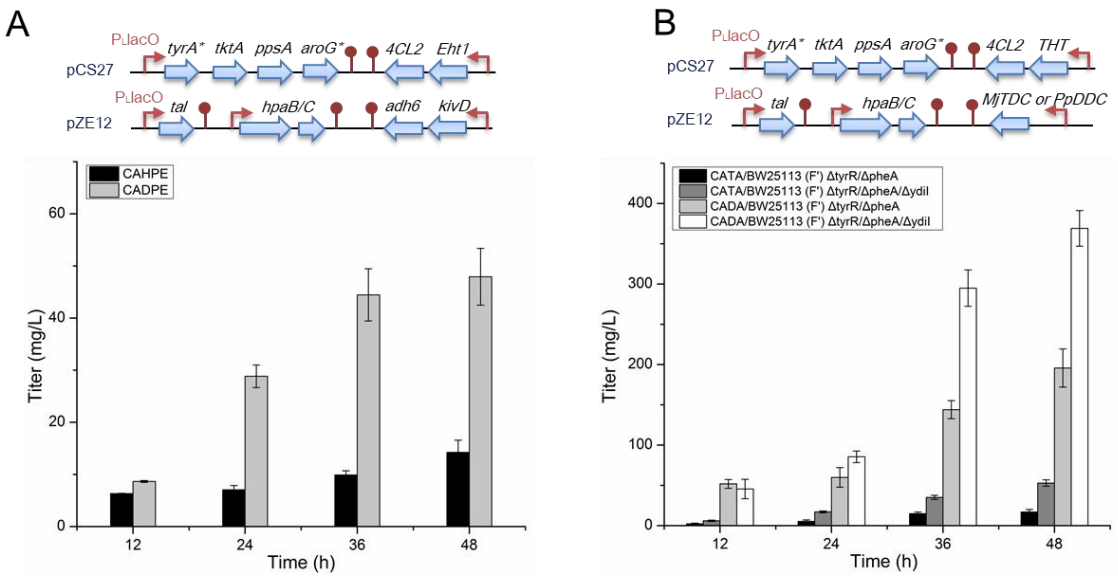


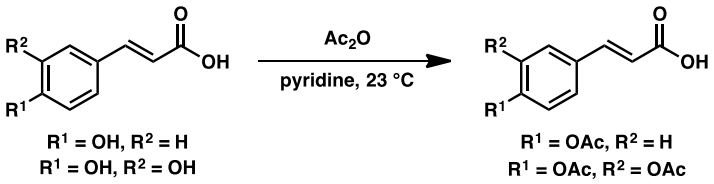

**Figure 2.6 *De novo* production of caffeic acid derived phenethyl ester and amide analogs.** (A) Production of CAHPE and CADPE in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA*/Δ*ydiI*. (B) Production of CATA and CADA in *E. coli* BW25113 (F') Δ*tyrR*/Δ*pheA* and Δ*tyrR*/Δ*pheA*/Δ*ydiI*.

## 2.7 Supplementary information

### 2.7.1 Supplementary Note

**Synthesis of caffeic acid and *p*-coumaric acid derived phenethyl esters and amides.** Esters and amides were synthesized according to standard protocols of carbodiimide couplings. If phenolic groups were present, protection as acetates was carried out prior to coupling, and deacetylations were performed afterward. Reactions were performed under an argon atmosphere unless otherwise noted. Dry acetonitrile (distilled from $CaH_2$) and $CH_2Cl_2$ (passed through activated alumina) were used. Anhydrous HCl gas was prepared by adding concentrated HCl to anhydrous $CaCl_2$. Qualitative TLC and preparatory plate chromatography was performed on 250 μm thick, 60 Å, glass backed, F254 silica (Silicycle, Quebec City, Canada). Visualization was accomplished with UV light and exposure to $KMnO_4$ stain solution followed by heating. Flash column chromatography was performed using Silicylce silica gel (230-400 mesh). $^1H$ NMR spectra were acquired on a Varian Mercury Plus (at 400 MHz) and are reported relative to $SiMe_4$ (δ 0.00).

**(1) Phenol Acetylation**

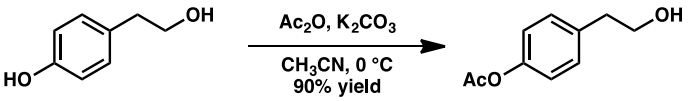



$R^1$ = OH, $R^2$ = H
$R^1$ = OH, $R^2$ = OH

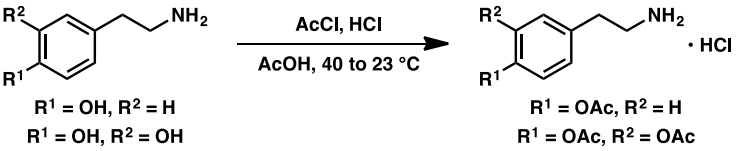
$R^1$ = OAc, $R^2$ = H
$R^1$ = OAc, $R^2$ = OAc

*Caffeic acid and p-coumaric acid*: The phenolic groups on caffeic acid and *p*-coumaric acid were acetylated using acetic anhydride in pyridine [241]. The protection of caffeic acid is described as an example procedure. To a solution of caffeic acid (251 mg, 1.39 mmol) in 12.5 mL of pyridine at 23 °C was added acetic anhydride (0.500 mL, 5.29 mmol), and the reaction was stirred at room temperature overnight. Upon completion as determined by TLC, the pyridine was removed by

rotary evaporation (using heptane for azeotropic removal), and the resulting carboxylic acid product was sufficiently pure to be carried directly to the subsequent reaction (334 mg, 91% yield, $R_f$ = 0.35 in 1:2 hexanes/EtOAc).



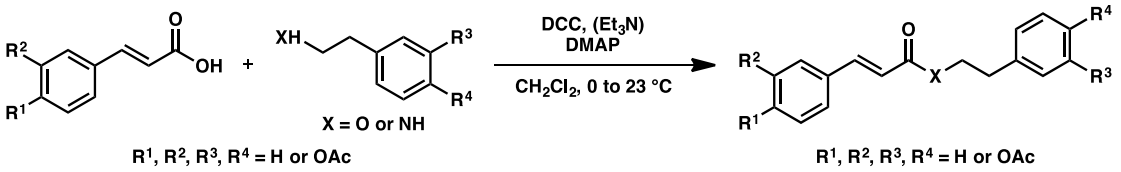

*Tyrosol*: The procedure of Banfi et al. was followed[242]. A solution of tyrosol (500 mg, 3.62 mmol) and $K_2CO_3$ (498 mg, 3.66 mmol) in dry acetonitrile (20 mL) was stirred at room temperature for 15 minutes. The solution was then cooled to 0 °C, and then acetic anhydride (0.700 mL, 7.41 mmol) was added dropwise, and the reaction mixture was stirred at 0 °C. Upon completion (5 h), the reaction mixture was diluted with water (20 mL) and extracted with ethyl acetate (3 x 30 mL). The organic layers were combined, washed with brine (30 mL), dried with $Na_2SO_4$, and concentrated in vacuo. The crude residue was purified by column chromatography (1:1 petroleum ether/EtOAc) to afford the alcohol (584 mg, 90% yield, $R_f$ = 0.57 in 1:2 hexanes/EtOAc) as a pale yellow oil.




*Tyramine and dopamine hydrochloride*: Phenolic amines were protected on the hydroxyl groups as acetates using AcCl and HCl gas in AcOH [243]. The protection of dopamine hydrochloride is described as an example procedure. To a solution of dopamine hydrochloride (1.01 g, 5.30 mmol) in AcOH (12 mL) at 40 °C was bubbled HCl gas (generated via addition of conc. HCl to

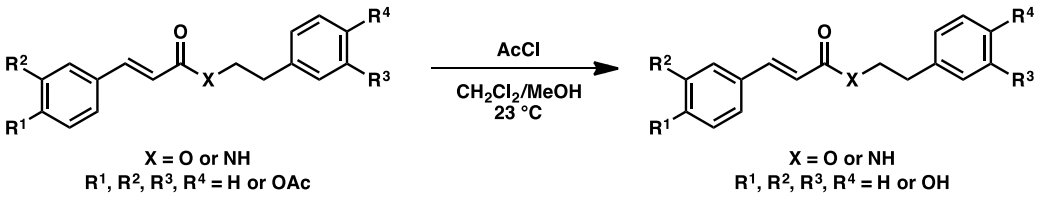
$CaCl_2$) for approx. 5 minutes. AcCl (7.00 mL, 98.1 mmol) was then added dropwise, and the solution was allowed to cool to room temperature and stirred overnight. The next morning, 30 mL $Et_2O$ was added to the reaction mixture, and the precipitated solid was collected by filtration (768 mg, 55% yield). The resulting amine hydrochloride salt as a white solid was carried directly to the subsequent reaction without further purification.

**(2) Synthesis of Amides and Esters**




The amides and esters were synthesized via carbodiimide coupling. One equivalent base ($Et_3N$) was used in the instances where an amine salt was used instead of an amine. The coupling of acetylated caffeic acid and acetylated dopamine hydrochloride is described as an example. To a stirred solution of diacetylated dopamine hydrochloride (93.4 mg, 0.341 mmol) and $Et_3N$ (0.0500 mL, 0.341 mmol) in dry $CH_2Cl_2$ (15 mL) at 0 °C was added diacetylated caffeic acid (90.0 mg, 0.341 mmol) and DMAP (4.5 mg, 0.0341 mmol) followed by DCC (77.3 mg, 0.375 mmol). The solution was stirred at 0 °C for 1 h and then allowed to warm to room temperature and stirred for 4 h. Upon completion, the reaction mixture was cooled to -10 °C and filtered through a plug of celite with cold $CH_2Cl_2$ to remove the urea side product. The filtrate was washed with aq. saturated $NaHCO_3$ (30 mL) followed by brine (30 mL). The organic phase was dried over $Na_2SO_4$ and concentrated in vacuo. The crude residue was purified by column chromatography (1:2 petroleum ether/EtOAc) to afford the corresponding amide (65 mg, 40% yield, $R_f$ = 0.22 in 1:2 hexanes/EtOAc) as a white solid.

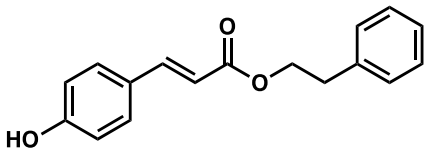
**(3) Deacetylation**




Phenolic acetates were deprotected using AcCl in $CH_2Cl_2$/MeOH. The deprotection of (*E*)-4-(3-((3,4-diacetoxyphenethyl) amino)-3-oxoprop-1-en-1-yl)-1,2-phenylene diacetate is described as an example procedure. To a solution of the tetra-acetylated amide (65.0 mg, 0.134 mmol) in 1:1 $CH_2Cl_2$/MeOH (6.00 mL) at 23 °C was added acetyl chloride (10.0 μl, 0.140 mmol) dropwise. The resulting mixture was stirred overnight at room temperature, and then neutralized with solid $NaHCO_3$ (118 mg, 1.40 mmol) following completion. The reaction mixture was concentrated in vacuo, and the residue was diluted with EtOAc and passed through a short plug of silica. The filtrate was concentrated in vacuo, and the resulting residue was purified by column chromatography to afford the product amide (39 mg, 95% yield, $R_f$ = 0.14 in 1:5 hexanes/EtOAc) as a white solid.

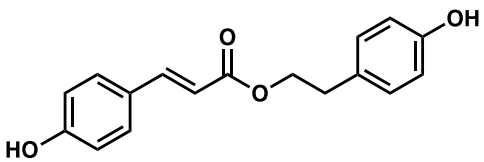
**(4) Characterization of synthesized compounds by NMR analysis**

(a) *p*-Coumaric acid phenethyl ester



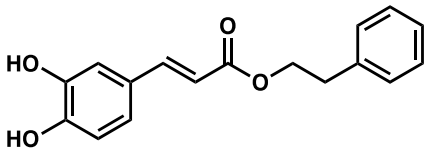
The spectroscopic data for this compound corresponded to previously reported data [244].

(b) *p*-Coumaric acid (4-hydroxyphenethyl) ester



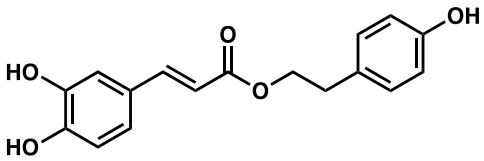
The spectroscopic data for this compound corresponded to previously reported data [245].

(c) Caffeic acid phenethyl ester (CAPE)



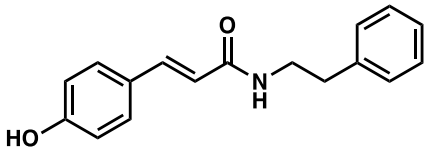

$^1$H NMR (400 MHz, $CDCl_3$): δ 7.57 (d, *J* = 15.9 Hz, 1H), 7.35-7.26 (m, 5H), 7.09 (s, 1H), 7.02 (d, *J* = 8.5 Hz, 1H), 6.88 (d, *J* = 8.5 Hz, 1H), 6.26 (d, *J* = 15.9 Hz, 1H), 5.64 (br s, 2H), 4.42 (t, *J* = 7.0 Hz, 2H), 3.02 (t, *J* = 7.0 Hz, 2H).

(d) Caffeic acid (4-hydroxyphenethyl) ester (CAHPE)



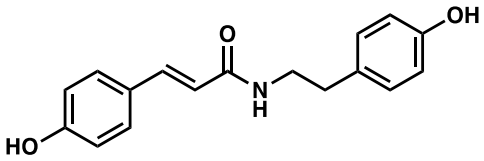

The spectroscopic data for this compound corresponded to previously reported data [246].

(e) *p*-Coumaroyl phenylethylamine



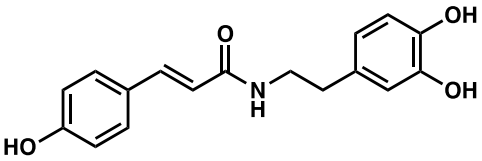

The spectroscopic data for this compound corresponded to previously reported data [247].

(f) *p*-Coumaroyl tyramine



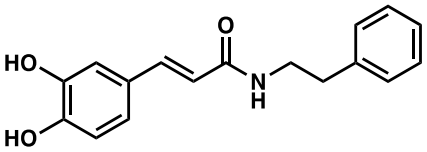

The spectroscopic data for this compound corresponded to previously reported data [246].

(g) *p*-Coumaroyl dopamine



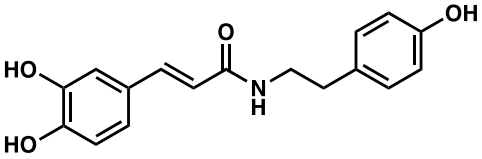

$^{1}$H NMR (400 MHz, DMSO-$d_6$): δ 9.38 (s, 1H (OH)), 8.77 (s, 1H (OH)), 8.66 (s, 1H (OH)), 8.01 (t, *J* = 5.2 Hz, 1H), 7.37 (d, *J* = 8.6 Hz, 2H), 7.30 (d, *J* = 15.8 Hz, 1H), 6.78 (d, *J* = 8.6 Hz, 2H), 6.63 (d, *J* = 8.0 Hz, 1H), 6.58 (d, *J* = 2.0 Hz, 1H), 6.45 (dd, *J* = 8.0, 2.0 Hz, 1H), 6.38 (d, *J* = 15.8 Hz, 1H), 3.30-3.26 (m, 2H), 2.54 (t, *J* = 7.2 Hz, 2H).

(h) Caffeic acid phenethyl amide (CAPA)



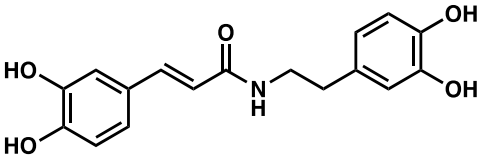

The spectroscopic data for this compound corresponded to previously reported data [246].

(i) Caffeoyl tyramine (CATA)




The spectroscopic data for this compound corresponded to previously reported data [248].

(j) Caffeoyl dopamine (CADA)




$^{1}$H NMR (400 MHz, DMSO-$d_6$): δ 9.37 (br s, 1H (OH)), 9.13 (br s, 1H, (OH)), 8.76 (br s, 1H, (OH)), 8.68 (br s, 1H (OH)), 8.02 (t, *J* = 6.1 Hz, 1H), 7.21 (d, *J* = 15.6 Hz, 1H), 6.93 (d, *J* = 2.2 Hz, 1H), 6.82 (dd, *J* = 8.3, 2.2 Hz, 1H), 6.73 (d, *J* = 8.3 Hz, 1H), 6.63 (d, *J* = 8.0 Hz, 1H), 6.59

(d, *J* = 2.2 Hz, 1H), 6.44 (dd, *J* = 8.0, 2.2 Hz, 1H), 6.30 (d, *J* = 15.6 Hz, 1H), 3.29-3.25 (m, 2H), 2.56 (t, *J* = 7.2 Hz, 2H).

### 2.7.2 Supplementary tables and figures

**Table S2.1 Strains and plasmids used in this study.**

| **Strains** | **Properties** | **Source** |
|---|---|---|
| *E. coli* XL1-Blue | *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 *supE*44 *relA*1 *lac F'* [*traD36 proAB lacI*$^q$*Z*Δ*M*15 *Tn*10 (*Tet*$^r$)] | Stratagene |
| *E. coli* BL21 Star (DE3) | F$^-$ *ompT hsdSB (rB-mB-) gal dcm rne131* (DE3) | Invitrogen |
| *E. coli* BW25113 (F') | *rrnBT*14 Δ*lacZWJ*16 *hsdR*514 *ΔaraBADAH*33 Δ*rhaBADLD*78 *F'* [*traD36 proAB lacI*$^q$*ZΔM15 Tn10(Tet*$^r$*)*] | [249] |
| *E. coli* BW25113 (F') *ΔtyrR* | *E. coli* BW25113 (F') with *tyrR* knockout | This study |
| *E. coli* BW25113 (F') *ΔydiI* | *E. coli* BW25113 (F') with *ydiI* knockout | This study |
| *E. coli* BW25113 (F') *ΔtyrR/ΔpheA* | *E. coli* BW25113 (F') with *tyrR* and *pheA* knockout | This study |
| *E. coli* BW25113 (F') *ΔtyrR/ΔpheA/ΔydiI* | *E. coli* BW25113 (F') with *tyrR*, *pheA* and *ydiI* knockout | This study |
| **Plasmids** | **Properties** | **Source** |
| pZE12-luc | $P_L$lacO1, *colE* ori, Amp$^r$ | [250] |
| pCS27 | $P_L$lacO1, *P*15*A* ori, Kan$^r$ | [251] |
| pSA74 | $P_L$lacO1, *pSC*101 ori, Cm$^r$ | [252] |

| | | |
|---|---|---|
| pETDuet-1 | T7 promoter, *pBR322* ori, $Amp^r$ | Novagen |
| pETDuet-*4CL2* | pETDuet-1 harboring *4CL2* from *Arabidopsis thaliana* | This study |
| pETDuet-*Eht1* | pETDuet-1 harboring *Eht1* from *Saccharomyces cerevisiae* S288C | This study |
| pZE-*Atf1-4CL2* | pZE12-luc harboring *Atf1* from *S. cerevisiae* and *4CL2* in one operon | This study |
| pZE-*Atf2-4CL2* | pZE12-luc harboring *Atf2* from *S. cerevisiae* and *4CL2* in one operon | This study |
| pZE-*Eeb1-4CL2* | pZE12-luc harboring *Eeb1* from *S. cerevisiae* and *4CL2* in one operon | This study |
| pZE-*Eht1-4CL2* | pZE12-luc harboring *Eht1* from *S. cerevisiae* and *4CL2* in one operon | This study |
| pZE-*CAT-4CL2* | pZE12-luc harboring *CAT* coding chloramphenicol acetyltransferase from pSA74 and *4CL2* in one operon | This study |
| pZE-*HCT-4CL2* | pZE12-luc harboring codon-optimized *HCT* from *Nicotiana tabacum* and *4CL2* in one operon | This study |
| pZE-*THT-4CL2* | pZE12-luc harboring codon-optimized *THT* from *Capsicum annuum* and *4CL2* in one operon | This study |
| pZE-*TH* | pZE12-luc harboring the codon-optimized | [208] |

| | | |
|---|---|---|
| | *tal* from *Rhodotorula glutinis* and *hpaBC* from *E. coli* BL21 Star (DE3) in separate operons | |
| pZE-*KA* | pZE-*TH* harboring *kivD* from *Lactococcus lactis* and *adh6* from *S. cerevisiae* in one operon | This study |
| pCS-*LbPDC* | pCS27 harboring *LbPDC* from *Lactobacillus brevis* | This study |
| pCS-*MjTDC* | pCS27 harboring *MjTDC* from *Methanocaldococcus jannaschii* | This study |
| pCS-*PpDDC* | pCS27 harboring *PpDDC* from *Pseudomonas putida* | This study |
| pZE-*TH-KA* | pZE-*TH* harboring *kivD* from *L. lactis* and *adh6* from *S. cerevisiae* in the other operon | This study |
| pZE-*TH-LbPDC* | pZE-*TH* harboring *LbPDC* from *L. brevis* | This study |
| pZE-*TH-MjTDC* | pZE-*TH* harboring *MjTDC* from *M. jannaschii* | This study |
| pZE-*TH-PpDDC* | pZE-*TH* harboring *PpDDC* from *P. putida* | This study |
| pCS-*TPTA* | pCS27 harboring *tyrA**, *pps*, *tktA* and *aroG** from *E. coli* | [208] |
| pCS-*TPTA-EC* | pCS-*TPTA* harboring *Eht1* from *S. cerevisiae* and *4CL2* from *A. thaliana* in the other operon | This study |

| | | |
|---|---|---|
| pCS-*TPTA-TC* | pCS-*TPTA* harboring the codon-optimized *THT* from *C. annuum* and *4CL2* from *A. thaliana* in the other operon | This study |
| pCS-P*lpp1.0-egfp* | pCS27 harboring *egfp* under control of P*lpp1.0* | This study |
| pCS-P*lpp2.0-egfp* | pCS27 harboring *egfp* under control of P*lpp2.0* | This study |
| pSA-P*lpp0.03-pheA** | pSA74 harboring *pheA** (coding a truncated feedback inhibition resistant PheA (1-303 aa) with two point mutants Thr259Ser and Tyr230Leu) under control of P*lpp0.03*, pheA* was amplified from previously constructed pCS-*PPTA* [253] | This study |
| pSA-P*lpp0.2-pheA** | pSA74 harboring *pheA** under control of P*lpp0.2* | This study |

**Table S2.2 Production of caffeic acid derived phenethyl esters and amides in feeding experiments.**

| Plasmids | CAPE[a] | CAHPE | CAPA | CATA | CADA |
|---|---|---|---|---|---|
| pZE-*Atf1-4CL2* | -[b] | - | - | - | - |
| pZE-*Atf2-4CL2* | - | - | - | - | - |
| pZE-*Eeb1-4CL2* | - | - | - | - | - |
| pZE-*Eht1-4CL2* | 52.9±0.3 | 35.0±5.0 | - | - | - |
| pZE-*CAT-4CL2* | - | - | - | - | - |
| pZE-*HCT-4CL2* | - | - | - | - | - |
| pZE-*THT-4CL2* | - | - | 117.5±7.6 | 132.8±2.9 | 79.9±8.5 |

[a], titer (mg/L) of each product when feeding corresponding precursors (500 mg/L) after 24 h.

[b], short dash denotes no production detected.

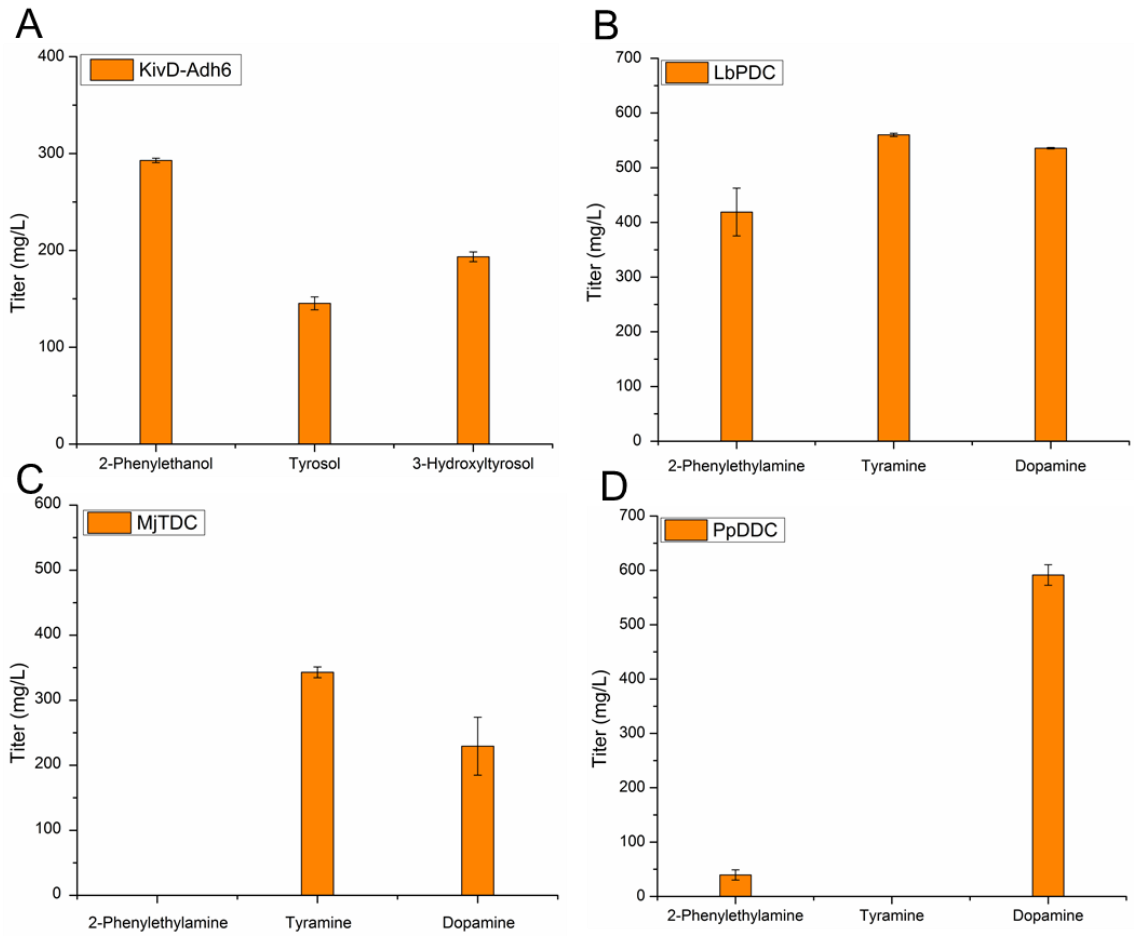



**Figure S2.1 Production of aromatic alcohols via 2-keto acid pathway and aromatic amines via aromatic amino acid decarboxylases.** (A) Phenylpyruvate, 4-hydroxyphenylpyruvate and 3, 4-hydroxyphenylpyruvate (500 mg/L) were fed into *E. coli* BW25113 (F') co-expressing KivD and Adh6. The production of 2-phenylethanol, tyrosol and 3-hydroxyltyrosol were measured at 24 h post substrate feeding and induction by 0.5 mM IPTG. Phenylalanine, tyrosine and L-Dopa (500 mg/L) were fed into *E. coli* BW25113 (F') expressing LbPDC (B), MjTDC (C) and PpDDC (D) for production of 2-phenylethylamine, tyramine and dopamine. The production was measured at 24 h post substrate feeding and induction by 0.5 mM IPTG.

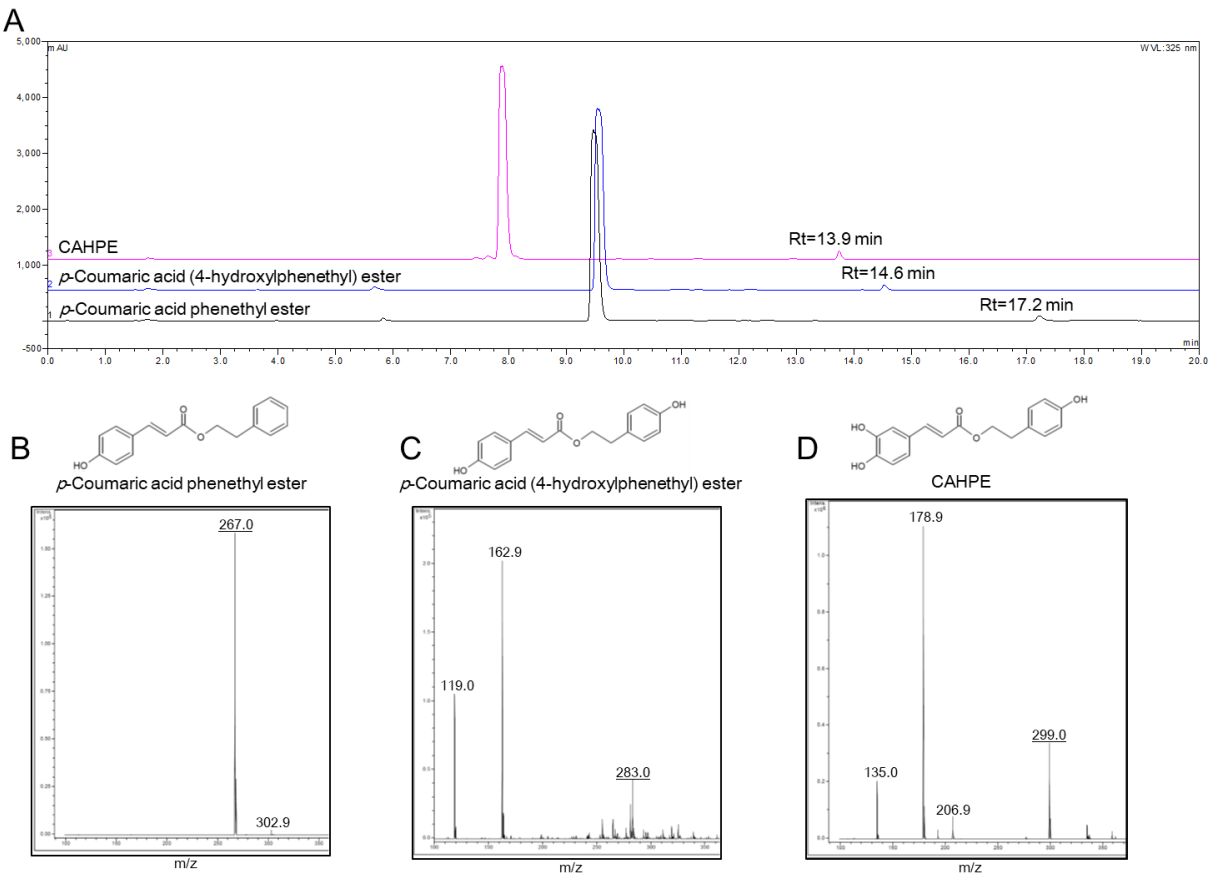
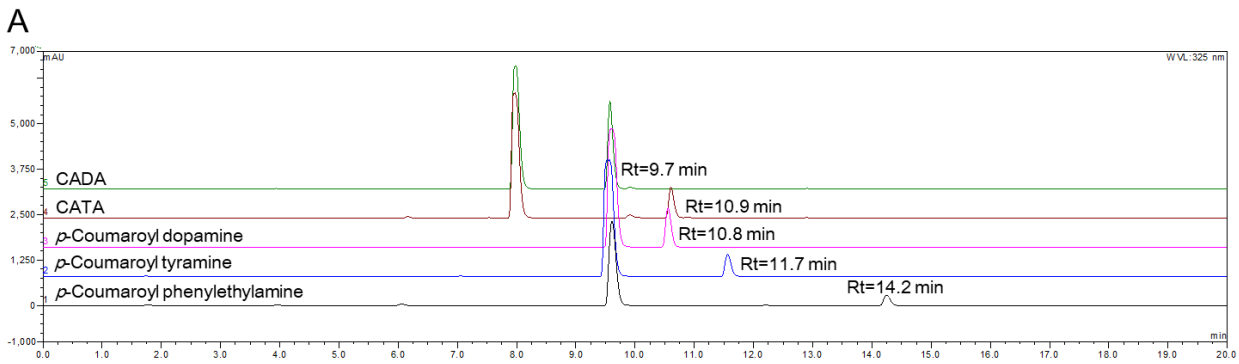



**Figure S2.2 HPLC and MS analysis of *p*-coumaric acid and caffeic acid derived esters.** (A) HPLC analysis of *p*-coumaric acid phenethyl ester (channel 1), *p*-coumaric acid (4-hydroxyphenethyl) ester (channel 2) and CAHPE (channel 3) when feeding 500 mg/L respective precursors (*p*-coumaric acid, caffeic acid, 2-phenylethanol or tyrosol) into *E. coli* BW25113 (F') co-expressing Eht1 and 4CL2 after 24 h. MS analysis of *p*-coumaric acid phenethyl ester (B), *p*-coumaric acid (4-hydroxyphenethyl) ester (C) and CAHPE (D). The underlined mass peaks corresponded to the molecular ion $[M-H]^-$ of each target compound.


A
mAU
WVL:325 nm
CADA
CATA
p-Coumaroyl dopamine
p-Coumaroyl tyramine
p-Coumaroyl phenylethylamine
Rt=9.7 min
Rt=10.9 min
Rt=10.8 min
Rt=11.7 min
Rt=14.2 min
min

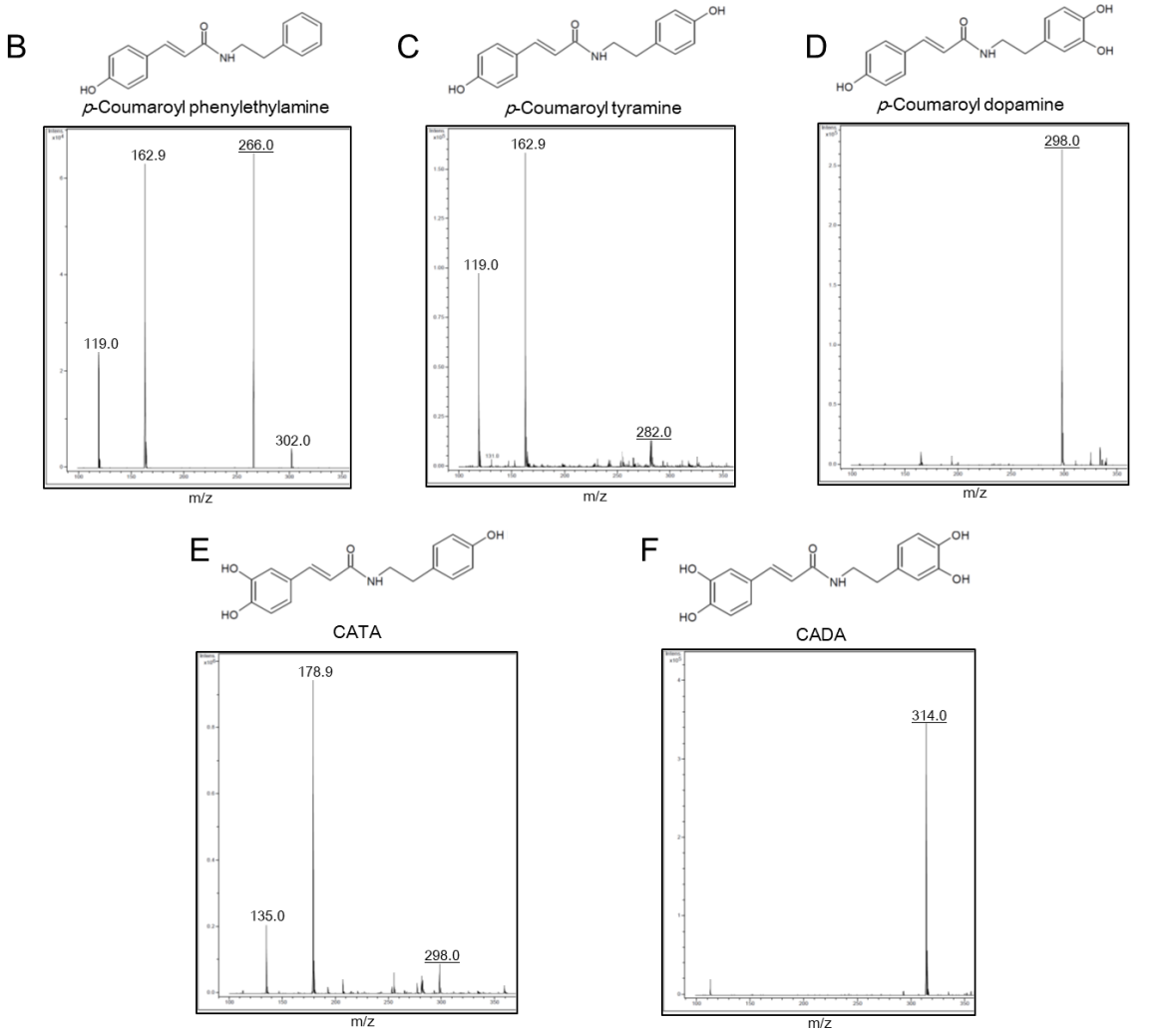

B
p-Coumaroyl phenylethylamine
162.9
266.0
119.0
302.0
m/z
C
p-Coumaroyl tyramine
162.9
119.0
282.0
m/z
D
p-Coumaroyl dopamine
298.0
m/z
E
CATA
178.9
135.0
298.0
m/z
F
CADA
314.0
m/z

**Figure S2.3 HPLC and MS analysis of *p*-coumaric acid and caffeic acid derived amides.** (A) HPLC analysis of *p*-coumaroyl phenylethylamine (channel 1), *p*-coumaroyl tyramine (channel 2), *p*-coumaroyl dopamine (channel 3), CATA (channel 4) and CADA (channel 5) when feeding 500 mg/L respective precursors (*p*-coumaric acid, caffeic acid, 2-phenylethylamine, tyramine or dopamine) into *E. coli* BW25113 (F') co-expressing THT and 4CL2 after 24 h. MS analysis of *p*-coumaroyl phenylethylamine (B), *p*-coumaroyl tyramine (C), *p*-coumaroyl dopamine (D), CATA (E) and CADA (F). The underlined mass peaks corresponded to the molecular ion $[M-H]^-$ of each target compound.

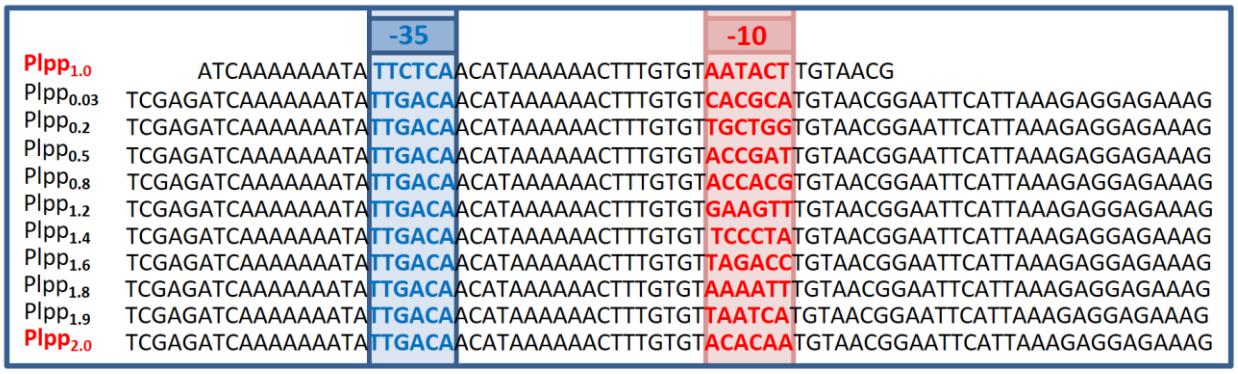



**Figure S2.4 Sequence alignment of P*lpp* promoter library.** The -35 box is indicated in blue and the -10 box is indicated in red.

Chapter 3

# DEVELOPING A PYRUVATE-DRIVEN METABOLIC SCENARIO FOR GROWTH-COUPLED MICROBIAL PRODUCTION[1]

[1]Jian Wang, Ruihua Zhang, Yan Zhang, Yaping Yang, Yuheng Lin, Yajun Yan. *Metabolic Engineering*, 2019. 55, 191-200.

## 3.1 Abstract

Microbial-based chemical synthesis serves as a promising approach for sustainable production of industrially important products. However, limited production performance caused by metabolic burden or genetic variations poses one of the major challenges in achieving an economically viable biomanufacturing process. To address this issue, one superior strategy is to couple the product synthesis with cellular growth, which renders production obligatory for cell survival. Here we create a pyruvate-driven metabolic scenario in engineered *Escherichia coli* for growth-coupled bioproduction, with which we demonstrate its application in boosting production of anthranilate and its derivatives. Deletion of a minimal set of endogenous pyruvate-releasing pathways engenders anthranilate synthesis as the salvage route for pyruvate generation to support cell growth, concomitant with simultaneous anthranilate production. Further introduction of native and non-native downstream pathways affords production enhancement of two anthranilate-derived high-value products including L-tryptophan and *cis, cis*-muconic acid from different carbon sources. The work reported here presents a new growth-coupled strategy with demonstrated feasibility for promoting microbial production.

## 3.2 Introduction

As a green alternative to petrol-based chemical synthesis or natural extraction, engineering microbial workhorses to produce natural products or industrially important commodity chemicals from cheap and renewable biomass is more economically and environmentally appealing [4,254-257]. Tailoring microbial hosts with natural or synthetic pathways, together with genetic optimization of endogenous metabolic networks, are conventional strategies that can readily confer microbes with novel or improved production capabilities [258-261]. However, the severe competition or conflict between cell growth and bioproduction may either lead to compromised cell fitness or loss-of-function phenotypes, ascribed to the adverse metabolic burden from intermediate or product toxicities and/or genetic variations [262-264].

To address these issues, several advanced strategies have been devised, including the sensor-regulator based dynamic regulation systems [265-267], the synthetic intermediate or product addiction systems [268,269], NADH or ATP driven systems [270,271], and quorum-sensing circuit based metabolic control systems [97,272]. The underlying principle of different strategies is to create a metabolic scenario whereby bioproduction process is adapted to cellular growth, achieving the balance between biomass formation and product synthesis.

To render a mandatory drive for bioproduction, one potentially superior strategy is to achieve growth-coupled bioproduction [58,273]. Generally, deletion of essential genes in microbial host forces rewiring of carbon flux and necessitates alternative salvage pathways [274,275]. When the target synthetic pathway serves as the sole compensation route, product synthesis will be obligatorily coupled to cellular growth and thus an integral part of the microbial metabolism. To enable growth-coupled production, one typical approach is to create auxotroph that can be rescued by synthetic pathways. For instance, introducing the citramalate pathway into an

isoleucine auxotrophic *Escherichia coli* enabled growth-based evolution of citramalate synthase and thus increased production of 1-propanol and 1-butanol [120]. An alternative coupling approach is to hijack enzymatic reactions in central carbon metabolism by enforcing the synthetic pathways for complementation; in this respect, the tricarboxylic acid (TCA) cycle was most frequently harnessed to drive bioproduction. For example, disruption of 2-ketoglutarate (2-KG) dehydrogenase gene (*sucA*) forced to the carbon flux to go through the deacetoxycephalosporin C synthase catalyzed reaction for 2-KG to succinate conversion, concomitant with obligatory conversion of penicillin G to 7-aminodeacetoxycephalosporanic acid [276]. Similarly, coupling proline 4-hydroxylase mediated conversion of 2-KG to succinate with an incomplete TCA cycle in *Corynebacterium glutamicum*, led to a 60 % increase in 4-hydroxy-L-proline production with no obvious impact on the cellular growth rate [277]. Interruption of the succinyl-CoA synthetase (*sucCD*) in *C. glutamicum* could enforce the lysine pathway as a compensation route for conversion of succinyl-CoA to succinate, via which the lysine yield was increased by 60 % [278]. With growth coupling, product synthesis is indispensably coined into the host metabolism, whose loss of function will always lead to growth retardation or disadvantage. Additionally, it is amenable to boost the production capabilities of a growth-coupled strain through adaptive laboratory evolution and via growth-based selection [58,279]. Despite its exceptional advantage and feasibility, growth-coupled strategy has been less exploited in practical metabolic engineering applications. This is probably due to the paucity of connecting reactions that can couple synthetic pathways with host cell metabolisms, or the existence of multiple known or unknown native metabolic repair pathways that may bypass or compromise the synthetic pathways [274,275].

In this work, we sought to establish a pyruvate-based driving force for microbial production. Pyruvate is a crucial central metabolite linking glycolysis and TCA cycle. Carbon dissimilation

via microbial central carbon metabolism always leads to pyruvate formation, whose further oxidation via TCA cycle releases energy and reducing power for microbial growth. Due to its essential role for cell survival in broad genetic contexts, we aim to create a pyruvate-driven metabolic scenario for growth-coupled bioproduction purposes. Rewriting the carbon metabolism and deleting endogenous pyruvate-releasing pathways in *E. coli* could enforce pyruvate-forming synthetic pathways to regenerate pyruvate for cell growth, achieving growth-coupled production of target chemicals. As proof-of-concept demonstration of our design, we chose anthranilate (AA) as our target as its biosynthesis involves pyruvate release; and more importantly, AA serves as a multi-functional platform chemical for either biosynthesis of complex value-added natural products or degradation to important commodity chemicals [211,280-283]. With anthranilate overproduction, further introduction of downstream biosynthetic pathways diverted anthranilate to two derivatives including L-tryptophan (L-Trp) and *cis, cis*-muconic acid (MA). Our work presented a platform for the rational rewriting of carbon metabolism to enable growth-coupled chemical production with superior capability.

## 3.3 Materials and methods

### 3.3.1 Bacterial strains, plasmids and chemicals

Bacterial strains and plasmids used in this study were listed in **Table 1**. *E. coli* XL1-Blue (Stratagene) was used for plasmid construction. *E. coli* ATCC 31884 [284] and QH4 [208] strains were starting strains for development of production strains. *E. coli* BW25113 (F′) was used for bioconversion experiments. Plasmids pZE12-luc [250], pCS27 [251], pSA74 [52] were used for gene expression, and pCP20 [285] was used for elimination of kanamycin resistance marker during gene disruption. Phusion DNA polymerase, restriction endonucleases and quick ligase were purchased

from NEB (New England Biolabs). Standard chemicals including anthranilate, L-Trp, catechol, and *cis, cis*-muconic acid were purchased from Sigma-Aldrich unless otherwise specified.

### 3.3.2 Plasmid construction

All manipulations of DNA were conducted referring the standard molecular cloning protocols [286]. The feedback-inhibition resistant mutant of *serA* from *E. coli* BW25113 (F′) with H344A/N364A mutation (*serA*$^{*}$) was created by overlapping PCR using primers containing mutations, and was cloned into the medium-copy-number plasmid pCS27 between *Acc65*I and *Sal*I sites, yielding pCS-*serA*$^{*}$. The wildtype *prsA* from *E. coli* BW25113 (F′) and its mutants obtained via overlapping PCR, were digested with *Sal*I and *Bam*HI, and constructed into pCS-*serA*$^{*}$ to generate pCS27-*serA*$^{*}$-*prsAwt*, pCS-*serA*$^{*}$-*prsA-A95T* and pCS-*serA*$^{*}$-*prsA-G226V*. The pCS-*trpE*$^{fbr}$*G* was obtained from our previous work [281]. The expression cassette of $P_L$*lacO1*-*trpE*$^{fbr}$*G* was then digested and cloned into pCS-*serA*$^{*}$-*prsA-A95T* between *Spe*I and *Sac*I sites, yielding pCS-*trpE*$^{fbr}$*G-serA*$^{*}$-*prsA-A95T*. The *trpDBCA* operon was placed on the low-copy-number plasmid pSA74, and the resulting plasmid pSA-*trpDBCA* was obtained from our previous work [287]. pCS-*trpE*$^{fbr}$*G-APTA* was obtained by amplifying and inserting the APTA module containing *aroL, ppsA, tktA, aroG*$^{fbr}$ (encoding feedback-inhibition resistant AroG with a D146N mutation) from previously constructed pCS-*APTA* [288] into pCS-*trpE*$^{fbr}$*G* between *Spe*I and *Sac*I sites. pZE-*PAPC* was obtained from previous work [281].

### 3.3.3 Gene disruption

The disruption of chromosomal genes in *E. coli* was conducted by P1 phage-based transduction [289]. Specifically, P1 lysates for target knockouts were prepared from the Keio collection strains

(CGSC) [290]. The transduced cells were screened on LB-agar plates with appropriate antibiotic and 100 mM sodium citrate. Removal of antibiotic resistance was conducted by electroporation of pCP20 into target strains [291]. Knock-out strains were confirmed by colony PCR.

### 3.3.4 Culture media and conditions

Luria-Bertani (LB) medium containing 10 g/L tryptone, 5 g/L yeast extract and 10 g/L sodium chloride was used for cell propagation. The M9 minimal medium (per liter) containing 6 g $Na_2HPO_4$, 0.5 g NaCl, 3 g $KH_2PO_4$, 1 g $NH_4Cl$, 246.5 mg $MgSO_4 \cdot 7H_2O$ and 14.7 mg $CaCl_2 \cdot 2H_2O$ was used as the basis medium for growth-related tests and shake flask cultivation. Antibiotics were added when necessary at the concentration of 100, 50 and 34 μg/mL for ampicillin (Amp), kanamycin (Kan) and chloramphenicol (Cl), respectively. For growth-related tests, M9 medium was supplemented with 20 g/L glycerol, 100 mg/L L-Phe, L-Tyr and L-Trp of each, and 0.5 or 5 g/L yeast extract when needed. For shake flask cultivation, M9 minimal medium was supplemented with 20 g/L glycerol or xylose and 5 g/L yeast extract (M9Y). Overnight cell culture was inoculated with 2 % into 20 mL M9Y medium in 125 mL flasks and incubated for 3 h at 37 °C and 280 rpm. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at a final concentration of 0.5 mM and cell cultures were switched to 30 °C for induction. For anthranilate to L-Trp bioconversion experiments, anthranilate was added along with IPTG to cell cultures at a final concentration of 2 g/L. Cell cultures were sampled every 12 h. The optical density at 600 nm ($OD_{600}$) of cell culture was measured in 96-well plates in Biotek Synergy HT Microplate Reader (Biotek). Products in the culture were quantified by high-performance liquid chromatography (HPLC) analysis after centrifugation at 13,523 × g for 10 min and filtration with 0.45 μm filter.

### 3.3.5 Growth test conditions and measurement

To investigate the pyruvate-dependent growth, seed cultures were inoculated in 3 mL LB culture in test tube at 37 °C for overnight. 100 μL of seed cultures were collected by centrifugation at 3,381 × g for 2 min, and the cell pellets were resuspended and inoculated into M9 minimal medium containing 20 g/L glycerol, 100 mg/L L-Phe, L-Tyr, and L-Trp of each. To test the effect of pyruvate supplementation, pyruvate was added with final concentrations ranging from 0 to 2 g/L. Cell cultures were sampled every 6 or 12 h for growth measurement ($OD_{600}$) in 96-well plates. The specific growth rate, $\mu$, was calculated from the $OD_{600}$ growth curve at exponential phase, as described previously ($\mu = \Delta \ln OD_{600}/\Delta t$, where $t$ is time) [292,293]. Similar approach was used for investigating the *trpE*-dependent growth of the engineered cells, with additional supplementation of 0.5 g/L yeast extract to support the initial cell growth.

### 3.3.6 HPLC analysis

Both standard chemicals and products in cell cultures including anthranilate, L-Trp, catechol, and *cis, cis*-muconic acid were quantified by a reverse phase HPLC system 1260 infinity II (Agilent technologies) equipped with a ZORBAX SB-C18 column and a 1260 infinity II Diode Array Detector WR. The mobile phase was consisted of 0.1 % trifluoroacetic acid (solvent A) and methanol (solvent B). The flow rate was set at 1 mL/min and temperature at 28 °C. The analyzing method was set as: 5 % solvent B for 2 min, from 5 to 50 % solvent B for 6 min, from 50 to 80 % solvent B at 2 min, from 80 to 5 % solvent B for 5 min, and 5 % solvent B for additional 2 min. The concentrations of anthranilate, L-Trp, catechol, and *cis, cis*-muconic acid were determined by the peak areas at absorbances of 320 nm, 276 nm, 276 nm and 260 nm, respectively. Glycerol and xylose concentrations were quantified by the Dionex HPLC system

with Coregel-64H column (Transgenomic). 4 mN $H_2SO_4$ was used as the mobile phase and the flow rate was set at 0.4 mL/min. The oven temperature was set at 45 °C.

## 3.4 Results

### 3.4.1 Metabolic design for pyruvate-driven growth-coupled bioproduction

To enable pyruvate-driven growth-coupled bioproduction, the centerpiece of our metabolic design is to rewire carbon metabolism and create a metabolic scenario of pyruvate-based coupling of cell growth and target synthetic pathways. Pyruvate is a key metabolite from central carbon metabolism, and can support cell growth alone via TCA cycle mediated oxidation [44]. Besides central carbon metabolism, pyruvate release is also a commonplace in synthetic pathways, especially those involving aromatic biosynthesis or catabolism [28,294,295]. We reason that it is possible to tightly couple cell growth with target synthetic pathways by: (1) deleting endogenous pyruvate-releasing pathways to create a pyruvate auxotrophic *E. coli*, (2) rendering the target biosynthetic pathway as the primary or even the sole pathway to regenerate pyruvate. We began our design by focusing on rewiring glycerol metabolism for growth-coupled AA production in *E. coli* (**Figure 3.1**). Glycerol is chosen as the carbon source due to its increased availability, low price, high degree of reduction, and most importantly, a higher maximal theoretical yield for aromatic compounds [33,296,297]. AA can be natively produced by *E. coli* via the shikimate pathway with one mole of pyruvate released per mole of AA produced, making it an ideal pyruvate-driven target pathway.

To render AA pathway as the main source for pyruvate generation, native pyruvate-releasing pathways in glycerol dissimilation and central carbon metabolism need to be blocked (**Figure 3.1**). Glycerol dissimilation into the central carbon metabolism in *E. coli* involves two pathways,

the major GlpK-GlpD/GlpABC mediated respiratory pathway and the minor GldA-DhaKLM mediated fermentative pathway [297]. To preclude pyruvate leaking via phosphoenolpyruvate (PEP)-dependent dihydroxyacetone kinase (DhaKLM), *gldA*, encoding glycerol dehydrogenase that catalyzes the first committed step of the fermentative pathway, will be deleted. In central carbon metabolism, two major pyruvate-releasing reactions are involved: the pyruvate kinases (PykA and PykF) mediated conversion of PEP to pyruvate, and malate dehydrogenase (MaeB) mediated decarboxylation of malate to pyruvate. In AA biosynthesis pathway, anthranilate synthase (TrpED) is responsible for conversion of chorismite to AA, concomitant with pyruvate release. Deletion of all corresponding genes would possibly make *E. coli* auxotrophic for pyruvate, or retaining *trpED* would make *E. coli* growth dependent on AA pathway. Moreover, deletion of *gldA* and *pykAF* would reserve PEP for shikimate pathway.

### 3.4.2 Engineering pyruvate-driven growth-coupled production of anthranilate

To create and validate an AA pathway dependent *E. coli*, we started knocking out endogenous pyruvate-releasing genes within strain *E. coli* QH4 (**Figure 3.2A**), a derivative of L-phenylalanine (L-Phe) overproducer *E. coli* ATCC31884 with disrupted L-Phe and L-tyrosine (L-Tyr) branches [208]. We first deleted the pyruvate kinase genes *pykA* and *pykF* in *E. coli* QH4, yielding strain JW1 (**Table 3.1**). Pyruvate kinases catalyze the conversion of PEP to pyruvate, and knocking out of *pykA* and *pykF* alleviates pyruvate formation but does not impair cell viability [267,298,299]. Subsequent stepwise deletion of *gldA* and *maeB* yielded two more strains, JW2 and JW3. When cultivated in minimal medium containing glycerol (20 g/L) and limited yeast extract (0.5 g/L), JW2 and JW3 showed 2.4- and 3.2-fold decrease in specific growth rate ($\mu$ = 0.08 and 0.06 $h^{-1}$, respectively) compared with their parental strain QH4 ($\mu$ = 0.19 $h^{-1}$)

(**Figure 3.2B**). When further knocking out anthranilate synthase gene *trpE* in JW2 or JW3, cell growth of the resultant strain JW5 or JW6 was arrested within 36 hrs. Noteworthy, the cell growth of *maeB* deleting strains, JW3 and JW6 ($\mu$ = 0.04 $h^{-1}$), were more severely impaired. In contrast, QH4 with only *trpE* knockout (JW4) exhibited no growth defect ($\mu$ = 0.19 $h^{-1}$) (**Figure 3.2B**), and could grow even with no yeast extract (**Figure 3.2C**). Taken together, these results established that disruption of the major pyruvate-releasing genes including *pykA*, *pykF*, *gldA* and *maeB* significantly compromised cell growth, possibly caused by limited pyruvate supply. Further knockout of *trpE* almost deprived cell of growth potential, as JW5 and JW6 could barely grow in glycerol minimal medium with limited yeast extract, and did not show detectable growth when further removing yeast extract from the medium (**Figure 3.2D, E**). This revealed that the cell growth of JW2 or JW3 was in part reliant on pyruvate released from AA pathway in glycerol minimal medium with limited yeast extract. When cultivated in glycerol minimal medium with surplus of yeast extract (5 g/L), all mutant strains showed improved growth while JW5 and JW6 still showed compromised growth rate and potential (**Figure S3.1**).

To investigate if exogenous pyruvate could restore cell growth of pyruvate auxotrophic *E. coli*, JW5 and JW6 were cultivated in glycerol minimal medium (containing100 mg/L L-Phe, L-Tyr and L-Trp but with no yeast extract) supplemented with different concentrations of pyruvate (0-2 g/L). As expected, the control strain JW4 showed no growth difference in different pyruvate concentrations, while the growth of JW5 and JW6 were restored by pyruvate and their growth potentials were dependent on the amount of pyruvate supplied (**Figure 3.2C, D, E**). This demonstrated that knocking out of *pykA*, *pykF*, *gldA* with or without *maeB* rendered *E. coli* growth pyruvate-dependent. This also implied that these genetic manipulations enabled high correlation of pyruvate supply and cell viability that the high-performers with pyruvate-forming

synthetic pathways will be potentially more advantageous. Of note, exogenous pyruvate addition facilitated glycerol utilization in JW5 and JW6 but not in JW4 (**Figure 3.2F**), as reflected by the difference in their growth potentials with 2 g/L pyruvate (**Figure 3.2C, D, E**).

Given the maintained but reduced cell growth attributed to single chromosomal copy of *trpE*, we reason that plasmid-based over-expression of anthranilate synthase in our engineered *E. coli* strains will promote not only cell growth but also AA production. To verify that, the feedback-inhibition resistant anthranilate synthase ($TrpE^{fbr}G$), composed of TrpE with an S40R mutation and the N-terminal domain of TrpD (designated TrpG) [281], was expressed under control of $P_L lacO1$ promoter on the medium-copy-number plasmid. The resultant plasmid pCS-$trpE^{fbr}G$ was transferred into *E. coli* QH4, JW1, JW2 and JW3, resulting in AA producer strains AA1 to AA4 (**Table 3.1**). All transformed strains were subjected to shake flask cultivation using M9 minimal medium containing 20 g/L glycerol with 5 g/L yeast extract. After 48 h of aerobic cultivation, all recombinant strains (AA2 to AA4) showed slightly decreased growth but increased AA production than the control strain AA1 (**Figure 3.2G**). Especially, AA4 with the most knockouts afforded the highest titer of 1.78 g/L AA, which is 2.1-fold to that produced by AA1 (0.85 g/L). In particular, AA4 achieved an AA yield of 0.143 g/g glycerol, which is 1.86-fold to that of strain AA3 (0.077 g/g glycerol) and 2.65-fold to that of control strain AA1 (0.054 g/g glycerol) (**Table 3.2**). This demonstrated that, with strong growth coupling, the engineered pyruvate-driven mechanism could significantly improve AA production, even without enhancing the upstream shikimate pathway flux.

### 3.4.3 Enhancing L-Trp production via the pyruvate-driven system

With establishment of the pyruvate-driven system for AA production, we next evaluated whether the engineered AA platform would translate into AA-derived bioproduction pathways. Natively, AA serves as the direct precursor for L-Trp, an essential aromatic-group amino acid for humans and a fundamental precursor to a variety of high-value compounds like neurotransmitters 5-hydroxytryptophan (5-HTP) and serotonin, and drugs like indolylglucosinolate and monoterpene indole alkaloids (MIAs) [8,257,300]. L-Trp biosynthesis is subjected to multi-level regulation, including TrpR-mediated transcriptional repression and TnaA mediated L-Trp degradation [301,302]. Therefore, to alleviate L-Trp degradation and also to deregulate TrpR mediated repression, both *tnaA* and *trpR* were knocked out in JW2 and JW3, with subsequent transduction of the F′ plasmid (harboring *lacI*$^{q}$) from *E. coli* XL-Blue to avoid leaky expression and to facilitate transformation. The resultant strains were designated JW7 and JW8 (**Table 3.1**).

To fully exploit the AA platform, we set to address rate-limiting steps during AA conversion to L-Trp. AA undergoes five enzymatic steps to L-Trp, during which supply of two co-substrates is of utmost importance: L-serine (L-Ser) and 5-phospho-α-D-ribose 1-diphosphate (PRPP) [303]. To assess the contribution of L-Ser supply, 2 g/L AA was fed to *E. coli* BW25113 (F′) expressing the feedback-inhibition resistant D-3-phosphoglycerate dehydrogenase SerA$^{*}$ (SerA H344A/N364A) and *trpDBCA* operon (pSA-*trpDBCA*). SerA$^{*}$ over-expression led to production of 1.27 g/L L-Trp in 48 h, a 6.2-fold increase over the control strain with no SerA$^{*}$ over-expression (**Figure 3.3A, Figure S3.2A**). Considering the residual of 0.93 g/L AA, SerA$^{*}$ overexpression resulted in an AA-to-L-Trp conversion ratio to 42.6 % of the theoretical maximum (**Figure S3.2B**). To further avoid AA build up, we sought to increase the PRPP pool by over-expressing the ribose-phosphate diphosphokinase (PrsA). However, co-expression of

wildtype *prsA* (*prsAwt*) with *serA*$^{*}$ and *trpDBCA* decreased L-Trp to 0.58 g/L, suggesting an adverse effect by PRPP surplus (**Figure 3.3A**). We next turned to two PrsA mutants, A95T and G226V, which were chosen because of their evolutionally beneficial role in producing a hemi-autotrophic *E. coli* that can grow on $CO_2$ and pyruvate [44,304]. Surprisingly, although with reduced ribose-phosphate diphosphokinase activities [44], L-Trp production was increased to 1.36 g/L by PrsA G26V (45.7 % of theoretical maximum), while increased to 2.02 g/L by PrsA A95T (68.0 % of theoretical maximum) (**Figure 3.3A, Figure S3.2A**). Meanwhile, with PrsA A95T expressing, the residual AA was significantly decreased to 0.45 g/L at 48 h. These results suggested that, by modulating L-Ser and PRPP pool, AA can be efficiently converted to L-Trp.

With engineered hosts and optimized L-Trp pathway, we then tested the ability of pyruvate-driven system in boosting L-Trp production from glycerol. The complete L-Trp pathway was reassembled on two plasmids, pCS-*trpE*$^{fbr}$*G-serA*$^{*}$*-prsA-A95T* and pSA-*trpDBCA*. They were respectively transferred into QH4, JW7 and JW8, with the resultant strains designated TP1 to TP3 (**Table 3.1**). After 60 h of aerobic cultivation in glycerol minimal medium, all recombinant strains reached peak L-Trp titers in 60 h with no detectable accumulation of AA. The parental strain TP1 could produce 0.73 g/L L-Trp, while the engineered TP2 produced the highest titer of L-Trp of 1.73 g/L, a 2.37-fold to that produced by TP1 (**Figure 3.3B, Figure S3.3**). The L-Trp yield by TP1 reached 0.037 g/g glycerol, whereas TP2 achieved a 3.57-fold increased L-Trp yield (0.132 g/g) (**Table 3.2**). Interestingly, although TP3 produced less L-Trp than TP2, it also consumed less glycerol, resulting in the highest L-Trp yield to 0.168 g/g glycerol. This demonstrated that our pyruvate-driven system could significantly improve both L-Trp titers and yields.

### 3.4.4 Enhancing cis, cis-muconic acid production via the pyruvate-driven system

To test the applicability of our AA platform for non-native bioproduction pathways, we implemented it into an AA derived non-native product MA. MA is microbially produced from renewable biomass as the synthetic precursor for adipic acid and terephthalic acid, which have been widely used as important platform chemicals for synthesis of nylon-6,6, polyurethane and polyethylene terephthalate (PET) [305-307]. Our lab previously established three independent MA synthetic pathways starting from chorismate: salicylic acid (SA) pathway, 2,3-dihydroxybenzoic acid (DHBA) pathway and AA pathway [281,308,309]. The AA-based MA pathway requires three enzymatic steps downstream of chorismate: conversion of chorismate to AA by TrpEG, conversion of AA to catechol by anthranilate 1,2-dioxygenase (ADO) and benzene ring cleavage of catechol to MA via catechol 1,2-dioxygenase (CDO). Although the AA-based MA pathway is shorter and less toxic, its production potential has been significantly limited by the AA supply [281,308].

To harness our AA platform for MA production, the synthetic pathway comprising ADO (*antABC* from *Pseudomonas aeruginosa*) and CDO (*catA* from *P. putida* KT2440), was introduced into AA overproducer strains on a high-copy-number plasmid (pZE-*PAPC*). Co-transformation of pZE-*PAPC* and pCS-*trpE*$^{fbr}$*G* into QH4, JW2 and JW3 yielded MA producer strains, MA1 to MA3. After cultivation in glycerol minimal medium for 72 h, MA2 and MA3 respectively afforded production of 1.82 and 1.60 g/L MA, along with 0.14 and 0.43 g/L residual catechol, while QH4 resulted in only 0.89 g/L MA with almost no accumulation of catechol (**Figure 3.4, Figure S3.4**). Interestingly, AA is barely accumulated in all production strains, suggesting that the downstream CDO is rate-limiting in MA production. To further increase carbon flux to shikimate pathway, we incorporated the APTA module (harboring *aroL*, *ppsA*,

*tktA*, and *aroG*$^{fbr}$) on the pCS-*trpE*$^{fbr}$*G* as an independent operon (pCS-*trpE*$^{fbr}$*G-APTA*). However, when transformed with pZE-*PAPC* and pCS-*trpE*$^{fbr}$*G-APTA* into JW2 and JW3, the resultant strains MA4 and MA5 showed substantially decreased MA production (0.86 and 0.71 g/L, respectively) (**Figure 3.4, Figure S3.5**). This may be explained by increased metabolic burden, possibly caused by excessive protein expression. Taking into account that the engineered pyruvate-driven strains are efficient in driving carbon flux to AA, minimal expression of downstream genes would allow optimal production. Compared with the previous report on AA-based MA production [281], our engineered strain increased MA titer by 4.7-fold with less over-expressed genes, demonstrating the robustness of our pyruvate-driven system for microbial production of AA-derived non-native products.

### 3.4.5 Extending the pyruvate-driven system to xylose metabolism

After demonstrating the application of the pyruvate-driven system in glycerol-based medium, we hypothesize that it can be potentially applied to utilizing other commonly used carbon sources like lignocellulosic sugar glucose and xylose. To extend the flexibility of carbon sources, the key point is to eliminate pyruvate release during carbon assimilation. Since glucose uptake primarily relies on PEP-dependent phosphotransferase system (PTS), to simplify the genetic manipulations, we therefore switch to xylose as proof-of-concept demonstration. Xylose is the second most abundant sugar in lignocellulosic biomass and has been widely used as a potential feedstock for microbial production of chemicals in recent decades [123,310,311]. In *E. coli*, xylose enters the pentose phosphate pathway via an ATP-dependent isomerase pathway with the combined action of xylose isomerase (XylA) and xylulokinase (XylB), leaving no additional pyruvate releasing reactions (**Figure 3.1**).

We first tested the applicability of our final pyruvate-driven strains for AA biosynthesis using xylose. Shake flask experiments were performed with strains AA1, AA3 and AA4 in minimal medium containing 20 g/L xylose. After 72 h of cultivation, the control strain AA1 only produced 0.56 g/L AA, while the engineered strain AA3 and AA4 produced 1.38 g/L and 1.90 g/L AA, respectively (**Figure 3.5, Figure S3.6**). Particularly, AA4 achieved the highest yield of 0.178 g/g xylose, which is 1.36-fold to that of AA3 (0.131 g/g) and 3.87-fold to that of AA1 (0.046 g/g) (**Table 3.2**). The production trend was in parallel with that observed with glycerol, indicating the applicability of the pyruvate-driven system for xylose-based chemical production.

Subsequently, the pyruvate-driven system was deployed for L-Trp and MA biosynthesis. When cultivating TP1, TP2 and TP3 in xylose minimal medium, TP2 afforded the highest L-Trp production with a peak titer of 1.31 g/L at 36 h, which is 1.35-fold and 1.6-fold increase to that produced by TP3 (0.97 g/L) and TP1 (0.82 g/L), respectively (**Figure 3.5 and Table 3.2**). The residual of AA (>0.5 g/L) in TP2 and TP3 cultures indicated that the downstream AA-to-L-Trp conversion in xylose-based condition is less efficient than that in glycerol-based condition (**Figure S3.7**). Considering the difference of substrate reluctance and metabolic infrastructure, the choice of carbon source might have different effects on downstream synthetic pathways. This effect is more notable in MA production, which did not show an increase trend in xylose medium as observed in glycerol medium. With the engineered pyruvate-driven system, MA2 and MA3 only afforded 0.46 and 0.47 g/L MA, which are less than that produced by the control strain MA1 (0.68 g/L) (**Figure 3.5 and Table 3.2**). The substantial accumulation of intermediates including AA (0.37-0.97 g/L) and catechol (0.43-0.80 g/L) in MA2 and MA3 cultures indicated that the downstream MA pathway is inefficient in conversion of AA to MA in xylose-based condition (**Figure S3.8**). Although the MA concentrations from xylose were lower than from

glycerol, the high accumulation of metabolites including AA and catechol still demonstrated the functionality of the pyruvate-driven system in driving carbon flux from xylose to MA pathway.

### 3.5 Discussion

Improving the microbial production performance is a long-term pursue in metabolic engineering fields. Efforts to repurpose microbes for chemical production have been largely focused on maximizing heterologous synthetic pathways and minimizing endogenous competing pathways [7,312]. The heavy metabolic burden, extensive genetic manipulations, metabolite toxicities or sophisticated metabolic regulations might compromise cell viability and production performance. To facilitate bioproduction, engineering inherent driving forces for metabolic engineering applications have been devised and employed mainly based upon cofactor-driven systems, like NADH-driven system in *E. coli* and ATP-driven system in cyanobacteria [270,271]. However, these cofactor-driven systems are conditionally functioning, like NADH-driven system only working in microaerobic or anaerobic conditions and ATP-driven system limited to photosynthetic organisms [313]. Here we report the development of a pyruvate-driven system to couple product synthesis with cellular growth, which enables the bioproduction coined as an integral part of host metabolism. Via growth coupling, product synthesis will be positively correlated with cell fitness, promoting high-performance variants to dominate the population and provoking adaptive evolution of cell robustness to lessen metabolic burden. Noteworthy, due to the crucial role of pyruvate in microbial carbon metabolism, the pyruvate-driven system can be readily adapted to any production microorganisms with no constraint of cultivation conditions. Yet, to fully deploy the pyruvate-driven system for microbial production, two key prerequisites

need to be fulfilled, removing major endogenous pyruvate-releasing reactions in carbon metabolism and anchoring pyruvate-forming synthetic pathways.

In *E. coli*, although there exist more than thirty genes that are involved in pyruvate formation, knocking out several major pyruvate-releasing genes can render *E. coli* pyruvate auxotrophic or dependent on specific pyruvate-forming pathways. When using glycerol as the carbon source, knocking out of a minimal set of only four genes including *pykA*, *pykF*, *gldA* and *maeB* rendered cellular growth dependent on anthranilate biosynthetic pathway, since further anthranilate synthase gene (*trpE*) knockout would endow *E. coli* pyruvate auxotrophic. Over-expression of feedback-inhibition resistant anthranilate synthase ($TrpE^{fbr}G$) in the engineered platform strain JW3 (QH4 Δ*gldA* Δ*pykA* Δ*pykF* Δ*maeB*) showed slightly decreased growth but 2.1-fold increase of AA production compared to the parental strain QH4, indicating the coupled effects of growth and bioproduction. The pyruvate-driven system can be theoretically applied to all commonly used carbon sources that are catabolized to pyruvate including lignocellulosic biomass derived glucose and xylose. To extend the feedstock flexibility of the pyruvate-driven system, the key principle of the metabolic design is to eliminate pyruvate formation during sugar catabolism. For glucose, replacing the PEP-dependent PTS transport system with galactose permease (GalP) and ATP-dependent glucokinase (Glk) would preclude pyruvate formation as has been demonstrated previously [294,295]; while for xylose, the native ATP-dependent pathway would result in no additional pyruvate release. Thus, as no additional genetic modifications needed, we directly shifted the carbon source to xylose with our engineered platform strain. The reproducible trend of titer and yield improvement from xylose demonstrated the applicability of pyruvate-driven system in different carbon sources.

The pyruvate-forming pathways are quite common in nature. Among those, the shikimate pathway is of remarkable industrial importance because of its involvement in production of a variety of high-value aromatic compounds. The branches of shikimate pathway that start from chorismate or its derivatives involve pyruvate release; for instances, one pyruvate is released during conversion of chorismate to AA or 4-hydroxybenzoate, isochorismate to 2,3-dihydroxy-2,3-dihydrobenzoate or salicylate, and 4-amino-4-deoxychorismate to 4-aminobenzoate [294]. These products could serve as precursors for an expanded spectrum of new products of interest. In the present study, we focused on establishing and demonstrating the pyruvate-driven system for AA and its derivative overproduction. We demonstrated that, with no bottlenecks in AA-derived synthetic pathways, the benefit of pyruvate-driven system was extended to AA-derived products including L-Trp and MA.

One key strength of the pyruvate-driven system is its capacity to improve the practical yield, a feature that can be seen in growth-coupled chemical production. With the advanced pyruvate-driven system, the carbon utilization, as well as product synthesis, were internally coupled to cellular growth. As we observed in the current study, installing the pyruvate-driven system increased the yield of AA, L-Trp and MA from different carbon source. From a practical perspective, the pyruvate-driven system could potentially afford high-yield and feedstock-efficient production by controlling carbon expenditure and flux redirection.

In summary, our research underlines the great potential of setting up pyruvate-driven growth-coupled metabolic design for microbial chemical production. With proof-of-concept demonstration with AA and its derivative pathways using different carbon sources, we conclude that the pyruvate-driven system could be readily achievable in different metabolic or genetic contexts. Current limitation of the pyruvate-driven system may be the low carbon flux toward the

synthetic pathways, which might restrict pyruvate supply for cellular growth or provoke the emergence of alternative metabolic routes for complementation [274]. By increasing the carbon flux through synthetic pathways and adaptive evolution of rationally engineered growth-coupled strains, we expect further increase of cell fitness and improvement of production performance in the future.

### 3.6 Tables and figures

**Table 3.1 Plasmids and bacterial strains used in this study.**

| Plasmid | Description | Source |
|---|---|---|
| pZE12-luc | $P_L$*lacO1*, *colE* ori, $Amp^r$ | [250] |
| pCS27 | $P_L$*lacO1*, *P15A* ori, $Kan^R$ | [251] |
| pSA74 | $P_L$*lacO1*, *pSC101* ori, $Cl^R$ | [52] |
| pCP20 | Flippase, $Amp^R$, and temperature-sensitive replicon | [285] |
| pZE-*PAPC* | pZE12-luc containing *antABC* from *Pseudomonas aeruginosa* (*paantABC*) and *catA* from *P. putida* KT2440 (*ppcatA*) | [281] |
| pCS-*trpE$^{fbr}$G* | pCS27 containing *trpE$^{fbr}$G* (encoding feedback-inhibition resistant TrpE with an S40R mutation and the N-terminal domain of TrpD (designated as TrpG)) from *E. coli* | [281] |
| pCS-*trpE$^{fbr}$G-APTA* | pCS-*trpE$^{fbr}$G* containing *aroL, ppsA, tktA, aroG$^{fbr}$* (encoding feedback-inhibition resistant AroG with a D146N mutation) from *E. coli* | This study |
| pCS-*serA** | pCS27 containing *serA** (encoding feedback-inhibition resistant SerA with H344A/N364A mutation) | This study |
| pCS-*serA*-prsAwt* | pCS27 containing $P_L$*lacO1-serA*-prsA* | This study |
| pCS-*serA*-prsA-* | pCS27 containing $P_L$*lacO1-serA*-prsA-A95T* | This study |

| | | |
|---|---|---|
| *A95T* | | |
| pCS-*serA*-prsA-G226V* | pCS27 containing $P_L$*lacO1-serA*-prsA-G226V* | This study |
| pCS-*trpE*$^{fbr}$*G-serA*-prsA-A95T* | pCS27 containing two operons: $P_L$*lacO1-trpE*$^{fbr}$*G* and $P_L$*lacO1-serA*-prsA-A95T* | This study |
| pSA-*trpDBCA* | pSA74 containing *trpDBCA* from *E. coli* | [287] |

| **Strain** | **Description** | **Source** |
|---|---|---|
| *E. coli* XL1-Blue | *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI*$^q$*ZΔM15Tn10 (Tet*$^R$*)]* | Stratagene |
| *E. coli* BW25113 (F′) | *rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78 F' [traD36 proAB lacIqZΔM15 Tn10(Tetr)]* | [249] |
| *E. coli* ATCC 31884 | A phenylalanine over-producing derivative of *E. coli* K-12 with *aroH367*, *tyrR366*, *tna-2*, *lacY5*, *aroF394*$^{fbr}$, *malT384*, *pheA101*$^{fbr}$, *pheO352*, *aroG397*$^{fbr}$ | ATCC |
| *E. coli* QH4 | *E. coli* ATCC31884 Δ*pheLA* Δ*tyrA* | [208] |
| *E. coli* JW1 | *E. coli* QH4 Δ*pykA* Δ*pykF* | This study |
| *E. coli* JW2 | *E. coli* QH4 Δ*pykA* Δ*pykF* Δ*gldA* | This study |
| *E. coli* JW3 | *E. coli* QH4 Δ*pykA* Δ*pykF* Δ*gldA* Δ*maeB* | This study |
| *E. coli* JW4 | *E. coli* QH4 Δ*trpE* | This study |
| *E. coli* JW5 | *E. coli* QH4 Δ*pykA* Δ*pykF* Δ*gldA* Δ*trpE* | This study |
| *E. coli* JW6 | *E. coli* QH4 Δ*pykA* Δ*pykF* Δ*gldA* Δ*maeB* Δ*trpE* | This study |
| *E. coli* JW7 | *E. coli* QH4 Δ*pykA* Δ*pykF* Δ*gldA* Δ*tnaA* Δ*trpR* with *F'* | This study |

| | | |
|---|---|---|
| | plasmid [traD36 *proAB lacI$^q$ZΔM15*] transduced from *E. coli* XL-1 Blue | |
| *E. coli* JW8 | *E. coli* QH4 with Δ*pykA* Δ*pykF* Δ*gldA* Δ*maeB* Δ*tnaA* Δ*trpR* with *F′* plasmid [traD36 *proAB lacI$^q$ZΔM15*] transduced from *E. coli* XL-1 Blue | This study |
| *E. coli* AA1 | *E. coli* QH4 with pCS-*trpE$^{fbr}$G* | This study |
| *E. coli* AA2 | *E. coli* JW1 with pCS-*trpE$^{fbr}$G* | This study |
| *E. coli* AA3 | *E. coli* JW2 with pCS-*trpE$^{fbr}$G* | This study |
| *E. coli* AA4 | *E. coli* JW3 with pCS-*trpE$^{fbr}$G* | This study |
| *E. coli* TP1 | *E. coli* QH4 with pCS-*trpE$^{fbr}$G-serA*-prsA-A95T* and pSA-*trpDBCA* | This study |
| *E. coli* TP2 | *E. coli* JW7 with pCS-*trpE$^{fbr}$G-serA*-prsA-A95T* and pSA-*trpDBCA* | This study |
| *E. coli* TP3 | *E. coli* JW8 with pCS-*trpE$^{fbr}$G-serA*-prsA-A95T* and pSA-*trpDBCA* | This study |
| *E. coli* MA1 | *E. coli* QH4 with pCS-*trpE$^{fbr}$G* and pZE-*PAPC* | This study |
| *E. coli* MA2 | *E. coli* JW2 with pCS-*trpE$^{fbr}$G* and pZE-*PAPC* | This study |
| *E. coli* MA3 | *E. coli* JW3 with pCS-*trpE$^{fbr}$G* and pZE-*PAPC* | This study |
| *E. coli* MA4 | *E. coli* JW2 with pCS-*trpE$^{fbr}$G-APTA* and pZE-*PAPC* | This study |
| *E. coli* MA5 | *E. coli* JW3 with pCS-*trpE$^{fbr}$G-APTA* and pZE-*PAPC* | This study |

**Table 3.2 Summary of production performance by the engineered pyruvate-driven *E. coli* strains.**

| Host strain | Glycerol | | | | | | Xylose | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AA | | L-Trp | | MA | | AA | | L-Trp | | MA | |
| | Titer [a] | Yield [b] | Titer | Yield | Titer | Yield | Titer | Yield | Titer | Yield | Titer | Yield |
| *E. coli* QH4 | 0.84 | 0.054 | 0.73 | 0.037 | 0.89 | 0.071 | 0.56 | 0.046 | 0.82 | 0.05 | 0.68 | 0.059 |
| *E. coli* JW2 | 1.00 | 0.077 | -[c] | - | 1.82 | 0.122 | 1.35 | 0.131 | - | - | 0.46 | 0.028 |
| *E. coli* JW3 | 1.78 | 0.143 | - | - | 1.60 | 0.125 | 1.90 | 0.178 | - | - | 0.47 | 0.029 |
| *E. coli* JW7 | - | - | 1.73 | 0.132 | - | - | - | - | 1.31 | 0.09 | - | - |
| *E. coli* JW8 | - | - | 1.32 | 0.168 | - | - | - | - | 0.97 | 0.08 | - | - |

[a] Peak titers were used for comparison with titer unit, g/L; [b] yields corresponding to peak titers were used for comparison with yield unit, g/g; [c], not tested.

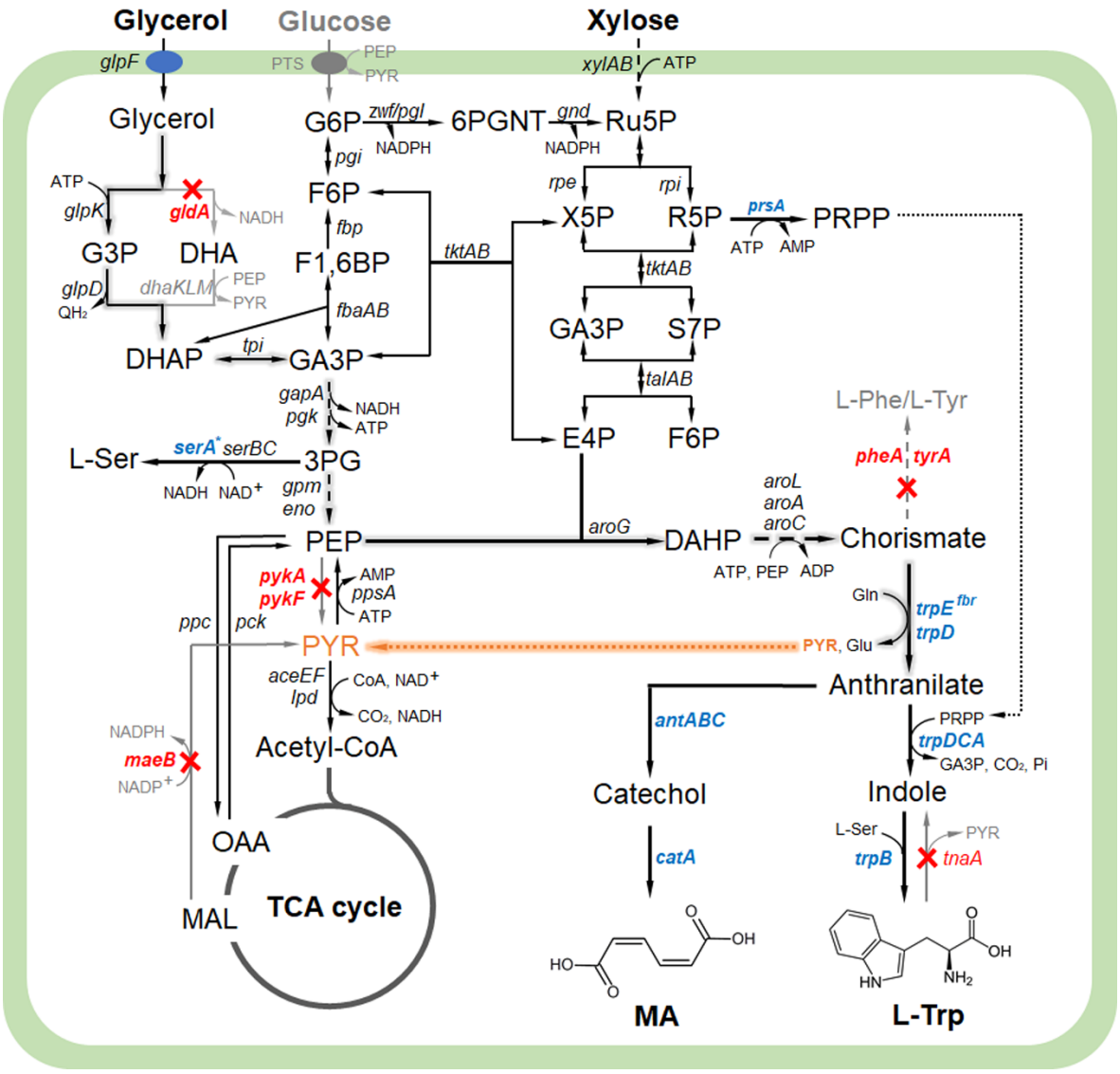

Glycerol
Glucose
Xylose
glpF
PTS
PEP
PYR
xylAB
ATP
Glycerol
G6P
zwf/pgl
6PGNT
gnd
Ru5P
NADPH
NADPH
pgi
ATP
glpK
gldA
NADH
F6P
rpe
rpi
X5P
R5P
prsA
PRPP
ATP
AMP
G3P
DHA
fbp
F1,6BP
tktAB
tktAB
glpD
QH₂
dhaKLM
PEP
PYR
fbaAB
GA3P
S7P
DHAP
tpi
GA3P
talAB
gapA
pgk
NADH
ATP
L-Phe/L-Tyr
E4P
F6P
pheA
tyrA
L-Ser
serA* serBC
3PG
NADH
NAD+
gpm
eno
aroL
aroA
aroC
aroG
DAHP
Chorismate
ATP, PEP
ADP
PEP
pykA
pykF
AMP
ppsA
ATP
Gln
trpE fbr
trpD
ppc
pck
PYR
PYR, Glu
aceEF
lpd
CoA, NAD+
CO₂, NADH
Anthranilate
antABC
PRPP
trpDCA
GA3P, CO₂, Pi
NADPH
maeB
NADP+
Acetyl-CoA
Catechol
Indole
L-Ser
PYR
trpB
tnaA
catA
OAA
MAL
TCA cycle
HO
OH
O
O
MA
HN
NH₂
L-Trp

**Figure 3.1 Engineering pyruvate-driven growth-coupled bioproduction in *E. coli*.** The scheme represents key metabolic pathways, metabolites and genes involved in carbon metabolism, and biosynthetic pathways for anthranilate, L-tryptophan and *cis, cis*-muconic acid. Deleted genes are indicated in red and over-expressed genes are indicated in blue. Metabolite abbreviations: G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; F1,6BP, fructose 1,6-bisphosphate; 6PGNL, 6-phosphogluconolactone; 6PGNT, 6-phosphogluconate; RU5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; X5P, xylulose-5-phosphate; S7P, seudoheptulose-7-phosphate; E4P, erythrose-4-phosphate; G3P, glycerol 3-phosphate; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; GA3P, glyceraldehyde-3-phosphate; 3PG, 3-phosphoglycerate; PEP, phosphoenolpyruvate; PYR, pyruvate; OAA, oxaloacetate; MAL, malate; TCA cycle, tricarboxylic acid cycle; DAHP, 3- deoxy-D-arabino-heptulosonate-7-phosphate; PRPP, 5-phospho-α-D-ribose 1-diphosphate; L-Ser, L-serine; L-Gln, L-glutamine; L-Glu, L-glutamate; L-Phe, L-phenylalanine; L-Tyr, L-tyrosine; L-Trp, L-tryptophan; MA, *cis, cis*-muconic acid.

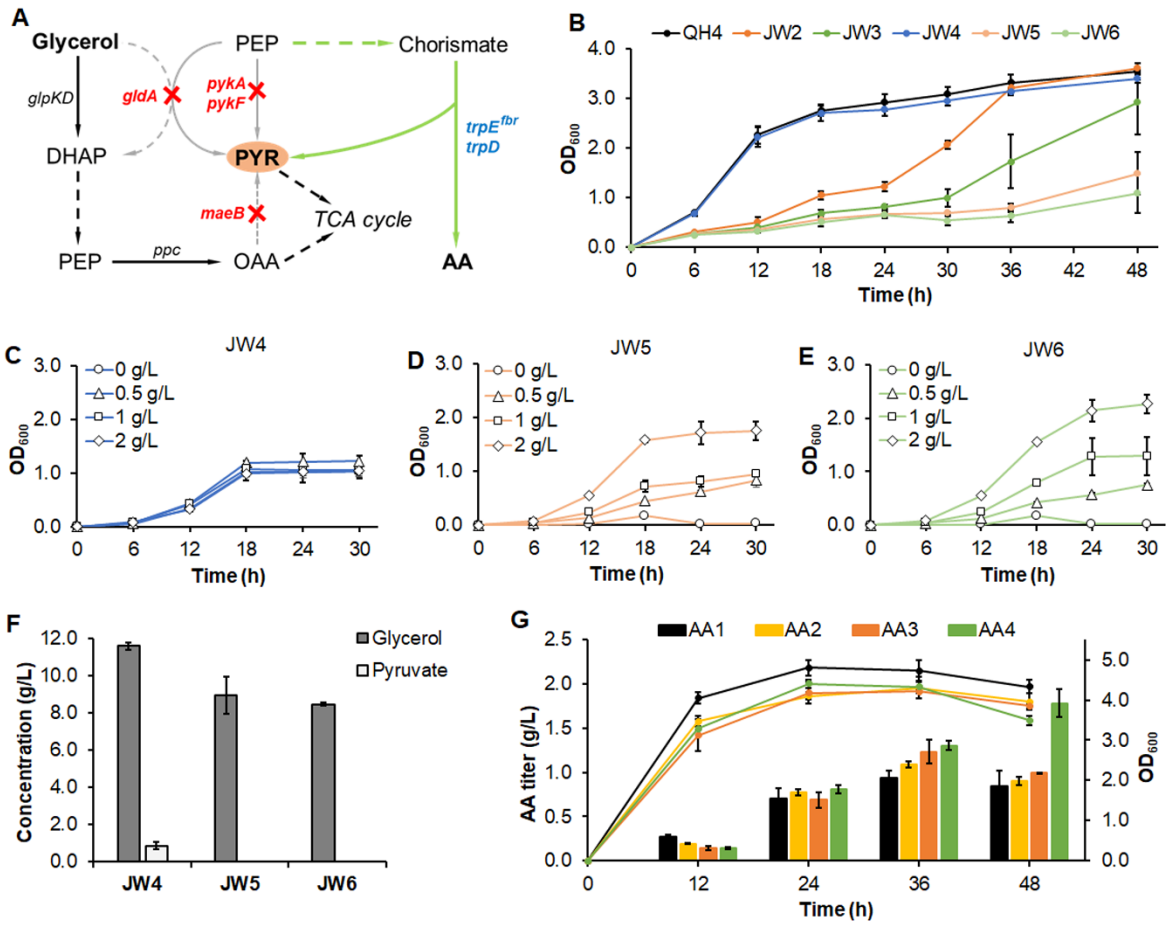



**Figure 3.2 Engineering pyruvate-driven *E. coli* strains for growth-coupled AA production.** (A) Deletion of endogenous pyruvate-releasing genes to enable pyruvate formation mainly via AA biosynthetic pathway. (B) Investigating the growth potential of *E. coli* mutants derived from QH4 in glycerol minimal medium with 0.5 g/L yeast extract. Growth curve of JW4 (C), JW5 (D) and JW6 (E) in glycerol minimal medium (no yeast extract) with gradient concentrations of pyruvate, from 0 (circles) to 0.5 (triangles), 1 (squares) and 2 g/L (diamonds). (F) Residual carbon source in the medium after 30 h cultivation of JW4, JW5 and JW6 with glycerol minimal medium (no yeast extract) containing 2 g/L pyruvate. (G) AA production profiles and growth curves of strains AA1, AA2, AA3 and AA4 in 48 h. Error bars represent s.d. (n = 3).

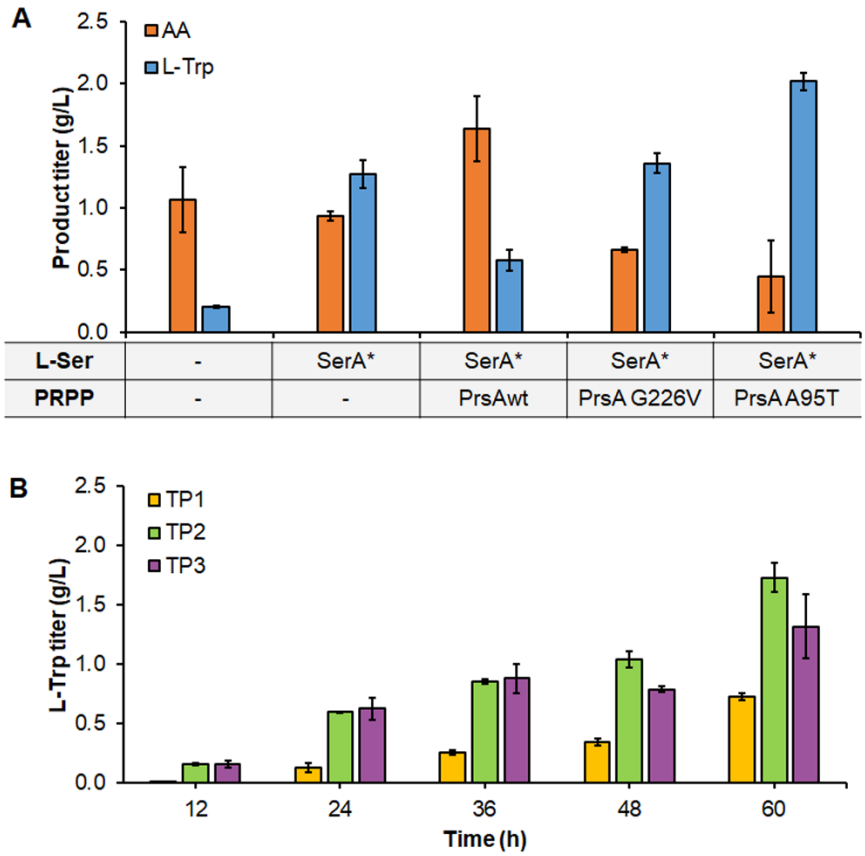



**Figure 3.3 Enhancing L-Trp production from glycerol.** (A) Improving bioconversion of anthranilate (orange) to L-Trp (blue) by increasing L-Ser supply (over-expressing the feedback-inhibition resistant mutant SerA$^{*}$) and PRPP supply (over-expressing PrsA or its variants) in *E. coli* BW25113 (F′). NC, negative control with neither SerA$^{*}$ nor PrsA over-expression. (B) *De novo* production of L-Trp from glycerol by TP1 (gold), TP2 (green) and TP3 (purple). Cell density and L-Trp titers were measured every 12 h for 60 h. Error bars represent s.d. (n = 3).

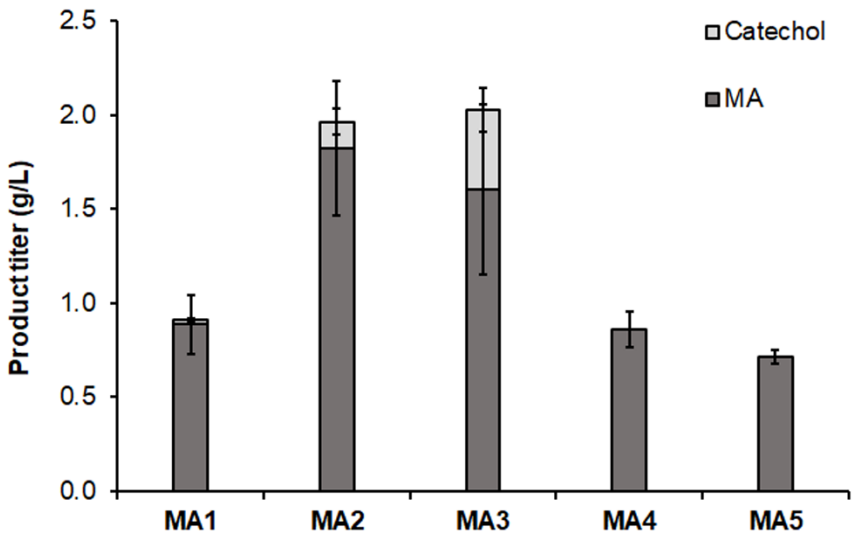



**Figure 3.4 MA production from glycerol using the engineered platform strains.** Comparison of MA production by MA1, MA2, MA3, MA4 and MA5 in glycerol minimal medium. Product profiles include: MA (dark grey) and catechol (light grey). All data points are reported as mean from three independent experiments and the peak titers were applied for comparison. Error bars represent s.d. (n = 3).

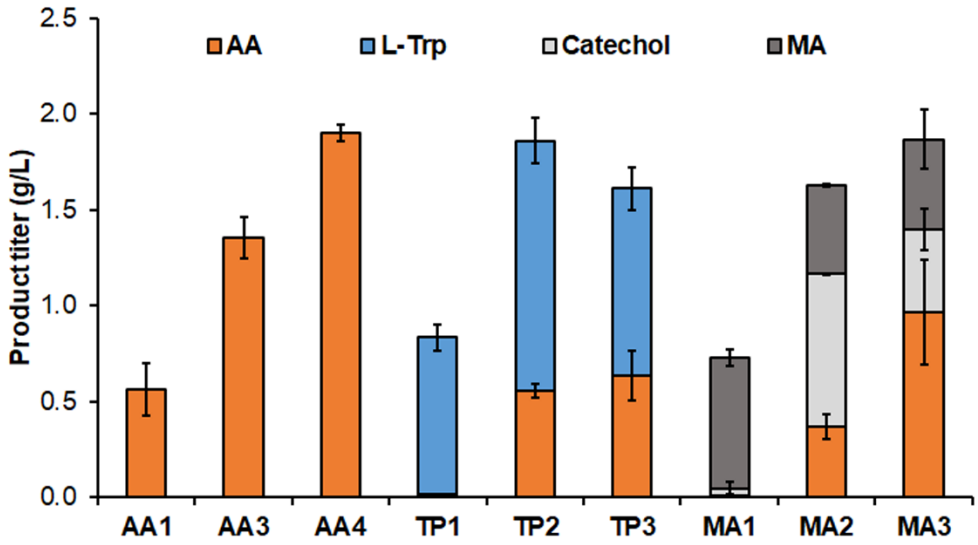



**Figure 3.5 Extending pyruvate-driven system for xylose-based bioproduction.** Production performance of AA producers (AA1, AA3 and AA4), L-Trp producers (TP1, TP2 and TP3) and MA producers (MA1, MA2 and MA3) in xylose minimal medium. Product profiles include: AA (orange), L-Trp (blue), MA (dark grey) and catechol (light grey). All data points are reported as mean from three independent experiments and the peak titers were applied for comparison. Error bars represent s.d. (n = 3).

### 3.7 Supplementary information

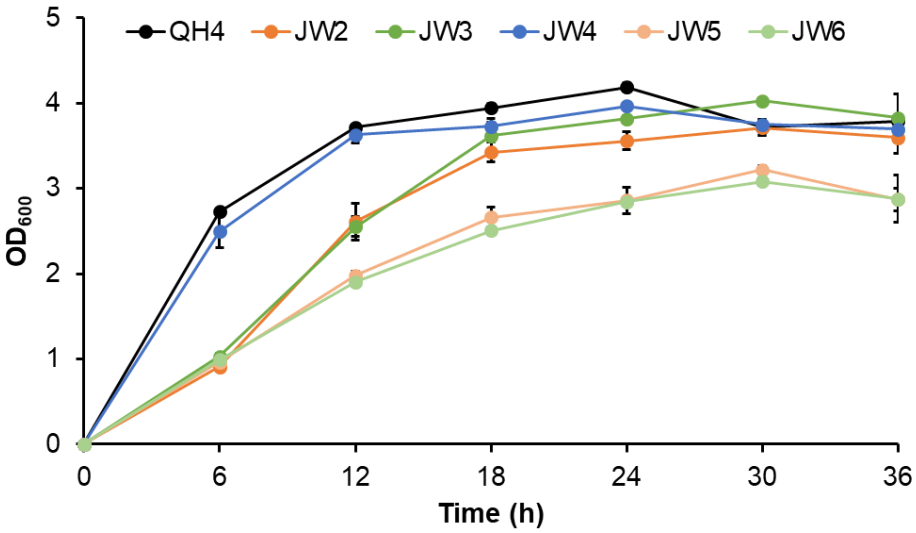



**Figure S3.1 Growth profiles of *E. coli* mutants in glycerol minimal medium with 5 g/L yeast extract in 36 h.**

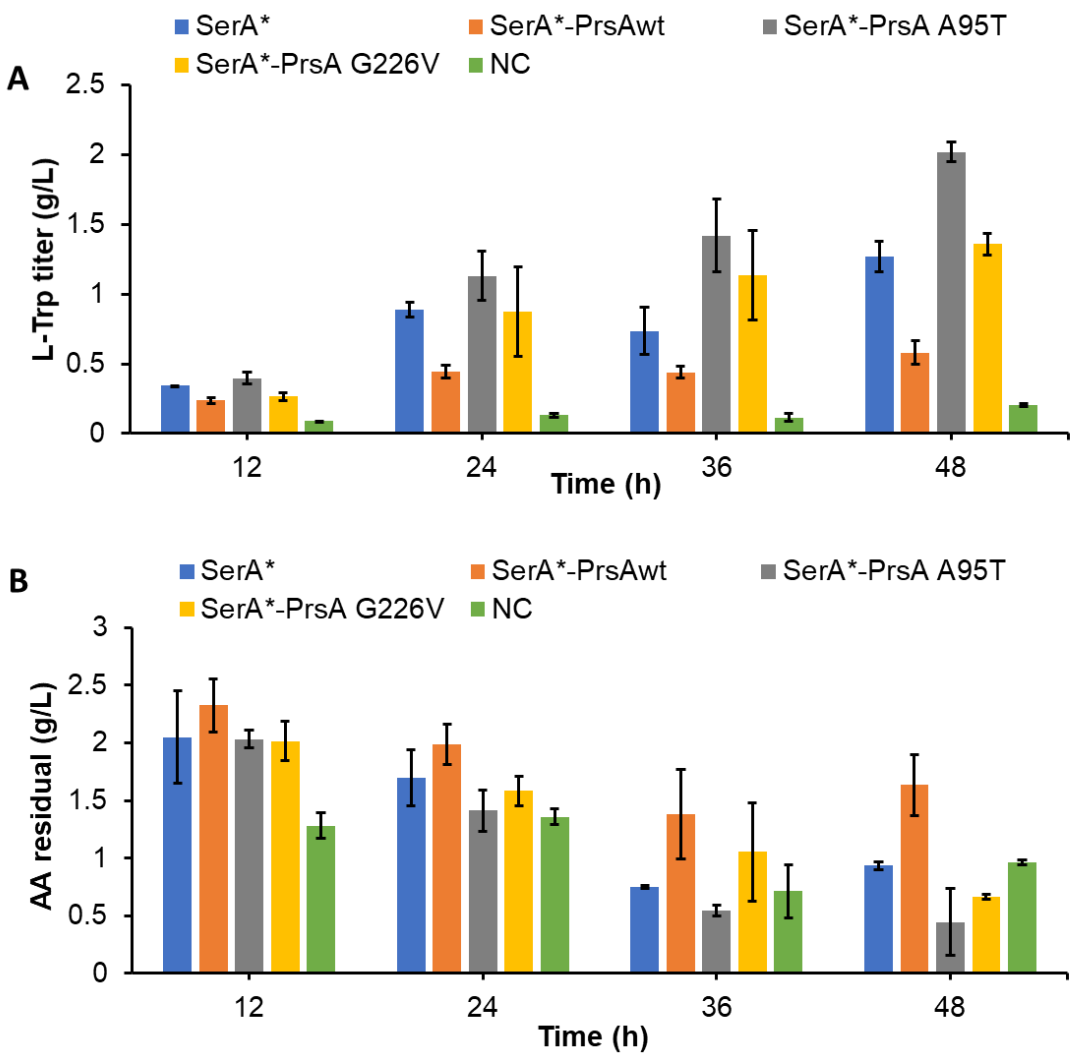



**Figure S3.2 Enhancing AA-to-L-Trp bioconversion by modulating L-Ser and PRPP supply.** AA of 2 g/L was fed into *E. coli* BW25113 (F′) expressing SerA$^{*}$ and PrsA variants. L-Trp production (A) and AA residual (B) in the medium were monitored in 48 h. Error bars represent s.d. (n = 3).

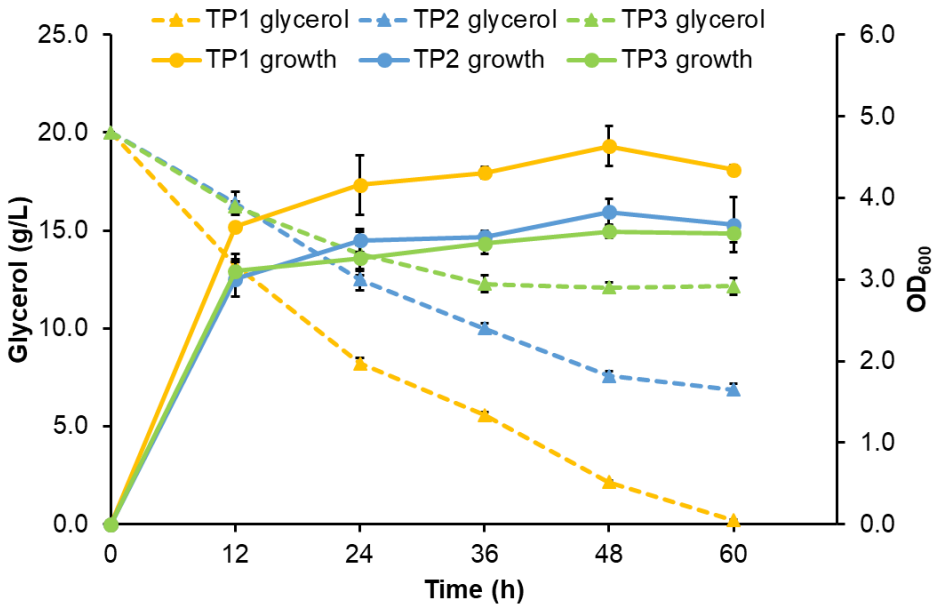



**Figure S3.3 Growth curve and glycerol consumption by L-Trp producers.** Growth curve and glycerol consumption by TP1, TP2 and TP3 in 60 h. Error bars represent s.d. (n = 3).

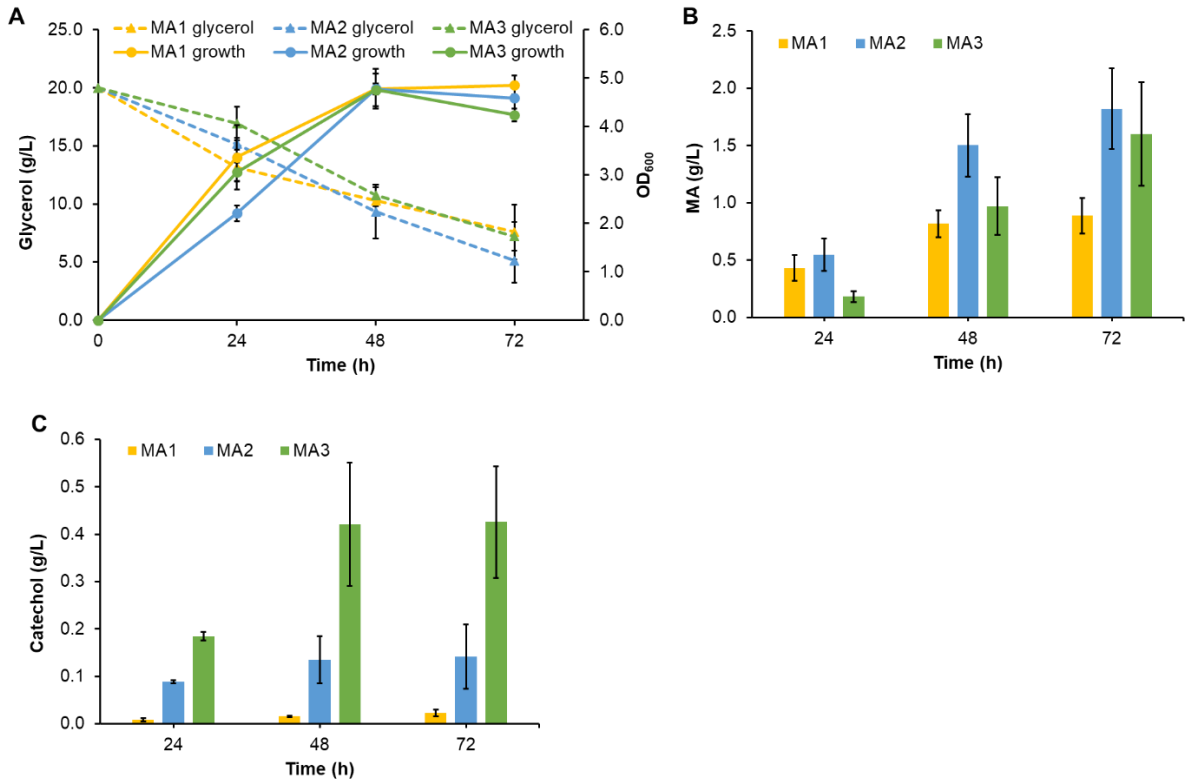



**Figure S3.4** ***De novo*** **MA production from glycerol.** MA production from glycerol via MA producers MA1, MA2 and MA3 in 72 h. Growth curve and glycerol consumption (A), MA production (B) and catechol accumulation (C) by MA1, MA2 and MA3. Error bars represent s.d. (n = 3).

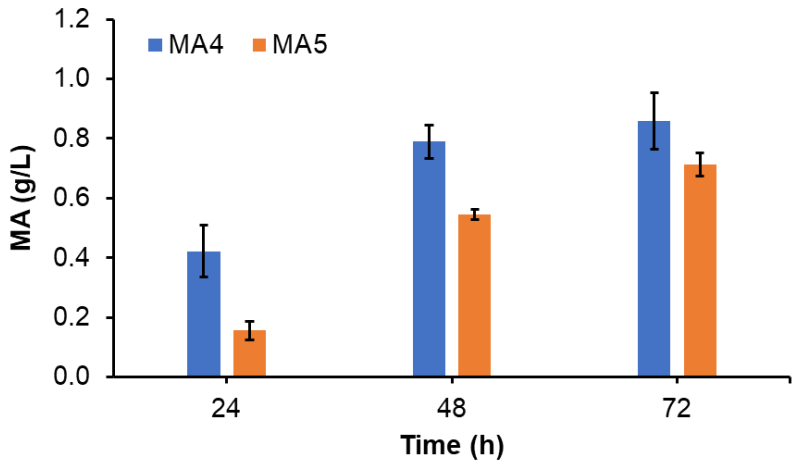



**Figure S3.5 MA production by MA4 and MA5.** MA production from glycerol by strains MA4 and MA5 in 72 h. Error bars represent s.d. (n = 3).

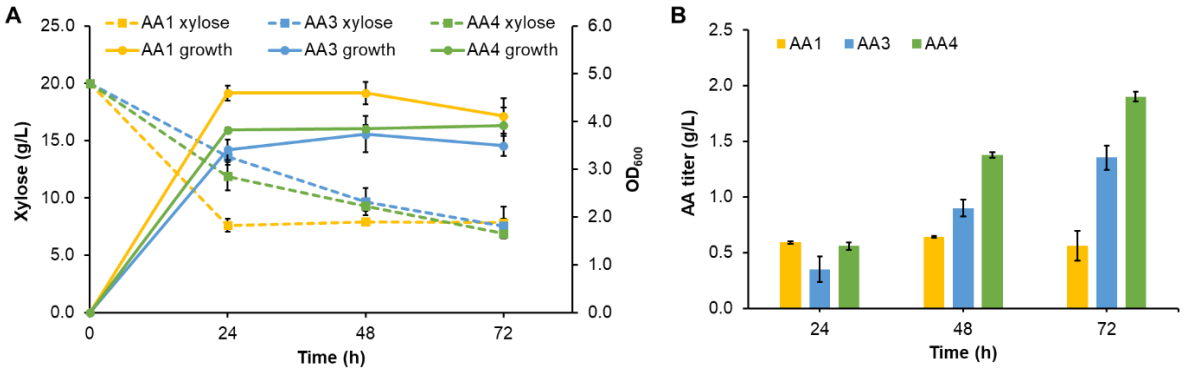



**Figure S3.6 *De novo* AA production from xylose using the engineered platform strains AA1, AA3 and AA4.** Growth curve and xylose consumption (A) and AA production (B) by AA1, AA3 and AA4 in 72 h. Error bars represent s.d. (n = 3).

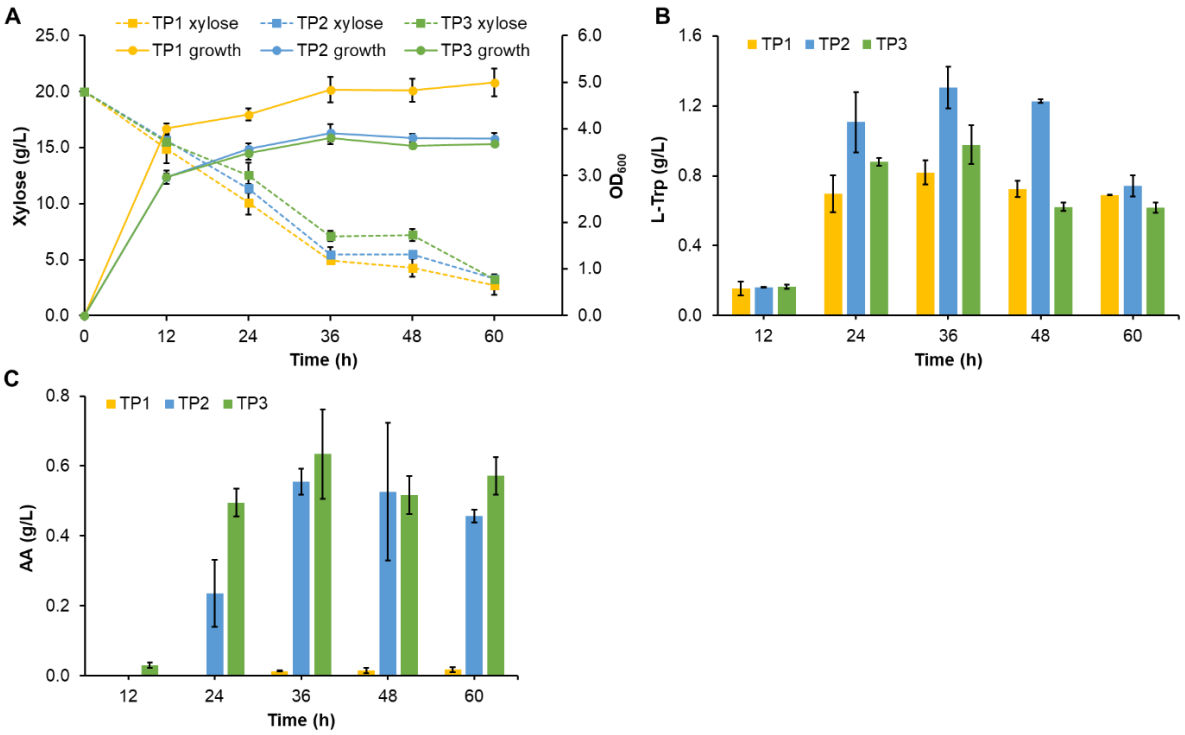



**Figure S3.7 *De novo* production of L-Trp from xylose by TP1, TP2 and TP3.** Growth curve and xylose consumption (A), L-Trp production (B) and AA residual (C) by TP1, TP2 and TP3 in 60 h. Error bars represent s.d. (n = 3).

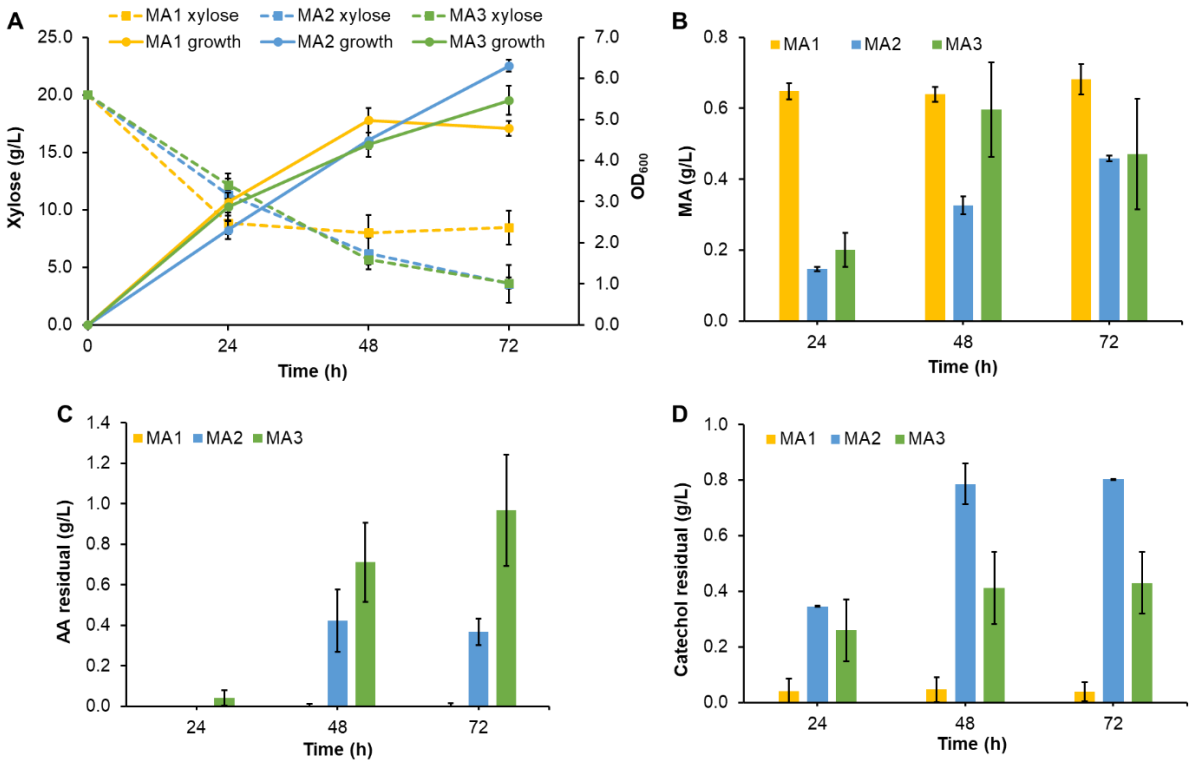



**Figure S3.8 *De novo* production of MA from xylose by MA1, MA2 and MA3.** Growth curve and xylose consumption (A), MA production (B), AA residual (C) and catechol residual (D) by MA1, MA2 and MA3 in 72 h. Error bars represent s.d. (n = 3).

# Chapter 4

# BACTERIAL SYNTHESIS OF C3-C5 DIOLS VIA EXTENDING AMINO ACID CATABOLISM[1]

[1]Jian Wang, Chenyi Li, Yusong Zou, Yajun Yan. Submitted to *Nature Communications*.

## 4.1 Abstract

Amino acids are naturally occurring and structurally diverse metabolites within all microbial systems, whose potentials for chemical expansion, however, have not been fully explored. Here, we devise a novel metabolic platform capable of producing industrially important C3-C5 diols from amino acids. The presented platform combines the natural catabolism of charged amino acids with a new catalytically efficient and thermodynamically favorable diol formation pathway, created by expanding the substrate scope of the carboxylic acid reductase towards noncognate ω-hydroxylic acids. Using the established platform as gateways, seven different diol-convertible amino acids can be converted to diols including 1,3 propanediol, 1,4-butanediol and 1,5-pentanediol. Particularly, we optimize the *de novo* 1,4-butanediol production from glucose and demonstrate the hitherto first synthetic pathway of 1,5-pentanediol, with titers reaching 1.41 and 0.97 g $l^{-1}$, respectively. Our work presents a new metabolic platform that enriches the pathway repertoire for non-natural diols from sugar and protein-derived feedstocks.

## 4.2 Introduction

Microbial hosts, like the most widely used *Escherichia coli* and *Saccharomyces cerevisiae*, can natively generate an exquisite precursor chemical pool and cofactors from renewable feedstocks, offering a microbial-based platform for production of a variety of natural or new-to-nature chemicals, including biofuels, biopolymers, commodity chemicals, and phytochemicals [8,60,314,315]. Extending from existing microbial biochemical infrastructure, carbon flux can be streamed into products of interest via purpose-driven biosynthetic pathways. However, most industrial applications are currently focused on utilization of only a limited subset of native metabolites or pathways; for instances, to produce short-chain alcohols or isoprenoids from glycolysis intermediates like pyruvate and acetyl-CoA, to produce long-chain fatty acids, alcohols or alkanes from fatty acid biosynthetic pathway, and to produce short-chain dicarboxylic acids from tricarboxylic acid (TCA) cycle [122,316,317]. Although substantial developments have been made in expanding the microbial capacity [261,318], the chemical space accessible from microbial-based system has been significantly restrained by the narrow scope of the precursor pools, as well as the scarcity of natural or synthetic pathways for new products.

Amino acids (AAs) are one of the major endogenous metabolites in almost all microbes, whose utilities as building blocks, however, have been overlooked. Currently, most of the 20 proteinogenic AAs are microbially produced as end and bulk biochemicals with around 5 million tons per year, and are mainly used as animal feed additives and flavor ingredients [319,320]. The structural diversity, ease of achieving high titers in microorganisms and ever-increasing availability from inexpensive protein sources, have substantiated AAs with potentials to serve as natural and versatile scaffolds for various industrially-relevant natural or synthetic products [52,321]. In recent decades, intensive efforts have been taken towards converting branched-chain AAs

(alanine, valine, leucine, isoleucine, etc.) to advanced alcohols or acids via the Ehrlich pathway [167,168,170], and aromatic AAs (phenylalanine, tyrosine and tryptophan) to polyphenolic phytochemicals via phenylpropanoic or alkaloid pathways [8,257]. Whereas, the development of efficient biocatalytic routes to utilize charged AAs (aspartate, glutamate and lysine, etc.), which account for > 40 % of all AAs in protein hydrolysates and > 80 % of all the amino acid market, are comparatively less tapped [144,319]. Charged AAs harbor bi-functional terminal groups, making them recalcitrant to be directly metabolized through central carbon metabolism. Additionally, the paucity of natural or synthetic pathways to utilize charged AAs significantly limited their valorization.

In the present study, we explored the potential of migrating the amino acid industry to chemical industry by establishing non-natural pathways capable of converting charged AAs to C3-C5 diols. C3-C5 diols, including 1,3 propanediol (1,3 PDO), 1,4-butanediol (1,4 BDO) and 1,5-pentanediol (1,5 PDO), have wide use in commodity and fine chemical industries such as fuels, solvents, polymer monomers and pharmaceutical precursors [322-324]. These short-chain diols represent an annual commercial market over $ 7 billion with a nearly 7 % annual growth rate [325]. 1,3 PDO can be naturally produced from anaerobic reduction of glycerol, which sometimes, however, involves supplementation of expensive vitamin $B_{12}$ for glycerol dehydratase [322]. Although a new glycerol-independent 1,3 PDO route has been developed from homoserine degradation, it suffered from limited pathway performance [326]. As for 1,4 BDO and 1,5 PDO, there is no natural pathway for their direct production. Until recently, two major artificial biocatalytic routes to 1,4 BDO were developed: one that involves reduction of succinyl-CoA or decarboxylation of α-ketoglutarate to 4-hydroxybutyrate (4HB) with subsequent CoA-dependent reduction of 4HB to 1,4 BDO, and a second that involves nonphosphorylative xylose metabolism

[129,311]. Notably, the former CoA-dependent pathway suffers from drawbacks including low enzyme kinetics and lactonization or hydrolysis of CoA intermediates that lead to incomplete conversion of 4HB to 1,4 BDO, while the latter is specialized for C5 sugar feedstocks (e.g. xylose and arabinose) and requires energy input from additional carbon sources [129,311,327]. With respect to 1,5 PDO, its petroleum-based synthesis has been restricted due to limited access to C5 petroleum feedstocks, while its biosynthesis from renewable feedstocks has not been developed to date [328]. Due to the linear infrastructures and the existence of catabolic machineries for charged AAs in nature, we presume that charged AAs can be potentially diverted to C3-C5 diols.

Here we describe the development of a metabolic platform that engenders direct conversion of charged AAs into C3-C5 diols. The centerpiece of the platform is to release the carbon skeletons from charged AAs into "C-1" ω-hydroxylic acids (ω-HAs) via native or heterologous amino acid catabolism, followed by exploiting the catalytic capabilities of carboxylic acid reductase (Car) to reduce noncognate ω-HA substrates to diols. The new synthetic diol pathways bypass either expensive co-enzyme supplementation or CoA-dependent chemistry, and is featured by broad host applicability and feedstock flexibility. The presented platform could convert a panel of seven AAs into C3-C5 diols, based upon which we were able to achieve and optimize the *de novo* production of 1,4 BDO from glucose and demonstrated a novel synthetic pathway to 1,5 PDO. Our platform highlighted the feasibility of expanding the chemical space of microbial systems by extending amino acid catabolism, which could potentially serve as a gateway for diol production from renewable biomass or protein-derived feedstocks.

## 4.3 Results

### 4.3.1 Design and validation of a synthetic platform for C3-C5 diols

The proposed platform of amino acid to diol conversion proceeds through the upstream degradation of charged AAs into corresponding "C-1" ω-HAs, followed by the downstream reduction of ω-HAs into counterpart diols (**Figure 4.1**). Typical charged AAs, including aspartate (Asp), glutamate (Glu) and lysine (Lys), share a common degradation process of sequential decarboxylation and transamination to counterpart "C-1" semialdehydes, which are further reduced to ω-HAs via aldehyde reductases (ALRs)/alcohol dehydrogenases (ADHs). With the wide existence of well-elucidated native or heterologous AA catabolism supporting our hypothesized platform, the ω-HA-to-diol conversion becomes the most pivotal step for diol production.

The classical pathway of ω-HA-to-diol conversion harnesses the CoA-dependent reduction route naturally existed for alcohol biosynthesis, which involves activation of carboxylate group by a CoA transferase, followed by the reduction the CoA intermediate to aldehyde and then diol by a bi-functional alcohol dehydrogenase (AdhE2) with input of 2 NAD(P)H (**Figure S4.1A**) [129,249]. In addition to low enzyme kinetics and low stability of thioester intermediate, this CoA-dependent route is oxygen-sensitive since AdhE2 is fully functional in anaerobic conditions, and suffers from significant acetate accumulation ascribed to the use of acetyl-CoA as CoA donor [129]. We propose an alternative direct reduction route that utilizes a promiscuous carboxylic acid reductase (Car) and endogenous ALRs/ADHs to reduce carboxylate group to alcohol group at the cost of 1 ATP and 2 NADPH (**Figure 4.2A**). Car was first characterized to catalyze the reduction of aromatic carboxylic acids, and was later revealed to accept a wide range of aliphatic fatty acids (C6-C18) [133]. We herein attempt to explore its catalytic promiscuity towards

noncognate substrates C3-C5 ω-HAs to produce counterpart short-chain diols. To compare the thermodynamics of the two reduction routes of 4HB to 1,4 BDO, we performed the Max-min Driving Force (MDF) analysis via eQuilibrator [329]. The Car-based reduction pathway operates at higher MDF (38.0 kJ/mol) than the CoA-dependent pathway (28.8 kJ/mol), indicating a higher pathway flux and/or a lower enzyme requirement under physiological conditions in *E. coli* (**Table S4.2**).

To practically evaluate the pathway performance, both reduction routes were expressed in *E. coli* and tested by feeding 3 g $l^{-1}$ of C3-C5 ω-HAs. Since 4HB and 5-hydroxyvalerate (5HV) are not commercially available, they were produced respectively from γ-butyrolactone and σ-valerolactone by saponification under alkaline conditions. The CoA-dependent route was borrowed from the previously established 1,4 BDO pathway that comprises the 4-hydroxybutyryl-CoA transferase (*cat2*) from *Porphyromonas gingivalis* and the codon-optimized alcohol dehydrogenase (*adhE2*) from *Clostridium acetobutylicum* [129]. For feeding experiments, *E. coli* strains were grown in M9 minimal media supplemented with 20 g $l^{-1}$ glucose and 5 g $l^{-1}$ yeast extract (hereafter referred to as M9Y media). When fed with C3-C5 ω-HAs including 3-hydroxypropionate (3HP), 4HB or 5HV, the recombinant *E. coli* strain HA-1 expressing *cat2* and *adhE2* on the high-copy plasmid (pZE-*cat2-adhE2*) produced peak titers of 17.4 mg $l^{-1}$ 1,3 PDO and 41.7 mg $l^{-1}$ 1,4 BDO in 36 h, respectively, while no production of 1,5 PDO was observed (**Figure S4.1B**). The Car-based pathway, comprised of *car* from *Mycobacterium marinum* and phosphopantetheine transferase (*sfp*) from *Bacillus subtilis*, were expressed in *E. coli* BW25113 (F′) on the high-copy-number plasmid (pCar), yielding strain HA-2. For the sake of simplicity, we excluded the expression of ALRs/ADHs as it has been reported that the endogenous ones could efficiently reduce various aldehydes to alcohols [330]. HA-2 showed limited conversion

capability towards 3HP, which produced 0.97 g $l^{-1}$ 1,3 PDO from 3 g $l^{-1}$ 3HP in 36 h (**Figure 4.2B**). HA-2 showed high conversion capability towards 4HB and 5HV, and produced 2.91 g $l^{-1}$ 1,4 BDO in 36 h and 2.70 g $l^{-1}$ 1,5 PDO in 12 h (**Figure 4.2B**). As a result, HA-2 achieved 3HP conversion to 38.6 % of the theoretical maximum, and achieved 4HB and 5HV conversion to 100 % of theoretical maximum. The bioconversion results underlined the thermodynamic and enzymatic superiority of the Car-based pathway over CoA-dependent pathway. Noteworthy, Car exhibited a catalytic preference towards longer chain ω-HAs as observed previously towards long-chain fatty acids [133,331]. This was further confirmed by the increase of the specificity enzyme activity of Car towards substrates with increased chain-lengths (0.18 μmol $min^{-1}$ $mg^{-1}$ for 3HP versus 1.25 μmol $min^{-1}$ $mg^{-1}$ for 5HV) during *in vitro* enzyme assay using crude enzyme extracts of HA-2 (**Figure 4.2C**).

### 4.3.2 Establishing the AA degradation pathways in *E. coli*

With validation of an efficient ω-HA-to-diol conversion pathway, our second goal is to assemble the upstream amino acid degradation pathways that could convert charged AAs into "C-1" ω-HAs. *E. coli* harbors native degradation pathways for Asp, Glu and Lys, which however, are committed to either CoA biosynthesis or acid resistance and do not naturally yield counterpart ω-HAs [332-334]. For Asp degradation, a synthetic 3HP pathway was constructed that consists of L-aspartate-α-decarboxylase (PanD), β-alanine pyruvate transaminase (BAPAT), and malonate semialdehyde (MSA) reductase [136]. The native Glu metabolism, also known as γ-aminobutyrate (GABA) pathway, produces GABA via glutamate decarboxylase (GadB). GABA can be trans-aminated to succinate semialdehyde (SSA) by γ-aminobutyrate aminotransferase (transaminase) (GabT), followed by reduction to 4HB via endogenous NAD(P)H-dependent

ADHs like YqhD [127,335]. The native Lys degradation pathway converts Lys to glutarate semialdehyde (GSA) via a four-step pathway comprising lysine decarboxylase (CadA), putrescine transaminase (PatA), 5-aminopentanal dehydrogenase (PatD) and GabT [336]. Alternatively, a heterologous three-step Lys degradation pathway from *Pseudomonas putida*, known as AMV pathway, converts Lys to GSA with lysine 2-monooxygenase (DavB), 5-aminovaleramidase (DavA) and 5-aminovalerate transaminase (DavT) [337]. Further reduction of GSA by endogenous ADHs will release 5HV.

To render conversion of AAs to C3-C5 ω-HAs, we constructed and expressed each ω-HA pathway in *E. coli*. The synthetic 3HP pathway, consisting of *panD* from *C. glutamicum*, *BAPAT* from *P. putida*, and MSA reductase (*ydfG*) from *E. coli* [136], was integrated into the medium-copy plasmid, yielding p3HP. Similarly, the native GABA pathway for Glu degradation consisting of *gadB*$^{*}$ (encoding GadB mutant E89Q Δ452-466) [338], *gabT* and *yqhD* from *E. coli*, was integrated into p4HB; while the heterologous AMV pathway for Lys degradation consisting of *davBA* from *P. putida*, *gabT* and *yqhD* from *E. coli*, was integrated into p5HV1. These three plasmids were transferred into *E. coli* BW25113 (F′), yielding strains HA-3 to HA-5 (**Figure 4.3A**). When fed with 3 g $l^{-1}$ corresponding AAs, HA-3 produced 247.3 mg $l^{-1}$ 3HP from Asp, HA-4 produced 382.5 mg $l^{-1}$ 4HB from Glu, and HA-5 produced 690.8 mg $l^{-1}$ 5HV from Lys in 48 h of cultivation (**Figure 4.3B**). These production demonstrated the functionality of all amino acid degradation pathways, with conversion efficiencies of 7.1, 12.8 and 23.0 %, respectively.

### 4.3.3 Direct conversion of charged AAs to C3-C5 diols

Validation of both upstream and downstream parts made it possible to achieve the direct conversion of charged AAs including Asp, Glu and Lys to diols. Considering the low

performance of 3HP and 4HB pathway on medium-copy plasmids, the 3HP and 4HB pathway were reassembled into the high-copy plasmid pCar, yielding p3PDO and p4BDO, respectively. *E. coli* strain 3PDO-1 harboring p3PDO could not produce any 1,3 PDO with no exogenous Asp feeding, but produced 276.6 mg $l^{-1}$ 1,3 PDO when fed with 3 g $l^{-1}$ of Asp in 48 h (**Figure 4.3C**). This confirmed the functional expression but low activity of 1,3 PDO pathway, which might be caused by the cumulative inefficiency of the 3HP pathway and Car towards 3HP. With no Glu or Lys feeding, strain 4BDO-1 containing p4BDO produced 588.3 mg $l^{-1}$ 1,4 BDO and 5PDO-1 containing pHV1 and pCar produced 123.9 mg $l^{-1}$ 1,5 PDO (**Figure 4.3C**). This indicated that both 1,4 BDO and 1,5 PDO pathways were functionally expressed, and were capable of direct conversion of intracellular Glu and Lys to counterpart diols. Further feeding of 3 g $l^{-1}$ Glu to strain 4BDO-1 slightly increased 1,4 BDO titer by 47.8 % (869.7 mg $l^{-1}$), accounting for an overall yield 0.29 g/g Glu (47.3 % of theoretical maximum) (**Figure 4.3C**). Considering low Glu to 4HB but high 4HB to 1,4 BDO conversion, Glu degradation was suggested to be a potential limiting step for 1,4 BDO production. This also clearly indicated that 1,4 BDO titer could not be significantly improved by simply imposing glutamate supply. Strikingly, feeding 3 g $l^{-1}$ Lys to strain 5PDO-1 drastically increased 1,5 PDO titer by 10.7-fold (1.33 g $l^{-1}$) (**Figure 4.3C**), accounting for an overall yield of 0.44 g/g Lys (62.2 % of theoretical maximum). This suggested that 1,5 PDO pathway is rather efficient as high-yield conversion was achieved even without optimization of the pathway and cultivation conditions. As such, lysine supply might be a major bottleneck for *de novo* 1,5 PDO production.

### 4.3.4 Bioconversion of diol-convertible AAs

On construction and validation of the C3-C5 diol production platform, we sought to broaden its substrate scope by rerouting all potential AAs into counterpart diols. *E. coli* natively harbors amino acid degradation network that enables convergence of four additional AAs into our platform (**Figure 4.4A**). Asparagine (Asn) is directly converted to Asp via asparaginases (AnsA and AnsB). Glutamine (Gln) can be converted to Glu via glutaminase (YbaS) and proline (Pro) can be converted to Glu via bi-functional proline utilization A (PutA). Arginine (Arg) can be degraded to GABA via four-step pathway comprising of arginine decarboxylase (AdiA), agmatinase (SpeB), PatA and PatD. All these AAs, hereafter referred as diol-convertible AAs, have previously remained unexplored for deriving new chemical entities due to structural complexity as well as no existing pathways.

We then adopted these native degradation pathways for diol-convertible AAs to enable diol release. To this end, we co-transferred each diol pathway plasmid and AA degradation pathway plasmid into *E. coli* BW25113(F′), and the resulting strains (DO-1 to DO-4) were cultured aerobically in M9Y media supplemented with 3 g $l^{-1}$ amino acid substrates for 48 h. When fed with Asn, strain DO-1 harboring p3PDO and pCS-*asnA* produced 145.3 mg $l^{-1}$ 1,3 PDO (**Figure 4.4B**). Regarding 1,4 BDO conversion, strain DO-2 expressing YbaS enabled a titer of 650.1 mg $l^{-1}$ from Gln, strain DO-3 expressing PutA enabled a titer of 903.1 mg $l^{-1}$ from Pro, and strain DO-4 expressing the four-step Arg degradation pathway enabled a titer of 648.8 mg $l^{-1}$ from Arg (**Figure 4.4B**). The bioconversion experiments demonstrated that all tested AAs were successfully converted to counterpart diols, albeit with relatively low conversion efficiencies. This also suggested that Glu degradation might be a major bottleneck limiting the utilization of 1,4 BDO convertible AAs. Since Asp is the direct precursor for Lys biosynthesis, we also test if

feeding Asp could lead to 1,5 PDO production. Unexpectedly, feeding Asp to strain 5PDO-1 could just slightly increase 1,5 PDO titer (138.7 mg $l^{-1}$), further suggesting that the downstream lysine pathway is limiting lysine biosynthesis, and hence the synthesis of 1,5 PDO.

### 4.3.5 Production enhancement of 1,4 BDO in *E. coli*

After achieving AA-to-diol conversion, we set out to demonstrate and optimize the viability of this new pathway for 1,4 BDO production from cheaper renewable carbon sources like glucose. 1,4 BDO is a large-volume commodity chemical with wide applications in chemical industry, and its bio-based production at the commercial scale has been achieved via the CoA-dependent pathway with extensive pathway and host engineering [339]. Given that glutamate is the most abundant intracellular metabolite serving as the major nitrogen donor in glucose-grown *E. coli* cells [340], and thus, it can readily be channeled into 1,4 BDO pathway. This is in line with the observation of 1,4 BDO production (588.3 mg $l^{-1}$) when only introducing the 1,4 BDO pathway into *E. coli* (**Figure 4.5A**), even without further optimization.

To further increase 1,4 BDO production, we sought to improve Glu decarboxylation efficiency. The wildtype glutamate decarboxylase of *E. coli* is only functional in acidic conditions, and its mutant GadB* (E89Q Δ452-466) showed improved activity at neutral pH [338]. Therefore, we resorted to improving its performance by increasing its cofactor supply. Glutamate decarboxylase $GadB^{*}$, as well as GABA aminotransferase GabT, are pyridoxal 5'-phosphate (PLP)-dependent enzymes. The native *de novo* PLP pathway in *E. coli* starts from erythrose-4-phosphate (E4P) via a seven-step deoxyxylulose-5-phosphate (DXP)-dependent pathway [341]. To determine the contribution of increased PLP supply to 1,4 BDO production, we instead over-expressed an alternative ribose 5-phosphate (R5P)-dependent PLP pathway from *Bacillus subtilis*

that involves only two enzymes, PdxS and PdxT. PdxS (PLP synthase subunit) and PdxT (glutamine hydrolase subunit) form a dimeric complex that utilizes glutamine, R5P and glyceraldehyde 3-phosphate (G3P) to synthesize PLP [342]. An additional benefit of PdxST pathway is the release of glutamate from glutamine, which can be taken up by 1,4 BDO pathway (**Figure 4.5A**). When cultivating strain 4BDO-2 harboring p4BDO and pCS-*pdxST* in M9Y medium, 675.4 mg $l^{-1}$ 1,4 BDO was produced in 48 h, accounting for 13.7 % increase than that produced by control strain 4BDO-1 (**Figure 4.5B**). When further fed with 3 g $l^{-1}$ glutamate, the 1,4 BDO titer was almost doubled (1.21 g $l^{-1}$), reaching an overall yield 0.40 g/g Glu (66.0 % of theoretical maximum) (**Figure S4.2**). These results implied that PLP surplus would significantly promote 1,4 BDO production, especially when glutamate flux was not limiting.

Hence, we next sought to increase carbon flux to Glu pool. To redirect carbon flux from glycolysis to TCA cycle, we over-expressed an NADH-insensitive citrate synthase mutant (*gltA*$^{*}$, R163L) [129,343], and phosphoenolpyruvate carboxylase (*ppc*) (**Figure 4.5A**). The resultant strain 4BDO-3 increased 1,4 BDO titer by 45.6 % (0.87 g $l^{-1}$) than that of control strain 4BDO-1 (**Figure 4.5B**). This confirmed that redirecting carbon flux to glutamate could augment Glu flux and thus 1,4 BDO production. When simultaneously introducing *pdxST*, the resultant strain BDO-4 afforded 1.41 g $l^{-1}$ 1,4 BDO in 48 h with a yield of 0.14 mol/mol glucose, representing approximately a 2.40-fold increase of titer in relative to its parental strain 4BDO-1 (**Figure 4.5B**). Noteworthy, only trace amounts of acetate were accumulated in the strain BDO-4 culture (**Figure S4.3**). As a demonstration, our results highlight the potential of using glutamate as an intermediate for 1,4 BDO overproduction from renewable sugars.

### 4.3.6 Optimization of the *de novo* 1,5 PDO production from glucose

1,5 PDO is another high-value diol and a safer replacement for 1,4 BDO with wide applications as a monomer for fibers and polyurethanes. However, its relatively high price and low production capacity limit its commercial use [325,344]. Currently, 1,5 PDO is chemically produced from biomass-derived furfural and tetrahydrofurfuryl alcohol via expensive metal catalysts [345,346], while direct biocatalytic routes from renewable carbon source to 1,5 PDO has not been established. With the lysine conversion platform, it is possible to fill the gap by achieving *de novo* production of 1,5 PDO from glucose.

As indicated by bioconversion experiments, lysine supply is rate-limiting for 1,5 PDO production. With no lysine supply, only 123.9 mg $l^{-1}$ 1,5 PDO could be produced by strain 5PDO-1 (**Figure 4.3C**). Given the need to increase flux of precursors to lysine, we set out to over-express essential genes in lysine pathway. The native lysine pathway involves ten enzymatic steps starting from oxaloacetate (OAA), within which its biosynthesis-related genes are transcriptionally repressed by lysine via ArgP while aspartate kinase (LysC) and dihydrodipicolinate synthetase (DapA) are feedback inhibited by lysine [347,348]. To increase carbon flux to lysine, we introduced *lysC* and *dapA* on p5HV1 and the resultant plasmid p5HV2 were co-transferred into *E. coli* BW25113 (F'), yielding strain 5PDO-2. Meanwhile, to desensitize lysine-mediated feedback inhibition [349,350], LysC mutant T253R and DapA mutant H56K or E84T were also recruited, yielding strain 5PDO-3 and 5PDO-4. The shake flask experiments were performed in M9Y media and samples were taken at 48 h. The HPLC analysis of the culture showed that strain 5PDO-2 expressing wildtype LysC and DapA produced 322.7 mg $l^{-1}$ 1,5 PDO, representing 2.6-fold increase to that of control strain 5PDO-1 (**Figure 4.6**). Whereas, strain 5PDO-3 and 5PDO-4 expressing LysC and DapA mutants could not further

increase 1,5 PDO production in comparison with strain 5PDO-2, with titers of 271.5 and 311.3 mg $l^{-1}$, respectively. This suggested that lysine degradation via 1,5 PDO pathway alleviated lysine mediated feedback inhibition, leading to enhanced 1,5 PDO production via wildtype pathway enzymes. Therefore, we seek to further enhance lysine degradation pathway to relieve feedback inhibition and improve 1,5 PDO production. We incorporated the *davB-davA-gabT-yqhD* operon into a high-copy-number plasmid pSC74 that was derived from pSA74 with *CloDF13 ori* (20-40 copies per cell), resulting in plasmid p5HV5. To test its efficacy in Lys conversion, 3 g $l^{-1}$ Lys was fed to strain 5PDO-5 containing pCar and p5HV5. Strikingly, 1,5 PDO titer was increased to 1.67 g $l^{-1}$, reaching an overall yield 0.56 g/g Lys (78.2 % of theoretical maximum) (**Figure S4.4**). This confirmed that increased expression of the Lys degradation pathway augmented Lys degradation and thus 1,5 PDO production. When co-transferring pCar, p5HV5 and pCS-*lysC-dapA* into *E. coli* BW25113 (F'), the resultant strain 5PDO-6 significantly increased 1,5 PDO to 0.86 g $l^{-1}$, representing a 6.94-fold increase than that of starter strain 5PDO-1 (**Figure 4.6**). Finally, to increase the supply of lysine pathway intermediate OAA, we knocked out *iclR*. IclR is a repressor regulator of the glyoxylate bypass that is an alternative mechanism of replenishing oxaloacetate [144,351]. The final 1,5 PDO titer produced by strain 5PDO-7 with *iclR* deletion reached 0.97 g $l^{-1}$ at 48 h with a yield of 0.08 mol/mol glucose. Interestingly, 5PDO-7 also accumulated the least acetate in the culture (0.24 g $l^{-1}$) (**Figure 4.6**). Taken together, our results demonstrated the functionality of a novel and efficient 1,5 PDO pathway in *E. coli*, and showed that the 1,5 PDO production could be further boosted with increased lysine supply.

## 4.4 Discussion

To expand the accessible chemical space for microbial production systems, exploiting new catalytic capabilities by extending existing metabolic networks can expedite the creation of non-natural pathways for new target products. As such, extending the existing microbial pathways took advantage of native metabolic highways and starting materials, which is superior over assembling long synthetic pathways from scratch, to produce new products of desired [168]. Native metabolic network provides naturally occurring backbones for cellular activity and biomass formation, which could otherwise be hijacked for rationally designed synthetic purposes. Amino acids are one of such primary and versatile metabolites naturally produced both in *E. coli* and in all other microorganisms. Their structural diversity and the existence of AA anabolism and natural degradation machinery make them ideal precursors to derive new chemicals [8,167,168,170,257]. Here we demonstrated using *E. coli* as a production host for extending AA catabolism to generate industrially important C3-C5 diols. The established platform unveiled a hitherto less explored path to expand the chemical space from charged AAs in microbial systems.

Charged AAs have been used to produce bi-functional chemicals, including short-chain ω-HAs, dicarboxylates, diamines, lactams, etc. [174,316]. However, production of highly reduced short-chain diols from AAs has remained unexplored. Herein, we first devised an enzymatically efficient and thermodynamically favorable Car-based reduction route for ω-HA conversion to diol. Then we adopted three native or heterologous catabolic pathways of major charged AAs (Asp, Glu and Lys) that led to release of C3-C5 ω-HAs. Final combination of the two segments successfully established a new platform capable of biorefining C3-C5 diols from amino acids (**Figure 4.1**). Considering the crosslinking of amino acid degradation network, we were able to convert four additional amino acids into C3-C5 diols using the established platform as entry routes.

Noteworthy, the carboxylic acid reductase exhibited rather high activity towards short-chain ω-HAs, leaving the upstream degradation of AAs as a major limiting step for diol production. Indeed, increasing cofactor supply increased Glu to 1,4 BDO conversion efficiency to 66.0 % of theoretical maximum, while enhancing expression of Lys degradation pathway enabled Lys to 1,5 PDO conversion efficiency to 78.2 % of theoretical maximum. Our results highlighted the potential of a new platform for diol production from inexpensive amino acids.

Taking advantage of the well-developed amino acid pathway in microbial systems could readily lead to sustainable production of diols from cheap sugar feedstocks, which is more economically appealing. We chose to demonstrate and optimize 1,4 BDO and 1,5 PDO production from glucose in *E. coli*. Glutamate is the most abundant metabolite in *E. coli* and synthesized from TCA α-ketoglutarate [340]. Even without host engineering and pathway optimization, 588.3 mg $l^{-1}$ 1,4 BDO could be produced. We endeavored to increase the cofactor PLP supply for glutamate degradation pathway and increase carbon flux to α-ketoglutarate, which collectively enhanced 1,4 BDO titer by 2.40-fold (1.41 g $l^{-1}$) (**Figure 4.5**). Lysine, however, could not accumulate in high concentrations without pathway enhancement. This was in line with the low production of 1,5 PDO titer without lysine feeding or with aspartate feeding in strain 5PDO-1. We herein demonstrated that, with elevated lysine degradation, lysine mediated transcriptional repression or allosteric inhibition of lysine biosynthesis pathway could be significantly relieved, thus driving 1,5 PDO production. With additional host engineering by deleting *iclR* to increase OAA supply, 0.97 g $l^{-1}$ 1,5 PDO was finally produced from glucose (**Figure 4.6**). Therefore, we established and demonstrated a new synthetic pathway for 1,4 BDO and the first synthetic pathway for 1,5 PDO in *E. coli*. Noteworthy, these new diol pathways offer several other potential benefits

including oxygen insensitivity and reduced acetate accumulation, in comparison with the CoA-dependent pathways.

In conclusion, the work outlined here describes the development of a novel bacterial platform for diol production via amino acid based biorefining. The new platform involves new CoA-independent chemistry, and affords the synthesis of C3-C5 diols from seven different amino acids or sugar feedstocks. Due to the low productivity of amino acids in wildtype *E. coli* host, the titers and yields of diols from sugars are still low to satisfy the commercialization benchmark. However, it is expected that additional metabolic engineering manipulations to enhancing amino acid flux will further improve diol production. We envision that this platform can also be readily implemented in other well-developed amino acid hyperproducers like *C. glutamicum*. Thus, we establish and demonstrate a new amino acid based platform for diol production, which shows high industrial potential with feedstock and host flexibility.

## 4.5 Materials and methods

### 4.5.1 Bacterial strains and chemicals

All strains and plasmids used in this study are listed in **Table S4.1**. *E. coli* strain XL1-Blue (Stratagene, La Jolla, CA) was used for plasmids construction and BW25113(F′) was used as production host. pSC74 that was derived from pSA74 by replacing its *pSC101* ori with *CloDF13 ori* from pCDFDuet-1. Plasmids pZE12-luc (high-copy), pCS27 (medium-copy), pSC74 (high-copy) were used for pathway construction and pCP20 was used for elimination of kanamycin resistance marker during gene disruption. Phusion High Fidelity DNA polymerase, restriction endonucleases and Quick Ligation kit were purchased from New England Biolabs (Ipswich, MA). Standard chemicals were purchased from Sigma-Aldrich (St. Louis, Missouri) unless

otherwise specified. Owing to the absence of commercially available standards, 4HB and 5HV standards used in this study were prepared respectively from γ-butyrolactone (GBL) and σ-valerolactone through saponification. Briefly, 664.5 μl GBL and 706.9 μl σ-valerolactone were added into 15 ml tubes and the pH was adjusted by 10 N NaOH to pH 12.0. The tubes were incubated at 37℃ for over 12 h and then the pH was adjusted to pH 7.0 by $H_2SO_4$. Pure water was added to a final volume of 3 ml, yielding a 300 g $l^{-1}$ solution for 4HB and 5HV.

### 4.5.2 Plasmid construction

All DNA manipulations were performed following the standard molecular cloning protocols [286]. Gene products amplified by PCR were digested with appropriate restriction enzymes and ligated into similarly digested vectors. The amplified *panD* from *C. glutamicum*, *Pp0596* from *P. putida* KT2440 and *ydfG* from *E. coli* were digested and then integrated into pCS27 to yield 3HP pathway plasmid p3HP. The mutant *gadB*$^*$(*gadB* from *E. coli* harboring E89Q with truncation of C-terminal 452-466) is obtained by overlap extension PCR. The *gadB*$^*$, along with *gabT* and *yqhD* from *E. coli,* were inserted into pCS27 in one operon, yielding 4HB pathway plasmid p4HB. The *davB* and *davA* from *P. putida* KT2440, *gabT* and *yqhD* from *E. coli*, were amplified and inserted into pCS27 in one operon, yielding 5HV pathway plasmid p5HV1. The *lysC* and *dapA* from *E. coli* were inserted into pCS27 in one operon, yielding pCS-*lysC-dapA*. Replacing *lysC* or *dapA* with corresponding mutant genes created two plasmids, pCS-*lysC* (T253R)-*dapA* (H56K) and pCS-*lysC* (T253R)-*dapA* (E84T). p5HV2 was created by inserting the $P_L$*lacO1-lysC-dapA* cassette into p5HV1. Similarly, inserting $P_L$*lacO1-lysC* (T253R)-*dapA* (H56K) or $P_L$*lacO1-lysC* (T253R)-*dapA* (E84T) cassette into pHV1 created p5HV3 or p5HV4. p5HV5 was created by inserting the *davB-davA-gabT-yqhD* cassette into pSC74. To construct 1,3 PDO and

1,4 BDO pathways, the *$P_L$lacO1-panD-Pp0596-ydfG* and *$P_L$lacO1-gadB*-gabT-yqhD* cassettes were amplified from p3HP and p4HB, and cloned into previously constructed pCar [352], yielding plasmids p3PDO and p4BDO. The *gltA** (*gltA* harboring R163L), obtained from overlap extension PCR, and *ppc* from *E. coli* were constructed into pCS27 in one operon, resulting pCS-*gltA*-ppc*. The *pdxS* and *pdxT* were amplified from genomic DNA of *Bacillus subtilis* 168 and cloned to pCS27 in one operon to generate pCS-*pdxST*. Insertion of the *$P_L$lacO1-pdxST* cassette into pCS-*gltA*-ppc* yielded plasmid pCS-*gltA*-ppc-pdxST*. The *asnA* and *ybaS* from *E. coli* were respectively constructed into pCS27, creating pCS-*asnA* and pCS-*ybaS*. The *putA* from *E. coli* was cloned into pCS27 to obtain pCS-*putA*. The *adiA* and *speB* from *E. coli*, were cloned into pCS27 to form pCS-*adiA-speB*. Similarly, the gene *patA* and *patD* from *E. coli* were cloned into pCS27 to yield pCS-*patAD*. The cassette *$P_L$lacO1-patAD* from pCS-*patAD* was then integrated into *pCS-adiA-speB* to obtain pCS-*adiA-speB-patAD*.

### 4.5.3 Culture media

Luria-Bertani (LB) medium containing 10 g $l^{-1}$ NaCl, 10 g $l^{-1}$ tryptone and 5 g $l^{-1}$ yeast extract was used for plasmid propagation and cell inoculation. The M9Y medium, prepared from M9 minimal medium (6.78 g $l^{-1}$ $Na_2HPO_4$, 0.5 g $l^{-1}$ NaCl, 3 g $l^{-1}$ $KH_2PO_4$, 1 g $l^{-1}$ $NH_4Cl$, 246.5 mg $l^{-1}$ $MgSO_4 \cdot 7H_2O$ and 14.7 mg $l^{-1}$ $CaCl_2 \cdot 2H_2O$) containing 20 g $l^{-1}$ glucose and 5 g $l^{-1}$ yeast extract, was used for feeding experiments and *de novo* production of 1,3-PDO, 1,4-BDO and 1,5-PDO. The antibiotics ampicillin (100 μg $ml^{-1}$), kanamycin (50 μg $ml^{-1}$) and chloramphenicol (34 μg $ml^{-1}$) were added into the medium when needed.

### 4.5.4 Shake flask experiments

Seed cultures of recombinant *E. coli* strains were first inoculated in 3 ml LB medium at 37 ℃ for 8-12 h, and then transferred to 125 ml shake flasks containing 20 ml fresh M9Y medium at an inoculation volume of 2 %. After inoculation at 37 ℃ until the optical density ($OD_{600}$) of cultures reached 0.6-0.8, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at a final concentration of 0.5 mM for protein induction and the shake flask cultures were then inoculated at 30 ℃ on a rotary shaker at 270 rpm. For bioconversion experiments, substrates (3HP, 4HB, 5HV or amino acids) were added at a final concentration of 3 g $l^{-1}$ into the cultures. The cultures were sampled every 12 h or 48 h post inoculation. The samples were subjected to centrifugation for 15 min at 12,000 rpm and filtration by a 0.22 μm filter membrane to remove any cell pellets. The processed samples were assayed by high performance liquid chromatography (HPLC) analysis. Shake flask experiments were performed in triplicates, and data are presented as the averages and s.d. (n=3).

### 4.5.5 *In vitro* enzyme assays

The seed culture of *E. coli* BW25113(F′) containing plasmid pCar was prepared in 3 ml LB medium at 37 ℃ for 8-12 h, and then transferred to 250 ml shake flasks containing 50 ml LB medium at an inoculation volume of 2 %. After inoculation at 37 ℃ until the optical density ($OD_{600}$) of cultures reached 0.6-0.8, 0.5 mM IPTG was added for protein induction and the shake flask cultures were then inoculated at 30 ℃ on a rotary shaker at 270 rpm for 9-12 h. The *E. coli* BW25113(F′) with pZE12-luc was used as the control. The induced cells were harvested by centrifugation at 5000 rpm at 4℃ for 10 min, and then resuspended by 1 ml distilled $H_2O$. The cells were lysed using mini bead beater (Biospec Products, Bartlesville, Oklahoma), and the

crude enzyme extracts were obtained by collecting the supernatants after centrifugation at 10,000 rpm for 10 min. The *in vitro* assay system contains 1 mM $MgCl_2$, 25 mM Tris-HCl (pH 7.5), 4 mM NADPH and 4 mM ATP, and varied concentration of substrates (3HP, 4HB and 5HV). The crude enzyme extracts were added to a final concentration of 0.14 mg $ml^{-1}$. The absorbance at 340 nm of the reaction was measured using a Genesys 10S UV-Vis Spectrophotometer (Thermo Scientific, Waltham, MA). Each assay was performed in triplicates. The changes of NADPH concentration (extinction coefficient 6220 $m^{-1}cm^{-1}$ at 340 nm) was used to determine the specific reaction rates of Car towards hydroxylic acid substrates.

### 4.5.6 HPLC analysis

Glucose, acetate, 3HP, 4HB, 5HV, 1,3 PDO, 1,4 BDO and 1,5 PDO, were quantified using corresponding standard by Dionex Ultimate 3000 (Ultimate 3000 Photodiode Array Detector) with a Coregel-64H column (Transgenomic, Omaha, NE). 4 mM $H_2SO_4$ was used as mobile phase at a flow rate of 0.40 ml $min^{-1}$. The oven temperature was set to 45 ℃. This HPLC method was modified from our previous research [353].

## 4.6 Tables and figures

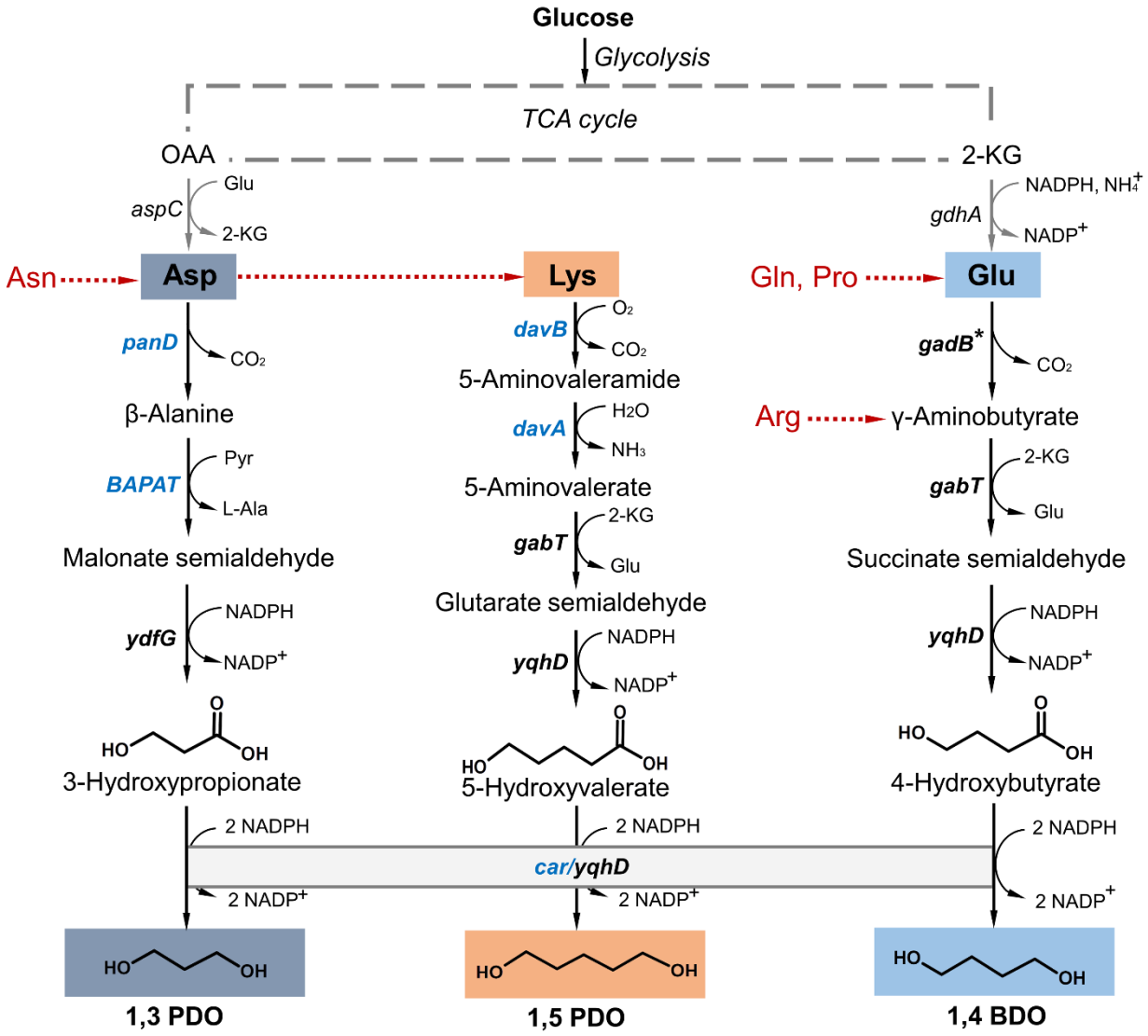


**Figure 4.1 Establishing a diol production platform via extending amino acid catabolism in *E. coli*.** The scheme represents native anabolism and designed catabolism of charged AAs for diol production. Asp, Glu and Lys are biosynthesized from TCA cycle intermediates. Via the synthetic platform, Asp can be converted to 1,3 PDO, Glu converted to 1,4 BDO and Lys converted to 1,5 PDO. Asn, Gln, Pro and Arg can be directed to the diol production platform via corresponding degradation pathways. Endogenous genes are indicated in black and heterologous genes are indicated in blue. *panD*, L-aspartate-α-decarboxylase; *BAPAT*, β-alanine pyruvate transaminase; *ydfG*, malonate semialdehyde reductase; *gadB*$^{*}$, mutant glutamate decarboxylase (GadB E89QΔ452-466); *gabT*, γ-aminobutyrate aminotransferase; *yqhD*, alcohol dehydrogenase; *davB*, lysine 2-monooxygenase; *davA*, 5-aminovaleramidase; *davT*, 5-aminovalerate transaminase; *car*, carboxylic acid reductase.

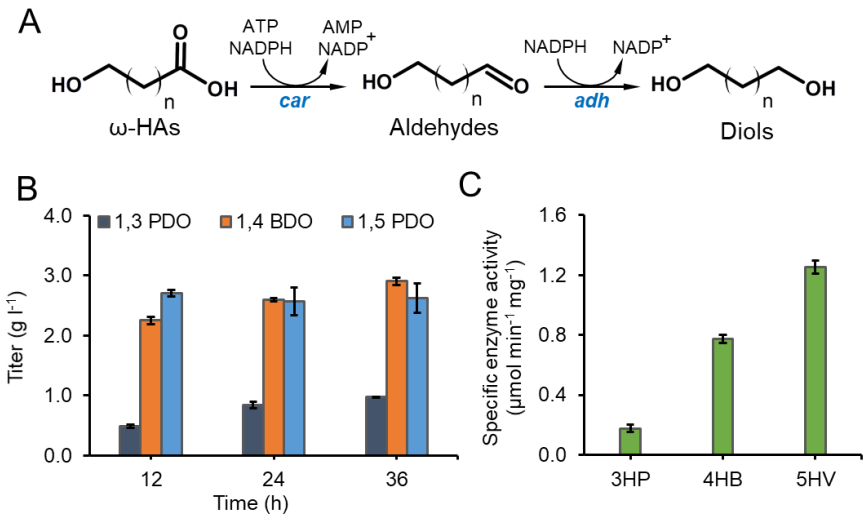



**Figure 4.2 Validation of Car-based diol production.** (A) Car-based biosynthetic route of ω-HAs to diols. (B) Time course conversion of 3 g $l^{-1}$ 3HP, 4HB and 5HV to 1,3 PDO, 1,4 BDO and 1,5 PDO respectively by HA-2 expressing *car* and *sfp*. (C) Enzymatic activity assay of Car towards 3HP, 4HB and 5HV. Error bars represent s.d. (n = 3).

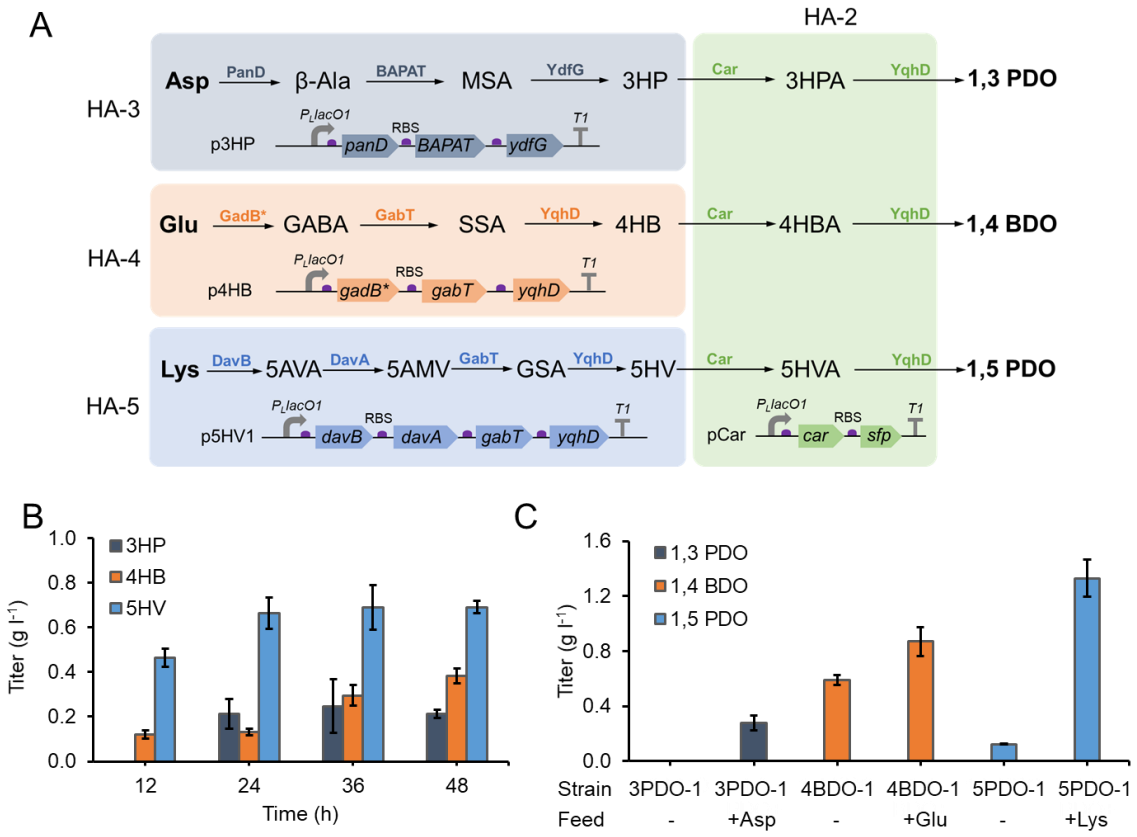



**Figure 4.3 Direct conversion of AAs to diols.** (A) Scheme of the bioconversion experiments via recombinant *E. coli* strains HA-3, -4 and -5. HA-3 (harboring p3HP), -4 (harboring p4HB) and -5 (harboring p5HV1) converted charged AAs to C3-C5 ω-HAs. (B) Time course conversion of 3 g $l^{-1}$ Asp, Glu and Lys to 3HP, 4HB and 5HV by HA-3, -4 and -5, respectively. (C) Bioconversion of 3 g $l^{-1}$ Asp, Glu and Lys to 1,3 PDO, 1,4 BDO and 1,5 PDO by strain 3PDO-1, 4BDO-1 and 5PDO-1, respectively. Error bars represent s.d. (n = 3).

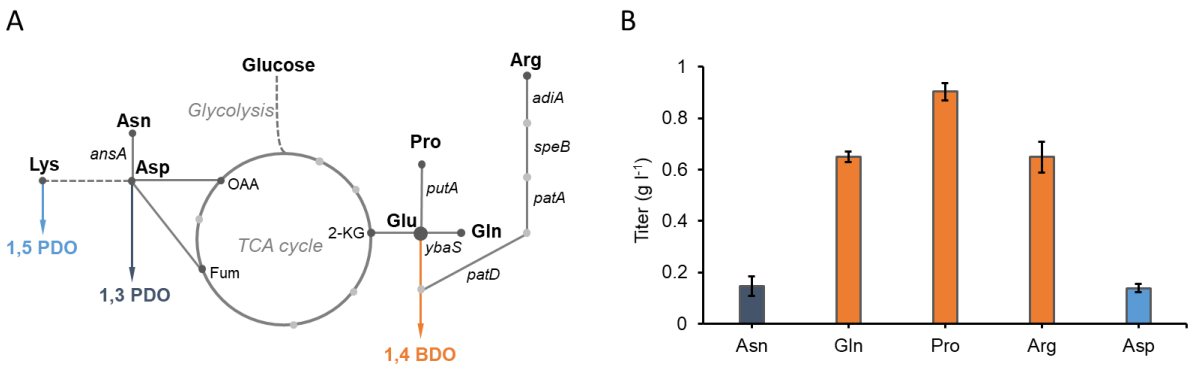



**Figure 4.4 Bioconversion of diol-convertible AAs.** (A) Engineering endogenous amino acid degradation pathways of different AAs to diols. (B) Conversion of diol-convertible AAs including Asn, Gln, Pro, Arg and Asp (3 g $l^{-1}$) to diols. Error bars represent s.d. (n = 3).

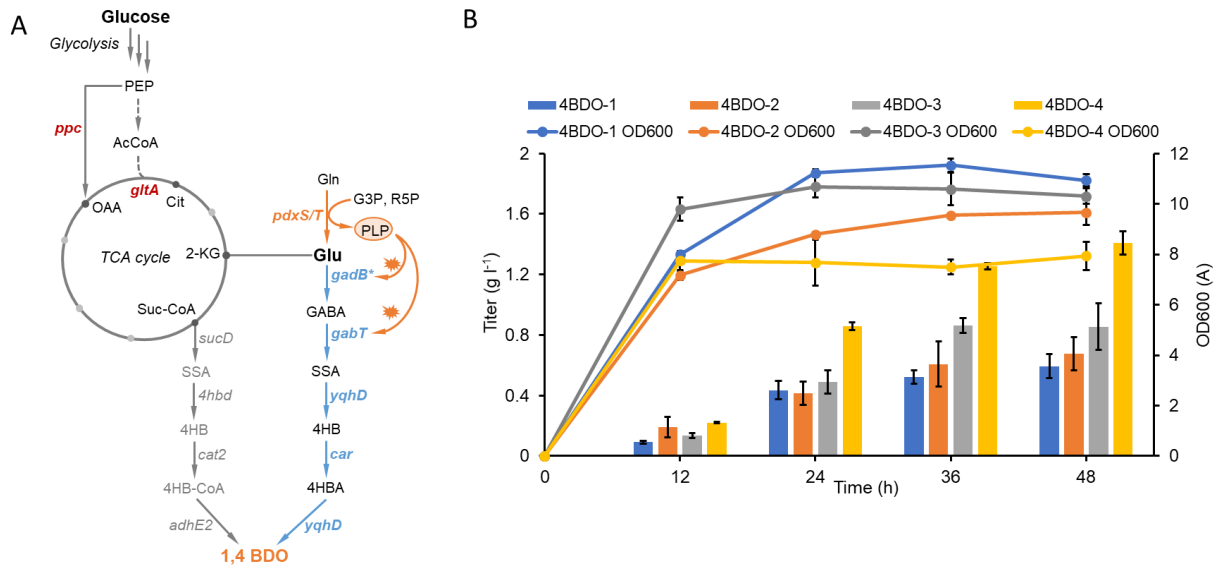



**Figure 4.5 *De novo* production of 1,4 BDO from glucose in engineered *E. coli*.** (A) *De novo* biosynthesis pathway of 1,4 BDO. Enzymes in gray indicated the CoA-dependent reduction pathway, and enzymes in blue indicated the Car-based pathway. PLP is a cofactor for GadB* and GabT. (B) 1,4 BDO production and growth profiles of strain 4BDO-1 to 4BDO-4 when cultivated in M9Y media. Samples were taken every 12 h for 48 h. All data points are reported as mean from three independent experiments. Error bars represent s.d. (n = 3).

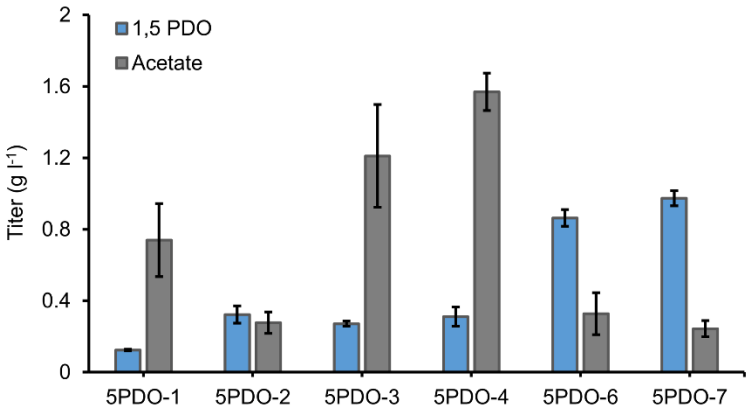



**Figure 4.6 Engineering and optimization of 1,5 PDO production from glucose.** Production performance of different 1,5 PDO producers in M9Y media. All data points are reported as mean from three independent experiments and the peak titers at 48 h were applied for comparison. Error bars represent s.d. (n = 3).

**4.7 Supplementary information**

**Table S4.1 Strains and plasmids used in this study.**

| Strains | Properties | Source |
|---|---|---|
| *E. coli* XL1-Blue | *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 *supE*44 *relA*1 *lac F′* [*traD36 proAB lacI*$^{q}$*ZΔM*15 *Tn*10 (*Tet*$^{r}$)] | Stratagene |
| *E. coli* BW25113 (F′) | *rrnBT*14 Δ*lacZWJ*16 *hsdR*514 *ΔaraBADAH*33 *ΔrhaBADLD*78 *F′* [*traD36 proAB lacI*$^{q}$*ZΔM15 Tn10(Tet*$^{r}$*)*] | [249] |
| *E. coli* BW25113 (F′) Δ*iclR* | *E. coli* BW25113 (F′) with deletion of *iclR* | This study |
| HA-1 | *E. coli* BW25113 (F′) harboring pZE-*cat2-adhE2* | This study |
| HA-2 | *E. coli* BW25113 (F′) harboring pCar | This study |
| HA-3 | *E. coli* BW25113 (F′) harboring p3HP | This study |
| HA-4 | *E. coli* BW25113 (F′) harboring p4HB | This study |
| HA-5 | *E. coli* BW25113 (F′) harboring p5HV1 | This study |
| 3PDO-1 | *E. coli* BW25113 (F′) harboring p3PDO | This study |
| 4BDO-1 | *E. coli* BW25113 (F′) harboring p4BDO | This study |
| 4BDO-2 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*pdxST* | This study |
| 4BDO-3 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*gltA*$^{*}$-*ppc* | This study |
| 4BDO-4 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*gltA*$^{*}$-*ppc-pdxST* | This study |

| | | |
|---|---|---|
| 5PDO-1 | *E. coli* BW25113 (F′) harboring pCar and p5HV1 | This study |
| 5PDO-2 | *E. coli* BW25113 (F′) harboring pCar and p5HV2 | This study |
| 5PDO-3 | *E. coli* BW25113 (F′) harboring pCar and p5HV3 | This study |
| 5PDO-4 | *E. coli* BW25113 (F′) harboring pCar and p5HV4 | This study |
| 5PDO-5 | *E. coli* BW25113 (F′) harboring pCar and p5HV5 | This study |
| 5PDO-6 | *E. coli* BW25113 (F′) harboring pCar, p5HV5 and pCS-*lysC-dapA* | This study |
| 5PDO-7 | *E. coli* BW25113 (F′) Δ*iclR* harboring pCar, p5HV5 and pCS-*lysC-dapA* | This study |
| DO-1 | *E. coli* BW25113 (F′) harboring p3PDO and pCS-*asnA* | This study |
| DO-2 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*ybaS* | This study |
| DO-3 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*putA* | This study |
| DO-4 | *E. coli* BW25113 (F′) harboring p4BDO and pCS-*adiA-speB-patAD* | This study |

| **Plasmids** | **Properties** | **Source** |
|---|---|---|
| pCDFDuet-1 | *PT7, CloDF13 ori*, Spec$^r$ | Novagen |
| pZE12-luc | $P_L$*lacO1*, *colE* ori, Amp$^r$ | [250] |
| pCS27 | $P_L$*lacO1*, *P15A* ori, Kan$^r$ | [251] |
| pSA74 | $P_L$*lacO1*, *pSC101* ori, Cl$^r$ | [52] |

| | | |
|---|---|---|
| pSC74 | *$P_L$lacO1, CloDF13 ori*, $Cl^r$, 20-40 copies; derived from pSA74 by replacing its replication origin with *CloDF13 ori* from pCDFDuet-1 | This study |
| pCar | pZE12-luc harboring *M. marinum* M and *sfp* amplified from *B. subtilis* 168 | [352] |
| pZE-*cat2-adhE2* | pZE12-luc harboring *cat2* from *P. gingivalis* and the codon-optimized *adhE2* from *C. acetobutylicum* | This study |
| p3HP | pCS27 harboring *panD* from *C. glutamicum*, *BAPAT* (*Pp0596*) from *P. putida* and *ydfG* from *E. coli* | This study |
| p4HB | pCS27 carrying *gadB**(*gadB E89Q*Δ452-466), *gabT* and *yqhD* from *E. coli* | This study |
| p5HV1 | pCS27 harboring *davB* and *davA* from *P. putida* KT2440, *gabT* and *yqhD* from *E. coli* | This study |
| p5HV2 | p5HV1 inserted with *$P_L$lacO1-lysC-dapA* cassette | This study |
| p5HV3 | p5HV1 inserted with *$P_L$lacO1-lysC* (T253R)-*dapA* (H56K) cassette | This study |
| p5HV4 | p5HV1 inserted with *$P_L$lacO1-lysC* (T253R)-*dapA* (E84T) cassette | This study |
| p5HV5 | pSC74 harboring *davB* and *davA* from *P. putida* KT2440, *gabT* and *yqhD* from *E. coli* | This study |
| p3PDO | pCar inserted with *$P_L$lacO1-panD-Pp0596-ydfG* cassette from p3HP | This study |

| | | |
|---|---|---|
| p4BDO | pCar inserted with $P_L$*lacO1-gadB*$^*$*-gabT-yqhD* cassette from p4HB | This study |
| pCS-*gltA*$^*$*-ppc* | pCS27 harboring *gltA*$^*$ (*gltA* R163L) and *ppc* from *E. coli* | This study |
| pCS-*pdxST* | pCS27 harboring *pdxS* and *pdsT* from *B. subtilis* 168 | This study |
| pCS-*gltA*$^*$*-ppc-pdxST* | pCS-*gltA*$^*$*-ppc* inserted with $P_L$*lacO1-pdxS-pdxT* cassette from pCS-*pdxS-pdxT* | This study |
| pCS-*asnA* | pCS27 harboring *asnA* from *E. coli* | This study |
| pCS-*ybaS* | pCS27 harboring *ybaS* from *E. coli* | This study |
| pCS-*putA* | pCS27 harboring *putA* from *E. coli* | This study |
| pCS-*adiA-speB-patAD* | pCS27 harboring *adiA, speB, patA and patD* from *E. coli* | This study |
| pCS-*lysC-dapA* | pCS27 harboring *lysC* and *dapA* from *E. coli* | This study |
| pCS-*lysC* (T253R)-*dapA* (H56K) | pCS27 harboring *lysC* (T253R) and *dapA* (H56K) | This study |
| pCS-*lysC* (T253R)-*dapA* (E84T) | pCS27 harboring *lysC* (T253R) and *dapA* (E84T) | This study |

**Table S4.2 Max-min Driving Force analysis of two alternative 4HB reduction pathways.**

| **Reaction name** | **Enzyme** | **Reaction** | **$\Delta_rG'°$** [a] **(kJ/mol)** | **MDF** [b] **(kJ/mol)** |
|---|---|---|---|---|
| CoA-dependent route | Cat2, AdhE2 | 4-hydroxybutyrate + Acetyl-CoA + 2 NADPH <=> 1,4 butanediol + 2 $NADP^+$ + CoA + acetate | *-14.5* ± 7.8 | 28.8 |
| Direct reduction route | Car, YqhD | 4-hydroxybutyrate + ATP + 2 NADPH <=> 1,4 butanediol + AMP + Ppi + 2 $NADP^+$ | *-18.9* ± 7.8 | 38.0 |

[a] $\Delta_rG'°$ is the change in Gibbs free energy for a reaction at a particular pH and ionic strength.

[b] Max-min Driving Force (MDF) is the smallest $-\Delta_rG'$ obtained by a reaction when metabolite concentrations are chosen to make the reaction as favorable as possible.

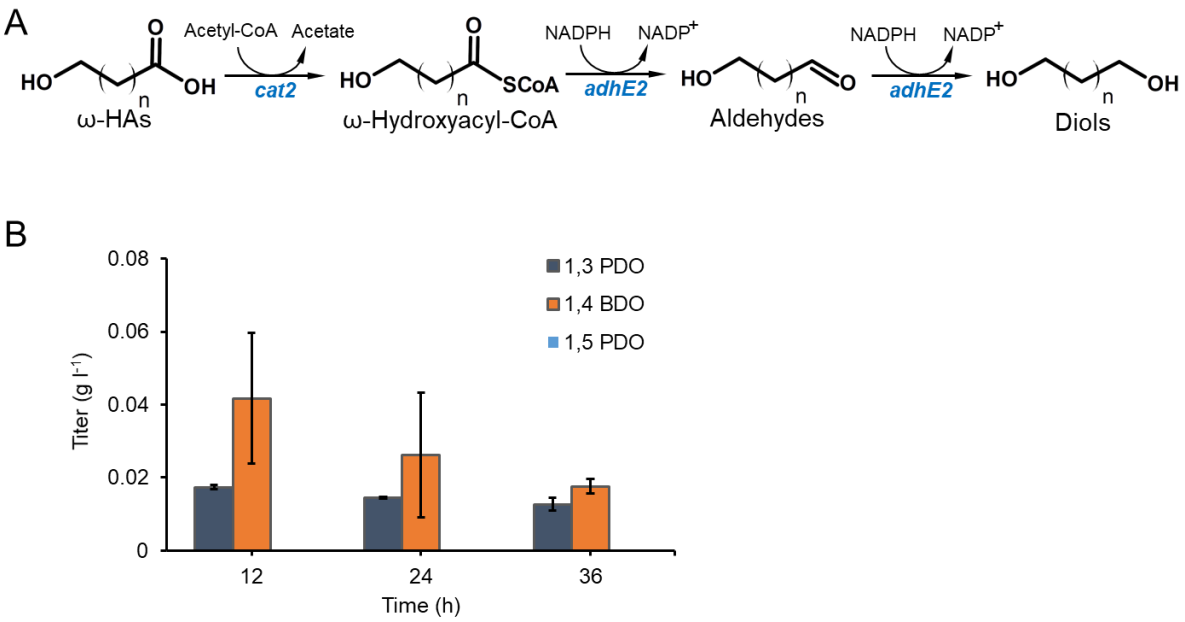



**Figure S4.1 Production of C3-C5 diols by CoA-dependent reduction route.** (A) CoA-dependent biosynthetic route of ω-HAs to diols. (B) Time course conversion of 3 g $l^{-1}$ 3HP, 4HB and 5HV to 1,3 PDO, 1,4 BDO and 1,5 PDO respectively by HA-4 co-expressing *cat2* and *adhE2*. Error bars represent s.d. (n = 3).

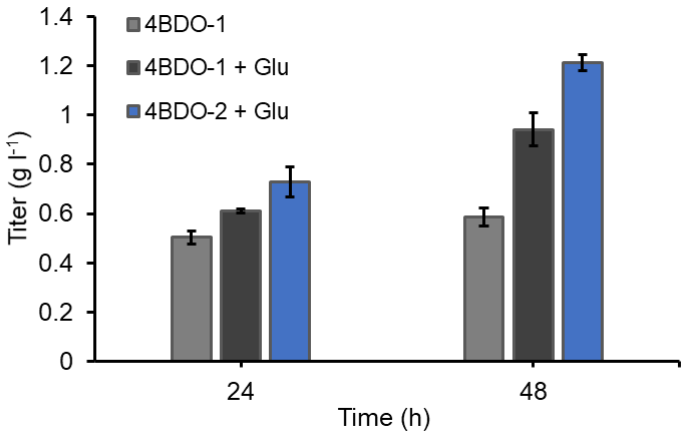



**Figure S4.2 Summary of 1,4 BDO production in bioconversion experiments.** C3-C5 ω-HAs Glu of 3 g $l^{-1}$ was fed into *E. coli* strain 4BDO-1 and 4BDO-2. 1,4 BDO production was detected at 24 and 48 h. Error bars represent s.d. (n = 3).

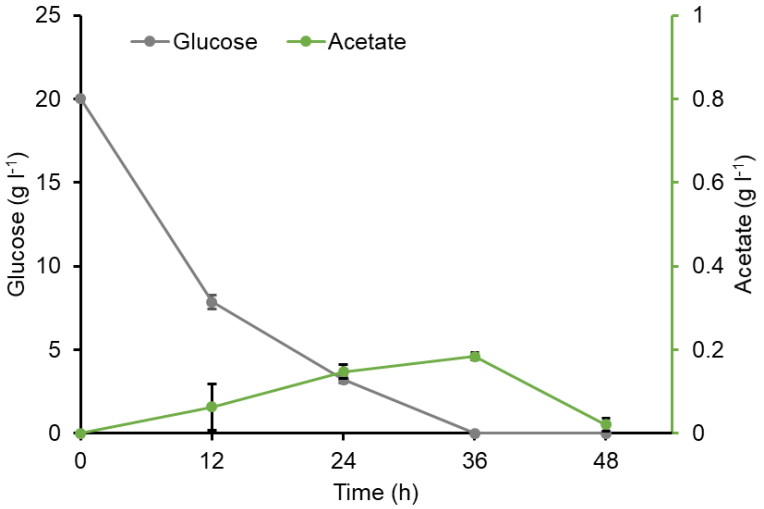



**Figure S4.3 Glucose and acetate concentrations during cultivation of strain 4BDO-4 in M9Y medium.** Strain 4BDO-4 was cultivated in M9Y medium and samples were taken every 12 h for 48 h for HPLC analysis. Error bars represent s.d. (n = 3).

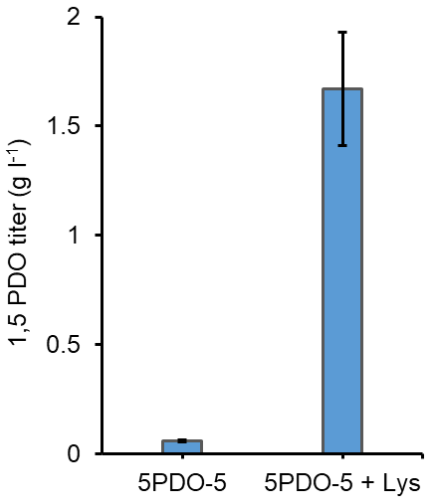



**Figure S4.4 Bioconversion of Lys to 1,5 PDO with strain 5PDO-5.** Strain 5PDO-5 was cultivated in M9Y medium fed with or without 3 g $l^{-1}$ Lys. 1,5 PDO production was detected at 48 h. Error bars represent s.d. (n = 3).

# Chapter 5

# *DE NOVO* BIOSYNTHESIS OF GLUTARATE VIA α-KETO ACID CARBON CHAIN EXTENSION AND DECARBOXYLATION PATHWAY IN *ESCHERICHIA COLI* [1]

[1]Jian Wang, Yifei Wu, Xinxiao Sun, Qipeng Yuan, Yajun Yan. *ACS Synthetic Biology*, 2017. 6 (10), 1922-1930.

Reprinted here with permission of the publisher.

## 5.1 Abstract

Microbial based bioplastics are promising alternatives to petroleum based synthetic plastics due to their renewability and economic feasibility. Glutarate is one of the most potential building blocks for bioplastics. The recent biosynthetic routes for glutarate were mostly based on the L-lysine degradation pathway from *Pseudomonas putida that required lysine either by feeding or lysine overproduction via genetic manipulations*. Herein, we established a novel glutarate biosynthetic pathway by incorporation of a "+1" carbon chain extension pathway from α-ketoglutarate (α-KG) in combination with α-keto acid decarboxylation pathway in *Escherichia coli*. Introduction of homocitrate synthase (HCS), homoaconitase (HA) and homoisocitrate dehydrogenase (HICDH) from *Saccharomyces cerevisiae* into *E. coli* enabled "+1" carbon extension from α-KG to α-ketoadipate (α-KA), which was subsequently converted into glutarate by a promiscuous α-keto acid decarboxylase (KivD) and a succinate semialdehyde dehydrogenase (GabD). The recombinant *E. coli* co-expressing all five genes produced 0.3 g/L glutarate from glucose. To further improve the titers, α-KG was rechanneled into carbon chain extension pathway via the clustered regularly interspersed palindromic repeats system mediated interference (CRISPRi) of essential genes *sucA* and *sucB* in tricarboxylic acid (TCA) cycle. The final strain could produce 0.42 g/L glutarate, which was increased by 40 % compared with the parental strain.

## 5.2 Introduction

Development of sugar-derived novel and sustainable bioplastics provides renewable and green alternatives to petroleum-derived products. Dicarboxylic acids, especially aliphatic dicarboxylic acids (C4-C10) such as succinate (C4), glutarate (C5), adipate (C6), pimelate (C7), suberate (C8), azelate (C9) and sebacate (C10), are important monomer building blocks for production of polymers such as polyurethanes, polyester polyols and polyamides.[354-356] Nylon, one of the most important polymers, is a class of common synthetic polymers produced by copolymerization of various dicarboxylic acids with diamines. Among these aliphatic dicarboxylic acids, succinate has been identified as one of the top 12 sugar-derived and value-added chemicals by the US Department of Energy (DOE) and adipate has been recognized as a super commodity chemical with a remarkable potential market.[357] However, except succinate that could be produced biologically with economically competitive productivity and yield, microbial biosynthesis of the other dicarboxylic acids has either not been established or has been achieved only at very low titers.[354,358] Very recently, dicarboxylic acids with chain length from C6 to C15 were produced in *E. coli* through either CoA-dependent non-decarboxylative Claisen condensation reactions in combination with subsequent β-reduction reactions or an engineered biotin-fatty acid biosynthetic pathway, which however, both suffered from low yield and productivity.[359,360]

Glutarate, or 1,5-pentanedioic acid, is an important C5 dicarboxylic acid building block for polyesters and polyamides like nylon-4,5 and nylon-5,5 (co-polymers of glutarate with putrescine and cadaverine, respectively).[361-364] It has the lowest melting point among all dicarboxylic acids, an interesting property that carries through to polymer derivatives.[365] Glutarate also serves as a precursor for 1,5-pentanediol, a common plasticizer used as a solder flux activator and a pharmaceutical intermediate. The common chemical methods to synthesize

glutarate are through the ring-opening of the butyrolactone or from the reaction of 1,3-dibromopropane with potassium cyanide, or through oxidative cleavage of cyclopentene with ionic liquids.[366,367] Owing to the cost and environmental concern of chemical methods, bio-based production of glutarate is much more desirable and renewable.

Bio-based production of glutarate, as well as its C5 derivatives cadaverine and 5-aminovalerate (5AVA), has been achieved in *Corynebacterium glutamicum* and *E. coli* through introducing a 4-step lysine degradation pathway from *Pseudomonas putida.*[68,361,368,369] Briefly, the so-called aminovalerate (AMV) pathway converts lysine into 5AVA by lysine 2-monooxygenase (DavB) and δ-aminovaleramidase (DavA). Then 5AVA is converted into glutarate by the enzymes aminovalerate aminotransferase (GabT) and succinate semialdehyde dehydrogenase (GabD). However, this degradation pathway involves lysine supply either by feeding or over-production via genetic manipulations, since *E. coli* is not a natural lysine overproducer. Meanwhile, the limited availability of co-factor α-ketoglutarate (α-KG) of aminovalerate aminotransferase (GabT), restricted the productivity of glutarate.[361,368] When 10 g/L lysine and 10 g/L α-KG were fed to recombinant *E. coli* WL3110 strain expressing the *davAB* and *gabTD* genes, 1.7 g/L of glutarate was produced.[361] The other potential glutarate pathway was proposed by bioreduction of glutaconate, an unsaturated C5-dicarboxylic acid that has been produced from α-KG in *E. coli.*[370] However, the low production of glutaconate (350 mg/L) and the lack of efficient enoate reductase presents a great challenge for glutarate production.[355,370,371]

Bypassing the development of lysine overproducers or screening for highly active enoate reductases, herein we engineered a novel and alternative glutarate biosynthetic route via α-keto acid mediated carbon chain extension from α-KG and oxidative decarboxylation in *E. coli*. Specifically, the α-aminoadipate (AAA) pathway from lysine biosynthesis in *Saccharomyces*

*cerevisiae* was harnessed to elongate α-KG into α-ketoadipate (α-KA), which was subsequently converted into glutarate via successive decarboxylation and dehydrogenation by promiscuous KivD and GabD. During this process, one acetyl-CoA was incorporated, with concomitant release of two $CO_2$ and two NADH. Assembling the novel pathway using genes from different microorganisms enabled the establishment of a total biosynthetic route for glutarate from glucose. Further pathway optimization achieved production of 0.42 g/L glutarate by channeling α-KG into glutarate pathway via CRISPRi mediated repression of essential genes in TCA cycle.

## 5.3 Results and discussion

### 5.3.1 Constitution of a novel biosynthetic pathway for glutarate in *E. coli*

The α-keto acid mediated "+1" carbon chain extension pathways are general metabolic pathways that widely exist in nature.[372] The synthetic recursive "+1" carbon chain extension pathway from leucine biosynthesis (LeuABCD) capable of extending the carbon chain from α-keto acids has been extensively recruited to produce carbon chain elongated alcohols and acids.[121,170,373,374] C4 dicarboxylic acid succinate can also be regarded as a product of "+1" carbon chain extension from oxaloacetate and decarboxylation of α-KG in the TCA cycle.[372] Similar to the leucine pathway, the α-keto acid elongation enzymes AksADEF, involved in coenzyme B biosynthesis in methanogenic archaea, have been harnessed to produce C6 dicarboxylic acid adipate in *E. coli*.[355,375,376]

Based upon these previous successes, we designed a novel glutarate biosynthetic pathway by harnessing a "+1" carbon chain extension pathway from α-KG in combination with an α-keto acid decarboxylation pathway: (1) the first half of the pathway is to generate α-KA from condensation of acetyl-CoA and α-KG by a 3-step pathway with homocitrate synthase (HCS),

homoaconitase (HA) and homoisocitrate dehydrogenase (HICDH); (2) the second half of the pathway is to produce glutarate via a successive decarboxylation of α-KA by a promiscuous α-keto acid decarboxylase and dehydrogenation of glutarate semialdehyde by a promiscuous semialdehyde dehydrogenase (**Figure 5.1**). The first half pathway has been well-characterized as portions of lysine biosynthesis pathway, known as the α-aminoadipate pathway, in yeast and fungi.[377,378] Different from AksADEF that could iteratively condense α-KG and acetyl-CoA, HCS-HA-HICDH specifically produced α-KA. Therefore, we recruited the well-characterized HCS (LYS20), HA (LYS4) and HICDH (LYS12) from *S. cerevisiae.*[377] However, the second half of the pathway especially the functional decarboxylase for α-KA has not been well documented.

### 5.3.2 Identification of a functional decarboxylase for glutarate production

To constitute the second half of the pathway, we resorted to a promiscuous α-keto acid decarboxylase KivD from *Lactococcus lactis* and a promiscuous succinate semialdehyde dehydrogenase GabD from *P. putida*. GabD was also known as a glutarate semialdehyde dehydrogenase that has been demonstrated to efficiently convert glutarate semialdehyde to glutarate with $NAD^+$ as the cofactor.[361,379] However, few decarboxylases have been reported to efficiently decarboxylate α-KA. Very recently, KivD and its homologue decarboxylase KdcA (88% identity with KivD) from *Lactococcus lactis* have been demonstrated capable of decarboxylating α-KA into glutarate semialdehyde.[376] To test KivD and GabD for glutarate production, the two enzymes were co-expressed in *E. coli* BW25113 (F') (strain GA1) and α-KA (1 g/L) was fed to strain GA1 for bioconversion. The high-performance liquid chromatography (HPLC) analysis indicated that glutarate was produced with a maximum titer of 0.40 g/L at 48 h

after induction in feeding experiments (**Figure 5.2A**). Glutarate in the cultures was confirmed by ESI-MS analysis (**Figure 5.2B**, upper panel). The presence of ion peak $[M-H]^-$ at m/z 130.9 corresponded to glutarate product ion, which was consistent with standard glutarate (**Figure 5.2B**, lower panel). Ions at m/z 113.0 and m/z 87.1 were characteristic fragment ions caused by in-source collision induced dissociation (CID) and respectively corresponded to the loss of one water molecule and one carboxyl group. The feeding experiments demonstrated that KivD decarboxylated α-KA and co-expression of KivD and GabD enabled conversion of α-KA into glutarate with a conversion ratio of 40 %. Considering the high activity of GabD,[368] we reasoned that KivD exerted limited activity towards α-KA, which might be rate-limiting for glutarate production. This was consistent with that its homologue KdcA that had a substrate conversion ratio of 51 % towards α-KA.[376]

Next, KivD and GabD were respectively over-expressed in *E. coli* BL21 Star (DE3) and purified to homogeneity for a coupled enzyme assay. KivD and GabD catalyzed a coupled reaction from α-KA to glutarate in the presence of cofactor thiamine diphosphate (ThDP), concomitant with the reduction of $NAD^+$ into NADH. The *in vitro* enzyme assay demonstrated that KivD was functional towards α-KA with a $K_m$ value of 3.12 mM and a $k_{cat}$ of 52.98 $min^{-1}$. Interestingly, KivD showed no detectable activity towards α-KG, which was in line with that its homologue KdcA also showed almost no activity towards α-KG.[380] These results supported that KivD showed relatively higher specificity and activity towards long chain α-keto acids as indicated previously.[381]

Interestingly, 40 mg/L glutarate was produced by *E. coli* BW25113 (F') at 48 h after induction when fed with α-KA (0.5 g/L) (**Figure 5.2C**), indicating the presence of functional endogenous decarboxylase(s) and promiscuous aldehyde dehydrogenase(s). The presence of promiscuous

aldehyde dehydrogenase(s) was not surprising, considering the prevalence of aldehyde dehydrogenases in *E. coli*.[382] 2-Oxoglutarate dehydrogenase multi-enzyme complex (OGDHC) catalyzes the conversion of α-KG to succinate and $CO_2$, during which SucA is responsible for the first-step 2-oxoglutarate decarboxylase activity. Considering the structural similarity of α-KA and α-KG, we hypothesized that OGDHC especially SucA might be involved in decarboxylation of α-KA. To testify the hypothesis, *sucA* was deleted in *E. coli* BW25113 (F') and α-KA (0.5 g/L) was fed to the engineered *E. coli* BW25113 (F') Δ*sucA* mutant for shake flask experiments. No production of glutarate was observed in *sucA* knockout strain, indicating that OGDHC could contribute to decarboxylation of α-KA into glutarate, albeit in relatively low activity compared with KivD (**Figure 5.2C**). Additionally, the growth profile showed that deletion of *sucA* significantly retarded cell growth.

### 5.3.3 *De novo* production of glutarate in *E. coli*

To establish the entire glutarate pathway, the upstream pathway genes were constructed into the high-copy number plasmid pZE12-luc and the downstream pathway genes were introduced into the medium-copy number plasmid pCS27, generating pZE-HCS-HA-HICDH and pCS-*kivD*-*gabD*. *E. coli* BW25113 (F') was transformed with these two plasmids, resulting in production strain GA2. Shake flask experiments with strain GA2 were performed in M9 medium supplemented with 20 g/L glucose and 5 g/L yeast extracts. HPLC analysis indicated that glutarate was successfully produced with a maximum titer of 0.3 g/L at 48 h after induction (**Figure 5.3**). This result experimentally demonstrated that assembly of the proposed pathway permitted the successful *de novo* production of glutarate in *E. coli*.

α-KG is an important metabolite in TCA cycle that also serves as a central carbon intermediate.[383] α-KG might be mostly channeled into TCA cycle in aerobic conditions. Indeed, α-KG availability in *E. coli* limited the production of other α-KG derived compounds including 4-hydroxybutyric acid (GHB), 1, 4-butanediol (1, 4-BDO), adipate and 6-aminocaproic acid (6-ACA).[128,376,384] As α-KG is the starting point for glutarate biosynthesis, we speculated that increase of α-KG availability might enhance glutarate titers. When feeding 1 g/L α-KG to strain GA2, glutarate titer was increased to 0.53 g/L, a 1.77-fold increase compared with the non-feeding control (**Figure 5.3**). However, increase of α-KG supply to 4 g/L did not significantly increase glutarate titer any further (0.6 g/L). These results indicated that increasing precursor pool of α-KG would enhance glutarate production, whereas the availability of α-KG is not the main rate-limiting factor when increased to certain level. These results also suggested that KivD was still the major rate-limiting enzyme in the glutarate pathway and excessively supplied α-KG could not be efficiently channeled to the end product glutarate.

### 5.3.4 Constructing CRISPRi system for gene repression

Since TCA is essential for cell viability, genetic disruption of TCA cycle to accumulate α-KG might lead to growth retardation or even cell death[385], as indicated above when knocking out *sucA*. Thus, we attempted to construct a CRISPRi based repression system to interfere TCA cycle by repressing rather than deleting certain essential genes. To create a functional CRISPRi system, we first constructed a dual plasmid system by inserting dCas9 into the low-copy number plasmid pSA74 and inserting sgRNAs into the medium-copy number plasmid pCS27. For construction of sgRNA plasmids, each synthesized sgRNA was inserted into pCS27 between *Xho*I and *Nde*I under control of $P_L$*lacO1* promoter and T1 terminator. Since *Sal*I and *Xho*I are a

pair of isocaudarners, sgRNAs could be iteratively incorporated into *Sal*I and *Bam*HI with independent promoters and terminators after digestion with *Xho*I and *Bam*HI (**Figure 5.4B**).

Next, to test the efficiency of such a CRISPRi system, a reporter strain GA0 was constructed by integrating $P_L$*lacO1-egfp* expression cassette into *E. coli* BW25113 (F') between *nupG* and *speC* loci. We designed sgRNAs targeting four locations on the non-template DNA strand of *egfp* reporter gene: promoter region at -35 box (sgRNA-*egfp*-promoter), ribosome binding site (sgRNA-*egfp*-RBS), start codon (sgRNA-*egfp*-start codon) and coding sequence (sgRNA-*egfp*-CDS) (**Figure 5.4A**). Interference efficiency was compared among multiple targeting positions, considering the sgRNA targeting was limited by the availability of PAM sequence (5'-NGG-3'). GA0 was transformed with pSA-dCas9 and pCS-sgRNAs and induced for fluorescence detection. The results showed that eGFP fluorescence was significantly reduced in all four target repression locations, among which sgRNAs targeting RBS and start codon displayed the highest repression efficiency (**Figure 5.3C**). This indicated that CRISPRi system could efficiently interfere gene expression level by either interfering transcriptional initiation or elongation.

### 5.3.5 Production enhancement of glutarate via CRISPRi

To channel α-KG into the glutarate pathway, CRISPRi mediated repression system was incorporated to repress *sucA* and *sucB*, whose products are responsible for converting α-KG into succinyl-CoA in the TCA cycle. Then sgRNAs targeting start codons of *sucA* and *sucB* were generated and assembled. The sgRNAs were constructed into the downstream pathway plasmid pCS-*kivD-gabD*. *E. coli* BW25113 (F') was transformed with three plasmids pZE-HCS-HA-HICDH, pCS-*kivD-gabD* containing sgRNAs and pSA-dCas9, resulting in strains GA3 and GA4. Shake flask cultivation of GA3 and GA4 were performed in M9 medium containing 20 g/L

glucose and 5 g/L yeast extracts and glutarate production was detected by HPLC analysis. The results showed that 0.27 g/L glutarate was produced by strain GA3 repressing *sucA*, while 0.42 g/L produced by strain GA4 repressing both *sucA* and *sucB*, a 40 % increase compared with that without repression (**Figure 5.5**). Notably, 11.6 g/L glucose was left unutilized by strain GA4, indicating a yield on glucose of 0.068 mol/mol (10 % of the theoretical yield 0.67 mol/mol). As acetyl-CoA was the other precursor for condensation with α-KG, we obtained a triple knockout strain *E. coli* BW25113 (F') Δ*poxB*Δ*ackA*Δ*pta* that would significantly reduce acetate accumulation. Unexpectedly, when transformed with the pathway plasmids containing sgRNAs targeting *sucA* and *sucB* (strain GA5), glutarate production was reduced to 0.19 g/L (**Figure 5.5**). We speculated that the growth retardation of strain GA5 might account for the decreased production of glutarate. Thus, simultaneous interference of *sucA* and *sucB* via CRISPRi mediated repression enabled glutarate production enhanced by 40 % in wild type strain, which however, is still far from industrial scale. To further increase α-KG supply, metabolic engineering using the glutamate overproducer *Corynebacterium glutamicum* as an alternative host or engineering the nonphosphorylative metabolism of C5 sugars in *E. coli* would be alternative choices for enhancing glutarate production.[386,387]

## 5.4 Conclusion

Dicarboxylic acids, especially linear aliphatic dicarboxylic acids, have important applications as monomers for sustainable bioplastics. Herein, we designed a novel pathway for glutarate by harnessing a "+1" carbon chain extension from α-KG combined with subsequent oxidative decarboxylation pathway and experimentally verified the production of glutarate in *E. coli*. This novel pathway for glutarate is favored by incorporation of one acetyl-CoA and release of two

$CO_2$ and two NADH. The theoretical yield of the novel glutarate pathway (0.67 mol/mol glucose) is slightly lower than that of the lysine-mediated glutarate pathway (0.75 mol/mol glucose), but it involves less enzymatic steps and reached a comparable practical yield (0.068 mol/mol).[368] Interestingly, the HCS-HA-HICDH mediated specific production of α-KA and substrate specificity of KivD towards α-KA channeled carbon flux from TCA cycle to glutarate. The titer of glutarate could be augmented either by increasing α-KG supply or improving the catalytic activity of the rate-limiting enzyme KivD. When CRISPRi was implemented to increase α-KG availability by repressing TCA cycle genes, the final engineered strain produced 0.42 g/L glutarate from glucose, which increased glutarate production by 40 %. The rate-limiting KivD has been extensively engineered to accommodate various α-keto acids with different chain-length via structure-based engineering of substrate binding pocket.[170,381,387] Future improvements of glutarate production might be made by increasing the catalytic turnover of KivD via rational protein engineering. Thus, this work established an alternative pathway that might be further manipulated for bioproduction of glutarate and its derived C5 chemicals in future.

## 5.5 Materials and methods

### 5.5.1 Bacteria strains, plasmids and chemical reagents

Bacterial strains and plasmids used in this study were listed in Table 5. 1. *E. coli* strain XL1-Blue was used as the host for plasmid construction, strain BL21 (DE3) for protein expression and strain BW25113 (F') for glutarate production. *E. coli* knockout strains including Δ*sucA* and Δ*poxB*Δ*ackA*Δ*pta* were generated from *E. coli* BW25113 (F') via P1 phage transduction method according to standard protocols.[206] Reporter strain GA0 was created by integrating $P_L$*lacO1-egfp* expression cassette into chromosome of *E. coli* BW25113 (F') between *nupG* and *speC* loci

using λ Red-mediated recombineering technology.[388] Plasmids pZE12-luc (high-copy number), pCS27 (medium-copy number) and pSA74 (low-copy number) were used for pathway construction. pETDuet-1 was used for gene expression and purification. Standard chemicals including glutarate, α-KG and α-KA were purchased from Sigma-Aldrich unless otherwise specified.

### 5.5.2 Culture media and conditions

Luria-Bertani (LB) medium containing 10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl was used for inoculant preparation and cell propagation. M9 minimal medium containing 20 g/L glucose, 5 g/L yeast extract, 1 g/L $NH_4Cl$, 6 g/L $Na_2HPO_4$, 3 g/L $KH_2PO_4$, 0.5 g/L NaCl, 1mM $MgSO_4$, 0.1mM $CaCl_2$, and 1.0 mg/L vitamin B1 was used for shake flask experiments. Ampicillin, kanamycin, and chloramphenicol were added to the medium when necessary at final concentrations of 100, 50, and 34 μg/mL, respectively.

### 5.5.3 Plasmid construction

All genes were amplified by polymerase chain reaction (PCR) using Phusion High-Fidelity DNA polymerase (New England Biolabs). Plasmid construction and DNA manipulation were performed following standard molecular cloning protocols.[207] Briefly, pETDuet-*kivD* and pETDuet-*gabD* were constructed by respectively amplifying and inserting *kivD* from *Lactococcus lactis* and *gabD* from *P. putida* into pETDuet-1 between *Bam*HI and *Sal*I. pZE-*kivD-gabD* was constructed by inserting the two genes into pZE12 between *Acc65*I and *Xba*I in a single operon, and pCS-*kivD-gabD* was constructed by amplifying *kivD-gabD* from pZE-*kivD-gabD* and inserting into pCS27 between *Spe*I and *Sac*I. pZE-HCS-HA-HICDH was constructed

by amplifying and inserting HCS, HA, and HICDH from *S. cerevisiae* into pZE12 between *Acc65*I and *Xba*I in a single operon. The pCS-sgRNA scaffold plasmid was constructed by linking $P_L$*lacO1* promoter and a synthesized sgRNA through overlapping PCR and then inserting it into pCS27 between *Xho*I and *Nde*I. The other pCS-sgRNAs targeting different positions were constructed by mutating the 20 bp target-specific sequence in the scaffold plasmid. The sgRNA array plasmids were constructed by digesting one sgRNA with *Xho*I and *Bam*HI and inserting into a second sgRNA plasmid between *Sal*I and *Bam*HI. pSA-dCas9 was constructed by creating a dCas9 gene (containing D10A and H841A mutations) via overlapping PCR and inserting it into pSA74 between *Acc65*I and *Pst*I.

### 5.5.4 Protein purification and *in vitro* enzyme assay

*E. coli* BL21 Star (DE3) was transformed with pETDuet-1 derived expression plasmids. Single colonies were picked for inoculation in 3 mL LB tubes at 37 °C on a rotary shaker. Overnight inoculants were transferred into 50 mL LB medium and induced with 0.5 mM β-D-1-thiogalactopyranoside (IPTG) when the optical density ($OD_{600}$) of cultures reached 0.6-0.8 and incubated at 30 °C on a rotary shaker for another 9-12 h. The cell cultures were harvested by centrifugation and protein purification were performed using His-Spin Protein Miniprep Kit (Zymo Research, Irvine, CA) according to manufactures' instructions. The purified protein was verified by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using 12% protein gel and the protein concentration was measured using a BCA Protein Assay Kit (Pierce) as manufactures' instructions.

The decarboxylation activity of KivD was measured at 30 °C, using a coupled enzymatic assay method.[170] Excess GabD was used to oxidize aldehyde into acid, and concomitantly, cofactor

$NAD^+$ was reduced to NADH. The assay mixture contained 0.25 mM $NAD^+$, 1.13 μM GabD, 0.85 μM KivD and 0-8 mM α-KA in assay buffer (50 mM potassium phosphate buffer, pH 6.8, 1 mM $MgSO_4$, 0.5 mM ThDP) with a total volume of 1 mL. The reactions were started by adding the substrate α-KA, and the formation of NADH was monitored at 340 nm (extinction coefficient, 6.22 $mM^{-1}cm^{-1}$). Kinetic parameters ($k_{cat}$ and $K_m$) were determined by fitting initial velocity data to the Michaelis-Menten equation, using Origin software.

### 5.5.5 Feeding experiments

Feeding experiments were conducted to examine the pathways for glutarate production. First, *E. coli* BW25113 (F') or its derived transformants were inoculated in 3 mL LB medium and grown for 8 h at 37 °C. Subsequently, 0.4 mL of the pre-inoculum was transferred to 20 mL fresh M9 medium and grown at 37 °C with shaking (270 rpm) for 3 h and then transferred into shakers at 30 °C for cultivation. Shake flask experiments were carried out in a rotary shaker (New Brunswick). For feeding experiments, 1 or 0.5 g/L α-KA or 1 or 4 g/L α-KG and 0.5 mM IPTG if needed were added when transferred to 30 °C for cultivation. 1 mL samples were harvested every 12 h after induction until 48 h. The $OD_{600}$ values were measured and glutarate in the supernatants was subjected to HPLC analysis. Triplicate transformants were used for feeding experiments.

### 5.5.6 *De novo* production of glutarate

To test *de novo* production of glutarate in *E. coli*, shake flask experiments were conducted in M9 medium using transformants of *E. coli* BW25113 (F') or its derivatives with different plasmids. Production strains were inoculated in 3 mL LB medium and grown at 37 °C for 8 h.

Subsequently, 0.4 mL of the pre-inoculum was re-inoculated into 20 mL fresh M9 medium and grown at 37 °C with shaking (270 rpm) for 3 h. Then 0.5 mM IPTG was added into the cultures and cultivation was performed at 30 °C for 48 h. For all the above shake flask experiments, samples were taken every 12 h after induction and glutarate production was analyzed by HPLC analysis.

### 5.5.7 HPLC and MS analysis

The analysis of the samples from shake flask experiments was performed by HPLC (Shimadzu) equipped with a Coregel-64H column (Transgenomic). Samples (1 mL) were centrifuged at 15,000 rpm for 10 minutes. The supernatants were filtered through 0.22 μm film and used for analysis. The mobile phase used was 4 mN $H_2SO_4$ with a flow rate of 0.6 mL/min. The oven temperature set at 55 °C. This HPLC method was modified from previous research.[389] Glutarate produced in the supernatants was detected by electrospray ionization (ESI)-mass spectrometry (MS) on a Thermo Scientific Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Glutarate standard was analyzed as a positive control.

### 5.5.8 Fluorescence assay

The analysis of CRISPRi based interference efficiency was performed by measuring the fluorescence intensity using BioTek micro-plate reader according to previously established methods with modifications.[156] Briefly, single colonies of GA0 transformed with empty pCS27 and pSA74 plasmids or pCS27 derived sgRNA plasmids and pSA-dCas9 were cultivated in 3 mL LB medium with appropriate antibiotics at 37 °C for 5 h. Then 10 μl inoculums were transferred into wells of a black 96-well plate containing 200 μl of LB media and 0.5 mM IPTG. The plate

was kept at 37 °C with continuous shaking and fluorescence intensity reading every 20 min. The eGFP fluorescence intensity was measured using excitation filter of 485 nm, and emission filter of 528 nm. Mean values were obtained from triplicate experiments.

## 5.6 Tables and figures

**Table 5.1 Strains and plasmids used in this study.**

| Strains | Properties | Source |
|---|---|---|
| *E. coli* XL1-Blue | *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 *supE*44 *relA*1 *lac F′* [*traD36 proAB lacI*$^{q}$Z$\Delta$*M*15 *Tn*10 (*Tet*$^{r}$)] | Stratagene |
| *E. coli* BL21 Star (DE3) | F$^{-}$ *ompT hsdSB (rB-mB-) gal dcm rne131* (DE3) | Invitrogen |
| *E. coli* BW25113 (F') | *rrnBT*14 $\Delta$*lacZWJ*16 *hsdR*514 *ΔaraBADAH*33 $\Delta$*rhaBADLD*78 *F′* [*traD36 proAB lacI*$^{q}$Z$\Delta$*M15 Tn10(Tet*$^{r}$*)*] | [249] |
| *E. coli* BW25113 (F') $\Delta$*sucA* | *E. coli* BW25113 (F') with *sucA deletion* | This study |
| *E. coli* BW25113 (F') $\Delta$*poxB*$\Delta$*ackA*$\Delta$*pta* | *E. coli* BW25113 (F') with *poxB, ackA* and *pta deletion* | This study |
| GA0 | *E. coli* BW25113 (F') with $P_L$*lacO1* promoter::*egfp* fusion chromosomally integrated between *nupG* and *speC* loci | This study |
| GA1 | *E. coli* BW25113 (F') transformed with pCS-*kivD-gabD* | This study |
| GA2 | *E. coli* BW25113 (F') transformed with pZE-HCS-HA-HICDH and pCS-*kivD-gabD* | This study |
| GA3 | *E. coli* BW25113 (F') transformed with pZE-HCS-HA-HICDH, pCS-*kivD-gabD*- | This study |

| | | |
|---|---|---|
| | sgRNA-*sucA* and pSA-*dCas9* | |
| GA4 | *E. coli* BW25113 (F') transformed with pZE-HCS-HA-HICDH, pCS-*kivD-gabD*-sgRNA-*sucA-sucB* and pSA-*dCas9* | This study |
| GA5 | *E. coli* BW25113 (F') Δ*poxB*Δ*ackA*Δ*pta* transformed with pZE-HCS-HA-HICDH, pCS-*kivD-gabD*-sgRNA-*sucA-sucB* and pSA-*dCas9* | This study |

| **Plasmids** | **Properties** | **Source** |
|---|---|---|
| pZE12-luc | $P_L$*lacO1*, *colE* ori, $Amp^r$ | [250] |
| pCS27 | $P_L$*lacO1*, *P15A* ori, $Kan^r$ | [251] |
| pSA74 | $P_L$*lacO1*, *pSC101* ori, $Cm^r$ | [252] |
| pETDuet-1 | T7 promoter, *pBR322* ori, $Amp^r$ | Novagen |
| pETDuet-*kivD* | pETDuet-1 harboring *kivD* from *Lactococcus lactis* | This study |
| pETDuet-*gabD* | pETDuet-1 harboring *gabD* from *Pseudomonas putida* KT2440 | This study |
| pZE-HCS-HA-HICDH | pZE12-luc harboring HCS, HA and HICDH from *Saccharomyces cerevisiae* Sc288 | This study |
| pZE-*kivD-gabD* | pZE12-luc harboring *kivD* and *gabD* | This study |
| pCS-*kivD-gabD* | pCS27 harboring *kivD* and *gabD* | This study |
| pSA-*dCas9* | pSA74 harboring *cas9* from *Streptococcus pyogenes* with mutation D10A/H841A | This study |

| | | |
|---|---|---|
| pCS-sgRNA-*egfp*-promoter | pCS27 harboring sgRNA targeting promoter region (at -35 box) of *egfp* gene | This study |
| pCS-sgRNA-*egfp*-RBS | pCS27 harboring sgRNA targeting ribosome binding site (RBS) of *egfp* gene | This study |
| pCS-sgRNA-*egfp*-start codon | pCS27 harboring sgRNA targeting start codon of *egfp* gene | This study |
| pCS-sgRNA-*egfp*-CDS | pCS27 harboring sgRNA targeting the middle region of the *egfp* coding sequence | This study |
| pCS-*kivD-gabD*-sgRNA-*sucA* | pCS-*kivD-gabD* harboring sgRNA targeting *sucA* gene | This study |
| pCS-*kivD-gabD*-sgRNA-*sucA-sucB* | pCS-*kivD-gabD* harboring sgRNA targeting *sucA* and *sucB* gene | This study |

**Table 5.2 Kinetic parameters of KivD.**

| **Substrate** | $K_m$ **(mM)** | $k_{cat}$ **(min$^{-1}$)** | $k_{cat}/K_m$ **(mM$^{-1}$min$^{-1}$)** |
|---|---|---|---|
| α-KA | 3.12 ± 1.26 | 52.98 ± 9.38 | 16.98 |
| α-KG | ND[a] | ND | ND |

[a] ND, no detection.

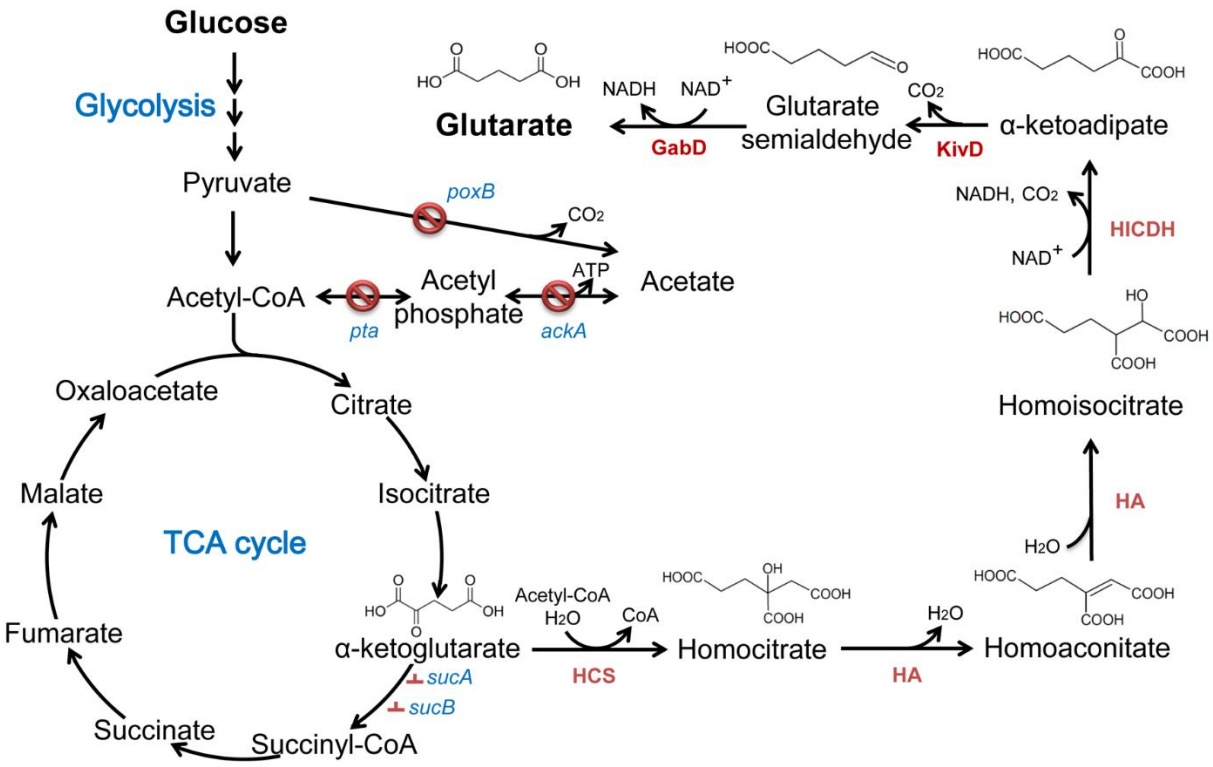



**Figure 5.1** ***De novo*** **biosynthetic pathway of glutarate in** ***E. coli*****.** The established glutarate pathway is initiated from α-KG by condensation with acetyl-CoA to form homocitrate by HCS. Homocitrate is then dehydrated and subsequently re-hydrated into homoisocitrate by HA. HICDH converts homoisocitrate into α-ketoadipate and $CO_2$ with the concomitant reduction of $NAD^+$. The α-ketoadipate is successively converted into glutarate via decarboxylation and dehydrogenation catalyzed by KivD and GabD. Native pathway genes of *E. coli* are shown in blue while heterologous enzymes shown in red. Acetate formation related genes including *poxB*, *pta* and *ackA* were knocking out while *sucA* and *sucB* were repressed in specific conditions in this study. HCS, homocitrate synthase from *S. cerevisiae*; HA, homoaconitase from *S. cerevisiae*; HICDH, homoisocitrate dehydrogenase from *S. cerevisiae*; KivD, α-keto acid decarboxylase from *L. lactis*; GabD, succinate semialdehyde dehydrogenase from *P. putida*.

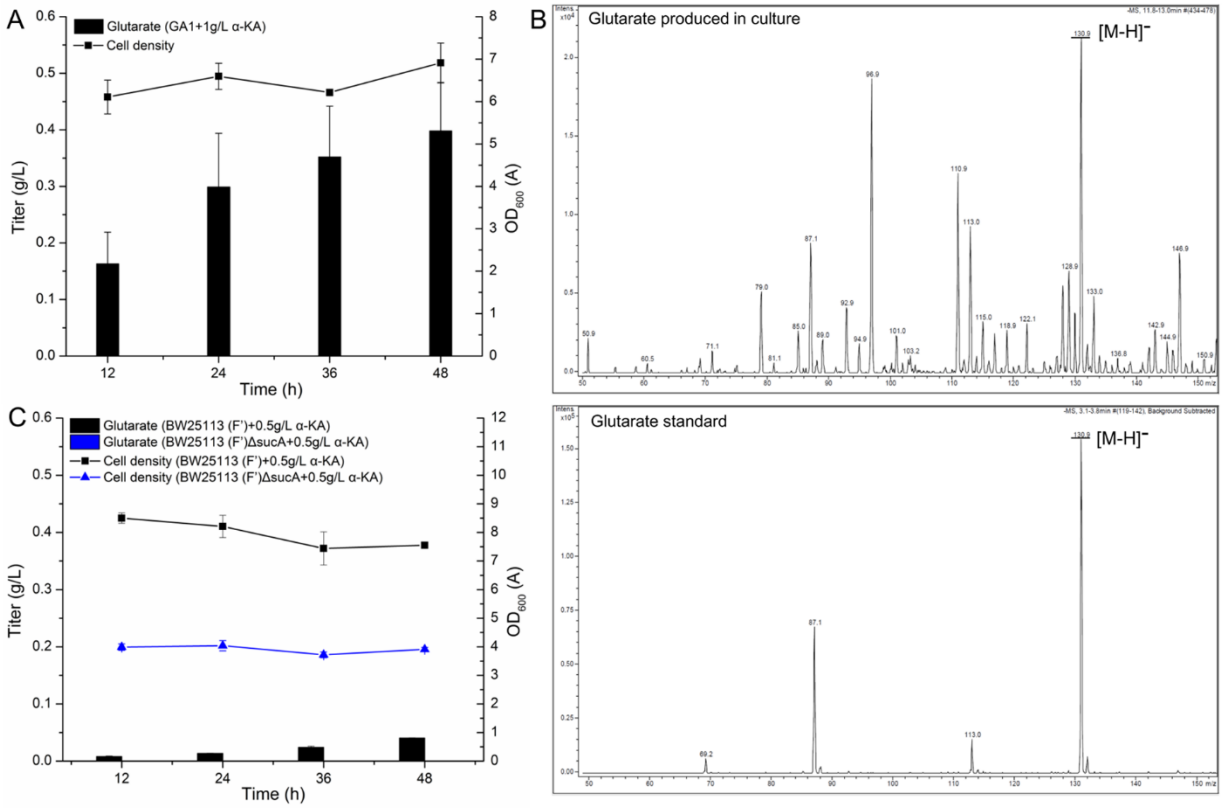



**Figure 5.2 Glutarate production by feeding α-KA in *E. coli*.** (A) α-KA (1 g/L) was fed into *E. coli* BW25113 (F') co-expressing KivD and GabD (strain GA1) for bioconversion. (B) ESI-MS analysis of glutarate produced in the culture medium from bioconversion experiments (upper panel) and glutarate standard (lower panel). (C) α-KA (0.5 g/L) was directly fed into *E. coli* BW25113 (F') (black) and its mutant Δ*sucA* (blue) for glutarate production. The cell density ($OD_{600}$) of *E. coli* BW25113 (F') (black square) or its mutant Δ*sucA* (blue triangle) and glutarate production were measured every 12 h after induction. The error bars represent the standard deviations of three independent experiments.

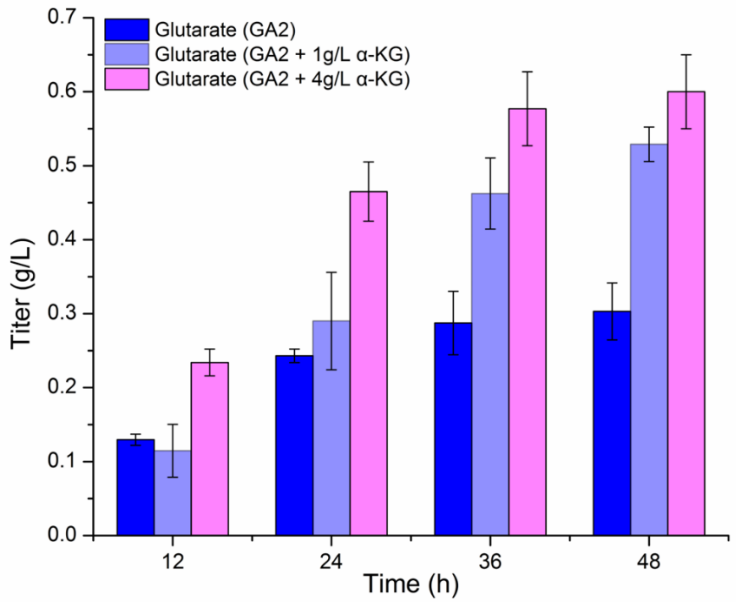



**Figure 5.3 *De novo* production of glutarate in *E. coli*.** *E. coli* BW25113 (F') was transformed with the pathway plasmids pZE-HCS-HA-HICDH and pCS-*kivD-gabD* (strain GA2) for glutarate production in M9 medium containing 20 g/L glucose and 5 g/L yeast extracts (blue columns). 1 g/L (light blue columns) or 4 g/L (pink columns) was added to enhance precursor supply. Samples were taken at four time points every 12 h after induction. The error bars represent the standard deviations of three independent experiments.

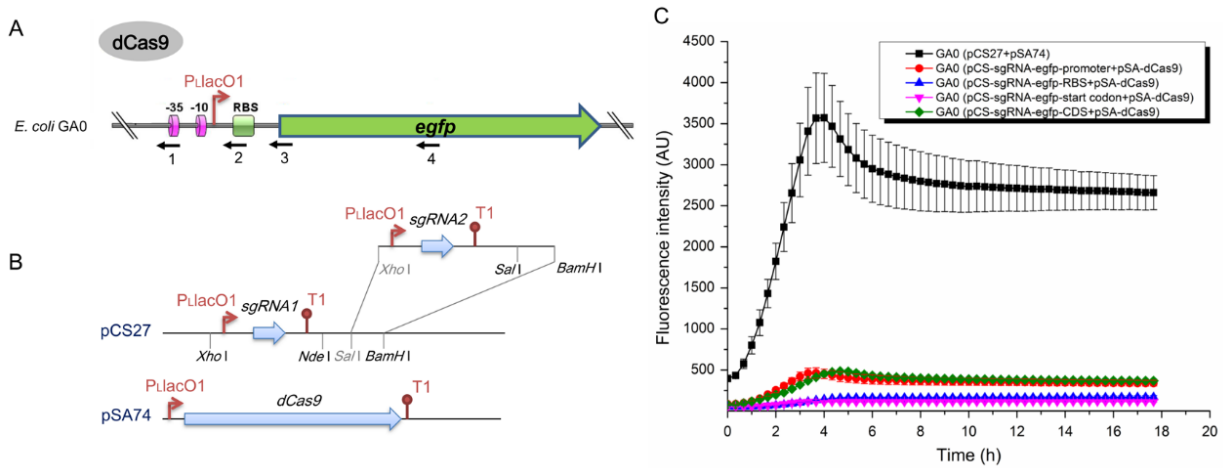



**Figure 5.4 CRISPRi based repression of eGFP expression.** (A) Representation of chromosomally integrated *$P_L$lacO1-egfp* expression cassette in strain GA0. Four sgRNA targeting sites were denoted: 1, promoter region at -35 box; 2, RBS; 3, start codon; 4, central position of coding sequence of *egfp*. (B) Scheme of CRISPRi plasmids construction. sgRNA was inserted into pCS27 and dCas9 was inserted into pSA74, both under control of *$P_L$lacO1* promoter with T1 terminator. (C) GA0 was transformed with CRISPRi plasmids for eGFP fluorescence detection. eGFP fluorescence intensity was recorded at emission wavelength of 528 nm and excitation wavelength of 485 nm. The error bars represent the standard deviations of three independent experiments.

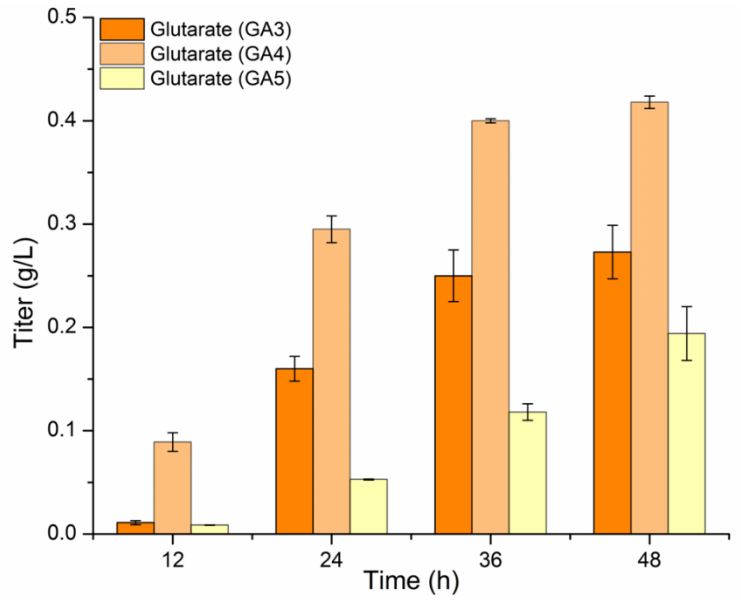



**Figure 5.5 Production enhancement of glutarate through CRISPRi mediated repression of *sucA* and/or *sucB*.** Production of glutarate when introducing CRISPRi mediated repression of *sucA* (strain GA3, orange columns), *sucA* and *sucB* (strain GA4, light orange columns) in *E. coli* BW25113 (F') or repression of *sucA* and *sucB* in *E. coli* BW25113 (F') Δ*poxB*Δ*ackA*Δ*pta* (strain GA5, light yellow columns). The error bars represent the standard deviations of three independent experiments.

Chapter 6

# MICROBIAL PRODUCTION OF BRANCHED-CHAIN DICARBOXYLATE 2-METHYLSUCCINIC ACID VIA ENOATE REDUCTASE-MEDIATED BIOREDUCTION[1]

[1]Jian Wang, Yaping Yang, Ruihua Zhang, Xiaolin Shen, Zhenya Chen, Qipeng Yuan, Yajun Yan. *Metabolic Engineering*, 2018. 45, 1-10.

Reprinted here with permission of the publisher.

## 6.1 Abstract

2-Methylsuccinic acid (2-MSA) is a C5 branched-chain dicarboxylate that serves as an attractive synthon for the synthesis of polymers with extensive applications in coatings, cosmetic solvents and bioplastics. However, the lack of natural pathways for 2-MSA biosynthesis has limited its application as a promising bio-replacement. Herein, we conceived a non-natural three-step biosynthetic route for 2-MSA, via employing the citramalate pathway in combination with enoate reductase-mediated bioreduction of the pathway intermediate citraconate. First, over-expression of codon-optimized citramalate synthase variant CimA* from *Methanococcus jannaschii*, endogenous isopropylmalate isomerase EcLeuCD and enoate reductase YqjM from *Bacillus subtilis* allowed the production of 2-MSA in *Escherichia coli* for the first time, with a titer of 0.35 g/L in shake flask experiments. Subsequent screening of YqjM-like enoate reductases of different bacterial origins enabled identification and characterization of a new NAD(P)H-dependent enoate reductase KpnER from *Klebsiella pneumoniae*, which exhibited higher activity towards citraconate than YqjM. Incorporation of KpnER into the 2-MSA biosynthetic pathway led to 2-MSA production improvement to a titer of 0.96 g/L in aerobic condition. Subsequent optimizations including cofactor regeneration, microaerobic cultivation and host strain engineering, boosted 2-MSA titer to 3.61 g/L with a molar yield of 0.36 in shake flask experiments. This work established a promising platform for 2-MSA bioproduction, which enabled the highest titer of 2-MSA production in microbial hosts so far.

## 6.2 Introduction

Dicarboxylic acids are important building blocks for the production of diols, polyurethanes, polyesters and polyamides with an ever-increasing market [354,355,390]. The roaring interests in developing novel bioplastics have prompted microbial metabolic engineering for large-scale production of natural or non-natural dicarboxylic acids from renewable biomass, which is both sustainable and environmentally friendly [390]. Current research mainly focused on the production of linear aliphatic dicarboxylic acids (C3-C10) such as succinate, glutarate and adipate [354,361,391-393]. Among those, only succinate production has achieved economically competitive productivities and yields in microbial hosts [358,394]. Recently, branched-chain or isomeric dicarboxylic acids like itaconate, mesaconate, malate and citramalate have drawn much more attention as alternative building block chemicals [395-399]. Itaconate has been recognized as one of the top 12 platform chemicals, serving as a precursor for value-added C5 chemicals like 2-methyl-1,4-butanediol (MBDO) and methyl-γ-butyrolactones (MGBLs) [357,400,401].

2-Methylsuccinic acid (2-MSA) is a methylated succinate and a reduced product of itaconate that has broad applications as a chemical synthon in the pharmaceutical and polymer industries [402-405]. It has been used as co-monomer with diols for the synthesis of biodegradable polyesters like poly (butylenes 2-methylsuccinate) (PBM), showing great applications for coatings and cosmetic solvents [402,403,406,407]. 2-MSA also serves as an alternative to itaconate for synthesis of C5 chemicals including MBDO and MGBLs. Generally, 2-MSA are produced from chemical reduction of itaconate by electro-catalytic valorization or chemo-catalytic hydrogenation via ruthenium-based catalysts, which however, are challenging owing to the costs of metal complexes and strict conditions like high temperature [408,409]. Recently, the direct formation of 2-MSA from citric acid was successfully achieved via one-step dehydration, decarboxylation and

hydrogenation, which also requires metal catalysts and harsh reaction conditions [410]. Bio-based production of 2-MSA has been only achieved from ethylmalonyl-coenzyme A pathway (EMCP) that comprises seven enzymatic steps including expressing a heterologous thioesterase to convert methylsuccinyl-CoA into 2-MSA in methylotroph *Methylobacterium extorquens* AM1 [406,411]. The efficiency of this CoA-dependent pathway was low and only 0.26 g/L 2-MSA was released from methylsuccinyl-CoA, which has to be significantly increased to allow a commercially attractive process [411].

Bioreduction of unsaturated C5 dicarboxylic acids including itaconate, mesaconate and citraconate by enoate reductases (ERs) might serve as potential pathways for 2-MSA [412,413]. The old yellow enzymes (OYEs) are a classic family of ERs that are flavin mononucleotide (FMN)-containing, NAD(P)H-dependent oxidoreductases. These OYE-like enzymes are capable of catalyzing the C=C bond reduction of α,β-unsaturated ketones, imides, nitroalkenes, aldehydes, carboxylic acids and their derivatives, affording hydrogenated products with a variety of industrial and pharmaceutical applications [414]. YqjM from *Bacillus subtilis* is one of the most extensively investigated OYE-like ER that has been utilized for bioreduction-based production of β-methyl-δ-valerolactone (βMδVL) and n-butanol [415,416]. YqjM has been reported to display *in vitro* reduction activity towards citraconate, which is probably ascribed to its structural similarity to the ideal substrate N-ethylmaleimide [412,417,418]. Citraconate is an intermediate metabolite in the well-characterized citramalate pathway that serves as an alternative isoleucine pathway in most methanogenic archaea such as *M. jannaschii* and certain bacteria such as *Leptospira interrogans* [120,419,420]. These imply the feasibility of establishing the ER-mediated reduction of citraconate for bio-production of 2-MSA.

Herein, to explore the possibility of using enoate reductase for 2-MSA production, we assembled a novel CoA-independent and short-cut 2-MSA biosynthetic pathway via harnessing *Escherichia coli* as a biological chassis. First, we established and validated the total biosynthetic pathway for 2-MSA mainly by introducing the optimized (R)-citramalate pathway and enoate reductase YqjM. Then we identified and characterized a superior YqjM-like enoate reductase from *Klebsiella pneumoniae* that enabled the improved production of 2-MSA. Final incorporation of a NADH regeneration system, cultivation optimization in microaerobic condition and host engineering permitted 2-MSA titer enhancement to 3.61 g/L in shake flask experiments. This work constituted and demonstrated a novel biosynthetic pathway for 2-MSA and achieved the highest bioproduction of 2-MSA so far, which might potentially serve as an industrial platform for bio-based synthesis of 2-MSA from renewable carbon sources.

## 6.3 Methods and materials

### 6.3.1 Strains, plasmids and chemicals

All bacterial strains and plasmids used in this study are summarized in Table 6. 1. *E. coli* strain XL-1 Blue (Stratagene) was used as the host for standard DNA cloning, BL21 Star (DE3) (Invitrogen) for protein expression and BW25113 (F') for 2-MSA production, respectively. *E. coli* strain BW25113 (F') is a BW25113 (*rrnBT14 ΔlacZWJ16 hsdR514 ΔaraBADAH33 ΔrhaBADLD78*) derivative strain with F' plasmid transduced from *E. coli* XL-1 Blue to supply $lacI_q$. *E. coli* BW25113 (F') Δ*adhE*Δ*ldhA*Δ*frdBC*Δ*fnr*Δ*pta* was created by deleting *adhE, ldhA, frdBC, fnr* and *pta* [249]. Compatible plasmids pZE12-luc (high-copy number) and pCS27 (medium-copy number) were used for the construction of 2-MSA biosynthetic pathway. Plasmid

pETDuet-1 was used for gene expression and purification. Standard chemicals including (*R*)-citramalate, citraconate and 2-MSA were purchased from Sigma-Aldrich.

### 6.3.2 Plasmid construction

All DNA manipulations were carried out following standard molecular cloning protocols [207]. Phusion high-fidelity DNA polymerase (New England Biolabs) was used for gene amplification and Quick ligase kit (New England Biolabs) was used for gene ligation. To construct pZE12-luc plasmids containing YqjM-like enoate reductase genes, *nemA* (NP_416167) from *E. coli*, *xenA* (AAN66878) from *Pseudomonas putida* and *reuER* (CAJ96943) from *Ralstonia eutropha* were respectively constructed into pZE12-luc between *Acc65*I/*Xba*I sites. Plasmids pZE12-*yqjM*, pZE12-*lmoER*, pZE12-*cglER*, pZE12-*kpnER*, pZE12-*oye2* and pZE12-*oye3* were constructed in our previous work [253]. For the purification purpose, *yqjM* and *kpnER* were cloned into pETDuet-1 between *Bam*HI/*Sal*I sites, obtaining pETDuet-*yqjM* and pETDuet-*kpnER*. pZE-*cimA**-*EcleuCD* was obtained by amplifying *cimA** from pZE-*cimA** and *EcleuCD* from *E. coli*, and inserting them into pZE12-luc by *Acc65*I/*Sal*I and *Sal*I/*Xba*I sites. To construct pZE-*cimA**-*EcleuCD-yqjM* and pZE-*cimA**-*EcleuCD-kpnER*, the cassettes $P_L$*lacO1-yqjM* and $P_L$*lacO1-kpnER* were respectively amplified from pZE12-*yqjM* and pZE12-*kpnER* and inserted into pZE-*cimA**-*EcleuCD* by *Spe*I/*Sac*I sites as independent operons. pZE-*cimA**-*MjleuCD-kpnER*, pZE-*cimA**-*VfleuCD-kpnER* and pZE-*cimA**-*Scleu1-kpnER* were constructed by amplifying *MjleuCD* from *Methanococcus jannaschii*, *VfleuCD* from *Vibrio fischeri* and *Scleu1* from *Saccharomyces cerevisiae and replacing EcleuCD in* pZE-*cimA**-*EcleuCD-kpnER* by *Sal*I/*Xba*I sites. pCS-*EcleuCD* was constructed by amplifying *EcleuCD* from *E. coli* and inserting into pCS27 between *Acc65*I and *Bam*HI. pCS27 containing a formate dehydrogenase gene from *Candida boidinii*

(pCS-*fdh*) was obtained from our previous work [421]. Plasmid pCS-*EcleuCD-fdh* was obtained by amplifying the cassette $P_L$*lacO1-fdh* from pCS-*fdh* and inserting into pCS-*EcleuCD* between *Spe*I and *Sac*I sites as an independent operon. All plasmids involved were listed in Table 6. 1.

### 6.3.3 Culture media and conditions

Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl) was used for cell propagation. 20 mL M9 minimal medium (20 g/L glucose, 6 g/L $Na_2HPO_4$, 0.5 g/L NaCl, 3 g/L $KH_2PO_4$, 1 g/L $NH_4Cl$, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 5 g/L yeast extract) was used for shake flask experiments in 125-mL shake flasks. Specifically, sodium formate was added to the M9 minimal medium if needed with a final concentration of 50 mM. When necessary, ampicillin and kanamycin were added to the medium at 100 and 50 μg/mL, respectively.

### 6.3.4 Protein expression, purification and in vitro enzyme assays

To express and purify enoate reductases, *E. coli* BL21 Star (DE3) was transformed with pETDuet-*yqjM* or pETDuet-*kpnER*. Single colonies were inoculated in 3 mL LB tubes at 37 °C. 500 μL of the overnight cultures were transferred to 50 mL fresh LB and grown for 2.5 to 3 h at 37°C. When the cell optical density at 600 nm ($OD_{600}$) of cultures reached 0.6-0.8, the cells were induced by isopropyl-β-D-1-thiogalactopyranoside (IPTG) with a final concentration of 0.5 mM and incubated at 30 °C on a rotary shaker for 9 h. The cultures were centrifuged at 10,000 rpm for 10 min to collect the cell pellets and the cells were lysed using Mini Bead Beater (Biospec). Protein purification was performed using His-Spin Protein Miniprep Kit (Zymo Research, Irvine, CA) following manufactures' instructions. The purified protein was verified by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using 12% protein gel and the protein

concentration was determined using a BCA Protein Assay Kit (Pierce) following manufactures' instructions.

The enzyme assay of enoate reductase was performed by measuring the oxidation of NADH/NADPH at 340 nm as previously described [417]. Briefly, the assays were performed at 25 °C in 1 mL of 100 mM Tris buffer (pH 7.5) containing either 500 μM NADH or NADPH. The substrate citraconate was added with varying concentrations (0, 5, 25, 50, 100, 200 μM) and the purified YqjM or KpnER was added at a concentration of 2.5 or 5 μM, respectively. The change of absorption values at 340 nm was monitored using UV spectrophotometer (Thermo Scientific). Kinetic constants ($K_m$ and $k_{cat}$) were determined through the Michaelis-Menten equation using Origin software.

### 6.3.5 Shake flask experiments

All shake flask experiments were performed in a rotary shaker (New Brunswick Scientific) at 30 °C with a speed of 270 rpm. Transformants of *E. coli* BW25113 (F') or its derived knockout strains were inoculated in 3 mL LB medium and grown at 37 °C for 8 h. Then 0.4 mL of the seed cultures were transferred to 20 mL fresh M9 medium and grown at 30 °C for cultivation. IPTG was added to the shake flasks with a final concentration of 0.5 mM when transferring the seed culture to M9 medium. Feeding experiments were conducted to test the bioconversion of citraconate to (R)-citramalate by *E. coli*. Briefly, citraconate was added to the medium along with IPTG (0.5 mM) to a final concentration of 0.5 g/L. For all the above shake flask experiments, samples were taken every 12 h after induction and 2-MSA production was analyzed by HPLC. For microaerobic cultivation, seed cultures (0.4 mL) were transferred into 125 mL screw cap bottles containing 20 mL M9 medium with appropriate antibiotics, 0.5 mM IPTG and

50 mM sodium formate. The cultures were grown at 30 °C for 72 h at 270 rpm. The screw cap bottles were tightened to make the oxygen-limited conditions. All experiments were performed in triplicate.

### 6.3.6 HPLC Analysis

The analysis of the samples from shake flask experiments was performed by HPLC (Shimadzu) equipped with a Coregel-64H column (Transgenomic). Samples (1 mL) were centrifuged at 15,000 rpm for 10 minutes. The supernatants were filtered through 0.22 μm film and used for analysis. The mobile phase used was 20 mN $H_2SO_4$ with a flow rate of 0.6 mL/min. The oven temperature set at 60 °C. This HPLC method was modified from previous research [422].

## 6.4 Results

### 6.4.1 Rational design and establishment of a non-natural 2-MSA biosynthetic pathway in *E. coli*

2-MSA is a non-natural specialty chemical that cannot be directly produced by any microorganism. However, considering that enoate reductase YqjM from B. *subtilis* exhibits high conversion efficiency towards unsaturated dicarboxylate citraconate [412], we set out to design an artificial biosynthetic pathway for 2-MSA through bioreduction of citraconate. The designed synthetic 2-MSA pathway consists of three steps: 1) generation of (R)-citramalate from direct condensation of pyruvate and acetyl-CoA via the citramalate synthase CimA, 2) dehydration of (R)-citramalate to citraconate by isopropylmalate isomerase (IPMI) LeuCD, and 3) bioreduction of citraconate to 2-MSA via enoate reductase (ER) (**Figure 6.1**). CimA is an essential enzyme in the citramalate pathway from methanogenic archaea or certain bacteria that catalyzes the first

step of leucine biosynthesis [120,423]. Directed evolution of CimA from *M. jannaschii* enabled generation of an isoleucine feedback inhibition resistant variant CimA3.7 (I47V/E114V/H126Q/T204A/L238S variant with additional truncation of the C-terminal domain from the 373rd residue) with improved activity [120]. IPMI LeuCD catalyzes the isomerization of α-isopropylmalate to β-isopropylmalate through β-isopropylmaleate in a reversible dehydration and trans-rehydration manner. In light of the structural similarity between (R)-citramalate and α-isopropylmalate, LeuCD of *E. coli* (EcLeuCD) was capable of converting (R)-citramalate to citraconate and further to β-methylmalate [120,419]. Thus, we proposed that final introduction of specific enoate reductases like YqjM from *B. subtilis* would eventually afford production of 2-MSA (**Figure 6.1**). Via this designed pathway, 2-MSA can be produced from glucose with a theoretical maximum yield of 1 mol/mol.

To establish and validate the total biosynthetic pathway, codon-optimized *cimA3.7* (named *cimA** hereafter) and *leuCD* from *E. coli* were constructed in one operon in a transcriptional order *cimA*-EcleuC-EcleuD*, while *yqjM* from *B. subtilis* was placed in an independent operon on the high-copy plasmid pZE12-luc. *E. coli* BW25113 (F') was transformed with the resultant plasmid pZE-*cimA*-EcLeuCD-yqjM*, resulting in strain MA-1. *E. coli* BW25113 (F') transformed with pZE-*cimA** (strain CMA-1) or pZE-*cimA*-EcleuCD* (strain CMA-2) were created as negative controls. All strains were subjected to shake flask experiments using M9 medium containing 20 g/L glucose. After 48 h of aerobic cultivation, strain CMA-1 and CMA-2 produced (R)-citramalate with the highest titer of 3.59 g/L and 3.96 g/L respectively, while there was negligible accumulation of citraconate or 2-MSA (**Figure 6.2A** and **B**). Acetate was not significantly accumulated when *cimA** was over-expressed (0.60 g/L in CMA-1 and 0.64 g/L in CMA-2 in 48 h), demonstrating the high catalytic activity of CimA* in producing (R)-

citramalate. Whereas, strain MA-1 expressing the full pathway allowed successful production of 2-MSA with a peak titer of 0.35 g/L at 36 h (**Figure 6.2C**). This demonstrated that the designed synthetic pathway of 2-MSA was functional in *E. coli* as anticipated.

An interesting observation was that (R)-citramalate produced by strain MA-1 (0.33 g/L) was significantly lower than that produced by the control strain CMA-2 (3.96 g/L). Meanwhile, there is no observable accumulation of citraconate in the cultures of strain MA-1 (**Figure 6.2C**). This indicated that either over-expression of YqjM impaired the (R)-citramalate pathway or YqjM was not efficient in redirecting the carbon flux to the final product 2-MSA.

### 6.4.2 Screening and characterization of superior ER for 2-MSA production

To screen superior ER for citraconate bioreduction, the protein-protein basic local alignment search tool (BlastP) was employed to search YqjM homologs from different microbial origins. We selected eight candidate YqjM homologs, four known ERs including NemA (NP_416167) from *E. coli,* XenA (AAN66878) from *Pseudomonas putida*, OYE2 (NP_012049) and OYE3 (DAA11263) from *Saccharomyces cerevisiae,* and four uncharacterized putative ERs including ReuER (CAJ96943) from *Ralstonia eutropha*, LmoER (AJA84118) from *Listeria monocytogenes*, CglER (AGT06739) from *Corynebacterium glutamicum* and KpnER (ABR77188) from *Klebsiella pneumoniae* (Table 6. 2). The protein sequence alignment showed that XenA, LmoER, CglER and KpnER shared relatively high similarity (>49 %) with YqjM, which were also conserved in amino acid residues that are involved in FMN and substrate binding (**Figure S6.1**).

Feeding citraconate into *E. coli* cells for enzyme screening might not be feasible, because citraconate will be hydrated either back to (R)-citramalate or to β-methylmalate by endogenous

LeuCD. This was indicated by the observation that negligible citraconate accumulated when *cimA** and *EcleuCD* were over-expressed in *E. coli* (**Figure 6.2B**). To further testify that, citraconate (0.5 g/L) was fed into wild type *E. coli* BW25113 (F') or strain CMA-2. In both cases, citraconate was almost completely converted into (R)-citramalate (**Figure 6.2D**). This suggested that hydration of citraconate to (R)-citramalate was preferred by EcLeuCD. Therefore, we set to perform an *in vivo* screening by co-expressing *cimA** and *EcleuCD* on pCS27 with each ER on pZE12-luc in *E. coli* BW25113 (F'). Shake flask experiments and HPLC analysis indicated that except OYE2 and OYE3 all the other five YqjM-like ERs enabled the production of 2-MSA with titers ranging from 0.02 to 0.24 g/L in 24 h (**Figure 6.3A**). Notably, KpnER from *K. pneumoniae* afforded the highest titer of 2-MSA (0.24 g/L), which was approximately 3.4-fold to that produced by YqjM (0.07 g/L). Interestingly, different ERs rendered varying decrease of (R)-citramalate titers, while YqjM enabled the lowest (R)-citramalate titer (0.14 g/L) (**Figure 6.3A**). This further indicated that expression of ERs especially YqjM significantly compromised (R)-citramalate production.

To verify the reductase activity towards citraconate, YqjM and KpnER were expressed with N-terminal $His_6$-tag and purified from *E. coli* BL21 Star (DE3). We tested both NADH and NADPH as the reductants in the *in vitro* enzyme assays. The results indicated that YqjM and KpnER could utilize both NADH and NADPH as reductants, while both ERs exhibited higher catalytic efficiency ($k_{cat}/K_m$) in the presence of NADPH (Table 6. 3). YqjM showed a $K_m$ value of 58.74 μM and a $k_{cat}$ value of 0.92 $min^{-1}$ when using NADH as the cofactor, which were 39.15 μM and 1.01 $min^{-1}$, respectively, with NADPH as the cofactor. As a comparison, KpnER showed a $K_m$ value of 11.57 μM and a $k_{cat}$ value of 16.45 $min^{-1}$ with NADH as the cofactor, which were 2.59 μM and 16.70 $min^{-1}$, respectively, when using NADPH as the cofactor. This clearly

demonstrated that KpnER showed both higher specificity and catalytic activity towards citraconate than YqjM in the presence of either NADH or NADPH.

For 2-MSA production, *cimA**, *EcleuCD* and *kpnER* were co-expressed on the high-copy plasmid pZE12-luc in *E. coli* BW25113 (F') (resultant strain MA-2). When cultured under the same condition, 2-MSA production was improved to a peak titer of 0.96 g/L at 36 h, which is 2.8-fold to that of strain MA-1 expressing *yqjM* (**Figure 6.3B**). Taken together, both *in vitro* assay and *in vivo* production demonstrated that KpnER exhibited improved activity towards citraconate, thus enabling the higher production of 2-MSA.

### 6.4.3 Optimization of IPMI for 2-MSA production

It is noteworthy that (R)-citramalate was significantly accumulated while there was almost no accumulation of the intermediate citraconate when over-expressing either *yqjM* or *kpnER*. This might be caused by the preferred re-hydration of citraconate to (R)-citramalate by endogenous EcLeuCD. To address this bottleneck, we resorted to screen alternative IPMIs from different bacterial contexts.

IPMIs are mostly two-subunit enzymes that require a [4Fe-4S] iron-sulfur cluster as a co-factor for catalytic function. Three different IPMIs were tested including MjLeuCD (Mj0499-Mj1277) from *M. jannaschii*, VfLeuCD (B5FGH2-B5FGH1) from *V. fischeri* and ScLeu1 (YGL009C) from *S. cerevisiae. VfLeuC and VfLeuD shared 87 % and 84 % similarity with EcLeuC and EcLeuD, respectively. MjLeuCD (*Mj0499-Mj1277*) was selected because of its native* role as isopropylmalate and (R)-citramalate isomerase in *M. jannaschii* [419]. *ScLeu1 is a unique and sing*ular IPMI that catalyz*es the second step of* isopropylmalate isomerization *in the leucine biosynthesis pathway in S. cerevisiae.*

To apply these alternate *IPMI*s for 2-MSA production, MjleuCD, VfleuCD and Scleu1 were incorporated into the 2-MSA pathway plasmid and transferred into *E. coli* BW25113 (F') for shake flask experiments. Strain MA-4 expressing *MjleuCD* produced 0.91 g/L 2-MSA (**Figure 6.4**), which is nearly comparable to that when expressing *EcleuCD* (0.96 g/L). Whereas, strain MA-5 expressing *VfleuCD* and MA-6 expressing *Scleu1* produced less 2-MSA, with maximum titers of 0.74 and 0.81 g/L, respectively (**Figure 6.4**). The control strain MA-3 that didn't over-express any IPMI only produced 0.11 g/L 2-MSA, which was probably due to the low-level basal expression of endogenous *EcleuCD* on the chromosome. These results indicated that the plasmid-borne expression of IPMI from *E. coli* is most beneficial to 2-MSA production.

Next, to test if further increasing the copy number of *EcleuCD* would improve 2-MSA production, pCS-*EcleuCD* was introduced into *E. coli* BW25113 (F') with pZE-*cimA**-*EcleuCD*-*kpnER* (strain MA-7) for shake flask experiments. Surprisingly, as a result of increased expression of *EcleuCD*, the titer of 2-MSA produced by strain MA-7 was slightly decreased to 0.88 g/L (**Figure 6.4**). This indicated that the intrinsic enzyme activity of IPMI rather than enzyme abundance inevitably limited further conversion of (R)-citramalate to citraconate.

### 6.4.4. Production enhancement via cofactor regeneration

During oxidative metabolism, NAD(P)H levels remain relatively low since oxygen serves as the major terminal electron acceptor [424]. To pull more carbon flux to 2-MSA, we next set out to increase endogenous NADH supply for KpnER via the heterologous introduction of the formate dehydrogenase in *E. coli* (**Figure 6.5A**). Briefly, FDH from *Candida boidinii* that catalyzes the oxidation of formate anion to $CO_2$ with concomitant reduction of $NAD^+$ to NADH, was expressed under control of $P_L$lacO1 promoter on the pCS27 (pCS-*fdh*). Strain MA-8 containing

pZE-*cimA*-EcleuCD-kpnER* and pCS-*fdh* was subjected to shake flask experiments using M9 medium supplemented with 20 g/L glucose and 50 mM sodium formate. As a result of co-expressing 2-MSA pathway and FDH, 1.12 g/L 2-MSA was produced in 48 h, which only amounted to a 16.7 % increase compared to that of control strain MA-2 (**Figure 6.5B**). To examine if simultaneous increase of reducing equivalents and *EcleuCD* expression might have a combined beneficial effect on 2-MSA production, *E. coli* BW25113 (F') was transformed with pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD-fdh* for shake flask experiments. The resultant strain MA-9 produced 1.34 g/L 2-MSA in 48 h, which was increased by 39.5 % in comparison with that of control strain MA-2 (**Figure 6.5B**). Interestingly, (R)-citramalate titer was reduced when 2-MSA titer was increased, demonstrating that increased cofactor supply enhanced carbon flux channeling to 2-MSA. However, the titer did not further improve when extending the cultivation time to 72 h (1.38 g/L) (**Figure 6.6A**).

### 6.4.5 Microaerobic cultivation for 2-MSA production enhancement

As a facultative anaerobe, *E. coli* might be more suitable for production of reduced chemicals in anaerobic or microaerobic conditions due to its favorable accumulation of NADH, and reduced carbon loss toward biomass and acetate [425,426]. The microaerobic condition has been commonly used considering its combinative features of both aerobic and anaerobic metabolism [128,427]. Thus, to further increase NADH supply, we turned to perform microaerobic cultivation using the best-performing strain MA-9 with M9 medium containing 20 g/L glucose and 50 mM sodium formate. Both the cell density ($OD_{600}$) and acetate accumulation of MA-9 in microaerobic cultivation were slightly lower than that in the aerobic cultivation (**Figure 6.6A** and **B**). However, metabolites analysis indicated that 1.64 g/L 2-MSA was produced within 48 h during

microaerobic cultivation, which was further increased to 1.81 g/L with a molar yield of 0.16 when extending the cultivation time to 72 h (**Figure 6.6B**). These results demonstrated that microaerobic condition is conducive to produce 2-MSA in *E. coli*.

To further alleviate carbon flux loss and increasing reducing power during microaerobic cultivation, we used *E. coli* BW25113 (F') Δ*adhE*Δ*ldhA*Δ*frdBC*Δ*fnr*Δ*pta* as the host strain for 2-MSA production. In this strain, *adhE*, *ldhA*, *frdBC*, *fnr* and *pta* were deleted to enable a dramatic decrease of ethanol, lactate, succinate, and acetate in anaerobic conditions [249,428]. The mutant host strain was transformed with pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD-fdh* (strain MA-10) and was subjected to shake flask experiments in microaerobic conditions. Remarkably, the titer of 2-MSA produced by strain MA-10 was boosted to 3.32 g/L in 48 h and further increased to 3.61 g/L in 72 h (**Figure 6.6C**), which amounted to 2.63-fold and 2.0-fold increases compared to that of strain MA-9 in aerobic condition and microaerobic condition, respectively. The molar yield of 2-MSA in strain MA-10 reached 0.36. Compared with strain MA-9 under aerobic condition, the biomass of strain MA-10 was slightly reduced, while the acetate was significantly reduced by 2.4-fold (**Figure 6.6 A** and **C**). This indicated that the carbon flux was significantly redirected to the 2-MSA pathway in the engineered host strain under microaerobic condition.

## 6.5 Discussion

2-MSA is a novel branched-chain dicarboxylic acid building block. Its bio-based production has been limited by the availability and efficacy of natural pathways. Potentially, 2-MSA can be produced from reduction of unsaturated C5 branched dicarboxylates including itaconate and mesaconate, whose microbial-based production have been well-developed with high titers in

engineered *E. coli* [396,429-431]. Enoate reductase mediated bioreduction of unsaturated compounds had shown promising applications that permitted developing new pathways or generating new chemical entities [253,416,432]. However, limited accessibility of functional enoate reductases towards these compounds restricted the feasibility through ER mediated bioreduction. Herein we developed a novel and shortcut route for 2-MSA via bioreduction of citraconate. Briefly, this pathway employed the well-established citramalate pathway to produce (R)-citramalate, which was successively dehydrated to citraconate by IPMI and reduced to 2-MSA by YqjM-like ER. Capitalization of the non-natural 2-MSA pathway in *E. coli* rendered successful production of 2-MSA from glucose. Based on that, we afforded to improve the 2-MSA titer to 3.61 g/L, stepwisely by enzyme mining, cofactor regeneration, host strain engineering and cultivation optimization.

First, enzyme mining permitted identification of a novel and superior YqjM-like enoate reductase KpnER from *K. pneumoniae* that exhibited higher specificity and activity towards citraconate. Sharing 51.8 % sequence similarity with YqjM, KpnER belongs to the OYE family of flavin oxidoreductase, considering its conserved FMN binding residues and yellow color when purified (**Figure S6.1**) [433]. Besides YqjM and KpnER showed even higher similarity (65 %) with chromate reductase (CrS) from the thermophile *Thermus scotoductus* SA-01, which has been demonstrated with higher turnover number than YqjM towards certain cyclic α, β-unsaturated carbonyl compounds [434]. In this work, we demonstrated that KpnER is a novel NAD(P)H dependent enoate reductase that exhibited higher catalytic efficiency with NADPH as the cofactor, as has also been observed in YqjM [417]. YqjM and its homologs exhibited a strong preference towards *cis*-α, β-unsaturated carboxylic acids or ketones [412,435], which may partially explain why KpnER showed high specificity and activity towards citraconate. When

incorporated into 2-MSA pathway, KpnER enabled 2-MSA production improvement to a final titer of 0.96 g/L, which was almost 3-fold to that of YqjM. Thus, KpnER implied great potential to expand the chemical repertoire of new value-added chemical entities from α, β-unsaturated precursors like citraconate or its analogs.

When expressing 2-MSA pathway in *E. coli*, it was of note that (R)-citramalate was significantly accumulated but no accumulation of citraconate was observed. To redistribute carbon flux to 2-MSA production, we next attempted to increase the push by screening superior IPMIs and enhance the pull by improving intracellular NADH supply. However, substitution of EcLeuCD with three IPMI homologs including MjLeuCD from *M. jannaschii*, VfLeuCD from *V. fischeri* and ScLeu1 from *S. cerevisiae didn't improve 2-MSA production.* Interestingly, rehydration of citraconate back to (R)-citramalate might be favored by IPMIs as observed here in EcLeuCD (**Figure 6.2D**) and demonstrated previously in MjLeuCD [419]. Increasing cofactor availability of NADPH or NADH might facilitate production of reduced products [436,437]. To enhance the pull while minimize genetic perturbations, we chose to improve cofactor supply via FDH-mediated NADH regeneration system, which was efficient when exogenously feeding formate sodium [421,426,438]. As expected, introducing NADH regenerating FDH from *C. boidinii* into *E. coli* led to 16.7 % increase of 2-MSA titer, which was further increased to 39.5 % when simultaneously increasing the copy number of *EcleuCD* (**Figure 6.5**). This indicated that increase of cofactor supply and expression of IPMI exerted a combinatorial effect in driving carbon flux to 2-MSA.

Final optimization was made via microaerobic cultivation that would allow simultaneous reduction of carbon loss and increase of reducing cofactors like NADH. Microaerobic cultivation of strain MA-9 almost doubled 2-MSA production (1.81 g/L), with a slight reduction of cell

growth and acetate accumulation (**Figure 6.6B**). To further alleviate carbon loss and eliminate NADH consumption pathway during the fermentative metabolism, host strain with deletions of *adhE, ldhA, frdBC, fnr* and *pta* was used. This would thereby increase precursor supply of pyruvate and acetyl-CoA, and create the surplus of NADH as the driving force for 2-MSA production. Considering KpnER functioning as the sole $NAD^+$ generator, the presence of limited oxygen might as well facilitate recycling excess NADH to balance internal redox [439,440]. As expected, 2-MSA titer was boosted to 3.61 g/L with a molar yield of 0.36 in the engineered strain MA-10, concomitant with slight reduction in cell growth and significant reduction in acetate accumulation. This was the highest titer and yield of 2-MSA achieved in microbial hosts so far. It is noteworthy that a significant amount of (R)-citramalate (around 3.0 g/L) was still accumulated in the culture medium of strain MA-10 (**Figure 6.6C**), which was probably caused by the preferred re-hydratase activity of IPMI. Thus, optimization of IPMI either through extensive enzyme mining or engineering might potentially further increase 2-MSA production in future.

In conclusion, this study first established a non-natural biosynthetic route for the sustainable production of branched-chain dicarboxylate 2-MSA from glucose in *E. coli*. Then enzyme screening permitted identification of one novel enoate reductase KpnER from *K. pneumoniae*, which enhanced 2-MSA production by approximately 3 folds. Final stepwise improvement via manipulation of cofactor regeneration and microaerobic cultivation boosted 2-MSA titer to 3.61 g/L with a final yield of 0.36 mol 2-MSA per mol glucose in engineered *E. coli* host. This work demonstrated the workflow of establishing non-natural pathways from scratch and engineering microbial chassis for production enhancement of non-natural chemicals. Especially, rational optimizations enabled driving of carbon flux to the 2-MSA synthetic pathway in a growth-

coupled manner, which can also be applied in general for high-level production of other value-added chemicals in microbial hosts.

## 6.6 Tables and figures

**Table 6.1 Strains and plasmids used in this study.**

| Strains | Properties | Source |
|---|---|---|
| *E. coli* XL1-Blue | *recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 *supE*44 *relA*1 *lac F′* [*traD36 proAB lacI*$^{q}$*ZΔM*15 *Tn*10 (*Tet*$^{r}$)] | Stratagene |
| *E. coli* BL21 Star (DE3) | F$^{-}$ *ompT hsdSB (rB-mB-) gal dcm rne131* (DE3) | Invitrogen |
| *E. coli* BW25113 (F') | *rrnBT*14 Δ*lacZWJ*16 *hsdR*514 *ΔaraBADAH*33 *ΔrhaBADLD*78 *F'* [*traD36 proAB lacI*$^{q}$*ZΔM15 Tn10(Tet*$^{r}$*)*] | [249] |
| *E. coli* BW25113 (F') Δ*adhE*Δ*ldhA*Δ*frdBC*Δ*fnr*Δ*pta* | *E. coli* BW25113 (F') with deletions of *adhE, ldhA, frdBC, fnr* and *pta* | [249] |
| CMA-1 | *E. coli* BW25113 (F') harboring pZE-*cimA** | This study |
| CMA-2 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD* | This study |
| MA-1 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-yqjM* | This study |
| MA-2 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* | This study |
| MA-3 | *E. coli* BW25113 (F') harboring pZE-*cimA*-KpnER* | This study |
| MA-4 | *E. coli* BW25113 (F') harboring pZE-*cimA*-MjleuCD-kpnER* | This study |

| | | |
|---|---|---|
| MA-5 | *E. coli* BW25113 (F') harboring pZE-*cimA*-VfleuCD-kpnER* | This study |
| MA-6 | *E. coli* BW25113 (F') harboring pZE-*cimA*-Scleu1-kpnER* | This study |
| MA-7 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD* | This study |
| MA-8 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*fdh* | This study |
| MA-9 | *E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD-fdh* | This study |
| MA-10 | *E. coli* BW25113 (F') Δ*adhE*Δ*ldhA*Δ*frdBC*Δ*fnr*Δ*pta* harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD-fdh* | This study |

| **Plasmids** | **Properties** | **Source** |
|---|---|---|
| pZE12-luc | $P_L$lacO1, *colE* ori, $Amp^r$ | [250] |
| pCS27 | $P_L$lacO1, *P*15*A* ori, $Kan^r$ | [251] |
| pETDuet-1 | T7 promoter, *pBR322* ori, $Amp^r$ | Novagen |
| pETDuet-*yqjM* | pETDuet harboring *yqjM* from *Bacillus subtilis* | This study |
| pETDuet-*kpnER* | pETDuet harboring *kpnER* from *Klebsiella pneumoniae* | This study |

| | | |
|---|---|---|
| pZE-*xenA* | pZE12-luc carrying *xenA* (AAN66878) from *Pseudomonas putida* KT2440 | This study |
| pZE-*nemA* | pZE12-luc carrying *nemA* (NP_416167) from *E. coli* | This study |
| pZE-*reuER* | pZE12-luc carrying the sequence encoding NADH:flavin oxidoreductase (CAJ96943) from *Ralstonia eutropha* H16 | This study |
| pZE-*yqjM* | pZE12-luc carrying *yqjM (*BAA12619*)* from *Bacillus subtilis* | [253] |
| pZE-*oye2* | pZE12-luc carrying *oye2* (NP_012049) from *Saccharomyces cerevisiae* | [253] |
| pZE-*oye3* | pZE12-luc carrying *oye3* (DAA11263) from *Saccharomyces cerevisiae* | [253] |
| pZE-*lmoER* | pZE12-luc carrying the sequence encoding NADH:flavin oxidoreductase (AJA84118) from *Listeria monocytogenes* | [253] |
| pZE-*cglER* | pZE12-luc carrying the sequence encoding NADH:flavin oxidoreductase (AGT06739) from *Corynebacterium glutamicum* | [253] |
| pZE-*kpnER* | pZE12-luc carrying the sequence encoding putative NADH:flavin oxidoreductase/NADH oxidase (ABR77188) from *Klebsiella pneumonia* | [253] |
| pZE-*cimA** | pZE12-luc harboring codon optimized *cimA3.7* (*cimA**) from *Methanococcus jannaschii* | [399] |

| | | |
|---|---|---|
| pCS-*cimA*-EcleuCD* | pCS27 harboring *cimA** and *EcleuCD* | This study |
| pZE-*cimA*-kpnER* | pZE-*cimA** containing *kpnER* | This study |
| pZE-*cimA*-EcleuCD-yqjM* | pZE12-*cimA*-EcleuCD* containing *yqjM* | This study |
| pZE-*cimA*-EcleuCD-kpnER* | pZE12-*cimA*-EcleuCD* containing *kpnER* | This study |
| pZE- *cimA*-MjleuCD-kpnER* | pZE12-*cimA*-MjleuCD* containing *kpnER* | This study |
| pZE- *cimA*-VfleuCD-kpnER* | pZE12-*cimA*-VfleuCD* containing *kpnER* | This study |
| pZE- *cimA*-Scleu1-kpnER* | pZE12-*cimA*-Scleu1* containing *kpnER* | This study |
| pCS-*fdh* | pCS27 containing formate dehydrogenase gene (*fdh*) from *Candida boidinii* | [421] |
| pCS-*EcleuCD* | pCS27 containing *EcleuCD* | This study |
| pCS-*EcleuCD-fdh* | pCS-*EcleuCD* containing *fdh* | This study |

**Table 6.2 Enoate reductase candidates screened in this study.**

| ER candidates | Accession no. | Seq. similarity (%) [a] | Microbial origin |
|---|---|---|---|
| YqjM | BAA12619 | 100/100 | *Bacillus subtilis* |
| XenA | AAN66878 | 49.2/37.4 | *Pseudomonas putida KT2440* |
| NemA | NP_416167 | 42.4/27.1 | *Escherichia coli* |
| OYE2 | NP_012049 | 33.8/22.5 | *Saccharomyces cerevisiae S288c* |
| OYE3 | DAA11263 | 34.6/21.4 | *Saccharomyces cerevisiae S288c* |
| ReuER | CAJ96943 | 36.8/25.2 | *Ralstonia eutropha H16* |
| LmoER | AJA84118 | 73.7/61.9 | *Listeria monocytogenes* |
| CglER | AGT06739 | 55.9/39.4 | *Corynebacterium glutamicum* |
| KpnER | ABR77188 | 53.8/41.0 | *Klebsiella pneumoniae* |

[a,] Sequence similarity/identity with YqjM.

**Table 6.3 Kinetic parameters of YqjM and KpnER towards citraconate.**

| Enzyme | NADH | | | NADPH | | |
|---|---|---|---|---|---|---|
| | $K_m$ (μM) | $k_{cat}$ ($min^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}min^{-1}$) | $K_m$ (μM) | $k_{cat}$ ($min^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}min^{-1}$) |
| YqjM | 58.74±0.11 | 0.92±0.01 | 15.64 | 39.15±9.35 | 1.01±0.03 | 25.69 |
| KpnER | 11.57±1.81 | 16.45±0.44 | 1421.48 | 2.59±1.10 | 16.70±0.43 | 6458.89 |

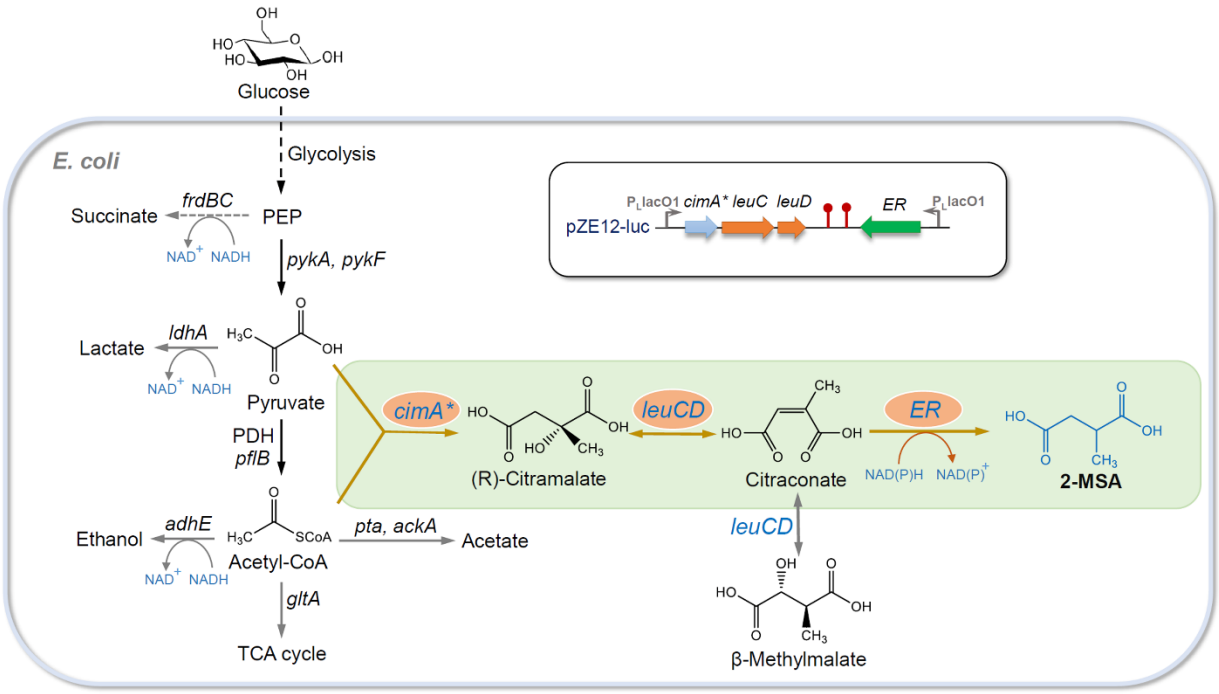



**Figure 6.1 Establishing a novel biosynthetic pathway for 2-methylsuccinic acid in *E. coli*.** Pyruvate and acetyl-CoA, two major metabolites derived from glycolysis pathway, are condensed by citramalate synthase (CimA*) to produce (R)-citramalate. (R)-Citramalate is dehydrated to citraconate by IPMI LeuCD, which is finally reduced to 2-methylsuccinic acid via enoate reductase (ER). Citraconate can also be rehydrated to β-methylmalate by LeuCD. The panel inside represents the construction of the full 2-MSA pathway in plasmid pZE12-luc. Genes involved in competing pathways or NADH consumptions are indicated in grey arrows.

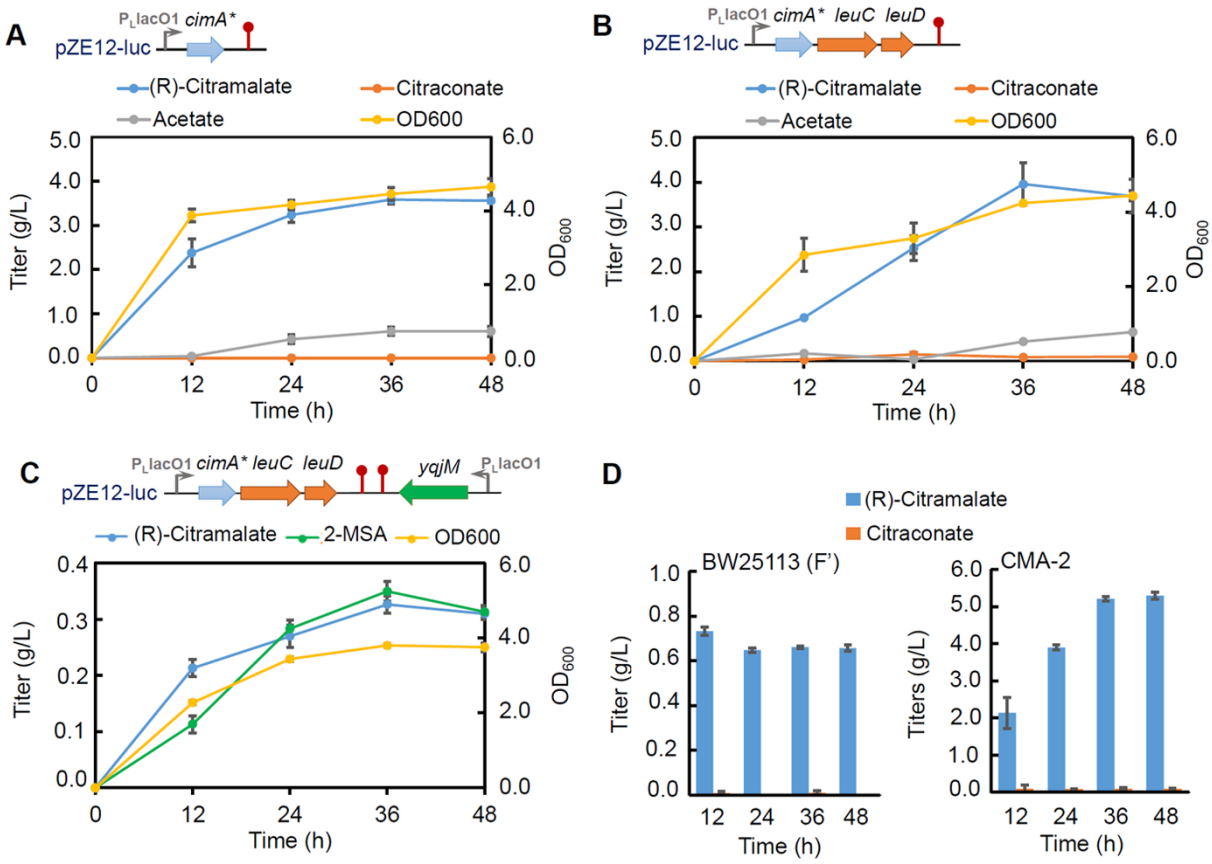



**Figure 6.2 *De novo* production of 2-MSA from glucose in *E. coli*.** Growth and metabolite profiles of strain CMA-1 (*E. coli* BW25113 (F') containing pZE-*cimA**) (A), CMA-2 (*E. coli* BW25113 (F') containing pZE-*cimA**-*EcleuCD*) (B), and MA-1 (*E. coli* BW25113 (F') containing pZE-*cimA**-*EcleuCD*-*yqjM*) (C). (D) Bioconversion test of citraconate to (R)-citramalate. Citraconate (0.5 g/L) was fed into *E. coli* BW25113 (F') and strain CMA-2 and cultivated for 48 h. The error bars represent the standard deviations of three independent experiments.

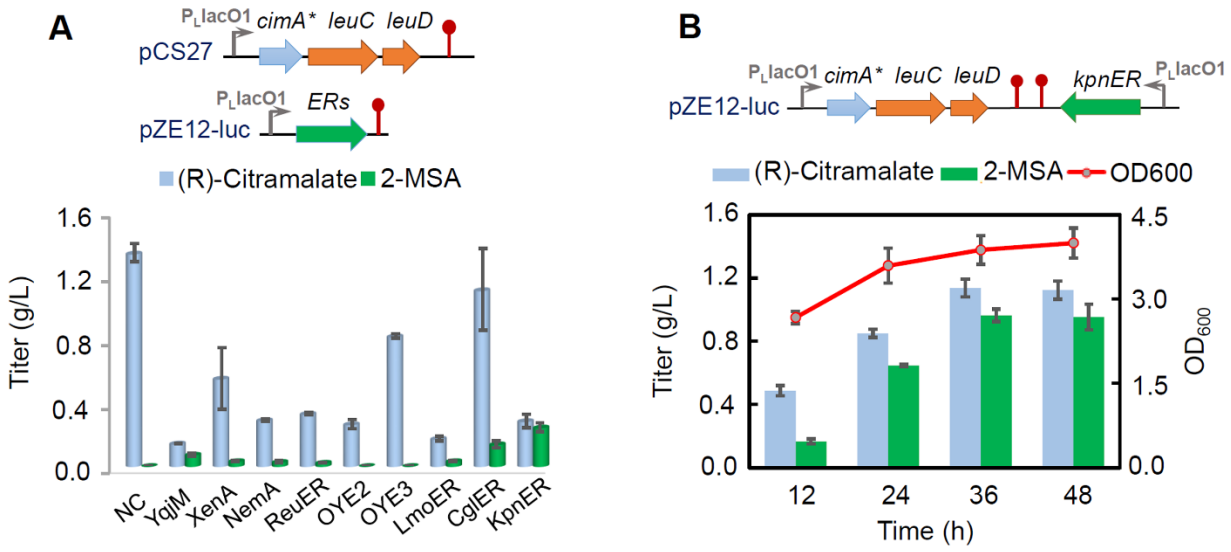



**Figure 6.3 Screening for superior enoate reductase for 2-MSA production.** (A) *In vivo* screening of enoate reductases. *E. coli* BW25113 (F') was transformed with pCS-*cimA**-*EcleuCD* and pZE12-luc containing each ER candidate genes for shake flask experiments. (R)-Citramalate and 2-MSA titers were measured from samples taken at 24 h. NC, negative control, *E. coli* BW25113 (F') transformed with pCS-*cimA**-*EcleuCD* and pZE12-luc empty plasmid. (B) 2-MSA production in strain MA-2 (*E. coli* BW25113 (F') containing pZE-*cimA**-*EcleuCD*-*kpnER*). The error bars represent the standard deviations of three independent experiments.

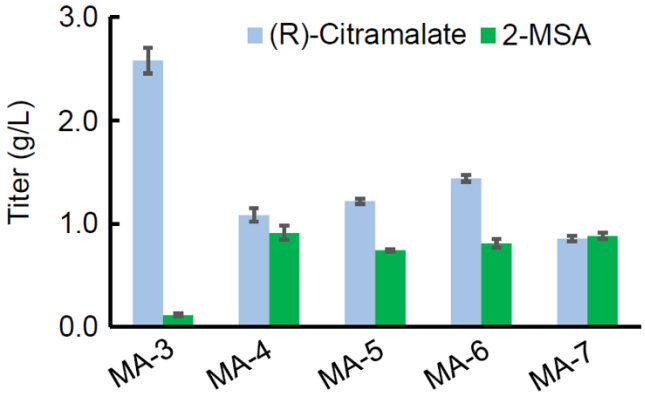



**Figure 6.4 Screening of IPMI for 2-MSA production.** *E. coli* BW25113 (F') transformed with pZE-*cimA*-kpnER* (strain MA-3), pZE-*cimA*-MjleuCD-kpnER* (strain MA-4), pZE-*cimA*-VfleuCD-kpnER* (strain MA-5), pZE-*cimA*-Scleu1-kpnER* (strain MA-6), pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD* (strain MA-7) were subjected to shake flask experiments. The final 2-MSA titers at 48 h were determined and compared. The error bars represent the standard deviations of three independent experiments.

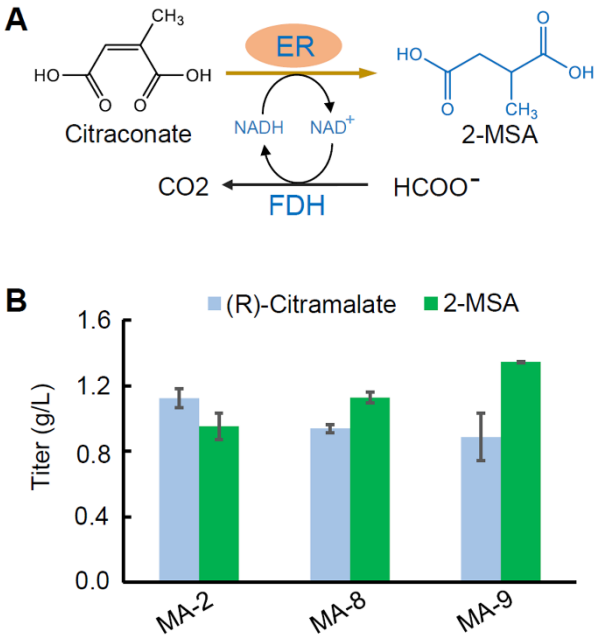
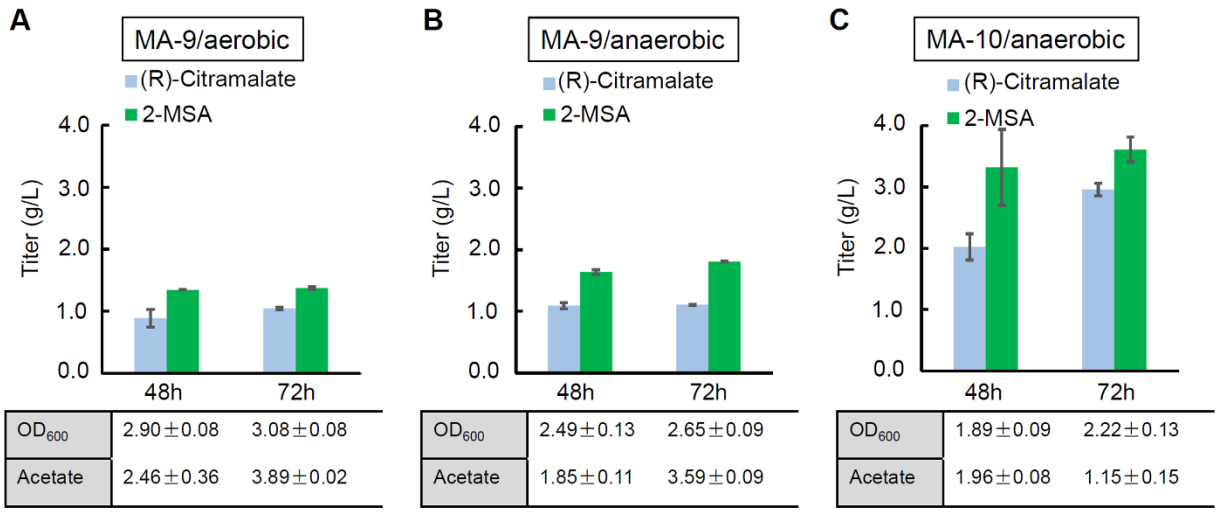



**Figure 6.5 Enhancing 2-MSA production via improving cofactor regeneration.** (A) Scheme of FDH mediated NADH regeneration. (B) The effects of over-expression of *fdh* on 2-MSA production. Strain MA-8 (*E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*fdh*) and strain MA-9 (*E. coli* BW25113 (F') harboring pZE-*cimA*-EcleuCD-kpnER* and pCS-*EcleuCD-fdh*) were used for shake flask experiments. The final 2-MSA titers at 48 h were determined and compared. The error bars represent the standard deviations of three independent experiments.




| | | |
|---|---|---|
| $OD_{600}$ | 2.90±0.08 | 3.08±0.08 |
| Acetate | 2.46±0.36 | 3.89±0.02 |

| | | |
|---|---|---|
| $OD_{600}$ | 2.49±0.13 | 2.65±0.09 |
| Acetate | 1.85±0.11 | 3.59±0.09 |

| | | |
|---|---|---|
| $OD_{600}$ | 1.89±0.09 | 2.22±0.13 |
| Acetate | 1.96±0.08 | 1.15±0.15 |

**Figure 6.6 Aerobic or microaerobic cultivation for 2-MSA production.** 2-MSA titers, cell densities and acetate accumulation of strain MA-9 under aerobic condition (A), strain MA-9 under microaerobic condition (B), and strain MA-10 under microaerobic condition (C) at 48 h and 72 h of cultivation. The error bars represent the standard deviations of three independent experiments.

## 6.7 Supplementary information

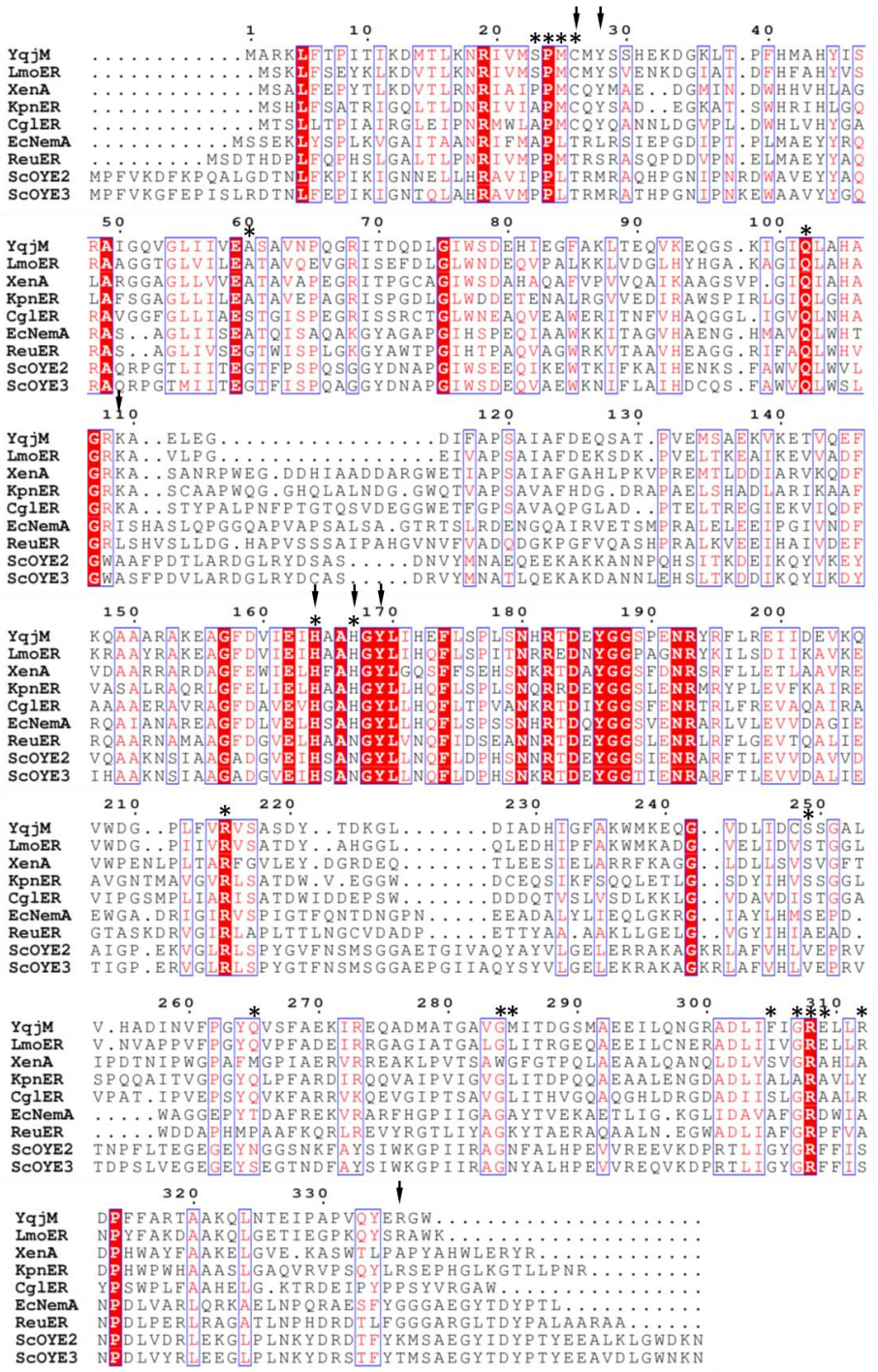



**Figure S6.1 Sequence alignment of YqjM-like enoate reductases.** FMN binding residues and putative substrate binding residues of YqjM are indicated by *asterisk* and *downward arrow*, respectively.

# Chapter 7

# CONCLUSION

Industrial microbes like *E. coli* harbor all amino acid biosynthetic pathways and can synthesize amino acids from cheap carbon feedstocks. With increasing understanding of amino acid biosynthetic machineries, it is possible to engineer microbial host for industrial production of amino acids in microbial hosts. However, most of the proteinogenic amino acids are currently produced as end and bulk biochemicals with around 5 million tons per year, and are mainly used as animal feed additives and flavor ingredients [319,320]. The utilization of amino acids as building blocks to expand chemical space remains underexplored. Considering the structural diversity and ease of achieving high titers in microbial hosts, amino acids are ideal starting materials for new value-added natural products or commodity chemicals. The work here describes the establishment of artificial pathways for caffeic acid derived esters and amides by extending aromatic amino acid pathways, the design of a pyruvate-driven system for overproduction of tryptophan and its derivatives, the demonstration of a short-chain diol (C3-C5) production platform by extending charged amino acid pathways, and the establishment of two different C5 dicarboxylate pathways via acetyl-CoA mediated carbon chain elongation pathways. These results have demonstrated the feasibility of expanding the chemical space of microbial systems by extending amino acid catabolism, and have laid the foundation in laboratory scale for further improvement of these new products, as outlined in **Table 7.1**.

**Developing artificial pathways for caffeic acid phenethyl esters and amides**

Hydroxycinnamic acid esters (HCEs) and amides (HCAAs) are a widely distributed and phytochemically important group of plant secondary metabolites with interesting pharmaceutical potentials. Among those HCEs and HCAAs, caffeic acid derived phenethyl esters and amides are of the most interest because of their potent anti-oxidative activity and pharmaceutical applications. To circumvent labor-intensive and low-yielding plant extraction, increasing efforts have been made to establish microbial cell factories for their production. However, the inaccessibility to their natural pathways have hindered the development of their microbial-based production. In chapter 2, we described the establishment of artificial biosynthetic pathways for caffeic acid phenethyl esters and amides by extending tyrosine and phenylalanine pathways in *E. coli*, which allowed the production a total of five valuable caffeic acid derived esters and amides from glucose, including the most well-known representatives caffeic acid phenethyl ester (CAPE) and caffeoyldopamine (CADA).

CAPE is a rare natural compound with multifaceted bioactivities, including antimicrobial, antioxidant, anti-inflammatory, anti-carcinogenic and immunomodulatory properties [441]. It can be either naturally extracted from propolis with low yields or enzymatically synthesized via lipase-catalyzed esterification of caffeic acid and 2-phenylethanol in organic solvents [200,442]. In this work, we first identified a promiscuous O-acyltransferase Eht1 from *S. cerevisiae* that catalyzes the esterification of caffeoyl-CoA and 2-phenylethanol to synthesize CAPE. Then, by introducing synthetic pathways for caffeic acid and 2-phenylethanol, we constructed the entire artificial pathway for CAPE in *E. coli*. Expression the complete CAPE pathway could only produce 5.8 mg/L CAPE from glucose. With further optimization by balancing tyrosine and phenylalanine branches and alleviating caffeoyl-CoA degradation, the CAPE titer was

significantly improved by 60 %, to 9.3 mg/L. However, this titer was still very low for industrial scale-up. There are two major factors that may limit the CAPE production. The first one is the low catalytic activity of Eht1 towards CAPE synthesis. It has been reported that Eht1 might have low activity towards aromatic acyl-CoA and/or aromatic alcohols, as implied by the reduced activity of Eht1 towards longer chain acyl-CoA (>C8) [233]. As observed in our study, when fed with 0.5 g/L of caffeic acid and 2-phenylethanol, only 10 % conversion yield was achieved (52.9 mg/L CAPE). This was further underlined by the results that Eht1 exhibited low esterification activity for CAPE formation and showed innate thioesterase activity towards caffeoyl-CoA (**Figure 2.3B**). To address this limitation, protein engineering to improve the performance of Eht1 or screening O-acyltransferase with higher activity are two possible solutions. Although the authentic enzyme or natural pathway for CAPE still remains unknown, a BAHD acyltransferase OsPMT from *Oryza sativa* was recently cloned and utilized for CAPE biosynthesis in *E. coli*, which allowed production of 23.8 mg/L CAPE from glucose using a similar pathway as we established here [443]. This implied the possibility to further increase CAPE titer with improved O-acyltransferase. The second factor is the low production level of phenylalanine and tyrosine in *E. coli*. With engineered *E. coli*, phenylalanine and tyrosine could only be produced to 2-3 g/L in shake flask level [62,444,445]. In fermentation level, 26.78 g/L phenylalanine (3-L bioreactor) and 43.14 g/L tyrosine (5-L bioreactor) have been achieved in *E. coli* so far [446,447]. It is expected that, with catalytically efficient O-acyltransferases and aromatic amino acid overproducers, CAPE could be further enhanced.

CADA is a clovamide-type phenylpropenoic acid amide found in various plants, including cocoa (*Theobroma cacao*). It shows strong antioxidant activity and can be served as a β-2 adrenoceptor agonist [448,449]. By using a promiscuous tyramine *N*-hydroxycinnamoyltransferase

(THT) cloned from pepper (*Capsicum annum*), we established a synthetic CADA pathway in *E. coli* and achieved the microbial production of CADA with a titer of 369.1 mg/L from glucose. Both titers and yields of CADA were higher than that of CAPE (**Table 7.1**), which is consistent with the higher activity of THT for caffeic acid derived amide formation (**Figure 2.3B**). Currently, our best engineered *E. coli* strain could only produce 766.68 mg/L caffeic acid, which is also the highest titer achieved in *E. coli* thus far [208,450]. Considering that both caffeic acid and dopamine are generated from tyrosine, this suggested that the CADA pathway is quite efficient and CADA production almost reached the upper limit in our engineered strain. Due to the high pathway efficiency, future improvement of CADA production could be just focused on increasing carbon flux to tyrosine pathway.

**Developing pyruvate-driven growth-coupled Trp production**

In chapter 3, to overcome metabolic burden or genetic variations during microbial production, we established and validated a superior strategy to couple the product synthesis with cellular growth, which could render production obligatory for cell survival. As a proof-of-concept demonstration, we developed a pyruvate-driven system to couple tryptophan biosynthesis with cellular growth, engendering tryptophan production an integral part of host metabolism. This was mainly achieved by deleting major endogenous pyruvate-releasing pathways and enforcing pyruvate released from tryptophan biosynthesis as the major pyruvate donor for cell growth. In our shake flask tests, Trp titer reached 1.73 g/L with a yield of 0.132 g/g glycerol (25 % of the theoretical maximum yield) (**Table 3.2**). By redirecting the carbon flux from the Trp pathway to a synthetic muconic acid pathway, muconic acid titer reached 1.82 g/L with a yield of 0.122 g/g glycerol (12 % of the theoretical maximum yield). Compared with their parental strains, the

engineered strains with the pyruvate driven system enhanced tryptophan titer by 2.37-fold, and muconic acid titer by 4.7-fold.

As outlined in chapter 1, glucose and glycerol are the most common carbon sources for microbial fermentation due to their low costs (around 40 cents $lb^{-1}$ for glucose and 3.5-10 cents $lb^{-1}$ for crude glycerol) [451]. *E. coli* strains have been engineered to be capable of producing Trp up to 48.68 g/L from glucose in a 30-L bioreactor with a product yield reaching up to 0.211 g/g glucose [452,453]. Very recently, fed-batch production of Trp from glycerol in engineered *E. coli* was conducted in a 15-L bioreactor, which achieved a titer of 12.5 g/L and a yield of 0.06 g/g glycerol [296]. Although bioreactor-based fermentation was not performed in our study, the final engineered strain could produce Trp from glycerol with a relatively higher yield (0.132 g/g glycerol) in shake flask test, implying its potential for scale-up fermentation. However, one noteworthy observation was that both the biomass formation and glycerol utilization of the engineered *E. coli* strains were impaired due to extensive knockouts of pyruvate-releasing genes (**Figure S3.3**). This altogether may limit the final titer and productivity of Trp. One obvious next step to improve Trp production is to apply adaptive laboratory evolution (ALE) to these engineered pyruvate-driven Trp producers to select fast-growing strains. We hypothesize that, via growth coupling, Trp synthesis will be positively correlated with cell fitness, promoting high-performance variants to dominate the population and provoking directed evolution of cell robustness to lessen metabolic burden.

The established pyruvate-based growth-coupled strategy could be readily applied to different metabolic pathways or genetic contexts. First, the pyruvate-forming pathways are quite common in nature. The branches of shikimate pathway that start from chorismate or its derivatives involve pyruvate release; for instances, one pyruvate is released during conversion of chorismate to

anthranilate or 4-hydroxybenzoate, isochorismate to 2,3-dihydroxy-2,3-dihydrobenzoate or salicylate, and 4-amino-4-deoxychorismate to 4-aminobenzoate [294]. These products could serve as precursors for an expanded spectrum of new products of interest. In addition, as outlined in chapter 1, *meta*-cleavage of lignin monomers like catechol and protocatechuate also involves pyruvate release [28]. These metabolic pathways could be introduced into our pyruvate-driven platform strains for purpose-driven applications, like chemical bioproduction or lignin valorization. Second, the pyruvate-driven system can be theoretically applied to other commonly used carbon sources including glucose and xylose. To extend the feedstock flexibility, the key principle of the metabolic design is to eliminate pyruvate formation during sugar catabolism. For glucose, replacing the PEP-dependent PTS transport system with galactose permease (GalP) and ATP-dependent glucokinase (Glk) would preclude pyruvate formation as has been demonstrated previously [294,295]; while for xylose, the native ATP-dependent pathway would result in no additional pyruvate release. When demonstrating its feedstock flexibility using xylose, 1.31 g/L Trp was produced with a yield of 0.09 g/g xylose, which is 1.8-fold increase compared to parental strain (**Table 3.2**).

**Establishing C3-C5 diol production platform from charged amino acids**

In chapter 4, we endeavored to establish a diol production platform by extending charged amino acid catabolism, whose utilization for highly reduced short-chain products has remained unexplored. C3-C5 diols, including 1,3 propanediol (1,3 PDO), 1,4-butanediol (1,4 BDO) and 1,5-pentanediol (1,5 PDO), have wide use in commodity and fine chemical industries such as fuels, solvents, polymer monomers and pharmaceutical precursors [322-324]. These short-chain diols represent an annual commercial market over $ 7 billion with a nearly 7 % annual growth rate [325]. In this work, by combining charged amino acid catabolism with an efficient carboxylic acid

reductase (Car) mediated diol formation pathway, we constructed and demonstrated a novel metabolic platform to produce industrially important C3-C5 diols from amino acids in *E. coli*. The platform could convert a panel of seven amino acids into C3-C5 diols, based upon which we achieved the *de novo* production of 1.41 g/L 1,4 BDO and 0.97 g/L 1,5 PDO from glucose.

The commercialized biosynthetic pathways for both C3 and C4 diols have been established and validated in engineered *E. coli* [322]. 1,3 PDO is produced by utilizing a natural glycerol reduction pathway from *K. pneumoniae*, which has allowed the highest 1,3 PDO production to a titer of 135 g/L and a yield of 0.6 mol/mol glucose by a metabolically engineered *E. coli* from DuPont and Genencor International Inc. [454,455]. Our Asp-degradation pathway could only produce 247.3 mg/L 3HP from 3 g/L Asp, which is obviously less competitive than the glycerol-reduction pathway due to multiple reasons. These primarily include the low flux to Asp in wildtype *E. coli* and low activity of Car towards 3-hydroxypropionic acid (3-HP) (**Figure 4.2**). Although increasing flux to Asp could lead up to 32.3 g/L β-alanine [171], the low activity of Car is a bottleneck that is not easy to overcome. One potential strategy is to screen and identify Car homologs with preference towards 3-HP. Therefore, we mainly focused on metabolic engineering of 1,4 BDO and 1,5 PDO in the present work.

1,4 BDO is produced via a synthetic pathway that involves CoA-dependent reduction of succinate to 4-hydroxybutyrate (4-HB) and subsequent reduction of 4-HB to 1,4 BDO by a bifunctional alcohol reductase [129]. With systematic engineering of enzyme, pathway, metabolic network, host organism, and fermentation process, the best 1,4 BDO-producing *E. coli* strain could produce more than 120 g/L with a yield of 0.40 g/g (80% of theoretical) [129,339]. A 1,4 BDO yield hitting 0.35 g/g glucose would be regarded as commercially-relevant [339]. With our established Car-based 1,4 BDO pathway, the current 1,4 BDO titer reached 1.41 g/L with a yield

of 0.14 mol/mol glucose (0.07 g/g glucose) by wildtype *E. coli* during shake flask test (**Figure 4.5B**). These suggested that further improvement of product yield is needed. Interestingly, during pathway establishment, the wild-type *E. coli* containing the commercialized 1,4 BDO pathway could only produce 117.16 mg/L (1,3 mM) 1,4 BDO in shake flask test [129]. Subsequent host engineering by deleting competing pathways and fed-batch fermentation optimization in microaerobic condition played a major role in improving 1,4 BDO production and significantly boosted 1,4 BDO titer to 18 g/L in 5 d in a 2-L bioreactor, which laid the foundation for following improvement and scale-up [129]. This suggested that optimizing microbial host and fermentation process might improve the production performance of Car-based 1,4 BDO pathway.

1,5 PDO serves as monomer for polyesters, thermoplastic polyurethanes, and plasticizers [344]. Because of its high price ($9,700/ton) than 1,4 BDO ($1,600-2,800/ton) and low production capacity (about 3000 tons per year), the application of 1,5 PDO was limited [328]. Currently, 1,5 PDO is chemically produced from biomass-derived furfural and tetrahydrofurfuryl alcohol via expensive metal catalysts [345,346]. Considering the high price of furfural ($900-1000/ton), direct conversion of cheaper and high-volume glucose to 1,5 PDO is more appealing. However, the lack of biosynthetic routes for 1,5 PDO hindered the direct conversion. In this work, by harnessing the lysine degradation pathway and Car-based reduction of 5-hydroxyvalerate, an efficient Lys-to-1,5 PDO pathway was established. The high lysine-to-1,5 PDO conversion (0.56 g/g Lys, 78.2 % of the theoretical maximum) suggested its potential for 1,5 PDO production, especially when lysine supply was not a limiting step (**Figure S4.4**). Given lysine production is subjected to multiple level repression in *E. coli*, we over-expressed the lysine biosynthetic pathway and increased carbon flux to lysine precursor oxaloacetate (*OAA*) by deleting the

glyoxylate cycle repressor *iclR*, which together significantly increased 1,5 PDO production to 0.97 g/L at 48 h with a yield of 0.08 mol/mol glucose (**Figure 4.6**). Although the commercial benchmark for industrial production of 1,5 PDO is yet unknown, the current titer and yield still need be further improved before scale-up fermentation.

The amino acid based diol production platform will present an opportunity for metabolic engineering of short-chain diols in *E. coli* and other industrial microbial hosts. Direct transfer of the diol pathways into an glutamate or lysine overproducer will permit the hijacking of the enhanced and optimized amino acid pathways for diol production. Since glutamate and lysine titer achieved in *E.coli* are relatively low (43.5 g/L and 12.23 g/L, respectively) [456], superior glutamate or lysine producers are more attractive. *Corynebacterium glutamicum* is one of the most investigated industrial over-producers for glutamate and lysine, whose titers could reach 100 g/L (with a yield of 0.60 g/g glucose) and 120 g/L (with a yield of 0.55 g/g glucose), respectively [142-144]. We envision that this diol platform demonstrated here in *E. coli* can be readily implemented into *C. glutamicum* and potentially can achieve industrially-relevant titer and yield for 1,5 PDO.

**Designing acetyl-CoA mediated carbon-chain extension for C5 dicarboxylates**

Among the monomeric feedstocks for polyester synthesis, dicarboxylic acids have received significant attention as platform chemicals. As outlined in chapter 1, examples of organic acids that are currently accessible via biotransformation or fermentation and are of commercial interest include glycolate (C2), 3-hydroxypropionate (C3), lactate (C3), succinate (C4), itaconate (C5), *cis,cis*-muconate (C6), and adipate (C6). Of these monomers, lactic acid is already produced on a scale of approximately 400,000 tons annually, led by companies such as Natureworks, Purac, and Galactic [457]. Bio-derived succinate, itaconate, and adipate have also been explored

commercially to produce poly(butylene succinate) (PBS), polyamide 54, polyitaconate and nylon-4,6, but to a much smaller extent [390,458]. Extending this molecular toolbox, especially in regard to odd numbered dicarboxylates, will afford the development of new polymers that have useful property profiles that will start to replace petroleum based plastics. In chapter 5 and 6, we designed novel pathways via acetyl-CoA mediated "+1" carbon chain extension for two C5 dicarboxylates, one that produced glutarate by extending from α-ketoglutarate (α-KG) and one that produced 2-methylsuccinate by extending from pyruvate. Both glutarate and 2-methylsuccinate are important C5 dicarboxylates with broad applications, as direct precursors for 1,5 PDO and 2-methyl-1,4-butanediol (2-methyl-1,4 BDO), or as chemical synthons for biodegradable polymers like nylon-5,5 and poly(butylene succinate-co-butylene 2-methyl succinate) [354,368,459,460].

The conventional biosynthetic route for glutarate resembles the upstream of the 1,5 PDO pathway, which involves degradation of lysine to glutarate semialdehyde (GSA) and subsequent conversion of GSA to glutarate [361]. In chapter 5, a novel six-step glutarate pathway was designed, which involved condensation of acetyl-CoA and α-KG to α-ketoadipate (α-KA) and subsequent decarboxylation of α-KA to glutarate (**Figure 5.1**). Although this novel glutarate pathway is shorter, its production performance was lower than the lysine based pathway in microbial chassis *E. coli* (0.3 g/L vs 0.82 g/L) [368]. Implementation of CRISPR/dCas9 mediated interference of essential genes (*sucAB*) in TCA cycle enhanced accumulation of α-KG, which increased glutarate titer by 40 % to 0.42 g/L (**Figure 5.5**). Recently, the lysine based glutarate pathway was significantly improved either by increasing lysine pathway in *E. coli* (54.5 g/L with a yield of 0.54 mol/mol glucose), or by introducing it into a lysine overproducing *C. glutamicum* (90 g/L with a yield of 0.70 mol/mol glucose) [71,461]. This underlined our previously proposed hypothesis

that introduction of 1,5 PDO pathway into a lysine overproduction platform could potentially boost 1,5 PDO production. Despite functional, the novel glutarate pathway needs further optimization for scale-up production. The first one is to increase the α-KG supply. As observed in our study, exogenously feeding α-KG or using CRISPRi to repress α-KG consumption both improved glutarate production (**Figure 5.3 and 5.5**). Since both α-KG and acetyl-CoA are abundant in *E. coli* [340], the second optimization is to increase the homocitrate synthase (HCS) activity that catalyzes the first committed step of condensation of acetylCoA with α-KG. Third, the promiscuous α-KA decarboxylase KivD could only enable 40 % conversion of α-KA into glutarate, which may limit glutarate production. Enzyme engineering of KivD is feasible, since both the structural information and substrate binding pocket of KivD are well-characterized [170]. Moreover, KivD mutant library with expanded substrate binding pocket are available, providing a basis for further engineering and screening [169,462].

In chapter 6, we designed a novel pathway for another branched chain C5 dicarboxylate, 2-methylsuccinate. A seven-step CoA-dependent 2-methylsuccinate pathway was previously established in *Methylobacterium extorquens*, which allowed production of 0.26 g/L 2-methylsuccinate [406,411]. In this work, we constructed a three-step pathway that involved condensation of acetyl-CoA and pyruvate to citramalate. Citramalate underwent isomerization and reduction to 2-methylsuccinate (**Figure 6.1**). Initial validation of the 2-methylsuccinate pathway using YqjM could only produce 0.35 g/L (**Figure 6.2**), suggesting the low catalytic activity of pathway enzymes. The first enzyme citramalate synthase (CimA*) is very efficient in citramalate production, which could afford a titer of 3.96 g/L in 48 h (**Figure 6.2**). This indicated that YqjM might be a bottleneck enzyme. Indeed, mining of YqjM homologs identified a superior enoate reductase KpnER from *K. pneumoniae*, with which 0.96 g/L 2-methylsuccinate

was produced from glucose (**Figure 6.3B**). However, citramalate was significantly accumulated with titers higher that 2-methylsuccinate, suggesting inefficient conversion of citramalate. To address the issue, combinatorial optimization via cofactor regeneration, host strain engineering and microaerobic cultivation eventually improved the 2-methylsuccinate titer to 3.61 g/L with a yield of 0.36 mol/mol glucose (**Figure 6.6**). 2-Methylsuccinate can be chemically produced from ruthenium-based hydrogenation of itaconate with 99 % conversion yield [408]. The market of itaconate could reach about 170,000 tons/year in 2025 and is sold at an average selling price of \$1.5/kg [463]. Considering that the overall conversion yield of itaconate to 2-methylsuccinate is 99 % (w/w), the estimated raw material cost in the chemical process is \$1.52/kg 2-methylsuccinate. When using the biosynthetic pathways, glucose is used as precursor and it is sold at a price of about \$0.39/kg [394]. With current yield of 0.36 mol/mol (i.e. 0.26 g/g), the raw material cost in the bioprocess is then \$1.5/kg 2-methylsuccinate, which is comparable with the chemical process. Thus, to establish a 2-methylsuccinate bioprocess that can compete with the chemical process, a product yield of at least 0.36 mol/mol glucose is needed. Further increase of product yield is possible, considering the significant amount of citramalate unconverted to 2-methylsuccinate (**Figure 6.6C**). This was probably caused by the preferred re-hydratase activity of isopropylmalate isomerase (LeuCD). Thus, optimization of LeuCD either through extensive enzyme mining or engineering might potentially further increase 2-MSA production in future.

**Conclusion and outlook**

In conclusion, we have developed several new synthetic pathways for both natural and non-natural chemical products by extending existing metabolic pathways, especially amino acid pathways in *E. coli*. Due to the structural diversity of amino acids, we designed counterpart synthetic pathways and applied different synthetic biology strategies to increase carbon flux to

target products. As outlined in **Figure 1.8** and **Table 7.1**, our target products covered both aromatic amino acids derived natural products, and charged amino acids or central metabolites derived commodity chemicals. In this work, we mainly focused on new pathway design and validation, and most of the work were conducted in shake flask levels. Our work either enriched the pathway repertoire or expanded chemical space in microbial system. To test the industrial potentials, we do believe bioreactor based fermentation is necessary. But before that, most of production pathways or strains still need to be further optimized.

Based upon current research, new products with high commercial values could be selected for further production optimization. These include CADA with pharmaceutical value, 1,5 PDO and 2-methylsuccinate with wide applications in polymer industry. Moreover, CADA and 1,5 PDO pathways are highly efficient with no bottleneck enzymes, while 2-methylsuccinate could be produced with relatively high titers by simply manipulating hosts and cultivation conditions. To further improve product titers and yields, the first stage will be to optimize metabolic network or select suitable microbial hosts. For products derived from amino acids, the best way is to construct or simply change to the amino acid overproducer. For CADA or 1,5 PDO, a tyrosine or lysine overproducer would drastically facilitate their production. Considering the need to express synthetic pathway genes on plasmids, a natural or mutant amino acid overproducer host strain would reduce the metabolic burden of expressing excessive enzymes. With amino acid overproducers and efficient synthetic pathways, the carbon flux can be "pulled" towards target products without constraint. For 2-methylsuccinate overproduction, the major challenge is to "push" carbon flux from citramalate to 2-methylsucinate with an optimal isopropylmalate isomerase or excess supply of NADH for enoate reductase. The second stage is to optimize culture conditions in shake flask levels, including medium composition, cultivation temperature,

inducer concentration and aerobic condition, etc. With completion of these two stages, the final stage of bioreactor fermentation will be considered. The synthetic pathways established in this work would undergo these three stages and work as a basis for future scale-up production.

**Table 7.1 Summary of major products produced in this work.**

| Product | Titer/yield | Industrial value |
|---|---|---|
| Caffeic acid phenethyl ester (CAPE) | 9.3 mg/L, 0.47 mg/g glucose | • Antimicrobial, antioxidant, anti-inflammatory, and cytotoxic properties<br>• Adjuvant to anti-tumor drugs, such as paclitaxel (Taxol) |
| Caffeoyldopamine (CADA) | 369.1 mg/L, 18.46 mg/g glucose | • Potent antioxidant activity<br>• β-2 adrenoceptor agonist |
| Tryptophan | 1.73 g/L, 0.13 g/g glycerol | • An essential aromatic-group amino acid for human<br>• A precursor to the neurotransmitter serotonin and niacin (vitamin B3)<br>• A precursor to monoterpene indole alkaloids (MIAs) |
| *cis, cis*-Muconic acid | 1.82 g/L, 0.12 g/g glycerol | • A platform chemical for polymer like nylon-6,6, polyurethane, and polyethylene terephthalate (PET) |
| 1,4 butanediol (1,4 BDO) | 1.41 g/L, 0.14 mol/mol glucose | • A precursor for elastic fibers and polyurethanes<br>• A precursor for tetrahydrofuran (THF) |

| | | |
|---|---|---|
| 1,5 pentanediol (1,5 PDO) | 0.97 g/L, 0.08 mol/mol glucose | • A monomer for polyesters, thermoplastic polyurethanes, and plasticizers |
| Glutarate | 0.42 g/L, 0.02 g/g glucose | • A monomer for polymers like nylon-5,5 |
| 2-Methylsuccinate | 3.61 g/L, 0.36 mol/mol glucose | • A precursor for 2-methyl-γ-butyrolactones (MGBLs), 2-methyl-1,4-butanediol (MBDO)<br>• A monomer for polymers like poly(butylene succinate-co-butylene 2-methyl succinate) |

**References**


1. Stöcker, M. Biofuels and biomass-to-liquid fuels in the biorefinery: Catalytic conversion of lignocellulosic biomass using porous materials. *Angewandte Chemie International Edition* **47**, 9200-9211 (2008).
2. Lucia, L.A. Lignocellulosic biomass: A potential feedstock to replace petroleum. *BioResources* **3**, 981-982 (2008).
3. Cherubini, F. The biorefinery concept: using biomass instead of oil for producing energy and chemicals. *Energy Conversion and Management* **51**, 1412-1421 (2010).
4. Bailey, J.E. Toward a science of metabolic engineering. *Science* **252**, 1668-1675 (1991).
5. Stephanopoulos, G., Aristidou, A.A. & Nielsen, J. *Metabolic engineering: principles and methodologies*, (Elsevier, 1998).
6. Keasling, J.D. Manufacturing molecules through metabolic engineering. *Science* **330**, 1355-1358 (2010).
7. Zhang, F., Rodriguez, S. & Keasling, J.D. Metabolic engineering of microbial pathways for advanced biofuels production. *Current opinion in biotechnology* **22**, 775-783 (2011).
8. Wang, J., Guleria, S., Koffas, M.A. & Yan, Y. Microbial production of value-added nutraceuticals. *Current opinion in biotechnology* **37**, 97-104 (2016).
9. Guerriero, G., Hausman, J.F., Strauss, J., Ertan, H. & Siddiqui, K.S. Lignocellulosic biomass: Biosynthesis, degradation, and industrial utilization. *Engineering in Life Sciences* **16**, 1-16 (2016).
10. Nanda, S., Mohammad, J., Reddy, S.N., Kozinski, J.A. & Dalai, A.K. Pathways of lignocellulosic biomass conversion to renewable fuels. *Biomass Conversion and Biorefinery* **4**, 157-191 (2014).

11. Lee, J. Biological conversion of lignocellulosic biomass to ethanol. *Journal of biotechnology* **56**, 1-24 (1997).

12. Sjostrom, E. *Wood chemistry: fundamentals and applications*, (Elsevier, 2013).

13. Corma, A., Iborra, S. & Velty, A. Chemical routes for the transformation of biomass into chemicals. *Chemical reviews* **107**, 2411-2502 (2007).

14. Hatakeyama, H. & Hatakeyama, T. Lignin structure, properties, and applications. in *Biopolymers* 1-63 (Springer, 2009).

15. Ragauskas, A.J. et al. Lignin valorization: improving lignin processing in the biorefinery. *Science* **344**, 1246843 (2014).

16. Zhao, Y., Deng, L., Liao, B., Fu, Y. & Guo, Q.-X. Aromatics production via catalytic pyrolysis of pyrolytic lignins from bio-oil. *Energy & Fuels* **24**, 5735-5740 (2010).

17. Leonowicz, A. et al. Biodegradation of lignin by white rot fungi. *Fungal genetics and biology* **27**, 175-185 (1999).

18. Arantes, V., Jellison, J. & Goodell, B. Peculiarities of brown-rot fungi and biochemical Fenton reaction with regard to their potential as a model for bioprocessing biomass. *Applied Microbiology and Biotechnology* **94**, 323-338 (2012).

19. Bugg, T.D., Ahmad, M., Hardiman, E.M. & Rahmanpour, R. Pathways for degradation of lignin in bacteria and fungi. *Natural product reports* **28**, 1883-1896 (2011).

20. Abdelaziz, O.Y. et al. Biological valorization of low molecular weight lignin. *Biotechnology Advances* **34**, 1318-1346 (2016).

21. Brebu, M. & Vasile, C. Thermal degradation of lignin—a review. *Cellulose Chemistry & Technology* **44**, 353 (2010).

22. Nowakowski, D.J., Bridgwater, A.V., Elliott, D.C., Meier, D. & de Wild, P. Lignin fast

pyrolysis: results from an international collaboration. *Journal of Analytical and Applied Pyrolysis* **88**, 53-72 (2010).

23. Radoykova, T., Nenkova, S. & Valchev, I. Black liquor lignin products, isolation and characterization. *Journal of Chemical Technology and Metallurgy* **48**, 524-529 (2013).
24. Choi, H.S. & Meier, D. Fast pyrolysis of Kraft lignin—Vapor cracking over various fixed-bed catalysts. *Journal of Analytical and Applied Pyrolysis* **100**, 207-212 (2013).
25. Kosa, M. & Ragauskas, A.J. Bioconversion of lignin model compounds with oleaginous Rhodococci. *Applied microbiology and biotechnology* **93**, 891-900 (2012).
26. Wu, W. et al. Lignin valorization: Two hybrid biochemical routes for the conversion of polymeric lignin into value-added chemicals. *Scientific reports* **7**, 8420 (2017).
27. Linger, J.G. et al. Lignin valorization through integrated biological funneling and chemical catalysis. *Proceedings of the National Academy of Sciences* **111**, 12013-12018 (2014).
28. Johnson, C.W. & Beckham, G.T. Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. *Metabolic engineering* **28**, 240-247 (2015).
29. Beckham, G.T., Johnson, C.W., Karp, E.M., Salvachúa, D. & Vardon, D.R. Opportunities and challenges in biological lignin valorization. *Current opinion in biotechnology* **42**, 40-53 (2016).
30. Dharmadi, Y., Murarka, A. & Gonzalez, R. Anaerobic fermentation of glycerol by Escherichia coli: a new platform for metabolic engineering. *Biotechnology and bioengineering* **94**, 821-829 (2006).
31. Da Silva, G.P., Mack, M. & Contiero, J. Glycerol: a promising and abundant carbon

source for industrial microbiology. *Biotechnology advances* **27**, 30-39 (2009).

32. Bagnato, G., Iulianelli, A., Sanna, A. & Basile, A. Glycerol production and transformation: a critical review with particular emphasis on glycerol reforming reaction for producing hydrogen in conventional and membrane reactors. *Membranes* **7**, 17 (2017).
33. Chen, Z. & Liu, D. Toward glycerol biorefinery: metabolic engineering for the production of biofuels and chemicals from glycerol. *Biotechnology for biofuels* **9**, 205 (2016).
34. Trinh, C.T. & Srienc, F. Metabolic engineering of Escherichia coli for efficient conversion of glycerol to ethanol. *Appl. Environ. Microbiol.* **75**, 6696-6705 (2009).
35. Blankschien, M.D., Clomburg, J.M. & Gonzalez, R. Metabolic engineering of Escherichia coli for the production of succinate from glycerol. *Metabolic engineering* **12**, 409-419 (2010).
36. Yazdani, S.S. & Gonzalez, R. Engineering Escherichia coli for the efficient conversion of glycerol to ethanol and co-products. *Metabolic engineering* **10**, 340-351 (2008).
37. Clomburg, J.M. & Gonzalez, R. Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. *Trends in biotechnology* **31**, 20-28 (2013).
38. Jung, W.S., Kang, J.H., Chu, H.S., Choi, I.S. & Cho, K.M. Elevated production of 3-hydroxypropionic acid by metabolic engineering of the glycerol metabolism in Escherichia coli. *Metabolic engineering* **23**, 116-122 (2014).
39. González-Pajuelo, M. et al. Metabolic engineering of Clostridium acetobutylicum for the industrial production of 1, 3-propanediol from glycerol. *Metabolic engineering* **7**, 329-336 (2005).
40. Maervoet, V.E., De Mey, M., Beauprez, J., De Maeseneire, S. & Soetaert, W.K.

Enhancing the microbial conversion of glycerol to 1, 3-propanediol using metabolic engineering. *Organic Process Research & Development* **15**, 189-202 (2010).

41. Duerre, P. & Eikmanns, B.J. C1-carbon sources for chemical and fuel production by microbial gas fermentation. *Current opinion in biotechnology* **35**, 63-72 (2015).
42. Gudmundsson, S. & Nogales, J. Cyanobacteria as photosynthetic biocatalysts: a systems biology perspective. *Molecular BioSystems* **11**, 60-70 (2015).
43. Carroll, A.L., Case, A.E., Zhang, A. & Atsumi, S. Metabolic engineering tools in model cyanobacteria. *Metabolic engineering* **50**, 47-56 (2018).
44. Antonovsky, N. et al. Sugar synthesis from CO2 in Escherichia coli. *Cell* **166**, 115-125 (2016).
45. Lee, O.K., Hur, D.H., Nguyen, D.T.N. & Lee, E.Y. Metabolic engineering of methanotrophs and its application to production of chemicals and biofuels from methane. *Biofuels, Bioproducts and Biorefining* **10**, 848-863 (2016).
46. Conrado, R.J. & Gonzalez, R. Envisioning the bioconversion of methane to liquid fuels. *Science* **343**, 621-623 (2014).
47. Bennett, R.K., Steinberg, L.M., Chen, W. & Papoutsakis, E.T. Engineering the bioconversion of methane and methanol to fuels and chemicals in native and synthetic methylotrophs. *Current opinion in biotechnology* **50**, 81-93 (2018).
48. Müller, J.E. et al. Engineering Escherichia coli for methanol conversion. *Metabolic engineering* **28**, 190-201 (2015).
49. Whitaker, W.B. et al. Engineering the biological conversion of methanol to specialty chemicals in Escherichia coli. *Metabolic engineering* **39**, 49-59 (2017).
50. Woolston, B.M., King, J.R., Reiter, M., Van Hove, B. & Stephanopoulos, G. Improving

formaldehyde consumption drives methanol assimilation in engineered E. coli. *Nature communications* **9**, 2387 (2018).

51. Yu, H. & Liao, J.C. A modified serine cycle in Escherichia coli coverts methanol and CO 2 to two-carbon compounds. *Nature communications* **9**, 3992 (2018).

52. Huo, Y.-X. et al. Conversion of proteins into biofuels by engineering nitrogen flux. *Nature biotechnology* **29**, 346 (2011).

53. Wernick, D.G. & Liao, J.C. Protein-based biorefining: metabolic engineering for production of chemicals and fuel with regeneration of nitrogen fertilizers. *Applied microbiology and biotechnology* **97**, 1397-1406 (2013).

54. Harun, R., Singh, M., Forde, G.M. & Danquah, M.K. Bioprocess engineering of microalgae to produce a variety of consumer products. *Renewable and Sustainable Energy Reviews* **14**, 1037-1047 (2010).

55. Harun, R., Danquah, M.K. & Forde, G.M. Microalgal biomass as a fermentation feedstock for bioethanol production. *Journal of Chemical Technology & Biotechnology* **85**, 199-203 (2010).

56. Choi, K.-Y., Wernick, D.G., Tat, C.A. & Liao, J.C. Consolidated conversion of protein waste into biofuels and ammonia using Bacillus subtilis. *Metabolic engineering* **23**, 53-61 (2014).

57. Fischer, C.R., Klein-Marcuschamer, D. & Stephanopoulos, G. Selection and optimization of microbial hosts for biofuels production. *Metabolic engineering* **10**, 295-304 (2008).

58. von Kamp, A. & Klamt, S. Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nature communications* **8**, 15956 (2017).

59. Pontrelli, S. et al. Escherichia coli as a host for metabolic engineering. *Metabolic*

*engineering* **50**, 16-46 (2018).

60. Lee, S.Y. et al. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis* **2**, 18 (2019).

61. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894 (2009).

62. Na, D. et al. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nature biotechnology* **31**, 170 (2013).

63. Sander, J.D. & Joung, J.K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology* **32**, 347 (2014).

64. Wendisch, V.F., Bott, M. & Eikmanns, B.J. Metabolic engineering of Escherichia coli and Corynebacterium glutamicum for biotechnological production of organic acids and amino acids. *Current opinion in microbiology* **9**, 268-274 (2006).

65. Vogt, M. et al. Pushing product formation to its limit: metabolic engineering of Corynebacterium glutamicum for L-leucine overproduction. *Metabolic engineering* **22**, 40-52 (2014).

66. Peters-Wendisch, P. et al. Metabolic engineering of Corynebacterium glutamicum for L-serine production. *Appl. Environ. Microbiol.* **71**, 7139-7144 (2005).

67. Park, S.H. et al. Metabolic engineering of Corynebacterium glutamicum for L-arginine production. *Nature communications* **5**, 4618 (2014).

68. Rohles, C.M., 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. *Microbial Cell Factories* **15**, 154 (2016).

69. Buschke, N. et al. Systems metabolic engineering of xylose-utilizing Corynebacterium glutamicum for production of 1, 5-diaminopentane. *Biotechnology journal* **8**, 557-570 (2013).

70. Takahashi, C. et al. Robust production of gamma-amino butyric acid using recombinant Corynebacterium glutamicum expressing glutamate decarboxylase from Escherichia coli. *Enzyme and microbial technology* **51**, 171-176 (2012).

71. Rohles, C.M. et al. A bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6, 5 using metabolically engineered Corynebacterium glutamicum. *Green chemistry* **20**, 4662-4674 (2018).

72. Krivoruchko, A. & Nielsen, J. Production of natural products through metabolic engineering of Saccharomyces cerevisiae. *Current opinion in biotechnology* **35**, 7-15 (2015).

73. Pompon, D., Louerat, B., Bronine, A. & Urban, P. [6] Yeast expression of animal and plant P450s in optimized redox environments. in *Methods in enzymology*, Vol. 272 51-64 (Elsevier, 1996).

74. Fossati, E. et al. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae. *Nature communications* **5**, 3283 (2014).

75. Hawkins, K.M. & Smolke, C.D. Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Nature chemical biology* **4**, 564 (2008).

76. Minami, H. et al. Microbial production of plant benzylisoquinoline alkaloids. *Proceedings of the National Academy of Sciences* **105**, 7393-7398 (2008).

77. Zhu, Q. & Jackson, E.N. Metabolic engineering of Yarrowia lipolytica for industrial

applications. *Current opinion in biotechnology* **36**, 65-72 (2015).

78. Ledesma-Amaro, R. et al. Metabolic engineering of Yarrowia lipolytica to produce chemicals and fuels from xylose. *Metabolic engineering* **38**, 115-124 (2016).

79. Qiao, K. et al. Engineering lipid overproduction in the oleaginous yeast Yarrowia lipolytica. *Metabolic engineering* **29**, 56-65 (2015).

80. Christen, S. & Sauer, U. Intracellular characterization of aerobic glucose metabolism in seven yeast species by 13C flux analysis and metabolomics. *FEMS yeast research* **11**, 263-272 (2011).

81. Xue, Z. et al. Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature biotechnology* **31**, 734 (2013).

82. Driouch, H., Sommer, B. & Wittmann, C. Morphology engineering of Aspergillus niger for improved enzyme production. *Biotechnology and bioengineering* **105**, 1058-1068 (2010).

83. Steiger, M.G., Rassinger, A., Mattanovich, D. & Sauer, M. Engineering of the citrate exporter protein enables high citric acid production in Aspergillus niger. *Metabolic engineering* **52**, 224-231 (2019).

84. Steiger, M.G., Blumhoff, M.L., Mattanovich, D. & Sauer, M. Biochemistry of microbial itaconic acid production. *Frontiers in microbiology* **4**, 23 (2013).

85. Englund, E., Andersen-Ranberg, J., Miao, R., Hamberger, B.r. & Lindberg, P. Metabolic engineering of Synechocystis sp. PCC 6803 for production of the plant diterpenoid manoyl oxide. *ACS synthetic biology* **4**, 1270-1278 (2015).

86. Yoshikawa, K., Toya, Y. & Shimizu, H. Metabolic engineering of Synechocystis sp. PCC 6803 for enhanced ethanol production based on flux balance analysis. *Bioprocess and*

*biosystems engineering* **40**, 791-796 (2017).

87. Lin, P.-C., Saha, R., Zhang, F. & Pakrasi, H.B. Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium Synechocysti s sp. PCC 6803. *Scientific reports* **7**, 17503 (2017).

88. Savakis, P. & Hellingwerf, K.J. Engineering cyanobacteria for direct biofuel production from CO2. *Current opinion in biotechnology* **33**, 8-14 (2015).

89. Carpine, R. et al. Genetic engineering of Synechocystis sp. PCC6803 for poly-β-hydroxybutyrate overproduction. *Algal research* **25**, 117-127 (2017).

90. Hansen, C.A., Dean, A.B., Draths, K. & Frost, J. Synthesis of 1, 2, 3, 4-tetrahydroxybenzene from D-glucose: exploiting myo-inositol as a precursor to aromatic chemicals. *Journal of the American Chemical Society* **121**, 3799-3800 (1999).

91. Yamaoka, M., Osawa, S., Morinaga, T., Takenaka, S. & Yoshida, K.-i. A cell factory of Bacillus subtilis engineered for the simple bioconversion of myo-inositol to scyllo-inositol, a potential therapeutic agent for Alzheimer's disease. *Microbial cell factories* **10**, 69 (2011).

92. Moon, T.S., Yoon, S.-H., Lanza, A.M., Roy-Mayhew, J.D. & Prather, K.L.J. Production of glucaric acid from a synthetic pathway in recombinant Escherichia coli. *Appl. Environ. Microbiol.* **75**, 589-595 (2009).

93. Werpy, T. & Petersen, G. Top value added chemicals from biomass: volume I--results of screening for potential candidates from sugars and synthesis gas. (National Renewable Energy Lab., Golden, CO (US), 2004).

94. Shiue, E. & Prather, K.L. Improving D-glucaric acid production from myo-inositol in E. coli by increasing MIOX stability and myo-inositol transport. *Metabolic engineering* **22**,

22-31 (2014).

95. Moon, T.S., Dueber, J.E., Shiue, E. & Prather, K.L.J. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. *Metabolic engineering* **12**, 298-305 (2010).

96. Brockman, I.M. & Prather, K.L. Dynamic knockdown of E. coli central metabolism for redirecting fluxes of primary metabolites. *Metabolic engineering* **28**, 104-113 (2015).

97. Gupta, A., Reizman, I.M.B., Reisch, C.R. & Prather, K.L. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature biotechnology* **35**, 273 (2017).

98. Banerjee, A. & Sharkey, T. Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. *Natural product reports* **31**, 1043-1055 (2014).

99. Liu, H. et al. MEP pathway-mediated isopentenol production in metabolically engineered Escherichia coli. *Microbial cell factories* **13**, 135 (2014).

100. Zhao, Y. et al. Biosynthesis of isoprene in Escherichia coli via methylerythritol phosphate (MEP) pathway. *Applied microbiology and biotechnology* **90**, 1915 (2011).

101. Ajikumar, P.K. et al. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* **330**, 70-74 (2010).

102. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nature biotechnology* **23**, 612 (2005).

103. Becker, J. & Wittmann, C. Systems metabolic engineering of Escherichia coli for the heterologous production of high value molecules—a veteran at new shores. *Current opinion in biotechnology* **42**, 178-188 (2016).

104. Ji, X.-J., Huang, H. & Ouyang, P.-K. Microbial 2, 3-butanediol production: a state-of-the-art review. *Biotechnology advances* **29**, 351-364 (2011).

105. Lan, E.I. & Liao, J.C. Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. *Bioresource technology* **135**, 339-349 (2013).

106. Baez, A., Cho, K.-M. & Liao, J.C. High-flux isobutanol production using engineered Escherichia coli: a bioreactor study with in situ product removal. *Applied microbiology and biotechnology* **90**, 1681-1690 (2011).

107. Azhar, S.H.M. et al. Yeasts in sustainable bioethanol production: A review. *Biochemistry and Biophysics Reports* **10**, 52-61 (2017).

108. Kennes, D., Abubackar, H.N., Diaz, M., Veiga, M.C. & Kennes, C. Bioethanol production from biomass: carbohydrate vs syngas fermentation. *Journal of Chemical Technology & Biotechnology* **91**, 304-317 (2016).

109. Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y. & Nielsen, J. Microbial acetyl-CoA metabolism and metabolic engineering. *Metabolic engineering* **28**, 28-42 (2015).

110. Uchino, K., Saito, T., Gebauer, B. & Jendrossek, D. Isolated poly (3-hydroxybutyrate)(PHB) granules are complex bacterial organelles catalyzing formation of PHB from acetyl coenzyme A (CoA) and degradation of PHB to acetyl-CoA. *Journal of bacteriology* **189**, 8250-8256 (2007).

111. Tseng, H.-C., Martin, C.H., Nielsen, D.R. & Prather, K.L.J. Metabolic engineering of Escherichia coli for enhanced production of (R)-and (S)-3-hydroxybutyrate. *Applied and environmental microbiology* **75**, 3137-3145 (2009).

112. May, A. et al. A modified pathway for the production of acetone in Escherichia coli. *Metabolic engineering* **15**, 218-225 (2013).

113. Atsumi, S. & Liao, J.C. Metabolic engineering for advanced biofuels production from Escherichia coli. *Current opinion in biotechnology* **19**, 414-419 (2008).

114. Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D. & Keasling, J.D. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature biotechnology* **21**, 796 (2003).

115. Withers, S.T. & Keasling, J.D. Biosynthesis and engineering of isoprenoid small molecules. *Applied microbiology and biotechnology* **73**, 980-990 (2007).

116. Lee, P. & Schmidt-Dannert, C. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. *Applied Microbiology and Biotechnology* **60**, 1-11 (2002).

117. Zheng, Y. et al. Metabolic engineering of Escherichia coli for high-specificity production of isoprenol and prenol as next generation of biofuels. *Biotechnology for biofuels* **6**, 57 (2013).

118. Tsuruta, H. et al. High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in Escherichia coli. *PLoS One* **4**, e4489 (2009).

119. Paddon, C.J. & Keasling, J.D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology* **12**, 355 (2014).

120. Atsumi, S. & Liao, J.C. Directed evolution of Methanococcus jannaschii citramalate synthase for biosynthesis of 1-propanol and 1-butanol by Escherichia coli. *Applied and environmental microbiology* **74**, 7802-7808 (2008).

121. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *nature* **451**, 86-89 (2008).

122. Vuoristo, K.S., Mars, A.E., Sanders, J.P., Eggink, G. & Weusthuis, R.A. Metabolic engineering of TCA cycle for production of chemicals. *Trends in biotechnology* **34**, 191-197 (2016).

123. Li, Z.-J., Hong, P.-H., Da, Y.-Y., Li, L.-K. & Stephanopoulos, G. Metabolic engineering of Escherichia coli for the production of L-malate from xylose. *Metabolic engineering* **48**, 25-32 (2018).

124. Song, C.W., Kim, D.I., Choi, S., Jang, J.W. & Lee, S.Y. Metabolic engineering of Escherichia coli for the production of fumaric acid. *Biotechnology and bioengineering* **110**, 2025-2034 (2013).

125. Jantama, K. et al. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. *Biotechnology and bioengineering* **99**, 1140-1153 (2008).

126. Yin, X. et al. Metabolic engineering in the biotechnological production of organic acids in the tricarboxylic acid cycle of microorganisms: advances and prospects. *Biotechnology advances* **33**, 830-841 (2015).

127. Choi, S., Kim, H.U., Kim, T.Y. & Lee, S.Y. Systematic engineering of TCA cycle for optimal production of a four-carbon platform chemical 4-hydroxybutyric acid in Escherichia coli. *Metabolic engineering* **38**, 264-273 (2016).

128. Yim, H. et al. Metabolic engineering of Escherichia coli for direct production of 1, 4-butanediol. *Nature chemical biology* **7**, 445-452 (2011).

129. Yim, H. et al. Metabolic engineering of Escherichia coli for direct production of 1, 4-butanediol. *Nature chemical biology* **7**, 445 (2011).

130. Klement, T. & Büchs, J. Itaconic acid–a biotechnological process in change. *Bioresource*

*technology* **135**, 422-431 (2013).

131. Liu, R. et al. Metabolic engineering of fatty acyl-ACP reductase-dependent pathway to improve fatty alcohol production in Escherichia coli. *Metabolic engineering* **22**, 10-21 (2014).

132. Schirmer, A., Rude, M.A., Li, X., Popova, E. & Del Cardayre, S.B. Microbial biosynthesis of alkanes. *Science* **329**, 559-562 (2010).

133. Akhtar, M.K., Turner, N.J. & Jones, P.R. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proceedings of the National Academy of Sciences* **110**, 87-92 (2013).

134. Sorigué, D. et al. An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**, 903-907 (2017).

135. Steen, E.J. et al. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562 (2010).

136. Song, C.W., Kim, J.W., Cho, I.J. & Lee, S.Y. Metabolic engineering of Escherichia coli for the production of 3-hydroxypropionic acid and malonic acid through β-alanine route. *ACS synthetic biology* **5**, 1256-1263 (2016).

137. Tong, W. et al. Biosynthetic pathway for acrylic acid from glycerol in recombinant Escherichia coli. *Applied microbiology and biotechnology* **100**, 4901-4907 (2016).

138. Chu, H.S. et al. Direct fermentation route for the production of acrylic acid. *Metabolic engineering* **32**, 23-29 (2015).

139. Kumar, V., Ashok, S. & Park, S. Recent advances in biological production of 3-hydroxypropionic acid. *Biotechnology advances* **31**, 945-961 (2013).

140. Rathnasingh, C. et al. Production of 3-hydroxypropionic acid via malonyl-CoA pathway

using recombinant Escherichia coli strains. *Journal of biotechnology* **157**, 633-640 (2012).

141. Young, V.R. Adult amino acid requirements: the case for a major revision in current recommendations. *The Journal of nutrition* **124**, 1517S-1523S (1994).

142. 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. *Metabolic engineering* **13**, 159-168 (2011).

143. Becker, J. & Wittmann, C. Bio-based production of chemicals, materials and fuels–Corynebacterium glutamicum as versatile cell factory. *Current opinion in biotechnology* **23**, 631-640 (2012).

144. Becker, J. & Wittmann, C. Systems and synthetic metabolic engineering for amino acid production–the heartbeat of industrial strain development. *Current opinion in biotechnology* **23**, 718-726 (2012).

145. Lee, K.H., Park, J.H., Kim, T.Y., Kim, H.U. & Lee, S.Y. Systems metabolic engineering of Escherichia coli for L-threonine production. *Molecular systems biology* **3**(2007).

146. Jain, N. & Ramawat, K.G. Nutraceuticals and antioxidants in prevention of diseases. in *Natural Products* 2559-2580 (Springer, 2013).

147. Marienhagen, J. & Bott, M. Metabolic engineering of microorganisms for the synthesis of plant natural products. *Journal of biotechnology* **163**, 166-178 (2013).

148. Lin, Y., Jain, R. & Yan, Y. Microbial production of antioxidant food ingredients via metabolic engineering. *Current opinion in biotechnology* **26**, 71-78 (2014).

149. Mora-Pale, M., Sanchez-Rodriguez, S.P., Linhardt, R.J., Dordick, J.S. & Koffas, M.A. Metabolic engineering and in vitro biosynthesis of phytochemicals and non-natural analogues. *Plant science* **210**, 10-24 (2013).

150. Leonard, E. & Koffas, M.A. Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by Escherichia coli. *Applied and environmental microbiology* **73**, 7246-7251 (2007).

151. Miyahisa, I. et al. Combinatorial biosynthesis of flavones and flavonols in Escherichia coli. *Applied microbiology and biotechnology* **71**, 53-58 (2006).

152. He, X.-Z., Li, W.-S., Blount, J.W. & Dixon, R.A. Regioselective synthesis of plant (iso) flavone glycosides in Escherichia coli. *Applied microbiology and biotechnology* **80**, 253-260 (2008).

153. Kim, S.Y., Lee, H.R., Park, K.-s., Kim, B.-G. & Ahn, J.-H. Metabolic engineering of Escherichia coli for the biosynthesis of flavonoid-O-glucuronides and flavonoid-O-galactoside. *Applied microbiology and biotechnology* **99**, 2233-2242 (2015).

154. Kim, B.-G., Kim, H.J. & Ahn, J.-H. Production of bioactive flavonol rhamnosides by expression of plant genes in Escherichia coli. *Journal of Agricultural and Food Chemistry* **60**, 11143-11148 (2012).

155. Xu, P., Ranganathan, S., Fowler, Z.L., Maranas, C.D. & Koffas, M.A. Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metabolic engineering* **13**, 578-587 (2011).

156. Yang, Y., Lin, Y., Li, L., Linhardt, R.J. & Yan, Y. Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. *Metabolic engineering* **29**, 217-226 (2015).

157. Wu, J., Du, G., Zhou, J. & Chen, J. Metabolic engineering of Escherichia coli for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metabolic engineering* **16**, 48-55 (2013).

158. van Summeren-Wesenhagen, P.V. & Marienhagen, J. Metabolic engineering of Escherichia coli for the synthesis of the plant polyphenol pinosylvin. *Applied and environmental microbiology* **81**, 840-849 (2015).

159. Kim, M.-J., Kim, B.-G. & Ahn, J.-H. Biosynthesis of bioactive O-methylated flavonoids in Escherichia coli. *Applied microbiology and biotechnology* **97**, 7195-7204 (2013).

160. Bhan, B.F.C., R. J. Linhardt, M.A.G. Koffas. Expanding the chemical space of polyketides through structure-guided mutagenesis of Vitis vinifera stilbene synthase. *Biochimie* (2015).

161. Bhan, N. et al. Enzymatic formation of a resorcylic acid by creating a structure-guided single-point mutation in stilbene synthase. *Protein Science* **24**, 167-173 (2015).

162. Nakagawa, A. et al. A bacterial platform for fermentative production of plant alkaloids. *Nature communications* **2**, 326 (2011).

163. Fossati, E. et al. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae. *Nature communications* **5**(2014).

164. Nakagawa, A. et al. (R, S)-Tetrahydropapaveroline production by stepwise fermentation using engineered Escherichia coli. *Scientific reports* **4**(2014).

165. Brown, S., Clastre, M., Courdavault, V. & O'Connor, S.E. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proceedings of the National Academy of Sciences* **112**, 3205-3210 (2015).

166. Mikkelsen, M.D. et al. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metabolic engineering* **14**, 104-111 (2012).

167. Hazelwood, L.A., Daran, J.-M., Van Maris, A.J., Pronk, J.T. & Dickinson, J.R. The

Ehrlich pathway for fusel alcohol production: a century of research on Saccharomyces cerevisiae metabolism. *Appl. Environ. Microbiol.* **74**, 2259-2266 (2008).

168. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *nature* **451**, 86 (2008).

169. Mak, W.S. et al. Integrative genomic mining for enzyme function to enable engineering of a non-natural biosynthetic pathway. *Nature communications* **6**, 10005 (2015).

170. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proceedings of the National Academy of Sciences* **105**, 20653-20658 (2008).

171. Song, C.W., Lee, J., Ko, Y.-S. & Lee, S.Y. Metabolic engineering of Escherichia coli for the production of 3-aminopropionic acid. *Metabolic engineering* **30**, 121-129 (2015).

172. Pei, W. et al. Molecular engineering of l-aspartate-α-decarboxylase for improved activity and catalytic stability. *Applied microbiology and biotechnology* **101**, 6015-6021 (2017).

173. Lacmata, S.T. et al. Enhanced poly (3-hydroxypropionate) production via β-alanine pathway in recombinant Escherichia coli. *PloS one* **12**, e0173150 (2017).

174. Chae, T.U., Ko, Y.-S., Hwang, K.-S. & Lee, S.Y. Metabolic engineering of Escherichia coli for the production of four-, five-and six-carbon lactams. *Metabolic engineering* **41**, 82-91 (2017).

175. Xu, P., Bhan, N. & Koffas, M.A. Engineering plant metabolism into microbes: from systems biology to synthetic biology. *Current opinion in biotechnology* **24**, 291-299 (2013).

176. Nakagawa, A. et al. Total biosynthesis of opiates by stepwise fermentation using engineered Escherichia coli. *Nature communications* **7**(2016).

177. Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F. & Smolke, C.D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095-1100 (2015).

178. Lin, Y. & Yan, Y. Biosynthesis of caffeic acid in Escherichia coli using its endogenous hydroxylase complex. *Microbial cell factories* **11**, 1 (2012).

179. Ni, J., Tao, F., Du, H. & Xu, P. Mimicking a natural pathway for de novo biosynthesis: natural vanillin production from accessible carbon sources. *Scientific reports* **5**(2015).

180. Medema, M.H., van Raaphorst, R., Takano, E. & Breitling, R. Computational tools for the synthetic design of biochemical pathways. *Nature Reviews Microbiology* **10**, 191-202 (2012).

181. Lin, Y., Shen, X., Yuan, Q. & Yan, Y. Microbial biosynthesis of the anticoagulant precursor 4-hydroxycoumarin. *Nature communications* **4**(2013).

182. Gülçin, İ. Antioxidant activity of caffeic acid (3, 4-dihydroxycinnamic acid). *Toxicology* **217**, 213-220 (2006).

183. Touaibia, M., Jean-Francois, J. & Doiron, J. Caffeic acid, a versatile pharmacophore: an overview. *Mini reviews in medicinal chemistry* **11**, 695-713 (2011).

184. Son, S. & Lewis, B.A. Free radical scavenging and antioxidative activity of caffeic acid amide and ester analogues: structure-activity relationship. *Journal of agricultural and food chemistry* **50**, 468-472 (2002).

185. Natarajan, K., Singh, S., Burke, T.R., Grunberger, D. & Aggarwal, B.B. Caffeic acid phenethyl ester is a potent and specific inhibitor of activation of nuclear transcription factor NF-kappa B. *Proceedings of the National Academy of Sciences* **93**, 9090-9095 (1996).

186. Murtaza, G. et al. Caffeic Acid Phenethyl Ester and Therapeutic Potentials. *BioMed*

*Research International* **2014**(2014).

187. Tolba, M.F., Azab, S.S., Khalifa, A.E., Abdel-Rahman, S.Z. & Abdel-Naim, A.B. Caffeic acid phenethyl ester, a promising component of propolis with a plethora of biological activities: A review on its anti-inflammatory, neuroprotective, hepatoprotective, and cardioprotective effects. *IUBMB life* **65**, 699-709 (2013).

188. Wu, J. et al. Caffeic acid phenethyl ester (CAPE), derived from a honeybee product propolis, exhibits a diversity of anti-tumor effects in pre-clinical models of human breast cancer. *Cancer letters* **308**, 43-53 (2011).

189. Onori, P. et al. Caffeic acid phenethyl ester decreases cholangiocarcinoma growth by inhibition of NF-κB and induction of apoptosis. *International Journal of Cancer* **125**, 565-576 (2009).

190. Lee, K.W. et al. Caffeic acid phenethyl ester inhibits invasion and expression of matrix metalloproteinase in SK-Hep1 human hepatocellular carcinoma cells by targeting nuclear factor kappa B. *Genes & nutrition* **2**, 319-322 (2008).

191. Chen, M.-J. et al. Caffeic acid phenethyl ester induces apoptosis of human pancreatic cancer cells involving caspase and mitochondrial dysfunction. *Pancreatology* **8**, 566-576 (2008).

192. Wu, C.-S., Chen, M.-F., Lee, I.-L. & Tung, S.-Y. Predictive role of nuclear factor-κB activity in gastric cancer: a promising adjuvant approach with caffeic acid phenethyl ester. *Journal of clinical gastroenterology* **41**, 894-900 (2007).

193. Xiang, D. et al. Caffeic acid phenethyl ester induces growth arrest and apoptosis of colon cancer cells via the [beta]-catenin/T-cell factor signaling. *Anti-cancer drugs* **17**, 753-762 (2006).

194. Kuo, H.-C. et al. Inhibitory effect of caffeic acid phenethyl ester on the growth of C6 glioma cells in vitro and in vivo. *Cancer letters* **234**, 199-208 (2006).

195. Chen, M.-F., Wu, C.-T., Chen, Y.-J. & Chen, W.-C. Cell killing and radiosensitization by caffeic acid phenethyl ester (CAPE) in lung cancer cells. *Journal of radiation research* **45**, 253-260 (2004).

196. Guo, X. et al. Antitumor activity of caffeic acid 3, 4-dihydroxyphenethyl ester and its pharmacokinetic and metabolic properties. *Phytomedicine* **20**, 904-912 (2013).

197. Yang, J., Kerwin, S.M., Bowman, P.D. & Stavchansky, S. Stability of caffeic acid phenethyl amide (CAPA) in rat plasma. *Biomedical Chromatography* **26**, 594-598 (2012).

198. SON, S., Lobkowsky, E.B. & Lewis, B.A. Caffeic Acid Phenethyl Ester (CAPE): Synthesis and X-Ray Crstallographic Analysis. *Chemical and pharmaceutical bulletin* **49**, 236-238 (2001).

199. Guyot, B., Bosquette, B., Pina, M. & Graille, J. Esterification of phenolic acids from green coffee with an immobilized lipase from Candida antarctica in solvent-free medium. *Biotechnology letters* **19**, 529-532 (1997).

200. Wang, J., Gu, S.-S., Cui, H.-S., Wu, X.-Y. & Wu, F.-A. A novel continuous flow biosynthesis of caffeic acid phenethyl ester from alkyl caffeate and phenethanol in a packed bed microreactor. *Bioresource technology* **158**, 39-47 (2014).

201. Rodriguez, G.M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in Escherichia coli. *Nature chemical biology* **10**, 259-265 (2014).

202. Zhang, F., Carothers, J.M. & Keasling, J.D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature biotechnology* **30**, 354-359 (2012).

203. Guo, D., Zhu, J., Deng, Z. & Liu, T. Metabolic engineering of Escherichia coli for production of fatty acid short-chain esters through combination of the fatty acid and 2-keto acid pathways. *Metabolic engineering* **22**, 69-75 (2014).

204. Xu, P., Qiao, K., Ahn, W.S. & Stephanopoulos, G. Engineering Yarrowia lipolytica as a platform for synthesis of drop-in transportation fuels and oleochemicals. *Proceedings of the National Academy of Sciences* **113**, 10848-10853 (2016).

205. Xu, P. et al. Modular optimization of multi-gene pathways for fatty acids production in E. coli. *Nature communications* **4**, 1409 (2013).

206. Thomason, L.C., Costantino, N. & Court, D.L. E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology*, 1.17. 1-1.17. 8 (2007).

207. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning*, (Cold spring harbor laboratory press New York, 1989).

208. Huang, Q., Lin, Y. & Yan, Y. Caffeic acid production enhancement by engineering a phenylalanine over-producing Escherichia coli strain. *Biotechnology and bioengineering* **110**, 3188-3196 (2013).

209. Shen, X. et al. Elevating 4-Hydroxycoumarin Production through Alleviating Thioesterase-mediated Salicoyl-CoA Degradation. *Metabolic Engineering* (2017).

210. Petersen, M. et al. Evolution of rosmarinic acid biosynthesis. *Phytochemistry* **70**, 1663-1679 (2009).

211. Eudes, A. et al. Production of hydroxycinnamoyl anthranilates from glucose in Escherichia coli. *Microbial cell factories* **12**, 62 (2013).

212. Liebgott, P.-P., Amouric, A., Comte, A., Tholozan, J.-L. & Lorquin, J. Hydroxytyrosol from tyrosol using hydroxyphenylacetic acid-induced bacterial cultures and evidence of

the role of 4-HPA 3-hydroxylase. *Research in microbiology* **160**, 757-766 (2009).

213. Kezmarsky, N.D., Xu, H., Graham, D.E. & White, R.H. Identification and characterization of a L-tyrosine decarboxylase in Methanocaldococcus jannaschii. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1722**, 175-182 (2005).

214. Zhang, K. & Ni, Y. Tyrosine decarboxylase from Lactobacillus brevis: soluble expression and characterization. *Protein expression and purification* **94**, 33-39 (2014).

215. Koyanagi, T. et al. Eukaryotic-type aromatic amino acid decarboxylase from the root colonizer Pseudomonas putida is highly specific for 3, 4-dihydroxyphenyl-L-alanine, an allelochemical in the rhizosphere. *Microbiology* **158**, 2965-2974 (2012).

216. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Combinatorial biosynthesis of plant-specific coumarins in bacteria. *Metabolic engineering* **18**, 69-77 (2013).

217. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *Journal of biological chemistry* **278**, 95-103 (2003).

218. Tang, C. & Sojinu, O.S. Simultaneous determination of caffeic acid phenethyl ester and its metabolite caffeic acid in dog plasma using liquid chromatography tandem mass spectrometry. *Talanta* **94**, 232-239 (2012).

219. Back, K. et al. Cloning and characterization of a hydroxycinnamoyl-CoA: tyramine N-(hydroxycinnamoyl) transferase induced in response to UV-C and wounding from Capsicum annuum. *Plant and Cell Physiology* **42**, 475-481 (2001).

220. Inouye, S. & Inouye, M. Up-promoter mutations in the lpp gene of Escherichia coli. *Nucleic acids research* **13**, 3101-3110 (1985).

221. Xu, P., Rizzoni, E.A., Sul, S.-Y. & Stephanopoulos, G. Improving metabolic pathway

efficiency by statistical model-based multivariate regulatory metabolic engineering. *ACS synthetic biology* **6**, 148-158 (2016).

222. Xu, P., Vansiri, A., Bhan, N. & Koffas, M.A. ePathBrick: a synthetic biology platform for engineering metabolic pathways in E. coli. *ACS synthetic biology* **1**, 256-266 (2012).

223. Báez-Viveros, J.L. et al. Metabolic engineering and protein directed evolution increase the yield of L-phenylalanine synthesized from glucose in Escherichia coli. *Biotechnology and bioengineering* **87**, 516-524 (2004).

224. Yang, S.-M., Shim, G.Y., Kim, B.-G. & Ahn, J.-H. Biological synthesis of coumarins in Escherichia coli. *Microbial cell factories* **14**, 1 (2015).

225. Lin, Y. & Yan, Y. Biotechnological production of plant-specific hydroxylated phenylpropanoids. *Biotechnology and bioengineering* **111**, 1895-1899 (2014).

226. Zhuang, Y. et al. Synthesis of rosmarinic acid analogues in Escherichia coli. *Biotechnology letters* **38**, 619-627 (2016).

227. Kim, B.-G., Jung, W.D., Mok, H. & Ahn, J.-H. Production of hydroxycinnamoyl-shikimates and chlorogenic acid in Escherichia coli: production of hydroxycinnamic acid conjugates. *Microbial cell factories* **12**, 15 (2013).

228. Cha, M.N., Kim, H.J., Kim, B.G. & Ahn, J.-H. Synthesis of chlorogenic acid and p-coumaroyl shikimates from glucose using engineered Escherichia coli. *J Microbiol Biotechnol* **24**, 1109-17 (2014).

229. Bloch, S.E. & Schmidt-Dannert, C. Construction of a chimeric biosynthetic pathway for the de novo biosynthesis of rosmarinic acid in Escherichia coli. *ChemBioChem* **15**, 2393-2401 (2014).

230. Eudes, A. et al. Production of tranilast [N-(3′, 4′-dimethoxycinnamoyl)-anthranilic acid]

and its analogs in yeast Saccharomyces cerevisiae. *Applied microbiology and biotechnology* **89**, 989-1000 (2011).

231. Sim, G.Y., Yang, S.-M., Kim, B.G. & Ahn, J.-H. Bacterial synthesis of N-hydroxycinnamoyl phenethylamines and tyramines. *Microbial cell factories* **14**, 162 (2015).

232. Eudes, A. et al. Exploiting members of the BAHD acyltransferase family to synthesize multiple hydroxycinnamate and benzoate conjugates in yeast. *Microbial Cell Factories* **15**, 198 (2016).

233. Knight, M.J., Bull, I.D. & Curnow, P. The yeast enzyme Eht1 is an octanoyl-CoA: ethanol acyltransferase that also functions as a thioesterase. *Yeast* **31**, 463-474 (2014).

234. Latham, J.A., Chen, D., Allen, K.N. & Dunaway-Mariano, D. Divergence of substrate specificity and function in the Escherichia coli hotdog-fold thioesterase paralogs YdiI and YbdB. *Biochemistry* **53**, 4775-4787 (2014).

235. Chen, M. et al. Identification of a hotdog fold thioesterase involved in the biosynthesis of menaquinone in Escherichia coli. *Journal of bacteriology* **195**, 2768-2775 (2013).

236. Juminaga, D. et al. Modular engineering of L-tyrosine production in Escherichia coli. *Applied and environmental microbiology* **78**, 89-98 (2012).

237. Santos, C.N.S., Xiao, W. & Stephanopoulos, G. Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in Escherichia coli. *Proceedings of the National Academy of Sciences* **109**, 13538-13543 (2012).

238. Lütke-Eversloh, T. & Stephanopoulos, G. Combinatorial pathway analysis for improved L-tyrosine production in Escherichia coli: identification of enzymatic bottlenecks by systematic gene overexpression. *Metabolic engineering* **10**, 69-77 (2008).

239. Lütke-Eversloh, T. & Stephanopoulos, G. L-tyrosine production by deregulated strains of Escherichia coli. *Applied microbiology and biotechnology* **75**, 103-110 (2007).

240. Na, D. et al. Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nature biotechnology* **31**, 170-174 (2013).

241. Liu, Y.-G. et al. Synthetic phenylethanoid glycoside derivatives as potent neuroprotective agents. *European journal of medicinal chemistry* **95**, 313-323 (2015).

242. Tassano, E. et al. Conjugation of Hydroxytyrosol with Other Natural Phenolic Fragments: From Waste to Antioxidants and Antitumour Compounds. *European Journal of Organic Chemistry* **2015**, 6710-6726 (2015).

243. Zhou, T. et al. Design, synthesis and biological evaluation of L-dopa amide derivatives as potential prodrugs for the treatment of Parkinson's disease. *European journal of medicinal chemistry* **45**, 4035-4042 (2010).

244. Shi, H., Xie, D., Yang, R. & Cheng, Y. Synthesis of caffeic acid phenethyl ester derivatives, and their cytoprotective and neuritogenic activities in PC12 cells. *Journal of agricultural and food chemistry* **62**, 5046-5053 (2014).

245. Maresca, A., Akyuz, G., Osman, S.M., AlOthman, Z. & Supuran, C.T. Inhibition of mammalian carbonic anhydrase isoforms I–XIV with a series of phenolic acid esters. *Bioorganic & medicinal chemistry* **23**, 7181-7188 (2015).

246. Yang, Y., Song, Z.-G. & Liu, Z.-Q. Synthesis and antioxidant capacities of hydroxyl derivatives of cinnamoylphenethylamine in protecting DNA and scavenging radicals. *Free radical research* **45**, 445-453 (2011).

247. Yamazaki, Y., Kawano, Y. & Uebayasi, M. Induction of adiponectin by natural and synthetic phenolamides in mouse and human preadipocytes and its enhancement by

docosahexaenoic acid. *Life sciences* **82**, 290-300 (2008).

248. Ma, X., Yan, J., Xu, K., Guo, L. & Li, H. Binding mechanism of trans-n-caffeoyltyramine and human serum albumin: Investigation by multi-spectroscopy and docking simulation. *Bioorganic chemistry* **66**, 102-110 (2016).

249. Atsumi, S. et al. Metabolic engineering of Escherichia coli for 1-butanol production. *Metabolic engineering* **10**, 305-311 (2008).

250. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic acids research* **25**, 1203-1210 (1997).

251. Shen, C.R. & Liao, J.C. Metabolic engineering of Escherichia coli for 1-butanol and 1-propanol production via the keto-acid pathways. *Metabolic engineering* **10**, 312-320 (2008).

252. Huo, Y.-X. et al. Conversion of proteins into biofuels by engineering nitrogen flux. *Nature biotechnology* **29**, 346-351 (2011).

253. Sun, J. et al. Aerobic biosynthesis of hydrocinnamic acids in Escherichia coli with a strictly oxygen-sensitive enoate reductase. *Metabolic engineering* **35**, 75-82 (2016).

254. Jullesson, D., David, F., Pfleger, B. & Nielsen, J. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnology advances* **33**, 1395-1402 (2015).

255. Alper, H. & Stephanopoulos, G. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology* **7**, 715 (2009).

256. Sun, X. et al. Synthesis of chemicals by metabolic engineering of microbes. *Chemical society reviews* **44**, 3760-3785 (2015).

257. Ehrenworth, A.M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nature chemical biology* **13**, 249 (2017).

258. Luo, Y. et al. Engineered biosynthesis of natural products in heterologous hosts. *Chemical Society Reviews* **44**, 5265-5290 (2015).

259. Lynch, M.D. Into new territory: improved microbial synthesis through engineering of the essential metabolic network. *Current opinion in biotechnology* **38**, 106-111 (2016).

260. Trantas, E.A., Koffas, M.A., Xu, P. & Ververidis, F. When plants produce not enough or at all: metabolic engineering of flavonoids in microbial hosts. *Frontiers in plant science* **6**, 7 (2015).

261. Smanski, M.J. et al. Synthetic biology to access and expand nature's chemical diversity. *Nature Reviews Microbiology* **14**, 135 (2016).

262. Wu, G. et al. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends in biotechnology* **34**, 652-664 (2016).

263. Tsoi, R. et al. Metabolic division of labor in microbial systems. *Proceedings of the National Academy of Sciences* **115**, 2526-2531 (2018).

264. McNerney, M.P., Watstein, D.M. & Styczynski, M.P. Precision metabolic engineering: The design of responsive, selective, and controllable metabolic systems. *Metabolic engineering* **31**, 123-131 (2015).

265. Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences* **111**, 11299-11304 (2014).

266. Zhang, F., Carothers, J.M. & Keasling, J.D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature biotechnology* **30**,

354 (2012).

267. Yang, Y. et al. Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. *Nature communications* **9**, 3043 (2018).

268. Rugbjerg, P., Sarup-Lytzen, K., Nagy, M. & Sommer, M.O.A. Synthetic addiction extends the productive life time of engineered Escherichia coli populations. *Proceedings of the National Academy of Sciences* **115**, 2347-2352 (2018).

269. Xiao, Y., Bowen, C.H., Liu, D. & Zhang, F. Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. *Nature chemical biology* **12**, 339 (2016).

270. Lan, E.I. & Liao, J.C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences* **109**, 6018-6023 (2012).

271. Shen, C.R. et al. Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Appl. Environ. Microbiol.* **77**, 2905-2915 (2011).

272. Williams, T. et al. Quorum-sensing linked RNA interference for dynamic metabolic pathway control in Saccharomyces cerevisiae. *Metabolic engineering* **29**, 124-134 (2015).

273. Klamt, S. & Mahadevan, R. On the feasibility of growth-coupled product synthesis in microbial strains. *Metabolic engineering* **30**, 166-178 (2015).

274. Pontrelli, S. et al. Metabolic repair through emergence of new pathways in Escherichia coli. *Nature chemical biology* **14**, 1005 (2018).

275. Long, C.P., Gonzalez, J.E., Feist, A.M., Palsson, B.O. & Antoniewicz, M.R. Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in Escherichia coli. *Proceedings of the National Academy of Sciences* **115**, 222-227 (2018).

276. Lin, B. et al. Reconstitution of TCA cycle with DAOCS to engineer Escherichia coli into

an efficient whole cell catalyst of penicillin G. *Proceedings of the National Academy of Sciences* **112**, 9855-9859 (2015).

277. Zhang, Y. et al. Reconstruction of tricarboxylic acid cycle in Corynebacterium glutamicum with a genome-scale metabolic network model for trans-4-hydroxyproline production. *Biotechnology and bioengineering* **116**, 99-109 (2019).

278. Kind, S., Becker, J. & Wittmann, C. Increased lysine production by flux coupling of the tricarboxylic acid cycle and the lysine biosynthetic pathway—metabolic engineering of the availability of succinyl-CoA in Corynebacterium glutamicum. *Metabolic engineering* **15**, 184-195 (2013).

279. Dinh, H.V., King, Z.A., Palsson, B.O. & Feist, A.M. Identification of growth-coupled production strains considering protein costs and kinetic variability. *Metabolic engineering communications* **7**, e00080 (2018).

280. Balderas-Hernández, V.E. et al. Metabolic engineering for improving anthranilate synthesis from glucose in Escherichia coli. *Microbial cell factories* **8**, 19 (2009).

281. Sun, X., Lin, Y., Huang, Q., Yuan, Q. & Yan, Y. A novel muconic acid biosynthesis approach by shunting tryptophan biosynthesis via anthranilate. *Appl. Environ. Microbiol.* **79**, 4024-4030 (2013).

282. Ikeda, M. Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering. *Applied microbiology and biotechnology* **69**, 615 (2006).

283. Pavlikova, M. et al. Novel pathway of 3-hydroxyanthranilic acid formation in limazepine biosynthesis reveals evolutionary relation between phenazines and pyrrolobenzodiazepines. *Scientific reports* **8**, 7810 (2018).

284. Tribe, D.E. Novel microorganism and method. (Google Patents, 1987).

285. Cherepanov, P.P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14 (1995).

286. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: a laboratory manual*, (Cold spring harbor laboratory press, 1989).

287. Sun, X., Lin, Y., Yuan, Q. & Yan, Y. Precursor-directed biosynthesis of 5-hydroxytryptophan using metabolically engineered E. coli. *ACS synthetic biology* **4**, 554-558 (2014).

288. Lin, Y., Shen, X., Yuan, Q. & Yan, Y. Microbial biosynthesis of the anticoagulant precursor 4-hydroxycoumarin. *Nature communications* **4**, 2603 (2013).

289. Thomason, L.C., Costantino, N. & Court, D.L. E. coli genome manipulation by P1 transduction. *Current protocols in molecular biology* **79**, 1.17. 1-1.17. 8 (2007).

290. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* **2**(2006).

291. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**, 6640-6645 (2000).

292. Lindqvist, R. & Barmark, G. Specific growth rate determines the sensitivity of Escherichia coli to lactic acid stress: implications for predictive microbiology. *BioMed research international* **2014**(2014).

293. Berney, M., Weilenmann, H.-U., Ihssen, J., Bassin, C. & Egli, T. Specific growth rate determines the sensitivity of Escherichia coli to thermal, UVA, and solar disinfection. *Appl. Environ. Microbiol.* **72**, 2586-2593 (2006).

294. Noda, S., Shirai, T., Oyama, S. & Kondo, A. Metabolic design of a platform Escherichia coli strain producing various chorismate derivatives. *Metabolic engineering* **33**, 119-129 (2016).

295. Noda, S., Shirai, T., Mori, Y., Oyama, S. & Kondo, A. Engineering a synthetic pathway for maleate in Escherichia coli. *Nature communications* **8**, 1153 (2017).

296. Tröndle, J., Trachtmann, N., Sprenger, G.A. & Weuster-Botz, D. Fed-batch production of l-tryptophan from glycerol using recombinant Escherichia coli. *Biotechnology and bioengineering* **115**, 2881-2892 (2018).

297. Durnin, G. et al. Understanding and harnessing the microaerobic metabolism of glycerol in Escherichia coli. *Biotechnology and bioengineering* **103**, 148-161 (2009).

298. Emmerling, M. et al. Metabolic flux responses to pyruvate kinase knockout in Escherichia coli. *Journal of bacteriology* **184**, 152-164 (2002).

299. Xu, Y.-F., Amador-Noguez, D., Reaves, M.L., Feng, X.-J. & Rabinowitz, J.D. Ultrasensitive regulation of anapleurosis via allosteric activation of PEP carboxylase. *Nature chemical biology* **8**, 562 (2012).

300. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Engineering bacterial phenylalanine 4-hydroxylase for microbial synthesis of human neurotransmitter precursor 5-hydroxytryptophan. *ACS synthetic biology* **3**, 497-505 (2014).

301. Chen, L., Chen, M., Ma, C. & Zeng, A.-P. Discovery of feed-forward regulation in l-tryptophan biosynthesis and its use in metabolic engineering of E. coli for efficient tryptophan bioproduction. *Metabolic engineering* **47**, 434-444 (2018).

302. Panichkin, V.B., Livshits, V.A., Biryukova, I.V. & Mashko, S.V. Metabolic engineering of Escherichia coli for L-tryptophan production. *Applied biochemistry and microbiology* **52**,

783-809 (2016).

303. Chen, L. & Zeng, A.-P. Rational design and metabolic analysis of Escherichia coli for effective production of L-tryptophan at high concentration. *Applied microbiology and biotechnology* **101**, 559-568 (2017).

304. Herz, E. et al. The genetic basis for the adaptation of E. coli to sugar synthesis from CO 2. *Nature communications* **8**, 1705 (2017).

305. Vardon, D.R. et al. cis, cis-Muconic acid: separation and catalysis to bio-adipic acid for nylon-6, 6 polymerization. *Green Chemistry* **18**, 3397-3413 (2016).

306. Kruyer, N.S. & Peralta-Yahya, P. Metabolic engineering strategies to bio-adipic acid production. *Current opinion in biotechnology* **45**, 136-143 (2017).

307. Barton, N. et al. Enabling the valorization of guaiacol-based lignin: Integrated chemical and biochemical production of cis, cis-muconic acid using metabolically engineered Amycolatopsis sp ATCC 39116. *Metabolic engineering* **45**, 200-210 (2018).

308. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in Escherichia coli. *Metabolic engineering* **23**, 62-69 (2014).

309. Sun, X., Lin, Y., Yuan, Q. & Yan, Y. Biological production of muconic acid via a prokaryotic 2, 3-dihydroxybenzoic acid decarboxylase. *ChemSusChem* **7**, 2478-2481 (2014).

310. Wang, J. et al. Establishing a novel biosynthetic pathway for the production of 3, 4-dihydroxybutyric acid from xylose in Escherichia coli. *Metabolic engineering* **41**, 39-45 (2017).

311. Tai, Y.-S. et al. Engineering nonphosphorylative metabolism to generate lignocellulose-

derived products. *Nature chemical biology* **12**, 247 (2016).

312. Lee, J.W. et al. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nature chemical biology* **8**, 536 (2012).

313. Wang, Y., San, K.-Y. & Bennett, G.N. Cofactor engineering for advancing chemical biotechnology. *Current opinion in biotechnology* **24**, 994-999 (2013).

314. Peralta-Yahya, P.P., Zhang, F., Del Cardayre, S.B. & Keasling, J.D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320 (2012).

315. Lee, S.Y., Kim, H.M. & Cheon, S. Metabolic engineering for the production of hydrocarbon fuels. *Current opinion in biotechnology* **33**, 15-22 (2015).

316. Chae, T.U. et al. Metabolic engineering for the production of dicarboxylic acids and diamines. *Metabolic engineering* (2019).

317. Choi, Y.J. & Lee, S.Y. Microbial production of short-chain alkanes. *Nature* **502**, 571 (2013).

318. Curran, K.A. & Alper, H.S. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metabolic engineering* **14**, 289-297 (2012).

319. Hirasawa, T. & Shimizu, H. Recent advances in amino acid production by microbial cells. *Current opinion in biotechnology* **42**, 133-146 (2016).

320. D'Este, M., Alvarado-Morales, M. & Angelidaki, I. Amino acids production focusing on fermentation technologies–a review. *Biotechnology advances* **36**, 14-25 (2018).

321. Lammens, T., Franssen, M., Scott, E. & Sanders, J. Availability of protein-derived amino acids as feedstock for the production of bio-based chemicals. *Biomass and Bioenergy* **44**, 168-181 (2012).

322. Zhang, Y., Liu, D. & Chen, Z. Production of C2–C4 diols from renewable bioresources:

new metabolic pathways and metabolic engineering strategies. *Biotechnology for biofuels* **10**, 299 (2017).

323. Zeng, A.-P. & Sabra, W. Microbial production of diols as platform chemicals: recent progresses. *Current Opinion in Biotechnology* **22**, 749-757 (2011).

324. Firlotte, N., Cooper, D.G., Marić, M. & Nicell, J.A. Characterization of 1, 5-pentanediol dibenzoate as a potential "green" plasticizer for poly (vinyl chloride). *Journal of Vinyl and Additive Technology* **15**, 99-107 (2009).

325. Huang, K. et al. Improving economics of lignocellulosic biofuels: An integrated strategy for coproducing 1, 5-pentanediol and ethanol. *Applied Energy* **213**, 585-594 (2018).

326. Zhong, W., Zhang, Y., Wu, W., Liu, D. & Chen, Z. Metabolic engineering of a homoserine-derived non-natural pathway for the de novo production of 1, 3-propanediol from glucose. *ACS synthetic biology* (2019).

327. Martin, C.H. et al. A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-γ-butyrolactone. *Nature communications* **4**, 1414 (2013).

328. Huang, K. et al. Conversion of furfural to 1, 5-pentanediol: process synthesis and analysis. *ACS Sustainable Chemistry & Engineering* **5**, 4699-4706 (2017).

329. Hädicke, O., von Kamp, A., Aydogan, T. & Klamt, S. OptMDFpathway: Identification of metabolic pathways with maximal thermodynamic driving force and its application for analyzing the endogenous CO2 fixation potential of Escherichia coli. *PLoS computational biology* **14**, e1006492 (2018).

330. Rodriguez, G.M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in Escherichia coli. *Nature chemical biology* **10**, 259 (2014).

331. Kramer, L. et al. Characterization of carboxylic acid reductases for biocatalytic synthesis

of industrial chemicals. *ChemBioChem* **19**, 1452-1460 (2018).

332. Richard, H. & Foster, J.W. Escherichia coli glutamate-and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *Journal of bacteriology* **186**, 6032-6041 (2004).

333. Moreau, P.L. The lysine decarboxylase CadA protects Escherichia coli starved of phosphate against fermentation acids. *Journal of bacteriology* **189**, 2249-2261 (2007).

334. Nozaki, S., Webb, M.E. & Niki, H. An activator for pyruvoyl-dependent l-aspartate α-decarboxylase is conserved in a small group of the γ-proteobacteria including Escherichia coli. *MicrobiologyOpen* **1**, 298-310 (2012).

335. Atsumi, S. et al. Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied microbiology and biotechnology* **85**, 651-657 (2010).

336. Knorr, S. et al. Widespread bacterial lysine degradation proceeding via glutarate and L-2-hydroxyglutarate. *Nature communications* **9**, 5071 (2018).

337. Revelles, O., Espinosa-Urgel, M., Fuhrer, T., Sauer, U. & Ramos, J.L. Multiple and interconnected pathways for L-lysine catabolism in Pseudomonas putida KT2440. *Journal of bacteriology* **187**, 7500-7510 (2005).

338. Ho, N.A.T., Hou, C.Y., Kim, W.H. & Kang, T.J. Expanding the active pH range of Escherichia coli glutamate decarboxylase by breaking the cooperativeness. *Journal of bioscience and bioengineering* **115**, 154-158 (2013).

339. Burgard, A., Burk, M.J., Osterhout, R., Van Dien, S. & Yim, H. Development of a commercial scale process for production of 1, 4-butanediol from sugar. *Current opinion in biotechnology* **42**, 118-125 (2016).

340. Bennett, B.D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature chemical biology* **5**, 593 (2009).

341. Rosenberg, J., Ischebeck, T. & Commichau, F.M. Vitamin B6 metabolism in microbes and approaches for fermentative production. *Biotechnology advances* **35**, 31-40 (2017).

342. Ma, W. et al. Engineering a pyridoxal 5'-phosphate supply for cadaverine production by using Escherichia coli whole-cell biocatalysis. *Scientific reports* **5**, 15630 (2015).

343. Stokell, D.J. et al. Probing the roles of key residues in the unique regulatory NADH binding site of type II citrate synthase of Escherichia coli. *Journal of Biological Chemistry* **278**, 35435-35443 (2003).

344. Brentzel, Z.J. et al. Chemicals from biomass: combining ring-opening tautomerization and hydrogenation reactions to produce 1, 5-Pentanediol from furfural. *ChemSusChem* **10**, 1351-1355 (2017).

345. Nakagawa, Y. & Tomishige, K. Production of 1, 5-pentanediol from biomass via furfural and tetrahydrofurfuryl alcohol. *Catalysis today* **195**, 136-143 (2012).

346. Xu, W. et al. Direct catalytic conversion of furfural to 1, 5-pentanediol by hydrogenolysis of the furan ring under mild conditions over Pt/Co 2 AlO 4 catalyst. *Chemical Communications* **47**, 3924-3926 (2011).

347. Contador, C.A., Rizk, M.L., Asenjo, J.A. & Liao, J.C. Ensemble modeling for strain development of L-lysine-producing Escherichia coli. *Metabolic engineering* **11**, 221-233 (2009).

348. Marbaniang, C.N. & Gowrishankar, J. Role of ArgP (IciA) in lysine-mediated repression in Escherichia coli. *Journal of bacteriology* **193**, 5985-5996 (2011).

349. Chen, Z., Rappert, S., Sun, J. & Zeng, A.-P. Integrating molecular dynamics and co-

evolutionary analysis for reliable target prediction and deregulation of the allosteric inhibition of aspartokinase for amino acid production. *Journal of biotechnology* **154**, 248-254 (2011).

350. Geng, F., Chen, Z., Zheng, P., Sun, J. & Zeng, A.-P. Exploring the allosteric mechanism of dihydrodipicolinate synthase by reverse engineering of the allosteric inhibitor binding sites and its application for lysine production. *Applied microbiology and biotechnology* **97**, 1963-1971 (2013).

351. Li, W. et al. Targeting metabolic driving and intermediate influx in lysine catabolism for high-level glutarate production. *Nature communications* **10**, 3337 (2019).

352. Chen, Z., Sun, X., Li, Y., Yan, Y. & Yuan, Q. Metabolic engineering of Escherichia coli for microbial synthesis of monolignols. *Metabolic engineering* **39**, 102-109 (2017).

353. Wang, J. et al. Microbial production of branched-chain dicarboxylate 2-methylsuccinic acid via enoate reductase-mediated bioreduction. *Metabolic engineering* **45**, 1-10 (2018).

354. Chung, H. et al. Bio-based production of monomers and polymers by metabolically engineered microorganisms. *Current opinion in biotechnology* **36**, 73-84 (2015).

355. Jambunathan, P. & Zhang, K. Novel pathways and products from 2-keto acids. *Current opinion in biotechnology* **29**, 1-7 (2014).

356. Gobin, M., Loulergue, P., Audic, J.-L. & Lemiègre, L. Synthesis and characterisation of bio-based polyester materials from vegetable oil and short to long chain dicarboxylic acids. *Industrial Crops and Products* **70**, 213-220 (2015).

357. Werpy, T. et al. Top value added chemicals from biomass. Volume 1-Results of screening for potential candidates from sugars and synthesis gas. (DTIC Document, 2004).

358. Okino, S. et al. An efficient succinic acid production process in a metabolically

engineered Corynebacterium glutamicum strain. *Applied microbiology and biotechnology* **81**, 459-464 (2008).

359. Cheong, S., Clomburg, J.M. & Gonzalez, R. Energy-and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. *Nature biotechnology* **34**, 556-561 (2016).

360. Haushalter, R.W. et al. Production of odd-carbon dicarboxylic acids in Escherichia coli using an engineered biotin-fatty acid biosynthetic pathway. (2017).

361. Park, S.J. et al. Metabolic engineering of Escherichia coli for the production of 5-aminovalerate and glutarate as C5 platform chemicals. *Metabolic engineering* **16**, 42-47 (2013).

362. Baldwin, A.F. et al. Poly (dimethyltin glutarate) as a prospective material for high dielectric applications. *Advanced Materials* **27**, 346-351 (2015).

363. Zhao, X. et al. Ultrasound enhances lipase-catalyzed synthesis of poly (ethylene glutarate). *Ultrasonics sonochemistry* **31**, 506-511 (2016).

364. Williams, J.B., Chapman, T.M. & Hercules, D.M. Synthesis of discrete mass poly (butylene glutarate) oligomers. *Macromolecules* **36**, 3898-3908 (2003).

365. Mishra, M.K., Varughese, S., Ramamurty, U. & Desiraju, G.R. Odd–even effect in the elastic modulii of α, ω-alkanedicarboxylic acids. *Journal of the American Chemical Society* **135**, 8121-8124 (2013).

366. Vafaeezadeh, M. & Hashemi, M.M. A non-cyanide route for glutaric acid synthesis from oxidation of cyclopentene in the ionic liquid media. *Process Safety and Environmental Protection* **100**, 203-207 (2016).

367. Paris, G., Berlinguet, L., Gaudry, R., English, J. & Dayan, J. Glutaric acid and

glutarimide. *Organic Syntheses*, 47-47 (2003).

368. Adkins, J., Jordan, J. & Nielsen, D.R. Engineering Escherichia coli for renewable production of the 5-carbon polyamide building-blocks 5-aminovalerate and glutarate. *Biotechnology and bioengineering* **110**, 1726-1734 (2013).

369. Liu, P. et al. Enzymatic production of 5-aminovalerate from L-lysine using L-lysine monooxygenase and 5-aminovaleramide amidohydrolase. *Scientific reports* **4**(2014).

370. Djurdjevic, I., Zelder, O. & Buckel, W. Production of glutaconic acid in a recombinant Escherichia coli strain. *Applied and environmental microbiology* **77**, 320-322 (2011).

371. Parthasarathy, A., Pierik, A.J., Kahnt, J.r., Zelder, O. & Buckel, W. Substrate specificity of 2-hydroxyglutaryl-CoA dehydratase from Clostridium symbiosum: toward a bio-based production of adipic acid. *Biochemistry* **50**, 3540-3550 (2011).

372. Felnagle, E.A., Chaubey, A., Noey, E.L., Houk, K.N. & Liao, J.C. Engineering synthetic recursive pathways to generate non-natural small molecules. *Nature chemical biology* **8**, 518-526 (2012).

373. Marcheschi, R.J. et al. A synthetic recursive "+ 1" pathway for carbon chain elongation. *ACS chemical biology* **7**, 689-697 (2012).

374. Marcheschi, R.J., Gronenberg, L.S. & Liao, J.C. Protein engineering for metabolic engineering: Current and next-generation tools. *Biotechnology journal* **8**, 545-555 (2013).

375. Howell, D.M., Harich, K., Xu, H. & White, R.H. α-Keto acid chain elongation reactions involved in the biosynthesis of coenzyme B (7-mercaptoheptanoyl threonine phosphate) in methanogenic Archaea. *Biochemistry* **37**, 10108-10117 (1998).

376. Turk, S.C. et al. Metabolic Engineering toward Sustainable Production of Nylon-6. *ACS synthetic biology* **5**, 65-73 (2015).

377. Xu, H., Andi, B., Qian, J., West, A.H. & Cook, P.F. The α-aminoadipate pathway for lysine biosynthesis in fungi. *Cell biochemistry and biophysics* **46**, 43-64 (2006).

378. Fazius, F., Shelest, E., Gebhardt, P. & Brock, M. The fungal α-aminoadipate pathway for lysine biosynthesis requires two enzymes of the aconitase family for the isomerization of homocitrate to homoisocitrate. *Molecular microbiology* **86**, 1508-1530 (2012).

379. Chang, Y. & Adams, E. Glutarate semialdehyde dehydrogenase of Pseudomonas. Purification, properties, and relation to L-lysine catabolism. *Journal of Biological Chemistry* **252**, 7979-7986 (1977).

380. Smit, B.A. et al. Identification, cloning, and characterization of a Lactococcus lactis branched-chain α-keto acid decarboxylase involved in flavor formation. *Applied and environmental microbiology* **71**, 303-311 (2005).

381. Mak, W.S. et al. Integrative genomic mining for enzyme function to enable engineering of a non-natural biosynthetic pathway. *Nature communications* **6**(2015).

382. Jo, J.-E. et al. Cloning, expression, and characterization of an aldehyde dehydrogenase from Escherichia coli K-12 that utilizes 3-hydroxypropionaldehyde as a substrate. *Applied microbiology and biotechnology* **81**, 51-60 (2008).

383. Yan, D., Lenz, P. & Hwa, T. Overcoming fluctuation and leakage problems in the quantification of intracellular 2-oxoglutarate levels in Escherichia coli. *Applied and environmental microbiology* **77**, 6763-6771 (2011).

384. Gulevich, A.Y., Skonechny, M., Sukhozhenko, A., Skorokhodova, A.Y. & Debabov, V. Study on aerobic biosynthesis of 4-hydroxybutyric acid by Escherichia coli cells upon heterologous expression of the 2-ketoglutarate decarboxylase gene. *Applied biochemistry and microbiology* **51**, 804-811 (2015).

385. Yu, B.J. et al. sucAB and sucCD are mutually essential genes in Escherichia coli. *FEMS microbiology letters* **254**, 245-250 (2006).

386. Spelberg, M. (2014).

387. Tai, Y.-S. et al. Engineering nonphosphorylative metabolism to generate lignocellulose-derived products. *Nature chemical biology* **12**, 247-253 (2016).

388. Gerlach, R.G., Hölzer, S.U., Jäckel, D. & Hensel, M. Rapid engineering of bacterial reporter gene fusions by using Red recombination. *Applied and environmental microbiology* **73**, 4234-4242 (2007).

389. Wu, Y. et al. Establishing a synergetic carbon utilization mechanism for non-catabolic use of glucose in microbial synthesis of trehalose. *Metabolic Engineering* **39**, 1-8 (2017).

390. Lee, J.W., Kim, H.U., Choi, S., Yi, J. & Lee, S.Y. Microbial production of building block chemicals and polymers. *Current opinion in biotechnology* **22**, 758-767 (2011).

391. Yu, J.L., Xia, X.X., Zhong, J.J. & Qian, Z.G. Direct biosynthesis of adipic acid from a synthetic pathway in recombinant Escherichia coli. *Biotechnology and bioengineering* **111**, 2580-2586 (2014).

392. Deng, Y., Ma, L. & Mao, Y. Biological production of adipic acid from renewable substrates: Current and future methods. *Biochemical Engineering Journal* **105**, 16-26 (2016).

393. Wang, J., Wu, Y., Sun, X., Yuan, Q. & Yan, Y. De novo Biosynthesis of Glutarate via α-Keto Acid Carbon Chain Extension and Decarboxylation Pathway in Escherichia coli. *ACS Synthetic Biology* (2017).

394. Song, H. & Lee, S.Y. Production of succinic acid by bacterial fermentation. *Enzyme and microbial technology* **39**, 352-361 (2006).

395. Wang, J. & Zhang, K. Production of mesaconate in Escherichia coli by engineered glutamate mutase pathway. *Metabolic engineering* **30**, 190-196 (2015).

396. Bai, W. et al. Engineering nonphosphorylative metabolism to synthesize mesaconate from lignocellulosic sugars in Escherichia coli. *Metabolic engineering* **38**, 285-292 (2016).

397. Blazeck, J. et al. Metabolic engineering of Yarrowia lipolytica for itaconic acid production. *Metabolic engineering* **32**, 66-73 (2015).

398. Dong, X. et al. Metabolic engineering of Escherichia coli W3110 to produce L-malate. *Biotechnology and Bioengineering* (2016).

399. Wu, X. & Eiteman, M.A. Production of citramalate by metabolically engineered Escherichia coli. *Biotechnology and Bioengineering* **113**, 2670-2675 (2016).

400. Spanjers, C.S. et al. Branched Diol Monomers from the Sequential Hydrogenation of Renewable Carboxylic Acids. *ChemCatChem* **8**, 3031-3035 (2016).

401. Liu, X. et al. A sustainable process for the production of 2-methyl-1, 4-butanediol by hydrogenation of biomass-derived itaconic acid. *Catalysis Today* (2016).

402. Xie, T., Gao, C., Wang, C., Shen, S.e. & Wu, Y. Application of Poly (butylenes 2-methylsuccinate) as Migration Resistant Plasticizer for Poly (vinyl chloride). *Polymer-Plastics Technology and Engineering* **53**, 465-471 (2014).

403. Loos, R., Mijolovic, D., Heimann, J. & Szarka, Z.J. Polyesters based on 2-methylsuccinic acid. (Google Patents, 2013).

404. Zhou, X. et al. Aspernigrins with anti-HIV-1 activities from the marine-derived fungus Aspergillus niger SCSIO Jcsw6F30. *Bioorganic & medicinal chemistry letters* **26**, 361-365 (2016).

405. Sortino, M., Postigo, A. & Zacchino, S. Effects of Chirality on the Antifungal Potency of

Methylated Succinimides Obtained by Aspergillus fumigatus Biotransformations. Comparison with Racemic Ones. *Molecules* **18**, 5669-5683 (2013).

406. Sonntag, F., Buchhaupt, M. & Schrader, J. Thioesterases for ethylmalonyl–CoA pathway derived dicarboxylic acid production in Methylobacterium extorquens AM1. *Applied microbiology and biotechnology* **98**, 4533-4544 (2014).

407. Takasu, A., Iio, Y., Mimura, T. & Hirabayashi, T. Room-temperature polycondensation of dicarboxylic acids and diols catalyzed by water-stable lewis acids. *Polymer journal* **37**, 946-953 (2005).

408. Huang, Q. et al. Preparing acid-resistant Ru-based catalysts by carbothermal reduction for hydrogenation of itaconic acid. *RSC Advances* **5**, 97256-97263 (2015).

409. Holzhäuser, F.J. et al. Electrocatalytic upgrading of itaconic acid to methylsuccinic acid using fermentation broth as a substrate solution. *Green Chemistry* (2017).

410. Verduyckt, J. & De Vos, D.E. Highly selective one-step dehydration, decarboxylation and hydrogenation of citric acid to methylsuccinic acid. *Chemical Science* (2017).

411. Sonntag, F. et al. High-level production of ethylmalonyl-CoA pathway-derived dicarboxylic acids by Methylobacterium extorquens under cobalt-deficient conditions and by polyhydroxybutyrate negative strains. *Applied microbiology and biotechnology* **99**, 3407-3419 (2015).

412. Stueckler, C. et al. Stereocomplementary bioreduction of α, β-unsaturated dicarboxylic acids and dimethyl esters using enoate reductases: enzyme-and substrate-based stereocontrol. *Organic letters* **9**, 5409-5411 (2007).

413. Guccione, E. et al. Reduction of fumarate, mesaconate and crotonate by Mfr, a novel oxygen-regulated periplasmic reductase in Campylobacter jejuni. *Environmental*

*microbiology* **12**, 576-591 (2010).

414. Toogood, H.S., Gardiner, J.M. & Scrutton, N.S. Biocatalytic reductions and chemical versatility of the old yellow enzyme family of flavoprotein oxidoreductases. *ChemCatChem* **2**, 892-914 (2010).

415. Reiße, S. et al. In vitro bioconversion of pyruvate to n-butanol with minimized cofactor utilization. *Frontiers in Bioengineering and Biotechnology* **4**(2016).

416. Xiong, M., Schneiderman, D.K., Bates, F.S., Hillmyer, M.A. & Zhang, K. Scalable production of mechanically tunable block polymers from sugar. *Proceedings of the National Academy of Sciences* **111**, 8357-8362 (2014).

417. Fitzpatrick, T.B., Amrhein, N. & Macheroux, P. Characterization of YqjM, an Old Yellow Enzyme homolog from Bacillus subtilis involved in the oxidative stress response. *Journal of Biological Chemistry* **278**, 19891-19897 (2003).

418. Wardrope, C., Mowat, C.G., Walkinshaw, M.D., Reid, G.A. & Chapman, S.K. Fumarate reductase: Structural and mechanistic insights from the catalytic reduction of 2-methylfumarate. *FEBS letters* **580**, 1677-1680 (2006).

419. Drevland, R.M., Waheed, A. & Graham, D.E. Enzymology and evolution of the pyruvate pathway to 2-oxobutyrate in Methanocaldococcus jannaschii. *Journal of bacteriology* **189**, 4391-4400 (2007).

420. Shen, C.R. & Liao, J.C. Synergy as design principle for metabolic engineering of 1-propanol production in Escherichia coli. *Metabolic engineering* **17**, 12-22 (2013).

421. Jain, R., Sun, X., Yuan, Q. & Yan, Y. Systematically engineering Escherichia coli for enhanced production of 1, 2-propanediol and 1-propanol. *ACS synthetic biology* **4**, 746-756 (2014).

422. Shen, X. et al. High-level De novo Biosynthesis of Arbutin in Engineered Escherichia coli. *Metabolic Engineering* (2017).

423. Howell, D.M., Xu, H. & White, R.H. (R)-citramalate synthase in methanogenic archaea. *Journal of bacteriology* **181**, 331-333 (1999).

424. Causey, T., Shanmugam, K., Yomano, L. & Ingram, L. Engineering Escherichia coli for efficient conversion of glucose to pyruvate. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 2235-2240 (2004).

425. Chen, X. et al. Microbial fed-batch production of 1, 3-propanediol by Klebsiella pneumoniae under micro-aerobic conditions. *Applied Microbiology and Biotechnology* **63**, 143-146 (2003).

426. Berríos-Rivera, S.J., Bennett, G.N. & San, K.-Y. Metabolic engineering of Escherichia coli: increase of NADH availability by overexpressing an NAD+-dependent formate dehydrogenase. *Metabolic engineering* **4**, 217-229 (2002).

427. Jain, R., Huang, J., Yuan, Q. & Yan, Y. Engineering microaerobic metabolism of E. coli for 1, 2-propanediol production. *Journal of industrial microbiology & biotechnology* **42**, 1049-1055 (2015).

428. Yan, Y. & Liao, J.C. Engineering metabolic systems for production of advanced fuels. *Journal of industrial microbiology & biotechnology* **36**, 471-479 (2009).

429. Zhang, X., Tervo, C.J. & Reed, J.L. Metabolic assessment of E. coli as a biofactory for commercial products. *Metabolic engineering* **35**, 64-74 (2016).

430. Vuoristo, K.S. et al. Metabolic engineering of itaconate production in Escherichia coli. *Applied microbiology and biotechnology* **99**, 221-228 (2015).

431. Harder, B.-J., Bettenbrock, K. & Klamt, S. Model-based metabolic engineering enables

high yield itaconic acid production by Escherichia coli. *Metabolic engineering* **38**, 29-37 (2016).

432. Joo, J.C. et al. Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chemical Science* **8**, 1406-1413 (2017).

433. Kitzing, K. et al. The 1.3 Å crystal structure of the flavoprotein YqjM reveals a novel class of old yellow enzymes. *Journal of Biological Chemistry* **280**, 27904-27913 (2005).

434. Opperman, D.J. et al. Crystal structure of a thermostable old yellow enzyme from Thermus scotoductus SA-01. *Biochemical and biophysical research communications* **393**, 426-431 (2010).

435. Schittmayer, M. et al. Old Yellow Enzyme-Catalyzed Dehydrogenation of Saturated Ketones. *Advanced Synthesis & Catalysis* **353**, 268-274 (2011).

436. Chemler, J.A., Fowler, Z.L., McHugh, K.P. & Koffas, M.A. Improving NADPH availability for natural product biosynthesis in Escherichia coli by metabolic engineering. *Metabolic engineering* **12**, 96-104 (2010).

437. Balzer, G.J., Thakker, C., Bennett, G.N. & San, K.-Y. Metabolic engineering of Escherichia coli to minimize byproduct formate and improving succinate productivity through increasing NADH availability by heterologous expression of NAD+-dependent formate dehydrogenase. *Metabolic engineering* **20**, 1-8 (2013).

438. Zhang, Y., Huang, Z., Du, C., Li, Y. & Cao, Z.a. Introduction of an NADH regeneration system into Klebsiella oxytoca leads to an enhanced oxidative and reductive metabolism of glycerol. *Metabolic Engineering* **11**, 101-106 (2009).

439. Shen, C.R. et al. Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Applied and environmental microbiology* **77**, 2905-2915 (2011).

440. Kim, J.-W. et al. Enhanced production of 2, 3-butanediol by engineered Saccharomyces cerevisiae through fine-tuning of pyruvate decarboxylase and NADH oxidase activities. *Biotechnology for biofuels* **9**, 265 (2016).

441. Murtaza, G. et al. Caffeic acid phenethyl ester and therapeutic potentials. *BioMed research international* **2014**(2014).

442. Widjaja, A., Yeh, T.-H. & Ju, Y.-H. Enzymatic synthesis of caffeic acid phenethyl ester. *Journal of the Chinese Institute of Chemical Engineers* **39**, 413-418 (2008).

443. Song, M.K., Cho, A.R., Sim, G. & Ahn, J.-H. Synthesis of diverse hydroxycinnamoyl phenylethanoid esters using Escherichia coli. *Journal of agricultural and food chemistry* **67**, 2028-2035 (2019).

444. Yakandawala, N., Romeo, T., Friesen, A. & Madhyastha, S. Metabolic engineering of Escherichia coli to enhance phenylalanine production. *Applied microbiology and biotechnology* **78**, 283-291 (2008).

445. Juminaga, D. et al. Modular engineering of L-tyrosine production in Escherichia coli. *Appl. Environ. Microbiol.* **78**, 89-98 (2012).

446. Ding, R. et al. Introduction of two mutations into AroG increases phenylalanine production in Escherichia coli. *Biotechnology letters* **36**, 2103-2108 (2014).

447. Kim, B., Binkley, R., Kim, H.U. & Lee, S.Y. Metabolic engineering of Escherichia coli for the enhanced production of l-tyrosine. *Biotechnology and bioengineering* **115**, 2554-2564 (2018).

448. Marinova, E., Georgiev, L., Totseva, I., Seizova, K. & Milkova, T. Antioxidant activity and mechanism of action of some synthesised phenolic acid amides of aromatic amines. *Czech Journal of Food Sciences* **31**, 5-13 (2013).

449. Park, J. N-coumaroyldopamine and its analogues found in cocoa (Theobroma cacao L.) are potent beta-adrenoceptor agonists suppressing platelet activation via increasing cAMP production. *Experimental Biology* **566**, 566 (2005).

450. Hernández-Chávez, G., Martinez, A. & Gosset, G. Metabolic engineering strategies for caffeic acid production in Escherichia coli. *Electronic Journal of Biotechnology* (2018).

451. Tan, H., Aziz, A.A. & Aroua, M. Glycerol production and its applications as a raw material: A review. *Renewable and Sustainable Energy Reviews* **27**, 118-127 (2013).

452. Wang, J. et al. Genetic engineering of Escherichia coli to enhance production of L-tryptophan. *Applied microbiology and biotechnology* **97**, 7587-7596 (2013).

453. Jing, K. et al. Overproduction of L-tryptophan via simultaneous feed of glucose and anthranilic acid from recombinant Escherichia coli W3110: Kinetic modeling and process scale-up. *Biotechnology and bioengineering* **115**, 371-381 (2018).

454. Nakamura, C.E. & Whited, G.M. Metabolic engineering for the microbial production of 1, 3-propanediol. *Current opinion in biotechnology* **14**, 454-459 (2003).

455. Saxena, R., Anand, P., Saran, S. & Isar, J. Microbial production of 1, 3-propanediol: recent developments and emerging opportunities. *Biotechnology advances* **27**, 895-913 (2009).

456. Kimura, E. Metabolic engineering of glutamate production. in *Microbial Production of l-Amino Acids* 37-57 (Springer, 2003).

457. Schneiderman, D.K. & Hillmyer, M.A. 50th Anniversary Perspective: There Is a Great Future in Sustainable Polymers. *Macromolecules* **50**, 3733-3749 (2017).

458. Becker, J., Lange, A., Fabarius, J. & Wittmann, C. Top value platform chemicals: bio-based production of organic acids. *Current Opinion in Biotechnology* **36**, 168-175 (2015).

459. Liu, X. et al. A sustainable process for the production of 2-methyl-1, 4-butanediol by hydrogenation of biomass-derived itaconic acid. *Catalysis Today* **274**, 88-93 (2016).

460. Zheng, L. et al. Multiblock copolymers composed of poly (butylene succinate) and poly (1, 2-propylene succinate): Effect of molar ratio of diisocyanate to polyester-diols on crosslink densities, thermal properties, mechanical properties and biodegradability. *Polymer Degradation and Stability* **95**, 1743-1750 (2010).

461. Li, W. et al. Targeting metabolic driving and intermediate influx in lysine catabolism for high-level glutarate production. *Nature communications* **10**, 1-8 (2019).

462. Chen, G.S., Siao, S.W. & Shen, C.R. Saturated mutagenesis of ketoisovalerate decarboxylase V461 enabled specific synthesis of 1-pentanol via the ketoacid elongation cycle. *Scientific reports* **7**, 11284 (2017).

463. De Carvalho, J.C., Magalhaes, A. & Soccol, C.R. Biobased itaconic acid market and research trends—Is it really a promising chemical. *Chim. Oggi-Chem. Today* **36**, 56-58 (2018).

# APPENDIX A

## LIST OF PUBLICATIONS

*Peer-reviewed journal papers*

1. **Jian Wang**, Chenyi Li, Yusong Zou, Yajun Yan. (2019) Bacterial synthesis of C3-C5 diols via extending amino acid catabolism. Submitted to *Nat. Commun.*
2. **Jian Wang**, Ruihua Zhang, Yan Zhang, Yaping Yang, Yuheng Lin, Yajun Yan. (2019) Developing a pyruvate-driven metabolic scenario for growth-coupled microbial production. *Metab. Eng.*, 55, 191-200.
3. Ruihua Zhang, Chenyi Li, **Jian Wang**, Yajun Yan. (2019) Microbial ligninolysis: Towards a bottom-up approach for lignin upgrading. *Biochemistry*, 58(11): 1501-1510.
4. Ruihua Zhang, Chenyi Li, **Jian Wang**, Yaping Yang, Yajun Yan. (2018) Microbial production of small medicinal molecules and biologics: From nature to synthetic pathways. *Biotechnol. Adv.*, 36 (8): 2219-2231.
5. Yaping Yang, Yuheng Lin, **Jian Wang**, Yifei Wu, Ruihua Zhang, Mengyin Cheng, Xiaolin Shen, Jia Wang, Zhenya Chen, Chenyi Li, Qipeng Yuan, Yajun Yan. (2018) Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. *Nat. Commun.*, 9 (1): 3043.
6. Jia Wang, Xiaolin Shen, **Jian Wang**, Yaping Yang, Qipeng Yuan, Yajun Yan. (2018) Exploring the promiscuity of phenol hydroxylase from *Pseudomonas stutzeri* OX1 for the biosynthesis of phenolic compounds. *ACS Syn. Biol.*, 7(5):1238-1243.
7. **Jian Wang**, Yaping Yang, Ruihua Zhang, Xiaolin Shen, Zhenya Chen, Qipeng Yuan,

Yajun Yan. (2018) Microbial production of branched-chain dicarboxylate 2-methylsuccinic acid via enoate reductase-mediated bioreduction. *Metab. Eng.*, 45, 1-10.

8. Narendran Sekar, **Jian Wang**, Yan Zhou, Yi Fang, Yajun Yan, Ramaraja P Ramasamy. (2018) Role of respiratory terminal oxidases in the extracellular electron transfer ability of cyanobacteria. *Biotechnol. Bioeng.*, 115(5): 1361-1366.
9. Yifei Wu, **Jian Wang**, Xiaolin Shen, Jia Wang, Zhenya Chen, Xinxiao Sun, Qipeng Yuan, Yajun Yan. (2018) Investigating the strategies for microbial production of trehalose from lignocellulosic sugars. *Biotechnol. Bioeng.*, 115(3):785-790.
10. **Jian Wang**, Monika Mahajani, Sheneika L Jackson, Yaping Yang, Mengyin Chen, Eric M Ferreira, Yuheng Lin, Yajun Yan. (2017) Engineering a bacterial platform for total biosynthesis of caffeic acid derived phenethyl esters and amides. *Metab. Eng.*, 44, 89-99.
11. Zhenya Chen, Xiaolin Shen, **Jian Wang**, Jia Wang, Ruihua Zhang, Justin Forrest Rey, Qipeng Yuan, Yajun Yan. (2017) Establishing an artificial pathway for *de novo* biosynthesis of vanillyl alcohol in *Escherichia coli*. *ACS Syn. Biol.*, 6 (9), 1784-1792.
12. **Jian Wang**, Yifei Wu, Xinxiao Sun, Qipeng Yuan, Yajun Yan. (2017) *De novo* biosynthesis of glutarate via α-keto acid carbon chain extension and decarboxylation pathway in *Escherichia coli*. *ACS Syn. Biol.*, 6 (10), 1922-1930.
13. Zhenya Chen, Xiaolin Shen, **Jian Wang**, Jia Wang, Qipeng Yuan, Yajun Yan. (2017) Rational engineering of p-hydroxybenzoate hydroxylase to enable efficient gallic acid synthesis via a novel artificial biosynthetic pathway. *Biotechnol. Bioeng.*, 114 (11), 2571-2580.
14. Xiaolin Shen, Jia Wang, **Jian Wang**, Zhenya Chen, Qipeng Yuan, Yajun Yan. (2017) High-level *de novo* biosynthesis of arbutin in engineered *Escherichia coli*. *Metab. Eng.*,

42, 52-58.

15. Jia Wang, Xiaolin Shen, Rachit Jain, **Jian Wang**, Qipeng Yuan, Yajun Yan. (2017) Establishing a novel biosynthetic pathway for the production of 3, 4-dihydroxybutyric acid from xylose in *Escherichia coli*. *Metab. Eng.*, 41: 39-45.
16. **Jian Wang**, Sanjay Guleria, Mattheos AG Koffas, Yajun Yan. (2016) Microbial production of value-added nutraceuticals. *Cur. Opin. Biotechnol.*, 37:97-104.
17. Xinxiao Sun, Xiaolin Shen, Rachit Jain, Yuheng Lin, **Jian Wang**, Jing Sun, Jia Wang, Yajun Yan, Qipeng Yuan. (2015) Synthesis of chemicals by metabolic engineering of microbes. *Chem. Soc. Rev.*, 44 (11): 3760-3785.

*Book chapters*

1. **Jian Wang**, Yaping Yang, Yajun Yan. Bioproduction of Resveratrol. Book: Biotechnology of Natural Products, 2018. Springer. Chapter 3: 61-79.
2. **Jian Wang**, Yajun Yan. Glycolysis and Its Metabolic Engineering Applications. Book: Engineering Microbial Metabolism for Chemical Synthesis, 2018. World Scientific. Chapter 1: 1-33.
3. Yaping Yang, **Jian Wang**, Ruihua Zhang, Yajun Yan. Antisense RNA Elements for Downregulating Expression, 2019. Microbial Metabolic Engineering, 1927: 23-35.