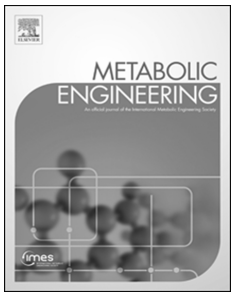
Journal Pre-proof



Building microbial factories for the production of aromatic amino acid pathway derivatives: From commodity chemicals to plant-sourced natural products

Mingfeng Cao, Meirong Gao, Miguel Suastegui, Yanzhen Mei, Zengyi Shao

PII: S1096-7176(19)30163-6

DOI: https://doi.org/10.1016/j.ymben.2019.08.008

Reference: YMBEN 1587

To appear in: *Metabolic Engineering*

Received Date: 15 April 2019

Revised Date: 3 August 2019

Accepted Date: 7 August 2019

Please cite this article as: Cao, M., Gao, M., Suastegui, M., Mei, Y., Shao, Z., Building microbial factories for the production of aromatic amino acid pathway derivatives: From commodity chemicals to plant-sourced natural products, *Metabolic Engineering* (2019), doi: https://doi.org/10.1016/j.ymben.2019.08.008.

Submitted to Metabolic Engineering on August 2nd, 2019

Building microbial factories for the production of aromatic amino acid pathway derivatives: from commodity chemicals to plant-sourced natural products

Mingfeng Cao[1, 2, †], Meirong Gao[1, 2, †], Miguel Suastegui[1, 2], Yanzhen Mei[5], and Zengyi Shao[1, 2, 3, 4, *]

[1] Department of Chemical and Biological Engineering
[2] NSF Engineering Research Center for Biorenewable Chemicals (CBiRC)
[3] Interdepartmental Microbiology Program
[4] The Ames Laboratory
4140 Biorenewables Research Laboratory,
Iowa State University, Ames, IA 50011, USA

[5]School of Life Sciences
No.1 Wenyuan Road, Nanjing Normal University,
Qixia District, Nanjing, China, 210023

[†]Contributed equally to this work

[*]Correspondence to: Dr. Zengyi Shao,
4140 Biorenewables Research Laboratory,
Iowa State University, Ames, IA 50011, USA
Phone: 515-294-1132
Email: zyshao@iastate.edu

Abstract

The aromatic amino acid biosynthesis pathway, together with its downstream branches, represents one of the most commercially valuable biosynthetic pathways, producing a diverse range of complex molecules with many useful bioactive properties. Aromatic compounds are crucial components for major commercial segments, from polymers to foods, nutraceuticals, and pharmaceuticals, and the demand for such products has been projected to continue to increase at national and global levels. Compared to direct plant extraction and chemical synthesis, microbial production holds promise not only for much shorter cultivation periods and robustly higher yields, but also for enabling further derivatization to improve compound efficacy by tailoring new enzymatic steps. This review summarizes the biosynthetic pathways for a large repertoire of commercially valuable products that are derived from the aromatic amino acid biosynthesis pathway, and it highlights both generic strategies and specific solutions to overcome certain unique problems to enhance the productivities of microbial hosts.

Keywords: aromatic amino acid biosynthesis, shikimate pathway, *de novo* biosynthesis, microbial production, flavonoids, stilbenoids, benzylisoquinoline alkaloids

Highlights

- The aromatic amino acid biosynthesis pathway yields commercially valuable products.
- The broad scope of products from aromatic amino acid biosynthesis is described.
- Biosynthetic pathways and interconnections among the sub-branches are presented.
- High-level organization of enzymes of the upstream module is illustrated.
- The exemplary sophisticated regulations in the upstream pathway at various levels are highlighted.

1. Introduction

Over the past two decades, technologies arising from synthetic biology and systems biology have revolutionized metabolic engineering to establish bio-based production in engineered microorganisms. The aromatic amino acid biosynthesis pathway, together with its downstream branches, represents one of the most valuable biosynthetic pathways. The products derived from this pathway range widely, from various high-volume low-value commodity chemicals to the realm of specialty chemicals and complex natural products (Suastegui and Shao, 2016; Thompson et al., 2015; Zhang et al., 2018).

The pathways leading up to the synthesis of L-tryptophan (L-Trp), L-phenylalanine (L-Phe), and L-tyrosine (L-Tyr) nodes are ubiquitous in all microorganisms, constituting the backbone reactions for the synthesis of a diverse class of compounds. This portion of the pathway can be subdivided into the shikimate pathway, the L-Tyr branch, the L-Phe branch, and the L-Trp branch. The model hosts *Escherichia coli* and *Saccharomyces cerevisiae* have been manipulated to produce this group of compounds in most studies, but other species such as *Scheffersomyces stipitis*, *Pseudomonas putida*, and *Corynebacterium glutamicum* have also been engineered as production hosts, and in many cases, they outperform *E. coli* and *S. cerevisiae* thanks to their special biochemical and metabolic features.

The downstream pathways beyond the biosynthesis of the three aromatic amino acids open a treasure trove of high-value molecules with extremely diverse structures, the majority of which are produced by plants. These natural products encompass a billion dollar market (Pandal, 2014; Rawat et al., 2013; Rinner and Hudlicky, 2012; Winter and Tang, 2012), but their extraction from plant tissues requires immense quantities of biomass and cumbersome separation processes. Open-field plant growth and the associated costs are also susceptible to environmental factors,

which are associated with variability in product yield and composition. In addition, the complex and delicate chemistry involved in forming these aromatics makes their *de novo* chemical synthesis very challenging and economically unviable at commercial scales. The implementation of microbial factories carrying the combination of tailored endogenous metabolism and heterologous biosynthetic pathways represents an encouraging solution to these problems. It potentially ensures sustainable and greener processes with higher yields, and also enables further derivatization for new drug development. Unlike the biosynthesis of the high-volume low-value chemicals, which has been attempted in various microbial hosts, strain engineering to produce these plant-derived natural products has been mainly performed in *E. coli* and *S. cerevisiae*. Between these two hosts, *S. cerevisiae* also has the advantage of expressing membrane-associated eukaryotic enzymes (e.g., cytochrome P450s), which are required for the synthesis of some complex natural products derived from plants (Chang and Keasling, 2006; Jung et al., 2011).

This review will focus on illustrating the broad scope of products that are derived from the aromatic amino acid biosynthesis, their biosynthetic pathways, and the interconnections among individual sub-branches. The article begins with the reaction skeleton of aromatic amino acid biosynthesis, the different enzyme organizations in prokaryotes and eukaryotes, and the sophisticated regulatory mechanisms at various levels. The products belonging to the relevant sub-branch reactions are color-coded and grouped together for effective discussion. The relatively generic strategies to enhance the upstream flux and the specific solutions to the individual problems appearing in the biosynthesis of downstream products are addressed separately. We provide future perspectives and recommendations of novel strategies to increase the productivities at the end.

2. The backbone reactions of aromatic biosynthesis

2.1 The shikimate pathway

The shikimate pathway initiates with the condensation of the intermediates of glycolysis and the pentose phosphate pathway (PPP), phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P), respectively (**Figure 1a**). In *S. cerevisiae*, the condensation reaction is catalyzed by the 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase isozymes Aro3 and Aro4, which are subjected to feedback inhibition caused by L-Phe and L-Tyr, respectively. In *E. coli*, three isozymes—AroF, AroG, and AroH—play a role in the condensation reaction and are regulated by L-Tyr, L-Phe, and L-Trp, respectively, via the transcriptional regulatory protein TyrR (Pittard and Yang, 2008). Strikingly, the enzymes in charge of the subsequent series of reactions from DAHP to 5-enolpyruvylshikimate-3-phosphate (EPSP) in different organisms have different genetic organizations. In lower eukaryotes, the subsequent five distinct catalytic domains are fused to form a pentafunctional enzyme that is encoded by a single open reading frame, *ARO1*, to complete the reactions whereas bacteria and plants direct the five reactions by separately encoded enzymes. Finally, the central branch point metabolite chorismic acid is produced from EPSP by chorismate synthase, Aro2, and it enters one of the two aromatic amino acid branches. The determination of these early reactions was mainly based on isolating the intermediates that accumulated in the respective gene disruption strains and analyzing the responses of the auxotrophs to the supplementation of the proposed metabolic intermediates. The detailed biochemical features of the enzymes involved in the upstream module have been summarized in two review articles (Herrmann, 1995; Pittard and Yang, 2008).

2.2 The L-Tyr branch and L-Phe branch

L-Tyr and L-Phe differ only at the 4-hydroxyl functional group, and their biosynthetic steps share the first one, the rearrangement of the enolpyruvyl side chain of chorismate to form prephenate (**Figure 1b and 1c**). In yeast and some bacteria (e.g., *Bacillus subtilis*), the monofunctional chorismate mutase Aro7 is responsible for this conversion, following which, the pathway diverges into L-Phe biosynthesis through prephenate dehydratase Pha2 and into L-Tyr biosynthesis through another prephenate dehydrogenase Tyr1, yielding the intermediates phenylpyruvate and 4-hydroxylphenylpyruvate, respectively. In contrast, *E. coli* uses two bifunctional enzymes to complete the rearrangement of chorismate to prephenate and the subsequent step. The gene *pheA* encodes chorismate mutase-prephenate dehydratase, which is feedback inhibited by L-Phe, whereas *tyrA* encodes chorismate mutase-prephenate dehydrogenase, which is feedback inhibited by L-Tyr. Interestingly, the N-terminal parts of these two bifunctional enzymes that have chorismate mutase activity do not share significant homology with each other or the monofunctional chorismate mutase found in yeasts (Braus, 1991). The last step in L-Phe and L-Tyr biosynthesis is a reversible transamination reaction, catalyzed by aromatic amino acid aminotransferases; the isozymes are encoded by *aro8* and *aro9* in fungi and mainly by *tyrB* in bacteria. Furthermore, some bacteria, such as cyanobacteria and *Pseudomonas* spp., perform the dehydrogenation reaction prior to the aminotransfer step by using arogenate as an intermediate, which is a preferred route in plants (Bonner and Jensen, 1987).

2.3 The L-Trp branch

The L-Trp branch is organized in a more complicated manner than the L-Tyr and L-Phe branches (**Figure 1b and 1c**). A series of multifunctional enzymes are utilized, but the enzyme association varies from species to species. In *S. cerevisiae*, six catalytic domains belonging to five enzymes (encoded by *TRP1–5*) are involved, with *TRP3* containing two independent catalytic domains. The first step involves the transfer of the amino group from glutamine to chorismate, yielding anthranilate, pyruvate, and glutamate as products. The reaction is catalyzed by a complex consisting of anthranilate synthase Trp2 (feedback inhibited by L-Trp) and the N-terminal glutamine amidotransferase domain of Trp3 (referred to as Trp3C). The C-terminus of Trp3 (referred to as Trp3B) is an indole-3-glycerol-phosphate (InGP) synthase catalyzing the fourth step in the L-Trp branch. In contrast, in *E. coli*, the glutamine amidotransferase domain is fused with the anthranilate phosphoribosyl transferase domain (catalyzing the second step), both of which are encoded by *trpD*. Note that in some literature, glutamine amidotransferase and anthranilate phosphoribosyl transferase are ambiguously specified separately as TrpG and TrpD because in some bacteria (e.g., *Serratia marcescens* and *Sulfolobus solfataricus* (Knochel et al., 1999; Spraggon et al., 2001)), the two functions are encoded by two separate genes. In *E. coli*, *trpG* only represents the N-terminal regions of the bifunctional gene *trpD*, which forms a heterotetramer with anthranilate synthase (encoded by *trpE*) to complete the first step of anthranilate synthesis. Following this reaction, the anthranilate phosphoribosyl transferase domain located at the C-terminal of TrpD proceeds to introduce a phosphoribosyl moiety from 5-phosphoribsyl pyrophosphate (PRPP) to the amino group of anthranilate, yielding phosphoribosyl anthranilate (PRA). In fungi, this step is carried out by the monofunctional enzyme Trp4.

The third step in the L-Trp branch entails an internal redox reaction to open the ring, forming ketone carboxyphenylamino-1-deoxyribulose 5-phosphate (CDRP), which is catalyzed by the monofunctional PRA isomerase Trp1 in *S. cerevisiae*, *Kluyveromyces lactis*, and *Candida maltosa* (Braus et al., 1985; Stark and Milner, 1989). Interestingly, in many other fungi (e.g., *Neurospora crassa*, *Aspergillus nidulans*, and *Aspergillus niger* (Kos et al., 1985; Mullaney et al., 1985; Schechtman and Yanofsky, 1983)), the enzyme in charge of this step is fused with the bifunctional enzyme Trp3 to form a trifunctional enzyme with the PRA isomerase domain arranged at the C-terminal. Fusion of a PRA isomerase domain with an N-terminal InGP synthase domain was observed in *E. coli*, together as a bifunctional enzyme TrpC, whereas the InGP synthase domain specified as Trp3B is located at the C-terminal of the glutamine amidotransferase specified as Trp3 in *S. cerevisi*ae.

The final reaction in the L-Trp branch is catalyzed by a bifunctional L-Trp synthase that contains two separate active sites in *S. cerevisiae*, encoded by *TRP5* (Miles, 1979; Yanofsky, 2005). One active site cleaves InGP into indole and glyceraldehyde-3-phosphate, and the other one condenses the indole and serine to L-Trp. In many other organisms, including *E. coli*, this function is specified by two separate genes, *trpA* and *trpB*, encoding the alpha and beta subunits of the tetrameric L-Trp synthase, respectively. Crystal structures revealed a hydrophobic tunnel connecting the two active sites and preventing the escape of indole during the movement from one subunit to the other (Hyde et al., 1988).

3. Regulation involved in the backbone reactions

Due to the essential role of amino acid biosynthesis in central metabolism, cells have evolved various mechanisms to maintain amino acid homeostasis. Akashi H. et al. calculated the energy

cost of synthesizing various amino acids in *E. coli* based on the consumption of ATP/GTP and NADH/NADPH/$FADH_2$ molecules. It was concluded that 74.3, 52.0, and 50.0 high-energy phosphate bonds contained in ATP/GTP molecules are required to synthesize one molecule of L-Trp, L-Phe, and L-Tyr, respectively (here the consumption of reducing agents is converted to that of ATP based on a ratio of 2 ATP per 1 NADH) (Akashi and Gojobori, 2002). It is not surprising that intrinsically complicated regulations are exerted as different control tiers, impeding the high-level production of any derivatives produced from this pathway especially considering that the corresponding number of energy cost for synthesizing one of the other amino acids is merely 11.7 to 38.3. Prokaryotes and lower eukaryotes share some common features in regulating enzyme activities but utilize distinct mechanisms to regulate the pathway genes at the transcriptional and translational levels. Two reviews have comprehensively summarized the regulations involved in aromatic amino acid biosynthesis in *S. cerevisiae* and *E. coli*, as well as the biochemical characteristics of the relevant enzymes (Braus, 1991; Pittard and Yang, 2008). The section below only highlights the most representative regulatory events that were detailed in these previous reviews. Recapitulating these examplary regulations will assist readers to understand the intrinsically complicated regulations behind aromatic amino acid biosynthesis as well as design solutions to overcome the hurdles constraining the production of many valuable compounds from this pathway.

3.1 Regulation of enzyme activity

Allosteric checkpoints exist at the entrance of the shikimate pathway, as well as the branching points between the L-Tyr/L-Phe and L-Trp branches. In yeast, the DHAP synthase isozymes Aro3 and Aro4 are sensitive to the feedback inhibition posed by L-Phe and L-Tyr,

respectively. The most common strategy to open the sluice to the shikimate pathway is to overexpress the mutant Aro4^{K229L} and/or Aro3^{K222L} to completely block the binding of L-Tyr and L-Phe. Feedback-insensitive mutants Aro4^{Q166K} and Aro4^{G226S} have also been reported in some studies (Koopman et al., 2012; Trenchard and Smolke, 2015). The overexpressed mutant typically leads to efficient condensation between PEP and E4P, and there is no need to remove the wild-type versions of the genes. In *E. coli*, it is believed that 20%, 80%, and a marginal percentage of the total activity of DHAP synthase is provided by L-Tyr-sensitive AroF, L-Phe-sensitive AroG, and L-Trp-sensitive AroH, respectively (Bongaerts et al., 2001; Tribe et al., 1976). The feedback inhibition-insensitive mutants AroG^{fbr} and AroF^{fbr} are therefore usually overexpressed to deregulate the feedback inhibition.

Chorismate mutase Aro7 initiates the L-Tyr/L-Phe branch by converting chorismate to prephenate in *S. cerevisiae* and is subjected to feedback inhibition by the end product L-Tyr. Similarly, L-Trp acts as a repression effector to the anthranilate synthase–glutamine amidotransferase complex (encoded by *TRP2-TRP3C*, **Figure 1b**), which is in charge of the first committed step of the L-Trp branch at the same time as the activator of Aro7, which is a dual control strategy utilized by *S. cerevisiae* to balance the fluxes flowing into the two branches. The overexpression of Aro7^{G141S} (Aro7^{T266I} or Aro7^{T227L}) has been conveniently used to allow the flux to enter the L-Tyr/L-Phe branch (Rodriguez et al., 2017a; Rodriguez et al., 2015). In *E. coli*, the equivalent feedback-insensitive mutants TyrA^{fbr}, PheA^{fbr}, and TrpE^{fbr}-TrpD$_N$ (or called TrpE^{fbr}-TrpG, **Figure 1c**) have been widely overexpressed to produce L-Trp-derived products.

3.2 Transcriptional regulation

Intrinsically complicated regulations have also been recognized at the transcriptional level (Hinnebusch, 1988; Lee and Hahn, 2013), some of which have not been fully elucidated. The regulatory mechanisms employed by eukaryotes significantly differ from those used by prokaryotes. Eukaryotes use global regulators to modulate amino acid biosynthesis as an entire pool in a non-specific manner, whereas prokaryotes generally rely on pathway-specific transcriptional regulators in response to the fluctuations of the three aromatic amino acids.

3.2.1 Transcriptional regulation in the lower eukaryotes

In contrast to bacteria, yeast and fungi maintain relatively large intracellular amino acid pools, and the corresponding biosynthetic genes are expressed at significant levels, which is hence named “basal control”. Under amino acid starvation, the expression of more than 30 genes involved in various amino acid biosynthetic pathways is enhanced, and the genes involved are not specifically from the pathway corresponding to the amino acid that is deficient. Such a cross-pathway regulation is conducted by global transcription regulators via the so-called “general control mechanism” (Hinnebusch, 1988). The general control mediated by the transcription regulator Gcn4 in *S. cerevisiae* is the best characterized system of the amino acid biosynthesis network. The Gcn4 recognition site has a consensus sequence of 5′-ATGA(C/G)TCAT-3′ (Braus, 1991), which has been found in the upstream activation sequences of the promoters of many amino acid biosynthesis genes. These promoters are typically regulated by both Gcn4-mediated general control and amino acid-independent basal control (Arndt et al., 1987; Struhl, 1986). Specific to aromatic amino acid biosynthesis, the regulatory events at the entrance of the shikimate pathway and the L-Trp branch are briefly discussed below.

In addition to the feedback inhibition, due to the important entry position of DHAP synthase in the entire pathway, *ARO3* and *ARO4* are also regulated at the transcriptional level. One Gcn4-binding site was found in the promoter of *ARO3*. A single mutation at this site or disruption of the *GCN4* gene results in a decreased basal level of *ARO3* expression (Braus, 1991; Paravicini et al., 1989). For *ARO3*, Gcn4 only maintains its basal-level transcription because no increase in *ARO3* expression is observed under amino acid starvation, which is different from the regulatory role of Gcn4 on *ARO4* (Braus, 1991). The basal expression level of *ARO4* is not affected by disrupting the Gcn4-recognition site in the upstream activation sequence, but amino acid starvation will no longer increase its transcription (Kunzler et al., 1992).

Gcn4-mediated transcription control is also often used in the L-Trp branch. The Gcn4-binding sites appear in the promoters of four out of the five genes (*TRP2–5*) in this branch, with *TRP1* as the only exception. Unlike the operon structure *trpEDCBA* employed by bacteria, the genes involved in amino acid biosynthesis are widely spread over on the yeast genome, indicating that a higher level of control must exist to coordinate the expression of the genes from a single pathway. For example, the heterodimeric enzyme complex consisting of anthranilate synthase (Trp2) and glutamine amidotransferase (Trp3C) catalyzes the first step in the L-Trp branch. Considering that equimolar amounts of the two monomeric polypeptide chains are needed, the expression of Trp2 and Trp3 must be coordinated (Braus, 1991). Each of the two promoters contains a Gcn4-binding site (Zalkin et al., 1984). A fascinating dual control was identified for Trp4. Three Gcn4-binding sites were found in its promoter region, one of which could also recruit another set of transcriptional regulators— Pho2/Bas2—involved in regulating Pi metabolism (Braus et al., 1989). Anthranilate phosphoribosyltransferase Trp4 requires PRPP as a substrate. Even under amino acid starvation, if Pi is limiting, *TRP4* expression is still

restricted by the occupancy of Pho2/Bas2 on the site that could otherwise bind to Gcn4 to fully activate *TRP4* expression, thereby preventing the unnecessary overproduction of the enzyme.

3.2.2 Transcriptional regulation in prokaryotes

In contrast to the employment of global regulators in eukaryotes, the response to the intracellular levels of the three aromatic amino acids in prokaryotes is much more localized. The transcription of only the genes in the specific amino acid biosynthetic pathway is enhanced when the relevant amino acid is in deficiency. This mechanism, named “metabolic control”, is mediated through one or both of the pathway-specific regulatory proteins, TyrR and TrpR, with intracellular levels that are delicately balanced to fulfill particular regulatory roles (Pittard and Yang, 2008).

TyrR is a large protein, composed of 513 amino acids arranged as an N-terminal domain, a central domain, and a C-terminal domain. The unique feature of TyrR is that the three domains are relatively independent, have different oligomerization preferences, and all have regulatory roles. The N- and C-terminal domains tend to form dimers, whereas the central domain prefers to form a hexamer. Depending on the presence of effector molecules and the positions of the TyrR-binding sites relative to the key regions (e.g., −35, −10, and the transcription initiation site) in a promoter, TyrR will take one oligomerization state over the other and can act either as a repressor or an activator, and in some cases, plays a crucial dual role in modulating the expression of a single gene.

The *tyrR* regulon has at least nine genes that are separately transcribed, and the palindromic sequence $TGTAAAN_6TTTACA$, together with two flanking nucleotides at each end, comprises the *tyrR* box. **Table 1** provides a snapshot of many detailed mechanistic studies summarized in

an earlier review (Pittard and Yang, 2008). At least two *tyrR* boxes were found in the transcription units for eight out of the nine genes, with the exception of *aroG*, which contains only one box. The binding of TyrR to one of these boxes requires the presence of ATP and L-Tyr, and the box is named a weak box; the other box, the strong box, is bound to TyrR in an effector-independent manner. In many cases, one of the *tyrR* boxes is close to or overlaps with the RNA polymerase-binding site, mediating the regulatory role of TyrR to the transcription of the gene. In special cases, DNA looping is required between distant binding sites, and the interaction might be mediated through TyrR oligomerization.

TrpR is a much smaller regulatory protein than TyrR, consisting of only 108 amino acids. At least five transcription units are involved in the *trpR* regulon, namely, the *trpEDCBA* operon, *aroH*, *trpR*, *Mtr*, and *aroL*. TrpR remains a dimer, and its binding to the operator requires prior binding with two L-Trp molecules, which will change the conformation of the HTH domains to facilitate the recognition. Similar to the mechanism described for TyrR, L-Trp-mediated repression is conducted by blocking the binding of RNA polymerase (−35 and −10 positions) or its progression (+1 position). TrpR also coordinates with TyrR to regulate the transcription of *Mtr* and *aroL* (**Table 1**) (Jeeves et al., 1999; Pittard and Davidson, 1991).

Lastly, because no nuclear membrane separates transcription and translation events in prokaryotes, another unique control is employed to modulate transcription in response to the level of specific charged aminoacyl-tRNAs. This mechanism, transcriptional attenuation, is found in the regulation of the *trpEDCBA* operon and *pheA* (Pittard and Yang, 2008). Briefly, a leader sequence encoding a short 14-amino acid peptide exists upstream of the *trp* operon. Two alternative RNA secondary structures can be assumed by the *trp* operon, and one of them causes the early termination of transcription. The leader sequence contains two adjacent L-Trp codons,

and the sliding of the ribosome will be stalled at this position if the charged tryptophanyl-tRNA is scarce. The resulting structure then allows RNA polymerase to continuously travel into the coding regions of the various genes in the *trp* operon. When tryptophanyl-tRNA is abundant, ribosomes will not stall, an early termination RNA structure is then formed, preventing the unnecessary transcription of the *trp* operon genes to continuously raise the intracellular L-Trp level (Landick et al., 1987). A similar transcriptional attenuation mechanism is used to regulate the expression of *pheA* in response to the concentration of phenylalanyl-tRNA (Landick and Jr., 1992). Considering that TrpE and PheA catalyze the first reaction in the L-Trp and L-Phe branches, respectively, having multiple mechanisms to respond to the fluctuation of the corresponding pathway product is critical to maintain amino acid homeostasis.

3.3. Translational regulations

In eukaryotes, the expression of the global amino acid biosynthesis regulator Gcn4 is modulated by amino acids at the translational level. This transcriptional regulator consists of 281 amino acids, but its mRNA is more than 1,500 nucleotides long. Interestingly, upstream of the start codon AUG, four additional extremely short open reading frames are found, each of which is only composed of two or three sense codons followed by a stop codon (Hinnebusch, 1988; Thireos et al., 1984). Deleting or introducing mutations in this extra leader sequence results in a significant increase in *GCN4* expression. It is believed that the appearance of these extra translation initiation signals substantially inhibits the ability of the ribosome to re-initiate the translation at the actual *GCN4* ORF (Mueller and Hinnebusch, 1986), which is a regulatory mechanism rarely used in *S. cerevisiae* but is quite popular in mammals. Under amino acid-

deficient conditions, the interference of these upstream AUG start codons is suppressed, and the translation of *GCN4* mRNA becomes most efficient for regulating amino acid biosynthesis.

4. Upstream products derived from the shikimate pathway

With a limited number of conversion steps, many commodity chemicals can be synthesized from the intermediates in the shikimate pathway (**Figure 2 and Table 2**). Recently, Averesch et al. and Huccetogullari et al. discussed metabolic engineering endeavors to produce a variety of compounds derived from this upstream module (Averesch and Kromer, 2018; Huccetogullari et al., 2019). To avoid redundancy, in the following sections, we will highlight the standard strategies for producing this group of compounds and expand upon the product derivatization strategies that have not been included in the previous reviews. The products are color-coded in **Figure 2** and organized by key nodes in the pathway. Most studies have been conducted in *E. coli* and *S. cerevisiae*, but a few other hosts have also demonstrated surprisingly high productivity. Regardless of the species, for *de novo* biosynthesis, the overexpression of the feedback-inhibited DAHP synthase mutant ($Aro3^{fbr}$ or $Aro4^{fbr}$ in *S. cerevisiae* and $AroF^{fbr}$ or $AroG^{fbr}$ in *E. coli*) and transketolase (Tkl1 in *S. cerevisiae* and TktA in *E. coli*) are the most common strategies used to enhance the flux entering the SA pathway and E4P availability, respectively. These strategies will not be repeatedly mentioned in the specific examples. Novel strategies and non-glucose feedstock that enhanced the production will be particularly highlighted.

4.1 The dehydroshikimate node

4.1.1 Shikimic acid and quinic acid

Shikimic acid (SA) is the precursor of oseltamivir phosphate, commercially known as Tamiflu®, an anti-influenza medicine (Federspiel et al., 1999). The commercial processes of SA synthesis include its extraction from the fruit of *Illicium verum* or microbial fermentation. The latter offers a more abundant supply in a timely manner, which is critical during a severe influenza outbreak, when about 30 billion doses of Tamiflu®, equivalent to 3.9 million kg of SA, might be needed (Sunder Rangachari, 2013). Since SA is a metabolic intermediate, its intracellular accumulation relies on the disruption of the downstream conversion. In *E. coli*, the disruption can be achieved by deleting the self-standing SA kinase I AroK and SA kinase II AroL. In the *S. cerevisiae* mutant, $Aro1^{D920A}$ needs to be overexpressed in the *aro1Δ* strain because Aro1 is a large pentafunctional enzyme, with distinct domains (Aro1C, E, D, B, and A) catalyzing the five steps from DAHP to EPSP.

In *E. coli*, glucose is transported into cells via the PEP:carbohydrate phosphotransferase (PTS) system, in which a phosphoryl group is transferred from PEP, converting glucose to glucose 6-phosphate (G6P). To avoid this PEP expenditure, two strategies are often tested. One is to overexpress PEP synthase encoded by *ppsA* to recycle pyruvate back to PEP, and the other is to heterologously express the *Zymomonas mobilis* glucose transporter and glucokinase, encoded by *glf* and *glk*, respectively, in a host deficient in the PTS system, which does not require PEP for glucose transportation but consumes one ATP. Although both approaches have been widely applied, the second one influences SA production more significantly (Chandran et al., 2003). In an engineered *E. coli* overexpressing 3-DHQ synthase encoded by *aroB* to enhance the conversion from DHAP to 3-DHQ, in combination with the aforementioned genetic manipulations, 87 g/L SA was produced by 10 L fed batch fermentation supplemented with 15 g/L yeast extract, with a yield of 0.32 g SA/g glucose and a productivity of 1.64 g/L/h. To avoid

the additional cost associated with aromatic amino acid supplementation for culturing the *aroK*-deletion strain, Lee et al. proposed to utilize a growth phase-dependent promoter to drive the expression of *aroK* (Lee et al., 2017). With this strategy, cells could still synthesize aromatic amino acids while accumulating SA. Furthermore, the enzymes catalyzing the conversion from 3-DHQ to DHS and then to SA were replaced by a bifunctional DHQ-SDH protein from the tree *Populus trichocarpa* to facilitate metabolic channeling. The combined strategy led to an accumulation of SA at 5.33 g/L.

Much higher production of SA has been achieved in *C. glutamicum*, a gram-positive soil bacterium that is used for producing L-glutamate and L-lysine at industrial scales (Kogure et al., 2016). In addition to the standard overexpression and deletion strategies, two interesting strategies were devised to increase the accumulation of SA. When the PTS system was interrupted and replaced by overexpressing the endogenous myo-inositol transporter (IolT1) and glucokinases (Glk1, Glk2, and Ppgk), glycerol and 1,3-dihydroxyacetone were unexpectedly built up at significant amounts. It was reasoned that the overflow of glycolytic intermediates caused this accumulation, and the issue was relieved by overexpressing *gapA* encoding glyceraldehyde-3-phosphate dehydrogenase (Gapdh) and blocking the formation of 1,3-dihydroxyacetone and glycerol. Besides, an aerobic conversion using growth-arrested high-density cells was developed, in which cells were grown in the rich medium for 18-20 h, then centrifuged, and added to the minimal medium with a ratio of 10% of wet cell weight per culture volume. Compared to the routine fed-batch fermentation, using growth-arrested cells reduced the carbon loss to cell growth and the byproduct formation, leading to the highest SA titer and yield attained by microbial biosynthesis reported so far (141 g/L and 0.49 g of SA per g of glucose).

Boosting the level of PEP has not been widely attempted in *S. cerevisiae*, which might be because the flux through E4P was identified as at least one order of magnitude less than that of the one through PEP (Suastegui et al., 2016a). Two unique genetic manipulations conducted in *S. cerevisiae* included the overexpression of ribose-5-phosphate ketol-isomerase encoded by *RKI1* and deletion of *RIC1* (Suastegui et al., 2017). The study suggested that Ric1 was a novel transcriptional repressor, and the corresponding binding sites were found in multiple promoters of the genes in the aromatic amino acid pathway; its deletion brought up the SA titer by 27%. *RKI1* overexpression was predicted *in silico* by OptForce, which suggested that the carbon flux in the PPP could be directed towards E4P without going back to glycolysis. This hypothesis was confirmed by an experiment in which the SA titer was increased from 0.8 g/L to 1.0 g/L (based on 2% glucose) and then 2.0 g/L (based on 4% glucose) to 2.4 g/L (based on 4% sucrose).

The metabolic flux analysis demonstrating that E4P is the rate-limiting substrate in *S. cerevisiae* (Suastegui et al., 2016a) motivated researchers to explore other non-model yeasts that have a high native xylose-utilization capacity. Xylose is efficiently assimilated through the PPP in CUG-clade yeasts, suggesting that these yeasts might have higher E4P availability compared to *S. cerevisiae*. With the classical manipulations such as overexpressing Tkt1, Aro4^{K220L}, and Aro1^{D900A}, the exemplary species *S. stipitis* produced SA at a titer of 3.11 g/L, which was 7-fold higher than that by *S. cerevisiae* under the same conditions (Gao et al., 2017). This represents the highest level of shikimate production in yeasts. Although this titer is still much lower than that of bacterial producers, these studies pave the way for engineering eukaryotic hosts that produce downstream products with nutraceutical and pharmaceutical value. The biosynthesis of some of these downstream compounds requires the participation of cytochrome P450 enzymes, which poses significant challenges for the bacterial production hosts.

Shikimate dehydrogenase encoded by *aroE* can also use 3-DHQ as the substrate. In the *E. coli* strain lacking *aroD* encoding 3DHQ dehydratase, quinic acid was accumulated at 49 g/L in fed-batch fermentation, with a yield of 0.21 g/g glucose (Ran et al., 2001). The same authors had also tested the heterologous quinic acid dehydrogenase (Qad) from *Klebsiella pneumoni*a, because this microorganism can naturally utilize quinic acid as its sole source of carbon. The heterologous expression of Qad in *E. coli* produced 4.8 g/L quinic acid in shake flasks (Draths et al., 1992). Quinic acid can be chemically converted to hydroquinone at high yields via oxidative decarboxylation and subsequent dehydration. Hydroquinone is used as an agent to develop photographs and an intermediate to synthesize some antioxidants and polymerization inhibitors, but in the past its synthesis heavily relied on the usage of carcinogenic benzene as the substrate. A recent study on the *de novo* biosynthesis of arbutin showed that hydroquinone could also be obtained by the hydrolysis of 4-hydroxylbenzoate (see section 4.2.2) (Shen et al., 2017).

4.1.2 *cis,cis*-Muconic acid

Muconic acid (MA) is an unsaturated dicarboxylic acid precursor of terephthalic and adipic acid, which are monomers of the plastics polyethylene terephthalate (PET) and nylon 66 (Lu et al., 2016; Niu et al., 2002). The traditional benzene-based chemical synthetic routes for adipic acid and terephthalic acid are detrimental to the environment, warranting the need to establish a sustainable green process. The microbial production route can be established via three steps, extending from DHS, as follows. DHS dehydratase (AroZ) and protocatechuate decarboxylase (AroY) both originate from *K. pneumonia*. DHS is converted to protocatechuate (PCA), and then into catechol, which undergoes an oxidative ring cleavage reaction catalyzed by *catA*-encoded catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus*, *Pseudomonas putida*, and *C.*

glutamicum. The expression of these three heterologous enzymes in combination with the aforementioned overexpression of three endogenous genes, *tktA*, *aroF*^{fbr}, and *aroB*, and the deletion of the downstream gene *aroE* in charge of converting DHS to SA, led to a high titer of 36.8 g/L in fed-batch fermentation (Niu et al., 2002).

The basic genetic manipulations necessary to create a *S. cerevisiae* MA producer are similar to those used to enable SA accumulation, with the mutation site in Aro1 switching from the SA kinase subunit to the SA dehydrogenase subunit enabled by Aro1^{D1409A}. The strain expressed AroZ, AroY, and Hqd2 from the ascomycete fungus *Podospora anserina*, the bacterium *K. pneumonia*, and the yeast *Candida albicans*, respectively (Curran et al., 2012), together with the native enzymes Tkl1, Rki1, Aro4^{K229L}, and Aro1^{D920A} in the *ric1*Δ *aro1*Δ deletion strain. This combination led to the production of 2.0 g/L PCA and 0.33 g/L MA based on 4% glucose shake-flask fermentation (Suastegui et al., 2016b; Suastegui et al., 2017). These findings led to the discussion of whether the deletion of *ZWF1*, encoding glucose-6-phosphate dehydrogenase would be beneficial for MA production. Researchers found that this deletion could force the carbon flux to enter through the non-oxidative branch of the PPP, thereby increasing E4P levels(Curran et al., 2012; Gold et al., 2015). It was found that while the *ZWF1* deletion indeed increased the DHS concentration before the incorporation of the three heterologous genes accounting for MA biosynthesis, it also decreased the overall production of PCA and MA by 31% in the final strain (Suastegui et al., 2017). Considering that the Zwf1-catalyzed reaction is the primary source of NADPH, the deletion of *ZWF1*, in combination with a higher metabolic burden imposed by expressing the three-gene heterologous MA pathway on a multi-copy plasmid, reduced the fitness. This is because major rearrangements occur in the cell to find new NADPH sources. Similarly, the interruption of NADPH cofactor regeneration also reduced SA

production because the step converting DHS to SA requires NADPH. The authors clarified that the deletion of *ZWF1* should not be used as a strategy to produce any compounds downstream of the DHS node unless NADPH can be efficiently replenished. For the products directly derived from DHS, such as MA, whether this step is beneficial depends on the tradeoff between the increase in E4P level and the negative impact on cell growth. Leavitt et al. created a hybrid promoter induced by aromatic amino acids and placed a *KanNeo* gene conferring a low level of resistance to geneticin (Leavitt et al., 2017). Ethyl methane sulfonate mutagenesis and adaptive laboratory evolution were implemented to create mutant strains with an increased flux into the aromatic amino acid biosynthesis. The flux was re-directed to create a strain producing MA at a titer of 2.1 g/L in fed-batch fermentation.

In all these engineered yeast strains, the PCA decarboxylase AroY seems to be the bottleneck step, causing significant accumulation of PCA and thus, low MA titers. Some studies have suggested that this is due to the O_2 sensitivity of this enzyme (Suastegui et al., 2016b; Weber et al., 2012), while others believe that its functionality requires the prenylated form of flavin mononucleotide as the cofactor, which is insufficient or even missing in some strains, such as the CEN.PK family of *S. cerevisiae* (Weber et al., 2017). Supplying the strain with an overexpressed flavin mononucleotide prenyltransferase (Pad1) addresses this bottleneck. Furthermore, with the strategies involving the overexpression of *E. coli* AroB and AroD to enhance the conversion up to DHS, and in-pulse triggering of the conversion of DHS to the flow downstream by destabilizing endogenous Aro1, the final strain produced 1.2 g/L MA without an additional supply of the aromatic amino acids and 5.1 g/L MA when amino acids were supplemented (Pyne et al., 2018).

The benefit of co-expressing AroY-associated proteins to overcome the accumulation of PCA was also observed in other hosts such as *P. putida* (Johnson et al., 2016) and *C. glutamicum* (Lee et al., 2018). Because both species have the native function of utilizing aromatics (e.g., *p*-coumarate, ferulate, and benzoate) via β-ketoadipate pathway, the genes encoding PCA dioxygenase (PcaHG), and the enzymes required for further metabolism of MA to muconolactone (CatB and CatC, together with their transcriptional regulator CatR) have to be deleted. AroY together with its associated protein EcdB from *Enterobacter cloacae* and dehydroshikimate dehydratase AsbF from *Bacillus cereus* were heterologously expressed in *P. putida* along with the overexpressed native CatA (Johnson et al., 2016). The resulting strain was able to directly synthesize MA from the lignin monomer *p*-coumarate with a titer of more than 15 g/L, and from glucose with a titer of 4.9 g/L in a fed-batch process. This additional co-expression of EcdB enabled the host to reach production at least 3-fold higher than the one that could not offer the prenylated flavin cofactor. For *C. glutamicum*, the fermentation was scaled up to 7-L and 50-L fed-batch processes, resulting in 38 g/L and 54 g/L of MA based on glucose (Lee et al., 2018).

A useful co-culture strategy was developed to split MA biosynthesis between two *E. coli* strains (Zhang et al., 2015b). The genes encoding the PTS glucose transport system (encoded by *ptsH*, *ptsI*, and *crr*) and shikimate dehydrogenase (encoded by *aroE* and *ydiB*) were deleted from the first strain, enabling it to utilize xylose and accumulate DHS. The xylose isomerase (encoded by *xylA*) in the second strain was then disrupted, but the DHS transporter gene (encoded by *shiA*), together with the three-step MA pathway (encoded by *aroZ*, *aroY*, and *catA*), was overexpressed. The resulting co-culture produced 4.7 g/L MA in the fed-batch bioreactor, with a yield of 0.35 g/g from a mixture of glucose and xylose. Splitting a long biosynthetic pathway into different

strains not only relieves the burden caused by extensive overexpression and deletion to a monoculture but also allows easy fine-tuning of the ratio between the modules accommodated by the distinct partners for the optimal outcome.

4.1.3 2-Pyrone-4,6-dicarboxylic acid and vanillin

2-Pyrone-4,6-dicarboxylic acid (PDC) is a dicarboxylic acid that can also be derived from PCA. Owing to its structural similarity with terephthalic acid, the monomeric building block of PET, PDC has been tested for polymerization with other diols or hydroxyacids to form polymers with interesting features. Luo et al. described a biosynthetic pathway in which dehydroshikimate dehydratase (encoded by *asbF* from *Bacillus thuringiensis*), PCA 4,5-dioxygenase (encoded by *pmdAB* from *Comamonas testosteroni*), and CHMS dehydrogenase (encoded by *pmdC* from *C. testosteroni*) were cloned in an *E. coli* strain carrying the *aroE* deletion for DHS accumulation (Luo et al., 2018). The endogenous genes *ppsA* and *shiA* (encoding the DHS importer) were overexpressed by the strong promoter pTrc to increase the availability of PEP and trap DHS intracellularly. The *sthA*-encoded pyridine nucleotide transhydrogenase in *E. coli* was also targeted to enhance the level of $NADP^+$ *via* two strategies: exchanging the promoter of *sthA* with pTrc or introducing three silent point mutations to the beginning of the coding region of *sthA* to deregulate the endogenous translational control. The combination of all these strategies allowed the production of 16.72 g/L PDC based on glucose fed-batch fermentation, with a yield and productivity of 0.201 g/g and 0.172 g/L/h, respectively.

In *S. cerevisiae* and the fission yeast *Schizosaccharomyces pombe*, DHS has been further extended with four-step conversions to synthesize vanillin β-D-glucoside (VG). Vanillin is one of the most important flavoring agents in the world (N., 2014). It is a natural product present in the

seedpods of *Vanilla planifolia*, but its current production mostly relies on a biochemical conversion involving lignin deconstruction or synthesis based on petrochemicals (Priefert et al., 2001). The *de novo* biosynthesis of vanillin involves the heterologous expression of *Podospora pauciseta* 3-dehydroshikimate dehydratase (3DSD), *Nocardia* aromatic carboxylic acid reductase (ACAR), *Homo sapiens O*-methyltransferase (OMT), and an additional phosphopantetheinyl transferase (PPTase) cloned from *C. glutamicum*. PPTase is only required by *S. cerevisiae* and not by *S. pombe*, because the latter possesses endogenous phosphopantetheinylation activity to help with the reduction step. The initial titers were fairly low in the two yeasts (45 mg/L and 65 mg/L), and a decent amount of product was reduced to vanillyl alcohol in *S. cerevisiae* unless the alcohol dehydrogenase *ADH6* was disrupted (Hansen et al., 2009). The low production was partially caused by the toxicity of vanillin and expression of an *Arabidopsis thaliana* family 1 uridine diphosphate (UDP) glycosyltransferase (UGT), which converted vanillin into the nontoxic VG with a production increase to 100 mg/L. It was also found that the knockout of *PDC1*, encoding pyruvate decarboxylase, could channel more flux to flow towards the shikimate pathway instead of the ethanol end using a continuous chemostat with a very low dilution rate, and the final *S. cerevisiae* strain produced 500 mg/L VG with a yield of 32 mg/g glucose (Brochado et al., 2010).

4.2 The chorismate node

Moving downstream beyond DHS and SA, chorismate becomes the next hub, from which various benzoic acids can be synthesized. Many health-related products, such as folate (vitamin B_9), CoQ_9, and aspirin, need benzoic acid as the precursor of their synthesis (Noda et al., 2016; Yao et al., 2018).

4.2.1 Salicylate, 4-aminobenzoate, 3-aminobenzoate, 2-aminobenzoate, 3-hydroxybenzoate, 4-hydroxybenzoate, and muconic acid

The chorismate lyase family of enzymes cleaves the C-C bond of chorismate and its close derivatives, releasing pyruvate and benzoic acid attached to either a hydroxyl or an amino group. Obtaining pyruvate as a product in these reactions motivated Noda et al. to design an innovative "pulling" strategy to boost the flux to synthesize various benzoic acids (Noda et al., 2016). They replaced the aforementioned endogenous PTS glucose transport system with galactose permease (GalP) and glucokinase (Glk) from *E. coli*, thereby disrupting the conversion from PEP to pyruvate during glucose transportation. In addition, the conversion of PEP to pyruvate during glycolysis was also eliminated by knocking out pyruvate kinase encoded by *pykF* and *pykA*. This increased the available PEP entering the shikimate pathway, and the cells were forced to replenish pyruvate *via* the chorismate lyase-catalyzed reaction. This strategy was first demonstrated in the biosynthesis of salicylate, which isomerizes chorismate to isochorismate (IsoCA) (IsoCA synthase encoded by *menF* from *E. coli*) before C-C bond cleavage (IsoCA-pyruvate lyase encoded by *pchB* from *Pseudomonas aeruginosa*). Without using the standard strategies to increase PEP and E4P levels (e.g., the overexpression of *tktA* and *ppsA*), the production of salicylate in *E. coli* already reached 11.5 g/L in the batch culture after 48 h. The strategy was subsequently generalized to produce seven different chorismate derivatives, including MA, 4-hydroxybenzoate (4-HBA; also known as *p*-HBA), 3-hydroxybenzoate (3-HBA), 2-aminobenzoate (2-ABA), 4-aminobenzoate (4ABA; also known as *p*-ABA), tyrosine, and phenol, at 1–3 g/L in test tube-scale culture.

Salicylate is a precursor in the synthesis of aspirin and the anti-HIV agent lamivudine and it can be combined with endogenously produced UDP-α-D-glucose, catalyzed by salicylate glucosyltransferase (UGT74F1) from *A. thaliana*, to synthesize the plant-derived salicylate 2-*O*-β-D-glucoside (SG), which has anti-inflammatory activity, as well as other glycosylated variants; once the heterologous deoxysugar pathways are introduced (Qi et al., 2018). In this study, the *irp9*-encoded bifunctional salicylate synthase from *Yersinia enterocolitica* was expressed to complete the two-step conversion from chorismate to salicylate.

Note that the enzymes that catalyze the synthesis of aminobenzoates are slightly different from those involved in the synthesis of hydroxybenzoates, because an amino group needs to be transferred from glutamine prior to the lyase-catalyzed reaction. The *de novo* biosynthesis of 4ABA was established in *S. cerevisiae* by overexpressing aminodeoxychorismate synthase (Abz1) and aminodeoxychorismate lyase (Abz2) from yeast (with a titer of 215 mg/L) using glycerol and ethanol as the feedstock in fed-batch fermentation (Averesch et al., 2016), and in *E. coli* (with a titer of 4.8 g/L) by overexpressing the bacterial analogous enzymes encoded by *pabABC* (Koma et al., 2014).

One-step extension from chorismate catalyzed by *E. coli* chorismate lyase (encoded by *ubiC*) leads to the synthesis of 4-HBA, which is currently used to prepare paraben preservatives, cosmetics, and pharmaceutical products. Using similar strategies as discussed previously in SA and MA production, including the overexpression of transketolase, the feedback-insensitive DAHP synthase mutant, and individual enzymes to enhance the conversion from DAHP to chorismate, 4-HBA production of 12 g/L was established in *E. coli* by fed-batch fermentation, with a yield of 0.19 g/g glucose, and a productivity of 0.17 g/L/h (Barker and Frost, 2001). In *S. cerevisiae* and *P. putida*, the downstream Tyr/Phe branch and the Trp branch were both disrupted,

and because *P. putida* has the endogenous PCA degradation pathway, additional knockout of *pobA* (encoding 4-HBA hydroxylase) was performed. Using fed-batch fermentation, titers of 2.9 g/L and 1.73 g/L were achieved in the two hosts, but the yield and productivity in *P. putida* were 5.8-fold and 1.7-fold higher than those obtained in *S. cerevisiae* (Averesch et al., 2017; Yu et al., 2016), respectively. This *P. putida* strain also included a deletion of a glucose catabolism repressor encoded by *hexR*, the expression of which increases the supply of E4P and NADPH. The aforementioned *E. coli* co-culture used to produce MA has also been extended to produce 4-HBA at 2.3 g/L and 3-aminobenzoic acid (3-ABA; also known as m-ABA) at a low titer of 48 mg/L, presumably due to the low activity of 3-ABA synthase (PctV) from *Streptomyces pactum* (Zhang et al., 2015b).

Section 2.3 summarizes the higher order of enzyme association occurring in the *trp* operon. In *E. coli*, TrpE and TrpD form a heterotetramer to mediate the synthesis of anthranilate and PRPP. The N-terminal of TrpD (designated as TrpG in some literature) with TrpE enables amino transfer from glutamine to complete anthranilate synthesis. Balderas-Hernández et al. identified a TrpD mutant with an early stop codon in the C-terminal anthranilate phosphoribosyl transferase domain (Balderas-Hernandez et al., 2009). This mutation led to the accumulation of anthranilic acid (2-ABA) in *E. coli* at 14 g/L in fed-batch fermentation.

MA can be synthesized via several pathways, which share the last conversion step of catechol to MA but differ in the reactions prior to catechol. For example, MA can be derived from at least four downstream branches from the chorismate node, including the pHBA route (*ubiC* and *pobA*) (Sengupta et al., 2015), the 2,3-dihydroxybenzoic acid (2,3-DHBA) route (*entCBA* and *bdc*) (Sun et al., 2014; Wang and Zheng, 2015), the salicylate route (*menF*, *pchB*, and *nahG*) (Lin et al., 2014a), and the anthranilate route (*trpE*, *trpG*, and *antABC*) (Sun et al.,

2013). The production levels of MA in these engineered *E. coli* strains ranged from 0.17 to 1.45 g/L in shake-flask fermentation using either glucose or glucose plus glycerol as the carbon source. Fujiwara et al. compared the *E. coli* strains producing MA through the DHS, pHBA, and salicylate routes and found that optimum performance was obtained by fusing chorismate synthase (AroC) and isochorismate synthase (MenF) to facilitate metabolite channeling (Fujiwara et al., 2018). The strain produced MA at 4.45 g/L, with a yield of 0.207 g/g glucose and a productivity of 2.77 g/L/day in a jar fermenter.

4.2.2 Arbutin, phenazine-1-carboxamide, gastrodin, gallic acid, and pyrogallic acid

Arbutin is a plant-derived glycosylated hydroquinone and has been widely used in cosmetic, healthcare, and medical industries because of its ability to block melanin synthesis for skin lightening, as well as its other anti-microbial, anti-inflammatory, and anti-oxidant activities (Shen et al., 2017; Yao et al., 2018). Extended from 4-HBA, the *de novo* biosynthesis of arbutin can be achieved by the expression of 4-HBA 1-hydroxylase (Mnx1) from *Candida parapsilosis* and glucosyltransferase (AS) from *Rauvolfia serpentina* at 4.19 g/L in *E. coli* strains engineered to accumulate 4-HBA (Shen et al., 2017). A similar idea has been extended to *Pseudomonas chlororaphis* (Wang et al., 2018c). The final strain carrying the genome-integrated genes encoding Mnx1, AS, XanB2 (chorismate lyase from *Xanthomonas oryzae pv. oryzae*), and PhzC (the *P. chlororaphis* version of DAHP synthase), and with deletions of *pykA* (pyruvate kinase converting PEP to pyruvate), *phzE* (converting chorismate for phenazine biosynthesis), and three negative regulatory genes (*lon*, *rpeA*, and *rsmE*) produced arbutin at 6.79 g/L and 0.094 g/L/h with mixed fed-batch fermentation of glucose and 4-HBA. Previously, the disruption of another set of pathway-specific transcriptional repressors encoded by *lon*, *parS*, and *psrA* enabled *P.*

chlororaphis to enhance its production of phenazine-1-carboxamide (PCN) at 4.1 g/L, with a yield of 0.23 g/g glycerol (Yao et al., 2018). PCN is naturally produced by *Pseudomonas* strains and displays broad-spectrum antibiotic properties.

Gastrodin is another plant-derived phenolic glycoside naturally synthesized by the herb *Gastrodia elata*, which has been used in East Asian medicine to treat diseases of the central nervous system and cardiovascular symptoms. Bai et al. successfully created a microbial biosynthetic route composed of carboxylic acid reductase (CAR) from *Nocardia iowensis*, phosphopantetheinyl transferase (Sfp) from *B. subtilis* to convert apoCAR to its holo form, and endogenous alcohol dehydrogenase (ADH) and glycosyltransferase (UGT73B6) from *Rhodiola sachalinensis* (Bai et al., 2016). The strain was first optimized to synthesize 4-hydroxybenzyl alcohol via the 4-HBA branch, and then, a more potent glycosyltransferase mutant was identified through directed evolution. The final strain produced gastrodin at a titer of 545 mg/L at 48 h in shake-flask glucose fermentation.

The previously mentioned *pobA* encoding 4-HBA hydroxylase in *Pseudomonas* spp. has a very high activity toward 4-HBA but low activity toward the product PCA. Gallic acid is a 3,4,5-trihydroxybenzoic acid, widely found in fruits, nuts, and tea, and it has strong anti-oxidant properties. To continue the hydroxylation of PCA to synthesize gallic acid, a structure-based rational design strategy was applied to *P. aeruginosa* PobA. The resulting mutant PobA containing mutations (Y385F and T294A) enabled the engineered *E. coli* strain to produce 1.27 g/L gallic acid in shake flasks (Chen et al., 2017). Gallic acid can be chemically converted to pyrogallol (other names: pyrogallic acid or 1,2,3-trihydroxybenzene) through thermal decarboxylation under high pressure and temperatures. Due to the highly active nature of each hydroxyl group on the benzene ring, pyrogallol has been used as the precursor to synthesize

many bioactive compounds such as bendiocarb (insecticide), trimethoprim (antibiotic), and gallamine triethiodide (muscle relaxant) (Wang et al., 2018b). To create an artificial biosynthetic pathway, the *entCBA* operon (encoding IsoCA synthase, isochorismatase, and 2,3-dihydro-2,3-dihydroxybenzoic acid dehydrogenase) was overexpressed to convert chorismate to 2,3-dihydroxybenzoic acid (2,3-DHBA). Then, five candidate oxygenases and hydroxylases were screened for the ability to transform the C1-carboxyl group of 2,3-DHBA to a hydroxyl group. Salicylate 1-monooxygase, encoded by *nahG* from *P. putida*, displayed a high *in vitro* conversion efficiency. Its heterologous expression, together with other optimization strategies, including the enhancement of the flux in the SA pathway, gene expression balancing, and alleviation of product autoxidation, led to the production of pyrogallic acid at ~1 g/L in *E. coli* in shake-flask culture (Wang et al., 2018b).

5. Products derived from the L-Trp branch

5.1 5-Hydroxytryptophan, serotonin, and melatonin

5-Hydroxytryptophan (5-HTP) is a natural non-canonical amino acid (**Figure 3 and Table 3**). It has been used for over 30 years as a precursor of the neurotransmitter serotonin to treat serotonin disorders that cause depression and a variety of other medical conditions such as chronic headaches, insomnia, and binge eating disorder and obesity (Lin et al., 2014b; Mora-Villalobos and Zeng, 2017; Mora-Villalobos and Zeng, 2018; Sun et al., 2015). Owing to its structural similarity to L-Trp, orally administered 5-HTP can easily pass through the blood–brain barrier and then be converted into serotonin in the central nervous system. The commercial production of 5-HTP currently relies on its extraction from the seeds of the African plant

Griffonia simplicifolia, whose season- and region-dependent supply strongly influences cost-effective production.

In both animals and plants, the biosynthesis of serotonin takes place in two steps but differs in the order. In animals, L-Trp is hydroxylated to 5-HTP by tryptophan 5-hydroxylase, and tryptophan decarboxylase (TDC) then converts 5-HTP into serotonin, whereas in plants, the decarboxylation occurs first, producing tryptamine as the intermediate, followed by the hydroxylation step catalyzed by tryptamine 5-hydroxylase. In the literature, both tryptophan 5-hydroxylase and tryptamine 5-hydroxylase are labeled as T5H, but the latter is a cytochrome P450 enzyme that is expressed in the endoplasmic reticulum membrane of plants (Fujiwara et al., 2018).

Tryptophan 5-hydroxylase in humans and animals uses 5,6,7,8-tetrahydrobiopterin (BH4) and Fe^{2+} as cofactors and O_2 as a co-substrate. During Trp hydroxylation, BH4 is oxidized to pterin-4α-carbinolamine (BH3OH) and is regenerated through pterin-4α-carbinolamine dehydratase (PCD) and dihydropteridine reductase (DHPR) (Hara and Kino, 2013). In humans, BH4 is converted from GTP via three steps, involving GTP cyclohydrolase I (GCHI), 6-pyruvate-tetrahydropterin synthase (PTPS), and sepiapterin reductase (SPR) (Yamamoto et al., 2003) (**Figure 3b**). *E. coli* does not naturally synthesize BH4, but it can use tetrahydromonapterin (MH4) as an alternative cofactor for tryptophan 5-hydroxylase (**Figure 3c**). However, the system required for MH4 regeneration is not fully available in *E. coli* and has to be heterologously incorporated.

Tryptophan 5-hydroxylase belongs to the family of pterin-dependent aromatic amino acid hydroxylases (AAAHs), which also includes phenylalanine-4-hydroxylases (P4Hs). The AAAHs from animals usually consist of the N-terminal regulatory domain, the central catalytic domain,

and the C-terminal domain involved in BH4 utilization and tetramer formation (Fitzpatrick, 2003). Prokaryotic P4Hs have only one catalytic domain corresponding to the central domain in animal AAAHs but show little activity to act on L-Trp. Because MH4 replaces BH4 as the major source of pterin in *E. coli*, P4Hs use MH4 instead of BH4 as their native coenzyme. Lin et al. implemented a protein-engineering strategy to create a mutant $P4H^{W179F}$ from *Xanthomonas campestris* (encoded by *phhA*), which reduced the substrate preference for L-Phe over L-Trp from 33.5 to 1.5 (Lin et al., 2014b). The final strain expressed $P4H^{W179F}$ and the MH4 regeneration pathway, including PCD (encoded by *phhB*) from *P. aeruginosa* and dihydromonapterin reductase (DHMR; encoded by *folM*) from *E. coli* (Figure 3c), in addition to the $trpE^{S40F}DCBA$ operon. The host was a previously well-developed L-Phe overproducer carrying *pheA* and *tyrA* deletions that prevented the synthesis of L-Phe and L-Tyr (Huang et al., 2013; Lin et al., 2014a), and in this study, its *tnaA* gene, encoding tryptophanase (TnaA) (catalyzing the degradation of tryptophan and 5-HTP to the corresponding indole), was knocked out. The engineered strain was fermented at 30 °C to prevent oxidative degradation and produced 5-HTP at 153 mg/L at the expense of 9.7 g/L glucose in shake flasks and at 1.1–1.2 g/L with the supplementation of 2 g/L L-Trp (Lin et al., 2014b).

Similar protein engineering based on *in silico* structure models was applied to another bacterial AAAH from *Cupriavidus taiwanensis*. The mutant $AAAH^{W192F}$, with enhanced substrate preference towards L-Trp, enabled an *E. coli* strain harboring the human BH4 (**Figure 3d**) regeneration pathway and tryptophanase disruption to produce 5-HTP at 0.55 g/L in 24 h with the supplementation of 1 g/L Trp. Here, the BH4 regeneration pathway, including PCD and DHPR, was used to regenerate the endogenous MH4 (Mora-Villalobos and Zeng, 2017). The same research group later developed a two-stage fermentation process to synthesize serotonin

through a decarboxylation reaction of 5-HTP catalyzed by the TDC from *Catharanthus roseus*. As mentioned earlier, in plants, TDC-catalyzed decarboxylation occurs prior to the hydroxylation step. The promiscuity of TDC caused it to recognize L-Trp as a substrate and decarboxylate it to tryptamine. To address this issue, the team split the production in two strains: the first strain synthesized 5-HTP at a titer of 962 mg/L, and the cell-free medium was fed to the second strain expressing TDC, leading to the production of 154.3 mg/L serotonin through fed-batch fermentation. The first strain was created based on a previous L-Trp overproducer, S028. Note that a different mutant, CtAAAH$^{F197L,\ E219C}$, was designed based on the *in silico* structure model and was verified to have a much larger k_{cat}/K_m than that of the initial variant CtAAAH^{W192F} towards L-Trp (Mora-Villalobos and Zeng, 2018). One study expressed both TDC and tryptamine-5-hydroxylase cloned from rice (*Oryza sativa*). Its expression in *E. coli* was enabled with the 37 amino acids at the N-terminal of tryptamine-5-hydroxylase replaced with the glutathione S-transferase tag to enable soluble expression. The resulting *E. coli* strain produced serotonin at 24 mg/L and 20 °C with the supplementation of L-Trp at 0.2 g/L in test tubes. Surprisingly, unlike the expression of other L-Trp hydrolases, neither cofactor biosynthesis nor regeneration pathways were needed for tryptamine-5-hydroxylase expression (Park et al., 2011).

Instead of using the native MH4 as the cofactor, Wang et al. recently reconstituted the BH4 synthesis and regeneration pathways in an *E. coli* strain expressing a truncated version of human TPH2 (NΔ145, CΔ24), which had a significantly better soluble expression than that of the wild type (Wang et al., 2018a). The GCHI used in the BH4 synthesis was encoded by *mtrA* from *B. subtilis*, which improved BH4 biosynthesis, and the four other genes were human-sourced (**Figure 3b**). The entire biosynthetic pathway was sectioned into L-Trp, hydroxylation, and BH4 modules. By modulating the plasmid copy number and promoter strength, the relative expression

levels among different modules were fine-tuned. The production of 5-HTP was optimized to 1.3 g/L in shake flasks and 5.1 g/L in a fed-batch fermenter using glycerol as the carbon source, representing the highest reported levels for the microbial *de novo* production of 5-HTP.

Zhang et al. compared the performance of L-Trp hydrolase from prokaryotic and eukaryotic sources, as well as the MH4 and BH4 systems, when using *S. cerevisiae* as a host (Zhang et al., 2016). The prokaryotic system involves an evolved P4H variant from *X. campestris*, *folM* (encoding DHMR) from *E. coli*, and *phhB* (encoding PCD) from *P. aeruginosa* for MH4 regeneration (**Figure 3c**), as discussed previously; the eukaryotic system involves tryptophan 5-hydroxylase from *Oryctolagus cuniculus* or tryptophan 5-hydroxylase from *Schistosoma mansoni*, together with a BH4 biosynthesis pathway, including PTPS and SPR from *Rattus norvegicus*, DHPR and GCH from *E. coli*, and PCBD from *P. aeruginosa* (**Figure 3b**). The results showed that the BH4 pathway enabled a 17-fold higher production of 5-HTP than its MH4 counterpart. It was also found that the overexpression of the native dihydrofolate reductase DFR1 to catalyze tetrahydrofolate synthesis could enhance the performance of the prokaryotic pathway by influencing MH4 generation.

To overcome the challenge of expressing L-Trp hydrolases in bacteria, Sun et al. devised a novel 5-HTP biosynthesis route via anthranilate, which omitted the requirement of the cofactors BH4 and MH4 (Sun et al., 2015). The rationale was that the enzymes encoded by *trpDCBA* could possibly accept 5-hydroxyanthranilate (5-HAA) as a substrate, considering that it only differs from the native substrate anthranilate by the 5-hydroxyl group. However, there is no reported anthranilate hydroxylase. Thus, they selected two salicylate 5-hydroxylases (S5Hs) cloned from *Ralstonia sp*. and *P. aeruginosa* characterized to hydroxylate salicylate into gentisate. As anthranilate is actually 2-ABA, whereas salicylate is 2-HBA, it was hypothesized

that S5H would also accept anthranilate as a substrate; subsequently, the researchers found a promising putative S5H gene cluster (designated as *salABCD*) from *Ralstonia eutropha*. These enzyme complexes require NAD(P)H instead of BH4 as the cofactor. Although the activities of TrpDCBA and SalABCD towards the nonnative substrates were proven separately in this study, the combined pathway did not accumulate 5-HTP at the beginning. It was suggested that the heterotetrameric complex formed by the native TrpE and TrpD may hinder the access of SalABCD to the substrate anthranilate, or that TrpDCBA preferred anthranilate over 5-HAA. A two-step strategy was then implemented to address this issue. One strain harboring $TrpE^{fbr}G$ and SalABCD produced 5-HAA, and after 24 h, the other strain harboring TrpDCBA converted 5-HAA in the collected supernatant of the first culture to 5-HTP, with the final titer reaching 98 mg/L at 12 h.

Melatonin, a natural hormone regulating the circadian rhythm, is produced by the pineal gland in the brain through acetylation and methylation steps. It has been sold as an over-the-counter dietary supplement, and studies have suggested its role in treating cancer and other neurological diseases (Hardeland, 2008; Leonardo-Mendonca et al., 2015; Xu et al., 2015). Unlike serotonin, melatonin can be transported across the brain–blood barrier and be absorbed by the central nervous system. At present, melatonin is primarily produced through chemical synthesis, which requires toxic substrates or catalysts.

Heterologous expression of *N*-acetyltransferase (AANAT) and *S*-adenosyl- L-methionine (SAM)-dependent *N*-acetylserotonin *O*-methyltransferase (ASMT) converts serotonin to melatonin. Germann et al. constituted a production pathway composed of *S. mansoni* TPH, *H. sapiens* DDC (5-HTP decarboxylase), *Bos taurus* AANAT, and *H. sapiens* ASMT in *S. cerevisiae* (Germann et al., 2016). The BH4 synthesis pathway was enabled by the expression of

PTS/SPR from *R. norvegicus* and basal-level expression of GTP cyclohydrolase I (GCH1 and FOL2); the BH4 regeneration pathway combined DHPR from *R. norvegicus* and PCBD from *P. aeruginosa*. All the genes were integrated into the genome except ASMT, which was beneficially placed on a high-copy plasmid. Supplementing 500 mg/L methionine, the precursor of SAM, did not increase production, indicating that the native SAM level was sufficient, but overexpression of the endogenous acetaldehyde dehydrogenase ALD6 increased production by 10%, presumably by improving the supply of acetyl-CoA. Melatonin production of 14.5 mg/L was obtained by using the synthetic feed-in-time (FIT) medium or mineral medium supplemented with glucose. The FIT medium mimics small-scale fed-batch fermentation because glucose was gradually released by the exogenously added glucoamylase from a solubilized glucose polymer.

5.2 Violacein and deoxyviolacein

Violacein and deoxyviolacein are naturally produced purple indolocarbazoles that are of great interest because of their antimicrobial, antileishmanial, antiviral, and antitumor properties. Several gram-negative bacteria produce them (mainly violacein), including *Alteromonas luteoviolacea*, *Chromobacterium violaceum*, *Pseudoalteromonas luteoviolacea*, and *Janthinobacterium lividum* (Rodrigues et al., 2014; Rodrigues et al., 2013; Sun et al., 2016a). Similar to many other secondary metabolites, their production in microorganisms is possibly triggered by self-defense against environmental danger. Violacein has also been used as a colorant for textile dyeing.

The accumulation of these pigments in some bacteria is enabled by the *vioABCDE* operon. L-Trp oxidase VioA first converts L-Trp to indole-3-pyruvic acid imine, and two of these

molecules are combined to form protodeoxyviolaceinic acid, catalyzed by iminophenyl-pyruvate dimer synthase (VioB) and violacein biosynthesis enzyme (VioE). Protodeoxyviolaceinic acid can be processed either directly by violacein synthase (VioC) or sequentially by tryptophan hydroxylase (VioD) and then by VioC, yielding deoxyviolacein or violacein as the final product. Rodrigues et al. placed the *vioABCE* cluster from *C. violaceum* under the control of an inducible *araC* promoter in *E. coli* (Rodrigues et al., 2013). To increase the availability of L-Trp, the feedback-insensitive mutants $TrpE^{N168D,\ C237R,\ M293T,\ and\ A478T}$ and $SerA^{T372D}$ (3-phosphoglycerate dehydrogenase) were created, and the latter increased the endogenous level of L-serine (L-Ser), which served as the co-substrate of indole in the TrpB-catalyzed condensation, the last step of the L-Trp biosynthesis pathway. Other overexpression targets included endogenous TktA (to increase the E4P level), AroFBL (to enhance the flux in the upstream shikimate pathway), and TrpDCBA (to increase the flux in the Trp biosynthesis branch). In addition, the genes *trpR*, *trpL*, *tnaA*, and *sdaA* were interrupted. TrpR and TrpL are involved in downregulating the genes in the L-Trp branch, whereas TnaA (tryptophanase) and SdaA (L-Ser deaminase) are involved in L-Trp and L-Ser degradation, respectively. In this study, 10 g/L arabinose was used as both the carbon source and the inducer. The combination of all these strategies led to an accumulation of deoxyviolacein at 324 mg/L in shake flasks. The addition of VioD from *J. lividum* led to the complete conversion of deoxyviolacein to violacein, reaching 710 mg/L in a 300 h fed-batch process based on arabinose. An immediate follow-up study indicated that the disruption of arabinose catabolism and switching to glycerol as a carbon source could further increase the production of deoxyviolacein to 1.6 g/L in fed-batch culture (Rodrigues et al., 2014).

The five genes in the *vioABCDE* operon actually use internal ribosome-binding sites, and the stop codon of the previous gene overlaps with the start codon of the next. Due to the high

performance of *C. glutamicum* as a host to produce aromatic amino acids, Sun et al. overexpressed the *vioABCDE* operon in an established industrial L-Trp hyperproducer, ATCC 21850, and inserted a *C. glutamicum* ribosome-binding site in front of each *vio* gene (Sun et al., 2016a). The resulting strain produced violacein at a titer of 5436 mg/L, with a productivity of 47 mg/L/h, representing the highest levels of production from both native and heterologous hosts. Some studies also focused on optimizing the fermentation processes of the native producers, e.g., *C. violaceum* (1.1 g/L), *Duganella sp.* (1.62 g/L), and *J. lividum* (3.5 g/L), and the heterologous hosts, e.g., *Citrobacter freundii* (4.13 g/L) (Duran et al., 2007; Lu et al., 2009; Wang et al., 2009; Yang et al., 2011).

5.3 Indole, hydroxyindole, indigo, and indirubin

Indole alkaloids (also called indigoids) such as indirubin and indigo are plant-derived secondary metabolites, many of which possess important pharmaceutical properties and have been employed as medicines and dyes since ancient times (Han et al., 2013). Indirubin, also known as couraupitine B, a red isomer of indigo, is the active ingredient in Danggui Longhui Wan, a traditional Chinese medicine for the treatment of chronic granulocytic leukemia (Berry et al., 2002). It also effectively treats diseases such as Alzheimer's disease (Hoessel et al., 1999). Indirubin and indigo compounds are currently extracted from plant materials, but due to their small quantities, downstream processing is very challenging and laborious.

Indole is oxidized by the heterologous naphthalene dioxygenase (NDO) from *Comamonas* sp. to form 3-hydroxyindole (interconvertible with 3-oxindole), 2-hydroxyindole (interconvertible with 2-oxindole), or isatin. The spontaneous dimerization of 3-hydroxyindole forms indigo in *E. coli*. When 0.5 g/L of 2-oxindole was added to L-Trp-containing media, it was dimerized with 3-

hydroxyindole, leading to the maximal production of indirubin at 58 mg/L. Further increases in 2-oxindol had no effect on the titer, but the proportion of indirubin in indigoids could be optimized to 48% when 1.5 g/L 2-oxindole was supplemented (Zhang et al., 2014). A compelling study found that cysteine supplementation at 0.36 g/L could increase the proportion of indirubin, presumably by influencing the regioselectivity of the flavin-containing monooxygenase (FMO) from *Methylophaga aminisulfidivorans*, thereby enhancing the synthesis of 2-hydroxyindole over 3-hydroxyindole as the major product in L-Trp-containing media. The resulting *E. coli* strain was capable of producing 223.6 mg/L indirubin and 6.8 mg/L indigo, as opposed to 5 mg/L indirubin and 920 mg/L indigo when no cysteine was supplemented (Han et al., 2011; Han et al., 2013). One study suggested that the L-Trp symporter encoded by *tnaB* (located 90 bp downstream of *tnaA*) was specifically responsible for transporting the exogenous L-Trp, which is converted to indole by TnaA (Li and Young, 2013).

To avoid supplementation of expensive precursors, the *de novo* biosynthesis of indigo and indirubin was established in *E. coli*; FMO from *M. aminisulfidivorans* was introduced into an *E. coli* strain overexpressing its native TnaA (catalyzing the β-elimination of L-Trp to indole), AroL (increasing the flux in the shikimate pathway), TktA (increasing E4P availability), and the two feedback-insensitive mutants $AroG^{D146N}$ and $TrpE^{S40F}$. Simultaneously, the pathway-specific negative regulator TrpR and the pyruvate kinases PykF and PykA (increasing PEP availability) were all interrupted. Fed-batch fermentation led to the production of indirubin (0.056 g/L) and indigo (0.64 g/L) using glucose as the carbon source (Du et al., 2018). In an earlier study, Berry et al. expressed *P. putida* NDO and the native $TrpE^{fbr}$DCB/26A in *E. coli* with a mutation introduced in the *trpB* gene to inactivate the conversion of indole and L-Ser to L-Trp (Berry et al., 2002). Of note, *in vitro* analysis suggested that AroG was inhibited by indoxyl or its downstream

intermediates. To address this issue, the AroG^{fbr} mutant was cloned in two co-transformed plasmids to increase the abundance of AroG^{fbr} and dilute the negative inhibitory impact. It was also found that preincubation of the enzyme with PEP provided partial protection against inactivation. The overexpression of TktA and disruption of *pykA* and *pykF* were correspondingly conducted to increase PEP availability. Combining all these strategies enabled a high production of 18 g/L in fed-batch fermentation with L-Trp fed at an increasing rate from 0.2 to 0.5 g/min 14 h after the start of glucose feed.

6. Products derived from the L-Phe branch

6.1 Phenylpyruvate node

6.1.1 Phenyllactic acid

E. coli has endogenous phenylpyruvate reductase (PPR) activity, which converts phenylpyruvate to phenyllactic acid (**Figure 4 and Table 4**); this is equivalent to incorporating an aromatic side chain to lactic acid, the precursor of the widely used polymer polylactic acid. The addition of a bulky functionality could improve the physical property of polylactic acid. To augment the production of phenyllactic acid in *E. coli*, Fujita et al. expressed a heterologous PPR encoded by *pprA* from *Wickerhamia fluorescens* in the *E. coli* strain that was mutated to prevent the synthesis of L-Tyr and L-Trp. Fed-batch fermentation was carried out with glucose supplemented with L-Tyr and L-Trp at concentrations of 0.15–0.2 g/L for 144 h, resulting in phenyllactic acid production at a titer of 29.2 g/L and a yield of 0.16 g/g glucose (Fujita et al., 2013).

6.1.2 2-Phenylethanol

Another endogenously synthesized compound from the phenylpyruvate node is obtained through the Ehrlich pathway to produce 2-phenylethanol. Due to its floral scent, 2-phenylethanol is one of the three most popular benzenoid flavors (the others being vanillin and cinnamyl alcohol) and widely applied in food and cosmetics. The Ehrlich pathway is initiated with a transamination reaction to convert amino acids to the corresponding α-keto acids, followed by a decarboxylation step to fusel aldehydes. Depending on the availability of reductases or oxidases, fusel aldehydes are converted to fusel acids or alcohols as the end products. In *S. cerevisiae*, the endogenous phenylpyruvate decarboxylase that is usually overexpressed for this conversion is Aro10, and the intermediate phenylacetaldehyde can be acted on by the alcohol dehydrogenase Adh2 to complete the Ehrlich pathway and generate fusel alcohols. Kim et al. chose to express Aro10 and Adh2 from *S. cerevisiae* in the fast-growing thermotolerant yeast *Kluyveromyces marxianus* (Kim et al., 2014b). Evolutionary engineering was then attempted to grow this strain during serial subcultures in media containing *p*-fluorophenylalanine, an L-Phe analog acting as a competitive inhibitor of protein synthesis. Previously, it was used to isolate an L-Phe-overproducing mutant *S. cerevisiae* strain that was relieved from feedback regulation. The evolved *K. marxianus* strain produced 1.3 g/L 2-phenylethanol without additional L-Phe supplementation. Surprisingly, overexpressing transketolase (TktA) and PEP synthetase (PPS) to increase the availability of the two precursors of the shikimate pathway did not enhance production (Kim et al., 2014b). The *de novo* biosynthesis of 2-phenylethanol has also been established in an *E. coli* strain expressing the heterologous 2-keto acid decarboxylase (Kdc4427) and alcohol dehydrogenase (Adh4428) enzymes from *Enterobacter* sp. The strain also carried those commonly used genetic manipulations to enhance the flux to the upstream phenylpyruvate pathway and precursor supply, but the titer (320 mg/L) and productivity (13.3 mg/L/h) were

lower than those achieved by the eukaryotic hosts. A toxicity assay suggested severe inhibition of cell growth when the concentration of 2-phenylethanol reached 2 g/L (Liu et al., 2018a).

Considering the toxicity of the product at economically viable concentrations, whole-cell bioconversion and in *situ* product removal were often applied in combination to achieve high production of 2-phenylethanol. Kim et al. converted 10 g/L L-Phe (used as the sole nitrogen source) to 2-phenylethanol in *S. cerevisiae* at 6.1 g/L (Kim et al., 2014a). Polypropylene glycol 1200 was added to the medium as an extractant for 2-phenylethanol. In addition to Aro9 and Aro10, the transcriptional activator Aro80 was also overexpressed because it could be activated by aromatic amino acids and aromatic alcohols through positive feedback regulation. The study also confirmed the role of Ald3, with its removal disrupting the competing step from phenylacetaldehyde to phenylacetate. Another study also indicated that the deletion of Aro8 encoding aromatic amino acid transaminase in *S. cerevisiae* enhanced the accumulation of 2-phenylethanol, suggesting that Aro8 and Aro9 are two transaminases catalyzing opposite reactions (Romagnoli et al., 2015). Combining biotransformation with organic extraction has been applied to *S. cerevisiae* and *K. marxianus* using oleic acid and polypropylene glycol 1200, respectively, as the extractants to improve the production to 12.6 g/L and 10.2 g/L; both systems relied on the endogenous Ehrlich pathway (Etschmann and Schrader, 2006; Stark et al., 2002).

6.1.3 *S*-mandelate and *R*-mandelate

Chiral mandelic acid and 4-hydroxymandelic acid are important fine-specialty chemicals in the synthesis of various pharmaceuticals (Chang et al., 2007; Reifenrath and Boles, 2018; Ward et al., 2007). 4-Hydroxymandelic acid can also be converted to ethyl vanillin, which is an important flavoring agent in food, beverages, and cosmetics (Reifenrath and Boles, 2018). 4-

Hydroxymandelate synthase (HmaS) from *Amycolatopsis orientalis* converts 4-hydroxyphenylpyruvate to 4-hydroxymandelate but also recognizes phenylpyruvate as a less-preferred substrate. Reifenrath et al. created two *S. cerevisiae* strains sharing a basic design foundation: increasing the flux in the shikimate pathway, preventing downstream decarboxylation and transamination (deleting *aro8, aro9, aro10*, and *pdc5*), and removing the L-Trp branch (deleting *trp2*) (Reifenrath and Boles, 2018). Depending on whether PheA2 or Tyr1 was deleted, the L-Phe or L-Tyr branch was disrupted and 1 g/L hydroxylmandelate or 0.236 g/L mandelate was produced, respectively.

Mandelate produced from the HmaS-catalyzed reaction is actually in the S-form. Two additional steps convert the S-form to the R-form: 4-hydroxymandelate oxidase (Hmo) from *Streptomyces coelicolor* can convert S-mandelate to phenylglyoxylate, and D-mandelate dehydrogenase (DMD) from *Rhodotorula graminis* when heterologously expressed catalyzes the unique stereo-reduction of phenylglyoxylate to *R*-mandelate in *E. coli*. Similar to engineering *S. cerevisiae* as a host, to trap phenylpyruvate for mandelate synthesis, the first steps entering the L-Trp branch and L-Tyr branch were disrupted by deleting *trpE* and *tyrA* genes, and the downstream enzymes (encoded by *tyrB* and *aspC*) catalyzing the conversion of phenylpyruvate to L-Phe were also removed. The resulting strains were able to produce *S*-mandelate at 1.02 g/L when HmaS was expressed together with $AroF^{fbr}$ and $PheA^{fbr}$ in shake flasks, and if Hmo and DMD were included, *R*-mandelate was produced at 0.88 g/L (Sun et al., 2011).

6.1.4 L-Phenylglycine and D-phenylglycine

Through a promising transaminase, HpgT, phenylglyoxylate can be converted to L-phenylglycine, which is an important side-chain building block for synthesizing the antibiotics

penicillin, virginiamycin S, and pristinamycin I, and the antitumor drug taxol (Croteau et al., 2006; Liu et al., 2014a; Mast et al., 2011). The aminotransfer step uses an unusual amino donor, L-Phe, and releases phenylpyruvate, which is the beginning substrate of this L-phenylglycine branch. Sets of enzymes (HmaS, Hmo, and HpgT) originating from *A. orientalis* and *S. coelicolor* were purified from *E. coli* and characterized *in vitro*. The set from *S. coelicolor* demonstrated higher specific activities and was then expressed in a previous L-Phe-overproducing *E. coli* strain (Liu et al., 2014b). This strain has its glucose transportation PTS system attenuated by disrupting the *crr* gene to avoid glucose overflow, which reduces acetic acid accumulation and thus enhances the yield of L-Phe. It also carried a thermostable mutant, $AroG^{D146N}$, which was beneficial to the fermentation setup, considering that thermosensitive promoters were used to drive gene expression. Besides, t*yrA* mutant was also created to eliminate the formation of 4-hydroxyphenylpyruvate; *tyrB* and *aspC* were both disrupted to save phenylpyruvate just for L-phenylglycine synthesis. The production was reported as 51.6 mg/g dry cell weight (Liu et al., 2014a). Thus, HpgT from *S. coelicolor* is an L-phenylglycine aminotransferase, and the product is in the L-form. There are also HpgATs in nature that specifically catalyze the amino transfer to produce D-phenylglycine. Similar to L-phenylglycine, D-phenylglycine is often used for preparing semi-synthetic pharmaceuticals such as cephalosporin and penicillin. D-Phenylglycine aminotransferase was obtained by screening the *P. putida* genome DNA library in *E. coli*. As the amino transfer reaction was reversible, D-hydroxyphenylglycine was used as the substrate, and the screening was based on the formation of 4-hydroxyphenylglyoxylate, which has a specific absorbance at 340 nm. The gene carried by the positive clone was found to strongly correlate with the previously confirmed D-phenylglycine aminotransferase HpgAT from *Pseudomonas stutzeri*. *In vitro* characterization supported the

stereoinverting capability of this HpgAT: only L-glutamate was accepted as the donor, and when D-glutamate was used, no D-phenylglycine could be produced. By assembling this gene with HmaS from *A. orientalis* and Hmo from *S. coelicolor* into the *E. coli* strain that was engineered to accumulate phenylpyruvate (by overexpressing $AroF^{fbr}$ and $PheA^{fbr}$ and deleting *pheA*, *tyrA*, *tyrR*, *tyrB*, and *aspC*), 146 mg/g dry cell weight D-phenylglycine was obtained in fed-batch fermentation (Muller et al., 2006).

6.2 The cinnamate node

6.2.1 *Trans*-cinnamic acid, cinnamaldehyde, cinnamyl alcohol, hydrocinnamyl alcohol, and benzoic acid

Trans-cinnamic acid is a phenylpropanoid that is widely used in food and cosmetics as a flavoring or fragrance agent, and it also possesses antibacterial and anticancer properties (De et al., 2011; Sova, 2012). It also serves as the precursor for synthesizing many other value-added compounds, such as flavonoids, styrene, *p*-hydroxycinnamic acids, cinnamaldehyde, and caffeic acid. A previously engineered L-Phe-overproducing *E. coli* strain was used to heterologously express phenylalanine ammonia lyase (PAL) from *Streptomyces maritimus*, which produced *trans*-cinnamic acid at 6.9 g/L when casamino acid was supplied as a feeding solution (Bang et al., 2018).

The three major cinnamon or sweet-spice aromas that have been widely used in food, cosmetic, and pharmaceutical industries are cinnamaldehyde, cinnamyl alcohol, and hydrocinnamyl alcohol. Cinnamaldehyde is also a popular natural nematicide (Kong et al., 2007; Ooi et al., 2006). Its current production mainly relies on tissue extraction from *Cinnamomum* species or chemical synthesis associated with difficult downstream processing. To enable

microbial production, Gottardi et al. heterologously expressed three genes encoding (1) phenylalanine ammonia lyase 2 (PAL2) from *A. thaliana*, converting L-Phe to *trans*-cinnamic acid; (2) ACAR from *Nocardia* sp., reducing *trans*-cinnamic acid to cinnamaldehyde; and (3) phosphopantetheinyl transferase (EntD) from *E. coli* to activate *acar*, because an endogenous phosphopantetheinylation activity is missing in *S. cerevisiae* (Gottardi et al., 2017). Cinnamaldehyde was not the final product, and it was reduced by the endogenous ADHs to cinnamyl alcohol, and then, unknown enzymes in *S. cerevisiae* continued reducing cinnamyl alcohol to hydrocinnamyl alcohol as the major product (113.1 mg/L). Toxicity assays indicated that *trans*-cinnamic acid and cinnamaldehyde exerted high toxicity at 100 mg/L, whereas cinnamyl alcohol and hydrocinnamyl alcohol exhibited mild cell growth inhibition at 200 mg/L.

Another biosynthetic route for cinnamaldehyde is through cinnamoyl-CoA. After the deamination of L-Phe into *trans*-cinnamic acid by PAL, an acid-thiol ligation of *trans*-cinnamic acid to cinnamoyl-CoA can be carried out by 4-coumarate:CoA ligase (4CL), and cinnamoyl-CoA reductase (CCR) will then reduce cinnamoyl-CoA to cinnamaldehyde. Bang et al. chose PAL from *S. maritimus*, 4CL from *S. coelicolor*, and CCR from *A. thaliana* and expressed them in an *E. coli* strain that was intensively engineered for L-Phe production (Bang et al., 2016). Besides the common strategies to improve L-Phe production, $PheA^{E159A, E232A, dm}$, which only had the first 300 amino acids of the wild-type PheA was used, which excluded the L-Phe regulatory domain. The two mutations (E159A, E232A) revived the substrate-binding affinity of the truncated PheA. The resulting strain produced cinnamaldehyde at 75 mg/L. The culture supernatant was found to kill 82% of *Bursaphelenchus xylophilu*s nematodes after 4 h treatment.

Multiple independent studies suggest that the PAL-catalyzed reaction is the rate-limiting step in *trans*-cinnamate production; therefore Masuo et al. designed an alternative route via D-

phenyllactate that did not involve the PAL reaction (Masuo et al., 2016). Some bacteria, such as *Clostridium* spp., possess phenyllactate dehydratase (Fld), which dehydrates D-phenyllactate to *trans*-cinnamate. Fld is uniquely composed of four proteins encoded by *fldABCI.* FldA is a cinnamoyl-CoA:phenyllactate CoA-transferase, transferring the CoA from cinnamoyl-CoA to D-phenyllacate and releasing D-phenyllactyl-CoA and *trans*-cinnamate as the products. FldBC is a dehydratase forming a heterotrimeric enzyme complex with FldA. FldI has the required ATPase activity for activating FldBC-catalyzed dehydration to recycle phenyllacetyl-CoA back to cinnamoyl-CoA; it also has a [4Fe-4S] cluster, and its sensitivity to oxygen disables the host from producing the expected product under regular aerobic conditions. To overcome this hurdle, the fermentation was split into (1) the aerobic process to accumulate D-phenyllacate and (2) a subsequent anaerobic process to convert the supernatant from the first-stage culture to *trans*-cinnamate at 1.7 g/L, which was higher than the titers achieved by using the PAL-dependent route in a batch process.

6.2.2 Styrene and hydroxystyrene

One of the sources for cloning the *pal* gene is *S. maritimus*, a gram-positive, mycelium-forming soil bacterium. *S. maritimus* possesses a plant-like β-oxidation amino acid modification that converts cinnamoyl-CoA into benzoyl-CoA. The wild-type strain accumulates benzoate (125–460 mg/L) after cell growth reaches the maximal level with various carbon sources (glucose, cellobiose, starch, and cellulose), although the detailed underlying mechanism remains unclear. As benzoyl-CoA is the key intermediate in the biosynthetic pathway of the natural product enterocin in *S. maritimus*, *encPHIMJN* were suggested as genes coding for the enzymes responsible for synthesizing benzoic acid. Benzoic acid is the precursor of terephthalic acid,

caprolactam, and phenol, which are the monomeric building blocks of PET, nylon 6, and polycarbonate, respectively (Noda et al., 2012).

Another very important compound that can be derived from the *trans*-cinnamate node is styrene. It is a versatile commodity chemical mainly used to produce various industrially important polymers. Its current production is derived from petroleum through the energy-intensive chemocatalytic dehydrogenation of ethylbenzene (McKenna and Nielsen, 2011; McKenna et al., 2014). Styrene is also synthesized naturally by *Styracaceae* trees, but the corresponding enzymes responsible for styrene biosynthesis have not been deconvoluted thus far. McKenna et al. screened PAL and *trans*-cinnamate decarboxylase (PADC) from bacterial, yeast, and plant sources, and established the first microbial styrene production from sustainable feedstocks (McKenna and Nielsen, 2011). PAL from two cyanobacteria, *Anabaena variabilis* and *Nostoc punctiforme*, and the plant *A. thaliana*, and PADC from the bacteria *Lactobacillus plantarum* and *B. subtilis* and the yeast *S. cerevisiae* were individually expressed, and the cell lysates were evaluated against the corresponding substrate. Although all these candidate PALs displayed the expected activity towards L-Phe, none of the proposed PADCs succeeded in converting *trans*-cinnamate to styrene. Instead, the enzymes from *L. plantarum* and *B. subtilis* displayed decarboxylase activity on *p*-coumaric acid, leading to the production of *p*-hydroxystyrene. Interestingly, the previously characterized ferulic acid decarboxylase (FDC1) from *S. cerevisiae* was shown to be capable of converting both *trans*-cinnamate or *p*-coumaric acid. Its expression with various PALs in an *E. coli* L-Phe-overproducing commercial strain, ATCC 31884, led to the maximal styrene production (260 mg/L) when PAL from *A. thaliana* was assembled with FDC1. Toxicity assays indicated that cell growth stopped when the styrene concentration reached ~300 mg/L. The same research group also attempted to engineer *S.*

cerevisiae as a host. They started with evolving L-Phe-overproducing strains by gradually increasing the amount of *m*-fluoro-DL-phenylalanine, the Phe analog that was previously used to evolve strains with enhanced flux through the L-Phe biosynthetic pathway, showing net extracellular accumulation of 2-phenylethanol and 2-phenylacetate. The resulting positive mutants were shown to have 1.8–9.3-fold upregulation of *ARO* genes, but the coding or the upstream noncoding regions of these genes remained intact. FDC1 expression was naturally induced by *trans*-cinnamate and the basal level expression was found to be sufficient to channel all the *trans*-cinnamate to styrene. Heterologous expression of PAL from *A. thaliana* with an Aro4 feedback-insensitive mutant gene with additional deletion of *ARO10* in the best evolved strain led to styrene production of 29 mg/L (McKenna et al., 2014).

7. Products derived from the L-Tyr branch

7.1 Phenylpropanoic acids: *p*-coumaric acid, caffeic acid, and ferulic acid

p-Coumaric acid (hydroxycinnamate), caffeic acid (3,4-dihydroxycinnamic acid), and ferulic acid are phenylpropanoic acids derived from the L-Tyr branch (**Figure 5 and Table 5**). These phenylpropanoic acids have attracted increasing attention as valuable monomers of liquid crystal polymers for electronics, as well as for their various pharmaceutical properties (Kaneko et al., 2006; Kang et al., 2012). *p*-Coumaric acid also serves as the pivotal precursor for the biosynthesis of a wide variety of flavonoids and stilbenoids in plants. There are two routes to establish *p*-coumaric acid biosynthesis, both starting from the deamination of an aromatic amino acid. One is carried out by tyrosine ammonia-lyase (TAL) to convert L-Tyr to *p*-coumaric acid directly, whereas the other involves the deamination of L-Phe to *trans*-cinnamate, catalyzed by PAL, followed by the hydroxylation of *trans*-cinnamate to yield *p*-coumaric acid, which is

catalyzed by the plant cytochrome P450 cinnamate 4-hydroxylase (C4H) and cytochrome P450 reductase. Various bacteria and eukaryotic microorganisms such as *Sporidiobolus pararoseus*, *Rhodotorula graminis*, *Saccharomycopsis fibuligera*, and *Rhodotorula glutinis* possess a PAL and/or TAL enzyme, which usually displays a preference for L-Phe over L-Tyr (Vannelli et al., 2007). The first over-gram-scale *de novo* production of *p*-coumaric acid was established in *S. cerevisiae* by Rodriguez et al (Rodriguez et al., 2015). Considering that the pathway intermediates in the L-Trp branch only differ from the corresponding ones in the L-Phe branch by the 4-hydroxy group, enzymes such as phenylpyruvate decarboxylase (Aro10), pyruvate decarboxylase (Pdc5), and transaminases (Aro8 and Aro9) can recognize both phenylpyruvate and 4-hydroxyphenylpyruvate as the substrates. When the TAL from *Flavobacterium johnsoniae* was heterologously expressed, the strain produced *p*-coumaric acid at 1.93 g/L using the aforementioned synthetic FIT medium. The same research group subsequently compared the performance of two yeast strain backgrounds, S288c and CEN.PK, with the identical set of genetic manipulations, in both batch and chemostat fermentation settings. The strain S288c is widely used in genetics and physiology studies, whereas the strain CEN.PK is often used for industrial biotechnology advancement (Rodriguez et al., 2017a). With the same genetic modifications, the production of *p*-coumaric acid by CEN.PK reached 2.4 g/L and 0.507 g/L for batch and chemostat fermentation, respectively, which were higher than those achieved by S288c. Through transcriptome analysis, the CEN.PK strain showed 652 genes to be significantly up/downregulated versus 1,927 genes being significantly affected by *p*-coumaric acid production in S288c. Metabolomics analysis also revealed that the changes in intracellular metabolite concentrations in CEN.PK were much lower than those in S288c, suggesting that CEN.PK is a more robust platform strain for producing aromatic compounds.

A hydroxylation step at the 3-position of the benzyl ring converts *p*-coumaric acid to caffeic acid, which is usually catalyzed by cytochrome P450 monooxygenases in plants. An additional OMT further converts caffeic acid to ferulic acid. The 4-coumarate 3-hydroxylase (C3H or Sam5) from *Saccharothrix espanaensis* and OMT from *A. thaliana* (COM) were co-expressed with TAL from *S. espanaensis* in an L-Tyr-overproducing *E. coli* strain, resulting in the production of *p*-coumaric acid at 974 mg/L, caffeic acid at 150 mg/L, and ferulic acid at 196 mg/L in shake-flask culture (Kang et al., 2012). *S. espanaensis* is an actinomycete species producing two heptadecaglycoside antibiotics, saccharomicins A and B, in which caffeic acid links a large oligosaccharide portion to the amino sulfonic acid taurine via an amide bond (Berner et al., 2006). Zhang et al. reconstituted two parallel biosynthesis routes in another L-Tyr-overproducing *E. coli* strain, rpoA14(DE3). One route consisted of TAL from *R. glutinis* and Sam5 from *S. espanaensis*, and the other route also included 4CL obtained from *Petroselinum crispum* to convert *p*-coumaric acid to coumaroyl-CoA and *E. coli* native CoA thioesterases (TE) to convert caffeoyl-CoA to caffeic acid in the last step. The conversion from coumaroyl-CoA to caffeoyl-CoA was catalyzed by Sam5. Batch fermentation showed that the first route rendered a better performance, with the production reaching 106 mg/L after 4 days (Zhang and Stephanopoulos, 2013).

The non-cytochrome P450 hydroxylase HpaBC in both *E. coli* and *Pseudomonas aeruginosa* can hydroxylate *p*-coumaric acid to caffeic acid. When *p*-coumaric acid was fed as a precursor, the *E. coli* strain overexpressing the native HpaBC converted *p*-coumaric acid more efficiently than the bacterial cytochrome P450 enzyme CYP199A2^{F185L} mutant from *Rhodopseudomonas palustris* did, although this mutant has been improved to display a higher conversion efficiency than that of the wild-type CYP199A2. *De novo* biosynthesis was established by expressing TAL

from *R. glutinis* and the native HpaBC in a commercial L-Phe-overproducing *E. coli* strain, ATCC31884, in which additional knockouts of *pheA* and *tyrA* were performed to convert the strain from an L-Phe overproducer to an L-Tyr overproducer (TyrA was replaced by the overexpressed $TyrA^{fbr}$). The combination of these manipulations yielded 767 mg/L caffeic acid after 72 h(Huang et al., 2013). A crude cell assay showed that TAL from *Rhodobacter capsulatus* expressed in *E. coli* has a slightly better activity toward L-Tyr than the widely used TAL from *R. glutinis* does (Lin and Yan, 2012).

Furuya et al. conducted a whole-cell conversion by expressing HpaBC from *P. aeruginosa* in *E. coli*, and 56.6 mM (10.2 g/L) caffeic acid was obtained in 24 h by repeatedly adding 20 mM *p*-coumaric acid four times during the reaction (Furuya and Kino, 2014). A similar whole-cell conversion was also performed with a mutant bacterial cytochrome P450, $CYP199A2^{F185L}$, from *R. palustris* expressed in *E. coli*, which exhibited 5.5 times higher hydroxylation activity for *p*-coumaric acid than that of the wild-type enzyme. The production of caffeic acid reached 15 mM (2.8 g/L) from 20 mM *p*-coumaric acid in 24 h, and in this case, the expression of $CYP199A2^{F185L}$ was accompanied by the redox partners consisting of putidaredoxin reductase (PdR) from *P. putida* and palustrisredoxin (Pux) from *R. palustris* (Furuya et al., 2012). For the conversion from *p*-coumaric acid to caffeic acid in *E. coli*, $CYP199A2^{F185L}$ was a better choice than C3H from *S. espanaensis.* Rodrigues et al. assembled each of these two enzymes with the codon-optimized TAL from *R. glutinis* and found that $CYP199A2^{F185L}$ displayed a higher catalytic activity than that of C3H, which enabled caffeic acid production of 1.56 mM (280 mg/L) from 3 mM L-Tyr as the supplemented precursor (Rodrigues et al., 2015). For all these whole-cell conversions, glucose or glycerol is normally supplemented as an energy source to regenerate the necessary cofactors.

7.2 Other compounds derived from the *p*-coumaric acid node: *p*-hydroxystyrene and phenylpropionic acids

Similar to styrene discussed in section 6.2.2, *p*-hydroxystyrene (4-vinylphenol) is widely used as a monomer for producing various polymers applied in coating, elastomers, and resins in photolithography and semiconductor manufacturing (Verhoef et al., 2009; W. et al., 1999). Its biosynthesis also perfectly mirrors the styrene pathway and relies on the non-oxidative decarboxylation of *p*-coumaric acid catalyzed by *p*-coumaric acid decarboxylase (PDC). Considering that *p*-hydroxystyrene as a solvent is very toxic to the production host, solvent-tolerant hosts such as *P. putida* strains were selected to express PDC from *Lactobacillus plantarum* and PAL/TAL from *Rhodosporidium toruloides*. The strains were previously designed for phenol and *p*-coumarate production, and the *fcs* gene encoding feruloyl-coenzyme A synthetase was additionally disrupted to prevent *p*-coumaric acid degradation. Relying on two-stage fed-batch fermentation in combination with *in situ* product extraction using 1-decanol, 21 mM (2.5 g/L) *p*-hydroxystyrene was obtained overall, while a high concentration of 147 mM product (17.6 g/L) was accumulated in the organic phase (Verhoef et al., 2009). The same pairs of heterologous genes synthesized *p*-hydroxystyrene at 0.4 g/L in an *E. coli* L-Phe overproducer in fed-batch fermentation, and the production toxicity led to a rapid decline in cellular respiration during fermentation (Qi et al., 2007).

Besides *p*-hydroxystyrene, 3-(4-hydroxyphenyl) propionic acid and 3-phenylpropionic acid are important commodity aromatic acids used as synthetic precursors in chemical, food, agriculture, and pharmaceutical industries. They are currently manufactured through chemical synthesis, which involves solvents and catalysts toxic to the environment. Since there are no

precedent pathways existing in nature for their synthesis, their *de novo* biosynthesis relied on finding appropriate enoate reductases (ERs) that can hydrogenate the C=C bond of cinnamic acid or hydroxycinnamic acid (i.e., *p*-coumaric acid). Eleven potential ERs from nine microorganisms were screened, and only the ER from *Clostridium acetobutylicum* displayed the expected activity under both anaerobic and aerobic conditions. By expressing this ER and TAL from *R. glutinis* in the aforementioned commercial L-Phe-overproducing *E. coli* strain ATCC31884 with the additional classical modifications to enhance L-Phe or L-Tyr production, the resulting strains could synthesize 366.77 mg/L 3-phenylpropionic acid and 225.10 mg/L 3-(4-hydroxyphenyl) propionic acid (Sun et al., 2016b).

7.3 Phenylethanoids: tyrosol, hydroxytyrosol, 3,4-dihydroxyphenylacetate, and salidroside

Phenylethanoids are naturally present in olive trees and green tea and exhibit powerful antioxidant activity and disease-preventing potential (Chung et al., 2017; Li et al., 2018a). Their *de novo* biosynthesis starts from the same decarboxylation step of 4-hydroxyphenylpyruvate to 4-hydroxyphenylacetaldehyde. The corresponding enzyme catalyzing this reaction is missing in *E. coli* but can rely on the heterologously expressed Aro10 from *S. cerevisiae*. To channel the carbon flux provided by the upper module shikimate pathway to tyrosol biosynthesis, PheA catalyzing the first and second steps in the L-Phe branch and phenylacetaldehyde dehydrogenase (FeaB) were knocked out, and the latter one is responsible for further converting 4-hydroxyphenylacetaldehyde to the corresponding 4-hydroxyphenylacetate. The native ADHs in *E. coli* were sufficient to convert 4-hydroxyphenylacetaldehyde to tyrosol. The resulting strain produced tyrosol at 573 mg/L from glucose. When 1.8 g/L L-Tyr was used as the substrate, the whole-cell conversion produced 1.2 g/L of tyrosol, wherein the aromatic amino acid

aminotransferase Aro8 from *S. cerevisiae* was expressed to enhance the conversion of L-Tyr to 4-hydroxyphenylpyruvate (Xue et al., 2017). In a later study, Li et al. incorporated both *ARO10* and *ADH6* from *S. cerevisiae* to *E. coli* to improve tyrosol production and extended the pathway for hydroxytyrosol production by overexpressing the *E. coli* native hydroxylase HpaBC (Li et al., 2018a). To improve the production, ascorbic acid was added to prevent hydroxytyrosol oxidation, and 1-dodecanol was added as an *in situ* extractant to overcome product toxicity. NH_4Cl was removed from the culture to force the transamination from L-Tyr to 4-hydroxyphenylpyruvate and release NH_4^+. The combined manipulations in an L-Tyr-overproducing *E. coli* strain with *feaB* knockout resulted in 647 mg/L tyrosol produced from glucose and 1,243 mg/L when L-Tyr (1 g/L) was fed at 0, 12, and 24 h after inoculation. In the same *E. coli* L-Tyr-overproducing strain to produce hydroxytyrosol, when *E. coli feaB* instead of *S. cerevisiae ADH6* was overexpressed, the acetaldehyde moiety was oxidized to the acetic acid moiety, and 3,4-dihydroxyphenylacetate became the final product because the native hydroxylase HpaBC could convert 4-hydroxyphenylacetate. Finally, the pyruvate kinase genes *pykA* and *pykF* were removed to conserve PEP, and the final strain produced 3,4-dihydroxyphenylacetate at 1,856 mg/L in the mixed glucose/glycerol media (Li et al., 2019).

Salidroside, a glucoside of tyrosol, is one of the major active ingredients of the medicinal herb rhodiola, which possesses various health-boosting activities, including bone loss reduction (Chung et al., 2017); neuro-, cardio-, and hepatoprotection; and the prevention of stress-induced impairments and disorders (Liu et al., 2018b). Chung et al. screened 12 uridine diphosphate-dependent glycosyltransferases (UGTs) from *A. thaliana* and found that only UGT85A1 could glycosylate tyrosol to form salidroside (Chung et al., 2017). In an *E. coli* strain expressing

UGT85A1 and a plant aromatic aldehyde synthase (AAS) from *P. crispum* and lacking *tyrR*, *pheA*, and *feaB*, salidroside was produced at 288 mg/L in shake-flask culture.

Recently, high-level production of salidroside was achieved in a syntrophic *E. coli* consortium. The two strains were obligatory cooperators through the cross-feeding of L-Tyr and L-Phe. The L-Phe-deficient strain also lost the ability to take up glucose (via *ptsG* and *manZ* deletions), and instead, it used xylose as the sole carbon source. It expressed 2-keto acid decarboxylase (KDC) from *Pichia pastoris* GS115 and used the endogenous ADH to complete the synthesis of tyrosol. The L-Tyr-deficient strain used glucose exclusively (via *xylA* deletion) and was able to convert glucose to UDP-glucose via phosphoglucomutase (Pgm) and UDP-glucose pyrophosphorylase (GalU). This strain expressed the above-mentioned glycosyltransferase UGT85A1 and also carried the deletion of *ushA* encoding UDP-sugar hydrolase to avoid the decomposition of UDP-glucose. The co-culture was optimized with different sugar concentrations and inoculation ratios of the two strains, with the highest production of salidroside reaching 6.03 g/L in fed-batch fermentation (Liu et al., 2018b).

7.4 Other compounds derived from 4-hydroxylphenylpyruvate or the L-Tyr node: salvianic acid A and L-DOPA

Salvianic acid A (3-(3′,4′-dihydroxyphenyl)-2-hydroxypropanoic acid), also called "danshensu", is the major bioactive ingredient of the Chinese herb *Salvia miltiorrhiza* (danshen). It is well known for its promising antioxidant activity and therapeutic potential in the treatment of vascular diseases (Yao et al., 2013; Zhou et al., 2017). Its derivatives, such as salvianolic acids B and A, and rosmarinic acid, have useful properties for applications in pharmaceutical and food industries. The content of salvianic acid A in crude *S. miltiorrhiza* was found to be only 0.045%

(Lam et al., 2007), which prevents large-scale applications. The *de novo* microbial biosynthesis of salvianic acid A can be achieved through the reduction of the ketone group and 3-hydroxylation of the phenyl ring in 4-hydroxyphenylpyruvate regardless of the order of the reactions. 4-Hydroxyphenylpyruvate reductases have been isolated from various plants, but they often showed broad substrate specificity. Yao et al. chose a mutant D-lactate dehydrogenase (D-LDH^{Y52A}) from *Lactobacillus pentosus*, which had been previously created to exhibit catalytic activity towards phenylpyruvate, and assembled it with the native hydroxylase HpaBC, which hydroxylates the phenyl ring in several L-Tyr pathway intermediates (Yao et al., 2013). The pathway was assembled in parallel using either plasmids or the genomic integration in an L-Tyr-overproducing *E. coli* strain. The plasmid version enabled salvianic acid A production of 7.1 g/L in fed-batch fermentation, which was slightly higher than the genomically integrated version (5.6 g/L), but the latter did not require the supplementation of any antibiotics or inducers (Yao et al., 2013; Zhou et al., 2017).

L-3,4-Dihydroxyphenylalanine (L-DOPA) is a phenolic compound obtained from *Mucuna pruriens* seeds. It has been used in the treatment of Parkinson's disease, caused by the deficiency of the neurotransmitter dopamine. Many microbial species possess tyrosinase, which can synthesize L-DOPA in one step, directly from L-Tyr. Feeding *A. oryzae*, *Yarrowia lipolytica*, *Bacillus sp. JPJ*, *Brevundimonas sp. SGJ*, *Pseudomonas melanogenum*, *Vibrio tyrosinaticus*, or *Acremonium rutilum* with L-Tyr (0.5–8.6 g/L) induced the expression of the endogenous tyrosinase, resulting in high-level conversion (0.497–7.7 g/L) to form L-DOPA (Ali and Ikram-ul-Haq, 2006; Ali et al., 2005; Ali et al., 2007; Ikram-ul-Haq et al., 2002; Krishnaveni et al., 2009; Yoshida et al., 1973; Yoshida et al., 1974). The *de novo* biosynthesis from glucose has also been established in *E. coli*, relying on the endogenous *p*-hydroxyphenylacetate 3-

hydroxylase HpaBC, which has a broad substrate specificity and thus can add a hydroxyl group to L-Tyr to form L-DOPA (Wei et al., 2016). Besides the deletion of TyrR (the L-Tyr pathway repressor), the carbon storage regulator A (CsrA), which regulates the expression levels of three enzymes directly participating in PEP metabolism, was disrupted too. The PTS-mediated glucose transport (*ptsH*Δ, *ptsI*Δ, and *crr*Δ) was replaced by the GalP/GlK-mediated system to further increase PEP availability. Besides, glucose-6-phosphate dehydrogenase encoded by *zwf* was knocked out to enhance the carbon flow in the PPP, whereas pyruvate was recycled to PEP via the overexpression of *ppsA*. Prephenate dehydratase and its leader peptide genes (*pheLA*) were also disrupted to eliminate the loss of prephenate to the competing L-Phe biosynthesis branch. Moreover, an innovative protein engineering strategy was applied to fuse $TyrA^{fbr}$, TyrB, and HpaBC *via* a $(G4S)_3$ linker to facilitate substrate channeling. With this foundation strain, L-DOPA production reached 307 mg/L. To further enhance the production, multiplex automated genome engineering (MAGE) was implemented, which targeted the ribosome-binding sites of 23 genes involved in L-Tyr biosynthesis and replaced them with random sequences. After 30 MAGE cycles, a mutant strain with 34% higher production was obtained. With the correction of the spontaneous mutation occurring in the 5′-UTR sequence of HpaC in the HpaBC operon, the final strain was able to produce 8.67 g/L L-DOPA based on fed-batch fermentation.

The above titer was much higher than when relying on endogenous tyrosinase for the conversion of L-Tyr, but it was still lower than that by another bioconversion route, which relies on a different enzyme, tyrosine phenol-lyase (TPL), to assemble L-DOPA or catechol, pyruvate, and ammonia salt. This enzyme has been found in *Erwinia herbicola*, a thermophilic *Symbiobacterium* species, and *C. freundii*, and the conversion has been demonstrated in *E. herbicola* and the heterologous hosts *P. aeruginosa* and *E. coli*, with the highest level of L-

DOPA reaching 29.8 g/L when pyrocatchol and sodium pyruvate were each supplemented at 50 mM and ammonia at 0.65 mM (Lee et al., 1996). When toluene dioxygenase (TDO) and *cis*-toluene dihydrodiol dehydrogenase (TCGDH) from *P. putida* were incorporated, benzene could replace catechol as the starting substrate, and *E. coli* and *P. aeruginosa* produced 592 mg/L and 2.7 g/L of L-DOPA, respectively (Park et al., 1998).

8. Benzylisoquinoline alkaloids

Benzylisoquinoline alkaloids (BIAs) are a diverse class of plant alkaloids derived from L-Tyr (**Figure 6 and Table 6**). They exhibit a broad range of pharmacological activities, including the antimicrobial and anticancer agents sanguinarine (Trenchard and Smolke, 2015) and berberine (Galanie and Smolke, 2015); the cough suppressant noscapine, which also displays anticancer activity (Li et al., 2018b); the analgesics morphine and codeine (Galanie et al., 2015); and magnoflorine, which prevents atherosclerotic disease and protects human lymphoblastoid cells (Minami et al., 2008). Their synthesis requires complex enzymatic and regulatory activities, rendering synthetic chemistry approaches ineffective due to the structural complexity. On the other hand, native plant hosts exhibit distinct alkaloid profiles, but the BIA products and intermediates are accumulated at low levels. Their structural similarity also complicates the downstream purification processes.

The reconstruction of plant-sourced BIA biosynthesis in microbial hosts has been extremely challenging. Many of the enzymes involved are plant endoplasmic reticulum membrane-bound proteins that are often not amenable to heterologous expression. When 20–30 genes sourced from plants, mammals, bacteria, and fungi are combined into one host, it is difficult to balance their expression levels, both among themselves and relative to the upstream enzymes in L-Tyr

biosynthesis, for efficient *de novo* biosynthesis. It is impossible to accurately predict the rate-limiting steps, but any impairment in the pathway will cause an intermediate buildup and thus, significantly decrease the final product yield.

The condensation of dopamine (obtained through the decarboxylation of L-DOPA) and 4-hydroxyphenylacetaldehyde (obtained through the decarboxylation of 4-hydroxyphenylpyruvate) serves as the first committed step in BIA biosynthesis in native plant hosts such as *Thalictrum flavum* and *Papaver somniferum*. This step is catalyzed by norcoclaurine synthase (NCS) to produce (*S*)-norcoclaurine. The subsequent chain of methylation reactions catalyzed by norcoclaurine 6-*O*-methyltransferase (6OMT), coclaurine-*N*-methyltransferase (CNMT), and 3′-hydroxy-*N*-methylcoclaurine 4′-*O*-methyltransferase (4′OMT), and a hydroxylation reaction catalyzed by cytochrome P450 (CYP80B1 or NMCH) lead to the synthesis of (*S*)-reticuline, the first key branch point intermediate, from which a wide variety of BIA metabolites can be derived. Before these plant-sourced enzymes were functionally expressed in microbial hosts, studies focused on using the commercially available substrate norlaudanosoline to yield (*R*,*S*)-reticuline. This alternative substrate differs from the natural substrate (*S*)-norcoclaurine at the 3′-hydroxyl group, and therefore, the CYP80B1-catalyzed hydroxylation can be skipped in the synthetic pathway. Smolke's group expressed 6OMT, CNMT, and 4′OMT enzyme variants from *T. flavum* and *P. somniferum* in all possible combinations in *S. cerevisiae* cells fed with 1 mM norlaudanosoline, after which the combinations with high conversion were selected (Hawkins and Smolke, 2008). Using an enzyme titration strategy that increased the expression level of each of the three methyltransferases using a galactose-inducible plasmid, the expression of 6OMT was found to be rate limiting. By additional expression of the berberine bridge enzyme (BBE) from *P. somniferum*, the *N*-methyl moiety of (*S*)-reticuline completed the oxidative cyclization and was

converted to (*S*)-scoulerine, which represents a second important branch point metabolite entering either the sanguinarine branch or the berberine branch.

For the berberine branch, in a subsequent study (Galanie and Smolke, 2015), scoulerine 9-*O*-methyltransferase (S9OMT) and cytochrome P450 canadine synthase (CAS) from *T. flavum*, *P. somniferum*, *Coptis japonica*, and *Argemone mexicana*, together with cytochrome P450 reductase (CPR) from *A. thaliana*, were expressed, leading to the production of 1.8 mg/L (*S*)-canadine and 6.5 μg/L berberine (spontaneous oxidization product from (*S*)-canadine) in *S. cerevisiae* in shake-flasks and 621 μg/L (*S*)-canadine and 27 μg/L berberine in a 0.2 L batch bioreactor. In both cases, 1–2 mM norlaudanosoline was supplied.

Reconstitution of the sanguinarine biosynthesis branch mainly encompasses overcoming the challenges associated with expressing multiple plant cytochrome P450s, including cheilanthifolinesynthase (CFS), stylopine synthase (STS), *cis*-*N*-methylstylopine14-hydroxylase (MSH), and protopine 6-hydroxylase (P6H). In another study conducted by Smolke's group (Trenchard and Smolke, 2015), all available enzyme variants from five species (*A. mexicana, Eschscholzia californica, P. somniferum, S. cerevisiae, and A. thaliana*) were tested. The results proved that it is not necessary to pair a cytochrome P450 with the CPR from the same species. Besides, high-level expression of endoplasmic reticulum membrane-bound cytochrome P450s is stressful to yeast, and thus, a balanced expression at an appropriate growth stage is important for high-level production. Other parameters that impacted cytochrome P450 expression and sanguinarine production included growing temperature and carbon source. With these strategies combined, the yeast strain fed with norlaudanosoline yielded 80 μg/L sanguinarine, together with various intermediates ranging from 76 to 548 μg/L.

Although feeding the commercial substrate norlaudanosoline simplifies part of the pathway design, the natural substrates for 6OMT and CNMT are not norlaudanosoline and 6-*O*-methylnorlaudanosoline, respectively. Transport of BIA substrates is also somewhat inefficient, warranting the consolidation of the entire *de novo* biosynthesis from a cheap feedstock. In plants and animals, L-DOPA is mainly synthesized from L-Tyr by a BH4-dependent tyrosine hydroxylase (TyrH), whereas *E. coli* and yeast do not have the native ability to synthesize BH4. To equip *E. coli* with the ability to synthesize (*S*)-reticuline from glucose or glycerol, a special copper-containing tyrosinase (TYR) from *Streptomyces castaneoglobisporus* was expressed in an L-Tyr-overproducing *E. coli* strain, leading to the production of 293 mg/L L-DOPA in the absence of reconsitutited BH4 biosynthesis (Nakagawa et al., 2011). L-DOPA decarboxylase (DODC) from *P. putida*, monoamine oxidase (MAO) from *Micrococcus luteus*, and NCS from *C. japonica* were incorporated into this strain. Unlike the condensation in *S. cerevisiae*, the substrates of NCS are dopamine and 3,4-dihydroxyphenylacetaldehyde (instead of 4-hydroxyphenylacetaldehyde), which is the hydroxylation product converted from dopamine catalyzed by MAO. Therefore, the product of this three-step reaction is (*S*)-norlaudanosoline, differing from (*S*)-norcoclaurine at the 3-hydroxyl group. Their expression together with 6OMT, CNMT, and 4′OMT, all from *C. japonica*, enabled *E. coli* to produce (*S*)-reticuline at 2.26 mg/L based on glucose and 6.24 mg/L when glycerol was used as the carbon source. It was found that L-DOPA was unintentionally converted to its quinone derivative (via the *o*-diphenolase activity of TYR), which was easily polymerized to form dark melanin pigment in the culture. To address this issue, TYR from *S. castaneoglobisporus* was replaced by TYR from *Ralstonia solanacearum* because the latter enzyme was reported to possess low *o*-diphenolase activity.

This replacement led to a 7-fold increase in the *de novo* biosynthesis of (*S*)-reticuline based on glycerol (46 mg/L).

Considering that yeast does not naturally synthesize BH4, in order to functionally express tyrosine hydroxylase originated from mammals, BH4 biosynthesis and regeneration pathway has to be reconstituted. The strategies have been discussed in section 5.1—activating tryptophan 5-hydroxylase, the enzyme converting L-Trp to 5-HTP and requiring the participation of BH4 (**Figure 3**). Briefly, GTP cyclohydrolase I (GCHI), 6-pyruvoyl-tetrahydropterin synthase (PTPS), and sepiapterin reductase (SepR) convert GTP to BH4; pterin-4alpha-carbinolamine dehydratase (PCD) and quinonoid dihydropteridine reductase (QDHPR) are required to recycle BH3OH back to BH4. Except GCHI, all the enzymes need to be heterologously expressed. Another challenge is that TyrH is subjected to both substrate and product inhibitions. A few mutations were reported to relieve such inhibitions (Briggs et al., 2011; Nakashima et al., 2002; Quinsey et al., 1998), among which the triple mutations (W166Y, R37E, R38E) created on TyrH from *R. norvegicus* led to the highest level of conversion (Trenchard et al., 2015). When coupled with DODC from *P. putida*, NCS from *C. japonica*, the three methyltransferases from *P. somniferum*, and the cytochrome P450 CYP80B1 (also known as NMCH) from *E. californica* with its reductase partner from *P. somniferum*, the mutant $TyrH^{W166Y, R37E, R38E}$ enabled the *S. cerevisiae* strain to produce 19.2 μg/L (*S*)-reticuline from glucose.

An alternative is to identify a BH4-independent tyrosine hydroxylase that could be functional in *S. cerevisiae*. DeLoache et al. implemented protein engineering strategies to cytochrome P450 tyrosine hydroxylase from sugar beet (*Beta vulgaris*) (DeLoache et al., 2015). To facilitate the screening, an enzyme-coupled biosensing platform was developed to convert the intermediate L-DOPA into a fluorescent pigment betaxanthin. Using this biosensor, a double mutant (TyrH*,

also known as CYP76AD1$^{W13L, F309L}$) was obtained by PCR mutagenesis, which improved the L-DOPA titer and avoided the side reaction from L-DOPA to L-dopaquinone. When coupled with other downstream enzymes described in the previous example when BH4-dependent TyrH was studied (Trenchard et al., 2015), this mutant BH4-independent TyrH* enabled the *de novo* biosynthesis of 80.6 μg/L (*S*)-reticuline from glucose in *S. cerevisiae*.

Although *E. coli* has been engineered to produce (*S*)-reticuline at a much higher level than that produced by *S. cerevisiae*, the production of diverse alkaloids beyond (*S*)-reticuline is still difficult in *E. coli* because several membrane-bound cytochrome P450s cannot be functionally expressed. To address this issue, an *E. coli–S. cerevi*siae co-culture was set up (Minami et al., 2008). *S. cerevisiae* cells expressing either BBE or the diphenyl ring bridging enzyme (CYP80G2) and CNMT from *C. japonica* were cultured for 20 h and then combined with *E. coli* cells containing (*S*)-reticuline biosynthetic genes, which had been cultured in the presence of 5 mM dopamine for 12 h. The co-culture was incubated for another 48–72 h. Magnoflorine and scoulerine were synthesized at 7.2 mg/L and 8.3 mg/L, respectively.

With the success in connecting (*S*)-reticuline biosynthesis to the utilization of a cheap carbon source in microbial hosts, pioneering studies have been conducted to demonstrate the reconstitution of complex natural product biosynthesis processes such as the complete biosynthesis of the pain-killing opioids thebaine (6.4 μg/L) and hydrocodone (0.3 μg/L) (Galanie et al., 2015) and the anticancer and cough-suppressing agent noscapine (2.2 mg/L) (Li et al., 2018b). Twenty to thirty enzymes from plants, bacteria, yeast, and mammals, including multiple additional challenging plant enzymes, were expressed in *S. cerevisiae*. Tremendous endeavors have been made in identifying the missing catalytic steps, thereby enabling (1) the functional expression of the challenging enzymes through mutagenesis and chimeric protein construction,

(2) the identification of the rate-limiting steps, (3) the tuning of the expression levels, and (4) adjustments in fermentation conditions. The current production levels in microbial systems have not satisfied the conditions for mass production yet. Considering the high price of this group of compounds, production at ~5 g/L would enable an economically competitive process compared to the plant-sourced production. Nevertheless, all these studies have pushed the frontier of tackling the challenges associated with extremely long multi-source pathways for the production of rare high-value molecules and have opened up potential pathways for synthesizing novel BIAs.

9. Flavonoids and stilbenoids

Flavonoids and stilbenoids comprise two other large families of secondary metabolites that are derived from aromatic amino acid biosynthesis in plants and display a wide variety of health-promoting properties (antioxidant, antiviral, antibacterial, and anticancer properties) (**Figure 7 and Table 7**). Similar to BIA biosynthesis, plant-sourced synthesis is constrained by the long cultivation cycle and complex recovery process. The limited genetic tools available to manipulate plant genomes also make plant-sourced production very challenging for mechanistic studies and the creation of novel derivatives with potentially improved functions. In comparison, although the productivities currently achieved by engineered microbial hosts have not met commercial expectations, the scope for improvement can be foreseen owing to the ease of manipulating microbial hosts. As shown by the examples discussed in the sections below, microbial platforms offer the unique advantage to achieve combinatorial assembly of enzymes from a wide variety of sources, which enlarge the mix-and-match space for seeking optimal outcomes. Among thousands of structures belonging to this group, we mainly selected the ones

with established *de novo* biosynthetic pathways in microbial hosts with the aim of illustrating their connections to the upstream aromatic amino acid biosynthesis.

p-Coumaric acid (derived from L-Tyr) and cinnamate (derived from L-Phe) are the starting precursors for synthesizing flavonoids containing linear C6–C3–C6 skeletons (e.g., naringenin and (*2S*)-pinocembrin), stilbenoids having C6–C2–C6 skeletons (e.g., resveratrol and pinosylvin), and a wide variety of derivatives (through methylation, alkylation, oxidation, *C*- and *O*-glycosylation, and hydroxylation). Cinnamate can be hydroxylated to *p*-coumaric acid through cinnamate 4-hydroxylase (C4H), a cytochrome P450 whose functional expression requires the assistance of cytochrome P450 reductase (CPR1). 4CL adds a CoA moiety to form *p*-coumaroyl-CoA. Depending on the availability of the next-step enzyme, *p*-coumaroyl-CoA can be condensed with three molecules of malonyl-CoA to form either resveratrol (catalyzed by a type III polyketide synthase STS) or naringenin (catalyzed by chalcone synthase CHS and chalcone isomerase CHI).

9.1 Representative flavonoids: naringenin and (*2S*)-pinocembrin

Santos et al. described the assembly of enzymes from two microbial sources and multiple plant species (Santos et al., 2011), including TAL (from *R. glutinis* or *Rhodobacter sphaeroides*), 4CL (from *S. coelicolor* or *P. crispum*), CHS (from *A. thaliana* or *Pericallis hybrida*), and CHI (from *Medicago sativa* or *Pueraria lobata*) in an L-Tyr-overproducing *E. coli* strain. Through extensive exploration of enzyme sources and fine-tuning of relative gene expression levels, the best combination produced naringenin at 29 mg/L from glucose without additional supplementation of any precursor from the aromatic amino acid biosynthetic pathway. Since malonyl-CoA is the other substrate for naringenin biosynthesis, when 20 mg/L cerulenin (an

inhibitor of fatty acid biosynthesis) was supplemented, the strain was capable of producing up to 84 mg/L naringenin, indicating the necessity of optimizing the availability of malonyl-CoA. A similar combination of PAL (instead of TAL) from *R. glutinis*, 4CL, CHS, and CHI was incorporated into a Phe-overproducing strain, in which *matC* (encoding *Rhizobium trifolii* malonate carrier protein) and *matB* (encoding *R. trifolii* malonate synthetase) were expressed to convert the supplemented malonate to malonyl-CoA (Wu et al., 2013). The counterpart product of naringenin, (*2S*)-pinocembrin, was produced at 40.02 mg/L.

With the same strain backgrounds, two independent *E. coli* strains were designed to expressing different portions of the biosynthetic pathway, one of which secreted both L-Tyr and *p*-coumaric acid to the co-cultured partner to synthesize naringenin (Ganesan et al., 2017). Owing to the flexibility to adjust the relative biosynthetic strength between the co-cultured specialists with reduced metabolic stress, the overall production of naringenin was improved to 41.5 mg/L without cerulenin addition. Meanwhile, significant amounts of L-Tyr and *p*-coumaric acid accumulated even when the initial ratio between the upstream strain and the downstream strain was made 1:5. However, the downstream strain presented significant growth disadvantage and its proportion was reduced to 16.7% over a cultivation course of 60 h. The discrepancy between the “pre-set” and final ratios was caused by the dynamic growth and the complex interaction between the two strains (Zhang et al., 2015a; Zhang et al., 2015b) but nonetheless indicated the necessity of flexibly increasing the portion of the rate-limiting steps with the rest of the pathway to improve the overall production.

The *de novo* biosynthesis of naringenin was also established in *S. cerevisiae*, in which the carbon fluxes from both L-Tyr and L-Phe were combined (Koopman et al., 2012). The enzymes were mainly cloned from *A. thaliana*, which has four isozymes for PAL and 4CL, and three

isozymes for CHS and CHI, as well as a number of genes mediating the further modifications of naringenin. To prevent the formation of 2-phenylethanol and tyrosol via the Ehrlich pathway, three out of the four 2-oxo acid decarboxylases, including Aro10, Pdc5, and Pdc6, were deleted, leaving only pyruvate decarboxylase (Pdc1) to enable appropriate strain growth on glucose. Increasing the copies of CHS3 integration was found to reduce the accumulation of *p*-coumaric acid and increase the production of naringenin to 109 mg/L in an aerated, pH-controlled batch reactor.

Regarding the limiting precursor malonyl-CoA, Wu et al. implemented the clustered regularly interspaced short palindromic repeats interference (CRISPRi) system to fine-tune central metabolism to improve the carbon channel towards malonyl-CoA without significantly altering biomass accumulation (Wu et al., 2015). Out of 30 target genes, mainly from the glycolysis, PPP, TCA cycle, and fatty acid biosynthesis pathways, 12 were found to increase malonyl-CoA availability after their expression levels were disrupted. Combinatorial genetic perturbation was conducted, and the highest production of naringenin (421.6 mg/L) was obtained in the *E. coli* strain with the knockdown of *fabF* (encoding 3-oxoacyl-acyl carrier protein synthase II), *fumC* (encoding fumarate hydratase), *sucC* (encoding succinyl-CoA synthetase), and *adhE* (encoding acetaldehyde dehydrogenase). Some of these knockdowns are not intuitive, suggesting that high-level carbon redistribution occurs if the expression of a primary metabolic pathway gene is altered.

9.2 Representative stilbenoids: resveratrol and pinosylvin

Resveratrol is naturally present in higher plants, such as blueberries and grapes, and is known for its antioxidant properties, which have drawn much attention from nutraceutical and cosmetic

industries. Its microbial biosynthetic pathway was established from the assembly of TAL from *Herpetosiphon aurantiacus*, 4CL from *A. thaliana*, and resveratrol synthase (also called stilbene synthase) VST1 or STS from *Vitis vinifera* (Li et al., 2015). A unique mutant of acetyl-CoA carboxylase ($ACC1^{S659A, S1157A}$) was incorporated. The enzyme converts cytosolic acetyl-CoA to the rate-limiting malonyl-CoA, and the double mutantations prevent its phosphorylation and subsequent degradation. Based on the hypothesis that the production of resveratrol is limited by the downstream enzyme expression, 8–11 copies of TAL, 4CL, and VST1 were integrated into the Ty4 retrotransposon sites in the genome of *S. cerevisiae*, resulting in the production of 235.57 mg/L resveratrol in batch fermentation. The production was further optimized by the addition of a fed-batch fermentation following the batch fermentation when glucose was depleted. Providing glucose or ethanol as the carbon source in the fed-batch fermentation stage improved the total production of resveratrol to 415.65 mg/L and 531.41 mg/L, respectively. When VST1 was expressed in *E. coli*, the enzyme can also act on cinnamoyl-CoA, leading to the synthesis of pinosylvin (or *trans*-3,5-dihydroxystilbene) as the counterpart of resveratrol derived from L-Tyr (Wu et al., 2017). The strain simultaneously expressed PAL from *Trichosporon cutaneum* and 4CL from *P. crispum*. CRISPRi revealed that repression the expression of six genes—*fabF*, *fabB*, *eno*, *adhE*, *fumC*, and *sucC*—enabled the highest production of pinosylvin (281 mg/L). Here, FabB catalyzes the same reaction as FabF (fatty acid biosynthesis), and Eno catalyzes the interconversion of 2-phosphoglycerate to PEP. Their repression presumably increases PEP availability.

9.3 Representative derivatives of the naringenin pathway: ponciretin, sakuranetin, kaempferol, liquiritigenin, resokaempferol, quercetin, and fisetin

Naringenin is a building block for synthesizing a wide variety of flavones, isoflavones, and flavonols. Other valuable modifications such as *O*-methylation and glycosylation also occur during flavonoid biosynthesis in plants and confer improved biological activities. For example, the synthesis of two *O*-methylflavonoids, sakuranetin (7-*O*-methylnaringenin), and ponciretin (4′-*O*-methylnaringenin), displaying antifungal and antibacterial activities, are mediated by naringenin *O*-methyltransferase (NOMT) from *O. sativa* and soybean *O*-methyltransferase (SOMT2) from *Glycine max* (Kim et al., 2013). Assembling each of these enzymes with TAL (from *S. espanaensis*), CHS (from *Populus euramericana*), and 4CL (from *O. sativa*) in an L-Tyr-overproducing *E. coli* strain led to the production of 42.5 mg/L ponciretin and 40.1 mg/L sakuranetin. Note that this strain did not carry the heterologously expressed CHI, indicating that the isomerization reaction might have occurred spontaneously. Besides, *icdA* encoding isocitrate dehydrogenase was deleted to increase the supply of CoA as the precursor for malonyl-CoA synthesis.

A *p*-coumaroyl-CoA-accumulating yeast strain carrying Aro4^{K229L}, Aro7^{G141S}, *F. johnsoniae* TAL, *P. crispum* 4CL, *P. hybrid* CHS, and *M. sativa* CHI, in the background with *Aro10* and *Pdc5* deletion, was constructed as the base strain (Rodriguez et al., 2017b). Extension with *Astragalus mongholicus* flavanone 3-hydroxylase (F3H) and *A. thaliana* flavonol synthase (FLS) led to the synthesis of kaempferol (26.57 mg/L) in synthetic fed-batch FIT media. When chalcone reductase (CHR) (from *A. mongholicus*) was added into the overexpression queue, isoliquiritigenin appeared as an intermediate, and resokaempferol was accumulated at 0.51 mg/L. In this strain, *p*-coumaric acid, kaempferol, naringenin, and liquiritigenin were also observed at much higher levels, indicating the incomplete reduction of *p*-coumaroyl-CoA mediated by CHR. Finally, cytochrome P450 (FMO) from *P. hybrida* was fused in-frame with cytochrome P450

reductase (CPR) from *C. roseus* and co-expressed in the resokaempferol-producing strain to synthesize fisetin (2.29 mg/L) and in the kaempferol-producing strain to synthesize quercetin (20.38 mg/L).

9.4 Representative curcuminoids: curcumin

Curcuminoids are another family of plant polyphenols that prevent oxidative damage in normal cells during the treatment of allergy, asthma, cancer, and Alzheimer's disease. Wang et al. described the assembly of enzymes from different sources as being akin to building Legos to generate several phenylpropanoid acids (cinnamic acid, *p*-coumaric acid, caffeic acid, and ferulic acid), stilbenoids (resveratrol, piceatannol and pinosylvin), and curcuminoids (curcumin, bisdemethoxycurcumin and dicinnamoylmethane) (Wang et al., 2015). Besides the pathways that have been discussed in the previous sections, piceatannol can be obtained from resveratrol through the specific C-3′-hydroxylation catalyzed by C3H from *S. espanaensis*. Co-expression of TAL from *S. espanaensis*, 4CL from *A. thaliana*, STS from *Arachis hypogaea*, and C3H from *S. espanaensis* in *E. coli* led to the production of 21.5 mg/L piceatannol. Bisdemethoxycurcumin is the condensed product of two molecules of *p*-coumaroyl-CoAs, whereas the symmetric reaction occurring at cinnamayl-CoA produces dicinnamoylmethane, both of which can be catalyzed by curcuminoid synthase (CUS) from *O. sativa.* It was found that 4CL from *A. thaliana* can also recognize ferulic acid and thus ligate it with CoA to yield feruloyl-CoA. The final assembly of the genes encoding TAL, C3H, COMT (from *M. sativa*), 4CL, and CUS established a complete biosynthetic pathway in the *E. coli* strain fed with 3 mM L-Tyr, producing curcumin at 0.6 mg/L.

9.5 Dihydrochalcones: phloretin, phlorizin, nothofagin, trilobatin, and naringenin dihydrochalcone

CHS has been found to be shared between flavonoid and dihydrochalcone biosynthesis in various fruits. Phloretin, the first committed dihydrochalcone, is formed by the CHS-catalyzed decarboxylative condensation of *p*-dihydrocoumaroyl-CoA with three units of malonyl-CoA and a subsequent cyclization. The pathway is initiated by the formation of *p*-dihydrocoumaroyl-CoA from *p*-coumaroyl-CoA, which is mediated by a double bond reductase (DBR). This reaction can be realized by the native enzyme TSC13 in *S. cerevisiae*. In plants, the biosynthesis of phloretin can be extended by many different UDP-dependent-glycosyltransferases (UGTs) that act on either the phloretin ring directly or the hydroxyl groups to form the antidiabetic molecule phlorizin, the antioxidant molecule nothofagin, and the sweet compound naringenin dihydrochalcone. The assembly of PAL2 (from *A. thaliana*), AmC4H (from *Ammi majus*), CPR1 (from *S. cerevisiae*), 4CL2 (from *A. thaliana*), CHS (from *Hypericum androsaemum*), and TSC13 (from *S. cerevisiae*) led to an initial production of phloretin at 42.7 mg/L in *S. cerevisiae* (Eichenberger et al., 2017). The incorporation of additional glycotransferase, including UGT88F2 (from *Pyrus communis*), CGT (from *O. sativa*), or UGT73B2 (from *A. thaliana*), led to the production of three glycosylated dihydrochalcones: phlorizin (65 mg/L), nothofagin (59 mg/L), and trilobatin (32.8 mg/L) in 24-well fermentation. Finally, extending the overexpressed enzyme set with the UDP-rhamnose transferases 1,2RHAT (from *Citrus maxima*) and RHM2 (from *A. thaliana*), responsible for converting the native UDP-D-glucose to UDP-L-rhamnose, resulted in the biosynthesis of naringenin dihydrochalcone at 11.6 mg/L. It is noteworthy that the host was not extensively engineered to overproduce any of the precursors, indicating the scope for further improvement in the production of these glycosylated products.

10. Conclusions, advanced strategies and future outlooks

Despite the importance of aromatic amino acid pathway derivatives, their current production levels based on microbial hosts are still far from meeting the criteria of commercialization. For commodity chemicals, even the pre-pilot scale requires titer, yield, and productivity to reach 10–100 g/L, 40–80%, and 0.5–2 g/L/h, respectively. Some compounds in this group can be produced at relatively high titers but their yields and/or productivities are still lower than the standards. For this reason, at least for now, high-value specialty aromatics might be more likely to make commercialization successes (Thompson et al., 2015).

To enable high-level biosynthesis of aromatic amino acid pathway derivatives, especially, plant-derived natural products whose synthesis typically requires 20–30 enzymatic reactions, the strategies that have been summarized in the previous sections at a high level include (1) partitioning a complex pathway into smaller modules, (2) independently addressing the outstanding issues arising from individual modules, and (3) overcoming the obstacles to connecting modules. This is because distinctly different challenges are recognized in individual parts of a pathway. Any impairment in the pathway will cause an intermediate buildup, thereby considerably decreasing the final product yield. The upstream pathway is native to the microbial hosts, being directly linked to their primary metabolism; therefore, the host employs sophisticated controls to minimize unnecessary energy expense (Braus, 1991; Hinnebusch, 1988; Lee and Hahn, 2013; Luttik et al., 2008). For the downstream foreign add-on steps, heterologous expression poses the major limitation (Jiang et al., 2012). Many plant-sourced enzymes are so delicate that the alteration of the intracellular expression environment would impair their functions. The modular design approach enables researchers to reach an optimal balance by

searching a relatively small combinatorial space, and using delicate “design-build-test-learn” cycles to improve the production by thousands of times over the starting levels (*e.g.*, 15,000-fold improvement in taxol production (Ajikumar et al., 2010)).

Throughout our research, we identified a few new directions that could potentially improve the production of aromatic amino acid pathway derivatives. Therefore, the following sections are dedicated to discussing the rationales of these strategies. We look forward to seeing the verification of their impacts on the biosynthesis of aromatic amino acid derivatives by the colleagues working in this area in the near future.

10.1 Reconstruct gluconeogenesis to optimize the availability of PEP

In contrast to the various successful engineering approaches that have been devised to enhance the accumulation of E4P, strategies to increase the availability of PEP have not been thoroughly investigated in yeast. The presence of organelles results in the compartmentalization of metabolites, which reduces the availability of the desired precursors in the cytoplasm and dilutes the possibility of reaching higher product titers. Our metabolic flux analysis showed that the rates of PEP conversion to pyruvate are at least one order of magnitude higher than the E4P flux in the pentose phosphate pathway (Suastegui et al., 2016a). This is expected because pyruvate has considerably more metabolic fates in the cell; hence it is in higher demand for production of essential metabolites in aerobic respiration, fatty acid production, and production of other amino acids. Future direction to improve the availability of PEP needs to consider a good metabolic balance for efficient biomass formation.

A new strategy for pyruvate recirculation could explore the expression of heterologous genes, namely *ppsA* from *E. coli*, and the pyruvate, orthophosphate dikinase *PDK* from the plant *A.*

thaliana. One of the advantages of expressing heterologous genes is that posttranslational modifications that inactivate proteins in yeast (such as phosphorylation and ubiquitylation) can be usually avoided. To ensure high pyruvate availability in the cytosol to serve as the substrate for the recirculation enzymes, the main carriers that transport pyruvate into the mitochondria need to be knocked out. This strategy has been previously established in *S. cerevisiae* to increase the cytosolic pools of pyruvate to elevate the production polyketides (Cardenas and Da Silva, 2016). In this study, two mitochondrial, outer membrane transporters (POR1 and POR2) and two mitochondrial pyruvate inner membrane carriers (MPC1 and MPC2) were deleted. Single-gene knockouts revealed that POR2 and MPC2 were good candidates to increase the pool of pyruvate in cytoplasm, resulting in at least a 3-fold improvement in triacetic acid lactone (TAL) production. The combination of these further increased the titers of TAL from ~300 mg L^{-1} to almost 1.0 g L^{-1}.

Gluconeogenesis is a metabolic pathway whose activation is triggered by a diauxic shift. When glucose is depleted, the glycolytic genes will undergo downregulation, while the gluconeogenic ones will be activated with the goal of regenerating primary metabolites. Because the gluconeogenic pathway is highly ATP-consuming, very strict regulations at all metabolic levels protect the cell from unnecessary energy expenditures in the presence of high concentrations of glucose. In *S. cerevisiae*, gluconeogenesis starts from the conversion of cytosolic malate to pyruvate through an $NADP^+$-dependent malic acid enzyme (pyruvic-malic enzyme, MAE1) or through the conversion of malate to OAA through the cytosolic malic acid dehydrogenase enzyme (MDH2) (**Figure 8**). OAA is further converted to PEP through the PCK1 enzyme. Similar to pyruvate, malate can also be transported into the mitochondria through the action of the specific carrier (OAC1). This is an inner membrane transporter that imports the

cytosolic OAA produced from pyruvate by the enzyme pyruvate carboxylase (PYC) (Palmieri et al., 1999). Therefore, the deletion of this transporter gene could be a potential strategy to increase the malate pool for the initiation of gluconeogenesis. In addition, considering that MDH2 and PCK1 are originated from the gluconeogenic pathway whose transcription is repressed in the exponential phase, equipping them directly with strong constitutive promoters might be problematic given the potential impact on the native cellular metabolism. It would be preferable to span a panel of promoters with different strengths to ensure an appropriate expression balance.

10.2 Engineering post-translational modifications

Post-translational modifications (PTMs) serve as another regulatory command to control the stability, localization, and activity of proteins in cells. Two of the most common posttranslational modifications in eukaryotic systems are phosphorylation and ubiquitylation. The former is usually associated with enzymatic activity, whereas the latter serves as a tag for protein degradation. PTMs are commonly orchestrated by a network of genes and the corresponding study has been determinant in understating major hierarchical regulatory mechanisms of key metabolic pathways. Hence, it is not unexpected that PTMs could play a crucial role in the regulation of highly energy-demanding pathways such as gluconeogenesis and the biosynthesis of aromatic amino acid.

Several studies have revealed that, indeed, the gluconeogenic genes undergo major PTMs as an additional control level to regulate their cellular activity and abundance. For example, the aforementioned cytosolic MDH protein (MDH2p) contains two phosphorylation residues at the proline 1 (P1) residue and the threonine 6 (T6) residue. Through biochemical characterizations, it

has been demonstrated that substituting P1 to serine (P1S), or removing the first 12 residues of the protein can reduce its degradation when cells are shifted from non-fermentable carbon sources (acetate) to growth on glucose(Hung et al., 2004; Minard and McAlister-Henn, 1992). Through global proteomic studies, four lysine residues (K42, K180, K254, and K259) subjected to ubiquitylation have also been identified in MDH2 (Swaney et al., 2013). It would be interesting to discern whether a higher activity of this enzyme can be achieved in the presence of glucose. Different mutant versions of the gene MDH2 can be constructed and incorporated into the new genetic platform to study the recirculation of malate to OAA.

To finalize the reconstruction of gluconeogenesis, the phosphoenolpyruvate carboxykinase gene, *PCK1*, can also be overexpressed, in order to ensure the conversion of OAA to PEP. Thus far there is no evidence of PCK1 undergoing PTMs such as phosphorylation or ubiquitylation. However, high sequence similarity with the protein Aro4 was observed in the first 12 amino acids, which is known to undergo phosphorylation at the serine residues 2 and 4 (Albuquerque et al., 2008). Therefore, new mutant versions of PCK1 with amino acid substitutions to avoid phosphorylation can be constructed and overexpressed to complete the production of PEP in the presence of glucose.

10.3 Creating a biochemical driving force

A potential pitfall of overexpressing MDH2 to increase the PEP availability is that the enzyme favors the reverse reaction, which is the conversion of OAA to malate (Zelle et al., 2008). An alternative strategy is to express a version of MDH2 with known irreversible activity, such as the malate:quinone oxidoreductase (Mqo) from *E. coli* (**Figure 8**).

This rational can also be complemented by linking the metabolic engineering strategies with a biochemical driving force (Shen et al., 2011) (Shen et al., 2011). Common driving forces are linked to energy exchange metabolites such as ATP, NADH, and NADPH. In this light, the activity of the enzymes of study can be elevated if its reaction can provide or complement a cofactor deficiency created by the deletion of selected genes. To enhance the production of aromatic compounds, such a driving force can be created by deleting the *ZWF1* gene in the oxidative pentose phosphate pathway. This would require a new source for replenishment of cytosolic NADPH levels that can be attained by expressing dehydrogenase enzymes that depend on the $NADP^+$ cofactor. Expression of heterologous MDH2 enzymes that prefer $NADP^+$ as the cofactor could help establish the metabolic driving force. One potential gene for this strategy is the NADPH-dependent MDH enzyme involved in pyruvate metabolism and carbon fixation pathways in several organisms such as *Zea mays* (Metzler et al., 1989), *A. thaliana*, and *Methanobacterium thermoautotrophicum* (Thompson et al., 1998) (**Figure 8**).

10.4 Global perturbations

Besides the targeted rational strategies that have been discussed in the above sections, dynamic approaches at a genome-scale need to be developed to allow engineering the partition of the flux between two of the most important pathways, namely, glycolysis and the pentose phosphate pathway for the accumulation of aromatic amino acid compounds. The intrinsically complicated regulations involved in the native upstream module prohibit the extra energy expenditure to create “foreign” molecules. With the rapid development of the CRISPR-associated technologies (e.g., CRISPR activation and CRISPR interference), creating a combinatorial library including both upregulation and downregulation at the genome-scale is

now feasible. This platform would enable the modulation of gene expression within its native genome context, bypassing the need to utilize heterologous or synthetic promoters. Rather than limiting the system to an “on” or “off” state (i.e., overexpression or deletion), it can provide a platform in which genes can be studied in a wide range of transcriptional strengths. This property enables the creation of “fluidic” genetic landscapes that otherwise would not be possible by single-gene deletion or overexpression (e.g., the glycolytic genes). Molecular sensors to detect the final product or the intermediates of aromatic amino acid biosynthesis in a high-throughput manner need to be developed. Success has been demonstrated in the colorimetric reaction for detecting L-DOPA and the fluorescence protein-based naringenin sensor (DeLoache et al., 2015; Skjoedt et al., 2016 (accepted)).

In addition, the current repertoire of microbial hosts needs to be broadened. Some nonconventional microbes possess special biochemical and physiological features that are beneficial to the production of aromatic amino acid pathway derivatives. Genome-scale transcriptome programming integrated with high-throughput screening might be the most desirable technology to further deepen the understanding of non-model organisms.

10.5 Enhancing the conversions in the downstream module

For the downstream module, ensuring successful gene expression in heterologous hosts is a complex process, without general rules to follow. It is impossible to accurately predict the rate-limiting steps when expressing enzymes from a wide variety of sources. A consensus drawn from pioneering work in opioid (Galanie et al., 2015; Thodey et al., 2014), artemisinic acid (Paddon et al., 2013; Ro et al., 2006; Westfall et al., 2012), and taxol (Ajikumar et al., 2010; Jiang et al., 2012) production is that it is important to explore the functional space defined by the

natural sequence diversity (Trenchard and Smolke, 2015); in each of these three platforms, a suite of enzymes from mammals, plants, bacteria, and fungi were assembled and tested to determine an optimal combination. Moreover, cytochrome P450 requires a partner enzyme, CPR, to shuttle electrons (Jung et al., 2011), and in some cases also requires interaction with cytochrome b_5, which provides additional redox support to achieve full activity (Paddon et al., 2013; Schenkman and Jansson, 2003; Zhang et al., 2007). Sufficient evidence has indicated that it is not necessary to match cytochrome P450 with the CPR from the same species; pairing multiple species-sourced enzymes has led to better turnovers (Paddon et al., 2013; Trenchard and Smolke, 2015).

Expressing membrane-bound enzymes can induce endoplasmic reticulum stress response because these enzymes have N-terminal signal peptides that are directly inserted into the endoplasmic reticulum (Sandig et al., 1999; Trenchard and Smolke, 2015; Zimmer et al., 1997). In some cases, altering the expression level of cytochrome P450 relative to CPR can increase the product yield because incomplete interactions between these two proteins can result in the release of reactive oxygen species, which reduce cell viability (Li et al., 2016; Paddon and Keasling, 2014; Paddon et al., 2013; Zangar et al., 2004). Besides, the expression of multiple cytochrome P450s could cause heme depletion, which can further induce cellular stress. In this scenario, a strategy of expressing the rate-limiting enzymes in endogenous heme biosynthesis should be considered, which has been shown to successfully alleviate such a stress (Michener et al., 2012).

Streamlining the performance of a process involving more than 20 steps requires the coordination of the corresponding genes. Synthesizing a desired compound often creates competition with cell growth, in which case, being able to easily fine-tune the expression of

individual genes is essential. This might be achieved by varying promoter strength, integrating additional copies of the rate-limiting step, and adjusting the temperature to optimize enzyme activity (Ajikumar et al., 2010; Fossati et al., 2014; Galanie and Smolke, 2015; Trenchard and Smolke, 2015). Once a high-production strain is established in shaking flasks, the production can be continuously improved in a bioreactor. By adjusting parameters such as temperature, pH, oxygen level, and sugar-feeding pattern, usually an increased yield of at least an order of magnitude can be expected.

Author Contributions

M. C., M. G, and Z. S. wrote the sections 1-9; M. S., Y. M, and Z. S. wrote the section 10.

Acknowledgements

This work was supported in part by the National Science Foundation Grants (MCB 1716837 and CBET 1749782).

References

Ajikumar, P. K., Xiao, W. H., Tyo, K. E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., Stephanopoulos, G., 2010. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. Science. 330**,** 70-4.

Akashi, H., Gojobori, T., 2002. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. 99**,** 3695-700.

Albuquerque, C. P., Smolka, M. B., Payne, S. H., Bafna, V., Eng, J., Zhou, H., 2008. A Multidimensional Chromatography Technology for In-depth Phosphoproteome Analysis. Mol. Cell. Proteomics. 7**,** 1389.

Ali, S., Ikram-ul-Haq, 2006. Kinetic basis of celite (CM 2:1) addition on the biosynthesis of 3,4-dihydroxyphenyl-L-alanine (L-DOPA) by *Aspergillus oryzae* ME2 using L-tyrosine as a basal substrate. World J. Microbiol. Biotechnol. 22**,** 347-353.

Ali, S., Ikram-ul-Haq, Qadeer, M. A., Rajoka, M. I., 2005. Double mutant of *Aspergillus oryzae* for improved production of L-dopa (3,4-dihydroxyphenyl-L-alanine) from L-tyrosine. Biotechnol. Appl. Biochem. 42**,** 143-149.

Ali, S., Shultz, J. L., Ikram-ul-Haq, 2007. High performance microbiological transformation of L-tyrosine to L-dopa by *Yarrowia lipolytica* NRRL-143. BMC Biotechnol. 7.

Arndt, K. T., Styles, C., Fink, G. R., 1987. Multiple global regulators control HIS4 transcription in yeast. Science. 237**,** 874-80.

Averesch, N. J., Winter, G., Kromer, J. O., 2016. Production of *para*-aminobenzoic acid from different carbon-sources in engineered *Saccharomyces cerevisiae*. Microb. Cell Fact. 15**,** 89.

Averesch, N. J. H., Kromer, J. O., 2018. Metabolic engineering of the shikimate pathway for production of aromatics and derived compounds-present and future strain construction strategies. Front. Bioeng. Biotechnol. 6**,** 32.

Averesch, N. J. H., Prima, A., Kromer, J. O., 2017. Enhanced production of *para*-hydroxybenzoic acid by genetically engineered *Saccharomyces cerevisiae*. Bioprocess. Biosyst Eng. 40**,** 1283-1289.

Bai, Y., Yin, H., Bi, H., Zhuang, Y., Liu, T., Ma, Y., 2016. *De novo* biosynthesis of gastrodin in *Escherichia coli*. Metab. Eng. 35**,** 138-147.

Balderas-Hernandez, V. E., Sabido-Ramos, A., Silva, P., Cabrera-Valladares, N., Hernandez-Chavez, G., Baez-Viveros, J. L., Martinez, A., Bolivar, F., Gosset, G., 2009. Metabolic engineering for improving anthranilate synthesis from glucose in *Escherichia coli*. Microb. Cell. Fact. 8**,** 19.

Bang, H. B., Lee, K., Lee, Y. J., Jeong, K. J., 2018. High-level production of *trans*-cinnamic acid by fed-batch cultivation of *Escherichia coli*. Process Biochem. 68**,** 30-36.

Bang, H. B., Lee, Y. H., Kim, S. C., Sung, C. K., Jeong, K. J., 2016. Metabolic engineering of *Escherichia coli* for the production of cinnamaldehyde. Microb. Cell Fact. 15**,** 16.

Barker, J. L., Frost, J. W., 2001. Microbial synthesis of *p*-hydroxybenzoic acid from glucose. Biotechnol. Bioeng. 76**,** 376-90.

Berner, M., Krug, D., Bihlmaier, C., Vente, A., Muller, R., Bechthold, A., 2006. Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. J. Bacteriol. 188**,** 2666-2673.

Berry, A., Dodge, T. C., Pepsin, M., Weyler, W., 2002. Application of metabolic engineering to improve both the production and use of biotech indigo. J. Ind. Microbiol. Biotechnol. 28**,** 127-133.

Bongaerts, J., Kramer, M., Muller, U., Raeven, L., Wubbolts, M., 2001. Metabolic engineering for microbial production of aromatic amino acids and derived compounds. Metab. Eng. 3**,** 289-300.

Bonner, C., Jensen, R., 1987. Prephenate aminotransferase. Methods Enzymol. 142**,** 479-87.

Braus, G., Furter, R., Prantl, F., Niederberger, P., Hutter, R., 1985. Arrangement of genes TRP1 and TRP3 of *Saccharomyces cerevisiae* strains. Arch. Microbiol. 142**,** 383-8.

Braus, G., Mosch, H. U., Vogel, K., Hinnen, A., Hutter, R., 1989. Interpathway regulation of the TRP4 gene of yeast. EMBO J. 8**,** 939-45.

Braus, G. H., 1991. Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: a model system for the regulation of a eukaryotic biosynthetic pathway. Microbiol. Rev. 55**,** 349-70.
Briggs, G. D., Gordon, S. L., Dickson, P. W., 2011. Mutational analysis of catecholamine binding in tyrosine hydroxylase. Biochemistry. 50**,** 1545-55.
Brochado, A. R., Matos, C., Moller, B. L., Hansen, J., Mortensen, U. H., Patil, K. R., 2010. Improved vanillin production in baker's yeast through *in silico* design. Microb. Cell Fact. 9.
Cardenas, J., Da Silva, N. A., 2016. Engineering cofactor and transport mechanisms in *Saccharomyces cerevisiae* for enhanced acetyl-CoA and polyketide biosynthesis. Metab. Eng. 36**,** 80-89.
Chandran, S. S., Yi, J., Draths, K. M., von Daeniken, R., Weber, W., Frost, J. W., 2003. Phosphoenolpyruvate availability and the biosynthesis of shikimic acid. Biotechnol. Prog. 19**,** 808-14.
Chang, M. C., Keasling, J. D., 2006. Production of isoprenoid pharmaceuticals by engineered microbes. Nat. Chem. Biol. 2**,** 674-81.
Chang, T. L., Teleshova, N., Rapista, A., Paluch, M., Anderson, R. A., Waller, D. P., Zaneveld, L. J., Granelli-Piperno, A., Klotman, M. E., 2007. SAMMA, a mandelic acid condensation polymer, inhibits dendritic cell-mediated HIV transmission. FEBS Lett. 581**,** 4596-602.
Chen, Z., Shen, X., Wang, J., Wang, J., Yuan, Q., Yan, Y., 2017. Rational engineering of *p*-hydroxybenzoate hydroxylase to enable efficient gallic acid synthesis via a novel artificial biosynthetic pathway. Biotechnol. Bioeng. 114**,** 2571-2580.
Chung, D., Kim, S. Y., Ahn, J. H., 2017. Production of three phenylethanoids, tyrosol, hydroxytyrosol, and salidroside, using plant genes expressing in *Escherichia coli*. Sci. Rep. 7.
Croteau, R., Ketchum, R. E., Long, R. M., Kaspera, R., Wildung, M. R., 2006. Taxol biosynthesis and molecular genetics. Phytochem. Rev. 5**,** 75-97.
Curran, K. A., Leavitt, J. M., Karim, A. S., Alper, H. S., 2012. Metabolic engineering of muconic acid production in *Saccharomyces cerevisiae*. Metab. Eng. 15**,** 55-66.
De, P., Baltas, M., Bedos-Belval, F., 2011. Cinnamic acid derivatives as anticancer agents-a review. Curr. Med. Chem. 18**,** 1672-703.
DeLoache, W. C., Russ, Z. N., Narcross, L., Gonzales, A. M., Martin, V. J., Dueber, J. E., 2015. An enzyme-coupled biosensor enables (*S*)-reticuline production in yeast from glucose. Nat. Chem. Biol. 11**,** 465-71.
Draths, K. M., Ward, T. L., Frost, J. W., 1992. Biocatalysis and nineteenth century organic chemistry: conversion of D-glucose into quinoid organics. J. Am. Chem. Soc. 114**,** 9725-9726.
Du, J., Yang, D., Luo, Z. W., Lee, S. Y., 2018. Metabolic engineering of *Escherichia coli* for the production of indirubin from glucose. J. Biotechnol. 267**,** 19-28.
Duran, N., Justo, G. Z., Ferreira, C. V., Melon, P. S., Cordi, L., Martins, D., 2007. Violacein: properties and biological activities. Biotechnol. Appl. Biochem. 48**,** 127-133.
Eichenberger, M., Lehka, B. J., Folly, C., Fischer, D., Martens, S., Simon, E., Naesby, M., 2017. Metabolic engineering of *Saccharomyces cerevisiae* for *de novo* production of dihydrochalcones with known antioxidant, antidiabetic, and sweet tasting properties. Metab. Eng. 39**,** 80-89.
Etschmann, M. M. W., Schrader, J., 2006. An aqueous-organic two-phase bioprocess for efficient production of the natural aroma chemicals 2-phenylethanol and 2-phenylethylacetate with yeast. Appl. Microbiol. Biotechnol. 71**,** 440-443.
Federspiel, M., Fischer, R., Hennig, M., Mair, H. J., Oberhauser, T., Rimmler, G., Albiez, T., Bruhin, J., Estermann, H., Gandert, C., Gockel, V., Gotzo, S., Hoffmann, U., Huber, G., Janatsch, G., Lauper, S., Rockel-Stabler, O., Trussardi, R., Zwahlen, A. G., 1999. Industrial synthesis of the key precursor in the synthesis of the anti-influenza drug oseltamivir phosphate (Ro 64-0796/002, GS-4104-02): Ethyl (3R,4S,5S)-4,5-epoxy-3-(1-ethyl-propoxy)-cyclohex-1-ene-1-carboxylate. Org. Process Res. Dev. 3**,** 266-274.
Fitzpatrick, P. F., 2003. Mechanism of aromatic amino acid hydroxylation. Biochemistry. 42**,** 14083-91.

Fossati, E., Ekins, A., Narcross, L., Zhu, Y., Falgueyret, J. P., Beaudoin, G. A., Facchini, P. J., Martin, V. J., 2014. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in *Saccharomyces cerevisiae*. Nat Commun. 5**,** 3283.

Fujita, T., Nguyen, H. D., Ito, T., Zhou, S. M., Osada, L., Tateyama, S., Kaneko, T., Takaya, N., 2013. Microbial monomers custom-synthesized to build true bio-derived aromatic polymers. Appl. Microbiol. Biotechnol. 97**,** 8887-8894.

Fujiwara, R., Noda, S., Tanaka, T., Kondo, A., 2018. Muconic acid production using gene-Level fusion proteins in *Escherichia coli*. ACS Synth. Biol. 7**,** 2698-2705.

Furuya, T., Arai, Y., Kino, K., 2012. Biotechnological production of caffeic acid by bacterial cytochrome P450 CYP199A2. Appl. Environ. Microbiol. 78**,** 6087-6094.

Furuya, T., Kino, K., 2014. Catalytic activity of the two-component flavin-dependent monooxygenase from *Pseudomonas aeruginosa* toward cinnamic acid derivatives. Appl. Microbiol. Biotechnol. 98**,** 1145-1154.

Galanie, S., Smolke, C. D., 2015. Optimization of yeast-based production of medicinal protoberberine alkaloids. Microb. Cell Fact. 14**,** 144.

Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., Smolke, C. D., 2015. Complete biosynthesis of opioids in yeast. Science. 349**,** 1095-100.

Ganesan, V., Li, Z., Wang, X., Zhang, H., 2017. Heterologous biosynthesis of natural product naringenin by co-culture engineering. Synth. Syst. Biotechnol. 2**,** 236-242.

Gao, M., Cao, M., Suastegui, M., Walker, J., Rodriguez Quiroz, N., Wu, Y., Tribby, D., Okerlund, A., Stanley, L., Shanks, J. V., Shao, Z., 2017. Innovating a nonconventional yeast platform for producing shikimate as the building block of high-value aromatics. ACS Synth. Biol. 6**,** 29-38.

Germann, S. M., Jacobsen, S. A. B., Schneider, K., Harrison, S. J., Jensen, N. B., Chen, X., Stahlhut, S. G., Borodina, I., Luo, H., Zhu, J. F., Maury, J., Forster, J., 2016. Glucose-based microbial production of the hormone melatonin in yeast *Saccharomyces cerevisiae*. Biotechnol. J. 11**,** 717-724.

Gold, N. D., Gowen, C. M., Lussier, F.-X., Cautha, S. C., Mahadevan, R., Martin, V. J. J., 2015. Metabolic engineering of a tyrosine-overproducing yeast platform using targeted metabolomics. Microb. Cell Fact. 14**,** 73.

Gottardi, M., Knudsen, J. D., Prado, L., Oreb, M., Branduardi, P., Boles, E., 2017. *De novo* biosynthesis of *trans*-cinnamic acid derivatives in *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 101**,** 4883-4893.

Han, G. H., Bang, S. E., Babu, B. K., Chang, M., Shin, H. J., Kim, S. W., 2011. Bio-indigo production in two different fermentation systems using recombinant *Escherichia coli* cells harboring a flavin-containing monooxygenase gene (*fmo*). Process Biochem. 46**,** 788-791.

Han, G. H., Gim, G. H., Kim, W., Seo, S. I., Kim, S. W., 2013. Enhanced indirubin production in recombinant *Escherichia coli* harboring a flavin-containing monooxygenase gene by cysteine supplementation. J. Biotechnol. 164**,** 179-187.

Hansen, E. H., Moller, B. L., Kock, G. R., Bunner, C. M., Kristensen, C., Jensen, O. R., Okkels, F. T., Olsen, C. E., Motawia, M. S., Hansen, J., 2009. *De novo* biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). Appl. Environ. Microbiol. 75**,** 2765-74.

Hara, R., Kino, K., 2013. Enhanced synthesis of 5-hydroxy-l-tryptophan through tetrahydropterin regeneration. AMB Express. 3**,** 70.

Hardeland, R., 2008. Melatonin, hormone of darkness and more: occurrence, control mechanisms, actions and bioactive metabolites. Cell Mol. Life Sci. 65**,** 2001-18.

Hawkins, K. M., Smolke, C. D., 2008. Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. Nat. Chem. Biol. 4**,** 564-73.

Herrmann, K. M., 1995. The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. Plant Cell. 7**,** 907-919.

Hinnebusch, A. G., 1988. Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. Microbiol. Rev. 52**,** 248-73.

Hoessel, R., Leclerc, S., Endicott, J. A., Nobel, M. E., Lawrie, A., Tunnah, P., Leost, M., Damiens, E., Marie, D., Marko, D., Niederberger, E., Tang, W., Eisenbrand, G., Meijer, L., 1999. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. Nat Cell Biol. 1**,** 60-7.

Huang, Q., Lin, Y., Yan, Y., 2013. Caffeic acid production enhancement by engineering a phenylalanine over-producing *Escherichia coli* strain. Biotechnol. Bioeng. 110**,** 3188-96.

Huccetogullari, D., Luo, Z. W., Lee, S. Y., 2019. Metabolic engineering of microorganisms for production of aromatic compounds. Microb. Cell Fact. 18**,** 41.

Hung, G. C., Brown, C. R., Wolfe, A. B., Liu, J., Chiang, H. L., 2004. Degradation of the gluconeogenic enzymes fructose-1,6-bisphosphatase and malate dehydrogenase is mediated by distinct proteolytic pathways and signaling events. J. Biol. Chem. 279**,** 49138-50.

Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., Davies, D. R., 1988. Three-dimensional structure of the tryptophan synthase α2β2 multienzyme complex from *Salmonella typhimurium*. J. Biol. Chem. 263**,** 17857-71.

Ikram-ul-Haq, Ali, S., Qadeer, M. A., 2002. Biosynthesis of L-DOPA by *Aspergillus oryzae*. Bioresour. Technol. 85**,** 25-29.

Jeeves, M., Evans, P. D., Parslow, R. A., Jaseja, M., Hyde, E. I., 1999. Studies of the *Escherichia coli* Trp repressor binding to its five operators and to variant operator sequences. Eur. J. Biochem. 265**,** 919-28.

Jiang, M., Stephanopoulos, G., Pfeifer, B. A., 2012. Downstream reactions and engineering in the microbially reconstituted pathway for Taxol. Appl. Microbiol. Biotechnol. 94**,** 841-9.

Johnson, C. W., Salvachua, D., Khanna, P., Smith, H., Peterson, D. J., Beckham, G. T., 2016. Enhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity. Metab. Eng. Commun. 3**,** 111-119.

Jung, S. T., Lauchli, R., Arnold, F. H., 2011. Cytochrome P450: taming a wild type enzyme. Curr. Opin. Biotechnol. 22**,** 809-17.

Kaneko, T., Thi, T. H., Shi, D. J., Akashi, M., 2006. Environmentally degradable, high-performance thermoplastics from phenolic phytomonomers. Nat. Mater. 5**,** 966-70.

Kang, S. Y., Choi, O., Lee, J. K., Hwang, B. Y., Uhm, T. B., Hong, Y. S., 2012. Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain. Microb. Cell Fact. 11.

Kim, B., Cho, B. R., Hahn, J. S., 2014a. Metabolic engineering of *Saccharomyces cerevisiae* for the production of 2-phenylethanol *via* Ehrlich pathway. Biotechnol. Bioeng. 111**,** 115-124.

Kim, M. J., Kim, B. G., Ahn, J. H., 2013. Biosynthesis of bioactive O-methylated flavonoids in *Escherichia coli*. Appl. Microbiol. Biotechnol. 97**,** 7195-204.

Kim, T. Y., Lee, S. W., Oh, M. K., 2014b. Biosynthesis of 2-phenylethanol from glucose with genetically engineered *Kluyveromyces marxianus*. Enzyme Microb. Technol. 61-62**,** 44-47.

Knochel, T., Ivens, A., Hester, G., Gonzalez, A., Bauerle, R., Wilmanns, M., Kirschner, K., Jansonius, J. N., 1999. The crystal structure of anthranilate synthase from Sulfolobus solfataricus: functional implications. Proc. Natl. Acad. Sci. U. S. A. 96**,** 9479-84.

Kogure, T., Kubota, T., Suda, M., Hiraga, K., Inui, M., 2016. Metabolic engineering of *Corynebacterium glutamicum* for shikimate overproduction by growth-arrested cell reaction. Metab. Eng. 38**,** 204-216.

Koma, D., Yamanaka, H., Moriyoshi, K., Sakai, K., Masuda, T., Sato, Y., Toida, K., Ohmoto, T., 2014. Production of *p*-aminobenzoic acid by metabolically engineered *Escherichia coli*. Biosci. Biotechnol. Biochem. 78**,** 350-7.

Kong, J. O., Lee, S. M., Moon, Y. S., Lee, S. G., Ahn, Y. J., 2007. Nematicidal activity of cassia and cinnamon oil compounds and related compounds toward *Bursaphelenchus xylophilus* (Nematoda: Parasitaphelenchidae). J. Nematol. 39**,** 31-6.

Koopman, F., Beekwilder, J., Crimi, B., van Houwelingen, A., Hall, R. D., Bosch, D., van Maris, A. J., Pronk, J. T., Daran, J. M., 2012. *De novo* production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. Microb Cell Fact. 11**,** 155.

Kos, A., Kuijvenhoven, J., Wernars, K., Bos, C. J., van den Broek, H. W., Pouwels, P. H., van den Hondel, C. A., 1985. Isolation and characterization of the *Aspergillus niger trpC* gene. Gene. 39**,** 231-8.

Krishnaveni, R., Rathod, V., Thakur, M. S., Neelgund, Y. F., 2009. Transformation of L-tyrosine to L-dopa by a novel fungus, *Acremonium rutilum*, under submerged fermentation. Curr. Microbiol. 58**,** 122-128.

Kunzler, M., Paravicini, G., Egli, C. M., Irniger, S., Braus, G. H., 1992. Cloning, primary structure and regulation of the *ARO4* gene, encoding the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae*. Gene. 113**,** 67-74.

Lam, F. F., Yeung, J. H., Chan, K. M., Or, P. M., 2007. Relaxant effects of danshen aqueous extract and its constituent danshensu on rat coronary artery are mediated by inhibition of calcium channels. Vascul. Pharmacol. 46**,** 271-7.

Landick, R., Carey, J., Yanofsky, C., 1987. Detection of transcription-pausing in vivo in the *trp* operon leader region. Proc. Natl. Acad. Sci. U. S. A. 84**,** 1507-11.

Landick, R., Jr., T. C., 1992. Transcriptional attenuation. Cold Spring Harbor Laboratory, Cold Spring, Harbor, NY**.,** 407-446.

Leavitt, J. M., Wagner, J. M., Tu, C. C., Tong, A., Liu, Y., Alper, H. S., 2017. Biosensor-enabled directed evolution to improve muconic acid production in *Saccharomyces cerevisiae*. Biotechnol. J. 12.

Lee, H. N., Shin, W. S., Seo, S. Y., Choi, S. S., Song, J. S., Kim, J. Y., Park, J. H., Lee, D., Kim, S. Y., Lee, S. J., Chun, G. T., Kim, E. S., 2018. *Corynebacterium* cell factory design and culture process optimization for muconic acid biosynthesis. Sci. Rep. 8**,** 18041.

Lee, K., Hahn, J. S., 2013. Interplay of Aro80 and GATA activators in regulation of genes for catabolism of aromatic amino acids in *Saccharomyces cerevisiae*. Mol. Microbiol. 88**,** 1120-34.

Lee, M. Y., Hung, W. P., Tsai, S. H., 2017. Improvement of shikimic acid production in *Escherichia coli* with growth phase-dependent regulation in the biosynthetic pathway from glycerol. World J. Microbiol. Biotechnol. 33**,** 25.

Lee, S. G., Ro, H. S., Hong, S. P., Kim, E. H., Sung, M. H., 1996. Production of L-DOPA by thermostable tyrosine phenol-lyase of a thermophilic *Symbiobacterium* species overexpressed in recombinant *Escherichia coli*. J. Microbiol. Biotechnol. 6**,** 98-102.

Leonardo-Mendonca, R. C., Martinez-Nicolas, A., de Teresa Galvan, C., Ocana-Wilhelmi, J., Rusanova, I., Guerra-Hernandez, E., Escames, G., Acuna-Castroviejo, D., 2015. The benefits of four weeks of melatonin treatment on circadian patterns in resistance-trained athletes. Chronobiol. Int. 32**,** 1125-34.

Li, C., Li, J., Wang, G., Li, X., 2016. Heterologous biosynthesis of artemisinic acid in *Saccharomyces cerevisiae*. J. Appl. Microbiol. 120**,** 1466-78.

Li, G., Young, K. D., 2013. Indole production by the tryptophanase TnaA in *Escherichia coli* is determined by the amount of exogenous tryptophan. Microbiology. 159**,** 402-410.

Li, M., Kildegaard, K. R., Chen, Y., Rodriguez, A., Borodina, I., Nielsen, J., 2015. *De novo* production of resveratrol from glucose or ethanol by engineered *Saccharomyces cerevisiae*. Metab. Eng. 32**,** 1-11.

Li, X. L., Chen, Z. Y., Wu, Y. F., Yan, Y. J., Sun, X. X., Yuan, Q. P., 2018a. Establishing an artificial pathway for efficient biosynthesis of hydroxytyrosol. ACS Synth. Biol. 7**,** 647-654.

Li, X. L., Shen, X. L., Wang, J., Ri, H. I., Mi, C. Y., Yan, Y. J., Sun, X. X., Yuan, Q. P., 2019. Efficient biosynthesis of 3, 4-dihydroxyphenylacetic acid in *Escherichia coli*. J. Biotechnol. 294**,** 14-18.

Li, Y., Li, S., Thodey, K., Trenchard, I., Cravens, A., Smolke, C. D., 2018b. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. Proc. Natl. Acad. Sci. U. S. A. 115**,** E3922-E3931.

Lin, Y., Sun, X., Yuan, Q., Yan, Y., 2014a. Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in *Escherichia coli*. Metab. Eng. 23**,** 62-9.

Lin, Y. H., Sun, X. X., Yuan, Q. P., Yan, Y. J., 2014b. Engineering bacterial phenylalanine 4-hydroxylase for microbial synthesis of human neurotransmitter precursor 5-hydroxytryptophan. ACS Synth. Biol. 3**,** 497-505.

Lin, Y. H., Yan, Y. J., 2012. Biosynthesis of caffeic acid in *Escherichia coli* using its endogenous hydroxylase complex. Microb. Cell Fact. 11.

Liu, C. Q., Zhang, K., Cao, W. Y., Zhang, G., Chen, G. Q., Yang, H. Y., Wang, Q., Liu, H. B., Xian, M., Zhang, H. B., 2018a. Genome mining of 2-phenylethanol biosynthetic genes from *Enterobacter* sp. CGMCC 5087 and heterologous overproduction in *Escherichia coli*. Biotechnol. Biofuels. 11.

Liu, S. P., Liu, R. X., El-Rotail, A. A. M. M., Ding, Z. Y., Gu, Z. H., Zhang, L., Shi, G. Y., 2014a. Heterologous pathway for the production of L-phenylglycine from glucose by *E. coli*. J. Biotechnol. 186**,** 91-97.

Liu, S. P., Liu, R. X., Xiao, M. R., Zhang, L., Ding, Z. Y., Gu, Z. H., Shi, G. Y., 2014b. A systems level engineered *E. coli* capable of efficiently producing L-phenylalanine. Process Biochem. 49**,** 751-757.

Liu, X., Li, X. B., Jiang, J. L., Liu, Z. N., Qiao, B., Li, F. F., Cheng, J. S., Sun, X. C., Yuan, Y. J., Qiao, J. J., Zhao, G. R., 2018b. Convergent engineering of syntrophic *Escherichia coli* coculture for efficient production of glycosides. Metab. Eng. 47**,** 243-253.

Lu, R., Lu, F., Chen, J., Yu, W., Huang, Q., Zhang, J., Xu, J., 2016. Production of diethyl terephthalate from biomass-derived muconic acid. Angew. Chem. Int. Ed. Engl. 55**,** 249-53.

Lu, Y., Wang, L. Y., Xue, Y., Zhang, C., Xing, X. H., Lou, K., Zhang, Z. D., Li, Y., Zhang, G. F., Bi, J. X., Su, Z. G., 2009. Production of violet pigment by a newly isolated psychrotrophic bacterium from a glacier in Xinjiang, China. Biochem. Eng. J. 43**,** 135-141.

Luo, Z. W., Kim, W. J., Lee, S. Y., 2018. Metabolic engineering of *Escherichia coli* for efficient production of 2-pyrone-4,6-dicarboxylic acid from glucose. ACS Synth. Biol. 7**,** 2296-2307.

Luttik, M. A., Vuralhan, Z., Suir, E., Braus, G. H., Pronk, J. T., Daran, J. M., 2008. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. Metab. Eng. 10**,** 141-53.

Mast, Y. J., Wohlleben, W., Schinko, E., 2011. Identification and functional characterization of phenylglycine biosynthetic genes involved in pristinamycin biosynthesis in *Streptomyces pristinaespiralis*. J Biotechnol. 155**,** 63-7.

Masuo, S., Kobayashi, Y., Oinuma, K. I., Takaya, N., 2016. Alternative fermentation pathway of cinnamic acid production *via* phenyllactic acid. Appl. Microbiol. Biotechnol. 100**,** 8701-8709.

McKenna, R., Nielsen, D. R., 2011. Styrene biosynthesis from glucose by engineered *E. coli*. Metab. Eng. 13**,** 544-554.

McKenna, R., Thompson, B., Pugh, S., Nielsen, D. R., 2014. Rational and combinatorial approaches to engineering styrene production by *Saccharomyces cerevisiae*. Microb. Cell Fact. 13.

Metzler, M. C., Rothermel, B. A., Nelson, T., 1989. Maize NADP-malate dehydrogenase: cDNA cloning, sequence, and mRNA characterization. Plant Mol Biol. 12**,** 713-22.

Michener, J. K., Nielsen, J., Smolke, C. D., 2012. Identification and treatment of heme depletion attributed to overexpression of a lineage of evolved P450 monooxygenases. Proc. Natl. Acad. Sci. U. S. A. 109**,** 19504-9.

Miles, E. W., 1979. Tryptophan synthase: structure, function, and subunit interaction. Adv. Enzymol. Relat. Areas Mol. Biol. 49**,** 127-86.

Minami, H., Kim, J. S., Ikezawa, N., Takemura, T., Katayama, T., Kumagai, H., Sato, F., 2008. Microbial production of plant benzylisoquinoline alkaloids. Proc. Natl. Acad. Sci. U. S. A. 105**,** 7393-8.

Minard, K. I., McAlister-Henn, L., 1992. Glucose-induced degradation of the MDH2 isozyme of malate dehydrogenase in yeast. J. Biol. Chem. 267**,** 17458-64.

Mora-Villalobos, J. A., Zeng, A. P., 2017. Protein and pathway engineering for the biosynthesis of 5-hydroxytryptophan in *Escherichia coli*. Eng. Life Sci. 17**,** 892-899.

Mora-Villalobos, J. A., Zeng, A. P., 2018. Synthetic pathways and processes for effective production of 5-hydroxytryptophan and serotonin from glucose in *Escherichia coli*. J. Biol. Eng. 12.

Mueller, P. P., Hinnebusch, A. G., 1986. Multiple upstream AUG codons mediate translational control of GCN4. Cell. 45**,** 201-7.

Mullaney, E. J., Hamer, J. E., Roberti, K. A., Yelton, M. M., Timberlake, W. E., 1985. Primary structure of the *trpC* gene from *Aspergillus nidulans*. Mol. Gen. Genet. 199**,** 37-45.

Muller, U., van Assema, F., Gunsior, M., Orf, S., Kremer, S., Schipper, D., Wagemans, A., Townsend, C. A., Sonke, T., Bovenberg, R., Wubbolts, M., 2006. Metabolic engineering of the *E. coli* L-phenylalanine pathway for the production of D-phenylglycine (D-Phg). Metab. Eng. 8**,** 196-208.

N., P., 2014. Global markets for flavors and fragrances. BCC Research, Wellesley, MA.

Nakagawa, A., Minami, H., Kim, J. S., Koyanagi, T., Katayama, T., Sato, F., Kumagai, H., 2011. A bacterial platform for fermentative production of plant alkaloids. Nat. Commun. 2**,** 326.

Nakashima, A., Kaneko, Y. S., Mori, K., Fujiwara, K., Tsugu, T., Suzuki, T., Nagatsu, T., Ota, A., 2002. The mutation of two amino acid residues in the N-terminus of tyrosine hydroxylase (TH) dramatically enhances the catalytic activity in neuroendocrine AtT-20 cells. J. Neurochem. 82**,** 202-6.

Niu, W., Draths, K. M., Frost, J. W., 2002. Benzene-free synthesis of adipic acid. Biotechnol. Prog. 18**,** 201-211.

Noda, S., Kitazono, E., Tanaka, T., Ogino, C., Kondo, A., 2012. Benzoic acid fermentation from starch and cellulose *via* a plant-like beta-oxidation pathway in *Streptomyces maritimus*. Microb. Cell Fact. 11.

Noda, S., Shirai, T., Oyama, S., Kondo, A., 2016. Metabolic design of a platform *Escherichia coli* strain producing various chorismate derivatives. Metab. Eng. 33**,** 119-129.

Ooi, L. S., Li, Y., Kam, S. L., Wang, H., Wong, E. Y., Ooi, V. E., 2006. Antimicrobial activities of cinnamon oil and cinnamaldehyde from the Chinese medicinal herb *Cinnamomum cassia* Blume. Am. J. Chin. Med. 34**,** 511-22.

Paddon, C. J., Keasling, J. D., 2014. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nat. Rev. Microbiol. 12**,** 355-67.

Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., Newman, J. D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 496**,** 528-32.

Palmieri, L., Vozza, A., Agrimi, G., De Marco, V., Runswick, M. J., Palmieri, F., Walker, J. E., 1999. Identification of the yeast mitochondrial transporter for oxaloacetate and sulfate. J. Biol. Chem. 274**,** 22184-90.

Pandal, N., Global markets for flavors and fragrances. BCC Research, Wellesley, MA, 2014.

Paravicini, G., Mosch, H. U., Schmidheini, T., Braus, G., 1989. The general control activator protein GCN4 is essential for a basal level of *ARO3* gene expression in *Saccharomyces cerevisiae*. Mol Cell Biol. 9**,** 144-51.

Park, H. S., Lee, J. Y., Kim, H. S., 1998. Production of L-DOPA(3,4-dihydroxyphenyl-*L*-alanine) from benzene by using a hybrid pathway. Biotechnol. Bioeng. 58**,** 339-343.

Park, S., Kang, K., Lee, S. W., Ahn, M. J., Bae, J. M., Back, K., 2011. Production of serotonin by dual expression of tryptophan decarboxylase and tryptamine 5-hydroxylase in *Escherichia coli*. Appl. Microbiol. Biotechnol. 89**,** 1387-1394.

Pittard, A. J., Davidson, B. E., 1991. TyrR protein of *Escherichia coli* and its role as repressor and activator. Mol. Microbiol. 5**,** 1585-92.

Pittard, J., Yang, J., 2008. Biosynthesis of the aromatic amino acids. EcoSal Plus. 3.

Priefert, H., Rabenhorst, J., Steinbuchel, A., 2001. Biotechnological production of vanillin. Appl. Microbiol. Biotechnol. 56**,** 296-314.

Pyne, M. E., Narcross, L., Melgar, M., Kevvai, K., Mookerjee, S., Leite, G. B., Martin, V. J. J., 2018. An engineered Aro1 protein degradation approach for increased *cis,cis*-muconic acid biosynthesis in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 84.

Qi, R. Q., Pfeifer, B. A., Zhang, G. J., 2018. Engineering heterologous production of salicylate glucoside and glycosylated variants. Front. Microbiol. 9.

Qi, W. W., Vannelli, T., Breinig, S., Ben-Bassat, A., Gatenby, A. A., Haynie, S. L., Sariaslani, F. S., 2007. Functional expression of prokaryotic and eukaryotic genes in *Escherichia coli* for conversion of glucose to *p*-hydroxystyrene. Metab. Eng. 9**,** 268-276.

Quinsey, N. S., Luong, A. Q., Dickson, P. W., 1998. Mutational analysis of substrate inhibition in tyrosine hydroxylase. J. Neurochem. 71**,** 2132-8.

Ran, N. Q., Knop, D. R., Draths, K. M., Frost, J. W., 2001. Benzene-free synthesis of hydroquinone. J. Am. Chem. Soc. 123**,** 10927-10934.

Rawat, G., Tripathi, P., Saxena, R. K., 2013. Expanding horizons of shikimic acid. Recent progresses in production and its endless frontiers in application and market trends. Appl. Microbiol. Biotechnol. 97**,** 4277-87.

Reifenrath, M., Boles, E., 2018. Engineering of hydroxymandelate synthases and the aromatic amino acid pathway enables *de novo* biosynthesis of mandelic and 4-hydroxymandelic acid with *Saccharomyces cerevisiae*. Metab. Eng. 45**,** 246-254.

Rinner, U., Hudlicky, T., 2012. Synthesis of morphine alkaloids and derivatives. Top. Curr. Chem. 309**,** 33-66.

Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C., Withers, S. T., Shiba, Y., Sarpong, R., Keasling, J. D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 440**,** 940-3.

Rodrigues, A. L., Becker, J., Lima, A. O. D. S., Porto, L. M., Wittmann, C., 2014. Systems metabolic engineering of *Escherichia coli* for gram scale production of the antitumor drug deoxyviolacein from glycerol. Biotechnol. Bioeng. 111**,** 2280-2289.

Rodrigues, A. L., Trachtmann, N., Becker, J., Lohanatha, A. F., Blotenberg, J., Bolten, C. J., Korneli, C., Lima, A. O. D., Porto, L. M., Sprenger, G. A., Wittmann, C., 2013. Systems metabolic engineering of *Escherichia coli* for production of the antitumor drugs violacein and deoxyviolacein. Metab. Eng. 20**,** 29-41.

Rodrigues, J. L., Araujo, R. G., Prather, K. L. J., Kluskens, L. D., Rodrigues, L. R., 2015. Heterologous production of caffeic acid from tyrosine in *Escherichia coli*. Enzyme Microb. Technol. 71**,** 36-44.

Rodriguez, A., Chen, Y., Khoomrung, S., Ozdemir, E., Borodina, I., Nielsen, J., 2017a. Comparison of the metabolic response to over-production of *p*-coumaric acid in two yeast strains. Metab. Eng. 44**,** 265-272.

Rodriguez, A., Kildegaard, K. R., Li, M. J., Borodina, I., Nielsen, J., 2015. Establishment of a yeast platform strain for production of *p*-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. Metab. Eng. 31**,** 181-188.

Rodriguez, A., Strucko, T., Stahlhut, S. G., Kristensen, M., Svenssen, D. K., Forster, J., Nielsen, J., Borodina, I., 2017b. Metabolic engineering of yeast for fermentative production of flavonoids. Bioresour. Technol. 245**,** 1645-1654.

Romagnoli, G., Knijnenburg, T. A., Liti, G., Louis, E. J., Pronk, J. T., Daran, J. M., 2015. Deletion of the *Saccharomyces cerevisiae ARO8* gene, encoding an aromatic amino acid transaminase, enhances phenylethanol production from glucose. Yeast. 32**,** 29-45.

Sandig, G., Kargel, E., Menzel, R., Vogel, F., Zimmer, T., Schunck, W. H., 1999. Regulation of endoplasmic reticulum biogenesis in response to cytochrome P450 overproduction. Drug. Metab. Rev. 31**,** 393-410.

Santos, C. N., Koffas, M., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. Metab. Eng. 13**,** 392-400.

Schechtman, M. G., Yanofsky, C., 1983. Structure of the trifunctional *trp-1* gene from *Neurospora crassa* and its aberrant expression in *Escherichia coli*. J. Mol. Appl. Genet. 2**,** 83-99.

Schenkman, J. B., Jansson, I., 2003. The many roles of cytochrome b5. Pharmacol. Ther. 97**,** 139-52.

Sengupta, S., Jonnalagadda, S., Goonewardena, L., Juturu, V., 2015. Metabolic engineering of a novel muconic acid biosynthesis pathway *via* 4-hydroxybenzoic acid in *Escherichia coli*. Appl. Environ. Microbiol. 81**,** 8037-8043.

Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., Liao, J. C., 2011. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. Appl. Environ. Microbiol. 77**,** 2905.

Shen, X. L., Wang, J., Wang, J., Chen, Z. Y., Yuan, Q. P., Yan, Y. J., 2017. High-level *de novo* biosynthesis of arbutin in engineered *Escherichia coli*. Metab. Eng. 42**,** 52-58.

Skjoedt, M. L., Snoek, T., Jensen, M. K., Keasling, J. D., 2016. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. Nat. Chem. Biol. 12(11): 951-958.

Sova, M., 2012. Antioxidant and antimicrobial activities of cinnamic acid derivatives. Mini Rev. Med. Chem. 12**,** 749-67.

Spraggon, G., Kim, C., Nguyen-Huu, X., Yee, M. C., Yanofsky, C., Mills, S. E., 2001. The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. Proc. Natl. Acad. Sci. U. S. A. 98**,** 6021-6.

Stark, D., Munch, T., Sonnleitner, B., Marison, I. W., von Stockar, U., 2002. Extractive bioconversion of 2-phenylethanol from L-phenylalanine by *Saccharomyces cerevisiae*. Biotechnol. Prog. 18**,** 514-523.

Stark, M. J., Milner, J. S., 1989. Cloning and analysis of the *Kluyveromyces lactis TRP1* gene: a chromosomal locus flanked by genes encoding inorganic pyrophosphatase and histone H3. Yeast. 5**,** 35-50.

Struhl, K., 1986. Constitutive and inducible Saccharomyces cerevisiae promoters: evidence for two distinct molecular mechanisms. Mol. Cell. Biol. 6**,** 3847-53.

Suastegui, M., Guo, W., Feng, X., Shao, Z., 2016a. Investigating strain dependency in the production of aromatic compounds in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 113**,** 2676-2685.

Suastegui, M., Matthiesen, J. E., Carraher, J. M., Hernandez, N., Rodriguez Quiroz, N., Okerlund, A., Cochran, E. W., Shao, Z., Tessonnier, J. P., 2016b. Combining metabolic engineering and electrocatalysis: application to the production of polyamides from sugar. Angew. Chem. Int. Ed. Engl. 55**,** 2368-73.

Suastegui, M., Shao, Z., 2016. Yeast factories for the production of aromatic compounds: from building blocks to plant secondary metabolites. J. Ind. Microbiol. Biotechnol. 43**,** 1611-1624.

Suastegui, M., Yu Ng, C., Chowdhury, A., Sun, W., Cao, M., House, E., Maranas, C. D., Shao, Z., 2017. Multilevel engineering of the upstream module of aromatic amino acid biosynthesis in *Saccharomyces cerevisiae* for high production of polymer and drug precursors. Metab. Eng. 42**,** 134-144.

Sun, H. N., Zhao, D. D., Xiong, B., Zhang, C. Z., Bi, C. H., 2016a. Engineering *Corynebacterium glutamicum* for violacein hyper production. Microb. Cell Fact. 15.

Sun, J., Lin, Y. H., Shen, X. L., Jain, R., Sun, X. X., Yuan, Q. P., Yan, Y. J., 2016b. Aerobic biosynthesis of hydrocinnamic acids in *Escherichia coli* with a strictly oxygen-sensitive enoate reductase. Metab. Eng. 35**,** 75-82.

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

Sun, X. X., Lin, Y. H., Yuan, Q. P., Yan, Y. J., 2014. Biological production of muconic acid *via* a prokaryotic 2,3-dihydroxybenzoic acid decarboxylase. Chemsuschem. 7**,** 2478-2481.

Sun, X. X., Lin, Y. H., Yuan, Q. P., Yan, Y. J., 2015. Precursor-directed biosynthesis of 5-hydroxytryptophan using metabolically engineered *E. coli*. ACS Synth. Biol. 4**,** 554-558.

Sun, Z. T., Ning, Y. Y., Liu, L. X., Liu, Y. M., Sun, B. B., Jiang, W. H., Yang, C., Yang, S., 2011. Metabolic engineering of the L-phenylalanine pathway in *Escherichia coli* for the production of *S*- or *R*-mandelic acid. Microb. Cell Fact. 10.

Sunder Rangachari, T. C. F., Greg Hartmann, Robert B. Weisenfeld, Processes for producing and recovering shikimic acid. Monsanto Technology Llc, US, 2013.

Swaney, D. L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., Krogan, N. J., Villen, J., 2013. Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. Nat. Methods. 10**,** 676-82.

Thireos, G., Penn, M. D., Greer, H., 1984. 5' untranslated sequences are required for the translational control of a yeast regulatory gene. Proc. Natl. Acad. Sci. U. S. A. 81**,** 5096-100.

Thodey, K., Galanie, S., Smolke, C. D., 2014. A microbial biomanufacturing platform for natural and semisynthetic opioids. Nat. Chem. Biol. 10**,** 837-44.

Thompson, B., Machas, M., Nielsen, D. R., 2015. Creating pathways towards aromatic building blocks and fine chemicals. Curr. Opin. Biotechnol. 36**,** 1-7.

Thompson, H., Tersteegen, A., Thauer, R. K., Hedderich, R., 1998. Two malate dehydrogenases in *Methanobacterium thermoautotrophicum*. Arch. Microbiol. 170**,** 38-42.

Trenchard, I. J., Siddiqui, M. S., Thodey, K., Smolke, C. D., 2015. *De novo* production of the key branch point benzylisoquinoline alkaloid reticuline in yeast. Metab. Eng. 31**,** 74-83.

Trenchard, I. J., Smolke, C. D., 2015. Engineering strategies for the fermentative production of plant alkaloids in yeast. Metab. Eng. 30**,** 96-104.

Tribe, D. E., Camakaris, H., Pittard, J., 1976. Constitutive and repressivle enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. J. Bacteriol. 127**,** 1085-97.

Vannelli, T., Qi, W. W., Sweigard, J., Gatenby, A. A., Sariaslani, F. S., 2007. Production of *p*-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. Metab. Eng. 9**,** 142-151.

Verhoef, S., Wierckx, N., Westerhof, R. G. M., de Winde, J. H., Ruijssenaars, H. J., 2009. Bioproduction of *p*-hydroxystyrene from glucose by the solvent-tolerant bacterium *Pseudomonas putida* S12 in a two-phase water-decanol fermentation. Appl. Environ. Microbiol. 75**,** 931-936.

W., F. L., K., S. V., G., W. C., 1999. Molecular model of phenolic polymer dissolution in photolithography. J. Polym. Sci. B. 37**,** 2103-2113.

Wang, H. J., Liu, W. Q., Shi, F., Huang, L., Lian, J. Z., Qu, L., Cai, J., Xu, Z. A., 2018a. Metabolic pathway engineering for high-level production of 5-hydroxytryptophan in *Escherichia coli*. Metab. Eng. 48**,** 279-287.

Wang, H. S., Jiang, P. X., Lu, Y., Ruan, Z. Y., Jiang, R. B., Xing, X. H., Lou, K., Wei, D., 2009. Optimization of culture conditions for violacein production by a new strain of *Duganella* sp B2. Biochem. Eng. J. 44**,** 119-124.

Wang, J., Shen, X. L., Yuan, Q. P., Yan, Y. J., 2018b. Microbial synthesis of pyrogallol using genetically engineered *Escherichia coli*. Metab. Eng. 45**,** 134-141.

Wang, J., Zheng, P., 2015. Muconic acid production from glucose using enterobactin precursors in *Escherichia coli*. J. Ind. Microbiol. Biotechnol. 42**,** 701-709.

Wang, S., Zhang, S., Xiao, A., Rasmussen, M., Skidmore, C., Zhan, J., 2015. Metabolic engineering of *Escherichia coli* for the biosynthesis of various phenylpropanoid derivatives. Metab. Eng. 29**,** 153-159.

Wang, S. W., Fu, C., Bilal, M., Hu, H. B., Wang, W., Zhang, X. H., 2018c. Enhanced biosynthesis of arbutin by engineering shikimate pathway in *Pseudomonas chlororaphis* P3. Microb. Cell Fact. 17.

Ward, M., Yu, B., Wyatt, V., Griffith, J., Craft, T., Neurath, A. R., Strick, N., Li, Y. Y., Wertz, D. L., Pojman, J. A., Lowe, A. B., 2007. Anti-HIV-1 activity of poly(mandelic acid) derivatives. Biomacromolecules. 8**,** 3308-16.

Weber, C., Bruckner, C., Weinreb, S., Lehr, C., Essl, C., Boles, E., 2012. Biosynthesis of *cis,cis*-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 78**,** 8421-30.

Weber, H. E., Gottardi, M., Bruckner, C., Oreb, M., Boles, E., Tripp, J., 2017. Requirement of a functional flavin mononucleotide prenyltransferase for the activity of a bacterial decarboxylase in a heterologous muconic acid pathway in *Saccharomyces cerevisiae*. Appl Environ Microbiol. 83.

Wei, T., Cheng, B. Y., Liu, J. Z., 2016. Genome engineering *Escherichia coli* for L-DOPA overproduction from glucose. Sci. Rep. 6.

Westfall, P. J., Pitera, D. J., Lenihan, J. R., Eng, D., Woolard, F. X., Regentin, R., Horning, T., Tsuruta, H., Melis, D. J., Owens, A., Fickes, S., Diola, D., Benjamin, K. R., Keasling, J. D., Leavell, M. D., McPhee, D. J., Renninger, N. S., Newman, J. D., Paddon, C. J., 2012. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. Proc. Natl. Acad. Sci. U. S. A. 109**,** E111-8.

Winter, J. M., Tang, Y., 2012. Synthetic biological approaches to natural product biosynthesis. Curr. Opin. Biotechnol. 23**,** 736-43.

Wu, J., Du, G., Chen, J., Zhou, J., 2015. Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. Sci Rep. 5**,** 13477.

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

Wu, J., Zhang, X., Zhu, Y., Tan, Q., He, J., Dong, M., 2017. Rational modular design of metabolic network for efficient production of plant polyphenol pinosylvin. Sci. Rep. 7**,** 1459.

Xu, J., Wang, L. L., Dammer, E. B., Li, C. B., Xu, G., Chen, S. D., Wang, G., 2015. Melatonin for sleep disorders and cognition in dementia: a meta-analysis of randomized controlled trials. Am. J. Alzheimers Dis. Other Demen. 30**,** 439-47.

Xue, Y. X., Chen, X. Z., Yang, C., Chang, J. Z., Shen, W., Fan, Y., 2017. Engineering *Eschericha coli* for enhanced tyrosol production. J. Agric. Food Chem. 65**,** 4708-4714.

Yamamoto, K., Kataoka, E., Miyamoto, N., Furukawa, K., Ohsuye, K., Yabuta, M., 2003. Genetic engineering of *Escherichia coli* for production of tetrahydrobiopterin. Metab. Eng. 5**,** 246-54.

Yang, C., Jiang, P. X., Xiao, S., Zhang, C., Lou, K., Xing, X. H., 2011. Fed-batch fermentation of recombinant *Citrobacter freundii* with expression of a violacein-synthesizing gene cluster for efficient violacein production from glycerol. Biochem. Eng. J. 57**,** 55-62.

Yanofsky, C., 2005. The favorable features of tryptophan synthase for proving Beadle and Tatum's one gene-one enzyme hypothesis. Genetics. 169**,** 511-6.

Yao, R., Pan, K., Peng, H., Feng, L., Hu, H., Zhang, X., 2018. Engineering and systems-level analysis of *Pseudomonas chlororaphis* for production of phenazine-1-carboxamide using glycerol as the cost-effective carbon source. Biotechnol. Biofuels. 11**,** 130.

Yao, Y. F., Wang, C. S., Qiao, J. J., Zhao, G. R., 2013. Metabolic engineering of *Escherichia coli* for production of salvianic acid A *via* an artificial biosynthetic pathway. Metab. Eng. 19**,** 79-87.

Yoshida, H., Tanaka, Y., Nakayama, K., 1973. Production of 3,4-dihydroxyphenyl-L-alanine (L-Dopa) by and its drivatives by *Vibrio tyrosinaticus*. Agric. Biol. Chem. 37**,** 2121-2126.

Yoshida, H., Tanaka, Y., Nakayama, K., 1974. Production of 3,4-dihydroxyphenyl-L-alanine (L-Dopa) by *Pseudomonas melanogenum*. Agr. Biol. Chem. 38**,** 455-462.

Yu, S. Q., Plan, M. R., Winter, G., Kromer, J. O., 2016. Metabolic engineering of *Pseudomonas putida* KT2440 for the production of *para*-hydroxy benzoic acid. Front. Bioeng. Biotechnol. 4.

Zalkin, H., Paluh, J. L., van Cleemput, M., Moye, W. S., Yanofsky, C., 1984. Nucleotide sequence of *Saccharomyces cerevisiae* genes TRP2 and TRP3 encoding bifunctional anthranilate synthase: indole-3-glycerol phosphate synthase. J. Biol. Chem. 259**,** 3985-92.

Zangar, R. C., Davydov, D. R., Verma, S., 2004. Mechanisms that regulate production of reactive oxygen species by cytochrome P450. Toxicol. Appl. Pharmacol. 199**,** 316-31.

Zelle, R. M., de Hulster, E., van Winden, W. A., de Waard, P., Dijkema, C., Winkler, A. A., Geertman, J. M., van Dijken, J. P., Pronk, J. T., van Maris, A. J., 2008. Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. Appl. Environ. Microbiol. 74**,** 2766-77.

Zhang, H., Im, S. C., Waskell, L., 2007. Cytochrome b5 increases the rate of product formation by cytochrome P450 2B4 and competes with cytochrome P450 reductase for a binding site on cytochrome P450 2B4. J. Biol. Chem. 282**,** 29766-76.

Zhang, H., Li, Z., Pereira, B., Stephanopoulos, G., 2015a. Engineering *E. coli-E. coli* cocultures for production of muconic acid from glycerol. Microb. Cell Fact. 14**,** 134.

Zhang, H., Pereira, B., Li, Z., Stephanopoulos, G., 2015b. Engineering *Escherichia coli* coculture systems for the production of biochemical products. Proc. Natl. Acad. Sci. U. S. A. 112**,** 8266-71.

Zhang, H. R., Stephanopoulos, G., 2013. Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars. Appl. Microbiol. Biotechnol. 97**,** 3333-3341.

Zhang, J. T., Wu, C. C., Sheng, J. Y., Feng, X. Y., 2016. Molecular basis of 5-hydroxytryptophan synthesis in *Saccharomyces cerevisiae*. Mol. Biosyst. 12**,** 1432-1435.

Zhang, R., Li, C., Wang, J., Yang, Y., Yan, Y., 2018. Microbial production of small medicinal molecules and biologics: From nature to synthetic pathways. Biotechnol. Adv. 36**,** 2219-2231.

Zhang, X. W., Qu, Y. Y., Ma, Q., Kong, C. L., Zhou, H., Cao, X. Y., Shen, W. L., Shen, E., Zhou, J. T., 2014. Production of indirubin from tryptophan by recombinant *Escherichia coli* containing naphthalene dioxygenase genes from *Comamonas* sp. MQ. Appl. Biochem. Biotechnol. 172**,** 3194-3206.

Zhou, L., Ding, Q., Jiang, G. Z., Liu, Z. N., Wang, H. Y., Zhao, G. R., 2017. Chromosome engineering of *Escherichia coli* for constitutive production of salvianic acid A. Microb. Cell Fact. 16.

Zimmer, T., Vogel, F., Ohta, A., Takagi, M., Schunck, W. H., 1997. Protein quality--a determinant of the intracellular fate of membrane-bound cytochromes P450 in yeast. DNA Cell Biol. 16**,** 501-14.

Table 1. The TyrR-mediated transcriptional regulations of L-Tyr regulon genes.

Gene	TyrR Box	Mechanisms and Special Features
tyrR (tyrosine pathway regulator)	A strong box and a weak box with the strong box overlapping with the -35 position.	1. The protein has an N-terminal domain, a central domain, and a C-terminal domain. 2. The presence of effector molecules and the positions of the TyrR-binding sites in a promoter determine the oligomerization states of TyrR, and its consequent role of being a repressor or an activator.
tyrB (tyrosine aminotransferase) *aroP* (aromatic amino acid transporter) *aroL* (shikimate kinase)	A strong box and a nearby weak box found immediately downstream of the transcription initiation site	1. L-Tyr-mediated hexamer binding (strong repression) or L-Phe-mediated dimer binding (weaker repression) prevents active transcription. 2. For *aroP*, another promoter upstream but in the opposite direction, occupies RNA polymerase and further lowers the availability of RNA polymerase to bind to the promoter of *aroP*. 3. For *aroL*, a TrpR binding site appears immediately downstream of the TyrR box. L-Trp-mediated repression requires TyrR binding to the TyrR box first, suggesting the potential spatial interaction of TrpR and TyrR. Besides, another strong box exists in the upstream region of the promoter, suggesting the possibility of TyrR-mediated DNA looping during oligomerization.
aroF (DAHP synthease)	1. A strong box and a nearby weak box found upstream with the weak box overlapping with the -35 sequence. 2. Another strong box exists in the further upstream region of the promoter.	1. L-Tyr-mediated repression conducted through TyrR hexamer binding to the weak box. 2. L-Phe-mediated repression was also observed when TyrR level is increased.
aroG (DAHP synthease)	A single strong box in the -35 position	1. Competition between TyrR and RNA polymerase leads to repression even in the absence of aromatic amino acids.
tyrP (Tyr-specific transporter)	A strong box and a nearby weak box found upstream with the weak box overlapping with the -35 sequence.	1. L-Tyr-mediated repression *via* TyrR hexamer binding to the weak box. 2. L-Phe (or L-Trp, to a lesser extent)-mediated activation conducted through TyrR dimer binding to the strong box and interacting with the α-subunit of RNA polymerase bound to the -35 position in the presence of L-Phe (or L-Trp) and absence of L-Tyr.
mtr (Trp-specific transporter)	1. A strong box and a weak box found upstream with the strong box close to the -35 position 2. A TrpR binding site overlaps with the -35 position.	1. L-Tyr (or L-Phe)-mediated activation via the interaction between the N-terminal domain of TyrR and RNA polymerase. 2. L-Trp-mediated repression (no longer need to transport L-Trp into the cells).
folA (dihydrofolate reductase)	A strong box and a weak box found in the far end upstream	1. Activation by TyrR in the presence of L-Tyr. 2. Activation requires all three domains of TyrR to be fully functional.

Table 2. The microbial production of the upstream products from the shikimate pathway.

Node	Target Compound	Host	Titer (g/L)	Yield (mg/g)	Productivity (mg/L/h)	Fermentation type	Unique strategies and genotype	Reference
Dehydroshikimate	Shikimic acid	*E. coli*	87	320	1640	Fed batch (53 h)	*Zm-glf, Zm-glk, Ec-aroF^{fbr}, Ec-tktA, Ec-aroE, Ec-serA,* and *ptsΔ*	Chandran et al., 2003
		E. coli	5.33	213	222	Shake flask (24 h)	Growth phase-dependent promoter for *aroK*; the plant-sourced DHQ-SDH bi-functional enzyme facilitated DHS channeling; *Ec-aroB, Ec-aroG^{fbr}, Ec-ppsA,* and *Ec-tktA*	Lee et al., 2017
		C. glutamicum	141	490	2938	Fed batch (48 h)	*Cg-tkt, Cg-tal, Ec-aroG^{S180F}, Cg-aroBDE, Cg-iolT1, Cg-glks (glk1, glk2, and ppsk), Cg-gapA, qsuDΔ, qsuBΔ, aroKΔ, ptsHΔ,* and *hdpAΔ*; Using growth-arrested high density cells	Kogure et al., 2016
		S. cerevisiae	1.98 (glucose) 2.5 (sucrose)	54.4 (glucose) 62.5 (sucrose)	27.8 (glucose) 33.3 (sucrose)	Shake flask (72 h)	*Sc-ARO1^{D920A}, Sc-ARO4^{K229L}, Sc-TKL, ScRKI1, ric1Δ,* and *aro1Δ*	Suastegui et al., 2017
		S. stipitis	3.11	77.8	25.9	Shake flask (120 h)	*S. stipitis* has native and superior xylose utilization capacity; *Ss-ARO4^{K220L}, Ss-TKT1,* and *Ss-ARO1^{D900A}*	Gao et al., 2017
	Quinic acid	*E. coli*	49	210	1020	Fed batch (48 h)	*Ec-aroE* and *aroDΔ*	Ran et al., 2001
		E. coli	4.8	333	66.7	Shake flask (72 h)	*Kp-qad* and *aroDΔ*	Draths et al., 1992
	Muconic acid	*E. coli*	36.8	174	767	Fed batch	*Kp-aroY, Kp-aroZ, Ac-catA, Ec-tktA, Ec-aroF^{fbr}, Ec-aroB,* and *aroEΔ*	Niu et al., 2002
		P. putida	4.92 (glucose) 15.59 (*p*-coumarate)	61 (glucose) 797 (*p*-coumarate	90 (glucose) 210 (*p*-coumarate)	Fed batch (54 h, glucose) (73 h, *p*-coumarate)	*Pp-catA, Ecl-aroY, Ecl-ecdB, Bc-asbF, catRBCΔ* and *pcaHGΔ;* The strain fermented on *p*-coumarate also carries *Ecl-ecdD*, another AroY-associated protein.	Johnson et al., 2016
		C. glutamicum	54	197	340	Fed batch (168 h)	*Kp-aroY, Kp-kpdBD, catBΔ, aroEΔ,* and *pcaHGΔ*	Lee et al., 2018
		E. coli coculture	4.7	350 (mixed glu/xyl)	65.3	Fed batch (72 h)	strain 1: *ptsHΔ, ptsIΔ, crrΔ, aroEΔ, ydiBΔ*; strain 2: *Kp-aroY, Kp-aroZ, Ac-catA, Ec-shiA,* and *xylAΔ*	Zhang et al., 2015a
		S. cerevisiae	0.32 (2.362 MA+PCA)	8 (59 MA+PCA)	4.4 (32.8 MA+PCA)	Shake flask (72 h)	*Kp-aroZ, Kp-aroY, Ca-HQD2, Sc-ARO1$^{D1409A, D920A}$, Sc-ARO4^{K229L}, Sc-TKL, Sc-RKI1, ric1Δ,* and *aro1Δ*	Suastegui et al., 2017
		S. cerevisiae	0.141	3.91	1.31	Shake flask (108 h)	*Sc-ARO4^{K229L}, Sc-TKL, Pa-aroZ, Ecl-aroY, Ca-HDQ2, aro3Δ,* and *zwf1Δ*	Curran et al., 2012

Table 2. The microbial production of the upstream products from the shikimate pathway (continued).

Node	Target Compound	Host	Titer (g/L)	Yield (mg/g)	Productivity (mg/L/h)	Fermentation type	Unique strategies and genotype	Reference
Dehydroshikimate	Muconic acid	*S. cerevisiae*	2.1	12.9	9	Fed batch (240 h)	Transcription factor-enabled aromatic amino acid biosensor and anti-metabolite-driven adaptive laboratory evolution; *Sc-ARO1* and *Sc-PAD1*	Leavitt et al., 2017
		S. cerevisiae	1.2 or 5.1 (with amino acid supplement)	13 (58)	7.14 (30.4)	Fed batch (168 h)	Coupling MA pathway to degron-tagged *ARO1*; *Sc-ARO1-CLN2$_{PEST}$*, *Sc-PAD1*, *Ca-HQD2*, *Sc-ARO4^{fbr}*, *Ec-aroB*, *Ec-aroD*, *Pa-aroZ*, *Kp-aroY*, *aro3*Δ, and *arp4*Δ	Pyne et al., 2018
	2-Pyrone-4,6-dicarboxylic acid	*E. coli*	16.72	201	172	Fed batch (96 h)	Enhancing DHS by overexpressing an importer, and manipulating NADPH/NADP$^+$ by modifying soluble pyridine nucleotide transhydrogenase; *Ec-aroG^{fbr}*, *Ec-ppsA*, *Ec-shiA*, *Bt-asbF*, *Ct-pmdAB*, *Ct-pmdC*, *Ec-sthA*, and *aroE*Δ	Luo et al., 2018
	Vanillin	*S. cerevisiae*	0.045	2.25	0.94	Shaking flask (48 h)	*Ppa-3DSD*, *Hs-OMT*, *Ng-ACAR*, *Cg-PPTase*, and *adh6*Δ	Hansen et al., 2009
		S. pombe	0.065	3.25	1.35	Shaking flask (48 h)	*Ppa-3DSD*, *Hs-OMT*, *Ng-ACAR*, and *At-UGT*	
	Vanillin β-D-glucoside	*S. cerevisiae*	0.5	32	N/A	chemostat	*In silico* design and target selection; *Ppa-3DSD*, *Hs-OMT*, *Ng-ACAR*, *Cg-PPTase*, *adh6*Δ, and *pdc1*Δ	Brochado et al., 2010
Chorismate	Salicylate	*E. coli*	11.5	320	240	Batch (48 h)	Replacing endogenous PTS system and disrupting conversion of PEP to PYR; *Ec-menF*, *Ppa-pchB*, *Ec-galP*, *Ec-glk*, *Ec-ppsA*, *Ec-maeB*, *Ec-ppc*, *pheA*Δ, *tyrA*Δ, *pykA*Δ, *pykF*Δ, *ptsH*Δ, and *ptsI*Δ	Noda et al., 2016
	Salicylate 2-O-β-D-glucoside	*E. coli*	~2	~160 (glucose + glycerol)	~42	Shaking flask (48 h)	*Ec-galU*, *Ec-pgm*, *Ye-irp9*, *At-Ugt74F1*, *pheA*Δ, and *tyrA*Δ	Qi et al., 2018
	4-Aminobenzoate	*E. coli*	4.8	160	100	Fed-batch (48 h)	*Ec-aroF^{fbr}*, *Ec-pabAB*, *and Ec-pabC*	Koma et al., 2014
		E. coli	2.88	112	60	Test tube (48 h)	*Ec-pabAB* and *Ec-pabC* (Referring salicylate production for additional genetic modifications)	Noda et al., 2016
		S. cerevisiae	0.215	2 (glycerol + ethanol)	2	Fed-batch (108 h)	*Sc-ABZ1*, *Sc-ABZ2*, *Sc-ARO4^{K229L}*, *aro7*Δ, and *trp3*Δ	Averesch et al., 2016
		E. coli	1.83	152	38	Test tube (48 h)	*Ec-pabA*, *Ec-pabB*, and *Ec-pabC* (Referring salicylate production for additional genetic modifications)	Noda et al., 2016

Table 2. The microbial production of the upstream products from the shikimate pathway (continued).

Node	Target Compound	Host	Titer (g/L)	Yield (mg/g)	Productivity (mg/L/h)	Fermentation type	Unique strategies and genotype	Reference
Chorismate	4-Hydroxybenzoate	*E. coli*	12	100	167	Fed-batch (72 h)	*Ec-ubic*, *Ec-tktA*, *Ec-aroF^{fbr}*, *Ec-aroB*, *Ec-aroL*, *Ec-aroA*, and *Ec-aroC*	Barker et al., 2001
		E. coli	1.82	391	19	Test tube (96 h)	*Ec-ubiC* (Referring salicylate production for additional genetic modifications)	Noda et al., 2016
		E. coli co-culture	2.3	110	24	Fed batch (96 h)	Strain 1: producing upstream module of DHS; Strain 2: *Ec-shiA*, *Ec-aroE*, *Ec-aroL*, *Ec-aroA*, *Ec-aroC*, and *Ec-ubiC*	Zhang et al., 2015b
		S. cerevisiae	2.9	2.2	29	Fed batch (100 h)	*Ec-aroL*, *Ec-ubiC*, *Ec-aro4^{K229L}*, *aro7*Δ, and *trp3*Δ	Averesch et al., 2017
		P. putida	1.73	139	54	Fed batch (44 h)	*Ec-ubiC*, *Pp-aroG^{D146N}*, *pobA*Δ, *pheA*Δ, *trpE*Δ, and *hexR*Δ	Yu et al., 2016
	3-Hydroxybenzoate	*E. coli*	2.18	192	30	Test tube (72 h)	*Sh-hyg5* (Referring salicylate production for additional genetic modifications)	Noda et al., 2016
		E. coli co-culture	0.048	2.4	0.33	Test tube (144 h)	Strain 1: containing upstream module of DHS; Strain 2: *Sp-pctV*	Zhang et al., 2015b
	Anthranilic acid	*E. coli*	14,000	200	412	Fed batch (34 h)	TrpD* has an early stop codon, and the protein only maintains the N-terminal glutamine amidotransferase activity but losts the anthranilate phosphoribosyl transferase at the C-terminal, leading to anthranilic acid accumulation; *Ec-trpD**, *Ec-aroGfbr*, and *Ec-tktA*	Balderas-Hernández et al., 2009
	Muconic acid	*E. coli*	3.1	136	32.3	Test tube (96 h)	*Ec-menF*, *Pae-pchB*, *Pp-nahG*, and *Pp-catA* (Referring salicylate production for additional genetic modifications)	Noda et al., 2016
		E. coli	0.17	21.3	2.36	Shake flask (72 h)	pHBA route: *Ec-ubiC*, *Ec-aroF^{fbr}*, *Ec-aroE*, *Ec-aroL*, *Pp-pobA*, *Kp-aroY*, *Ac-catA*, *ptsH*Δ, *ptsI*Δ, *crr*Δ, and *pykF*Δ	Sengupta et al., 2015
		E. coli	0.48	38.4 (glucose+ glycerol)	10	Shake flask (48 h)	2,3-DHBA route: *Ec-aroL*, *Ec-ppsA*, *Ec-tktA*, *Ec-aroG^{fbr}*, *Ec-entCBA*, *Kp-bdc*, *Pp-cat*, and *entE*Δ	Sun et al., 2014
		E. coli	0.605	168	8.4	Shake flask (72 h)	2,3-DHBA route: *Ec-aroG^{fbr}*, *Ec-aroL*, *Ec-entCBA Ec-entX*, and *Pp-catA*	Wang et al., 2015
		E. coli	1.45	116 (glucose+ glycerol)	30.2	Shake flask (48 h)	salicylate route: *Ec-aroL*, *Ec-ppsA*, *Ec-tktA*, *Ec-aroG^{fbr}*, *Pa-pchB*, *Ec-entC*, *Pp-nahG^{opt}*, *Pp-catA*, *pheA*Δ, and *tyrA*Δ	Lin et al., 2014a
		E. coli	0.39	31.2 (glucose+ glycerol)	122	Shake flask (32 h)	anthranilate route: *Ec-ppsA*, *Ec-tktA*, *Ec-aroG^{fbr}*, *Ec-aroB*, *Ec-aroE*, *Ec-aroL*, *Ec-trpE*, *Ec-trpG*, *Pa-antABC*, *Pp-catA*, *Ec-glnA*, and *trpD*Δ	Sun et al., 2013
		E. coli	4.45	0.207	61.8	Test Tube (72 h)	Fusion of AroC and MenF to facilitate metabolite channeling; *Ec-galP*, *Ec-glk*, *Pp-nahG*, *Pp-catA*, *Pa-pchB*, *Ec-aroC/Ec-menF*, *ptsH*Δ, *ptsI*Δ, *pheA*Δ, *tyrA*Δ, *pykA*Δ, and *pykF*Δ	Fujiwara et al., 2018

Table 2. The microbial production of the upstream products from the shikimate pathway (continued).

Node	Target Compound	Host	Titer (g/L)	Yield (mg/g)	Productivity (mg/L/h)	Fermentation type	Unique strategies including genotype	Reference
Chorismate	Arbutin	*E. coli*	4.19	140	87.3	Shake flask (48 h)	*Cp-MNX1*, *Rs-AS*, *Ec-aroL*, *Ec-ppsA*, *Ec-tktA*, *Ec-aroG*^{fbr}, and *Ac-ubiC*	Shen et al., 2017
		P. chlororaphis	6.79	N/A	94.3	Fed batch (72 h)	*Pc-ppsA*, *Pc-tktA*, *Pc-phzC*, *Pc-aroB*, *Pc-aroD*, *Pc-aroE*, *Xo-xanB2*, *Cp-MNX1*, *Rs-AS*, *pykA*Δ, *pykF*Δ, *phzE*Δ, *pobA*Δ, *lon*Δ, *rpeA*Δ, and *rsmE*Δ	Wang et al., 2018c
	Phenazine-1-carboxamide (PCN)	*P. chlororaphis*	4.1	230 (glycerol)	137	Shake flask (30 h)	Disrupting pathway specific negative regulatory genes; *Pc-phzEDFABGH*, *lon*Δ, *parS*Δ, and *prsA*Δ	Yao et al., 2018
	Gastrodin	*E. coli*	0.545	27.3	11.4	Shake flask (48 h)	*Ec-aroG*^{fbr}, *Ec-ubiC*, *Ni-car*, *Bs-sfp*, and *Rsa-UGT73B6F389S*	Bai et al., 2016
	Gallic acid	*E. coli*	1.27	102 (glucose+ glycerol)	35.3	Shake flask (36 h)	*Pae-pobA*$^{Y385F, T294A}$, *Ec-aroL*, *Ec-ppsA*, *Ec-tktA*, *Ec-aroG*^{fbr}, and *Ec-ubiC*	Chen et al., 2017
	Pyrogallic acid	*E. coli*	1.036 (with 2 mM ascorbic acid supplement)	46 (glucose+ glycerol)	43.2	Shake flask (24 h)	*Ec-aroL*, *Ec-ppsA*, *Ec-tktA*, *Ec-aroGfbr*, *Ec-entCBA*, and *Pp-nahG*	Wang et al., 2018b

Table 3. The microbial production of compounds derived from L-Trp.

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type (total hours)	Reference
5-Hydroxy-tryptophan	*E. coli*	153 mg/L (Glu); 1.1–1.2 g/L (Glu with 2 g/L L-Trp supplement)	*Xc-p4h*^{W179F}, MH4 regeneration pathway (*Pa-pcd and Ec-dhmr*), *Ec-trpE*^{S40F}*DCBA, pheLAΔ, tyrAΔ,* and *tnaAΔ* Fermentation at 30°C to prevent oxidative degradation.	Shake flask (48 h)	Lin et al., 2014b
	E. coli	550 mg/L (Glu with 1 g/L L-Trp supplement)	*Ct-aaah*^{W192F}, MH4 regeneration pathway (*Hs-PCD* and *Hs-DHPR*), and *tnaAΔ* Fermentation at 30°C to prevent oxidative degradation.	Shake flask (24 h)	Mora-Villalobos et al., 2017
	E. coli	5.1 g/L (glycerol) (1.3 g/L in shake flasks)	*Hs-TPH2*$^{NΔ145, CΔ24}$, BH4 synthesis pathway (*Bs-mtrA, Hs-PTPS, Hs-SPR*), BH4 regeneration pathway (*Hs-PCD, Hs-DHPR*), and *tnaAΔ* Modulation of plasmid copy number and promoter strength.	Fed batch (40 h)	Wang et al., 2018a
	E. coli	98 mg/L (Glu)	Strain 1 (*Re-salABCD* and *Ec-trpE*^{fbr}*G*); Strain 2 (*Ec-trpDCBA*) SalABCD uses NAD(P)H as the cofactor, omitting the usage of the BH4 or MH4 system.	Two-stage shake flask (stage 1: 24 h; stage 2: 12 h)	Sun et al., 2015
	S. cerevisiae	0.255 μg/L/OD (Glu with 2 g/L L-Trp supplement)	Xc-p4h^{W179F}, MH4 regeneration pathway (*Pa-pcd, Ec-dhmr,* and *Sc-DFR1*) *Sm-T5H*, BH4 synthesis pathway (*Ec-gch, Rn-PTPS,* and *Rn-SPR*), BH4 regeneration pathway (*Ec-dhpr* and *Pa-pcbd*) The eukaryotic system enabled 17-fold higher production than the prokaryotic system	Shake flasks (72 h)	Zhang et al., 2016
Serotonin	*E. coli*	154 mg/L (Glu); (962 mg/L 5HTP produced by strain 1)	Strain 1 containing *Ct-aaah*^{F197L,E219C}, MH4 regeneration pathway (*Hs-PCD* and *Hs-DHPR*), and *trpRΔ*; Strain 2 containing *cr-TDC* and *tnaAΔ*. Strain 1 was created from the strain S028, a previous L-Trp overproducer.	Two-stage fed batch fermentation (60 h stage 1 and 52 h stage 2)	Mora-Villalobos et al., 2018
	E. coli	24 mg/L (Glu with 0.2 g/L L-Trp supplement)	*Os-GST-Δ37T5H,* and *Cr-TDC*	Test tube (24 h post IPTG induction)	Park et al., 2011
Melotonin	*S. cerevisiae*	14.5 mg/L (FIT medium or mineral medium supplemented with glucose)	*Sm-TPH, Hs-DDC, Bt-AANAT, Hs-ASMT, Sc-ALD6* BH4 synthesis pathway (*Rn-PTS* and *Rn-SPR*), and BH4 regeneration pathway (*Rn-DHPR* and *Pa-pcbd1*) Basal levels: *SAM, Sc-GCH1,* and *Sc-FOL2* Genome integration of all the genes except Hs-ASMT in high-copy plasmid; Special feed-in-time (FIT) medium.	Fed batch (small-scale in microplates)	Germann et al., 2016

Table 3. The microbial production of compounds derived from L-Trp (continued).

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type	Reference
Indigo	*E. coli*	920 mg/L (2 g/L L-Trp)	*Ma-fmo* 920 mg/L Indigo and 5 mg/L indirubin	Batch (24 h)	Han et al., 2011
	E. coli	640 mg/L (Glu)	*Ma-fmo, Ec-tnaA, Ec-tktA, Ec-aroG^{D146N}, Ec-trpE^{S40F}, Ec-aroL, trpRΔ, pykFΔ,* and *pykAΔ* *De novo* biosynthesis	Fed batch (72 hr)	Du et al., 2018
	E. coli	18 g/L (Glu with 0.2-0.5 g/min L-Trp)	*Pp-ndo, Ec-aroG^{fbr},* and *Ec-trpE^{fbr}DCB/26A* AroG^{fbr} was provided by two plasmids; Indoxyl or its downstream intermediates significantly inhibited AroG; PEP was found to protect AroG from inhibition.	Fed batch (72 h)	Berry et al., 2002
Indirubin	*E. coli*	58 mg/L (3.28 g/L L-Trp and 0.5 g/L 2-oxindole)	*Cs-ndo* Supplementation of 2-oxindole played a more favorable influence than isatin due to the inhibitory effect of isatin; Increasing 2-oxindole from 0 to 1.5 g/L improved the proportion of indirubin from 5% to 48%.	Batch	Zhang et al., 2014
	E. coli	223.6 mg/L (2 g/L L-Trp and 0.36 g/L Cys)	*Ma-fmo* Cysteine influenced the regioselectivity of FMO and enhanced the synthesis of 2-hydroxindole (6.8 mg/L indigo)	Batch (48 h)	Han et al., 2013 Han et al., 2011
	E. coli	56 mg/L (Glu)	*Ma-fmo, Ec-tnaA, Ec-tktA, Ec-aroG^{D146N}, Ec-trpE^{S40F}, Ec-aroL, trpRΔ, pykFΔ,* and *pykAΔ* *De novo* biosynthesis	Fed batch (72 h)	Du et al., 2018
Violacein	*E. coli*	710 mg/L (Glu)	*Ji-vioD, Cv-vioABCE* operon, *Ec-trpE^{fbr}, Ec-serA^{fbr}, Ec-tktA, Ec-aroFBL, Ec-trpDCBA, trpRΔ, trpLΔ, tnaAΔ,* and *sdaAΔ* The expression of *vioABCE* was induced by arabinose, which was also used as the carbon source; *Ji-vioD* was genome-integrated.	Fed batch (300 h)	Rodrigues et al., 2013
	C. glutamicum	5.4 g/L (Glu) 54 mg/g glucose 47 mg/L/hr	*Cv-vioBAECD* Inserting rbs in front of each *vio* gene; exploiting the established industrial L-Trp hyper producer strain ATCC 21850	Fed batch (115 h)	Sun et al., 2016a
Deoxy-violacein	*E. coli*	324 mg/L (10 g/L arabinose)	*Cv-vioABCE* operon, *Ec-trpE^{fbr}, Ec-serA^{fbr}, Ec-tktA, Ec-aroFBL, Ec-trpDCBA, Ec-trpRΔ, trpLΔ, tnaAΔ,* and *sdaAΔ* The expression of *vioABCE* was induced by arabinose, which was also used as the carbon source.	Shake flasks (228 h)	Rodrigues et al., 2013
	E. coli	1.6 g/L (glycerol) 20 mg/g glycerol 8 mg/L/hr	*Cv-vioABCE* operon, *Ec-trpE^{fbr}, Ec-serA^{fbr}, Ec-tktA, Ec-aroFBL, Ec-trpDCBA, trpRΔ, trpLΔ, tnaAΔ, sdaAΔ,* and *araBADΔ* Arabinose catabolism was disrupted; Arabinose was used only as an inducer.	Fed-batch (200 h)	Rodrigues et al., 2014

Table 4. The microbial production of compounds derived from L-Phe.

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type (total hours)	Reference
D-phenyllactic acid	*E. coli*	29.2 g/L (Glu with 0.2 g/L L-Tyr and 0.15 g/L L-Trp); 0.16 g/g glucose	*Ec-aroG*^{fbr}, *Ec-pheA*^{frb}, and *Wf-pprA* The starting strain have been mutated to prevent the synthesis of L-Tyr and L-Trp	Fed-batch (144 h)	Fujita et al., 2013
2-Phenylethanol	*K. marxianus*	1.3 g/L (Glu)	*Sc-ARO10, Sc-ADH2,* and *Km-AroG*^{fbr} *K. marxianus* is a thermal tolerance yeast with a short doubling time.	Batch (72 h)	Kim et al., 2014b
	E. coli	320 mg/L 13.3 mg/L/h	*Es-kdc4427, Es-adh4428, Ec-aroF, Ec-pheA, Ec-tktA,* and *Ec-ppsA* Severe product inhibition was observed.	Batch (24 h)	Liu et al., 2018a
	S. cerevisiae	6.1 g/L (10g/L L-Phe as the sole N source)	*Sc-ARO9, Sc-ARO10, Sc-ARO80,* and *ald3Δ* Whole cell conversion using PPG1200 as the in situ extractant; Aro80 is a transcriptional activator regulator whose function can be further enhanced by aromatic amino acids.	Two-phase batch (72 h)	Kim et al., 2014a
	K. marxianus	10.2 g/L (with high conc. of L-Phe)	Whole cell conversion using PPG1200 as the *in situ* extractant	Fed-batch (30 h)	Etschmann et al., 2006
	S. cerevisiae	12.6 g/L (with high conc. of L-Phe)	Whole cell conversion using oleic acid as the *in situ* extractant	Fed-batch (153 h)	Stark et al., 2002
Mandelate and 4-hydroxy-mandelate	*E. coli*	1.02 g/L (*S*-mandelate) 0.88 g/L (*R*-mandelate)	(*S*-mandelate) *Ao-hmaS, Ec-aroF*^{fbr}, *Ec-pheA*^{fbr}, *trpEΔ, tyrAΔ, tyrBΔ,* and *aspCΔ*; (*R*-mandelate) *Sco-hmo* and *Rg-dmd* in addition The biosynthesis of all three aromatic amino acids was disrupted.	Shake flasks (84 h)	Sun et al., 2011
	S. cerevisiae	1 g/L (4-hydroxy-mandelate)	*Ao-hmaS, Sc-ARO1, Sc-ARO3*^{fbr}, *Sc-ARO4*^{fbr}, *aro10Δ, pdc5Δ, trp2Δ, aro8Δ, aro9Δ,* and *pha2Δ*	Shake flasks (120 h)	Reifenrath and Boles, 2018
		0.236 g/L (mandelate)	*Ao-hmaS, Sc-ARO1, Sc-ARO3*^{fbr}, *Sc-ARO4*^{fbr}, *aro10Δ, pdc5Δ, trp2Δ, aro8Δ, aro9Δ,* and *tyr1Δ*		
Phenylglycine	*E. coli*	51.6 mg/g DCW (L-phenylglycine)	*Sco-hmaS, Sco-hmo, Sco-hpgT, Ec-tyrA*^{S121L}, *Ec-aroG*^{D146N}, *Ec-pheA*^{fbr}, *Ec-ydiB, Ec-aroK, crrΔ, tyrBΔ,* and *aspCΔ* Using temperature-sensitive promoters to induce the expression at 38°C; Sco-HpgT is a L-hydroxyphenylglycine aminotransferase.	Shake flasks (48 h)	Liu et al., 2014a
	E. coli	146 mg/g DCW (D-phenylglycine)	*Ao-hmaS, Sco-hmo, Pp-hpgT, Ec-aroF*^{fbr}, *Ec-pheA*^{fbr}, *pheAΔ, tyrAΔ, tyrRΔ, tyrBΔ,* and *aspCΔ* PpHpgT is a D-hydroxyphenylglycine aminotransferase.	Fed-batch (24 h)	Müller et al., 2006

Table 4. The microbial production of compounds derived from L-Phe (continued).

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type (total hours)	Reference
Trans-cinnamate, cinnamaldehyde, cinnamylalcohol, and hydro-cinnamylalcohol	*E. coli*	6.9 g/L (*trans*-cinnamate) 0.028 g/g glu 0.138 g/L/h	*Sm-pal, crrΔ, tyrRΔ, trpEΔ, tyrAΔ,* and *pykAΔ* Casamino acid was supplied as a feeding solution.	Fed-batch (86 h)	Bang et al., 2018
	S. cerevisiae	0.3 mg/L (cinnamaldehyde) 27.8 mg/L (cinnamylalcohol) 113.1 mg/L (hydrocinnamylalcohol)	*At-PAL2*, *Ns-acar*, *Ec-entD,* and *Sc-ADHs* Cinnamate and cinnamaldehyde are more toxic to cell growth than cinnamylalcohol and hydro cinnamylalcohol.	Shake flasks (26 h)	Gottardi et al., 2017
	E. coli	75 mg/L (cinnamaldehyde)	*Sm-pal*, *Sco-4cl*, *At-CCR*, *Ec-aroG*$^{D146N, A202T}$, *Ec-ydiB*, *Ec-aroK*, *Ec-pheA*$^{E159A, E232A, dm}$, *Ec-galP*, *Ec-glk*, *crrΔ*, *tyrRΔ*, *trpEΔ*, *tyrAΔ*, and *pykAΔ* PheA$^{E159A, E232A, dm}$ has the first 300 amino acids of the wild-type PheA, which excludes the L-Phe regulatory domain; the two mutations revive the substrate binding affinity of the truncated PheA.	Shake flasks (48 h)	Bang et al., 2016
	E. coli	1.7 g/L (trans-cinnamate)	*Ec-aroG*^{fbr}, *Ec-phe*^{frb}, *Wf-pprA,* and *Cs-fldABCI* Aerobic phase to accumulate D-phenyllactate and anaerobic phase to convert D-phenyllactate to trans-cinnamate by the O_2-sensitive FldABCI.	Two-stage batch (172 h)	Masuo et al., 2016
Benzoic acid	*S. maritimus*	257 mg/L (glucose) 337 mg/L (cellobiose) 460 mg/L (starch) 125 mg/L cellulose	Native plant-like β-oxidation pathway composed of *encPHIMJN*; can utilize cellobiose due to the endogenous β-glucosidase; Secreting the heterologous endo-glucanase tfu0901 cloned downstream of the phospholipase D promoter with the signal peptide from phospholipase D from *Streptoverticillium cinnamoneum*	Shake-flask (4-6 days)	Noda et al., 2012
Styrene	*E. coli*	260 mg/L (styrene)	*At-PAL2* and *Sc-FDC1* in ATCC31884 ATCC 31884 is a commercial L-Phe overproducing strain.	Shake-flask (39 h)	McKenna et al., 2011
	S. cerevisiae	29 mg/L (styrene) 1.44 mg/g glucose	*At-PAL2* (*Sc-FDC1*), *Sc-ARO4*^{K229L}, and *aro10Δ* Using the evolved strain after m-fluoro-DL-phenylalanine treatment; transcription analysis revealed several Aro genes in the upper module were up-regulated; *fdc1* was endogenously expressed with the native promoter.	Shake-flask (48 h)	McKenna et al., 2014

Table 5. The microbial production of compounds derived from the L-Tyr branch.

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type (total hours)	Reference
p-Coumaric acid, caffeic acid, and ferulic acid	*S. cerevisiae*	1.93 g/L (*p*-coumaric acid)	*Fj-tal, Sc-ARO4*^{K229L}*, Sc-ARO7*^{G141S}*, Ec-aroL, pdc5Δ,* and *aro10Δ* 96-deep well plate with FIT media	Fed batch (72 h)	Rodrigues et al., 2015
	S. cerevisiae	2.0 g/L (*p*-coumaric acid, S288c, batch) 2.4 g/L (*p*-coumaric acid, CEN.PK, batch)	*Sc-ARO4*^{K229L}*, Sc-ARO7*^{G141S}*, Ec-aroL, Fj-tal, pdc5Δ,* and *aro10Δ* CSTR gave 0.410 g/L (S288c) and 0.507 g/L (CEN.PK) production at the steady state; CEN.PK strain showed fewer changes in the transcription profile and intracellular metabolites concentrations.	Batch and CSTR	Rodriguez et al., 2017a
	E. coli	974 mg/L (*p*-coumaric acid) 150 mg/L (caffeic acid) 196 mg/L (ferulic acid)	*Se-tal, Ec-aroG*^{fbr}*, Ec-tyrA*^{fbr}, and *tyrRΔ* *Se-tal, Se-sam5, Ec-aroG*^{fbr}*, Ec-tyrA*^{fbr}, and *tyrRΔ* *Se-tal, Se-sam5, At-COM, Ec-aroG*^{fbr}*, Ec-tyrA*^{fbr}, and *tyrRΔ*	Shake flask (36 h)	Kang et al., 2012
	E. coli	106 mg/L (caffeic acid)	Route 1: *Rg-TAL and Se-Sam5* in strain rpoA14(DE3); Route 2: *Rg-TAL, Se-Sam5, Pc-4CL,* and the native *Ec-te* Strain rpoA14(DE3) was a previously constructed L-Tyr overproducer.	Batch (4 days)	Zhang et al., 2013
	E. coli	767 mg/L (caffeic acid)	*Rg-TAL, Ec-hpaBC, Ec-tyrA*^{fbr}*, Ec-aroG*^{fbr}*, Ec-ppsA, Ec-tktA, pheAΔ,* and *tyrAΔ* The host E. coli ATCC31884 is a commercial L-Phe overproducer.	Shake flask (72 h)	Huang et al., 2013
	E. coli	10.2 g/L (caffeic acid based on 4-time supplementation of 20 mM p-coumaric acid)	*Pa-hpaBC* Repeatedly adding 20 mM of *p*-coumaric acid 4 times during the reaction.	Whole cell conversion (24 h)	Furuya et al., 2014
	E. coli	2.8 g/L (caffeic acid based on supplementation of 20 mM p-coumaric acid)	*CYP199A2*^{F185L}*, Pp-pdR,* and *Rp-pux* Pp-PdR, and Rp-pux serve as the redox partners for electron transfer.	Whole cell conversion (24 h)	Furuya et al., 2012
	E. coli	280 mg/L (caffeic acid based on supplementation of 3 mM L-Tyr)	*Rg-TAL, Se-C3H or Rp-CYP199A2*$^{F-185L}$*, Pp-pdR,* and *Rp-pux* CYP199A2^{F185L} displayed a higher catalytic activity than C3H; Pp-PdR, and Rp-Pux serve as the redox partners of CYP199A2 for electron transfer.	Whole cell conversion (63 h)	Rodrigues et al., 2015
Hydroxystyrene	*P. putida*	2.5 g/L 90 mg/L/h	*Rt-PAL/TAL, Lp-pdc,* and *fcsΔ* Product toxicity was overcome by using 1-decanol as the *in situ* extractant; The host previously optimized for phenol and *p*-coumaric acid production.	Two-stage fed batch (50 h)	Verhoef et al., 2009
	E. coli	0.4 g/L	*Rt-PAL/TAL* and *Lp-pdc* in strain NST74 NST74 is a previously constructed L-Phe overproducer.	Fed batch (56 hr)	Qi et al., 2007
3-Phenylpropionic acid 3-(4-hydroxyphenyl) propionic acid	*E. coli*	366.77mg/L (3-phenylpropionic acid) 225.10mg/L (3-(4-hydroxyphenyl) propionic acid)	*Rg-TAL, Ca-er, Ec-tyrA*^{fbr}*, Ec-aroG*^{fbr}*, Ec-ppsA,* and *Ec-tktA*, in ATCC31884 (3-phenylpropionic acid); *Rg-TAL, Ca-er, Ec-tyrA*^{fbr}*, Ec-aroG*^{fbr}*, Ec-ppsA, Ec-tktA, pheAΔ,* and *tyrAΔ* in ATCC31884 (3-(4-hydroxyphenyl) propionic acid). Screened 11 ERs from 9 microorganisms; the host E. coli ATCC31884 is a commercial L-Phe overproducer.	Shake flask (72 h)	Sun et al., 2016b

Table 5. The microbial production of compounds derived from the L-Tyr branch (continued).

Target Compound	Host	Reported titer, yield, and/or productivity (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type (total hours)	Reference
Tyrosol	*E. coli*	573 mg/L	*Sc-ARO10, pheAΔ,* and *feaBΔ* (*de novo* biosynthesis); *Sc-ARO10, Sc-ARO8, pheAΔ,* and *feaBΔ* (whole cell conversion). 1203 mg/L tyrosol was produced in the whole cell conversion when 1.8 g/L L-Tyr was used as the substrate.	Shake flask (48 h) Whole cell conversion (20 h)	Xue et al., 2017
Hydroxytyrosol	*E. coli*	647 mg/L (from glucose) 1243 mg/L (1 g/L L-Tyr was fed three times)	*Sc-ARO10, Sc-ADH6, Ec-hpaBC, Ec-tyrA, Ec-ppsA, Ec-tktA, Ec-aroG^{fbr},* and *feaBΔ* Ascorbic acid was added to prevent hydroxytyrosol oxidation; 1-dodecanol was used as in situ extractant to overcome product toxicity; NH_4Cl was removed from the culture to enhance transamination.	Shake flask (48 h)	Li et al., 2018a
3,4-Dihydroxyphenyl acetate	*E. coli*	1856 mg/L 0.17 g/g carbon source	*Sc-ARO10, Ec-feaB, Ec-hpaBC, Ec-tyrA, Ec-ppsA, Ec-tktA, Ec-aroG^{fbr}, pykAΔ,* and *pykFΔ* Mixed glucose and glycerol as the carbon sources resulted in best production.	Shake flask (48 h)	Li et al., 2019
Salidroside	E. coli	288 mg/L	*Pc-AAS, At-UGT86A1, tyrRΔ, pheAΔ,* and *feaBΔ* UGT85A1 was one of the 12 UGTs from *A. thaliana* that showed a glycosylation activity towards tyrosol.	Shake flask (48 h)	Chung et al., 2017
	E. coli co-culture	6.03 g/L (based on glucose and xylose mixed culture) 0.047 g/L/hr	Strain 1: *Pp-kdc4, Ec-aroG^{fbr}, Ec-tyrA^{fbr}, Ec-aroE, tyrRΔ, pykAΔ, pykFΔ, pheAΔ, feaBΔ, ptsGΔ,* and *manZΔ* Strain 2: *At-UGT85A1, Ec-galU, Ec-pgm, xylAΔ, tyrAΔ,* and *ushAΔ*	Fed batch (129 h)	Liu et al., 2018b
Salvianic acid	*E. coli*	7.1 g/L (plasmid version) 5.6 g/L (genome integration)	*Ec-aroG^{fbr}, Ec-tyrA^{fbr}, Ec-aroE, Ec-glk, Ec-tktA. Ec-ppsA, Ec-hpaBC, Lp-D-ldh^{Y52A}, ptsGΔ, pykFΔ, pykAΔ, pheAΔ,* and *tyrRΔ* The genome integration version did not require antibiotics and inducers.	Fed batch (60-70 h)	Yao et al., 2013 Zhou et al., 2017
L-DOPA	*E. coli*	8.67 g/L	*Ec-tyrA^{fbr}, Ec-tyrB, Ec-hpaBC, Ec-galP, Ec-glK, tyrRΔ, csrAΔ, zwfΔ, pheLAΔ, ptsHIΔ,* and *crrΔ* TyrA^{fbr}, TyrB, and HpaBC were fused to form a fusion protein chimera	Fed batch (60 h)	Wei et al., 2016
	E. coli	29.8 g/L	*Ss-tpl* 50 mM pyrocatechol and sodium pyruvate, and 0.65 mM ammonia were fed.	Whole-cell conversion (6 h)	Lee et al., 1996
	E. coli *P. aeruginosa*	592 mg/L (*E. coli*) 2761 mg/L (*P. aeruginosa*)	*Cf-tpl, Pp-tdo,* and *Pp-tcgdh* Benzene and pyruvate were added intermittently as the substrates.	Whole-cell conversion (14 h)	Park et al., 1998

Table 6. Microbial production of selected benzylisoquinoline alkaloids derived from L-Tyr.

Product	Host	Titer	The genetic manipulations in the final producer	Fermentation type	Reference
Bioconversion					
(S)-canadine	*S. cerevisiae* (supplemented with 1 mM norlaudanosoline)	621 μg/L	*Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *At-CPR*, *Ps-BBE*, *Ps-S9OMT*, *Tf-S9OMT*, *Cj-CAS*, and *Bw-STOX*	Batch (72 h)	Galanie and Smolke, 2015
Berberine		27 μg/L	1.8 mg/L canadine and 6.5 μg/L berberine in the shake flasks; 39 μg/L berberine in fed-batch.		
(S)-sanguinarine	*S. cerevisiae* (supplemented with 2 mM norlaudanosoline)	80 μg/L	*Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *Ps-BBE*, *Eca-CFS*, *Eca-STS*, *Ps-TNMT*, *Ps-MSH*, *Eca-P6H*, and *At-CPR* (*ATR1*)	Tube or 96-well plate (10 day)	Trenchard and Smolke, 2015
Magnoflorine	*E. coli and S. cerevisiae co-culture* (supplemented with 5 mM dopamine)	7.2 mg/L	*E. coli*: *Ml-mao*, *Cj-PR10A* (*NCS*), *Cj-6OMT*, *Cj-CNMT*, and *Cj-4'OMT* *S. cerevisiae*: *Cj-BBE*	Separately growing two strains for 12-20 h, and then mixing for 48-72 h	Minami et al., 2008
(S)-scoulerine		8.3 mg/L	*E. coli*: *Ml-mao*, *Cj-PR10A* (*NCS*), *Cj-6OMT*, *Cj-CNMT*, and *Cj-4'OMT* *S. cerevisiae*: *Cj-CYP80G2* and *Cj-CNMT*		
De novo biosynthesis					
(*S*)-Reticuline	*E. coli* (from glycerol)	46 mg/L	*Rs-tyr*, *Pp-dodc*, *Ml-mao*, *Cj-NCS*, *Cj-6OMT*, *Cj-CNMT*, *Cj-4'OMT*, *Ec-aroG*^{fbr}, *Ec-tyrA*^{fbr}, *Ec-ppsA*, *Ec-tktA*, and *tyrRΔ*	Fed batch (80 h)	Nakagawa et al., 2011
	S. cerevisiae (from glucose)	19.2 μg/L	*Rn-TyrH*$^{W166Y, R37E, R38E}$, *Pp-dodc*, *Cj-NCS*, *Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *Eca-CYP80B1*, *Ps-CPR*, *Sc-TKL1*, *Rn-SepR*, *Rn-PTPS*, *Rn-QDHPR*, *Rn-PCD*, *Sc-ARO4*^{Q166K}, and *zwf1Δ*	96-well plates (96 h)	Trenchard et al., 2015
	S. cerevisiae (from glucose)	80.6 μg/L	*Bv-CYP76AD1*$^{W13L, F309L}$, *Pp-dodc*, *Ps-NCS*, *Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *Eca-CYP80B1*, and *Sc-ARO4*^{K229L}	Shake flask (96 h)	DeLoache et al., 2015
Thebaine	*S. cerevisiae* (from glucose)	6.4 μg/L	*Sc-ARO4*^{Q166K}, *Sc-TKL*, *Sc-ARO10*, *Sc-ARO7*^{T266I}, *Rn-SepR*, *Rn-PTPS*, *Rn-QDHPR*, *Rn-PCD*, *Rn-TyrH*$^{R37E, R38E, W166Y}$, *Rn-DHFR*, *Pp-dodc*, *Cj-NCS*, *Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *Ps-CPR*, *Eca-NMCH*, *Pb-DRS-DRR*, *Eca-CFS*$^{1-83}$*-PbSalSyn*$^{92-504}$, *Pb-SalR*, and *Ps-SalAT*	Tube or 96-well plate (120 h)	Galanie et al., 2015
Hydrocodone		0.3 μg/L	*Ps-T6ODM* and *Pp-MORB* in addition to all the above manipulations		
Noscapine	*S. cerevisiae* (from glucose and glycerol)	2.2 mg/L	*Sc-ARO4*^{Q166K}, *Sc-TKL*, *Sc-ARO10*, *Sc-ARO7*^{T266I}, *Sc-ALD6*, *Sc-TYR1*, *Rn-SepR*, *Rn-PTPS*, *Rn-QDHPR*, *Rn-PCD*, *Rn-TyrH*$^{R37E, R38E, W166Y}$, *Rn-DHFR*, *Pp-dodc*, *Cj-NCS*, *Ps-6OMT*, *Ps-CNMT*, *Ps-4'OMT*, *Ps-CPR*, *Eca-NMCH*, *Ps-BBE*, *Ps-S9OMT*, *Cj-CAS*, *Ps-CYP82Y1A*, *Ps-TNMT*, *Ps-CYP82X1*, *Ps-CXE1*, *Ps-MT2*, *Ps-CYP82X2*, *Ps-AT1*, *Ps-SDR1*, and *Ps-MT3*	Shake flask (72 h)	Li et al., 2018b

Table 7. Microbial production of selected flavonoids and stilbenoids derived from the L-Tyr and L-Phe branches.

Product	Host	Titer (carbon source with additional supplements)	The genetic manipulations in the final producer Metabolic engineering strategies	Fermentation type	Reference
Naringenin	*E. coli*	29 mg/L 84 mg/L (with addition of cerulenin)	two copies of (*Ec-tyrA*^{fbr} and *Ec-aroG*^{fbr}), *Rg-TAL, Pc-4CL, Ph-CHS, Ms-CHI, pheAΔ,* and *tyrRΔ* Cerulenin is an inhibitor to fatty acid biosynthesis; malonyl-CoA is the rate-limiting precursor.	Shake flask (48 h)	Santos et al., 2011
	E. coli co-culture	41.5 mg/L	Strain 1: two copies of (*Ec-tyrA*^{fbr} and *Ec-aroG*^{fbr}), *Rg-TAL, pheAΔ,* and *tyrRΔ*; Strain 2: *Rg-TAL, Pc-4CL, Ph-CHS,* and *Ms-CHI* L-Tyr and *p*-coumaric acid are the two connecting molecules.	Shake flask (36 h)	Ganesan et al., 2017
	S. cerevisiae	109 mg/L	*Sc-ARO4*^{G226S}*, At-PAL1, At-C4H, At-CPR1, At-4CL3, At-CHS3, At-CHI1, Rc-tal1, aro3Δ, aro10Δ, pdc5Δ,* and *pdc6Δ* Two additional copies of *At-CHS3* integration increased the titer by 2.5-fold.	Batch fermenter (38 h)	Koopman et al., 2012
	E. coli	421.6 mg/L (supplementation of 3 mM L-Tyr)	*Anti-fabF/fumC/fabB/sucC/adhE, Rg-TAL*, *Pc-4CL*, *Ph-CHS*, and *Ms-CHI* CRISPRi was applied to target central metabolic pathway genes to increase malonyl-CoA level.	Shake-flask (48 h)	Wu et al., 2015
(*2S*)-Pinocembrin	*E. coli*	40.02 mg/L	*Ec-aroF*^{wt}*, Ec-phe*^{fbr}*, Rg-PAL, Pc-4CL, Ph-CHS, Ms-CHI, Rt-matB,* and *Rt-matC* MatB and MatC were overexpressed to increase malonyl-CoA level.	Shake-flask (48 h)	Wu et al., 2013
Resveratrol	*S. cerevisiae*	415.65 mg/L (from glucose) 531.41 mg/L (from ethanol)	*Sc-ARO4*^{K229L}*, Sc-ARO7*^{G141S}*, Sc-ACC1*$^{S659A, S1157A}$*, Ha-tal, At-4CL1,* and *Vv-VST1* (multiple copies of *tal*, *4CL1*, and *VST1* were integrated.) Batch fermentation on glucose followed by fed-batch on glucose or ethanol	Two-stage fermentation (~100 h)	Li et al., 2015
Pinosylvin	*E. coli*	281 mg/L	*Ec-aroF*^{wt}*, Ec-pheA*^{fbr}*, Tc-PAL, Pc-4CL, Vv-STS,* and *anti-fabF/fabB/eno/adhE/fumC/sucC* sgRNA (CRISPRi)	Shake-flask (48 h)	Wu et al., 2017
Ponciretin	*E. coli*	42.5 mg/L	*Ec-aroG*^{fbr}*, Ec-ppsA, Ec-tktA, Ec-tyrA*^{fbr}*, Se-tal, Os-4CL, Pe-CHS, Gm-SOMT2, tyrRΔ, pheAΔ,* and *icdAΔ*	Shake-flask (60 h)	Kim et al., 2013
Sakuranetin		40.1 mg/L	*Ec-aroG*^{fbr}*, Ec-ppsA, Ec-tktA, Ec-tyrA*^{fbr}*, Se-tal, Os-4CL, Pe-CHS, Os-NOMT, tyrRΔ, pheAΔ,* and *icdAΔ*	Shake-flask (48 h)	
Kaempferol	*S. cerevisiae*	26.57 mg/L	*Sc-ARO4*^{K229L}*, Sc-ARO7*^{G141S}*, Fj-tal, Pc-4CL, aro10Δ,* and *pdc5Δ* (base strain) plus *Ph-CHS, Ms-CHI, Am-F3H,* and *At-FLS*	96-well synthetic FIT fed-batch fermentation (72 h)	Rodriguez et al., 2017b
Liquiritigenin		5.31 mg/L	plus *Ph-CHS, Ms-CHI,* and *Am-CHR*		
Resokaempferol		0.51 mg/L	plus *Ph-CHS, Ms-CHI, Am-CHR, Am-F3H,* and *At-FLS*		
Quercetin		20.38 mg/L	plus *Ph-CHS, Ms-CHI, Am-F3H, At-FLS, Ph-FMO,* and *Cr-CPR*		
Fisetin		2.29 mg/L	plus 2 copies of *Ph-CHS/Am-CHR, Ms-CHI, Am-F3H, At-FLS, Ph-FMO,* and *Cr-CPR*		
Curcumin	*E. coli*	0.6 mg/L (supplemented with 3 mM L-Tyr)	*Tp-TAL, At-4CL, Ms-COMT, Se-c3h,* and *Os-CUS* Bisdemethoxycurcumin (*Tp-TAL, At-4CL,* and *Os-CUS*) and dicinnamoylmethane (*Tp-PAL, At-4CL,* and *Os-CUS*) can be produced when 3 mM L-Tyr and L-Phe were supplied, respectively.	Shake-flask (48 h)	Wang et al., 2015
Phloretin	*S. cerevisiae*	42.7 mg/L	*At-PAL2, Am-C4H, Sc-CPR1, At-4CL2, Han-CHS,* and *Sc-TSC13*	24-well plate batch (72 h)	Eichenberger et al., 2017
Phlorizin		65 mg/L	*At-PAL2, Am-C4H, Sc-CPR1, At-4CL2, Han-CHS, Sc-TSC13,* plus *Pc-UGT88F2* (for phlorizin), *Os-CGT* (for nothofagin), *At-UGT73B2* (for trilobatin), or the combination of *At-UGT73B2*, *Cm-1,2RHAT*, and *At-RHM2* (for naringin dihydrochalcone)		
Nothofagin		59 mg/L			
Trilobatin		32.8 mg/L			
Naringin dihydrochalcone		11.6 mg/L			

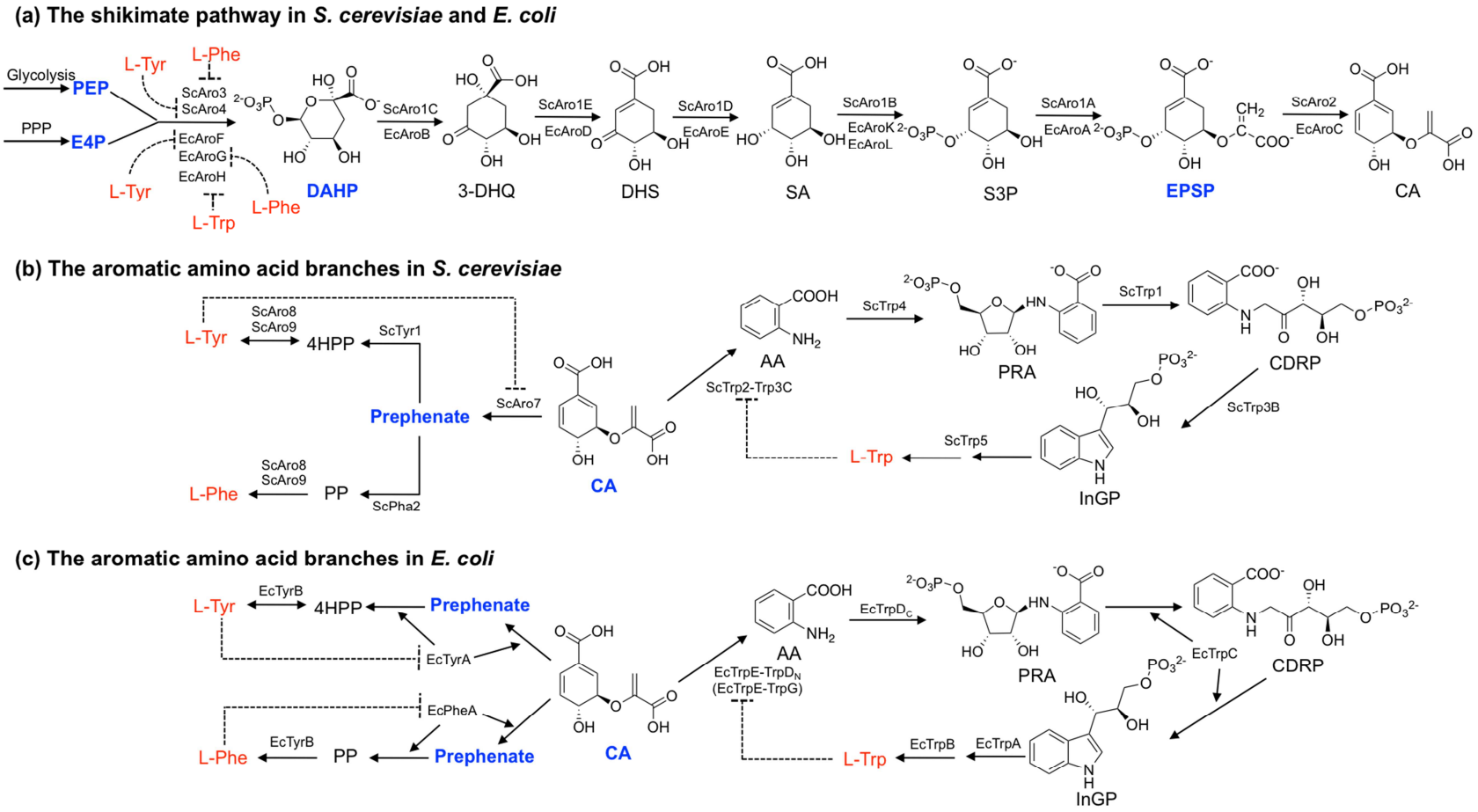


Figure 1. The backbone reactions of aromatic biosynthesis in *S. cerevisiae* and *E. coli*. Some parts of the pathway are organized differently in the two hosts with distinct higher-order enzyme associations. (a) The upstream shikimate pathway; (b) The three aromatic amino acid branches in *S. cerevisiae*; (c) The three aromatic amino acid branches in *E. coli*.

Abbreviations of compounds: 4-HPP, 4-hydroxyphenylpyruvate; AA, anthranilate; 3-DHQ, 3-dehydroquinic acid; CA, chorismate; CDRP, 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-phosphate; DHS: dehydroshikimate; DAHP, 3-deoxy-D-arabino-heptulosonic acid 7-phosphate; E4P, erythrose 4-phosphate; InGP, indole-3-glycerol-phosphate; PEP, phosphoenolpyruvate; SA, shikimic acid; S3P, shikimate 3-phosphate; EPSP, 5-enolpyruvylshikimate-3-phosphate; L-Phe, L-phenylalanine, PP, Phenylpyruvate; PRA, phosphoribosyl anthranilate; L-Trp, L-tryptophan; L-Tyr, L-Tyrosine.

Abbreviations of enzymes: Aro1, pentafunctional arom protein; Aro1A/AroA, EPSP synthase; Aro1B, SA kinase; Aro1E/AroD, 3DHQ dehydratase; Aro1D/AroE, shikimate dehydrogenase; Aro2/AroC, chorismate synthase; Aro3/Aro4/AroF/AroG/AroH, DAHP synthase; Aro7, chorismate mutase; Aro8/9, aromatic amino acid transaminase; AroK, SA kinase I; AroL, SA kinase II; PheA/Pha2, prephenate dehydratase; Trp1, PRA isomerase; Trp2/3c, anthranilate synthase; Trp3B, InGP synthase; Trp4, anthranilate phosphoribosyl transferase; Trp5, bi-functional Trp synthase; TrpA, alpha subunit of Trp synthase; TrpB, beta subunit of Trp synthase; TrpC, bi-functional InGP synthase; $TrpD_C$, C-terminal of TrpD, anthranilate phosphoribosyl transferase; TrpE, anthranilate synthase; $TrpG/TrpD_N$, N-terminal of TrpD, glutamine amidotransferase; Tyr1/TyrA, prephenate dehydrogenase; TyrB, aromatic amino acid transaminase.

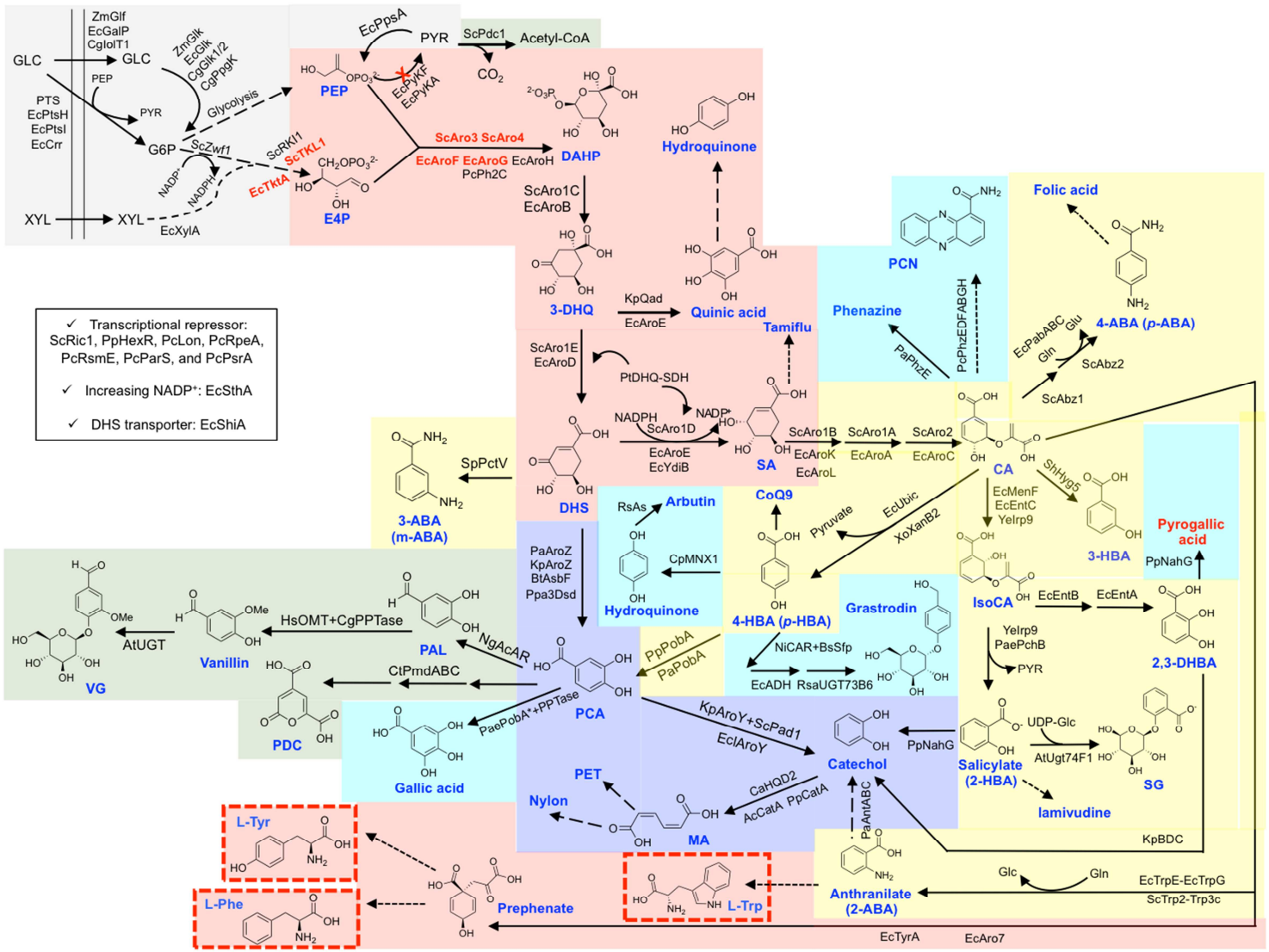


Figure 2. The upstream products derived from the shikimate pathway and their biosynthetic pathways.

Abbreviations of compounds: 2,3-DHBA, 2,3-dihydroxybenzoic acid; 2-ABA, anthranilate; 2-HBA, salicylate; 3-DHQ, 3-dehydroquinic acid; 3-ABA, 3-aminobenzoic acid; 3-HBA, 3-hydroxybenzoate; 4-ABA, 4-aminobenzoate; 4-HBA, 4-hydroxybenzoate; CA, chorismate; DHS: dehydroshikimate; DAHP, 3-deoxy-D-arabino-heptulosonic acid 7-phosphate; E4P, erythrose 4-phosphate; G6P, glucose 6-phosphate; GLC, glucose; Gln, glutamine; Glu, glutamate; IsoCA, isochorismate; MA, muconic acid; PAL, protocatechuic aldehyde; PEP, phosphoenolpyruvate; PET, polyethylene terephthalate; PCA, protocatechuate; PCN, phenazine-1-carboxamide; PDC, 2-pyrone-4,6-dicarboxylic acid; PYR, pyruvate; SA, shikimic acid; SG, salicylate 2-O-β-D-glucoside; VG, vanillin β-D-glucoside; XYL, xylose.

Abbreviations of enzymes: 3DSD, 3-dehydroshikimate dehydratase; Abz1, aminodeoxychorismate synthase; Abz2, aminodeoxychorismate lyase; ACAR, aromatic carboxylic acid reductase; ADH, alcohol dehydrogenase; AntABC, anthranilate 1,2-dioxygenase; Aro1, pentafunctional arom protein; Aro3/Aro4/AroF/AroG/AroH, DAHP synthase; AroC, chorismate synthase; AroD, 3DHQ dehydratase; AroE, shikimate dehydrogenase; AroK, SA kinase I; AroL, SA kinase II; AroY, protocatechuate decarboxylase; AroZ, DHS dehydratase; AS, glucosyltransferase; AsbF, dehydroshikimate dehydratase; BDC, 2,3-DHBA decarboxylase; CAR, carboxylic acid reductase; CatA, catechol 1,2-dioxygenase; DHQ-SD, dehydroquinate dehydratase-shikimate dehydrogenase; EntA, 2,3-dihydro-2,3-DHBA dehydrogenase; EntB, isochorismatase; EntC, isochorismate synthase; GalP, galactose permease; Glf, glucose transporter; Glk, glucokinase; HexR, glucose catabolism repressor; HQD2, atechol-1,2-dioxygenase; Hyg5, chorismatase; Irp9, bifunctional salicylate synthase; MenF, IsoCA synthase; MNX1, 4-HBA 1-hydroxylase; NahG, Salicylate 1-monoxygase;Pad1, flavin mononucleotide prenyltransferase; OMT, *O*-methyltransferase; PabAB, aminodeoxychorismate synthase; PabC, 4-amino-4-deoxychorismate lyase; PchB, IsoCA-pyruvate lyase; PctV, 3-ABA synthase; PmdAB, PCA 4,5-dioxygenase; PmdC, CHMS dehydrogenase; PobA, 4-HBA hydroxylase; PhzC, DAHP synthase; PhzE, anthranilate synthase; PpsA, phosphoenolpyruvate synthase; PPTase, phosphopantetheinyl transferase; PTS, phosphoenolpyruvate: carbohydrate phosphotransferase; PtsH, PtsI, and Crr, PTS glucose transport system; PykA/F, pyruvate kinase; Qad, quinic acid dehydrogenase; RKI1, ketol-isomerase; Sfp, phosphopantetheinyl transferase; ShiA, DHS transporter; SthA, pyridine nucleotide transhydrogenase; Tkl1/TktA, transketolase; Trp2, anthranilate synthase; Trp3C, the N-terminal glutamine amidotransferase domain of TRP3; TrpE, anthranilate synthase; TrpG, glutamine amidotransferase; UbiC, chorismate lyase; UGT, UDP glycosyltransferase; Ugt74F1, salicylate glucosyltransferase; UGT73B6, glycosyltransferase; XanB2, chorismate lyase; XylA, xylose isomerase; YdiB, shikimate dehydrogenase; Zwf1, glucose-6-phosphate dehydrogenase.

Abbreviations of species: Ac, *Acinetobacter calcoaceticus*; At, *Arabidopsis thaliana*; Bs, *Bacillus subtilis*; Bt, *Bacillus thuringiensis*; Ca, *Candida albicans*; Cg, *Corynebacterium glutamicum*; Cp, *Candida parapsilosis*; Ct, *Comamonas testosterone*; Ec: *E. coli*; Ecl, *Enterobacter cloacae*; Hs, *Homo sapiens*; Kp, *Klebsiella pneumonia*; Ng, *Nocardia genus*; Ni, *Nocardia iowensis*; Pa, *Podospora anserine*; Pae, *Pseudomonas aeruginosa*; Pc, *Pseudomonas chlororaphis*; Pp, *Pseudomonas putida*; Ppa, *Podospora pauciseta*; Pt, *Populus trichocarpa*; Rs, *Rauvolfia serpentine*; Rsa, *Rhodiola sachalinensis*; Sc, *S. cerevisiae*; Sp, *Streptomyces pactum*; Xo, *Xanthomonas oryzae pv. oryzae*; Ye, *Yersinia enterocolitica*; Zm, *Zymomonas mobilis*.

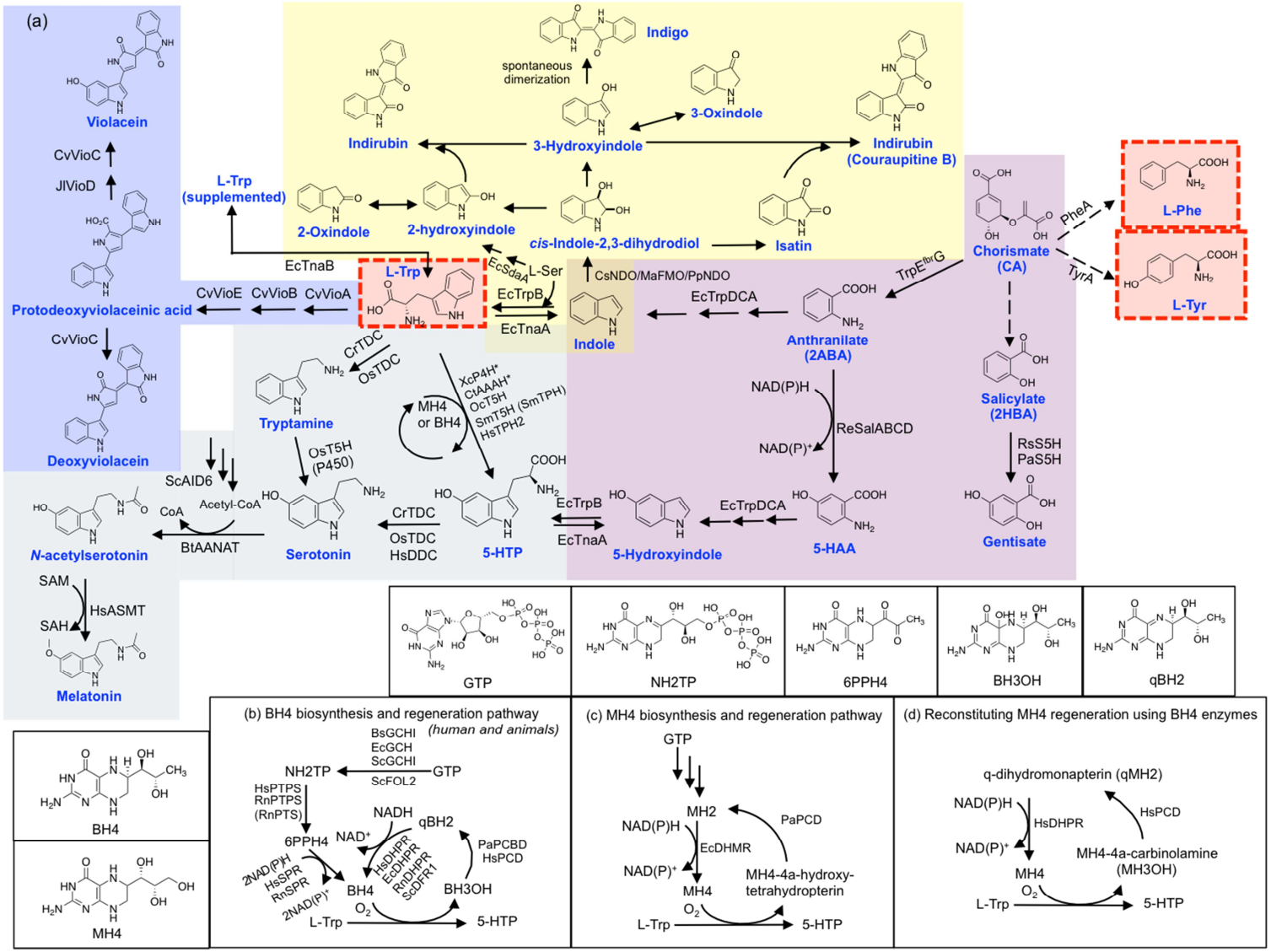

(a)
Violacein
CvVioC
JlVioD
Protodeoxyviolaceinic acid
CvVioE
CvVioB
CvVioA
Deoxyviolacein
L-Trp (supplemented)
EcTnaB
L-Trp
EcSdaA
L-Ser
EcTrpB
EcTnaA
Indole
CsNDO/MaFMO/PpNDO
Indigo
spontaneous dimerization
Indirubin
3-Hydroxyindole
3-Oxindole
Indirubin (Couraupitine B)
2-Oxindole
2-hydroxyindole
cis-Indole-2,3-dihydrodiol
Isatin
EcTrpDCA
Anthranilate (2ABA)
TrpE^fbr G
Chorismate (CA)
PheA
TyrA
L-Phe
L-Tyr
CrTDC
OsTDC
Tryptamine
MH4 or BH4
XcP4H*
CtAAAH*
OcT5H
SmT5H (SmTPH)
HsTPH2
NAD(P)H
NAD(P)+
ReSalABCD
Salicylate (2HBA)
RsS5H
PaS5H
Gentisate
OsT5H (P450)
ScAID6
Acetyl-CoA
CoA
BtAANAT
N-acetylserotonin
Serotonin
CrTDC
OsTDC
HsDDC
5-HTP
EcTrpB
EcTnaA
5-Hydroxyindole
EcTrpDCA
5-HAA
SAM
SAH
HsASMT
Melatonin
GTP
NH2TP
6PPH4
BH3OH
qBH2
BH4
MH4
(b) BH4 biosynthesis and regeneration pathway
BsGCHI
EcGCH
ScGCHI
(human and animals)
NH2TP
GTP
ScFOL2
HsPTPS
RnPTPS
(RnPTS)
6PPH4
NADH
NAD+
qBH2
PaPCBD
HsPCD
2NAD(P)H
HsSPR
RnSPR
2NAD(P)+
HsDHPR
EcDHPR
RnDHPR
ScDFR1
BH4
BH3OH
O2
L-Trp
5-HTP
(c) MH4 biosynthesis and regeneration pathway
GTP
MH2
NAD(P)H
PaPCD
EcDHMR
NAD(P)+
MH4-4a-hydroxy-tetrahydropterin
MH4
O2
L-Trp
5-HTP
(d) Reconstituting MH4 regeneration using BH4 enzymes
q-dihydromonapterin (qMH2)
NAD(P)H
HsDHPR
HsPCD
NAD(P)+
MH4-4a-carbinolamine (MH3OH)
MH4
O2
L-Trp
5-HTP

Figure 3.

Figure 3. The products derived from the L-Trp branch and their biosynthetic pathways (a). BH4 biosynthesis and regeneration pathway in human and animals (b). MH4 biosynthesis in *E. coli* and the heterologous regeneration pathway (c). Reconstituting MH4 regeneration in *E. coli* using BH4 regeneration enzymes (d).

Abbreviations of compounds: 5-HAA, 5-hydroxyanthranilate; BH3OH, pterin-4α-carbinolamine; BH4, 5,6,7,8-tetrahydrobiopterin; L-Ser, L-serine; MH4, tetrahydromonapterin;

Abbreviations of enzymes: AAAHs, pterin-dependent aromatic amino acid hydroxylases; AANAT, *N*-acetyltransferase; ALD6, acetaldehyde dehydrogenase; ASMT, *N*-acetylserotonin *O*-methyltransferase; DFR1, dihydrofolate reductase; DHMR, dihydromonapterin reductase; DHPR, dihydropteridine reductase; GCHI/FOL2, GTP cyclohydrolase I; FMO, flavin-containing monooxygenase; NDO, naphthalene dioxygenase; P4Hs, phenylalanine 4 hydroxylases; PCBD/PCD, pterin-4-alpha-carbinolamine dehydratase; PheA, prephenate dehydratase; PTPS, 6-pyruvate-tetrahydropterin synthase; S5Hs, salicylate 5 hydroxylases; SalABCD, salicylate 5-hydroxylase; SAM, *S*-adenosyl-L-methionine; SdaA, L-serine deaminase; SerA, 3-phosphoglycerate dehydrogenase; SPR, sepiapterin reductase; T5H, tryptophan 5-hydroxylase or tryptamine 5-hydroxylase; TDC, tryptophan decarboxylase; TnaA, tryptophanase; TnaB, L-Trp symporter; TPH2, tryptophan hydroxylase 2; TrpB, beta subunit of Trp synthase; TrpA, alpha subunit of Trp synthase; TrpC, bi-functional InGP synthase; TrpD, anthranilate phosphoribosyl transferase; TrpE, anthranilate synthase; TyrA, prephenate dehydrogenase; VioA, L-Trp oxidase; VioB, iminophenyl-pyruvate dimer synthase; VioE, violacein biosynthesis enzyme; VioC, violacein synthase; VioD, tryptophan hydroxylase.

Abbreviations of species: Bs, *Bacillus subtilis*; Bt, *Bos taurus*; Cs, *Comamonas sp.*; Cr, *Catharanthus roseus*; Cv, *Chromobacterium violaceum*; Ct, *Cupriavidus taiwanensis*; Ec, *Escherichia coli*; Hs, *Homo sapiens*; Jl, *Janthinobacterium lividum*; Ma, *Methylophaga aminisulfidivorans*; Os, *Oryza sativa*; Oc, *Oryctolagus cuniculu*; Pa, *Pseudomonas aeruginosa*; Pp, *Pseudomonas putida*; Re, *Ralstonia eutropha*; Rn, *Rattus norvegicus*; Sc, *Saccharomyces cerevisiae*; Sm, *Schistosoma mansoni*; Xc, *Xanthomonas campestris.*

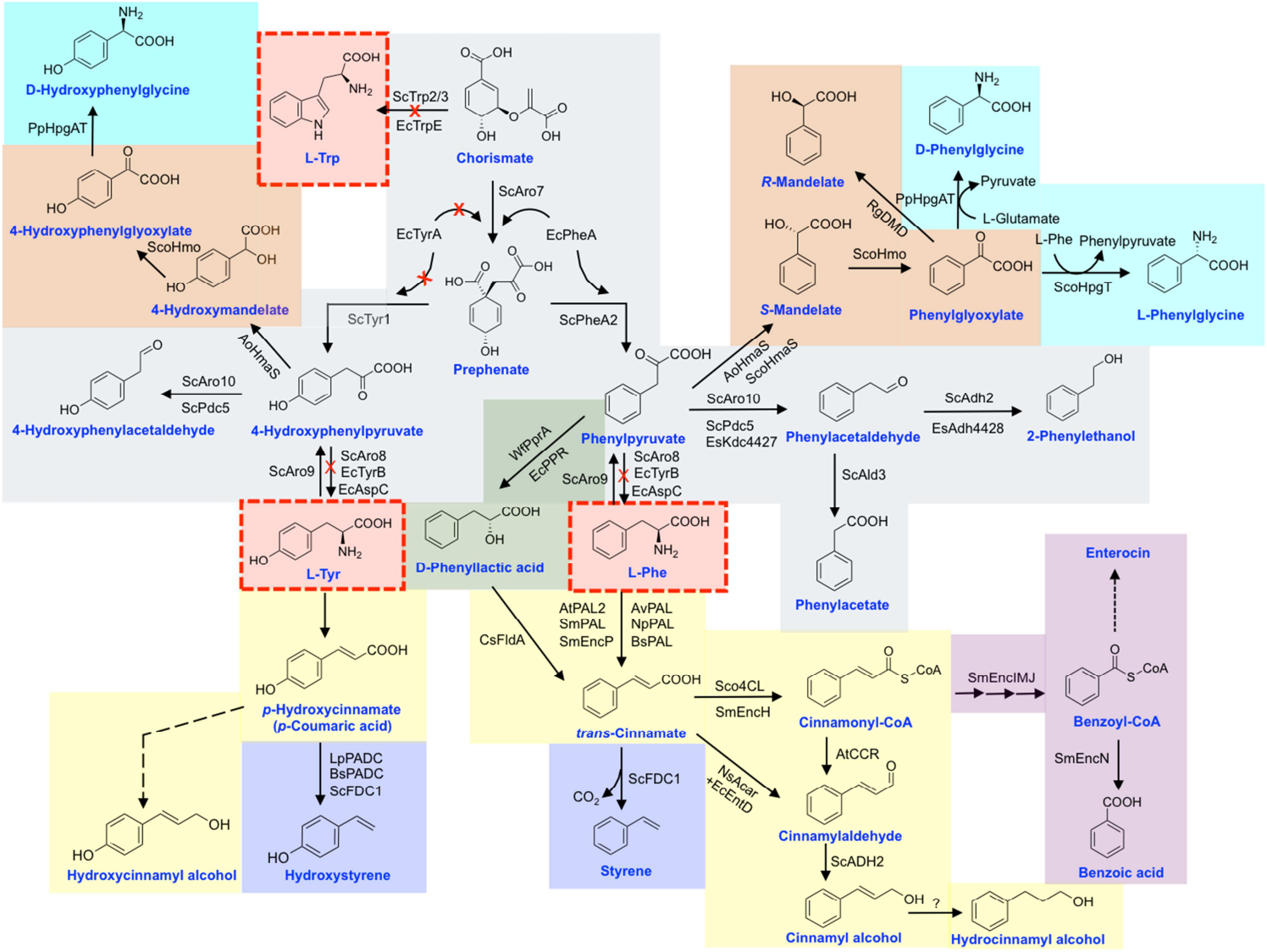


Figure 4. The products derived from the L-Phe branch and their biosynthetic pathways.

Abbreviations of enzymes: 4CL, 4-coumarate:CoA ligase; Acar, arylcarboxylic acid reductase; ADH2/Adh4428, alcohol dehydrogenase; Ald3, aldehyde dehydrogenases; Aro7, chorismate mutase; Aro8/9, aromatic amino acid transaminase; Aro10, phenylpyruvate decarboxylase; AspC, aspartate aminotransferase; CCR, cinnamoyl-CoA reductase; DMD, D-mandelate dehydrogenase; EntD, phosphopantetheinyl transferase; EncH, cinnamate-CoA ligase; EncI, cinnamoyl-CoA hydratase; EncJ, 3-keto-3-phenylpropionyl-CoA thiolase; EncM, β-hydroxyacyl-CoA dehydrogenase; EncN, benzoate-CoA ligase; EncP, phenylalanine ammonia lyase;FDC1, ferulic acid decarboxylase; FldA: cinnamoyl-CoA:phenyllactate CoA-transferase; HmaS, 4-hydroxymandelate synthase; Hmo, 4-hydroxymandelate oxidase; HpgT, L-phenylglycine aminotransferase; HpgAT, D-phenylglycine aminotransferase; Kdc4427, 2-keto acid decarboxylase; PADC, *trans*-cinnamate decarboxylase; PAL, phenylalanine ammonia lyase; PAL2, phenylalanine ammonia lyase 2; PheA/PheA2, prephenate dehydratase; Pdc5, pyruvate decarboxylase; PPR/PprA, phenylpyruvate reductase; Trp2/3, anthranilate synthase; TrpE, anthranilate synthase; Tyr1/TyrA, prephenate dehydrogenase; TyrB, aromatic amino acid aminotransferase.

Abbreviations of species: Ao, *Amycolatopsis orientalis*; At, *Arabidopsis thaliana*; Av, *Anabaena variabilis*; Bs, *Bacillus subtilis*; Cs, *Clostridium sp.*; Ec, *Escherichia coli*; Es, *Enterobacter sp.*; Lp, *Lactobacillus plantarum*; Np, *Nostoc punctiforme*; Ns, *Nocardia sp.*; Pp, *Pseudomonas putida*; Rg, *Rhodotorula graminis*; Sc, *S. cerevisiae*; Sco, *Streptomyces coelicolor*; Sm, *Streptomyces maritimus*; Wf, *Wickerhamia fluorescens*.

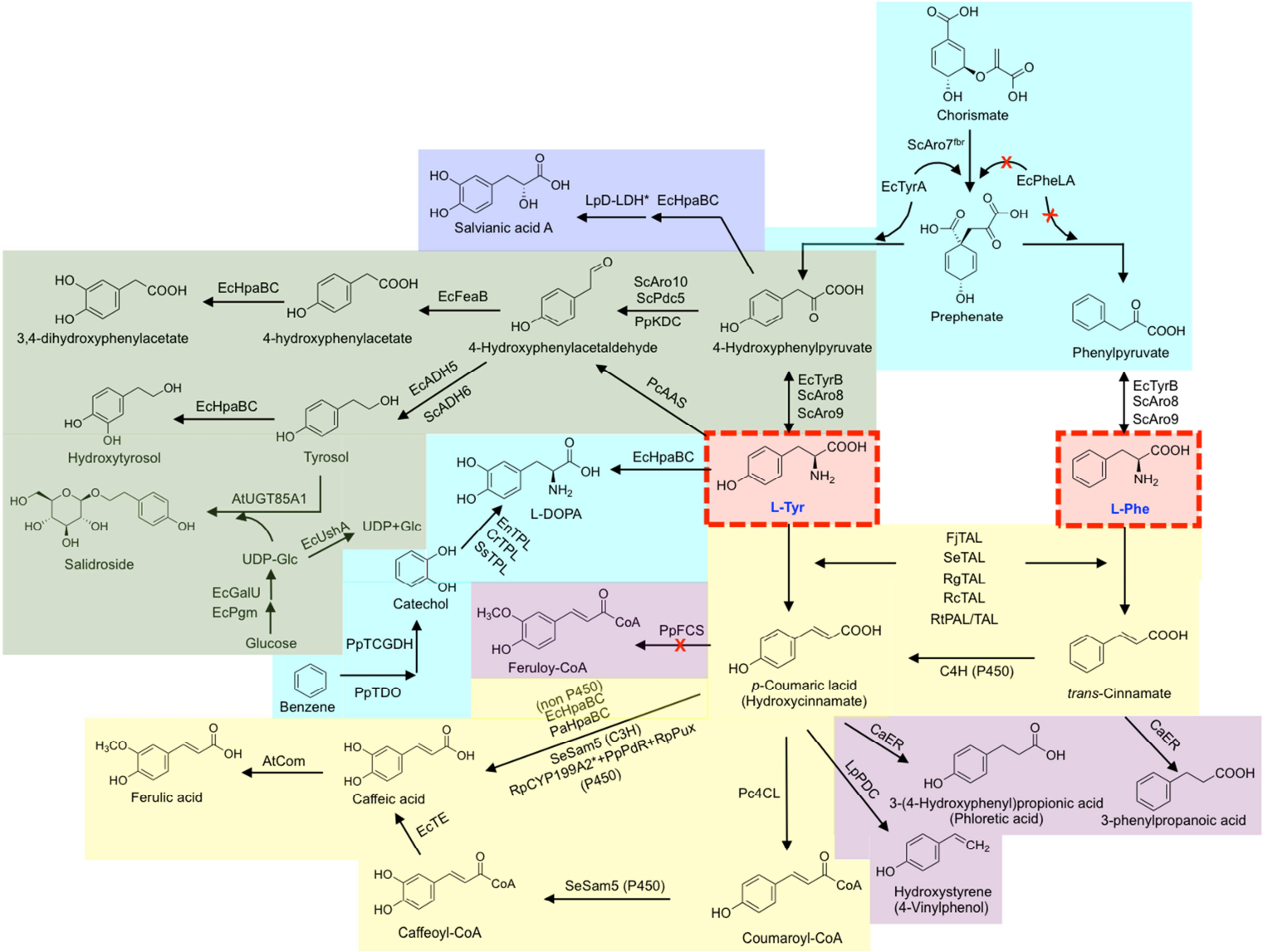


Figure 5. The products derived from the L-Tyr branch and their biosynthetic pathways.

Abbreviations of compounds: Glc, glucose; UDP, uridine diphosphate.

Abbreviations of enzymes: 4CL, 4-coumarate:CoA ligase; AAS, aromatic aldehyde synthase; ADH, alcohol dehydrogenase; Aro7, chorismate mutase; Aro8/9, aromatic amino acid transaminase; Aro10, phenylpyruvate decarboxylase; C4H, cinnamate 4-hydroxylase; C3H/Sam5, P450 4-coumarate 3-hydroxylase; Com, *O*-methyltransferase; CYP199A2, P450 hydroxylase; D-LDH, D-lactate dehydrogenase; ER, enoate reductases; FCS, feruloyl-coenzyme A synthetase; FeaB, phenylacetaldehyde dehydrogenase; GalU, UDP-glucose pyrophosphorylase; HpaBC, nonP450 hydroxylase; KDC, 2-keto acid decarboxylase; PAL, phenylalanine ammonia-lyase; PDC, *p*-coumaric acid decarboxylase; Pdc5, pyruvate decarboxylase; PdR, putidaredoxin reductase; Pgm, phosphoglucomutase; PheLA, prephenate dehydratase and its leader peptide genes; Pux, palustrisredoxin; TAL, tyrosine ammonia-lyase; TCGDH, *cis*-toluene dihydrodiol dehydrogenase; TDO, toluene dioxygenase; TE, CoA thioesterases; TPL, tyrosine phenol-lyase; TyrA, prephenate dehydrogenase; TyrB, aromatic amino acid aminotransferase; UGT85A1, uridine diphosphate-dependent glycosyltransferase; UshA, UDP-sugar hydrolase.

A**bbreviations of species:** At, *Arabidopsis Thaliana*; Ca, *Clostridium acetobutylicum*; Cf, *Citrobacter freundii*; Ec, *Escherichia coli*; Fj, *Flavobacterium johnsoniaeu*; Lp, *Lactobacillus plantarum*; Lpe, *Lactobacillus pentosus*; Pa, *Pseudomonas aeruginosa*; Pc, *Petroselinum crispus*; Rc, *Rhodobacter capsulatus*; Rg, *Rhodotorula glutinis*; Rp, *Rhodopseudomonas palustris*; Rt, *Rhodosporidium toruloides*; Pc, *Petroselinum crispum*; Pp, *Pseudomonas putida*; Ppa, *Pichia pastoris*; Sc, *Saccharomyces cerevisiae*; Se, *Saccharothrix espanaensis*; Ss, *Symbiobacterium* species.

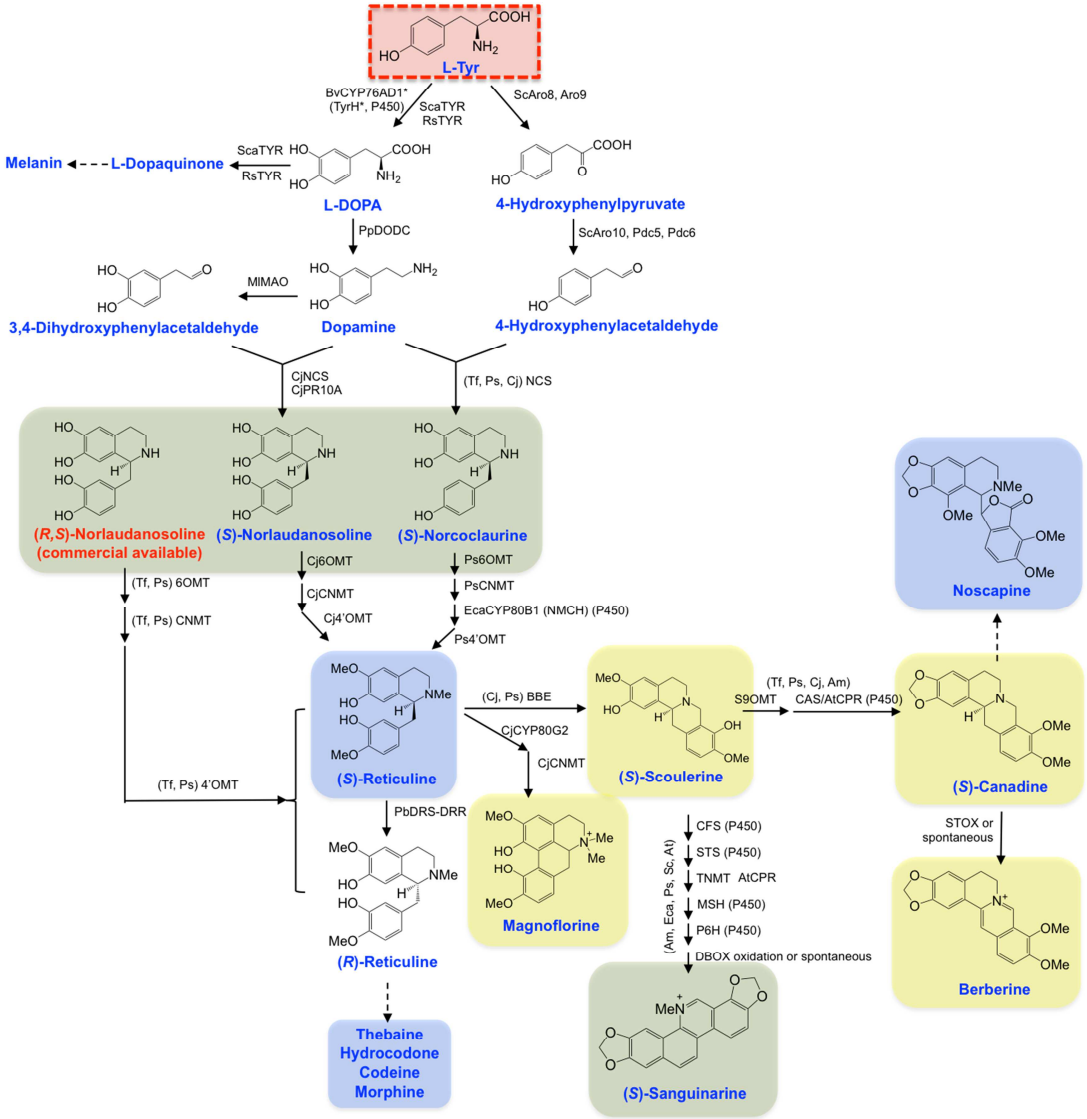


Figure 6. The microbial production of benzylisoquinoline alkaloids derived from L-Tyr.

Abbreviations of enzymes: Aro8, aromatic aminotransferase I; Aro9, aromatic aminotransferase II; Aro10, phenylpyruvate decarboxylase; BBE, berberine bridge enzyme; CAS, canadine synthase; CFS, cheilanthifoline synthase; CNMT, coclaurine-*N*-methyltransferase; CPR, cytochrome P450 reductase; CYP76AD1*(TyrH*), tyrosine hydroxylase mutant; CYP80B1 (NMCH), *N*-methylcoclaurine hydroxylase; CYP80G2, diphenyl ring bridging enzyme; DBOX, dihydrobenzophenanthridine oxidase; DODC, L-DOPA decarboxylase; DRS-DRR, 1,2-dehydroreticuline synthase-1,2-dehydroreticuline reductase; MAO, monoamine oxidase; MSH, *cis*-*N*-methylstylopine 14-hydroxylase; NCS, norcoclaurine synthase; 4′OMT, 3′-hydroxy-*N*-methylcoclaurine 4′-*O*-methyltransferase; 6OMT, 6-*O*-methyltransferase; P6H, protopine 6-hydroxylase; Pdc5 and Pdc6, pyruvate decarboxylase; PR10A, pathogenesis-related 10 protein; S9OMT, scoulerine 9-*O*-methyltransferase; STOX, (*S*)-tetrahydroprotoberberine oxidase; STS, stylopine synthase; TNMT, tetrahydroprotoberberine *N*-methyltransferase; TYR, tyrosinase.

Abbreviations of species: Am, *Argemone mexicana*; At, *Arabidopsis thaliana*; Bv, *Beta vulgaris*; Bw, *Berberis wilsonae*; Cj, *Coptis japonica*; Eca, *Eschscholzia californica;* Ml, *Micrococcus luteus*; Pb, *Papaver bracteatum*; Pp, *Pseudomonas putida*; Ps, *Papaver somniferum*; Rn, *R. norvegicus*; Rs, *Ralstonia solanacearum*; Sca, *Streptomyces castaneoglobisporus*; Sc, *Saccharomyces cerevisiae*; Tf, *Thalictrum flavum.*

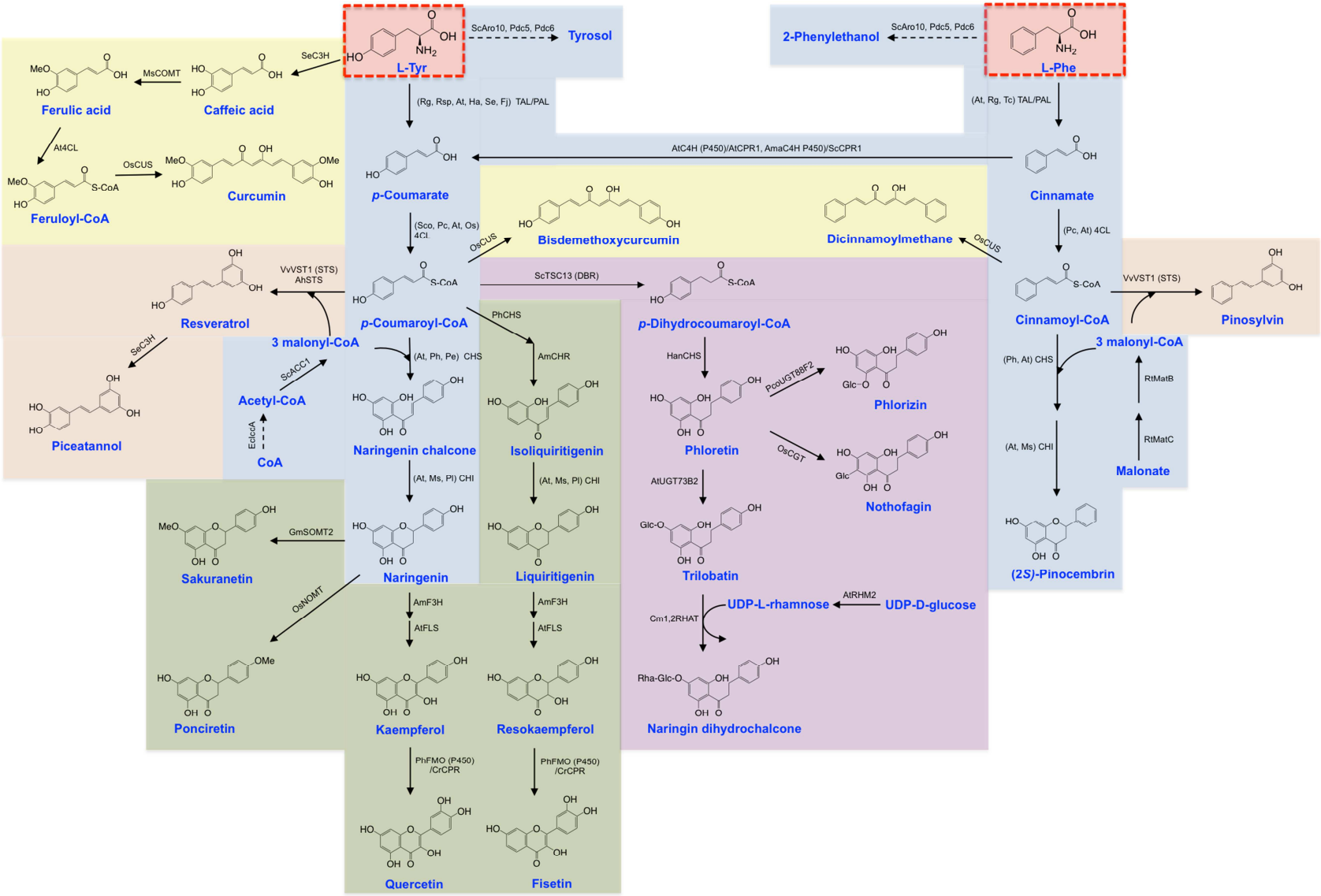


Figure 7. Microbial production of flavonoids and stilbenoids derived from L-Tyr and L-Phe.

Abbreviations of enzymes: ACC1, acetyl-CoA carboxylase; Aro10, phenylpyruvate decarboxylase; 4CL, 4-coumarate-CoA ligase; C3H, 4-coumarate 3-hydroxylase; C4H, cinnamic acid hydroxylase; CHI, chalcone isomerase; CHR, chalcone reductase; CHS, chalcone synthase; COMT, caffeic acid 3-*O*-methyltransferase; CPR, cytochrome P450 reductase; CUS, curcumin synthase; DBR, double bond reductase; F3H, flavanone 3-hydroxylase; FLS, flavonol synthase; FMO, flavonoid 3'-monooxygenase; IcdA, isocitrate dehydrogenase; matB, malonyl-CoA synthetase; matC, malonate carrier protein; NOMT, naringenin 7-*O*-methyltransferase; PAL, phenylalanine ammonia lyase; Pdc5 and Pdc6, pyruvate decarboxylase; 1,2RHAT, 1,2 rhamnosyltransferase; RHM2, UDP-glucose 4,6-dehydratase; SOMT2, soybean *O*-methyltransferase; TAL, tyrosine ammonia lyase; TSC13, very long chain enoyl-CoA reductase; UGT88F2, UGT73B2, CGT, UDP-dependent-glycosyltransferase; VST1 (STS), resveratrol synthase or stilbene synthase.

Abbreviations of species: Ama, *Ammi majus*; Am, *Astragalus mongholicus*; Ah, *Arachis hypogaea*; At, *Arabidopsis thaliana*; Cm, *Citrus maxima*; Cr, *Catharanthus roseus*; Ec, *Escherichia coli*; Fj, *Flavobacterium johnsoniae*; Gm, *Glycine max*; Ha, *Herpetosiphon aurantiacus*; Han, *Hypericum androsaemum*; Ms, *Medicago sativa*; Os, *Oryza sativa*; Pc, *Petroselinum crispum*; Pco, *Pyrus communis*; Pe, *Populus euramericana*; Ph, *Pericallis hybrida*; Pl, *Pueraria lobata*; Rc, *Rhodobacter capsulatus*; Rg, *Rhodotorula glutinis*; Rsp, *Rhodobacter sphaeroides*; Rt, *Rhizobium trifolii*; Sc, *Saccharomyces cerevisiae*; Sco, *Streptomyces coelicolor*; Se, *Saccharothrix espanaensis*; Tc, *Trichosporon cutaneum*; Tp, *Trifolium pretense*; Vv, *Vitis vinifera*.

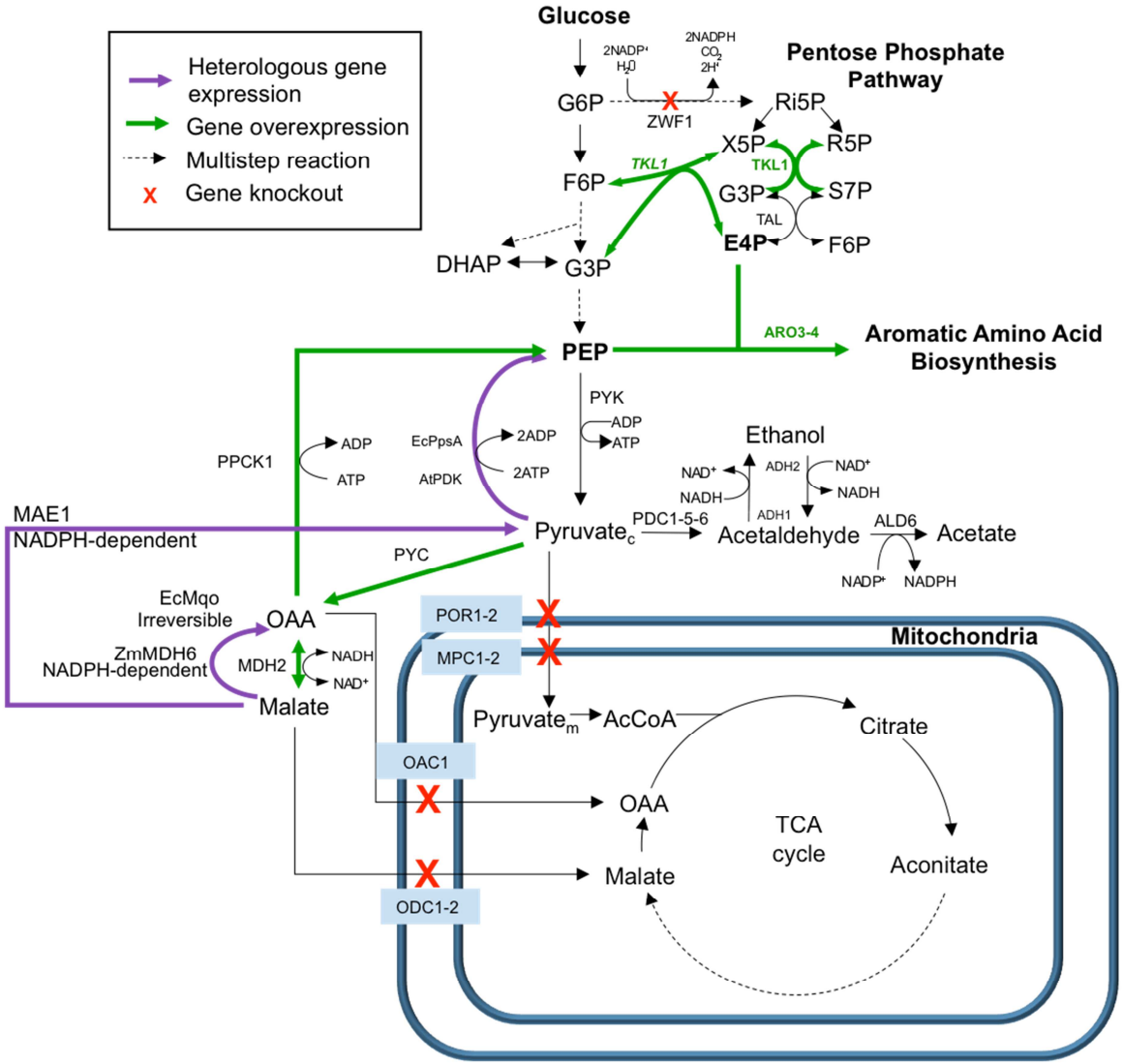


Figure 8. Proposed strategies that could potentially enhance the *de novo* production of aromatic amino acid pathway derivatives in *S. cerevisiae*.

Abbreviations of enzymes and proteins: MAE1, pyruvic-malic enzyme; MDH, malate dehydrogenase; MPC1-2, mitochondrial pyruvate inner membrane carriers; Mqo, malate:quinone oxidoreductase; OAC1, mitochondrial oxaloacetate transport protein; ODC1-2, mitochondrial carriers; PDK, orthophosphate dikinase; POR1-2, outer membrane transporters; PPCK, phosphoenolpyruvate carboxykinase; PpsA, Phosphoenolpyruvate synthase; PYC, pyruvate carboxylase; PYK, pyruvate kinase.

Abbreviations of species: At, *Arabidopsis thaliana*; Ec, *Escherichia coli*; *Zm, Zea mays*.

Highlights

- The aromatic amino acid biosynthesis pathway yields commercially valuable products.
- The broad scope of products from aromatic amino acid biosynthesis is described.
- Biosynthetic pathways and interconnections among the sub-branches are presented.
- High-level organization of enzymes of the upstream module is illustrated.
- The sophisticated endogenous regulations in the upstream pathway at various levels are discussed.