METABOLIC ENGINEERING OF THE DEOXYXYLULOSE PHOSPHATE PATHWAY FOR ISOPRENOID PRODUCTION IN *E. COLI*

ZHANG CONGQIANG

(M.Eng. Tianjin University)

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN CHEMICAL AND PHARMACEUTICAL ENGINEERING (CPE)

SINGAPORE-MIT ALLIANCE

NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

张聪强

Zhang Congqiang

Date: 26 Mar 2014

ACKNOWLEDGEMENT

Time flies, more than 4 years has passed since I started my PhD study. It is more than study and research, but a colorful and thrilling journey. No doubt my future life will benefit a lot from the four-year training and experience.

Without Prof Too, the thesis would be the “mission impossible” for me. Prof Too is a great mentor for me both in research and life. He is the man of vision, I remember he asked me that do you know what a PhD is and who you want to be when I just joined this lab and these two questions go through my whole research career. He is always supportive and inspiring, when I am frustrated with my failed experiments, when I lose self-confidence and when I am lost, he is always there, chatting and enlightening me patiently. He taught me everything, from global self management, like how to organize your time and how to develop good habits, to details in research such as how to design experiments, how to write proposals and how to keep a good lab book, to how to be an entrepreneur and open a startup; from philosophy to sports, he introduced me to the wonderful karate club I love… He spares no trouble to correct my countless mistakes like my bad habits of doing experiment and my English grammar mistakes of scientific writing. I wish I could learn and work with him much earlier, and I am very grateful to Prof Too. Without him, I would not be as mature as I am now.

My sincere thanks also go to my co-thesis supervisor, Prof. Gregory Stephanopoulos (MIT ChemEng) who gives me insight advices in research that I benefit a lot. And I am also blessed to have Prof. Saif Khan, Prof. Li Zhi, Prof Patrick Doyle and Prof. Raj Rajagopalan as my thesis advisors who gave me many valuable comments.

My heartfelt gratitude to Dr Zhou Kang and Dr Zou Ruiyang, who constantly assist me in the challenges I met in research and in life and give me enormous precious comments, and their successes inspirit me. And my earnest appreciation to Ms. Chen Xixian, my amazing partner, for the time we share ideas, have discussions and have laugh and fun. I also want to thank all my lab mates, especially Dr. Wan Guoqiang, Dr. Zhou Lihan, Dr. Tang Yew-Chung, Mr Lim Qing En, Ms. Sha Lanjie, Ms Sew Hui Yin, Mr Sew Kok Huei, Ms Ho Yoon Khei, Ms Wong Longhui, Ms Chen Meiyi, Mr Cheng He, Ms Chan Christine and Mr Tan Justin for lab discussions and fun and laughter we shared these years. Besides, I also like to thank research/administration officers in SMA-CPE and NUS ChemEng, especially Dr. Yang Liming.

Lastly, and most importantly, I wish to thank my parents. They support me, teach me and love me. And also to my wife Lina and our little daughter Linda who I love and am loved by. To them I dedicate this thesis.

TABLE OF CONTENT

SUMMARY

For more than a decade, the deoxylxylulose phosphate (DXP) pathway has been used to produce many commercially valuable isoprenoids. This thesis is focused on engineering both the pathway and the global environment to maximize isoprenoid production and attempts were made to gain a better understanding of the underlying biochemical mechanisms.

In the first part of the study (Chapter 2), the hypothesis that the availability of a vital precursor (phosphoenolpyruvate, PEP) of the DXP pathway played a significant role in the production of isoprenoids was tested. The underlying rationale was that PEP is known to be consumed by the phosphotransferase system (PTS) when carbohydrates are imported and hence deletion of the PTS should increase intracellular PEP concentration thereby resulting in the enhancement of isoprenoid production. As expected, the replacement of the PTS with the ATP-consuming glucose transport system (galactose permease and glucose kinase) significantly improved production yield of various isoprenoids. In addition, the systematic optimization of the expression of glucose uptake pathway and the DXP pathway further improved the product yield by more than 190 fold over the control, demonstrated the power of global optimization.

Building on the PTS knockout strain, the critical media compositions were screened in lycopene-producing strain using factorial design of experiments. Glycerol and KH_2PO_4 were found to be critical components for high yield lycopene production. Aided by response surface

methodology, the lycopene production was maximized at a particular optimal concentration of glycerol and KH_2PO_4. Consequently, our optimized medium significantly outperformed all the commonly used complex and defined media (Chapter 3). Using real-time qPCR, the transcriptional analyses of the expression of genes in the DXP and the lycopene biosynthetic pathways revealed that transcripts in the latter pathway were highly expressed due to an increase in plasmid copy number in our optimized media. These results highlight the significance of the media optimization and provide insight into the underlying mechanisms of media optimization.

Using the tools we developed, the product titer increased significantly. However, product inhibition becomes a significant challenge as the product titer increases and this limits further improvement and scaling up of the process. To address this issue, we screened the native multidrug resistant transporters in *E. coli* (Chapter 4). Using single gene knockout mutants, we found that *tolC* was closely involved in the efflux of amorphadiene. The deletion of *tolC* entrapped the majority of AD inside the cells and significantly decreased extracellular AD. This was the first time that AD was demonstrated not to be passively secreted through the membrane of *E. coli* but effluxed by the aid of a transporter. Furthermore, it was found that supplementation of *tolC-msbA*, *tolC-emrAB*, *tolC-emrKY* or *tolC-macAB* can enhance the AD titer by more than 3 fold. The use of

transporters may be useful for the production of cytotoxic compounds in *E. coli.*

Flux determination is one of the major challenges in metabolic engineering and the key to quantitatively understand enzymes and pathways and for the design of *de novo* pathways. We developed a semi-closed dual-flux platform to monitor the metabolite intermediates in the DXP pathway (Chapter 5). Using this platform, we managed to capture the kinetic behaviors of the pathway and found that the efflux rates were dependent on the intracellular concentration of the metabolites, almost linearly correlated. This platform which we have developed to gain an understanding of the kinetics of DXP pathway and the observation of the correlation of the efflux rate and intracellular concentration pave the road to understand the DXP pathway quantitatively in the future.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

AD	amorpha-4,11-diene
LYC	lycopene
TRY	titer, rate (or productivity) and yield
IPP	isopentenyl diphosphate
DMAPP	dimethylallyl diphosphate
idi	IPP isomerase
DXP	1-deoxy-D-xylulose-5-phosphate
PTS	the carbohydrate phosphotransferase system
PEP	phosphoenolpyruvate
RBS	ribosome binding site
MAGE	multiplex automated genome engineering
TRMR	Trackable multiplex recombineering
gTME	Global transcription machinery engineering
MVA	mevalonate
GAP	glyceraldehyde-3-phosphate
dxs	DXP synthase
dxr	DXP reductoisomerase
MEP	methylerythritol phosphate
CDPME	4-diphosphocytidyl-2-C-methyl-D-erythritol
ispD	CDPME synthase
ispE	CDPME kinase
ispF	CDPMEP synthase
ispG	HMBPP synthase
ispH	HMBPP reductase
HMGS(erg13)	HMG-CoA synthase
MK(erg12)	mevalonate kinase
PMK (erg8)	phosphomevalonate kinase
PMD (erg19)	mevalonate diphosphate decarboxylase
CCR	carbon catabolite repression
AC	adenylate cyclase
cAMP	cyclic AMP
G6P	glucose-6-phosphate
R5P	ribose-5-phosphate
X5P	xylulose-5-phosphate
HPr	histidine protein
PYR	pyruvate

CRP	cAMP receptor protein
CAP	catabolite gene-activator protein
PPP	pentose phosphate pathway
galP	galactose permease
glk	glucokinase
MMME	multivariate modular metabolic engineering
GPP	geranyl diphosphae
GGPP	geranylgeranyl pyrophosphate
FPP	farnesyl pyrophosphate
ispA	FPP synthase
OD	optical density
GMA	Generalized Mass Action
LCMS	liquid chromatography-mass spectrometry
GGS	galactose permease and glucose kinase system
μ-UNEICS	Univariant Extrinsic Initiator Control System for microbes
GLC	glucose
IPTG	isopropyl β-D-1-thiogalactopyranoside
G3P	glycerol-3-phosphate
DHAP	dihydroxyacetone phosphate
G13P	1,3-biphospho-glycerate
3PG	3-phospho-glycerate
2PG	2-phospho-glycerate
glpF	glycerol facilitator
glpK	glycerol kinase
glpD	glycerol-3-phosphate dehydrogenase
tpi	triose phosphate isomerase
gapA	glyceraldehyde-3-phosphate dehydrogenase A
pgk	phosphoglycerate kinase
gpmM	phosphoglycero mutase III
gpmA	phosphoglyceromutase I
eno	enolase
ppc	PEP carboxylase
pck	PEP carboxykinase
ppsA	phosphoenolpyruvate synthetase
pykFA	pyruvate kinase type I and II
crtE	GGPP synthase
crtB	phytoene synthase
crtI	phytoene desaturase
ADS	amorpha-4,11-diene synthase

RSM	response surface methodology
MDR	multidrug-resistant
ABC	ATP-binding cassette
MFS	the major facilitator superfamily
MATE	multidrug and toxic-compound extrusion
SMR	small multidrug resistance
RND	resistance nodulation division
xylB	xylulokinase
DX	1-deoxy-D-xylulose
13C glu	D-glucose-$^{13}C_6$
BST	Biochemical Systems Theory
NMR	nuclear magnetic resonance

Chapter 1. Thesis description

1.1 Challenges in metabolic engineering

Metabolic engineering offers a sustainable approach for the production of various compounds from inexpensive and renewable starting materials, including bulk chemicals, fine chemicals, drugs and biofuels. The objective of metabolic engineering is that developing cell/cell-free systems and fermentation/scaling-up processes to obtain products/processes with high TRY (high **T**iter, high production **R**ate and high **Y**ield) from renewable resources, thus both economic viable and environment friendly. In order to obtain high TRY, we need to optimize both strains and processes. First, strain improvement includes: enhancing substrate uptake, providing more precursor supply, cofactor balancing, minimizing side reactions, blocking intermediate secretion from cells, coping the toxicity of both intermediates and products and blocking unwanted allosteric regulations by understanding the biological mechanisms. (Figure 1-1). Second, the global environment, including medium and process controlling, has to be optimized specifically for the improved hosts. Although strain improvement is the foundation of metabolic engineering, the global environment control is non-trivial but very critical in minimizing the product costs and determining the economic viability of certain process or product. Therefore, it is urgently required an integrated platform for combining strain optimization and environment optimization.

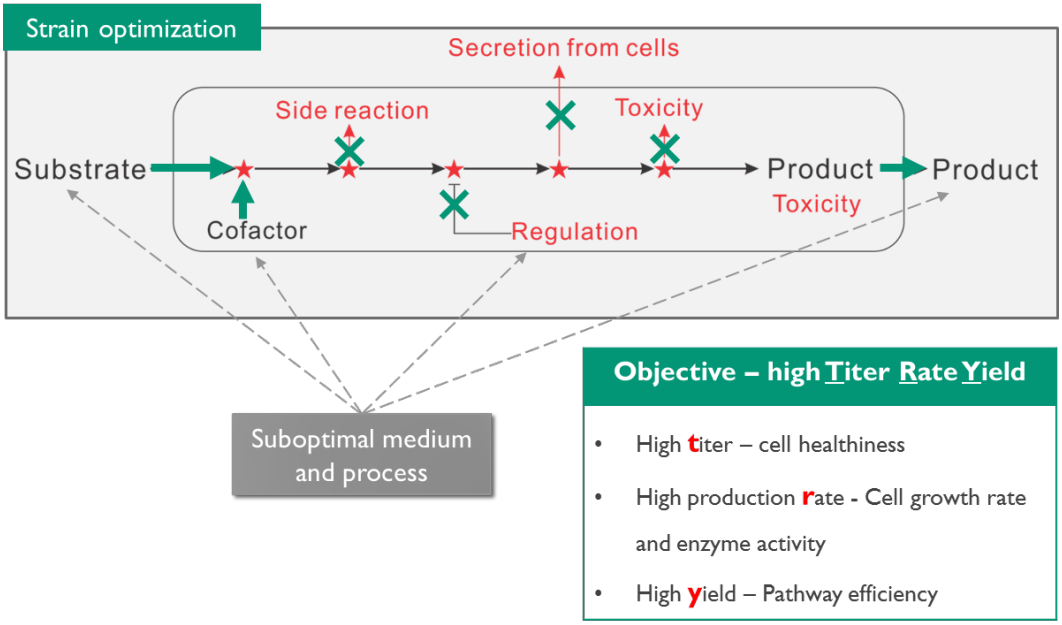


Figure 1-1. Objectives and challenges in metabolic engineering.
Multiple enzymatic reactions introduced into hosts can have undesirable interactions with these engineered pathways, including metabolite toxicity, intermediate loss via secretion or side reactions, unwanted allosteric regulation and product toxicity due to accumulation of high concentration products. On top of strain optimization, the medium and process has to be developed for the optimized strain. And an integrated platform is required to combine strain optimization and environment optimization.

1.2 Objective of this thesis

In this thesis, multiple approaches/platforms were developed to address the challenges listed in Figure 1-1 in metabolic engineering. Although these approaches and platforms were demonstrated in the production of two isoprenoids (lycopene, the antioxidant and amorphadiene, the precursor of antimalarial drug artemisinin) in *E. coli*, the approaches/platforms developed in the thesis are generally applicable for optimizing other pathways and the production of metabolites/proteins and underlying mechanisms discovered in the thesis are applicable for

understanding regulations and interactions in cells in metabolic engineering and synthetic biology.

First, the precursor PEP was identified for the first time as a limiting factor in the DXP pathway and its availability was increased through engineering carbon uptake pathway. To address the toxicity of accumulated intermediates and to enhance the carbon fluxes in the DXP pathway, the approach of experimental design aided systematic pathway optimization (EDASPO) was developed to systematically optimize large number of modules simultaneously. Second, to enhance the product titer and to decrease the product cost, an integrated platform was developed by improving strain using EDAPSO and statistically optimizing the medium and cultural conditions. Third, to deal with the product toxicity, a platform was developed for screening multidrug resistant transporters and engineering transporters for cytotoxic isoprenoid production. Lastly, a semi-closed dual-flux platform (SCDFP) was established to monitor the concentration of intracellular and extracellular metabolite intermediates. This platform and the observed correlation of the efflux and intracellular concentration allow us to understand quantitatively the isoprenoid pathway and the generated mathematical model is invaluable in predicting the pathway behaviors.

1.3 Description of the thesis

In the thesis, multiple approaches were explored and established to enhance the yield of isoprenoids production (lycopene and AD) in *E. coli* using the DXP pathway and to understand the kinetic behaviors of the DXP pathway *in vivo*. In chapter 3, substrate engineering by deleting the PTS was combined with pathway optimization. Replacement of the PTS with *galP-glk* improved the yield of AD significantly (~6 fold). The EDASPO approach was developed and applied to optimized a four-module system by controlling the expression of glucose-uptake enzymes (*galP-glk*), the rate-limiting enzymes of the DXP pathway (dxs and *ispDF-idi*) and the downstream farnesyl diphosphate (FPP) synthase and AD synthase (*ispA-ads*), resulted in more than 550 fold increase in AD production. In chapter 4, the integrated platform was developed to optimize lycopene production by optimizing the growth conditions (media composition, temperature and oxygen feeding) with strain optimization (PTS knockout and pathway optimization). Reverse engineering was applied to gain insight into how media optimization increased the production of these isoprenoids. In chapter 5, native multidrug resistant efflux pumps in *E. coli* were screened using single knockout mutants to identify transporters involved in export of AD and the identified pumps were overexpressed and combined to improve the production of AD. In chapter 6, a well-designed experiment captured the kinetic data of both intracellular and extracellular intermediates of the DXP pathway using a SCDFP developed in this thesis

and the S-system was attempted to model the dynamics of the system. Chapter 7 described the conclusion and future works.

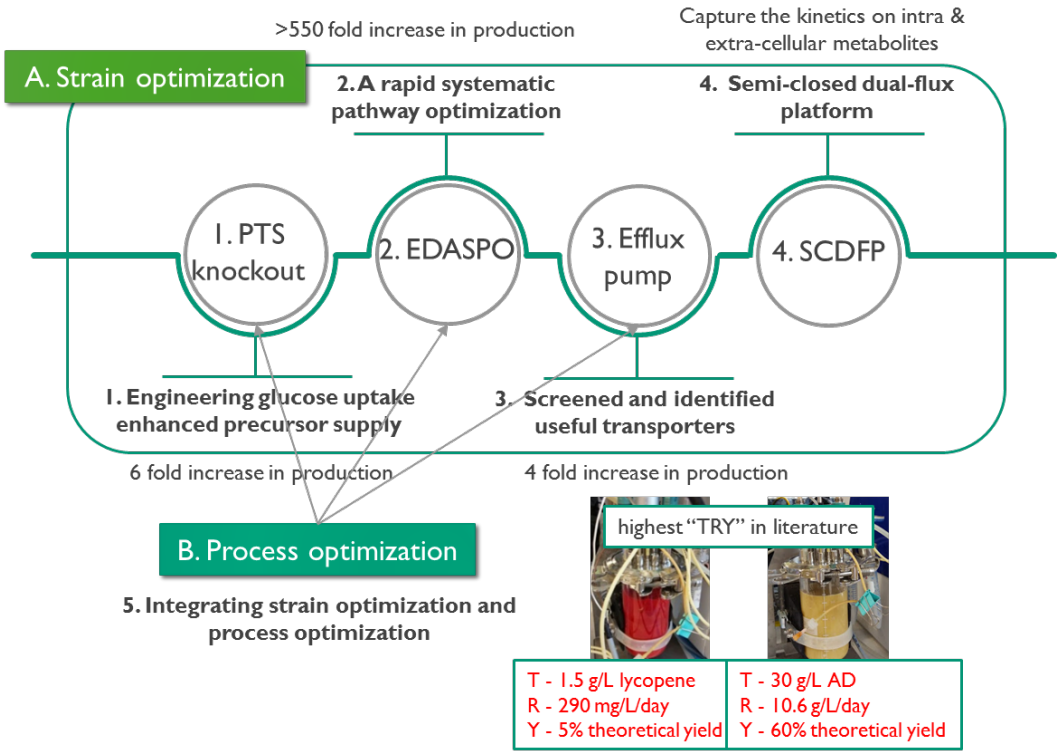


Figure 1-2. Overview of the thesis.
Part I PTS knockout and engineering glucose uptake pathway. Part II, experimental design aided systematic pathway optimization (EDASPO). Part III, Efflux pumps was screened and engineered to enhance amorphadiene titer. Part IV, a semi-closed dual-flux platform (SCDFP). Part V, An integrated platform by integrating strain optimization and process optimization. The approached developed in the thesis was able to produce AD and lycopene at the highest "TRY" in literature (details were not put into the thesis due to undisclosed patent).

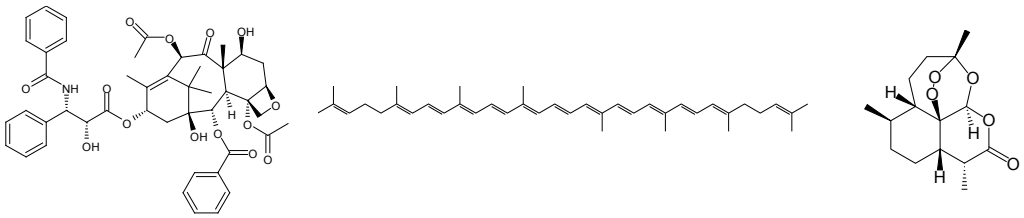
Chapter 2. Literature review

2.1 Literature review

Isoprenoids (or terpenoids) constitute one of the most diverse classes of secondary metabolites in nature, with more than 55,000 distinct compounds. This class of compounds include numerous flavors, fragrances, pharmaceuticals, pheromones and even biofuels [1]. Isoprenoids could serve various functions such as membrane structure (sterols, hopanoids), hormones (gibberellins), light harvesting and photoprotection (chlorophylls, carotenoids), electron carriers (ubiquinone, plastoquinone), and mediators of polysaccharide assembly, as well as communication and defense mechanisms [2, 3]. The various biological functions of isoprenoids are related to their diversified structures.

Some isoprenoids are long-chain linear hydrocarbons (e.g. lycopene), while most of them exhibit complex structures having chiral, cyclic skeletons (e.g. Taxol). Based on the number of carbon atoms, isoprenoids are classified as hemi- (C5), mono- (C10), sesqui- (C15), di- (C20), ses- (C25), tri- (C30), tetra- (C40) and poly- (> C40) isoprenoids. The demand for some isoprenoids (e.g. Taxol as anticancer drug, artemisinin as anti-malaria drug and lycopene as nutraceutical, Figure 2-1) is on the rise and there is a global shortage due to limited supply from natural sources like plants. For example, it would require the entire bark of a 100-year old *Taxus brevifola* tree to extract a single dose of Taxol [4]. Market

fluctuation and environmental changes have further exacerbated the problem and this resulted in a growing demand for a sustainable supply of these isoprenoids [5, 6]. Total chemical syntheses of isoprenoids have been successful but face significant challenges including low yield due to their highly complicated structures (especially with compounds that have multiple chiral centers). Alternatively, metabolic engineering of microbes has emerged as a solution that addresses these challenges [7].



Taxol Lycopene Artemisinin

Figure 2-1. Chemical structure of lycopene, artemisinin and taxol

2.1.1 Metabolic engineering

Metabolic engineering is defined as the directed modulation of metabolic pathways using methods of recombinant technology for the purpose of producing biofuels, bulk and fine chemicals and pharmaceuticals [8]. An important objective in metabolic engineering is to be able to produce compounds (drugs, bulk chemicals, fuel and biopolymers) from inexpensive carbon source (e.g. sucrose, starch and cellulosic biomass) in a sustainable, renewable and economical way (Figure 2-2). To be cost effective and commercially viable, three important parameters, namely titer, rate (or productivity) and yield (TRY), have to be optimized [9].

With the development of next-generation sequencing technology and 'omics', previously unrecognized metabolic reactions have now been identified and the number of pathways to be explored has increased tremendously. Databases of genes, metabolic reactions and enzymes have broadened our capabilities to develop *de novo* design of biosynthetic pathways using computational tools [10]. The recent development of high-throughput analytic tools has enabled us to decipher "–omics" (genomes, transcriptomes, proteomes, metabolomes and fluxomes) at system level and has led to the appreciation of "system metabolic engineering" [11]. In addition, reverse engineering, integrated with specific perturbation or random genetic modifications, have helped to probe the causal network that links genotype and phenotype and to develop better metabolic engineering strategies [12]. To summarize, the strategies of metabolic engineering in literature are categorized in Figure 2-2, ranging from synthetic biology to protein engineering to pathway engineering.

Table 2-1. Intuitive rational approaches and systematic combinatorial approaches.

Intuitive rational approaches	Systematic combinatorial approaches
Precursor enrichment and balancing	Multivariate-modular pathway engineering
Elimination of the competing and/or degrading pathways	RNA or protein scaffold-based spatial optimization
Rerouting metabolic pathway	Global perturbation of genome or transcriptome
Engineering of carbon source utilization	Directed evolution and enzymatic evolution
Transporter engineering	
Cofactor optimization	

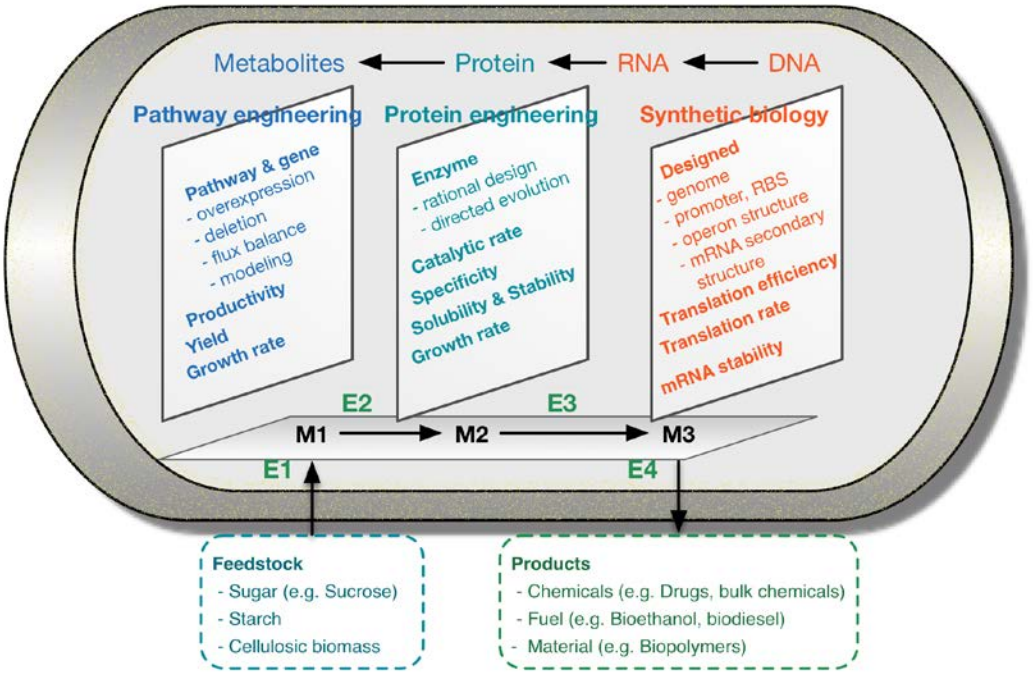


Figure 2-2. Metabolic engineering as a synthesis of synthetic biology and protein and pathway engineering. Reproduced with permission of the copyright owner [9].

From another perspective, these strategies can be grouped into two categories: Intuitive rational approaches and systematic combinatorial approaches [11]. Intuitive rational approaches are mainly used to solve well-defined problems and thus more intuitive. While systematic combinatorial approaches are applied to improve the overall performance of cells that are treated as a global network with high complexity (Table 2-1).

2.1.1.1 Intuitive rational approaches in metabolic engineering

1. Precursor enrichment and balancing

A common approach in metabolic engineering to increase the availability of precursors in cells is by overexpressing rate-limiting enzymes (e.g.

Taxol [13], levopimaradiene [14], plant alkaloids [15], lycopene [16], L-valine [17], artemisinin [18]). In addition, balancing the supplys and demands of precursors are also very important when more than one precursor is required. For example, all the isoprenoids are built from two units, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP isomerase (*idi*) is the enzyme which catalyzes the inter-conversion of relatively un-reactive IPP to more-reactive electrophile DMAPP. *Idi* gene was shown to play an important role in balancing the ratio of IPP and DMAPP and improving the isoprenoid production [19-21]. Similarly, balancing of the fluxes of the two precursors of the 1-deoxy-D-xylulose-5-phosphate (DXP) pathway (the pyruvate and glyeraldehyde-3-phosphate) was shown to increase the lycopene production significantly [22-24].

2. Elimination of the competing and/or degrading pathways

Elimination of competing pathway to enhance the flux through the desired pathway is another commonly used approach in metabolic engineering. There are numerous successful examples using this approach, especially in the production of primary metabolites (e.g., succinic acid [25], lactic acid [26] and amino acids [27]). Two recent examples where the deletion of acyl-CoA synthetase genes improved the production of fatty acid-derived biofuels and chemicals in *Saccharomyces cerevisiae* [28], and the deletion of *aspA* gene which blocked the conversion of fumaric acid into L-aspartic acid increased the production of fumaric acid [29].

3. Rerouting metabolic pathways

This approach contains two aspects: the first is to search for more efficient or substrate-specific enzyme via heterologous expression of enzymes from various sources. For example, to increase the yield of plant alkaloids, the l-DOPA-specific decarboxylase from *Pseudomonas putida* strain KT2440, which exhibited greater than 103-fold preference for l-DOPA in comparison to other aromatic amino acids, was selected to replace the less specific monoamine oxidase [15]. The second aspect is to construct new biosynthetic pathways. For instance, 1-Butanol titer was increased 10 fold by engineering a modified 1-butanol pathway with an irreversible reaction catalyzed by trans-enoyl-coenzyme A (CoA) reductase that created NADH and acetyl-CoA as driving forces to direct the flux [30].

4. Engineering of carbon source utilization

To reduce the cost, different approaches have been applied to enhance the uptake efficiency of carbon sources in microbes. A bi-level optimization algorithm, SIMUP, was developed to find metabolic engineering strategies to efficiently use lignocellulosic biomass which contains both pentose (xylose) and hexose (glucose) [31]. Likewise, the inactivation of the carbohydrate phosphotransferase system (PTS) has been demonstrated to increase the aromatic products by preserving more phosphoenolpyruvate (PEP) [32].

5. Transporter engineering

Removal of products from reaction system can alter the equilibrium of chemical reactions. Export of products into extracellular space through native/engineered transporters can play a similar role. In addition, this export process can minimize feedback inhibition and accumulation of products to potentially toxic levels intracellularly. Engineering of efficient transporters successfully improved the tolerance of hosts to toxic and/or high-yield products (e.g. biofuels) [33-35] and further increased production (e.g. carotenoids) [36].

6. Cofactor optimization

Many redox enzymes catalyze synthetic reactions which commonly require precursors and cofactors, such as NADPH, NADH and ATP. The synthesis of certain products may be limited by the availability of cofactors or the imbalance of redox. Thus, cofactors engineering has been investigated and successfully applied to increase the yield of various products (polyphenols [37], lycopene [38], thymidine [39], succinate [40]). For example, NADPH pool was increased by deletion or attenuation of glycolytic enzymes which redirected metabolic flux through the pentose phosphate pathway [37, 41], or the replacement of central metabolic enzyme that produced NADPH instead of NADH [42, 43].

2.1.1.2 Systematic combinatorial approaches

1. Multivariate-modular pathway engineering

Overexpression of single rate-limiting enzyme cannot result in optimal production for most of the metabolic pathways, and global balancing of the

expressions of multiple enzymes is a useful approach [44]. To search for the optimal balance between pathway expression and cell viability, synthetic promoters libraries [45] or tunable intergenic regions libraries (e.g. RBS, mRNA secondary structures) has been explored [46, 47]. A good example is that of taxadiene production which was increased ~15,000 fold through a systematic optimization of upstream and downstream modules [13].

2. RNA or protein scaffold-based spatial optimization

Multi-enzymatic pathways are often spatially organized into scaffolds, clusters or microcompartments [48]. Three-dimensional organization and optimization can concentrate reactants to drive unfavorable reactions [49], minimize loss of intermediates due to diffusion, protect unstable intermediates from solvent and/or decrease the transit times of intermediates [50]. Synthetic RNA or protein scaffold has been designed to optimize pathway through spatial optimization, which resulted in 48-fold increase in hydrogen production [51] and 77-fold increase in mevalonate production [50].

3. Global perturbation of genome or transcriptome

Compared to intuitive rational approaches, combinatorial approaches can be more effective through exploiting biodiversity and high throughput screening [52]. The approach of random mutagenesis (introduced by chemical mutagens or ultraviolet light followed by screening of strains) has been employed extensively for the generation of industrial strains for many

decades [53]. Extending this approach, reverse engineering (or inverse engineering) was introduced which involves genome sequencing the improved phenotypes and finding out the mutation sites and then to engineer the genotypes. The information obtained from sequencing is critical to decipher regulatory mechanisms and to find multiple knockout/overexpression targets, or mutational site of enzymes. Understanding the underlying biochemical mechanism can direct future engineering work to further improve the performances of phenotypes.

Nowadays, many approaches have been developed to generate large-scale libraries of genomic mutants in a relative short time (a few days), which greatly facilitate the subsequent screening and sequencing analyses. For example, multiplex automated genome engineering (MAGE) was developed for large-scale programming and evolution of cells by simultaneously targeting many locations on the chromosome. Using MAGE, over 4.3 billion combinatorial genomic variants were created per day which enormously expanded the genotype library [54]. Trackable multiplex recombineering (TRMR), which generated a mixture of barcoded mutants, was successfully applied to identify genes that improved the tolerance of lignocellulosic hydrolysate [55]. Global transcription machinery engineering (gTME) was another method developed to reprogram transcriptome of prokaryote and eukaryote. gTME has been successfully applied to increase ethanol tolerance and lycopene production of *Escherichia coli* (*E. coli*) through mutating the sigma factor

[56] and the ethanol tolerance of *S. cerevisiae* via mutating TATA-binding protein (encoded by SPT15) and one of the TATA-binding protein-associated factors (TAF25) [57].

4. Directed evolution and enzymatic evolution

Directed evolution has been used for years to improve solvent tolerance of strains [33], growth rate of mutated or stressed strains [41], and enzyme performance (e.g. turnover number k_{cat} and/or catalytic efficiency k_{cat}/K_m [14], substrate and/or product specificity [58] and thermostability [59]). Improvement of enzyme activity via enzyme evolution has the advantage of introducing less stress to cells with lower expression but higher activity. In contrast, overexpression of enzyme may lead to the formation of inclusion bodies which have no or little activity and thus affects the cell growth and performance [60].

Most of the aforementioned approaches are technically distinct and may be combined for possible use to synergistically improve productivity.

2.1.2 Isoprenoid synthesis pathway and *Escherichia coli* platform

In nature, isoprenoids are derived from two basic units, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP and DMAPP are synthesized from two pathways, the mevalonate (MVA) and the DXP pathway (or MEP pathway) (Figure 2-3). The MVA pathway is present in eukaryotes (all mammals, the cytosol and mitochondria of plants, and fungi) archaea, and a few eubacteria [61]. While the DXP

pathway exists mainly in eubacteria, cyanobacteria, green algae, *apicomplexan* parasites and higher plants [62]. Interestingly, plants have both pathways, located in different cellular compartments.

E. coli is a favorable platform for recombinant proteins [63], chemicals and biofuels [64, 65], due to numerous advantages, e.g., well-characterized genetics and metabolism, rapid growth rate, lots of genetic tools and databases and easy to scale up in production. In *E. coli*, IPP and DMAPP are synthesized through the DXP pathway. DXP pathway starts from two precursors, pyruvate and glyceraldehyde-3-phosphate (GAP). Pyruvate and GAP are catalyzed by DXP synthase (*dxs*) to produce DXP and CO_2. DXP reductoisomerase (*dxr*) converts DXP into methylerythritol phosphate (MEP) using NADPH as the cofactor. The gene *lspD*, *ispE* and *ispF* encode 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) synthase (CMS, which catalyzes CDP-ME from MEP and CTP), CDP-ME kinase (CMK, which converts the ATP-dependent phosphorylation of CDP-ME to 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate (CDP-MEP)) and 2-C-methyl-D-erythritol-2,4-diphosphate (MEC) synthase (MCS, which produces CDP-MEP into MEC), respectively. *lspG* gene encodes 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate (HMBPP) synthase (HDS) that catalyzes the formation of HMBPP from MEC. *lspH* encodes 4-hydroxy-3-methyl-2-(*E*)-butenyl-4-diphosphate reductase (IDS) that transforms HMBPP into a 5:1 mixture of IPP and DMAPP. Lastly, IPP isomerase (coded by gene *idi*) interconverts IPP and DMAPP in *E. coli*

(Figure 2-3). The MVA pathway uses seven enzymatic reactions to transform the precursor acetyl-CoA to IPP and DMAPP, through five intermediates, acetoacetyl-CoA, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), mevalonic acid, mevalonate-5-phosphate, mevalonate-5-diphosphate (Figure 2-3).

Both MVA and DXP pathways have been exploited industrially for the production of isoprenoids in yeast and bacteria. Currently, research groups using the MVA pathway have produced high titer of isoprene [66]. The MVA pathway was completely described in 1967 [67]. In contrast, the DXP pathway was discovered relatively recently [62] and has not been well studied, especially with respect to the iron-sulfur cluster enzymes ispG and ispH which are poorly characterized [68-71]. Though optimization of the DXP pathway is challenging, the DXP pathway is worth pursuing because of the more efficient use of carbon, with 30.2% theoretical yield when using glucose versus 25.2% for the MVA pathway, as demonstrated in Table 2-2.

Table 2-2. Comparison of theoretical yields of the DXP and MVA pathway

Pathways	Biochemical equation	Theoretical yield
MVA	1.5 Glucose + 2 O2 → IPP + 4 CO2 + 5 H2O	25.2%
DXP	1.25 Glucose + 0.5 O2 → IPP + 2.5 CO2 + 3.5 H2O	30.2%

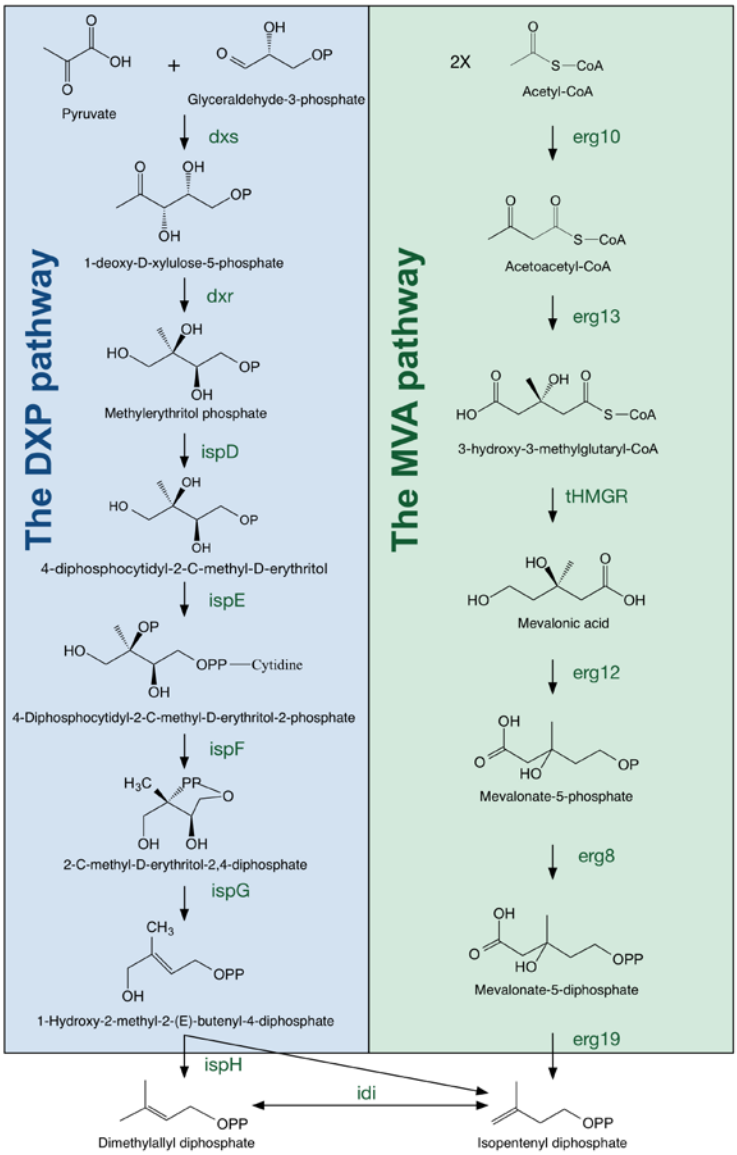


Figure 2-3. 1-deoxy-D-xylulose-5-phosphate (DXP) and mevalonate (MVA) pathway. Enzymes in the DXP pathway, *dxs*, DXP synthase; *dxr*, DXP reductase; *ispD*, CDPME synthase; *ispE*, CDPME kinase; *ispF*, CDPMEP synthase; *ispG*, HMBPP synthase; *ispH*, HMBPP reductase; *idi*, IPP isomerase. Enzymes in the MVA pathway, AAS (erg10), acetoacetyl-CoA synthase; HMGS (erg13), HMG-CoA synthase; MK(erg12), mevalonate kinase; PMK (erg8), phosphomevalonate kinase; PMD (erg19), mevalonate diphosphate decarboxylase.

2.1.2.1 Lycopene and amorphadiene

In this thesis, two particular isoprenoids, lycopene (non-volatile metabolite) and amorphadiene (volatile metabolite) were chosen as examples to demonstrate the platforms developed in this thesis.

The recent changes in the food regulation and labelling rule has resulted in the increase in demand of natural colorant as food additives and nutraceuticals. As a natural reddish pigment, lycopene can potentially replace many colours traditionally produced by chemical synthesis (*e.g.* carmine). As an antioxidant compound, lycopene is also a nutraceutical with various health benefits and can be used as food and drink additives. As a powerful free radical quencher, lycopene is also applied as skin protectant in cosmetics to reduce the damage from skin exposure to sunlight [66]. In addition, there is evidence from phase I and II clinical trials indicates that lycopene is a potential chemoprotective agent for prostate cancer [67-70]. Recent study indicates that lycopene supplementation and tomato uptake can lower the risk of cardiovascular disease [71]. Therefore, the global market lycopene is expected to increase to $84M in 2018 from $66M in 2010, according to a recent BCC research.

Amorphadiene is the precursor of artemisinin. Artemisinin is naturally produced by a plant, *Artemisia annua*, and it was identified as the active compound responsible for antimalarial properties of *Artemisia annua* extracts by Chinese scientists in the 1970s [72]. It was estimated that there were 219 million cases of malaria reported and 660,000 malaria-

associated deaths occurred in 2010, and about 90% of malaria death occurred in Africa [73]. Many malaria-endemic developing nations have a per capita healthcare expenditure of under $100. In other words, a much cheaper and sustainable supply of artemisinin is required. Semi-synthetic artemisinin project funded by a grant from the Bill and Melinda Gates Foundation are under developing.

2.1.3 Substrate utilization engineering

Lignocellulose biomass is viewed as potential feedstock for future production of chemicals and fuels [74]. Lignocellulose hydrolysate contains glucose and xylose as the two major carbon sources. The problem of using lignocellulose is that microbe such as *Saccharomyces cerevisiae* is unable to utilize the xylose or that microorganism such as *E. coli* uses this sugar mixture sequentially and hence the yield and productivity are limited [75]. To tackle this problem, substrate utilization engineering is introduced in two aspects, one is for the engineering of the strains to use unfavorable carbon sources, and the other is to make more efficient use of carbon resource.

1. Engineering microbes to use xylose and pentose simultaneously

The sequential use of sugar mixture in *E. coli* is closely related to the phenomenon of carbon catabolite repression (CCR). CCR is regulated by EIIA in the carbohydrate phosphotransferase system (PTS). As shown in Figure 2-4, glucose is internalized mainly by the PTS using PEP as phosphate donor. In this process, the ratio of PEP/Pyruvate decreases,

thus EIIA is mainly dephosphorylated and inhibits the expression of protein adenylate cyclase (AC). The repression of AC prevents the production of cyclic AMP (cAMP) and thus the cAMP-CRP complex cannot be formed, resulting in the inhibition of the promoters of many catabolic genes and operons including xylose uptake genes. To disrupt this regulation and to co-utilize sugar mixture, *ptsG* (coding EIIC) can be deleted and this resulted in the simultaneous utilization of glucose and lactose [76]. Inspired by this work, a *ptsG* mutant was applied to co-utilize glucose, xylose and arabinose and achieved a 94% theoretical yield of ethanol [77]. In addition to *ptsG*, deletion of *mgsA* [78] or *crp* [79] was also reported to facilitate the co-utilization of glucose and xylose. An approach to co-culture two kinds of engineered *E. coli* was applied where one strain used glucose but not xylose and the other used xylose but not glucose [80].

Recently, a rational bi-level optimization algorithm, SIMUP, was developed to predict the knockout mutant strains which have to use xylose and glucose simultaneously [31]. Based on the prediction of SIMUP, seven solutions were found and grouped into three strategies, A, B and C (Figure 2-5). In strategy A, the synthesis of only glucose-6-phosphate (G6P) was from glucose, and all the remaining precursors were from xylose. LMSE1, the mutant with three genes (*pgi*, *gnd* and *eda*) deleted, was constructed to validate the strategy A (Figure 2-5A). However, the predicted mutant LMSE1 did not co-utilize glucose and xylose, but consumed xylose alone,

and this could be due to either G6P was not a biomass precursor or the mutant could synthesize G6P from xylose. Strategy B was a group of five solutions in which synthesis of only ribose-5-phosphate (R5P) was from glucose and only xylulose-5-phosphate (X5P) was from xylose. A three-gene (*pgi*, *rpe* and *eda*) mutant, LMSE2, was constructed to validate the strategy (Figure 2-5B). As predicted, LMSE2 co-utilized glucose and xylose very efficiently. Strategy C is not intuitive as strategy A and B. The mutant LMSE5 was constructed based on the solution with four reaction disrupted by deletion of the following genes, *pgi*, *gnd*, *fbp*, and *pfk* (Figure 2-5C). Deletion of *pgi* and *fbp* prevented gluconeogenic flux, and disruption of *gnd* stopped entry of glucose into pentose phosphate pathway. In agreement with the SIMUP prediction, the mutant co-utilized glucose and xylose in batch cultivation.

The advantage of SIMUP is that co-utilization phenotype was achieved without any prior knowledge of the regulatory mechanisms involved in CCR. This is especially useful when the targeted mutations in the regulatory pathways is yet to be identified.

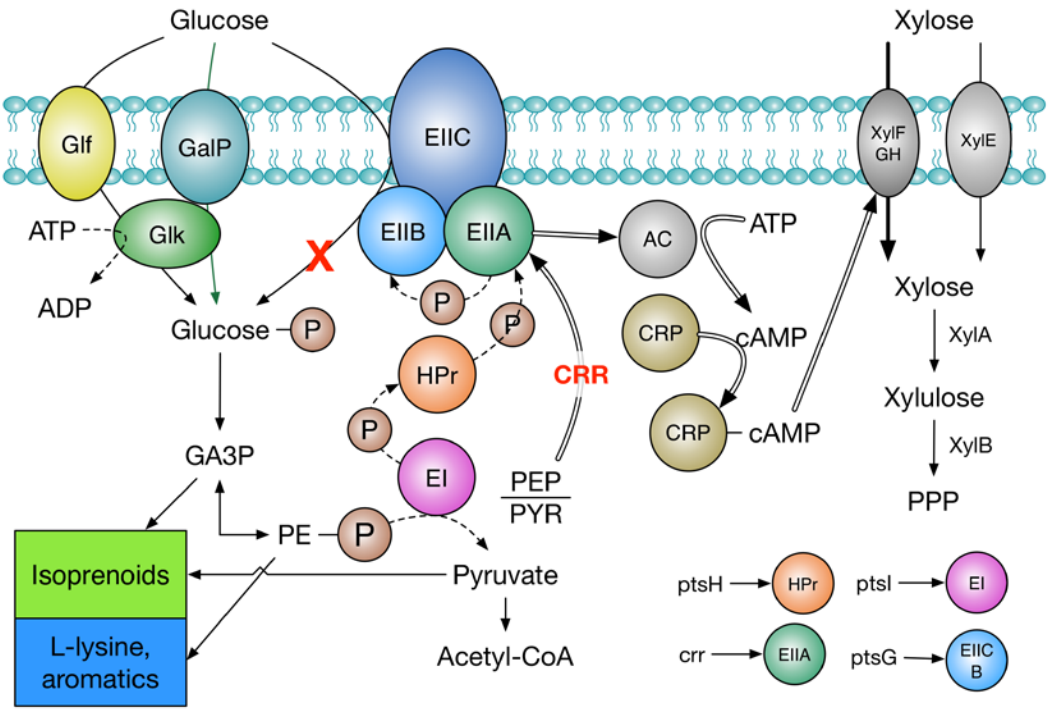


Figure 2-4. Glucose and xylose uptake pathway in *E. coli*
In wild type *E. coli*, glucose is mainly transported by the PTS. The PTS is composed of three distinct proteins: enzyme I (EI), histidine protein (HPr) and enzyme II (EII). EI initiates the phosphorylation chain by autophosphorylating with phosphoenolpyruvate (PEP). The phosphoralated EI transfers the phosphyl group to Hpr, and the phosphryl group is subsequently donated to EIIA, finally to EIIB and from there to the carbohydrate through the membrane domain EIIC. EIIA is also named as catabolite repression resistance (crr), which regulates the carbon catabolite repression (CCR). When glucose is abundant, the phosphryl group is drained from PEP to glucose through EIIA, thus the concentration ratio between PEP and pyruvate (PYR) is low, and EIIA is dominantly dephosphorated. On the contrary, when glucose is absent, the ratio PEP/PYR is high, and EIIA is mainly phosphoralated. Phosphoralated EIIA activates membrane-bound protein adenylate cyclase (AC), and AC converts ATP into cyclic AMP (cAMP). The produced cAMP binds cAMP receptor protein (CRP, also called catabolite gene-activator protein (CAP)) and forms cAMP-CRP complex. The cAMP-CRP complex activates the promoters of many catabolic genes and operons (e.g. xylose uptake gene) [81]. Two pathways exists in *E. coli* for the uptake of xylose, xylFGH and xylE. XylFGH is ATP-binding cassette transporter with high affinity and it's the major transporter for xylose; xylE is D-xylose proton symporter which uses the proton gradient as a source of energy [82]. The expression of XylFGH and xykE is under the regulation of CRR. The abbreviations in the figure are, PTS, The phosphotransferase system; XylE, D-xylose transport system (low affinity); XylFGH, D-xylose transport system (high affinity); XylA, Xylose isomerase; XylB, Xylulokinase; PPP, Pentose phosphate pathway; AC, Adenylate cyclase; Glf, Glucose facilitator from *Z. mobilis*; GalP, galactose permease; Glk, glucose kinase; cAMP receptor protein; CRR, Catabolite repression resistance

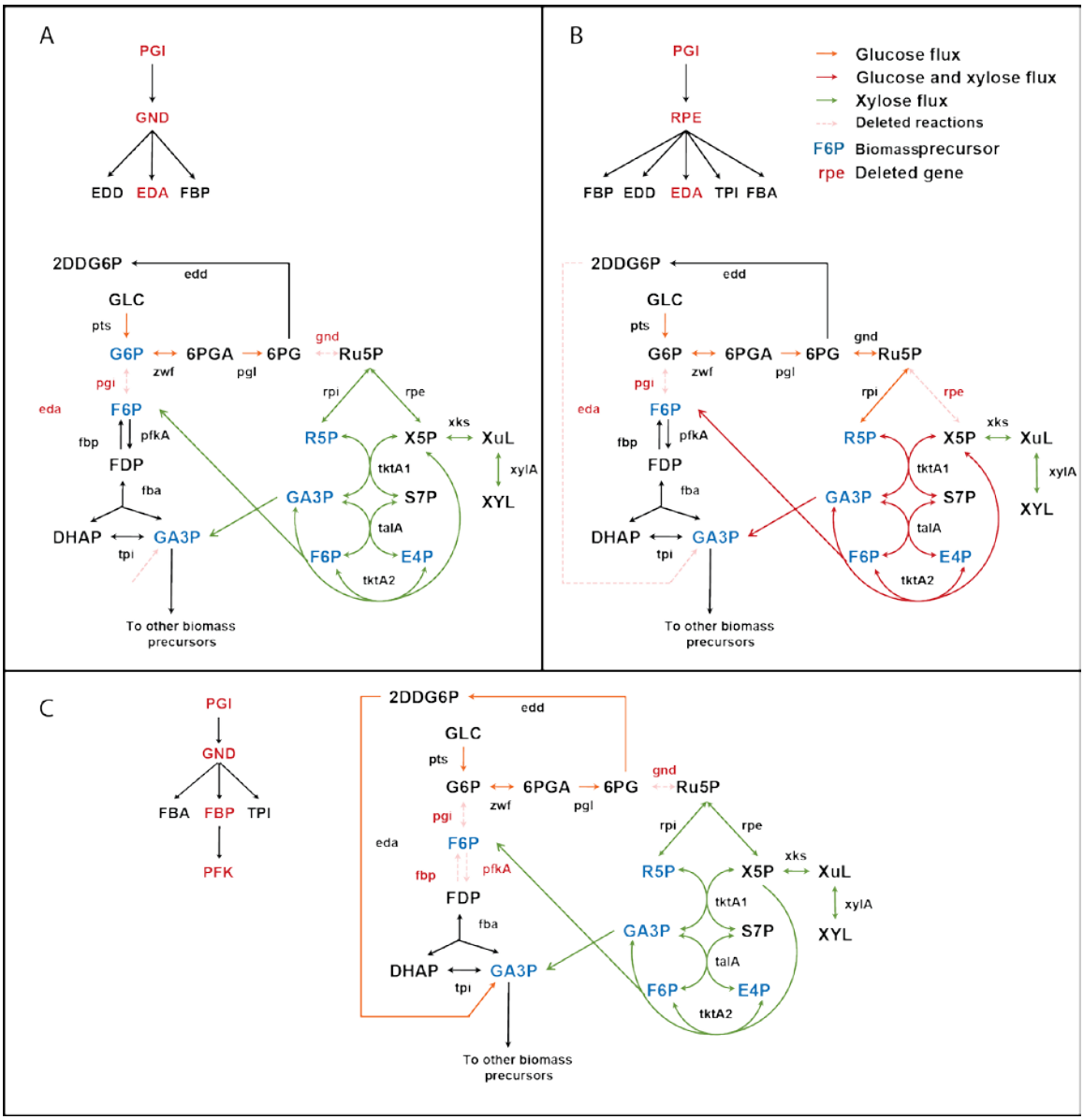


Figure 2-5. Predicted solutions by the SIMUP algorithm.
Based on how each solution splits the metabolic network, they have been divided into three strategies: (A) Strategy A in which glucose is used to synthesize glucose-6-phosphate and all other metabolites are synthesized from xylose. Among the three gene combinations shown, the highlighted (red) solution was used to construct mutant LMSE1. Deletion of *pgi* blocked glucose to enter glycolysis. Disruption of *gnd* prevented glucose from entering pentose phosphate pathway (PPP) and the synthesis of glucose from PPP. Finally, deletion of *eda* stopped glucose from synthesizing glyceraldehyde-3-phosphate (GAP), which could be channeled into glycolysis. (B) Strategy B has five solutions. The highlighted solution was used to construct the mutant LMSE2. Glucose was used to synthesize ribose-5-phosphate and xylose was used to synthesize xylulose-5-phosphate, and using these two metabolites all the remaining biomass precursors were synthesized. (C) Strategy C contains three solutions. The four gene deletion strategy highlighted was used to construct the mutant LMSE5. [31]. Reproduced with permission of the copyright owner.

2. Efficient use of glucose with the PTS knockout

The uptake of glucose by the PTS requires equal amount of PEP as the phosphate donor, and this leads to inefficient use of glucose for the production of certain compounds (e.g. recombinant proteins [83], aromatics [32], lysine [84], L-phenylalanine [85] and succinate [86]) as PEP is the important precursor (Figure 2-4). To address the limited availability of PEP, PEP synthase (ppsA) was overexpressed to recycle PEP from pyruvate [87]. Another way was to disrupt the PTS while other pathways were used to transport glucose [32, 38, 83, 85, 88, 89]. PTS disruption resulted in slower growth of *E. coli*. To increase growth rate, adaptive evolution was explored to screen fast-growing strains lacking the *ptsHIcrr* operon [32, 38]. Another solution was to overexpress galactose permease (*galP*) and glucokinase (*glk*) which facilitated the internalization and phosphorylation of glucose, respectively [90, 91].

PEP is a vital precursor for the DXP pathway, and the increase in PEP concentration by genetically rerouting central metabolism has been shown to enhance the production of lycopene [22]. Since PEP is consumed by the PTS when carbohydrates are imported [92], it is worth investigating whether deletion of PTS would increase intracellular PEP concentration which may result in the enhancement of isoprenoid production.

2.1.4 Optimization of metabolic pathways

In metabolic engineering, it is well known that overexpression of single rate-limiting enzyme cannot lead to globally optimized pathway and thus

balanced expression of multiple enzymes is required [93, 94]. Fine-controlled pathway can regulate the concentration of intermediates thus eliminate their toxicity/inhibition and maximize the production [46, 95]. Pathway balancing can be viewed as the extension of traditional method of overexpressing rate-limiting genes. The rate-limiting gene is actually context-dependent, e.g., when the rate-limiting enzyme is overexpressed, another one may serve as the new rate-limiting step. Therefore, identification of the rate-limiting steps often requires multiple steps which is time consuming [19]. Combinatorial/modular approach can solve this issue by simultaneously controlling the expression of multiple enzymes thus expediting the process.

Two recent studies have demonstrated the usefulness of this approach. One was the sequential optimization of the MVA pathway to produce amorpha-4,11-diene (AD) [96]. The whole pathway was grouped into three parts, MevT operon which controls the upstream genes, *AtoB*, *HMGS* and *HMGR*, MBIS operon which controls the downstream genes (*MK*, *PMK*, *PM*, *idi* and *ispA*) and AD synthase (*ADS*) (Figure 2-6A).

The upstream pathway was firstly optimized by codon optimization and replacement with a stronger promoter. Replacing the promoter of *ADS* with a stronger one further increased the AD production. Next, rate-limiting step (*MK*) in the MBIS operon was identified by overexpressing each gene individually, and the codon-optimized *MK* was incorporated into the system, resulting in higher productivity.

The other example was the production of Taxol precursor using multivariate modular metabolic engineering (MMME) [13]. The authors assembled the DXP pathway (including *dxs*, *ispD*, *ispF* and *idi*) into the upstream module and the geranylgeranyl pyrophosphate (GGPP) synthase and taxadiene synthase into the downstream module (Figure 2-6B). Simultaneously controlling the three different-strength promoters (Trc, T5 and T7) or different gene copy numbers (1 copy in chromosome, ~5 copies, ~10 copies and ~20 copies in pSC101, p15A or pBR322 plasmid, respectively), the upstream pathway and downstream pathway were well balanced that a global maximum exhibited a 15,000-fold increase in the titer of taxadiene over control (Figure 2-6C). This approach also facilitated the discovery of indole accumulation that was later shown to be detrimental to isoprenoid production. This observation demonstrated that MMME could also be applied to survey the regulations in the expression of heterologous pathways.

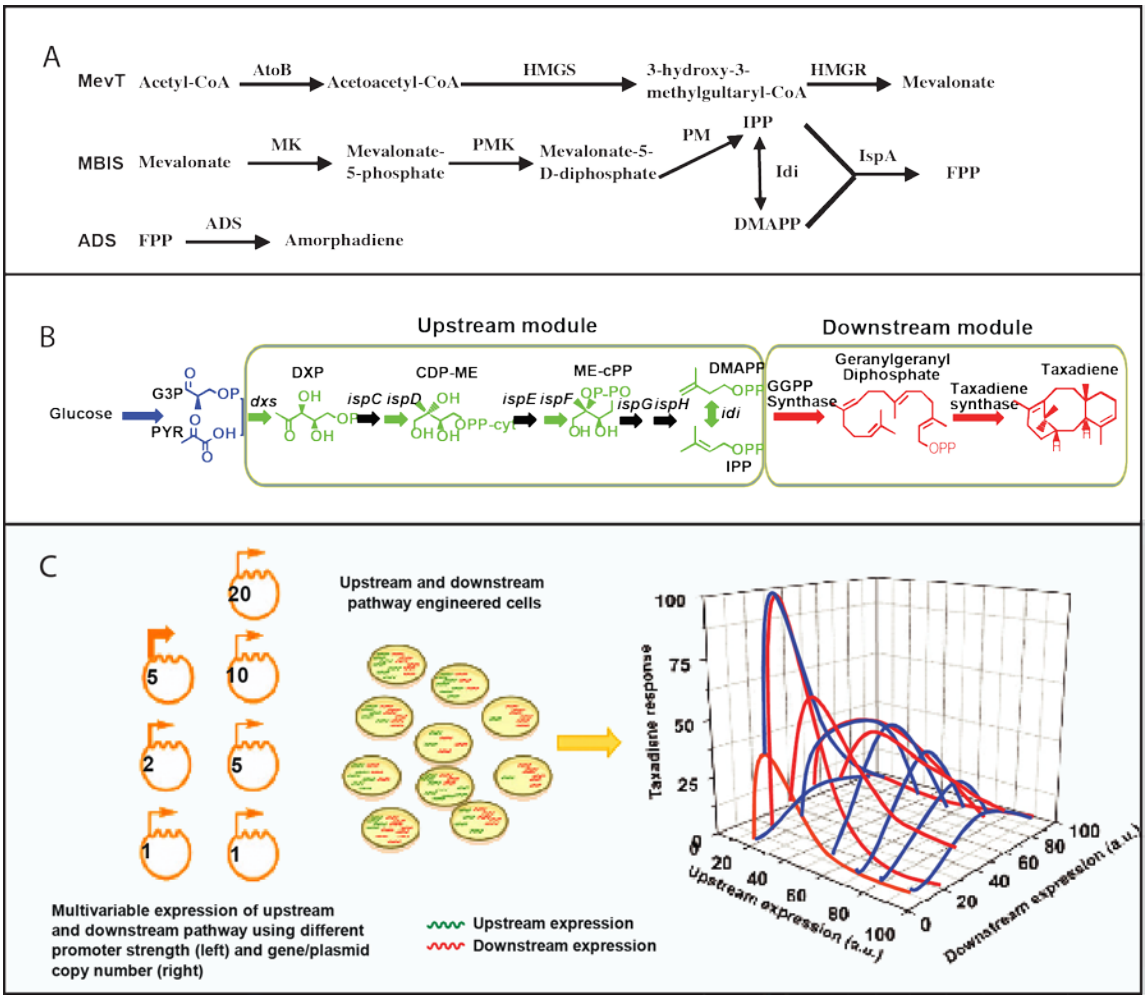


Figure 2-6. Illustration of modules of the MVA and DXP used for production of AD and taxadiene.

(A) Production of AD via the mevalonate pathway. The abbreviations for enzymes and pathway intermediates are as follows: AtoB, acetoacetyl-CoA thiolase; HMGS, HMG-CoA synthase; tHMGR, truncated HMG-CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; PMD, mevalonate pyrophosphate decarboxylase; Idi, IPP isomerase; IspA, FPP synthase; ADS, AD synthase; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate [96]. (B) Schematic of the two modules, the native upstream MEP isoprenoid pathway (green) and synthetic taxadiene pathway (red). (C) Schematic of the multivariate-modular isoprenoid pathway engineering approach for probing the non-linear response in terpenoid accumulation from upstream and downstream pathway engineered cells. Expression of upstream and downstream pathways is modulated by varying the promoter strength [1 (Trc), 2 (T5) and 5 (T7)] or gene/plasmid copy number (right) [97]. Reproduced with permission of the copyright owner.

2.1.5 Transporter engineering

Many heterologous products (e.g. biofuel) are toxic to microbes, and this product toxicity is one of the major problems that prevent successful commercialization by microbial production [36]. To enhance tolerance of the strains and improve the yield of products, one efficient strategy is to recruit membrane transport protein pumps to export/secrete intracellular biofuels out of the cells [98]. Successful efflux of products can improve cell viability and enhance the production and productivity by alleviating product inhibition. The successful application of efflux pumps depends severely on successful identification of the appropriate pumps for the desired products. For example, the *ygaZH* genes encoded a protein which shows significant homology to the known *Corynebacterium* L-valine exporter and could be a potential pump for L-valine in *E. coli*. And co-overexpression of the *lrp* (which can activate the expression of *ygaZH*) and *ygaZH* genes led to a 2.13-fold increase in L-valine production [17]. Transcriptome analyses revealed that the expression of *rhtC* gene (encoding a threonine exporter) in threonine-producing strain was increased as compared to control strain. Overexpression of *rhtC* resulted in 50.2% increase in theronine production [27].

However, most of the pumps involved in the export of the desired products are currently unknown. To facilitate screening of pumps involved, different screening approaches have been applied. One example is the bioinformatics-assisted screening approach where 43 heterologous pumps

were screened using competition-based strategy Figure 2-7A and B. Efflux pump-expressed strains were grown individually and then pooled in equal proportion. The pooled culture was grown in the presence of biofuel for 96 h (controls are without biofuel) and was serial diluted every 10–14h. At each dilution time point, plasmids of each efflux pump from the culture were isolated and quantified using a custom microarray. The strains expressing pumps that export biofuel have growth advantage, therefore, these strains dominated the culture after a few generation of growth. Using the screening method, a previously uncharacterized pump from *Alcanivorax borkumensis* was identified and the strain expressing this efflux pump increased the limonene production by more than 50% [35].

The ATP-binding cassette (ABC) exporters can export a diverse range of hydrophobic molecules such as lipids, drugs, and steroids. And ABC transporters are widely found in all five kingdoms (bacteria, cyanobacteria, diatoms, yeast, and algae) [99], thus they have the potential for cross-species application. Recently, a study screened 19 bacterial ABC transporters to enhance the export of carotenoids. It was proposed that the localization of heterologous carotenoids was akin to that of lipid A, thus ABC transporter could help to secrete carotenoids out of the inner membrane (Figure 2-7C). The more efficient transporters for zeaxanthin and canthaxanthin were *Salmonella enterica* Serovar *typhymurium MsbA* and *Escherichia coli MsbA*, respectively. When transporters were overexpressed, the secreted zeaxanthin and canthaxanthin (concentration

in the organic layer) improved ~2.4 fold and ~4.4 fold, respectively. However, the total yield (in organic layer plus inside the cells) did not improve much.

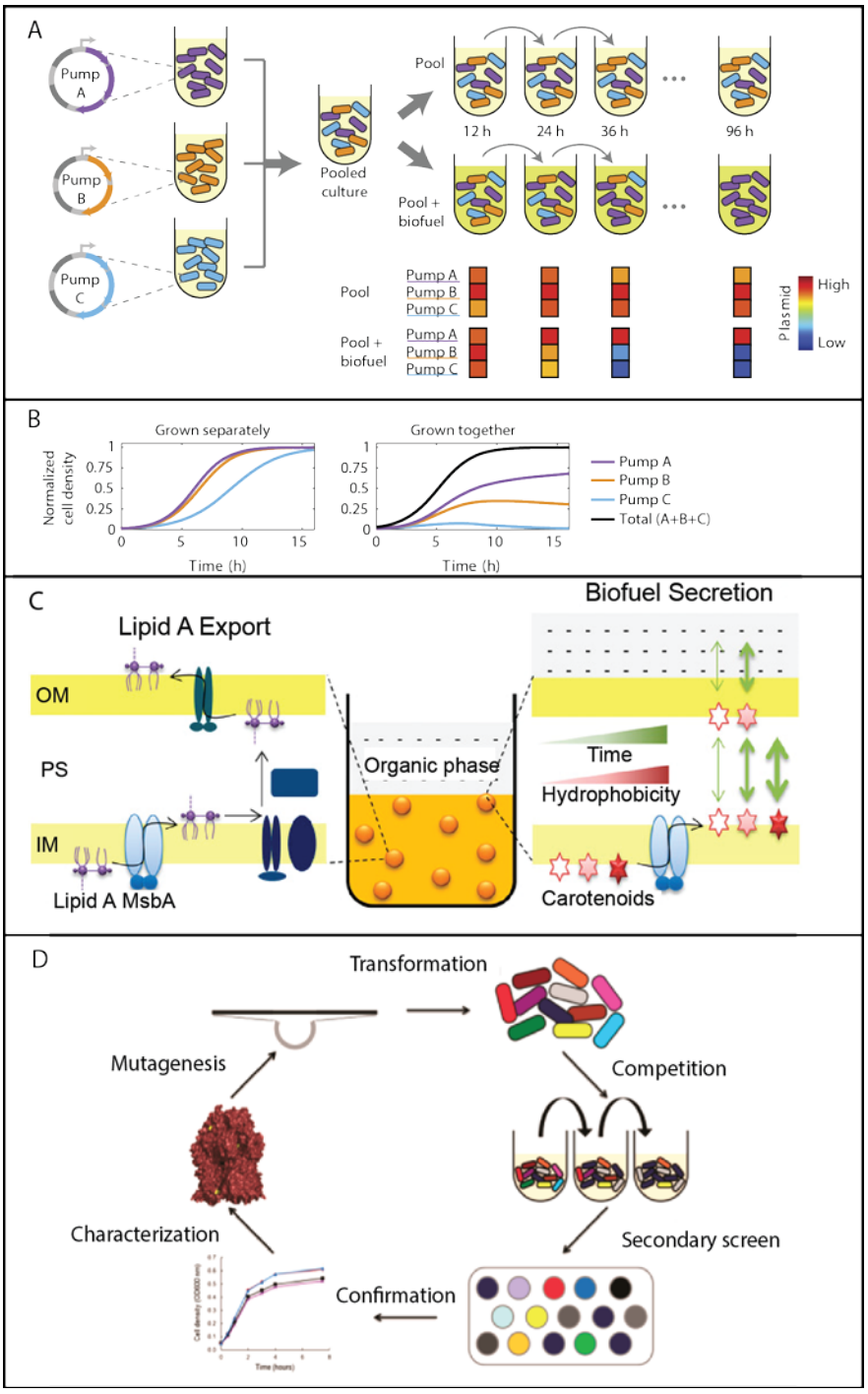


Figure 2-7. Screening approaches used for transporter engineering in literature. (A) Schematic of competition-based strategy [35]. Different strains with individual efflux pump were grown independently and pooled in equal proportion. The pooled culture was grown both with and without biofuel, with a few steps of dilution into fresh media. When plasmids isolated, the plasmids providing growth advantage became overrepresented in the culture. (B) Simulations of competitive exclusion [35]. Strains with growth advantage dominate the population during co-culture. In contrast, the growth rate of each strain won't be affected when grown separately. (C) Schematic representation of the biofuel secretion system [36]. The authors hypothesized that the export of carotenoids cross inner membrane (IM) to periplasmic space (PS) is similar to the lipid A using msbA. And the carotenoids partitions from the PS to organic layer through the out membrane (OM). (D) Directed-evolution approach used to generate tolerance-conferring variants of the AcrB efflux pump [34]. *E. coli* expressing the libraries of AcrB variants were subjected to growth competition in the presence or absence of 0.5% n-butanol. Following that, cells from both pools were plated and individual clones were tested for growth performance in n-butanol. After confirming AcrB variant-conferred growth, another round of mutation and selection was performed. Reproduced with permission of the copyright owner.

Screening can identify pump candidates for various products. However, the pumps may not be highly efficient. Thus, a directed evolution-based method was used to improve the efficiency of transporters towards certain compounds [33, 34]. Indeed, variants of *acrB* were found to confer improved efficiency of efflux of n-octane [33] and n-Butanol [34]. By sequencing, the mutated sites of the *acrB* variant were identified and correlated to the improved activities.

2.1.6 Kinetics modeling

Flux determination is one of the useful tools in metabolic engineering. Reliable flux estimation is very informative and can be used to identify the rate-limiting enzymes in a pathway [9]. Successful flux analysis is highly dependent on the availability of reliable (less noisy and relatively complete) experimental data. Isotopic tracer-based approach has been used to understand the flux distribution in the network [100, 101]. When

detailed information about the kinetics of metabolic reactions is available, we may be able to describe the dynamics in Equation 1-1.The crucial step of combining stoichiometric property with kinetic features is the search for appropriate function to represent *v*. There are three major categories to represent *v*, (1) mechanistically based functions (e.g., law of mass action, Michaelis–Menten law); (2) *ad hoc* approaches; and (3) different types of canonical models (e.g., Generalized Mass Action (GMA), BST and lin-log representations), and details about these approaches were described in the review [102].

The equation of mass action is shown in Equation 1-2, and the method is typically used to describe reaction networks consisting of elementary reactions. However, most biochemical reactions are not elemental but composites of several elementary equations [103].

S-system structures, a form of Biochemical Systems Theory (BST), is one of promising model in metabolic engineering. The model is constructed by approximating flux with products of power-law functions. The generic form of S-system formulation is shown in Equation 1-3. There are numerous publications using this model to understand gene regulatory networks [104-109] and numerous novel algorithms were developed to estimate the parameters [110-115]. For example, Liu *et al* [111] successfully used the S-system model to describe a batch fermentation process. Gautam *et al* [116] and Jia *et al* [117] developed new algorithms for the S-system model based on the NMR data from the *L. lactis* pathway.

$$\frac{dS}{dt} = N \cdot v, \qquad \text{Equation 2-1}$$

where *S* is a vector of metabolite concentrations and *v* is a vector of fluxes. The sstoichiometric matrix *N* is the core of stoichiometric models that show how metabolite concentrations change over time.

$$v = k\prod_{g=1}^{n} X_i^{g_i},\ i = 1,2,..,n, \qquad \text{Equation 2-2}$$

where *k* is the rate constant, which is always positive, g_i are kinetic orders which are non-negative integer numbers that reflect the numbers of molecules involved in the reaction, and X_i denotes the concentration or amount of a variable or variable pool and *n* is the total number of time dependent variables in the system.

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}},\ i = 1,2,..,n, \qquad \text{Equation 2-3}$$

where X_i represents a time-dependent variable (metabolite) and *n* denotes the number of variables in the system. The non-negative multipliers α_i and β_i are rate constants which quantify the turnover rate of the production or degradation, respectively. The real numbers g_{ij} and h_{ij} are kinetic orders that reflect the strengths of the effects that the corresponding variables X_j have with a given flux term.

Chapter 3. OPTIMIZATION OF GLUCOSE UPTAKE AND THE DXP PATHWAY FOR ISOPRENOID PRODUCTION

3.1 Introduction

Glucose is the most abundant and widely used carbon and energy source for microbial processes. In *E. coli*, glucose is transported into cytoplasm mainly through the carbohydrate phosphotransferase system (PTS). The PTS is composed of enzyme I (EI), histidine protein (HPr), sugar-specific enzyme IIA (EIIA) and enzyme IICB (EIICB), coded by *ptsHIcrr* operon and *ptsG* gene, respectively. A phosphoryl group is transferred from PEP to EI and HPr, then to EIIA, EIIB and EIIC. Finally, the phosphoryl group is transferred to glucose by the integral membrane permease EIIC [81]. In *E. coli*, half of the PEP produced from glucose is further metabolized into pyruvate during the glucose uptake by the PTS. PEP is an important precursor to produce certain industrially interesting compounds, such as succinate [86, 118], aromatic compounds via shikimate-pathway [32].

Inactivation of the PTS has been used to improve the yield of desired products via increasing the availability of intracellular PEP [32, 86, 118]. However, disruption of the PTS resulted in decreased glucose uptake and extremely slow cell growth rate (from 0.7 to 0.1 hr^{-1}) when glucose was the only carbon source used in fermentation [38]. To address this issue,

adaptive evolution was used to select a fast-growing strain [32]. Transcriptome analysis of these fast-growing strain showed that transcriptional levels of the gene *galP* (coding for galactose permease), *glk* (coding for glucose kinase) and *pgi* (coding for phosphoglucose isomerase) were increased when compared to the wild-type strain, suggesting that these genes were important to enhance glucose uptake [91, 119]. Therefore, *galP* and *glk* were overexpressed to enhance the growth rate of microbes growing in glucose medium [91, 120] and the yields of valuable products such as aromatics [121], recombinant protein [122] and ethanol [90]. Combinatorial modulation of *galP* and *glk* gene expressions have been used to improve growth rates [120], but it is unknown if such an approach will increase the production of isoprenoids.

Optimization of the activities and expressions of enzymes in a pathway is essential to improve product yield. Optimized pathway can minimize the accumulation of intermediate metabolites with potential cytotoxicity and thus decrease the metabolic burdens and increase the final product titer [46, 96]. Multivariate-modular approach to balance metabolic pathway has been successfully used to systematically search for the optimal conditions for increasing productivity [13]. Tuning plasmid copy number or engineering different types of promoters [123] has been used to control and balance the expression of multiple genes. However, these reported approaches are time-consuming and resource demanding, often do not cover the whole search space for identifying optimal conditions which

results in suboptimal productivity. To address this unmet need, our group has developed another approach, the Univariant Extrinsic Initiator Control System for microbes (μ-UNEICS), which controls multiple modules (metabolic pathway) simultaneously using T7 variant promoters with different transcriptional efficiencies [124].

In this study, to test the hypothesis that AD production from glucose was limited by the availability of PEP. PEP synthase (*ppsA*) was overexpressed and this resulted in the improvement of AD yield by increasing the intracellular concentration of PEP. Subsequently, the PTS was deleted and the glucose uptake pathway comprising the GGS was overexpressed. Unlike the use of PTS, the use of the GGS did not divert PEP for other side reactions and can be directly shuttled to improve AD production. Furthermore, systematic pathway optimization using μ-UNEICS approach to globally optimize the four modules enabled the production of high amounts of AD.

3.2 Results

3.2.1 Increasing AD yield by *ppsA* overexpression

In this study, the native DXP pathway in *E. coli* was used to produce AD (Figure 3-1). The DXP pathway starts from pyruvate and glyceraldehyde-3-phosphate (GAP). It was reported that GAP was limiting in the biosynthesis of lycopene via the DXP pathway, and redirecting flux from pyruvate to GAP can enhance the lycopene yield [23]. To test the

hypothesis that overexpression of *ppsA* can help in the production of AD by converting pyruvate back into GAP through PEP, endogenous *ppsA* was overexpressed in MG1655 DE3 strain with T7 promoter. High expression of *ppsA* is toxic and negatively affected the growth of *E. coli*, thus moderate amount of the inducer IPTG was added to induce the expression of *ppsA* (IPTG concentration was not greater than 0.03 mM). As shown in Figure 3-2, the AD titer (mg/L) increased about 2 fold (from 10 mg/L to 20 mg/L) when *ppsA* was overexpressed in all the conditions (different concentrations of IPTG). And the specific yield of AD increased from about 2.5 mg/L/OD to around 3.5 mg/L/OD. The result strongly suggested that increasing intracellular PEP concentration was likely to be beneficial to the AD production, since *ppsA* can catalyze pyruvate back to PEP.

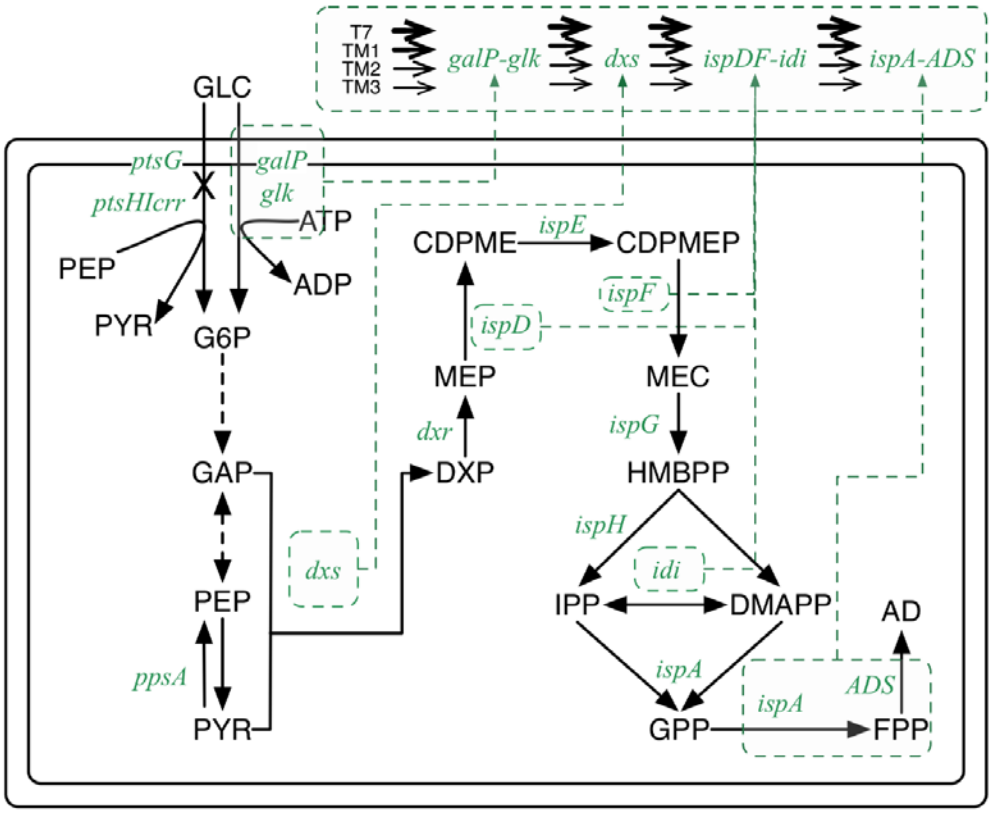


Figure 3-1. Global control and balancing of metabolic pathway for AD production. T7, TM1, TM2 and TM3 promoters were used to fine controlling the expression of the four different modules (*galP-glk*, *dxs*, *ispDF-idi* and *ispA-ADS*). The transcriptional efficiencies of TM1, TM2 and TM3 were 92%, 37% and 16% of that of T7 promoter, respectively. The abbreviations for metabolites in the figure: glucose (GLC), glucose-6-phosphate (G6P), glyceraldehyde 3-phosphate (GAP), phosphoenolpyruvate (PEP), pyruvate (PYR), 1-deoxy-D-xylulose 5-phosphate (DXP), 2C-methyl-D-erythritol 4-phosphate (MEP), 4-diphosphocytidyl-2C-methyl D-erythritol (CDPME), 4-diphosphocytidyl-2C-methyl D-erythritol 2-phosphate (CDPMEP), 2C-methyl-D-erythritol 2,4-diphosphate (MEC), hydroxylmethylbutenyl diphosphate (HMBPP), isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), geranyl diphosphate (GPP), farnesyl pyrophosphate (FPP), amorpha-4,11-diene (AD). The abbreviations for enzyme-coding genes in the figure: PTS enzyme IIBC (*ptsG*), histidine protein (*ptsH*), PTS enzyme I (*ptsI*), PTS enzyme IIA(*crr*), galactose permease (*galP*), glucose kinase (*glk*), phosphoenolpyruvate synthetase (*ppsA*), DXP synthase (*dxs*), DXP reductase (*dxr*), CDPME synthase (*ispD*), CDPME kinase (*ispE*), CDPMEP synthase (*ispF*), HMBPP synthase (*ispG*), HMBPP reductase (*ispH*), IPP isomerase (*idi*), farnesyl pyrophosphate (*FPP*) synthase (*ispA*) and *ADS* synthase (*ADS*).

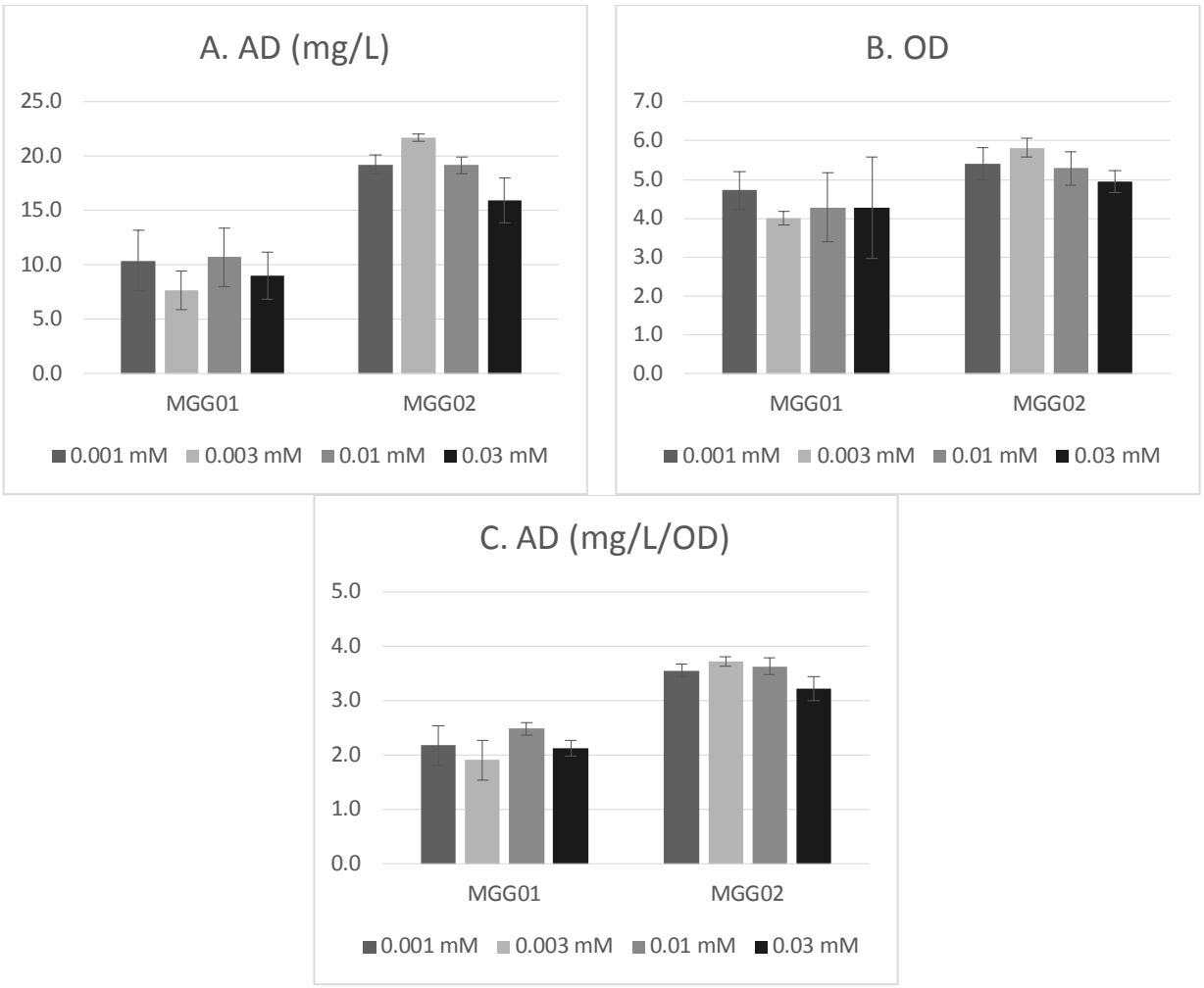
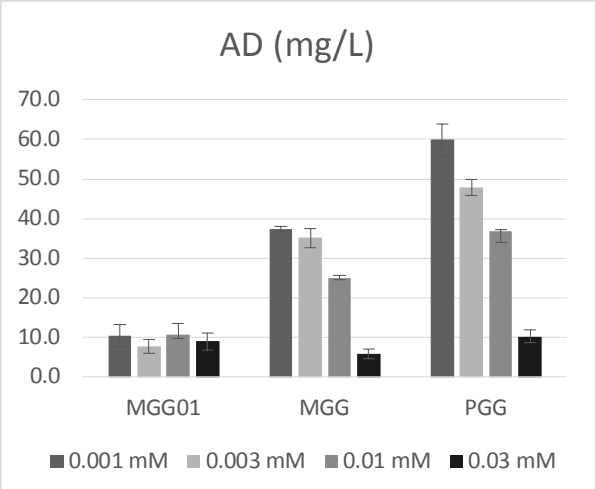


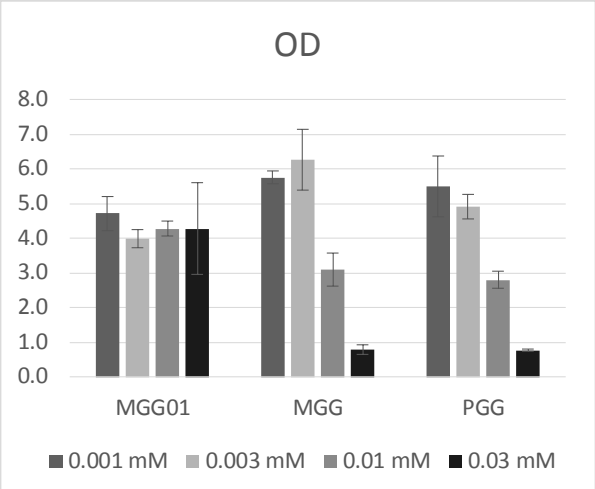
Figure 3-2. Effect of enhanced expression of *ppsA* on the AD production. MGG01 strain is the control without overexpressing *ppsA*, while *ppsA* is overexpressed in MGG02 strain. (A) AD titer (mg/L). (B) Cell density of AD-producing strains. (C) Specific yield of AD (mg/L/DO).

3.2.2 Increasing AD yield by overexpressing *galP* and *glk*

PEP enrichment by overexpressing *ppsA* increased AD production. And glucose uptake via the PTS consumes half of PEP produced from glucose. This inspired us to study the effects of using alternative glucose uptake pathways on the AD production. Unlike the PTS which uses PEP as phosphate donor, the GGS transports and phosphorylates glucose using ATP as phosphate donor. As shown in Figure 3-3, overexpression of

the GGS in the wild type strain (MGG) resulted in an increase in AD titer by more than 3-fold (from 10 mg/L to 37 mg/L) as compared to wild type strain (MGG01). With overexpression of the GGS in the PTS knockout mutant (PGG), AD titer was further increased by 70% to ~60 mg/L. And the specific yield of AD increased from about 7 mg/L/OD (MGG) to around 12 mg/L/OD (PGG). The results indicated that the route of glucose uptake did affect the AD production, and the ATP-dependent GGS (PGG) was better than the PEP-dependent PTS (MGG01) for the AD production in *E. coli*. Noteworthy, the cell density was similar for the three strains in most of the conditions except that the cell density of MGG and PGG was exceptionally low when high concentration IPTG was used. This suggests that the high expression of the GGS could be toxic to cells.





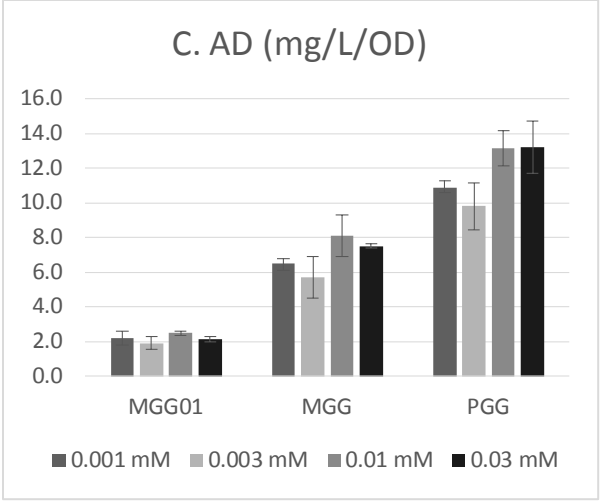
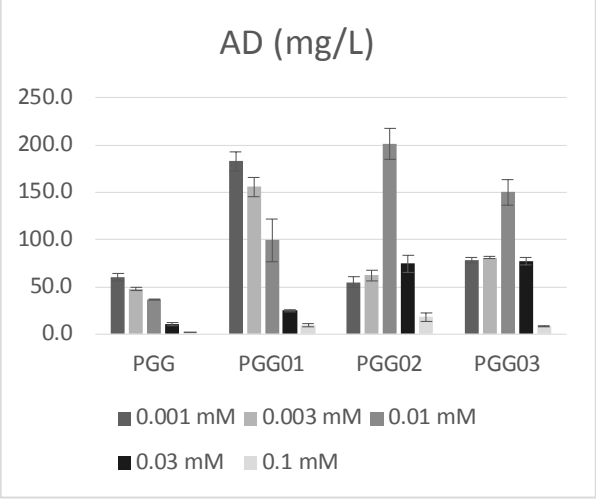


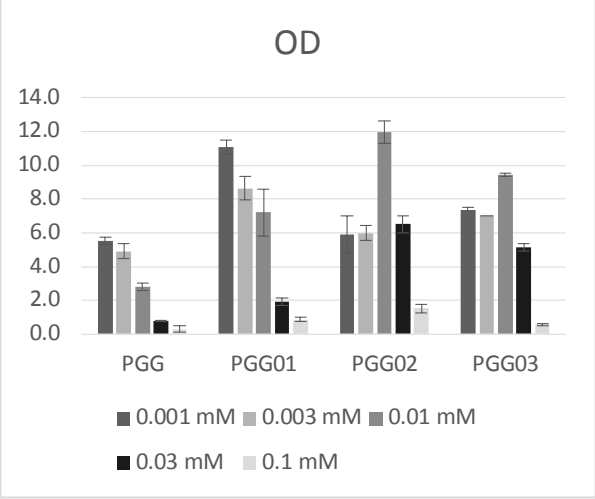
Figure 3-3. The effect of glucose uptake pathway (the GGS and/or the PTS) on AD production.
MGG01 is the wild type strain which used the PTS to uptake glucose. MGG is the GGS-overexpressed strain which used both the PTS and the GGS to uptake the glucose. While PGG is the PTS knockout mutant with the GGS overexpressed, thus the glucose was uptaken solely by the GGS. (A) AD titer (mg/L). (B) Cell density of different strains. (C) Specific yield of AD (mg/L/DO).

3.2.3 Optimization of the expression of the GGS

The replacement of the PTS with the GGS increased the AD production. Because the high expression of the GGS resulted in slightly better specific production but was shown to be toxic to cells, it is worth investigating whether tuning the expression of the GGS can further increase the cell density and thus the AD production. Therefore, the T7 promoter of the GGS in strain PGG was replaced with weaker promoters, TM1, TM2 or TM3, where transcriptional strength was about 92%, 37% and 16% of the strength of T7 promoter, respectively. TM1, TM2 and TM3 promoters were generated as described in the patent [125]. In the strains PGG, PGG01, PGG02 and PGG03, all the genes *dxs idi*, *ispA* and *ADS* were controlled by T7 promoter. Operon *galP-glk* was controlled by T7, TM1, TM2 and

TM3 in strain PGG, PGG01, PGG02 and PGG03, respectively. As a result, the AD titers in strains PGG01 and PGG02 increased by more than 3 fold as compared to that in PGG (with T7), from 60 mg/L to 200 mg/L (Figure 3-4). In addition, the specific yield was also slightly increased significantly from ~13 mg/L/OD to ~18 mg/L/OD, which means that better tuning of the GGS benefits both cell density and the performance of the DXP pathway. However, similar to strain PGG, although the expression levels of the GGS was decreased using the weaker promoters, the cell density of strain PGG01, PGG02 and PGG03 was still very low when higher concentration of IPTG was used.





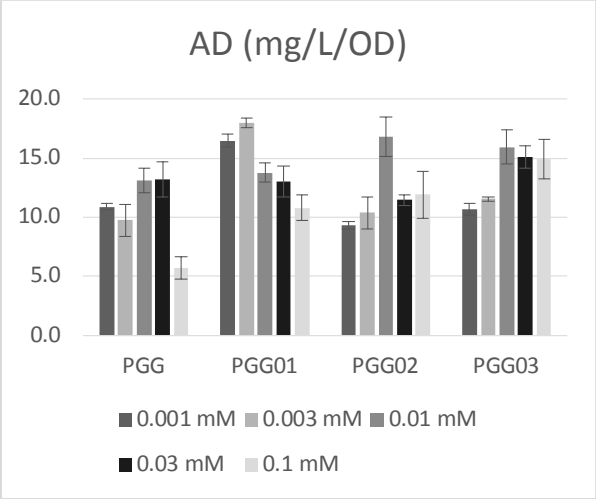


Figure 3-4. Further increase of AD production based on the model prediction. In strain PGG01, PGG02 and PGG03, the *galP-glk* was placed under control of TM1, TM2 and TM3 promoter, respectively. And all the others genes were placed under the control of T7 promoter. (A) AD titer (mg/L). (B) Cell density of different strains. (C) Specific yield of AD (mg/L/DO).

Table 3-1. Reported AD production in literature

#	AD titer	AD yield	AD Productivity	Host	Literature
	g/L	Cmol%	g/L/h		
1	0.02		0.0004	Bacillus subtilis	Zhou, K., et al. (2013)
2	27.4		0.17	Escherichia coli	Tsuruta, H., et al. (2009)
3	41	17.0%	0.35	Saccharomyces cerevisiae	Westfall, P. J., et al. (2012)
4	4	3%	0.1	Escherichia coli	This study

3.2.4 Differential control of gene expressions in the multi-modules encoding enzymes in the glucose uptake, DXP and AD synthesis pathways

A two-module optimization of the upstream DXP pathway and heterologous downstream terpenoid–forming pathway increased Taxol precursor titer significantly [13]. As the tuning of the expression of the

GGS increased the AD yield, we hypothesized that the global pathway optimization of the GGS, the DXP and AD synthesis pathways may further increase AD production. The genes *dxs*, *ispDF* and *idi* were identified to encode enzymes that are rate-limiting in the DXP pathway [19], and the overexpression resulted in improved yields of carotenoids [16, 19] and Taxol precursor [13]. In addition, *ADS* was reported as the rating-limiting step for the AD production [44, 96]. All these rate-limiting enzymes with glucose-uptake enzymes were divided into four distinct modules. The first is *galP-glk*, which regulates the uptake of glucose. The second is *dxs*, which diverts the carbon flux from central metabolic pathway into the DXP pathway. The third is *ispDF-idi*, in this module *ispDF* facilitates the conversion of intermediates of DXP pathway and *idi* is critical for the balancing of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). The last is *ispA-ADS*, and *ispA* and *ADS* are the last two steps which carry out the conversion of IPP and DMAPP into AD via farnesyl pyrophosphate (FPP).

All the four modules were controlled simultaneously under the T7 or T7 variant promoters (Figure 3-1). There would be 4^4=256 constructions in total if we did the full combination and there would be 256x4=1024 conditions with 4 different concentration of IPTG. To be cost and time effective, a limited combination of constructions was chosen based on the following rationale. To cover a broad range, each module was controlled by three different promoters (TM1, TM2 and TM3). And each promoter for

each module was repeated at least three times (for example, TM1 for the first module was used in four different constructions). As a result, 14 constructions were chosen (Figure 3-5D). As shown in Figure 3-5A, through the fine tuning of the four modules, AD titer increased significantly from about 0.36 mg/L (strain P1717 with 0.03 mM IPTG) to 68.9 mg/L (strain P2121 with 0.01 mM IPTG). However, the highest titer (~68.9 mg/L) was much lower than that of the PGG02 (~200 mg/L), and the highest specific production (~12.0 mg/L/OD) was also lower than that of PGG02 (~16.9 mg/L/OD). And a noteworthy phenomenon was that cell density decreased significantly when the inducer IPTG concentration was 0.03 mM, which suggested cells were under stress when more IPTG was used.

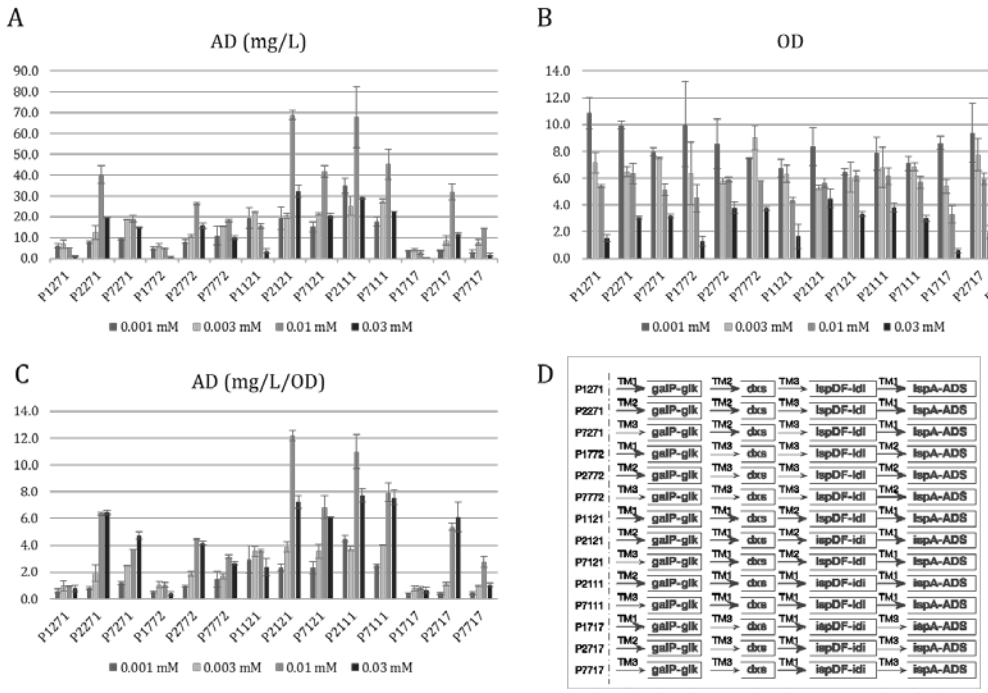


Figure 3-5. Optimization of AD production by controlling 4 modules.
TM1, TM2 and TM3 promoters were used to control the expression of the four modules: *galP-glk*, *dxs*, *ispDF-idi* and *ispA-ADS*. The strains were named according to their sequence of promoters, *e.g.*, P1271 refers to a construction in which the four modules

galP-glk, *dxs*, *ispDF-idi* and *ispA-ADS* are under the control of TM1, TM2, TM3 and TM1 promoter, respectively. (A) AD titer (mg/L). (B) Cell density of different strains. (C) Specific yield of AD (mg/L/DO). (D) Illustration of different constructions (The thickness of arrows represents the relative strength of promoters).

3.2.5 Regression analysis of modular controlling

To identify the global optimal constructions and understand the behavior of the four-module-controlling system, a linear regression model was applied to the experimental data in Figure 3-5. The AD titers in all the different conditions were plotted based on the relative expression (Table 3-1) of each module (Figure 3-6A). The relative expression of each module was calculated from the strength of promoters and copy numbers of plasmids, as described in materials and methods. For the ease of visualization and analyses, the data were shown as heat maps using a linear regression model (Figure 3-6B). Based on the calculations, the optimal ranges for the first (*galP-glk*) and second (*dxs*) modules were successfully covered by the 14 chosen constructions. However, for the third (*ispDF-idi*) and fourth (*ispA-ADS*) modules, the optimal conditions may not have been suitably covered in the collection of constructs. As shown in Figure 3-6B, the optimal ranges of relative expression levels for the four module *ispDF-idi* and *ispA-ADS* were 0.05~0.2, 0.2~0.4, below 0.2 and above 0.4, respectively. A relatively low expression of first module (*galP-glk*) were beneficial to AD production. This was reasonable, because excessive expression of the first module inhibited cell growth and

very low expression was not sufficient to provide enough carbon sources thus decreased the final yield. Low expression level of the third module was beneficial to AD production. Further experiments indicated that the gene *ispDF* in the third module were less important (data not shown) and the overexpression of *ispDF* may overburden the cells. This could be due to the enzymes encoded by *ispDF* were more soluble [60] and possessed higher activities than the other rate-limiting steps thus this may require relatively lower expression levels. Higher expression levels for the second module and the fourth module were required. The expressed *dxs* gene in the second module was found to be less soluble than *ispDF*, and hence to achieve higher activity of *dxs*, a higher expression was required. The fourth module also required higher expression level as the activity of heterologous *ADS* from *Artemisia annua* is known to be poor in *E. coli* [126, 127]. Hence, a higher expression of this module was required.

Table 3-2. Relative expression levels of each module in the 14 constructions

#	Strain	Relative expression level of each module			
		galP-glk	*dxs*	*ispDF-idi*	*ispA-ADS*
1	P1271	0.17	0.21	0.09	0.52
2	P2271	0.08	0.24	0.10	0.58
3	P7271	0.04	0.25	0.11	0.61
4	P1772	0.31	0.16	0.16	0.37
5	P2772	0.15	0.20	0.20	0.45
6	P7772	0.07	0.22	0.22	0.50
7	P1121	0.12	0.37	0.15	0.37
8	P2121	0.05	0.39	0.16	0.39
9	P7121	0.02	0.41	0.16	0.41
10	P2111	0.04	0.32	0.32	0.32
11	P7111	0.02	0.33	0.33	0.33
12	P1717	0.20	0.10	0.59	0.10
13	P2717	0.09	0.12	0.67	0.12

14	P7717	0.04	0.12	0.71	0.12

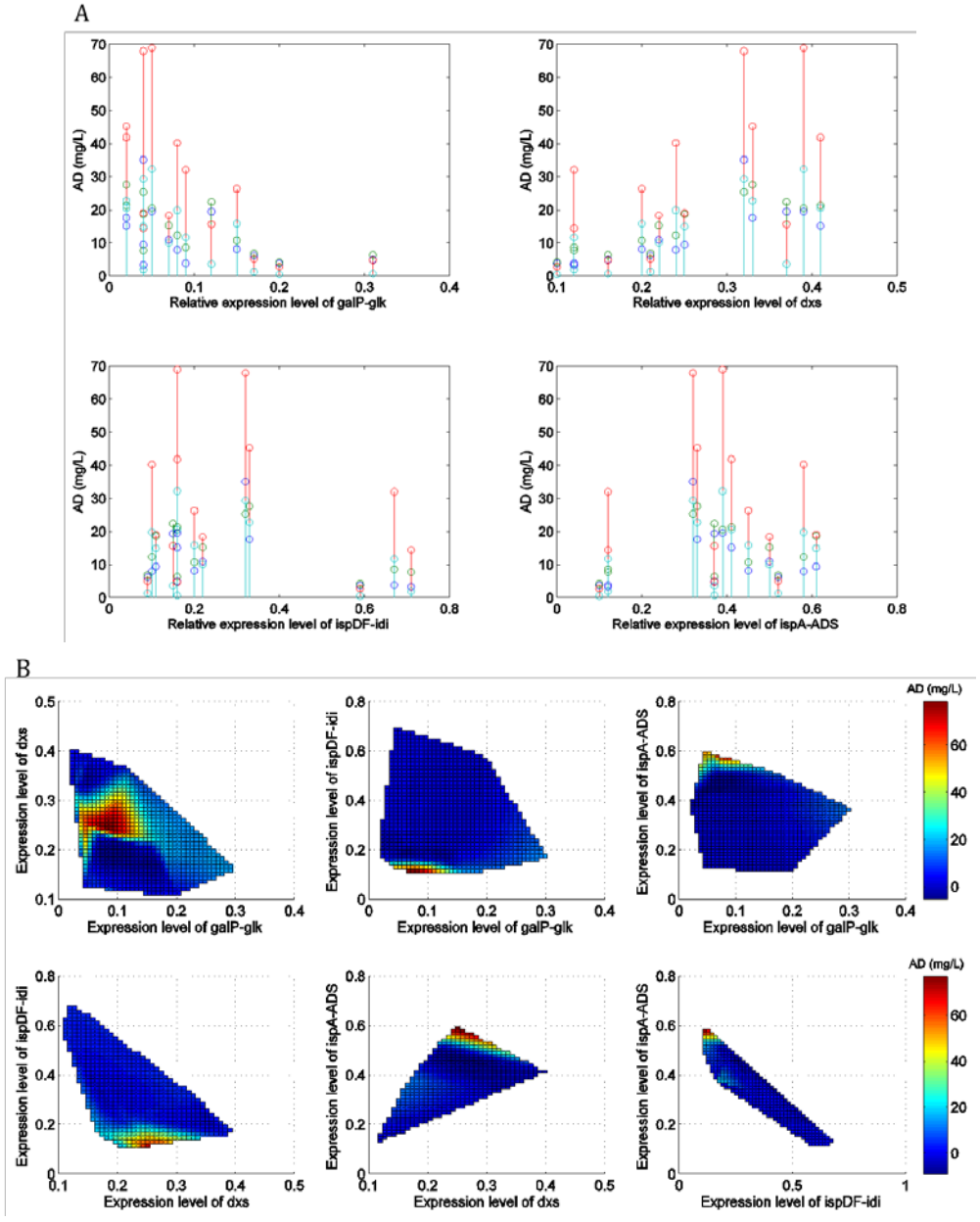


Figure 3-6. Regression and prediction of the optimal expression levels of the four modules.
(A) AD titers (mg/L) versus different expression levels of each module. (B) Regression of AD titers (mg/L) on the relative expression levels of every two modules. Relative expression levels were calculated based on the promoter strength and plasmid copy number (described in Materials and Methods).

3.3 Discussion

PEP synthase (*ppsA*) catalyzes pyruvate to PEP, and it has been used to enhance the production of lycopene (a C40 isoprenoids) [22, 23]. Consistent with the literature results, the AD production was predictably increased when *ppsA* was overexpressed (Figure 3-2). This inspired us to explore the productivity of the PTS knockout strain, as inactivating the PTS can potentially redirect PEP produced from glucose to the synthesis of downstream isoprenoids [32, 121]. In *E. coli*, the PTS is the main glucose transport system and is also responsible for carbon catabolite repression (CCR) and chemotaxis [90]. Thus deletion of the PTS is also pleiotropic and its mutant has several distinct characteristics over the wild type strain, such as enhancement of intracellular PEP concentration, coexistence of glycolytic and gluconeogenic pathway and the simultaneous utilization of different carbon sources [119]. The PTS mutant is able to co-utilize xylose and glucose [77] and produces higher titers of various compounds such as recombinant proteins [83], aromatics [32], lysine [84], L-phenylalanine [85] and succinate [86]. However, the uptake rate of glucose in PTS mutant is very slow and thus the mutant grows very slowly. Therefore, in this study, the GGS was overexpressed in the PTS deficient mutant to enhance the glucose transport.

In line with our hypothesis that the production of AD was limited by the availability of PEP, the use of GGS (consuming ATP) for glucose uptake was more beneficial for AD production than that of the PTS (consuming

PEP) (Figure 3-3). However, it was observed that higher expression of the GGS improved specific yield of AD but inhibited cell growth in strain PGG and MGG (Figure 3-3). Therefore, a fine-tuning of the expression of the GGS was tested, as expected, this indeed improved the AD titer through improving the biomass. In addition, the specific production (mg/L/OD) was also improved by 38%, resulting in an increase in AD titer by more than 3 fold (Figure 3-4).

Production of taxadinene, an isoprenoid, has been shown to be significantly improved by balancing the DXP pathway and taxadiene synthesis pathway [13]. Considering the benefit of using the GGS glucose uptake pathway, we hypothesized that global pathway balancing of glucose uptake pathway, DXP pathway and AD synthesis pathway may further improve the production. Although the AD yield varied by more than 190 fold in the panel of constructs and conditions, the maximal yield was 69 mg/L (Figure 3-5) which was slightly higher than the yield when using PGG strain (60 mg/L) and significantly lower than PGG01 (182 mg/L) or PGG02 (201 mg/L). It is noteworthy that the cell density declined significantly when inducer dosage (IPTG concentration) was increased to 0.03mM. The decreased cell density at high induction level is likely to be due to the stress of high expression of galactose permease and glucose kinase [120], the toxicity of accumulated high-concentration of glucose-6-phosphate or inhibition from high concentration IPTG [90], and further study is required to address this issue.

To better understand the behavior of the four-module-controlling system, we analyzed the experimental data of the four-module system using a linear regression model. Based on the regression analysis (Figure 3-6), the optimal ranges of expression level for the first and second module were within the panel of constructs. However, none of the constructs managed to capture the optimal range of expression levels for the third and fourth modules. More importantly, combining the regression model of the fourth-module constructs with the results of PGG01, PGG02 and PGG03 in Figure 3-4 (in which *ispDF* were not overexpressed), we concluded that the *ispDF* in the second module was less important in contributing to AD production. Excluding *ispDF* from the system is beneficial to the system possibly because part of limited resources were wasted for producing *ispDF*. Furthermore, it was predicted in the regression model that relatively lower expression of *galP-glk* and higher expression of *dxs* and *ispA-ADS* were beneficial for AD production. And all of these predictions were just consistent with the expression levels in strain PGG01 (TI1-*galP-glk*, T7-*dxs-T7-idi-ispA-ADS*) and PTS02 (TI2-*galP-glk*, T7-*dxs-T7-idi-ispA-ADS*). That explained the observation that PGG01 and PGG02 produced the higher titer of AD than all the chosen four-module-controlling constructs.

3.4 Conclusion

In this study, the increased availability of PEP was shown to be important in improving AD production via the DXP pathway. To increase intracellular

PEP concentration, the PEP-consuming PTS was replaced with the GGS that uses ATP for glucose uptake, consequently, this strategy increased AD titer by ~ 6 fold. Furthermore, fine tuning of the expression levels of the GGS further improved AD titer by another 3 fold by increasing the biomass and the specific yield. A global and systematic optimization was performed to simultaneously control the expression of glucose uptake pathway, DXP pathway and AD synthesis pathway. The optimal ranges for each module were analyzed and provided possible routes to improve the productivity of AD.

3.5 Materials and Methods

3.5.1 Bacteria strains and plasmids

E. coli K12 MG1655 ΔrecAΔendA DE3 was used as the parental strain [13]. *MG1655 ΔPTS::FRT* (PTS- for short) strain was obtained by replacing operon *ptsHIcrr* with a kanamycin-resistant gene using the λ Red recombination system [24]. And the kanamycin resistant gene was subsequently excised by the Flp/FRT site-specific recombination system [128]. Chromosomal deletions were verified by testing for antibiotic markers and colony PCR analysis.

Plasmid pET-*galP-glk* was constructed from three parts, pETaa vector, gene *galP* and gene *glk* by Cross-lapping In Vitro Assembling (CLIVA) method using I-pETR(galPF), I-galPF(pETR), I-galPR(glkF), I-glkF(galPR), I-pETF(glkR) and I-glkR(pETR) (Table 3-2) [129]. Plasmids pETA-

TM1/2/3-*galP-glk* were constructed from pET-*galP-glk* by mutating T7 promoter to TM1/TM2/TM3 promoter using the oligo TM1, TM2 and TM3. Plasmid p15A-spec-T7-*ppsA* was constructed from two parts 15A-spec vector and *ppsA* gene by CLIVA method using I-*ppsA*(t7pr)-F, I-*ppsA*(t7tf)-R, I-T7P(-)-R and I-T7T(-)-F (Table 3-2). Plasmid pBAD-galP-glk was constructed from pBAD-B (Invitrogen) and pET-*galP-glk* by CLIVA methods using I-gg-F(pBADR), I-gg-R(pBADF), I-pBAD-F(ggR) and I-pBAD-R(ggF).

All the information about the strains, plasmids and primers is in Table 3-2 and Table 3-3.

3.5.2 Media and culture conditions

3.5.2.1 Defined media composition

The defined media without carbon contained 4 g/L $(NH_4)_2HPO_4$, 3.50 g/L KH_2PO_4 and 9.28 g/L K_2HPO_4, 1.7 g/L citric acid, 0.5 g/L $MgSO_4$, 2.5 mg/L $CoCl_2 \cdot 6H_2O$, 15.0 mg/L $MnSO_4 \cdot 4H_2O$, 1.5 mg/L $CuSO_4 \cdot 2H_2O$, 3 mg/L H_3BO_3, 2.5 mg/L $Na_2MoO_4 \cdot 2H_2O$, 13 mg/L $Zn(CH_3COO)_2$, 60 mg/L Fe(III) citrate and 8.4 mg/L EDTA. And 20 g/L glucose was added into the defined media before use.

All the AD titers reported here were based on the strains grown in the defined media supplemented 20 g/L with glucose.

3.5.2.2 2xpy media composition

The 2xpy media contained 20 g/L peptone, 10 g/L yeast extract and 10 g/L NaCl. The pH of the media was adjusted to 7.0 and the media was autoclaved at 121 °C for 20 mins.

3.5.3 Culture conditions

New transformed strains were rescued and grown in 2xpy media overnight, and 1% (v/v) overnight grown cell culture was inoculated into 1 mL defined media in a 14mL BD Falcon™ tube. Two hundred microliter dodecane with 100 mg/L beta-caryophyllene was added on top of the culture. Cells were grown at 28°C with 300-rpm shaking speed, and induced with IPTG (for T7 or mutated T7 promoter) or L-arabinose (for araBAD promoter) when OD600 reached the range of 0.5~1.0. The media were selectively supplemented with 100 mg/L ampicillin, 34 mg/L chloramphenicol, 50 mg/L kanamycin and 100 mg/L spectinomycin to maintain the pET-11a derived plasmids, pAC-LYC derived plasmids, pETK-lacI and p15A-spec-T7-*ppsA* respectively.

3.5.4 Quantification of AD

Amopha-4,11-diene was extracted by diluting 5 µL dodecane phase into 495 µL ethyl acetate and analyzed on an Agilent 7980A gas chromatograph equipped with an Agilent 5975C mass spectrometer (GC/MS). Beta-caryophyllene (Sigma-Aldrich) was used as an equivalent and the GC-MS condition was essentially as the method reported [130].

3.5.5 Calculation of relative expression of modules

Relative expression level of each module was calculated based on the promoter strength and plasmid copy number. It is calculated based on the following equation.

$$E_i(\mathrm{C}) = \frac{M_i \cdot N_i}{\sum_{j=1}^{4} M_j \cdot N_j}$$

Where, E_i is the relative expression level of the module *i* in the construction C, M_i is the promoter strength of module *i* and the N_i is the plasmid copy number of module *i*.

For example, the strengths of TM1, TM2 and TM3 were 92%, 37% and 16% of that of T7 promoter, respectively; the first module was on the pAC vector (p15a origin) with around 10 copies, and the other three modules were on pET vector (pBR322 origin) with about 30 copies, respectively. Therefore, the relative expression level of the first module (*galP-glk*) in strain P1271 was

$$E_1(\mathrm{P1271}) = \frac{92\% \times 10}{92\% \times 10 + (37\% + 16\% + 92\%) \times 30} = 0.17$$

Table 3-3. Strains, plasmids and primers used in this study

Strains/Plasmids/Primers	Description/sequence	References
	Strains	
MG1655 DE3	*E. coli* K12 MG1655 ΔrecAΔendA DE3	[13]
MG1655 DE3 ΔPTS::FRT	MG1655 DE3 derived, ΔpstHIcrr::FRT	This study
	Plasmids	
p15A-spec-T7-*ppsA*	pAC-LYC derived [131], replacing crtEBI with *ppsA* under the control of T7 promoter, and chloramphenicol resistant gene was replaced with spectinomycin resistant gene.	This study
pETA-galP-glk	pET-11a (Novagen) derived, *galP-glk* operon under the control of T7 promoter.	This study

pETA-TM1-*galP-glk*	pETA-*galP-glk* derived, mutating T7 promoter to TM1 promoter, which is 92% strength of T7 promoter.	This study
pETA-TM2-*galP-glk*	pETA-*galP-glk* derived, mutating T7 promoter to TM2 promoter, which is 37% strength of T7 promoter.	This study
pETA-TM3-*galP-glk*	pETA-*galP-glk* derived, mutating T7 promoter to TM3 promoter, which is 16% strength of T7 promoter.	This study
pBAD-galP-glk	pBAD-B (Invitrogen) derived, *galP-glk* operon were put under the control araBAD promoter.	This study
pACM-T7-*dxs*-T7-*ADS*-ispA-*idi*	pAC-LYC derived, *dxs-ADS-ispA-idi* operon is under the control of T7 promoter.	This study
pETK-lacI	pET-11a derived, ampicillin resistant gene was replaced with kanamycin resistant gene using CLIVA.	This study
pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF* and *ADS* are under the control of TM2, TM3 and TM1 promoter, respectively.	[RY]
pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF* and *ADS* are under the control of TM3, TM3 and TM2 promoter, respectively.	[RY]
pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF* and *ADS* are under the control of TM1, TM2 and TM1 promoter, respectively.	[RY]
pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF* and *ADS* are under the control of TM1, TM1 and TM1 promoter, respectively.	[RY]
pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF* and *ADS* are under the control of TM3, TM1 and TM3 promoter, respectively.	[RY]
	Primers	
TM1 promoter	TAATACGACTCACTAATGGGGA	This study
TM2 promoter	TAATACGACTCACTCGAGGGGA	This study
TM3 promoter	TAATACGACTCACTATAGAAGA	This study
I-pETR(*galP*F)	AGCGTCAGGCATA*TGTATATCTCCTTC*TTAAAGTT	This study
I-*galP*F(pETR)	AAGGAGATATA*CATATGCCTGACGCT*AAAAAAC	This study
I-galPR(glkF)	GATCTGCCTCCT*TTAATCGTGAGCG*CCTA	This study
I-glkF(galPR)	CGCTCACGATTA*AAGGAGGCAGATC*AAATGACAAAGTATGCA TTAGTC	This study
I-pETF(*glk*R)	GGTCACATTCT*GTAAAAGCTGAGTTGGC*TGC	This study
I-*glk*R(pETR)	GCCAACTCAGCTT*TTACAGAATGTGACC*TAAGG	This study
I-gg-F(pBADR)	ATTAACCA*TGCCTG*ACGCTAAAAAAC	This study
I-gg-R(pBADF)	AACAGCCT*TACAGA*ATGTGACCTAAGG	This study
I-pBAD-F(ggR)	TCTGTAAG*GCTGTT*TTGGCGGATGAG	This study
I-pBAD-R(ggF)	CAGGCATG*GTTAAT*TCCTCCTGTTAGCCC	This study
I-*ppsA*(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*TTTCTCAAACCGTTCATTTA	This study
I-*ppsA*(t7tf)-R	CCAACTCA*GCTTCCTTT*AACGCAGGATGTCTGTGAA	This study
I-T7P(-)-R	AACAAAATTATT*TCTAGAGGG*GAA	This study
I-T7T(-)-F	AAAGGAAG*CTGAGTTGG*CTG	This study

* The oligos asterisked are phosphorothioate-modified oligos used for CLIVA cloning [129].

Table 3-4. Name list for strains carrying different plasmids used in this study

Strain	Plasmid	Short Name	Reference
MG1655 DE3 ΔPTS::FRT	pETA-*galP-glk* and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	PGG	This study
MG1655 DE3	pETA-*galP-glk* and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	MGG	This study
MG1655 DE3	pETK-lacI and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	MGG01	This study
MG1655 DE3	p15A-spec-T7-*ppsA*, pETK-lacI and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	MGG02	This study

MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	PGG01	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	PGG02	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-T7-*dxs*-T7-*ADS-ispA-idi*	PGG03	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk*, pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P1271	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk*, pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P2271	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk*, pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P7271	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk*, pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P1772	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk*, pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P2772	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk*, pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P7772	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk*, pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P1121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk*, pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P2121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk*, pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P7121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk*, pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	P2111	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk*, pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	P7111	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk*, pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P1717	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk*, pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P2717	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk*, pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P7717	This study

Chapter 4. STATISTICAL MEDIA OPTIMIZATION

4.1 Introduction

Optimization of growth medium is an effective method to increase cell density and production of recombinant proteins [132, 133] and bulk chemicals [134, 135]. However, growth medium optimization has not been systematically investigated in metabolic engineering of microbes for the production of isoprenoids. A variety of standard media are commonly used for isoprenoid production without major modifications, including both complex media (LB [19, 54], 2xYT [136, 137], TB [96, 138]) and defined media (M9 [19, 139], 2xM9 [16, 140]). In medium optimization, it is usually impractical to test the full combination of all the ingredients at different concentrations due to the relatively large number of possibilities. It is also a challenge to construct a mechanistic model to predict the optimal medium composition for the production of isoprenoids, due to the lack of understanding of the biological functions of each medium ingredient and their interactions. Statistical design of experiments to screen critical components and optimize the concentration of the screened components by response surface methodology (RSM) offers a unique opportunity to determine the optimal concentrations of these components in the media with the minimal number of experiments.

In this study, the enhancement of lycopene production in *E. coli* was investigated by increasing the intracellular concentration of PEP using a PTS knocked out mutant. PTS knockout resulted in low cell growth in media containing only

glucose as the carbon source. In order to increase cell density, the growth medium was then systematically optimized with the aid of factorial experiment design. Both of cell growth and productivity of lycopene was enhanced by the medium optimization, the mechanism of which was subsequently investigated by metabolite and transcription profiling. In addition to the significant enhancement of lycopene production, amorpha-4,11-diene production was similarly enhanced by the use of the approach described herein, suggesting the approach has a broader utility for production of other isoprenoids.

4.2 Results

4.2.1 Increasing lycopene production and biomass of PTS mutant strain

PEP is a glycolytic intermediate and a major phospho-donor for many cellular processes. PEP can be converted to phosphoglycerates and then to glyceraldehyde 3-phosphate (GAP), the limiting precursor of the DXP pathway (Figure 4-1) [23]. It is reasonable that isoprenoid production can be enhanced through increasing the supply of PEP by deleting selective competing pathways, such as the phosphotransferase system (PTS). To test this hypothesis, the *ptsHIcrr* operon (encoding the components of the PTS) in *E. coli* genome was deleted, and the production of lycopene, a C40 isoprenoid, was compared in both the wild-type strain (MG01) and the knockout strain (PTS01). In line with the hypothesis, lycopene yield (g/g DCW) was increased by about three fold in PTS01 as compared to MG01 (Figure 4-2). However, the cell density was

exceptionally low in the glucose minimal medium, consistent with previous reports [32]. This was likely due to the disruption of glucose uptake by PTS [38].

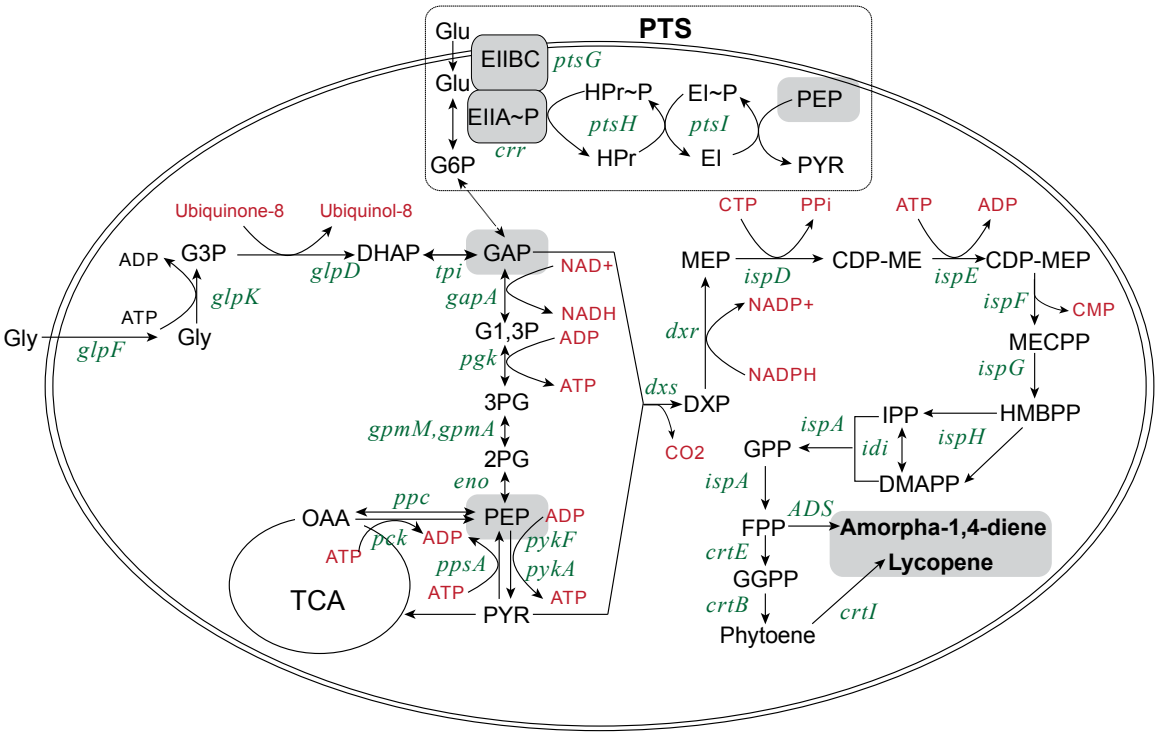


Figure 4-1. Relationship of the central metabolic and DXP pathways
The abbreviations for metabolites in the figure were as follows: glycerol (Gly), glucose (Glu), glucose-6-phosphate (G6P), glycerol-3-phosphate (G3P), dihydroxyacetone phosphate (DHAP), glyceraldehyde 3-phosphate (GAP), 1,3-biphospho-glycerate (G1,3P), 3-phospho-glycerate (3PG), 2-phospho-glycerate (2PG), phosphoenolpyruvate (PEP), the phosphotransferase system (PTS), pyruvate (PYR), oxaloacetate (OAA), tricarboxylic acid cycle (TCA), 1-deoxy-D-xylulose 5-phosphate (DXP), 2C-methyl-D-erythritol 4-phosphate (MEP), 4-diphosphocytidyl-2C-methyl D-erythritol (CDP-ME), 4-diphosphocytidyl-2C-methyl D-erythritol 2-phosphate (CDP-MEP), 2C-methyl-D-erythritol 2,4-diphosphate (MEC), hydroxylmethylbutenyl diphosphate (HMBPP), isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), farnesyl pyrophosphate (FPP), geranylgeranyl pyrophosphate (GGPP), phosphate (PPi), carbohydrate phosphotransferase system (PTS). The abbreviations for enzyme-coding genes in the figure are as follows: PTS enzyme IIBC (*ptsG*), histidine protein (*ptsH*), PTS enzyme I (*ptsI*), PTS enzyme IIA(*crr*), glycerol facilitator (*glpF*), glycerol kinase (*glpK*), glycerol-3-phosphate dehydrogenase (*glpD*), triose phosphate isomerase (*tpi*), glyceraldehyde-3-phosphate dehydrogenase A (*gapA*), phosphoglycerate kinase (*pgk*), phosphoglycero mutase III (*gpmM*), phosphoglyceromutase I (*gpmA*), enolase (*eno*), PEP carboxylase (*ppc*), PEP carboxykinase (*pck*), phosphoenolpyruvate synthetase (*ppsA*), pyruvate kinase type I and II (*pykFA*), DXP synthase (*dxs*), DXP reductase (*dxr*), CDPME synthase (*ispD*), CDPME kinase (*ispE*), CDPMEP synthase (*ispF*), HMBPP synthase (*ispG*), HMBPP reductase (*ispH*), IPP isomerase (*idi*), farnesyl pyrophosphate (*FPP*) synthase (*ispA*), GGPP synthase (*crtE*), phytoene synthase (*crtB*) and phytoene desaturase (*crtI*) and amorpha-4,11-diene synthase (*ADS*).

In order to improve the cell growth of PTS01, glycerol was added into the media to replace glucose. Cell density was found to increase with the addition of glycerol into the media, and the best growth was obtained when glycerol was used as the sole carbon source. More importantly, the specific productivity of lycopene (g lycopene / g DCW) in glycerol media was comparable to that of the media that had glucose as the sole carbon source. In addition, PTS01 produced more lycopene than MG01 in the glycerol media (Figure 4-2).

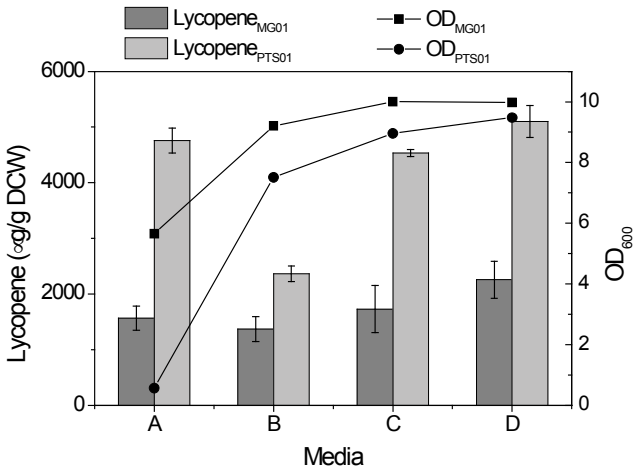


Figure 4-2. Comparison of lycopene yield and cell density of MG01 and PTS01 strains grown in different media compositions
Lycopene production and OD were assayed after 24 h for MG01 and PTS01 strains grown in Riesenberg media [141] containing different concentrations of glucose and glycerol. The carbon concentrations in media A, B, C, D were 10 g/L glucose + 0g/L glycerol, 7 g/L glucose + 3 g/L glycerol, 3 g/L glucose + 7 g/L glycerol and 0 g/L glucose + 10 g/L glycerol, respectively. All other components in the Riesenberg media were the same. The strains grew in the media at 37°C and 225 rpm).

4.2.2 Statistical optimization of growth media

A systematic optimization of the growth medium for PTS01 was performed to increase lycopene productivity. The contributions of five ingredients to lycopene

productivity were characterized using a minimal run resolution IV fractional factorial design (Table 4-1 and Table 4-2), based on the glycerol minimal medium optimized for high cell density [142]. The five factors chosen for optimization are: Glycerol, KH_2PO_4, $(NH_4)_2HPO_4$, sodium pyruvate, and L-arabinose. Glycerol was the major carbon source in media; KH_2PO_4 served as a buffer, ionic agent and phosphorus source; $(NH_4)_2HPO_4$ was the source of both nitrogen and phosphorus; sodium pyruvate was used as the substrate of the DXP pathway for lycopene production; and L-arabinose was used to induce the pBAD vector for the expression of *dxs-idi-ispDF* operon.

Table 4-1. Summary of selected medium ingredients for screening (Details of experiment design are shown in Table S1 in File S1).

Factor	Name	Unit	Low (-)	High (+)	Function in media
A	Glycerol	g/L	8	20	C source
B	KH_2PO_4	g/L	8	20	P source, buffer
C	$(NH_4)_2HPO_4$	g/L	2	5	N source, buffer
D	Na(Pyruvate)	g/L	2	5	Precursor
E	L-Arabinose	mM	0	0.1	Inducer

Table 4-2. Experimental design of Min Run Res IV for the production of lycopene of PTS01 strain

No.	Glycerol	KH_2PO_4	$(NH_4)_2HPO_4$	Pyruvate	Inducer	Lycopene
Unit	g/L	g/L	g/L	g/L	mM	µg/g DCW
1	8	8	5	5	0	1992
2	8	8	5	2	0	1755
3	20	8	2	2	0	2705
4	8	20	5	5	0.1	4688
5	20	8	5	5	0	4694
6	20	20	2	2	0.1	7304
7	20	8	5	2	0.1	4337
8	8	20	2	5	0	4851
9	20	20	5	2	0	6473
10	8	20	2	2	0.1	5444
11	8	8	2	5	0.1	3578
12	20	20	2	5	0.1	7568

The results of experimental design showed that glycerol and KH_2PO_4 were the most important factors for lycopene production (Figure 4-3A) and the relationships can be described by Equation 3-1 (Factor A is glycerol and factor B is KH_2PO_4). The equation indicated that lycopene production was enhanced by increased concentrations of glycerol and KH_2PO_4.

$$Lycopene = -836 + 150 * A + 240 * B \qquad \text{Equation 4-1}$$

After identifying the two critical ingredients in media for lycopene production, the optimal concentrations were further determined by applying RSM with a central composite design (Table 4-3). The experimental results were analyzed statistically using ANOVA. The mean predicted and observed responses, and the details of the statistical analysis, are presented in Table 4-3 and Table 4-4. The values of the regression coefficients were calculated and the fitted equation (in terms of coded values) for the prediction of lycopene production (µg/g DCW) are shown in Equation 3-2, where Factor A is glycerol and factor B is KH_2PO_4.

$$Lycopene = 11944 + 2524 * A - 35 * B - 1216 * A * B - 496 * A^2 - 2795 * B^2 \qquad \text{Equation 4-2}$$

Table 4-3. Central composite design of RSM design for production of lycopene of PTS01 strain with corresponding results

Order	Glycerol/ (g/L)	KH_2PO_4/ (g/L)	Observed / (µg/g DCW)	Predicted / (µg/g DCW)
1	10	20	5602	4947
2	30	20	11365	12429
3	10	50	8184	7308
4	30	50	9083	9925
5	5.9	35	6258	7380
6	34.1	35	15831	14522
7	20	13.8	6654	6404
8	20	56.2	6240	6303
9	20	35	13331	12222
10	20	35	12142	12222

11	20	35	12107	12222
12	20	35	11733	12222
13	20	35	11797	12222

Based on this equation, optimal concentrations of glycerol and KH_2PO_4 for production of lycopene (16,261 µg/g DCW) were predicted to be 50 g/L and 24 g/L, respectively (Figure 4-3B). To verify the model, lycopene production was measured from *E. coli* grown in these predicted optimal concentrations of glycerol and KH_2PO_4. The measured lycopene yield was 17542±1105 µg/g DCW (Table 4-5), similar to the predicted optimized yield of 16,261 µg/g DCW, thereby validating the model.

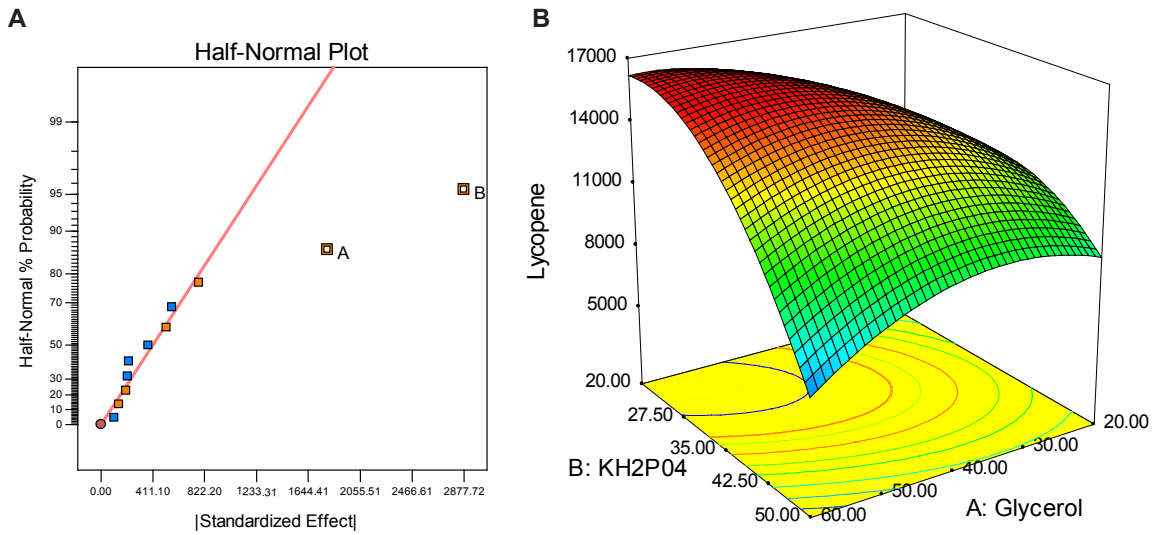


Figure 4-3. Identification and optimization of critical media composition for the production of lycopene in PTS01 strain.
(A) Half-normal probability plot of the results of minimal resolution IV experiment design (The details of the experiment design were shown in Table S1 in File S1). In the half-normal probability plot, the line was called the near zero line. The estimated effect of an unimportant factor will typically be on or close to a near-zero line, while the estimated effect of an important factor will typically be displaced well off the line. (B) RSM plot of lycopene production versus concentrations of glycerol and KH_2PO_4 of PTS01.

Table 4-4. Analysis of RSM design of PTS01 strain for the production of lycopene

Source	Coefficient Estimate	p-value
Model		0.001
A-Glycerol	2525	0.0003
B-KH2P04	-36	0.9261
A*B	-1216	0.0540
A^2	-635	0.1551
B^2	-2934	0.0002

Table 4-5. Model prediction and experimental validation of RSM

Item	Glycerol (g/L)	KH2PO4 (g/L)	Lycopene (ppm)
Model Prediction	50	24	16261
Verifying Experiment	50	25	17542±1105
R-Squared	0.95		
Adj R-Squared	0.90		

As shown in Figure 4-4A, PTS01 grown in this optimized media (designated OPT1) produced more than twice as much lycopene as MG01. PEP concentrations were consistently higher in PTS01 than MG01 at all growth phase (Figure 4-4B). There is a correlation between lycopene production and PEP concentration, suggesting that increased lycopene production can be attributed to higher amounts of PEP available.

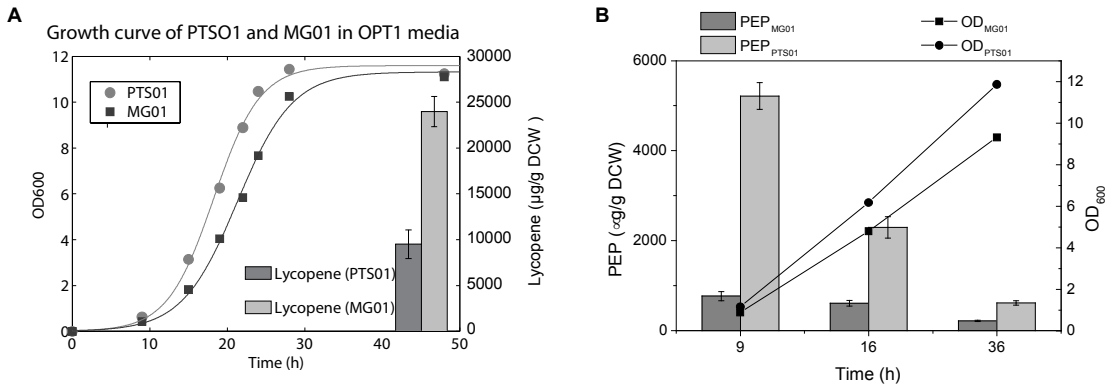


Figure 4-4. Comparison of lycopene production, growth curve and PEP levels of MG01 and PTS01

(A) Growth curve and lycopene productions of PTS01 and MG01 in optimal glycerol medium (OPT1) at 37°C, with a shaking speed of 300 rpm. (B) Comparison of PEP levels in MG01 and PTS01 at different growth stages (early-log phase, pre-induction, 9 h; mid-log phase, 16 h post-

induction; stationary phase, 36 h post-induction). Both of the strains were grown in OPT1 at 37°C, with a shaking speed of 300 rpm).

The culture conditions of temperature and O_2 feeding (controlled by rotation speed of incubation) were optimized. It was found that by increasing rotation speed to 300 rpm from 225 rpm, and keeping culture temperature at 37°C (earlier experiments were conducted at 225 rpm and 37°C), lycopene yield was increased to around 20,000 µg/g DCW for PTS01 (Figure 4-5A).

The optimized medium (OPT1) and culture conditions were used in further experiments to investigate the effect of PTS deletion on the accumulation of intracellular PEP and the concomitant increase in lycopene productivity.

The growth rate and lycopene productivity of cells grown in OPT1 were compared with other commonly used complex medium (2xPY, LB, 2xYT) and defined medium (2xM9 with 10 g/L glycerol, R-media with 10 g/L glycerol). It was found that the optimized medium, OPT1, was superior in lycopene production (Figure 4-5B) compared to all other media tested. Although the strain grew faster in complex media (LB, 2xYT and 2xPY), the biomass yields were lower when compared to the defined media, including OPT1 (Figure 4-5C). By using this approach to optimize lycopene production using the DXP pathway, we achieved the highest lycopene yield reported so far, compared to previous studies [16, 143].

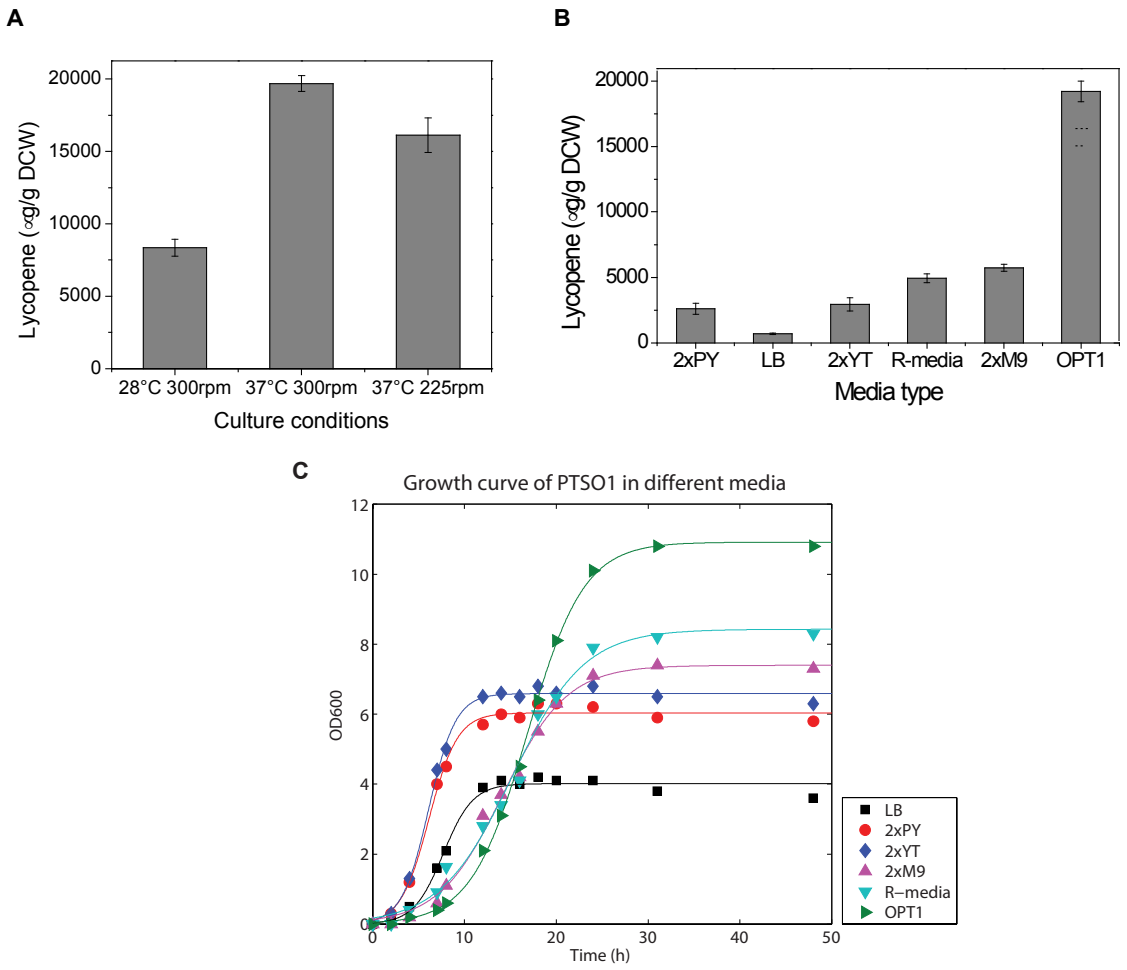


Figure 4-5. Optimization of culture conditions for PTS01 and comparison of OPT1 with other commonly used media.
(A) Lycopene production of PTS01 at different temperatures and shaking speeds. (B) Comparison of lycopene production of PTS01 in 2xPY, LB, 2xYT, 2xM9, R-media and OPT1 media. The strain was grown at 37°C, with a shaking speed of 300 rpm. (C) The growth curves of PTS01 in 2xPY, LB, 2xYT, 2xM9, R-media and OPT1 media.

Table 4-6. Reported lycopene titer in literature

#	Lycopene titer	Lycopene yield	Productivity	Media	Host	Literature
Unit	mg/L	ppm	mg/L/h			
1	775	64900	8.1		Blakeslea trispora	Choudhari, S. M., et al. (2008).
2	260	5400	4.333	2YT+gly	Escherichia coli	Kim, S.-W., et al. (2009)
3	79	33430		2YT+gly	Escherichia coli	Chen, Y. Y., et al. (2013)
4	102	22000		2YT+gly	Escherichia coli	Yoon, S., et al. (2006).
5		18,000		2XM9+glc	Escherichia coli	Alper, H., et al. (2005)

6	80	20,000	1.7	Def+gly	Escherichia coli	Zhang, C., et al. (2013)
7	198			2YT+gly	Escherichia coli	Rad, S. A., et al. (2012)
8	220	27,000	8.8	2XM9+glc	Escherichia coli	Alper, H., et al. (2006)
9	**1350**	**32,000**	**33.8**	**AA+Def+gly**	**Escherichia coli**	**Kim, Y. S., et al. (2011)**
10	**1480**	**30,300**	**32.2**	**Def+glc**	**Escherichia coli**	**This study**
11	54	5,400	0.45		Mucor circinelloides	Nicolas-Molina, F. E., et al. (2008)
12	74	4,600	1.8		Pichia pastoris	Bhataya, A., et al. (2009)
13		780		YPD	Pichia pastoris X-33	Araya-Garay, J. M., et al. (2012)
14	15		0.083		Rhodospirillum rubrum	Wang, G. S., et al. (2012).

4.2.3 Transcriptional investigation and metabolites profiling of PST01 grown in OPT1 medium

Transcriptional analysis by qPCR has been used to gain mechanistic insights into metabolic engineering [85, 144]. The transcription of the genes involved in lycopene biosynthesis was investigated in order to better understand the underlying mechanisms affecting lycopene production in the OPT1 medium, and how it compares to complex media. The PTS01 strain harbored two plasmids (pBAD-SIDF and pAC-LYC) which overexpressed four rate-limiting genes of the DXP pathway (*dxs*, *idi* and *ispDF*) and three heterologous lycopene synthetic genes (*crtE*, *crtB* and *crtI*), respectively (Figure 4-1).

PTS01 grown in complex media (2xPY) and the OPT1 medium was measured and compared using the expression level of one gene from each plasmid (*dxs* and *crtE*) and an endogenous DXP pathway gene (*ispE*). Interestingly, at different growth phases, the transcriptional levels of the plasmid-encoded genes

(*dxs* and *crtE*) were found to be up-regulated by approximately 4-8 times in the OPT1 medium, compared to those in 2xPY (Figure 4-6). Similarly, the plasmid copy numbers of both of these plasmids were up-regulated in the OPT1 medium (Figure 4-7). However, the expression level of the endogenous gene (*ispE*) was similar in both strains when grown in either media (Figure 4-6). These results indicate that increased transcriptions of *dxs* and *crtE* were probably due to the increase in plasmid copy number. Therefore, it was hypothesized that the OPT1 medium increased the plasmid copy numbers of pBAD-SIDF and pAC-LYC in greater quantities than the 2xPY medium, and this led to the increased expression levels of the genes in those plasmids, and finally contributed to the increased production of lycopene.

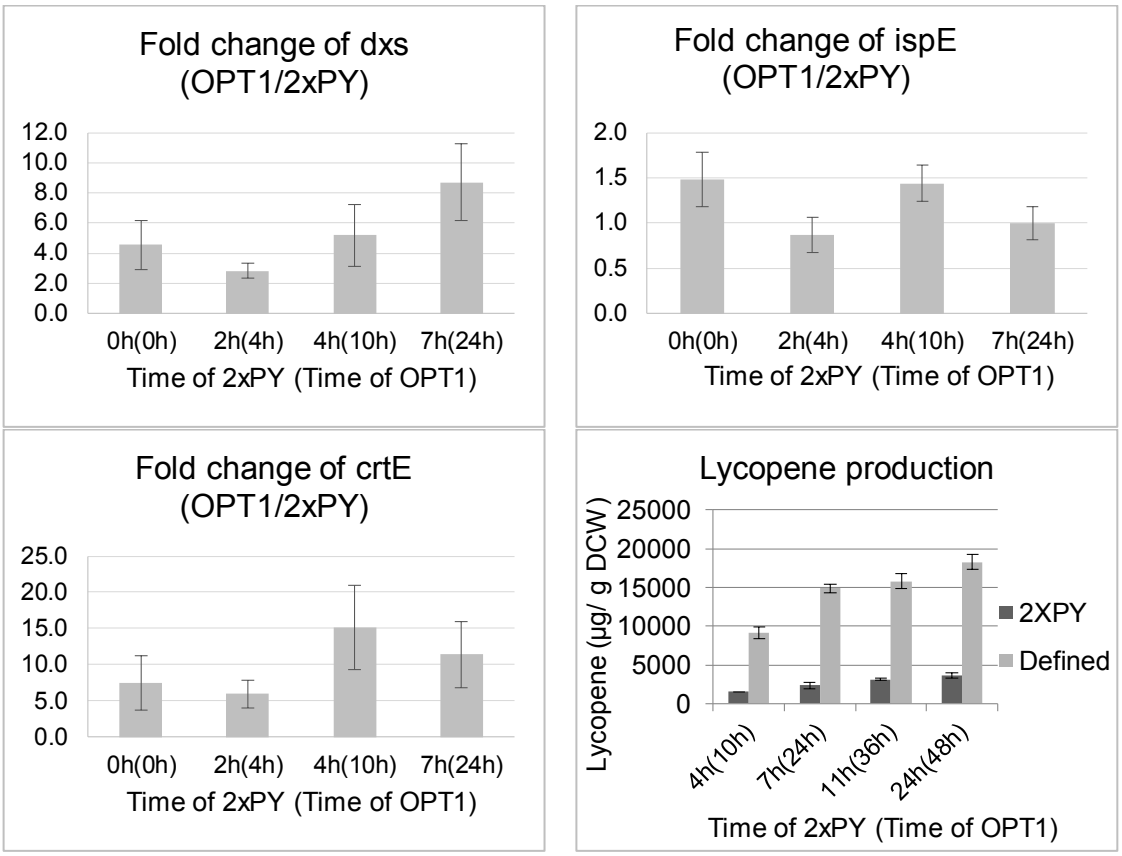


Figure 4-6. Fold change of transcriptional levels of dxs, ispE and crtE in PTS01 strain grown in OPT1 as compared to those in 2xPY medium and lycopene production.

The transcriptional levels measured in 2xPY and OPT1 were compared at four different growth stages, late lag phase (0h for 2xPY and 0h for OPT1, the time of induction was set as time 0h), early log phase (2h for 2xPY and 4h for OPT1), middle log phase (4h for 2xPY and 10h for OPT1) and late log phase (4h for 2xPY and 10h for OPT1). The sampling time in each growth stage was due to differences in growth rates of PTS01 which was faster in 2xPY than in OPT1 (As shown in Figure 4-5C). All the measurements were normalized to the expression of *cysG*.

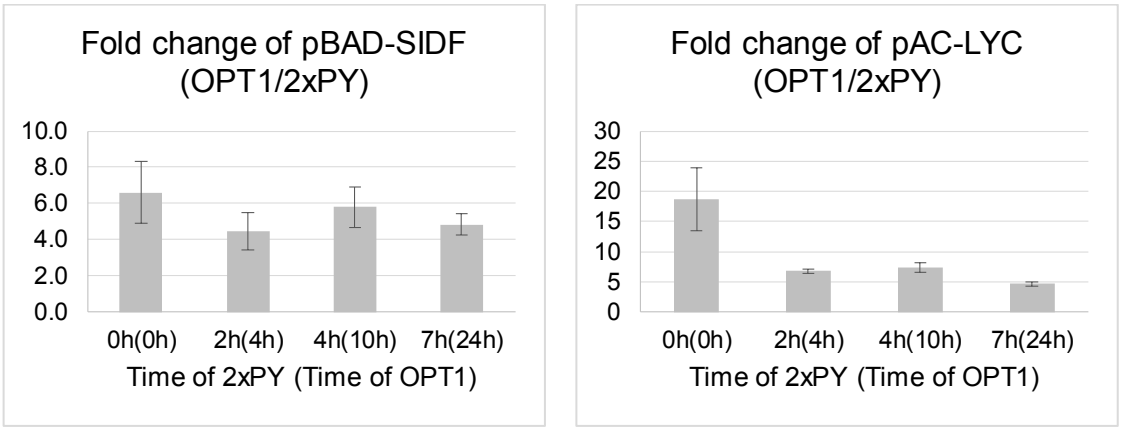


Figure 4-7. Fold change of plasmid copy number in PTS01 grown in OPT1 and 2xPY media. The x-axis described parameters identical to Figure 4-6. Plasmid copy number was calculated using copy number of plasmid resistant gene (ampicillin resistant gene for pBAD-SIDF and chloramphenicol resistant gene for pAC-LYC) normalized by the copy number of chromosomal gene *ispE*.

To test the hypothesis that co-upregulations of these genes were major contributors to the enhancement of lycopene production in the OPT1 medium, both the operon *dxs-idi-ispDF* and the operon *crtEBI* were overexpressed in 2xPY. Using the pBAD vector, the operon *dxs-idi-ispDF* was put under the control of the araBAD promoter. The lycopene biosynthetic genes (*crtEBI)* were placed under the control of the araBAD promoter in the pAC vector (designated pACM-LYC, which was constructed from plasmid pAC-LYC by replacing the constitutive promoter with the araBAD inducible promoter). PTS02, a MG1655 PTS^- strain carrying pACM-LYC and pBAD-SIDF was then generated and grown in 2xPY media. By gradually increasing the concentration of L-arabinose, the transcriptional levels of *dxs-idi-ispDF* and *crtEBI* were induced to levels greater

than or similar to that found in PTS01 in the OPT1 medium (Figure 4-8). Lycopene production was found to increase gradually from 500 to 6000 µg/g DCW with increasing amounts of L-arabinose, indicative of the enhancement of lycopene production by the co-upregulation of *dxs-idi-ispDF* and *crtEBI.* However, despite achieving transcriptional levels of *dxs-idi-ispDF* and *crtEBI* of PTS02 (grown in 2xPY) that were higher or at least comparable with those of PTS01 (grown in OPT1 medium), the maximal lycopene production of PTS02 in 2xPY (~6000 µg/g DCW) was nonetheless found to be significantly lower than that of PTS01 in OPT1 (~18,000 µg/g DCW) (Figure 4-8). These results indicated that the co-upregulation of *dxs-idi-ispDF* and *crtEBI* was essential but not sufficient to enhance lycopene production to the levels observed in PTS01 grown in the OPT1 medium.

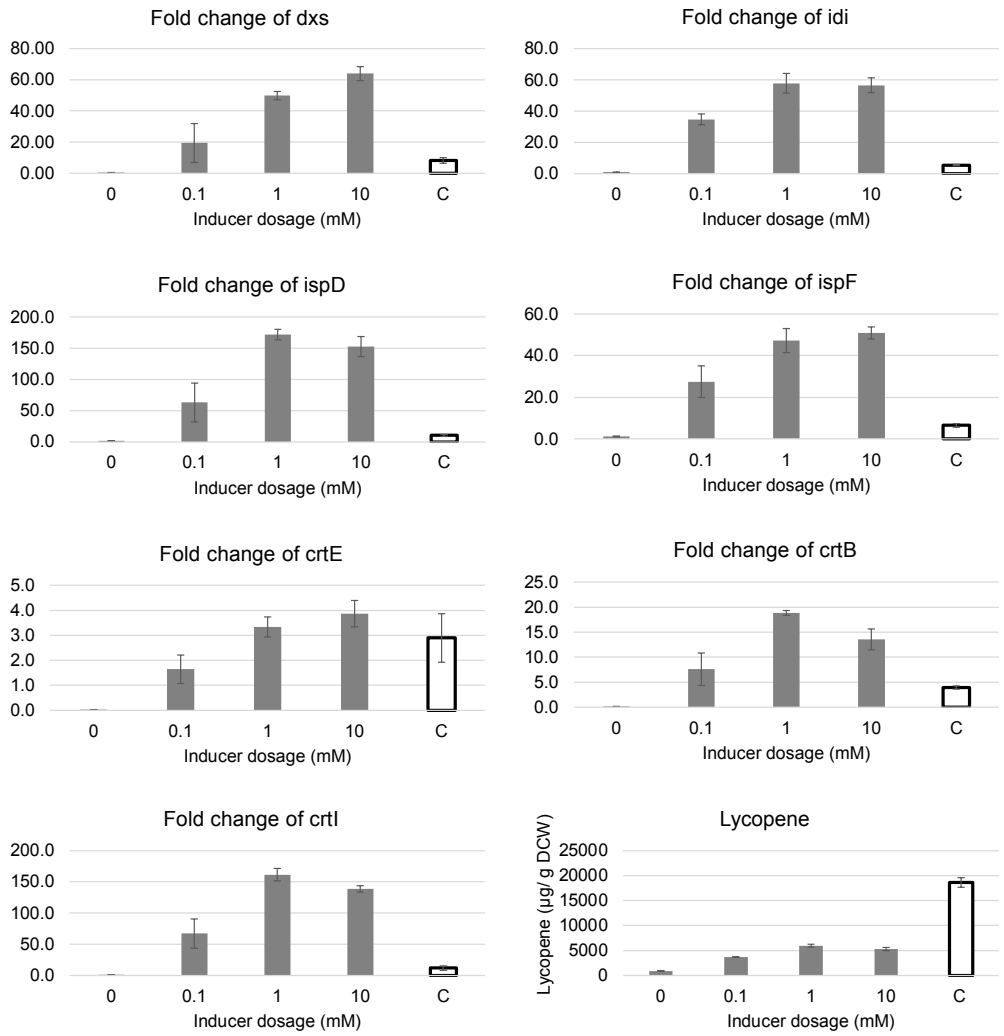
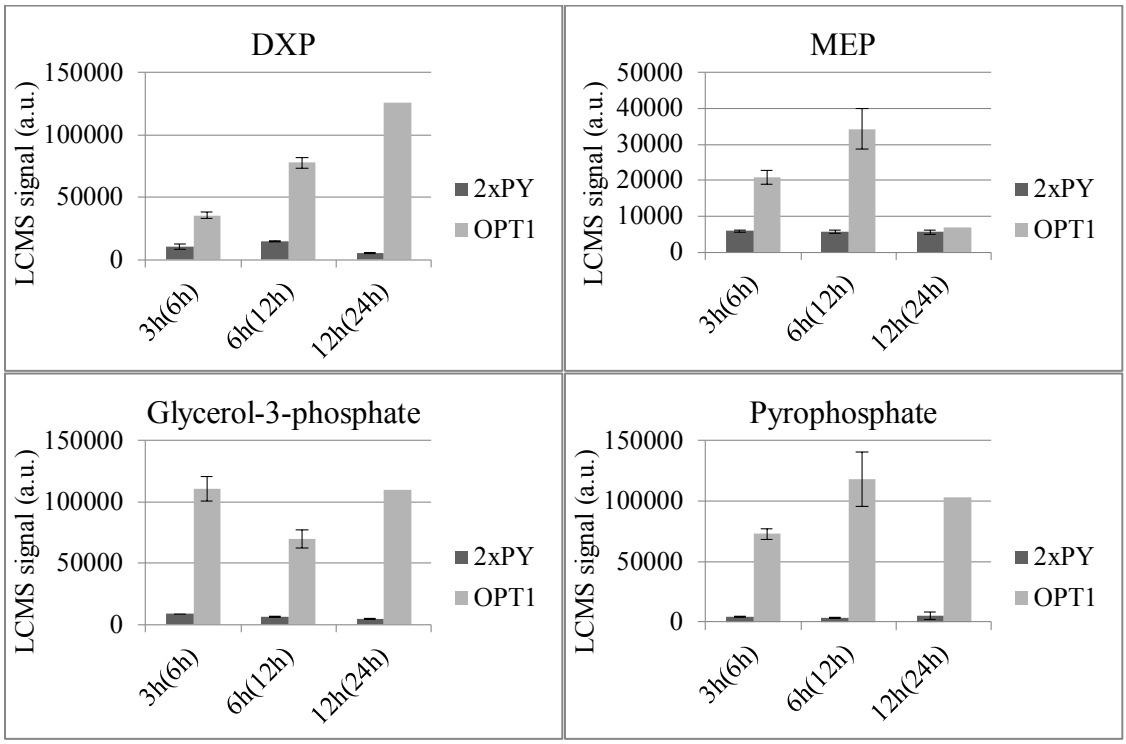


Figure 4-8. Transcriptional analysis of *dxs*, *idi*, *ispD ispF*, *crtE*, *crtB* and *crtI.*
PTS02 was grown in 2xPY and induced with different concentrations of L-arabinose (0.1, 1 and 10 mM). Columns labelled "C" represent the expression levels of PTS01 grown in OPT1 in the absence of the inducer (L-arabinose). Fold changes of transcriptional levels of these genes were normalized to the expression levels of corresponding genes of PTS01 grown in 2xPY in the absence of the inducer (L-arabinose).

In addition to transcriptional analysis, metabolite profiling also provides useful information in attempts to understand underlying biochemical mechanisms [145]. The intracellular metabolites of the PTS01 cells grown in the 2xPY and OPT1 media were profiled and compared. A panel of metabolites were found to be significantly enriched in the cells grown in the OPT1 medium, which included

glycerol-3-phosphate (G3P), DXP and pyrophosphate (Figure 4-9). G3P and DXP are critical intermediates in the isoprenoid synthetic pathway when routing from glycerol (the carbon source of the defined medium), and the increases in exponential state concentrations were consistent with enhanced lycopene production. The upregulated level of pyrophosphate was also expected in the OPT1 medium, as the medium contained large amounts of phosphate. Major cellular co-factors ATP and NADH were detected in different growth stages and most of their concentrations were comparable in the 2xPY and OPT1 media, suggesting that the increase in lycopene production was not due to changes in these cofactor supplies in the OPT1 medium.



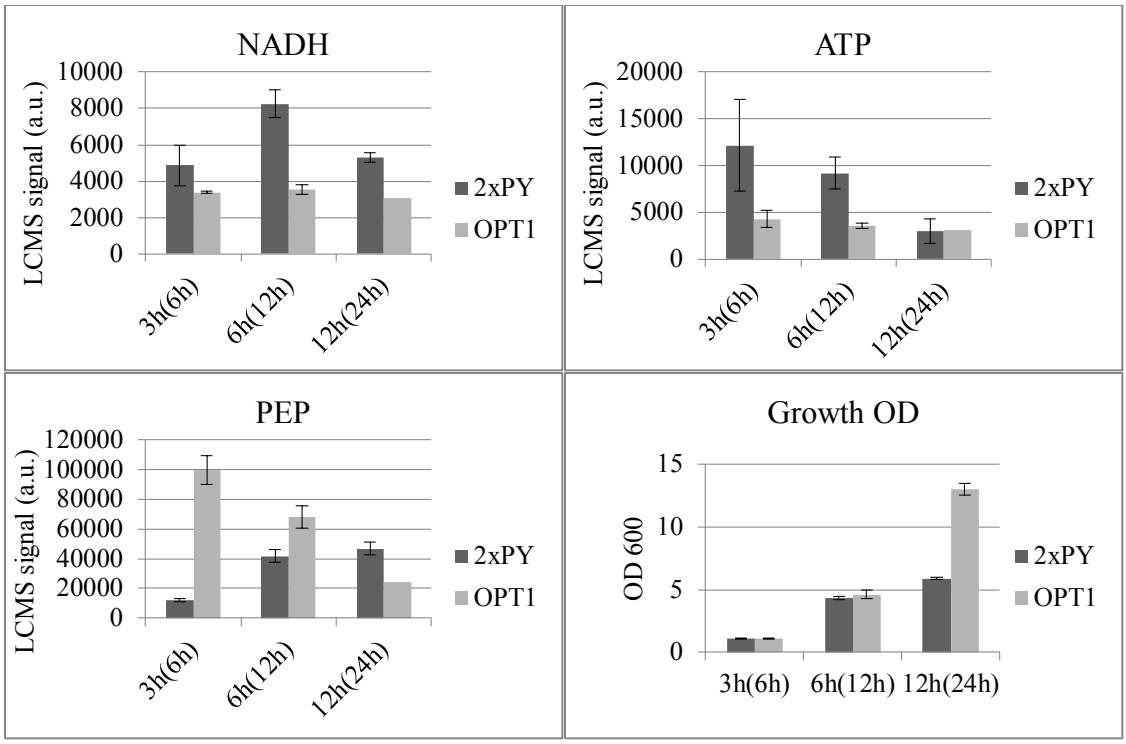


Figure 4-9. Concentrations of metabolites and cofactors in PTS01 strain grown in 2xPY or OPT1 media.

The metabolites in 2xPY and OPT1 were compared at three different growth stages, early log phase (3h for 2xPY, 6h for OPT1) , middle log phase (6h for 2xPY, 12h for OPT1) and early stationary phase (12h, 24h). The time of induction was defined as time 0h.

4.2.4 Production of amorpha-4,11-diene in PTS03 strain in defined media

Using PTS deletion and media development, we have established a rapid and effective method to increase lycopene production in *E. coli*. In order to demonstrate the broader utility of this approach, the production of amorpha-4,11-diene, a distinct C15 isoprenoid, was attempted. The growth medium was optimized using the same statistical method as described before, where three important factors A, B and E (which are glycerol, KH_2PO_4 and L-arabinose, respectively) were screened by fractional factorial design (Figure 4-10A and Table 4-6). The optimal concentrations of the three critical factors were determined by RSM, and the results are shown in Figure 4-10B, Figure 4-10C and Table 4-7. This optimized medium for amorpha-4,11-diene was designated OPT2 and an enhancement of amorpha-4,11-diene production was observed (182 mg/L), which was significantly higher than the yield derived from 2xPY complex media (35.3 mg/L). Similarly to lycopene, the production of amorpha-4,11-diene (182 mg/L) was better in the PTS^- strain (PTS03) than in the wide type (MG02) in OPT2 medium (136 mg/L, Figure 4-10D).

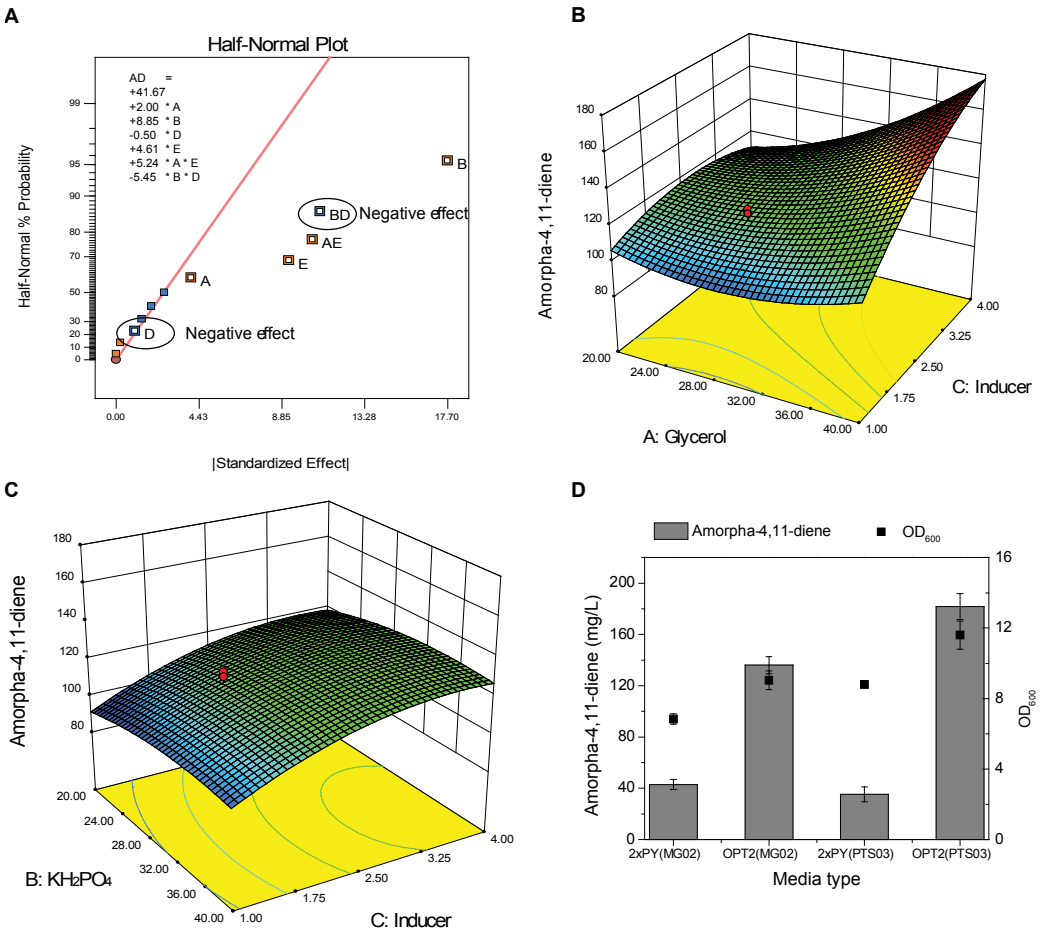


Figure 4-10. Media optimization of amorpha-4,11-diene production in PTS03 strain. (A) Screening of key factors for amorpha-4,11-diene production (The factors share the same denotations as those described in Figure 4-3A.In the figure, A, B and E are dominant factors with positive effects on the production of amorpha-1.4-diene, therefore they were selected as key factors for further optimization). (B) Production of amorpha-4,11-diene versus glycerol concentration and inducer dosage. (C) Production of amorpha-4,11-diene versus KH_2PO_4 concentration and inducer dosage. (D) Comparison of amorpha-4,11-diene production in 2xPY and OPT2 media for the MG02 and PTS03 strains. The details of the experiment design are described in Supplementary Table 4-6 and Table 4-7.

Table 4-7. Experimental design of Min Run Res IV for the production of amorpha-4,11-diene of PTS03 strain

Glycerol/ (g/L)	KH2PO4/ (g/L)	(NH4)2HPO4/ (g/L)	Pyruvate (g/L)	Inducer/ mM	AD / (mg/L)
8	8	5	5	0.1	33.6
8	8	5	2	0.1	26.3
20	8	2	2	0.1	19.6
8	20	5	5	0.3	40.4
20	8	5	5	0.1	29.4

20	20	2	2	0.3	69.2
20	8	5	2	0.3	40.4
8	20	2	5	0.1	46.3
20	20	5	2	0.1	49.5
8	20	2	2	0.3	52.1
8	8	2	5	0.3	38.4
20	20	2	5	0.3	55.0

Table 4-8. Central composite design of RSM design for the production of amorpha-4,11-diene of PTS03 strain

Order	Glycerol/ (g/L)	KH2PO4 / (g/L)	Inducer/ mM	Observed AD/ (mg/L)	Predicted AD / (mg/L)
1	27.5	27.5	2	118.9	115.92
2	35	35	3	158.9	147.25
3	27.5	14.9	2	101.7	100.64
4	35	20	3	139.0	136.4
5	35	35	1	102.2	105.72
6	27.5	27.5	2	98.5	115.92
7	20	20	1	90.0	99.56
8	27.5	27.5	3.7	111.3	126.64
9	35	20	1	92.8	94.87
10	27.5	27.5	2	118.9	115.92
11	40.1	27.5	2	140.2	143.03
12	27.5	40.1	2	107.8	115.85
13	20	20	3	122.8	111.46
14	20	35	3	122.8	118.7
15	20	35	1	112.1	106.79
16	27.5	27.5	2	119.1	115.92
17	14.9	27.5	2	118.7	122.97
18	27.5	27.5	2	121.6	115.92
19	27.5	27.5	2	119.6	115.92
20	27.5	27.5	0.3	90.0	81.71

4.3 Discussion

In this study, the deletion of PTS and the use of statistical medium optimization enhanced the production of a C40 isoprenoid (lycopene) and a C15 isoprenoid (amorpha-4,11-diene), suggesting that the approaches developed herein may be used for optimizing the production of other isoprenoids.

Increasing intracellular concentration of PEP in PTS knockout mutant when grown in glucose media has been reported to enhance the yield of aromatics [32, 38]. However, it is not known if PTS deletion can also enrich intracellular PEP when cells are grown in glycerol media. Glycerol uptake genes (*glpF*, *glpK* and *glpD*) are known to be regulated by carbon catabolite repression in *E. coli* and are intimately linked to Enzyme IIA (gene *ccr,* part of the *ptsHIcrr* operon) and the ratio of PEP to pyruvate [81, 146, 147]. Although it is expected that PTS deletion will affect transcription of PTS related genes (e.g. *pck*, *ppc*, *ppsA*, *pykFA* and *eno,* Figure 4-1), no significant difference was found in the expression levels of these PEP related genes in PTS01 and MG01 (Figure 4-11). This does not rule out the possibility that PTS disruption may modulate post-translational events resulting in PEP accumulation, an interesting possibility yet to be examined. PEP is thought to transfer phosphate to EI (gene *ptsI*), a PEP-dependent protein-kinase. The phosphate group is subsequently transferred from EI~P to HPr (gene *ptsH*), from HPr~P to the soluble EIIAGlc (sometimes also called EIIACrr), and finally from

EIIAGlc~P to the glucose-specific membrane protein EIICBGlc (gene *ptsG*), resulting in the uptake and phosphorylation of glucose (Figure 4-1) [81]. As there are more than 20 different Enzymes II in *E. coli* [148], the deletion of the *ptsHIcrr* operon in this study may cause a decrease in phosphorylation of these enzymes resulting in the accumulation of PEP. This hypothesis is the subject of future investigations.

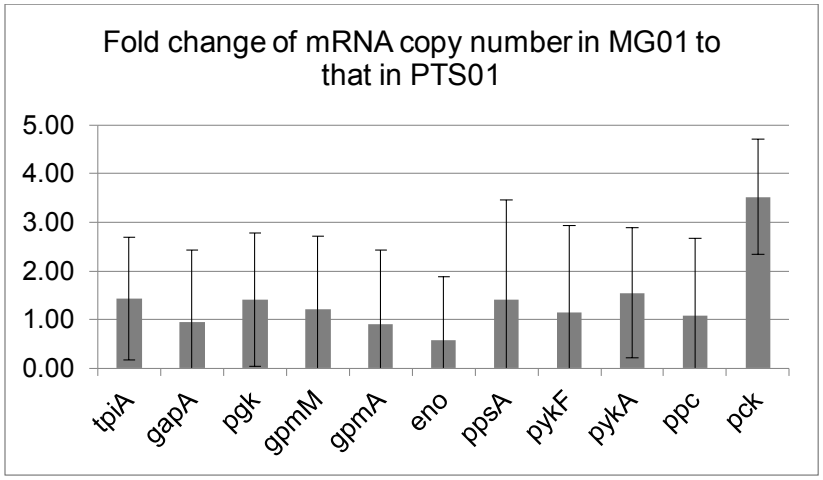


Figure 4-11. Fold change of transcriptional levels of genes involved in the metabolic pathway of glycerol in PTS01 and MG01 grown in OPT1 media.
The growth media used in this study was OPT1 and all the samples of MG01 and PTS01 were collected at early log phase. The expression levels of all the genes were normalized to *cysG*.

PTS knockout not only affects the uptake of glucose but also plays an important role in carbon catabolite repression. This is because phosphorylated $EIIA^{Glc}$ binds and activates adenylate cyclase (AC), which leads to cyclic AMP (cAMP) synthesis, and high cAMP concentrations trigger the formation of cAMP–CRP (cyclic AMP (cAMP) receptor protein; also called catabolite gene-activator protein (CAP)) complexes, which bind

and activate the promoters of many catabolic genes [81]. It was also reported that the araBAD promoter was affected by CAP [149]. To understand the effects of PTS knockout on the activities of the araBAD promoter in the pBAD-SIDF and a constitutive promoter in pAC-crtEBI, the expression levels of genes in the pBAD-SIDF plasmid and pAC-crtEBI plasmid for the MG01 strain and PTS01 strain, both cultivated in OPT1 media, were compared. However, the results indicated that the expression levels of all the genes in the two plasmids were not affected significantly when PTS was deleted (Figure 4-12). This could be because there was no arabinose added into the system, so the activity of the araBAD promoter was relatively low under such conditions and the effects of a PTS knockout was insignificant.

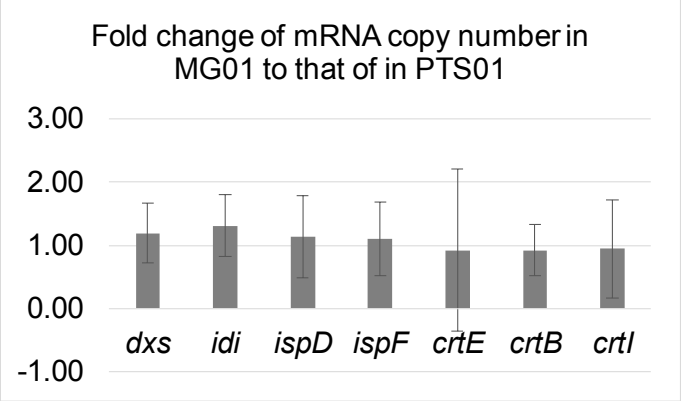


Figure 4-12. Fold change of transcriptional levels of genes in the pBAD-SIDF and pAC-crtEBI in PTS01 and MG01 grown in OPT1 media.
The growth media used in this study was OPT1 and all the samples of MG01 and PTS01 were collected at early log phase. The expression levels of all the genes were normalized to *cysG*.

The optimization of a growth medium is an effective method to increase cell density and production of recombinant proteins [132, 133] and bulk chemicals [134, 135]. To optimize the media, it was impractical to test the full combinations of components at different concentrations. Furthermore, it is a challenge to construct a mechanistic model to predict the interactions of the medium compositions with respect to the production of isoprenoids. Instead, we explored the use of statistically designed experiments [150] to screen critical components and optimize the concentrations rapidly, with a minimal number of experiments. From such studies, glycerol and KH_2PO_4 were identified as the most dominant factors in lycopene production, while pyruvate, $(NH_4)_2HPO_4$ and inducer dosage were found to be less critical. The KH_2PO_4 was further validated by single-factor experiment (Figure 4-13). It is likely that glycerol and KH_2PO_4 have global metabolic effects. Limiting the nitrogen supply increased the production of amorpha-4,11-diene [130], which is consistent with our results that $(NH_4)_2HPO_4$ is not critical for the production of both lycopene and amorpha-4,11-diene. Pyruvate was determined to be less critical to the production of such isoprenoids, perhaps because of an abundance of intracellular pyruvate.

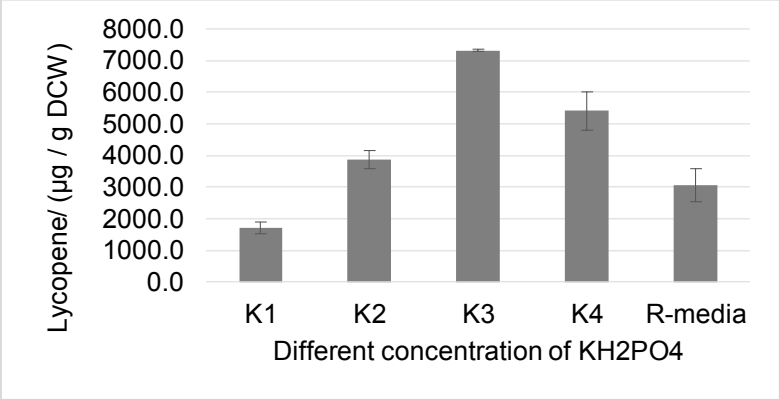


Figure 4-13. The effect of KH_2PO_4 concentration on lycopene production. The concentration of KH_2PO_4 in K1, K2, K3, K4 and R-media were 73, 173.5, 273.5, 473.5 and 97.7 mM, respectively. All the other all the other components were kept at the same levels as the R-media.

Interestingly, the effects of inducer dosage were different for the production of lycopene and amorpha-4,11-diene. For lycopene, low inducer dosage did not affect production and higher inducer dosage was detrimental to production in PTS01 (Figure 4-14). While there was an optimal inducer dosage for amorpha-4,11-diene production, too much or too little inducer had a negative impact. In addition to media composition, temperature and oxygen supply are also important factors that were tuned to optimize the final yield of lycopene.

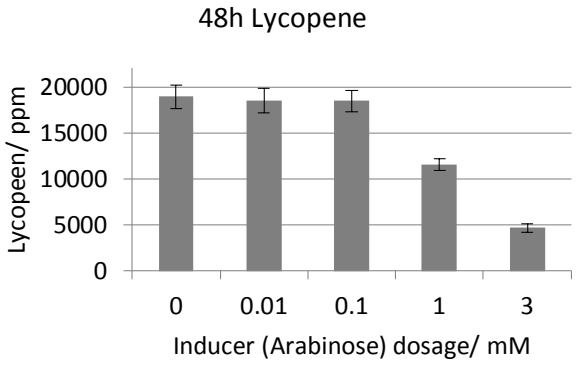


Figure 4-14. The effect of Inducer dosage on lycopene production in the OPT1 media

Our transcriptional analysis provided useful insights into how optimization of media resulted in increased lycopene production. The increase in *dxs* and *crtE* expressions was observed in OPT1 when compared to 2xPY (Figure 4-6). Further investigation indicated that the up-regulation of these genes was due to greater plasmid copy numbers in OPT1 (Figure 4-7). It is known that media composition can affect the quality and yield of plasmids [151]. Similarly, the alteration of carbon-to-nitrogen (C:N) ratio [152] and amino acid starvation [153] have also been shown to increase plasmid copy numbers. By lowering the specific growth rate, plasmid DNA yields can be increased [141]. This inverse correlation between plasmid DNA synthesis and growth rate was thought to be due to higher plasmid stability as well as increased plasmid replication in reduced growth rates [154]. The PTS deleted strain (PTS01) grew slowly in the OPT1 medium, which is consistent with the earlier proposal, and it may share similar

mechanisms to increase the plasmid copy number, thereby increasing transcription.

Although lycopene production levels increased significantly, paralleling an increase in the transcription of *dxs-idi-ispDF* and *crtEBI* in 2xPY, the highest lycopene yield from 2xPY was still much less than in OPT1. These results suggest that the co-upregulation of *dxs-idi-ispDF* and *crtEBI* was required to enhance lycopene production, but was still insufficient to reach the levels observed in PTS01 grown in an OPT1 medium. Metabolite profiling results showed that major cellular co-factors ATP and NADH were comparable in the 2xPY and OPT1 media, suggesting that the availability of cofactors was unlikely to be a limitation and may not account for the increase in lycopene production in the OPT1 medium. In addition to the transcriptional upregulation of *crtEBI*, the prolonged exponential phase of OPT1 could be beneficial in diverting more carbon fluxes into DXP pathway, thus increasing the overall yield.

4.4 Conclusion

To our knowledge, this is the first study that demonstrated the enhancement of isoprenoid production (lycopene and amorpha-4,11-diene) using the phosphotransferase system deficient strain of *E. coli.* The growth medium was rapidly optimized for the knockout mutant with the aid of statistical experiment design. This concurrently overcome cell growth retardation, a side effect of PTS disruption on a glucose medium, and

increased the production of lycopene, a C40 isoprenoid. The highest lycopene yield in this study was 20,000 µg/g DCW, which was a significant yield as compared to published reports. An attempt to understand the underlying mechanisms in cells grown in the optimized defined media demonstrated that the transcription of the lycopene biosynthetic genes was essential but insufficient to account for the observed enhancements in lycopene production. Deletion of PTS in combination with statistical medium optimization was also successfully applied to enhance the production of amorpha-4,11-diene, a distinct C15 isoprenoid, suggesting that the approach developed herein has broad utility beyond the scope of this study.

4.5 Materials and methods

4.5.1 Statistical experimental design

Fractional factorial designs of experiment and response surface methodology were calculated using Design Expert® V8 Software, Stat-Ease, Inc.

4.5.2 Bacteria strains and plasmids

E coli. K-12 MG1655 [F- lambda- *ilvG*- *rfb*-50 *rph*-1] was used as the parental strain, and MG1655 PTS$^-$ (PTS$^-$ for short) was used in the following experiments for lycopene production with the pAC-LYC plasmid [131]. PTS$^-$ strain was obtained by replacing operon *ptsHIcrr* with a kanamycin-resistant gene using the λ Red recombination system

developed by Datsenko et al. [128]. The details of the oligos (EcoRB-PTSF-KanF and EcoRBR-PTSR-KanR) and the plasmid pKD46 used for this work was shown in Table S6 in File S1, and the correct mutant of PTS^- was verified by PCR with the wild-type strain as the control. Plasmid pBAD-SIDF was constructed via ligation of *dxs-idi-ispDF* operon amplified by polymerase chain reaction (PCR) from p20T7MEP [13] into a modified pBAD-B vector (Invitrogen), which controlled the expression of the *dxs-idi-ispDF* operon using the araBAD promoter. Plasmid pACM-LYC was constructed by modifying the pAC-LYC plasmid so that three genes *crtE*, *crtB* and *crtI* were under the control of the araBAD promoter. Plasmid pAC-ADS was constructed based on pAC-LYC by replacing the crtEBI operon with the amorpha-4,11-diene synthase gene (ADS) using the ligase-independent cloning (LIC) method [155]. All the information about the strains, plasmids and primers is in Table S6 in File S1.

4.5.3 Culture media and growth conditions

1. Defined media recipe for optimization

Different concentrations of glycerol, $(NH_4)_2HPO_4$, KH_2PO_4, sodium pyruvate and citric acid were added into the media according to the experimental design. In addition, the media also contained 0.5 g/L $MgSO_4$, 2.5 mg/L $CoCl_2 \cdot 6H_2O$, 15.0 mg/L $MnSO_4 \cdot 4H_2O$, 1.5 mg/L $CuSO_4 \cdot 2H_2O$, 3 mg/L H_3BO_3, 2.5 mg/L $Na_2MoO_4 \cdot 2H_2O$, 13 mg/L $Zn(CH_3COO)_2$, 60 mg/L Fe(III) citrate, 4.5 mg/L thiamine·HCl and 8.4 mg/L EDTA.

The final optimized defined media for PTS01 consisted of 50 g/L Glycerol, 24 g/L KH_2PO_4, 4 g/L $(NH_4)_2HPO_4$, 1.7 g/L citric acid, and all the other components were kept at the same levels as the Riesenberg media [142]. This optimized defined medium was named OPT1.

The final optimized defined media for PTS03 consisted of 45 g/L Glycerol, 27.5 g/L KH_2PO_4, 4 g/L $(NH_4)_2HPO_4$, 1.7 g/L citric acid, and all the other components were the same as the Riesenberg media [142]. This optimized defined medium was named OPT2. The concentration of L-arabinose (inducer) in OPT2 was 5mM.

All the defined media were adjusted to pH7.0 using a NaOH solution, and filtered with a 0.2 micron membrane.

2. Riesenberg media (R-media) composition [142]

The R-media contained 10 g/L glycerol, 4 g/L $(NH_4)_2HPO_4$, 13.3 g/L KH_2PO_4 ,1.7 g/L citric acid, 2.5 mg/L $CoCl_2 \cdot 6H_2O$,15.0 mg/L $MnSO_4 \cdot 4H_2O$, 1.5 mg/L $CuSO_4 \cdot 2H_2O$, 3 mg/L H_3BO_3, 2.5 mg/L $Na_2MoO_4 \cdot 2H_2O$, 13 mg/L $Zn(CH_3COO)_2$, 60 mg/L Fe(III) citrate, 4.5 mg/L thiamine·HCl and 8.4 mg/L EDTA.

Glycerol, (NH4)2HPO4, KH2PO4 and citric acid were autoclaved separately. Sterile solutions of glucose, MgSO4, and thiamine-HCl were added afterwards, the pH of the media was adjusted to 7.0, and the media was autoclaved at 121 °C for 20 mins.

3. 2xM9 defined media composition [16]

The 2xM9 media contained 25.6 g/L $Na_2HPO_4 \cdot 7H_2O$, 6 g/L KH_2PO_4, 1 g/L NaCl, 2 g/L NH_4Cl , 2 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 10 g/L glycerol. The pH of the media was adjusted to 7.0, and the media was autoclaved at 121 °C for 20 mins.

4. 2xPY, LB and 2xYT complex media compositions

The 2xPY media contained 20 g/L peptone, 10 g/L yeast extract and 10 g/L NaCl. The pH of the media was adjusted to 7.0 and the media was autoclaved at 121 °C for 20 mins.

The LB media contained 10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl. The pH of the media was adjusted to 7.0 and the media was autoclaved at 121 °C for 20 mins.

The 2xYT media contained 16 g/L tryptone, 10 g/L yeast extract and 5 g/L NaCl. The pH of the media was adjusted to 7.0 and the media was autoclaved at 121 °C for 20 mins.

1% (v/v) overnight grown cell culture was inoculated into 1mL 2xPY or defined medium in a 14mL BD Falcon™ tube. Cells were grown at 37°C (or 28°C) with 300 rpm (or 225 rpm) shaking, and induced with L-arabinose when OD600 reached the range of 0.5~1.0. The media were supplemented with 100 mg/L Ampicillin and 34 mg/L chloramphenicol to maintain the plasmids pBAD-SIDF and pAC-LYC (or pBAD-LYC) respectively. Besides, 25 mg/L kanamycin was added into the PTS01, PTS02 and PTS03 strain. For the strain that was producing amorpha-

4,11-diene, the cells were grown at 28 °C in order to decrease the evaporation of dodecane.

Cell density was monitored spectrophotometrically at 600 nm.

4.5.4 Lycopene and amorpha-4,11-diene assay

Intracellular lycopene content was extracted from 10-100 µL bacterial culture (depending on the content of lycopene in the cells) in the stationary phase (24 h for LB, 2xYT and 2xPY, 48 h for 2xM9 and other defined media after induction). Cell pellets were washed with PBS, and lycopene was extracted by acetone. The lycopene content was quantified through absorbance at 472 nm using a 96-well microplate reader (Spectra Max 190, Molecular Devices) and concentrations were calculated by interpreting the standard curve. Cell mass was calculated by correlating DCW with a reading of OD600 and used for calculating specific productivity (µg lycopene/g DCW).

The strains producing amopha-4,11-diene were cultured in defined media with the addition of another 25% (v/v) organic dodecane phase to extract amopha-4,11-diene [138]. Amopha-4,11-diene was extracted by diluting 5 µL dodecane phase into 495 µL ethyl acetate and analyzed on an Agilent 7980A gas chromatograph equipped with an Agilent 5975C mass spectrometer (GC/MS), by scanning at only 189 m/z ion (single-ion monitoring, SIM mode), using trans-caryophyllene as the internal standard. The temperature program for the GC/MS analysis was essentially as described [130].

4.5.5 RNA purification and cDNA synthesis

Total RNA from *E. coli* was prepared using TRIzol® reagent (Invitrogen) according to the manufacturer's instructions. Total RNA was collected from samples in quadruplicate at each treatment time point. RNA concentration was quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific). Eight hundred ng of total RNA were reverse transcribed in a total volume of 20 µL containing ImpromII (Promega) for 60 min at 42°C according to the manufacturer's instructions. The reaction was terminated by heating at 70°C for 10 min.

4.5.6 Quantitative real-time PCR

The cDNA levels were analyzed using a BioRad iCycler 4 Real-Time PCR Detection System (Bio-Rad) with SYBR Green I detection. Each sample was measured in duplicate in a 96-well plate (Bio-Rad) in a reaction mixture (25 µL final volume) containing 1x XtensaMix-SG (BioWORKS, Singapore), 200 nM primer mix, 2.5 mM $MgCl_2$, 0.75 U of iTaq DNA polymerase (iDNA). Real time PCR was performed with an initial denaturation of 3 min at 95ºC, followed by 40 cycles of 20s at 95ºC, 20s at 60ºC, and 20s at 72ºC. The primers used for real time PCR were given in supplementary data, and the reference gene used for normalization of real time PCR data was *cysG* [156].

The plasmid copy number was obtained by normalizing the copy number of plasmid resistant genes (ampicillin resistant gene for pBAD-SIDF and

chloramphenicol resistant gene for pAC-LYC) to the copy number of chromosomal gene *ispE*. Standard curves were constructed using linearized plasmids.

4.5.7 Metabolites assay by LC-MS

Metabolites were analyzed on the UPLC (Waters ACQUITY UPLC) – (TOF) MS (Bruker micrOTOF II) platform. Cell pellets were collected at a density of ~ 1 OD by centrifugation at 16,000 *g* for 1 min. The supernatant was then removed and the cell pellet was re-suspended in 30uL of double distilled water. 120 uL of 100% methanol was then added to the cell suspension and incubated at room temperature for 10 min. The solution was centrifuged at 16,000 *g* for 2 min, and 50 uL of the clear supernatant was collected for analysis. Five microliter of the clarified supernatant was injected into the reverse phase UPLC C18 column (Waters CSH C18 1.7 µm, 2.1 mm x 50 mm) and separation was carried out at a flow rate of 0.15 mL/min with an aqueous mobile phase consisting of 15 mM acetic acid and 10mM tributylamine. Elution was carried out with methanol, in an increasing concentration gradient, as shown in Table 4-8. Electrospray ionization was used and (TOF) mass spectrometry setting was essentially as described [157].

Table 4-9. Mobile phase gradient used for the separation of DXP intermediates.

Step	Accumulative Time (mins)	Aqueous solution*	Methanol
1	1.8	100%	0
2	3.1	60%	40%
4	4.9	60%	40%
5	5.4	10%	90%

6	9.5	10%	90%
7	10	100%	0

*Aqueous solution: 15mM acetic acid and 10mM tributylamine

Chapter 5. TRANSPORTER ENGINEERING FOR THE PRODUCTION OF ISOPRENOIDS

5.1 Introduction

Microbial production is a sustainable approach to produce various chemicals and biofuels. Using tools developed in metabolic engineering, significant improvements have been achieved in titers and yields of these products. However, the high-titer production of the heterologous compounds often leads to inhibition of cell growth [35]. In addition, it was reported that for bio-based products, the cost of recovery and refinement is substantial. For example, in the manufacturing of carboxylic acid, 50% to 60% of production costs are attributed to recovery and refinement [158]. Bacteria have evolved the capability to secrete various compounds, native or heterologous, through the multidrug resistant (MDR) efflux pumps system [159]. There are five families of efflux pumps in *E. coli* that are associated with MDR. They are the ATP-binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the multidrug and toxic-compound extrusion (MATE) family, the small multidrug resistance (SMR) family and the resistance nodulation division (RND) family (Table 5-1) [160]. The structures of efflux pumps can be made up of either a single component or multiple components. For example, the putative fosmidomycin efflux protein *fsr* has only a single component. Whereas a*crAB-tolC* complex, the most studied RND family pumps in *E. coli*,

consists of three components - the inner membrane protein *acrB*, the periplasmic adapter (or membrane fusion protein) *acrA* and outer membrane protein *tolC* (Figure 5-1). The substrates of the MDR efflux pumps are not restricted to antibiotics, but include many environmental and host-derived xeno-metabolites. Thus, these pumps have a general detoxification function in bacteria [161, 162] and have the potential to efflux various heterologous products.

Table 5-1. Five categories of membrane transporters

Acronym	Full name
MFS	Major Facilitator Superfamily
SMR	Small Multidrug Resistance
RND	Resistance Nodulation Cell Division
ABC	ATP-Binding Cassette
MATA	Multidrug and Toxic Compound Extrusion

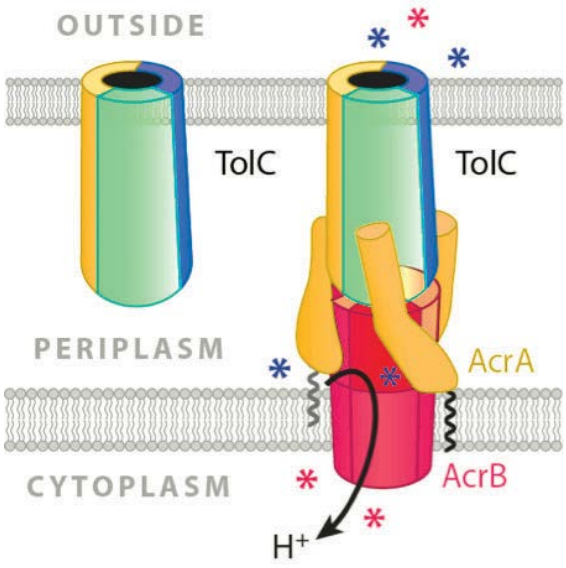


Figure 5-1. Illustration of *acrAB-tolC* RND pump system.
TolC (colourful), the outer membrane protein; *acrA* (yellow), the periplasmic adaptor (or membrane fusion protein); *acrB* (pink), the inner membrane protein. The acrAB-tolC pump system have various substrates (the asterisks with different colors).

Recently, there are increasing reports of the manipulations of different transporters to improve tolerance of the hosts to different adverse conditions and to increase yields of various products, summarized in Table 5-2. The use of transporters can also benefit the downstream recovery process by effluxing the products into the media that can then be *in situ* extracted using additional phase (for example, organic phase was used to extract AD [138]). In these studies, some studies are focused on the RND family transporters [35, 163] and some on the ABC family transporters [36], and as yet, there is no study that systematically explored different families in combinations. Although some heterologous efflux pumps were investigated in *E. coli*, the native transporters (*acrAB-tolC* [35], *msbA* [36]) were found to be most useful in *E. coli*. This could simply be due to the poor expression of functional heterologous pumps in different hosts. *E. coli* has many native MDR transporters which are known to handle highly diversified substrates and the potential use of these MDR transporters in efflux of AD has not been fully investigated.

In this study, the effects of single deletion of genes of 21 selected MDR efflux pump on the AD production were investigated. It was found that the deletion of *tolC* significantly decreased the extracellular concentration of AD and concomitantly increased the intracellular concentration of AD. Furthermore, overexpression of *tolC* together with ABC family (*msbA* or *macAB*) or MFS family transporters (*emrAB* or *emrKY*) enhanced AD yield significantly. In addition, we also investigated the effects of single deletion

of these MDR genes on the efflux of phosphorylated compounds such as, 1-deoxy-xylulose-5-phosphate (DXP) and 2-Methyl-D-erythritol-2,4-cyclodiphosphat (MEC), two intermediates in the DXP pathway.

Table 5-2. Recent publications using transporters to enhance the microbial productions

Products	Transporter gene/protein	Performances when overexpressed	Performances when deleted	Reference
n-Butanol	Engineered acrB through directed evolution	The overexpression of variants of the native acrB enhanced growth rates of E. coli in the presence of n-butanol by up to 25%.	/	[34]
Carotenoids (zeaxanthin, Canthaxanthin, β-carotene)	msbA	A ~4.4-fold, significant ($P < 0.03$) increase in secreted canthaxanthin was observed mediated by overexpressed EcoMsbA compared with the control following 96 h of aerobic incubation at 30 °C. A ~2.4- fold increase in secreted zeaxanthin mediated by StMsbA was achieved compared with the nonexpressing control.	/	[36]
Limonene	A previously uncharacterized pump from Alcanivorax borkumensis	The strain that overexpressed the pump produced significantly more limonene than those with no pump.	/	[35]
Cadaverine	CadB/cadaverine-lysine antiporter	The strain overexpressing the efflux pumps showed substantially retarded cell growth and reduced cadaverine production titer.	/	[164]
L-threonine	tdcC/Thr transporter rhtC/threonine exporter	The strain which overexpressed the rhtC gene obtained 50.2% higher production of L-threonine than that obtained without rhtC amplification. And further amplification of the rhtA and rhtB genes resulted in a small increase in L-threonin production.	A 15.6% increase in L-threonine was observed with the deletion of tdcC which uptakes extracellular L-threonine into the cell.	[27]
Amorphadiene and kaurene	AcrAB-tolC, MdtEF-tolC from E. coli and MexAB-OprM from Pseudomonas aeruginosa	Overexpression of heterologous MexB increased kaurene production by 70 %. The combination of tolC and tolC and AcrB improved amorphadiene yield with 118 %, AcrA and tolC and AcrB improved kaurene yield with 104 %.	/	[163]
L-Valine	YgaZH/valine exporter	The overexpression of the ygaZH genes enhanced the production of L-valine. And co-overexpression of the lrp and ygaZH genes was synergistic and led to a 2.13-fold increase in L-valine production.	The level of L-valine production did not change in the ygaZH-deleted strain.	[17]

Putrescine	PotE/ putrescine-ornithine antiporter	Overexpression of potE gene in a low-copy plasmid led to retarded cell growth and lower putrescine production; While overexpression by replacement of native promoter of the speF-potE operon with a strong trc promoter resulted in enhanced putrescine production.	[165]

5.2 Results

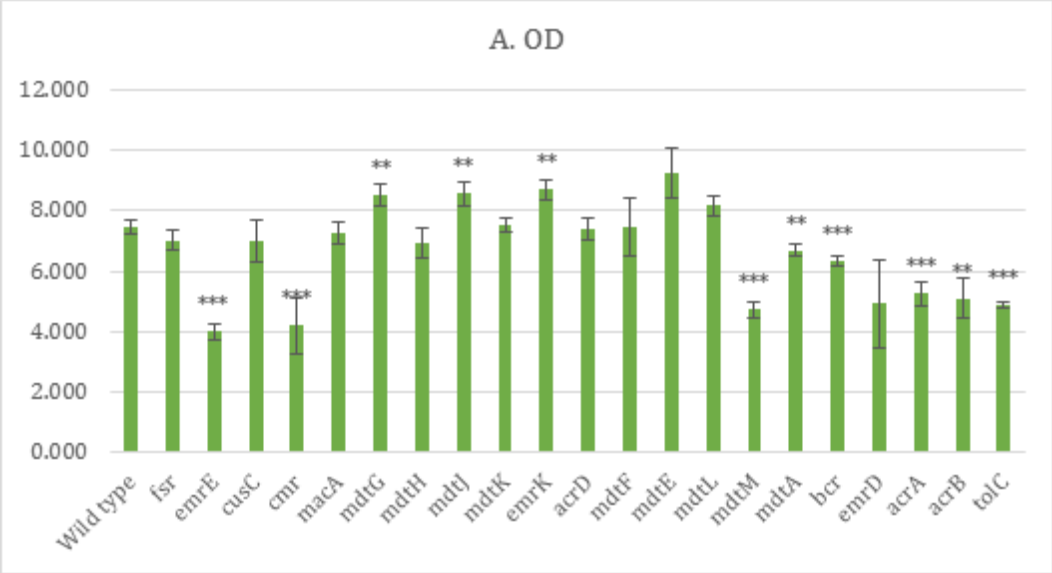
5.2.1 Identifying the pump candidates using single-gene knockout mutant

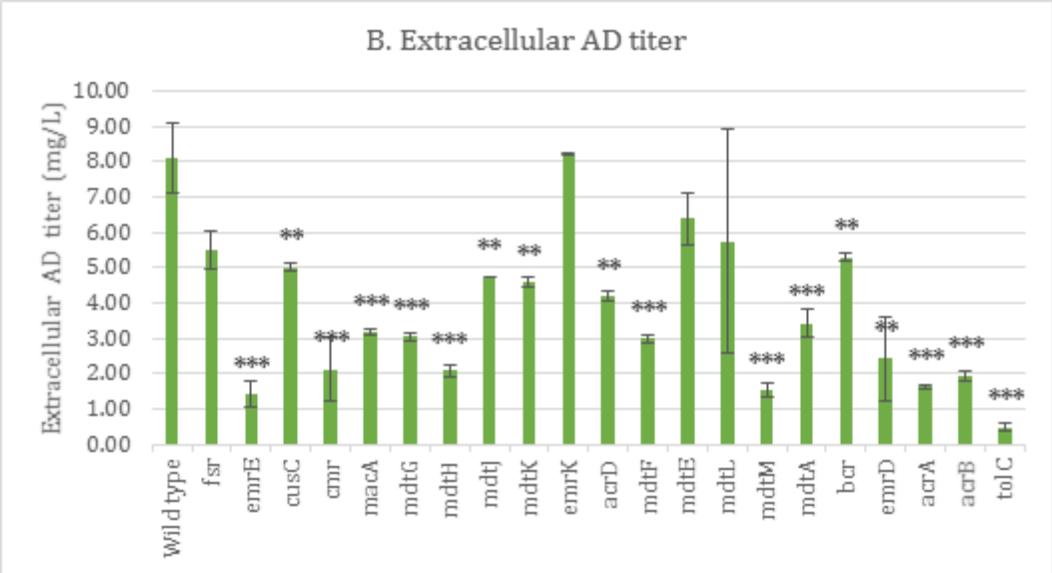
The *BW25113* strain and its 21 single gene knockout mutants were obtained from the Keio collection of gene knockout mutants [166]. These mutants were transformed with two plasmids, pBAD-dxs-idi-ispDF and pACM-ADS. The information about functions and categories of transporter genes was summarized in supplemental data (Table 5-6). The 21 pump encoding genes covered all the five families of efflux pumps, MFS family (*bcr*, *cmr* (or *mdfA*), *emrD*, *emrK*, *fsr*, *mdtG*, *mdtH*, *mdtL*, *mdtM*), RND family (*mdtA* (or *yegM*), *cusC*, *acrD*, *mdtF*, *mdtE*, *acrA* and *acrB*), SMR family (*emrE* and *mdtJ*), MATE (*mdtK*) and ABC family (*macA*), in addition, the outer membrane protein *tolC* was also included in the study.

With single deletion of pump encoding genes, most of the extracellular AD titer was decreased significantly and this could be due to various reasons. As compared to the wild type, the final ODs of the strains with single

deletion of *emrE*, *cmr*, *mdtM*, *mdtA*, *bcr*, *acrA*, *acrB* and *tolC* were significantly reduced (P<0.005) (Figure 5-2A). As a result, the extracellular yield of AD similarly decreased accordingly (P<0.005) (Figure 5-2B). In contrast, cell density was not affected significantly with the single deletion of *cusC*, *macA*, *mdtH*, *mdtK*, *acrD*, *mdtF* or *emrD*. However, the extracellular production of AD of these mutants decreased significantly (P<0.005). Although the ODs of the strains with single deletion of *mdtG* and *mdtJ* were increased (P<0.005), the extracellular titer of AD decreased significantly (P<0.005).

In addition, most of the gene deletions affected the AD productions both intra- and extra-cellularly. The single deletion of *emrE*, *macA*, *mdtG*, *mdtH*, *mdtJ*, *mdtF*, *mdtA*, *bcr* or *acrA* decreased both intracellular and extracellular concentration of AD significantly (P<0.005, Figure 5-3B&C). The knockout of *tolC* increased the intracellular AD production by 8.2 fold and concomitantly decreased extracellular AD production by about 90%. These data strongly supported the notion that *tolC* was involved in the secretion of AD. Interestingly, the copy number of pBAD-SIDF in the mutant strains (*cusC, macA, mdtG, mdtH, mdtJ, acrD, mdtE, mdtL* and *mdtM*) decreased significantly as compared to that in wild type strain (Figure 5-3A). However, the copy number of pACM-ADS remained unchanged to any significance in all the mutant strains (Figure 5-3B).


A. OD
12.000
10.000
8.000
6.000
4.000
2.000
0.000
Wild type
fsr
emrE
cusC
cmr
macA
mdtG
mdtH
mdtJ
mdtK
emrK
acrD
mdtF
mdtE
mdtL
mdtM
mdtA
bcr
emrD
acrA
acrB
tolC


B. Extracellular AD titer
Extracellular AD titer (mg/L)
10.00
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
Wild type
fsr
emrE
cusC
cmr
macA
mdtG
mdtH
mdtJ
mdtK
emrK
acrD
mdtF
mdtE
mdtL
mdtM
mdtA
bcr
emrD
acrA
acrB
tolC

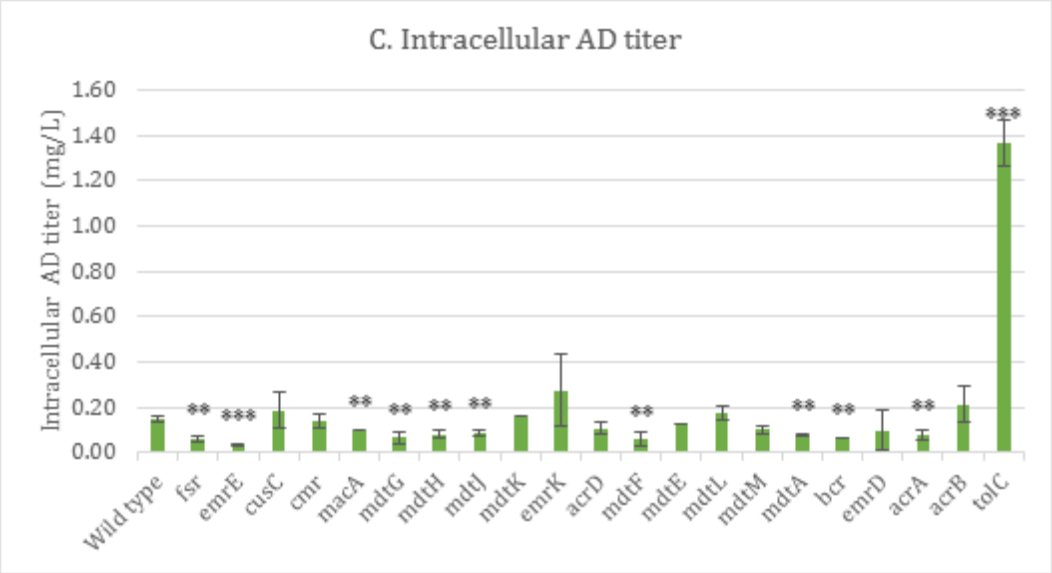


Figure 5-2. The effect of single-MDR-gene deletion on the cell density and the AD production.
(A) OD at 48 h in single-MDR-gene knockout mutants. All the strains carried the two plasmids, pBAD-SIDF and pACM-ADS. (B) Extracellular AD production in single-MDR-gene knockout mutants. (C) Intracellular extracellular AD production in single-MDR-gene knockout mutants. Stars indicated significance in two tailed Student's t-test, and all the t-tests were calculated between knockout mutant and the wild type, **P<0.005 and ***P<0.0005.

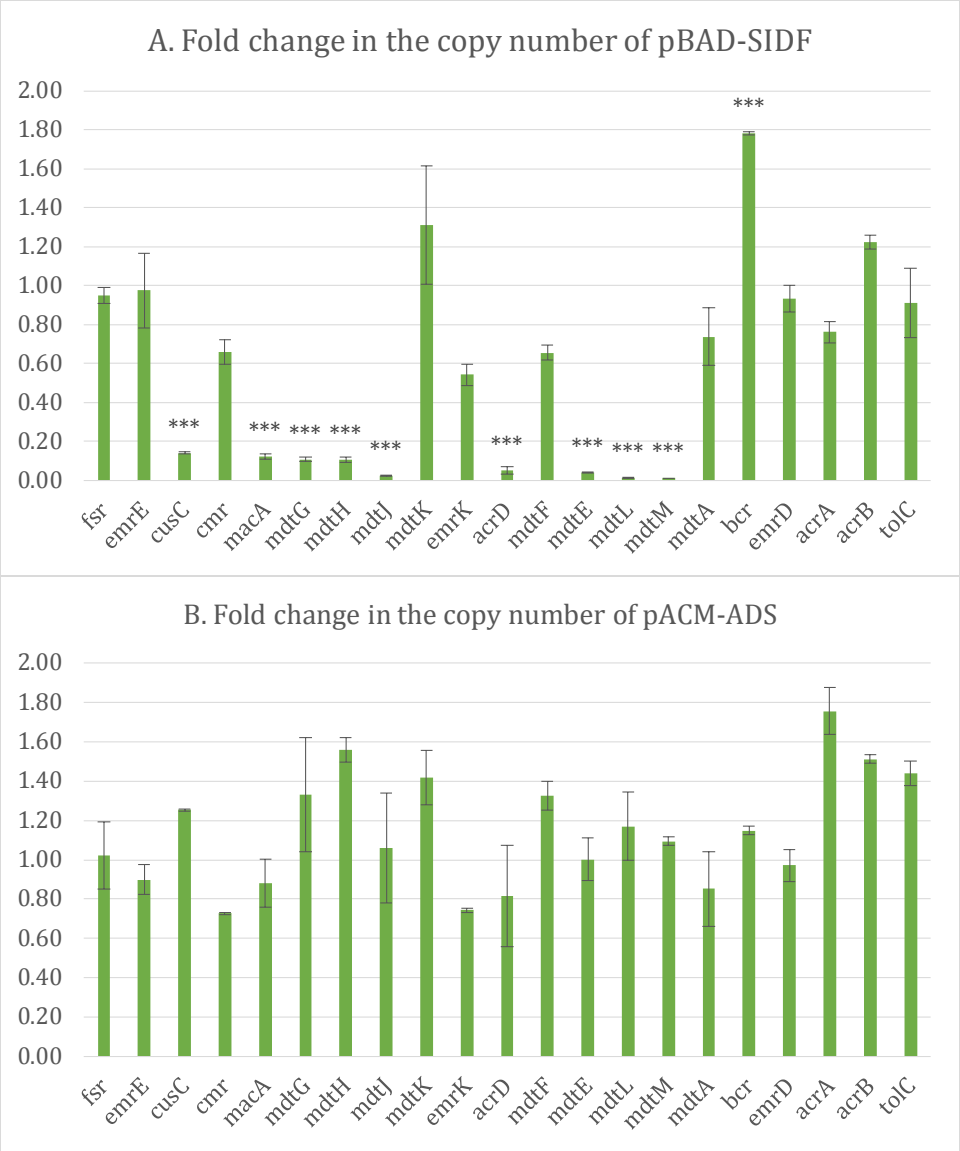


Figure 5-3. The effect of single-MDR-gene knockout on plasmid copy number. (A) Fold change in the copy number of plasmid pBAD-SIDF. (B) Fold change in the copy number of plasmid pACM-ADS. The plasmid copy number were calculated as described in methods and materials. And they are presented as the fold change compared to that in the wild type strain.

5.2.2 Reconstitution of tolC increased the efflux of AD in tolC mutant

If *tolC* was involved with the efflux of AD, we speculated that reconstitution of *tolC* into the *tolC* mutant should increase the efflux in the mutant. The

tolC gene was subcloned into RK2 vector and the plasmid was transformed into the *tolC* knockout mutant (tolC+ strain). In line with the hypothesis, the reconstitution of *tolC* in the mutant restored the level of effluxed AD from cells (Figure 5-4). Similarly, the knockout of *tolC* in *BW25113* strain resulted in the extracellular AD concentration decreased from 2.31 g/L (wild type strain) to 0.65 g/L (tolC strain) and the intracellular AD concentration increased from 0.07 g/L (wild type strain) to 2.06 g/L (tolC strain). In the presence of RK2-tolC plasmid, the extracellular AD increased from 0.65 g/L (tolC strain) to 2.37 g/L (tolC+ strain), a level similar to that in the wild type strain. And the intracellular concentration of AD was decreased from 2.06 g/L (tolC+ strain) to 0.22 g/L when *tolC* was overexpressed. Interestingly, the reconstitution of *tolC* also restored cell growth as evident by the comparable cell density of the *tolC* mutant (TolC+) and the wild type. These results suggested that AD was not passively diffused out of the cell but effluxed with the aid of *tolC*.

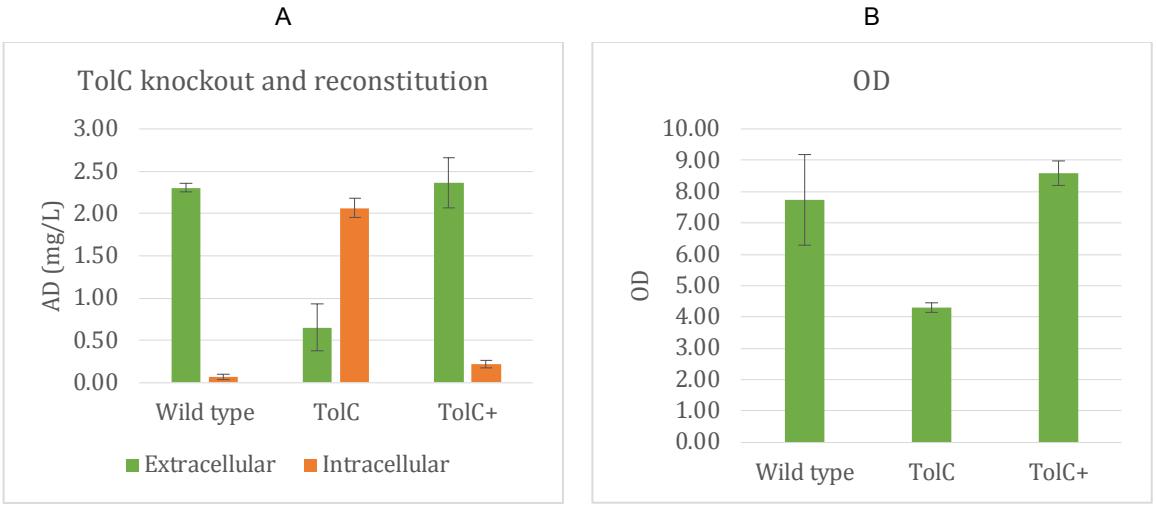


Figure 5-4. Supplementation of *tolC* in *tolC* knockout mutant.

"tolC" is the *tolC* knockout mutant, "tolC+" is the tolC mutant supplemented with *tolC* gene in the RK2 plasmid. All the wild type, tolC and tolC+ strains carried plasmid pACM-ADS, in addition, the tolC+ also carried plasmid RK2(amp)-T7-tolC. One fifth (v/v) dodecane was added into the media to capture the extracellular AD.

5.2.3 Effects of reconstitution of efflux pumps in AD-producing strains

As deletion of *tolC* resulted in decreased efflux of AD, it was worth investigating in greater detail of the effects of reconstitution of *tolC* on AD production. *TolC* is an outer membrane protein and can form complexes wtih three types of transport systems: ABC family (*macAB*), RND family (*acrAB*, *acrEF*, *mdt*ABC, *acr*AD and *mdtE*F) and MFS family (*emrAB*, *emrKY* and *entS*) (Table 5-3) [167]. During the export of AD into the medium, AD has to traverse both inner membrane and outer membrane. Since the outer membrane channel *tolC* does not possess any substrate specificity, there may be some inner membrane proteins that regulate the export of AD across the inner membrane. Hence, it will be highly desirable to identify these inner membrane pumps involved in AD export. It is predicted then that the overexpression of these pumps together with *tolC* will likely to increase the yield or productivity of AD by driving the reversible reactions (*dxr* [168], *ispG* [169] and *ispH* [170]) forward and reducing product inhibition [98] simply by shuttling AD out of cells.

In addition to the *tolC*-related transporters, we also investigated a few MDR transporters (*emrE*, *emrD*, *fsr*, *cmr*, *mdtM* and *mdtH*) and the transporters in lipopolysaccharide (LPS) transport system (*lptD*,

lptCABFG, and *msbA*). *LptD* is an outer membrane protein that is required for LPS transport to the cell surface in *E. coli*. In addition, *LptCBFG*, an ABC transporter, is responsible for the release of LPS from the inner membrane [171]. *LptA* is a periplasmic protein which acts as a LPS chaperone or a bridge between the inner membrane and the outer membrane [172]. *MsbA* facilitates flipping of the lipid A-core structure across the inner membrane [173]. The reason we chose to examine LPS transporters was that ABC exporters are known to export extremely hydrophobic compounds, such as lipids, drugs and steroids, and AD was hydrophobic compound [174]. Furthermore, the overexpression of *msbA* has been reported to enhance the secretion of hydrophobic carotenoids [36].

Table 5-3. *TolC*-related transporter system

Transporters	Full name	Categories	Substrates
macAB-tolC	*MacAB-tolC* macrolide efflux transport system	ABC	Macrolide
AcrAB-tolC	*AcrAB-tolC* multidrug efflux transport system	RND	Organic solvents , dyes and detergents as well as lipophilic antibiotics including novobiocin, erythromycin, fusidic acid and cloxacillin
EmrAB-tolC	*EmrAB-tolC* multidrug efflux transport system	MFS	Hydrophobic compounds such as carbonyl cyanide m-chlorophenylhydrazone, tetrachlorosalicylanilide, organomercurials and nalidixic acid
AcrIC	*AcrEF-tolC* multidrug efflux transport system	RND	Highly homologous to the *acrAB* multidrug efflux system
EmrKY-tolC	multidrug efflux transport system	MFS	Tetracyclin, chloramphenicol, or salicylate
MdtABC (yegMNO)-tolC	*MdtABC-tolC* multidrug efflux transport system	RND	Salt derivatives, sodium dodecyl sulfate (SDS) and novobiocin
EntS-tolC	*EntS-tolC* Enterobactin Efflux Transport System	MFS	Mitomycin C, enterobactin

AcrAD-tolC	*AcrAD-tolC* multidrug efflux transport system	RND	Deoxycholate, SDS, novobiocin, kanamycin, tetracycline, nalidixic acid, norfloxacin, fosfomycin, bile acids, fusidic acid and progesterone
MdtEF-tolC	*MdtEF-tolC* multidrug efflux transport system	RND	Deoxycholate

As shown in Figure 5-5, AD production in both early stage (8 h) and final stage (24 h) increased by more than 3 fold in the strains that overexpressed *tolC-msbA*, *tolC-emrAB*, *tolC-emrKY* or *tolC-macAB*. In addition, the overexpression of *tolC*, *lptCABEF*, *tolC-mdtE*F, *tolC-acrEF* or *msbA* increased AD yield by more than 2 fold. Noteworthy, the inner membrane LPS transport protein *msbA* has not been reported to interact with the outer membrane *tolC*, yet the co-overexpression of *tolC* and *msbA* have synergistic effect resulting in more AD produced than their single overexpression. AD production was also increased noticeably when *lptD*, *tolC-acrD* or *tolC-acrAB* was overexpressed. However, when *emrE*, *emrD*, *fsr*, *cmr*, *mdtM* or *mdtH* was overexpressed, the AD production was nearly abolished. The cell density of all the strains was similar, except that the OD of the strain with overexpression of *mdtEF-tolC*, *acrEF-tolC* and *acrAB-tolC* was lower. The results indicated that unexpectedly, many *tolC*-related transporter system and LPS transporter *msbA* have positive effects on the AD production, especially ABC transporters (*msbA* and *macAB*) and MFS transporters (*emrAB* and *emrKY*) significantly increased the AD production.

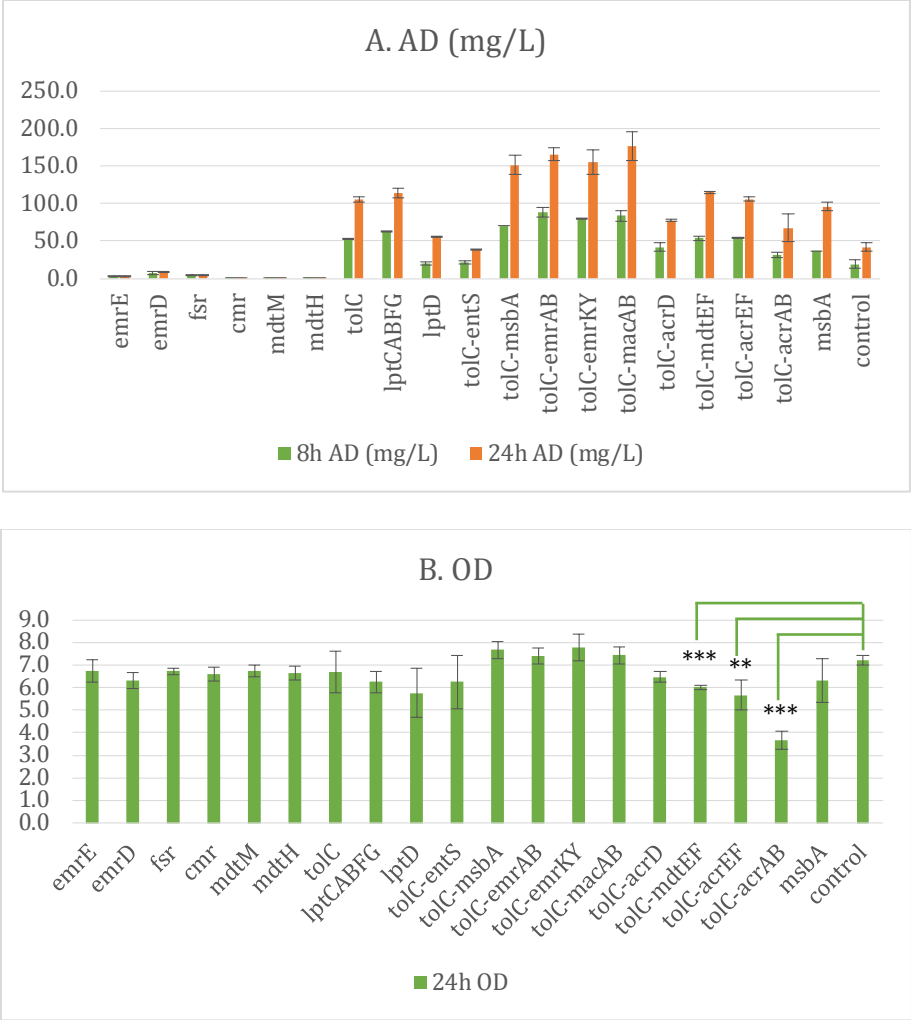


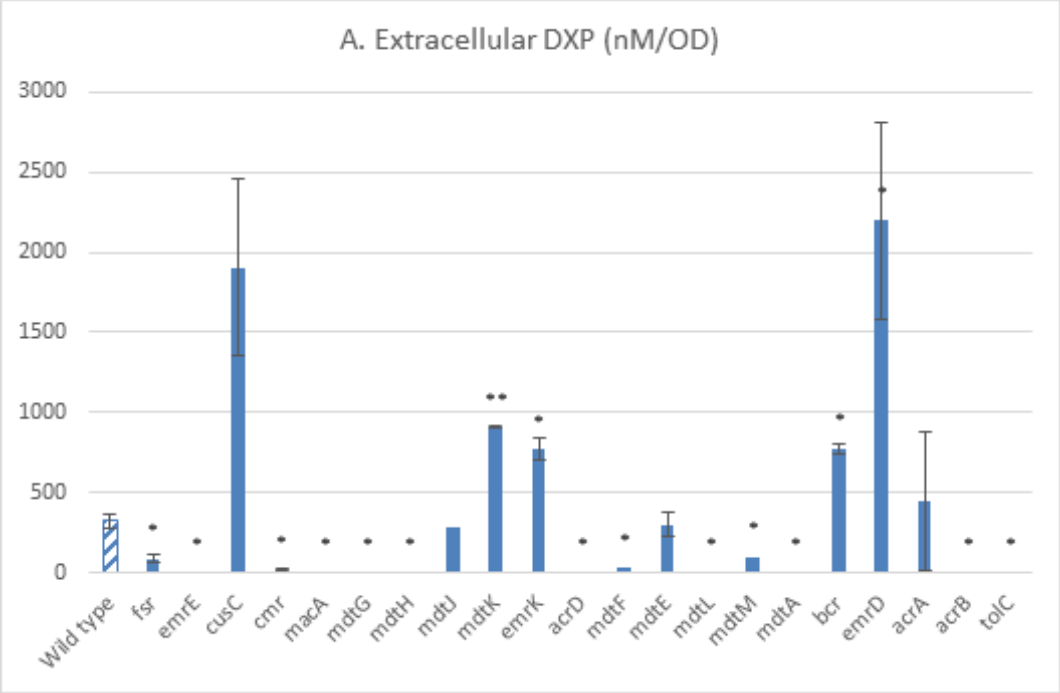
Figure 5-5. AD production in strains with supplemental expression of efflux pumps. (A) AD titers in different strains. (B) Cell density of different strains at 24 h. Stars indicated significance in two tailed Student's t-test, and all the t-tests were calculated between the strain with the overexpression of pump-coding gene and control strain, **P<0.005 and ***P<0.0005.

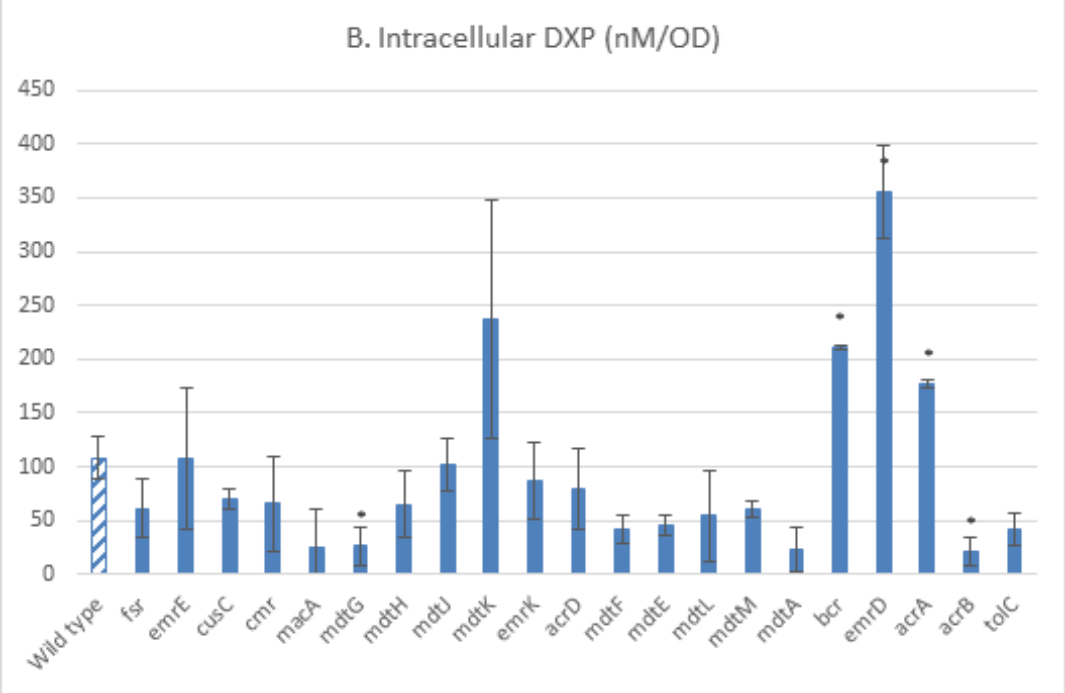
5.2.4 Preliminary attempts to identify pump candidates for the efflux of DXP and MEC

Recently, our group found that the intermediates DXP and MEC in the DXP pathway were exported and accumulated in the growth medium [129, 157]. As DXP and MEC are phosphorylated compounds that do not

passively diffuse across cell membrane, it is likely that some efflux pump is responsible for the export out of cells. Identifying efflux pump of MEC is useful for both fundamental study of membrane transporters of *E. coli* and potential applications in metabolic engineering and synthetic biology. For instance, knockout of the pump would force these compounds to stay inside the cell, and thus the flux in the DXP pathway may be enhanced resulting in higher yields of final products.

In order to systematically evaluate the potential roles of a variety of transporters in the efflux of DXP, gene knock-out strains of the same genetic background were obtained from the Keio collection of gene knockout mutants [166]. Preliminary analyses (Figure 5-6) demonstrated that the accumulation of extracellular DXP was less than the control strain in some of the knock-out mutants (e.g., *fsr, emrE, cmr,* etc). However, the corresponding intracellular DXP levels in the corresponding strains were either similar or even higher than the controls (eg. *bcr*, *emrD*, etc). As this is a preliminary study, further work to evaluate if these particular strains that showed both an increase in efflux and the simultaneous accumulation of intracellular DXP (e.g., *bcr*, *emrD*) also show an increase in productivity of downstream isoprenoids.


A. Extracellular DXP (nM/OD)
3000
2500
2000
1500
1000
500
0
Wild type
fsr
emrE
cusC
cmr
macA
mdtG
mdtH
mdtJ
mdtK
emrK
acrD
mdtF
mdtE
mdtL
mdtM
mdtA
bcr
emrD
acrA
acrB
tolC


B. Intracellular DXP (nM/OD)
450
400
350
300
250
200
150
100
50
0
Wild type
fsr
emrE
cusC
cmr
macA
mdtG
mdtH
mdtJ
mdtK
emrK
acrD
mdtF
mdtE
mdtL
mdtM
mdtA
bcr
emrD
acrA
acrB
tolC

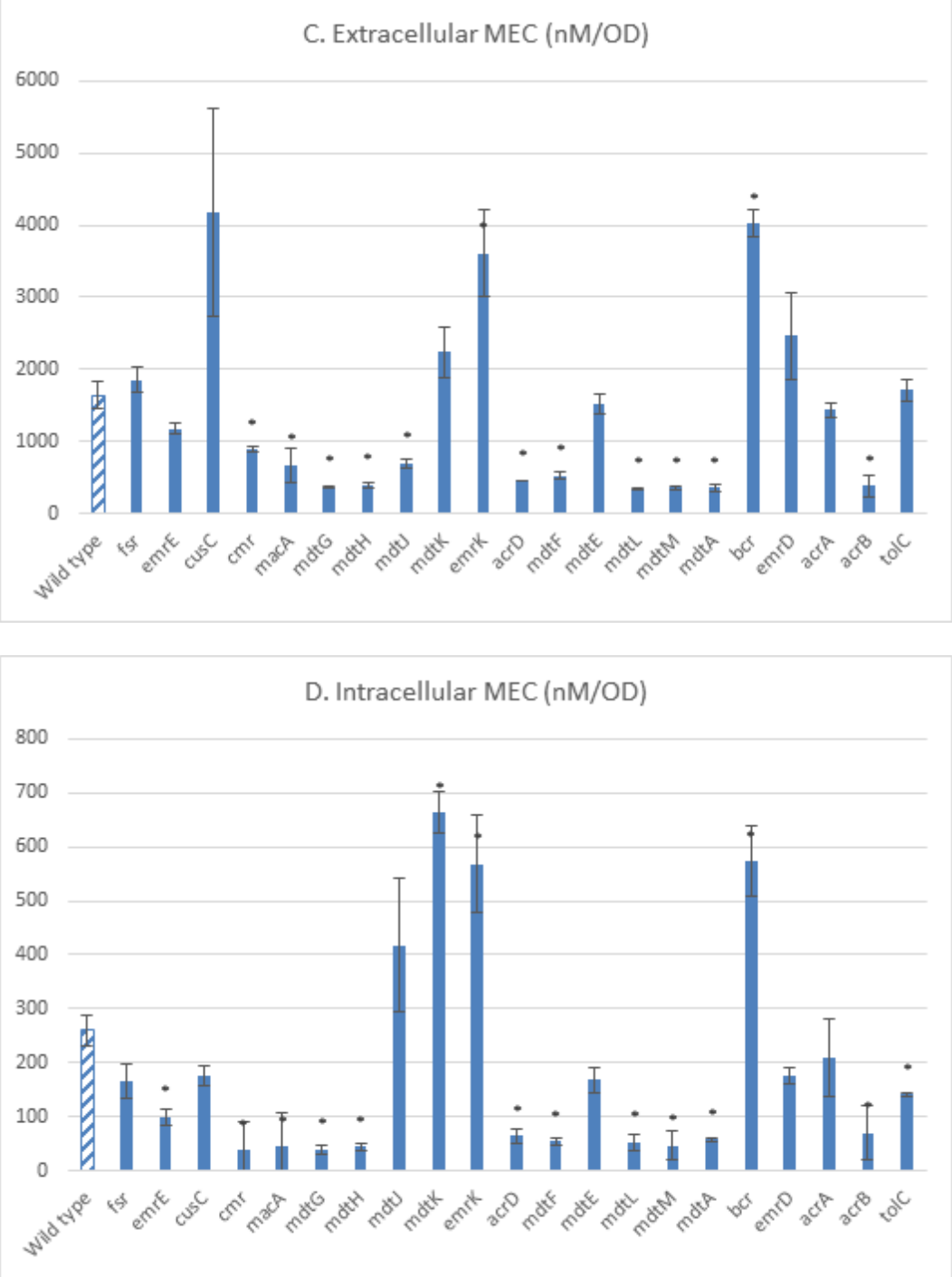


Figure 5-6. Fold change of the concentration of DXP in the DXP pathway. (A) Extracellular DXP (nM/OD), (B) Intracellular DXP (nM/OD) (C) Extracellular MEC (nM/OD), (D) Intracellular MEC (nM/OD). Stars indicated significance in two tailed Student's t-test, and all the t-tests were calculated between knockout mutant and the wild type, *P<0.05 and **P<0.005.

5.3 Discussion

Toxicity is a common and significant challenge with cell-based production [30, 35, 175]. As molecules have to traverse the inner and outer membranes of *E. coli*, the simple strategies of reducing the permeability of the outer membrane through decreasing porin expression or LPS modifications are not very useful in removing such toxic products in cells [175]. Microbial efflux pumps may serve as an alternative route to improving the export of native and heterologous compounds as they endow microbes with the ability to extrude a wide range of compounds such as antibiotics, solvents, cationic and lipophilic products [176-179]. With this, we hypothesized that *E. coli* has the capability to export AD or other similar heterologous hydrophobic compounds via native efflux pumps or other transport mechanisms.

For a very long while, it was assumed that the volatile heterologous AD passively diffuses out of cells through cell membrane in *E. coli*. However, this is proven incorrect as we now show for the first time, that *tolC* is closely involved with AD efflux. Without *tolC*, even adding an organic layer into the cell culture to extract AD, the majority of AD (more than 80%) was found to be entrapped intracellularly. The outer membrane protein *tolC* is known to be involved in the extrusion of various harmful substances such as antibiotics [180], dyes [181], organic solvents [182] and detergents [183]. Furthermore, it was found that the export of some metabolites (porphyrin [182], cysteine [183], indole [183] and siderophores [184])

involved tolC. TolC is also involved in the extrusion of intracellular metabolites such as signaling molecules such as cyclic AMP [185]. Thus, *tolC* mutant has pleiotropic effects, such as increasing susceptibility to antibiotics, cell division defects, changes in the expression of outer membrane porins, sensitivity to acid and decreased virulence [184].

Consistent with the *tolC* deletion results, the overexpression of *tolC* transport complexes, *tolC-emrAB*, *tolC-emrKY* or *tolC-macAB,* significantly improved the AD production. AD production was also increased when single gene *tolC*, *lptD* or *msbA* was overexpressed. *MsbA* is an essential ABC transporter that exports charged lipid A across the inner membrane in *E. coli* [185]. It was reported that the overexpression of *msbA* was able to enhance the secretion of some carotenoids [36]. Interestingly, co-overexpression of *msbA* and *tolC* has a synergetic effect on the AD production. These results indicated that multiple transporters are likely to be involved in the export of AD from cytoplasm. The intrinsic property of MDR transporters is that they are able to extrude various types of toxic heterologous compounds (e.g. antibiotics, biocides) [186]. Therefore, it is possible that more than one native transporter in *E. coli* is involved in the export of AD. Another interesting finding was that the overexpression of *emrE*, *emrD*, *fsr*, *cmr*, *mdtM* and *mdtH* nearly abolished the AD production due to unknown mechanisms, which could be due to their role in extruding the inducer IPTG. We have also overexpressed these transporters in lycopene-producing strains (Figure 5-5). However, none of these

transported was able to export this extremely hydrophobic compound, lycopene (data not shown). Similarly, the overexpression of another efflux pump, *msbA,* did not enhance the export of lycopene out of *E. coli* [36].

Interestingly, our previous study demonstrated the efflux of the metabolic intermediates in the DXP pathway, DXP and MEC [129, 157]. The efflux of these two intermediates reduces the carbon fluxes in the pathway and thus decreases the final yield and productivity of downstream products. Thus, the identification and subsequent knock out or knock down of responsible efflux pumps should extend previous observations.

Preliminary studies showed that the absence of some pumps may affect the levels of extracellular and intracellular levels of DXP. These observations must be further confirmed and extended before any conclusions can be confidently deduced. Furthermore, in the next chapter, we attempted to use mathematical simulation of the DXP pathway to study the efflux and the kinetic behavior of the intracellular pathway.

5.4 Conclusion

Through single gene deletion, we found that outer membrane protein *tolC* was closely involved in the efflux of AD in *E. coli*. And the overexpression of *tolC-emrAB*, *tolC-emrKY* or *tolC-macAB* significantly improved the AD production. This study indicated that pump engineering is an effective approach in metabolic engineering to construct high-yield AD producers. Beyond this, the strategy herein may be used to enhance the production of other kinds of heterologous compounds.

5.5 Methods and materials

5.5.1 Strains, plasmids and transporter genes

BW25113 (Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ- rph-1 Δ(rhaD-rhaB)568 hsdR514) and its single gene knockout mutant were used in the screening experiments. All of them were from the Keio collection of gene knockout mutants [187]. *K12 MG1655 ΔrecAΔendA DE3* strain [13] was used for AD production with overexpression of transporters.

Selected genes and operons were amplified by PCR from *MG1655 ΔrecAΔendA DE3* genomic DNA with their putative upstream ribosome binding sites. All the genes and operons were inserted into modified RK2 (amp) plasmid [188] with the Pm promoter replaced with TM1 (T7 variant) promoter. *TolC* was inserted into RK2 first to generate RK2-tolC. Selected operons (*macAB*, *acrAB*, *emrAB*, *arcEF*, *emrKY*, *entS*, *acrD*, *mdtEF* and *msbA*) were also inserted into RK2-tolC to generate artificial operons with *tolC*. The genes (*MdtM*, *mdtH*, *fsr*, *cmr*, *emrD*, *emrE*, *msbA* and *lptD*) were inserted directly into RK2 plasmid. The operons *lptCAB* and *lptFG* were co-inserted into RK2 plasmid. The gene ADS was inserted into modified pAC plasmid with the araBAD promoter to obtain pACM-ADS. All the constructions used CLIVA methods [129] and the phosphorothioated primers used were shown in Table 5-5. The transporter genes studied in this chapter and their main functions are listed in Table 5-6.

Table 5-4. Bacterial strains, plasmids and primers used in this study

Strains, plasmids or primers	Descriptions	Source or references
Strains		
BW25113	Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ- rph-1 Δ(rhaD-rhaB)568 hsdR514	[166]
fsr	BW25113 Δfsr::kan	[166]
emrE	BW25113 ΔemrE::kan	[166]
cusC	BW25113 ΔcusC::kan	[166]
cmr	BW25113 Δcmr::kan	[166]
macA	BW25113 ΔmacA::kan	[166]
mdtG	BW25113 ΔmdtG::kan	[166]
mdtH	BW25113 ΔmdtH::kan	[166]
mdtJ	BW25113 ΔmdtJ::kan	[166]
mdtK	BW25113 ΔmdtK::kan	[166]
emrK	BW25113 ΔemrK::kan	[166]
acrD	BW25113 ΔacrD::kan	[166]
mdtF	BW25113 ΔmdtF::kan	[166]
mdtE	BW25113 ΔmdtE::kan	[166]
mdtL	BW25113 ΔmdtL::kan	[166]
mdtM	BW25113 ΔmdtM::kan	[166]
mdtA	BW25113 ΔmdtA::kan	[166]
bcr	BW25113 Δbcr::kan	[166]
emrD	BW25113 ΔemrD::kan	[166]
acrA	BW25113 ΔacrA::kan	[166]
acrB	BW25113 ΔacrB::kan	[166]
tolC	BW25113 ΔtolC::kan	[166]
MG1655	K12 MG1655 ΔrecAΔendA DE3	[13]
Plasmids		
RK2-emrE	RK2 carrying emrE under TM1 promoter control	This study
RK2-emrD	RK2 carrying emrD under TM1 promoter control	This study
RK2-fsr	RK2 carrying fsr under TM1 promoter control	This study
RK2-cmr	RK2 carrying cmr under TM1 promoter control	This study
RK2-mdtM	RK2 carrying mdtM under TM1 promoter control	This study
RK2-mdtH	RK2 carrying mdtH under TM1 promoter control	This study
RK2-tolC	RK2 carrying tolC under TM1 promoter control	This study
RK2-lptCABFG	RK2 carrying lptCABFG artificial operon under TM1 promoter control	This study
RK2-lptD	RK2 carrying lptD under TM1 promoter control	This study
RK2-msbA	RK2 carrying msbA under TM1 promoter control	This study
RK2-tolC-entS	RK2 carrying tolC-entS artificial operon under TM1 promoter control	This study
RK2-tolC-msbA	RK2 carrying tolC-msbA artificial operon under TM1 promoter control	This study
RK2-tolC-emrKY	RK2 carrying tolC-emrKy artificial operon under TM1 promoter control	This study
RK2-tolC-acrD	RK2 carrying tolC-acrD artificial operon under TM1 promoter control	This study
RK2-tolC-acrAB	RK2 carrying tolC-acrAB artificial operon under TM1 promoter	This study

	control	
RK2-tolC-emrAB	RK2 carrying tolC-emrAB artificial operon under TM1 promoter control	This study
RK2-tolC-macAB	RK2 carrying tolC-macAB artificial operon under TM1 promoter control	This study
RK2-tolC-mdtEF	RK2 carrying tolC-mdtEF artificial operon under TM1 promoter control	This study
RK2-tolC-acrEF	RK2 carrying tolC-acrEF artificial operon under TM1 promoter control	This study
pBAD-dxs-idi-ispDF	pBAD carrying dxs-idi-ispDF artificial operon under araBAD promoter control	[24]
pACM-ADS	pAC carrying ADS under araBAD promoter control	This study
pACM-SAAI	pAC carrying dxs-idi-ispA-ADS under T7 promoter control	[129]
pET-lacI	pET plasmid carrying lacI	[129]

Table 5-5. Phosphorothioated oligos used in this study

Oligos	Sequences
I-T7P(-)-R	AACAAAATTATT*TCTAGAGGG*GAA
I-tolC(-)-R	TCAGTTACG*GAAAGGG*TTAT
I-T7T(-)-F	AAAGGAAG*CTGAGTTGG*CTG
I-tolc(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*AAATACTGCTTCACCACAAG
I-tolc(t7tf)-R	CCAACTCA*GCTTCCTTT*TCAGTTACGGAAAGGGTTAT
I-mdtM(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*CTCACCTGCCTTTCGTCA
I-mdtM(t7tf)-R	CCAACTCA*GCTTCCTTT*AAATCACTGCTCCTCCACTA
I-mdtH(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*CGGCAATGCAGAATAGAAA
I-mdtH(t7tf)-R	CCAACTCA*GCTTCCTTT*CAGTAAGGGCAGTGATCGTAA
I-fsr(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*TAAGATCATTCTTTACATCGCAACC
I-fsr(t7tf)-R	CCAACTCA*GCTTCCTTT*AACGATGCAGCCAAAGTTTG
I-cmr(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*CTTTCTCGTTAGCTGCGCTT
I-cmr(t7tf)-R	CCAACTCA*GCTTCCTTT*ACAAAGCAGTCAGGCATTTT
I-emrD(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*TTTATGGGAATGCGTAGTG
I-emrD(t7tf)-R	CCAACTCA*GCTTCCTTT*GGAGCTGATGACGATGCT
I-emrE(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*AAACTGTAGTACAATTCTC
I-emrE(t7tf)-R	CCAACTCA*GCTTCCTTT*CATATTTTAGTCGTTTAG
I-msbA(tolcr)-F	CCCTTTCC*GTAACTGA*TTTCAACAAATGCTGGTT
I-msbA(t7tf)-R	CCAACTCA*GCTTCCTTT*GGATTCACCAGACCAGAT
I-msbA(t7pr)-F	CCCTCTAGAAAT*AATTTTGTT*TTTCAACAAATGCTGGTT
I-lptCAB(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*CTGGATGAAGCCAAAGGG
I-lptCAB(fgf)R	TCATGGCTTAA*ACGTCGCAAAC*TTCTACCCTA
I-lptFG(cabr)F	GTTTGCGACGT*TTAAGCCATGA*AACAAGCT
I-lptFG(t7tf)-R	CCAACTCA*GCTTCCTTT*GAAGGTAAGTCAGGCGAGAT
I-lptD(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*CAGGATTAACACTAGCGTAGA
I-lptD(t7tf)-R	CCAACTCA*GCTTCCTTT*GGCGATACCGAGAAGCAG
I-macAB(tolcr)-F	CCCTTTCC*GTAACTGA*GGTTTTGTTGATATTTCGTTG
I-macAB(t7tf)-R	CCAACTCA*GCTTCCTTT*CAGTCGCATAGCTCTTCTGT
I-acrAB(tolcr)-F	CCCTTTCC*GTAACTGA*AACGCAGCAATGGGTTTAT

I-acrAB(t7tf)-R	CCAACTCA*GCTTCCTTT*GTGGTTCAATTACTCCTTAATGTTC
I-emrAB(tolcr)-F	CCCTTTCC*GTAACTGA*GTTAAGAAGATCGTGGAGAACA
I-emrAB(t7tf)-R	CCAACTCA*GCTTCCTTT*CAGGAACTGCACATCTAGTCAG
I-*acrEF*(tolcr)-F	CCCTTTCC*GTAACTGA*ATTTGTAGGATAGCGAACTGT
I-*acrEF*(t7tf)-R	CCAACTCA*GCTTCCTTT*CACTGGAAATAATAAAGGCAC
I-emrKY(tolcr)-F	CCCTTTCC*GTAACTGA*TTGATATGAAGAATGAATGCTC
I-emrKY(t7tf)-R	CCAACTCA*GCTTCCTTT*ATCCTTTCCCTGGGTGAG
I-mdtABC(tolcr)-F	CCCTTTCC*GTAACTGA*CTCCCTTATTGGCTGGCTAC
I-mdtABC(t7tf)-R	CCAACTCA*GCTTCCTTT*CGGTGTTTACGATGGTGGT
I-entS(tolcr)-F	CCCTTTCC*GTAACTGA*TGATGGCAAGGCATTGTA
I-entS(t7tf)-R	CCAACTCA*GCTTCCTTT*GTTTTAAGCATTAACTGTCGG
I-acrD(tolcr)-F	CCCTTTCC*GTAACTGA*GCAAGTCAAGCCTACAACG
I-acrD(t7tf)-R	CCAACTCA*GCTTCCTTT*CTGAGCAGGTTCTTAATCGGAAAGT
I-mdtEF(tolcr)-F	CCCTTTCC*GTAACTGA*CGCTGGTCTGTAAATCCC
I-mdtEF(t7tf)-R	CCAACTCA*GCTTCCTTT*GAATGGTTAGCAGGAAAGAGT
I-ppsA(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*TTTCTCAAACCGTTCATTTA
I-ppsA(t7tf)-R	CCAACTCA*GCTTCCTTT*AACGCAGGATGTCTGTGAA

Table 5-6. Transporters and their functions

Gene	Category	Length/ Subunit composition	Full name	Function
fsr	MFS	1221 bp / 406 aa	Fosmidomycin efflux transporter	Expression of *fsr* has been shown to confer resistance to the isoprenoid fosmidomycin
cmr	MFS	1233 bp / 410 aa	Multidrug efflux transporter MdfA	Overexpression of *mdfA* has demonstrated that it confers resistance to tetracycline, chloramphenicol, erythromycin, some aminoglycosides and fluoroquinolones, and organic cations such as ethidium bromide. Deletion of *mdfA* resulted in increased susceptibility to ethidium bromide and benzalkonium chloride. Overexpression of *mdfA* also results in spectinomycin sensitivity and isopropyl-B-D-thiogalactopyranoside (IPTG) exclusion due to unknown mechanisms.
emrD	MFS	1185 bp / 394 aa	Multidrug efflux transporter *EmrD*	*EmrD* is a multidrug efflux protein involved in adaptation to low energy shock (exposure to uncouplers of oxidative phosphorylation) in *E. coli*
emrE	SMR	333 bp / 110 aa	Multidrug / betaine / choline efflux transporter *emrE*	Overexpression of the gene in *E. coli* has shown that it confers resistance to a range of toxic compounds including ethidium, methyl viologen, tetracycline, tetraphenylphosphonium (TPP+), erythromycin, sulfadiazine, acriflavin, crystal violet, and benzalkonium. *EmrE* also transports the osmoprotectants betaine and choline and plays a role in osmotic regulation of *E. coli* K-12.
mdtH	MFS	1209 bp / 402 aa	Polypeptide: YceL MFS transporter	Overexpression of the cloned *yceL* gene in a drug-sensitive background strain resulted in a two-fold increase in resistance to norfloxacin and enoxacin
mdtM	MFS	1233 bp / 410 aa	Transporter: multidrug efflux transporter *MdtM*	Overexpression of the cloned *mdtM* gene in a drug-sensitive background strain results in a four-fold increase in resistance to acriflavine and a two-fold increase in resistance to chloramphenicol, ethidium bromide and

				tetraphenylphosphonium bromide (TPP). Strains lacking *mdtM* show impaired growth in the presence of ethidium bromide and chloramphenicol compared to wild type. Overexpression of cloned *mdtM* in both a wild type background and in an *mdtM* null background results in increased resistance to ethidium bromide and chloramphenicol. *MdtM* is functional in whole cell ethidium bromide transport assays and efflux is disrupted by the addition of the ionophore, carbonyl cyanide 3-chlorophenylhydrazone (CCCP)
msbA	ABC	1749 bp / 582 aa	Lipopolysaccharide ABC transporter	A member of the ATP Binding Cassette (ABC) Superfamily of transporters, is essential for the translocation of the lipid A-core from the inner leaflet to the outer leaflet of the inner membrane
lptD	OMP	2355 bp / 784 aa	Outer membrane lipopolysaccharide transport and assembly complex - LptD subunit	The LptD (lipopolysaccharide transport) protein is an essential outer membrane (OM) protein which, in complex with LptE, functions to assemble lipopolysaccharides at the surface of the OM. LptD is involved in determining the level of organic solvent tolerance in *E. coli*
lptE	Lipoproteins	582 bp / 193 aa	Outer membrane lipopolysaccharide transport and assembly complex - LptE subunit	LptE (formerly RlpB) is involved with lipopolysaccharide assembly in the outer membrane as part of the outer membrane LPS assembly complex with LptD. LptE is one of two (along with LptD) so-called rare lipoproteins in *E. coli*. Purified *lptE* binds to LPS specifically.
lptCABFG	ABC	330.0 kD [*LptG*][*LptB*]$_2$[*LptF*]	Lipopolysaccharide transport system	LptCBFG is an ATP Binding Cassette (ABC) transporter which facilitates the release of LPS from the inner membrane. LptA is a periplasmic protein initially thought to be an LPS chaperone but now believed to act as a bridge between the inner membrane and the outer membrane. LptA and LptC interact in vivo and form a stable complex.
MacAB-tolC	ABC	[(*tolC*)$_3$][*MacB*]$_2$[*MacA*]	Macrolide efflux transport system	*MacAB*, also known as YbjYZ, probably forms a complex with outer membrane protein *tolC* to confer macrolide resistance via active drug efflux
AcrAB-tolC	RND	[*AcrA*][(*tolC*)$_3$][*AcrB*]$_3$	*AcrAB-tolC* multidrug efflux transport system	*AcrAB* and *tolC* make up a three-component proton motive force-dependent multidrug efflux system which confers resistance to multiple antimicrobial ag*EntS*. The complex is the major contributor to the intrinsic resistance of *E. coli* to organic solv*EntS*, dyes and deterg*EntS* as well as lipophilic antibiotics including novobiocin, erythromycin, fusidic acid and cloxacillin
emrAB-tolC	MFS	[*EmrA*][*EmrB*][(*tolC*)$_3$]	EmrAB-*tolC* multidrug efflux transport system	Expression of *emrB* has been shown to confer resistance to hydrophobic compounds such as carbonyl cyanide m-chlorophenylhydrazone, tetrachlorosalicylanilide, organomercurials and nalidixic acid
AcrEF-tolC	MFS	[*AcrE*][*AcrF*][(*tolC*)$_3$]	*AcrEF-tolC* multidrug efflux transport system	*AcrEF* is a multidrug efflux system in *E. coli* that is highly homologous to the *AcrAB* multidrug efflux system. Plasmid overexpression of *AcrEF* suppresses the hyper-drug sensitivity of *AcrAB* deletion mutants however the *AcrEF* system is not thought to play a large role in drug resistance as *AcrEF* deletion mutants do not display a drug sensitive phenotype. The level of expression of *AcrEF* is thought to be low under laboratory conditions. Plasmid overexpression of *AcrF* is able to complement an *AcrB* deletion mutation.
EmrKY-tolC	MFS	[*EmrK*][*EmrY*][(*tolC*)$_3$]	*EmrKY-tolC* multidrug efflux transport system	*EmrK* and *EmrY* show sequence similarity to *EmrA* and *EmrB* respectively. Cloning of the promoter along with the creation of an *EmrK*-lacZ fusion revealed that expression was increased in the prescence of a sub inhibitory

				concentration of tetracyclin, chloramphenicol, or salicylate but not by carbonylcyanide m-chlorophenylhydrazone, nalidixic acid, or kanamycin
MdtABC-tolC	RND	[(*tolC*)$_3$][*MdtC*][*MdtB*]$_2$[*MdtA*]	*MdtABC-tolC* multidrug efflux transport system	*MdtA* is a membrane fusion protein functioning as a component of the *MdtABC* RND-type drug exporter. Membrane fusion proteins serve as 'adaptor' proteins that connect an inner membrane transporter to an outer membrane channel such as *tolC* - they are believed to be largely periplasmic. These genes code for an RND-type drug export distribution (RND) complex, *MdtABC*, which confers resistance against bile salt derivatives, sodium dodecyl sulfate (SDS) and novobiocin.
EntS-tolC	MFS	[*EntS*][(*tolC*)$_3$]	*EntS-tolC* Enterobactin Efflux Transport System	*EntS* was identified in a screen for genes that reduce the lethal effects of stress. An *EntS* insertion mutant is more sensitive than wild type to mitomycin C and other stresses. *EntS* has been implicated in arabinose efflux. The *EntS* protein is a member of the major facilitator superfamily (MFS) of transporters. Based on sequence similarity, it functions as a proton-driven efflux system. Siderophore nutrition assays have shown that an *EntS* mutant is unable to export enterobactin efficiently to alleviate iron deprivation, though some export does still occur through another mechanism. Deletion of *tolC* abolishes enterobactin export completely. Enterobactin appears to have more than one mechanism for export to the periplasm, one involving *EntS*, but *tolC* is the only outer membrane channel that can export enterobactin from the periplasm.
MdtEF-tolC	RND	[(*tolC*)$_3$][*MdtE*]$_3$[*MdtF*]$_3$	*MdtEF-tolC* multidrug efflux transport system	Studies performed using overexpression of the response regulator EvgA conferred multiple drug resistance (MDR) to *E. coli* cells lacking the *AcrAB* MDR transporter. The plasmid-containing cells showed drug resistance against deoxycholate (>32-fold compared with control level), doxorubicin (64-fold), rhodamine 6G (16-fold), erythromycin (8-fold), crystal violet (8-fold), benzalkonium (8-fold), and sodium dodecyl sulfate (SDS) (4-fold). EvgA is known to positively regulate the gene expression of the drug resistance system *EmrKY*, but cells in which only *EmrKY* is overproduced acquire resistance only to deoxycholate (8-fold). Deletion mutation studies were conducted in which *MdtEF* was removed by chromosomal in-frame mutation. The *MdtEF* deletion strain showed no increased drug resistance (except for deoxycholate) when evgA was overexpressed. Since overexpression of EvgA normally confers an MDR phenotype, these results strongly suggest that the EvgA-induced MDR is due to stimulation of *MdtEF* gene expression.
cusC	RND	1374 bp / 457 aa CusCFBA	The copper/silver transporting efflux system, CusCFBA	The copper/silver transporting efflux system, CusCFBA, is one of at least three systems involved in copper resistance. CusB is a member of the membrane fusion protein (MFP) family. CusC is the outer membrane factor which forms a channel in the outer membrane. CusA is the resistance-nodulation-division (RND) permease. CusF is the periplasmic copper binding protein. CusCBA is a tripartite complex that spans both the inner and outer membrane and, along with the periplasmic chaperone CusF, functions to export copper and silver ions from both the the cytoplasm and the periplasm to the extracellular environment.
mdtG (yceE)	MFS	1227 bp / 408 aa	MdtG drug MFS transporter, or	Overexpression of the cloned *mdtG* gene in a drug-sensitive background strain resulted in a two-fold increase in resistance to deoxycholate and a four-fold increase in

			MdtG multidrug efflux pump	resistance to fosfomycin.
mdtJ (ygF)	SMR	366 bp / 121 aa	MdtJI spermidine SMR transporter	Episomal expression of *mdtJI* resulted in excretion of spermidine in transport assays and allowed recovery of cells from spermidine toxicity due to over-accumulation. Drug-resistance studies indicate that overexpression of *mdtJI* in an expression vector conferred increased resistance to deoxycholate (4-fold), nalidixic acid (2-fold), fosfomycin (2-fold), and SDS (2-fold). Episomal expression of *mdtJ* causes inhibition of cell growth, compared to wild type
mdtL (yidY)	MFS	1176 bp / 391 aa	YidY drug MFS transporter, MdtL multidrug efflux pump	Overexpression of the cloned *yidY* gene in a drug sensitive background strain resulted in a four-fold increase in resistance to chloramphenicol, ethidium bromide and TPP.
mdtK	MATE	1374 bp / 457 aa	MdtK multidrug efflux transporter	Expression of the cloned *mdtK* gene conferred increased resistance to norfloxacin, ciprofloxacin, acriflavine, and tetraphenylphosphonium. *E. coli* strains expressing*norM* or *mdtK* showed elevated resistance to several antimicrobial agents, including fluoroquinolones, ethidium bromide, rhodamine 6G, acriflavine, crystal violet, berberine, doxorubicin, novobiocin, enoxacin, and tetraphenylphosphonium chloride.
acrD	RND	3114 bp / 1037 aa	*AcrA*D-*tolC* multidrug efflux transport system	AcrD is believed to form a tripartite complex with *AcrA* (an MFP) and *tolC* for multidrug transport. AcrD shows amino acid sequence similarity with the *E. coli* RND multidrug efflux proteins *AcrB* and AcrF. Experiments using proteoliposomes show the *AcrA*D-*tolC* complex is able to transport drugs from both the cytoplasm and the periplasm to the extracellular space. Disruption of the *acrD* gene did not result in hypersusceptibility to lipophilic and amphiphilic drugs, but did result in hypersusceptibility to a variety of aminoglycosides including amikacin, gentamicin, tobramycin, kanamycin, neomycin, erythromycin, and polymyxin B. Expression of *acrD* on a multicopy plasmid from native or IPTG inducible promoters in an *acrAB* mutant resulted in increased resistance to deoxycholate, SDS, novobiocin, kanamycin, tetracycline, nalidixic acid, norfloxacin, fosfomycin, bile acids, fusidic acid, and progesterone.
bcr	MFS	1191 bp / 396 aa	Multidrug efflux transporter Bcr	Bcr is a probable multidrug efflux protein, expression of which has been shown to confer resistance to bicyclomycin and sulfathiazole.

5.5.2 Quantification of intracellular and extracellular AD and MEC

The quantification of intracellular and extracellular AD and MEC was illustrated in Figure 5-7. One milliliter fresh medium in 14mL BD Falcon™ tube was inoculated with 1% overnight grown cell culture. And the medium was covered by 200 μL dodecane (with 100 mg/L β-caryophyllene). When

the OD reaches about 5, five microliter dodecane was transferred into 495 µL ethyl acetate (Extracellular AD). And 400 µL cell culture was transferred into 1.5 mL tube and centrifuged to separate the cell pellets from the media. The obtained supernatant was diluted 20X by 100% methanol and centrifuged to remove the salts (Extracellular MEC). The cell pellets were re-suspended into 100 µL double distilled water. Sixty microliter ethyl acetate and 10 mg 0.1 mm BioSpec glass beads were added into the cell suspension and was shaken for 10 min at 4 °C to lyse the cells and extract the intracellular AD (ethyl acetate phase) and MEC (water phase). MEC was quantified on the UPLC (Waters ACQUITY UPLC) – (TOF) MS (Bruker micrOTOF II) platform [24]. AD was measured using Agilent 7980A gas chromatograph equipped with an Agilent 5975C mass spectrometer (GC/MS) as previously reported [130].

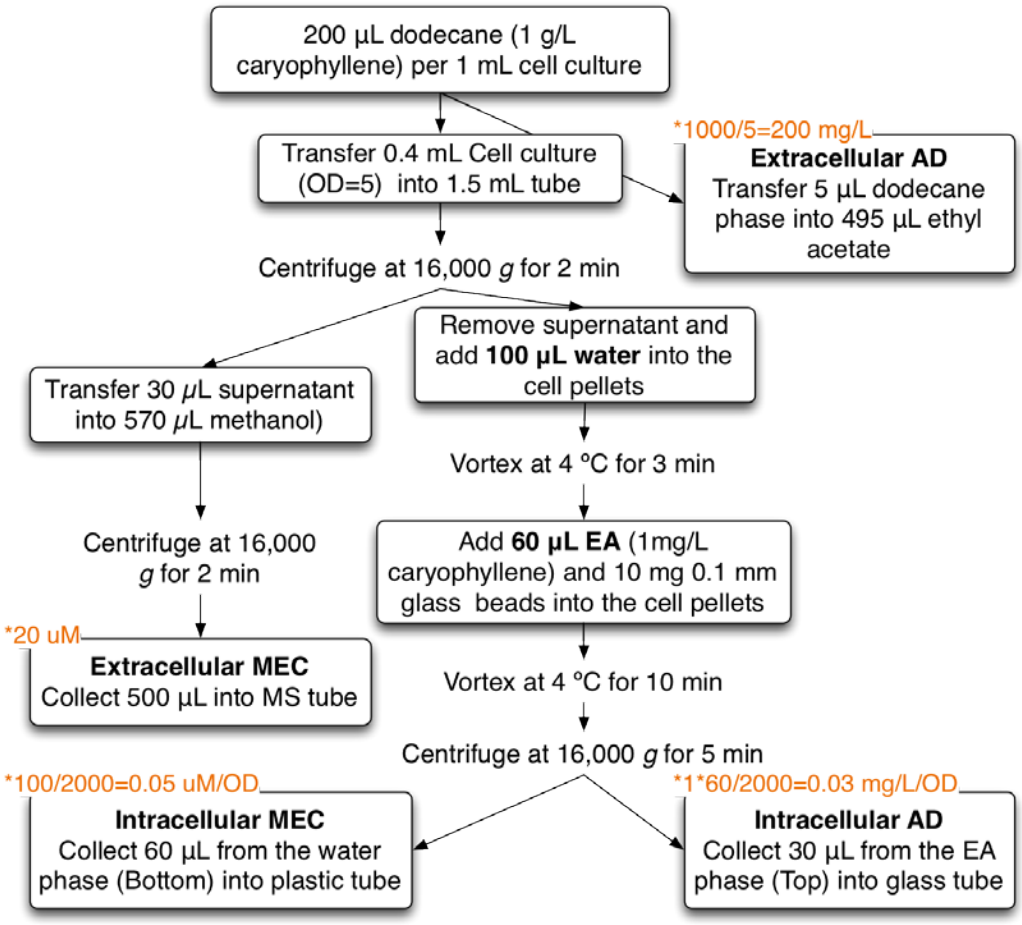


Figure 5-7. Protocol for the quantification of intracellular and extracellular AD and MEC.

5.5.3 Quantification of plasmid copy number

The plasmid copy number was obtained based on real-time quantitative PCR data. It was calculated by normalizing the copy number of plasmid resistant genes (ampicillin resistant gene for pBAD-SIDF and chloramphenicol resistant gene for pACM-ADS) to the copy number of chromosomal gene *cysG*.

Chapter 6. KINETIC STUDY OF THE DXP PATHWAY

6.1 Introduction

To identify the rate-limiting steps in the DXP pathway *in vivo*, a sequential method was previously used where single gene in the pathway was individually overexpressed to identify gene targets, followed by screening of two-gene combinations using the identified genes, and then three-gene combinations, until all the combinations of identified targets were tested [19]. Obviously, this method was very tedious and time-consuming. *In silico* modeling offers a rapid and predictive way to optimize a pathway or a network. To obtain such a computational model, a robust and controllable system is required for the generation of reliable experimental data. The DXP pathway is a linear pathway and starts from the conversion of pyruvate (PYR) and glyceraldehyde-3-phosphate (GAP) to DXP catalyzed by DXP synthase (encoded by gene *dxs*). Glucose is used to provide energy and to produce biomass and metabolites. As the glycolytic network is complex, it is very difficult to calculate or measure the ratio of carbon that enters into the DXP pathway from glucose. On the other hand, it was reported that D-xylulokinase (encoded by gene *xylB*) was able to phosphorylate 1-deoxy-D-xylulose (DX) to DXP [187]. With the overexpression of *xylB*, it is possible to construct an artificial semi-closed system in which DX is exclusively used to produce carbon fluxes in the

DXP pathway. And using ^{13}C labeled glucose and ^{12}C DX, it is possible to differentiate the carbon flux from either the central pathway or from DX.

In addition to the dynamics of the intracellular metabolites, we were also interested in studying the kinetics of extracellular metabolites and their relationship with the dynamics of the intracellular metabolites. DXP and MEC were found to be effluxed into the media when they were over produced [129]. The efflux of DXP and MEP not only affect the kinetics of the pathway and they may play a regulatory role as a signaling molecule. For example, it was also hypothesized that the carbon flux in the DXP pathway mainly accumulated as MEC, due simply to the inefficient activity of HMBPP synthase or limited availability of cofactors for the enzyme [157]. An established kinetic model of both intracellular and extracellular metabolites can facilitate validation of such hypothesis.

In this study, we first constructed and tested the semi-closed dual-flux platform using ^{13}C glucose and ^{12}C DX. We then applied this platform to study the kinetics of the intracellular and extracellular metabolites for two genotypes. Finally, the S-system model was attempted to simulate the kinetic data.

6.2 Results

6.2.1 Experimental design

6.2.1.1 Construction of the dual flux system

As shown in Figure 6-1, there are two routes of carbon to flux into the DXP pathway: to be converted from PYR and GAP (central pathway) by D-deoxyxylulose-5-phosphate synthase (*dxs*) or from DX (synthetic pathway) which is catalyzed by D-xylulokinase (*xylB*). To study the kinetics of the DXP pathway, the expressions of *dxs* and *xylB* were controlled independently. The cells were grown in the defined media with D-glucose-$^{13}C_6$ (^{13}C Glu) as the only carbon source, thus all the DXP intermediates from central pathway were ^{13}C labeled. Whereas all the carbon fluxes from DX were ^{12}C.

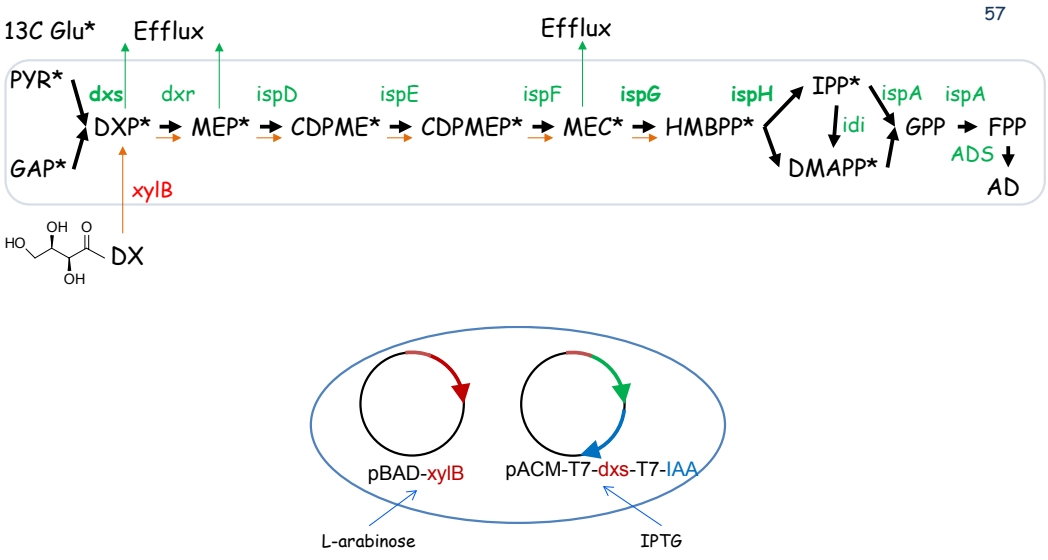


Figure 6-1. The semi-closed dual-carbon-flux system of the DXP pathway. DX, 1-deoxy-D-xylulose; PYR, pyruvate; GAP, glyceraldehyde-3-phosphate; DXP, 1-deoxy-D-xylulose-5-phosphatec; MEP, methylerythritol phosphate; CDPME, 4-

diphosphocytidyl-2-C-methyl-D-erythritol; CDPMEP, 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate; MEC, 2-C-methyl-D-erythritol-2,4-diphosphate; HMBPP, 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; AD, amorpha-4,11-diene. RK2A uses the P1 promoter induced by m-toluic acid, and pACM uses the T7 promoter induced by IPTG. IAA refers to the operon *idi-ispA-ADS*.

Our initial intention was to maintain the original system as much as possible; therefore, the ^{12}C carbon flux from DX was minimized. It was assumed that the ^{12}C flux acted as a tracer for the kinetic behavior of the system and did not perturb the endogenous system significantly. To minimize the ^{12}C carbon flux from DX, a low-copy plasmid RK2A with the weaker P1 promoter was used to carry the gene *xylB* (plasmid RK2-xylB). While the *dxs* was overexpressed in the higher copy vector pET, controlled by the stronger T7 promoter. To facilitate studies on the effect of the kinetics of the pathway on the production of AD, *idi-ispA-ADS* was co-overexpressed with *dxs* in the pET vector (plasmid pACM-SIAA).

Using this system, a preliminary study was first carried out. As shown in Figure 6-2, both the amounts of ^{13}C and ^{12}C DXP and MEC in the DXP pathway were successfully detected. However, MEP in the DXP pathway was not detected. The MS peak of ^{13}C MEP coincided with a background MS peak, therefore, it was not accurately quantified. Failure to detect ^{12}C MEP was likely to be due to the concentration of MEP being too low to be detected in this condition. To tackle this issue, the RK2A vector was replaced with higher copy-number pBAD vector with araBAD promoter (plasmid pBAD-xylB). As a result, the ^{12}C MEP was then detected despite its concentration (0.3 nmol/g DCW) was one order of magnitude lower

than DXP (5 nmol/ g DCW) and MEC (2 nmol/ g DCW) (Figure 6-3). The concentration of ^{12}C DXP, MEP and MEC were increased but not of ^{13}C labeled metabolites, indicative that the expression of *xylB* was successfully controlled by araBAD promoter without affecting the DXP pathway significantly.

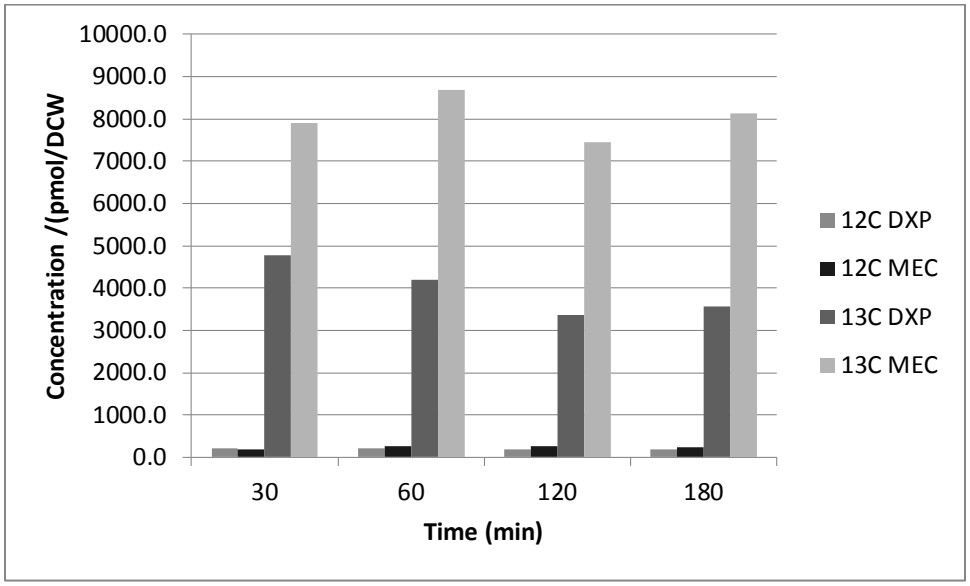
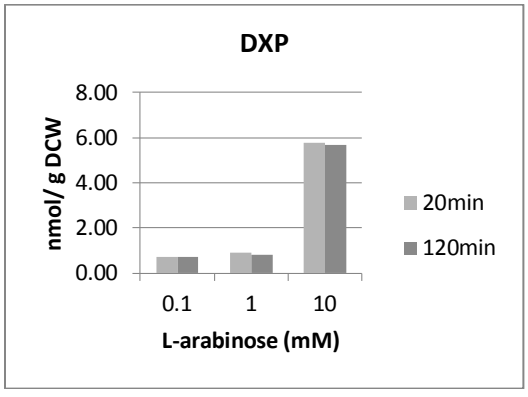
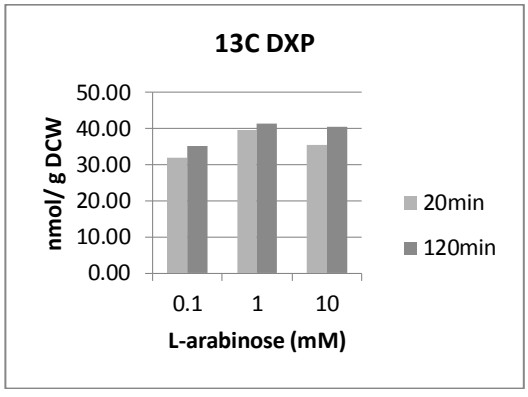


Figure 6-2. The kinetic data of the dual-flux system.
Samples were collected at 30 min, 60 min, 120 min and 180 mins after feeding of DX. The medium used in this study was MC defined media, and the cells were incubated at 28 °C at the shaking speed of 300 rpm. Plasmid RK2A-xylB was used in this study.





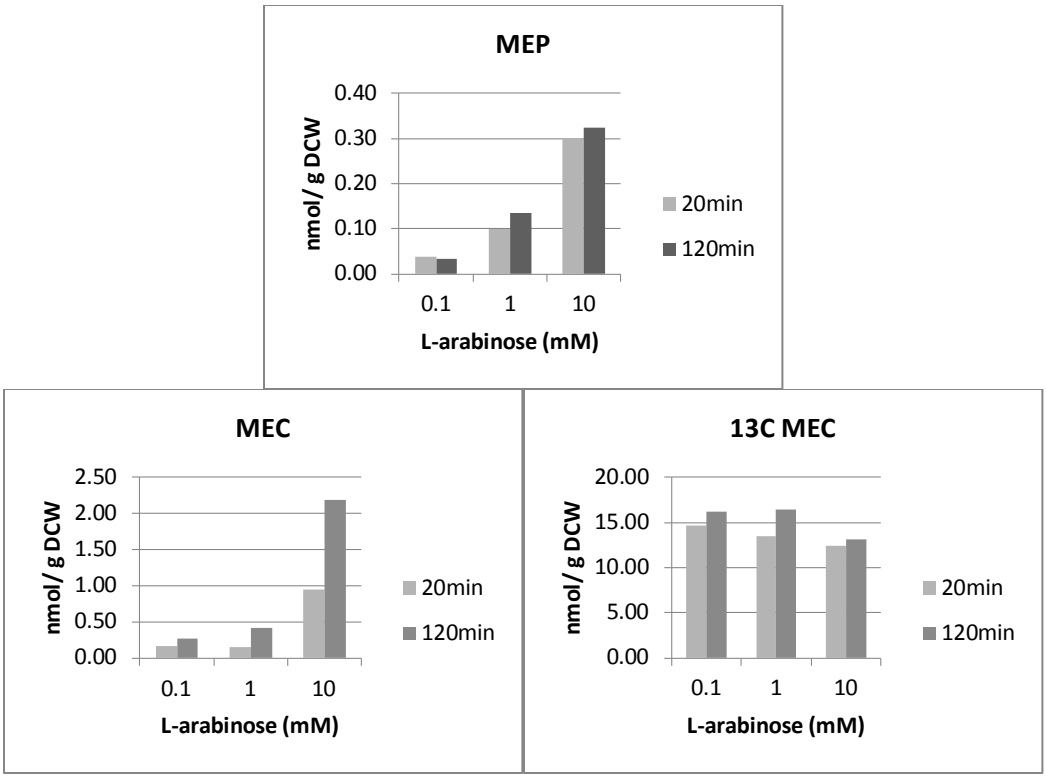


Figure 6-3. The DXP intermediates in the dual-flux system with the plasmid pBAD-xylB.

6.2.1.2 Quantification of DX

^{13}C glucose and DX were quantified using GCMS. In order to make glucose and DX volatile for quantification by GCMS, these compounds were derivatized using methoxime-trimethylsilylation. As shown in Figure 6-4, the ions m/z 323 and ions m/z 205 have large mass numbers and relatively high intensities, thus they were chosen as monitor ions for ^{13}C glucose and DX. Using these ions intensity and the concentration of glucose and DX, we established the standard curves for ^{13}C glucose and DX (Figure 6-5).

(A)

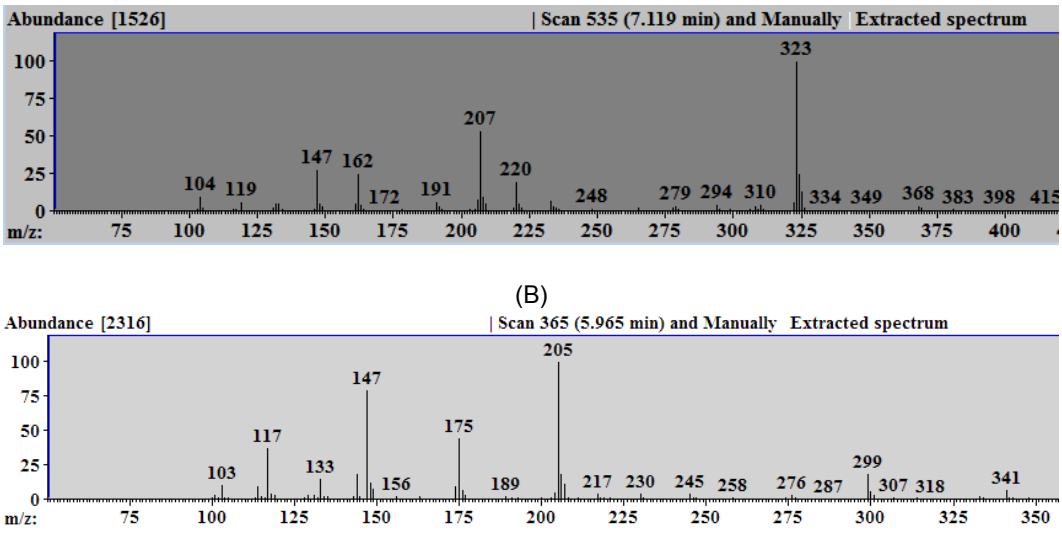


Figure 6-4. The mass spectra of methoxime trimethylsilylated ^{13}C glucose (A) and DX (B)

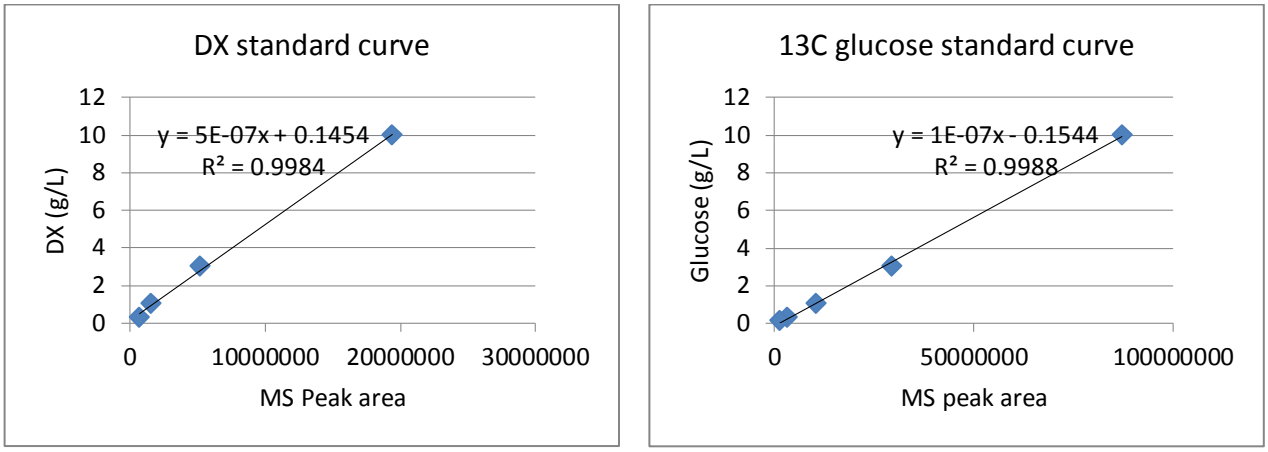


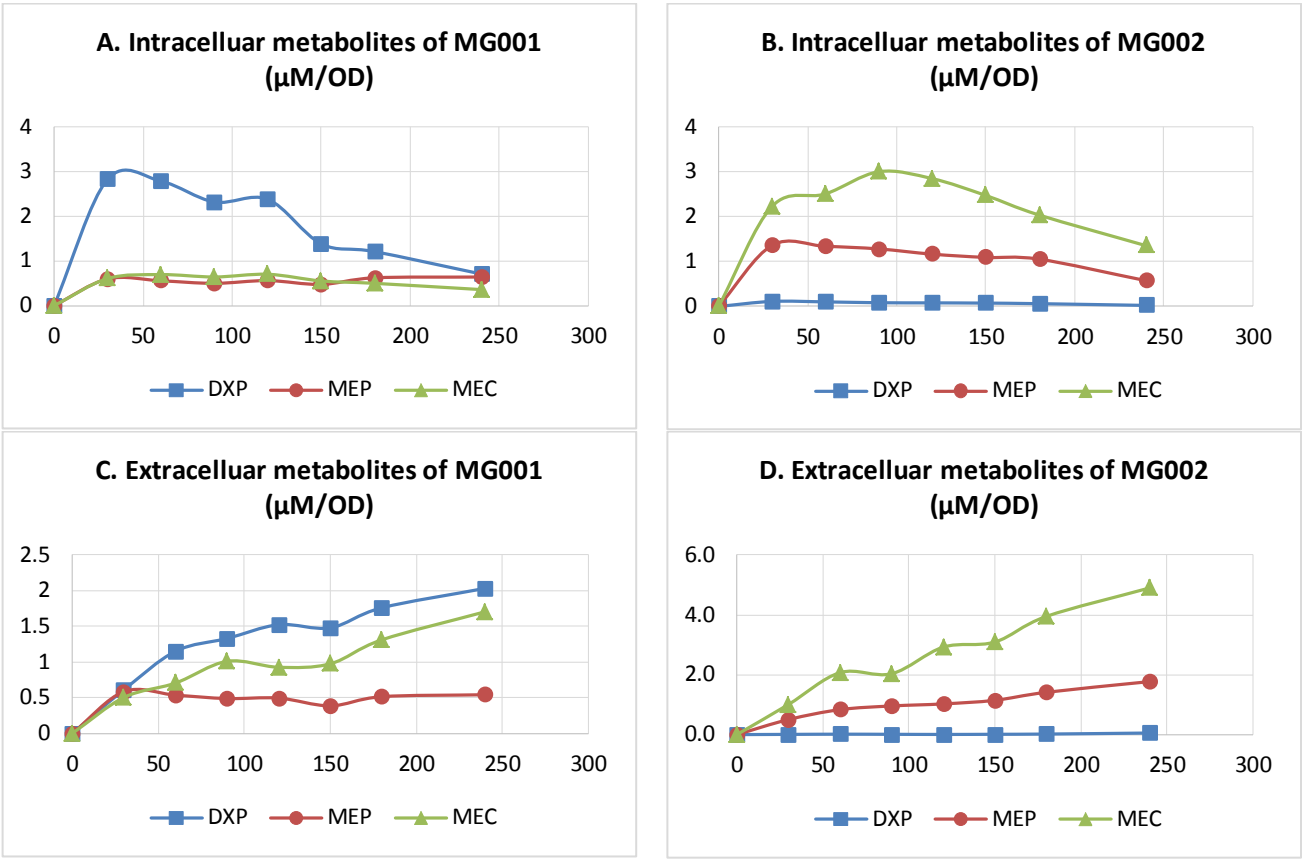
Figure 6-5. The standard curve for ^{13}C glucose and DX

6.2.2 Data collection by LCMS and GCMS

Two experimental conditions were chosen to study the kinetics of the DXP pathway. One was the overexpression of *dxs* with *idi-ispA-ADS* (strain MG001), and the other one was the overexpression of *dxs-dxr* with *idi-ispA-ADS* (strain MG002). As shown in Figure 6-1, *dxs* is the first enzyme and *dxr* is the second enzyme in the DXP pathway. Therefore, perturbing

the system by *dxs* or *dxs* and *dxr* can be a good starting point to understand how the overexpression of *dxs* and *dxr* affect the kinetics of the pathway.

The kinetic data for studies using MG001 and MG002 were shown in Figure 6-6. DXP, MEP and MEC in the DXP pathway were detected, but CDPME and CDPMEP were not detected. This was likely due to the efficient conversion of CDPME and CDPMEP to MEC [189]. Because the cells were dividing during the sampling period, all the concentrations of intracellular and extracellular metabolites were normalized to the cell density (μM/OD).



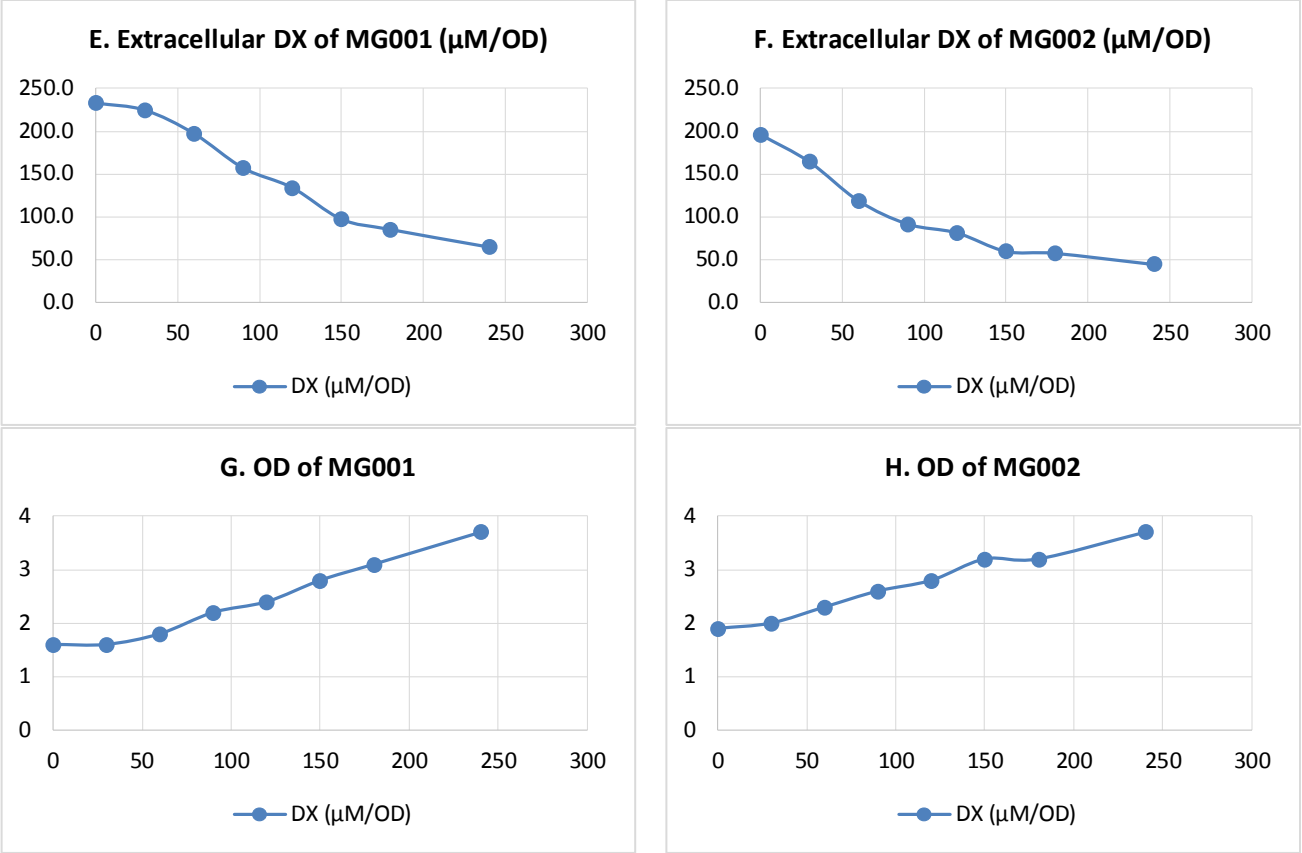


Figure 6-6. The kinetic data of the intermediates in the DXP pathway. A, C, E and G are the kinetic data of intracellular metabolite, extracellular metabolites, extracellular DX and cell density for strain MG001, respectively; and B, D, F and H are the kinetics data of intracellular metabolite, extracellular metabolites, extracellular DX and cell density for strain MG002, respectively. The time of feeding DX was set as the starting point (0 min).

In MG001, the overexpression of *dxs* resulted in the gradual accumulation of intracellular DXP and peaked at around 40 min and then started to decrease. Intracellular MEP and MEC increased at the first 40 min and maintained over the period of the study. As expected, the extracellular DXP and MEC accumulated over the time, whereas the extracellular MEP increased at the first 30 min and plateaued. In MG001, the intracellular DXP accumulated to higher concentrations than that of MEP and MEC.

This is reasonable as only *dxs* in the pathway was overexpressed so that more DXP was produced than consumed.

When *dxs* and *dxr* were co-overexpressed (MG002), the intracellular concentration of DXP was almost undetectable over the period of the study. As a result, extracellular DXP did not accumulate. The intracellular concentration of MEP was increased initially and started to decrease after about 40 min. Similar to MEP, intracellular MEC increased initially and peaked at 90 min (later than MEP) and then started to drop. The extracellular MEP and MEC were found to increase gradually but their rates of increase were decreased at later stage.

As compared to MG001, the intracellular concentration of DXP was much lower while the concentrations of MEP and MEC were higher in MG002. This was likely to be due to the overexpression of *dxr* along with *dxs* in MG002 and thus DXP was efficiently converted into MEP. And the *ispDEF* complex [189] efficiently converted MEP to MEC without accumulating high concentrations of CDPME and CDPMEP.

6.2.3 Kinetic Modeling of the DXP pathway

6.2.3.1 S-system model of the DXP pathway

The S-system structure, a form within Biochemical Systems Theory (BST), is one of the most promising canonical nonlinear models in metabolic modeling [188]. We attempted to use the S-system model to simulate the kinetic data in Figure 6-6. The model is constructed by approximating flux with products of power-law functions. In this model, two assumptions were

made. First, we assumed all the ^{12}C carbon from DX was efficiently converted into DXP intracellularly, thus the reaction rate of DXP could be calculated from the levels of DX in the media. Second, we assumed that the efflux rate was the function of intracellular metabolites, and to be specific, they also followed the power-law structure of chemical reaction. From these two assumptions, the S-system equations for the DXP pathway were as follows:

$$\begin{aligned}
\dot{X}_1 &= -\alpha_1 X_1^{p11} \\
\dot{X}_2 &= \alpha_1 X_1^{p11} - \beta_1 X_2^{p21} - \alpha_2 X_2^{p22} \\
\dot{X}_3 &= \beta_1 X_2^{p21} - \beta_2 X_3^{p31} - \alpha_3 X_3^{p32} \\
\dot{X}_4 &= \beta_2 X_3^{p31} - \beta_3 X_4^{p41} - \alpha_4 X_4^{p42} \\
\dot{X}_5 &= \alpha_2 X_2^{p22} \\
\dot{X}_6 &= \alpha_3 X_3^{p32} \\
\dot{X}_7 &= \alpha_4 X_4^{p42}
\end{aligned} \qquad \text{Equation 6-1,}$$

where X1 represented the DX concentration in the media, and it was determined by the conversion rate of DX into DXP. X2 represented the intracellular concentration of DXP which was determined by the conversion of DX into DXP, the efflux of DXP and consumption of DXP by *dxs*. X3 and X4 represented the intracellular concentrations of MEP and MEC, respectively, and they were calculated similarly to that of intracellular DXP. X5, X6 and X7 represent the extracellular concentrations of DXP, MEP and MEC, respectively, and they were determined based on their own efflux rate.

6.2.3.2 Data regression

Although two parameter-estimating approaches [111, 117] were used initially, the experimental data (MG001 in Figure 6-6) could not be successfully fitted using the S-system model (Equation 1-1), especially the intracellular MEP (red line in Figure 6-7). This could be due to ill condition of the non-linear dynamic system caused by the data structure or the parameter estimation approaches employed. In addition, the total experimental data points were relatively few (only 7 points) where a single or two inaccurate data points would have produced unacceptably large deviations thus, limiting the modeling.

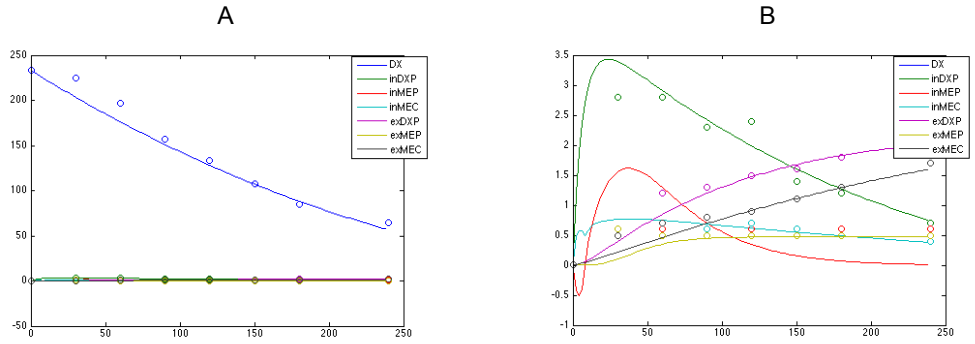


Figure 6-7. Parameter estimation for the kinetic data of MG001.
(A) The global fitting results including DX. (B) The fitting results of intermediates without DX.

6.3 Discussion

Various valuable products can be produced through the DXP pathway [2], however, the capability of the pathway was limited by our poor understanding of the kinetic behavior of the DXP pathway. Therefore, in this chapter, a MS-based offline quantitative approach was developed.

Using D-xylulokinase, we created a controllable artificial branch entry into the DXP pathway. Using unlabeled and labeled carbons to differentiate the carbon flux from central pathway or from artificial branch, the relatively good time-course data for the DXP pathway were obtained in Figure 6-6. As the successful parameter estimation is dependent on the availability of good quality experimental data sets, the experimental data obtained in this study may be used to develop a model to describe the pathway kinetics.

Based on the kinetic data obtained in Figure 6-6, the efflux of extracellular metabolites was likely to be dependent on the concentration of the intracellular metabolites. In our model, we hypothesized that the efflux rate was the power law function of the intracellular concentration. This hypothesis may be incorrect. As the data shown in Figure 6-8, the initial efflux rate of DXP and MEC was zero when the intracellular concentration was low. Furthermore, the efflux rate of MEC gradually increased with increasing intracellular MEC concentrations and they were highly correlated (R^2 of DXP = 0.877, R^2 of MEC = 0.923). Eventually, the efflux rate reached the maximal when the intracellular concentration was above certain level, where for MEC, when the intracellular MEC was above 0.8 μM, the efflux rate did not increase but decreased a bit further as the intracellular MEC increased. This observation was unexpected and more evidence is required to support and validate the hypothesis. And because of the significant effect of the efflux on the intracellular carbon flux and

pathway kinetics, once the relationship is confirmed, it should be integrated to refine the kinetic model.

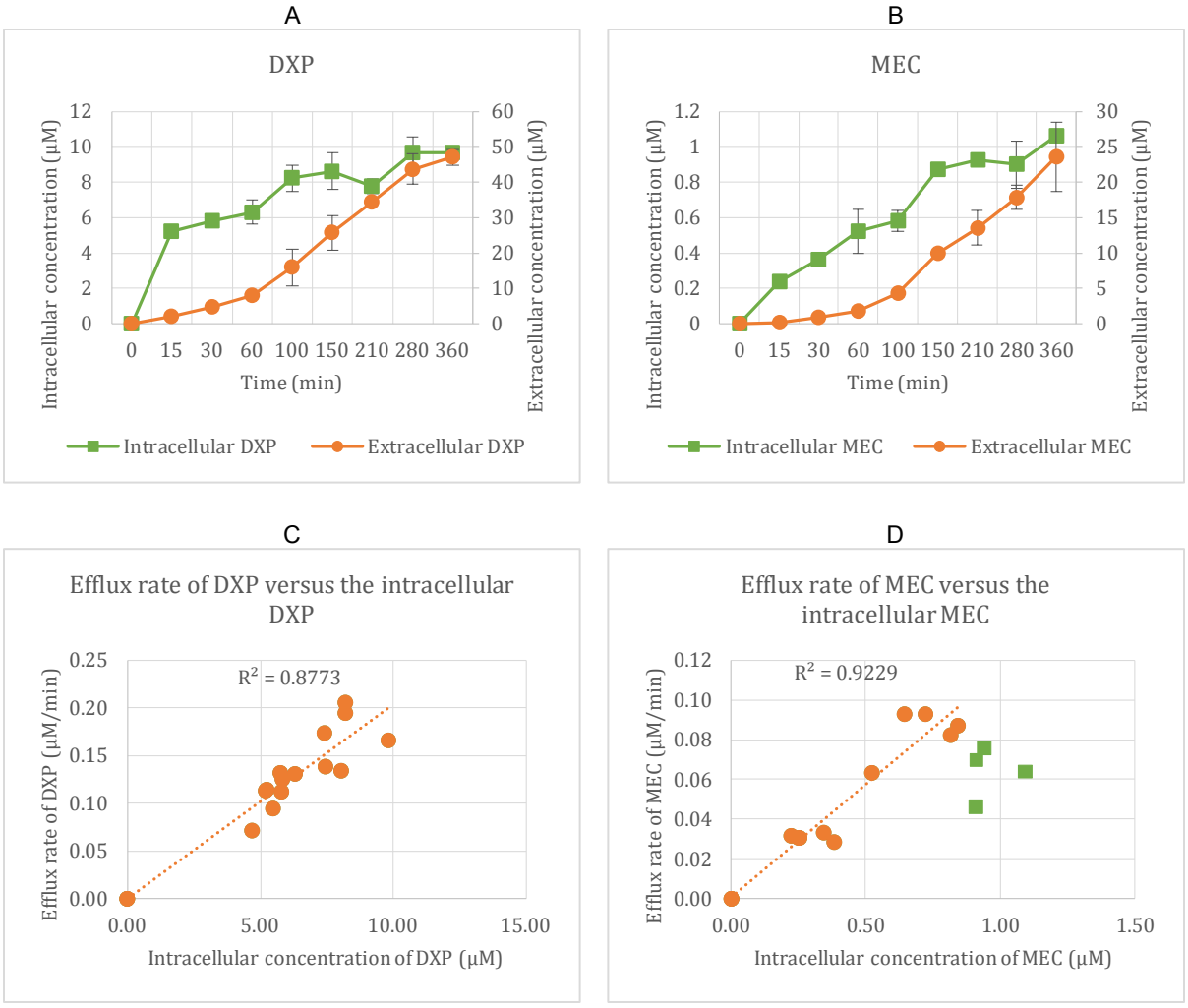


Figure 6-8. Efflux study of DXP and MEC.
(A) The time course of intracellular and extracellular DXP; (B) The time course of intracellular and extracellular MEC; (C) The efflux rate of DXP versus its intracellular concentration; (D) The efflux rate of MEC versus its intracellular concentration. The strain used in the experiment was MG1655 DE3 with the operon *dxs-idi-ispDF* knocked into its genomic DNA. And pBAD-xylB was transformed into the strain.

Metabolic steady state is the prerequisite for kinetics study, which means that all intermediate concentrations are constant and all the fluxes are in balance throughout the experiment. In chemostat with continuous cultivation, a metabolic steady state is readily obtained, while in batch

process, a metabolic steady state only exist during the exponential growth phase, when the growth rate is constant [190]. Therefore, the DX feeding and sampling were conducted at exponential phase.

The interval between two time points was 30 min, which was not ideal for kinetics study. This time intervals were simply due to technical limitations where it took about 15 min to process the samples (OD measuring, collection of cells, centrifugation and separation). To enhance the precision of data obtained at these time points, it may be better to develop an online monitoring system. For example, Matthias et al [191] designed an MS-base quantitative real-time analysis system and successfully studied *in vitro* network using a cell free system. The advantage of the system is that it can pinpoint changes in metabolite concentration after perturbation to a certain enzyme in the context of the system. Another approach to design online monitoring system is to use metabolite biosensor. The *in vivo* ^{13}C nuclear magnetic resonance (NMR) is another promising method to investigate the *in vivo* flux distribution through the catabolic pathways [192] . Recently, novel biosensors based on flavonoid-responsive transcriptional regulators were developed [193], and there was a linear correlation between the fluorescence intensity and metabolite concentration using the biosensors. Hence, the use of an online monitoring system using the metabolite biosensor may circumvent the technical issues faced in this preliminary study in the future.

6.4 Conclusion

In an attempt to gain an understanding of the kinetics of the DXP pathway, an alternative supply of the first precursor (DXP) in the pathway was explored by the addition of DX. The carbon fluxes in the pathway were quantified by MS based methods. Using this system, we managed to capture some kinetic data of the DXP pathway when *dxs* alone or *dxs* and *dxr* were overexpressed. It was found that the efflux rates of DXP and MEC were dependent on their intracellular concentrations.

6.5 Methods and materials

6.5.1 Media and strains

The MC media used in the modeling work was defined media revised from Rersenberg media [142], in which citric acid and thiamine were removed from the media so that the only carbon source was glucose. The media contained 5 g/L ^{13}C glucose, 4 g/L $(NH_4)_2HPO_4$, 2.8 g/L KH_2PO_4, 7.4 g/L K_2HPO_4, 0.5 g/L $MgSO_4$, 2.5 mg/L $CoCl_2 \cdot 6H_2O$, 15.0 mg/L $MnSO_4 \cdot 4H_2O$, 1.5 mg/L $CuSO_4 \cdot 2H_2O$, 3 mg/L H_3BO_3, 2.5 mg/L $Na_2MoO_4 \cdot 2H_2O$, 13 mg/L $Zn(CH_3COO)_2$, 60 mg/L Fe(III) citrate and 8.4 mg/L EDTA.

E coli. K-12 MG1655 *ΔrecAΔendA DE3* (MG1655) were used as the parental strain in the experiments. The plasmid pBAD-xylB and pACM-dxs-idi-ispA-ADS were transformed into the MG1655 strain (MG001). The plasmid pBAD-xylB and pACM-dxs-dxr-idi-ispA-ADS were transformed into MG1655 strain (MG002).

6.5.2 Cultivation and sampling

The MC media was inoculated by 1% cell culture that was grown in ^{13}C glucose media overnight. The cells were cultivated at 37 °C with 300 rpm and the culture temperature was shifted to 28 °C after induction. The cells were induced when the OD reached ~0.5 by 0.01 mM IPTG and 10 mM m-toluic acid (or 10 mM L-arabinose). About two hours later, when the OD reached ~2.0 and the enzymes DXP synthase and D-xylulokinase were expressed, 0.2 g/L ^{12}C DX was added into the culture. After about 20 minutes, seven samples were collected every 30 min. The concentration of samples was quantified by LCMS or GCMS (Figure 6-9).

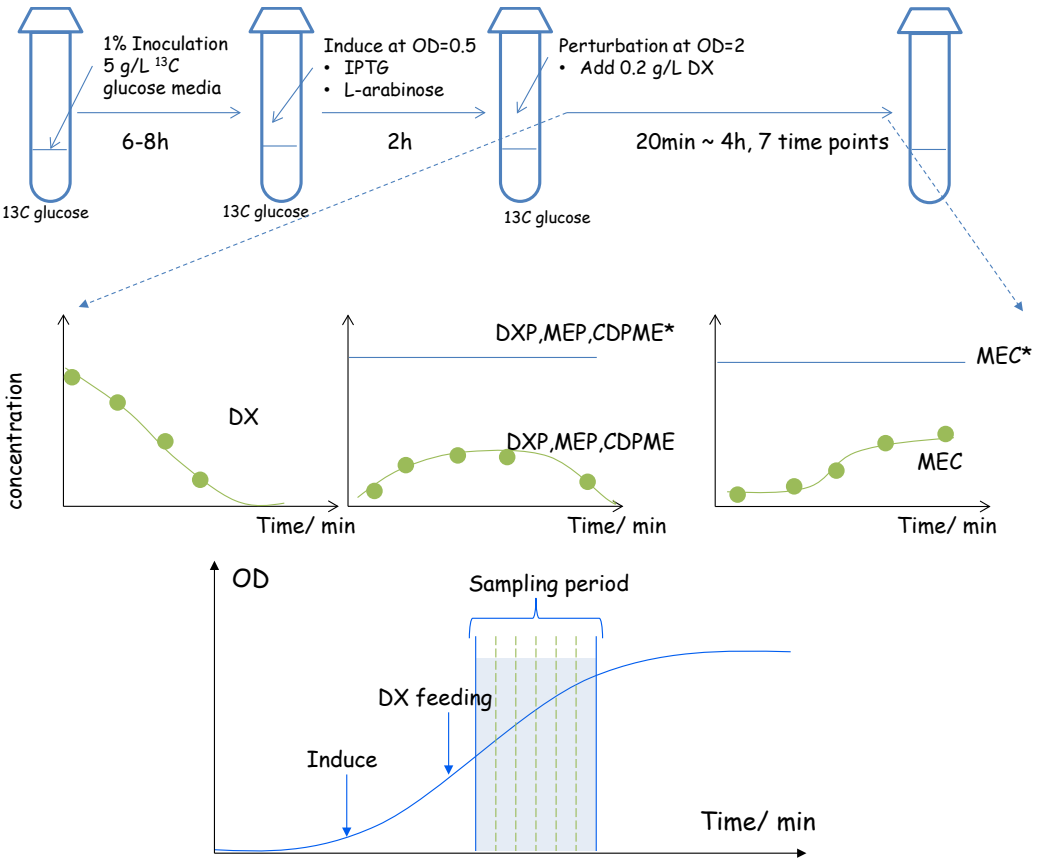


Figure 6-9. The experiment procedure for the kinetics study.
The samples were collected between mid-log to late-log phase when the cells was at steady state.

6.5.3 Quantification of 13C Glucose and DX in the media

The culture media was centrifuged at 16,000 *g* for 1 min and 10 µL supernatant was transferred into 90 µL methanol. The solution was well mixed and centrifuged to remove the salts in the media, and 80 µL of the desalted solution was dried and incubated with 30 µL MOX reagent (Thermo Scientific) at 37 °C for 1 hour. Thirty microliter MFTFA + 1% TMCS (Thermo Scientific) was subsequently added and incubated at 37 °C for 1 hour. Samples were centrifuged at 16,000 *g* for 1 min and 1 µL sample was injected into the GC-MS in splite mode with 9:1 split ratio at 270 °C. Agilent 7890 GC equipped with HP-5ms GC column and Agilent MS were used at flow rate 1 mL/min. Oven temperature was held at 100 °C for 2 mins, and elevated to 300 °C in 4 mins and held at 300 °C for 3 mins. MS source and quadrupole temperature were 230 °C and 150 °C respectively. MS data was acquired in scan mode from 100 to 500 m/z.

6.5.4 Quantification of intracellular and extracellular metabolites

Cell culture media equivalent to 1 mL OD600 =1.0 cells was withdrawn and centrifuged at 16,000 for 2 mins. One hundred microliter supernatant (extracellular metabolites) was transferred into 800 µL methanol with 100 µL double distilled water. The cell pellets (intracellular metabolites) were

re-suspened in 200 µL water. The suspension was mixed with 800 µL acidic extraction buffer (acetonitrile/ methanol/water 40:40:20 +0.1 M formic acid) and incubated at room temperature for 5 mins. All the solutions (intracellular and extracellular) were centrifuged at 16,000 *g* for 1 min. Seven milligram LC-NH_2 resin (SUPELCO) was placed into the resin cartridges and the resin was conditioned with 100 µL extraction buffer. All the supernatant was transferred into the resin cartridges and was centrifuged at 2000 *g* for 1 min. The metabolites in the resin were eluted with 56 µL 1% NH_4OH solution. And pH of eluent was subsequently adjusted to acidic (pH = 5) by 3 mL acetic acid. The samples were analyzed using the UPLC (Waters ACQUITY UPLC) – (TOF) MS (Bruker micrOTOF II) platform. UPLC setting and (TOF) mass spectrometry setting was essentially as described [157].

Chapter 7. CONCLUSION AND RECOMMENDATION OF FUTURE WORKS

7.1 General conclusion

Glucose is the preferred substrate for various fermentation processes, however the DXP pathway suffers from low efficiency when using glucose. In this study, the PTS was modified which resulted in higher productivity of downstream isoprenoids. Half of the PEP, the precursor of the DXP pathway, produced from glucose is consumed during glucose internalization by the PTS. Replacing the PTS with *galP* and *glk* increased the intracellular concentration of PEP and significantly enhanced the amorpha-4,11-diene (AD) production. Further a systematical pathway optimization method was developed EDASPO, and it was demonstrated by systematic optimization of the expression of *galP-glk*, *dxs*, *idi-ispDF* and *ispA-ADS*, resulted in more than 550 fold increase in AD production from glucose.

Next, the PTS knockout strategy was combined with statistical media optimization to enhance the production of lycopene. Two critical ingredients in the media were screened experimentally using factorial design. Using response surface method (RSM), the optimal concentrations of the two critical ingredients were rapidly identified and the lycopene yield was increased by more than 3 fold. The optimized medium was superior to all the commonly used complex media or defined media

for lycopene production. In addition, it was found that the copy number of the plasmid encoding the lycopene biosynthetic genes and the expression were increased in the optimized media. The results provided insight into the underlying mechanisms of media optimization.

Many biochemicals and biofuels produced at high concentrations are toxic to host organisms. Efflux pump engineering is an effective way to cope with the issue. In this thesis, native efflux pumps were screened for the AD-producing strain. The outer membrane *tolC* was found to be closely involved in the secretion of AD. The deletion of *tolC* significantly decreased the extracellular production of AD and increased the intracellular accumulation of AD. Supplementing certain combinations of *tolC*-related transporter systems increased AD titer significantly. The pump engineering approach developed herein can be readily applied to other products.

Lastly, the kinetics of the DXP pathway was studied in this thesis. The semi-closed dual -flux platform (SCDFP) was developed and optimized. Using this platform, the kinetic data of the DXP pathway were successfully obtained. It was found that the effluxes of DXP and MEC were closely related to their intracellular concentrations. The platform and knowledge obtained here is invaluable to understand the regulatory mechanisms of the pathway and to better engineer the DXP pathway in the future.

7.2 Recommendation of future works

In this thesis, the native transporters were studied for the efflux of two isoprenoids, lycopene and AD. The *tolC*-related transporters were able to enhance the efflux and the yield of AD. However, none of them was able to export lycopene, a very hydrophobic compound. As the accumulation of high concentration lycopene is very toxic to *E. coli* and this severely limits further improvement of lycopene production, it would be useful to screen or engineer a novel transporter for lycopene. We propose to evolve the high lycopene-producing strains in microfluidic devices with organic layer such as decane which can capture the exported lycopene. The microfluidic screening and assay system developed by Stephanopoulos's lab is perfect for this kind of study in which individual cells was compartmentalized for growth and analysis in monodisperse nanoliter aqueous droplets surrounded by an immiscible fluorinated oil phase [194]. Following that, a microarray or sequencing study on the mRNA or gene expression profiling will be carried out on those cells capable of secreting lycopene, and gained information can lead to novel discovery of efflux pumps for lycopene and other carotenoids.

The kinetics of the DXP pathway was investigated for the first time using the semi-closed platform designed in the thesis. Whereas the offline quantification method was not ideal for metabolic flux analysis as the time points interval was relatively protracted. In order to improve on the platform, online monitoring approach can be integrated in the future. Such

in vivo ^{13}C nuclear magnetic resonance or the MS-base quantitative real-time analysis system [191] have already been successfully used in other studies and may be useful for this endeavor. For example, bacterial cells will be grown and disrupted online, and the lysed cells were centrifuged and filtered through automation devices, and the resulting cell free extract is analyzed MS-based system.

In addition, the kinetic model can be further refined as the relationship between the efflux rates of metabolites and their intracellular concentrations are different from the relationship between the fluxes and chemical reactions. The complexity of the kinetics model makes the equations stiff and hence, it is very difficult to accurately estimate the parameters without deconvoluting the equations. Therefore, in the future, the decoupling strategy such as slope-estimation-decoupling or global stochastic search algorithm such as evolutionary computation technique can be investigated. The constructed mathematical model will be invaluable for model analyses and generation and testing of new hypotheses, such as the effect of DXP and MEC efflux on the pathway optimization.

Chapter 8. BIBLIOGRAPH

1. Breitmaier, E., *Terpenes: Importance, general structure, and biosynthesis.* Terpenes: Flavors, Fragrances, Pharmaca, Pheromones, 2006: p. 1-9.
2. Ajikumar, P.K., K. Tyo, S. Carlsen, O. Mucha, *et al.*, *Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms.* Molecular Pharmaceutics, 2008. **5**(2): p. 167-90.
3. Rodriguez-Concepcion, M., *The MEP pathway: A new target for the development of herbicides, antibiotics and antimalarial drugs.* Current pharmaceutical design, 2004. **10**(19): p. 2391-2400.
4. Kirby, J. and J.D. Keasling, *Metabolic engineering of microorganisms for isoprenoid production.* Nat Prod Rep, 2008. **25**(4): p. 656-61.
5. Van Noorden, R., *Demand for malaria drug soars.* Nature, 2010. **466**(7307): p. 672.
6. Kirby, J. and J.D. Keasling, *Biosynthesis of plant isoprenoids: perspectives for microbial engineering.* Annual Review of Plant Biology, 2009. **60**: p. 335-55.
7. Kolewe, M.E., V. Gaurav, and S.C. Roberts, *Pharmaceutically active natural product synthesis and supply via plant cell culture technology.* Molecular Pharmaceutics, 2008. **5**(2): p. 243-56.
8. Bailey, J.E., *Toward a science of metabolic engineering.* Science, 1991. **252**(5013): p. 1668-75.
9. Stephanopoulos, G., *Synthetic biology and metabolic engineering.* ACS synthetic biology, 2012. **1**(11): p. 514-525.

10. Medema, M.H., R. van Raaphorst, E. Takano, and R. Breitling, *Computational tools for the synthetic design of biochemical pathways.* Nature Publishing Group, 2012. **10**(3): p. 191-202.

11. Lee, J.W., D. Na, J.M. Park, J. Lee*, et al.*, *Systems metabolic engineering of microorganisms for natural and non-natural chemicals*. Nat Chem Biol, 2012. **8**(6): p. 536-546.

12. Rockman, M.V., *Reverse engineering the genotype-phenotype map with natural genetic variation.* Nature, 2008. **456**(7223): p. 738-44.

13. Ajikumar, P.K., W.H. Xiao, K.E. Tyo, Y. Wang*, et al.*, *Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli.* Science, 2010. **330**(6000): p. 70-4.

14. Leonard, E., P.K. Ajikumar, K. Thayer, W.H. Xiao*, et al.*, *Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control.* Proc Natl Acad Sci U S A, 2010. **107**(31): p. 13654-9.

15. Nakagawa, A., H. Minami, J.S. Kim, T. Koyanagi*, et al.*, *A bacterial platform for fermentative production of plant alkaloids.* Nat Commun, 2011. **2**: p. 326.

16. Alper, H., K. Miyaoku, and G. Stephanopoulos, *Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets.* Nature Biotechnology, 2005. **23**(5): p. 612-6.

17. Park, J.H., K.H. Lee, T.Y. Kim, and S.Y. Lee, *Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome*

analysis and in silico gene knockout simulation. Proc Natl Acad Sci U S A, 2007. **104**(19): p. 7797-802.

18. Paddon, C.J., P.J. Westfall, D.J. Pitera, K. Benjamin*, et al.*, *High-level semi-synthetic production of the potent antimalarial artemisinin.* Nature, 2013. **496**(7446): p. 528-32.
19. Yuan, L.Z., P.E. Rouviere, R.A. Larossa, and W. Suh, *Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in E. coli.* Metabolic Engineering, 2006. **8**(1): p. 79-90.
20. Albrecht, M., N. Misawa, and G. Sandmann, *Metabolic engineering of the terpenoid biosynthetic pathway of Escherichia coli for production of the carotenoids β-carotene and zeaxanthin.* Biotechnology Letters, 1999. **21**(9): p. 791-795.
21. Wang, C., S.-H. Yoon, A.A. Shah, Y.-R. Chung*, et al.*, *Farnesol production from Escherichia coli by harnessing the exogenous mevalonate pathway.* Biotechnology and Bioengineering, 2010: p. n-a-n-a.
22. Farmer, W.R. and J.C. Liao, *Precursor balancing for metabolic engineering of lycopene production in Escherichia coli.* Biotechnology Progress, 2001. **17**(1): p. 57-61.
23. Farmer, W.R. and J.C. Liao, *Improving lycopene production in Escherichia coli by engineering metabolic control.* Nature Biotechnology, 2000. **18**(5): p. 533-7.

24. Zhang, C., X. Chen, R. Zou, K. Zhou, *et al.*, *Combining Genotype Improvement and Statistical Media Optimization for Isoprenoid Production in E. coli.* PLoS One, 2013. **8**(10): p. e75164.

25. Hong, S. and S. Lee, *Importance of redox balance on the production of succinic acid by metabolically engineered Escherichia coli.* Applied microbiology and biotechnology, 2002. **58**(3): p. 286-290.

26. Zhou, S., T.B. Causey, A. Hasona, K.T. Shanmugam, *et al.*, *Production of optically pure D-lactic acid in mineral salts medium by metabolically engineered Escherichia coli W3110.* Appl Environ Microbiol, 2003. **69**(1): p. 399-407.

27. Lee, K.H., J.H. Park, T.Y. Kim, H.U. Kim, *et al.*, *Systems metabolic engineering of Escherichia coli for L-threonine production.* Mol Syst Biol, 2007. **3**: p. 149.

28. Runguphan, W. and J.D. Keasling, *Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals.* Metab Eng, 2013.

29. Song, C.W., D.I. Kim, S. Choi, J.W. Jang, *et al.*, *Metabolic engineering of Escherichia coli for the production of fumaric acid.* Biotechnol Bioeng, 2013. **110**(7): p. 2025-34.

30. Shen, C.R., E.I. Lan, Y. Dekishima, A. Baez, *et al.*, *Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli.* Appl Environ Microbiol, 2011. **77**(9): p. 2905-15.

31. Gawand, P., P. Hyland, A. Ekins, V.J. Martin*, et al.*, *Novel approach to engineer strains for simultaneous sugar utilization.* Metab Eng, 2013. **20C**: p. 63-72.

32. Flores, N., J. Xiao, A. Berry, F. Bolivar*, et al.*, *Pathway engineering for the production of aromatic compounds in Escherichia coli.* Nat Biotechnol, 1996. **14**(5): p. 620-3.

33. Foo, J.L. and S.S. Leong, *Directed evolution of an E. coli inner membrane transporter for improved efflux of biofuel molecules.* Biotechnol Biofuels, 2013. **6**(1): p. 81.

34. Fisher, M.A., S. Boyarskiy, M.R. Yamada, N. Kong*, et al.*, *Enhancing Tolerance to Short-Chain Alcohols by Engineering the Escherichia coli AcrB Efflux Pump to Secrete the Non-native Substrate n-Butanol.* ACS Synth Biol, 2013.

35. Dunlop, M.J., Z.Y. Dossani, H.L. Szmidt, H.C. Chu*, et al.*, *Engineering microbial biofuel tolerance and export using efflux pumps.* Mol Syst Biol, 2011. **7**: p. 487.

36. Doshi, R., T. Nguyen, and G. Chang, *Transporter-mediated biofuel secretion.* Proc Natl Acad Sci U S A, 2013. **110**(19): p. 7642-7.

37. Chemler, J.A., Z.L. Fowler, K.P. McHugh, and M.A. Koffas, *Improving NADPH availability for natural product biosynthesis in Escherichia coli by metabolic engineering.* Metab Eng, 2010. **12**(2): p. 96-104.

38. Martinez, K., R. de Anda, G. Hernandez, A. Escalante*, et al.*, *Coutilization of glucose and glycerol enhances the production of aromatic compounds in*

an Escherichia coli strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. Microb Cell Fact, 2008. **7**(1): p. 1.

39. Lee, H.C., J.S. Kim, W. Jang, and S.Y. Kim, *High NADPH/NADP+ ratio improves thymidine production by a metabolically engineered Escherichia coli strain.* J Biotechnol, 2010. **149**(1-2): p. 24-32.
40. Singh, A., K. Cher Soh, V. Hatzimanikatis, and R.T. Gill, *Manipulating redox and ATP balancing for improved production of succinate in E. coli.* Metab Eng, 2011. **13**(1): p. 76-81.
41. Auriol, C., G. Bestel-Corre, J.B. Claude, P. Soucaille*, et al.*, *Stress-induced evolution of Escherichia coli points to original concepts in respiratory cofactor selectivity.* Proc Natl Acad Sci U S A, 2011. **108**(4): p. 1278-83.
42. Wang, Y., K.Y. San, and G.N. Bennett, *Improvement of NADPH bioavailability in Escherichia coli by replacing NAD-dependent glyceraldehyde-3-phosphate dehydrogenase GapA with NADP -dependent GapB from Bacillus subtilis and addition of NAD kinase.* J Ind Microbiol Biotechnol, 2013.
43. Martinez, I., J. Zhu, H. Lin, G.N. Bennett*, et al.*, *Replacing Escherichia coli NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a NADP-dependent enzyme from Clostridium acetobutylicum facilitates NADPH dependent pathways.* Metab Eng, 2008. **10**(6): p. 352-9.
44. Pitera, D.J., C.J. Paddon, J.D. Newman, and J.D. Keasling, *Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli.* Metab Eng, 2007. **9**(2): p. 193-207.

45. Alper, H., C. Fischer, E. Nevoigt, and G. Stephanopoulos, *Tuning genetic control through promoter engineering.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12678-12683.

46. Pfleger, B.F., D.J. Pitera, C.D. Smolke, and J.D. Keasling, *Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes.* Nat Biotechnol, 2006. **24**(8): p. 1027-32.

47. Babiskin, A.H. and C.D. Smolke, *A synthetic library of RNA control modules for predictable tuning of gene expression in yeast.* Molecular Systems Biology, 2011. **7**.

48. Burack, W.R. and A.S. Shaw, *Signal transduction: hanging on a scaffold.* Curr Opin Cell Biol, 2000. **12**(2): p. 211-6.

49. Agapakis, C.M., P.M. Boyle, and P.A. Silver, *Natural strategies for the spatial optimization of metabolism in synthetic biology.* Nat Chem Biol, 2012. **8**(6): p. 527-35.

50. Dueber, J.E., G.C. Wu, G.R. Malmirchegini, T.S. Moon*, et al.*, *Synthetic protein scaffolds provide modular control over metabolic flux.* Nat Biotechnol, 2009. **27**(8): p. 753-9.

51. Delebecque, C.J., A.B. Lindner, P.A. Silver, and F.A. Aldaye, *Organization of intracellular reactions with rationally designed RNA assemblies.* Science, 2011. **333**(6041): p. 470-4.

52. Kim, H.J., T.L. Turner, and Y.S. Jin, *Combinatorial genetic perturbation to refine metabolic circuits for producing biofuels and biochemicals.* Biotechnol Adv, 2013. **31**(6): p. 976-85.

53. Rowlands, R., *Industrial strain improvement: mutagenesis and random screening procedures.* Enzyme and Microbial Technology, 1984. **6**(1): p. 3-10.

54. Wang, H.H., F.J. Isaacs, P.A. Carr, Z.Z. Sun*, et al.*, *Programming cells by multiplex genome engineering and accelerated evolution.* Nature, 2009. **460**(7257): p. 894-8.

55. Warner, J.R., P.J. Reeder, A. Karimpour-Fard, L.B.A. Woodruff*, et al.*, *Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides.* Nature Biotechnology, 2010. **28**(8): p. 856-862.

56. Alper, H. and G. Stephanopoulos, *Global transcription machinery engineering: a new approach for improving cellular phenotype.* Metab Eng, 2007. **9**(3): p. 258-67.

57. Alper, H., J. Moxley, E. Nevoigt, G.R. Fink*, et al.*, *Engineering yeast transcription machinery for improved ethanol tolerance and production.* Science, 2006. **314**(5805): p. 1565-8.

58. Notomista, E., R. Scognamiglio, L. Troncone, G. Donadio*, et al.*, *Tuning the specificity of the recombinant multicomponent toluene o-xylene monooxygenase from Pseudomonas sp. strain OX1 for the biosynthesis of tyrosol from 2-phenylethanol.* Appl Environ Microbiol, 2011. **77**(15): p. 5428-37.

59. Zhao, H. and F.H. Arnold, *Directed evolution converts subtilisin E into a functional equivalent of thermitase.* Protein engineering, 1999. **12**(1): p. 47-53.

60. Zhou, K., R. Zou, G. Stephanopoulos, and H.P. Too, *Enhancing solubility of deoxyxylulose phosphate pathway enzymes for microbial isoprenoid production.* Microb Cell Fact, 2012. **11**: p. 148.

61. Goldstein, J.L. and M.S. Brown, *Regulation of the mevalonate pathway.* Nature, 1990. **343**(6257): p. 425-30.

62. Rohmer, M., *The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants.* Nat Prod Rep, 1999. **16**(5): p. 565-74.

63. Ferrer-Miralles, N., J. Domingo-Espin, J.L. Corchero, E. Vazquez*, et al.*, *Microbial factories for recombinant pharmaceuticals.* Microb Cell Fact, 2009. **8**: p. 17.

64. Atsumi, S. and J.C. Liao, *Metabolic engineering for advanced biofuels production from Escherichia coli.* Curr Opin Biotechnol, 2008. **19**(5): p. 414-9.

65. Chen, X., L. Zhou, K. Tian, A. Kumar*, et al.*, *Metabolic engineering of Escherichia coli: A sustainable industrial platform for bio-based chemical production.* Biotechnol Adv, 2013.

66. Pandel, R., #x17e, Polj, #x161*, et al.*, *Skin Photoaging and the Role of Antioxidants in Its Prevention.* ISRN Dermatology, 2013. **2013**: p. 11.

67. Kucuk, O., F.H. Sarkar, W. Sakr, Z. Djuric*, et al.*, *Phase II randomized clinical trial of lycopene supplementation before radical prostatectomy.* Cancer Epidemiology Biomarkers & Prevention, 2001. **10**(8): p. 861-868.

68. Kucuk, O., F.H. Sarkar, Z. Djuric, W. Sakr, *et al.*, *Effects of lycopene supplementation in patients with localized prostate cancer.* Experimental Biology and Medicine, 2002. **227**(10): p. 881-885.

69. Schwarz, S., U.C. Obermüller-Jevic, E. Hellmis, W. Koch, *et al.*, *Lycopene inhibits disease progression in patients with benign prostate hyperplasia.* The Journal of nutrition, 2008. **138**(1): p. 49-53.

70. Dahan, K., M. Fennal, and N.B. Kumar, *Lycopene in the prevention of prostate cancer.* Journal of the Society for Integrative Oncology, 2007. **6**(1): p. 29-36.

71. Burton-Freeman, B.M. and H.D. Sesso, *Whole Food versus Supplement: Comparing the Clinical Evidence of Tomato Intake and Lycopene Supplementation on Cardiovascular Risk Factors.* Advances in Nutrition: An International Review Journal, 2014. **5**(5): p. 457-485.

72. Klayman, D.L., *Qinghaosu (artemisinin): an antimalarial drug from China.* Science, 1985. **228**(4703): p. 1049-1055.

73. World Health Organisation, *World Malaria Report 2012.* WHO, 2012.

74. Ragauskas, A.J., C.K. Williams, B.H. Davison, G. Britovsek, *et al.*, *The path forward for biofuels and biomaterials.* Science, 2006. **311**(5760): p. 484-9.

75. Kim, J.H., D.E. Block, and D.A. Mills, *Simultaneous consumption of pentose and hexose sugars: an optimal microbial phenotype for efficient fermentation of lignocellulosic biomass.* Appl Microbiol Biotechnol, 2010. **88**(5): p. 1077-85.

76. Kimata, K., H. Takahashi, T. Inada, P. Postma, *et al.*, *cAMP receptor protein–cAMP plays a crucial role in glucose–lactose diauxie by activating the major glucose transporter gene in Escherichia coli.* Proceedings of the National Academy of Sciences, 1997. **94**(24): p. 12914-12919.

77. Nichols, N., B. Dien, and R. Bothast, *Use of catabolite repression mutants for fermentation of sugar mixtures to ethanol.* Applied microbiology and biotechnology, 2001. **56**(1-2): p. 120-125.

78. Yomano, L.P., S.W. York, K.T. Shanmugam, and L.O. Ingram, *Deletion of methylglyoxal synthase gene (mgsA) increased sugar co-metabolism in ethanol-producing Escherichia coli.* Biotechnol Lett, 2009. **31**(9): p. 1389-98.

79. Yao, R., Y. Hirose, D. Sarkar, K. Nakahigashi, *et al.*, *Catabolic regulation analysis of Escherichia coli and its crp, mlc, mgsA, pgi and ptsG mutants.* Microb Cell Fact, 2011. **10**: p. 67.

80. Eiteman, M.A., S.A. Lee, and E. Altman, *A co-fermentation strategy to consume sugar mixtures effectively.* J Biol Eng, 2008. **2**: p. 3.

81. Gorke, B. and J. Stulke, *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nat Rev Microbiol, 2008. **6**(8): p. 613-24.

82. Desai, T.A. and C.V. Rao, *Regulation of Arabinose and Xylose Metabolism in Escherichia coli.* Appl Environ Microbiol, 2009. **76**(5): p. 1524-1532.

83. Chou, C.H., G.N. Bennett, and K.Y. San, *Effect of modified glucose uptake using genetic engineering techniques on high-level recombinant protein*

production in escherichia coli dense cultures. Biotechnology and Bioengineering, 1994. **44**(8): p. 952-60.

84. Lindner, S.N., G.M. Seibold, A. Henrich, R. Kramer*, et al.*, *Phosphotransferase system-independent glucose utilization in corynebacterium glutamicum by inositol permeases and glucokinases.* Appl Environ Microbiol, 2011. **77**(11): p. 3571-81.
85. Baez-Viveros, J.L., N. Flores, K. Juarez, P. Castillo-Espana*, et al.*, *Metabolic transcription analysis of engineered Escherichia coli strains that overproduce L-phenylalanine.* Microb Cell Fact, 2007. **6**: p. 30.
86. Zhang, X., K. Jantama, J.C. Moore, L.R. Jarboe*, et al.*, *Metabolic evolution of energy-conserving pathways for succinate production in Escherichia coli.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20180-5.
87. Patnaik, R. and J.C. Liao, *Engineering of Escherichia coli central metabolism for aromatic metabolite production with near theoretical yield.* Appl Environ Microbiol, 1994. **60**(11): p. 3903-8.
88. Gosset, G., *Production of aromatic compounds in bacteria.* Curr Opin Biotechnol, 2009. **20**(6): p. 651-8.
89. Lin, H., G.N. Bennett, and K.Y. San, *Metabolic engineering of aerobic succinate production systems in Escherichia coli to improve process productivity and achieve the maximum theoretical succinate yield.* Metab Eng, 2005. **7**(2): p. 116-27.
90. Hernandez-Montalvo, V., A. Martinez, G. Hernandez-Chavez, F. Bolivar*, et al.*, *Expression of galP and glk in a Escherichia coli PTS mutant restores*

glucose transport and increases glycolytic flux to fermentation products. Biotechnol Bioeng, 2003. **83**(6): p. 687-94.

91. Flores, N., L. Leal, J.C. Sigala, R. de Anda*, et al.*, *Growth recovery on glucose under aerobic conditions of an Escherichia coli strain carrying a phosphoenolpyruvate:carbohydrate phosphotransferase system deletion by inactivating arcA and overexpressing the genes coding for glucokinase and galactose permease.* J Mol Microbiol Biotechnol, 2007. **13**(1-3): p. 105-16.
92. Lee, P.C., B.N. Mijts, and C. Schmidt-Dannert, *Investigation of factors influencing production of the monocyclic carotenoid torulene in metabolically engineered Escherichia coli.* Applied microbiology and biotechnology, 2004. **65**(5): p. 538-46.
93. Koffas, M.A., G.Y. Jung, and G. Stephanopoulos, *Engineering metabolism and product formation in Corynebacterium glutamicum by coordinated gene overexpression.* Metab Eng, 2003. **5**(1): p. 32-41.
94. Santos, C.N. and G. Stephanopoulos, *Combinatorial engineering of microbes for optimizing cellular phenotype.* Curr Opin Chem Biol, 2008. **12**(2): p. 168-76.
95. Glick, B.R., *Metabolic load and heterologous gene expression.* Biotechnol Adv, 1995. **13**(2): p. 247-61.
96. Anthony, J.R., L.C. Anthony, F. Nowroozi, G. Kwon*, et al.*, *Optimization of the mevalonate-based isoprenoid biosynthetic pathway in Escherichia coli for production of the anti-malarial drug precursor amorpha-4,11-diene.* Metabolic Engineering, 2009. **11**(1): p. 13-9.

97. Yadav, V.G., M. De Mey, C.G. Lim, P.K. Ajikumar*, et al.*, *The future of metabolic engineering and synthetic biology: towards a systematic practice.* Metab Eng, 2012. **14**(3): p. 233-41.

98. Dunlop, M.J., *Engineering microbes for tolerance to next-generation biofuels.* Biotechnol Biofuels, 2011. **4**: p. 32.

99. Higgins, C.F., *ABC transporters: from microorganisms to man.* Annu Rev Cell Biol, 1992. **8**: p. 67-113.

100. Antoniewicz, M.R., J.K. Kelleher, and G. Stephanopoulos, *Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements.* Metab Eng, 2006. **8**(4): p. 324-37.

101. Antoniewicz, M.R., J.K. Kelleher, and G. Stephanopoulos, *Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions.* Metab Eng, 2007. **9**(1): p. 68-86.

102. Gombert, A.K. and J. Nielsen, *Mathematical modelling of metabolism.* Curr Opin Biotechnol, 2000. **11**(2): p. 180-6.

103. Schulz, A.R., *Enzyme kinetics: from diastase to multi-enzyme systems*. 1994: Cambridge University Press.

104. Chowdhury, A.R., M. Chetty, and N.X. Vinh, *Incorporating time-delays in S-System model for reverse engineering genetic networks.* BMC Bioinformatics, 2013. **14**: p. 196.

105. Nakayama, T., S. Seno, Y. Takenaka, and H. Matsuda, *Inference of S-system models of gene regulatory networks using immune algorithm.* J Bioinform Comput Biol, 2011. **9 Suppl 1**: p. 75-86.

106. Kimura, S., D. Araki, K. Matsumura, and M. Okada-Hatakeyama, *Inference of S-system models of genetic networks by solving one-dimensional function optimization problems.* Math Biosci, 2012. **235**(2): p. 161-70.

107. Calcada, D., S. Vinga, A.T. Freitas, and A.L. Oliveira, *Quantitative modeling of the Saccharomyces cerevisiae FLR1 regulatory network using an S-system formalism.* J Bioinform Comput Biol, 2011. **9**(5): p. 613-30.

108. Liu, L.Z., F.X. Wu, and W.J. Zhang, *Inference of biological S-system using the separable estimation method and the genetic algorithm.* IEEE/ACM Trans Comput Biol Bioinform, 2012. **9**(4): p. 955-65.

109. Lee, Y., P.W. Chen, and E.O. Voit, *Analysis of operating principles with S-system models.* Math Biosci, 2011. **231**(1): p. 49-60.

110. Vilela, M., I.C. Chou, S. Vinga, A.T. Vasconcelos*, et al.*, *Parameter optimization in S-system models.* BMC Syst Biol, 2008. **2**: p. 35.

111. Liu, P.K. and F.S. Wang, *Inference of biochemical network models in S-system using multiobjective optimization approach.* Bioinformatics, 2008. **24**(8): p. 1085-92.

112. Kutalik, Z., W. Tucker, and V. Moulton, *S-system parameter estimation for noisy metabolic profiles using newton-flow analysis.* IET Syst Biol, 2007. **1**(3): p. 174-80.

113. Gonzalez, O.R., C. Kuper, K. Jung, P.C. Naval, Jr.*, et al.*, *Parameter estimation using Simulated Annealing for S-system models of biochemical networks.* Bioinformatics, 2007. **23**(4): p. 480-6.

114. Hasegawa, T. and J. Yoshimura, *An effective method to increase solvability in biochemical systems using S-system.* Math Biosci, 2006. **201**(1-2): p. 125-42.

115. Voit, E.O., *Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists*. 2000: Cambridge University Press.

116. Goel, G., I.C. Chou, and E.O. Voit, *System estimation from metabolic time-series data.* Bioinformatics, 2008. **24**(21): p. 2505-11.

117. Jia, G., G.N. Stephanopoulos, and R. Gunawan, *Parameter estimation of kinetic models from metabolic profiles: two-phase dynamic decoupling method.* Bioinformatics, 2011. **27**(14): p. 1964-70.

118. Chatterjee, R., C.S. Millard, K. Champion, D.P. Clark*, et al.*, *Mutation of the ptsG Gene Results in Increased Production of Succinate in Fermentation of Glucose by Escherichia coli.* Applied and environmental microbiology, 2001. **67**(1): p. 148-154.

119. Flores, S., G. Gosset, N. Flores, A.A. de Graaf*, et al.*, *Analysis of Carbon Metabolism in Escherichia coli Strains with an Inactive Phosphotransferase System by 13C Labeling and NMR Spectroscopy.* Metabolic Engineering, 2002. **4**(2): p. 124-137.

120. Lu, J., J. Tang, Y. Liu, X. Zhu*, et al.*, *Combinatorial modulation of galP and glk gene expression for improved alternative glucose utilization.* Applied Microbiology and Biotechnology, 2012. **93**(6): p. 2455-2462.

121. Yi, J., K.M. Draths, K. Li, and J.W. Frost, *Altered glucose transport and shikimate pathway product yields in E. coli.* Biotechnology Progress, 2003. **19**(5): p. 1450-1459.

122. De Anda, R., A.R. Lara, V. Hernandez, V. Hernandez-Montalvo*, et al.*, *Replacement of the glucose phosphotransferase transport system by galactose permease reduces acetate accumulation and improves process performance of Escherichia coli for recombinant protein production without impairment of growth rate.* Metab Eng, 2006. **8**(3): p. 281-90.

123. Jones, K.L., S.-W. Kim, and J.D. Keasling, *Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria.* Metabolic Engineering, 2000. **2**(4): p. 328.

124. Too;, H.-P., R. Zou;, and G. Stephanopoulas, *Univariant Extrinsic Initiator Control System for microbes and an in vitro assembly of large recombinant dna molecules from multiple components*. 2012: US.

125. HP, T., Z. R, and G. Stephanopolous, *UNivariant Extrinsic Initiator Control System for Microbes (μ-UNIECS) and An In Vitro Assembly of Large Recombinant DNA Molecules from Multiple Components (CLIVA)* 2013: USA.

126. Chen, X., C. Zhang, R. Zou, K. Zhou*, et al.*, *Statistical Experimental Design Guided Optimization of a One-Pot Biphasic Multienzyme Total Synthesis of Amorpha-4,11-diene.* PLoS One, 2013. **8**(11): p. e79650.

127. Green, S., C.J. Squire, N.J. Nieuwenhuizen, E.N. Baker*, et al.*, *Defining the potassium binding region in an apple terpene synthase.* J Biol Chem, 2009. **284**(13): p. 8661-9.

128. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.

129. Zou, R., K. Zhou, G. Stephanopoulos, and H.P. Too, *Combinatorial Engineering of 1-Deoxy-D-Xylulose 5-Phosphate Pathway Using Cross-Lapping In Vitro Assembly (CLIVA) Method.* Plos one, 2013. **8**(11): p. e79557.

130. Tsuruta, H., C.J. Paddon, D. Eng, J.R. Lenihan*, et al.*, *High-Level Production of Amorpha-4,11-Diene, a Precursor of the Antimalarial Agent Artemisinin, in Escherichia coli.* PloS one, 2009. **4**(2): p. e4489.

131. Cunningham, F.X., Jr., Z. Sun, D. Chamovitz, J. Hirschberg*, et al.*, *Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium Synechococcus sp strain PCC7942.* Plant Cell, 1994. **6**(8): p. 1107-21.

132. Islam, R.S., D. Tisi, M.S. Levy, and G.J. Lye, *Framework for the rapid optimization of soluble protein expression in Escherichia coli combining microscale experiments and statistical experimental design.* Biotechnology Progress, 2007. **23**(4): p. 785-93.

133. Iyer, P.V. and R.S. Singhal, *Screening and Selection of Marine Isolate for l-Glutaminase Production and Media Optimization Using Response Surface*

Methodology. Applied Biochemistry and Biotechnology, 2009. **159**(1): p. 233-250.

134. Mokhtarihosseini, Z., E. Vasheghanifarahani, A. Heidarzadehvazifekhoran, S. Shojaosadati*, et al.*, *Statistical media optimization for growth and PHB production from methanol by a methylotrophic bacterium.* Bioresource Technology, 2009. **100**(8): p. 2436-2443.
135. Zhou, Z., G. Du, Z. Hua, J. Zhou*, et al.*, *Optimization of fumaric acid production by Rhizopus delemar based on the morphology formation.* Bioresource Technology, 2011. **102**(20): p. 9345-9.
136. Kang, M.J., Y.M. Lee, S.H. Yoon, J.H. Kim*, et al.*, *Identification of genes affecting lycopene accumulation in Escherichia coli using a shot-gun method.* Biotechnology and bioengineering, 2005. **91**(5): p. 636-42.
137. Kim, S.W. and J.D. Keasling, *Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in Escherichia coli enhances lycopene production.* Biotechnology and bioengineering, 2001. **72**(4): p. 408-15.
138. Newman, J.D., J. Marshall, M. Chang, F. Nowroozi*, et al.*, *High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered Escherichia coli.* Biotechnology and bioengineering, 2006. **95**(4): p. 684-91.
139. Alper, H., Y.S. Jin, J.F. Moxley, and G. Stephanopoulos, *Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli.* Metabolic Engineering, 2005. **7**(3): p. 155-64.

140. Jin, Y.S. and G. Stephanopoulos, *Multi-dimensional gene target search for improving lycopene biosynthesis in Escherichia coli.* Metabolic Engineering, 2007. **9**(4): p. 337-47.

141. Ryan, W. and S.J. Parulekar, *Recombinant protein synthesis and plasmid instability in continuous cultures of Escherichia coli JM103 harboring a high copy number plasmid.* Biotechnol Bioeng, 1991. **37**(5): p. 415-29.

142. Riesenberg, D., V. Schulz, W.A. Knorre, H.D. Pohl*, et al.*, *High cell density cultivation of Escherichia coli at controlled specific growth rate.* J Biotechnol, 1991. **20**(1): p. 17-27.

143. Alper, H., K. Miyaoku, and G. Stephanopoulos, *Characterization of lycopene-overproducing E. coli strains in high cell density fermentations.* Appl Microbiol Biotechnol, 2006. **72**(5): p. 968-74.

144. Brynildsen, M.P. and J.C. Liao, *An integrated network approach identifies the isobutanol response network of Escherichia coli.* Mol Syst Biol, 2009. **5**: p. 277.

145. Nguyen, Q.T., M.E. Merlo, M.H. Medema, A. Jankevics*, et al.*, *Metabolomics methods for the synthetic biology of secondary metabolism.* FEBS Lett, 2012. **586**(15): p. 2177-83.

146. Weissenborn, D.L., N. Wittekindt, and T.J. Larson, *Structure and regulation of the glpFK operon encoding glycerol diffusion facilitator and glycerol kinase of Escherichia coli K-12.* J Biol Chem, 1992. **267**(9): p. 6122-31.

147. Freedberg, W.B. and E.C. Lin, *Three kinds of controls affecting the expression of the glp regulon in Escherichia coli.* J Bacteriol, 1973. **115**(3): p. 816-23.

148. Gabor, E., A.K. Gohler, A. Kosfeld, A. Staab*, et al.*, *The phosphoenolpyruvate-dependent glucose-phosphotransferase system from Escherichia coli K-12 as the center of a network regulating carbohydrate flux in the cell.* Eur J Cell Biol, 2011. **90**(9): p. 711-20.

149. Zhang, X. and R. Schleif, *Catabolite gene activator protein mutations affecting activity of the araBAD promoter.* J Bacteriol, 1998. **180**(2): p. 195-200.

150. Mandenius, C.-F. and A. Brundin, *Bioprocess optimization using design-of-experiments methodology.* Biotechnology Progress, 2008. **24**(6): p. 1191-1203.

151. Carnes, A.E., *Fermentation design for the manufacture of therapeutic plasmid DNA.* BioProcess Int, 2005. **3**(9): p. 36-44.

152. O'Kennedy, R.D., C. Baldwin, and E. Keshavarz-Moore, *Effects of growth medium selection on plasmid DNA production and initial processing steps.* J Biotechnol, 2000. **76**(2-3): p. 175-83.

153. Hofmann, K., P. Neubauer, S. Riethdorf, and M. Hecker, *Amplification of pBR322 plasmid DNA in Escherichia coli relA strains during batch and fed-batch fermentation.* J Basic Microb, 1990. **30**(1): p. 37-41.

154. Prather, K.J., S. Sagar, J. Murphy, and M. Chartrain, *Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design,*

production, and purification. Enzyme and Microbial Technology, 2003. **33**(7): p. 865-883.

155. Blanusa, M., A. Schenk, H. Sadeghi, J. Marienhagen*, et al.*, *Phosphorothioate-based ligase-independent gene cloning (PLICing): An enzyme-free and sequence-independent cloning method.* Anal Biochem, 2010. **406**(2): p. 141-6.
156. Zhou, K., L. Zhou, Q.E. Lim, R. Zou*, et al.*, *Novel reference genes for quantifying transcriptional responses of Escherichia coli to protein overexpression by quantitative PCR.* BMC Mol Biol, 2011. **12**(1): p. 18.
157. Zhou, K., R. Zou, G. Stephanopoulos, and H.P. Too, *Metabolite profiling identified methylerythritol cyclodiphosphate efflux as a limiting step in microbial isoprenoid production.* PLoS One, 2012. **7**(11): p. e47513.
158. Orjuela, A., C.T. Lira, and D.J. Miller, *A novel process for recovery of fermentation-derived succinic acid: process design and economic analysis.* Bioresour Technol, 2013. **139**: p. 235-41.
159. Piddock, L.J.V., *Multidrug-resistance efflux pumps - not just for resistance.* Nature Reviews Microbiology, 2006. **4**(8): p. 629-636.
160. Li, X.Z. and H. Nikaido, *Efflux-mediated drug resistance in bacteria.* Drugs, 2004. **64**(2): p. 159-204.
161. Allen, H.K., J. Donato, H.H. Wang, K.A. Cloud-Hansen*, et al.*, *Call of the wild: antibiotic resistance genes in natural environments.* Nat Rev Microbiol, 2010. **8**(4): p. 251-9.

162. Zhang, Y., M. Xiao, T. Horiyama, X. Li, *et al.*, *The multidrug efflux pump MdtEF protects against nitrosative damage during the anaerobic respiration in Escherichia coli.* J Biol Chem, 2011. **286**(30): p. 26576-84.

163. Wang, J.-F., Z.-Q. Xiong, S.-Y. Li, and Y. Wang, *Enhancing isoprenoid production through systematically assembling and modulating efflux pumps in Escherichia coli.* Applied microbiology and biotechnology, 2013. **97**(18): p. 8057-8067.

164. Qian, Z.G., X.X. Xia, and S.Y. Lee, *Metabolic engineering of Escherichia coli for the production of cadaverine: a five carbon diamine.* Biotechnology and bioengineering, 2011. **108**(1): p. 93-103.

165. Qian, Z.G., X.X. Xia, and S.Y. Lee, *Metabolic engineering of Escherichia coli for the production of putrescine: a four carbon diamine.* Biotechnology and bioengineering, 2009. **104**(4): p. 651-662.

166. Baba, T., T. Ara, M. Hasegawa, Y. Takai, *et al.*, *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Mol Syst Biol, 2006. **2**: p. 2006 0008.

167. Paulsen, I.T., J.H. Park, P.S. Choi, and M.H. Saier, Jr., *A family of gram-negative bacterial outer membrane factors that function in the export of proteins, carbohydrates, drugs and heavy metals from gram-negative bacteria.* FEMS Microbiol Lett, 1997. **156**(1): p. 1-8.

168. Hoeffler, J.F., D. Tritsch, C. Grosdemange-Billiard, and M. Rohmer, *Isoprenoid biosynthesis via the methylerythritol phosphate pathway.*

Mechanistic investigations of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. Eur J Biochem, 2002. **269**(18): p. 4446-57.

169. Xiao, Y., D. Rooker, Q. You, C.L. Meyers*, et al.*, *IspG-catalyzed positional isotopic exchange in methylerythritol cyclodiphosphate of the deoxyxylulose phosphate pathway: mechanistic implications.* Chembiochem, 2011. **12**(4): p. 527-30.
170. Rohdich, F., S. Hecht, K. Gartner, P. Adam*, et al.*, *Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1158-63.
171. Ruiz, N., D. Kahne, and T.J. Silhavy, *Transport of lipopolysaccharide across the cell envelope: the long road of discovery.* Nat Rev Microbiol, 2009. **7**(9): p. 677-83.
172. Sperandeo, P., R. Cescutti, R. Villa, C. Di Benedetto*, et al.*, *Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli.* J Bacteriol, 2007. **189**(1): p. 244-53.
173. Polissi, A. and C. Georgopoulos, *Mutational analysis and properties of the msbA gene of Escherichia coli, coding for an essential ABC family transporter.* Mol Microbiol, 1996. **20**(6): p. 1221-33.
174. Pohl, A., P.F. Devaux, and A. Herrmann, *Function of prokaryotic and eukaryotic ABC proteins in lipid transport.* Biochim Biophys Acta, 2005. **1733**(1): p. 29-52.

175. Nicolaou, S.A., S.M. Gaida, and E.T. Papoutsakis, *A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation.* Metab Eng, 2010. **12**(4): p. 307-31.

176. Nishino, K. and A. Yamaguchi, *Analysis of a complete library of putative drug transporter genes in Escherichia coli.* J Bacteriol, 2001. **183**(20): p. 5803-12.

177. Ramos, J.L., E. Duque, M.T. Gallegos, P. Godoy*, et al.*, *Mechanisms of solvent tolerance in gram-negative bacteria.* Annu Rev Microbiol, 2002. **56**: p. 743-68.

178. Ramos, J.L., E. Duque, J.J. Rodriguez-Herva, P. Godoy*, et al.*, *Mechanisms for solvent tolerance in bacteria.* J Biol Chem, 1997. **272**(7): p. 3887-90.

179. Brown, M.H. and R.A. Skurray, *Staphylococcal multidrug efflux protein QacA.* J Mol Microbiol Biotechnol, 2001. **3**(2): p. 163-70.

180. Nikaido, H., *Multidrug efflux pumps of gram-negative bacteria.* J Bacteriol, 1996. **178**(20): p. 5853-9.

181. Tegos, G.P. and M.R. Hamblin, *Phenothiazinium antimicrobial photosensitizers are substrates of bacterial multidrug resistance pumps.* Antimicrob Agents Chemother, 2006. **50**(1): p. 196-203.

182. Aono, R., N. Tsukagoshi, and M. Yamamoto, *Involvement of outer membrane protein TolC, a possible member of the mar-sox regulon, in maintenance and improvement of organic solvent tolerance of Escherichia coli K-12.* J Bacteriol, 1998. **180**(4): p. 938-44.

183. Sulavik, M.C., C. Houseweart, C. Cramer, N. Jiwani, *et al.*, *Antibiotic susceptibility profiles of Escherichia coli strains lacking multidrug efflux pump genes.* Antimicrob Agents Chemother, 2001. **45**(4): p. 1126-36.

184. Zgurskaya, H.I., G. Krishnamoorthy, A. Ntreh, and S. Lu, *Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria.* Frontiers in microbiology, 2011. **2**.

185. Karow, M. and C. Georgopoulos, *The essential Escherichia coli msbA gene, a multicopy suppressor of null mutations in the htrB gene, is related to the universally conserved family of ATP-dependent translocators.* Mol Microbiol, 1993. **7**(1): p. 69-79.

186. Krulwich, T.A., O. Lewinson, E. Padan, and E. Bibi, *Do physiological roles foster persistence of drug/multidrug-efflux transporters? A case study.* Nat Rev Microbiol, 2005. **3**(7): p. 566-72.

187. Wungsintaweekul, J., S. Herz, S. Hecht, W. Eisenreich, *et al.*, *Phosphorylation of 1-deoxy-D-xylulose by D-xylulokinase of Escherichia coli.* Eur J Biochem, 2001. **268**(2): p. 310-6.

188. Chou, I.C. and E.O. Voit, *Recent developments in parameter estimation and structure identification of biochemical and genomic systems.* Math Biosci, 2009. **219**(2): p. 57-83.

189. Gabrielsen, M., C.S. Bond, I. Hallyburton, S. Hecht, *et al.*, *Hexameric assembly of the bifunctional methylerythritol 2,4-cyclodiphosphate synthase and protein-protein associations in the deoxy-xylulose-dependent pathway*

of isoprenoid precursor biosynthesis. J Biol Chem, 2004. **279**(50): p. 52753-61.

190. Quek, L.E., C. Wittmann, L.K. Nielsen, and J.O. Kromer, *OpenFLUX: efficient modelling software for 13C-based metabolic flux analysis.* Microb Cell Fact, 2009. **8**: p. 25.
191. Bujara, M., M. Schumperli, R. Pellaux, M. Heinemann*, et al.*, *Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis.* Nat Chem Biol, 2011. **7**(5): p. 271-7.
192. Fonseca, C., A.R. Neves, A.M. Antunes, J.P. Noronha*, et al.*, *Use of in vivo 13C nuclear magnetic resonance spectroscopy to elucidate L-arabinose metabolism in yeasts.* Appl Environ Microbiol, 2008. **74**(6): p. 1845-55.
193. Siedler, S., S.G. Stahlhut, S. Malla, J. Maury*, et al.*, *Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into Escherichia coli.* Metab Eng, 2013. **21C**: p. 2-8.
194. Wang, B.L., A. Ghaderi, H. Zhou, J. Agresti*, et al.*, *Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption.* Nat Biotechnol, 2014. **32**(5): p. 473-8.

LIST OF RELATED PUBLICATIONS AND INVENTION DISCLOSURE

1. **Zhang C**, Chen X, Zou R, Zhou K, Stephanopoulos G, Too, HP. (2013) Combining Genotype Improvement and Statistical Media Optimization for Isoprenoid Production in E. coli. PLoS One 8: e75164. (Chapter 3).
2. **Zhang C.**, Chen X, Zou R, Zhou K, Stephanopoulos G, Too HP. (2014) Experimental design aided systematic pathway optimization of glucose uptake and deoxyxylulose phosphate pathway to enhance the production of amorphadiene. (Under review, Biotechnol Bioeng) (Chapter 2)
3. **Zhang C**., Chen X, Stephanopoulos G, Too HP. Transporter engineering for the production of isoprenoids. Manuscript in preparation. (Chapter 4)
4. Zhang C., Chen X, Zou R, Too HP. High-level production of lycopene and amorphadiene in Escherichia coli. (Manuscript in preparation)
5. Chen X, **Zhang C**, Zou R, Zhou K, Stephanopoulos G, et al. (2013) Statistical Experimental Design Guided Optimization of a One-Pot Biphasic Multienzyme Total Synthesis of Amorpha-4,11-diene. PLoS One 8: e79650.
6. Zhou K, Zou R, **Zhang C**, Stephanopoulos G, Too HP (2013) Optimization of amorphadiene synthesis in bacillus subtilis via

transcriptional, translational, and media modulation. Biotechnol Bioeng 110: 2556-2561.

7. Chen X, **Zhang C**, Zou R, Stephanopoulos G, Too HP. Unravel the regulatory behavior of the in vitro reconstituted amorpha-4,11-diene synthesis pathway by Lin-log approximation. (Manuscript in preparation)
8. Chen X, Zou R, **Zhang C**, Stephanopoulos G, Too HP. Hybrid in vivo and in vitro production of artemisinic acid. (Manuscript in preparation)
9. In vitro synthetic multi-biocatalytic system for the total synthesis of isoprenoids and isoprenoid precursors. US Provisional Application No. 61/871,940. Inventor: Too Heng Phon, Chen Xixian, **Zhang Congqiang**, Zou Ruiyang.