Bioproduction of L-tyrosine and L-tyrosine derivatives by biosensing and modular co-culture engineering approaches

By

ZHENGHONG LI

A dissertation submitted to the

School of Graduate Studies

Rutgers, The state university of New Jersey

In partial fulfilment of the requirements

For the degree of

Doctor of philosophy

Graduate program in Chemical and Biochemical Engineering

Written under the direction of

Haoran Zhang

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May 2020

ABSTRACT OF THE THESIS

Bioproduction of L-tyrosine and L-tyrosine derivatives by biosensing and modular

co-culture engineering approaches

By ZHENGHONG LI

Thesis Director:

Dr. Haoran Zhang

Producing aromatic compounds, especially by using sustainable and environmentally friendly methods, is of great research and application significances. L-Tyrosine is one of 20 standard amino acids and is a key precursor for biosynthesis of a wide range of valuable biochemicals. This thesis research focuses on constructing a microbial L-tyrosine producer and utilizing it as a versatile platform for bioproduction of value-added L-tyrosine derivatives. For developing a L-tyrosine overproducer, key L-tyrosine biosynthesis pathway enzymes were first over-expressed in *E. coli*. Subsequently, a biosensor-assisted cell selection system was established which, via utilization of a tyrosine biosensor protein TyrR, maintained the growth of high performing cells in an isogenic population and repressed the growth of the low performing cells. The experimental results showed that this method resulted in a 5.9-fold improvement of L-tyrosine production. On the other hand, overproduction of tyrosine derivatives, including phenol, 4-hydroxystyrene, caffeic acid and rosmarinic acid, were also investigated. Specifically, modular co-culture

engineering approaches were utilized for high-efficiency biosynthesis of these products. The biosynthetic pathways for these products were divided into separate modules, each of which was contained in one specialized E. coli strain. By using this approach, phenol, 4-hydroxystyrene, caffeic acid, and rosmarinic acid production was improved for 5.3, 2.5, 1.2, and 38 folds, respectively. Moreover, selected biosensors were used in a growth regulation strategy in co-culture system, which was designed to automatically adjust the cell growth behavior based on the tyrosine availability change. For 4-hydroxystyrene and caffeic acid, the integrated use of biosensors and modular co-culture engineering resulted in 2.7 folds and 2.5 folds production enhancement for 4-hydroxystrene and caffeic acid, respectively, compared with co-culture systems without biosensor, and 6.9 folds and 2.9 folds improvement compared with the monoculture controls. The accomplishments of this thesis study demonstrate that biosensing and modular co-culture engineering are valuable tools for future development of metabolic engineer and microbial biosynthesis.

Acknowledgements

Contents

List of figures

List of tables

Chapter 1 Introduction

1.1 Biosynthetic methods for chemical production

Nowadays, production of daily chemicals, especially for the aromatic compounds, highly relies on the petroleum industry, which is not renewable or sustainable. Traditional methods of producing these compounds is also costly, because they often involve high temperature and/or pressure. Also, the solvent and byproducts from the production processes can be detrimental to the environment and people's heath. As such, microbial biosynthesis are considered as a sustainable tool for producing important chemicals, including fuels, commodity chemicals, specialty chemicals, and pharmaceutical chemicals. Actually, it has been a long history for people to use microbes to produce alcohols, cheeses and sauces even before the microbes were discovered in the scientific sense. It is therefore of great significance to take advantage of these cell factories and develop biosynthetic systems better than traditional petroleum-based processes.
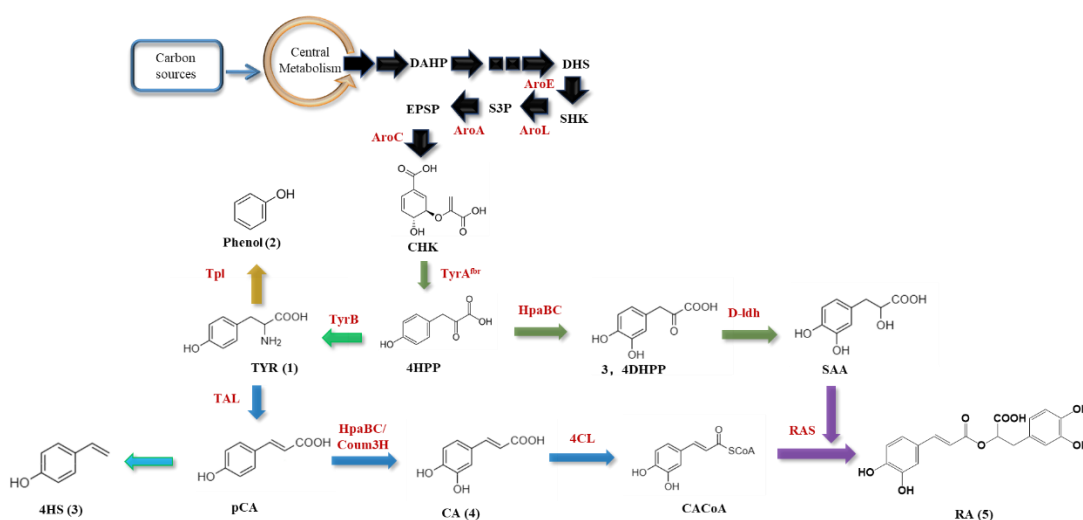
Figure 1.1 Schematic illustration of biosynthesis pathways for the compounds of interest.
DAHP (3-deoxy-D-arabino heptulosonate-7-phosphate); DHS (3-dehydroshikimate); SHK
(shikimate); S3P (SHK-3-phosphate); EPSP (5-enolpyruvyl-shikimate 3-phosphate); CHK
(chorismate); 4HPP (4-hydroxyphenolicpyruvate); 3,4DHPP (3,4-
Dihydroxyphenolicpyruvate); SAA (salvianic acid A); TYR (L-tyrosine); pCA (*p*-coumaric
acid); 4HS (4-hydroxystyrene); CA (caffeic acid); CACoA (caffeoyl CoA); RA (rosmarinic
acid); phenol (phenol).

Among the commonly used microbes, *E. coli* is considered one of the best biosynthetic tools due to the clear genetic background and the fast growth rate. Also, *E. coli* is easy to manipulate and there are many available tools to engineer *E. coli* for a specific biosynthetic system. This thesis investigated the biosynthesis of a variety of important chemicals including simple compounds such as 1) tyrosine, 2) phenol, 3) hydroxystyrene 4) caffeic acid, and complicated nature products 5) rosmarinic acid. (Figure 1.1).

1.2 Thesis objectives

The overarching goal of this thesis research is to develop robust biosynthetic systems for overproducing desired compounds with high performance. For this, traditional metabolic engineering tools as new engineering strategies, such as modular co-culture engineering and biosensing, were introduced for enhancing the biosynthetic ability. To this end, a tyrosine over-producing platform will be first established by constructing bacterium *E. coli* strains with an engineered biosynthesis pathway and a biosensor assisted cell selection system. A series of downstream enzymatic steps were then introduced to *E. coli* to convert tyrosine to desired products. Namely, this study reconstituted the biosynthesis pathways for heterologous production of value-added biochemicals shown in Figure 1.1.
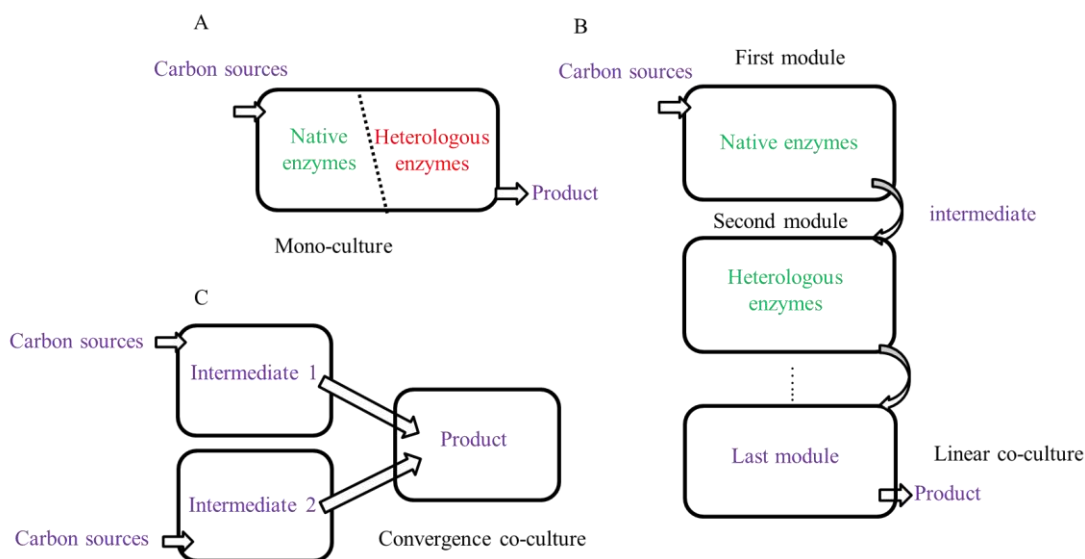
Figure 1.2 Monoculture and co-culture design. A) Monoculture design. B) Co-culture engineering in a linear form.C) Co-culture design in convergence form. Different host strains were constructed for different modules of the biosynthesis pathway

1.2.1 Modular co-culture engineering

Modular co-culture engineering is a newly developed approach for microbial biosynthesis, which utilizes microbial co-cultures to accommodate modularized biosynthesis pathways. For biosynthesis of the selected products of this study [1, 2], all the involved biosynthesis pathways were divided into separate modules, respectively (Figure 1.2B and C). Each module was only responsible for one portion of the bioproduction labor. Accordingly, a series of co-cultures were rationally designed and constructed to accommodate the modularized pathways. Each co-culturecontains multiple specialized *E. coli* strains that were engineered to harbor the assigned biosynthetic pathway modules. The bioproduction by the co-cultures was systematically optimized by changing several factors, including inoculation ratio between the constituent co-culture strains, co-culture cultivation conditions, the limiting step of the pathway, etc. The engineered co-cultures were characterized to gain insights for the co-culture growth and biosynthesis

behaviors. Specifically, the time profiles for strain-to-strain ratio, overall growth, concentration of pathway substrates, intermediates, and products, overall production yield, were analyzed and compared.
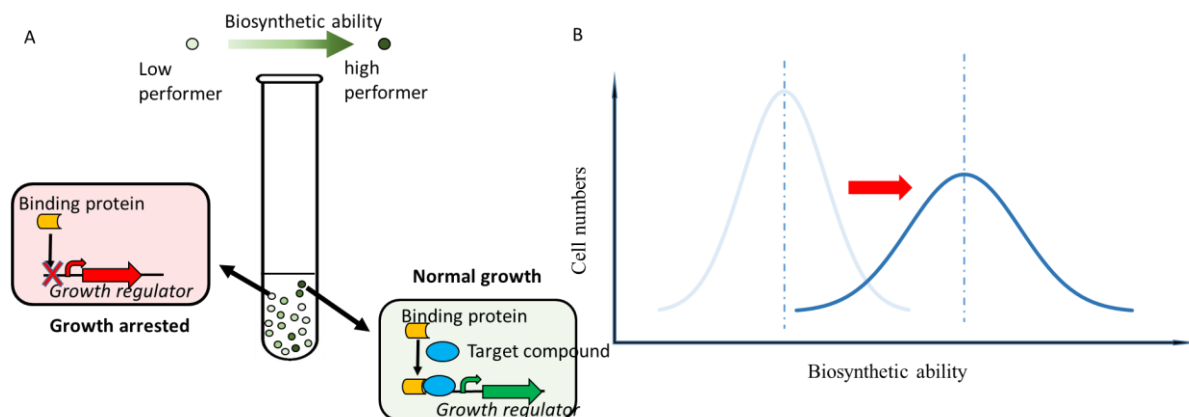
Figure 1.3 Biosensor assisted high performing cell selection system. A) Mechanism of biosensor assisted cell selection system on high and low performing cells. B) Population shift for biosensor assisted selection system.

1.2.2 Biosensor-assisted high performing cell selection

Using biosensors to select high performers among a cell population is another emerging strategy that can effectively improve the production. The utilization of biosensor is based on the fact that cells have different biosynthetic abilities even though they have the same genome configuration (the isogenic cells are naturally different in terms of bioproduction performance). This can be explained by the following reasons [3]: 1) The gene copy numbers can be different from cell to cell; 2) The epigenetic modifications of each cell is not same; 3) The stability and activity of mRNA are varied; 4) There is stochastic gene expression. To select the cells with high biosynthetic ability, the biosensor systems are designed to maintain the growth and propagation of high performing cells and repress the low performing cells. They consist of three parts: signals, signal processing systems and responses. In the biosensor assisted selection systems, signals are the

concentrations of the target compounds that bind with biosensors, and responses are usually the cell growth conditions. For signal processing, a biosensor binds with the target metabolite and acts on a promoter that can specifically turn on/off the expressionof an antibiotic resistance gene (on-switch promoter) or a toxin gene (off-switch promoter). Based on this design, the high performing cells with higher concentration of the target compound can turn on the expression of antibiotic resistance gene or turn off the expression of toxin to maintain a normal growth status. Instead, low performing cells are repressed for growth by the antibiotic or the toxin, which leads into a population shift in favor of the high performing cells, as suggested in Figure 1.3B [3, 4].

1.2.3 Biosensor-based growth regulation system in co-culture engineering

A biosensor can also be used as the growth regulator in co-culture engineering for dynamic balancing of the biosynthetic pathway. Specifically, pathway intermediate concentration is used as the signal and downstream strain growth is used as response. The signal processing system is the similar to the biosensor-assisted selection system. When the pathway intermediate concentration is lower, only upstream strains can grow and produce intermediate. Downstream strain only grows and converts intermediate after the accumulated concentration of the target intermediate is high enough to de-repress the cell growth via the sensing and gene expression regulation function of the selected biosensor. When the intermediate concentration is lower due to the downstream strain consumption, upstream cell gains growth priority and starts to accumulate intermediate again, as suggested in Figure 1.4.
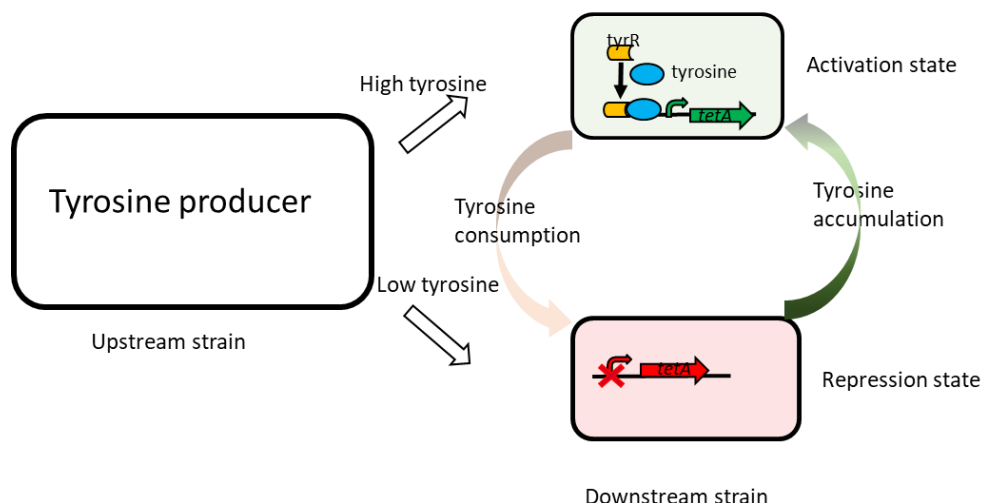
Figure 1.4 Mechanism of biosensor-assisted cell growth regulation system within a co-culture.

This growth regulation system can maintain the upstream strain advantageous when the concentration of intermediate is low and promote the growth of the downstream strain when intermediate concentration is high. As such, the biosynthetic pathway can be dynamically balanced for bioproduction optimization.

1.2.4 In-situ removal of product

As some biosynthesis products are toxic to cell growth, an in-situ extraction method was also tested in this research. There are many advantages for using this in-situ extraction method [5, 6]. Firstly, the organic solvent can reduce the concentration of the product compound in aqueous phase and remove the toxicity of product (e.g. phenol. Secondly, in-situ removal of product reduces the product concentration in the aqueous phase, which pushes the enzymatic conversion equilibrium toward the final compounds to improve the overall biosynthetic

ability. Also, this in-situ product removal strategy is effective for alleviating the feedback control of enzymatic reactions .

1.3 Significance of the thesis

The significance of this research are several folds. First, the investigation of tyrosine-derived biochemical production generated new knowledge for microbial biosynthesis of these valuable product with high efficiency. In particular, the inclusion of heterologous enzymes and their functional expression in *E. coli* promoted the understanding about the performance of individual pathway enzymes and their collective behaviors in heterologous host.

Second, from the perspective of practical application, the success of this study paved the way for large scale bioproduction of the involved products. Especially, the bioreactor production study offered important protocols for high cell density, high substrate consumption and high product production operations. This in turn improved the availability of these products using a renewable, sustainable, and cost-effective method.

Third, this study provides critical knowledge for using modular co-culture engineering to address the challenges of conventional mono-culture engineering. To this end, the production advantages associated with engineered co-cultures were highlighted and the corresponding results showed how the co-cultures led to better biosynthesis performance than the mono-culture controls.

Lastly, this research crosses several research areas such as metabolic engineering, biosensing, synthetic biology and bioprocess engineering, and leverages the power of particular areas to others. As such, the pursuit of this work

promoted synergistic advances of these research areas, which is highly significant for the future development of new engineering tools and methodologies.

1.4 Thesis organization

Chapter 2 describes the construction of a tyrosine overproducing platform using the biosensor assisted selection systems. Different biosensors and toxins were used to establish effective selectin systems.

Chapter 4 and 5 disucss biosynthesis of tyrosine derivative phenyl. The tyrosine producer constructed in Chapter 2 was used for phenol production using a co-culture strategy. Another pre-constructed tyrosine producer P2H was used for phenol production from glycerol in monoculture by in situ product removal strategy with resins, followed by the catalytic reaction to produce alkylated phenol.

In Chapter 6, the aforementioned tyrosine producer P2H was used as the upstream strain for producing caffeic acid and 4-hydroxystyrene using a growth regulated biosensor in co-culture engineering. Downstream strains with the tyrosine biosensor can adjust the growth automatically according to the concentration change in the culture and thus facilitate the biosynthesis optimization.

Biosynthetic pathways in Chapter 4 and 6 are all linear pathways using L-tyrosine as an intermediate for the co-culture systems, while some natural products involve more complicated structure and convoluted pathways. In Chapter 7, biosynthesis of rosmarinic acid, an ester of caffeic acid and salvianic acid A, was accomplished. The involved divergence-convergence pathway is highly challenging engineered with traditional metabolic engineering tools but

0offers a great platform to demonstrate the power of modular co-culture engineering.

Finally, Chapter 8summarizes the impacts of the experimental results and discusses future study directions.

Chapter 2 L-Tyrosine biosynthesis

2.1 Introduction

2.1.1 Backgrounds

L-Tyrosine (*(2S)-2-amino-3-(4-hydroxyphenyl) propionic acid*) is one of 20 standard amino acids that are used for protein biosynthesis in cells. It has broadened use in multiple fields due to its good bioactivity. As a dietary supplement, L-tyrosine was found advantageous for people suffering stress, cold, fatigue and sleep deprivation symptoms. It is also a vitiligo alleviator as well as an analogue to neurotransmitter and hormone, which leads to high medical value.

Besides, L-tyrosine is also a versatile precursor for a series of aromatic compounds that are widely applied in many different industries [7]. Traditionally, those aromatic derivatives mainly rely on the production of petroleum-based industry. To this end, this thesis research aims to establish modular co-culture systems for production of L-tyrosine in *E. coli*, which provides us a more sustainable and environmentally friendly access to those aromatic compounds.

Industrial L-tyrosine production relies on the protein hydrolysis in early times. In 1820, Braconnot first extracted L-glycine and L-tyrosine from lamp muscle hydrolysis solutions. Proteins were treated with acid for amino acids and then the amino acids were extracted by ion-exchange resins. This method is limited by the availability of raw material, complexity of reaction and separation technology and long manufacturing period. Three substitute methods are widely used right now.

1) Enzymatic reaction method.

This method utilizes tyrosine phenol lyase (TPL) from microbes to convert phenol, pyruvate, ammonia or phenol, L-serine to L-tyrosine. TPL with high enzymatic reaction activity from *Erwinia herbicola*, *Citrobacter intermedius* and *Citrobactet freundii* was well studied. *Klebsiella aerogenes* and *Erwinia herbicola* were first used by Lee and Hsio to produce L-tysoine from a two-step reaction from L-glycine. After 16 h of reaction, 26.3 g L-tyrosine was produced at a 61.4% conversion rate of L-glycine. However, this production system fluctuated a lot due to the high inhibition effect of L-glycine to TPL. Given the low activity and stability of TPL, molecular biology tools for modifying the DNA drew a lot research focus. Eugene form KRIBB improved the activity of TPL by high throughput screening from the library generated by random mutation and DNA shuffling [8]. DNA sequencing results suggested the mutation T129I and T451A occurred on the functional region of the enzyme and A13V, E83K and T407A helped to improve the thermostability. In vitro experiments were performed using the cell culture supernatant achieved 130 g/L of L-tyrosine production and 94% conversion of phenol.

2) Fermentation

This method used selected microbes to convert carbon sources such as glycerol, glucose and xylose to L-tyrosine by fermentation [9]. Early research involved induced mutation for high L-tyrosine producing strains by screening strains with feedback control resistance to L-tyrosine and L-phenylalanine. However, most microbes lacked the ability of large-scale accumulation of L-tyrosine and traditional induced mutation methods were not able to modify all L-tyrosine biosynthesis pathways. Recent research utilized metabolic engineering

tools to redesign the metabolic pathways of L-tyrosine in *Escherichia coli*, *Corynebacterium glutamicum*, *Brevibacterium flavum* and *Bacillus subtillis*. This is similar to our research for building L-tyrosine overproduction platform.

3) Chemical synthesis

Although chemical synthesis strategy of L-tyrosine was design in 19th century, the method was widely used after 1950s. Organic synthesis [10] of amino acids was not limited to L-tyrosine; it can also produce unnatural amino acid with special structures. Chemical synthesis produced both D-tyrosine and L-tyrosine, although further separation effort isneeded. This method is still widely used today and produces million tons per year.

2.1.2 L-Tyrosine biosynthesis pathway

E. coli has been proved one of the most robust heterologous host for L-tyrosine overproduction due to the amenability in DNA manipulation and high versatility in suiting the need of various gene expression. L-tyrosine production in *E. coli* involves multiple native pathways. Figure 2.1 shows the biosynthesis of L-tyrosine from different carbon sources. In this study, L-tyrosine biosynthetic pathways enzymes were modified and overexpressed in vivo.
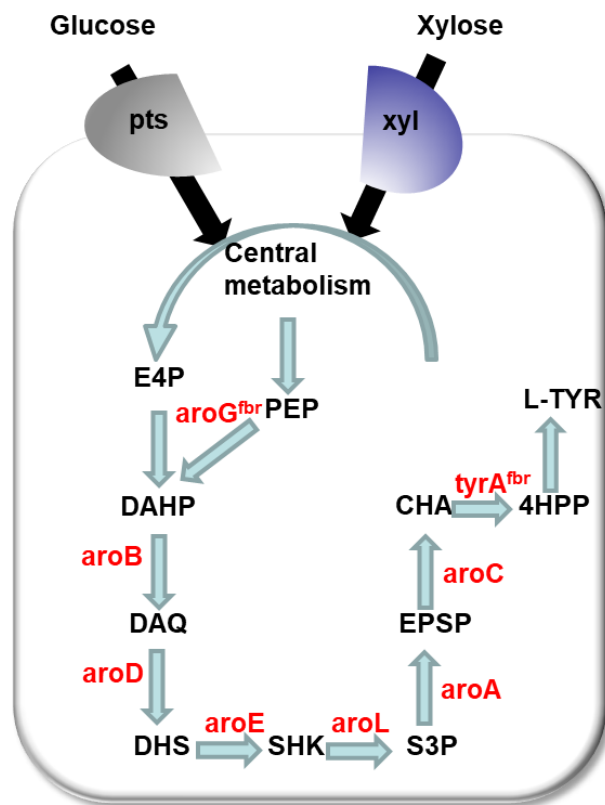
Figure 2.1 Schematic presentation of L-tyrosine biosynthesis pathway. E4P (erythrose-4-P); PEP (phosphoenolpyruvate); DAHP (3-deoxy-D-arabino heptulosonate-7-phosphate); DHQ(3-dehydroquinate); DHS (3-dehydroshikimate); SHK (shikimate); S3P (SHK-3-phosphate); EPSP (5-enolpyruvyl-shikimate 3-phosphate); CHA (chorismate); TYR (L-tyrosine). PTS: glucose uptake system, xyl: xylose uptake system.

E. coli is naturally capable of converting several different carbon sources to aromatic amino acids via the shikimate pathway. This thesis research used D-glucose and D-xylose as carbon sources. D-Glucose and D-xylose can be obtained by hydrolysis of lignocellulosic biomass, so they are available in large quantities and low cost. More importantly, use of these renewable carbon substrates generates less pollution compared to the petroleum industry. For catabolism of D-glucose and D-xylose, these sugars are first converted to D-glucose 6-phosphate (G6P) and D-xylulose 5-phosphate (X5P), which are subsequently introduced to both glycolysis pathway and pentose phosphate pathway (PPP). The resulting

intermediates phosephoenolpyruvate (PEP) and D-erythose 4-phosphate (E4P) combine with each other and enter the shikimate pathway, forming chorismate. From chorismate, it branches to 3 aromatic amino acids L-tyrosine, L-tryptophan and L-phenylalanine.

2.1.3 Previous work for L-tyrosine biosynthesis in *E. coli*

There have been extensive efforts for L-tyrosine overproduction in *E. coli* for in the past. Significant success has been achieved for converting renewable feedstocks to L-tyrosine employing various engineering strategies.

Santos et al. developed a high-throughput screening method for L-tyrosine production, resulting in 0.204 g L-tyrosine/g glucose [11]. Further adaption of the global transcription machinery engineering improved the production to 9.7 g/L [12]. Na et al. used an sRNA inhibition method to repress the competing pathway of L-tyrosine, which, in combination with the conventional methods for overexpressing L-tyrosine biosynthesis pathway genes, achieved a yield of 0.1 g L-tyrosine/g glucose [13]. Juminaga et al. obtained a high L-tyrosine yield by overexpressing nearly all L-tyrosine pathway genes and optimization of the promoters and copy numbers for the involved genes, leading to 2.17 g/L L-tyrosine production and 0.43 g L-tyrosine/g glucose yield [14]. Xiao et al. established a novel biosensor based population quality control (PopQC) method using selection pressure to repress the growth of low L-tyrosine producers in the population, resulting in a 0.05 g L-tyrosine/g glucose yield [3].

2.2 Experimental design

To establish a L-tyrosine producer as a platform for biosynthesis of other L-tyrosine derivatives, this study overexpressed a series of genes including *aroB, aroD, aroE, aroL, aroA, aroC*. Also, *aroG* and *ayrA* genes were modified to generate feedback control resistance to yield *aroG^{fbr}* and *tyrA^{fbr}* [15]. To further elevate the production of L-tyrosine, the biosensorassisted selection system was introduced.

In previous design [3], tetracycline was added as the selection pressure and the *E. coli* subpopulation with high L-tyrosine production can activate the expression of tetracycline exporter gene *tetA*. This leads to better survival for the high performing cells compared with low performing cells as shown in Figure 2.2A. However, this design can be problematic when applied to larger scale production, because 1) the addition of tetracycline results in high process cost; 2) tetracycline concentration per cell is decreased when the cell culture grows into a high cell density level and thus reduces the selection pressure and 3) when L-tyrosine concentration is too high and oversaturates the biosensor, the biosensor switch are kept on at the max level and loses the function of regulating the gene expression. This thesis work utilizes the similar concept of the PopQC method but proposes a new approach to achieve L-tyrosine bioproduction in *E. coli*.

exporter[18]. The purpose here is to reduce the intracellular concentration of L-tyrosine so that it can be maintained within the sensing range of the biosensor. Even when the extracellular concentration is high, the exporter can keep the intracellular concentration of L-tyrosine at a low level, which alleviate the oversaturation issue of biosensor.

2.3 Materials and methods

2.3.1 Plasmids and strains

All *E. coli* strains as well as plasmids used in this study are presented in Table 2.1. Primers used in this study were listed in Appendix.

Table 2.1 Plasmids and strains used in Chapter 2

Plasmids	Description
pB1	pACYCDuet-1 carrying the *E. coli aroB* gene under the control of the proD promoter (PproD)
pBD	pACYCDuet-1 carrying the *E. coli aroB* and *aroD* genes under the control of the proD promoter (PproD)
pBDE	pACYCDuet-1 carrying the *E. coli aroB, aroD* and *ph*pCAT genes under the control of the proD promoter (PproD)

pSE1	pET21c carrying the *hipA* gene under the control of the *mtr* promoter (P*mtr*)
pBS2	pET28a carrying the proD promoter (PproD) and the *aroE*, *aroL*, *aroA*, *aroC, tyrA^{fbr}* and *aroG^{fbr}* genes
pBS8	pET21c carrying the *aroP* promoter
pBS9	pBS8 carrying the *E. coli hipA* gene

Strains	Description
TM2	*E. coli* BL21(DE3) carrying pET21c
BST	*E. coli* BL21(DE3) carrying pBS9
TPS1	*E. coli* BL21(DE3) carrying pBS2 and pBS9
TPR1	*E. coli* BL21(DE3) carrying pBS2 and pET21c
TPS2	*E. coli* BL21(DE3) carrying pBS2, pBD and pBS9
TPS3	*E. coli* BL21(DE3) carrying pBS2, pBDE and pBS9

To construct an L-tyrosine producer, a strong constitutive promoter proD was used [19]. A previously constructed plasmid pPH0-1 [4] was adapted for over-expression of *aroE, aroL, aroA* and *aroC* genes under the control of promoter proD. A DNA fragment containing the genes *tyrA^{fbr}* and *aroG^{fbr}* was PCR

amplified with primers ZLPR1TA and ZLPR2TA using the *E. coli* P2H chromosomal DNA as the template. The PCR product was digested with SpeI/HindIII followed by ligation with pPH0-1 treated with the same enzymes to make plasmid pBS2. For pBDE construction, a commercial synthesized DNA fragment of gene aromatic amino acid *phpCAT* was digested with HindIII and XhoI and ligated to plasmid pBD treated with the same restriction enzymes.

To select for the high L-tyrosine producers, the promoter of the *E. coli aroP* gene was utilized. The promoter fragment was PCR amplified with primers ZLPR1AP and ZLPR2AP using K12(DE3) [12] chromosome as the template and assembled to pET21c by SphI/NdeI sites to generate pBS8. Plasmid pBS9 was a pBS8 derivative with inclusion of the *E. coli hipA* gene by digesting both pBS8 and a previously constructed plasmid pSE1[4] using NdeI and XhoI sites.

2.3.2 Cultivation conditions

All *E. coli* strains were cultivated in 3 mL MY1 medium in 37 $^\circ$C at 250 rpm. 1 L MY1 medium was comprised of 5g glucose, 0.5 g yeast extract, 2.0 g NH_4Cl, 5.0 g $(NH_4)_2SO_4$, 3.0 g KH_2PO_4, 7.3 g K_2HPO_4, 8.4 g MOPS, 0.5 g NaCl, 0.24 g $MgSO_4$, 40 mg L-tyrosine, 40 mg phenylalanine, 40 mg tryptophan, 10 mg 4-hydroxybenzate and trace elements. The working concentrations of trace elements were 0.4 mg/L Na_2EDTA, 0.03 mg/L H_3BO_3, 1 mg/L thiamine, 0.94 mg/L $ZnCl_2$, 0.5 mg/L $CoCl_2$, 0.38 mg/L $CuCl_2$, 1.6 mg/L $MnCl_2$, 3.77 mg/L $CaCl_2$, and 3.6 mg/L $FeCl_2$ [20, 21]. The antibiotics were used in the following concentration: 50 mg/L kanamycin, 34 mg/L chloramphenicol and 100 mg/L ampicillin.

For L-tyrosine producer's cultivation, 2 % (v/v) overnight LB cultures of the desired *E. coli* strains were inoculated in MY1 medium with necessary antibiotics

and incubated in 37 °C for 10 h. The cells were then harvested through centrifugation and re-suspended in the fresh MY1 medium with an initial OD_{600} of 0.6. After 48 h cultivation, the culture samples were taken for HPLC analysis.

For the L-tyrosine biosensor-assisted cell selection system characterization, strain BST and TM2 were constructed by transformation of plasmid pBS9 and pET21c into BL21(DE3), respectively. To test the growth regulation without the biosensor system, overnight culture with OD_{600} of 0.3 was inoculated into fresh MY1 medium containing 2 g/L glucose. OD_{600} was measured after 18 h cultivation.

2.3.3 Metabolites quantification

Quantification of the pathway metabolites was conducted using Angilent 1100 HPLC with a DAD detector. 1.0 mL culture sample was centrifuged at 10000 rpm for 5 min, and the supernatant was filtered through 0.45 μm polytetrafluoroethylene membrane syringe filters (VWR International). 10 μL of filtered sample was injected into a column from ES Industries Inc. (HyperSelect ODS Plus C18 column 4.6 × 150 mm, 5 μm) for L-tyrosine quantification. The following gradient was utilized for elution: 0 min, 100% solvent A; 5 min, 95 % solvent A; 6 min, 75% solvent A; 10 min, 10% solvent A; 11-16 min 100% solvent A.

2.4 Results and discussion

2.4.1 Biosensor-assisted selection system for L-tyrosine over-production

E. coli strain TPR1 was constructed as L-tyrosine over-producer. This strain was engineered to over-express key enzymes of the tyrosine biosynthesis pathway

from a medium copy plasmid pBS2. The TPR1 culture was grown on 5 g/L glucose for L-tyrosine bioproduction. As shown in Figure 2.3A, 136 mg/L of L-tyrosine was produced after 48 h cultivation. On top of this efforts, a biosensor-assisted high performing cell selection system was utilized. Specifically, an *E. coli* toxin gene *hipA* was placed under the control of the *E. coli*'s *aroP* gene's promoter. This promoter represses gene expression in the presence of L-tyrosine through the regulation by transcriptional regulator TyrR. The activation of the hipA gene expression generate toxic product and inhibit the growth of the host strain. The constructed P*aroP-hipA* operon was introduced into the BL21(DE3) strain with an intact chromosomal *tyrR* gene. Based on this design, high concentration of L-tyrosine in the high-performing cells represses the toxic *hipA* gene expression and thus do not disrupt normal cell growth. In comparison, the low-performing cells' growth should be inhibited due to the unrestricted expression of toxic *hipA* gene. As a result, the population of the engineered upstream strain would be dominated by high-performing cells for enhancing production of L-tyrosine.

After the establishment of the biosensor-assisted cell selection system, an L-tyrosine responses test was performed for the resulting *E. coli* BST. Specifically, cell growth in the presence of different concentration of L-tyrosine was analyzed. *E. coli* strain TM2 without the biosensor-assisted selection system was constructed as the control group. As shown in Figure 2.3B, L-tyrosine concentration had no significant influence on cell growth in the control strain without the biosensor. In comparison, for the strain with the biosensor, the cell density exhibited an increasing trend as the L-tyrosine concentration increase. These results clearly confirmed that the constructed biosensor-assisted cell

selection system indeed had the desired L-tyrosine sensing and growth regulation functions.
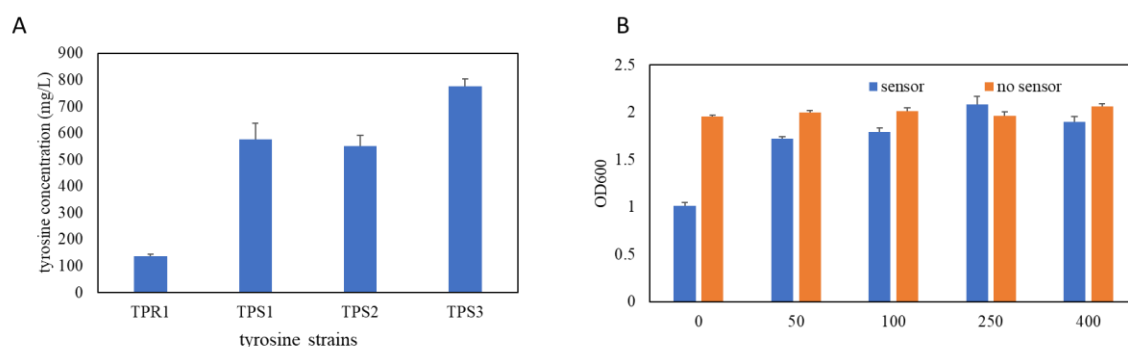
Figure 2.3 Engineering E. coli for L-tyrosine production. A) L-Tyrosine production of different *E. coli* strains. B) Correlation between cell growth and different concentrations of L-tyrosine. *E. coli* strains without and with the biosensor-assisted cell selection system were compared.

In order to examine the functionality of biosensor-assisted cell selestion for supporting L-tyrosine production, plasmid pBS9 harboring the biosensor system was used to generate the new L-tyrosine producing strain TPS1. It was found that the production of L-tyrosine was significantly improved to 577 mg/L, demonstrating the strength of the adopted cell selection strategy. Compared to TPR1, L-tyrosine production in strain TPS1 was improved by 4.4 folds. Further modifications of the TPS1 strain was also attempted by over-expressing another 2 pathway genes *aroB* and *aroD*. The resulting TPS2 strains produced 552 mg/L of L-tyrosine, indicating no significance production improvement. The reasons that L-tyrosine production was not improved can be the following. First, biosensor-assisted cell selection system was saturated at high tyrosine concentration in the culture. As suggested in Figure 2.3B, when L-tyrosine concentration was beyond 400 mg/L, the cell growth showed no obvious response to higher L-tyrosine concentration. The possible explanation is that biosensor was already activated to the maximum level due to the high tyrosine concentration,

and further increase of tyrosine concentration did not make any positive impact on cell growth regulation. On the other hand, enzymatic conversion catalyzed by AroB and AroD might not be the rate limiting steps compared to the other pathway genes. Overexpression of these genes thus did not benefit the overall production.

To address the biosensor saturation issue at high L-tyrosine concentration, the aromatic amino acid exporter *PhpCAT* [18] was overexpressed using plasmid pBD. Plasmid pBDE was then transformed into *E. coli* and the resulting strain TPS3 was used for tyrosine production. It was found that the L-tyrosine production by this strain reached 775 mg/L which is 5.9 folds higher of the starting strain TPS3. In fact, the aromatic amino acid exporter enabled the cell to maintain a lower level of intracellular concentration of L-tyrosine. Thus, the biosensor was not saturated, and the cell selection system could work as expected to select for high tyrosine producing cells. Notably, further analysis of the intracellular L-tyrosine concentration change before/after the exporter was introduced will be helpful to characterize the effectiveness of the adopted strategy.

2.4.2 investiate the use of other tyrosine biosensor and growth regulators

The biosensor-assisted selection method was proved a new methodology in metabolic engineering. To expand its application, other metabolite-responsive gene promoters and growth regulators were examined for the use in tyrosine over-production. The promoters and regulators tested are listed in Table 2.2.

Table 2.2 Summary of the promoters and growth regulators that can be used for biosensor-assisted cell selection system

Sensing system	target molecules	Effectiveness

arop-TyrR	L-Tyrosine	Effective
arop-TyrR	L-Phenylalanine	Effective
mtr-TyrR	L-Tyrosine	Effective
mtr-TrpR	L-Tryptophan	Effective
mphA-MphR	Erythromycin	Effective
dmp-DmpR	Phenol	Effective
nah-NahR	Anthranilic acid	Effective
Growth regulators	Inhibition mechanism	Effectiveness
sRNA (pyrH/dnaE/fabA)	Providing oligonucleotides binding on transcriptional initial region	Not effective
sRNA (rpoC)	Providing oligonucleotides binding on transcriptional initial region	Slightly effective
asRNA (rpoC)	Providing oligonucleotides binding on transcriptional initial region	Not effective
gltX	Gene mischarging *E. coli* tRNA$_1^{Gln}$ with glutamate	Slightly effective
hipA	Gene inhibiting the macromolecular synthesis	Effective
mltB	Gene causing cell lysis	Effective

Arop-TyrR sensing system has good response to L-tyrosine and L-phenylalanine as an off-switch biosensor [16, 22]. Promoter of gene mtr, which is an L-tyrosine transporter, is a L-tyrosine-responsive and acts as an on-switch

for tyrosine and an off-switch for L-tryptophan with the presence of TyrR and TrpR. respectively [16]. Promoters of regulons *mph* [23, 24] and *dmp* [25, 26] were verified to be responsive to erythromycin and phenol with MphR and DmpR, respectively. The *nah* regulon promoter [27, 28] can be used to design the on-switch biosensor selection system for anthranilic acid. For on-switch biosensors, antibiotics resistance genes can be used as a growth regulator. In the cases of off-switch biosensors, toxins can be used as growth inhibitors.

The small RNA (sRNA) strategy [13] uses a DNA fragment that can express oligonucleotides to bind with the transcription initial region (TIR) of the target gene to prevent the expression. Notably, a scaffold structure following the anti-sense sRNA facilitate the combination of the Hfq protein, which increases the affinity of sRNA and target DNA strands. We used the sRNA to downregulate some essential genes to interrupt the cell growth. Small RNA of PyrH (uridine monophosphate kinase, gene for pyridine biosynthesis) [29], FabA (β-hydroxyacyl-acyl carrier protein dehydratase/isomerase, gene for fatty acid biosynthesis) [30], RpoC (gene for RNA polymerase subunit β') [31, 32] and DnaE (gene for DNA polymerase III subunit α) [33]. However, only sRNA for *rpoC* gene showed a little growth inhibition and others showed no significant repression when sRNAs were induced. Similarly, in an asRNA design [34], RNA fragment is also expressed to interact the TIR for gene downregulation . The difference is that two arms located on both sides of anti-sense RNA fragment were assembled. The complimentary two arms can form hair pin structure, which can stabilize the anti-sense binding and prevent the rapid degradation of the asRNA. The asRNA targeting the *rpoC* was tested but no obvious growth inhibition was observed. Gene *gltX* was then tried as a toxin. This gene can cause the mischarge

effect of the amino acid glutamate onto the tRNA1^{Gln} and the regular protein synthesis will be perturbed [35]. Gene *hipA* [17] can convert the cells to dormant state to make cells maintain a low metabolic rate without growth or propagation. Gene *mltB* is also a toxin that can kill cell by causing cell lysis. All the three genes were proved to be working to a variable extent.

To summarize the use of the biosensor assisted selection system, all the sensing promoters and toxins provide abundant tools in the biosensor system. By changing the sensing promoter, we can work on different target compounds for production improvement. By varying the growth inhibitor, we can achieve different levels of inhibition. For example, HipA and MltB are both toxins to cell but the influence on cell is different due to the different inhibition mechanism. HipA toxin enables the cells to a persister, which means the cells are not killed but kept alive with a low metabolic rate. On the contrary, MltB kills cells by lysing the membrane structure and dead cells will release the intracellular nutrients and metabolites to feed other living cells [36]. These toxins can largely broaden our tools for metabolic engineering by adjusting the cell growth conditions.

2.5 Discussion

In this chapter, a L-tyrosine producing system was established. The L-tyrosine biosynthesis pathway involves only endogenous enzymes, which are easier to manipulate for expression in *E. coli*. For this, L-Tyrosine pathways genes were cloned and overexpressed on plasmid. Specifically, *aroB*, *aroD*, *aroE*, *aroL*, *aroA*, *aroC*, *tyrA* and *aroG* were all cloned into the plasmids and introduced into the host strain. Also, *tyrA* and *aroG* genes were modified to resist feedback control,

providing large metabolic flux to L-tyrosine. Besides these conventional metabolic engineering strategies, a biosensor-assisted approch was used to select high L-tyrosine producer to enhance the overall biosynthetic ability of L-tyrosine. With the biosensor assisted selection system, L-tyrosine concentration was improved from 136 mg/L to 577 mg/L.

However, the biosensor could be saturated with high concentration of L-tyrosine in the cell culture. Also, the metabolite mass transfer from cytosol to the intracellular environment can affect the biosensor performance. For example, when the concentration in the cell culture is high, low performing cells could assimilate the environmental L-tyrosine so that the intracellular concentration is high enough to activate the biosensor switch . As a result, low producing cells can also maintain regular growth and propagation. To further improve the L-tyrosine production, an aromatic amino acid exporter was adopted to reduce the intracellular concentration of tyrosine. The exporter can maintain the intracellular concentration of L-tyrosine to a lower level, stimulating the cells to produce more L-tyrosine. As such, only those cells with truly high biosynthetic ability can survive in the culture. It was shown in this chapter that 775 mg/L of L-tyrosine could be produced by the exporter producing strain, which was 5.9 folds improvement compared to the starting L-tyrosine producing strain. These results demonstrated the great potential of this method for metabolic engineering.

On the other hand, the biosensor-assisted cell selection system was not limited to producing native metabolites. Previous literature indicated various systems that can be employed for application the biosensor-assisted cell selection system. We have constructed a series of biosensor system in *E. coli*, and the results were shown in Table 2.2. Among those, biosensor of phenol was discussed in a later

chapter. Moreover, the off-switch biosensor can also be adapted to use with the growth inhibitor.

Expression of toxin gene *hipA* can convert the cells to dormant persisters, which can be used as a good tool for growth repressor instead of killing cells. On the contrary, gene *mltB* expression can cause cell lysis. Experimental results suggested a cell density drop after *mltB* was induced. These growth inhibitors provide a versatile toolbox to exploit for different level of cell growth regulation, which will facilitate the use of biosensor-assisted cell selection system in the future

Chapter 3 Phenol biosynthesis

3.1 Introduction

3.1.1 Background

Phenol is an important commodity chemical with well-recognized industrial values and enormous global market. Current production is reported to be 8.9 million tons a year all over the world. Phenol was also named carbolic acid because it was first discovered by German chemist Runge F. from coal tar. Phenol is also well known for its use as a sanitizer. Phenol is also a good precursor for many industrial products. Phenol and its derivatives are important for manufacturing resins, polycarbonates, epoxies, bakelite, nylon, detergents, herbicides such as phenoxy herbicides, and numerous pharmaceutical drugs [37].

Currently, phenol production relies heavily on utilization of petrochemicals, which often raises economical, environmental, and sustainability concerns. To produce phenol from renewable feedstocks, microbes can be engineered to convert L-tyrosine to phenol by decarboxylation carried out by enzyme tyrosine phenol decarboxylase (TPL).

3.1.2 Industrial production of phenol

Phenol was first extracted from coal tar followed by the chemical synthesis methods. In the middle 1960s, cumene process was adopted for producing phenol and acetone. After decades of development, 90% of phenol is synthesized from this method [38]. Other industrial phenol production methods of involve

oxidation of benzene and toluene, hydrolysis of chlorobenzene and hydrolysis of benzenesulfonate.

1) Hydrolysis of benzenesulfonate.

As an early commercial process developed by Bayer and Monsanto in 1900s, benzenesulfonate is first synthesized with benzene and sulfate. The benzenesulfonate is then reacted with a strong base. The conversion is represented below [39]. However, the benzenesulfonate pathway required high consumption of sulfate and sodium hydroxide. This method is gradually eliminated due to the high environmental hazard.
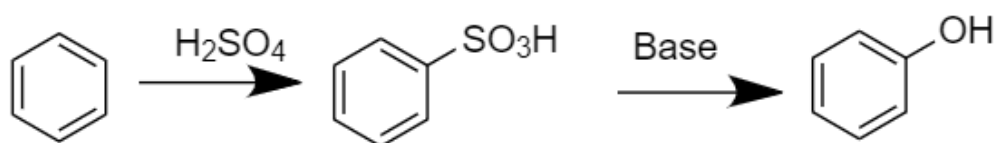
Figure 3.1 Hydrolysis of benzenesulfonate

2) Cumene process

In this method, cumene is synthesized from benzene and propene in the presence of $AlCl_3$. The resulting product is then treated with acid to form phenol and acetone.
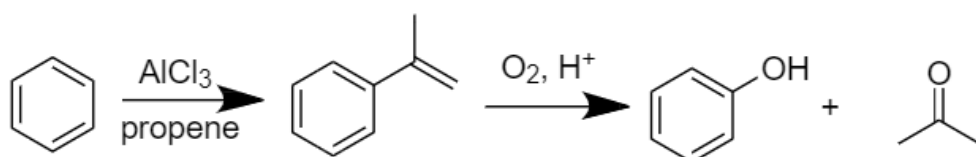
Figure 3.2 Cumene process

3) Hydrolysis of chlorobenzene

This method is similar to hydrolysis of benzenesulfonate. Chlorobenzene is hydrolyzed to phenol using either base or steam [40].
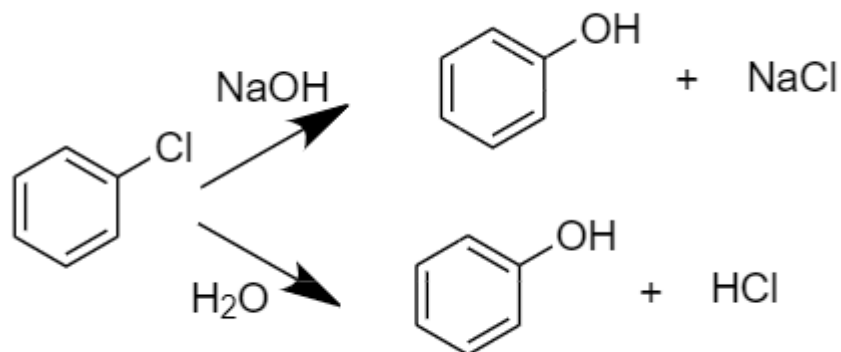
Figure 3.3 hydrolysis of chlorobenzene

4) Oxidation of benzene and toluene

The direct oxidation of benzene to phenol is theoretically possible but has not been commercialized [41, 42]. Using the toluene instead of benzene is considered a more sustainable method for producing phenol.
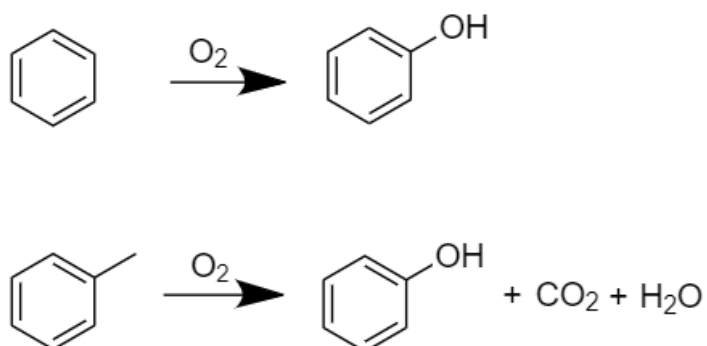
Figure 3.4 oxidation of benzene and toluene

5) Coal pyrolysis

Phenol can also be recovered from the byproduct of coal pyrolysis. This process relies on the coal production industry and isn't used anymore.

3.1.3 Biosynthesis pathways of phenol

Chemical synthesis of phenol soften involves chemicals that are not considered 'green' or sustainable. Producing phenol and other aromatic compounds from biomass has received increasing interest. Using microbes such as *E. coli* to produce phenol is considered as a potential alternative.
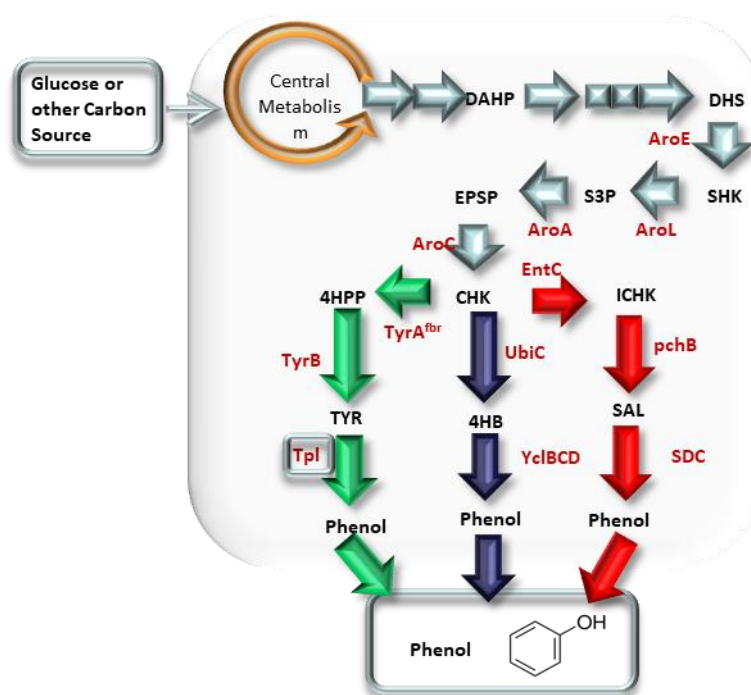
Figure 3.5 Phenol biosynthesis in *E. coli* through 3 pathways. DAHP (3-deoxy-D-arabino heptulosonate-7-phosphate); DHS (3-dehydroshikimate); SHK (shikimate); S3P (SHK-3-phosphate); EPSP (5-enolpyruvyl-shikimate 3-phosphate); CHK (chorismate); 4HPP (4-hydroxyphenolicpyruvate); TYR (L-tyrosine); 4HB (4-hydroxybenzoic acid); ICHK (iso-chorismate); SAL (salicylate).

There are 3 pathways for biosynthesis of phenol in *E. coli*, as shown in Figure 3.5 [43]. Carbon sources is first assimilated from the medium and converted to chorismate by shikimate pathway. Then the chorismate can be converted to L-tyrosine and then phenol by tyrosine phenol lyase (TPL). Alternatively, chorismite can be converted to phenol via 4-hydroxybenzoic acid or salicylate, respectively. In this thesis, the biosynthesis pathway through L-tyrosine were adopted for phenol bioproduction.

3.1.4 Previous work for phenol biosynthesis in *E. coli*

All the three metabolic pathways for phenol biosynthesis in *E. coli* have been studied in previous literature. Kim *et al.* construct a biosynthesis pathway from glucose to phenol via intermediate L-tyrosine in a sRNA regulated L-tyrosine producer [5]. With the assistant of an in-situ product removal method for phenol extraction, 3.79 g/L of phenol was produced with a yield of 0.02 g/g glucose in biphasic fed-batch bioreactor. Noda *et al.* established a chorismite-producing platform and used it to produce 1.1 g/L phenol through tyrosine phenol lyase (TPL) [44]. Miao *et al.* developed the pathway from 4-hydroxybenzoic acid and produced 9.51 g/L phenol from glucose and yeast extract with a yield of 0.06 g/g glucose using high density cultivation and a bi-phase phenol extraction strategy [45]. Moreover, Ren *et al.* produced phenol from through salicylate and achieved 472 mg/L phenol production using glucose, glycerol and yeast extract as the carbon substrates [46]. Thompson *et al.* compared all three pathways of phenol producing and concluded that the salicylate-dependent pathway had a higher production yield (35.7 mg/g) than the other two pathways under the analog cultivation conditions [43].

3.2 Experimental design

An L-tyrosine producer was constructed as described in Chapter 2. Genes *aroB, aroD, aroE, aroL, aroA, aroC* were over-expressed in *E. coli*Also, *aroG* and *ayrA* genes were modified to generate feedback control resistance to yield *aroG^{fbr}* and *ayrA^{fbr}*. To further elevate the provision of L-tyrosine, the biosensor-assisted selection system was introduced. For tyrosine-phenol conversion, the *tpl* gene can be conveniently added to such a system for phenol production from glucose. Specifically, a plasmid containing the *tpl* gene was constructed for functional expression of the tyrosine phenol lyase from *P. multocida* in *E. coli*.

Also, phenol bioproduction by emerging co-culture engineering approach was also tested. Engineered co-culture has been shown to be a robust platform for overcoming the challenges in recent metabolic engineering research. For comparison, conventional monoculture approached and adopted co-culture engineering approach was both adopted, as designed in Figure 3.6.
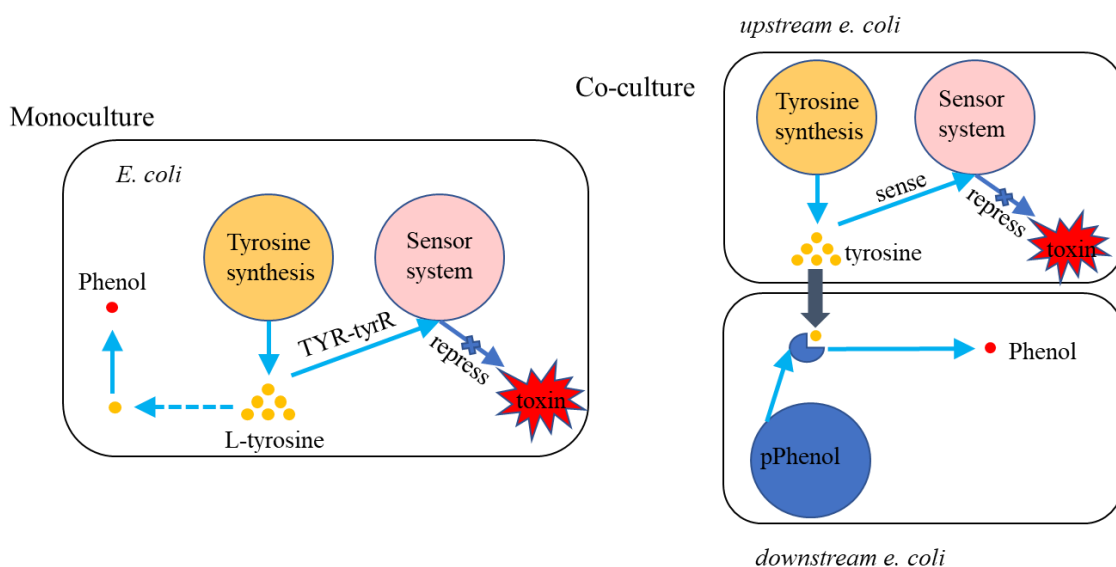
Figure 3.6 Monoculture and co-culture designs for phenol production.

3.3 Material and methods

3.3.1 Plasmids and strains

All *E. coli* strains as well as plasmids used in this study are presented in Table 3.1. Primers used in this study were listed in Appendix.

Table 3.1 Plasmids and strains used in Chapter 3

Plasmids	Description
pB1	pACYCDuet-1 carrying the *E. coli aroB* gene under the control of the proD promoter (PproD)
pBD	pACYCDuet-1 carrying the *E. coli aroB* and *aroD* genes under the control of the proD promoter (PproD)
pBS2	pET28a carrying the proD promoter (PproD) and the *aroE*, *aroL*, *aroA*, *aroC, tyrA^{fbr}* and *aroG^{fbr}* genes
pBS5	pET28a carrying the codon-optimized *Tpl* gene
pBS6	pUC57 carrying the codon-optimized *Tpl* gene with a constitutive *Zymomonas mobilis* pyruvate decarboxylase promoter (Ppdc)
pBS7	pET28a carrying the proC promoter and the *Tpl* gene
pBS8	pET21c carrying the *aroP* promoter

pBS9	pBS8 carrying the *E. coli hipA* gene
pBS10	pBS2 carrying a Cm^R replacing Kan^R
pBS14	pET21c carrying *dmpR* regulon and *tetA* gene under the control of *Pdmp* promoter
pBS15	pET21c carrying *tetA* gene under the control of *Pmtr* promoter
pRA	pET21c carrying *pobR* operon and *tetA* gene under the control of *Ppob* promoter
pSP2	pUC57$_{pdc}$ (kan) carrying the *E. coli aroG^{fbr}, aroE, aroL, aroA, aroC* and *ubiC* genes under the control of the constitutive Zymomonas mobilis pyruvate decarboxylase promoter (Ppdc)
pBR322	Amp^R and Tet^R
pYCL	pET28a carrying the *E. coli* W *yclBCD* genes that is under the control of the proD promoter (PproD)

Strains	Description
BH2	*E. coli* BL21(DE3) $\Delta xylA$ $\Delta tyrA$ $\Delta pheA$
TM2	*E. coli* BL21(DE3) carrying pET21c

BST   *E. coli* BL21(DE3) carrying pBS9

TPS1   *E. coli* BL21(DE3) carrying pBS2 and pBS9

TPR1   *E. coli* BL21(DE3) carrying pBS2 and pET21c

TPS2   *E. coli* BL21(DE3) carrying pBS2, pBD and pBS9

TPR2   *E. coli* BL21(DE3) carrying pBS2, pBD and pET21c

YPD1   *E. coli* BL21(DE3) carrying pET21c, pBS5

YPD2   *E. coli* BL21(DE3) carrying pET21c, pBS6

YPD3   *E. coli* BL21(DE3) carrying pET21c, pBS7

YPD4   YPD3 carrying pACYCDuet-1

MPS1   BL21(DE3) carrying pBS5, pBS9 and pBS10

MRR1   BL21(DE3) carrying pBS5, pET21c and pBS10

DPS1   DH5α carrying pBS14

LSDU   BL21(DE3) carrying pBS2, and pBS15

LSRU   BL21(DE3) carrying pBS2, and pBR322

LSDD   BL21(DE3) carrying pBS7, and pBS14

LSRD   BL21(DE3) carrying pBS7, and pBR322

YU3R   BH2 carrying pSP2, pRA and pBD

| YU33 | BH2 carrying pSP2, pBR322 and pBD |
| DY3-D | BL21 carrying pYCL, pBS14, and pACYCDuet |

The construction of tyrosine related plasmids is described in Chapter 2. For the tyrosine-phenol pathway, the *Tpl* gene from *P. multocida* was codon-optimized and synthesized by Bio Basic Inc, USA. Plasmids pBS5 and pBS6 were constructed by inserting the *Tpl* gene to NdeI/XhoI digested pET28a and pdc-VS, respectively. The *Tpl* gene was then PCR amplified by primers ZLPR1TL and ZLPR2TL and ligated to pET28a-proC vector by SpeI and XhoI to generate pBS7. Plasmid pBS10 was constructed by replacing the Kan^R gene of a previously constructed plasmid pBS2 with the Cm^R gene. Specifically, the Cm^R fragment was amplified by primers ZLPR1KC and ZLPR2KC using pACYCDuet-1 as template and inserted into pBS2 using XhoI and EcoNI.

4.3.2 Cultivation conditions

All *E. coli* strains were cultivated in 3 mL MY1 medium in 37 $^{\circ}$C at 250 rpm. 1 L MY1 medium was comprised of 5g glucose, 0.5 g yeast extract, 2.0 g NH$_4$Cl, 5.0 g (NH$_4$)$_2$SO$_4$, 3.0 g KH$_2$PO$_4$, 7.3 g K$_2$HPO$_4$, 8.4 g MOPS, 0.5 g NaCl, 0.24 g MgSO$_4$, 40 mg L-tyrosine, 40 mg phenylalanine, 40 mg tryptophan, 10 mg 4-hydroxybenzate and trace elements. The working concentrations of trace elements were 0.4 mg/L Na$_2$EDTA, 0.03 mg/L H$_3$BO$_3$, 1 mg/L thiamine, 0.94 mg/L ZnCl$_2$, 0.5 mg/L CoCl$_2$, 0.38 mg/L CuCl$_2$, 1.6 mg/L MnCl$_2$, 3.77 mg/L CaCl$_2$, and 3.6 mg/L FeCl$_2$. The antibiotics were used in the following concentration: 50 mg/L kanamycin, 34 mg/L chloramphenicol and 100 mg/L ampicillin.

For phenol monoculture strains' cultivation, glycerol stock of the desired *E. coli* strains was inoculated in LB medium with necessary antibiotics in 37 $^{\circ}$C for. The cells in overnight culture were then harvested through centrifugation and re-suspended in the fresh MY1 medium with an initial OD_{600} of 0.6. After 48 h cultivation, the culture samples were taken for HPLC analysis.

For phenol production using *E. coli–E. coli* co-cultures, glycerol stock of the desired *E. coli* strains was inoculated in LB medium with necessary antibiotics in 37 $^{\circ}$C. The upstream and downstream cells in overnight culture were then harvested through centrifugation and re-suspended in the fresh MY1 medium according to inoculum ratio with an initial OD_{600} of 0.6, followed by 48 h cultivation at 37 $^{\circ}$C.

To test the growth response to phenol, the Overnight *E. coli* strain DPS1 cultures was centrifuged and re-suspended in 2ml fresh M9 medium with 0.2 initial inoculation OD_{600} at different concentrations of phenol, the culture was then subjected to optical density analysis at 600 nm after 14 h incubation at 37 $^{\circ}$C.

3.3.3 Metabolites quantification

Quantification of the pathway metabolites was conducted using Angilent 1100 HPLC with a DAD detector. 1.0 mL culture sample was centrifuged at 10000 rpm for 5 min, and the supernatant was filtered through 0.45 µm polytetrafluoroethylene membrane syringe filters (VWR International). 10 µL of filtered sample was injected into a column from ES Industries Inc. (HyperSelect ODS Plus C18 column 4.6 × 150 mm, 5 µm) for both L-tyrosine and phenol quantification. The following gradient was utilized for elution: 0 min, 100%

solvent A; 5 min, 95 % solvent A; 6 min, 75% solvent A; 10 min, 10% solvent A; 11-16 min 100% solvent A.

3.4 Results and discussion

3.4.1 Phenol monoculture construction

For phenol production through the tyrosine-dependent pathway, only one heterologous enzyme is needed for converting tyrosine to phenol. Therefore, plasmid pBS5 containing tyrosine phenol lyase gene (*tpl*) was constructed and transformed to both L-tyrosine producing strains without and with the biosensor-assisted cell selection system, respectively. The resulting strains, MPS1 and MPR1, were then cultivated in MY1 medium for 48 h. As shown in Figure 3.7, the phenol production was detected in both mono-culture strains. Interestingly, MPS1 showed lower phenol production than MPR1. On the other hand, the overall flux through L-tyrosine in both strains was lower than the original L-tyrosine producer strain. In fact, the previous result in Chapter 2 indicated that more than 4 folds increase of tyrosine flux was achieved using the biosensor-assisted cell selection system (577 mg/L vs. 130 mg/L). However, in phenol production system, the total flux of towards phenol in the strain without the selection system was comparable to the strain with the cell selection system. This can be explained as follows. The introduction of the tyrosine-sensing cell selection system in fact favored the growth of the cells with better tyrosine accumulation capability (high tyrosine intracellular concentration can help the cells to grow better in the presence of the cell selection mechanism). However, this does not necessarily help select for high phenol producing cells. From this

point of view, the phenol production is not improved by introduction of the biosensor-assisted cell selection system using tyrosine as the sensing target.
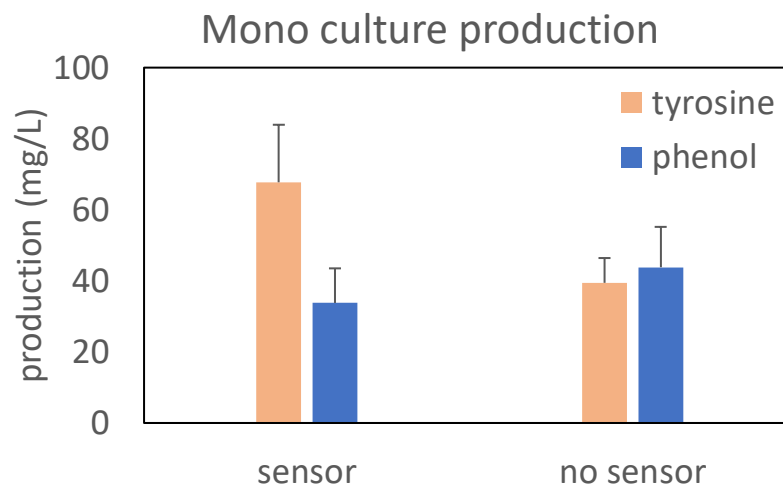
Figure 3.7 phenol production by E. coli monocultures without (MPR1) and with (MPS1) the biosensor-assisted cell selection system.

Also, in *E. coli* strain harboring the biosensor-assisted cell selection system, use of L-tyrosine in the cells is two-folds. First, L-tyrosine is the precursor of phenol biosynthesis and is consumed by the enzymatic reaction to produce phenol. On the other hand, L-tyrosine binds with TyrR sensor protein and the resulting hexamer complex interacts with the corresponding gene promoter and represses the HipA toxin expression to support normal cell growth. Therefore, the introduction of the tyrosine-sensing cell selection system in fact favored the growth the cells with better tyrosine accumulation capability (high tyrosine intracellular concentration can help the cells to grow better in the presence of the cell selection mechanism). However, this does not necessarily help select for high phenol producing cells. From this point of view, the phenol production is not improved by introduction of the biosensor-assisted cell selection system using

tyrosine as the sensing target.There is subtle tradeoff between L-tyrosine formation and consumption under the selection pressure, as shown in Figure 3.8.
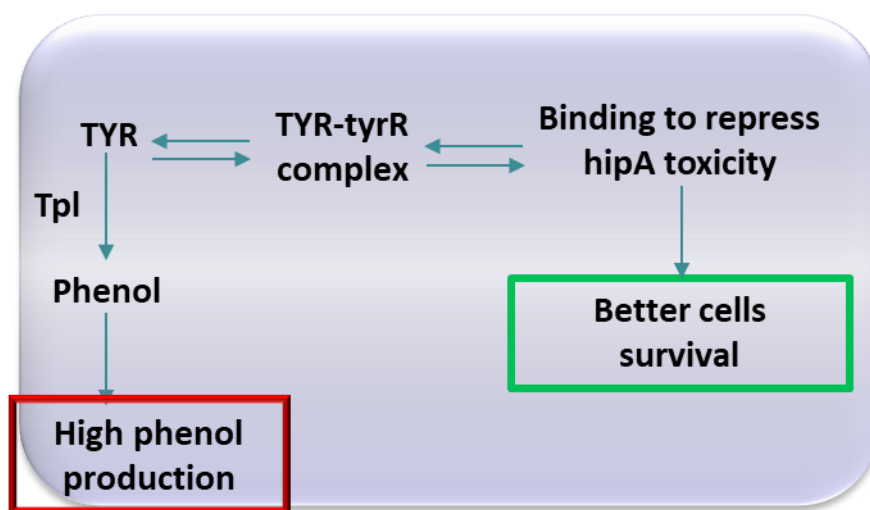
Figure 3.8 schematic illustration of the tyrosine's roles in phenol production and in biosensor-assisted cell selection system

4.4.2 Co-culture engineering for phenol production

Since the use of the biosensor-assisted cell selection strategy was not suitable for phenol production in the context of the monoculture, the biosynthetic system was redesigned. To this end, the modular co-culture engineering strategy was adopted for phenol biosynthesis. Specifically, two *E. coli* strains were recruited to accommodate the whole phenol biosynthesis pathway. The upstream strain is responsible for producing tyrosine from carbon source, and the downstream strain was engineered to convert tyrosine to phenol. Such a design provides the following advantages [47, 48]. 1) The biosensor-assisted cell selection system can be used in the upstream strain for enhancing tyrosine formation; the downstream is solely responsible for phenol production without the use of the biosensor. This avoids the issues of favoring the formation of tyrosine accumulator cells but not the phenol producing cells, as encountered in

the mono-culture ; 2) The metabolic burden associated with over-expression of the entire pathway is divided between two strains, which improves each strain's fitness and biosynthesis performance; 3) Using two E. coli strains in one system offers a straightforward method for pathway balancing by adjusting the subpopulations of the co-culture strains; 4) The downstream strain is dedicated to functional expression of the heterologous *tpl* gene, which improves the bioconversion efficiency of the associated reaction.

To implement the strategy of co-culture engineering, plasmid pBS5 expressing the *tpl* gene was transformed into downstream strain. The resulting strain YPD1 was co-cultivated with the upstream strains TPR1 and TPS1, respectively. Different inoculum ratios were used for adjusting the biosynthetic capabilities of the pathway modules and bioproduction optimization.
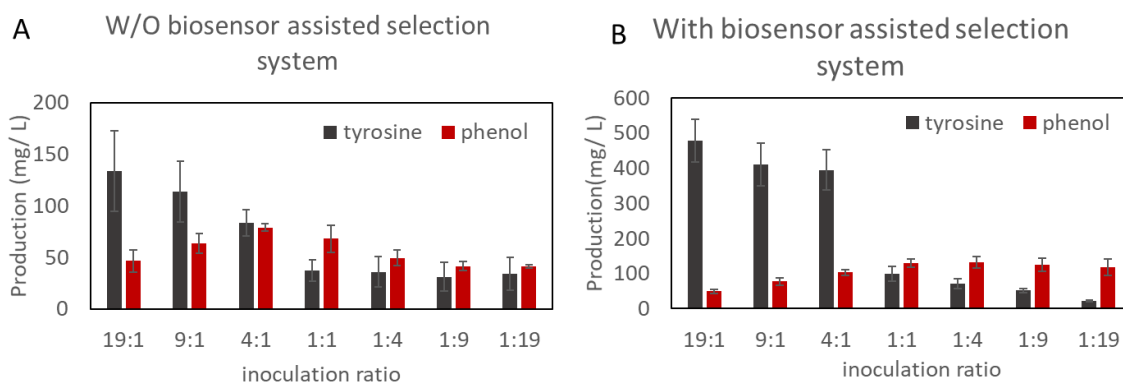
Figure 3.9 Phenol and tyrosine concentrations for TPR1:YPD1 and TPS1:YPD1 co-cultures

As shown in Figure 3.9A, overly high inoculation of the upstream stain led to excessive accumulation of the pathway intermediate tyrosine. High inoculation of the downstream strain generated strong tyrosine consumption power and low tyrosine accumulation; but it also limited the upstream strain's subpopulation size

and cause insufficient supply of tyrosine. Based on these two effects, the phenol concentration increased and then decreased with the inoculation ratio variation between the co-culture strains. The highest phenol production for TPR1:YPD1 co-culture without the biosensor-assisted cell selection system was 75 mg/L, which was achieved at inoculum ratio of 4:1. As mentioned in Chapter 2, there is an reaction equilibrium for L-tyrosine conversion to phenol. The phenol production can be largely improved if the L-tyrosine provision is improved. To this end, a more powerful tyrosine producer with biosensor-assisted selection system was used as the upstream strain. As shown in Figure 3.9B, after the biosensor system was incorporated, the tyrosine accumulation was increased tremendously, and the overall phenol production was improved at most of the inoculation ratios. The highest phenol production was improved to 121 mg/L. In addition, due to the enhanced tyrosine supply capability of the upstream strain, less upstream strain cells are needed for phenol production. As such, the optimal inoculum ratio shifted from 4:1 to 1:4. These results confirm that more powerful upstream tyrosine provider enabled the downstream consumer to make more the final product.
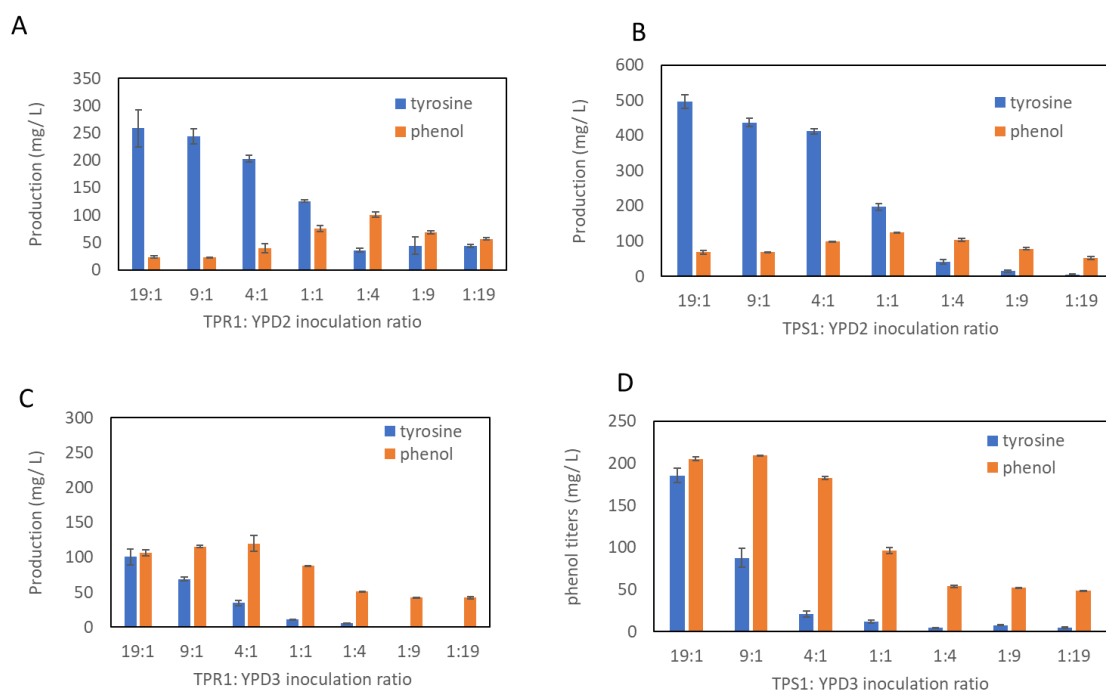
Figure 3.10 Phenol bioproduction by co-culture (A) TPR1:YPD2, (B) TPS1:YPD2, (C) TPR1:YPD3 and (D) TPS1:YDP3. TPR1 and TPS1 are the strains without and with the biosensor-assisted cell selection system, respectively. YPD2 and YPD3 are the strain using the pdc and proC promoter (with different strengths) to express *tpl* gene, respectively.

Co-culture engineering can also be combined with the other traditional metabolic engineering tools, such as optimization of the gene promoters and copy number, and gene knockouts etc., for biosynthesis improvement. In this chapter, high L-tyrosine accumulation of 450 mg/L was observed, indicating the rate limiting step was no more the upstream provision of L-tyrosine but the conversion to phenol. To this end, two constitutive promoters *Ppdc* and *PproC* were used to control the tpl gene expression, respectively, yielding two new downstream strains, YPD2 and YPD3. As shown in Figure 3.10 (A) and (B), for the co-culture using the downstream strain YPD2 with the *Ppdc* promoter, the phenol production results were similar to using T7 promoter (Figure 3.9). Both co-cultures without/with the biosensor produced 100-120 mg/L phenol, suggesting the

downstream conversion rate limited the overall production of phenol regardless of tyrosine provision. As shown in Figure 3.10 (A) an (B), for the co-cultures using the downstream strain YPD3 with the *PproC* promoter [19], the phenol production was largely elevated, and the highest phenol concentration was 210 mg/L at the inoculum ratio of 9:1. Also, the accumulation of tyrosine was reduced tremendously (500mg/L to 180mg/L) compared with the co-cultures using the *Ppdc* promoter. The L-tyrosine accumulation was kept a low level, indicating that its supply was the bottleneck step for the co-cultures using the *PprocC* promoter.

A

B
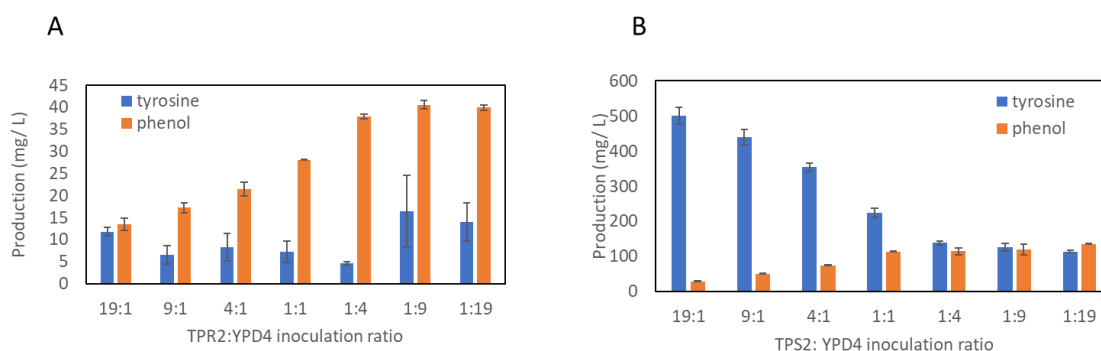
Figure 3.11 Phenol biosynthesis using the TPR2:YPD4 and TPS2:YPD4 co-culture systems. Genes *aroB* and *aroD* of the upstream shkimate pathway were over-expressed in the upstream strains TPR2 and TPS2.

Efforts for improving L-tyrosine flux was made to further improve phenol production by introducing additional copies of the shikimate pathway genes *aroB* and *aroD*. New upstream strains TPR2 andTPS2 strains carrying genes *aroB* and *aroD* was then co-cultivated with YPD4 strains. As shown in Figure 3.11, the over-expression of the *aroB* and *aroD* genes generated no improvement for phenol production. In fact, the final product concentrations were lowered, comparing Figure 3.11A and Figure 3.10C. This was also due to the limited availability of L-tyrosine concentration. After the introduction of the biosensor-

assisted cell selection (TRS2:YPD4), the phenol production was improved to 200 mg/L at the optimal inoculation ratio, which is similar to the results without *aroB* and *aroD* (Figure 3.10D and Figure 3.11B). On the other hand, the tyrosine accumulation was found to be higher than TPS1:YPD4 co-culture at most inoculation ratios, indicating that the *aroB* and *aroD* over-expression improve tyrosine production. However, the higher tyrosine provision did not result in higher phenol production. The reason for such production performance could be that product inhibition on the TPL enzyme due to high phenol concentration prevented efficient tyrosine conversion, even when tyrosine was over-accumulated. Also, the optimal inoculation ratio shifted from 9:1 to 1:19, suggesting a large demand of downstream subpopulation compared to the system without *aroB* and *aroD*. This can be explained by the hampered biosynthetic ability after an empty plasmid pACYCDuet-1 was introduced to the downstream strain to balanced antibiotic resistance from the new introduced plasmid in upstream strain.

Notably, the accumulation of L-tyrosine of TPS2 and YPD4 co-culture was as high as 600 mg/L under 19:1 ratio, which was even comparable to the production of TPS2 cultivated separately. If both tyrosine and phenol were taken into consideration for calculation together, the total flux was around 700 mg/L compared to 553 mg/L for cultivating TPS2 only. The result showed a good indication of how co-culture engineering was able to increase the driving force to adjust and enhance the total flux of the holistic biosynthetic ability.

3.4.3 Employment of both tyrosine and phenol sensors for phenol production

So far, it has been clearly shown that the biosensor selection system played an important role for improving the biosynthetic ability. For the previous studies, a biosensor was only used to sense either the final product in monoculture or the intermediate in co-culture system. However, employment of the biosensors for both pathway product and intermediate has not been studied before. In this section, a downstream strain containing the phenol-targeted cell selection system was constructed for phenol production.

As stated in Chapter 1, *dmp* regulating system for phenol was also tested in this thesis [26]. The mechanism of the sensor is similar. Phenol can combine with DmpR protein and interact with the promoter *Pdmp*. The ON switch sensor *Pdmp* will be stimulated and the expression of TetA is enhanced. As a result, the cells with higher production of phenol can maintain the normal metabolism under the selection pressure tetracycline.
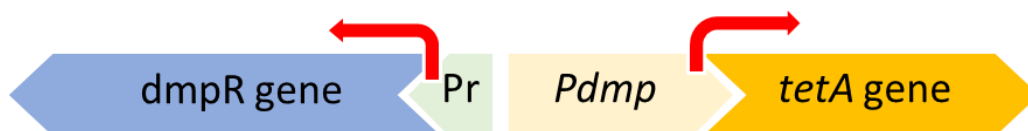
Figure 3.12 Design of phenol-targeted biosensor-assisted cell seletion system. DmpR is the phenol sensor protein. *Pdmp* promoter can be activated by DmpR-phenol complex to upregulate tetA gene expression.

The phenol biosensor was first accessed for the response to various concentration of phenol. Specifically, the E. coli strain containing the phenol-targeted biosensor-assisted cell selection system was grown in the presence of various concentrations of phenol. Tetracycline was also added to the cell culture to impose the selection pressure. Cell culture without the use of exogenous tetracycline was adapted as the control group. As shown in Figure 3.13A, for the

control group, cell density went down as the phenol concentration increased, which was caused by the toxicity of phenol at high concentration. When tetracycline was added to impose the selection pressure, the cell density was reduced at all phenol concentration, as the cell was forced to use some metabolic resource to resist the tetracycline toxicity. Also, the cell density of the culture with tetracycline went up at low phenol concentration and then went down when phenol concentration was high. To better indicate the effectiveness of the biosensor system, the normalized OD_{600} (ratios of OD_{600} with tetracycline/without tetracycline) was calculated to evaluate the growth of cell harboring the biosensor-assisted cell selection system. As shown in Figure 3.13B, the normalized cell density exhibited an increasing trend, indicating the biosensor system was indeed able to upregulate cell growth with the increasing phenol concentration.
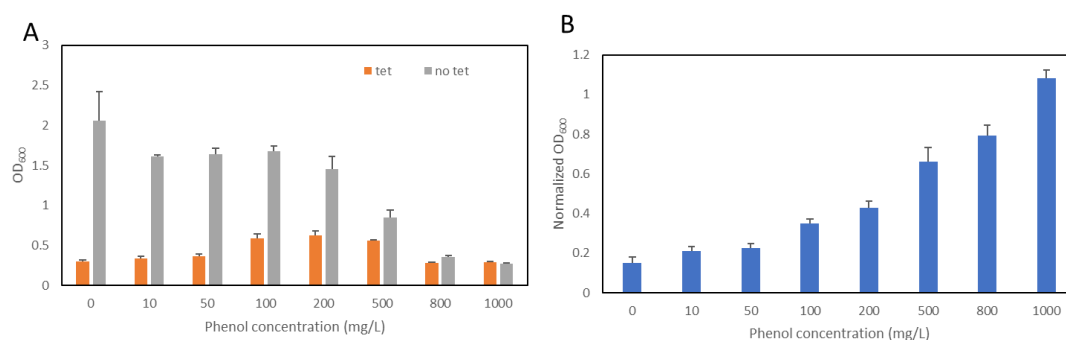
Figure 3.13 Growth of cell cultures in response to various phenol concentrations. A) Cell density of the constructed strain with the biosensor-assisted cell selection system grown without and with the exogenous tetracycline. B) Normalized cell growth of the cell culture without and with tetracycline.

Notably, the constructed phenol-sensing system cannot be directly used in the tyrosine-phenol producing system, because the tyrosine producer used an off-

switch biosensor (TyrR sensor protein downregulating the *ParoP* promoter), but the phenol producer used an on-switch sensor. The addition of tetracycline to the co-culture containing both tyrosine producer and phenol producer can thus be lethal to the tyrosine-producer strain. To address this issue, two alternative approaches for using both biosensors were tested.

First, an on-switch biosensor for tyrosine was adopted. For this, *Pmtr* promoter derived from the *E. coli mtr* gene was used to control the expression of the *tetA* gene. *Pmtr* promoter is upregulated by the TyrR sensor protein and it serves as an on-switch sensor for L-tyrosine. Two upstream strains, LSDU and LSRU, with/without the tyrosine biosensor were constructed, respectively. Also, two downstream strains, LSDD and LSRD, with/without biosensor were constructed, respectively. Four combinations of the four strains were utilized for phenol production on glucose.
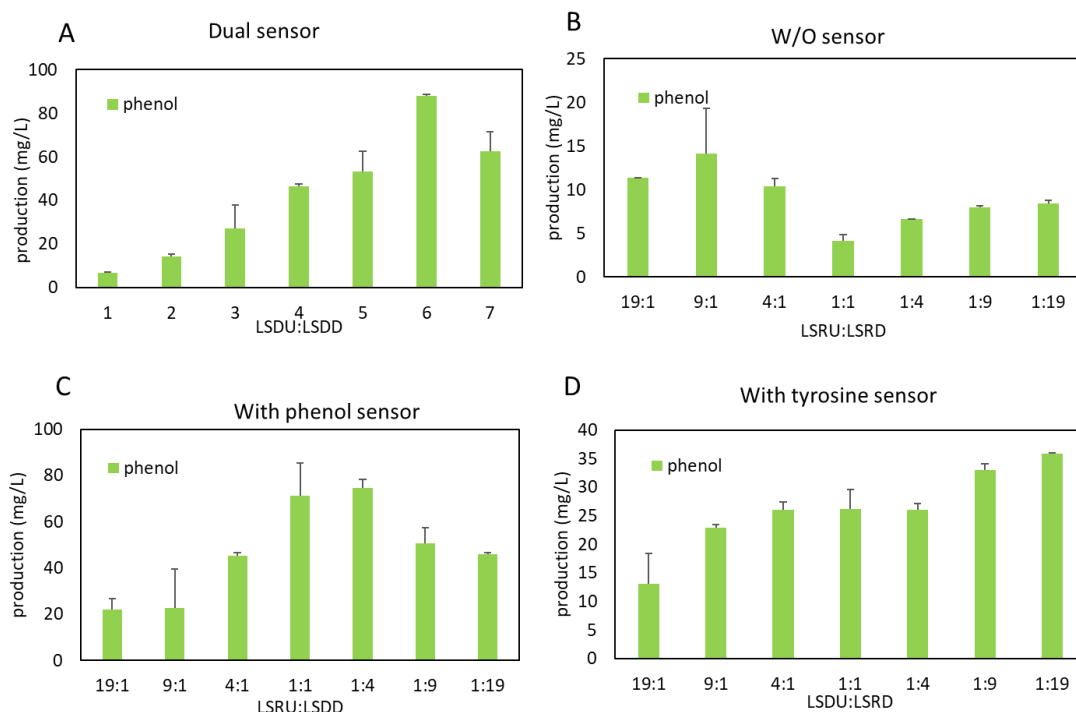
Figure 3.14 Phenol production using co-cultures containing A) both tyrosine and phenol sensors, B) no sensors, C) only downstream phenol sensor, and D) only upstream tyrosine sensor.

As shown in Figure 3.14, the phenol production performance suggested a clear rank: dual biosensors >only phenol biosensor in the downstream strain >only tyrosine biosensor in the upstream strain >no biosensor. The results hereby are in agreement with the expectations. However, the overall phenol production was as low as 90 mg/L (compared to 210 mg/L for previous design). The exact reason for this is unknown.

The other method for dual sensor system was from another phenol producing pathway. The precursor chorismate was first converted to 4-hydroxybenzoic acid and then catalyzed by genes *YclBCD* to yield phenol. Guo et al. constructed the 4HB to phenol system before using a 4HB biosensor system for phenol production and producing around 167 mg/L phenol for no sensor system and 283 mg/L phenol for sensor system [49]. After introducing the downstream sensor, the production was shown in Figure 3.15.
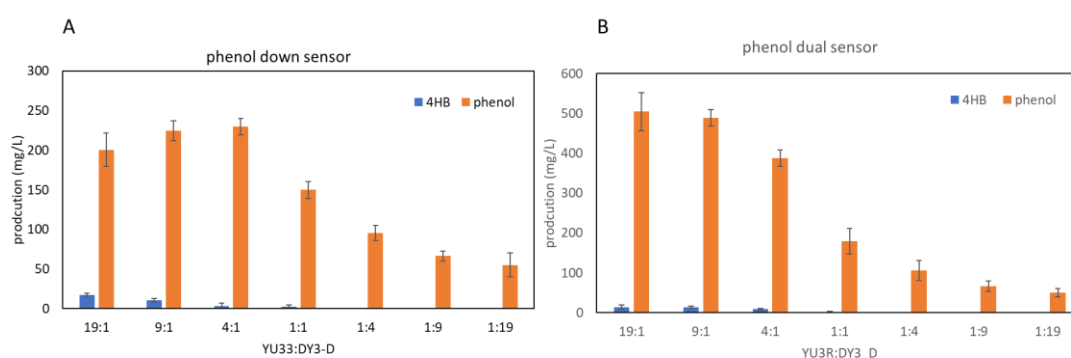
Figure 3.14 Dual biosensor production results for 4HB-phenol system. A) *E. coli* **co-culture with phenol sensor YU33: DY3-D. B)** *E. coli* **co-culture with dual sensors YU3R: DY3-D.**

For sensor system with only phenol biosensor, all intermediate 4HB was consumed even at the ratio of 19: 1, which means the 4HB provision is the rate limiting step. Compared with no sensor system, introducing the phenol sensor helped the production to improve from 167 mg/L to 230 mg/L, indicating a similar production level compared to the system with 4HB biosensor only. After inserting the 4HB biosensor system to form the dual sensor system, the phenol production was improved to 505 mg/L, which was almost doubled compared to either upstream 4HB sensor system or downstream phenol system. The dual sensor system can improve both upstream intermediate availability and downstream conversion ability, and the total flux to phenol was also largely improved, achieving the highest production yield for all phenol producing systems so far.

3.5 Summary

In this chapter, the L-tyrosine producer constructed was used for phenol production. The key pathway enzymes were over-expressed in E. coli, which enabled the production of phenol from simple carbon substrate glucose. When the biosensor-assisted cell selection was introduced, the phenol production was not improved in the context of the monoculture. In fact, the L-tyrosine biosensor-assisted cell selection system can largely enhance the production of tyrosine. However, when the L-tyrosine was used to produce phenol, the consumption of tyrosine was in conflict with the desired function of the L-tyrosine biosensor-assisted cell selection system. Namely, the biosensor-assisted cell selection favored the growth of tyrosine accumulating cells but not phenol high-producing cells.

Next, modular co-culture engineering strategy was used for the phenol production system Modularization of the whole metabolic pathway between two strains largely reduces the metabolic burden on each strain, compared with than monoculture approach. Also, a host strain was dedicated to heterologous over-expression of the *tpl* gene. More importantly, co-culture engineering provided a straightforward method for balancing the biosynthetic pathway by changing the initial strain-to-strain inoculation ratio.

For phenol biosynthesis via tyrosine, the activity of the enzyme TPL is considered the bottleneck step of bioconversion process. As such, it is important to increase the downstream conversion ability. Using modular co-culture engineering, the L-tyrosine biosynthetic ability and the downstream conversion ability can be independently engineered and enhanced. After optimization, best production of phenol was achieved for 210 mg/L at 9:1 inoculation ratio, which was around 2 folds higher than the control mono-culture strain. Further enhancement of the tyrosine pathway by additional expression of the *aroB* and *aroD* genes didn't lead to more production of phenol.

On top of the efforts using modular co-culture engineering to produce phenol, this chapter also investigated the integration of biosensor-assisted cell selection system with co-culture engineering. This was accomplished by establishing the tyrosine biosensor-assisted cell selection mechanism for the upstream co-culture strain. Such design avoided the issue encountered by using the biosensor in the monoculture and successfully improved the phenol production. Moreover, a phenol biosensor-assisted cell selection system was added to the downstream strain to enable a dual biosensor system for phenol production. The results

demonstrated that the phenol biosynthesis was further improved by this approach, indicating the effectiveness of the integration of biosensor-assisted cell selection and modular co-culture engineering for advancing microbial biosynthesis.

Chapter 4. Phenol biosynthesis scale-up and application in production of alkylated phenols

4.1 Introduction

4.1.1 Background

Phenol bioproduction systems have been characterized in Chapter 4. However, larger scale phenol production using engineered *E. coli* has not yet been explored. In fact, it has been reported that phenol is toxic to cells at the concentration of gram per liter level. It was therefore important to control the phenol concentration in aqueous phase to alleviate the toxicity to the cell. In-situ extraction of phenol is a good strategy to address this issue. On the other hand, the phenol removal is also beneficial to push the enzymatic tyrosine-phenol conversion equilibrium toward phenol formation. Some extractants have been utilized in previous reports. Miao *et al.* tested a series of different solvents and concluded that tributyrin was the best candidate for in situ removal with high efficiency and biocompatibility. Some polymeric resins also provide good phenol affinity. All these materials were applied in this chapter for facilitating phenol removal and production.

For phenol bioproduction scale-up, a fed-batch bioreactor was adopted for establishing high cell density, high substrate consumption and high phenol production process. The selected phenol extractant was also used in the bioreactor cultivation to achieve in situ phenol removal. Moreover, the phenol bioproduction system was integrated with a catalytic system to produce alkylated phenols, which are known to have high industrial values. The addition of alkylated phenol in

rubbers, soap and fibers can stabilize the products due to the anti-oxidation property. The compounds can also be used in pesticides and paint coating because it is a good UV adsorbent. A process from simple carbon substrate like glucose, xylose and glycerol to alkylated phenols was designed and established in this chapter, which cannot be achieved by either chemical synthesis or biosynthesis alone. Specifically, alkylated aromatics production often involves catalytic reactions where homogeneous mineral acids such as H2SO4, H3PO4 or metal salt catalysts like FeCl3, AlCl3 are heavily used. Therefore, extensive efforts are needed for separation and purification of the final products , which often involves generation of high amounts of hazardous waste [50]. Although industrial manufacturing of phenol derivatives was well developed, studies aiming at a more sustainable way is of great research and application significance.

4.1.2 Industrial production of alkylated phenol

Industrial production of alkylated phenols is based on reaction between phenol and olefins by Friedel–Crafts alkylation [51].
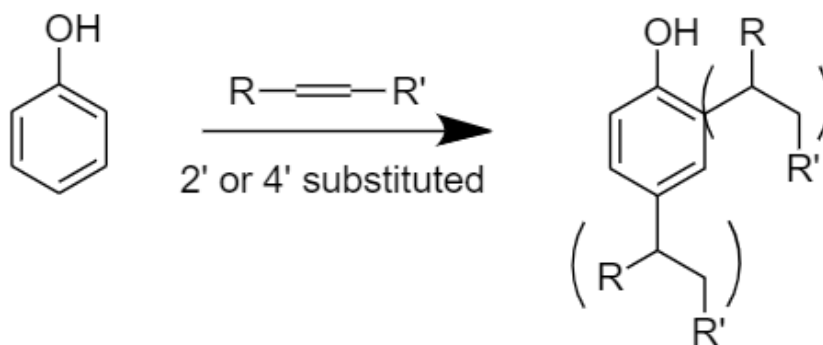
Figure 4.1 industrial production of alkylated phenol

The alkylated phenols produced in this thesis are *tert*-butyl substituents of phenol. The resulting products was a mixture of 2-*tert*-butylphenol (2tBP), 4-*tert*-butylphenol (4tBP), 2,4-di-*tert*-butylphenol (2,4tBP), 2,6-di-*tert*-butylphenol (2,6tBP) and 2,4,6-tri-*tert*-butylphenol (2,4,6tBP).

4.1.3 Experimental design

The whole synthesis process from simple carbon substrate glycerol to the final products is divided into three main modules [52]. In the first module, phenol was produced through cultivation of a metabolically engineered *E. coli* strain. [5, 45, 49, 53]. For phenol production, cell cultivation was first performed in shake flask and then scaled up to fed-batch bioreactor level.

In the second module, three polymeric resins Amberlite IRA 400 (Cl), Amberlite CG-50 and Amberlyst 15 and an organic solvent tributyrin are selected and compared for phenol extraction. The extraction efficiency and biocompatibility will be tested for all materials with chemical phenol and the better performed one will be selected for the in-situ extraction system for batch bioreactor production of phenol.

For the last module, catalytic reaction for phenol alkylation will be employed. The resulting phenol attached to resin will be mixed with *tert*-butanol and subsequently utilized for catalytic reaction [54, 55]. Polymeric resins were also used as the catalyst, as reported to be effective for alkylation reactions in literature [56]. Therefore, the resins were evaluated for alkylation reaction efficiency and the best one will be chosen for the holistic production of alkylphenol.

4.2 Material and methods

4.2.1 Fermentation

E. coli strain PGP was constructed for phenol biosynthesis from glycerol. This strain was derived from a previously constructed strain P2H [57] engineered for tyrosine overproduction [12]. Plasmid pBS7 carrying a constitutive *proC* promoter and the codon-optimized *tpl* gene [49]and plasmid pBR322 were transformed into P2H to generate strain PGP.

LB medium was used for seed culture preparation. M9Y1 medium was used for shake flask cultivation of the engineered strain PGP. One-liter M9Y1 medium was comprised of 20 g glycerol, 0.5 g yeast extract, 1 g of NH_4Cl, 3 g of KH_2PO_4, 6.8 g of Na_2HPO_4, 0.5 g of NaCl, 0.24 g of $MgSO_4$, 1 ml trace elements and 50 mg kanamycin. The working concentrations of trace elements were: 0.4 mg/L Na_2EDTA, 0.03 mg/L H_3BO_3, 1 mg/L thiamine, 0.94 mg/L $ZnCl_2$, 0.5 mg/L $CoCl_2$, 0.38 mg/L $CuCl_2$, 1.6 mg/L $MnCl_2$, 3.77 mg/L $CaCl_2$, and 3.6 mg/L $FeCl_2$.

M9Y2 medium was used for fed-batch bioreactor cultivation. M9Y2 medium had the same composition with M9Y1 medium except that it contained an initial concentration of 5 g/L glycerol and 2 g/L yeast extract.

For bio-phenol production using shake flask, the overnight LB culture of *E. coli* strain PGP was harvested by centrifugation at 10,000 rpm for 1 min and re-suspended in MYG1 medium with an initial OD_{600} of 0.25. The inoculated culture was grown at 37 °C and 250 rpm.

For bio-phenol production using fed-batch bioreactor, the overnight LB culture of strain PGP was added to a 2.5 L bioreactor (Eppendorf Bioflo 120) containing 1 L MYG2 medium. The initial OD_{600} of the culture after inoculation

was 0.125. Cultivation was carried out at 37 °C with an air flow of 2 L/min. Agitation speed was cascaded to maintain DO levels between 5 % and 10 %. The pH was maintained at 7.0 by automatic addition of 5 M sodium hydroxide. Antifoam B (Silicone Emulsion) was added periodically to suppress the foam. 550 ml of 100 g/L glycerol solution was supplemented to the bioreactor at a rate of 0.15 mL/min from 24 h to 84 h. For in situ extraction, 100 g Amberlite IRA-400 (Cl) resin was added to the culture at 72 h. Samples were taken every 24 h for quantitative analysis. For samples taken after 72 h, 70-80 mg of Amberlite IRA-400 (Cl) resin was taken and immersed in 1 mL TBA for 24 h incubation, and the phenol-TBA solution was then analyzed by HPLC. The phenol bioproduction was calculated based on the following equation. The actual titer was calculated by the equation $C = (C_{aq} \cdot V_{aq} + Q_{resin} \cdot M_{resin})/V_{aq}$, where C_{aq} (mg/L) and Q_{resin} (mg/g) are the phenol concentration in the aqueous phase and resin, respectively, and V_{aq} and M_{resin} are the volume of the remaining aqueous phase and weight of resin at the end of cultivation, respectively.

To recover the extracted bio-phenol, the resin particles were separated from cell culture using filter paper and air-dried in a fume hood before mixed with 300 mL TBA. After 24 h of incubation, the resin was removed by filtration to generate bio-phenol-TBA solution. TBA was then distilled at 120 °C to adjust the phenol-TBA ratio for the subsequent alkylation reaction.

4.2.2 Phenol adsorption/extraction

For phenol adsorption experiment, all adsorption experiments for Amberlite IRA-400 (Cl), Amberlite CG-50 and Amberlyst 15 [58, 59] resins were performed by mixing the adsorbent particles with MYG1 medium containing various

concentrations of phenol. The resins were added to the culture with a concentration of 2.5% w/w. After 24 h of incubation at 37 °C and 250 rpm, resin particles were centrifugated, and the supernatant was diluted for HPLC analysis. Data from the adsorption experiment was fit into the Langmuir isotherm model equation. Data simulation was conducted using Matlab where $1/Ce$ was plotted against $1/Qe$ to calculate the model equation parameters.

For phenol recovery experiment, adsorption experiment for Amberlite IRA 400 (Cl) resins and extraction experiment for tributyrin were performed by mixing the adsorbent particles or solvent tributyrin with MYG1 medium containing various concentrations of phenol. The resins were added to the culture with 10% w/w and tributyrin was added with 25% v/v. After 24 h of incubation at 37°C and 250 rpm, culture with resin particles were transferred to Eppendorf tube and centrifugated, and the supernatant was removed. The remaining resin was air dried and methanol was added for another 24h incubation in 37°C. The resulting mixture was centrifuged again and supernatant (phenol methanol solution) was diluted for HPLC analysis. For tributyrin cultures, samples were centrifuged and both phases were taken for dilution before sent to HPLC analysis.

5.2.3 Quantification of metabolites

Quantification of the phenol biosynthesis pathway metabolites was conducted using Agilent 1100 HPLC equipped with a DAD detector. 1.0 mL culture sample was first centrifuged at 10,000 rpm for 5 min, and the resulting supernatant was filtered through 0.45 µm polytetrafluoroethylene membrane syringe filters (VWR International). 10 µL of filtered sample was then injected into a

column(HyperSelect ODS Plus C18 column 4.6 × 150 mm, 5 μm) and eluted using solvent B (0.05 % acetic acid in HPLC water) and solvent A (99.9 % acetonitrile) run at a flow rate of 0.7 mL/min. The following gradient was utilized for elution: 0 min, 100% solvent B; 7 min, 80% solvent B plus 20% solvent A; 9 min, 100% solvent B at a flow rate of 0.7 ml/L. Total elution time was 15 minutes. Tyrosine and phenol were both measured using absorbance at 280 nm.

5.2.4 Catalytic alkylation reaction

Catalytic reactions were performed in Dr. Tsilomelekis lab by Adam Zuber. *Tert*-butylation of phenol was carried out in batch using a 15 mL pressure vessel from Ace Glass Incorporated, equipped with a thermowell to accommodate a thermocouple for precise control of reaction temperature. In a typical reaction, a mixture of *tert*-butyl alcohol (TBA) and phenol in a 1:10 molar ratio, respectively, was added to the reaction vessel with an appropriate stir bar; no solvent was used for reaction. Approximately 0.313 g of catalyst, corresponding to 3 wt% catalyst loading, was carefully added to the vessel, which was subsequently sealed tightly with the provided bushing and O-ring. A thermocouple was placed in the thermowell with silicon oil; the vessel was submerged into an oil bath on top of a heating plate. The reaction was then carried out with 1050 rpm mixing at temperatures ranging between 80–120°C. Initial samples were taken prior to addition of catalyst and analyzed by gas chromatography (Agilent Technologies 7890B GC System), equipped with a flame ionization detector (FID) and capillary column (HP-5, 30 m × 0.320 mm × 0.25 μm), to confirm initial concentration. Toluene was used as a solvent throughout GC analysis. After reaction, the mixture underwent centrifugation to separate the liquid products from solid catalyst

particles; final samples were subsequently taken to confirm concentration and product distribution by GC.

5.3 Results and discussion

5.3.1 Phenol adsorption by resins

To remove the phenol toxicity against *E. coli*, three resins, including Amberlite IRA400(Cl), Amberlite GC-50 and Amberlyst 15, were tested for the phenol adsorption. To this end, resins were added to the cell culture containing different concentrations of phenol. After incubation for phenol adsorption, the remaining concentration of phenol in cell culture was measured by HPLC. Based on the measured concentrations, the data was used to analyzed using the empirical Langmuir isotherm adsorption shown below.

$$Q_e = Q_m \frac{K_{eq}C_e}{1+K_{eq}C_e} \quad (1)$$

where Q_e is the equilibrium concentration on adsorbent (mg/g), Q_m is the maximum capacity of adsorbent (mg/g), K_{eq} is the equilibrium constant (L/mg) and C_e is the equilibrium concentration (mg/L). The Langmuir isotherm can be written as,

$$\frac{1}{Q_e} = \frac{1}{Q_m K_{eq}} \times \frac{1}{C_e} + \frac{1}{Q_m} \quad (2)$$

Adsorption equilibrium data were fit into the equation above, and the model parameters, K_{eq} and Q_m, were obtained based on linear regression. The values of K_{eq} and Q_m for the three polymeric resins are summarized in Table 5.1. Based on

the Q_m value, it was found that Amberlite IRA 400(Cl) demonstrated the highest adsorption capacity from the Langmuir isothermal equation, with approximately 136.99 mg/g of aqueous phenol adsorbed onto the resin at equilibrium.

Table 4.1. Comparison between phenol adsorption parameters of three resins

resins	Amberlite IRA-400 (Cl)	Amberlyst 15	Amberlite CG-50
K_{eq} (L/mg)	0.00043	0.00070	0.00073
Qm (g/mg)	136.99	9.21	14.49

4.3.2 Phenol recovery from resin by organic solvent

The Amberlite IRA 400 (Cl) resin was demonstrated to be the best adsorbent among the selected resins. However, it is also important to investigate the phenol extraction by organic solvent. To this end, phenol recovery experiments were performed to quantify and compare the amount of phenol that can be extracted from medium by resin and solvent tributyrin which has been used for phenol extraction from microbial cell culture before [5].
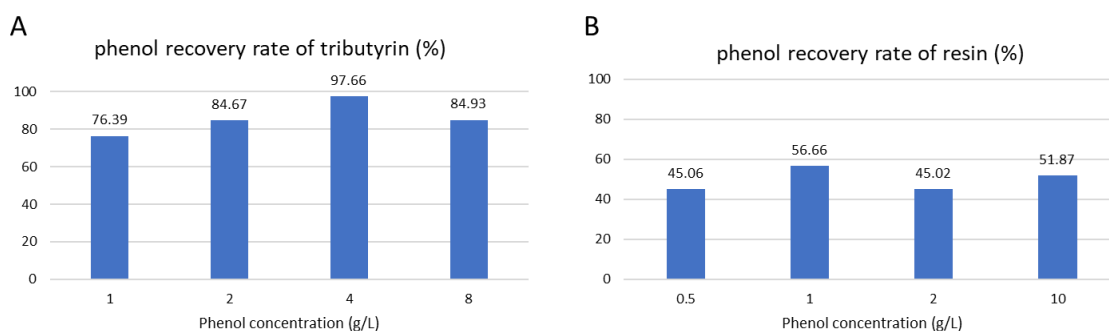
Figure 4.2 Comparison of phenol recovery from cell culture by A) solvent tributyrin and

B) resin Amberlite IRA400(Cl).

As shown in Figure 4.2, advantages and disadvantages for phenol extraction are summarized below.

Tributyrin,

Pros

1) High recovery efficiency (average 90%)

2) High in situ removal ability (high g phenol/g solvent)

3) Easy to reuse

Cons

1) Hard to use for separation (density close to water)

2) Hard to separate with phenol (high boiling point)

3) Low toxicity to cells

Amberlite IRA400(Cl),

Pros

1) Easy for separation (filtration)

2) Easy to separate with phenol (using low boiling point solvent to wash phenol down from resin)

3) No toxicity on cell

Cons

1) Low recovery efficiency

2) Low in situ removal ability (low g phenol/g solvent)

3) Cannot be reused (low capacity for 2nd use)

Although tributyrin was more efficient in phenol removal, its high boiling point prevented the further separation with phenol. Also, the resin was reported to have catalytic activity and the disadvantages of low recovery efficiency can be compensated by using a relatively large amount. Therefore, resin Amberlite IRA400(Cl) was selected for the in situ extraction in the following sections

4.3.3 Phenol production using shake flasks

To produce phenol from glycerol, *E. coli* PGP was cultivated in the MYG1 medium containing 20 g/L glycerol in shake flasks. As shown in Figure 5.3, the cell density of the PGP culture quickly increased to $OD_{600}=2$ at 24 h and remained steady for the rest of the cultivation period.
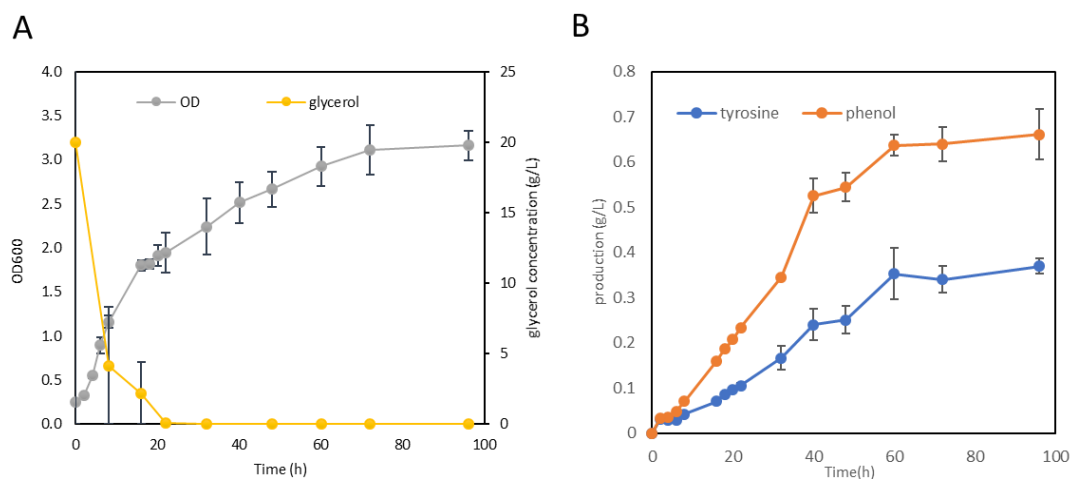
Figure 4.3 Characterization of E. coli PGP growth and phenol production in shake flasks. A) Time profiles of glycerol concentration (g/L) and OD_{600} during cell cultivation. B) Time profiles of tyrosine (g/L) and phenol concentration (g/L) during cell cultivation.

Phenol concentration increased concurrently with cell growth, as most of the production was achieved in the exponential growth phase. The highest phenol production of 661 mg/L and yield of 0.033 g/g glycerol was achieved after 60 h cultivation. In addition, the concentration of the pathway intermediate tyrosine was kept at around 350 mg/L after 60 h of cultivation. The incomplete conversion of tyrosine to phenol suggested that there is equilibrium for the reaction catalyzed by tyrosine phenol lyase (TPL), indicating the potential of further improvement of phenol production by in situ product removal.

4.3.4 Phenol bioproduction in a fed-batch bioreactor

Next, the phenol biosynthesis system was scaled up using a fed-batch bioreactor. With established bioreactor techniques, the cell cultivation process parameters, including pH, dissolved oxygen, were better controlled. As a result, cell density of the engineered strain was improved, and correspondingly, phenol bioproduction capability was elevated. However, tyrosine conversion equilibrium controlled by tyrosine phenol lyase (TPL) still limited further phenol product when phenol concentration was high. Moreover, it has been found that high concentration of phenol (> 2 g/L) is toxic for *E. coli* growth [5]. To address the two issues listed above, 10 % (wt/wt) polymeric resin amberlite IRA-400 (Cl), was added to the cell culture for in situ extraction of phenol.

Using the fed-batch bioreactor assisted by in situ phenol extraction, the engineered *E. coli* strain PGP was cultivated under 37 $^{\circ}$C for 144 h. As shown in Figure 4.4A, the cell density increased rapidly over the first 72 h and stabilized at around OD=4.5 until the end of cultivation. The high cell density hereby indicated that the phenol toxicity against cell was overcome by in situ phenol removal using

the resin. Glycerol initially added in the medium was depleted within 24 h. Additional glycerol was fed to the culture, but the glycerol concentration remained zero throughout cultivation, indicating that all supplemented glycerol from the external feed was consumed quickly.
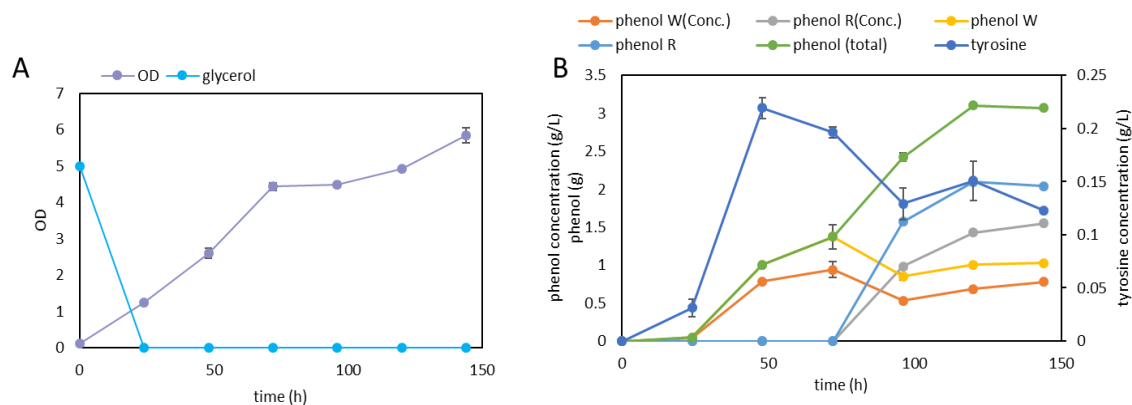
Figure 4.4 Phenol time profile in bioreactor for 144 h fermentation. A) glycerol concentration (g/L) and OD$_{600}$. B) tyrosine (d/L), phenol concentration in water (g/L), phenol in water (g), phenol concentration on resin (g/L), phenol on resin (g) and total phenol (g)

Phenol concentration change with time is shown in Figure 4.4B. To minimize the impact of resin on cell growth, Amberlite IRA-400 (Cl) was added to the bioreactor culture at 72 h. Throughout the whole cultivation process, tyrosine concentration was maintained below 0.22 g/L, indicating a better conversion rate than shake flask cultivation. Phenol concentration in the cell culture reached 0.94 g/L at 72 h. After supplementation of the resin, tyrosine concentration decreased in the culture and phenol was enriched in the resin, as the phenol concentration on resin increased. Due to the phenol adsorption by the resin, phenol concentration in the cell culture was well maintained under 1 g/L

throughout the cultivation, greatly alleviating the phenol toxicity to the cells and the TPL equilibrium issue. The total phenol production (combining the mass of phenol in the culture and adsorbed on the resin) showed a quick increase between 24 and 120 h, with the highest production reaching 3.1 g. The overall yield of bio-phenol production reached 0.052 g/g glycerol at the end of cultivation, which is a 58% higher compared to the production without resin. The findings of the fed-batch bioreactor experiments provide phenol materials for the catalytic phenol alkylation in the following step.

4.3.5 Phenol alkylation results

After phenol was produced from the fed-batch bioreactor, the next step was to verify the catalytic activity of the resin for converting phenol and *tert*-butanol (TBA) to alkylphenol. This experiment was performed in Dr. Tsilomelekis lab by Ph.D. candidate Adam Zuber. Two approaches were used for the phenol collection and alkylation. In the first approach, resin Amberlite IRA400(Cl) was used for both adsorption and reaction. The resin was supplemented with the cell culture in the fed-batch bioreactor for in situ adsorption. The resin particles were then separated from the culture by filtration using filter paper. The wet resin particles were first air-dried and mixed with tert-butanol (TBA) which serves as an alkylating agent. The alkylation reaction was then carried out to convert phenol to alkylated phenols. However, catalytic reaction results suggested the main product for the system was ethel product instead of alkylphenol, and the conversion rate was low (data not shown). Then the second approach was employed.

In the second approach, the in-situ extraction and resin particle separation steps were the same. But when the resin particles were mixed with TBA, TBA was used as an extractant to remove phenol from the resin. The phenol-TBA mixture was separated from the resin particle by filter paper filtration and was taken to 120 °C to remove water and concentrate phenol. The resulting phenol-TBA mixture was added to a container with resin Amberlyst 15 (it was found to have stronger catalytic activity for alkylphenol production) and underwent the alkylation reaction in which TBA was used as the alkylating agent and the resin was used as the catalyst.
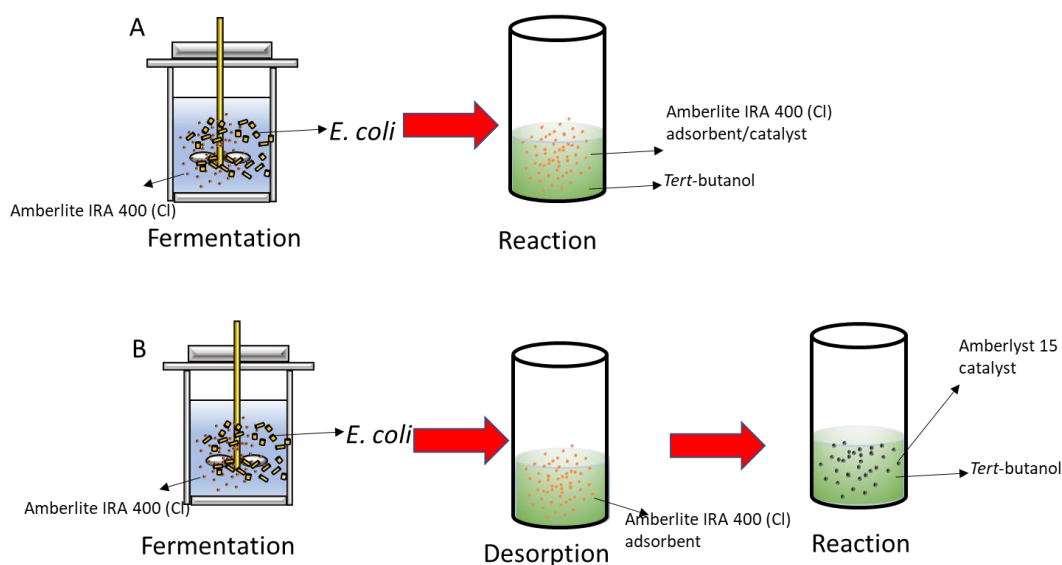
Figure 4.5 Process designed for alkylphenol production from glycerol. A) Resin Amberlite IRA 00 (Cl)used as both a phenol adsorbent and a catalyst for alkylation reaction. B) Amberlite IRA 400(Cl) was used as an adsorbent and Amberlyst 15 was used as a catalyst for alkylation reaction.

The catalytic alkylation of biophenol (phenol produced from cell culture and prepared using the method above) was conducted by Dr. George Tsilomelekis'

group at Department of Chemical and Biochemical Engineering at Rutgers University. The highlight of the results is indicated in Figure 4.6. Overall, it can be seen that biophenol alkylation occurs at both 80 and 120°C. At 80°C, TBA conversion for the biophenol reaction shows an approximate 20% enhancement as compared to that for chemical phenol (phenol standard). The product selectivity of 4-tert-butylphenol (4-TBP) selectivity was higher for biophenol alkylation at both 80 °C and 120 °C. During microbial biosynthesis for the production of phenol, a large number of bio-impurities such as DNA, RNA, organic acids, etc. are present in the cell-culture. This complex background of impurities derived from the bioproduction system can directly impact the catalyst's ability to selectively react, since many of these impurities can potentially compete for adsorption and reaction at the same active sites. Generally, the alkylation reaction results suggest that there are acidic species present in the biophenol that are derived from E. coli cell culture, which augment the reaction rate. The nature and identity of these compounds is currently unknown but would be of research interest for future efforts.
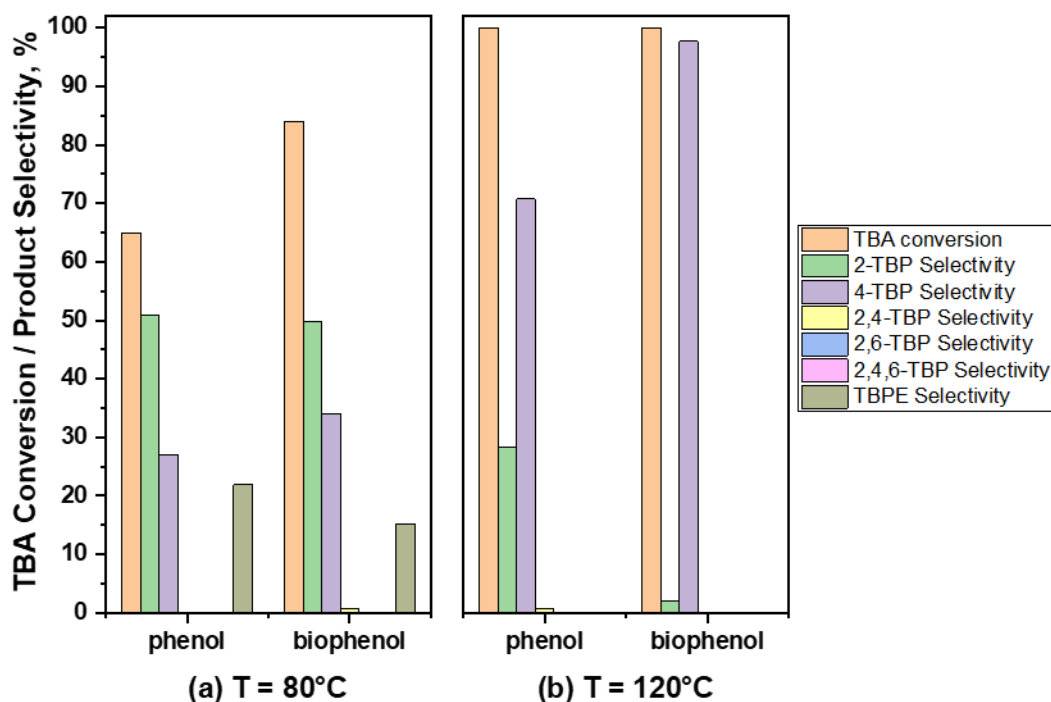
Figure 4.6 Comparison of tert-butylation of chemical phenol and biophenol in batch mode using Amberlyst® 15 catalyst. Reaction conditions: t = 4 h, catalyst loading =3 wt%, Phenol:TBA molar ratio = 10:1, T = (a) 80°C and (b) 120°C. (Courtesy of Adam Zuber and Dr. George Tsilomelekis)

4.4 Summary

In this chapter, phenol bioproduction on glycerol was studied in shake flask and fed-batch bioreactor. In shake flask experiments, 661 mg/L of phenol was produced from 20 g/L of glycerol with 350 mg/L intermediate L-tyrosine accumulated in cell culture. As highlighted before, the tyrosine-phenol equilibrium strongly constrained the conversion of the L-tyrosine. The

accumulation of tyrosine presents a challenge for further phenol bioproduction. Therefore, an in-situ phenol removal method was exploited. On the one hand, this method reduced the toxicity of phenol against the cells , which enabled the cells tostay healthy even at a high concentration of phenol. On the other hand, removal of phenol from aqueous phase pushed the tyrosine-phenol equilibrium toward phenol formation and generated a driving force to produce more phenol. As the bioreactor results showed, 3.1 g/L phenol was produced from 60 g/L glycerol using this method, indicating a 0.052 g/g glycerol yield compared to 0.033 g/g glycerol in the shake flask.

For microbial biosynthesis, aqueous products such as phenol are hard to extract for downstream use. The in-situ removal method could not only increase the metabolites production but facilitate the downstream bioseparation effort. As discussed in this chapter, the adsorbed phenol on resin using this method can subsequently react with other reactants (e.g. TBA) using resin as a catalyst. Alternatively, the adsorbed phenol can be extracted from the resin and further purified for more downstream use.

It is interesting to note that the biophenol produced from the cell culture was found to be effective for alkylation reaction and exhibited a superior performance compared to chemical phenol standard. This was presumably due to the impurities carried over from the cultivation process. Since a lot of chemical species were present in the medium, it is challenging to identify exactly which species facilitated the alkylation process. More in-depth experiments are needed to gain insights about the biological impurities' role on alkylation reaction performance.

Chapter 5 Biosynthesis of 4-Hydroxystyrene and caffeic acid

5.1 Introduction

5.1.1 Background

In Chapter 2 and 3, a biosensor was used to construct a high performing cell selection system for improving the overall tyrosine biosynthetic ability. On the other hand, in Chapter 4, this method was also integrated into a co-culture system to enhance the intermediate provision for better phenol production, resulting in a 4.2 folds improvement of phenol production. In this chapter, the biosensor was used as a growth regulator within the co-culture system to automatically coordinate the upstream provision and downstream conversion for biosynthesis optimization.

For metabolic engineering, balancing the pathway remains a bottleneck for optimizing the production. For conventional strategies, gene copy number and promoter can be changed to modulate the intermediate provision and consumption. However, this requires tremendous effort to test different combinations. Also, the exact copy number and promoter strength for gene expression are hard to determine quantitatively, leaving the whole optimization process an inefficient trial-and-error process.

On the other hand, co-culture engineering provides a new perspective for pathway engineering. The supply and consumption of the intermediate can be balanced by varying the subpopulation ratio between the strains carrying the corresponding pathway modules. In this chapter, a new biosensor-regulated system was developed for dynamic self-adjustment of upstream supply and

downstream conversion for optimization of hydroxystyrene and caffeic acid biosynthesis.

Both hydroxystyrene and caffeic acid are tyrosine derivatives and widely used in many industries. 4-hydroxystyrene is an important commodity chemical that has been widely used as medicine precursor and monomer of resin material. For example, the emerging photoresist technology uses 4-hydroxystyrene to make poly-4-hydroxystyrene derivatives as acid-sensitive resins, which can act as an anticorrosive reagent. The poly-4-hydroxystyrene-based polymer photoresists have become the key materials of photolithographic etch lines for producing 0.11 μm chips [60] polymers. Caffeic acid is an organic acid that can be used as a precursor for pharmaceutical industry. It also possesses a wide range of antibacterial and antivirus activities. Also, the compound can be used in cosmetics due to its ability for UV absorption and enhancing the color for hair dye.

5.1.2 Industrial production of 4-hydroxystyrene and caffeic acid

For industrial production of 4-hydroxystyrene, nucleophilic addition reaction of malonic acid and aldehyde has been developed. t-hydroxybenzaldehyde is used to react with malonic acid, the resulting product undergoes decarboxylation reaction to produce 4-hydroxystyrene. The reaction is isothermal in an aprotic high-boiling solvent, and the reaction solution is extracted and washed at specific pH value.

$$HO-\underset{}{\bigcirc}-CHO + CH_2(COOH)_2 \xrightarrow[110\sim160\,°C]{NH(C_2H_5)_2} HO-\underset{}{\bigcirc}-CH=CH_2$$
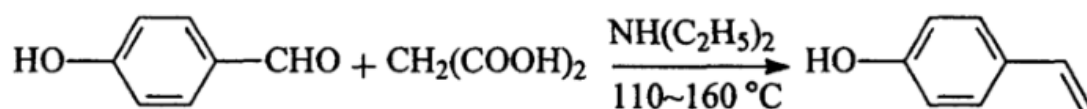
Figure 5.1 Chemical synthesis of 4-hydroxystyrene in industry

Industrial production of caffeic acid involves no chemical synthesis due to the complicated structure of the compound. Caffeic acid is currently produced by extraction from plants. For production of CA, the rhizomes of *Cimicifuga* are extracted with methanol and concentrated under reduced pressure to remove methanol. The residue is then dissolved by heated water . After filtration, benzene is added for extraction, followed by washing with a 1% aqueous $NaHCO_3$. Diluted hydrochloric acid was added to the collected washing solution to acidify the organic acid, which help enrich CA in benzene. Finally, CA molecule is then concentrated under reduced pressure to remove benzene.

5.1.3 Biosynthesis pathway of 4-hydroxystyrene and caffeic acid

Biosynthesis of 4-hydroxystyrene is an extension of the L-tyrosine pathway. Tyrosine ammonia Lyase (TAL) can convert L-tyrosine to *p*-coumaric acid [61, 62]. Then the resulting *p*-coumaric acid undergoes decarboxylation reaction to yield 4-hydroxystyrene via enzyme Ferulic acid decarboxylase 1 (FDC1) [63, 64](Figure 5.2).

Similarly, caffeic acid is also a derivative of *p*-coumaric acid. Enzyme 4-coumarate 3-hydroxylase encoded by gene *Coum3H* [61] from *Saccharothrix espanaensis* can catalyze the conversion of *p*-coumaric acid to caffeic acid by adding a hydroxy group on the 3' carbon on the benzene ring. The L-tyrosine platform constructed in the previous chapter is hereby utilized for both 4-hydroxystyrene and caffeic acid production from renewable carbon sources.
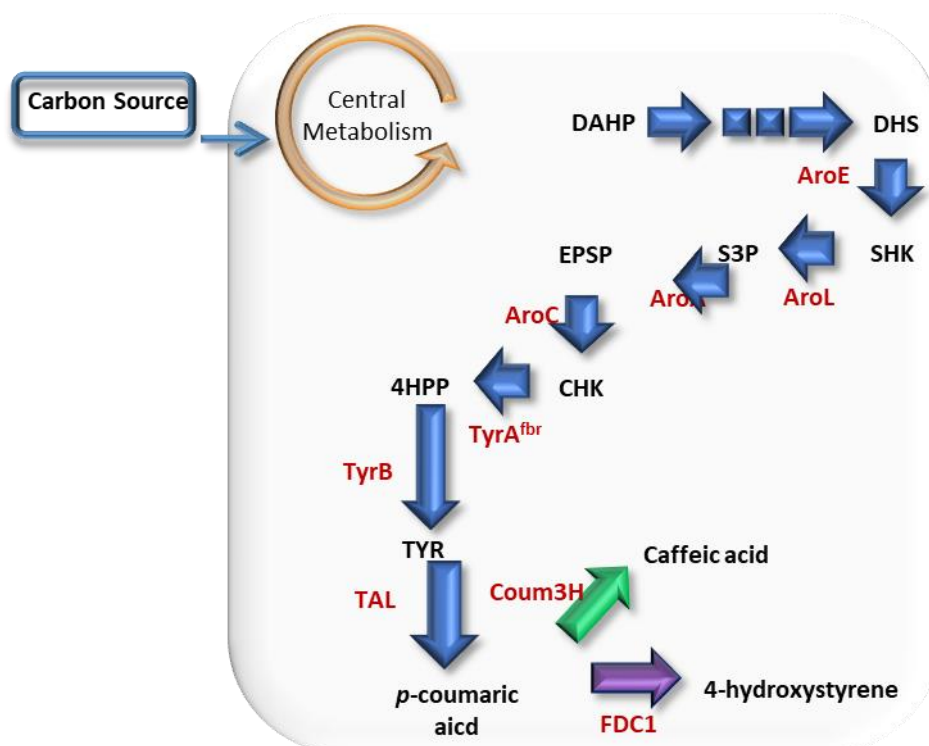
Figure 5.2 Biosynthesis pathway of 4-hydroxystyrene and caffeic acid. DAHP (3-deoxy-D-arabino heptulosonate-7-phosphate); DHS (3-dehydroshikimate); SHK (shikimate); S3P (SHK-3-phosphate); EPSP (5-enolpyruvyl-shikimate 3-phosphate); CHK (chorismate); 4HPP (4-hydroxyphenolicpyruvate); TYR (L-tyrosine).

5.2 Experimental design

In this chapter, the tyrosine producing strain developed in Chapter 4 was used as the baseline strain for 4-hydroxystyrene and caffeic acid production. After the production for monoculture was tested, the co-culture system was constructed for comparison. Moreover, a biosensor-based system that could automatically regulate cell growth for co-culture biosynthesis optimization was also introduced and tested for its effectiveness.

This biosensor-based cell regulation system is different from the previously constructed cell selection system. Specifically, the tyrosine-responsive biosensor

TyrR is used to control the expression the tetracycline resistance gene tetA via the promoter Pmtr. High tyrosine concentration triggers the tetA expression and thus support the cell growth in the presence of exogenous antibiotic tetracycline. The TyrR-Pmtr-tetA system is introduced into the downstream strain of the co-culture system. At the initial growth state of cell culture, there is no sufficient tyrosine accumulation for activating the tetA expression in the downstream strain. As a result, this strain's growth is inhibited, which frees up more metabolic resources to support the growth of the upstream strain and promote the production of L-tyrosine. After L-tyrosine is accumulated to a certain level, the biosensor-based growth regulation system is activated, and the downstream strain's growth is promoted, which in turn help consume the accumulated L-tyrosine. Through such a dynamic self-regulation design, the biosynthesis pathway is balanced for biosynthesis optimization. For characterization of the functionality of biosensor-based cell regulation system, the tyrosine accumulation and final product titers were monitored and the co-culture population composition was analyzed to provide insights for the engineered co-culture's growth and biosynthesis behaviors.

5.3 Material and methods

5.3.1 Plasmids and strains

All strains and plasmids used in this study are summarized in Table 5.1. Primers used for PCR amplification were listed in Appendix.

Table 5.1 Plasmids and strains used in Chapter 5

Plasmids	Description
pET28a	T7 promoter, Kan^R
pET21c	T7 promoter, Amp^R
pRSFDuet-1	double T7 promoters, Kan^R
pCDFDuet-1	double T7 promoters, Sp^R
pBR322	Amp^R, Tet^R
pTrcHis2B-*RgTAL*	pTrcHis2B carrying the codon-optimized *RgTAL* gene
pRSF-*Coum3H*	pRSFDuet-1 carrying the codon-optimized *Coum3H* gene
pRSF-*FDC1*	pRSFDuet-1 carrying the codon-optimized *FDC1* gene
pCDF-*TAL*	pCDFDuet-1 carrying the codon-optimized *RgTAL* gene with a trc promoter
pBS2	pET28a carrying the proD promoter and the *aroE*, *aroL*, *aroA*, *aroC*, *tyrA^{fbr}* and *aroG^{fbr}* genes
pMtr-*tetA*	pET21c carrying the *tetA* gene under the control of *Pmtr* promoter

pMtr-*TAL*	pET21c carrying the *TAL* gene under the control of *Pmtr* promoter
pMtr-*FDC1*	pET21c carrying the *FDC1* gene under the control of *Pmtr* promoter

Strains	Description
BL21(DE3)	F⁻ *omp*T *hsd*S$_B$ (r$_B$⁻, m$_B$⁻) *gal dcm* (DE3)
K12(DE3)	F- lambda- *ilvG- rfb*-50 *rph*-1 (DE3)
P2	*E. coli* K12(DE3) Δ*pheA* Δ*tyrR lacZ*::P$_{LtetO-1}$-*tyrA^{fbr}aroG^{fbr} tyrR*::P$_{LtetO-1}$-*tyrA^{fbr}aroG^{fbr}
P2H	P2 *hisH*(L82R) (DE3)
DPT1	DH5α carrying pMtr-*tetA* plasmid
HSMK	K12(DE3) carrying pMtr-*RgTAL*, pMtr-*FDC1* and pBS10
HSMB	BL21(DE3) carrying pMtr-*RgTAL*, pMtr-*FDC1* and pBS10
HSMS	P2H carrying pTrcHis2B-*RgTAL* and pRSF-*FDC1*
CAMR	P2H carrying pTrcHis2B-*RgTAL* and pRSF-*Coum3H*
PMPR	P2H carrying pBS5 and pBR322

COHDR	BL21(DE3) carrying pTrcHis2B-*RgTAL* and pRSF-*FDC1*
COU	P2H carrying pBR322, pET28a and pCDFDuet-1
COHDS	BL21(DE3) carrying pMtr-*tetA*, pTrcHis2B-*RgTAL* and pRSF-*FDC1*
COCDS	K12(DE3) carrying pMtr-*tetA*, pTrcHis2B-*RgTAL* and pRSF-*Coum3H*
COCDR	K12(DE3) carrying pBR322, pTrcHis2B-*RgTAL* and pRSF-*Coum3H*
COPU	P2H carrying pET28a and pBR322
COPDR	BL21(DE3) carrying BR322 and pBS5
COPDS	BL21(DE3) carrying pMtr-*tetA* and pBS5

5.3.2 Cultivation conditions

All *E. coli* strains were cultivated in 2 mL M9Y medium in 37 °C at 250 rpm. 1 L M9Y medium was comprised of 5 g glucose, 0.5 g yeast extract, 1 g of NH_4Cl, 3 g of KH_2PO_4, 6.8 g of Na_2HPO_4, 0.5 g of NaCl, 0.24 g of $MgSO_4$, 0.2mM IPTG, 1 ml trace elements and antibiotics, accordingly. The working concentration of antibiotics was 50 mg/L kanamycin, 100 mg/L ampicillin, 50 mg/L streptomycin and 20 mg/L tetracycline, accordingly. The working concentrations of trace elements were: 0.4 mg/L Na_2EDTA, 0.03 mg/L H_3BO_3, 1

mg/L thiamine, 0.94 mg/L $ZnCl_2$, 0.5 mg/L $CoCl_2$, 0.38 mg/L $CuCl_2$, 1.6 mg/L $MnCl_2$, 3.77 mg/L $CaCl_2$, and 3.6 mg/L $FeCl_2$.

For monoculture strains' cultivation, glycerol stock of the desired *E. coli* strains was inoculated in LB medium with necessary antibiotics in 37 °C for. The cells in overnight culture were then harvested through centrifugation and re-suspended in the fresh MY1 medium with an initial OD_{600} of 0.6. After 48 h cultivation, the culture samples were taken for HPLC analysis.

For production using *E. coli-E. coli* co-cultures, glycerol stock of the desired *E. coli* strains was inoculated in LB medium with necessary antibiotics in 37 °C. The upstream and downstream cells in overnight culture were then harvested through centrifugation and re-suspended in the fresh MY1 medium according to inoculum ratio with an initial OD_{600} of 0.6, followed by 48 h cultivation at 37 °C. For regulation pressure test, co-culture system was cultivated under different concentration of tetracycline. Production was measure using HPLC.

To test the growth response to tyrosine, the Overnight *E. coli* strain DPS2 cultures was centrifuged and re-suspended in 2 ml fresh M9 medium with 0.2 initial inoculation OD_{600} at different concentrations of L-tyrosine, the culture was then subjected to optical density analysis at 600 nm after 14 h incubation at 37 °C.

5.3.3 Metabolites quantification

Quantification of the pathway metabolites was conducted using Angilent 1100 HPLC with a DAD detector. 1.0 mL culture sample was centrifuged at 10000 rpm for 5 min, and the supernatant was filtered through 0.45 µm polytetrafluoroethylene membrane syringe filters (VWR International). For 4-hydroxystyrene, phenol and tyrosine quantification, 10 µL of filtered sample was

injected into a column from ES Industries Inc. (HyperSelect ODS Plus C18 column 4.6 × 150 mm, 5 μm). The following gradient was utilized for elution (solvent A water with 0.5% acetic acid and solvent B acetonitrile): 0 min, 100% solvent A; 5 min, 95 % solvent A; 6 min, 75% solvent A; 10 min, 10% solvent A; 11-16 min 100% solvent A. For caffeic acid and tyrosine quantification, the following gradient was utilized for elution: 0 min, 100% solvent A; 7 min, 80% solvent A plus 20% solvent B; 9 min, 100% solvent A. Total elution time is 12 minutes. SAA was measured using absorbance at 280 nm. CA and RA were both measured at 320 nm.

5.4 Results and discussion

5.4.1 4-hydroxystyrene biosynthesis

To construct 4-hydroxystyrene producer, *E. coli* P2H strain was used for tyrosine provision. Plasmids pCDF-*TAL* and pRSF-*FDC1* containing the required pathway genes were introduced into P2H for converting L-tyrosine to 4-hydroxystyrene. The resulting strain HSMS was cultivated in M9Y1 medium and the production was shown in Figure 5.3A. The production of 4-hydroxystyrene was around 70 mg/L, and accumulation of both L-tyrosine and *p*-coumaric acid were observed. These finding confirmed that the established pathway in E. coli has been designed biosynthesis capability.

Next, the genes TAL and FDC1 were cloned under the control of *Pmtr* promoter. The resulting strain DPT1 was used for a biosensor response experiment to confirm the function of the biosensor system. Strain DPT1 was

cultivated in M9 medium under different L-tyrosine concentration in the presence of 20 mg/L tetracycline. The results were show in Figure 5.3B. When the tyrosine concentration increased, the cell density increased, indicating a good biosensing response to tyrosine.

The tyrosine pathway and 4-hydroxystyrene producing genes under the control of *Pmtr* promoter were transformed in both *E. coli* K12(DE3) and BL21(DE3) strain for testing the biosensor's effect on 4-hydroxystyrene production in monoculture. The production of 4-hydroxystyrene in both resulting strains, HSMB and HSMK, was as low as 10 mg/L and accumulation of L-tyrosine and *p*-coumaric acid were both very limited. This was presumably caused by the high metabolic burden of incorporating too many genes (9 genes) into the strain and also unbalanced biosynthetic ability between the tyrosine provision and consumption. In comparison, HSMS has an engineered tyrosine pathway by chromosomal modification, which possess stronger capability for tyrosine provision and high 4-hydroxystyrene production. However, since HSMS does not have the *tyrR* gene for making the tyrosine sensor protein, this strain cannot be used for the tyrosine biosensor-based gene expression system (TyrR-Pmtr-TAL-FDC1).
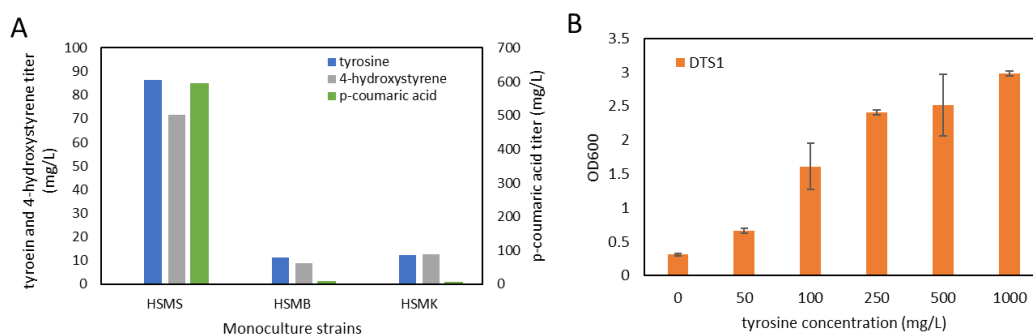
Figure 5.3 Monoculture production of 4-hydorxystyrene. A) Production of 4-hydroxystyrene in monoculture; B) Growth response of E. coli containing the tyrosine biosensor-based regulation mechanism to different concentration of tyrosine.

Then the effectiveness of the tyrosine biosensor-based cell growth system within the context of a co-culture was studied. First, a control co-culture was developed. *E. coli* strain P2H was assigned for the upstream production of tyrosine and a BL21(DE3) strain containing tyrosine to 4-hydroxystyrene pathway was adapted as the downstream strain. The resulting COHU and COHD strains were co-cultivated in M9Y1 medium, and the production results were shown in Figure 5.4A.
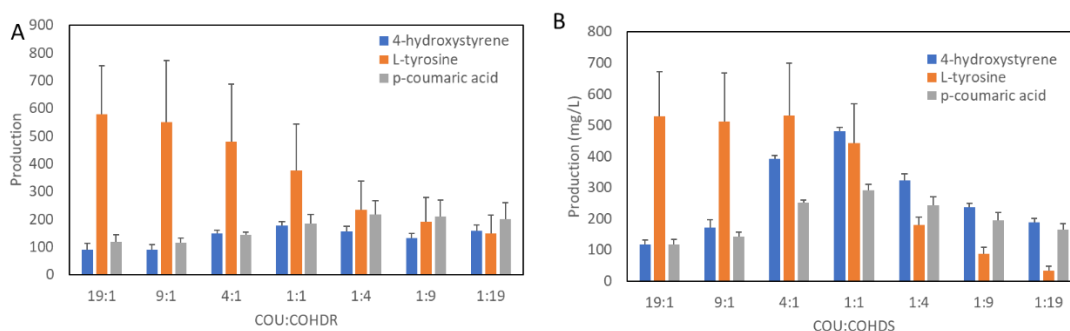
Figure 5.4 Co-culture production of 4-hydroxystyrene. A) Production of L-tyrosine, 4-hydroxystyrene and *p*-coumaric acid using the COHU: COHD co-culture; B) Production of L-tyrosine, 4-hydroxystyrene and *p*-coumaric acid using the COHUs:COHDs co-culture.

At 19:1 inoculation ratio, over 550 mg/L of tyrosine was accumulated but the downstream strain was not strong for tyrosine conversion to 4-hydroxystyrene. As the downstream strain's proportion increased in the culture, tyrosine accumulation decreased and 4-hydroxystyrene production increased. For the ratios beyond 1:4, the production was not improved any more due to limited

provision of tyrosine. The highest production of177 mg/L was achieved at the initial inoculation ratio 1:1. Yet, the production yield (mass of product over mass of substrate) was low, indicating that most of the cell resources was not utilized for but the cell metabolism but not production. In fact, when the tyrosine accumulation was low, it is wasteful to use metabolic resources to grow the downstream strain. For addressing this issue, the tyrosine biosensor-based growth regulation system (TyrR-Pmtr-tetA) was introduced into the downstream strain [16]. When tyrosine concentration is low, the promoter Pmtr is shut down for tetA expression, so the downstream strain was inhibited for growth by tetracycline added to the culture. When tyrosine concentration is high, the *Pmtr* promoter is activated and TetA was expressed for resisting tetracycline toxicity so that the strain could restore its growth and consume tyrosine. As a result, the cell resources can be better allocated between the co-culture strains in dynamic fashion and the production can be improved.

The production results using this new approach were shown in Figure 5.4B. The tyrosine accumulation was comparable to the control in Figure 5.4A. However, the highest 4-hydroxystyrene production was improved to 480 mg/L and the best inoculum ratio remained 1:1. Compared with the control co-culture without the biosensor, the COU:COHDS co-culture produced a similar tyrosine accumulation level but much higher 4-dydroxystyrene production, indicating a large increase of total flux. Further experiments will be performed to characterize the dynamics for tyrosine concentration, 4-hydroxystyrene concentration and TetA expression for offering critical insights of the engineered co-culture Growth and biosynthesis behaviors.

5.4.2 Caffeic acid

Similarly, experiment using the biosensor system for overproduction of caffeic acid was performed. Monoculture production of caffeic acid was first tested by introducing the plasmids pCDF-*TAL* and pRSF-*Coum3H* into *E. coli* P2H strain. The resulting strain CAMR was cultivated in M9Y1 medium. As shown in Figure 5.5, 86 mg/L of caffeic acid was produced from 5 g/L glucose.
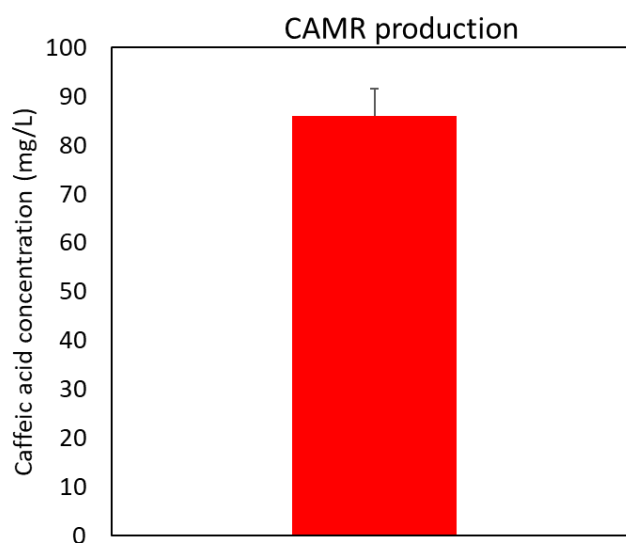
Figure 5.5 Monoculture production of caffeic acid.

Next, the metabolic pathway was divided into two moduels. The *TAL* and *Coum3H* genes were transformed into K12(DE3) strain to develop a downstream strain COCDR. The upstream strain is similar to the strain used for 4-hydroxystyrene co-culture biosynthesis. The COU: COCDR co-culture was cultivatedin M9Y1 medium for 48 h. As shown in Figure 5.6A, for inoculation ratios with high upstream percentage, tyrosine accumulation was high because the downstream conversion ability was limited. As the proportion of downstream strain increased, the downstream biosynthetic ability was improved but the provision of tyrosine became limited. Due to these effects, the best production of

100 mg/L was achieved at 1:1 inoculation ratio. To increase to the production of final product, biosensor plasmid containing the TyrR-Pmtr-tetA system was introduced into the downstream strain, and the production results of the resulting co-culture COU:COCDS were shown in Figure 5.6B. The highest production of caffeic acid was improved to 247 mg/L. The optimal inoculum ratio shifted 1:9, which can be explained as follows. Since the downstream COCDS strain with biosensor was repressed during the cell growth, more COCDS strain was needed at initial stage for pathway balancing.
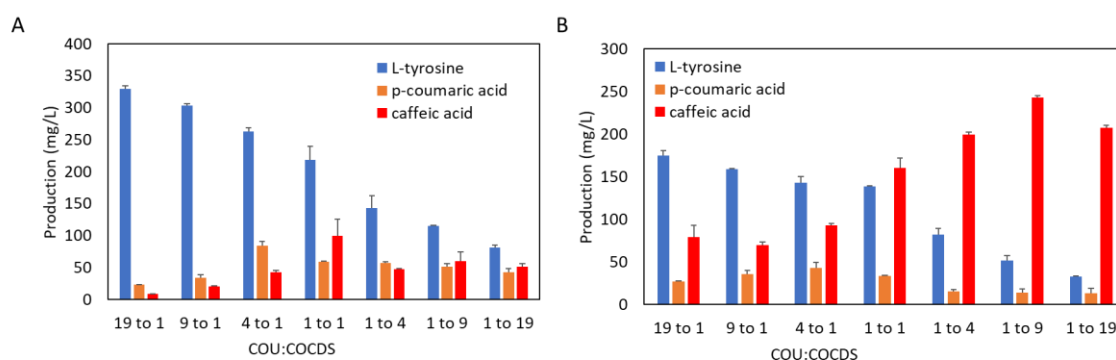
Figure 5.6 Co-culture production of caffeic acid. A) production by COU: COCDR co-culture without biosensor; B production by COU: COCDS co-culture with biosensor.

After the use of the biosensor-based cell growth regulation system, the caffeic acid production was improved by 2.5 folds, indicating great potential of this method. Also, the overall accumulation of tyrosine was much lower than the control group without biosensor. More in-depth studies are needed for better understanding the dynamic behaviors of the engineered co-culture system.

5.4.3 Phenol

Although phenol production was well studied in aforementioned chapters, this biosensor method was still applied in phenol production system. A T7 promoter controlled *tpl* gene was introduced into *E. coli* P2H and the resulting strain PMPR was cultivated in M9Y1 medium for phenol production. As shown in Figure 5.7, the phenol production was around 26 mg/L and 200 mg/L tyrosine remained unconverted. These results are in good agreement with the findings in Chapter 4 that T7 promoter wasn't a good promoter for phenol production in *E. coli*.
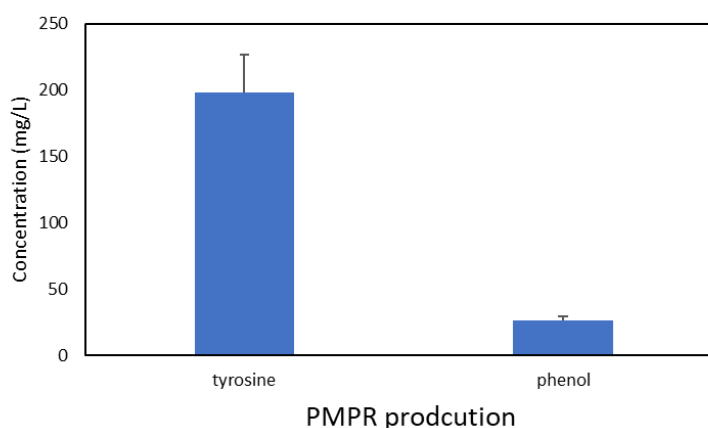
Figure 5.7 phenol production by the monoculture of PMPR strain

For phenol biosynthesis using a co-culture, the plasmid with T7-*tpl* operon was transformed into BL21(DE3) for generating downstream strain COPDR. Co-cultivating COPU and COPDR strains was carried out in M9Y1 medium for 48 hours. Utilization of co-culture engineering is able to reduce the metabolic burden for each strain and improve the biosynthesis performance. Also, the co-culture design provides the flexibility of using BL21(DE3), a commonly used expression host, for functional expression of the heterologous gene *tpl*. The production results of the co-culture were shown in Figure 5.8A. Phenol production of the

COPU:COPDR co-culture system reached 94 mg/L at the inoculation ratio of 1:4, which is significantly higher than the monoculture.
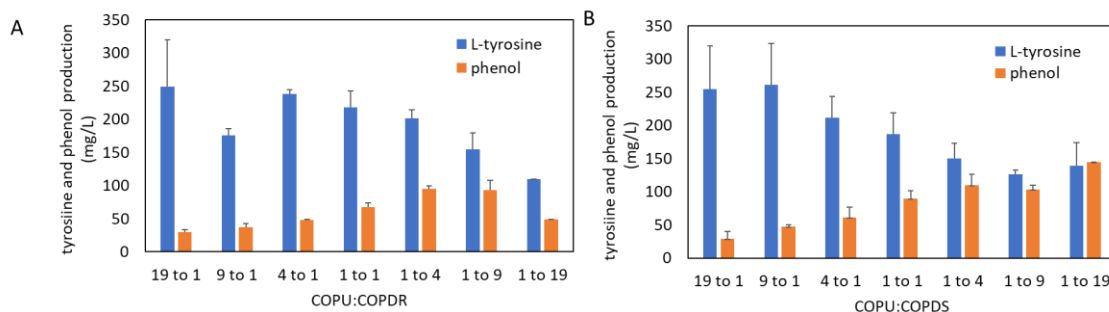
Figure 5.8 Phenol production by co-culture. A) COPU: COPDR without the biosensor, B) COPU: COPDS with the biosensor. 20 mg/L of tetracycline was added to the culture.

Further optimization of the production system involved introduction of biosensor-based cell growth regulation mechanism into downstream strain to yield strain COPDS. The newly constructed strain was co-cultivated with COPU in the presence of 20 mg/L tetracycline. It was found that the production was improved to 144 mg/L and the optimized inoculation ratio was 1:19 (Figure 5.8B). The shift of the optimal ratio of inoculation was similar to the findings of the caffeic acid biosynthesis using the engineered co-culture. Also, tyrosine concentration was similar between the co-cultures without and with the biosensor.

Next, two other concentrations of tetracycline, 5 mg/L and 10 mg/L, were used in phenol producing system. Tetracycline was used as the growth inhibitor in this design and higher concentration of tetracycline requires higher expression level of *tetA* gene, which was regulated by the promoter *Pmtr*. Theoretically, downstream strain needed higher concentration of tyrosine to resist higher concentration of tetracycline.
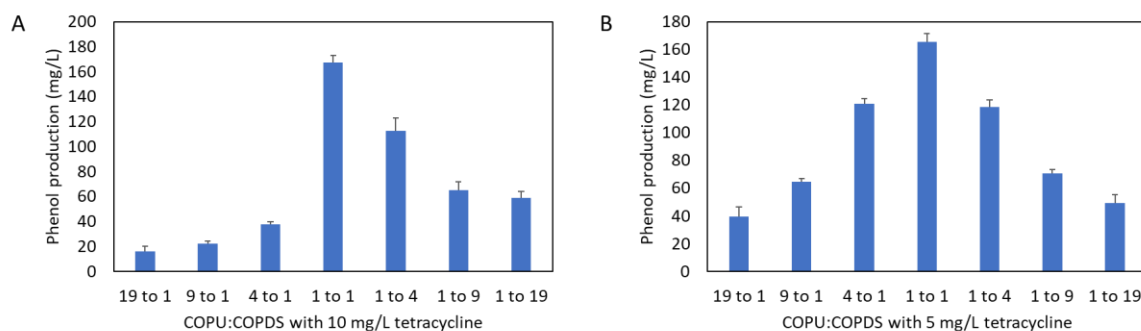
Figure 5.8 Phenol production by co-culture under different concentration of tetracycline.

A) 10 mg/L tetracycline; B) 5 mg/L of tetracycline.

As shown in Figure 5.9, the highest phenol production was 160 mg/L for both 10 mg/L and 5 mg/L tetracycline. Also, the optimal inoculum ratio of COPU:COPDS were both shifted compared to when 20 mg/L tetracycline was used. As explained above, less concentration of tetracycline meant that lower tyrosine accumulation was required for downstream strain growth upregulation. As a result, smaller initial population of downstream strains was needed. However, the production of phenol couldn't be further improved, indicating that the bottleneck of the system was still the downstream conversion efficiency.

5.5 Summary

In previous chapters, a tyrosine biosensor was used to select for the high performing cells among the microbial population. The biosensor assisted selection systems were demonstrated to be very effective in term of improving the production. In this chapter, the biosensor was used for regulating the cell growth, which led to facilitate the pathway balancing according to intermediate accumulation. When the intermediate concentration was low, the growth of

downstream strain could be inhibited to free up more resources for the upstream strain to produce the intermediate. When the intermediate concentration was accumulated to a certain level, the biosensor in downstream strain was activated and the tetracycline exporter gene *tetA* was overexpressed to resist antibiotic tetracycline. As a result, the downstream strain restored the growth and better consume intermediate tyrosine. Also, the intermediate provision from the upstream strain was slowed down due to the competition of metabolic resource with the downstream stream. By using this method, 4-hydroxylstyrene, caffeic acid, and phenol productions were improved by 2.7 folds, 2.5 folds and 1.44 folds, respectively, in comparison to the control groups. The results indicated the strong ability of this method to balance the metabolic pathway and increase the biosynthesis performance.

Chapter 6. Biosynthesis of rosmarinic acid

6.1 Introduction

6.1.1 Background

Compound (R)-α-[3-(3,4-Dihydroxyphenyl)-1-oxo-2E-propenyl] oxy-3,4-dihydroxy-benzenepropanoic acid, also known as rosmarinic acid ($C_{18}H_{16}O_8$, CAS: 20283-92-5), is one of the hydroxycinnamic acid esters (HCEs). HCEs are a class of natural products with strong antioxidant and anti-inflammatory activities that widely distributed in different plants. Rosmarinic acid is chemically an ester of caffeic acid and salvianic acid A and has been found to possess important activities. For example, RA is used for cataract and recovered the transparency of sonicated human cataract [65]. Moreover, RA is an active agent for attenuating neurological disorder diseases (e.g. Parkinson's disease and Alzheimer's disease) [66]. Also, cancel cells' growth can be inhibited by rosemary extracts, in which RA and carnosic acid are the main ingredients. At present, RA has shown its versatile application value in the pharmaceutical, food, cosmetic and other industries.

Romarinic acid was first found in rosemary (*Rosmarinus officinalis*) by Italian chemists M.L. Scarpatti and G.Oriente and then confirmed to be widely accumulated by a variety of different plants (e.g. *Lamiaceae, Marantaceae* and *Anthocerotophyta* etc.) [67, 68]. Even in the same plant, the RA concentration differs a lot in different parts. For example, concentration of RA is only 0.009% accumulation in the leaf while 4.574% occurrence in the root for *Menthaspicata*.

For the same part of plant, RA concentration varies in different season. Plant harvested in summer has higher accumulation of RA than autumn or winter.

6.1.2 Industrial production of rosmarinic acid

Rosmarinic acid has the complicated chemical structure and traditional chemical synthesis usually leads to low production yield. Currently, RA production mainly relies on extraction from plants

The natural accumulation of RA in the plants is low, so a concentrated extract of plant is recommended for use. The dried and powdered plant material was boiling water for 1 hour for extraction. Then, the solution was transferred into another beaker and pH was adjusted by HCl to pH=2. The resulting precipitation was removed, and clear supernatant solution was transferred for extraction. After performing liquid-liquid extraction with the filtered and acidified water and diethylether (40:100), the organic phase is evaporated for generation of purified RA.

6.1.3 Chemical synthesis of rosmarinic acid

Chemical synthesis of rosmarinic acid often involves long synthesis pathway and low overall production yield. Eicher *et al.* designed a total synthesis pathway for rosmarinic acid and its derivatives from piperonyl aldehyde, as shown in Figure 6.1 [69].

Figure 6.1 RA chemical synthesis designed by Theophil et al. The overall yield is 5%.

This synthesis pathway was designed in 1996 and the yield is 5%. To improve the RA synthesis yield, Davi et al used (S)-tyrosine as raw material and after 7 steps of reaction, (S)- (-)-rosmarnic acid with optical activity can be synthesis, as shown in Figure 6.2 [70]. The overall yield using this strategy is 9%.

Figure 6.2 RA synthesis designed by Davi et al. The overall yield is 9%.

Han et al. exploited methyl group as the protective group of hydroxyl groups. They reported a total synthesis method using methylvanillin as the precurosor. A series of steps, including Erlenmeyer reaction, hydrolysis, reduction, addition of protective groups, condensation and removal of protective groups, the final product rosmarinic acid can be produced. This pathway offers a better overall conversion rate using inexpensive materials as the precursor (Figure 6.3) [71].
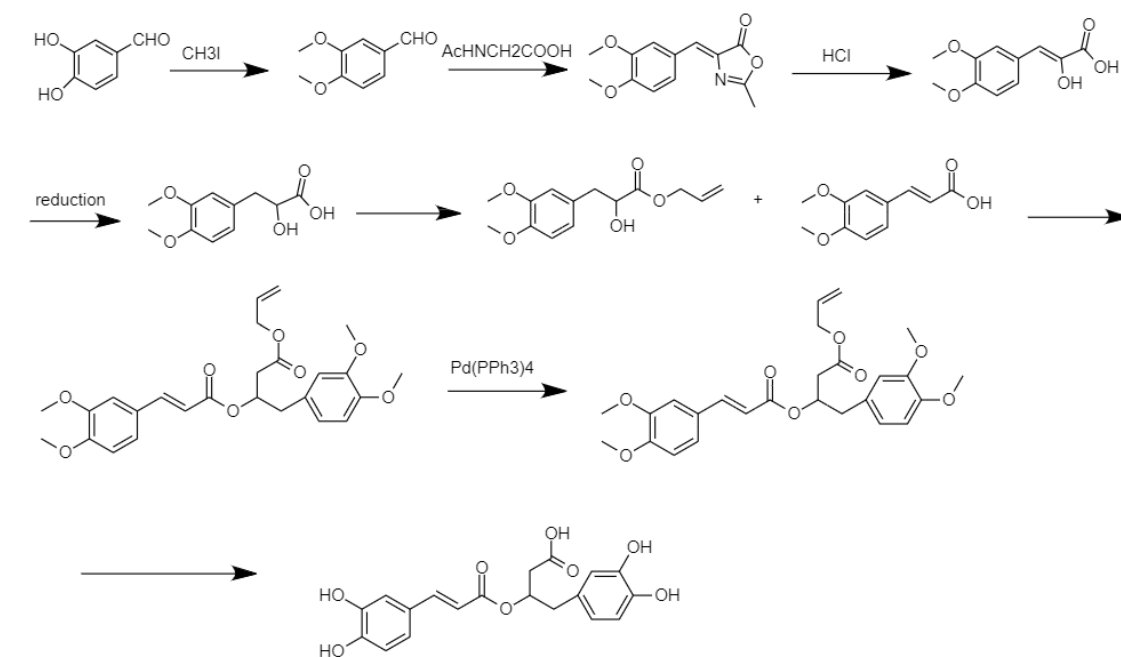
Figure 6.3 RA synthesis designed by Yuan *et al*. The overall yield is 30%.

6.1.4 Biosynthesis of rosmarinic acid

Due to the high complexity and low overall yield of RA chemical synthesis as well as the limited efficiency from extraction methods, metabolic engineering effort for RA biosynthesis has drawn great attention. RA is the ester of two precursors, caffeic acid (CA) and salvianic acid A (SAA). (). The RA biosynthesis pathway is shown in Figure 6.4. For caffeic acid biosynthesis, carbon substrates are first converted to L-tyrosine, and then catalyzed by TAL and HpaBC/Coum3Hto produce CA via p-coumaric acid. On the other hand, 4-hydroxyphenylpyruvate (4HPP), a precursor of tyrosine, is converted to 3,4-dihdroxyphenylpyruvate and SAA by HpaBC and D-ldh, respectively [72-74]. CA is then converted to caffeoyl CoA, which then reactswith SAA to produce the final production, rosmarinic acid.

Notably, the biosynthesis of RA involved a non-linear (divergence-convergence) pathways instead of a linear pathway. The whole pathway was split at the point of 4HPP. One pathway branch goes through intermediate L-tyrosine toCA. The other branch is for producing SAA. For traditional metabolic engineering method, it's hard to balance the two pathway intermediates. Also, more than 6 genes were overexpressed in the system besides the tyrosine producing pathway. As a result, biosynthetic ability may be hampered by high metabolic burden. Moreover, gene *hpaBC* has promiscuous activities, which means it can act on both substrates pCA and 4HPP for making undesired products [75]. The two pathway modules are difficult to balance for best precursor provisions. In this chapter, the whole RA biosynthesis pathway was divided into three modules. The CA and SAA modules were used for providing the precursors, while the RA module was used for making RA from the precursors.
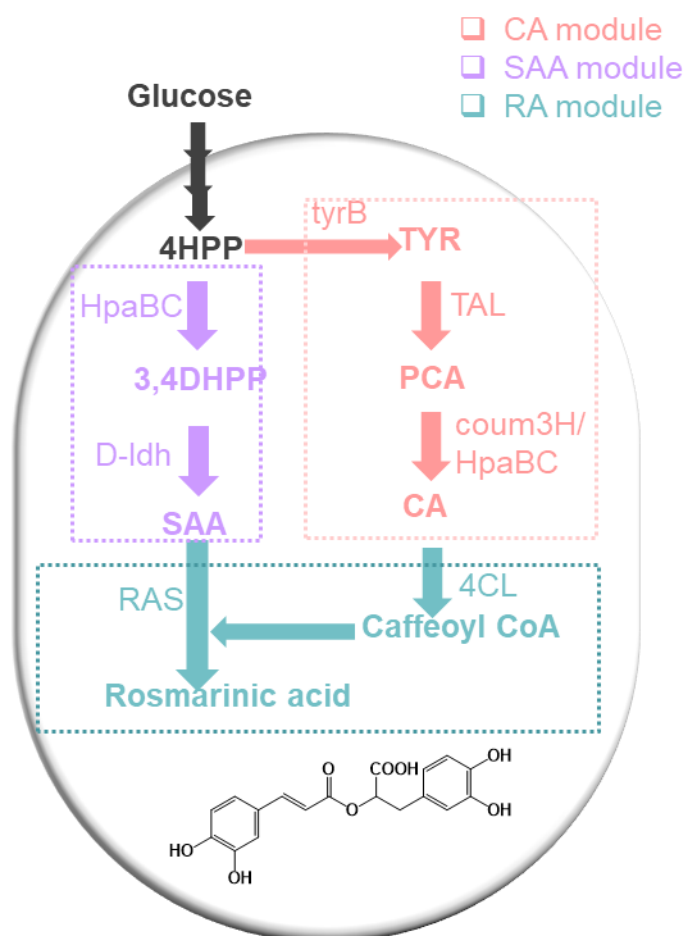
Figure 6.4 RA biosynthesis pathway. 4HPP (4-hydroxyphenylpyruvate); 3,4DHPP (3,4-dihdroxyphenylpyruvate); SAA (salvianic acid A); TYR (L-tyrosine); pCA (*p*-coumaric acid); CA (caffeic acid).

6.2 Experimental design

E. coli strain P2H was used as the host strain for tyrosine production. Rosmarinic acid production experiment was performed under 3 different temperatures, 25 $^{\circ}$C, 30 $^{\circ}$C and 37 $^{\circ}$C. However, using traditional monoculture engineering approach cannot fully unleash the potential of RA production because of the following reasons.

1) Monoculture design incorporated too many genes in a single host strain, causing high metabolic burden during cell growth and repressing the overall performance of enzymatic reactions.

2) Gene *hpaBC* had promiscuous activity for substrates pCA and 4HPP. The biosynthesis of CA might be interfered by conversion of 4HPP.

3) The divergence-convergence form of biosynthesis pathway made it difficult to adjust the CA and SAA module to balance the provision of the precursors by traditional metabolic engineering tools.

Therefore, co-culture engineering was utilized for further improving RA production. Two original designs are shown in Figure 6.5 by changing the allocations of 3 modules into two strains. In Figure 6.5A, CA and RA modules are overexpressed in the same strain, and SAA module is overexpressed in an independent strain. In Figure 6.5B, RA module is combined with SAA module, but CA module is accommodated in a separated host strain. Five two strain co-culture designs will be discussed.
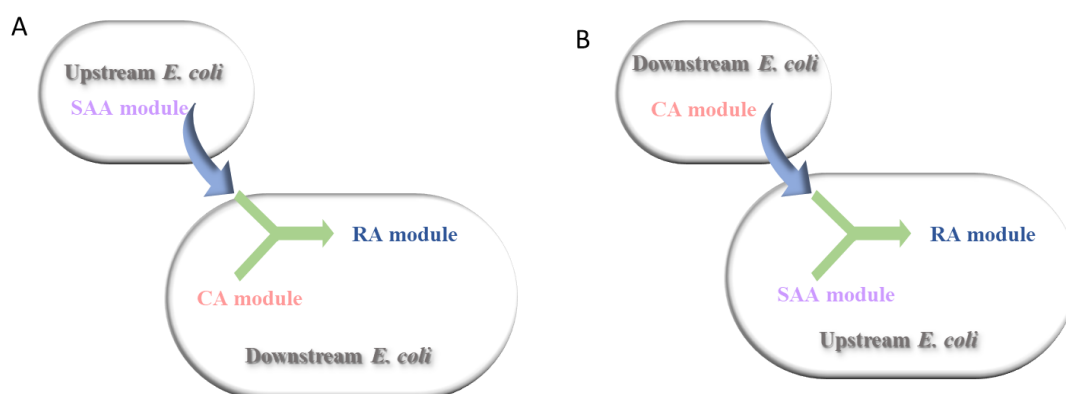
Figure 6.5 Modular designs for RA producing system. A) CA and RA module are overexpressed in the same strain, and SAA module is overexpressed in an independent strain.

B) SAA and RA module are overexpressed in the same strain, and CAA module is overexpressed in an independent strain.

6.3 Material and methods

6.3.1 Plasmids and strains

All strains and plasmids used in this study are summarized in Table 6.1. Primers used for PCR amplification were listed in Appendix.

Table 6.1 Plasmids and strains used in Chapter 6

Plasmids	Description
pTrcHis2B	trc promoter, pBR322 ori, Amp^R
pET28a	T7 promoter, Kan^R
pET21c	T7 promoter, Amp^R
pACYCDuet-1	double T7 promoters, Cm^R
pRSFDuet-1	double T7 promoters, Kan^R
pCDFDuet-1	double T7 promoters, Sp^R
pBR322	Amp^R, Tet^R
pPH0-1	pET28a carrying the *aroE, aroL, aroA* and *aroC* genes under the control of the *proD* promoter (P*proD*)
pTE2	pET28a carrying the *trpE^{fbr}, aroG^{fbr}, aroE, aroL, aroA* and *aroC* gene under the control of T7 promoter

pHACM-rpoA14	a gTME plasmid carrying a mutated alpha subunit of RNA polymerase for enhancing the shikimate pathway
pTrcHis2B-*RgTAL*	pTrcHis2B carrying the codon-optimized *RgTAL* gene
pRSF-*Coum3H*	pRSFDuet-1 carrying the codon-optimized *Coum3H* gene
pCDF-trc-*RgTAL*	pCDFDuet-1 carrying the codon-optimized *RgTAL* gene with a trc promoter
cELACU	pET28a carrying the *E. coli aroE, aroL, aroA, aroC* and *ubiC* genes under the control of the constitutive *Zymomonas mobilis* pyruvate decarboxylase (pdc) promoter
pBA3	pET28a carrying the *E. coli aroE*, *aroL*, *aroA*, *aroC* and *ubiC* genes
pUC57-PDC-VS	pUC57 carrying the *pctV* and *shiA* genes with a constitutive *Zymomonas mobilis* pyruvate decarboxylase promoter
pBS1	pET28a carrying the proD promoter and the *aroE*, *aroL*, *aroA*, and *aroC* genes
pBS2	pET28a carrying the proD promoter and the *aroE*, *aroL*, *aroA*, *aroC*, *tyrA^{fbr}* and *aroG^{fbr}* genes

pBS3	pACYCDuet-1 carrying the *aroE, aroL, aroA, aroC* genes
pBS4	pACYCDuet-1 carrying the *aroE, aroL, aroA, aroC tyrA^{fbr}* and *aroG^{fbr}* genes
pRP1	pET28a carrying the codon-optimized *Lpd-ldh* gene
pRP2	pET21c carrying the codon-optimized *MoRAS* gene
pRP3	pET28a carrying the *hpaBC* gene
pRP4	pET28a carrying the *hpaBC* gene and codon-optimized *Lpd-ldh* gene
pRP5	pET21c carrying the codon-optimized *Pc4CL* gene
pRP6	pET21c carrying the codon-optimized *Pc4CL* and *MoRAS* genes
pRP7	pET21c carrying the *hpaBC* genes
pRP8	pET28a carrying the codon-optimized *MoRAS* gene
pRP9	pAYCYDuet-1 carrying the codon-optimized *Lpd-ldh* gene
pRP10	pAYCYDuet-1 carrying the *tetA* gene

pRP11	pUC57 carrying the codon-optimized *RgTAL* gene with a constitutive *Zymomonas mobilis* pyruvate decarboxylase promoter
pRP12	pCDFDuet-1 carrying the codon-optimized *Pc*4CL and *MoRAS* genes
pRP13	pET28a carrying the *hpaBC* and a*roE* genes

Strains	Description
BL21(DE3)	F$^-$ *omp*T *hsd*S$_B$ (r$_B^-$, m$_B^-$) *gal dcm* (DE3)
K12(DE3)	F- lambda- *ilvG- rfb*-50 *rph*-1 (DE3)
BW25113	F-, Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), λ - , rph-1, Δ(rhaD-rhaB)568, hsdR514
JW4014-2	BW25113 *tyrB::kan*
P2	*E. coli* K12(DE3) Δ*pheA* Δ*tyrR lacZ*::P$_{LtetO-1}$-*tyrA^{fbr}aroG^{fbr} tyrR*::P$_{LtetO-1}$-*tyrA^{fbr}aroG^{fbr}*
P2H	P2 *hisH*(L82R) (DE3)
P6	P2H *ΔptsH ΔptsI Δcrr ΔaroE ΔydiB*
BX	*E. coli* BL21(DE3) *ΔxylA*
P2I	P2H Δ*tyrB*

RAM	P2H carrying pTrcHis2B-*RgTAL*, pRP4 and pRP12
P2HB	P2H carrying pET21c, pET28a and pCDFDuet-1
RAU1	P2H carrying pTrcHis2B-*RgTAL* and pRP3
RAD1	P2H carrying pRP4 and pRP6
RAU2	P2H carrying pTrcHis2B-*RgTAL* and pRSF-*Coum3H*
RAU3	P2H carrying pTrcHis2B-*RgTAL*, pRSF-*Coum3H* and pRP12
RAD2	P2H carrying pRP4 and pET21c and pCDFDuet-1
RAD3	BL21(DE3) carrying pBS4, pRP4 and pRP6
RAD4	P2I carrying pRP4 and pRP6
CAL1	RAU2 carrying pHACM-rpoA14
CAL2	RAU1 carrying pHACM-rpoA14
CAL3	P2H carrying pCDF-trc-*RgTAL* and pRSF-*Coum3H*
CAL4	P2H carrying pCDF-trc-*RgTAL* and pRP3
CAL5	P2H carrying pRP11 and pRP7
CAL6	BL21(DE3) carrying pBS4, pCDF-trc-RgTAL and pRP7

CAL7	BL21(DE3) carrying pBS4, pCDF-trc-RgTAL and pRSF-*Coum3H*
CAL8	BL21(DE3) carrying pBS4, pTrcHis2B-*RgTAL* and pRP7
CAL9	BL21(DE3) carrying pBS4, pTrcHis2B-*RgTAL* and pRSF-*Coum3H*
CAL10	BL21(DE3) carrying pBS2, pCDF-trc-*RgTAL* and pRP7
CAL11	P6 carrying pTrcHis2B-*RgTAL*, pRP13and pHACM-rpoA14
SAL1	P2I carrying pRP9 and pRP7
SAL2	P2H carrying pRP4
SAL3	P2I carrying pRP1 and pRP7
SAL4	P2I carrying pRP4 plasmid
SAL5	P2I carrying pBS3, pRP1 and pRP7
SAL6	P2I carrying pBS4, pRP1 and pRP7
SAL7	P2I carrying pBS1, pRP9 and pRP7
SAL8	P2I carrying pBS2, pRP9 and pRP7

SAL9	BL21(DE3) carrying pBS2, pRP9 and pRP7
SAL10	BL21(DE3) carrying pBS4, pRP1 and pRP7
SAL11	BX carrying pBS2, pRP9 and pRP7
MAM1	BL21(DE3) carrying pRP5, pRP8 and pRP10
MAM2	K12(DE3) carrying pRP5, pRP8 and pRP10
MAM3	P6 carrying pRP5, pRP8 and pRP10

For plasmids construction, genes $d\text{-}ldh^{Y52A}$ [76, 77] and *MoRAS* [78] were codon-optimized and synthesized by Bio Basic Inc, USA. $d\text{-}ldh^{Y52A}$ gene was cloned to pET28a and pACYC-Duet plasmids using NdeI and XhoI restriction sites to yield pRP1 and pRP9. *MoRAS* gene was cloned to pET21c and pET28a using NdeI and XhoI sites to form pRP2 and pRP8. *Pc*4CL gene was PCR amplified with plasmid pCDF-trc-*RgTAL- Pc4CL* using primers ZLPR1CL and ZLPR2CL and inserted into pET21c using NdeI and XhoI sites to generate plasmid pRP5. The *hpaBC* genes were amplified from the BL21(DE3) chromosome using primers ZLPR1HP and ZLPR2HP. The PCR product was digested with NdeI and XhoI and then ligated to pET21c and pET28a treated with the same enzymes to generate plasmids pRP3 and pRP7, respectively. Plasmid pRP4 was constructed by digesting pRP1 with XbaI and XhoI to get the *Lpd-ldh* gene, which was then ligated with pRP3 treated by SpeI and XhoI. Plasmid pRP6 was constructed by digesting pRP2 using XbaI and XhoI sites to get the codon-

optimized *MoRAS* gene [78], which was ligated with pRP5 treated by SpeI and XhoI.

To construct plasmid pRP10, plasmid pBR322 was used as the template for PCR amplification of the *tetA* gene using primers ZLPR1TA and ZLPR2TA. After NdeI/XhoI digestion of the PCR product and pACYCDuet-1, two fragments were ligated to generate plasmid pRP10. Plasmid pRP11 was constructed by digesting pCDF-trc-*RgTAL* with NcoI/SalI and the resulting fragment was ligated with the NcoI/XhoI treated fragment of plasmid pUC57-PDC-VS. Plasmid pRP12 was constructed by digesting pRP6 with BglII/XhoI and inserting the *pc4CL* and *MoRAS* fragment into BamHI/XhoI treated pCDFDuet-1.

For construction of tyrosine overproduction plasmids, a strong constitutive promoter proD [19] and an inducible T7 promoter were used. A DNA fragment containing genes *aroE, aroL, aroA* and *aroC* was isolated from plasmid pPH0-1 (unpublished data) with HindIII/XhoI and then cloned to an engineered pET28a plasmid containing the proD promoter (unpublished data) treated with the same enzymes to generate plasmid pBS1. The same cloning strategy was used to insert the *aroE, aroL, aroA* and *aroC* genes into pACYCDuet-1 to generate pBS3. A DNA fragment containing the genes *tyrA^{fbr}* and *aroG^{fbr}* was PCR amplified with primers ZLPR1AG and ZLPR21AG using the *E. coli* P2H chromosomal DNA as the template. The PCR product was digested with SpeI/HindIII followed by ligation with pBS1 treated with the same enzymes to make plasmid pBS2. The PCR product was also digested with SalI/HindIII followed by ligation with pBS3 treated with the same enzymes to make plasmid pBS4. For plasmid pRP13, an *aroE* fragment was first amplified from plasmid pBA3 using primers ZLPR1AE

and ZLPR2AE and cloned to the XbaI/XhoI sites of plasmid pUC57-pdc-VS [79]. The resulting plasmid pUC57-pdc-*aroE* was digested by XbaI/XhoI to transfer the *aroE* fragment to plasmid pRP3 treated by SpeI/XhoI, generating plasmid pRP13.

For constructing strain P2I, primers ZLPR1PT and ZLPR2PT was used to generate DNA fragment with Keio collection JW4014 template for DNA insertion into strain P2H. The plasmid PCP20 was transformed to remove the kanamycin resistance gene to yield P2I.

6.3.2 Cultivation conditions

LB medium was used for seed culture preparation. M9Y1 medium was used for shake flask cultivation of the engineered strain PGP. One-liter M9Y1 medium was comprised of 5 g glucose, 0.5 g yeast extract, 1 g of NH_4Cl, 3 g of KH_2PO_4, 6.8 g of Na_2HPO_4, 0.5 g of NaCl, 0.24 g of $MgSO_4$, 1 ml trace elements and antibiotics, accordingly. The working concentration of antibiotics was 50 mg/L kanamycin, 100 mg/L ampicillin, 34 mg/L chloramphenicol, 50 mg/L streptomycin. The working concentrations of trace elements were: 0.4 mg/L Na_2EDTA, 0.03 mg/L H_3BO_3, 1 mg/L thiamine, 0.94 mg/L $ZnCl_2$, 0.5 mg/L $CoCl_2$, 0.38 mg/L $CuCl_2$, 1.6 mg/L $MnCl_2$, 3.77 mg/L $CaCl_2$, and 3.6 mg/L $FeCl_2$.

For monoculture, two-strain co-culture and three-strain co-culture biosynthesis in test tube, the seed cultures were first cultivated overnight at 37 ^{o}C in LB medium. The overnight seed cultures were then collected by centrifugation and re-suspended in fresh M9Y1 medium. After OD measurement, desired amounts of seed cultures were collected and inoculated into the M9Y medium to reach a total initial OD_{600} of 0.6. IPTG was added at the beginning of the

cultivation. To identify the optimal cultivation temperature for RA production, cultures of RAM (for monoculture), RAU1 and RAD1 (for two-strain co-culture) were grown at 25 $^{\circ}$C, 30 $^{\circ}$C and 37 $^{\circ}$C for 48 hours. To maintain the same initial OD_{600} of 0.6, RAU1 and RAD1 were inoculated to reach OD_{600} of 0.3, respectively (RAU1:RAD1 =1:1). For the co-culture system under other different ratios, the needed initial OD_{600} for individual strains was calculated based on the inoculation ratio. Proper amounts of cell cultures were then added in the M9Y medium to a total OD_{600} of 0.6. After 48 hours of cultivation, samples were taken for HPLC analysis.

For shake flask cultivation, seed cultures of the involved strains were cultivated in LB medium, respectively. After overnight growth, individual cultures were centrifuged and re-suspended in fresh M9Y medium. After the OD measurement, desired amounts of re-suspended cell cultures were added to 100 mL fresh M9Y medium at different ratios to make a total initial OD_{600} of 0.6. The co-culture was then grown at 37 $^{\circ}$C for 48 hours. The medium used for the shake flask experiments contained 5 g/L glucose for the CAL2:SAL9:MAM2 co-culture and 5 g/L sugar mixtures for the CAL11:SAL11:MAM3 co-culture, respectively. Samples at different time points of the cultivation were taken from the culture for OD measurement, strain-to-strain ratio analysis and HPLC quantification.

7.3.3 Strain to strain ratio determination

The strain-to-strain ratio of the three-strain co-culture was analyzed by the combination of a blue-white screening method and an antibiotic selection method. Specifically in the experiment of using glucose as sole carbon source, 10 μL of the CAL2:SAL9:MAM2 co-culture sample was diluted 10^5 to 10^6-fold before

being spread onto an LB agar plate containing IPTG, X-Gal, ampicillin, kanamycin and chloramphenicol. After 24 h of incubation, the CAL2 strain carrying the disrupted *lacZ* gene generated white colonies while the SAL9 and MAM2 strains carrying the intact *lacZ* gene generated blue colonies (Figure 6.6A). The numbers of blue and white colonies were counted, respectively.30 ~40 blue colonies were then picked and re-streaked on a second plate containing 50 μg/mL tetracycline. Since MAM2 contained the *tetA* gene and SAL9 did not, only MAM2 could make new colonies on the second plate, which was used to distinguish MAM2 and SAL9 (Figure 6.6B). All three co-culture strains' colony numbers were counted separately for calculating their ratio in the co-culture population.
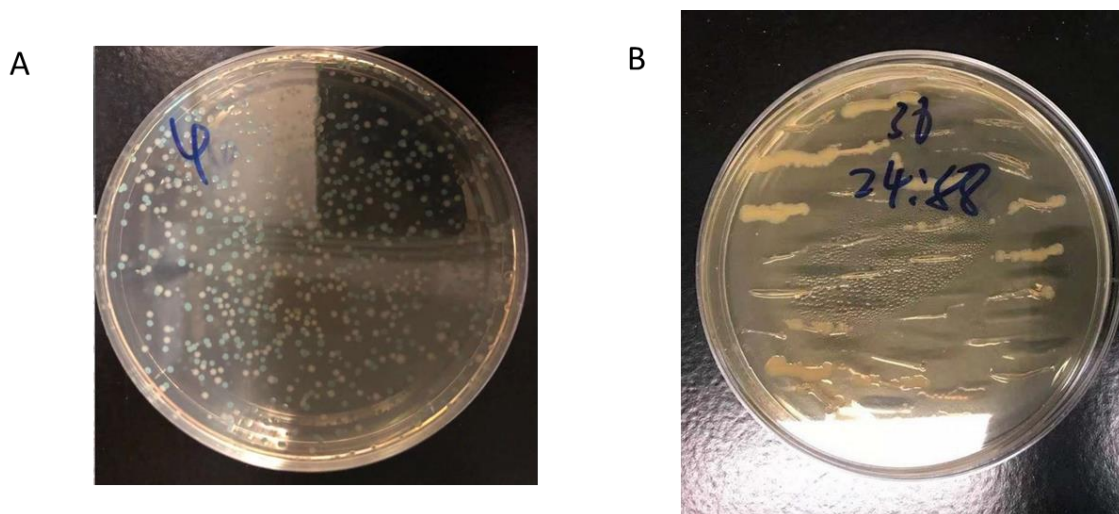
Figure 6.6 Determination of strain to strain ratios in the co-cuture by A) bluewhite colony counting, B) antibiotics resistance differentiation.

For the CAL11:SAL11:MAM3 co-culture grown on xylose/glucose mixture, CAL11 and MAM3 strains formed white colonies on the IPTG and X-Gal plates, while SAA11 formed blue colonies. After counting the numbers of the blue or white colonies, 30~40 white colonies were re-streaked on the tetracycline plates.

New colonies on the tetracycline plates represented the MAM3 strain, which was then used to distinguish SAL11 and MAM3. All three co-culture strains' colony numbers were counted separately for calculating their ratio in the co-culture population.

6.3.4 Metabolites quantification

LC-MS/MS was used for confirmation of the RA biosynthesis. 1 mL cell culture was mixed with 1 mL ethyl acetate by vortex for 30 seconds (10 seconds for three times). The mixture was then centrifuged at 11,000 rpm for 1 min. The supernatant (ethyl acetate phase) was transferred to a clean Eppendorf tube and air dried overnight. The aired samples were dissolved in 1 mL acetonitrile and injected to Agilent 1100 Series HPLC connected with Thermo-Finigan LTQ Mass-Spectrometer. Samples were run through a Waters C18 column using 90% acetonitrile and 10% water for 20 minutes at a flow rate of 0.6 mL/min. Positive-mode ESI was used for ionization, and MS/MS scanning events were set up for the parent ion mass of RA (361 m/z) using 50% ionization energy for fragmentation.

CA, SAA and RA concentrations were determined by HPLC quantification. Culture samples were first collected and centrifuged at 10,000 rpm for 2 minutes. The supernatant was filtered by 0.45 μm PTFE membrane (VWR, Radnor, PA) before subjected to analysis by Agilent 1100 Series HPLC with Photodiode Array detector. The analysis was performed on a Waters C18 column using acetonitrile (solvent A) and water (solvent B) as the mobile phase at a total flow rate of 0.8 mL/min. The following gradient was utilized for elution: 0 min, 100% solvent B; 7 min, 80% solvent B plus 20% solvent A; 9 min, 100% solvent B. Total elution

time is 12 minutes. SAA was measured using absorbance at 280 nm. CA and RA were both measured at 320 nm.

6.4 Results and discussion

6.4.1 RA bioproduction by monoculture and metabolic burden investigation

RA biosynthesis in *E. coli* needs expression of endogenous as well as heterologous enzymes. To construct the RA biosynthesis system, this chapter employed the previously constructed L-tyrosine producer *E. coli* P2H as the baseline strain for mono-culture biosynthesis. P2H is capable of providing a strong tyrosine pathway flux for supporting the CA biosynthesis, which is a RA precursor [12]. A codon-optimized gene encoding D-lactate dehydrogenase (D-ldh) from *Lactobacilluc pentosus* [77]and the *E. coli* native *hpaBC* genes encoding 4-hydroxyphenylacetate 3-hydroxylase were used for converting 4HPP to salvianic acid A (SAA). Notably, the *d-ldh*^{Y52A} mutation was used for its better enzymatic activity. On the other hand, the L-tyrosine was converted to caffeic acid (CA) using the *hpaBC* gene and an engineered *Rhodotorula glutinis TAL* gene encoding tyrosine ammonia lyase. After the two precursors were produced, 4-coumarate:CoA ligase (4CL) from *Petroselinum crispus* [62] and rosmarinic acid synthase (RAS) from *Melissa officinalis* [78] were recruited to combine CA and SAA to form RA (Figure 6.4).
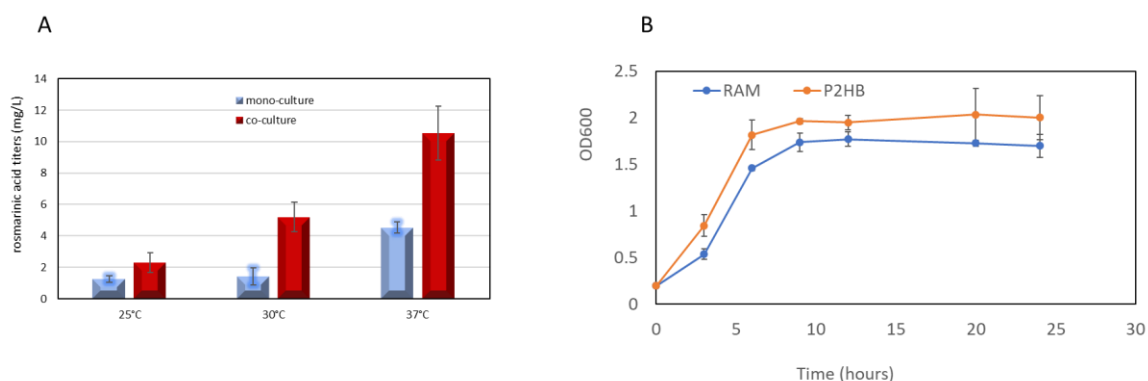
Figure 6.7 RA production by engineered *E. coli*. A) Comparison of monoculture strain RAM and RAU1/RAD1 co-culture inoculated at 1:1 ratio for RA production at three different temperatures; B) Growth comparison of strain RAM harboring the RA pathway and control strain P2HB with the RA pathway.

All the selected pathway genes were cloned into plasmid vectors and subsequently introduced into *E. coli* for reconstitution of the RA pathway. The resulting strain RAM harboring the entire pathway was cultivated on 5 g/L glucose as a monoculture for the RA biosynthesis. To verify the RA was indeed produced, culture samples was analyzed by LC/MS/MS. The LC/MS/MS chromatograms and mass spectra of the RA standard and the *E. coli*-produced RA matched well, indicating that the desired RA product was produced by *E. coli* (Figure 6.8). Therefore, the constructed heterologous RA pathway in *E. coli* was confirmed to be functional as desired.

To investigate the effect of temperature on RA bioproduction, 3 different temperatures 25 °C, 30 °C and 37 °C were tested. As shown in Figure 6.7, the highest RA production was achieved at 37 °C. This result shows that the RA producing system didn't require a lower temperature to help protein folding. The highest production by the mono-culture at the optimal temperature 4.5 mg/L.

Another phenomenon observed in this experiment was the cell growth change. The P2HB strain has the same metabolic baseline with RAM, but it contains only empty plasmids without the RA pathway genes. P2HB was cultivated side by side with RAM strain under the same condition. The result in Figure 6.7B showed a 25% lower cell density of RAM than P2HB strain, indicating an obvious growth disadvantage caused by the metabolic burden associated with expressing the RA pathway.

6.4.2 Use two strains co-culture system to improve the RA produciton

To apply modular co-culture engineering strategy for the RA biosynthesis, a two strain co-culture system was designed using two *E. coli* strains to accommodate the modularized pathway (Figure 6.9A). In this co-culture design, the upstream strain RAU1 was only used for producing precursor CA, while the downstream strain RAD1 containing both the SAA and RA modules was responsible for both SAA production and RA assembly. This design allowed each co-culture strain to undertake only part of the biosynthetic labor and thus reduced the associated metabolic burden. Moreover, the two strains' ratio inside the co-culture could be manipulated for the purpose of pathway balancing.
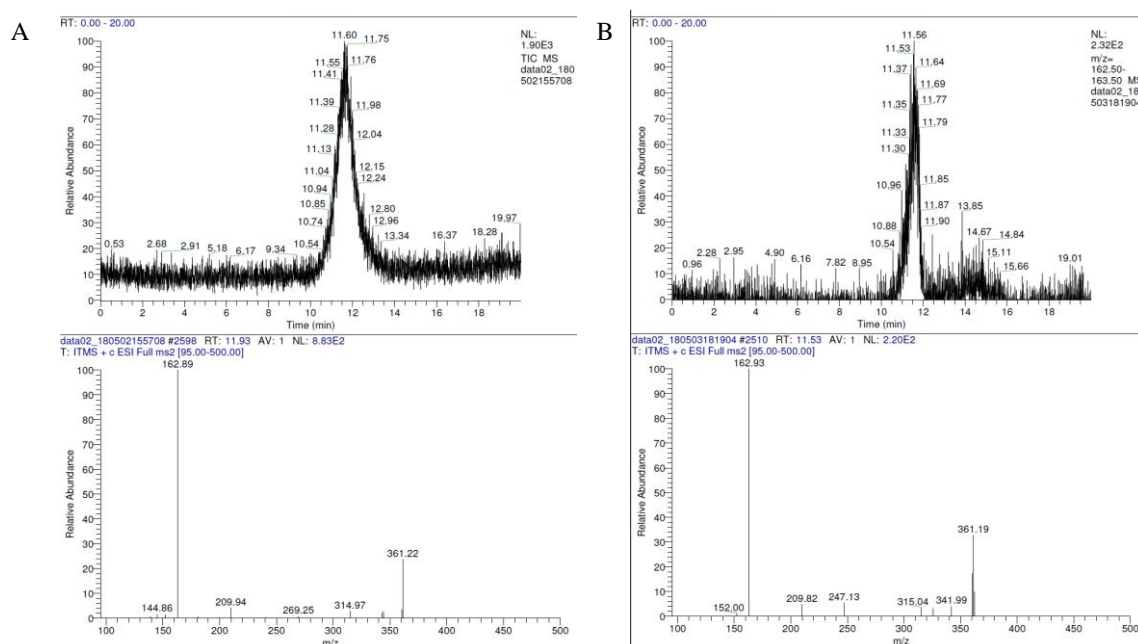
Figure 6.8 Qualitative analysis of RA produced by engineered *E. coli*. A) LC/MS/MS chromatogram and speactrum for RA standard; B) LC/MS/MS chromatograms and spectrum for RA produced by *E. coli*.

The constructed RAU1 and RAD1 strains were co-cultivated with the inoculation ratio of 1:1 for RA production. As shown in Figure 6.7A, the production suggested that the RA production by co-culture also preferred a higher temperature, which is in good agreement with the mono-culture bioproduction results. The production at 37 °C by the co-culture was 10.5 mg/L, which was 2.3 folds higher than monoculture strain.

Next, the RA biosynthesis was optimized by changing the initial inoculation ratio of the two co-culture strains, which enabled the flexible adjustment of the biosynthetic strengths between the corresponding pathway modules. As shown in Figure 6.9B, the increase of the RAU1:RAD1 inoculation ratio (more RAU1 and less RAD1) resulted in the higher biosynthetic ability of the upstream CA module and higher strength of the SAA and RA modules. With the increase of the inoculation ratio, the corresponding pathway modules' biosynthetic capabilities changed, which resulted in decreased CA concentration and increased SAA

concentration. The CA and SAA provision was best optimized at the inoculation

ratio of 3:1, which led to the RA production of 12 mg/L.
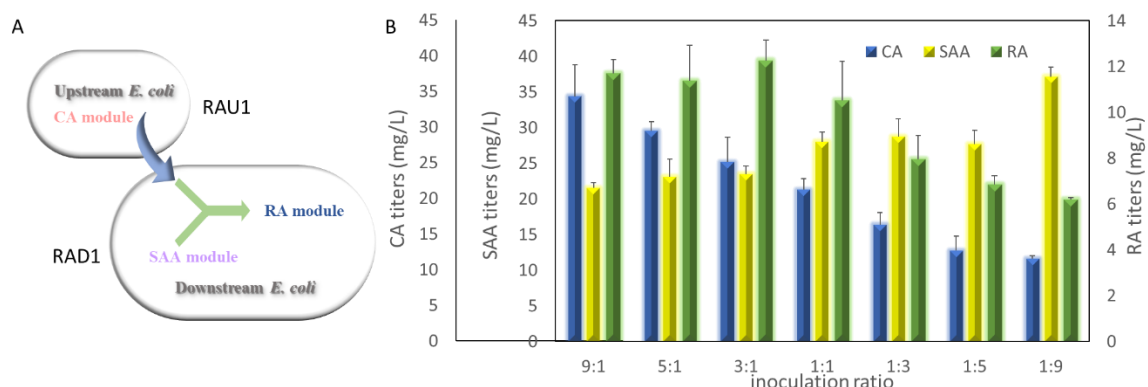
Figure 6.9 Two-strain co-culture design 1 of RA production. A) Co-culture design for RAU1 and RAD1; B) RA, CA and SAA concentration of co-culture RAU1 and RAD1.

RA biosynthesis was further modified by adapting traditional metabolic engineering strategies in the context of the two-strain co-culture. The pathway enzyme responsible for converting pCA to CA was changed from *E. coli* 4-hydroxyphenylacetate 3-hydroxylase complex HpaBC to *Saccharothrix espanaensis* 4-coumarate 3-hydroxylase Coum3H for better CA biosynthesis, which has been reported before [61]. When the new co-culture RAU2:RAD1 harboring the Coum3H from *S. espanaensis* was used, the RA biosynthesis increased under all inoculation ratios, indicating a large CA provision improvement by this strategy. The SAA concentrations were still similar to that of the RAU1:RAD1 co-culture, as the SAA provision capability was not changed in the RAD1 strain. The highest RA biosynthesis of 48 mg/L by the RAU2:RAD1 co-culture was 4 times higher than the RAU1:RAD1 co-culture. Notably, the optimal inoculation shifted from 3:1 to 9:1, suggesting that the condition for pathway balancing was changed due to the biosynthesis capability enhancement

of the CA module. Coum3H was thus used for the CA module in the following co-culture studies.
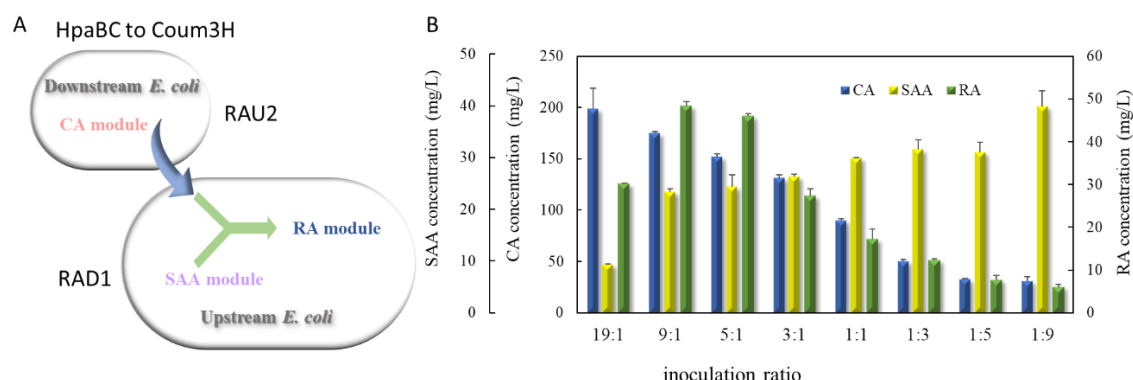
Figure 6.10 Two-strain co-culture design 2 of RA production. A) Co-culture design for RAU2 and RAD1; B) RA, CA and SAA concentration of co-culture RAU2 and RAD1.

For the first two designs of co-culture system, both strains used P2H as the host strain. Previous researches suggested *d-ldh* was better performed in *E. coli* BL21(DE3) which was commonly used for expression of heterologous enzyme (*pc*4CL and *mo*RAS) in *E. coli*. So, the third co-culture strategy was developed (Figure 6.11A). Surprisingly, the resulting co-culture RAU2: RAD3 produced merely 20 mg/L RA from 5 g/L glucose after the inoculation ratio optimization (Figure 6.11B). This RA concentration was much lower than that of the RAU2:RAD1 co-culture, although the CA and SAA accumulation was still similar. This finding indicated that the RA module's activity in BL21(DE3) strain was not well reconstituted. Also, the SAA provision was also not improved, revealing another bottle neck for SAA biosynthesis. SAA used 4HPP as precursor. However, 4HPP could also be converted to tyrosine by gene *tyrB*, which limited the formation of SAA due to pathway competition.
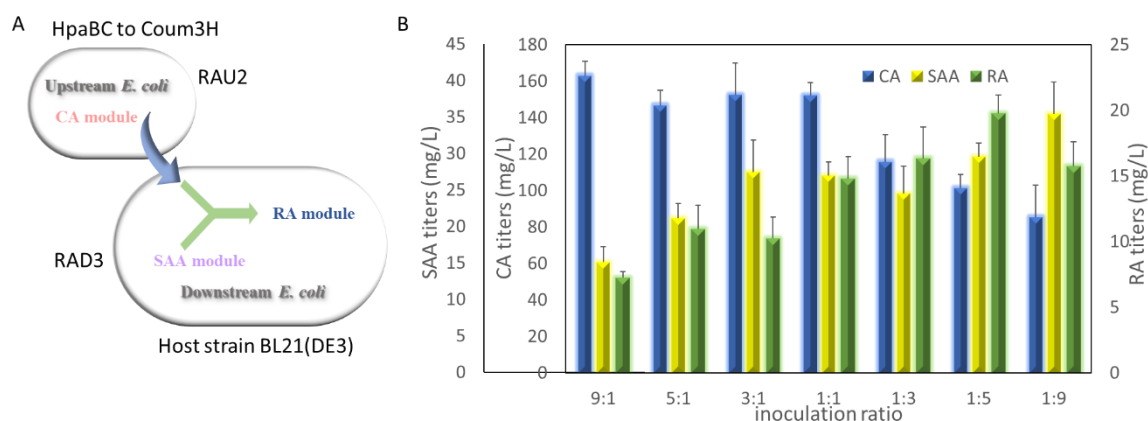
Figure 6.11 Two-strain co-culture design 3 of RA production. A) Co-culture design for RAU2 and RAD3; B) RA, CA and SAA concentration of co-culture RAU2 and RAD3.

Since the third design was not improving the RA production, we refocus on the second co-culture system. For next step, we reconstituted the biosynthesis system by moving the RA module to the strain harboring the CA module and the SAA module was overexpressed in an independent strain (Figure 6.12A). The resulting co-culture strains RAU3 containing the CA and RA modules and RAD2 containing the SAA module were inoculated at different ratios for the RA biosynthesis. As shown in Figure 6.12B, RA bioproduction by this RAU3:RAD2 co-culture was dramatically reduced under all inoculation ratios. In fact, for this design of (CA+RA):SAA co-culture, the strain RAU3 harboring the CA and RA module was imposed with excessive metabolic stress, as indicated by the significantly lowered CA concentrations compared with the previous CA:(SAA+RA) design (200 mg/L vs. 30 mg/L). Also, the SAA provision was improved to around 80 mg/L, as expected. Therefore, the RA bioproduction reduction was a result of the insufficient supply of CA and an excess amount of SAA provision. On the other hand, although the SAA and CA modules can be relatively balanced by changing the inoculation ratio in the new co-culture, the relative biosynthesis strengths between the CA and RA modules were fixed in

strain RAU3. Hence, the lack of the CA and RA module balancing also contributed to the low RA bioproduction in the RAU3:RAD2 co-culture.

Figure 6.12 Two-strain co-culture design 4 of RA production. A) Co-culture design for RAU3 and RAD2; B) RA, CA and SAA concentration of co-culture RAU3 and RAD2.
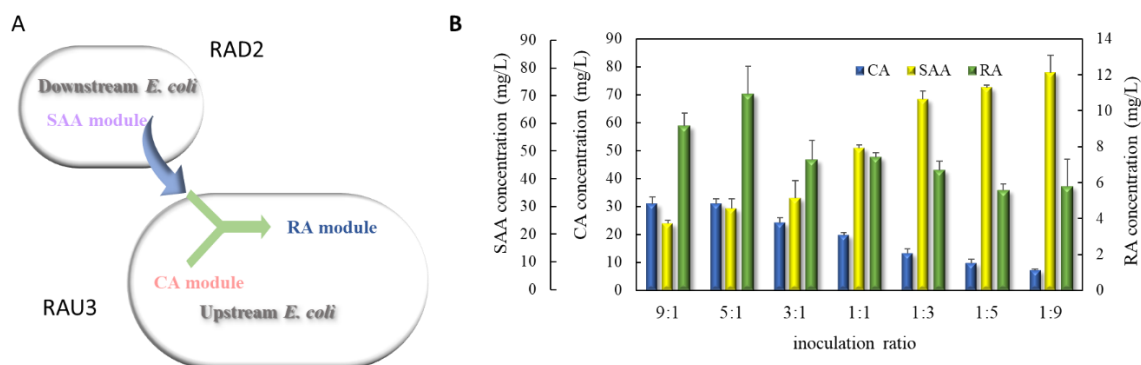
The strategy to integrate RA and CA module together didn't help improving RA production, either. We tried to use the metabolically engineered strains to improve the SAA availability. In the second and third strategy, the *tyrB* gene encoding the tyrosine aminotransferase wasn't deleted from the downstream strain's chromosome, which enabled some metabolic flux towards tyrosine biosynthesis. So, the downstream strain RAD1 was engineered to delete the *tyrB* gene [76]. As shown in Figure 6.13A, the resulting co-culture RAU2:RAD4 showed higher accumulation of SAA at most inoculation ratios, compared with the RAU2:RAD1 co-culture above. Accordingly, the RA biosynthesis was improved to 60 mg/L after the inoculation ratio optimization. It was also observed that with the inoculation ratio decrease (less inoculum for the upstream strain), the CA accumulation decreased, and the SAA accumulation increased. This finding was consistent with the relative biosynthetic capability change for the corresponding pathway modules. It should be noted that the *tyrB* deletion strategy cannot be employed in the mono-culture design, as it will eliminate the tyrosine

provision for the CA module and undermine the overall RA biosynthesis. This situation highlighted an outstanding advantage of module co-culture engineering for physically segregating pathway modules in separate strains to individually satisfy their different biosynthesis needs, which is challenging to achieve by mono-culture engineering.
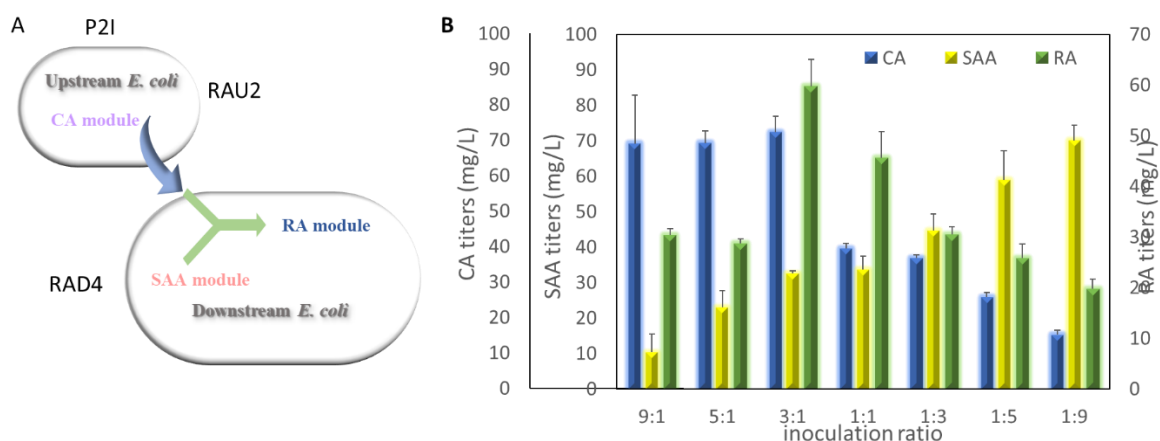
Figure 6.13 Two-strain co-culture design 5 of RA production. A) Co-culture design for RAU2 and RAD4; B) RA, CA and SAA concentration of co-culture RAU2 and RAD4.

Although RA production was improved to a decent level, there is a lot of potential for further development of this biosynthesis system. For the RAU2:RAD1 and RAU3:RAD2 co-cultures. They were similar system with exactly the same host strains but different in that RA module was combined with either CA or SAA module. From the perspective of precursors accumulation, when RA was combined with SAA module, the SAA provision dropped from 80 mg/L to 35 mg/L. On the contrary, CA concentration dropped from around 200 mg/L to 30 mg/L if CA and RA modules were integrated together. The challenge was a system with better availability of both precursors was needed. To overcome this challenge, a third co-culture strain was introduced, and a three strains co-culture system was established, as discussed next

6.4.3 RA production by a three-strain co-culture

The RA biosynthesis improvement by the two-strain co-culture design in last section showed the promising prospect of engineering microbial co-cultures for advancing microbial biosynthesis of complex natural products. However, two strains co-culture engineering still raised issues such as the precursors provision. Either co-culture design RA+SAA with CA or RA+CA with SAA wasn't a good solution to addressing this issue. A third strain was considered necessary for further unleashing the biosynthetic power of modular co-culture engineering.
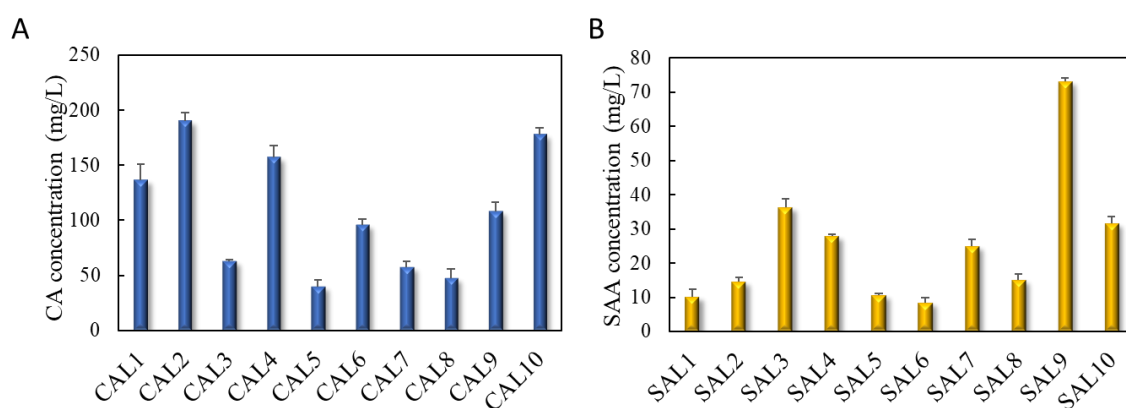
Figure 6.14 selection of strains for accommodating the CA and SAA modules. A) Comparison of 10 CA module strains constructed for CA production; B) Comparison of 10 SAA module strains constructed for SAA production.

To this end, a series of *E. coli* strains was constructed and screened to identify the best performers for expressing individual pathway modules. Specifically, 10 *E. coli* strains (combinations of different background strains and the CA module plasmids) were screened for the CA biosynthesis capability. As shown in Figure 6.14A, strain CAL2 produced the highest concentration of CA (190 mg/L) from 5 g/L glucose. Similarly, another 10 *E. coli* strains were constructed and screened for the SAA biosynthesis capability. As shown in Figure

6.14B, strain SAL9 was identified to be the highest producer of SAA which produced 80 mg/L SAA.

Moreover, RA module was introduced to two *E. coli* strains, BL21(DE3) and K12(DE3), with different genotypic characteristics, and resulted in strains MAM1 and MAM2, respectively. Two three-strain co-cultures CAL2:SAL9:MAM1 and CAL2:SAL9:MAM2 were then constructed to produce RA from glucose.
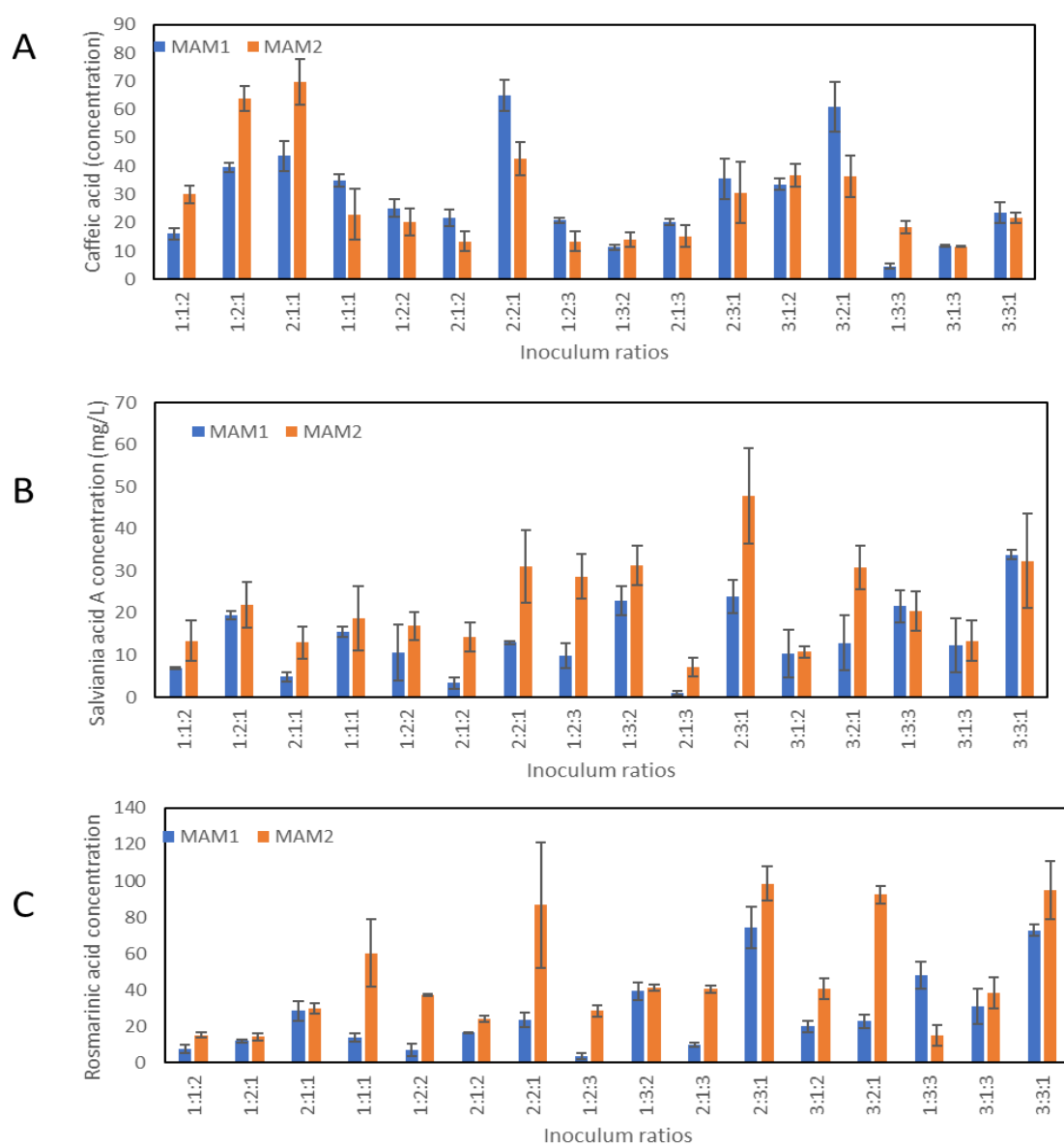
Figure 6.15 RA production by 3-strain co-cultureCAL2:SAL9:NAM1 and CAL2:SAL9:MAM2. A) CA concentration for three strain co-culture system; B) SAA

concentration for three strain co-culture system; C) RA concentration for three strain co-culture system.

The interaction between the co-culture members in the three-strain system was more complex and therefore required more sophisticated balancing. To optimize the RA biosynthesis using these new co-cultures, the inoculation ratio between all three constituent strains were adjusted to coordinate the biosynthetic strengths of the three modules. As shown in Figure 6.15, the inoculation ratio change tremendously influenced the RA production. The RA titer fluctuated significantly with the inoculation ratio change. The increase of any strain's inoculation led to the strengthening of corresponding pathway modules and the corresponding production performance change.

For the CAL2:SAL9:MAM1 co-culture, when three strains were inoculated at 1:1:2 ratio, RA was produced the lowest titer of 8 mg/L. In contrast, the RA concentration was improved to 74 mg/L at the optimal inoculation ratio of 2:3:1. Also, relatively higher RA production was achieved when the SAA module strain (SAL9) was inoculated at higher ratios. Less CA module and more SAA module led to better RA production. It was therefore indicated that there was strong biosynthetic capabilities imbalance between three pathway modules, which could be compensated through changing the strain-to-strain ratio in the co-culture population.

For the CAL2:SAL9:MAM2 co-culture, the change of the RA production with the inoculation ratio showed different RA production levels. RA concentrations were overall higher than CAL2:SAL9:MAM1 at most inoculation ratios, which was consistent with our previous finding that BL21(DE3) strain was not a good expression host for the RA module (Figure 6.11). CA and SAA accumulation in the CAL2:SAL9:MAM1 and CAL2:SAL9:MAM2 co-cultures

also showed varied profiles, as shown in Figure 6.15A and B. These findings showed that the use of different strains (MAM1 and MAM2) for accommodating the RA module changed the relative biosynthetic strengths between individual pathway modules and thus resulted in varied co-culture biosynthesis behavior. However, the accumulation of CA and SAA fluctuated significantly with the changes of the initial inoculation ratios. Further investigation of dynamic ratio change may offer a good explanation for this. Nonetheless, the highest production of 98 mg/L and the optimal inoculation ratio of CAL2:SAL9:MAM2 co-culture was 2:3:1, a 63% higher than the optimal production by the two-strain co-cultures (60 mg/L by RAU2:RAD4), indicating that the RA biosynthesis ability was better balanced in the context of three-strain co-culture.
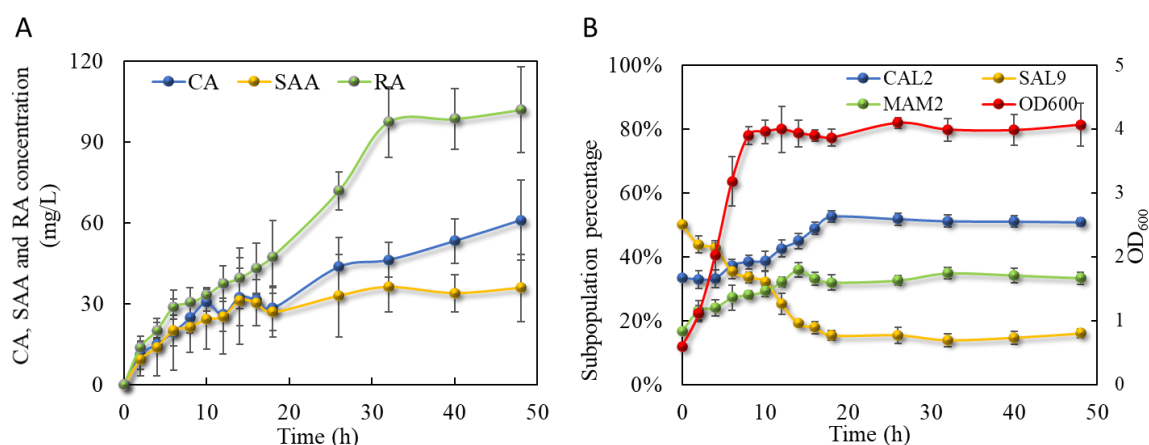
Figure 6.16 RA production dynamics by the three strain co-culture CLA2:SAL9:MAM2. A) Time profiles of CA, SAA and RA concentrations ; B) Time profiles of cell density and subpopulation percentage of the three strains.

Next, the dynamics of the three-strain co-culture biosynthesis was further analyzed by growing CAL2:SAL9:MAM2 on 5g/L glucose in shake flasks with the inoculation ratio of 2:3:1. As shown in Figure 6.16B, the cell density of the co-culture developed with time and reached a plateau at around 8 h. Interestingly, the co-culture population composition showed a highly dynamic change over time.

The ratio between co-culture strains' sub-populations started to change immediately after the inoculation. The relatively population size of the SAA module strain SAL9 reduced from 50% to 34% at 8 h. After the co-culture growth entered the stationary phase, this percentage kept declining until it stabilized at around 15% after 18 h. In contrast, the population percentages of the strains harboring the CA and RA modules increased over time from 33% and 16.7% and stabilized at the plateaus of 53% and 32%, respectively. At 18 h. The delay of the co-culture population stabilization compared with cell density stabilization indicated that the dynamic change of the strain-to-strain ratio and overall cell growth were not synchronized. It was also clearly shown that the SAA module strain was at growth disadvantage in this co-culture, and thus a higher amount of inoculation of this strain was required to compensate its growth decline during the cultivation.

The concentration profiles of CA, SAA and RA was shown in Figure 6.16A. Both CA and SAA accumulation was observed before 18h and stabilized after 28 h. Despite the large difference in the sub-population sizes of the corresponding strains, CA and SAA concentrations were maintained at comparable levels, which was consistent with the need of even provision of these precursors for the downstream conversion. The RA concentration increased gradually beyond the exponential growth phase and plateaued in the middle of the stationary phase. The RA production reached 102 mg/L after 32 h. Overall, the concentration profile of CA, SAA and RA suggested that the production of pathway metabolites was not entirely synchronized with the co-culture strains' growth.

6.4.4 Three strains co-cultivation with two carbon substrates

Next, we attempted to further stabilize the co-culture population composition in order to achieve additional RA biosynthesis improvement. To this end, the co-culture strains were engineered to grow on separate carbon substrates to reduce the growth competition against each other. Specifically, a previously constructed *E. coli* strain P6 with disrupted glucose uptake system (deletion of genes *ptsH*, *ptsI*, *crr*) was used to accommodate the CA and RA modules, respectively, to generate strains CAL11 and MAM3. Meanwhile, *E. coli* strain BX with disabled xylose metabolic pathway (deletion of gene *xylA*) was used to accommodate the SAA module to generate strain SAL11. The CAL11:SAL11:MAM3 co-culture was then cultivated on a sugar mixture of glucose and xylose. In this system, xylose was the preferred carbon substrate for CAL11 and MAM3, whereas glucose was the preferred carbon substrates for SAL11. Such a design allowed us to better manipulate the growth of individual co-culture strains.
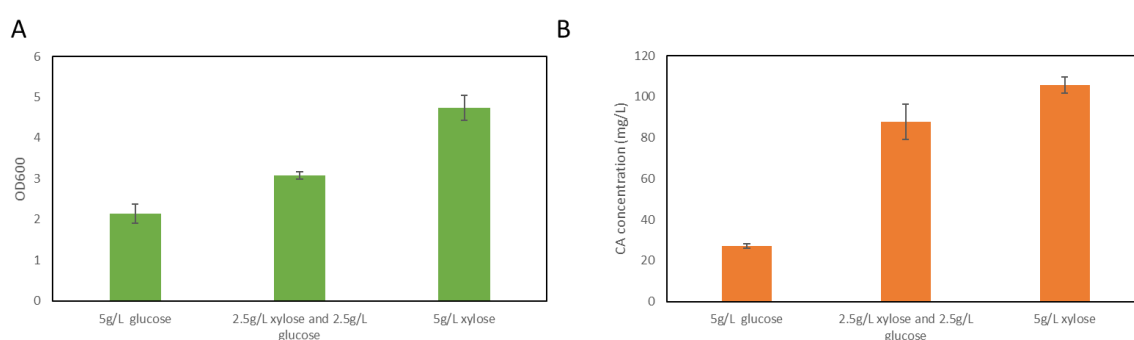
Figure 6.17 CA production with different carbon sources. 5g/L glucose, 2.5 g/L glucose/2.5g/L xylose and 5 g/L xylose, respectively.

Before the RA production experiment, the CA production ability of CAL11 was tested. P6 strain was designed for overexpress DHS instead of tyrosine and the genes *aroE* and *ydiB* were deleted. The *aroE* gene was therefore cloned into

pACYCDuet-1 to recover the tyrosine production. The CA producing experiment for the new constructed strains was performed in 5g/L glucose, 2.5 g/L glucose/2.5g/L xylose and 5 g/L xylose, respectively. Results are shown in Figure 6.17.

For 5 g/L glucose, xylose strain exhibited the lowest cell density after 48 h growth (OD_{600} from 0.6 to 1.8 around) and produced lowest amount of CA. When half glucose and half xylose were used as carbon sources, the cell growth improved marginally while the production reached 90 mg/L. When 5g/L xylose was used, the production of CA reached 120 mg/L. The production performance was not comparable to the CAL2 strain (190 mg/L), as the production ability might be limited by the xylose consumption rate. With decent amount of CA provision, the strain could be used in three strains co-culture system with two carbon substrates.
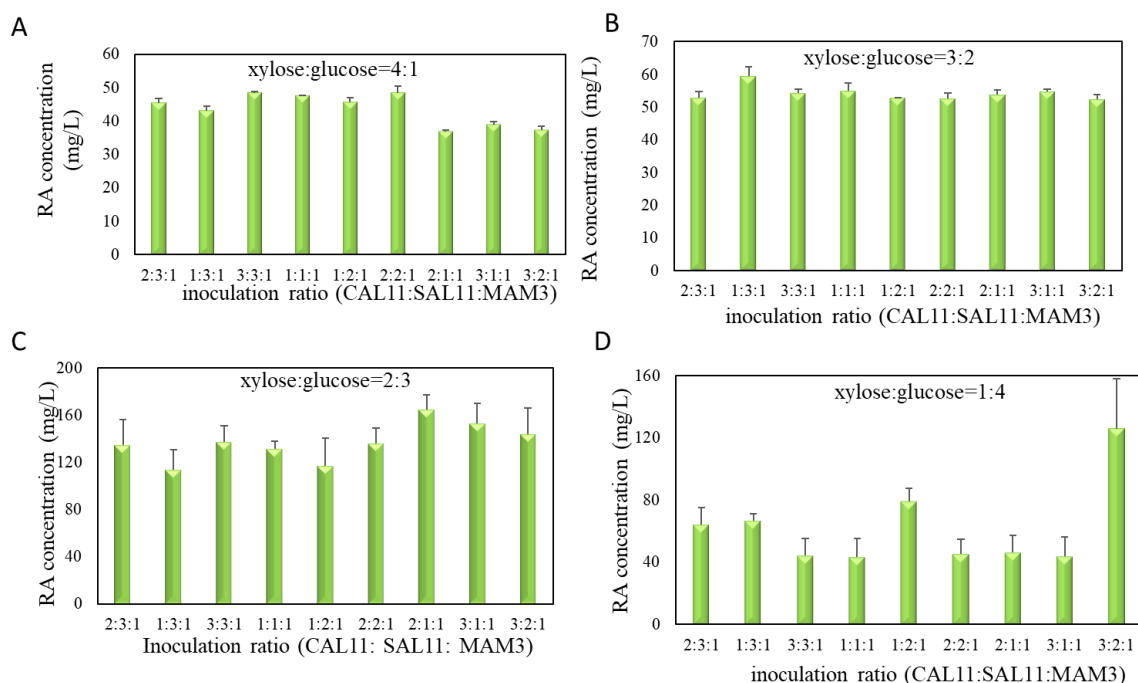
Figure 6.18 RA production with different carbon substrate ratios and inoculum ratios.

For the RA biosynthesis using different carbon sources, composition of sugar mixture was first optimized. Specifically, four compositions, including xylose to glucose mass ratio of 4:1, 3:2, 2:3 and 4:1 (5 g/L sugar in total), were used for growing the CAL11:SAL11:MAM3 co-culture. As shown in Figure 6.18, the RA concentration showed very different profiles at these conditions. Overall, low glucose content (xylose: glucose=4:1, Figure 6.18A) only produced low RA biosynthesis, as there was no enough carbon substrate glucose for supporting the SAA module strain's biosynthetic activity, which has been found to be a potential limiting factor for the RA biosynthesis in one sugar cultivation (Figure 6.16B). Similar production was also observed when the xylose:glucose ratio was 2:3 (Figure 6.18C). On the other hand, low xylose content (xylose:glucose=1:4, Figure 6.18D) was not enough for providing enough carbon substrate for the CA and RA module strains and thus generated sub-optimal production performance, although the highest production reached 125 mg/L at the inoculation ratio of 3:2:1.
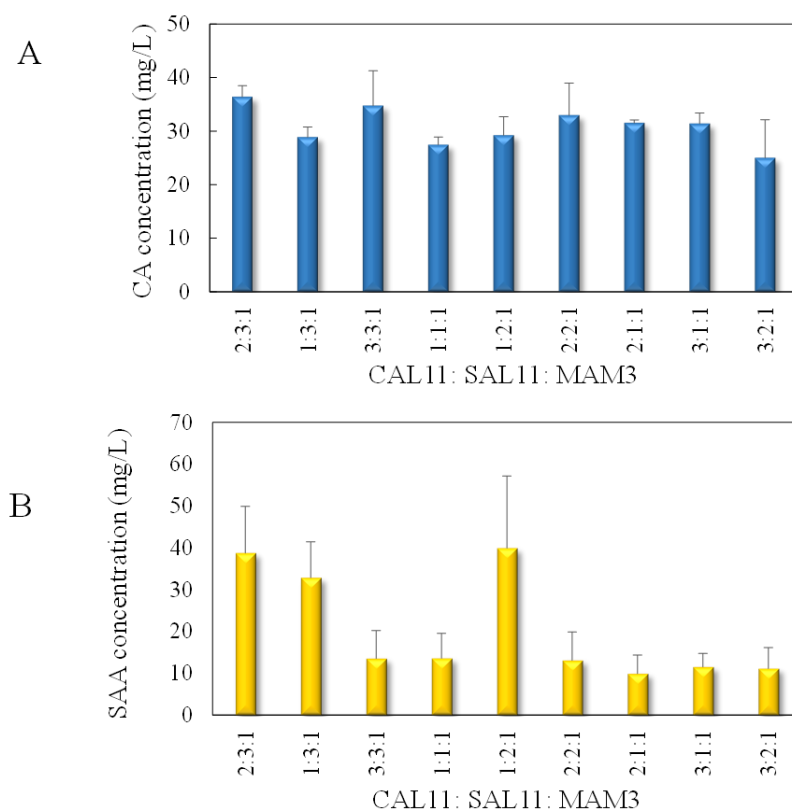
Figure 6.19 CA and SAA accumulations with xylose: glucose=2: 3 at different inoculum ratios.

In comparison, when 2 g/L xylose and 3 g/L glucose was used for cultivation, the CAL11:SAL11:MAM3 produced greater than 110 mg/L at all tested inoculation ratios (Fig. 6C). At the optimal ratio of 2:1:1, the RA concentration reached 165 mg/L, 1.7-folds higher than the co-culture grown on sole carbon source of 5 g/L glucose. Furthermore, it was found that, although the CA accumulation was relatively stable with the change of the inoculation ratio, the SAA accumulation varied to a large degree (Figure 6.18). Low SAA build-up (10 mg/L) at the optimal inoculation ratio of 2:1:1 suggested relatively thorough SAA bioconversion to RA under this condition, although there was still 32 mg/L CA remain un-converted. It is noteworthy that the optimal inoculation ratio shifted from 2:3:1 for one-sugar cultivation to 2:1:1 for two-sugar cultivation. This

suggested that the SAA module strain SAL11 had better growth in the CAL11:SAL11:MAM3 co-culture (as the easy carbon substrate glucose was solely assimilated by SAL11 without interference with the other two strains) and thus reduced the need of high inoculum for growth coordination with the other two module strains. On the other hand, it was shown that when carbon substrate was switched from glucose to xylose (CAL2:SAL9:MAM2 vs CAL11:SAL11:MAM3), the CA and RA module strains were still able to utilize the new carbon substrate xylose for meeting the growth and biosynthesis needs. These results clearly demonstrated that engineering co-culture strains to grow on separate carbon substrates was a viable strategy to improve their growth compatibility and the overall co-culture biosynthesis performance.

Notably, inoculation ratios didn't change the RA production dramatically as observed in three strain co-culture system using two carbon substrates (Figure 6.18). The rationale of this was that when carbon substrate composition was fixed, the competition between SAA and CA/RA strain was limited as they consumed respective sugar without any interference. Also, the competition between RA and CA module was also reduced because the low consumption rate of xylose. Further modification can incorporate a third carbon substrate and each module consumes the specific carbon sources. By such a design, the three modules can be further balanced.
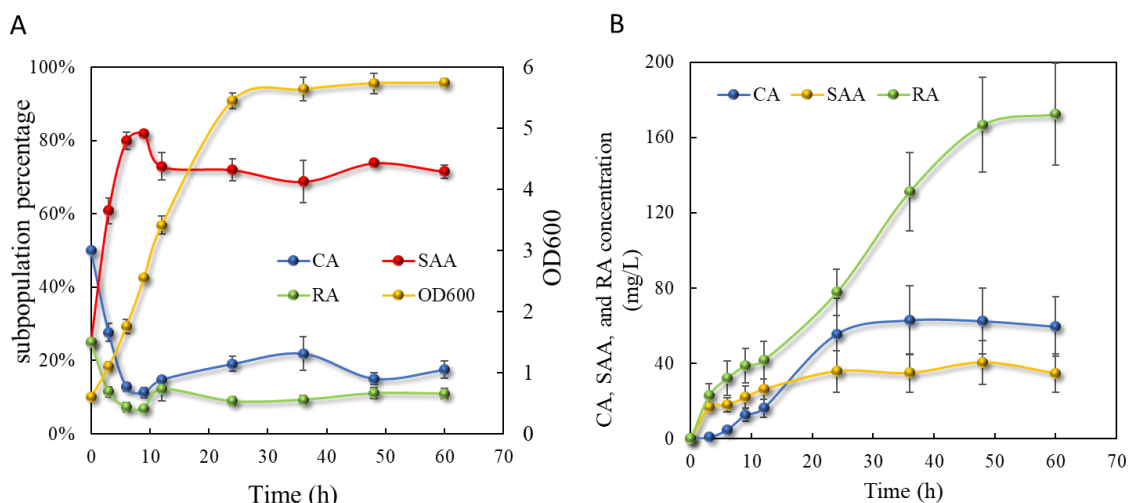
Figure 6.20 RA production dynamics for three strains co-culture grown on two carbon substrates. A) Time profiles of cell density and subpopulation percentage of the three strains. B) Time profiles of CA, SAA and RA concentrations for three strain co-culture system.

After the best carbon substrate composition and inoculation ratio were identified in the three strains co-culture grown on 2 carbon sources, efforts were made to further investigate the dynamics of the CAL11:SAL11:MAM3 co-culture by analyzing time samples of the co-culture grown in shake flasks. As shown in Fig. 6.20A, the cell density of the co-culture increased with time and plateaued after 24 h. Compared with the co-culture grown on single sugar glucose, it took longer for the co-culture grown on glucose/xylose mixture to enter the stationary phase. This was because the uptake of xylose was slower and the biomass was accumulated in a lower rate.

The time profiles of the individual strains' sub-population size variation showed different patterns. SAL11 percentage quickly increased from 25% to around 80% within 6 h and leveled off at around 70% for the rest of the cultivation period. Compared with its low percentage in the one-sugar cultivation (Figure 6.19A), the SAA module strain in the two-sugar cultivation showed much better

growth profile. This is again due to its better growth on glucose without the competition of the other co-culture strains, which helped to enhance the SAA provision for the pathway and thus contributed to the final RA production improvement. The percentage of CAL11 decreased rapidly after the inoculation and fluctuated between 10% and 20% toward the end of cultivation. Similar trend was found for MAM3 whose percentage dropped from 25% and stabilized at around 10% after 12 h. These findings clearly showed that the use of the two sugars to support the co-culture strain growth effectively changed the co-culture population composition's dynamic development with time, which in turn generated a new bioproduction behavior different from the one-sugar cultivation case.

The concentration change of CA, SAA and RA is shown in Figure 6.20B. It was observed that CA was accumulated to around 60 mg/L at 24 h and stabilized at this level toward the end of cultivation. In comparison, SAA accumulation was only higher than CA in the first 12 h, but overall, it fluctuated in a relatively small range throughout the cultivation. Although SAL11 occupied the majority of the co-culture population, the SAA concentration was similar to what was observed for the one sugar cultivation. This suggested that the two-sugar strategy did not necessarily increase the pool size of the SAA precursor; instead, it improved the carbon flux through SAA (and CA) for the RA biosynthesis. It was also found that RA concentration steadily increased over time until it leveled off at 48 h. 172 mg/L RA was produced at the end of the cultivation. Interestingly, occurrence of the RA biosynthesis was not limited to the exponential phase. In fact, a significantly fraction of the RA bioproduction took place in the stationary phase, which suggested that there was still stable carbon flux going into the RA pathway

at this stage. The dynamic analysis of the co-culture cultivation also showed that the strain-to-strain ratio in the co-culture was not necessarily consistent with the ratio between biosynthesis capabilities of the corresponding pathway modules. For example, although CAL11 subpopulation was much bigger than SAL11, the CA and SAA supplies by these two strains were maintained at a compared level, which facilitated the RA bioproduction optimization. Nonetheless, these results revealed the dynamic behaviors of the co-culture's growth and biosynthesis and validated the rationale of the co-culture design for addressing the needs of non-linear RA biosynthetic pathway.

6.5 Summary

For metabolic engineering, balancing the metabolic flux is a critical issue. Traditional metabolic engineering tool uses strategies such as changing promoters, gene copy numbers and deletion of competitive pathways. However, it requires a lot of effort and may generate only suboptimal results for each particular case. Even for simple linear metabolic pathways such as phenol and 4-hydroxystyrene, precursor L-tyrosine provision and downstream enzymatic conversion are still hard to balance. Also, the choice of promoters and plasmids backbones are limited, and expression level are hard to control delicately in each host strains. For RA, the system is even more challenging for balancing due to the complexity of the two precursors formation and conversion. The CA and SAA supply cannot be easily balanced because they share the same precursor 4HPP. Moreover, the HpaBC can catalyze the byproduct formation using 4HPP and pCA. Expression of a large number of genes in a system can also be problematic for cell fitness.

The led to a low cell density of cell growth, indicating the overall lower metabolic vitality of cell.

The co-culture strategy, however, offers another level of perspective for optimizing microbial bioproduction systems. The complicated systems can be simplified by allocating the two precursors formation pathway into two different host strains. Each strain is only responsible for producing one precursor and the metabolic burden for each cell was greatly reduced. Also, the provision for each precursor can be easily tuned by changing the initial inoculation ratio to find the best combination, which is more straightforward compared to traditional metabolic engineering strategies. Also, the co-culture systems provide physical barrier to segregate enzymes in each module and prevent the undesired promiscuous activity.

The advantage of co-culture design was testified by RA bioproduction. Monoculture for RA production only produced 4.5 mg/L RA, which was far lower than the 60 mg/L production by the engineered co-culture. Importantly, biosynthesis coordination between CA and SAA+RA modules, or between CA+RA and SAA modules was achieved through straightforward manipulation of the inoculation ratio of co-culture strains. It is noteworthy that the use of the co-culture design offered the unique opportunity for implementing *tyrB* deletion to enhance the SAA module. This cannot be achieved in the context of mono-culture due to the CA module's reliance on the TyrB enzyme. Therefore, it serves as a great example that the pathway module segregation by modular co-culture engineering possesses unparalleled advantages over mono-culture engineering.

The two-strain co-culture approach mainly addressed the imbalance between CA and SAA+RA modules. Further recruitment of the three-strain co-

cultures enabled more flexible balancing between all three individual modules. Although CA and SAA compounds needed to be provided at 1:1 molar ratio for the downstream RA formation, the specific biosynthetic strengths of the CA and SAA modules (CA or SAA biosynthesis per cell) were not necessarily equal. As such, the two harboring strains where be inoculated at uneven ratios to coordinate the two compounds' supply for the RA bioproduction optimization. In the meantime, the RA module's bioconversion capability also needed to be matched with the provision of the CA and SAA precursors. This was reflected in the optimized inoculation ratio of 2:3:1, which best satisfied the pathway modules' different needs for pathway balancing. Moreover, growing co-culture strains on different carbon substrates further helped coordinate their growth profiles for improving the population stability and the biosynthesis performance. It should be noted that the number of possible inoculation ratios increases dramatically as more strains are recruited to constitute the co-cultures. For the three-strain co-cultures of this study, we only investigated the production profiles using 9 inoculation ratios, due to the tremendous workload for testing all possible ratios. It is therefore likely that even higher RA biosynthesis may be achieved using other inoculation conditions.

In conclusion, 172 mg/L RA was produced using a rationally designed three-strain co-culture, which is 38-fold higher than the original mono-culture strain developed in this study. The RA concentration is the highest among all reported studies for de novo RA biosynthesis. Hence, the accomplishment of this study marks an important progress towards unleashing the power of modular co-culture engineering for advancing microbial biosynthesis of complex natural products, especially for those involving non-linear pathways.

Chapter 7. Conclusion and future work

7.1 Summary of work

The research work in this thesis was centered around engineering microbial biosynthesis system for producing L-tyrosine and L-tyrosine derivatives compounds. These compounds are mainly derived from the shikimate pathway using metabolite PEP and E4P as the precursors. In pursuit of the thesis research, a L-tyrosine bioproduction platform was constructed and the other metabolic pathways enzymes are integrated into the bioproduction system for bioproduction of L-tyrosine derivatives. During this process, several new metabolic engineering methodologies were exploited to enhancing the bioproduction ability. These include modular co-culture engineering, biosensor assisted high performing cell selection system and biosensor mediated growth regulation system.

In Chapter 2, the biosensor-assisted cell selection systems for L-tyrosine over-production was studied. Wild type *E. coli* is unable to overproduce tyrosine due to low growth demand of the compound. Therefore, the L-tyrosine pathway genes were overexpressed on plasmids and resulted in 136 mg/L L-tyrosine production. For further improvement of L-tyrosine and anthranilic acid production, the biosensor-assisted selection system was used. The biosensor can effectively enhance the biosynthetic ability by promoting the growth of high performing cells and inhibiting the growth of low performing cells. As a result, the overall production ability can be tremendously improved.

Next, modular co-culture engineering was discussed for tyrosine derivatives biosynthesis in Chapter 3 and Chapter 6. Co-culture system showed stronger

biosynthetic ability than monoculture in all tested systems. For phenol bioproduction, it is a simple one step conversion from L-tyrosine for linear form. The final titer was improved by 5.3 folds compared to the monoculture control. For rosmarinic acid, it involved a more complicated divergent-convergent pathway. The RA production in monoculture was as low as 4.5 mg/L due to high metabolic burden and unbalanced biosynthetic ability of both precursors. The modular co-culture engineering provided a good platform for RA production and the production was 38 folds higher than the monoculture control.

To further investigate the potential of co-culture engineering, the biosensor was introduced to automatically coordinate the intermediate accumulation and consumption. The production of 4-hydroxystrene and caffeic was improved for 2.7 folds and 2.5 folds in comparison with co-culture systems without biosensor.

7.2 Advantages and disadvantages of the methodologies

The methodologies discussed in this thesis were able to largely increase the biosynthetic ability. However, there are also weaknesses for these methods to use in large scale production processes. The pros and cons of the methods are discussed below.

Firstly, for biosensor- assisted selection system, some challenges of the original design were evaluated in Chapter 2. Although the toxin system can be applied to enable the use of the off-switch biosensor, there are more on-switch biosensors in nature. For on-switch biosensor, use of antibiotic resistance appears to be a more convenient choice for the selection system. However, large scale

production using this strategy will suffer from the issues of high antibiotic costs and uneven selection pressure during the cultivation, as mentioned in Chapter 2. Also, the biosensor saturated issue occurred in L-tyrosine production systems and present a challenge for biosynthesis improvement. On the other hand, in situ extraction using solvents or resins can be used to address the biosensor saturation problem but may also require extra optimization effort for biocompatibility and extraction efficiency. Moreover, L-tyrosine production only reached gram per liter level. Since industrial production of the compound often reaches 50-100 g/L, the application of the biosensor system still needs to be studied for further improvement.

Secondly, the advantages for modular co-culture engineering, are several folds. Co-culture can reduce metabolic burden for each module and enhance the biosynthetic performance for the system. Also, the use of varied host strains for different co-culture modules provides diversified cellular environment for functional expression of genes, particularly, the heterologous genes. As a result, some genes with promiscuous activity can be designed to be allocated into a separated module in case of the interference of the undesired enzymatic reaction. Moreover, the co-culture engineering offers the flexibility to coordinate the biosynthetic ability for different module to achieve best metabolic flux distribution. In RA production system, two carbon substrate xylose and glucose were used for different co-culture modules. This strategy can reduce the competition between different modules and coordinate their biosynthesis contributions. Lastly, the co-culture engineering provides the plug-and play biosynthesis. For example, the L-tyrosine module can be used to produce phenol, 4-hydroxystyrene and caffeic acid by simply swapping the downstream pathway
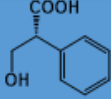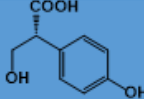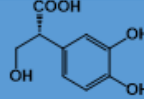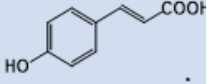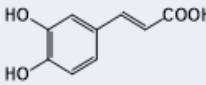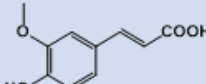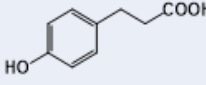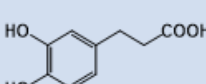
modules. The biosynthetic pathway can be optimized by changing inoculation ratios. Those 6 advantages discussed above indicates the strong potential of co-culture engineering. However, there are also some issues for co-culture engineering. Although most of the cases shown in this thesis suggested a positive result for co-culture engineering, some unreported results indicated a lower production compared to monoculture. One of the most important issue of co-culture engineering is to define a criterion regarding which biosynthetic system is suitable for co-culture engineering. Also, co-culture engineering requires the stability of all modules. For monoculture, the cell growth and biosynthesis can be maintained to a certain level with relatively high reproducibility. However, the stability of the co-culture is more challenging to control. Once one module of the co-culture systems failed to perform as expected, the biosynthetic strength balanced before can be interrupted and lead to reduced production. More importantly, due to different carbon source assimilation ability of different modules, the optimized inoculation ratio in small scale production system cannot be directly applied to large scale production and further optimization is required. Also, the use of co-culture system can incorporate some empty plasmid for maintaining the upstream and downstream cells with the same antibiotic resistance. This may somehow hurt the biosynthetic ability as shown in the thesis.

7.3 Future work and recommendations

The projects on bioproduction of tyrosine and tyrosine derivatives discussed in this thesis are successful in in terms of production improvement. However, future research works are also necessary to advance the research in these areas.

For tyrosine producing project, an amino acid exporter was used to reduce the intracellular concentration of L-tyrosine. Although the production was improved to a higher level, the exportation process needs to be characterized to gain better understanding of the biosynthesis improvement. The intracellular concentration of L-tyrosine can be studied by taking time samples to monitor the intracellular concentration change of L-tyrosine. For phenol project, phenol production can further improved using more advanced larger scale bioreactor techniques such as perfusion bioreactor. For 4-hydroxystyrene and caffeic acid production using biosensor-based growth regulation in the context of co-cultures, real time change of cell growth in response to tyrosine accumulation throughout the bioproduction process can be studied to confirm that the mechanism of the biosensor-based system indeed work as designed for production improvement.

Table 7.1 RA derivatives feeding experiment conversion rate.

	PL	HPL	DHPL
p-coumaric	92.72%	29.39%	17.46%
caffeic acid	93.9%	28.9%	9.7%
ferulic acid	30.8%	11.58%	8.17%
HPPA	96.25%	22.75%	14.91%
DHPPA	44.5%	8.2%	10.24%

The current findings in RA project demonstrated that modular co-culture engineering is a powerful tool to handle different biosynthesis systems and complicated pathways. Notably, the established RA producing platform can also be used for other HCEs' production. Zhuang et al. reported the conversion of a series of HCEs by feeding different precursors using a strain similar to the third module strain of this project [80]. This indicates that the de novo biosynthesis platform can be easily employed for producing other HCE compounds (Table 7.1) by swapping the desired genes in the co-culture system.

Overall, this thesis explores the potential of different metabolic engineering strategies for microbial biosynthesis of L-tyrosine and its derivatives in the

context of both microbial monoculture and co-culture. The results confirm that microbial biosynthesis system can be rationally designed, constructed, and optimized to suit the need of production of various products. The findings of this work provide new knowledge and experiences that pave the way for the future studies in these areas.

References

1. Zhang, H. and X. Wang, *Modular co-culture engineering, a new approach for metabolic engineering.* Metabolic engineering, 2016. **37**: p. 114-121.

2. Zhou, K., et al., *Distributing a metabolic pathway among a microbial consortium enhances production of natural products.* Nature biotechnology, 2015. **33**(4): p. 377.

3. Xiao, Y., et al., *Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis.* Nature chemical biology, 2016. **12**(5): p. 339.

4. Wang, X., et al., *Biosensor-assisted high performing cell selection using an E. coli toxin/antitoxin system.* Biochemical Engineering Journal, 2019.

5. Kim, B., et al., *Metabolic engineering of Escherichia coli for the production of phenol from glucose.* Biotechnology journal, 2014. **9**(5): p. 621-629.

6. Luo, Z.W., J.S. Cho, and S.Y. Lee, *Microbial production of methyl anthranilate, a grape flavor compound.* Proceedings of the National Academy of Sciences, 2019. **116**(22): p. 10749-10756.

7. Lütke-Eversloh, T., C.N.S. Santos, and G. Stephanopoulos, *Perspectives of biotechnological production of L-tyrosine and its applications.* Applied microbiology and biotechnology, 2007. **77**(4): p. 751-762.

8. Choi, S.-L., et al., *High throughput screening and directed evolution of tyrosine phenol-lyase.* Microbiology and Biotechnology Letters, 2006. **34**(1): p. 58-62.

9. Zhang, H., et al., *Engineering E. coli–E. coli cocultures for production of muconic acid from glycerol.* Microbial cell factories, 2015. **14**(1): p. 134.

10. Jung, M.E. and J.C. Rohloff, *Organic chemistry of L-tyrosine. 1. General synthesis of chiral piperazines from amino acids.* The Journal of Organic Chemistry, 1985. **50**(24): p. 4909-4913.

11. Santos, C.N.S. and G. Stephanopoulos, *Melanin-based high-throughput screen for L-tyrosine production in Escherichia coli.* Appl. Environ. Microbiol., 2008. **74**(4): p. 1190-1197.

12. Santos, C.N.S., *Combinatorial search strategies for the metabolic engineering of microorganisms.* 2010, Massachusetts Institute of Technology.

13. Na, D., et al., *Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs.* Nature biotechnology, 2013. **31**(2): p. 170.

14. Juminaga, D., et al., *Modular engineering of L-tyrosine production in Escherichia coli.* Appl. Environ. Microbiol., 2012. **78**(1): p. 89-98.

15. Kikuchi, Y., K. Tsujimoto, and O. Kurahashi, *Mutational analysis of the feedback sites of phenylalanine-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase of Escherichia coli.* Appl. Environ. Microbiol., 1997. **63**(2): p. 761-762.

16. Pittard, J., H. Camakaris, and J. Yang, *The TyrR regulon.* Molecular microbiology, 2005. **55**(1): p. 16-26.

17. Bokinsky, G., et al., *HipA-triggered growth arrest and β-lactam tolerance in Escherichia coli are mediated by RelA-dependent ppGpp synthesis.* Journal of bacteriology, 2013. **195**(14): p. 3173-3182.

18. Widhalm, J.R., et al., *Identification of a plastidial phenylalanine exporter that influences flux distribution through the phenylalanine biosynthetic network.* Nature communications, 2015. **6**(1): p. 1-11.

19. Davis, J.H., A.J. Rubin, and R.T. Sauer, *Design, construction and characterization of a set of insulated bacterial promoters.* Nucleic acids research, 2010. **39**(3): p. 1131-1141.

20. Zhang, H. and G. Stephanopoulos, *Co-culture engineering for microbial biosynthesis of 3-amino-benzoic acid in Escherichia coli.* Biotechnol J, 2016. **11**(7): p. 981-7.

21. Zhang, H., et al., *Engineering Escherichia coli coculture systems for the production of biochemical products.* Proceedings of the National Academy of Sciences, 2015: p. 201506781.

22. Pittard, A. and B. Davidson, *TyrR protein of Escherichia coli and its role as repressor and activator.* Molecular microbiology, 1991. **5**(7): p. 1585-1592.

23. Zheng, J., et al., *Structure and function of the macrolide biosensor protein, MphR (A), with and without erythromycin.* Journal of molecular biology, 2009. **387**(5): p. 1250-1260.

24. Meng, H.L., et al., *Construction of polyketide overproducing Escherichia coli strains via synthetic antisense RNAs based on in silico fluxome analysis and comparative transcriptome analysis.* Biotechnology journal, 2016. **11**(4): p. 530-541.

25. Choi, S.-L., et al., *Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits.* ACS synthetic biology, 2014. **3**(3): p. 163-171.

26. Shingler, V., M. Bartilson, and T. Moore, *Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators.* Journal of Bacteriology, 1993. **175**(6): p. 1596-1604.

27. van Sint Fiet, S., J.B. van Beilen, and B. Witholt, *Selection of biocatalysts*

for chemical synthesis. Proceedings of the National Academy of Sciences, 2006. **103**(6): p. 1693-1698.

28.   Cebolla, A., C. Sousa, and V. de Lorenzo, *Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading Pseudomonas define protein sites involved in binding of aromatic inducers.* Journal of Biological Chemistry, 1997. **272**(7): p. 3986-3992.

29.   Serina, L., et al., *Escherichia coli UMP kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP.* Biochemistry, 1995. **34**(15): p. 5066-5074.

30.   Joshi, V. and S.J. Wakil, *Studies on the mechanism of fatty acid synthesis: XXVI. Purification and properties of malonyl-coenzyme A—Acyl carrier protein transacylase of Escherichia coli.* Archives of biochemistry and biophysics, 1971. **143**(2): p. 493-505.

31.   Chenchik, A., et al., *Contacts of Escherichia coli RNA polymerase subunits with nucleotides of lacUV5 promoter.* Molekuliarnaia biologiia, 1982. **16**(1): p. 35-46.

32.   Simpson, R.B., *The molecular topography of RNA polymerase-promoter interaction.* Cell, 1979. **18**(2): p. 277-285.

33.   Maki, H. and A. Kornberg, *The polymerase subunit of DNA polymerase III of Escherichia coli. II. Purification of the alpha subunit, devoid of nuclease activities.* Journal of Biological Chemistry, 1985. **260**(24): p. 12987-12992.

34.   Yang, Y., et al., *Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products.* Metabolic engineering, 2015. **29**: p. 217-226.

35.   Baick, J.-W., et al., *Growth Inhibition of Escherichia coli during Heterologous Expression ofBacillus subtilis Glutamyl-tRNA Synthetase that Catalyzes the Formation of Mischarged Glutamyl-tRNA 1 Gln.* The Journal of Microbiology, 2004. **42**(2): p. 111-116.

36.   Ehlert, K., J.V. Höitje, and M.F. Templin, *Cloning and expression of a murein hydrolase lipoprotein from Escherichia coli.* Molecular microbiology, 1995. **16**(4): p. 761-768.

37.   Chadwick, S.S., *Ullmann's encyclopedia of industrial chemistry.* Reference Services Review, 1988.

38.   Schmidt, R.J., *Industrial catalytic processes—phenol production.* Applied Catalysis A: General, 2005. **280**(1): p. 89-103.

39.   Elliott, J.H., *Industrial organic chemicals in perspective. Part one: Raw materials and manufacture, Harold A. Wittcoff and Bryan G. Reuben, Wiley-Interscience, New York, 1980, 298 pp.* Journal of Polymer Science: Polymer Letters Edition, 1980. **18**: p. 751-751.

40. Stadelhofer, J., *Industrial Aromatic Chemistry*. 1988, Springer-Verlag, Berlin.

41. Lee, B., et al., *Alternating-Current Electrolysis for the Production of Phenol from Benzene.* Angewandte Chemie International Edition, 2012. **51**(28): p. 6961-6965.

42. Parmon, V., et al., *Nitrous oxide in oxidation chemistry and catalysis: application and production.* Catalysis Today, 2005. **100**(1-2): p. 115-131.

43. Thompson, B., M. Machas, and D.R. Nielsen, *Engineering and comparison of non-natural pathways for microbial phenol production.* Biotechnology and bioengineering, 2016. **113**(8): p. 1745-1754.

44. Noda, S., et al., *Metabolic design of a platform Escherichia coli strain producing various chorismate derivatives.* Metabolic engineering, 2016. **33**: p. 119-129.

45. Miao, L., et al., *Construction of a novel phenol synthetic pathway in Escherichia coli through 4-hydroxybenzoate decarboxylation.* Applied microbiology and biotechnology, 2015. **99**(12): p. 5163-5173.

46. Ren, Y., et al., *Microbial production of phenol via salicylate decarboxylation.* Rsc Advances, 2015. **5**(112): p. 92685-92689.

47. Chen, T., et al., *Advances in heterologous biosynthesis of plant and fungal natural products by modular co-culture engineering.* Biotechnology letters, 2018: p. 1-8.

48. Jones, J.A. and X. Wang, *Use of bacterial co-cultures for the efficient production of chemicals.* Current opinion in biotechnology, 2018. **53**: p. 33-38.

49. Guo, X., et al., *De novo phenol bioproduction from glucose using biosensor-assisted microbial coculture engineering.* Biotechnology and bioengineering, 2019. **116**(12): p. 3349-3359.

50. Kaiser, M.J., *A review of refinery complexity applications.* Petroleum Science, 2017. **14**(1): p. 167-194.

51. Dorothea, G., E. Barbara, and H. Stephen, *Phenol Derivatives, Ullmann's Encyclopedia of Industrial Chemistry*. 1991, Barbara, E.

52. Khanna, S., A. Goyal, and V.S. Moholkar, *Microbial conversion of glycerol: present status and future prospects.* Critical reviews in biotechnology, 2012. **32**(3): p. 235-262.

53. Carmona, M., et al., *Combined adsorption and ion exchange equilibrium of phenol on Amberlite IRA-420.* Chemical Engineering Journal, 2006. **117**(2): p. 155-160.

54. Ma, Q., et al., *Alkylation of phenol: a mechanistic view.* The Journal of

Physical Chemistry A, 2006. **110**(6): p. 2246-2252.

55. Krishnan, V., K. Ojha, and N.C. Pradhan, *Alkylation of phenol with tertiary butyl alcohol over zeolites.* Organic process research & development, 2002. **6**(2): p. 132-137.

56. Lachter, E.R., et al., *Use of Ion-Exchange Resins in Alkylation Reactions*, in *Applications of Ion Exchange Materials in Chemical and Food Industries*. 2019, Springer. p. 35-74.

57. Santos, C.N.S., W. Xiao, and G. Stephanopoulos, *Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in Escherichia coli.* Proceedings of the National Academy of Sciences, 2012. **109**(34): p. 13538-13543.

58. Kunin, R., et al., *Characterization of amberlyst 15. macroreticular sulfonic acid cation exchange resin.* Industrial & Engineering Chemistry Product Research and Development, 1962. **1**(2): p. 140-144.

59. Pal, R., T. Sarkar, and S. Khasnobis, *Amberlyst-15 in organic synthesis.* ARKIVOC: Online Journal of Organic Chemistry, 2012.

60. Lin, Q., *Properties of photoresist polymers*, in *Physical Properties of Polymers Handbook*. 2007, Springer. p. 965-979.

61. Zhang, H. and G. Stephanopoulos, *Engineering E. coli for caffeic acid biosynthesis from renewable sugars.* Applied microbiology and biotechnology, 2013. **97**(8): p. 3333-3341.

62. Santos, C.N.S., M. Koffas, and G. Stephanopoulos, *Optimization of a heterologous pathway for the production of flavonoids from glucose.* Metabolic engineering, 2011. **13**(4): p. 392-400.

63. McKenna, R. and D.R. Nielsen, *Styrene biosynthesis from glucose by engineered E. coli.* Metabolic engineering, 2011. **13**(5): p. 544-554.

64. Liu, C., et al., *A systematic optimization of styrene biosynthesis in Escherichia coli BL21 (DE3).* Biotechnology for biofuels, 2018. **11**(1): p. 14.

65. Chemerovski-Glikman, M., et al., *Rosmarinic acid restores complete transparency of sonicated human cataract ex vivo and delays cataract formation in vivo.* Scientific reports, 2018. **8**(1): p. 1-11.

66. Wang, J., et al., *Neurorescue effect of rosmarinic acid on 6-hydroxydopamine-lesioned nigral dopamine neurons in rat model of Parkinson's disease.* Journal of molecular Neuroscience, 2012. **47**(1): p. 113-119.

67. Scarpati, M.L., *Oriente G (1958) Isolamento e costituzione dell'acido rosmarinico (dal rosmarinus off.).* Ric Sci, 1958. **28**: p. 2329-2333.

68. Pedersen, J.A., *Distribution and taxonomic implications of some phenolics in the family Lamiaceae determined by ESR spectroscopy.* Biochemical systematics and Ecology, 2000. **28**(3): p. 229-253.

69. Eicher, T., M. Ott, and A. Speicher, *Bryophyte constituents; 7: new synthesis of (+)-rosmarinic acid and related compounds.* Synthesis, 1996. **1996**(06): p. 755-762.

70. Bogucki, D.E. and J.L. Charlton, *A non-enzymatic synthesis of (S)-(−)-rosmarinic acid and a study of a biomimetic route to (+)-rabdosiin.* Canadian journal of chemistry, 1997. **75**(12): p. 1783-1794.

71. 袁虎, et al., *(±)-迷迭香酸 (Rosmarinic acid) 的全合成.* 化学学报, 2011. **69**(8): p. 945-948.

72. Huang, Q., Y. Lin, and Y. Yan, *Caffeic acid production enhancement by engineering a phenylalanine over-producing Escherichia coli strain.* Biotechnology and bioengineering, 2013. **110**(12): p. 3188-3196.

73. Lin, Y. and Y. Yan, *Biosynthesis of caffeic acid in Escherichia coli using its endogenous hydroxylase complex.* Microbial cell factories, 2012. **11**(1): p. 42.

74. Wang, J., et al., *Engineering a bacterial platform for total biosynthesis of caffeic acid derived phenethyl esters and amides.* Metabolic engineering, 2017. **44**: p. 89-99.

75. Wang, J., et al., *Exploring the promiscuity of phenol hydroxylase from Pseudomonas stutzeri OX1 for the biosynthesis of phenolic compounds.* ACS synthetic biology, 2018. **7**(5): p. 1238-1243.

76. Yao, Y.-F., et al., *Metabolic engineering of Escherichia coli for production of salvianic acid A via an artificial biosynthetic pathway.* Metabolic engineering, 2013. **19**: p. 79-87.

77. Zhu, Y., et al., *Enhancement of phenyllactic acid biosynthesis by recognition site replacement of D-lactate dehydrogenase from Lactobacillus pentosus.* Biotechnology letters, 2015. **37**(6): p. 1233-1241.

78. Bloch, S.E. and C. Schmidt-Dannert, *Construction of a chimeric biosynthetic pathway for the de novo biosynthesis of rosmarinic acid in Escherichia coli.* ChemBioChem, 2014. **15**(16): p. 2393-2401.

79. Zhang, H. and G. Stephanopoulos, *Co-culture engineering for microbial biosynthesis of 3-amino-benzoic acid in Escherichia coli.* Biotechnology journal, 2016. **11**(7): p. 981-987.

80. Zhuang, Y., et al., *Synthesis of rosmarinic acid analogues in Escherichia coli.* Biotechnology letters, 2016. **38**(4): p. 619-627.

Appendix

Primers used in the thesis

Primers	DNA sequence	Description
ZLPR1 AG	ATGTCGACACTAGTATGGTTGCTGAA TTGACCGCATTACG	Cloning *tyrA^fbr* and *aroG^fbr* genes
ZLPR2 AG	CGAAGCTTTTACCCGCGACGCGCTTT TACT	Cloning *tyrA^fbr* and *aroG^fbr* genes
ZLPR1 HP	GCCCATATGAAACCAGAAGATTTCC GCG	Cloning *hipA* gene
ZLPR2 HP	GACTCGAGACTAGTTTAAATCGCAG CTTCCAT	Cloning *hipA* gene
ZLPR1 CL	GCACTAACATATGGGTGACTGCGTTG CCCC	Cloning *4CL* gene
ZLPR2 CL	GCACTCGAGATACTAGTTTACTTCGG CAGGTCGCCG	Cloning *4CL* gene
ZLPR1TA	ACCATATGAACAAATAGGGGTTCCG C	Cloning *tetA* gene with *Ptet* promoter
ZLPR2TA	TGCTCGAGTTCCATTCAGGTCGAGGT	Cloning *tetA* gene with *Ptet* promoter
ZLPR1PT	CTTATTACGCGCCTGACT	Gene knockout for tyrB
ZLPR2PT	AGTCACAGGCAATAAGGC	Gene knockout for tyrB
ZLPR1 AE	TAATACGACTCACTATAGGG	Cloning *aroE* gene
ZLPR2 AE	GCCTCGAGCGACTAGTTCACGCGGA CAATTCCTC	Cloning *aroE* gene
ZLPR1 AP	ATGCATGCATGAGTTCCTGTCTTAAG CC	Cloning *Parop* promoter

ZLPR2 AP	ATTCTAGAGATATCTGCGGCGCAGG	Cloning *Parop* promoter
ZLPR1TL	CGAAGCTTTTACCCGCGACGCGCTTT TACT	Cloning *tpl* gene
ZLPR2TL	GCCTCGAGTCACGCTTTCGGTTCGAA GC	Cloning *tpl* gene
ZLPR1 KC	ATGCGGCCGCCCCTCGAGTCTGGTAA AG	Cm^R from Kan^R
ZLPR2 KC	ATCCTGATTCAGGAGGGACAGCTGA TAGAAA	Cm^R from Kan^R
ZLPR17 M	GGTACCTCTAGAAATAATTTTGTTT	Cloning *TAL* gene under *Pmtr*
ZLPR27 M	atCTCGAGTTATGCCAGCATCTTCAG	Cloning *TAL* gene under *Pmtr*
ZLPR1 BS	GCGCATGCTCGCATTCTCAACAAGCC	*Pmtr* from T7
ZLPR2 BS	CGTCTAGATGCATTGCACTGTACCAG	*Pmtr* from T7
ZLPR1 MF	ATCCATGGCCGGCTTCCACTTTTTCC C	T7 to *Pmtr* for *FDC1*
ZLPR2 MF	TTCATATGTATATCTCCTTCTTAAAG	T7 to *Pmtr* for *FDC1*
ZLPR1T D	ACGTCGACATGGCTGACATTCTGCTG C	Cloning *trpGD* gene
ZLPR2T D	ATGCGGCCGCTTACAGAATCGGTTGC AGCGTGTTG	Cloning *trpGD* gene

Heterologous genes sequences

Tpl gene

ATGCGTAACTACCCAGCTGAGCCTTACAAGGTCAAGGCA
GTGGAGCCTATCGCTATGACTACTCGTGAGCAACGTGAAG
CATACATGAAGAAGGCTGGTTACAACACCTTCCTGCTGAA
CTCCGAGGAGGTGTACATCGATCTGCTGACTGACTCCGGT
ACTAGCGCAATGTCTGACAAACAGTGGGCAGGTCTGATG
ATCGGTGACGAGGCTTACGCAGGTTCTCGTAACTTCATGC
ACCTGCAAGACGTCGTCCGTGAATACTACGGTTTCAAATA
CGTCGTGCCGACCCACCAAGGTCGTGGTGCAGAAAACCT
GCTGTCTACTATCATGATCAAGCCAGGCGATTACGTGCCA
GGTAACATGTACTTCACCACCACGCGTGCACACCAGGAA
CGTAACGGTGCAACTTTCGTTGATATCATTATCGACGAGG
CACACGATTCCCAGATCGACCTGCCGTTTAAAGGTAACGT
GGACGTGAAAAAACTGCAGAAACTGATCGACGAAGTGGG
CGCCGATAAAATCCCGTACATCTGTCTGGCGGTGACCGTG
AACCTGGCAGGTGGTCAGCCGGTATCTATGGCTAATATGC
GTGAAGTAAAAGCGCTGTGCTCTAAACACGGTATCAAAG
TAATGTTCGACGCCACGCGCTGCGTAGAAAACGCCTACTT
TATCAAAGAACGCGAAGCGGAATACAAAGACGCTACCAT
CAAAGACATCCTGAAAGAAATGATGAGCTACGCCGACGG
CTGCACCATGTCCGGCAAAAAAGACTGCCTGGTTAACATC
GGCGGCTTCCTGTGCATCAACGACGATGATCTGTACCAGC
AGGCTTGTGAACTGGTAGTTCTGTTCGAAGGCATGCCGAG
CTATGGCGGCCTGGCTGGTCGTGATATGGAAGCGATGGCT
ATCGGTATCACTGAAAGCGTTGACTTCCACTATATCCAGC
ACCGCGTAGCCCAGTGTTATTATCTGGCGGATAAGCTGGA
AGCGGCTGGTGTTCCGATTGTTAAACCGGTTGGTGGCCAT
GCTGTATTTCTGGATGCTAAAAAATTTCTGCCGCACATTC
CGCAGGAACAGTTCCCGGCCCAGATGCTGGCGGCGCAGA
TTTATATTGAAGGCGGCGTTCGCTCTATGGAACGTGGCAT
TGTTTCCGCGGGCCGTGATAAAAAAACGGGCGCCAATCAT
ACCCCGAAACTGGAACTGGTTCGTCTGACCATTCCGCGTC
GCGTTTATACCTATGCGCATCTGGATCATGTTGCGGATAC
CATTATTAAACTGTTCAAACACCGCGACGACATTAAAGGC
CTGGATATGGTTTATGAACCGAAGCTGCTGCGCTTCTTTA
CCGCGCGCTTCGAACCGAAAGCGTGA

Pc4CL gene

ATGGGTGACTGCGTTGCCCCGAAAGAGGATCTGATCTTCC
GCAGCAAACTGCCGGACATTTACATTCCAAAGCATCTGCC
GCTGCATACGTATTGTTTTGAGAATATCAGCAAGGTTGGC
GACAAGAGCTGTCTGATCAACGGCGCAACCGGCGAAACG
TTTACCTACAGCCAGGTCGAGCTGCTGTCCCGTAAAGTTG
CCAGCGGCCTGAACAAGCTGGGCATTCAACAAGGTGATA
CCATTATGCTGTTGCTGCCGAATTCCCCGGAGTACTTTTTC
GCTTTCCTGGGTGCGAGCTATCGCGGTGCAATCAGCACCA
TGGCGAATCCATTCTTTACCAGCGCAGAAGTGATCAAGCA
ACTGAAAGCGAGCCAAGCGAAGCTGATTATCACCCAGGC
ATGCTATGTTGACAAGGTCAAGGACTACGCAGCGGAGAA
AAACATCCAGATCATTTGTATTGACGATGCACCGCAGGAT

TGCCTGCACTTTAGCAAGCTGATGGAAGCGGATGAGAGC
GAAATGCCGGAAGTGGTCATTAACAGCGATGATGTGGTG
GCATTGCCGTACAGCTCTGGCACCACCGGCCTGCCGAAAG
GCGTTATGCTGACCCACAAGGGTCTGGTTACGAGCGTTGC
ACAACAGGTGGATGGTGATAACCCGAACCTGTATATGCA
CTCCGAGGATGTCATGATCTGCATCCTGCCACTGTTCCAT
ATCTATAGCCTGAACGCTGTTCTGTGTTGTGGTCTGCGTGC
GGGCGTCACCATTCTGATCATGCAAAAGTTCGACATTGTG
CCGTTTCTGGAGCTGATTCAGAAGTATAAGGTTACGATTG
GTCCGTTTGTCCCGCCGATCGTGCTGGCCATCGCGAAAAG
CCCGGTCGTTGACAAGTACGACTTGTCTAGCGTGCGCACC
GTCATGAGCGGTGCAGCGCCGCTGGGTAAAGAGTTGGAG
GACGCTGTCCGTGCGAATTCCCGAACGCGAAGCTGGGTC
AAGGCTATGGCATGACCGAAGCCGGTCCGGTCCTGGCGA
TGTGTCTGGCGTTCGCCAAAGAGCCGTATGAGATTAAGTC
TGGCGCATGCGGTACCGTTGTGCGTAATGCCGAGATGAAA
ATCGTTGACCCAGAAACGAATGCGTCTCTGCCGCGTAATC
AGCGTGGTGAGATTTGCATCCGTGGTGATCAGATTATGAA
AGGTTACCTGAATGACCCGGAAAGCACCCGCACCACGAT
CGACGAAGAGGGTTGGTTGCACACGGGTGACATTGGTTTC
ATCGACGATGACGATGAACTGTTCATTGTCGATCGTTTGA
AAGAAATCATTAAGTACAAAGGTTTTCAAGTTGCTCCGGC
GGAGTTGGAAGCACTGCTGCTGACGCACCCGACGATCAG
CGATGCCGCGGTGGTTCCGATGATTGACGAGAAAGCGGG
TGAAGTGCCAGTGGCGTTTGTCGTGCGTACCAATGGTTTT
ACCACGACCGAAGAAGAAATCAAACAATTTGTGAGCAAA
CAGGTCGTGTTCTACAAACGTATCTTCCGCGTCTTCTTCGT
TGACGCTATTCCGAAATCCCCGAGCGGCAAGATTTTGCGT
AAGGATCTGCGCGCTCGTATTGCGAGCGGCGACCTGCCGA
AGTAA

<table>

RgTAL gene	ATGGCGCCTCGCCCGACTTCGCAAAGCCAGGCCCGCACTT
	GCCCGACGACGCAGGTTACCCAAGTTGATATCGTTGAGAA
	AATGTTGGCGGCTCCTACTGATAGCACGCTGGAGCTGGAC
	GGTTATAGCCTGAATCTGGGTGATGTCGTGAGCGCTGCGC
	GTAAGGGTCGTCCTGTCCGTGTCAAAGATAGCGATGAAAT
	CCGCAGCAAAATCGACAAGAGCGTTGAATTCCTGCGCAG
	CCAACTGAGCATGTCGGTTTACGGTGTGACGACCGGCTTT
	GGCGGCTCCGCGGACACGCGCACGGAGGACGCAATTAGC
	CTGCAAAAGGCGTTGCTGGAACACCAGCTGTGTGGTGTGT
	TGCCGAGCAGCTTCGACAGCTTTCGCTTGGGTCGTGGTCT
	GGAGAATAGCCTGCCGTTGGAAGTCGTTCGCGGTGCAATG
	ACCATTCGTGTGAATTCGCTGACCCGTGGCCATAGCGCTG
	TTCGTCTGGTTGTTCTGGAAGCACTGACGAACTTTCTGAA
	CCACGGTATTACCCCGATTGTTCCGCTGCGCGGTACGATC
	TCCGCGAGCGGCGATCTGTCTCCACTGTCGTACATTGCAG
	CGGCGATTAGCGGTCACCCGGATAGCAAAGTTCACGTGGT
	CCATGAAGGCAAAGAGAAGATCCTGTACGCGCGCGAAGC
	GATGGCGCTGTTTAACCTGGAGCCGGTGGTTTTGGGTCCG
	AAGGAGGGCCTGGGTCTGGTGAATGGTACGGCAGTCTCC
	GCGAGCATGGCAACGCTGGCACTGCACGACGCGCATATG

</table>

TTGAGCCTGTTGAGCCAATCGCTGACCGCGATGACCGTGG
AGGCGATGGTCGGTCACGCGGGCAGCTTCCATCCATTCCT
GCACGATGTTACGCGTCCGCACCCGACGCAAATCGAGGT
CGCGGGTAACATTCGCAAACTGCTGGAGGGCTCGCGCTTC
GCGGTCCACCACGAGGAAGAGGTTAAGGTCAAGGATGAT
GAAGGCATTTTGCGTCAGGATCGTTATCCGTTGCGCACGA
GCCCGCAATGGTTGGGTCCGCTGGTGTCCGACCTGATTCA
CGCTCATGCCGTCTTGACGATCGAAGCGGGTCAAAGCACC
ACCGATAACCCACTGATCGATGTTGAGAATAAGACCAGC
CATCACGGTGGCAACTTTCAAGCGGCAGCGGTTGCCAACA
CGATGGAAAAGACCCGTCTGGGCTTGGCCCAAATCGGTA
AACTGAATTTCACCCAGCTGACGGAGATGCTGAACGCGG
GCATGAATCGTGGCTTGCCGAGCTGCCTGGCGGCTGAAGA
CCCATCCCTGAGCTATCATTGCAAAGGTCTGGACATTGCG
GCGGCTGCATATACGAGCGAACTGGGCCACCTGGCTAAC
CCGGTCACCACCCACGTCCAACCGGCTGAAATGGCAAAC
CAGGCGGTGAATAGCTTGGCGTTGATTAGCGCACGTCGTA
CCACGGAATCTAACGACGTTCTGTCCCTGCTGCTGGCAAC
GCACCTGTACTGCGTGCTGCAGGCGATCGACCTGCGTGCG
ATTGAGTTCGAGTTCAAGAAACAGTTTGGTCCTGCCATTG
TTAGCCTGATCGACCAACACTTTGGTAGCGCGATGACGGG
TAGCAATCTGCGTGATGAGCTGGTTGAAAAGGTCAATAA
GACTCTGGCCAAGCGTTTGGAGCAAACCAATAGCTACGAT
CTGGTTCCGCGCTGGCACGACGCTTTTAGCTTCGCTGCAG
GCACTGTTGTCGAGGTTCTGTCCAGCACGAGCCTGAGCTT
GGCGGCCGTGAACGCATGGAAGGTTGCGGCAGCCGAGAG
CGCGATCTCCTTGACGCGCCAGGTCCGTGAAACGTTTTGG
TCCGCTGCAAGCACCTCCAGCCCGGCGTTGTCTTACTTGA
GCCCGCGCACGCAGATCCTGTACGCATTTGTGCGTGAGGA
ACTGGGTGTCAAAGCCCGCCGTGGTGACGTCTTCTTGGGT
AAACAAGAAGTTACCATCGGCAGCAACGTTAGCAAGATT
TACGAAGCCATCAAGAGCGGCCGTATCAACAATGTTCTGC
TGAAGATGCTGGCATAA

Coum3
H gene

ATGACCATCACAAGCCCGGCACCGGCCGGTCGTTTAAAC
AACGTGCGTCCGATGACCGGCGAAGAATATCTGGAAAGT
CTGCGTGACGGTCGTGAGGTTTACATCTACGGTGAGCGTG
TGGATGATGTGACCACACATCTGGCCTTTCGCAACAGCGT
GCGTAGCATCGCCCGCCTGTACGATGTGTTACACGACCCT
GCAAGCGAAGGTGTGTTACGTGTGCCTACCGATACCGGCA
ACGGCGGTTTCACACACCCGTTCTTTAAGACAGCCCGTAG
CAGCGAGGACCTGGTTGCAGCCCGTGAAGCCATCGTGGG
CTGGCAACGCCTGGTGTACGGTTGGATGGGTCGCACCCCG
GATTATAAGGCAGCCTTCTTCGGCACACTGGATGCCAACG
CCGAGTTCTATGGTCCGTTCGAGGCAAATGCACGTCGCTG
GTACCGCGACGCACAAGAGCGCGTGCTGTACTTCAATCAT
GCCATCGTGCACCCTCCTGTTGATCGTGACCGTCCTGCAG
ATCGCACCGCAGATATCTGTGTGCACGTGGAAGAGGAGA
CAGACAGCGGCCTGATCGTTAGCGGCGCCAAAGTTGTGG
CAACCGGCAGTGCCATGACCAACGCCAACTTAATCGCCC
ATTATGGCCTGCCGGTTCGCGATAAGAAGTTCGGCCTGGT

GTTCACCGTGCCGATGAACAGCCCGGGTCTGAAGCTGATC
TGCCGCACAAGCTATGAACTGATGGTGGCCACCCAAGGT
AGCCCGTTTGACTACCCGCTGAGCAGTCGCCTGGACGAGA
ACGACAGCATCATGATCTTCGACCGCGTTCTGGTGCCTTG
GGAGAATGTGTTCATGTACGACGCAGGTGCCGCAAACAG
CTTTGCCACCGGCAGTGGTTTCCTGGAACGCTTCACCTTTC
ACGGTTGTACCCGCCTGGCAGTGAAACTGGACTTCATCGC
CGGTTGTGTGATGAAGGCCGTGGAAGTTACCGGCACCACC
CATTTCCGTGGCGTGCAGGCACAGGTGGGCGAAGTGCTG
AATTGGCGCGATGTTTTCTGGGGCCTGAGCGACGCAATGG
CAAAGAGTCCGAACAGTTGGGTTGGCGGCAGTGTTCAGC
CGAACCTGAACTACGGTCTGGCCTATCGCACCTTTATGGG
TGTGGGCTACCCGCGCATCAAGGAAATTATCCAGCAAACC
CTGGGCAGTGGCCTGATCTATCTGAATAGCAGCGCCGCCG
ACTGGAAGAACCCTGACGTTCGTCCGTATCTGGACCGCTA
CTTACGTGGCAGTCGTGGCATCCAGGCCATCGACCGCGTT
AAGCTGTTAAAGCTGCTGTGGGATGCCGTGGGTACAGAAT
TCGCAGGTCGCCACGAACTGTATGAGCGTAACTACGGTGG
CGACCATGAGGGTATCCGTGTGCAGACCCTGCAAGCCTAC
CAAGCAAACGGTCAAGCCGCCGCCCTGAAAGGCTTCGCA
GAGCAATGCATGAGCGAGTACGATTTAGACGGCTGGACC
CGCCCTGATCTGATTAATCCGGGCACCTGA

MoRA *S* gene	ATGCGTATTGATATCAAAGATAGCACCATGGTTAAACCGG
CGGCTGAAACTCCGGGCGGTTCTGTTTGGCTGACCAACCT
GGATCTGCTGAGCCCGGCGAACTACCACACCCTGTCTGTG
CACTTCTATCACCACGATGGCTCTGAAAACTTCTTTGATG
CGGCGGCGCTGAAAGAAGCGCTGAGCCGTGCTCTGGTTG
ATTTCTACCCGTACGCGGGTCGTCTGAAACTGAAAGATAA
CCGTCTGGAAATCGACTGCAACGGTGAAGGTGTTCTGCTG
GTTGAAGCGGAAAGCGATGGCGCGCTGGCGGAACTGGGT
GAATTTGCGCCGCGTCCGGATCTGAACCTGATCCCGCAGG
TTGATTATGCGAAAGGTATCTCTACCTATCCGCTGATGCT
GTTCCAGCTGACCCGCTTCAAATGCGGCGGCGTTGGTCTG
GGTGTTGCTAACGAACACCACCTGTCTGATGGCGTTGCAG
CGCTGCACTTCATTAACACCTGGGCGCACCTGGCGCGTGG
CGTTCCGGCGCCGTCTCCGCCGCCGGTTTTCGATCGTCGC
AGCCTGTCCGCGCGTAACCCGCCGAAACCGCAGTTTTCTC
ACGCGGAATATCAGCCGCCGCCGACTCTGCCGACCCCGCT
GACCGATACCGCGATCGCTTATTCCAAACTGAAAGTGACC
CGCGATCAGCTGGGCGCGCTGAAAGCGAAATGCCTGGCA
GGCGACCCGTCTGGCAAACCGCGTAGCACCTTCGAAGTTC
TGGCGGGCCACATTTGGCGTTGTGTTTGCGCGGCGCGTGG
CCTGCCGGAAGATCAGGAAACCAAACTGCACATCCCGTT
CGATGGTCGTGCGAAACTGCGTCTGCCGCCGGGCTACTTC
GGTAACGCGATCTTCTTCGCGACCCCGGTTGCGACCTGCG
GCGAAATCGAATCTAACAGCCTGGCGCACGCGGTGAAAC
GTGTTGGTGACGCGATCGCGCGTCTGGACGAAGATTACCT
GCGTAGCTCCATCGATTTCCTGGAACTGCAGGAAGATATC
TCCAAACTGGCGCAGGGCGCGCACTCTTTCCGTTGCCCGA
ACCTGTGGGTTATCTCTTGGGTTAGACTGCCGGTTTACGA

ACCGGATTTTGGCTGGGGCAAAGCCGTTTATATGGGCCCA
TGGGCTGCACCTTTTGAAGGTAAATCCTACCTGCTGCCGA
ACCCGGATAATGATGGCTCTCTTTTTGTAGCGATCACGCT
CCACACCCAGCACATGGAACGTTTTGAAAAGCTGTTCTAT
GAAATCTAA

LpD-ldh gene	ATGAAAATCATCGCGTACGCAGTTCGCGATGATGAACGTC CGTTCTTCGACACCTGGATGAAAGAAAACCCAGACGTTGA GGTTAAACTGGTGCCGGAACTGTTGACAGAGGACAACGT TGACCTGGCTAAAGGTTTTGACGGTGCCGACGTGGCCCAG CAGAAAGATTATACCGCGGAAGTCCTGAACAAACTGGCC GACGAAGGTGTCAAAAACATCTCGCTGCGTAACGTAGGT GTCGATAACCTGGATGTTCCGACCGTTAAAGCGCGTGGCC TGAACATCAGCAACGTTCCGGCGTACAGCCCGAACGCGA TCGCGGAACTGTCTGTTACCCAGCTGATGCAGCTGCTGCG TCAGACCCCGATGTTCAACAAAAAACTGGCGAAACAGGA TTTCCGTTGGGCGCCGGATATCGCGAAAGAACTGAACACC ATGACCGTTGGTGTTATCGGCACCGGCCGTATCGGCCGTG CGGCGATCGATATCTTCAAAGGTTTCGGCGCGAAAGTTAT CGGCTACGATGTTTACCGTAACGCGGAACTGGAAAAAGA AGGCATGTACGTTGATACCCTGGATGAACTGTACGCGCAG GCGGATGTTATCACCCTGCACGTTCCGGCGCTGAAAGATA ACTACCACATGCTGAACGCGGATGCGTTCAGCAAAATGA AGATGGTGCGTACATCCTGAACTTCGCGCGTGGTACCCT GATCGATTCTGAAGATCTGATCAAAGCACTGGATAGCGGC AAAGTTGCGGGCGCGGCGCTGGTTACCTACGAATACGAA ACTAAAATCTTCAACAAAGATCTGGAAGGTCAGACCATC GATGATAAAGTTTTCATGAACCTGTTCAACCGTGATAACG TTCTGATCACCCCGCACACCGCGTTTTACACCGAAACCGC GGTTCACAACATGGTTCACGTGAGCATGAACTCTAACAAA CAGTTCATCGAAACCGGTAAAGCCGATACCCAGGTTAAAT TCGATTAA
FDC1 gene	ATGCGTAAACTGAACCCGGCTCTGGAATTTCGTGACTTCA TCCAGGTTCTGAAAGATGAAGATGATCTGATTGAAATCAC CGAAGAAATCGACCCGAACCTGGAAGTTGGTGCTATCAT GCGTAAAGCGTATGAATCTCACCTGCCGGCACCGCTGTTC AAAAACCTGAAAGGCGCGTCTAAAGACCTGTTCTCCATCC TGGGTTGCCCGGCGGGTCTGCGTTCTAAAGAAAAAGGTG ACCACGGTCGTATCGCTCACCACCTGGGTCTTGATCCGAA AACTACCATCAAAGAAATTATCGATTATCTGCTGGAATGC AAAGAAAAAGAACCGCTGCCGCCGATCACCGTTCCGGTT AGCTCCGCGCCGTGCAAAACCCATATCCTGAGCGAAGAA AAAATCCACCTGCAGTCTCTGCCGACTCCGTACCTGCACG TGTCTGACGGTGGTAAATACCTGCAGACCTACGGCATGTG GATTCTGCAAACCCCGGACAAAAAATGGACCAACTGGTC TATCGCTCGTGGCATGGTTGTTGATGATAAACACATTACT GGTCTGGTAATCAAACCGCAGCACATCCGTCAGATCGCAG ATTCTTGGGCTGCGATCGGTAAAGCAAACGAAATCCCGTT CGCTCTGTGCTTCGGTGTTCCGCCGGCAGCAATCCTGGTT AGCTCTATGCCGATTCCAGAAGGTGTGAGCGAATCCGATT

ACGTTGGTGCTATCCTGGGTGAATCTGTGCCGGTGGTTAA
ATGTGAAACTAACGACCTGATGGTGCCGGCGACCTCTGAA
ATGGTGTTCGAAGGTACTCTGTCTCTGACCGACACTCACC
TGGAAGGCCCGTTCGGTGAAATGCACGGTTACGTGTTCAA
ATCCCAGGGTCACCCGTGCCCGCTGTACACCGTTAAAGCT
ATGTCCTACCGTGATAACGCCATCCTGCCAGTTTCCAACC
CAGGTCTGTGTACCGATGAAACTCACACCCTGATCGGCTC
TCTGGTAGCAACTGAAGCAAAAGAACTGGCTATCGAATCT
GGCCTGCCGATCCTGGATGCTTTCATGCCGTATGAAGCAC
AGGCACTGTGGCTGATCCTGAAAGTAGATCTGAAAGGTCT
GCAGGCTCTGAAAACTACCCCGGAAGAGTTCTGCAAAAA
AGTTGGTGACATCTACTTCCGTACCAAAGTTGGTTTCATC
GTTCATGAAATCATTCTGGTAGCGGACGACATCGATATCT
TCAACTTCAAAGAAGTGATTTGGGCTTACGTTACCCGTCA
CACCCCGGTTGCTGACCAGATGGCTTTCGACGATGTTACT
TCTTTCCCGCTGGCTCCGTTCGTTAGCCAGAGCTCTCGTTC
TAAAACCATGAAAGGCGGTAAATGTGTTACCAACTGCATC
TTCCGTCAGCAGTACGAACGTTCTTTCGACTACATCACCT
GTAACTTCGAAAAAGGTTACCCGAAAGGTCTGGTTGACA
AAGTTAACGAAAACTGGAAACGCTACGGCTACAAATAA

PhpCA T gene

ATGTCTCTGCGTTGGTACGATCTGGTTGGTTTCGGTGTTGG
TGGTATGGTGGGTGCGGGTGTTTTCGTGACTTCTGGTCGT
GCGTCTAGCCACTGCGCGGGCCCGGCGGTAGTTCTGAGCT
ACGCGATCGCGGGTTTCTGTGCTCTGCTGTCTGCTTTCTGC
TACACTGAATTCGCTGTTGACATGCCGGTTGCAGGTGGTG
CATTCAGCTACATCCGTATCACTTTTGGTGAATTCCTGGCG
TTTCTGACCGGTGCAAACCTGATCATCGACTATGTTCTTTC
TAACGCTGCGGTTGCGCGTTCTTTCACCGGCTACCTGTGC
ACCGCGCTGGGCATCGAATCCAAACTGCGTATCACCGTGA
ACGGTCTGCCGGATGGCTTCAACGAAATCGATGTTGTTGC
TGTGCTGGTAGTTCTGGCGCTGACTGTTATCATCTGCTACT
CTACCCGTGAATCTAGCGTTCTGAACATGGTGCTGACCGT
GCTGCACATCGTTTTCATCGTTTTCGTTATCGTTATCGGCT
TCACCCGCGGTGACACCAAAAACTTCACCAAAGCTGGTG
ATTCTAACCACGCTTCCGGTTTCTTCCCGTTCGGCGCATCC
GGTGTCTTCAACGGTGCGGCGATGGTTTACCTGAGCTACA
TCGGTTACGACGCCGTTAGCACCATGGCCGAAGAAGTTAA
AAACCCGGTGAAAGACATCCCGGTTGGTGTTTCTGGTTCT
GTGATTCTGGTTACCGTTCTGTACTGTCTGATGGCAGCGTC
CATGAGCATGCTCCTGCCGTATGATATGATCGATCCGGAT
GCTCCGTTCTCTGGTGCGTTCATGGGTTCCGATGGCTGGC
GCTGGGTTTCTAACGTTATCGGCGTTGGTGCAGGCTTCGG
TATTCTGACCTCTCTGCTGGTTGCAATGCTGGGCCAGGCG
CGTTACATGTGCGTAATCGGTCGTTCTTCTGTTGTGCCGGC
GTGGTTCGCTAAAGTTCACCCGAAAACTTCTACTCCGGTT
AACGCATCCGCGTTCCTGGGTATCTGCACCGCTGCAATCG
CGCTGTTCACCGACCTGCAGATCCTTCTGAACCTGGTTAG
CATCGGTACCCTGTTCGTTTTCTACATGGTGGCCAACGCG
GTTATCTACAAACGCTACGTTAGCGTTGGTGTTACCAACC
CGTGGCCGACCCTGTCCTACCTGTTCTGCTTCTCTCTGACG

TCTATCCTGTTCACCCTGCTGTGGCAGTTCGCGCCGCCGG
GTAAACCGAAAGCATTCATGCTGGGCGCATGCACCGCCAT
CGCGATCGGCGTTCTGCAGCTGTTCCACTACATGGTTCCG
CAGGCGCGTAAACCGGAATTCTGGGGTGTTCCCCTGATGC
CGTGGATCCCGTCCATCTCTATCTTCCTGAACATCTTCCTG
CTGGGTTCCCTGGATAAACCGTCTTACGTTCGTTTCGGTTT
CTTCTCTGCACTGGCTGTTCTGGTTTACGTTCTGTATAGCG
TTCACGCTTCCTTCGATGCTGAAGAAGATGGTACCCTGAG
CCAGAAAAACATCGAACTGGTTAAAGAATCTATCGAAAA
CCAGGACCACACCCTGAAAGTTTAA