

*De novo* synthesis and derivatization of norlaudanosoline in yeast

Lauren Narcross

A Thesis

In the Department

of

Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Biology) at Concordia University

Montreal, Quebec, Canada

July 2021

© Lauren Narcross, 2021

**CONCORDIA UNIVERSITY**  
**SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Lauren Narcross

Entitled: *De novo* synthesis and derivatization of norlaudanosoline in yeast

and submitted in partial fulfillment of the requirements for the degree of  
Doctor Of Philosophy (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

\_\_\_\_\_ Chair  
Dr. David Mumby

\_\_\_\_\_ External Examiner  
Dr. Mattheos Koffas

\_\_\_\_\_ External to Program  
Dr. Peter Pawelek

\_\_\_\_\_ Thesis Supervisor  
Dr. Vincent Martin

\_\_\_\_\_ Examiner  
Dr. David Walsh

\_\_\_\_\_ Examiner  
Dr. Malcolm Whiteway

Approved by \_\_\_\_\_  
Dr. Robert Weladji, Graduate Program Director

8/26/2021 \_\_\_\_\_  
Dr. Pascale Sicotte, Dean  
Faculty of Arts and Science

## Abstract

### *De novo* synthesis and derivatization of norlaudanosoline in yeast

Lauren Narcross, Ph.D.

Concordia University, 2021

Benzylisoquinoline alkaloids (BIAs) are a class of plant secondary metabolites with interesting pharmaceutical properties, such as analgesic, anti-cancer, anti-spasmodic, anti-microbial, anti-tussive, vasodilation, and muscle relaxant. The BIA family includes the opiates morphine and codeine. Several other BIAs are also commercially available; relevant to this thesis is sanguinarine, which is marketed as a growth enhancer for animal feed by the company Phytobiotics. To date, commercial production of almost all BIAs relies on their extraction from plants. However, the rarity of most BIAs in nature and the cost and difficulty of their chemical synthesis limits research and potential valorization of most of these compounds. An alternative route to BIA synthesis could enable sustainable production of BIAs currently extracted from plants while also expanding access to rarer alkaloids. Synthesis in a microbial host like baker's yeast is an attractive option due to cheap feedstock, genetic tractability, and ease of scaleup. The highest BIA titer produced in a microbial host is currently 4.6 g/L of (*S*)-reticuline, the last common pathway intermediate to several commercially relevant alkaloids, including sanguinarine. This success has not yet been carried forward to downstream targets, whose reported titers remain at the microgram to low milligram scale.

This thesis is focused on the development of a microbial platform for the synthesis of the BIA dihydrosanguinarine, the direct precursor of sanguinarine. In 2014 our group demonstrated the reconstruction of a partial synthesis pathway in yeast, resulting in 1.5% conversion of the supplemented BIA norlaudanosoline to dihydrosanguinarine. Here, we improve conversion of

norlaudanoline to dihydrosanguinarine to 10% by eliminating a pathway bottleneck. Next, we introduce *de novo* norlaudanoline synthesis to yeast, which improves yields and reduces side products compared to the route previously used for BIA synthesis. Finally, we extend reticuline synthesis to dihydrosanguinarine, demonstrating its first *de novo* production in a microbial host and culminating in 630 mg/L of dihydrosanguinarine and sanguinarine in fed-batch fermentation. This represents a >300-fold increase over the previous highest reported titer of a commercial BIA target – a major milestone in the quest to make the entire BIA family commercially accessible.

## Acknowledgements

First and foremost, thanks to Vincent Martin for many years of supervision. I've spent the better part of a decade working with you in various capacities and your lab has been a wonderful place to grow as a scientist. Thanks as well to my committee members David Walsh and Malcolm Whiteway for critical guidance during the development of this project. Thanks also to John Dueber, who is always up for a conversation and whose trainees have been so friendly and so critically helpful.

Thank you to Audrey Morin and Chris Plenzich, without whom our lab would dissolve into complete chaos.

Our lab's BIA work in yeast was initially established by Elena Fossati and Andy Ekins. Introduction of *de novo* BIA synthesis in yeast was a collaborative effort between our lab and William deLoache and Perry Grewal in John Dueber's lab. Michael Pyne and Kaspar Kevvai came into the lab, took the system and turned it into one of the best case-studies in heterologous metabolite production in yeast (in my non-biased opinion). It's been a pleasure working with each and every one of you. I'm truly standing on the shoulders of giants.

I've been lucky enough to learn how to operate a number of machines thanks to the people who took the time to teach me. A huge thanks to Jean-Pierre Falgueyret and Marcos diFalco for maintenance of and training on the HPLC-FT-ICR-MS and the HPLC-qTOF-MS, and thanks especially to Marcos for pushing countless dewars of nitrogen through the GE basement. Thanks to Meghan Davies for the training on the GC-MS, and Kaspar Kevvai for the training on the bioreactors.

Thanks to everyone in the Martin lab and Concordia over the years who have been there for support, guidance, and after-hours beer, especially Shoham Mookerjee, Meghan Davies, Mindy

Melgar, David Colatriano, Arthi Ramachandran, James Perry, James Dhaliwal, Mohammed Nasr, and Bjorn Bean. Thanks also to my friends Natalie Rachel and Alec Sweet, who were always there for me in thick and thin. Thank you to my sister Hannah Narcross for insisting that there is light at the end of the PhD tunnel, and my mom Louise Fletcher for being a Woman in Science before people cared about that sort of thing. And to both of you for not pestering me too badly about when I was going to graduate.

Finally, to Kaspar: you came into the lab and changed the course of my thesis and more importantly changed my life. Thanks for patiently explaining and re-explaining basic concepts in fermentation, handling panicked bioreactor phone calls, many Friday night beers, picnics, putzes, road trips, and especially for cooking and cleaning and generally shouldering all responsibilities while I stared at my thesis for 14 hours straight for days on end. I love you and I'm looking forward to what comes next.

## **Contribution of authors**

I am the principal author of this thesis. Additionally, I am primarily responsible for all conception of experiments, experimental design, data collection, data analysis, and manuscript preparation. I also co-ordinated the assistance of co-authors for each manuscript reproduced in this thesis, whose contributions are listed below.

Chapter 2: Elena Fossati generated Figure 2.2 and contributed text to Section 2.10. Leanne Bourgeois generated Table 2.1. All co-authors provided feedback on the manuscript.

Chapter 3: The design, purchase, and handling of the synthetic DNA described in Chapter 3 was a collaborative effort between Elena Fossati, Leanne Bourgeois, Euan Burton and myself. Euan Burton conceptualized and constructed the pBOT vector system. Leanne Bourgeois performed the experiments that are presented in Figure 3.5. All co-authors provided feedback on the manuscript.

Chapter 4: Michael Pyne constructed strain LP524. Kaspar Kevvai provided training on Applikon 3L bioreactors, established the fermentation conditions described in the chapter, and consulted on troubleshooting and data analysis. Ka-Hui Siu was responsible for the conception and performance of experiments that are presented in Figure 4.7 and Figure 4.8. Smita Amarnath and Nicholas Gold constructed vector pGC1899. Michael Pyne, Kaspar Kevvai, and Vincent Martin provided feedback on the manuscript.

Finally, Mohammed Nasr, Michael Pyne, Kaspar Kevvai, and Vincent Martin provided feedback on the thesis during its preparation.

## Table of Contents

List of Figures .....	x
List of Tables.....	xi
List of Abbreviations .....	xii
<b>Chapter One: Introduction .....</b>	<b>1</b>
1.1 Research objectives.....	3
1.2 Thesis organization.....	4
<b>2 Chapter Two: Literature Review .....</b>	<b>6</b>
2.1 Trends.....	6
2.2 Abstract .....	6
2.3 Microbial synthesis of pharmaceuticals to enhance drug discovery .....	7
2.4 Pharmaceutical properties of benzyloquinoline alkaloids .....	7
2.5 BIA diversity is derived from a single scaffold .....	8
2.6 Current status of microbial aromatic amino acid (AA) production .....	12
2.6.1 Update in 2021 .....	12
2.7 Aldehyde scavenging limits the available 4-HPAA pool .....	13
2.7.1 Update in 2021 .....	13
2.8 Multiple strategies for dopamine synthesis in microbial hosts.....	14
2.9 BIA scaffolds can be synthesized <i>de novo</i> from simple sugars .....	17
2.9.1 Norlaudanoline synthesis in <i>E. coli</i> .....	18
2.9.2 Norcoclaurine synthesis in <i>S. cerevisiae</i> .....	21
2.10 Synthesis and derivatization of morphinan alkaloids in <i>S. cerevisiae</i> .....	23
2.10.1 Derivatization of morphinans from supplemented precursors .....	26
2.10.2 Discovery of reticuline epimerase enables <i>de novo</i> synthesis of opiates .....	29
2.10.3 Update in 2021 .....	30
2.11 Derivatization of other BIA alkaloids in <i>S. cerevisiae</i> .....	31
2.11.1 Present and future diversity of backbone synthesis.....	31
2.11.2 Update in 2021: Non-natural BIA synthesis in microbial hosts.....	32
2.11.3 BIA derivatization in microbes requires knowledge of synthesis pathways .....	33
2.12 Biosecurity of opiate production in microbes .....	33
2.13 Addressing promiscuity in heterologous pathway reconstitution .....	34
2.14 Growth conditions for BIA synthesis.....	37
2.15 Concluding remarks .....	38
<b>3 Chapter Three: Optimization of Dihydrosanguinarine Pathway in Yeast.....</b>	<b>39</b>
3.1 Abstract .....	39
3.2 Introduction .....	40
3.3 Results.....	45
3.3.1 Generation of CYP719 and NMT enzyme libraries.....	45
3.3.2 Selection of replacement Ring A-closing CYP719s .....	50

3.3.3	Engineering of a non-natural stylophine synthesis pathway.....	56
3.3.4	Combinatorial testing of Ring A- and Ring B-closing CYP719s .....	56
3.3.5	Selection of replacement TNMT .....	59
3.3.6	Retrofitting and testing an optimized dihydrosanguinarine producing strain .....	61
<b>3.4</b>	<b>Discussion .....</b>	<b>63</b>
<b>3.5</b>	<b>Materials and Methods .....</b>	<b>66</b>
3.5.1	Transcriptome data analysis.....	66
3.5.2	Construction of plasmids .....	67
3.5.3	Chromosomal integration of genes and multi-gene pathways .....	68
3.5.4	BIA culture supplementation assays .....	69
3.5.5	Liquid chromatography-mass spectrometry.....	69
<b>3.6</b>	<b>Supplemental materials for Chapter 3 .....</b>	<b>70</b>
3.6.1	Expression analysis of CYP719s by flow cytometry.....	70
3.6.2	Supplemental materials and methods.....	73
3.6.3	Supplemental tables.....	75
<b>4</b>	<b>Chapter 4: <i>De novo</i> Production of Norlaudanosoline for Dihydrosanguinarine Synthesis in Yeast.....</b>	<b>95</b>
<b>4.1</b>	<b>Abstract .....</b>	<b>95</b>
<b>4.2</b>	<b>Introduction .....</b>	<b>96</b>
<b>4.3</b>	<b>Results.....</b>	<b>100</b>
4.3.1	<i>ARO10</i> knockout eliminates <i>de novo</i> 4'-dehydroxynorcoclaurine synthesis .....	100
4.3.2	Expression of <i>HsMAO-A</i> for norlaudanosoline synthesis.....	103
4.3.3	Fed-batch fermentation with <i>HsMAO-A</i> driven by <i>pTEF2</i> .....	107
4.3.4	Optimization of norlaudanosoline synthesis through branch point balancing .....	111
4.3.5	<i>De novo</i> dihydrosanguinarine synthesis <i>via</i> norlaudanosoline .....	113
<b>4.4</b>	<b>Discussion .....</b>	<b>120</b>
<b>4.5</b>	<b>Materials and Methods .....</b>	<b>123</b>
4.5.1	Yeast and <i>E. coli</i> growth conditions .....	123
4.5.2	Strain construction .....	124
4.5.3	Growth curves and determination of maximum growth rate .....	125
4.5.4	Fed-batch fermentation.....	125
4.5.5	High pressure liquid chromatography analysis by ultraviolet absorbance (HPLC-UV) and mass spectrometry (HPLC-MS).....	126
<b>4.6</b>	<b>Supplemental materials for Chapter 4 .....</b>	<b>129</b>
4.6.1	Supplemental figures .....	129
4.6.2	Supplemental materials and methods.....	135
4.6.3	Supplemental tables.....	137
<b>5</b>	<b>Chapter Five: Conclusion and Perspectives .....</b>	<b>145</b>
<b>5.1</b>	<b>The necessity of multi-copy gene integration .....</b>	<b>147</b>
<b>5.2</b>	<b>The impacts of enzyme promiscuity.....</b>	<b>150</b>
<b>5.3</b>	<b>Perspectives .....</b>	<b>152</b>
<b>6</b>	<b>References .....</b>	<b>155</b>

## List of Figures

Figure 2.1. Microbial synthesis of BIAs and BIA diversity .....	11
Figure 2.2. Synthesis of morphinan alkaloids and derivatives in <i>Saccharomyces cerevisiae</i> .....	25
Figure 2.3. Strategies for the reduction of side-products in pathways with promiscuous enzymes. .....	35
Figure 3.1. The dihydrosanguinarine pathway as introduced to yeast.....	43
Figure 3.2. Phylogenetic tree of NMTs and CYP719s .....	48
Figure 3.3. Screening of CYP719 library for protoberberine Ring A- and Ring B-closing activity .....	53
Figure 3.4. CYP719-catalyzed stylophine synthesis from scoulerine, cheilanthifoline or nandine .....	55
Figure 3.5. Combinatorial expression of selected CYP719s with and without TNMT co- expression .....	58
Figure 3.6. Activities of NMTs on select BIAs .....	60
Figure 3.7. Reconstruction of a complete dihydrosanguinarine pathway with different combinations of CYP719s .....	62
Figure 3.8. The pBOT vector system.....	71
Figure 3.9. Qualitative and comparative analysis of gene expression and enzyme activity.....	72
Figure 4.1. Synthesis of benzyloquinoline alkaloids and analogs in yeast .....	99
Figure 4.2. Effects of <i>ARO10</i> knockout in yeast synthesizing benzyloquinoline alkaloids.....	102
Figure 4.3. Promoter titration of <i>HsMAO-A</i> in an <i>ARO10</i> knockout background.....	106
Figure 4.4. Optimization of reticuline production in fed-batch fermentation.....	110
Figure 4.5. <i>De novo</i> dihydrosanguinarine synthesis in yeast.....	115
Figure 4.6. Unidentified metabolites in fermentation samples of strain LN1084 .....	119
Figure 4.7. Metabolite profile of strains expressing <i>HsMAO-A</i> from various promoters.....	129
Figure 4.8. Visualization of <i>HsMAO-A</i> localization by fluorescence microscopy .....	131
Figure 4.9. Assessment of <i>HsMAO-A</i> peroxisomal compartmentalization .....	132
Figure 4.10. Fed-batch fermentation of strain LP524.....	133
Figure 4.11. Detection of dihydrosanguinarine and sanguinarine in fermentation broth and supernatant .....	134

## List of Tables

Table 2.1. Diversity of BIAs in drug development.....	9
Table 2.2. Summary of <i>de novo</i> reticuline yields .....	14
Table 2.3. Synthesis of reticuline in microbial hosts.....	19
Table 2.4. Derivatization of BIA backbones in microbial hosts.....	28
Table 3.1. Characterized NMTs.....	46
Table 3.2. Characterized CYP719s.....	46
Table 3.3. Enzyme list for Chapter 3 .....	75
Table 3.4. Primer list for Chapter 3 .....	77
Table 3.5. Plasmid list for Chapter 3 .....	83
Table 3.6. Strain list for Chapter 3.....	88
Table 4.1. Strain list for Chapter 4.....	137
Table 4.2. Plasmid list for Chapter 4 .....	139
Table 4.3. Primer list for Chapter 4 .....	139
Table 4.4. Synthetic gene list for Chapter 4.....	144

## List of Abbreviations

3,4-dHPAA	3,4-Dihydroxyphenylacetaldehyde
3,4-dHPAC	3,4-Dihydroxyphenylacetic acid
4-HPAA	4-Hydroxyphenylacetaldehyde
4-HPAC	4-Hydroxyphenylacetic acid
4-HPP	4-Hydroxyphenylpyruvate
4'OMT	3'-Hydroxy- <i>N</i> -methylcoclaurine 4'- <i>O</i> -methyltransferase
6OMT	Norcoclaurine 6- <i>O</i> -methyltransferase
AA	Amino acid
AADC	Amino acid decarboxylase
AAS	Aromatic aldehyde synthase
AAT	Amino acid transferase
AcN	Acetonitrile
ADH	Alcohol dehydrogenase
ALD	Aldehyde dehydrogenase
BBE	Berberine bridge enzyme
BH <sub>4</sub>	Tetrahydrobiopterin
BIA	Benzylisoquinoline alkaloid
BLAST	Basic Local Alignment Search Tool
BRENDA	BRaunschweig ENzyme DAtabase
CFS	Cheilanthisfoline synthase
CNMT	Coclaurine <i>N</i> -methyltransferase
CODM	Codeine 6- <i>O</i> -demethylase
COR	Codeinone reductase
CPR	Cytochrome P450 reductase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CYP	Cytochrome P450
DBOX	Dihydrobenzophenanthridine oxidase
DEA	Drug Enforcement Agency
DODC	Dopamine decarboxylase
E4P	Erythrose 4-phosphate
ePTS	Enhanced peroxisomal targeting sequence
EV	Empty vector
FA	Formic acid
FT ICR	Fourier-transform ion cyclotron resonance
GFP	Green fluorescent protein
HIV	Human Immunodeficiency Virus
HPAH	Hydroxyphenylacetic acid hydroxylase
HPLC	High-pressure liquid chromatography

LC	Liquid chromatography
MAO	Monoamine oxidase
MS	Mass spectrometry
MeOH	Methanol
MSH	<i>N</i> -Methylstylopine hydroxylase
MT	Methyltransferase
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
NCS	Norcoclaurine synthase
NMCH	<i>N</i> -Methylcoclaurine hydroxylase
NMT	<i>N</i> -Methyltransferase
OMT	<i>O</i> -Methyltransferase
ORF	Open reading frame
P6H	Protopine 6-hydroxylase
PBS	Phosphate-buffered saline
PDV	Prodeoxyviolacein
PEP	Phosphoenolpyruvate
PPDC	Phenylpyruvate decarboxylase
RBS	Ribosome binding site
REP	Reticuline epimerase
RuBisCO	Ribulose biphosphate carboxylase/oxygenase
SAR	Salutaridine reductase
SAS	Salutaridine synthase
SAT	Salutaridinol 7- <i>O</i> -acetyltransferase
SC	Synthetic complete media
SPS	Stylopine synthase
T6ODM	Thebaine 6- <i>O</i> -demethylase
TFA	Trifluoroacetic acid
TH	Tyrosine hydroxylase
TNMT	Tetrahydroprotoberberine <i>N</i> -methyltransferase
TS	Thebaine synthase
TyDC	Tyrosine decarboxylase
TYR	Tyrosinase
UV	Ultraviolet
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose media
YTK	Yeast Toolkit

## Chapter One: Introduction

Benzyloquinoline alkaloids (BIAs) are a class of over 2,500 plant secondary metabolites with broad applications in human health and agriculture. Several BIAs are commercially available as pharmaceuticals, most notably the opiates including morphine and codeine, but also others including noscapine and glaucine (cough suppressants) and papaverine (vasodilator)<sup>1</sup>. Properties of naturally occurring BIAs have also informed chemical synthesis of BIA-like molecules. For example, the muscular relaxant properties of tubocurarine, used in poison darts, inspired the chemical synthesis of atracurium, which is routinely used in surgery during mechanical ventilation<sup>2</sup>. New uses are still being discovered for BIAs, such as the emerging potential of noscapine as an anticancer drug<sup>3</sup>, and the recent finding that papaverine can safely sensitize tumors prior to chemotherapy<sup>4</sup>. In agriculture, sanguinarine and chelerythrine are commercially available as probiotics for use in animal feed<sup>5</sup>.

Industrial-scale sources for BIAs are fairly limited. They tend to be difficult to synthesize chemically, due to the presence of multiple chiral centers, requirement for expensive precursors, and number of reactions<sup>6</sup>. Instead, most of the aforementioned compounds (morphine, codeine, and noscapine) are extracted from opium poppy<sup>7,8</sup>. Several others (glaucine, sanguinarine, chelerythrine) are abundant in other plants, but most are not. It is not a coincidence that the BIAs that are commercial products are a subset of those that accumulate to an appreciable degree in plants. An alternative route for BIAs would provide access to the rest of the BIA family for research and commercialization.

Microbial production is an attractive option for BIA synthesis. Microbes such as *Escherichia coli* and *Saccharomyces cerevisiae* are genetically tractable, grow more rapidly than plants, and can be scaled-up to grow in large (20,000-2,000,000L) fermenters. A standout

commercial success in *E. coli* includes the production of 1,4-butanediol (Genomatica). Yeast, especially, shines as a host for plant natural product synthesis, with commercial platforms for stevia (Cargill, Amyris), Vitamin E (DSM), and beta-farnesene (Amyris) among others<sup>9</sup>. It has been estimated that a titer of 5 g/L opioids would be sufficient for commercial viability in yeast<sup>10</sup>.

There is a well-established history of BIA synthesis in microbes. In 2008, it was demonstrated that enzymes involved with BIA synthesis were functional when expressed in both *E. coli*<sup>11</sup> and in yeast<sup>12</sup>. Both groups chose to supplement the commercially available norcoclaurine analog norlaudanoline, and both groups chose to convert it into reticuline, which is the last common intermediate between morphine, noscapine, and sanguinarine synthesis. *De novo* synthesis of BIAs from sugar was achieved in *E. coli* in 2011<sup>13</sup> and in yeast in 2015<sup>14</sup>. Since then, complete synthesis of the opiate hydrocodone from sugar has been achieved in both *E. coli*<sup>15</sup> and yeast<sup>10</sup>. Yeast was also used for *de novo* synthesis of noscapine<sup>7</sup>. Recently, BIA titers approaching those of commercial viability (4.6 g/L reticuline) were reached in yeast<sup>16</sup>.

This thesis is organized around synthesis of the BIA dihydrosanguinarine. Dihydrosanguinarine is the reduced, less toxic form of the quaternary ammonium antibiotic sanguinarine<sup>17</sup>. Sanguinarine biosynthesis has been suggested as a general model for BIA synthesis *in planta* due to the high degree of characterization, down to subcellular localization, of all enzymes involved<sup>18</sup>. Further, the sanguinarine biosynthetic pathway also serves as a model for the challenges inherent in reconstitution of any BIA synthesis in a microbial host, including long pathway length, a number of membrane-bound cytochromes P450, and numerous accounts of enzyme promiscuity<sup>19–22</sup>.

This work builds upon a foundational publication from our group from 2014 in which the dihydrosanguinarine pathway was first reconstituted in yeast, yielding 34 µg/L

dihydrosanguinarine from a supplemented precursor<sup>23</sup>. The ultimate result of this thesis is the *de novo* synthesis of 630 mg/L dihydrosanguinarine and sanguinarine in fed-batch fermentation from sucrose, representing a 20,000-fold increase in titers. This work also represents a >300-fold increase in the highest-reported titers of a BIA end-product, which is a major milestone in the quest to make the entire BIA family commercially accessible.

## 1.1 Research objectives

In nature, BIAs are derived from the common scaffold (*S*)-norcoclaurine, which is then diversified through a network of enzyme-catalyzed modifications and rearrangements into the entire family. With the appropriate downstream pathway, a chassis strain making (*S*)-norcoclaurine can be used to make any number of BIAs. BIA synthesis from simple sugars may thus be conceptualized as two complementary challenges: efficient synthesis of the common scaffold, and efficient synthesis of a target BIA from the scaffold. Each of these challenges is addressed as a research objective in this thesis, with the target BIA of choice being dihydrosanguinarine.

### *1. Optimize a BIA synthesis pathway in yeast through in vitro supplementation of precursors.*

This objective expands upon a 2014 study in which the dihydrosanguinarine pathway was first introduced into yeast. The pathway was functional but inefficient, resulting in just 1.5% conversion of supplemented norlaudanosoline to dihydrosanguinarine. In this work, conversion is improved to 10%.

### *2. Improve synthesis of BIAs from simple sugar in yeast.*

In this objective, BIA synthesis is improved in a strain of yeast making reticuline. We identified that synthesis of the common scaffold norcoclaurine was also accompanied by the synthesis of unwanted analogous condensation products. These side products were significantly

reduced by re-routing BIA synthesis through the alternative scaffold norlaudanosoline. The result was a 54% improvement of reticuline yield from sugar (17.5 to 27 mg reticuline/g sucrose).

### 3. *Demonstrate the de novo synthesis of dihydrosanguinarine in yeast.*

The optimized dihydrosanguinarine pathway from Objective 1 is introduced to the improved reticuline synthesis strain from Objective 2. New bottlenecks are identified and partially resolved, with suggestions for future work. The end result is a strain of yeast that can produce dihydrosanguinarine and sanguinarine from simple sugar at a titer of 630 mg/L in fed-batch fermentation.

## 1.2 Thesis organization

Spontaneous condensation of norcoclaurine results in mix of (*R*)- and (*S*)-enantiomers. Norcoclaurine synthase, however, catalyzes the enantiospecific condensation of (*S*)-norcoclaurine<sup>1</sup>. Throughout this thesis, when this chiral centre is present, the (*S*)-enantiomer is referred to unless otherwise specified.

This thesis comprises three manuscripts, organized into three chapters.

Chapter 2 summarizes the area of microbial BIA synthesis. This chapter includes a review published in 2016 (Narcross *et al.* 2016a, reference 1) highlighting the challenges associated with establishing *de novo* BIA synthesis in both *E. coli* and *S. cerevisiae*. The successes of heterologous BIA pathway reconstruction are also detailed. Additional text has been incorporated to reflect relevant works that have been published since 2016.

Chapter 3 (Narcross *et al.* 2016b, reference 193) describes the optimization of the dihydrosanguinarine pathway through combinatorial co-expression of pathway enzymes. A bottleneck was identified between the pathway intermediates scoulerine and stylophine. After an

extensive screen of alternative enzyme homologs, several enzymes were found that improved flux through the entire pathway to 10%.

Chapter 4 (Narcross *et al.* 2021, manuscript in preparation) details the development of a strain of yeast synthesizing reticuline *via* the norcoclaurine analog norlaudanosoline. Compared to the norcoclaurine-based approach, the norlaudanosoline route to reticuline in yeast results in a slightly improved titer while using a simpler cultivation protocol and 35% less sugar. Finally, reticuline is converted into dihydrosanguinarine, using the optimized pathway described in Chapter 3, resulting in 630 mg/L dihydrosanguinarine from sucrose. This represents the highest titers of any BIA final product by a factor of >300.

Finally, Chapter 5 summarizes the major findings, discusses the synthesis of Chapters 3 and 4 into a single dihydrosanguinarine-synthesizing strain of yeast, and provides suggestions for future work in this area.

## 2 Chapter Two: Literature Review

**Adapted from:** Narcross, L., Fossati, E., Bourgeois, L., Dueber, J.E., and Martin, V.J.J. (2016). Microbial Factories for the Production of Benzyloquinoline Alkaloids. *Trends in Biotechnology*, **34**, 3, 228-241.

### 2.1 Trends

- Both *Escherichia coli* and *Saccharomyces cerevisiae* have been engineered to convert a simple carbon source such as glucose to complex BIAs.
- The variety of BIA scaffolds synthesized in microbial hosts continues to increase, now encompassing benzyloquinolines, aporphines, protoberberines, protopines, benzophenanthridines, pro-morphinans, and morphinans.
- Key challenges for future work have been identified, including pathway bottlenecks and the generation of side-products from promiscuous enzymes.

### 2.2 Abstract

Benzyloquinoline alkaloids (BIAs) are a family of ~2500 alkaloids with both potential and realized pharmaceutical value, including most notably the opiates such as codeine and morphine. Only a few BIAs accumulate readily in plants, which limits the pharmaceutical potential of the family. Shifting BIA production to microbial sources could provide a scalable and flexible source of these compounds in the future. This review details the current status of microbial BIA synthesis and derivatization, including rapid developments in 2015 culminating in the synthesis of opioids from glucose in a microbial host.

### **2.3 Microbial synthesis of pharmaceuticals to enhance drug discovery**

Plant secondary metabolites are a valuable source of natural products with pharmaceutical properties<sup>24,25</sup>. Until recently pharmaceutical companies had phased out screening of natural products as potential drug leads in favor of chemically synthesized libraries, citing “poor yields of chemical synthesis” and “impracticality of scale-up”<sup>26,27</sup>. While these challenges remain, there has been a resurgence of natural products in the drug discovery process owing to their potent biological activity and untapped potential, and recent improvements in screening technology<sup>27,28</sup>.

Biosynthesis of plant metabolites in microbes provides an opportunity to advance the drug discovery process. Microbial production circumvents the need for the cultivation of source plants<sup>29</sup>. Not all products accumulate to high levels in plants, and hence microbes provide a way to scale up production of these interesting compounds to relieve supply limitations<sup>30</sup>. The successful industrial production of the antimalarial artemisinin acid in yeast highlights the capability of microbes to act as natural product factories<sup>31</sup>. Finally, microbial production facilitates the possibility to design new compounds with novel activity or improve the clinical profile of existing drugs through combinatorial chemistry<sup>32</sup>.

### **2.4 Pharmaceutical properties of benzyloquinoline alkaloids**

Benzyloquinoline alkaloids (BIAs) are a class of molecules that would benefit from microbial synthesis. With over 2500 family members, BIAs exhibit diverse pharmaceutical properties, with a history of human use dating back thousands of years<sup>33</sup>. In addition to their prominent role in traditional medicine, BIAs have a wide variety of pharmacological applications, acting as analgesics, antitussives, antimicrobials, and antispasmodics, and several members are on the World Health Organization list of essential medicines (Table 2.1). Recently, preliminary studies have uncovered new potential in treating cancer, malaria, HIV, and psychosis (Table 2.1).

Despite their clinical applications, it is argued that BIAs (among other alkaloids) are not proportionately represented in modern medicine and drug development, in large part owing to difficulties in supply<sup>34</sup>. To overcome this barrier, several groups have been working to synthesize and derivatize BIAs in the microbial hosts *Escherichia coli* and *Saccharomyces cerevisiae*.

## 2.5 BIA diversity is derived from a single scaffold

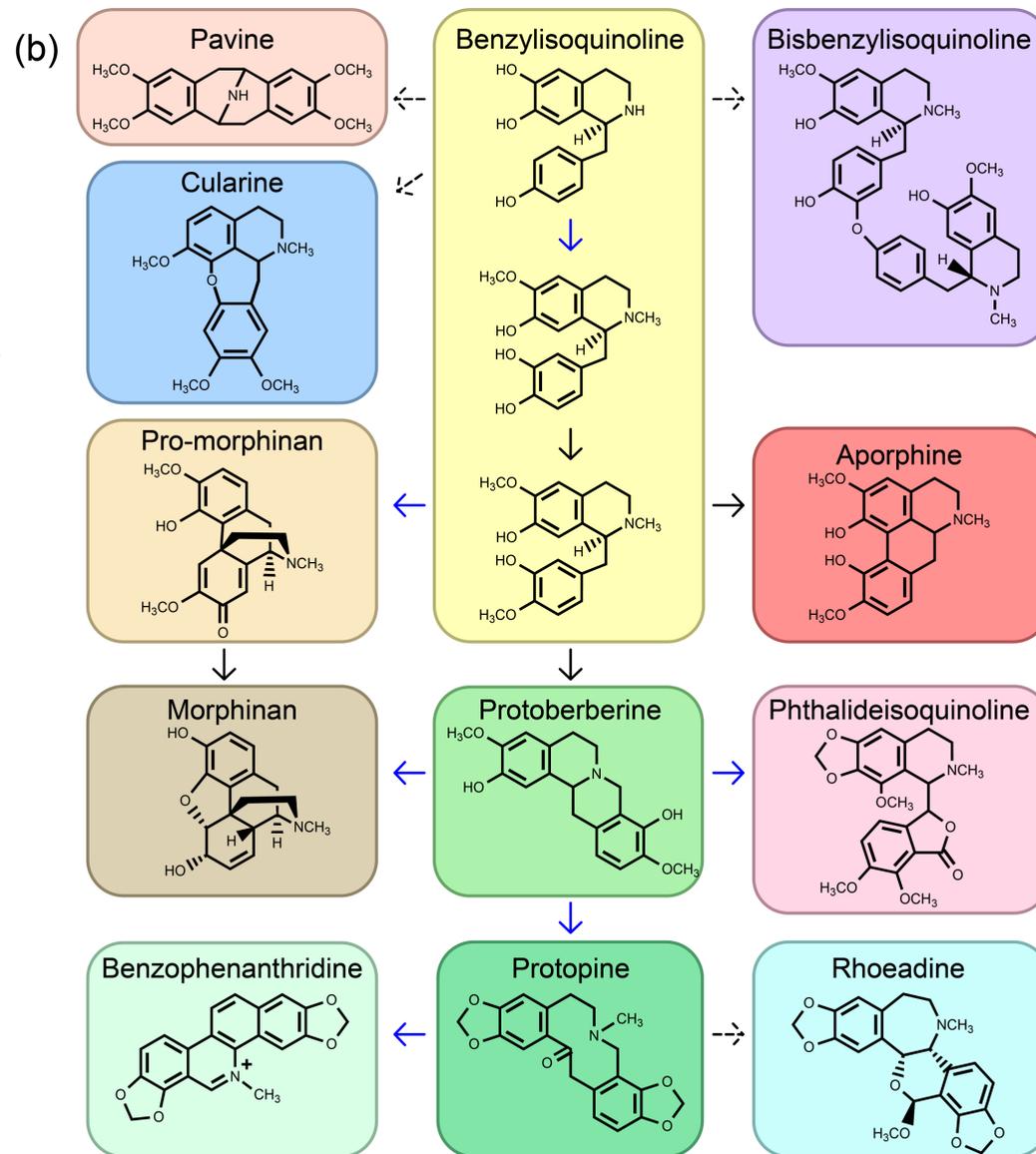
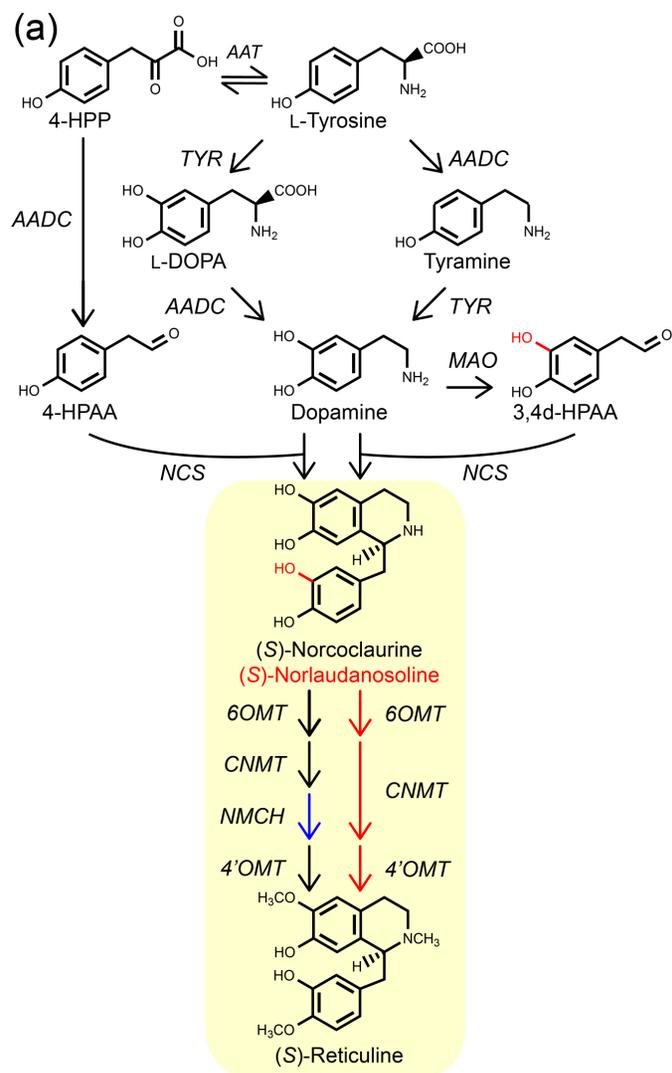
BIA synthesis in plants begins with the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to form (*S*)-norcoclaurine (Figure 2.1A). (*S*)-norcoclaurine is the scaffold from which over 2500 other BIAs can be produced (examples of the structural diversity of BIAs are provided in Figure 2.1B)<sup>35</sup>. This means that a microbial platform strain capable of producing the BIA scaffold from simple carbon sources could then be engineered to produce any BIA of interest, provided the necessary enzymes have been elucidated.

Ten enzyme families add functional groups and catalyze the rearrangement of BIA scaffolds<sup>33,36</sup>. Many of these enzymes are cytochromes P450 that require endomembranes for optimal activity (Figure 2.1B, blue arrows). For this reason, *S. cerevisiae* has been used extensively for heterologous BIA pathway development, although ample successes have been achieved in *E. coli*. Enzymes in BIA synthesis tend to have broad substrate ranges, allowing the same reactions to be performed even as BIAs diverge in structure (see<sup>37-41</sup> for examples). Although BIA synthesis pathways are commonly drawn in straight lines, the substrate acceptance profiles of enzymes in a single pathway can overlap, meaning that these pathways are often more of a web resulting in a multitude of products (an example is the morphine pathway, shown in Figure 2.2). Metabolic engineering strategies will be necessary to overcome promiscuity for the reconstitution of BIA synthesis pathways in heterologous hosts (see Section 2.13).

**Table 2.1. Diversity of BIAs in drug development**

<b>BIA Subfamily</b>	<b>Compound</b>	<b>Pharmaceutical Applications</b>	<b>Clinical Status</b>	<b>Ref</b>
Benzylisoquinoline	Norcoclaurine	Cardiac Stimulant	Clinical trials (Phase 1)	42–44
	Drotaverine	Antispasmodic	Approved	42,45
	Papaverine	Vasodilator	Approved	42,46,47
Smooth muscle relaxant		Approved		
Bisbenzylisoquinoline	Atracurium <sup>a</sup>	Neuromuscular blocker	Approved	45,48
	Mivacurium <sup>a</sup>	Neuromuscular blocker	Approved	49,50
Muscle relaxant		Approved		
Phthalidisoquinoline	Noscapine	Antitussive	Approved	47,51 drugs.com/international/noscapine.html
		Potential Anticancer	Clinical trials (Phase 2)	
		Antimalarial	N/A	
Aporphine	Glaucine	Antitussive	Approved	42,52 drugs.com/international/glaucine.html
		Anticancer	N/A	
Protoberberine	Berberine	HIV treatment	N/A	42,45,47,53–55
		Antibacterial	N/A	
		Antiparasitic	Experimental	
		Antifungal	Experimental	
		Antidiarrheal	Experimental	
	Type II diabetes	Clinical trials (Phase 3)		
	Stepholidine	Psychosis	N/A	56
Benzophenanthridine	Sanguinarine	Anti-microbial	Not approved	42,49
Morphinan	Codeine	Analgesic	Approved	42,45,47
		Antitussive	Approved	
	Morphine	Analgesic	Approved	42,45,47
Morphinan semisynthetic derivatives	Oxycodone <sup>a</sup>	Analgesic	Approved	42,45,57
	Naloxone <sup>a</sup>	Opioid antagonist	Approved	
	Naltrexone <sup>a</sup>	Opioid antagonist	Approved	

<sup>a</sup>Synthetic or semi-synthetic BIAs



## Figure 2.1. Microbial synthesis of BIAs and BIA diversity

(A) 4-HPP is the microbial precursor to the substrates of norcoclaurine/norlaudanosoline condensation. Dopamine and 4-HPAA condense to form norcoclaurine (black), while dopamine and 3,4-dHPAA condense to form norlaudanosoline (red). The extra hydroxyl group on 3,4-dHPAA and norlaudanosoline is indicated in red, and reactions that accept norlaudanosoline are indicated with red arrows. Metabolite abbreviations: 4-HPP, 4-hydroxyphenylpyruvate; 4-HPAA, 4-hydroxyphenylacetaldehyde; 3,4-dHPAA, 3,4-dihydroxyphenylacetaldehyde. Enzyme abbreviations: AADC, amino acid decarboxylase; AAT, amino acid transferase; CNMT, coclaurine *N*-methyltransferase; NCS, norcoclaurine synthase; NMCH, *N*-methylcoclaurine 3'-*O*-hydroxylase; 4'OMT, 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase; 6OMT, norcoclaurine 6'-*O*-methyltransferase; TYR, any enzyme which hydroxylates tyrosine. (B) Diversity of BIA scaffolds. Featured BIAs are examples of indicated scaffold types. The number of arrows is not indicative of pathway length. Dashed lines indicate unknown pathways; blue lines indicate one or more cytochrome P450-catalyzed reactions.

## 2.6 Current status of microbial aromatic amino acid (AA) production

The (*S*)-norcoclaurine precursors 4-HPAA and dopamine are derived from the aromatic AA pathway, and hence aromatic AA overproduction is a key goal for the development of microbial sources of BIAs. Entry of carbon into the aromatic AA pathway begins with the condensation of the glycolysis intermediate phosphoenolpyruvate and the pentose phosphate pathway intermediate erythrose-4-phosphate. This committed step is transcriptionally and allosterically regulated, which is a common theme throughout the aromatic AA pathway, especially for enzymes at key metabolic branch-points<sup>58</sup>. Strategies for improving flux towards tyrosine in *E. coli* and *S. cerevisiae* are similar, but the highest published yields of aromatic AAs or aromatic AA-derived compounds are currently an order of magnitude higher in *E. coli* than in yeast, both in shake-flask (~ 300 mg/L vs 2 g/L)<sup>59-61</sup> and fermentation (2 g/L vs 55 g/L) conditions<sup>62,63</sup>. Notably, the shift from laboratory shake-flask to industrial fermentation can improve total titers, although not necessarily yield, by one to two orders of magnitude. A striking example of this is the development of fermentation conditions for a yeast strain to synthesize artemisinic acid, in which titers of the intermediate amorpha-4,11-diene increased from 160 mg/L to 40 g/L<sup>64</sup>. Although BIA titers reported throughout this review are low, much of the work was done in shake flasks, suggesting that optimized fermentation conditions combined with the latest developments in strain engineering will easily improve upon these titers.

### 2.6.1 Update in 2021

Since 2016, published titers of aromatic AAs or aromatic AA-derived compounds have increased in both *E. coli* and yeast. In yeast, especially, an increased focus on the use of bioreactors has led to a greater number of papers reporting gram-scale production of such compounds<sup>65</sup>. Currently, the highest reported titer for aromatic AA production in *E. coli* is 72.9 g/L L-

phenylalanine in fed-batch fermentation<sup>66</sup>. The work combined metabolomics, transcriptomics, and proteomics analysis to accurately identify enzyme bottlenecks and precisely target genes for overexpression. The authors comment that this approach is likely to have applications beyond phenylalanine and beyond *E. coli*. Yeast, too, has seen a boost in maximum titers, with the phenylalanine-derived coumaric acid produced in bioreactor at 13 g/L<sup>67</sup>. Among other innovations, this elegant paper introduced a novel method for balancing carbon flux between the glycolytic and pentose phosphate pathways, which has already proven successful in other projects<sup>68</sup>.

## 2.7 Aldehyde scavenging limits the available 4-HPAA pool

Because the aldehyde 4-HPAA is a substrate of norcochlorine synthesis, its *in-vivo* stability is necessary for BIA overproduction in microbes. However, aldehydes are rapidly scavenged to limit cellular toxicity<sup>69</sup>. Aldehydes can be reduced or oxidized depending on the redox status of the cell<sup>70</sup>. No less than six aldehyde dehydrogenases (ALDs) and 16 alcohol dehydrogenases (ADHs) can participate in aldehyde reduction/oxidation in yeast, with significant redundancy<sup>70</sup>. *E. coli* also harbors multiple ALDs and ADHs, whose knockout has improved heterologous production of aldehydes<sup>71,72</sup>. Preliminary data suggests that ALD and ADH knockouts may improve *de novo* BIA production in yeast<sup>73</sup>.

### 2.7.1 Update in 2021

Oxidoreductase knockouts do indeed improve *de novo* BIA production in yeast. An extremely thorough combinatorial knockout approach was undertaken in yeast to reduce the number of oxidoreductases acting on 4-HPAA in order to improve BIA titers<sup>16</sup>. Seven functionally redundant oxidoreductases, most of which had no previously identified activity on aromatic aldehydes, were identified and subsequently knocked out. In combination with engineering to improve flux through aromatic AA synthesis, this yeast strain was capable of producing 4.6 g/L

reticuline from sucrose in fed-batch fermentation (Table 2.2)<sup>16</sup>. This represents a 30-fold improvement over the previous highest reticuline titers in *E. coli*, and a 2,300-fold improvement over the previous highest BIA titers demonstrated in yeast.

**Table 2.2. Summary of *de novo* reticuline yields**

Year	Organism	Reticuline (mg/L)	Media	Growth Conditions	Fermentation Protocol	Yield (mg/g sugar)	Ref.
2011	<i>E. coli</i>	40.5	Complex (Turbo)	Bioreactor	Manual sugar additions	0.810	13
2018	<i>E. coli</i>	160	Complex (Terrific)	Bioreactor	Batch	5.33	74
2015	<i>S. cerevisiae</i>	0.081	2x Synthetic complete	Shake flask	n/a	0.00202	14
2016	<i>S. cerevisiae</i>	0.019	Synthetic complete (-tyrosine) 2 mM ascorbic acid	96-well plate	n/a	0.000950	75
2018	<i>S. cerevisiae</i>	3	Synthetic complete 10 mM ascorbic acid	Shake flask	n/a	0.15	7
2020	<i>S. cerevisiae</i>	4580	Minimal media	Bioreactor	Feedback-controlled pulsed fed-batch	17.5	16
2021	<i>S. cerevisiae</i>	4810	Minimal media	Bioreactor	Exponential fed-batch	27.0	Chp. 4

Chp.4 refers to Chapter 4 of this work

## 2.8 Multiple strategies for dopamine synthesis in microbial hosts

The formation of dopamine from tyrosine requires one hydroxylation and one decarboxylation event (Figure 2.1A). Depending on enzyme specificity, these reactions could occur in either order; when decarboxylation occurs first the intermediate is tyramine, while if hydroxylation occurs first the intermediate is L-DOPA. Thus far, the decarboxylation-first pathway has been avoided through the use of a decarboxylase that has a strong preference for L-

DOPA<sup>13</sup>. This is historically because production of L-DOPA, as opposed to tyramine, made downstream pathway engineering easier in *E. coli*<sup>13</sup>. L-DOPA has remained the intermediate of choice as *de novo* BIA synthesis has been introduced to yeast.

Many options for heterologous L-DOPA synthesis exist, with no one enzyme standing out as being clearly superior to other enzymes. Currently, enzyme selection for L-DOPA synthesis requires choosing between undesired side-activities and the requirement for a cofactor. There are two types of side-activities relating to L-DOPA synthesis: broad substrate range and L-DOPA oxidation to dopaquinone (diphenolase activity). Tyrosinases (TYR) and hydroxyphenylacetic acid hydroxylases (HPAH) have broad substrate ranges<sup>76,77</sup>, while TYRs and the cytochrome P450 hydroxylase CYP76AD1 have diphenolase activity on L-DOPA<sup>78,79</sup>. While catalyzing two types of side activity, TYR has the lowest cofactor requirement (some require only inorganic copper). By contrast, tyrosine hydroxylase (TH) has the lowest amount of side-activity but the highest cofactor requirements because it uses the cofactor tetrahydrobiopterin (BH<sub>4</sub>). BH<sub>4</sub> is not native to *E. coli* and *S. cerevisiae*, and hence the functional expression of TH also requires the heterologous expression of a BH<sub>4</sub> synthesis and regeneration pathway. THs have an additional disadvantage in that they are heavily regulated by allosteric inhibition and post-translational modification reflecting their role as the rate-limiting step of catecholamine synthesis in neurons<sup>80</sup>.

The disadvantages of various enzyme families have been addressed during the introduction of L-DOPA synthesis in both *E. coli* and *S. cerevisiae*. A TYR with an unusually low level of diphenolase activity (*R*sTYR) was expressed in *E. coli* for heterologous BIA synthesis, resulting in the production of 2.5 g/L dopamine<sup>81</sup>. However, *R*sTYR still possessed some diphenolase activity and its broad substrate range allowed it to oxidize downstream BIAs. This has since been addressed by using multiple strains of *E. coli* to sequester *R*sTYR from downstream BIA synthesis

enzymes<sup>81</sup>. However, this multi-strain system is not without drawbacks (see Section 2.14). While HPAH also has a broad substrate range, it does not oxidize its products, and hence HPAH may be a better option than *RsTYR* for L-DOPA synthesis in *E. coli*.

HPAH has not been demonstrated to be functional in yeast, and TYR activity in yeast is low<sup>14</sup>, pointing towards TH and CYP76AD1 as better options for L-DOPA production in *S. cerevisiae*. The side-activity of CYP76AD1 was reduced by subjecting it to mutagenesis followed by screening with a color-based biosensor that can detect and distinguish between L-DOPA and dopaquinone synthesis<sup>14</sup>. Alternatively, the BH<sub>4</sub> synthesis pathway has now been introduced into *S. cerevisiae*<sup>75,82</sup>, permitting functional TH expression for BIA synthesis<sup>75</sup>. The use of either enzyme to produce dopamine, regardless of strain engineering strategy, currently results in yields of ~10 - 25 mg/L. These titers are lower than the 2.5 g/L dopamine produced in *E. coli*, but are not uncommon for heterologous products derived from aromatic AAs in yeast<sup>61,83</sup>, which points to the necessity to improve the synthesis of precursors to achieve higher dopamine levels. Recent titers of 2 and 3 g/L of aromatic AA pathway derivatives in yeast indicate that engineering strategies are being developed that should further improve BIA yields<sup>63,84</sup>.

#### 2.8.1.1 Update in 2021

The original strategy for norlaudanosoline synthesis in *E. coli* required the pathway to be split across multiple strains to prevent the enzyme used for L-DOPA synthesis, *RsTYR*, from oxidizing norlaudanosoline<sup>81</sup>. In 2018, *RsTYR* was swapped for tyrosine hydroxylase from *Drosophila melanogaster* (*DmTH*), which does not act on norlaudanosoline, thus allowing the entire pathway to be expressed in a single host strain<sup>74</sup>. Using *DmTH* also required the introduction of a three-gene biosynthetic pathway for the non-native cofactor BH<sub>4</sub>. The *DmTH*-driven approach enabled the synthesis of 160 mg/L reticuline, a 4-fold improvement over previous titers (Table

2.2). The authors noted that a further 1 g/L dopamine accumulated during the experiment, indicating that *DmTH* was not the bottleneck in this pathway.

In contrast to *E. coli*, tyrosine hydroxylase was determined to be a bottleneck in the TH-enabled synthesis of norcoclaurine in yeast<sup>7</sup>. The Smolke group swapped the wild-type tyrosine hydroxylase from *Rattus norvegicus* (*RnTH*) with a feedback-resistant mutant, which improved yields ~2.5-fold. In total, 3 mg/L *de novo* reticuline synthesis could be achieved (Table 2.2). It is not reported whether feedback resistant *RnTH* continues to be a bottleneck in this pathway. However, since this strain lacks any aldehyde-scavenging knockouts, and even contains an aldehyde dehydrogenase overexpression, it is likely that titers are currently limited due to insufficient aldehyde.

An alternative strategy for L-DOPA synthesis in yeast was to use the plant cytochrome P450 CYP76AD1, found in beets<sup>79</sup>. This enzyme natively oxidizes tyrosine to L-DOPA and L-DOPA to dopaquinone. While a mutagenesis-based approach successfully reduced the over-oxidation by ~90%, the enzyme still retained some diphenolase activity<sup>14</sup>. In 2016, another tyrosine hydroxylating cytochrome P450 from beets was identified, CYP76AD5, which has no native dopaquinone-producing activity<sup>85</sup>. The Martin group swapped CYP76AD1 with CYP76AD5 and saw a 2-fold improvement in titer, ultimately achieving 4.6 g/L reticuline synthesis<sup>16</sup>. The CYP76AD5 approach stands out as a superior means for dopamine synthesis in yeast.

## 2.9 BIA scaffolds can be synthesized *de novo* from simple sugars

Pictet–Spengler condensation of an amine and aldehyde is a reaction mechanism common to the committed step of several alkaloid families<sup>86</sup>. Spontaneous condensation generates racemic (*R,S*) mixtures, whereas enzymatic condensation is enantio-specific. The committed step of BIA synthesis is enzymatically catalyzed by norcoclaurine synthase (NCS), which condenses dopamine

and 4-HPAA to generate (*S*)-norcoclaurine. Because BIAs are derived from (*S*)-norcoclaurine, the synthesis of (*R*)-norcoclaurine is unproductive, and hence enzymatic condensation is preferable to spontaneous condensation for microbial production of BIA scaffolds. NCS has a broad substrate range for aldehydes. In addition to 4-HPAA, NCS can accept the double-hydroxylated 3,4-dHPAA, which when condensed with dopamine generates (*S*)-norlaudanosoline (Figure 2.1A). Norlaudanosoline has been used extensively for BIA derivatization in *E. coli* because the extra hydroxyl group on 3,4-dHPAA negates the need for later cytochrome P450-catalyzed hydroxylation of the BIA scaffold (Figure 2.1A)<sup>87</sup>. Both norcoclaurine and norlaudanosoline are unstable end-products because they are subject to enzymatic oxidation as well as spontaneous oxidation at alkaline pHs. Therefore, the key branch-point intermediate reticuline (Figure 2.1B), derived from norcoclaurine/norlaudanosoline (Figure 2.1A), is frequently used as a readout for *de novo* synthesis of norcoclaurine/norlaudanosoline.

### **2.9.1 Norlaudanosoline synthesis in *E. coli***

3,4-dHPAA can be generated from dopamine via monoamine oxidase (MAO), making dopamine the source of both amine and aldehyde for norlaudanosoline synthesis (Figure 2.1A). As a first proof of concept for the synthesis of BIAs in a microbial host, reticuline was produced from supplemented dopamine in *E. coli*<sup>11</sup>. Initially, 1.3% of supplemented dopamine was converted to reticuline, which was improved in later studies after optimizing both fermentation conditions and gene copy number (Section 2.14, Table 2.3)<sup>88</sup>. The introduction of endogenous dopamine synthesis to *E. coli* enabled synthesis of reticuline from glycerol, with an initial dopamine-to-reticuline conversion of 4% (Table 2.3)<sup>13</sup>.

Enzymatic oxidation of norlaudanosoline was identified as a side-activity of TYR that reduced conversion of dopamine to reticuline. To prevent TYR activity on norlaudanosoline,

norlaudanosoline synthesis was divided between two strains of *E. coli* (glycerol-to-dopamine and dopamine-to-norlaudanosoline), which were cultured sequentially (see Section 2.13). Preventing diphenolase activity on norlaudanosoline improved its accumulation 300-fold, resulting in 16% conversion from dopamine. However, the addition of a third strain to convert norlaudanosoline to reticuline did not improve total dopamine-to-reticuline yields compared to previous results (Table 2.3). It is possible that this discrepancy is due to the extended time that norlaudanosoline was exposed to culture supernatant, which promotes spontaneous oxidation (see Section 2.14). Norlaudanosoline oxidation in supernatant may be avoided in the future by re-engineering the multi-strain system such that dopamine produced by the first strain is converted directly to reticuline by a second strain expressing MAO as well as the reticuline synthesis pathway.

**Table 2.3. Synthesis of reticuline in microbial hosts**

Strain	Dopamine	Norlaudanosoline	Reticuline	% Yield <sup>a</sup>	Comments	Ref
<i>E. coli</i>	Supplement, 5 mM	-	33 $\mu$ M	1.3	First synthesis of reticuline in a microbial host	11
<i>E. coli</i>	Supplement, 3 mM	-	165 $\mu$ M	11	Improved fermentation conditions	88
<i>E. coli</i>	<i>De novo</i> , 7 mM <sup>b</sup>	-	140 $\mu$ M	4	First <i>de novo</i> BIA synthesis. Fed-batch.	13
<i>E. coli</i>	<i>De novo</i> , 14 mM <sup>b</sup>	-	145 $\mu$ M	2	Three-step fermentation. Fed-batch.	81
Yeast	-	Supplement, 4 mM	455 $\mu$ M <sup>c</sup>	10	First derivatization of norlaudanosoline in yeast	12
Yeast	-	Supplement, 0.01 mM	2 $\mu$ M	20	Improved yield	23
Yeast	<i>De novo</i> , 0.155 mM <sup>b</sup>	-	0.2 $\mu$ M	0.13	First <i>de novo</i> BIA synthesis in yeast. Shake-flask.	14
Yeast	<i>De novo</i> , 0.065 mM	-	0.2 $\mu$ M	0.31	Improved yield. Shake-flask.	10

<sup>a</sup>Molar yield.

<sup>b</sup>Dopamine quantified in supernatant of strains expressing no downstream enzymes.

<sup>c</sup>Reticuline concentration estimated based off of closest available standard.

### 2.9.1.1 Update in 2021

In 2018, BIA synthesis in *E. coli* was migrated from a step-wise multi-strain approach to an integrated system, enabled by the use of an alternative route to tyrosine hydroxylation (described in Section 2.8.1.1)<sup>74</sup>. Consequently, reticuline titers from glycerol improved 4-fold to 160 mg/L. Reticuline synthesis was accompanied by a further 1 g/L dopamine. The authors note that synthesis of 3,4-dHPAA from dopamine by monoamine oxidase from *Micrococcus luteus* (*MI*MAO) was the bottleneck in their BIA synthesis strategy and suggest that an alternative enzyme may be a better path forward. Aromatic aldehyde synthases (AAS) are capable of catalyzing two reactions: first, an amino acid is decarboxylated to form an amine; next, the amine is de-aminated to form an aldehyde. Thus, tyrosine may be converted through tyramine into 4-HPAA, and L-DOPA into 3,4-dHPAA.

In 2019, the Kondo group tested the AAS from *Bombyx mori* (*BmAAS*) in BIA synthesis<sup>89</sup>. While AASs typically release only aldehydes as final products, a mutagenesis approach was developed to release a portion of the reaction intermediate (dopamine). This allows L-DOPA to be the branch-point between dopamine and aldehyde synthesis, and mutant *BmAAS* (*BmAASmut*) to be the sole enzyme responsible for both products. Using *BmAASmut*, the group was capable of synthesizing 10  $\mu$ M norlaudanoline and 1  $\mu$ M reticuline from 1 mM supplemented L-DOPA (Table 2.2).

In a follow-up work, researchers revealed that two opium poppy enzymes, tyrosine decarboxylase (*PsTyDC*) and phenylpyruvate decarboxylase (*PsPPDC*), also have previously unidentified AAS activity on tyrosine and L-DOPA, respectively<sup>90</sup>. In addition, the mutagenesis previously performed on *BmAASmut* was used to inform the mutagenesis of the dopamine decarboxylase from *Pseudomonas putida* (*PpDDC*) to develop it into an AAS-capable enzyme.

With co-expression of *PsPPDC* and *PpDDCmut*, researchers were able to synthesize 70  $\mu\text{M}$  reticuline from 4.5 mM supplemented tyrosine and 2.5 mM supplemented L-DOPA (Table 2.2). While intriguing, AAS-directed aldehyde synthesis currently lags behind that of the *M/MAO*-based approach.

### 2.9.2 Norcoclaurine synthesis in *S. cerevisiae*

In 2015, two groups achieved BIA synthesis from simple carbon sources in yeast<sup>10,14</sup>. Neither group used MAO to generate 3,4-dHPAA from dopamine. Instead, endogenous cytosolic 4-HPAA was the source of the aldehyde. Both groups reported comparable reticuline yields several orders of magnitude lower than concurrent yields published for *E. coli* (Table 2.3). In particular, the drop in titers from dopamine (~10–25 mg/l) to norcoclaurine (~80–100  $\mu\text{g/l}$ ) indicates that norcoclaurine synthase is a key bottleneck in the *de novo* synthesis of BIAs in yeast<sup>14</sup>.

Supplementation of dopamine for spontaneous condensation to norcoclaurine in yeast was much less efficient than in *E. coli* (0.0025% vs 16%)<sup>75</sup>. Endogenously produced dopamine was also converted to reticuline at lower levels in yeast than in *E. coli* (0.3% vs 4%) (Table 2.3). Considering that endogenously produced dopamine is found in the supernatant<sup>14</sup>, low norcoclaurine yields could in part be due to dopamine secretion occurring more readily than norcoclaurine synthesis can occur. The  $K_M$  for dopamine for the NCS of *Thalictrum flavum* is 25 mM<sup>91</sup>, which is approximately double the concentration of dopamine produced in *E. coli*, but 170-fold of that currently produced in yeast. It is possible that dopamine values in yeast are currently too low for efficient synthesis of norcoclaurine. Local dopamine concentrations could be improved further via strain engineering to improve titers, reduce efflux, or sequester dopamine in subcellular compartments.

Currently, as a percentage of dopamine, reticuline yields from groups working in *S. cerevisiae* are similar to the first numbers published in *E. coli* (Table 2.3). Increases in *E. coli* yields were incremental, and required optimization of strain design and fermentation conditions<sup>11,13,81,88</sup>. Fermentation conditions developed for BIA derivatization in yeast will likely improve *de novo* synthesis as well (see Section 2.14)<sup>92</sup>. Further strain engineering will be necessary for competitive titers of *de novo* BIA synthesis in yeast.

#### 2.9.2.1 Update in 2021

Since 2016, two publications have expanded upon the titers of BIAs produced in yeast. In 2018, the Smolke group reported the production of 3 mg/L reticuline in shake-flasks, which could be converted into 2 mg/L of the BIA end-product noscapine<sup>7</sup>. Several engineering approaches were combined to achieve this increase in titers. Notably, the first 24 amino acids of NCS were removed, which improved BIA titers 10-fold. Additionally, due to the number of NADPH-utilizing cytochromes P450 in the noscapine pathway, the authors overexpressed three genes that produce NADPH: *ZWFI*, *TYRI*, and the aldehyde dehydrogenase *ALD6*.

In 2020, the Martin group published the synthesis of 4.6 g/L reticuline in fed-batch fermentation, an achievement which firmly establishes yeast as the current best host for heterologous BIA production. Similarly to the Smolke group, the authors removed the first 35 amino acids of NCS, which improved BIA titers. The authors further identified that 4-HPAA, not dopamine, was the limiting factor in high-level BIA synthesis (see Section 2.7.1). The authors knocked out 7 genes, including *ALD6*.

## 2.10 Synthesis and derivatization of morphinan alkaloids in *S. cerevisiae*

Synthesis of naturally-occurring morphinan alkaloids proceeds through (*R*)-reticuline (Figure 2.2)<sup>57,93</sup>. The identification of the epimerase that catalyzes the conversion of (*S*)-reticuline to (*R*)-reticuline in 2015<sup>41,94</sup> marks the complete characterization of all the genes involved in morphinan biosynthesis in planta since the first isolation of morphine from opium poppy in 1806<sup>33</sup>. Parts of the morphine pathway have been reconstituted in *S. cerevisiae* by supplementation of intermediates. In 2015, the first synthesis of opioids from simple carbon sources in a microbial system was achieved<sup>10</sup>.



**Figure 2.2. Synthesis of morphinan alkaloids and derivatives in *Saccharomyces cerevisiae***

Epimerization of (*S*)-reticuline to (*R*)-reticuline and spontaneous rearrangement of salutaridinol-7-*O*-acetate to thebaine mark the starting points for the synthesis of pro-morphinan and morphinan alkaloids, respectively. Colored lines represent the pathways leading to opiate and opiate derivatives with pharmaceutical properties. Pink, synthesis of morphine from thebaine via codeine; purple, synthesis of morphine from thebaine via oripavine; yellow, synthesis of hydromorphone via morphine; green, synthesis of hydromorphone via oripavine; blue, synthesis of hydrocodone and oxycodone from thebaine. Promiscuous enzymes are indicated in colored boxes: thebaine-6-*O*-demethylase (T6ODM, blue), codeine demethylase (CODM, red), *Pseudomonas putida* morphinone reductase (morB, purple), codeinone reductase (COR), and *P. putida* morphine dehydrogenase (morA), which catalyze most of the same reactions, are in green. Spontaneous (Sp) rearrangements, which favor synthesis of additional substrates for promiscuous enzymes, are indicated in yellow. Other abbreviations: CPR, cytochrome P450 reductase; REP, reticuline epimerase; SAS, salutaridine synthase (CYP719B1); SAR, salutaridine reductase; SAT, salutaridinol 7-*O*-acetyltransferase; TS, thebaine synthase; NISO, neopinone isomerase.

### 2.10.1 Derivatization of morphinans from supplemented precursors

Thebaine is a primary feedstock for the chemical synthesis of naturally-occurring and semi-synthetic opioids<sup>57</sup>. As a proof of concept, *S. cerevisiae* expressing morphinan synthesis genes has been used to provide a microbial alternative to chemical derivatization of thebaine<sup>95</sup>. Synthesis of any single morphinan, and in particular morphine, from thebaine is an engineering challenge owing to the complex array of products that can be generated from a small number of enzymes. Morphine synthesis from thebaine requires the co-expression of three enzymes (Figure 2.2). Each enzyme has broad substrate specificity<sup>37,96,97</sup> and the reactions can occur in multiple orders. Two of these orders can result in morphine synthesis (pink and purple pathways in Figure 2.2) while other orders result in the synthesis of side-products (e.g., neopine and 14-hydroxycodeine). Spontaneous rearrangements (within the pink pathway in Figure 2.2) add another level of complexity because some are productive for morphine synthesis while others are not. In short, the array of morphinans produced is entirely dependent on the relative rates of enzyme activities and spontaneous reactions (see Section 2.13).

Conversion of thebaine to morphine by yeast expressing the necessary enzymes was limited to 1.5%, with another 12.7% of thebaine being converted to intermediates and side-products, mainly neopine and 14-hydroxycodeine (Table 2.4)<sup>95</sup>. If flux travels through neopinone, these side-products will be difficult to avoid because neopine is generated when codeinone reductase (COR) activity outcompetes a spontaneous reaction, while 14-hydroxycodeine is generated when spontaneous reactions outcompete COR activity (Figure 2.2). Favoring synthesis through oripavine (purple pathway in Figure 2.2) is likely to reduce the number of side-products in morphine synthesis (see Section 2.13). To achieve the reported 1.5% conversion of thebaine to morphine, some metabolic engineering strategies to favor synthesis of codeinone have already

been employed. High-yield synthesis of morphine in microbes, although possible, will require the development of innovative strategies (see Section 2.13 for examples) to funnel the carbon flux towards morphine.

By contrast, the right combination of enzymes and spontaneous reactions can prove to be very effective at generating a particular morphinan of interest. For example, a pathway to hydrocodone and oxycodone was developed using genes isolated from a soil bacterium growing on industrial poppy waste (Figure 2.2, blue pathway)<sup>95</sup>. Without COR expression, many of the side-products of morphine synthesis were avoided, and almost half of supplemented thebaine was converted to the intended products (Table 2.4). Avoiding particular enzymes with high promiscuity, or identifying/engineering enzymes with desired substrate specificities, may prove to be a more general strategy for morphinan and/or BIA synthesis *in vivo*. For example, hydromorphone synthesis through neopinone (yellow pathway in Figure 2.2) resulted in only 0.4% hydromorphone (Table 2.4), whereas a hypothetical pathway through oripavine (indicated in green in Figure 2.2) could reduce side-products substantially because it avoids COR (Figure 2.2).

Thebaine itself has been derivatized from supplemented norlaudanosoline in *S. cerevisiae* expressing the appropriate enzymes<sup>10,98</sup>. Another spontaneous reaction presents an engineering problem unique to this portion of the pathway: the intermediate salutaridinol-7-*O*-acetate spontaneously rearranges to thebaine at pH 8–9, but rearranges to an undesired side-product at pH 6–7 (Figure 2.2)<sup>99</sup>. To address this issue, a two-step fermentation system was used in which yeast biomass was allowed to accumulate and subsequently switched to pH-buffered media supplemented with either (*R*)-reticuline or salutaridine. The highest production of thebaine was observed at alkaline pHs (pH 8.5-9)<sup>98</sup>. This study highlights pH-adaptable fermentation conditions as an additional challenge for BIA synthesis in yeast (see Section 2.14).

**Table 2.4. Derivatization of BIA backbones in microbial hosts**

<b>Initial compound</b>	<b>Heterologous genes</b>	<b>Final compound(s)</b>	<b>% Yield<sup>a</sup></b>	<b>Backbone</b>	<b>Ref</b>
Reticuline 100 $\mu$ M	SAS, CPR, SAR, SAT	Salutaridine, 15 $\mu$ M Thebaine, 1 $\mu$ M	15 1	Pro-morphinan Morphinan	98
Thebaine 1000 $\mu$ M	COR, T6ODM, CODM	Codeine, 27 $\mu$ M Morphine, 15 $\mu$ M Other opiates, 100 $\mu$ M	2.7 1.5 10	Morphinan	95
Thebaine 1000 $\mu$ M	T6ODM, MorB	Hydrocodone, 180 $\mu$ M Oxycodone, 220 $\mu$ M Other opiates, 30 $\mu$ M	18 22 3	Morphinan	95
Norlaud. 1000 $\mu$ M	6OMT, CNMT, 4'OMT, NMCH, REP, SAS, CPR, SAR, SAT	Thebaine, 0.2 $\mu$ M	0.02	Morphinan	10
Glucose 110000 $\mu$ M	BH4 pathway, TH, DODC, NCS, 6OMT, CNMT, NMCH, 4'OMT, REP, SAS, CPR, SAR, SAT	Thebaine, 0.02 $\mu$ M	<0.001% <sup>b</sup>	Morphinan	10
Glucose 110000 $\mu$ M	BH4 pathway, TH, DODC, NCS, 6OMT, CNMT, NMCH, 4'OMT, REP, SAS, CPR, SAR, SAT, T6DOM, morB	Hydrocodone, 0.001 $\mu$ M	<0.001% <sup>b</sup>	Morphinan	10
Reticuline 33 $\mu$ M	BBE	Scoulerine, 25 $\mu$ M	76	Protoberberine	11
Reticuline 33 $\mu$ M	CTS, CNMT	Magnofluorine, 21 $\mu$ M	64	Aporphine	11
Norlaud. 4000 $\mu$ M	6OMT, CNMT, 4'OMT	Reticuline <sup>c</sup> , 455 $\mu$ M	10	Benzylisoquinoline	12
Norlaud. 4000 $\mu$ M	6OMT, CNMT, 4'OMT, BBE	Scoulerine, 160 $\mu$ M	4	Protoberberine	12
Norlaud. 4000 $\mu$ M	6OMT, CNMT, 4'OMT, BBE, CPR, SOMT, CAS	Canadine, 88 $\mu$ M	2.2	Protoberberine	12
Norlaud. 10 $\mu$ M	6OMT, CNMT, 4'OMT, BBE, CPR, CFS, SPS, TNMT, MSH, P6H	Dihydrosanguinarine, 0.15 $\mu$ M	1.5	Benzophenanthridine	23
1	6OMT, CNMT, 4'OMT, BBE, CPR, SPS, CFS, TNMT, MSH, P6H	Sanguinarine, 0.24 $\mu$ M	0.012	Benzophenanthridine	92

<sup>a</sup>Molar yield.

<sup>b</sup>Yield calculated from glucose, involving multiple steps not included in other strains in this table.

<sup>c</sup>Reticuline concentration estimated based off of closest available standard.

Norlaud: norlaudanosoline.

### 2.10.1.1 Update in 2021

Two enzymes have been identified that catalyze steps in opiate synthesis previously described as spontaneous, which will be useful to researchers looking to guide flux towards codeine and morphine. In 2018, an enzyme was discovered that catalyzes the conversion of salutaridinol-7-*O*-acetate to thebaine, which is called thebaine synthase (TS) (Figure 2.2)<sup>100</sup>. Yeast strains expressing TS are capable of producing 24-fold more thebaine compared to those relying on the spontaneous reaction. An enzyme has also been identified that catalyzes the conversion of neopinone to codeinone, called neopinone isomerase (NISO)<sup>101</sup> (Figure 2.2). When supplemented with thebaine, yeast expressing *T6ODM* and *COR-B* alone produced 100-fold less codeine than yeast additionally expressing *NISO*. These achievements have not yet been followed by improvements in *de novo* titers.

### 2.10.2 Discovery of reticuline epimerase enables *de novo* synthesis of opiates

Until 2015 the fermentation of opiates from simple sugars was not possible owing to the inability to produce (*R*)-reticuline *in vivo*. Initially, the *in vivo* production of (*R,S*)-reticuline from spontaneously condensed (*R,S*)-norlaudanosoline was proposed as a source of (*R*)-reticuline<sup>11,12</sup>. However, more recent reports demonstrate that only (*S*) enantiomers can be accepted by enzymes of the norlaudanosoline-to-reticuline pathway that have been assayed thus far<sup>98</sup>. While enzymes capable of accepting (*R*) enantiomers may exist<sup>18</sup>, epimerization of (*S*)-reticuline to (*R*)-reticuline is currently crucial for opiate biosynthesis. The enzyme catalyzing this stereochemical conversion is reticuline epimerase (REP), a cytochrome P450-reductase fusion protein discovered in 2015<sup>41,94,102</sup>. REP activity has been demonstrated *in vitro*<sup>41,94</sup> and *in vivo*, bridging the upper glucose-to-(*S*)-reticuline and lower (*R*)-reticuline-to-opiate sections of the pathway<sup>10,14,75,98</sup>.

Yeast strains capable of converting supplemented (*S*)-norlaudanoline or endogenously synthesized (*S*)-norcoclaurine to thebaine and hydrocodone have now been engineered. While yields are currently low (Table 2.4), areas of pathway improvement have been highlighted. In yeast capable of converting (*S*)-norlaudanoline to (*R*)-reticuline, total reticuline levels were 100-fold lower than yields in other engineered *S. cerevisiae* strains lacking REP (Table 2.4)<sup>12,23</sup>. This could be due to the promiscuity of REP, which has been demonstrated to accept norlaudanoline-to-reticuline pathway intermediates *in vitro*<sup>41</sup>. *In vivo*, not all available (*S*)-reticuline was converted to (*R*)-reticuline, indicating that there is room to improve REP activity as well. Promoting REP activity on (*S*)-reticuline while limiting its activity on pathway intermediates will be an engineering challenge for the future (see Section 2.13).

### 2.10.3 Update in 2021

In 2016, opioid synthesis was achieved in *E. coli*<sup>15</sup>. Although the identity of REP became known during preparation of the manuscript, REP is a cytochrome P450 and thus the authors reported difficulty in its functional expression in *E. coli*. Instead, the authors focused on the spontaneous condensation of *de novo* dopamine and 3,4-dHPAA into (*R,S*)-norlaudanoline, which provided a pool of (*R*)-norlaudanoline to convert to (*R*)-reticuline. While the methyltransferases between norlaudanoline and reticuline are largely enantiospecific<sup>98</sup>, some (*R*)-reticuline was obtained for downstream derivatization. In total, 2 mg of thebaine was synthesized from 283 mg/L (*R,S*)-norlaudanoline, which could be converted to 0.36 mg/L hydrocodone.

## 2.11 Derivatization of other BIA alkaloids in *S. cerevisiae*

### 2.11.1 Present and future diversity of backbone synthesis

In addition to morphinans, other BIA scaffolds have been produced in yeast (Table 2.4). In combination with (*S*)-coclaurine-*N*-methyltransferase, corytuberine synthase was used to generate the aporphines corytuberine and magnoflorine from reticuline<sup>11</sup>. Alternatively, the protoberberine scoulerine can be synthesized from reticuline by the berberine bridge enzyme (Figure 2.1B)<sup>11,12,23,92</sup>. Scoulerine is a precursor of the protopine, benzophenanthridine, phthalideisoquinoline, and rhoeadine scaffolds. Of these, protopine, benzophenanthridines, and phthalideisoquinolines have been synthesized in yeast<sup>7,23,92</sup>. Some of the phthalideisoquinoline noscapine pathway has been reconstituted, with scoulerine being converted to the downstream protoberberine canadine via the expression of scoulerine *O*-methyltransferase and canadine synthase<sup>12</sup>. Most of the noscapine synthesis pathway has been elucidated as of 2015, raising the possibility of its reconstitution in yeast<sup>36,103</sup>. Bisbenzylisoquinoline scaffolds, too, can likely be achieved *in vivo* because reticuline epimerase can also epimerize the (*R*)-benzylisoquinoline enantiomer required for berbamunine synthesis (see Section 2.13 and Figure 2.3). The synthesis of rhoeadines from protopines is still unknown, as are the enzymes responsible for synthesizing other backbones shown in Figure 2.1B. Enzymes responsible for the addition of functional groups to many of these structures are also unknown.

#### 2.11.1.1 Update in 2021

The final steps of the noscapine biosynthetic pathway have now been elucidated<sup>104,105</sup>, which has enabled reconstitution of the entire pathway in yeast<sup>7</sup>. In total, 17 enzymes downstream of norcoclaurine synthesis are required to achieve noscapine synthesis. Total BIA synthesis in this strain is relatively low compared to the highest BIA titers achieved in yeast (Table 2.2; 3 mg/L

reticuline vs 4.8 g/L reticuline in this work). However, flux through the BIA pathway itself was very efficient considering the number of cytochromes P450 and promiscuous enzymes in the pathway, a feat the authors attribute to a careful selection of carbon source (see Section 2.14). For example, with 10% trehalose, berberine bridge enzyme activity was low, which blocked the pathway at reticuline. With 10% glycerol, on the other hand, reticuline did not accumulate and there was almost no trace of any BIA other than the dominant product, noscapine.

### **2.11.2 Update in 2021: Non-natural BIA synthesis in microbial hosts**

Thanks to the broad substrate acceptance range of norcoclaurine synthase and BIA methyltransferases, it is possible to generate non-natural BIAs and BIA-like scaffolds *via* precursors other than dopamine and 4-HPAA.

Supplementation of 3-X-tyrosine to yeast, where X is fluorine, chlorine or iodine, results in hydroxylation of these compounds by tyrosine hydroxylase and decarboxylation by DODC to yield 3-X-dopamine<sup>7</sup>. NCS condenses these with 4-HPAA resulting in halogenated norcoclaurine analogues; these can be further methylated to produce halogenated reticuline.

In yeast, the enzyme responsible for generating 4-HPAA from 4-HPP, Aro10, is also capable of catabolizing other 2-oxoacids to produce analogous aldehydes, as part of the Ehrlich pathway of amino acid catabolism<sup>70</sup>. In addition to tyrosine, the Ehrlich pathway can also catabolize leucine, isoleucine, tryptophan, phenylalanine, methionine, and valine. When supplemented with leucine, tryptophan, phenylalanine, or methionine as sole nitrogen sources, yeast expressing a BIA synthesis pathway can not only produce other BIA-like scaffolds from them but also methylate them<sup>16</sup>.

Human sulphotransferases may also be expressed in order to modify BIAs after condensation by NCS. In *E. coli*, the human sulphotransferases hSULT1A3 and hSULT1E3 were used to

sulphate specific hydroxyl groups on reticuline<sup>74</sup>. It was further demonstrated that these sulphated reticulines have biological activity in human cells.

### **2.11.3 BIA derivatization in microbes requires knowledge of synthesis pathways**

The diversity of backbones highlights the flexibility of microbial systems for BIA production, while the many cytochrome P450-catalyzed reactions point to yeast as an ideal host (Figure 2.1B). However, the successes and absences of BIA scaffold synthesis in microbes demonstrate the underlying need for information about the enzymes responsible for their production. As genetic techniques for microbe manipulation rapidly improve the turnaround time and throughput of strain development<sup>84,106</sup>, a limiting factor in microbial BIA diversification will become pathway elucidation. Plant biologists have been the traditional source of knowledge through metabolite and transcript profiling of induced and mutagenized plants<sup>107,108</sup>. In addition, the emergence of publicly-available transcriptome databases such as the 1000 Plants and PhytoMetSyn collections<sup>109,110</sup>, combined with advanced techniques in strain engineering, will accelerate the functional discovery of unknown genes and the reconstitution of complex synthetic pathways in microbes.

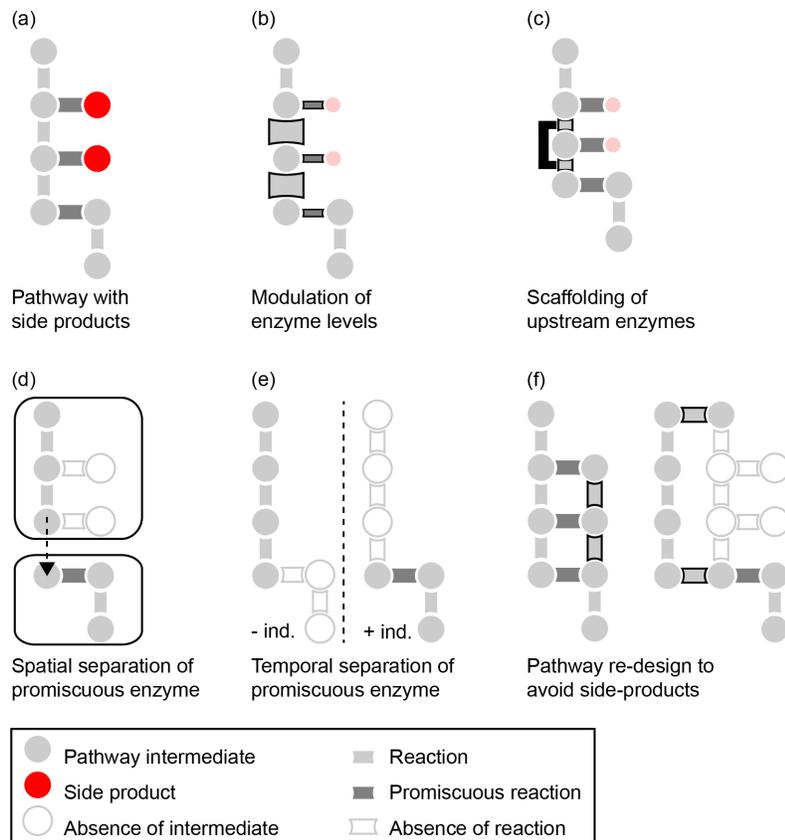
## **2.12 Biosecurity of opiate production in microbes**

When *de novo* synthesis of BIAs in yeast was first reported in 2015, a single step remained to be elucidated that prevented the synthesis of opiates directly from simple carbon sources<sup>14</sup>. Reflecting on this, Oye *et al.*<sup>111</sup> accompanied their report with a list of engineering and policy recommendations that groups seeking to create opiates in microbes should consider incorporating into their practices. Recently, the final step in morphine synthesis has not only been revealed<sup>41,94</sup> but also successfully introduced into yeast to synthesize opioids from sugar<sup>10</sup>. Access to this strain is limited to those with approval from the US Drug Enforcement Administration (DEA), as

recommended by Oye *et al.* While titers remain low, it has been proposed to incorporate biosafety features into the high opiate producing strains to add an extra layer of security to dissuade their theft and malicious use in the future. Suggested features could include (non-disclosed) methods of preventing growth of the strain outside its intended use, as well as methods of identifying strains that have been stolen, such as introducing unique DNA signatures, or “watermarks”, into the strain.

### **2.13 Addressing promiscuity in heterologous pathway reconstitution**

At least 37% of *E. coli* enzymes can accept more than one substrate<sup>112</sup>, and this percentage increases for enzymes involved with secondary metabolism<sup>113</sup>. The broad substrate range of enzyme families in BIA synthesis has been demonstrated *in vitro* (examples include <sup>37-41</sup>). Consequently, reconstitution of BIA synthesis pathways has resulted in the generation of side-products in microbial hosts (Figure 2.3) for both the morphine<sup>95</sup> and sanguinarine pathways<sup>23</sup>.



**Figure 2.3. Strategies for the reduction of side-products in pathways with promiscuous enzymes.**

(A) An initial pathway that results in the accumulation of side-products. (B) DNA copy number, ribosome binding site (RBS) strength, and promoter strength can be adjusted to improve flux by modulating enzyme expression (thinner pipe, less enzyme; fatter pipe, more enzyme; small red dot, less side-product). (C) Enzyme scaffolding (black line and shorter pipes) can prevent access of intermediates to promiscuous enzymes. (D) Enzymes can be physically separated into subcellular compartments or between microbial strains (black boxes). (E) Expression of a promiscuous enzyme can be delayed (- ind) until an inducer is added or growth conditions result in transcription (+ ind). (F) Alternative enzymes can be expressed to adjust the pathway such that side-products are not produced.

Promiscuity presents both a challenge and an opportunity. As a positive, promiscuous enzymes can be used to catalyze reactions for which a dedicated enzyme has not been identified (reviewed in <sup>114</sup>). For example, a promiscuous *N*-methyltransferase was used to methylate BIA structures other than those for which it was characterized<sup>11</sup>. While not always ideal, initial successes can be a starting point for mutagenesis or directed evolution to promote the desired activity<sup>115,116</sup>. However, directed evolution begins with an effective screening strategy. A colorimetric sensor has been used to improve the synthesis of the BIA precursor L-DOPA<sup>14</sup>, but an effective screen remains to be developed for downstream BIA derivatization.

The effects of promiscuity can be mediated by improving flux through the appropriate pathway, or through spatial or temporal sequestration of enzymes away from pathway intermediates. General flux improvement strategies (reviewed in <sup>117</sup>), such as modulation of gene copy number<sup>88,95</sup>, and promoter strength (Figure 2.3B)<sup>12,92</sup>, have improved yield in BIA synthesis pathways. Enzyme scaffolding could also push flux through the intended pathway (Figure 2.3C). Sequestration of promiscuous enzymes into other organelles<sup>95</sup> or separate engineered microbes<sup>81</sup> has improved synthesis of BIAs (Figure 2.3D). Alternatively, temporal control at the level of transcriptional regulation could allow buildup of the desired intermediate before the expression of a promiscuous enzyme (Figure 2.3E).

Re-engineering the pathway itself can also be used to avoid promiscuous side-reactions. For example, non-productive side-products could be brought back into the main pathway through the co-expression of other promiscuous enzymes (Figure 2.3F). Alternatively, the substrate acceptance profiles of each enzyme could be matched to avoid the generation of side-products, recently demonstrated by the heterologous synthesis of >90% unique carotenoids using only promiscuous enzymes<sup>118</sup>. Finally, the use of enzymes to protect functional groups from unwanted side-activities

has been described in opium poppy<sup>103</sup>. This raises the possibility of re-engineering pathways to include blocking steps to avoid promiscuous enzymes, followed by later removal of the group (Figure 2.3F), a method commonly used in synthetic organic chemistry.

## 2.14 Growth conditions for BIA synthesis

The pH of fermentation affects BIA yields, both synthesis and derivatization, in *E. coli* and in yeast. *E. coli* fermentations are usually performed at pH 7, which is sufficiently alkaline for the non-productive spontaneous oxidation of L-DOPA, dopamine, and norlaudanosoline<sup>13</sup>. pH 6 was found to be the best compromise between *E. coli* growth and BIA synthesis<sup>88</sup>. However, the continued disappearance of norlaudanosoline in supernatant indicates that conditions are still not optimal<sup>81</sup>.

The pH of yeast cultures (3-6) is lower than *E. coli* cultures, which reduces spontaneous oxidation of norlaudanosoline and precursors in supernatant. However, BIA (and other alkaloid)<sup>82</sup> derivatization from supplemented precursors is more efficient at higher pHs. As pH is increased from 3 to 8, a greater fraction of supplemented BIAs are associated with cell extracts, and conversion to downstream products is higher<sup>119</sup>. While supplemented precursors represent an intermediate step to a final production strain, endogenously produced dopamine, norcoclaurine, and reticuline are primarily found in yeast supernatant<sup>14</sup>, and typically do not re-enter cells efficiently once outside<sup>75</sup>. Knockout of transporters to prevent secretion is an alternative method for controlling ratios of BIA fractionation<sup>120</sup>. However, secretion of end-products would make downstream industrial processing easier. Importantly, the fractionation of BIAs as well as the rate of flux through BIA synthesis pathways must be balanced for the greatest pathway efficiency.

Groups working with *E. coli* and yeast have explored other fermentation conditions for *de novo* synthesis and derivatization of BIAs *in vivo*. Growth at lower temperature improves BIA

conversion in both species<sup>88,92</sup>. Alternative carbon sources can improve the production of precursors<sup>13</sup> and derivatization of downstream BIAs<sup>92</sup>. Conditions such as pH, temperature, and carbon sources can be adjusted throughout fermentative production. An initial accumulation of biomass before heterologous compound production, such as has been done for production of 1,3-propanediol in *E. coli*<sup>121</sup>, has also improved both the synthesis of aromatic AA pathway derivatives<sup>84</sup> as well as the derivatization of BIAs<sup>23,92,98</sup>.

## 2.15 Concluding remarks

While the opiates are perhaps the most famous members of the benzyloisoquinoline alkaloid family, many other members have potential and realized pharmaceutical value. The scalability and flexibility of microbes has encouraged their development as factories for BIA synthesis. *De novo* synthesis of (*S*)-reticuline in *E. coli*, achieved in 2011, is currently at 160 mg/L and will continue to grow, while *de novo* synthesis in yeast has achieved near-commercial titers of 4.6 g/L. There still exist challenges at all levels of pathway development from improving the intracellular availability of precursors to preventing promiscuity in the final steps of a pathway. Nevertheless, the pace towards industrial microbial fermentation has been impressive, mirroring recent advances in synthetic biology and metabolic engineering, and providing optimism for enabling the microbial production of this large family of natural products.

## 3 Chapter Three: Optimization of Dihydrosanguinarine Pathway in Yeast

**Adapted from:** Narcross, L., Bourgeois, L., Fossati, E., Burton, E., and Martin, V.J.J. (2016). Mining enzyme diversity of transcriptome libraries through DNA synthesis for benzyloquinoline alkaloid pathway optimization in yeast. *ACS Synthetic Biology*, **5**, 12, 1505-1518.

### 3.1 Abstract

The ever-increasing quantity of data deposited to GenBank is a valuable resource for mining new enzyme activities. Falling costs of DNA synthesis enables metabolic engineers to take advantage of this resource for identifying superior or novel enzymes for pathway optimization. Previously, we reported synthesis of the benzyloquinoline alkaloid dihydrosanguinarine in yeast from norlaudanosoline at a molar conversion of 1.5%. Molar conversion could be improved by reduction of the side-product *N*-methylcheilanthifoline, a key bottleneck in dihydrosanguinarine biosynthesis. Two pathway enzymes, an *N*-methyltransferase and a cytochrome P450 of the CYP719A subfamily, were implicated in the synthesis of the side-product. Here, we conducted an extensive screen to identify enzyme homologs whose co-expression reduces side-product synthesis. Phylogenetic trees were generated from multiple sources of sequence data to identify a library of candidate enzymes that were purchased codon-optimized and pre-cloned into expression vectors designed to facilitate high-throughput analysis of gene expression as well as activity assay. Simple *in vivo* assays were sufficient to guide the selection of superior enzyme homologs that ablated the synthesis of the side-product, and improved molar conversion of norlaudanosoline to dihydrosanguinarine to 10%.

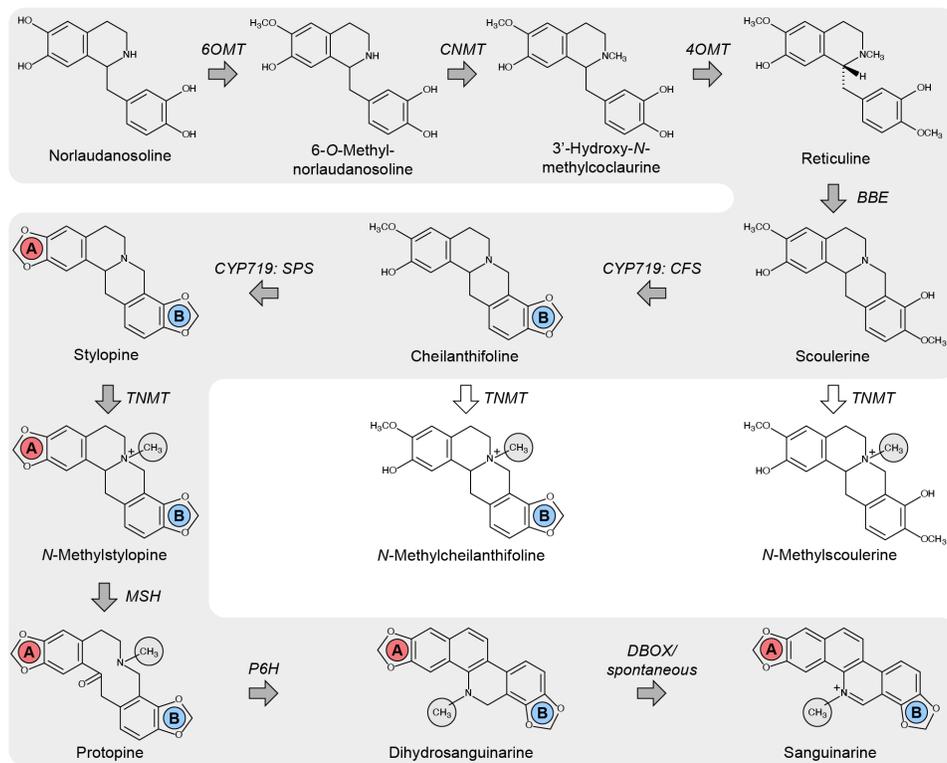
## 3.2 Introduction

A vast amount of genome and transcriptome data is deposited in publicly available resources such as GenBank, which reached a milestone of one trillion base pairs of sequence data in January 2015<sup>122</sup>. More targeted databases like the Thousand Plant<sup>109</sup> and PhytoMetaSyn Projects<sup>110,123</sup> provide further sources of sequence information. Such *in silico* resources are valuable for evolutionary analysis but have traditionally presented few opportunities for metabolic engineers due to the lack of physical DNA available to them<sup>124</sup>. Until recently, the RNA used to generate transcriptome sequence databases was also the source of cDNA used for the targeted amplification of putative ORFs and gene discovery<sup>125–132</sup>. With the cost of DNA synthesis falling from \$1/bp in 2006 to \$0.12/bp in 2014<sup>133</sup>, digital sources of DNA sequences are becoming broadly-accessible primary resources of unique enzymes for the purposes of pathway optimization and the identification of novel activities. This information represents an attractively simple alternative to more traditional methods of pathway optimization through protein engineering approaches such as directed evolution<sup>116,134</sup> or rational modification<sup>135,136</sup>. For example, heterologous synthesis of methyl halides was enabled through the screening of 89 putative methyl halide transferases from metagenomics data deposited to NCBI<sup>137</sup>. More recently, an enzyme bottleneck in the heterologous synthesis of coumarate was alleviated through the screening of a library of both putative and published enzymes purchased entirely from GenBank<sup>138</sup>.

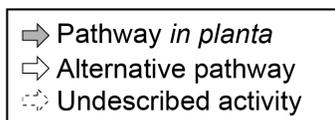
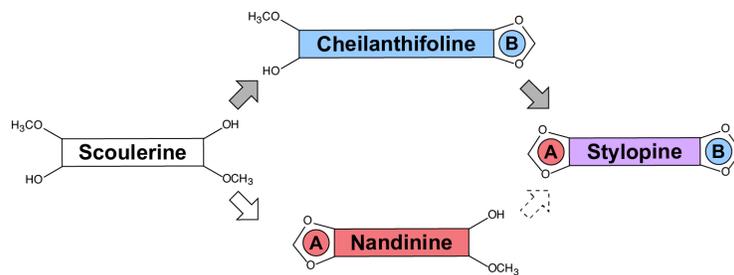
Previously, we reported the reconstitution in *Saccharomyces cerevisiae* of a 10-gene pathway for the synthesis of the benzylisoquinoline alkaloid (BIA) dihydrosanguinarine – a reduced form of the antimicrobial sanguinarine – from the precursor norlaudanosoline (Figure 3.1A)<sup>23</sup>. *De novo* synthesis of benzylisoquinolines in yeast is currently at the microgram/liter level due to low precursor titers and poor performance of the enzyme catalyzing the committed step

<sup>1,14,75</sup>, although the conversion of norlaudanosoline to downstream products can also be inefficient. When supplemented to culture medium, the highest-reported molar conversion of the substrate norlaudanosoline to the key branch point reticuline in yeast is 20%<sup>1,23</sup>. The additional co-expression of the 7 enzymes necessary for dihydrosanguinarine synthesis from reticuline drops molar conversion to 1.5%<sup>23</sup>, with the accumulation of the dead-end intermediate *N*-methylcheilanthifoline contributing to the drop in yield (Figure 3.1A)<sup>23</sup>.

A



B



**Figure 3.1. The dihydrosanguinarine pathway as introduced to yeast**

(A) Dihydrosanguinarine synthesis from (*S*)-norlaudanosoline. Grey background and arrows represent pathway progression. White background and arrows represent undesired side-products.

(B) Schematic representation of CYP719-catalyzed reactions of the dihydrosanguinarine pathway.

The *in planta* pathway, proceeding through cheilanthifoline, is indicated with grey arrows, while an alternative pathway proceeding through nandinine is indicated with white arrows. A previously undescribed step is indicated with dashed lines. Abbreviations: 6OMT, norcoclaurine 6-*O*-methyltransferase; CNMT, coclaurine *N*-methyltransferase; 4'OMT, 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase; BBE, berberine bridge enzyme; CYP719, cytochrome P450 family 719; CFS, cheilanthifoline synthase; SPS, stylophine synthase; TNMT, tetrahydroprotoberberine *N*-methyltransferase; MSH, *N*-methylstylophine hydroxylase; P6H, protophine 6-hydroxylase; DBOX, dihydrobenzophenanthridine oxidase.

Conversion of the dihydrosanguinarine pathway intermediates scoulerine to stylophine is catalyzed by two cytochrome P450 enzymes of the CYP719A subfamily: cheilanthifoline synthase (CFS) converts scoulerine to cheilanthifoline, and stylophine synthase (SPS) converts cheilanthifoline to stylophine (Figure 3.1A). However, in *in vitro* and heterologous *in vivo* systems, stylophine synthase activity is insufficient and leads to cheilanthifoline accumulation<sup>23,139</sup>. Since cheilanthifoline is also a substrate for the promiscuous *N*-methylating enzyme tetrahydroprotoberberine *N*-methyltransferase (TNMT), accumulation of cheilanthifoline results in the synthesis of the non-productive intermediate *N*-methylcheilanthifoline.

Enzyme promiscuity leading to non-productive intermediates is a common problem in heterologous pathway reconstitution<sup>1,95,98,140–142</sup>. Many strategies for reducing side-reactions, such as compartmentalization of competing reactions or enzyme engineering for improved specificity, work within the constraints of enzymes currently in use<sup>1,118,143–145</sup>. Here, using the dihydrosanguinarine pathway, we demonstrate the power of mining transcriptome libraries combined with gene synthesis as an effective strategy for pathway engineering. We postulated that either an SPS that is able to outcompete TNMT for cheilanthifoline, or a TNMT with narrower substrate selectivity, or both, would prevent *N*-methylcheilanthifoline synthesis and greatly improve current dihydrosanguinarine yields.

In this work, two enzyme libraries, one of TNMTs and one of CYP719s were purchased as codon-optimized synthetic genes and screened individually and in combinations. In assaying these 73 enzymes, a new activity was discovered, which inspired a new route to stylophine synthesis (Figure 3.1B). Consequently, synthesis of *N*-methylcheilanthifoline was ablated. The newly engineered dihydrosanguinarine pathway now reaches 10% conversion in yeast cultures

supplemented with the precursor (*R,S*)-norlaudanosoline. The strategy described here is a simple alternative to more rational methods that can be applied to any pathway that requires optimization.

### 3.3 Results

#### 3.3.1 Generation of CYP719 and NMT enzyme libraries

Synthesis of the BIA dihydrosanguinarine from (*S*)-norlaudanosoline requires nine enzymatic reactions (Figure 3.1A). Two of these are catalyzed by *N*-methyltransferases: conversion of 6-*O*-methylnorlaudanosoline to 3'-hydroxy-*N*-methylcoclaurine by coclaurine *N*-methyltransferase (CNMT), and conversion of stylophine to *N*-methylstylophine by tetrahydroprotoberberine *N*-methyltransferase (TNMT). NMTs from BIA-producing plants can accept a variety of BIAs as substrates (Table 3.1). Nevertheless, NMTs have also been demonstrated to differentiate between BIAs that differ by a single methyl group or methylenedioxy bridge<sup>126</sup>. Thus, one approach to reducing *N*-methylcheilanthifoline synthesis was to identify a TNMT that accepted stylophine but not cheilanthifoline. CNMTs were also considered, as some can *N*-methylate downstream dihydrosanguinarine pathway intermediates (Table 3.1). The reverse-translated PhytoMetaSyn transcriptome database was queried using a conserved TNMT/CNMT motif. Putative ORFs selected from the transcriptome database were aligned with published TNMTs/CNMTs using MUSCLE and a phylogenetic tree was generated with the program MEGA6 (Figure 3.2A)<sup>146</sup>. The phylogenetic tree served as a guide for the choice of enzyme candidates to be screened. A total of 15 published and putative NMTs were purchased.

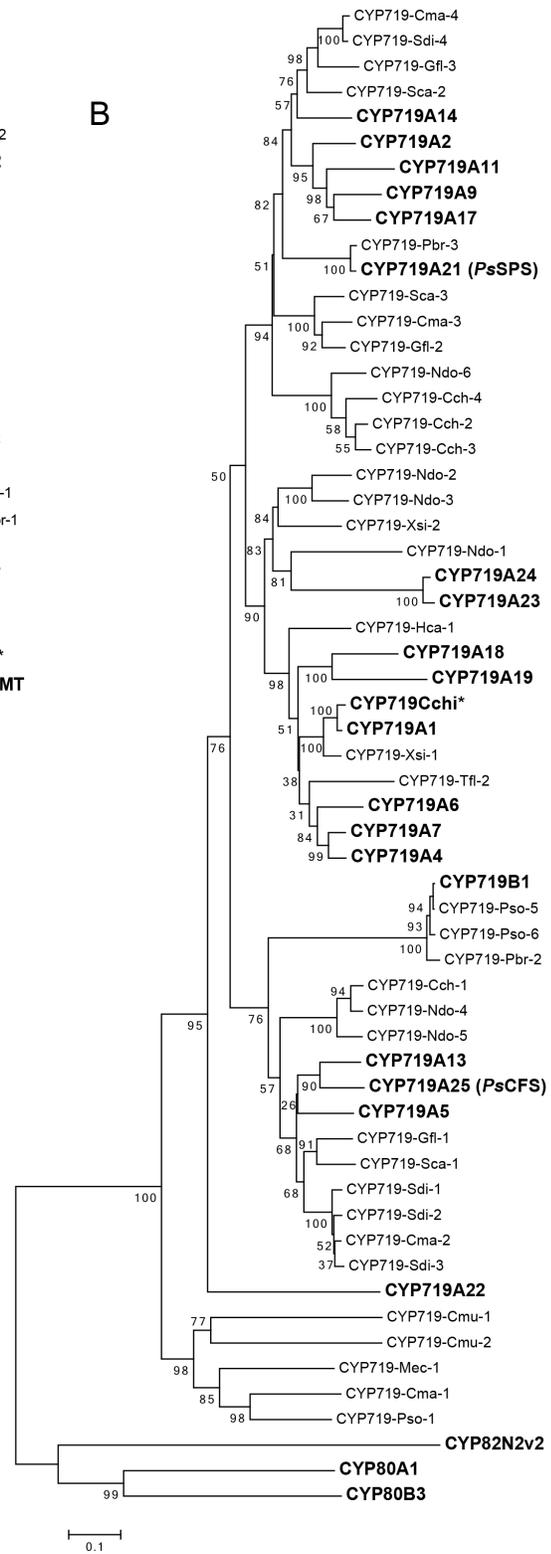
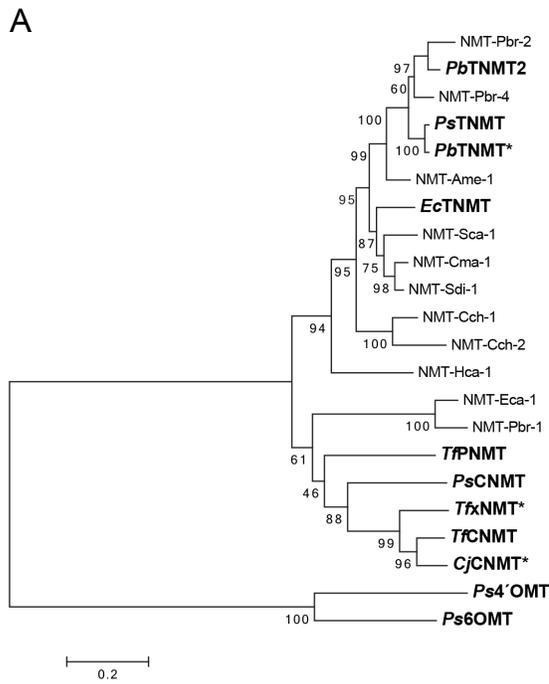
**Table 3.1. Characterized NMTs**

Name	Organism	Identified substrates	Backbone	Ref
CjCNMT	<i>Coptis japonica</i>	Coclaurine Norreticuline Norlaudanosoline 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline 1-methyl-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline	Benzylisoquinoline Benzylisoquinoline Benzylisoquinoline Isoquinoline Isoquinoline	147,148
TfCNMT	<i>Thalictrum flavum</i>	Norreticuline Dimethoxytetrahydroisoquinoline Pavine Scoulerine Tetrahydropalmatine Norlaudanosoline	Benzylisoquinoline Benzylisoquinoline Pavine Protoberberine Protoberberine Benzylisoquinoline	40,149
PsTNMT	<i>Papaver somniferum</i>	Tetrahydropalmatine Canadine Stylopine Scoulerine Cheilanthifoline	Protoberberine Protoberberine Protoberberine Protoberberine Protoberberine	23,40
EcTNMT	<i>Eschscholzia californica</i>	Stylopine Tetrahydropalmatine Canadine Scoulerine	Protoberberine Protoberberine Protoberberine Protoberberine	126
PbTNMT	<i>Papaver bracteatum</i>	Tetrahydropalmatine Stylopine	Protoberberine Protoberberine	126
TfPNMT	<i>Thalictrum flavum</i>	Pavine Stylopine Scoulerine Tetrahydropalmatine	Pavine Protoberberine Protoberberine Protoberberine	126

**Table 3.2. Characterized CYP719s**

Name	Organism	Ring	Published substrates	Backbone	Ref.
CYP719A1	<i>Coptis japonica</i>	A	Tetrahydrocolumbamine	Protoberberine	150
CYP719A2	<i>Eschscholzia californica</i>	A	Scoulerine Cheilanthifoline	Protoberberine Protoberberine	151
CYP719A3	<i>Eschscholzia californica</i>	A	Scoulerine Cheilanthifoline Tetrahydrocolumbamine	Protoberberine Protoberberine Protoberberine	151
CYP719A5	<i>Eschscholzia californica</i>	B	Scoulerine	Protoberberine	152
CYP719A9	<i>Eschscholzia californica</i>	A	(S)-reticuline	Benzylisoquinoline	152
CYP719A13	<i>Argemone mexicana</i>	A	Scoulerine Cheilanthifoline	Protoberberine Protoberberine	139
CYP719A14	<i>Argemone mexicana</i>	B	Scoulerine	Protoberberine	139
CYP719A20	<i>Papaver somniferum</i>	A	Cheilanthifoline	Protoberberine	23
CYP719A21	<i>Papaver somniferum</i>	A	Tetrahydrocolumbamine	Protoberberine	153
CYP719A23	<i>Sinopodophyllum hexandrum</i>	A	Matairesinol	*	128
CYP719A24	<i>Podophyllum peltatum</i>	A	Matairesinol	*	128
CYP719A25	<i>Papaver somniferum</i>	B	Scoulerine	Protoberberine	23
CYP719B1	<i>Papaver somniferum</i>	C	(R)-Reticuline **	Benzylisoquinoline	154

\* not a BIA; \*\* not a methylenedioxy ring bridge closure



### **Figure 3.2. Phylogenetic tree of NMTs and CYP719s**

Phylogenetic tree of **(A)** NMTs or **(B)** CYP719s reported in the scientific literature or from the PhytoMetaSyn database. Alignments were generated with MUSCLE. Phylogenetic trees were generated using the neighbor-joining method with the program MEGA6. Confidence values for each branch, generated with a bootstrap value of 1000, are indicated. Published (**bold**) and putative enzymes were screened for activity with the exception of those indicated with a “\*”. Accession numbers for enzymes are listed in Supplemental Table S1. Outgroups: 4’OMT (AKH61488.1); 6OMT (AAQ01669.1); CYP82N2v2 (F2Z9C1.1); CYP80A1 (P47195.1); CYP80B3 (Q9SP06.1).

Conversion of the dihydrosanguinarine pathway intermediate scoulerine to stylophine requires the formation of two methylenedioxy bridges, indicated by “A” and “B” in Figure 1B. While theoretically the reactions could occur in either order, it has been experimentally determined *in planta* that Ring B closure (catalyzed by CFS) occurs before Ring A closure (catalyzed by SPS)<sup>155</sup>. Both CFS and SPS are cytochrome P450s in the CYP719A subfamily. This subfamily also includes other members that catalyze methylenedioxy bridge formations on other BIAs and other alkaloids (Table 3.2), and still other methylenedioxy bridge-containing alkaloids have been identified for which the appropriate methylenedioxy bridge-forming enzymes are still unknown. Diversity amongst methylenedioxy-bridge containing alkaloids and CYP719 substrate acceptance profiles suggests that the CYP719 enzyme family is extensive and may include SPS enzyme homologs that are more appropriate for heterologous reconstitution of the dihydrosanguinarine pathway. Traditionally, the naming scheme for CYP719s is based on an identified product (i.e., stylophine synthase and cheilanthifoline synthase). However, this naming scheme becomes untenable when the same enzyme can synthesize multiple products. Here, we refer to CYP719s by the location of methylenedioxy bridge formation: Ring A-closing CYP719s and Ring B-closing CYP719s.

Reverse-translated transcriptome data from the PhytoMetaSyn database was queried for a conserved heme-binding cytochrome P450 motif and an N-terminal motif conserved amongst published CYP719s. Queries were narrowed down using BLASTclust. BLASTclust sorts sequences into groups using the criterion of percent sequence identity, which is convenient for the study of CYPs because CYP families and subfamilies are defined based on this criterion (45% amino acid identity defines a family, 55% amino acid identity defines a subfamily)<sup>156</sup>. Stringency was set to 50% in order to include the CYP719B subfamily, which also has activity on BIAs<sup>154</sup>.

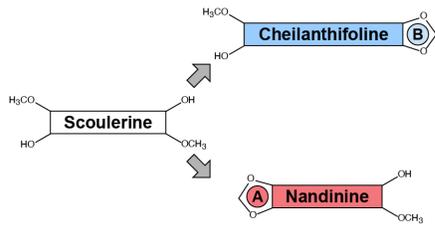
Putative CYP719s that clustered with published CYP719s were aligned with MUSCLE, and a phylogenetic tree was generated using MEGA6 (Figure 3.2B)<sup>146</sup>. Three clades were observed, which were assigned predicted activities based on the co-alignment with characterized CYP719s: Ring A-closing CYP719s, further segregated into two subclades of CYP719s predicted to act on cheilanthifoline (stylophine synthases) or on the BIA tetrahydrocolumbamine (canadine synthases); Ring B-closing CYP719s, further segregated into one subclade of cheilanthifoline synthases and one subclade of CYP719Bs; and CYP719s with unknown activities. A total of 54 characterized and putative CYP719s were purchased for screening.

### **3.3.2 Selection of replacement Ring A-closing CYP719s**

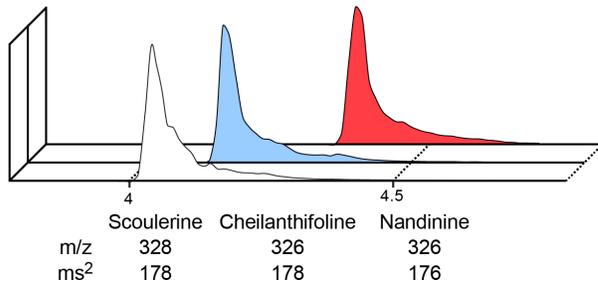
The library of CYP719s included enzymes with characterized activity on relevant BIAs (scoulerine and cheilanthifoline), other BIAs or other alkaloids, as well as enzymes with predicted activity or no predicted activity (Table 3.2, Figure 3.2B). Following a qualitative assessment of CYP719 expression through comparison of the fluorescence of CYP719-GFP fusion proteins (see Supplemental Results & Discussion), an initial activity screen was performed to validate predicted Ring A- and Ring B-closing activities within the CYP719 library. CYP719s were expressed in yeast and supplemented with the dihydrosanguinarine pathway intermediate scoulerine. Scoulerine is a substrate for both Ring A- and Ring B-closing CYP719s, forming nandinine and cheilanthifoline, respectively (Figure 3.3A). Although nandinine and cheilanthifoline have the same mass and similar structures, they are distinguishable by HPLC-MS both through elution time and by MS/MS profile (Figure 3.3B). Analysis of nandinine and cheilanthifoline synthesis confirmed that predicted Ring A- and Ring B-closing activities were generally accurate with no cases of Ring B-closure where Ring A-closure was predicted, or vice versa (Figure 3.3C and Figure 3.3D). In addition, no conversion of scoulerine was detected for predicted CYP719Bs, CYP719s

with characterized activity on non-BIAs, or putative CYP719s with no predicted activity. Scoulerine was widely accepted amongst Ring B-closing CYP719s with 10 of 12 candidates converting >95% of the scoulerine to cheilanthifoline (Figure 3.3C and Figure 3.3D). Conversely, a greater range of nandinine synthesis was observed amongst predicted Ring A-closing CYP719s. Of these, 7 of 18 predicted stylophine synthases and 3 of 16 predicted canadine synthases converted >95% of the scoulerine to nandinine (Figure 3.3C and Figure 3.3D). These Ring A-closing CYP719s were considered for further characterization.

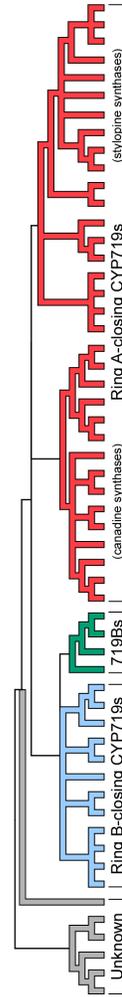
A



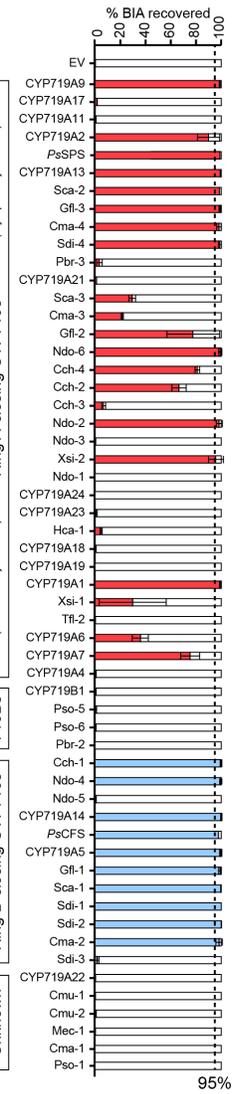
B



C



D



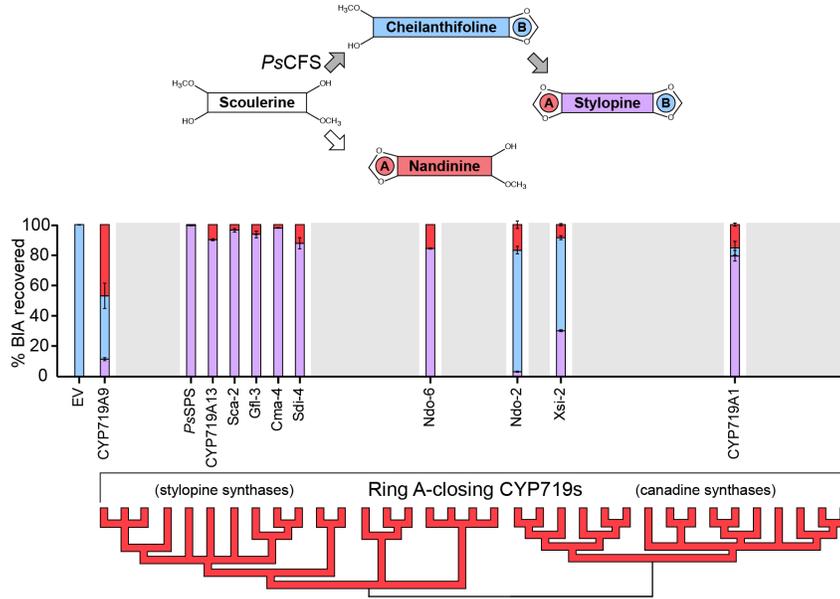
**Figure 3.3. Screening of CYP719 library for protoberberine Ring A- and Ring B-closing activity**

Yeast with an integrated *PsCPR* gene were transformed with individual plasmids harboring the CYP719 library and screened for activity on scoulerine, which is a substrate for both Ring A- and Ring B-closure. **(A)** LC-MS analysis of scoulerine, cheilanthifoline, and nandinine showing distinct elution profiles as well as parent ( $m/z$ ) and daughter ( $m/z$ ) ions. **(B)** The relative percentage of each BIA produced from scoulerine in culture supplementation assays. **(C)** Condensed phylogenetic tree (80% cutoff) for comparison of predicted activity with actual activity. **(D)** Predicted Ring A-closing CYP719s are in red, predicted Ring B-closing CYP719s are in blue, CYP719Bs are in green, and enzymes with no predicted activity are in gray. EV: empty vector control.

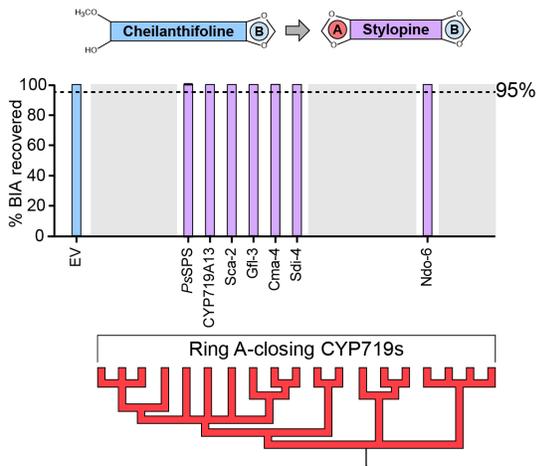
Next, selected Ring A-closing CYP719s were co-expressed with the Ring B-closing CYP719 *PsCFS*, previously used in the heterologous reconstitution of dihydrosanguinarine synthesis<sup>23,92</sup>. When supplemented with scoulerine, the expected product is stylophine (Figure 3.4A). In each yeast strain, scoulerine was entirely consumed but different proportions of cheilanthifoline, nandinine, and stylophine were observed depending on the co-expressed Ring A-closing CYP719 (Figure 3.4A). Expression of Ring A-closing CYP719s predicted to be canadine synthases resulted in residual cheilanthifoline. These candidates were not considered for further screening. With one exception, cheilanthifoline was not detected when the Ring A-closing CYP719 was predicted to be a stylophine synthase.

Unlike previously reported combinations of cheilanthifoline and stylophine synthases, nandinine was a product observed in each combination of Ring A-closing CYP719s and *PsCFS*<sup>23,92,139</sup>. Nandinine synthesis resulted from Ring A-closing CYP719s out-competing *PsCFS* for scoulerine. Accumulation of nandinine also indicated that it is not a preferred substrate of *PsCFS*. Improved activity of Ring A-closing CYP719s relative to *PsCFS* is a desired quality, but the generation of a new side product is not. Nandinine accumulation could be avoided by limiting the pool of potential SPS's to those with activity on cheilanthifoline but not scoulerine. Because many Ring A-closing CYP719s in the library can accept scoulerine, this is not an ideal limitation. Alternatively, if a Ring B-closing CYP719 could be identified that also accepts nandinine as substrate, then nandinine could be re-captured into the main pathway, shifting from a side-product to a pathway intermediate (Figure 3.1B).

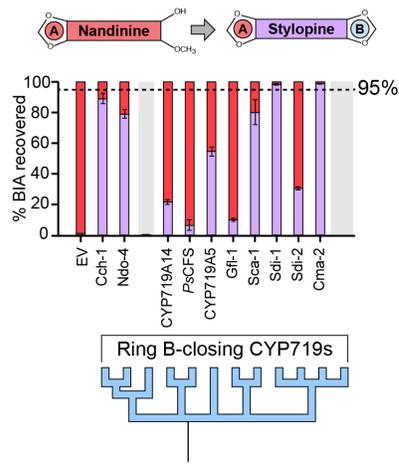
A



B



C



**Figure 3.4. CYP719-catalyzed stylophine synthesis from scoulerine, cheilanthifoline or nandinine**

(A) BIA production profile from scoulerine for yeast strains harboring *PsCPR*, *PsCFS*, and Ring A-closing CYP719s. (B) BIA production profile from scoulerine for yeast strains harboring *PsCPR* and Ring B-closing CYP719s. (C) BIA production profile from scoulerine for yeast strains harboring *PsCPR* and Ring A-closing CYP719s. The amount of each BIA produced is reported as a percentage of total BIAs. EV: empty vector control.

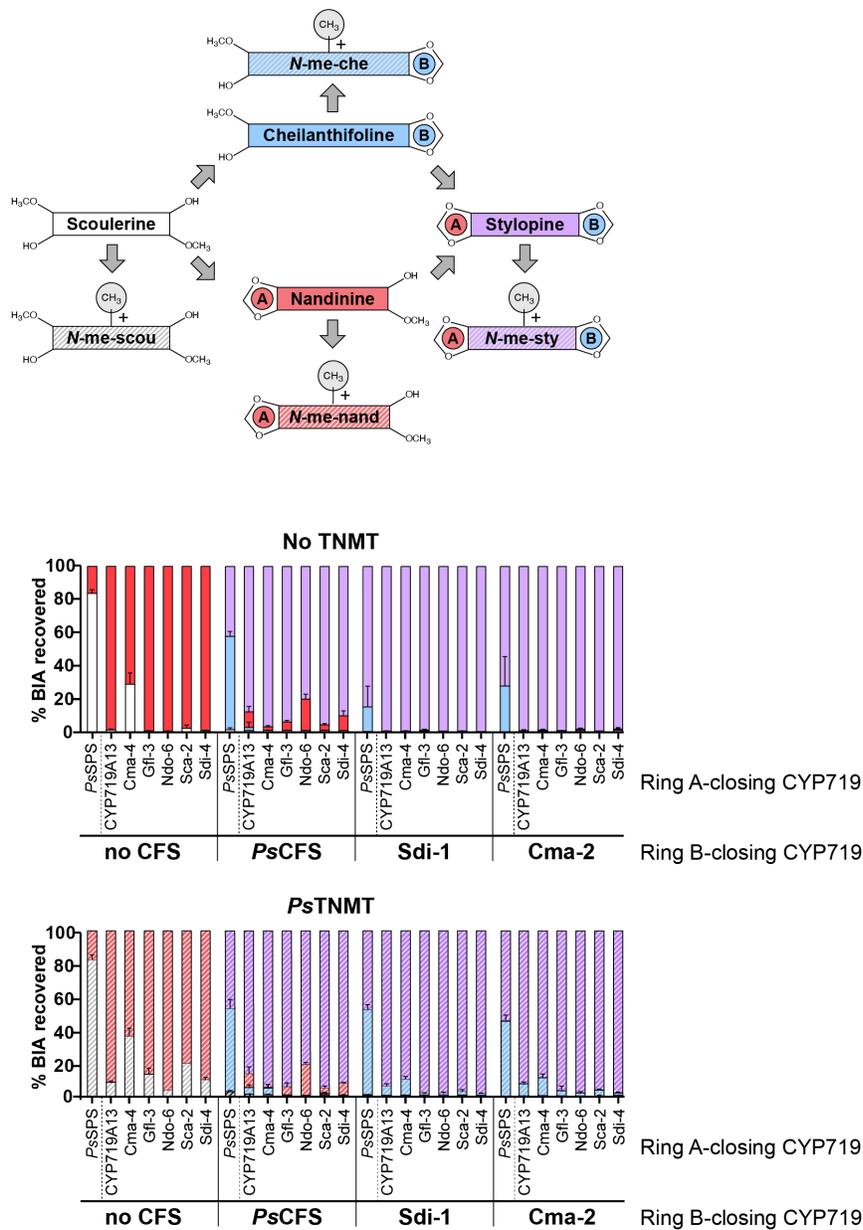
### 3.3.3 Engineering of a non-natural stylopine synthesis pathway

To simplify further assessment of Ring A- and Ring B-closing CYP719 activity, the appropriate substrates were supplied directly to yeast strains. As cheilanthifoline and nandinine were not available commercially, they were generated from scoulerine by incubation with yeast expressing an appropriate CYP719. Supernatant containing cheilanthifoline was used as substrate to test the activity of Ring A-closing CYP719s (Figure 3.4B), while supernatant containing nandinine was used to test Ring B-closing CYP719s (Figure 3.4C). As was suggested in co-expression analysis, all selected Ring A-closing CYP719s were able to convert >95% of cheilanthifoline to stylopine. Many Ring B-closing CYP719s had some activity on nandinine, with 2 of 10 candidates (Sdi-1 and Cma-2) converting >95% of supplemented nandinine to stylopine. Acceptance of nandinine by the two Ring B-closing CYP719s enables the non-natural “Ring A first” pathway for stylopine synthesis (Figure 3.1B). Hence, the six Ring A-closing CYP719s and two Ring B-closing CYP719s that converted >95% of their supplemented BIA to stylopine were selected for combinatorial testing in the presence of TNMT.

### 3.3.4 Combinatorial testing of Ring A- and Ring B-closing CYP719s

Combinations of Ring A- and Ring B-closing CYP719s were next co-expressed in the presence and absence of TNMT to measure production of downstream *N*-methylated BIAs (Figure 3.5). When supplemented with scoulerine, stylopine should be the product in the absence of TNMT and *N*-methylstylopine should be the product in the presence of TNMT. Any accumulated cheilanthifoline, nandinine, or their *N*-methylated derivatives, would indicate an undesired combination of CYP719s. Nandinine was observed during co-expression of *Ps*CFS with Ring A-closing CYP719s, but not with either of the two selected Ring B-closing CYP719s. Residual cheilanthifoline (and *N*-methylcheilanthifoline in the presence of TNMT) was observed in samples

expressing 2 of the 6 Ring A-closing CYP719s, which was not expected because these enzymes previously converted >95% of cheilanthifoline to stylophine (Figure 3.4B). Between experiments, Ring A-closing CYP719s had been placed under the control of a new promoter / terminator pair in order to allow homology-mediated cloning of a double CYP719 gene cassette. The other 4 Ring A-closing CYP719s, when co-expressed with either of the 2 Ring B-closing CYP719s, resulted in >95% conversion of scoulerine to stylophine in the absence of TNMT, and >95% conversion of scoulerine to *N*-methylstylophine in the presence of TNMT. These combinations were selected for integration into the dihydrosanguinarine pathway.

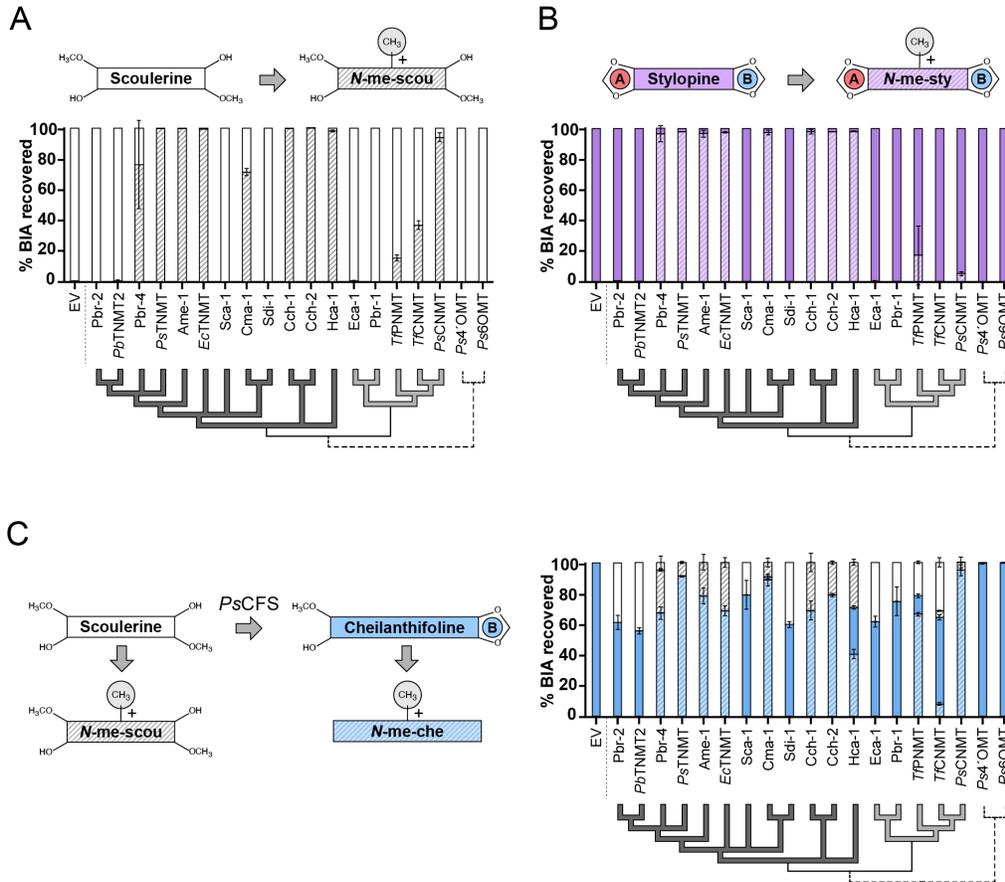


**Figure 3.5. Combinatorial expression of selected CYP719s with and without TNMT co-expression**

BIA-producing profiles from scoulerine for yeast strains harboring *PsCPR* or *PsCPR* and *PsTNMT* genes and co-transformed with plasmids harboring Ring A- and Ring B-closing CYP719s. The amount of each BIA produced is reported as a percentage of total BIAs. Abbreviations: CPR: cytochrome P450 reductase; TNMT: tetrahydroprotoberberine *N*-methyltransferase.

### 3.3.5 Selection of replacement TNMT

Before retrofitting a final dihydrosanguinarine production strain with the newly discovered CYP719s, the NMT library was screened for NMTs with greater substrate preference for stylophine. Yeast strains harboring the NMT library were supplemented with scoulerine (Figure 3.6A) and stylophine (Figure 3.6B) in order to assess relevant substrate acceptance profiles. In addition, the activity of NMTs on cheilanthifoline was also assayed through the supplementation of scoulerine to yeast strains co-expressing the NMT library and the Ring B-closing CYP719 *PsCFS* (Figure 3.6C). Included, as negative controls, were two *O*-methyltransferases (OMTs) in the dihydrosanguinarine pathway that were not expected to methylate the supplemented BIAs<sup>157</sup>. In general, if an NMT *N*-methylated stylophine, it also *N*-methylated cheilanthifoline and scoulerine (Figure 3.6). These NMTs all aligned with published TNMTs, but not every enzyme aligning with TNMTs had activity on the BIAs tested here. Amongst CNMTs, *PsCNMT* is uniquely able to *N*-methylate scoulerine and cheilanthifoline, but no CNMT was able to *N*-methylate stylophine. Furthermore, when co-expressed with *PsCFS*, NMTs with no measurable activity on the BIAs tested appeared to interfere with cheilanthifoline synthesis, as conversion of scoulerine to cheilanthifoline was lower than either empty vector or OMT control strains. Since we did not identify an NMT that methylated stylophine without also methylating scoulerine and/or cheilanthifoline, the strategy of identifying an NMT with a different substrate acceptance profile was not pursued further.



**Figure 3.6. Activities of NMTs on select BIAs**

Cultures of yeast strains harboring NMT-expressing plasmids were supplemented with (A) scoulerine or (B) stylopine. (C) Cultures of yeast strains harboring *PsCPR* and *PsCFS* genes were transformed with plasmids expressing the NMTs and supplemented with scoulerine. Yeast were incubated for 16 hrs before total BIAs were extracted and analyzed. Below each graph is a condensed phylogenetic tree for comparison of predicted and actual activities. Predicted TNMTs are in dark grey, predicted CNMTs are in light grey, and OMTs are indicated with a dashed line. Abbreviations: EV: empty vector.

### 3.3.6 Retrofitting and testing an optimized dihydrosanguinarine producing strain

Through CRISPR-directed homologous recombination, most of the norlaudanosoline-to-dihydrosanguinarine pathway was chromosomally integrated into a single yeast strain. Missing were the CYP719 Ring A- and Ring B-closing enzymes, which were combinatorially co-expressed from plasmids. Cultures of the resulting strains were then supplemented with (*R,S*)-norlaudanosoline at increasing concentrations, and levels of *N*-methylated intermediates and dihydrosanguinarine were measured. In the absence of CYP719s, *N*-methylscoulerine was produced as a result of the accumulation of scoulerine (Figure 3.7A). Normalized *N*-methylscoulerine levels remained constant across increasing substrate concentrations, which indicated that there were no measurable bottlenecks in the pathway from norlaudanosoline to scoulerine. *N*-methylscoulerine was not observed when CYP719s were expressed. As expected, BIAs extracted from yeast cultures expressing the original CYP719 combination (*PsSPS* and *PsCFS*) showed *N*-methylcheilanthifoline accumulation at every norlaudanosoline concentration. Further, as norlaudanosoline concentration increased, *N*-methylcheilanthifoline levels rose relative to normalized values. In contrast, *N*-methylcheilanthifoline did not accumulate with any combination of the newly selected CYP719s, at any concentration of norlaudanosoline. When the strain co-expressing *PsSPS* and *PsCFS* was supplemented with 10  $\mu$ M norlaudanosoline, 5% was converted to dihydrosanguinarine and sanguinarine (Figure 7B). Multiple combinations of Ring A- and Ring B-closing CYP719s resulted in improved levels of dihydrosanguinarine and sanguinarine, reaching ~10% conversion of norlaudanosoline to dihydrosanguinarine with no intermediate or side product observed. At higher concentrations of norlaudanosoline, the pathway intermediate *N*-methylstylopine accumulated, increasing relative to normalized norlaudanosoline conversion values.



### 3.4 Discussion

With the cost of next-generation DNA sequencing less than \$1/million base pairs<sup>133</sup>, transcriptome databases can be cheaply generated and used for enzyme discovery<sup>125,127,158–160</sup>. Historically, harnessing the power of transcriptome databases was difficult without access to the physical RNA used to generate the libraries. Now, advances in DNA synthesis technologies<sup>161</sup> and molecular biology techniques for efficient heterologous gene expression<sup>84,162–164</sup> can accelerate and improve the enzyme discovery process. Here, we demonstrate the power of combined accessibility to transcriptome data and affordable gene synthesis for pathway engineering and optimization. Not only did our synthetic gene library contain multiple enzymes capable of improving the pathway, it also included a novel activity enabling pathway redesign to prevent the synthesis of side products. As the cost of DNA synthesis continues to drop, we foresee this strategy becoming increasingly common for pathway engineering and optimization.

The work presented here is an example of the power and limits of predictive search for enzymatic activities from sequence data. Predicted activities for putative ORFs were assigned based on co-alignment of characterized enzymes within clades. For both NMTs and CYP719s, broad activities were accurate (i.e., *N*-methylase activity or Ring A- vs. Ring B-closure), but enzymes co-aligned within a sub-clade did not necessarily have the same substrate acceptance profiles. This is not an uncommon phenomenon<sup>125,165,166</sup>, which highlights the value in using large libraries of orthologous enzymes to increase the chances of finding an activity of interest.

For both published NMTs and CYP719s, our analysis of substrate preferences included more positive hits than previous characterization. For instance, TNMTs have been demonstrated to have much lower activity on scoulerine than stylophine (0-10% of relative activity depending on the homolog)<sup>40,126,167,168</sup>, whereas no difference was observed here (Figure 3.6A, Figure 3.6C).

Similarly, some CYP719s such as CYP719A1 have been shown to have little to no activity on scoulerine<sup>150</sup> but here were able to convert >95% of supplemented scoulerine to nandinine. Unlike traditional biochemical assays, supplementation and bioconversion assays often occur over a longer duration (1 hr vs. >16 hrs) and are a measure of reaction progression, not reaction speed<sup>11,169</sup>. For the identification of enzymes with novel activities for heterologous pathway engineering, longer incubation times in *in vivo* conditions may flag candidates that may otherwise be discarded.

In this work, all CYP719s were co-expressed with a single cytochrome P450 reductase (CPR) from *P. somniferum* (*PsCPR*). This setup was sufficient to identify both Ring A- and Ring B-closing CYP719s from multiple plant species with activity on supplemented BIAs. However, *PsCPR* may not be an ideal partner for all the CYPs in the library. It has long been recognized that any one CPR cannot support the activity of all CYPs<sup>170</sup>. While a CPR from the same species is often used if available<sup>171–173</sup>, the existence of multiple CPR in plants complicates the selection process<sup>61</sup>. Further, the actual relationship between CYP and CPR is unpredictable, with CPRs from other plants like *Arabidopsis thaliana*<sup>92,154,174</sup> and even the native yeast CPR<sup>22</sup> supporting heterologous activity of some CYPs. Therefore we presume that combinatorial co-expression of the CYP719 library with a putative CPR library from other organisms may improve the activity of some of the CYPs in our library.

Enzyme promiscuity presents serious challenges for the reconstitution of heterologous pathways<sup>23,118,175</sup>, which can compound as a pathway increases in size. For example, scoulerine is a substrate for 4 of 9 enzymatic steps in the dihydrosanguinarine pathway (Table 3.1, Table 3.2). Nevertheless, *in vivo* combinatorial screens identified multiple enzyme combinations in which no scoulerine side-products were found to accumulate. Initially, nandinine synthesis was an undesired

activity. Although nandinine acceptance by Ring B-closing CYP719s had not been characterized (Table 3.2), and although published Ring B-closing CYP719s had little activity on nandinine when assayed here, our enzyme library contained CYP719s that could synthesize stylophine from nandinine. Thus, a side-product generated through enzyme promiscuity was re-introduced back into the pathway through the activity of a second promiscuous enzyme. Using complementary substrate acceptance profiles to achieve a single product of interest is still new<sup>118</sup>, although we expect this approach will become more prominent as large enzyme libraries become increasingly accessible through gene synthesis.

The conversion of norlaudanoline to dihydrosanguinarine in this work compares favorably to other reports of heterologous dihydrosanguinarine/sanguinarine synthesis. Norlaudanoline supplementation in this work is lower than the system published by Trenchard *et al.* (10  $\mu$ M vs. 2 mM), but yield is higher (10% conversion to dihydrosanguinarine and sanguinarine vs. 0.012% conversion to sanguinarine)<sup>92</sup>. Following our identification of a bottleneck at SPS<sup>23</sup>, Trenchard *et al.* performed a 2x2 combinatorial search for a new CFS and SPS, ultimately selecting CYP719A5 and CYP719A2, respectively<sup>92</sup>. While both enzymes were included in our screen, neither enzyme was ultimately selected; CYP719A5 did not have sufficient activity on nandinine, and CYP719A2 did not have sufficient activity on scoulerine. CYP719A2 also displays low fluorescence as a GFP fusion protein, whereas enzymes selected for integration into the dihydrosanguinarine pathway tended to display the highest levels of fluorescence amongst enzymes with any particular desired activity (see Section 3.6.1).

While the system here represents a 10-fold improvement over previous work<sup>23</sup>, the pathway can be improved further. Conversion of norlaudanoline to dihydrosanguinarine was 10% at 10  $\mu$ M norlaudanoline. However, as supplemented norlaudanoline concentrations increased, so

did buildup of the intermediate *N*-methylstylopine. This points to the downstream enzyme *N*-methylstylopine hydroxylase (MSH), a member of the CYP82 family, as the next target for improvement. A single MSH has been identified thus far<sup>176</sup>, but other members of the same family have been demonstrated to have activity on BIAs<sup>38,177</sup>. The strategy of enzyme selection and screening employed here may be able to improve this next step as well.

Commonly-cited strategies for heterologous pathway optimization include control of enzyme transcription and translation, as well as spatial control of enzymes through the use of scaffolds or targeting signals<sup>117</sup>. The work presented here adds to the growing number of studies that highlight the potential of transcriptome libraries to provide new solutions to these problems. As the cost of DNA synthesis continues to drop, the screening of enzyme homolog libraries should be considered an integral part of pathway engineering.

## **3.5 Materials and Methods**

### **3.5.1 Transcriptome data analysis**

The PhytoMetaSyn database ([www.phytometasyn.ca](http://www.phytometasyn.ca)) of assembled transcriptome data from BIA-producing plants<sup>3</sup> was reverse-translated into putative ORFs using OrfPredictor<sup>178</sup>. Translated ORFs were scanned for motifs of interest generated from sequence alignments of published proteins and candidate sequences kindly provided by Dr. Peter Facchini (University of Calgary) (NMTs: ERAQI(K/Q)DG; CYP719s: FxxGxxxCxG, PxIGN). Putative NMTs identified from the library were aligned with published sequences and candidate sequences and phylogenetic trees were generated and a subset was manually selected for testing. Putative CYP719s, published CYP719s, and CYP719s deposited online on GenBank and the Cytochrome P450 Homepage<sup>179</sup> were grouped by BLASTclust into groups with 55% sequence identity at the amino acid level of over 95% of the sequence. These groups were then aligned and phylogenetic trees were generated

in the same manner as for NMTs, and a subset was manually chosen for testing. All alignments were performed with MUSCLE and phylogenetic trees were built by the neighbor-joining method using the program MEGA6.0<sup>146</sup> with a bootstrap value of 1000. When indicated, branches with less than 80% confidence values were condensed to build condensed phylogenetic trees. Enzymes selected for study were codon-optimized for expression in yeast and synthesized by Gen9 (Cambridge, MA).

### 3.5.2 Construction of plasmids

All cloning was performed *via* yeast homologous recombination<sup>162</sup> using regions of homology added to DNA during PCR (see primer list, Table 3.4). PCR of DNA to be cloned was performed with Phusion polymerase. When appropriate, *E. coli* was cultivated in LB medium at 37°C with shaking at 200 rpm with supplementation of 100 µg/mL ampicillin as necessary. A series of vectors designated as pBOT (Figure 3.8, Supplemental Figures) were designed to facilitate gene expression and enzyme activity assays. Each pBOT vector has a unique combination of yeast selection marker, promoter, and terminator (Table 3.5). Details on the construction of the pBOT vectors is outlined in Section 3.6.2. To switch selectable markers of Ring B-closing CYP719s, the promoter-gene-terminator cassette from pBOT-Trp was introduced to pBOT-Leu by digesting both vectors with *NotI/AscI*, gel purifying the pBOT-Trp insert and the pBOT-Leu vector, and ligating both fragments.

Dihydrosanguinarine pathway genes other than those purchased in this study were cloned into either pGREG or pYES vectors (Table 3.5), where they could be used for activity assays and/or genomic integration. Promoters, genes, and terminators introduced into pGREG or pYES vectors were amplified with overlapping homology regions (Table 3.4) and cloned by homologous

recombination. Heterologous DNA (linkers C1, C6, H1, and H2) were added as previously described<sup>23</sup> (indicated in bold in Table 3.4, Table 3.5, and Table 3.6).

### 3.5.3 Chromosomal integration of genes and multi-gene pathways

To facilitate activity assays and to build a dihydrosanguinarine synthesis strain, some genes and multi-gene pathways were integrated into the genome of *S. cerevisiae* into sites previously determined to allow high levels of gene expression<sup>180</sup> (sites and genes are indicated in Table 3.6). Strains built to facilitate activity assays (strains GCY1333, GCY1270, and GCY1317) were built using homologous recombination and selected with antibiotic resistance to 200 µg/L geneticin and/or hygromycin using the antibiotic markers *kanMX* and *hphNT1*, respectively<sup>181,182</sup>. The dihydrosanguinarine synthesis strain (strain GCY1440) was built using homologous recombination and CRISPR-Cas9 (Table 3.5). Regions of DNA (~500 bp) upstream and downstream of the integration site (UP and DOWN regions, respectively) were amplified with homology to heterologous DNA to guide gene integration. Dihydrosanguinarine synthesis genes used to build strains GCY133, GCY1270 and GCY1317 were excised from pGREG vectors using *AscI/NotI* and gel purified, while dihydrosanguinarine synthesis genes used to build strain GCY1440 were amplified from plasmids and heterologous linkers were added (LV3, LTP1, LTP2, LV5). When applicable, Cas9 was directed to the 5' and 3' ends of integration sites using two guide RNAs (gRNAs). The 20-bp targeting regions of gRNAs were introduced through splice overlap extension using the primers indicated in Table 3.4. Both gRNAs, along with linearized vector containing Cas9 (pCAS-Tyr)<sup>183</sup> were co-transformed into *S. cerevisiae*. Successful gene integration by both methods was confirmed by PCR. All primers used are listed in Table 3.4.

### 3.5.4 BIA culture supplementation assays

Activity of CYP719s and NMTs on BIAs was analyzed using culture substrate supplementation assays. Yeast cultures were grown in yeast nitrogen base with 2% glucose and amino acid dropout media as appropriate (YNB) at 30°C and 200 rpm. Yeast cells harboring enzymes of interest were inoculated in triplicate into 100 µL of media in 96-well 2 mL deep-well plates and incubated overnight. The following day, 900 µL of fresh media was added (1:10 dilution) and cultures were incubated for an additional 6 hrs. Cells were pelleted by centrifugation for 5 min at 3,200 g and supernatants were aspirated. Cell pellets were suspended in 300 µL TE (10 mM Tris, 1 mM EDTA, pH 8) containing BIAs as appropriate, at a concentration of 5 µM unless otherwise specified and incubated overnight at 30°C with shaking. The following day, cells were pelleted at 3,200 g. Supernatants were transferred to 96-well microtiter plates, diluted 1:1 in 100% methanol and clarified at 3,200 g prior to analysis by LC-MS. To extract BIAs from cells, pellets were suspended in 300 µL methanol, vortexed at 1,000 rpm and 4°C for 30 min and centrifuged at 3,200 g. The resulting extracts were analyzed by LC-MS as described below. To stay in the linear range of the LC-MS, samples supplemented with >10 µM of a BIA were diluted before analysis. (*S*)-Scoulerine and (*S*)-stylophine were purchased from ChromaDex (Irvine, CA, USA) and (*R,S*)-norlaudanosoline was purchased from Enamine Ltd. (Kiev, Ukraine). Dihydrosanguinarine was prepared from sanguinarine by NaBH<sub>4</sub> reduction<sup>184</sup>.

### 3.5.5 Liquid chromatography-mass spectrometry

Using a Perkin Elmer SERIES 200 Micropump, 5 µL of samples were injected onto an Agilent Zorbax Rapid Resolution HT C18 2.1\*30mm, 1.8 micron column and analytes were separated using reverse-phase HPLC using the following gradient: Solvent A, 0.1% formic acid; Solvent B, 100% acetonitrile, 0.1% formic acid; 0-1 min 95% A, 1-8 min 5 to 100% B (linear

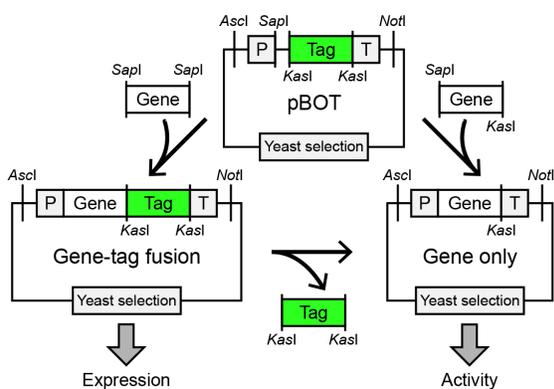
gradient), 8-9 min 100% B; 9-9.1 min 95% A, followed by a 2 min equilibration at 95% A. HPLC-grade methanol and acetonitrile were purchased from Fischer Scientific, and HPLC-grade water and formic acid were purchased from Fluka. Sample elution was followed by injection into the 7T-LTQ FT ICR mass spectrometer (Thermo Scientific) under the following conditions: resolution, 50000 at 400  $m/z$ ; scanning range, 150-500 AMU; source voltage, 4.9 kV; source temperature, 380°C; AGC target for full mass spectrum was set to  $1 \times 10^6$  ions. Retention time, exact mass (<2 ppm) of the  $[M+H]^+$  protonated monoisotopic ion, and in the case of nandinine and cheilanthifoline  $MS^2$  spectra were used to identify alkaloids. (*S*)-Scoulerine, (*S*)-stylophine, and sanguinarine were quantified using commercially available standards. (*S*)-Nandinine and (*S*)-cheilanthifoline were assumed to have similar ionization efficiencies to (*S*)-scoulerine and (*S*)-stylophine. *N*-methylated (*S*)-scoulerine, (*S*)-nandinine, (*S*)-cheilanthifoline and (*S*)-stylophine were also assumed to have similar ionization efficiencies to each other.

## 3.6 Supplemental materials for Chapter 3

### 3.6.1 Expression analysis of CYP719s by flow cytometry

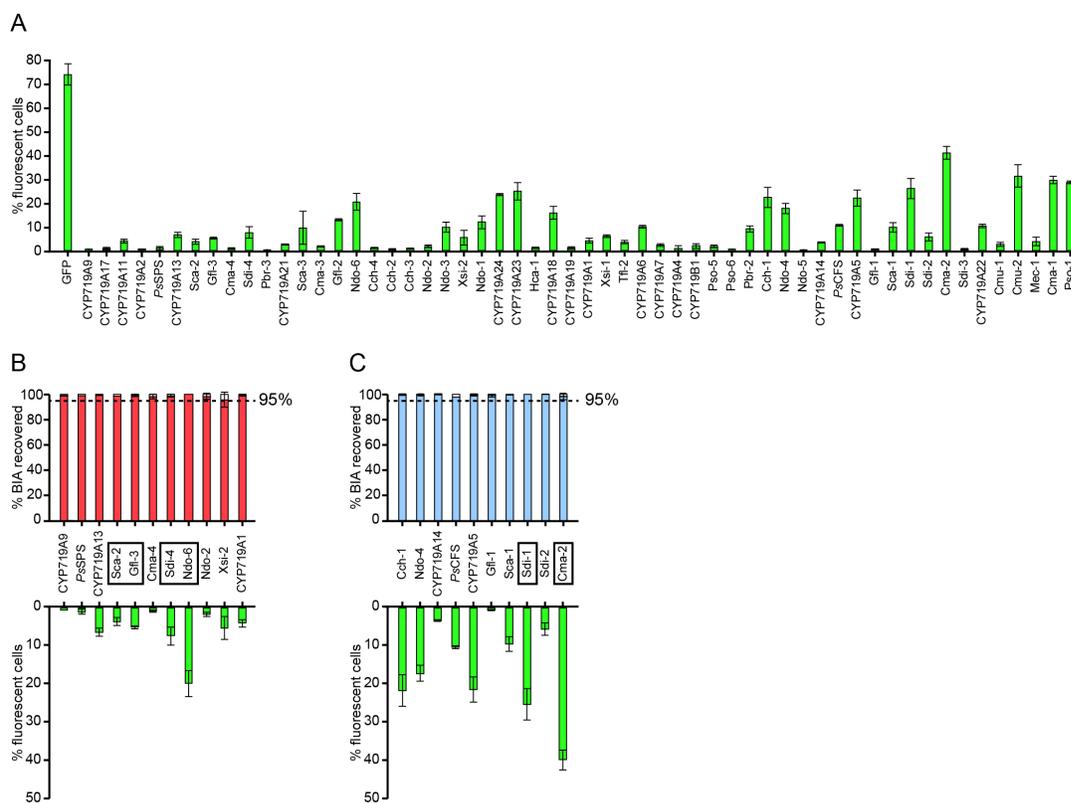
The pBOT vector system (Figure 3.8) made it possible to quickly assess gene expression in parallel to activity assay. CYP719-GFP fusions exhibited fluorescence as determined by percentage of fluorescent cells above background (Figure 3.9A). Fluorescence values varied from barely detectible to about half that of a GFP positive control. Across the entire set of CYP719s, fluorescence values were not correlated with enzyme activity on scoulerine. However, since not all CYP719s were predicted to have activity on scoulerine (Figure 3.3C), this is not unexpected. Comparison of CYP719s that converted >95% of supplemented scoulerine to nandinine or cheilanthifoline also resulted in a range of fluorescence values (Figure 3.9B and Figure 3.9C, respectively). CYP719s exhibiting low fluorescence as GFP fusions included PsSPS, the Ring A-

closing CYP719 used in our previous study that we sought to replace<sup>23</sup>. Conversely, high fluorescers also tended to be the sub-set of CYP719s selected for integration into the dihydrosanguinarine pathway (Figure 3.9B and Figure 3.9C, boxes) through the previously-described activity screens. This suggests that fluorescence values cannot be used in the absence of an activity assay, but are suggestive of overall performance in a pathway of interest. This also indicates that other fluorescent enzymes in the library with no detected activity on supplemented substrates are likely to have activities unrelated to the pathway of interest, which should prove useful for the optimization of other heterologous BIA synthesis pathways<sup>15,153</sup> and development of novel pathways.



**Figure 3.8. The pBOT vector system**

pBOT vectors, each with unique promoter, terminator and yeast selectable marker (grey) can be used for expression of tagged or untagged genes of interest depending on the restriction enzymes used to digest and ligate gene and vector (*SapI* or *SapI/KasI*, respectively). The tag (green) can also be removed *via* restriction digest.



**Figure 3.9. Qualitative and comparative analysis of gene expression and enzyme activity**

**(A)** The fluorescence of yeast transformed with plasmids harboring CYP719-GFP fusions was analyzed by flow cytometry. Yeast expressing GFP alone were included as a positive control, while yeast expressing no enzymes were used as a negative control for gating. **(B)** Comparison of CYP719 fluorescence and activity on scoulerine for CYP719s converting >95% of scoulerine to nandinine. **(C)** Comparison of CYP719 fluorescence and activity on scoulerine for CYP719s converting >95% of scoulerine to cheilanthifoline. CYP719s selected for combinatorial integration into the dihydrosanguinarine pathway are indicated in boxed text. Enzyme activity data is from Figure 3.3C (CYP719 activity on scoulerine).

## 3.6.2 Supplemental materials and methods

### 3.6.2.1 *The pBOT vector series*

A series of four yeast expression vectors, designated as pBOT, was designed to facilitate high-throughput gene expression analysis as well as tag cleavage for tag-free activity assays (Figure 3.8)<sup>185</sup>. Each gene expression cassette contains *SapI* sites between promoter and terminator. Genes are cloned without a stop codon into the *SapI* restriction site, whose overhangs were designed to avoid the introduction of extraneous DNA. Amplification of a gene of interest flanked by the sequences 5'-GCTCTTCTACA...GGCTGAAGAGC-3' (*SapI* sites underlined), followed by *SapI* digestion and ligation of gene and vector, results in scarless formation of the Kozak sequence 5'-AAACA-3' upstream of the gene's start codon. Downstream of the gene is a GFP tag separated from the gene of interest by a 12 amino acid flexible linker (GAAGSAAGSGEF). The GFP tag can be excised by restriction digest with *KasI*. pBOT vectors have unique selectable markers to allow co-transformation in yeast. If two genes to be co-transformed do share a selectable marker, the promoter-gene-terminator cassette can be moved to a different pBOT vector by restriction digest with *AscI/NotI*. Finally, multi-gene constructs can be constructed from pBOT vectors in a sequence-independent manner by the amplification of entire promoter-gene-terminator cassettes. Each pBOT vector has a unique combination of promoter/terminator flanking the gene of interest (Table 3.5) to allow the cloning of multi-gene constructs with methods like DNA Assembler<sup>162</sup> or Gibson Assembly<sup>186</sup> that require homology between DNA parts.

### 3.6.2.2 *Flow cytometry for expression analysis of CYP719s*

GFP translational fusions of CYP719s were used to qualitatively assess gene expression. Yeast cells harboring GFP translational fusions were inoculated in triplicate into 100  $\mu$ L of media in 96-well 2 mL deep-well plates (Grenier Bio-One) and incubated overnight. The following day,

900  $\mu$ L of fresh media was added (1:10 dilution) and cultures were incubated for an additional 6 hrs. Cells were pelleted by centrifugation for 5 min at 3,200 g and suspended in 1 mL dH<sub>2</sub>O, transferred to microcentrifuge tubes, and incubated at room temperature for 1 hr. Cells were vortexed briefly prior to analysis. Controls were cells not expressing GFP (negative) and cells expressing GFP without an enzyme fusion (positive). Cells were analyzed on the Accuri C6 flow cytometer (BD Biosciences) with excitation 488nm and emission 533nm. Single, live cells were gated by forward scatter height vs. side scatter height. A gate for fluorescence was determined based on the negative control.

### 3.6.3 Supplemental tables

**Table 3.3. Enzyme list for Chapter 3**

<b>Enzymes purchased as synthetic codon-optimized genes</b>				
<b>Name</b>	<b>Organism</b>	<b>Source</b>	<b>Accession #</b>	<b>Ref.</b>
CYP719A1	<i>Coptis japonica</i>	Genbank; Q948Y1.1	KX528927	150
CYP719A2	<i>Eschscholzia californica</i>	Genbank; Q50LH3.1	KX528938	151
CYP719A4	<i>Thalictrum flavum</i>	Genbank; AAU20771.1	KX528970	149
CYP719A5	<i>Eschscholzia californica</i>	Genbank; B5UAQ8.1	KX528936	152
CYP719A6	<i>Aquilegia formosa</i>	Cytochrome P450 Homepage	KX528921	179
CYP719A7	<i>Aquilegia formosa</i>	Cytochrome P450 Homepage	KX528922	179
CYP719A9	<i>Eschscholzia californica</i>	Genbank; BAG75114.1	KX528937	152
CYP719A11	<i>Eschscholzia californica</i>	Genbank; Q948Y1.1	KX528939	152
CYP719A13	<i>Argemone mexicana</i>	Genbank; B1NF19.1	KX528919	139
CYP719A14	<i>Argemone mexicana</i>	Genbank; B1NF20.1	KX528920	139
CYP719A17	<i>Eschscholzia californica</i>	Genbank; BAG75116.1	KX528940	152
CYP719A18	<i>Coptis japonica</i>	Genbank; BAF98470.1	KX528928	Unp
CYP719A19	<i>Coptis japonica</i>	Genbank; AB374408.1	KX528929	Unp
CYP719A21	<i>Papaver somniferum</i>	Genbank; AFB74615.1	KX528957	153
CYP719A22	<i>Nelumbo nucifera</i>	Cytochrome P450 Homepage	KX528952	187
CYP719A23	<i>Sinopodophyllum hexandrum</i>	Genbank; AGC29953.1	KX528969	128
CYP719A24	<i>Podophyllum peltatum</i>	Genbank; AGC29954.1	KX528955	128
CYP719A25	<i>Papaver somniferum</i>	PhytoMetaSyn; ADB89213.1	KX528958	23
CYP719B1	<i>Papaver somniferum</i>	Genbank; B1NF18.1	KX528959	154
CYP719-Cch-1	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528923	Chp. 3
CYP719-Cch-2	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528924	Chp. 3
CYP719-Cch-3	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528925	Chp. 3
CYP719-Cch-4	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528926	Chp. 3
CYP719-Cma-1	<i>Chelidonium majus</i>	PhytoMetaSyn	KX528930	Chp. 3
CYP719-Cma-2	<i>Chelidonium majus</i>	PhytoMetaSyn	KX528931	Chp. 3
CYP719-Cma-3	<i>Chelidonium majus</i>	PhytoMetaSyn	KX528932	Chp. 3
CYP719-Cma-4	<i>Chelidonium majus</i>	PhytoMetaSyn	KX528933	Chp. 3
CYP719-Cmu-1	<i>Cissampelos mucronata</i>	PhytoMetaSyn	KX528934	Chp. 3
CYP719-Cmu-2	<i>Cissampelos mucronata</i>	PhytoMetaSyn	KX528935	Chp. 3
CYP719-Gfl-1	<i>Glaucium flavum</i>	PhytoMetaSyn	KX528941	Chp. 3
CYP719-Gfl-2	<i>Glaucium flavum</i>	PhytoMetaSyn	KX528942	Chp. 3
CYP719-Gfl-3	<i>Glaucium flavum</i>	PhytoMetaSyn	KX528943	Chp. 3
CYP719-Hca-1	<i>Hydrastis canadensis</i>	PhytoMetaSyn	KX528944	Chp. 3
CYP719-Mec-1	<i>Menispermum canadense</i>	PhytoMetaSyn	KX528945	Chp. 3
CYP719-Ndo-1	<i>Nandina domestica</i>	PhytoMetaSyn	KX528946	Chp. 3
CYP719-Ndo-2	<i>Nandina domestica</i>	PhytoMetaSyn	KX528947	Chp. 3
CYP719-Ndo-3	<i>Nandina domestica</i>	PhytoMetaSyn	KX528948	Chp. 3
CYP719-Ndo-4	<i>Nandina domestica</i>	PhytoMetaSyn	KX528949	Chp. 3
CYP719-Ndo-5	<i>Nandina domestica</i>	PhytoMetaSyn	KX528950	Chp. 3
CYP719-Ndo-6	<i>Nandina domestica</i>	PhytoMetaSyn	KX528951	Chp. 3
CYP719-Pbr-2	<i>Papaver bracteatum</i>	PhytoMetaSyn	KX528953	Chp. 3
CYP719-Pbr-3	<i>Papaver bracteatum</i>	PhytoMetaSyn	KX528954	Chp. 3
CYP79-Pso-1	<i>Papaver somniferum</i>	PhytoMetaSyn	KX528956	Chp. 3
CYP719-Pso-5	<i>Papaver somniferum</i>	PhytoMetaSyn	KX528960	Chp. 3
CYP719-Pso-6	<i>Papaver somniferum</i>	PhytoMetaSyn	KX528961	Chp. 3
CYP719-Sca-1	<i>Sanguinaria canadensis</i>	PhytoMetaSyn	KX528962	Chp. 3
CYP719-Sca-2	<i>Sanguinaria canadensis</i>	PhytoMetaSyn	KX528963	Chp. 3

CYP719-Sca-3	<i>Sanguinaria canadensis</i>	PhytoMetaSyn	KX528964	Chp. 3
CYP719-Sdi-1	<i>Stylophorum diphyllum</i>	PhytoMetaSyn	KX528965	Chp. 3
CYP719-Sdi-2	<i>Stylophorum diphyllum</i>	PhytoMetaSyn	KX528966	Chp. 3
CYP719-Sdi-3	<i>Stylophorum diphyllum</i>	PhytoMetaSyn	KX528967	Chp. 3
CYP719-Sdi-4	<i>Stylophorum diphyllum</i>	PhytoMetaSyn	KX528968	Chp. 3
CYP719-Tfl-2	<i>Thalictrum flavum</i>	PhytoMetaSyn	KX528971	Chp. 3
CYP719-Xsi-1	<i>Xanthorhiza simplicissima</i>	PhytoMetaSyn	KX528972	Chp. 3
TjCNMT	<i>Thalictrum flavum</i>	Genbank; AAU207766.1	KX528986	126,149
EcTNMT	<i>Eschscholzia californica</i>	Genbank; ACO90222.1	KX528978	126
PbTNMT2	<i>Papaver bracteatum</i>	Genbank; ACO90236.1	KX528982	126
TjPNMT	<i>Thalictrum flavum</i>	Genbank; ACO90251.1	KX528987	126
NMT-Ame-1	<i>Argemone mexicana</i>	PhytoMetaSyn	KX528973	Chp. 3
NMT-Cch-1	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528974	Chp. 3
NMT-Cch-2	<i>Corydalis cheilanthifolia</i>	PhytoMetaSyn	KX528975	Chp. 3
NMT-Cma-1	<i>Chelidonium majus</i>	PhytoMetaSyn	KX528976	Chp. 3
NMT-Eca-1	<i>Eschscholzia californica</i>	PhytoMetaSyn	KX528977	Chp. 3
NMT-Hca-1	<i>Hydrastis canadensis</i>	PhytoMetaSyn	KX528979	Chp. 3
NMT-Pbr-1	<i>Papaver bracteatum</i>	PhytoMetaSyn	KX528980	Chp. 3
NMT-Pbr-2	<i>Papaver bracteatum</i>	PhytoMetaSyn	KX528981	Chp. 3
NMT-Pbr-4	<i>Papaver bracteatum</i>	PhytoMetaSyn	KX528983	Chp. 3
NMT-Sca-1	<i>Sanguinaria Canadensis</i>	PhytoMetaSyn	KX528984	Chp. 3
NMT-Sdi-1	<i>Stylophorum diphyllum</i>	PhytoMetaSyn	KX528985	Chp. 3
<b>Other enzymes used in this study</b>				
Ps6OMT	<i>Papaver somniferum</i>		KF554144.1	
Ps4'OMT2	<i>Papaver somniferum</i>		KF661327.1	
PsCNMT	<i>Papaver somniferum</i>		KF661326.1	
PsBBE	<i>Papaver somniferum</i>		AF025430.1 <sup>a</sup>	
CYP719A20 (PsSPS)	<i>Papaver somniferum</i>		KF481962.1	
PsTNMT	<i>Papaver somniferum</i>		AA79117.1	
PsMSH	<i>Papaver somniferum</i>		KC154003.1	
EcP6H	<i>Eschscholzia californica</i>		AB598834.1	
PsCPR	<i>Papaver somniferum</i>		KF661328.1	

Unp.: deposited online

Chp.3 refers to Chapter 3 of this work

a: First 12 amino acids are truncated, as described in <sup>23</sup>

**Table 3.4. Primer list for Chapter 3**

For construction of the pBOT vectors		
Name	Sequence (5'-3')	Description
URA3_SDM_F	TAAAGCCGATAACAAAATCTTTGTCGCTCTTG	Site-directed mutagenesis of SapI site in <i>URA3</i> cassette
URA3_SDM_R	GGGTACTGTTGACATTGCCAAGAGCGACAA	
CEN6ARS4_SDM_F	GCATTAATACCTGAGAGCAGGAAGAGG	Site-directed mutagenesis of SapI site in CEN6/ARS4 <sup>ori</sup>
CEN6ARS4_SDM_R	TACTACCTTTTATCTTCCTCTTCC	
LV5_(KanMX)_F	CCTCTTTATATTACATCAAATAAGAAAATAA TTATAACACAGATCCGCGGCCGC	Amplification of <i>Kan</i> <sup>R</sup> cassette from pGC964
LX_(KanMX)_R	TGACCTAGACTGGCTTTGATCTTTAATTACAC ACTTATCCTCGACAACCCTTAATATAACTT	
LX_(His)_F	GGATAAGTGTGTAATTAAGATCAAAGCCAG TCTAGGTCATATGCGTCGAGTTCAAGAGAAA	Amplification of <i>HIS3</i> cassette from pGC964
LY_(His)_R	CTGTTGCCTGACGTGAGTGGTGCCTTTGATGA TGAGATACCGTTTTAAGAGCTTGGTG	
LX_(Trp)_F	GGATAAGTGTGTAATTAAGATCAAAGCCAG TCTAGGTCATATGAGGCAAGTGCACAAACAA T	Amplification of <i>TRP1</i> cassette from pGC965
LY_(Trp)_R	CTGTTGCCTGACGTGAGTGGTGCCTTTGATGA TGAGATAACAACGACATTACTATATATAATA TAGGAAGC	
LX_(Leu)_F	GGATAAGTGTGTAATTAAGATCAAAGCCAG TCTAGGTCATATGCCTGATTCAAGAAATATCT TG	Amplification of <i>LEU2</i> cassette from pGC966
LY_(Leu)_R	CTGTTGCCTGACGTGAGTGGTGCCTTTGATGA TGAGATACTAGTTTCATGATTTTCTGTTACAC	
LX_(Ura)_F	GGATAAGTGTGTAATTAAGATCAAAGCCAG TCTAGGTCATATGGGTAATAACTGATATAATT AAATTGAAG	Amplification of <i>URA3</i> cassette from pGC967
LY_(Ura)_R	CTGTTGCCTGACGTGAGTGGTGCCTTTGATGA TGAGATACATTCATCATTTTTTTTTTATTCTT	
LY_(CEN6ARS4)_F	GTATCTCATCATCAAAGGCCACTCACGTCA GGCAACAGGGACGGATCGCTTG	Amplification of CEN6/ARS4 <sup>ori</sup> from pGC964
LZ_(CEN6ARS4)_R	CTGACGTCGGTAAAGTAGGAGTGCTGCAATA GGTCTTAAGGTCCTTTTCATCACGT	
LZ_(Ecoli_unit)_F	TTAAGACCTATTGCAGACACTCCTACTTTACC GACGTCAGCAGGTGGCACTTTTCG	Amplification of <i>pMB1</i> <sup>ori</sup> and <i>Amp</i> <sup>R</sup> from pBluescript
LV3_(Ecoli_unit)_R	GCATTTTTATTATATAAGTTGTTTTATTTCAGAG TATTCCTGGCGCGCCCGGCTTGCTGGCGTT	
LV3_(P <sub>TEF1</sub> )_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCATAGCTTCAAAATGTTTCTACTC	Amplification of P <sub>TEF1</sub> with 5' linker and 3' insertion of <i>SapI</i> site and Kozak sequence
P <sub>TEF1</sub> _(SapIstuffer)_R	TTAATTATTTCTCTTCCTTTTATAATAAATTTT CTAGGCTCTTCATGTTTATTAACACTTAGATT AGATTGCTATG	
LV3_(P <sub>TDH3</sub> )_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCTCGAGTTTATCATTATCAATACT	Amplification of P <sub>TDH3</sub> with 5' linker and 3' insertion of <i>SapI</i> site and Kozak sequence
P <sub>TDH3</sub> _(SapIstuffer)_R	ACCAAGAACTTAGTTTCGAAAACATGAAGAG CCTAGAAAATTTATTATAAAAGGAAGAGAAA TAATTAA	
LV3_(P <sub>FBA1</sub> )_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCATCCAACCTGGCACCG	Amplification of P <sub>FBA1</sub> with 5' linker and 3'

P <sub>FBA1</sub> _(SapIstuffer)_R	TTAATTATTTCTCTTCCTTTTATAATAAATTTT CTAGGCTCTTCATGTTTTATGTATTACTTGGTT ATGGTTATATAT	insertion of <i>SapI</i> site and Kozak sequence
LV3_(P <sub>PMAl</sub> )_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCACAGGCATTGCTGG	Amplification of P <sub>PMAl</sub> with 5' linker and 3' insertion of <i>SapI</i> site and Kozak sequence
P <sub>PMAl</sub> _(SapIstuffer)_R	TTAATTATTTCTCTTCCTTTTATAATAAATTTT CTAGGCTCTTCATGTTTTTTGATAATTAATCT TTCTTATCTT	
SapIstuffer_(GFP)_F	AGCCTAGAAAATTTATTATAAAAGGAAGAGA AATAATTAACAGCTCTTCTGGCGCCGCTGGC TCCGCTG	Insertion of GFP downstream of Kozak sequence
(GFP) R	TTAGGCGCCTTATTTGTATAGTTCATCCATG	
GFP_(T <sub>PGII</sub> )_F	CACATGGCATGGATGAACTATACAAATAAGG CGCCTAAAACAAATCGCTCTTAAATATATAC	Amplification of T <sub>PGII</sub> with linker
LV5_(T <sub>PGII</sub> )_R	TGTTATAATTATTTTCTTATTTTGATGTAATAT AAAGAGGGGTATACTGGAGGCTTCAT	
GFP_(T <sub>CYC1</sub> )_F	CACATGGCATGGATGAACTATACAAATAAGG CGCCTAATCATGTAATTAGTTATGTCACG	Amplification of T <sub>CYC1</sub> with linker
LV5_(T <sub>CYC1</sub> )_R	TGTTATAATTATTTTCTTATTTTGATGTAATAT AAAGAGGGCAAATTAAGCCTTCG	
GFP_(T <sub>ADH1</sub> )_F	CACATGGCATGGATGAACTATACAAATAAGG CGCCTAATGGACTTCTTCGCCA	Amplification of T <sub>ADH1</sub> with linker
LV5_(T <sub>ADH1</sub> )_R	TGTTATAATTATTTTCTTATTTTGATGTAATAT AAAGAGGGCATGCCGGTAGAG	
GFP_(T <sub>TPH1</sub> )_F	CACATGGCATGGATGAACTATACAAATAAGG CGCCTAAGATTAATATAATTATATAAAAATAT TATCTTCTTT	Amplification of T <sub>TPH1</sub> with linker
LV5_(T <sub>TPH1</sub> )_R	TGTTATAATTATTTTCTTATTTTGATGTAATAT AAAGAGGCTATATAACAGTTGAAATTTGAAT AAG	
<b>For construction of the pYES vectors harboring dihydrosanguinarine pathway enzymes</b>		
<b>Name</b>	<b>Sequence (5'-3')</b>	<b>Description</b>
pYES F	CCTGCATTAATGAATCGGC	Amplification of pYES
pYES R	ACTAGTGGATCATCCCCAC	
pYES2:C1 F	CCGCCGCGCTTAATGGGGCGCTACAGGGCGC GTGGGGATGATCCACTAGTGAGACTGCAGCA TTACTTTGAGAAG	Amplification of constructs with C1 and C6 linkers with homology to pYES
pYES2:C6 R	CAATACGCAAACCGCCTCTCCCCGCGCGTTGG CCGATTCATTAATGCAGGACAACCTCATGGTGA TGTGATTGCC	
C1 TDH3p F	GAGACTGCAGCATTACTTTGAGAAGTCGAGTT TATCATTATCAATAC	Amplification of P <sub>TDH3</sub> with C1 linker and homology to CNMT
TDH3p-K-CNMT R	GCTCTTCCTTTGCTTTCAGTTGCATTGTTTTTC GAAACTAAGTTCCTGGTGTTTTAAAC	
TDH3p-K-CNMT F	GTTTTAAAACACCAAGAAGCTTAGTTTCGAAAA ACAATGCAACTGAAAGCAAAGGAAGAGC	Amplification of CNMT with homology to P <sub>TDH3</sub> and T <sub>PGII</sub>
CNMT-PGII t R	GGTATATATTTAAGAGCGATTTGTTTTACTTTT TCTTAAAGAGTAGATGAGTTAAC	
CNMT-PGII t F	CTCATCTACTCTTTAAGAAAAAGTAAAACAAA TCGCTCTTAAATATATACCTAAAG	Amplification of T <sub>PGII</sub> with C6 linker and homology to CNMT
C6 PGII t R	ACAATCATGGTGATGTGATTGCCGGTATACT GGAGGCTTCATGAG	
C1 TDH3p F	GAGACTGCAGCATTACTTTGAGAAGTCGAGTT TATCATTATCAATAC	

TDH3p-K-4OMT2 R	GGCTTCGCATCCAAGGAACCCATTGTTTTTCG AACTAAGTTCTTGGTGTTTTAAACT	Amplification of P <sub>TDH3</sub> with C1 linker and homology to 4'OMT2
TDH3p-K-4OMT2 F	AGTTTTAAACACCAAGAAGCTTAGTTTCGAAA ACAATGGGTTCTTGGATGCG	Amplification of 4'OMT2 with homology to P <sub>TDH3</sub> and T <sub>ADH1</sub>
4OMT2-ADH1t R	GACCAAACCTCTGGCGAAGAAGTCCATTATG GAAAAGCTTCTATAACAGATTGTATTG	
4OMT2-ADH1t F	GAGCAATACAATCTGTTATAGAAGCTTTTCCA TAATGGACTTCTTCGCCAGAGGTTTG	Amplification of T <sub>AHD1</sub> with C6 linker and homology to 4'OMT2
C6 ADH1t R	ACAACCTCATGGTATGTGATTGCCGCATGCCG GTAGAGGTGTGG	
C1 PMA1p	GAGACTGCAGCATTACTTTGAGAAGACAGGC ATTGCTGGGATCAC	Amplification of P <sub>PMA1</sub> with C1 linker and homology to 6OMT
PMA1p-K-6OMT R	TTGATACTGTTTCCATTGTTTTTTGATAATTA AATCTTTCTTATCTTCTTATTCTTTTC	
PMA1p-K-6OMT F	GAAGATAAGAAAGATTTAATTATCAAAAAA CAATGGAAACAGTATCAAAGATCGAC	Amplification of 6OMT with homology to P <sub>PMA1</sub> and T <sub>CYC1</sub>
6OMT-CYC1t R	GAATGTAAGCGTGACATAACTAATTACATGAT TAATATGGATAGGCTTCGATCACG	
6OMT-CYC1t F	CGTGATCGAAGCCTATCCATATTAATCATGTA ATTAGTTATGTCACGCTTACATTC	Amplification of T <sub>CYC1</sub> with C6 linker and homology to 6OMT
C6 CYC1t R	ACAACCTCATGGTATGTGATTGCCGCAAATTA AAGCCTTCGAGCGTC	
<b>For construction of pGREG vectors harboring dihydrosanguinarine pathway enzymes</b>		
<b>Name</b>	<b>Sequence (5'-3')</b>	<b>Description</b>
C1:506	TAACCCTCACTAAAGGGAACAAAAGCTGGAG CTCGTTTAAACGGCGCGCCGAGACTGCAGCAT TACTTTGAGAAG	Amplification of constructs with C1 and C6 linkers with homology to pGREG
C6:506	ATAACTTCGTATAATGTATGCTATACGAAGTT ATTAGGTACCGCGGCCGCACAACCTCATGGTGA TGTGATTGCC	
C1 TDH3 F	GAGACTGCAGCATTACTTTGAGAAGTCGAGTT TATCATTATCAATAC	Amplification of P <sub>TDH3</sub> with C1 linker and homology to PsCFS
TDH3p-CFS R	CAACCAAAATGTCACCTCCATTGTTTTTCGAA ACTAAGTTCTTGGTGTTTTAAAC	
TDH3p-CFS F	TTTTAAACACCAAGAAGCTTAGTTTCGAAAAA CAATGGAGGTGACATTTTGGTTGATAAC	Amplification of PsCFS with homology to P <sub>TDH3</sub> and T <sub>CYC1</sub>
CFS-CYC1t R	GTAAGCGTGACATAACTAATTACATGATTAAT GGATACGAGGAGTAATTTTGGC	
CFS-CYC1t F	GCCAAAATTACTCCTCGTATCCATTAATCATG TAATTAGTTATGTCACGCTTAC	Amplification of T <sub>CYC1</sub> with C6 linker and homology to PsCFS
C6 CYC1t R	ACAACCTCATGGTATGTGATTGCCGCAAATTA AAGCCTTCGAGCGTC	
C1 TDH3p F	GAGACTGCAGCATTACTTTGAGAAGTCGAGTT TATCATTATCAATAC	Amplification of P <sub>TDH3</sub> with C1 linker
TDH3p R	TCGAAACTAAGTTCTTGGTGTTTTAAAC	
TDH3p-P450R F	GTTTTAAACACCAAGAAGCTTAGTTTCGAAAA ACAATGGGGTCAAACAACCTGGC	Amplification of CPR with homology to P <sub>TDH3</sub> and T <sub>CYC1</sub>
P450R-CYC1t R	GTAAGCGTGACATAACTAATTACATGATTACC ATACATCTCTCAAGTATCTCTC	
P450R-CYC1t F	GAGAGATACTTGAGAGATGTATGGTAATCAT GTAATTAGTTATGTCACGCTTAC	Amplification of T <sub>CYC1</sub> with C6 linker and homology to PsCPR
C6 CYC1t R	ACAACCTCATGGTATGTGATTGCCGCAAATTA AAGCCTTCGAGCGTC	

C1:H1 F	CTCATGGCGGGGGTCGGAATGATTAAGAAA GGGGCTGTGGGCGAGATTGGAGACTGCAGCA TACTTTGAGAAG	Amplification of C1,C6 flanked cassettes with linkers for multi-cassette assembly
C6:H1 R	CAATCTCGCCACAGCCCCTTTCTTTAATCATT CCGACCCCCGCCATGAGACAACTCATGGTGAT GTGATTGCC	
C1 PMA1p F	GAGACTGCAGCATTACTTTGAGAAGACAGGC ATTGCTGGGATCAC	Amplification of P <sub>PMA1</sub> with C1 linker and homology to P <sub>PMA1</sub>
PMA1p-BBE R	CATTAACATCACCCATTGTTTTTTGATAATTA AATCTTTCTTATCTTCTTATTCTTTTC	
PMA1p-BBE F	GAAGATAAGAAAGATTTAATTATCAAAAAA CAATGGGTGATGTTAATGATAATCTCCTC	Amplification of BBE with homology to P <sub>PMA1</sub> and T <sub>PGII</sub>
BBE-PGI1t R	CTTAGGTATATATTTAAGAGCGATTGTCT ACAATTCCTTCAACATGTAAATTTCC	
BBE-PGI1t F	GGAAATTTACATGTTGAAGGAATTGTAGAAC AAATCGCTCTTAAATATATACCTAAAG	Amplification of T <sub>PGII</sub> with C6 linker and homology to PsBBE
C6 PGI1t R	CAACTCATGGTGATGTGATTGCCGGTATACT GGAGGCTTCATGAG	
<b>Moving Ring A-closing CYP719s from pBOT-Trp to pBOT-TPP</b>		
<b>Name</b>	<b>Sequence (5'-3')</b>	<b>Description</b>
P <sub>PMA1</sub> _(CYP719A14)_F	GAATAAGAAGATAAGAAAGATTTAATTATCA AAAAACAATGGAAGAAAGATTATGACC	Amplification of CYP719A14 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719A14)_R	TCATGAATTCGCCAGAACCAGCAGCGGAGCC AGCGGCGCCATTCTAGGAACAATACCG	
P <sub>PMA1</sub> _(CYP719Cma4)_F	AATAAGAAGATAAGAAAGATTTAATTATCAA AAAACAATGGAAATGAATCCTTTATTG	Amplification of CYP719- Cma-4 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719Cma4)_R	CATGAATTCGCCAGAACCAGCAGCGGAGCCA GCGGCGCCAACATTTCTACCAGCGATTCC	
P <sub>PMA1</sub> _(CYP719Gfl3)_F	AATAAGAAGATAAGAAAGATTTAATTATCAA AAAACAATGTTATTGATCTTGATCATGG	Amplification of CYP719- Gfl-3 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719Gfl3)_R	CATGAATTCGCCAGAACCAGCAGCGGAGCCA GCGGCGCCGCATCTGGAGGTAATTCTAG	
P <sub>PMA1</sub> _(CYP719Ndo6)_F	AATAAGAAGATAAGAAAGATTTAATTATCAA AAAACAATGGAAATGAAGTCTTCTGTC	Amplification of CYP719- Ndo-6 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719Ndo6)_R	ATGAATTCGCCAGAACCAGCAGCGGAGCCAG CGGCGCCACATCTTGGAGTAATACGGG	
P <sub>PMA1</sub> _(CYP719Sca2)_F	AATAAGAAGATAAGAAAGATTTAATTATCAA AAAACAATGGAAATTCGAAAAAATTATG	Amplification of CYP719- Sca-2 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719Sca2)_R	CTCATGAATTCGCCAGAACCAGCAGCGGAGC CAGCGGCGCCGGATCTTGGGACGATTCC	
P <sub>PMA1</sub> _(CYP719Sdi4)_F	AATAAGAAGATAAGAAAGATTTAATTATCAA AAAACAATGGAAATCCTTGATTGATTTC	Amplification of CYP719- Sdi-4 for expression under P <sub>PMA1</sub> and T <sub>PGII</sub>
GFP_(CYP719Sdi4)_R	ATGAATTCGCCAGAACCAGCAGCGGAGCCAG CGGCGCCACAACGTGGAGTGATTCTAG	
P <sub>PMA1</sub> _(PsSPS)_F	TAAGAAGATAAGAAAGATTTAATTATCAAAA ACAATGGAAATTATCGATTACTACCTC	Amplification of PsSPS for expression under P <sub>PMA1</sub> and T <sub>PGI</sub>
GFP_(PsSPS)_R	ATGAATTCGCCAGAACCAGCAGCGGAGCCAG CGGCGCCAACCTTGGGACTATCCTCG	
<b>Generating the Flagfeldt-targeted pCAS-tyr vectors</b>		
<b>Name</b>	<b>Sequence (5'-3')</b>	<b>Description</b>
pCas-tyr_F	TAGGTCTAGAGATCTGTTTAGCTTG	Amplification of pCas-tyr with no gRNA cassette
pCas-tyr_R	GCATTTAAGCATAAACACGC	
gRNA_F	CACCTATATCTGCGTGTTGC	Outer gRNA cassette primers
gRNA_R	TCAAGACTGTCAAGGAGG	

16UPgRNA_F	TTTTCCGACAATCAAATATTGTTTTAGAGCTA GAAATAGCAAGT	Inner gRNA cassette primers for 16 UP
16UPgRNA_R	AATATTTGATTGTCGGAAAAAAGTCCCATTC GCCACC	
16DOWNgRNA_F	TCATCAAAGAGACATTTTTGTTTTAGAGCTA GAAATAGCAAGT	Inner gRNA cassette primers for 16 DOWN
16DOWNgRNA_R	AAAAATGTCTCTTTTGATGAAAAGTCCCATTC GCCACC	
20UPgRNA_F	AAAATTCTCTCTGAGGATATGTTTTAGAGCTA GAAATAGCAAGT	Inner gRNA cassette primers for 20 UP
20UPgRNA_R	ATATCCTCAGAGAGAATTTTAAAGTCCCATTC GCCACC	
20DOWNgRNA_F	GTTAGAGCTGTTACAAGTTAAAAATTCTCTCT GAGGATATGTTTTAGAGCTAGAAATAGCAAG T	Inner gRNA cassette primers for 20 DOWN
20DOWNgRNA_R	TAACTTGTAACAGCTCTAACAAAGTCCCATTC GCCACC	
21UPgRNA_F	TTTATACATTTACATGTACGTTTTAGAGCTA GAAATAGCAAGT	Inner gRNA cassette primers for 21 UP
21UPgRNA_R	GTACATGTGAAATGTATAAAAAAGTCCCATTC GCCACC	
21DOWNgRNA_F	TATCCATGGCCTCTTAGTTGTTTTAGAGCTAG AAATAGCAAGT	Inner gRNA cassette primers for 21 DOWN
21DOWNgRNA_R	AACTAAGAGGCCATGGAATAAAAGTCCCATTC GCCACC	
<b>For amplifying gene cassettes for CRISPR-directed genomic integration</b>		
(16UP) F	TTCGTGAAACACGTGGGATACC	Amplification of Flagfeldt site 16 UP with linkers
LV3_(16UP)_R	GCATTTTTATTATATAAGTTGTTTTATTCAGAG TATTCCTGGCGCGCCTCCGTTAATTCGGGTT	
LV5_(16DOWN)_F	CCTCTTTATATTACATCAAATAAGAAAATAA TTATAACACAGATCCGCGGCCGCTGCCTACGC AACACTTTAG	Amplification of Flagfeldt site 16 DOWN with linkers
(16DOWN) R	TTGTTGGGATTCCATTGTGATTAAGG	
(20UP) F	GCCAGGCGCCTTTATATCAT	Amplification of Flagfeldt site 20 UP with linkers
LV3_(20UP)_R	GCATTTTTATTATATAAGTTGTTTTATTCAGAG TATTCCTGGCGCGCCTTTGCGAAACCCTATGC	
LV5_(20DOWN)_F	CCTCTTTATATTACATCAAATAAGAAAATAA TTATAACACAGATCCGCGGCCGCAATGGAAG GTCGGGAT	Amplification of Flagfeldt site 20 DOWN with linkers
(20DOWN) R	ATAAAGCAGCCGCTACCAA	
(21UP) F	AAAGGAGGTGCACGCATTATGG	Amplification of Flagfeldt site 21 UP with linkers
LV3_(21UP)_R	GCATTTTTATTATATAAGTTGTTTTATTCAGAG TATTCCTGGCGCGCCTTCCAAGGAGGTGAAGA AC	
LV5_(21DOWN)_F	CCTCTTTATATTACATCAAATAAGAAAATAA TTATAACACAGATCCGCGGCCGCGATGGGAC GTCAGCAC	Amplification of Flagfeldt site 21 DOWN with linkers
(21DOWN) R	TCAAGACACTCCGGTATTAC	
LV3_(6OMT)_F	AGGAATACTCTGAATAAAACAATTATATAAT AAAAATGCTCGAGTTTATCATTATCAATACT	Amplification of 6OMT gene cassette with linkers
6OMT_(LTP1)_R	CATCTGTCCACTGTATGTTTATGCGTTTTCTA AGACCGGGCAAATTAAGCCTTCG	
LTP1_(4'OMT2)_F	CCGGTCTTAGAAAACGCATAAACATACAAGT GGACAGATGATCCAAGTGGCACCG	Amplification of 4'OMT2 gene cassette with linkers

4'OMT2_(LTP2)_R	TGATTGTTAAGAATGTTTAGTGATGAGTATGT TATTTAAGGCATGCCGGTAGAG	
LTP2_(CNMT)_F	CTTAAATAACATACTCATCACTAAACATTCTT ACAATCAACATGCGACTGGGT	Amplification of CNMT gene cassette with linkers
CNMT_(LV5)_R	TGTTATAATTATTTTCTTATTTTGTATGTAATAT AAAGAGGGGTATACTGGAGGCTTCAT	
LV3_(CPR)_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCTCGAGTTTATCATTATCAATACT	Amplification of CPR gene cassette with linkers
CPR_(LTP1)_R	CATCTGTCCACTTGTATGTTTATGCGTTTTCTA AGACCGGGCAAATTAAGCCTTCG	
LTP1_(BBE)_F	CCGGTCTTAGAAAACGCATAAACATACAAGT GGACAGATGACAGGCATTGCTGG	Amplification of BBE gene cassette with linkers
BBE_(LV5)_R	TGTTATAATTATTTTCTTATTTTGTATGTAATAT AAAGAGGGGTATACTGGAGGCTTCAT	
LV3_(P6H)_F	AGGAATACTCTGAATAAAACAACCTTATATAAT AAAAATGCACATGCGACTGGGT	Amplification of P6H gene cassette with linkers
P6H_(LTP1)_R	CATCTGTCCACTTGTATGTTTATGCGTTTTCTA AGACCGGGCAAATTAAGCCTTCG	
LTP1_(MSH)_F	CCGGTCTTAGAAAACGCATAAACATACAAGT GGACAGATGTGCGAGTTTATCATTATCAATACT	Amplification of MSH gene cassette with linkers
MSH_(LTP2)_R	TGATTGTTAAGAATGTTTAGTGATGAGTATGT TATTTAAGGCATGCCGGTAGAG	
LTP2_(TNMT)_F	CTTAAATAACATACTCATCACTAAACATTCTT ACAATCAATCCAACCTGGCACCG	Amplification of TNMT gene cassette with linkers
TNMT_(LV5)_R	TGTTATAATTATTTTCTTATTTTGTATGTAATAT AAAGAGGGGTATACTGGAGGCTTCAT	

**Table 3.5. Plasmid list for Chapter 3**

<b>Base vectors</b>				
<b>Number</b>	<b>Name</b>	<b>Description</b>	<b>Ref.</b>	
pGC964	pGREG-His	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>HIS3</i>	23	
pGC965	pGREG-Trp	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>TRP1</i>	23	
pGC966	pGREG-Leu	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>LEU2</i>	23	
pGC967	pGREG-Ura	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>URA3</i>	23	
pGC1312	pBOT-His	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>HIS3</i> , P <sub>TEF1</sub> -GFP-T <sub>PGII</sub>	Chp. 3	
pGC1313	pBOT-Trp	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>TRP1</i> , P <sub>TDH3</sub> -GFP-T <sub>CYC1</sub>	Chp. 3	
pGC1314	pBOT-Leu	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>LEU2</i> , P <sub>FBA1</sub> -GFP-T <sub>ADH1</sub>	Chp. 3	
pGC1315	pBOT-Ura	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>URA3</i> , P <sub>PMA1</sub> -GFP-T <sub>TPH1</sub>	Chp. 3	
pGC1381	pBOT-TPP	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>TRP1</i> , P <sub>PMA1</sub> -GFP-T <sub>PGII</sub>	Chp. 3	
pGC1441	pBOT-LTC	CEN6/ARS4 <sup>ori</sup> , <i>pMB1</i> <sup>ori</sup> , <i>Amp</i> <sup>R</sup> , <i>Kan</i> <sup>R</sup> , <i>LEU2</i> , P <sub>TDH3</sub> -GFP-T <sub>CYC1</sub>	Chp. 3	
pGC84	pYES2	2 $\mu$ <sup>ori</sup> , <i>pUC</i> <sup>ori</sup> , <i>URA3</i> , <i>Amp</i> <sup>R</sup> , P <sub>GAL1</sub> -T <sub>CYC1</sub>	Invitrogen	
pGC588	pYES-Trp	2 $\mu$ <sup>ori</sup> , <i>pUC</i> <sup>ori</sup> , <i>TRP1</i> , <i>Amp</i> <sup>R</sup> , P <sub>GAL1</sub> -T <sub>CYC1</sub>	142	
pGC587	pYES-Leu	2 $\mu$ <sup>ori</sup> , <i>pUC</i> <sup>ori</sup> , <i>LEU2</i> , <i>Amp</i> <sup>R</sup> , P <sub>GAL1</sub> -T <sub>CYC1</sub>	20	
pGC1421	pCas-Tyr	2 $\mu$ <sup>ori</sup> , <i>ColEI</i> <sup>ori</sup> , <i>Kan</i> <sup>R</sup> , P <sub>RNR2</sub> -Cas9-T <sub>CYC1</sub> , tRNA <sup>tyr</sup> -3'HDV-gRNA-T <sub>SNR52</sub>	183	
<b>NMT assays</b>				
<b>Number</b>	<b>Base vector</b>	<b>Gene in expression cassette</b>	<b>Description</b>	<b>Ref.</b>
pGC633	pYES-Leu	Ps6OMT	<i>pGC587::C1-P<sub>PMA1</sub>-6OMT-T<sub>CYC1</sub>-C6</i>	Chp. 3
pGC635	pYES-Trp	Ps4' OMT2	<i>pGC588::C1-P<sub>TDH3</sub>-4' OMT2-T<sub>ADH1</sub>-C6</i>	Chp. 3
pGC652	pYES-Trp	PsCNMT	<i>pGC588::C1-P<sub>TDH3</sub>-CNMT-T<sub>PGII</sub>-C6</i>	Chp. 3
pGC717	pGREG-His	PsTNMT, CPR	<i>pGC964::C1-P<sub>FBA1</sub>-CPR-T<sub>CYC1</sub>-C6-H1-C1-P<sub>TDH3</sub>-TNMT-T<sub>PGII</sub>-C6</i>	23
pGC1442	pBOT-His	NMT-Ame-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Ame-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1443	pBOT-His	NMT-Cch-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cch-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1444	pBOT-His	NMT-Cch-2	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cch-2-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1445	pBOT-His	NMT-Cma-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cma-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1446	pBOT-His	NMT-Eca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Eca-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1447	pBOT-His	NMT-Hca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Hca-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1448	pBOT-His	NMT-Pbr-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1449	pBOT-His	NMT-Pbr-2	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-2-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1450	pBOT-His	NMT-Pbr-4	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-4-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1451	pBOT-His	NMT-Sca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Sca-1-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1452	pBOT-His	TfCNMT	<i>pGC1312::P<sub>TEF1</sub>-TfCNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1453	pBOT-His	TfPNMT	<i>pGC1312::P<sub>TEF1</sub>-TfPNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1454	pBOT-His	EcTNMT	<i>pGC1312::P<sub>TEF1</sub>-EcTNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1455	pBOT-His	PbTNMT	<i>pGC1312::P<sub>TEF1</sub>-PbTNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1456	pBOT-His	PbTNMT2	<i>pGC1312::P<sub>TEF1</sub>-PbTNMT2-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1457	pBOT-His	TfxNMT	<i>pGC1312::P<sub>TEF1</sub>-TfxNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1458	pBOT-His	CjCNMT	<i>pGC1312::P<sub>TEF1</sub>-CjCNMT-GFP-T<sub>PGII</sub></i>	Chp. 3
pGC1459	pBOT-His	NMT-Ame-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Ame-1-T<sub>PGII</sub></i>	Chp. 3
pGC1460	pBOT-His	NMT-Cch-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cch-1-T<sub>PGII</sub></i>	Chp. 3
pGC1461	pBOT-His	NMT-Cch-2	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cch-2-T<sub>PGII</sub></i>	Chp. 3
pGC1462	pBOT-His	NMT-Cma-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Cma-1-T<sub>PGII</sub></i>	Chp. 3
pGC1463	pBOT-His	NMT-Eca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Eca-1-T<sub>PGII</sub></i>	Chp. 3
pGC1464	pBOT-His	NMT-Hca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Hca-1-T<sub>PGII</sub></i>	Chp. 3
pGC1465	pBOT-His	NMT-Pbr-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-1-T<sub>PGII</sub></i>	Chp. 3
pGC1466	pBOT-His	NMT-Pbr-2	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-2-T<sub>PGII</sub></i>	Chp. 3

pGC1467	pBOT-His	NMT-Pbr-4	<i>pGC1312::P<sub>TEF1</sub>-NMT-Pbr-4-T<sub>PGII</sub></i>	Chp. 3
pGC1468	pBOT-His	NMT-Sca-1	<i>pGC1312::P<sub>TEF1</sub>-NMT-Sca-1-T<sub>PGII</sub></i>	Chp. 3
pGC1469	pBOT-His	TfCNMT	<i>pGC1312::P<sub>TEF1</sub>-TfCNMT-T<sub>PGII</sub></i>	Chp. 3
pGC1470	pBOT-His	TfPNMT	<i>pGC1312::P<sub>TEF1</sub>-TfPNMT-T<sub>PGII</sub></i>	Chp. 3
pGC1471	pBOT-His	EcTNMT	<i>pGC1312::P<sub>TEF1</sub>-EcTNMT-T<sub>PGII</sub></i>	Chp. 3
<b>CYP719 assays</b>				
Number	Base vector	Gene in expression cassette	Description	Ref.
pGC1472	pBOT-Trp	CYP719-Cch-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1473	pBOT-Trp	CYP719-Cch-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1474	pBOT-Trp	CYP719-Cch-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1475	pBOT-Trp	CYP719-Cch-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-4-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1476	pBOT-Trp	CYP719-Cma-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1477	pBOT-Trp	CYP719-Cma-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1478	pBOT-Trp	CYP719-Cma-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1479	pBOT-Trp	CYP719-Cma-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-4-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1480	pBOT-Trp	CYP719-Cmu-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cmu-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1481	pBOT-Trp	CYP719-Cmu-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cmu-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1482	pBOT-Trp	CYP719-Gfl-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1483	pBOT-Trp	CYP719-Gfl-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1484	pBOT-Trp	CYP719-Gfl-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1485	pBOT-Trp	CYP719-Hca-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Hca-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1486	pBOT-Trp	CYP719-Mec-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Mec-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1487	pBOT-Trp	CYP719-Ndo-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1488	pBOT-Trp	CYP719-Ndo-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1489	pBOT-Trp	CYP719-Ndo-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1490	pBOT-Trp	CYP719-Ndo-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-4-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1491	pBOT-Trp	CYP719-Ndo-5	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-5-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1492	pBOT-Trp	CYP719-Ndo-6	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-6-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1493	pBOT-Trp	CYP719-Pbr-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pbr-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1494	pBOT-Trp	CYP719-Pbr-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pbr-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1495	pBOT-Trp	CYP719-Pso-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1496	pBOT-Trp	CYP719-Pso-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1497	pBOT-Trp	CYP719-Pso-5	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-5-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1498	pBOT-Trp	CYP719-Pso-6	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-6-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1499	pBOT-Trp	CYP719-Sca-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1500	pBOT-Trp	CYP719-Sca-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1501	pBOT-Trp	CYP719-Sca-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1502	pBOT-Trp	CYP719-Sdi-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1503	pBOT-Trp	CYP719-Sdi-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1504	pBOT-Trp	CYP719-Sdi-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-3-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1505	pBOT-Trp	CYP719-Sdi-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-4-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1506	pBOT-Trp	CYP719-Tfl-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Tfl-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1507	pBOT-Trp	CYP719-Xsi-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Xsi-1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1508	pBOT-Trp	CYP719-Xsi-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Xsi-2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1509	pBOT-Trp	CYP719A1	<i>pGC1313::P<sub>TDH3</sub>-CYP719A1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1510	pBOT-Trp	CYP719A2	<i>pGC1313::P<sub>TDH3</sub>-CYP719A2-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1511	pBOT-Trp	CYP719A4	<i>pGC1313::P<sub>TDH3</sub>-CYP719A4-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1512	pBOT-Trp	CYP719A5	<i>pGC1313::P<sub>TDH3</sub>-CYP719A5-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1513	pBOT-Trp	CYP719A6	<i>pGC1313::P<sub>TDH3</sub>-CYP719A6-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1514	pBOT-Trp	CYP719A7	<i>pGC1313::P<sub>TDH3</sub>-CYP719A7-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1515	pBOT-Trp	CYP719A9	<i>pGC1313::P<sub>TDH3</sub>-CYP719A9-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1516	pBOT-Trp	CYP719A11	<i>pGC1313::P<sub>TDH3</sub>-CYP719A11-GFP-T<sub>CYC1</sub></i>	Chp. 3

pGC1517	pBOT-Trp	CYP719A13	<i>pGC1313::P<sub>TDH3</sub>-CYP719A13-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1518	pBOT-Trp	CYP719A14	<i>pGC1313::P<sub>TDH3</sub>-CYP719A14-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1519	pBOT-Trp	CYP719A17	<i>pGC1313::P<sub>TDH3</sub>-CYP719A17-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1520	pBOT-Trp	CYP719A18	<i>pGC1313::P<sub>TDH3</sub>-CYP719A18-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1521	pBOT-Trp	CYP719A19	<i>pGC1313::P<sub>TDH3</sub>-CYP719A19-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1522	pBOT-Trp	CYP719A21	<i>pGC1313::P<sub>TDH3</sub>-CYP719A21-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1523	pBOT-Trp	CYP719A22	<i>pGC1313::P<sub>TDH3</sub>-CYP719A22-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1524	pBOT-Trp	CYP719A23	<i>pGC1313::P<sub>TDH3</sub>-CYP719A23-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1525	pBOT-Trp	CYP719A24	<i>pGC1313::P<sub>TDH3</sub>-CYP719A24-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1526	pBOT-Trp	CYP719A25	<i>pGC1313::P<sub>TDH3</sub>-CYP719A25-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1527	pBOT-Trp	CYP719B1	<i>pGC1313::P<sub>TDH3</sub>-CYP719B1-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1528	pBOT-Trp	LsSPS-GFP	<i>pGC1313::P<sub>TDH3</sub>-LsSPS-GFP-T<sub>CYC1</sub></i>	Chp. 3
pGC1529	pBOT-Trp	CYP719-Cch-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1530	pBOT-Trp	CYP719-Cch-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1531	pBOT-Trp	CYP719-Cch-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1532	pBOT-Trp	CYP719-Cch-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cch-4-T<sub>CYC1</sub></i>	Chp. 3
pGC1533	pBOT-Trp	CYP719-Cma-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1534	pBOT-Trp	CYP719-Cma-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1535	pBOT-Trp	CYP719-Cma-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1536	pBOT-Trp	CYP719-Cma-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cma-4-T<sub>CYC1</sub></i>	Chp. 3
pGC1537	pBOT-Trp	CYP719-Cmu-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cmu-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1538	pBOT-Trp	CYP719-Cmu-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Cmu-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1539	pBOT-Trp	CYP719-Gfl-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1540	pBOT-Trp	CYP719-Gfl-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1541	pBOT-Trp	CYP719-Gfl-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Gfl-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1542	pBOT-Trp	CYP719-Hca-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Hca-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1543	pBOT-Trp	CYP719-Mec-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Mec-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1544	pBOT-Trp	CYP719-Ndo-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1545	pBOT-Trp	CYP719-Ndo-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1546	pBOT-Trp	CYP719-Ndo-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1547	pBOT-Trp	CYP719-Ndo-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-4-T<sub>CYC1</sub></i>	Chp. 3
pGC1548	pBOT-Trp	CYP719-Ndo-5	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-5-T<sub>CYC1</sub></i>	Chp. 3
pGC1549	pBOT-Trp	CYP719-Ndo-6	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Ndo-6-T<sub>CYC1</sub></i>	Chp. 3
pGC1550	pBOT-Trp	CYP719-Pbr-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pbr-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1551	pBOT-Trp	CYP719-Pbr-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pbr-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1552	pBOT-Trp	CYP719-Pso-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1553	pBOT-Trp	CYP719-Pso-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1554	pBOT-Trp	CYP719-Pso-5	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-5-T<sub>CYC1</sub></i>	Chp. 3
pGC1555	pBOT-Trp	CYP719-Pso-6	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Pso-6-T<sub>CYC1</sub></i>	Chp. 3
pGC1556	pBOT-Trp	CYP719-Sca-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1557	pBOT-Trp	CYP719-Sca-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1558	pBOT-Trp	CYP719-Sca-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sca-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1559	pBOT-Trp	CYP719-Sdi-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1560	pBOT-Trp	CYP719-Sdi-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1561	pBOT-Trp	CYP719-Sdi-3	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-3-T<sub>CYC1</sub></i>	Chp. 3
pGC1562	pBOT-Trp	CYP719-Sdi-4	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Sdi-4-T<sub>CYC1</sub></i>	Chp. 3
pGC1563	pBOT-Trp	CYP719-Tfl-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Tfl-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1564	pBOT-Trp	CYP719-Xsi-1	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Xsi-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1565	pBOT-Trp	CYP719-Xsi-2	<i>pGC1313::P<sub>TDH3</sub>-CYP719-Xsi-2-T<sub>CYC1</sub></i>	Chp. 3
pGC1566	pBOT-Trp	CYP719A1	<i>pGC1313::P<sub>TDH3</sub>-CYP719A1-T<sub>CYC1</sub></i>	Chp. 3
pGC1567	pBOT-Trp	CYP719A2	<i>pGC1313::P<sub>TDH3</sub>-CYP719A2-T<sub>CYC1</sub></i>	Chp. 3
pGC1568	pBOT-Trp	CYP719A4	<i>pGC1313::P<sub>TDH3</sub>-CYP719A4-T<sub>CYC1</sub></i>	Chp. 3
pGC1569	pBOT-Trp	CYP719A5	<i>pGC1313::P<sub>TDH3</sub>-CYP719A5-T<sub>CYC1</sub></i>	Chp. 3

pGC1570	pBOT-Trp	CYP719A6	<i>pGC1313::P<sub>TDH3</sub>-CYP719A6-T<sub>CYC1</sub></i>	Chp. 3
pGC1571	pBOT-Trp	CYP719A7	<i>pGC1313::P<sub>TDH3</sub>-CYP719A7-T<sub>CYC1</sub></i>	Chp. 3
pGC1572	pBOT-Trp	CYP719A9	<i>pGC1313::P<sub>TDH3</sub>-CYP719A9-T<sub>CYC1</sub></i>	Chp. 3
pGC1573	pBOT-Trp	CYP719A11	<i>pGC1313::P<sub>TDH3</sub>-CYP719A11-T<sub>CYC1</sub></i>	Chp. 3
pGC1574	pBOT-Trp	CYP719A13	<i>pGC1313::P<sub>TDH3</sub>-CYP719A13-T<sub>CYC1</sub></i>	Chp. 3
pGC1575	pBOT-Trp	CYP719A14	<i>pGC1313::P<sub>TDH3</sub>-CYP719A14-T<sub>CYC1</sub></i>	Chp. 3
pGC1576	pBOT-Trp	CYP719A17	<i>pGC1313::P<sub>TDH3</sub>-CYP719A17-T<sub>CYC1</sub></i>	Chp. 3
pGC1577	pBOT-Trp	CYP719A18	<i>pGC1313::P<sub>TDH3</sub>-CYP719A18-T<sub>CYC1</sub></i>	Chp. 3
pGC1578	pBOT-Trp	CYP719A19	<i>pGC1313::P<sub>TDH3</sub>-CYP719A19-T<sub>CYC1</sub></i>	Chp. 3
pGC1579	pBOT-Trp	CYP719A21	<i>pGC1313::P<sub>TDH3</sub>-CYP719A21-T<sub>CYC1</sub></i>	Chp. 3
pGC1580	pBOT-Trp	CYP719A22	<i>pGC1313::P<sub>TDH3</sub>-CYP719A22-T<sub>CYC1</sub></i>	Chp. 3
pGC1581	pBOT-Trp	CYP719A23	<i>pGC1313::P<sub>TDH3</sub>-CYP719A23-T<sub>CYC1</sub></i>	Chp. 3
pGC1582	pBOT-Trp	CYP719A24	<i>pGC1313::P<sub>TDH3</sub>-CYP719A24-T<sub>CYC1</sub></i>	Chp. 3
pGC1583	pBOT-Trp	CYP719A25	<i>pGC1313::P<sub>TDH3</sub>-CYP719A25-T<sub>CYC1</sub></i>	Chp. 3
pGC1584	pBOT-Trp	CYP719B1	<i>pGC1313::P<sub>TDH3</sub>-CYP719B1-T<sub>CYC1</sub></i>	Chp. 3
pGC1585	pBOT-Trp	LsSPS	<i>pGC1313::P<sub>TDH3</sub>-LsSPS-T<sub>CYC1</sub></i>	Chp. 3
pGC1586	pBOT-TPP	CYP719-Cma-4	<i>pGC1381::P<sub>PMA1</sub>-CYP719-Cma-4-T<sub>PGII</sub></i>	Chp. 3
pGC1587	pBOT-TPP	CYP719-Gfl-3	<i>pGC1381::P<sub>PMA1</sub>-CYP719-Gfl-3-T<sub>PGII</sub></i>	Chp. 3
pGC1588	pBOT-TPP	CYP719-Ndo-6	<i>pGC1381::P<sub>PMA1</sub>-CYP719-Ndo-6-T<sub>PGII</sub></i>	Chp. 3
pGC1589	pBOT-TPP	CYP719-Sca-2	<i>pGC1381::P<sub>PMA1</sub>-CYP719-Sca-2-T<sub>PGII</sub></i>	Chp. 3
pGC1590	pBOT-TPP	CYP719-Sdi-4	<i>pGC1381::P<sub>PMA1</sub>-CYP719-Sdi-4-T<sub>PGII</sub></i>	Chp. 3
pGC1591	pBOT-TPP	CYP719A14	<i>pGC1381::P<sub>PMA1</sub>-CYP719A14-T<sub>PGII</sub></i>	Chp. 3
pGC1592	pBOT-TPP	PsSPS	<i>pGC1381::P<sub>PMA1</sub>-PsSPS-T<sub>PGII</sub></i>	Chp. 3
pGC1593	pBOT-LTC	CYP719-Cma-1	<i>pGC1441::P<sub>TDH3</sub>-CYP719-Cma-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1594	pBOT-LTC	CYP719-Sdi-1	<i>pGC1441::P<sub>TDH3</sub>-CYP719-Sdi-1-T<sub>CYC1</sub></i>	Chp. 3
pGC1595	pBOT-LTC	CYP719A25	<i>pGC1441::P<sub>TDH3</sub>-CYP719A25-T<sub>CYC1</sub></i>	Chp. 3
<b>Templates for dihydrosanguinarine pathway gene integration</b>				
<b>Number</b>	<b>Base vector</b>	<b>Gene in expression cassette</b>	<b>Description</b>	<b>Ref.</b>
pUG6	pFA6	KanMX	<i>pFA6::loxP-P<sub>AgTEF1</sub>-kanMX-T<sub>AgTEF1</sub>-loxP</i>	181
pZC3	pFA6	HphNT1	<i>pFA6::loxLE-P<sub>AgTEF2</sub>-hphNT1-T<sub>CYC1</sub>-loxRE</i>	182
pGC557	pGREG-His	PsCPR	<i>pGC964::C1-P<sub>TDH3</sub>-CPR-T<sub>CYC1</sub>-C6</i>	23
pGC552	pGREG-Ura	CYP719A25 <sup>a</sup>	<i>pGC967::C1-P<sub>TDH1p</sub>-CYP719A25-T<sub>CYC1</sub>-C6</i>	Chp. 3
pGC1062	pGREG-Leu	Ps6OMT, Ps4'OMT2, PsCNMT	<i>pGC966::C1-P<sub>TDH3</sub>-6OMT-T<sub>CYC1</sub>-C6-H1-C1-P<sub>FBA1</sub>-4'OMT2-T<sub>ADH1</sub>-C6-H2-C1-P<sub>PDC1</sub>-CNMT-T<sub>PGII</sub>-C6</i>	23
pGC997	pGREG-Ura	PsP6H, PsMSH, PsTNMT	<i>pGC967::C1-P<sub>PDC1</sub>-P6H-T<sub>CYC1</sub>-C6-H1-C1-P<sub>TDH3</sub>-MSH-T<sub>ADH1</sub>-C6-H2-C1-P<sub>FBA1</sub>-TNMT-T<sub>PGII</sub>-C6</i>	23
pGC717	pGREG-His	PsCPR, PsTNMT	<i>pGC964::C1-P<sub>FBA1</sub>-CPR-T<sub>CYC1</sub>-C6-H1-C1-P<sub>TDH1</sub>-TNMT-T<sub>ADH1</sub>-C6</i>	23
pGC556	pGREG-His	PsCPR, PsBBE <sup>a</sup>	<i>pGC964::C1-P<sub>TDH3</sub>-CPR-T<sub>CYC1</sub>-C6-H1-C1-P<sub>PMA1</sub>-BBE-T<sub>PGII</sub>-C6</i>	Chp. 3
<b>CRISPR vectors</b>				
<b>Number</b>	<b>Base vector</b>	<b>gRNA target</b>	<b>Description</b>	<b>Ref.</b>
pGC1596	pCas-Tyr	Flagfeldt 16 UP	<i>pGC1421::gRNA</i> 5'-TTTTCCGACAATCAAATATT-3'	Chp. 3
pGC1597	pCas-Tyr	Flagfeldt 16 DOWN	<i>pGC1421::gRNA</i> 5'-TCATCAAAAGAGACATTTTT-3'	Chp. 3
pGC1598	pCas-Tyr	Flagfeldt 20 UP	<i>pGC1421::gRNA</i> 5'-AAAATTCTCTCTGAGGATAT-3'	Chp. 3
pGC1599	pCas-Tyr	Flagfeldt 20 DOWN	<i>pGC1421::gRNA</i> 5'-GTTAGAGCTGTTACAAGTTA-3'	Chp. 3
pGC1600	pCas-Tyr	Flagfeldt 21 UP	<i>pGC1421::gRNA</i>	Chp. 3

			5'-TTTATACATTTCACATGTAC-3'	
pGC1601	pCas-Tyr	Flagfeldt 21 DOWN	<i>pGC1421::gRNA</i> 5'-TATTCCATGGCCTCTTAGTT-3'	Chp. 3

Chp.3 refers to Chapter 3 of this work

a: Gene source is described in Ref. 3.

**Table 3.6. Strain list for Chapter 3**

<b>Base strain</b>				
<b>Name</b>	<b>Brief description</b>		<b>Genotype</b>	<b>Ref.</b>
CEN.PK2-1D	Quadruple auxotroph (his, trp, leu, ura)		<i>MATa ura3-52 trp1-289 leu2-3,112 his3-Δ1 MAL2-8C SUC2</i>	188
<b>Genomic integrations</b>				
<b>Name</b>	<b>Brief description</b>		<b>Genotype</b>	<b>Ref.</b>
GCY1333	<i>PsCPR</i> integrant into site 16		CEN.PK2-1D <i>YNRCΔ9(ChrXIV)::kanMX-C1-P<sub>TDH3</sub>-PsCPR-T<sub>CYC1</sub>-C6</i>	Chp. 3
GCY1270	<i>PsCPR</i> , <i>PsTNMT</i> integrant into site 18		CEN.PK2-1D <i>YORWΔ17(ChrXV)::kanMX-C1-P<sub>FBA1</sub>-PsCPR-T<sub>CYC1</sub>-C6-H1-C1-P<sub>TDH3</sub>-PsTNMT-T<sub>ADH1</sub>-C6</i>	Chp. 3
GCY1317	<i>PsCPR</i> integrant into site 16 <i>PsCFS</i> integrant into site 18		CEN.PK2-1D <i>YNRCΔ9(ChrXIV)::hphNT1-C1-P<sub>TDH3</sub>-PsCPR-T<sub>CYC1</sub>-C6, YORWΔ17(ChrXV)::kanMX-C1-P<sub>TDH3</sub>-PsCFS-T<sub>CYC1</sub>-C6</i>	Chp. 3
GCY1440	<i>Ps6OMT</i> , <i>Ps4'OMT2</i> , <i>PsCNMT</i> integrant into site 21; <i>PsCPR</i> , <i>PsBBE</i> integrant into site 16; <i>PsTNMT</i> , <i>PsMSH</i> , <i>PsP6H</i> integrant into site 20		CEN.PK2-1D <i>YNRCΔ9(ChrXIV)::LV3-P<sub>TDH3</sub>-PsCPR-T<sub>CYC1</sub>-LTP1-P<sub>PMA1</sub>-PsBBE-T<sub>PGI1</sub>-LV5, YPRC<i>d15(ChrXVI)::-LV3-P<sub>PDC1</sub>-PsP6H-T<sub>CYC1</sub>-LTP1-P<sub>TDH3</sub>-PsMSH-T<sub>ADH1</sub>-LTP2-P<sub>FBA1</sub>-PsTNMT-T<sub>PGI1</sub>-LV5, YPRC<i>x3(ChrXVI)::LV3-P<sub>TDH3</sub>-Ps6OMT-T<sub>CYC1</sub>-LTP1-P<sub>FBA1</sub>-Ps4'OMT2-T<sub>ADH1</sub>-LTP2-P<sub>PDC1</sub>-PsCNMT-T<sub>PGI1</sub>-LV5</i></i></i>	Chp. 3
<b>NMT assays</b>				
<b>Name</b>	<b>Genotype</b>	<b>Plasmid(s)</b>	<b>Description</b>	<b>Ref.</b>
GCY1602	CEN.PK2-1D	pGC964	CEN.PK2-1D with empty vector	Chp. 3
GCY1603	CEN.PK2-1D	pGC1442	CEN.PK2-1D with NMT-Ame-1-GFP	Chp. 3
GCY1604	CEN.PK2-1D	pGC1443	CEN.PK2-1D with NMT-Ceh-1-GFP	Chp. 3
GCY1605	CEN.PK2-1D	pGC1444	CEN.PK2-1D with NMT-Ceh-2-GFP	Chp. 3
GCY1606	CEN.PK2-1D	pGC1445	CEN.PK2-1D with NMT-Cma-1-GFP	Chp. 3
GCY1607	CEN.PK2-1D	pGC1446	CEN.PK2-1D with NMT-Eca-1-GFP	Chp. 3
GCY1608	CEN.PK2-1D	pGC1447	CEN.PK2-1D with NMT-Hca-1-GFP	Chp. 3
GCY1609	CEN.PK2-1D	pGC1448	CEN.PK2-1D with NMT-Pbr-1-GFP	Chp. 3
GCY1610	CEN.PK2-1D	pGC1449	CEN.PK2-1D with NMT-Pbr-2-GFP	Chp. 3
GCY1611	CEN.PK2-1D	pGC1450	CEN.PK2-1D with NMT-Pbr-4-GFP	Chp. 3
GCY1612	CEN.PK2-1D	pGC1451	CEN.PK2-1D with NMT-Sea-1-GFP	Chp. 3
GCY1613	CEN.PK2-1D	pGC1452	CEN.PK2-1D with TfcNMT-GFP	Chp. 3
GCY1614	CEN.PK2-1D	pGC1453	CEN.PK2-1D with TfpNMT-GFP	Chp. 3
GCY1615	CEN.PK2-1D	pGC1454	CEN.PK2-1D with EcTNMT-GFP	Chp. 3
GCY1616	CEN.PK2-1D	pGC1455	CEN.PK2-1D with PbTNMT-GFP	Chp. 3
GCY1617	CEN.PK2-1D	pGC1456	CEN.PK2-1D with PbTNMT2-GFP	Chp. 3
GCY1618	CEN.PK2-1D	pGC1457	CEN.PK2-1D with NMT-Ame-1	Chp. 3
GCY1619	CEN.PK2-1D	pGC1458	CEN.PK2-1D with NMT-Ceh-1	Chp. 3
GCY1620	CEN.PK2-1D	pGC1459	CEN.PK2-1D with NMT-Ceh-2	Chp. 3
GCY1621	CEN.PK2-1D	pGC1460	CEN.PK2-1D with NMT-Cma-1	Chp. 3
GCY1622	CEN.PK2-1D	pGC1461	CEN.PK2-1D with NMT-Eca-1	Chp. 3
GCY1623	CEN.PK2-1D	pGC1462	CEN.PK2-1D with NMT-Hca-1	Chp. 3
GCY1624	CEN.PK2-1D	pGC1463	CEN.PK2-1D with NMT-Pbr-1	Chp. 3
GCY1625	CEN.PK2-1D	pGC1464	CEN.PK2-1D with NMT-Pbr-2	Chp. 3
GCY1626	CEN.PK2-1D	pGC1465	CEN.PK2-1D with NMT-Pbr-4	Chp. 3
GCY1627	CEN.PK2-1D	pGC1466	CEN.PK2-1D with NMT-Sea-1	Chp. 3

GCY1628	CEN.PK2-1D	pGC1467	CEN.PK2-1D with TfcNMT	Chp. 3
GCY1629	CEN.PK2-1D	pGC1468	CEN.PK2-1D with TfpNMT	Chp. 3
GCY1630	CEN.PK2-1D	pGC1469	CEN.PK2-1D with EcTNMT	Chp. 3
GCY1631	CEN.PK2-1D	pGC1470	CEN.PK2-1D with PbTNMT	Chp. 3
GCY1632	CEN.PK2-1D	pGC1471	CEN.PK2-1D with PbTNMT2	Chp. 3
GCY1633	CEN.PK2-1D	pGC717	CEN.PK2-1D with PsTNMT	Chp. 3
GCY1634	CEN.PK2-1D	pGC633	CEN.PK2-1D with Ps6OMT	Chp. 3
GCY1635	CEN.PK2-1D	pGC635	CEN.PK2-1D with Ps4'OMT2	Chp. 3
GCY1636	CEN.PK2-1D	pGC652	CEN.PK2-1D with PsCNMT	Chp. 3
GCY1637	GCY1317	pGC964	CPR, CFS integrant with empty vector	Chp. 3
GCY1638	GCY1317	pGC1457	CPR, CFS integrant with NMT-Ame-1	Chp. 3
GCY1639	GCY1317	pGC1458	CPR, CFS integrant with NMT-Cch-1	Chp. 3
GCY1640	GCY1317	pGC1459	CPR, CFS integrant with NMT-Cch-2	Chp. 3
GCY1641	GCY1317	pGC1460	CPR, CFS integrant with NMT-Cma-1	Chp. 3
GCY1642	GCY1317	pGC1461	CPR, CFS integrant with NMT-Eca-1	Chp. 3
GCY1643	GCY1317	pGC1462	CPR, CFS integrant with NMT-Hca-1	Chp. 3
GCY1644	GCY1317	pGC1463	CPR, CFS integrant with NMT-Pbr-1	Chp. 3
GCY1645	GCY1317	pGC1464	CPR, CFS integrant with NMT-Pbr-2	Chp. 3
GCY1646	GCY1317	pGC1465	CPR, CFS integrant with NMT-Pbr-4	Chp. 3
GCY1647	GCY1317	pGC1466	CPR, CFS integrant with NMT-Sca-1	Chp. 3
GCY1648	GCY1317	pGC1467	CPR, CFS integrant with <i>Tfc</i> NMT	Chp. 3
GCY1649	GCY1317	pGC1468	CPR, CFS integrant with <i>Tfp</i> NMT	Chp. 3
GCY1650	GCY1317	pGC1469	CPR, CFS integrant with <i>Ec</i> TNMT	Chp. 3
GCY1651	GCY1317	pGC1470	CPR, CFS integrant with <i>Pb</i> TNMT	Chp. 3
GCY1652	GCY1317	pGC1471	CPR, CFS integrant with <i>Pb</i> TNMT2	Chp. 3
GCY1653	GCY1317	pGC717	CPR, CFS integrant with <i>Ps</i> TNMT	Chp. 3
GCY1654	GCY1317	pGC633	CPR, CFS integrant with <i>Ps</i> 6OMT	Chp. 3
GCY1655	GCY1317	pGC635	CPR, CFS integrant with <i>Ps</i> 4'OMT2	Chp. 3
GCY1656	GCY1317	pGC652	CPR, CFS integrant with <i>Ps</i> CNMT	Chp. 3
<b>CYP719 assays</b>				
<b>Name</b>	<b>Genotype</b>	<b>Plasmid(s)</b>		<b>Ref.</b>
GCY1657	GCY1333	pGC965	CPR integrant with empty vector	Chp. 3
GCY1658	GCY1333	pGC1472	CPR integrant with CYP719-Cch-1-GFP	Chp. 3
GCY1659	GCY1333	pGC1473	CPR integrant with CYP719-Cch-2-GFP	Chp. 3
GCY1660	GCY1333	pGC1474	CPR integrant with CYP719-Cch-3-GFP	Chp. 3
GCY1661	GCY1333	pGC1475	CPR integrant with CYP719-Cch-4-GFP	Chp. 3
GCY1662	GCY1333	pGC1476	CPR integrant with CYP719-Cma-1-GFP	Chp. 3
GCY1663	GCY1333	pGC1477	CPR integrant with CYP719-Cma-2-GFP	Chp. 3
GCY1664	GCY1333	pGC1478	CPR integrant with CYP719-Cma-3-GFP	Chp. 3
GCY1665	GCY1333	pGC1479	CPR integrant with CYP719-Cma-4-GFP	Chp. 3
GCY1666	GCY1333	pGC1480	CPR integrant with CYP719-Cmu-1-GFP	Chp. 3
GCY1667	GCY1333	pGC1481	CPR integrant with CYP719-Cmu-2-GFP	Chp. 3
GCY1668	GCY1333	pGC1482	CPR integrant with CYP719-Gfl-1-GFP	Chp. 3
GCY1669	GCY1333	pGC1483	CPR integrant with CYP719-Gfl-2-GFP	Chp. 3
GCY1670	GCY1333	pGC1484	CPR integrant with CYP719-Gfl-3-GFP	Chp. 3
GCY1671	GCY1333	pGC1485	CPR integrant with CYP719-Hca-1-GFP	Chp. 3
GCY1672	GCY1333	pGC1486	CPR integrant with CYP719-Mec-1-GFP	Chp. 3
GCY1673	GCY1333	pGC1487	CPR integrant with CYP719-Ndo-1-GFP	Chp. 3
GCY1674	GCY1333	pGC1488	CPR integrant with CYP719-Ndo-2-GFP	Chp. 3
GCY1675	GCY1333	pGC1489	CPR integrant with CYP719-Ndo-3-GFP	Chp. 3
GCY1676	GCY1333	pGC1490	CPR integrant with CYP719-Ndo-4-GFP	Chp. 3
GCY1677	GCY1333	pGC1491	CPR integrant with CYP719-Ndo-5-GFP	Chp. 3
GCY1678	GCY1333	pGC1492	CPR integrant with CYP719-Ndo-6-GFP	Chp. 3

GCY1679	GCY1333	pGC1493	CPR integrant with CYP719-Pbr-2-GFP	Chp. 3
GCY1680	GCY1333	pGC1494	CPR integrant with CYP719-Pbr-3-GFP	Chp. 3
GCY1681	GCY1333	pGC1495	CPR integrant with CYP719-Pso-1-GFP	Chp. 3
GCY1682	GCY1333	pGC1496	CPR integrant with CYP719-Pso-3-GFP	Chp. 3
GCY1683	GCY1333	pGC1497	CPR integrant with CYP719-Pso-5-GFP	Chp. 3
GCY1684	GCY1333	pGC1498	CPR integrant with CYP719-Pso-6-GFP	Chp. 3
GCY1685	GCY1333	pGC1499	CPR integrant with CYP719-Sca-1-GFP	Chp. 3
GCY1686	GCY1333	pGC1500	CPR integrant with CYP719-Sca-2-GFP	Chp. 3
GCY1687	GCY1333	pGC1501	CPR integrant with CYP719-Sca-3-GFP	Chp. 3
GCY1688	GCY1333	pGC1502	CPR integrant with CYP719-Sdi-1-GFP	Chp. 3
GCY1689	GCY1333	pGC1503	CPR integrant with CYP719-Sdi-2-GFP	Chp. 3
GCY1690	GCY1333	pGC1504	CPR integrant with CYP719-Sdi-3-GFP	Chp. 3
GCY1691	GCY1333	pGC1505	CPR integrant with CYP719-Sdi-4-GFP	Chp. 3
GCY1692	GCY1333	pGC1506	CPR integrant with CYP719-Tfl-2-GFP	Chp. 3
GCY1693	GCY1333	pGC1507	CPR integrant with CYP719-Xsi-1-GFP	Chp. 3
GCY1694	GCY1333	pGC1508	CPR integrant with CYP719-Xsi-2-GFP	Chp. 3
GCY1695	GCY1333	pGC1509	CPR integrant with CYP719A1-GFP	Chp. 3
GCY1696	GCY1333	pGC1510	CPR integrant with CYP719A2-GFP	Chp. 3
GCY1697	GCY1333	pGC1511	CPR integrant with CYP719A4-GFP	Chp. 3
GCY1698	GCY1333	pGC1512	CPR integrant with CYP719A5-GFP	Chp. 3
GCY1699	GCY1333	pGC1513	CPR integrant with CYP719A6-GFP	Chp. 3
GCY1700	GCY1333	pGC1514	CPR integrant with CYP719A7-GFP	Chp. 3
GCY1701	GCY1333	pGC1515	CPR integrant with CYP719A9-GFP	Chp. 3
GCY1702	GCY1333	pGC1516	CPR integrant with CYP719A11-GFP	Chp. 3
GCY1703	GCY1333	pGC1517	CPR integrant with CYP719A13-GFP	Chp. 3
GCY1704	GCY1333	pGC1518	CPR integrant with CYP719A14-GFP	Chp. 3
GCY1705	GCY1333	pGC1519	CPR integrant with CYP719A17-GFP	Chp. 3
GCY1706	GCY1333	pGC1520	CPR integrant with CYP719A18-GFP	Chp. 3
GCY1707	GCY1333	pGC1521	CPR integrant with CYP719A19-GFP	Chp. 3
GCY1708	GCY1333	pGC1522	CPR integrant with CYP719A21-GFP	Chp. 3
GCY1709	GCY1333	pGC1523	CPR integrant with CYP719A22-GFP	Chp. 3
GCY1710	GCY1333	pGC1524	CPR integrant with CYP719A23-GFP	Chp. 3
GCY1711	GCY1333	pGC1525	CPR integrant with CYP719A24-GFP	Chp. 3
GCY1712	GCY1333	pGC1526	CPR integrant with CYP719A25-GFP	Chp. 3
GCY1713	GCY1333	pGC1527	CPR integrant with CYP719B1-GFP	Chp. 3
GCY1714	GCY1333	pGC1528	CPR integrant with <i>Ps</i> SPS-GFP	Chp. 3
GCY1715	GCY1333	pGC1529	CPR integrant with CYP719-Cch-1	Chp. 3
GCY1716	GCY1333	pGC1530	CPR integrant with CYP719-Cch-2	Chp. 3
GCY1717	GCY1333	pGC1531	CPR integrant with CYP719-Cch-3	Chp. 3
GCY1718	GCY1333	pGC1532	CPR integrant with CYP719-Cch-4	Chp. 3
GCY1719	GCY1333	pGC1533	CPR integrant with CYP719-Cma-1	Chp. 3
GCY1720	GCY1333	pGC1534	CPR integrant with CYP719-Cma-2	Chp. 3
GCY1721	GCY1333	pGC1535	CPR integrant with CYP719-Cma-3	Chp. 3
GCY1722	GCY1333	pGC1536	CPR integrant with CYP719-Cma-4	Chp. 3
GCY1723	GCY1333	pGC1537	CPR integrant with CYP719-Cmu-1	Chp. 3
GCY1724	GCY1333	pGC1538	CPR integrant with CYP719-Cmu-2	Chp. 3
GCY1725	GCY1333	pGC1539	CPR integrant with CYP719-Gfl-1	Chp. 3
GCY1726	GCY1333	pGC1540	CPR integrant with CYP719-Gfl-2	Chp. 3
GCY1727	GCY1333	pGC1541	CPR integrant with CYP719-Gfl-3	Chp. 3
GCY1728	GCY1333	pGC1542	CPR integrant with CYP719-Hca-1	Chp. 3
GCY1729	GCY1333	pGC1543	CPR integrant with CYP719-Mec-1	Chp. 3
GCY1730	GCY1333	pGC1544	CPR integrant with CYP719-Ndo-1	Chp. 3
GCY1731	GCY1333	pGC1545	CPR integrant with CYP719-Ndo-2	Chp. 3

GCY1732	GCY1333	pGC1546	CPR integrant with CYP719-Ndo-3	Chp. 3
GCY1733	GCY1333	pGC1547	CPR integrant with CYP719-Ndo-4	Chp. 3
GCY1734	GCY1333	pGC1548	CPR integrant with CYP719-Ndo-5	Chp. 3
GCY1735	GCY1333	pGC1549	CPR integrant with CYP719-Ndo-6	Chp. 3
GCY1736	GCY1333	pGC1550	CPR integrant with CYP719-Pbr-2	Chp. 3
GCY1737	GCY1333	pGC1551	CPR integrant with CYP719-Pbr-3	Chp. 3
GCY1738	GCY1333	pGC1552	CPR integrant with CYP719-Pso-1	Chp. 3
GCY1739	GCY1333	pGC1553	CPR integrant with CYP719-Pso-3	Chp. 3
GCY1740	GCY1333	pGC1554	CPR integrant with CYP719-Pso-5	Chp. 3
GCY1741	GCY1333	pGC1555	CPR integrant with CYP719-Pso-6	Chp. 3
GCY1742	GCY1333	pGC1556	CPR integrant with CYP719-Sca-1	Chp. 3
GCY1743	GCY1333	pGC1557	CPR integrant with CYP719-Sca-2	Chp. 3
GCY1744	GCY1333	pGC1558	CPR integrant with CYP719-Sca-3	Chp. 3
GCY1745	GCY1333	pGC1559	CPR integrant with CYP719-Sdi-1	Chp. 3
GCY1746	GCY1333	pGC1560	CPR integrant with CYP719-Sdi-2	Chp. 3
GCY1747	GCY1333	pGC1561	CPR integrant with CYP719-Sdi-3	Chp. 3
GCY1748	GCY1333	pGC1562	CPR integrant with CYP719-Sdi-4	Chp. 3
GCY1749	GCY1333	pGC1563	CPR integrant with CYP719-Tfl-2	Chp. 3
GCY1750	GCY1333	pGC1564	CPR integrant with CYP719-Xsi-1	Chp. 3
GCY1751	GCY1333	pGC1565	CPR integrant with CYP719-Xsi-2	Chp. 3
GCY1752	GCY1333	pGC1566	CPR integrant with CYP719A1	Chp. 3
GCY1753	GCY1333	pGC1567	CPR integrant with CYP719A2	Chp. 3
GCY1754	GCY1333	pGC1568	CPR integrant with CYP719A4	Chp. 3
GCY1755	GCY1333	pGC1569	CPR integrant with CYP719A5	Chp. 3
GCY1756	GCY1333	pGC1570	CPR integrant with CYP719A6	Chp. 3
GCY1757	GCY1333	pGC1571	CPR integrant with CYP719A7	Chp. 3
GCY1758	GCY1333	pGC1572	CPR integrant with CYP719A9	Chp. 3
GCY1759	GCY1333	pGC1573	CPR integrant with CYP719A11	Chp. 3
GCY1760	GCY1333	pGC1574	CPR integrant with CYP719A13	Chp. 3
GCY1761	GCY1333	pGC1575	CPR integrant with CYP719A14	Chp. 3
GCY1762	GCY1333	pGC1576	CPR integrant with CYP719A17	Chp. 3
GCY1763	GCY1333	pGC1577	CPR integrant with CYP719A18	Chp. 3
GCY1764	GCY1333	pGC1578	CPR integrant with CYP719A19	Chp. 3
GCY1765	GCY1333	pGC1579	CPR integrant with CYP719A21	Chp. 3
GCY1766	GCY1333	pGC1580	CPR integrant with CYP719A22	Chp. 3
GCY1767	GCY1333	pGC1581	CPR integrant with CYP719A23	Chp. 3
GCY1768	GCY1333	pGC1582	CPR integrant with CYP719A24	Chp. 3
GCY1769	GCY1333	pGC1583	CPR integrant with CYP719A25	Chp. 3
GCY1770	GCY1333	pGC1584	CPR integrant with CYP719B1	Chp. 3
GCY1771	GCY1333	pGC1585	CPR integrant with <i>Ps</i> SPS	Chp. 3
GCY1772	GCY1317	pGC965	CPR, CFS integrant with empty vector	Chp. 3
GCY1773	GCY1317	pGC1585	CPR, CFS integrant with <i>Ps</i> SPS	Chp. 3
GCY1774	GCY1317	pGC1536	CPR, CFS integrant with CYP719-Cma-4	Chp. 3
GCY1775	GCY1317	pGC1541	CPR, CFS integrant with CYP719-Gfl-3	Chp. 3
GCY1776	GCY1317	pGC1545	CPR, CFS integrant with CYP719-Ndo-2	Chp. 3
GCY1777	GCY1317	pGC1549	CPR, CFS integrant with CYP719-Ndo-6	Chp. 3
GCY1778	GCY1317	pGC1557	CPR, CFS integrant with CYP719-Sca-2	Chp. 3
GCY1779	GCY1317	pGC1562	CPR, CFS integrant with CYP719-Sdi-4	Chp. 3
GCY1780	GCY1317	pGC1565	CPR, CFS integrant with CYP719-Xsi-2	Chp. 3
GCY1781	GCY1317	pGC1566	CPR, CFS integrant with CYP719A1	Chp. 3
GCY1782	GCY1317	pGC1572	CPR, CFS integrant with CYP719A9	Chp. 3
GCY1783	GCY1317	pGC1575	CPR, CFS integrant with CYP719A14	Chp. 3
GCY1784	GCY1333	pGC1592, pGC1314	CPR integrant with <i>Ps</i> SPS	Chp. 3

GCY1785	GCY1333	pGC1591, pGC1314	CPR integrant with CYP719A14	Chp. 3
GCY1786	GCY1333	pGC1586, pGC1314	CPR integrant with CYP719-Cma-4	Chp. 3
GCY1787	GCY1333	pGC1587, pGC1314	CPR integrant with CYP719-Gfl-3	Chp. 3
GCY1788	GCY1333	pGC1588, pGC1314	CPR integrant with CYP719-Ndo-6	Chp. 3
GCY1789	GCY1333	pGC1589, pGC1314	CPR integrant with CYP719-Sca-2	Chp. 3
GCY1790	GCY1333	pGC1590, pGC1314	CPR integrant with CYP719-Sdi-4	Chp. 3
GCY1791	GCY1333	pGC1592, pGC1595	CPR integrant with PsSPS, CYP719A25	Chp. 3
GCY1792	GCY1333	pGC1591, pGC1595	CPR integrant with CYP719A14, CYP719A25	Chp. 3
GCY1793	GCY1333	pGC1586, pGC1595	CPR integrant with CYP719-Cma-4, CYP719A25	Chp. 3
GCY1794	GCY1333	pGC1587, pGC1595	CPR integrant with CYP719-Gfl-3, CYP719A25	Chp. 3
GCY1795	GCY1333	pGC1588, pGC1595	CPR integrant with CYP719-Ndo-6, CYP719A25	Chp. 3
GCY1796	GCY1333	pGC1589, pGC1595	CPR integrant with CYP719-Sca-2, CYP719A25	Chp. 3
GCY1797	GCY1333	pGC1590, pGC1595	CPR integrant with CYP719-Sdi-4, CYP719A25	Chp. 3
GCY1798	GCY1333	pGC1592, pGC1593	CPR integrant with PsSPS, CYP719-Cma-2	Chp. 3
GCY1799	GCY1333	pGC1591, pGC1593	CPR integrant with CYP719A14, CYP719- Cma-2	Chp. 3
GCY1800	GCY1333	pGC1586, pGC1593	CPR integrant with CYP719-Cma-4, CYP719-Cma-2	Chp. 3
GCY1801	GCY1333	pGC1587, pGC1593	CPR integrant with CYP719-Gfl-3, CYP719- Cma-2	Chp. 3
GCY1802	GCY1333	pGC1588, pGC1593	CPR integrant with CYP719-Ndo-6, CYP719-Cma-2	Chp. 3
GCY1803	GCY1333	pGC1589, pGC1593	CPR integrant with CYP719-Sca-2, CYP719- Cma-2	Chp. 3
GCY1804	GCY1333	pGC1590, pGC1593	CPR integrant with CYP719-Sdi-4, CYP719- Cma-2	Chp. 3
GCY1805	GCY1333	pGC1592, pGC1594	CPR integrant with PsSPS, CYP719-Sdi-1	Chp. 3
GCY1806	GCY1333	pGC1591, pGC1594	CPR integrant with CYP719A14, CYP719- Sdi-1	Chp. 3
GCY1807	GCY1333	pGC1586, pGC1594	CPR integrant with CYP719-Cma-4, CYP719-Sdi-1	Chp. 3
GCY1808	GCY1333	pGC1587, pGC1594	CPR integrant with CYP719-Gfl-3, CYP719- Sdi-1	Chp. 3
GCY1809	GCY1333	pGC1588, pGC1594	CPR integrant with CYP719-Ndo-6, CYP719-Sdi-1	Chp. 3
GCY1810	GCY1333	pGC1589, pGC1594	CPR integrant with CYP719-Sca-2, CYP719- Sdi-1	Chp. 3
GCY1811	GCY1333	pGC1590, pGC1594	CPR integrant with CYP719-Sdi-4, CYP719- Sdi-1	Chp. 3

GCY1812	GCY1270	pGC1592, pGC1314	CPR, TNMT integrant with PsSPS	Chp. 3
GCY1813	GCY1270	pGC1591, pGC1314	CPR, TNMT integrant with CYP719A14	Chp. 3
GCY1814	GCY1270	pGC1586, pGC1314	CPR, TNMT integrant with CYP719-Cma-4	Chp. 3
GCY1815	GCY1270	pGC1587, pGC1314	CPR, TNMT integrant with CYP719-Gfl-3	Chp. 3
GCY1816	GCY1270	pGC1588, pGC1314	CPR, TNMT integrant with CYP719-Ndo-6	Chp. 3
GCY1817	GCY1270	pGC1589, pGC1314	CPR, TNMT integrant with CYP719-Sca-2	Chp. 3
GCY1818	GCY1270	pGC1590, pGC1314	CPR, TNMT integrant with CYP719-Sdi-4	Chp. 3
GCY1819	GCY1270	pGC1592, pGC1595	CPR, TNMT integrant with PsSPS, CYP719A25	Chp. 3
GCY1820	GCY1270	pGC1591, pGC1595	CPR, TNMT integrant with CYP719A14, CYP719A25	Chp. 3
GCY1821	GCY1270	pGC1586, pGC1595	CPR, TNMT integrant with CYP719-Cma-4, CYP719A25	Chp. 3
GCY1822	GCY1270	pGC1587, pGC1595	CPR, TNMT integrant with CYP719-Gfl-3, CYP719A25	Chp. 3
GCY1823	GCY1270	pGC1588, pGC1595	CPR, TNMT integrant with CYP719-Ndo-6, CYP719A25	Chp. 3
GCY1824	GCY1270	pGC1589, pGC1595	CPR, TNMT integrant with CYP719-Sca-2, CYP719A25	Chp. 3
GCY1825	GCY1270	pGC1590, pGC1595	CPR, TNMT integrant with CYP719-Sdi-4, CYP719A25	Chp. 3
GCY1826	GCY1270	pGC1592, pGC1593	CPR, TNMT integrant with PsSPS, CYP719- Cma-2	Chp. 3
GCY1827	GCY1270	pGC1591, pGC1593	CPR, TNMT integrant with CYP719A14, CYP719-Cma-2	Chp. 3
GCY1828	GCY1270	pGC1586, pGC1593	CPR, TNMT integrant with CYP719-Cma-4, CYP719-Cma-2	Chp. 3
GCY1829	GCY1270	pGC1587, pGC1593	CPR, TNMT integrant with CYP719-Gfl-3, CYP719-Cma-2	Chp. 3
GCY1830	GCY1270	pGC1588, pGC1593	CPR, TNMT integrant with CYP719-Ndo-6, CYP719-Cma-2	Chp. 3
GCY1831	GCY1270	pGC1589, pGC1593	CPR, TNMT integrant with CYP719-Sca-2, CYP719-Cma-2	Chp. 3
GCY1832	GCY1270	pGC1590, pGC1593	CPR, TNMT integrant with CYP719-Sdi-4, CYP719-Cma-2	Chp. 3
GCY1833	GCY1270	pGC1592, pGC1594	CPR, TNMT integrant with PsSPS, CYP719- Sdi-1	Chp. 3
GCY1834	GCY1270	pGC1591, pGC1594	CPR, TNMT integrant with CYP719A14, CYP719-Sdi-1	Chp. 3
GCY1835	GCY1270	pGC1586, pGC1594	CPR, TNMT integrant with CYP719-Cma-4, CYP719-Sdi-1	Chp. 3
GCY1836	GCY1270	pGC1587, pGC1594	CPR, TNMT integrant with CYP719-Gfl-3, CYP719-Sdi-1	Chp. 3
GCY1837	GCY1270	pGC1588, pGC1594	CPR, TNMT integrant with CYP719-Ndo-6, CYP719-Sdi-1	Chp. 3
GCY1838	GCY1270	pGC1589, pGC1594	CPR, TNMT integrant with CYP719-Sca-2, CYP719-Sdi-1	Chp. 3

GCY1839	GCY1270	pGC1590, pGC1594	CPR, TNMT integrant with CYP719-Sdi-4, CYP719-Sdi-1	Chp. 3
GCY1840	GCY1440	pGC965, pGC1314	DHS pathway integrant with empty vector	Chp. 3
GCY1841	GCY1440	pGC1592, pGC1595	DHS pathway integrant with PsSPS, CYP719A25	Chp. 3
GCY1842	GCY1440	pGC1587, pGC1593	DHS pathway integrant with CYP719-Gfl-3, CYP719-Cma-2	Chp. 3
GCY1843	GCY1440	pGC1588, pGC1593	DHS pathway integrant with CYP719-Ndo-6, CYP719-Cma-2	Chp. 3
GCY1844	GCY1440	pGC1589, pGC1593	DHS pathway integrant with CYP719-Sca-2, CYP719-Cma-2	Chp. 3
GCY1845	GCY1440	pGC1590, pGC1593	DHS pathway integrant with CYP719-Sdi-4, CYP719-Cma-2	Chp. 3
GCY1846	GCY1440	pGC1587, pGC1594	DHS pathway integrant with CYP719-Gfl-3, CYP719-Sdi-1	Chp. 3
GCY1847	GCY1440	pGC1588, pGC1594	DHS pathway integrant with CYP719-Ndo-6, CYP719-Sdi-1	Chp. 3
GCY1848	GCY1440	pGC1589, pGC1594	DHS pathway integrant with CYP719-Sca-2, CYP719-Sdi-1	Chp. 3
GCY1849	GCY1440	pGC1590, pGC1594	DHS pathway integrant with CYP719-Sdi-4, CYP719-Sdi-1	Chp. 3

Chp.3 refers to Chapter 3 of this work

## 4 Chapter 4: *De novo* Production of Norlaudanosoline for Dihydrosanguinarine Synthesis in Yeast

**Adapted from:** Narcross, L., Pyne, M., Kevvai, K., Siu, K.H., Dueber, J.E., and Martin, V.J.J. Human monoamine oxidase A improves benzyloquinoline alkaloid yields in yeast. Manuscript in preparation.

### 4.1 Abstract

The benzyloquinoline alkaloid (BIA) family of natural products comprises over 2,500 members, including the pharmaceuticals morphine, codeine, noscapine, glaucine, and papaverine as well as the antibiotics sanguinarine and chelerythrine used in animal husbandry. Agricultural cultivation can currently supply demand for the BIAs that accumulate in plants, but broader access to the entire BIA family would open new avenues of research and commercialization. Microbial synthesis presents an attractive option due to cheap feedstock, genetic tractability, and ease of scale-up; platforms for BIA synthesis have been developed using both *Escherichia coli* and *Saccharomyces cerevisiae* as hosts. Previously, we reported titers of the key branch-point BIA reticuline of 4.6 g/L in yeast. Here, we identify that the strain also produced large quantities of side products due to the concerted promiscuity of the 2-oxoacid decarboxylase Aro10 and norcoclaurine synthase. We circumvent this issue by leveraging human monoamine oxidase A in place of Aro10, resulting in the synthesis of the BIA norlaudanosoline. The norlaudanosoline route to reticuline synthesis is more selective and more efficient, enabling titers of 4.8 g/L reticuline while improving yields from 17 mg/g sucrose to 27 mg/g sucrose. Finally, reticuline synthesis is extended to dihydrosanguinarine through stepwise pathway construction. A titer of 630 mg/L dihydrosanguinarine and sanguinarine was achieved *de novo* in fed-batch fermentation, the highest reported titer of a BIA end-product by a factor of >300.

## 4.2 Introduction

Benzylisoquinoline alkaloids (BIAs) are a large class of plant secondary metabolites with broad applications across human health and agriculture. While some BIAs accumulate to a sufficient degree in plants to allow for commercial-scale production, most do not. A sustainable, scalable source of BIAs would expand access to this valuable class of natural products. One promising option is the introduction of BIA synthesis to a microbial host.

Microbial BIA synthesis from simple carbon sources has been established in *Escherichia coli*<sup>13</sup> and *Saccharomyces cerevisiae*<sup>14,75</sup>. (*S*)-Reticuline is a common target for *de novo* BIA synthesis, as it is the last common pathway intermediate in the morphine, sanguinarine, and noscapine pathways<sup>1</sup>. To date, the highest BIA titers reported are 0.16 g/L (*S*)-reticuline in *E. coli*<sup>74</sup> and 4.6 g/L (*S*)-reticuline in yeast<sup>16</sup>. The reticuline titers in yeast are particularly noteworthy because they approach the target set for commercial development of opioids in microbes: 5 g/L<sup>10</sup>.

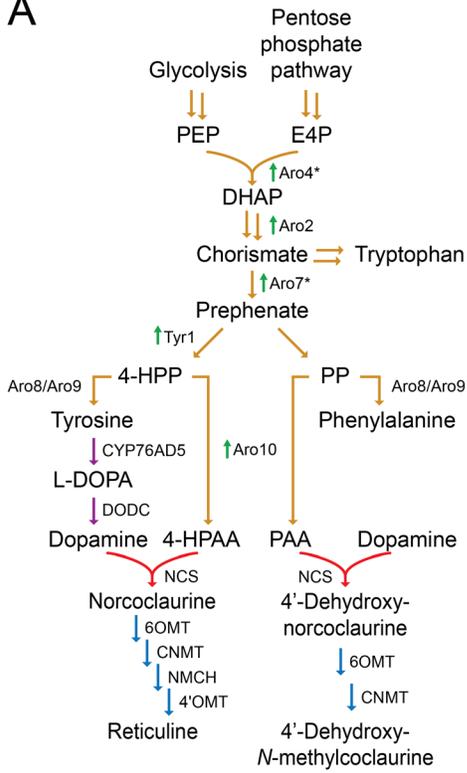
The committed step of BIA synthesis in plants is the formation of norcoclaurine through the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA), both derivatives of the aromatic amino acid (AA) pathway (Figure 4.1)<sup>1</sup>. Conveniently, 4-HPAA is a native metabolite in yeast, derived from decarboxylation of the tyrosine precursor 4-hydroxyphenylpyruvate (4-HPP) by the 2-oxoacid decarboxylase Aro10. Thus, overexpression of *ARO10* has been the preferred method of aldehyde generation in yeast<sup>7,16</sup>. However, the use of Aro10 for 4-HPAA synthesis in yeast is problematic because it catabolizes not only the 2-oxoacid that is the precursor of tyrosine, but also those of phenylalanine, tryptophan, methionine, leucine, isoleucine, and valine<sup>189,190</sup>. Compounding the effects of this promiscuity, norcoclaurine synthase (NCS) can accept a variety of aldehydes in addition to 4-HPAA<sup>39</sup>. Thus, *ARO10* expression can result in the synthesis of alternative BIA-like scaffolds. Although this is a promising source of novel

compounds, when the intention is to synthesize BIAs, these compounds are a drain on productivity. An alternative, more selective mechanism for aldehyde synthesis could thus further improve BIA production in yeast.

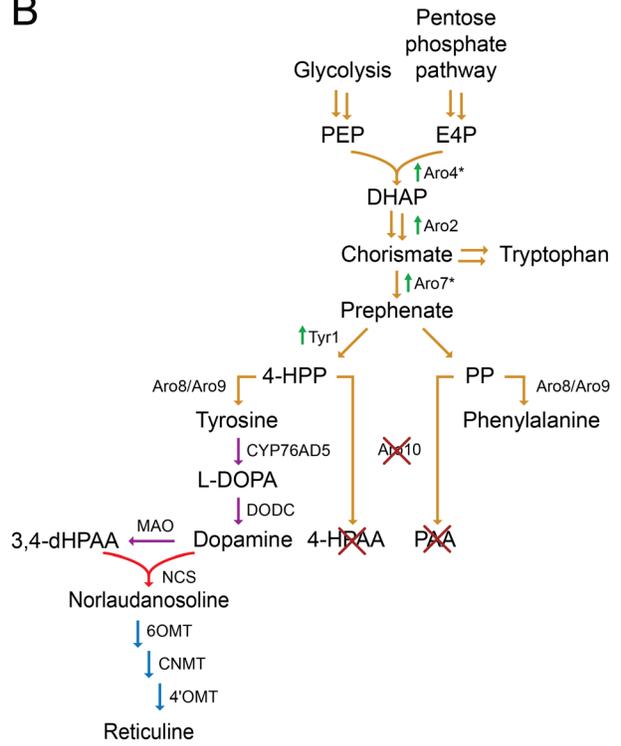
BIA synthesis in *E. coli* avoids norcoclaurine in favor of the norcoclaurine analog norlaudanosoline. Norlaudanosoline, made from dopamine and 3,4-dihydroxyphenylacetaldehyde (3,4-dHPAA), contains an extra hydroxyl group at the 3' position which eliminates the need for a cytochrome P450-catalyzed step in reticuline synthesis (Figure 4.1A). 3,4-dHPAA is generated by the oxidative deamination of dopamine. This dopamine-specific route to aldehyde synthesis presents an opportunity to bypass the promiscuity of Aro10. Thus far, 3,4-dHPAA synthesis has been catalyzed in *E. coli* by the enzyme monoamine oxidase from *Micrococcus luteus* (*MMAO*). Unfortunately, *MMAO* has been identified as a major bottleneck in *de novo* BIA synthesis in *E. coli*<sup>74</sup>, and other proposed alternatives have even less activity on dopamine<sup>89,90</sup>.

In this work, a yeast strain engineered to synthesize reticuline *via* norcoclaurine is retrofitted to synthesize reticuline *via* norlaudanosoline. 3,4-dHPAA synthesis from dopamine is achieved using human monoamine oxidase A (*HsMAO-A*). The norlaudanosoline route to BIA synthesis in yeast is capable of producing equivalent reticuline titers to the norcoclaurine route at a significantly higher yield while almost completely eliminating undesirable condensation products. Reticuline synthesis is then extended to the BIA dihydrosanguinarine, marking the first *de novo* synthesis of this BIA. This work highlights the remarkable speed at which microbial BIA synthesis is maturing as a technology.

A



B



- Native yeast enzyme
- Heterologous: dopamine synthesis
- Heterologous: committed step
- Heterologous: BIA derivitization

### Figure 4.1. Synthesis of benzyloquinoline alkaloids and analogs in yeast

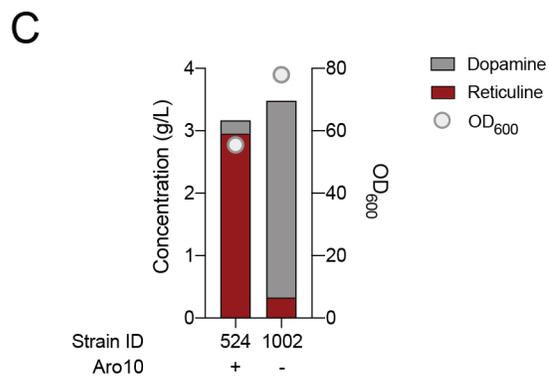
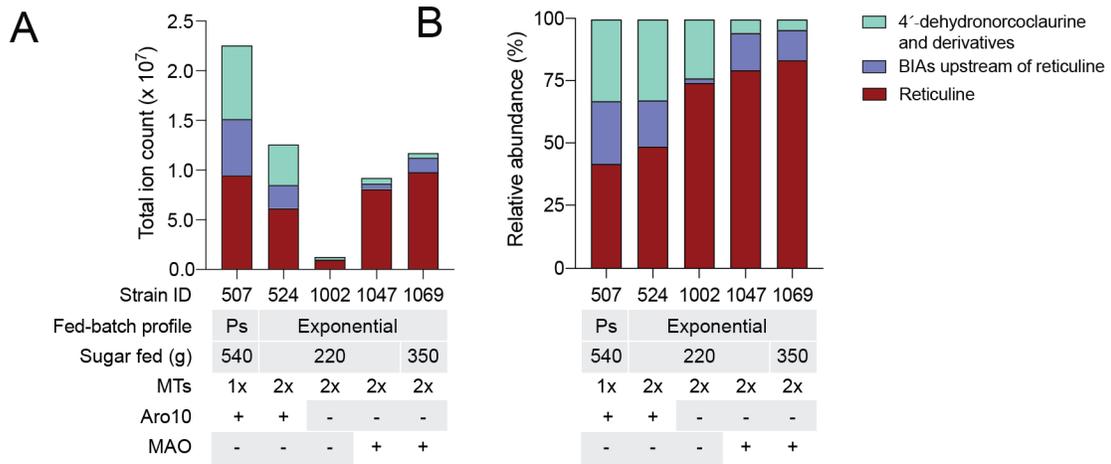
**(A)** Norcoclaurine and 4'-dehydronorcoclaurine synthesis in yeast. Norcoclaurine is derived from the condensation of dopamine and 4-HPAA, which are both derived from the tyrosine precursor 4-HPP. Flux is pushed down the aromatic amino acid synthesis pathway through overexpression of *ARO4*, *ARO2*, *ARO7*, and *TYR1*, indicated in green arrows. (Feedback-resistant enzyme variants are indicated with \*). *ARO10* overexpression increases 4-HPAA production from 4-HPP. Dopamine is synthesized through heterologous expression of *CYP76AD5* and *DODC*. Dopamine and 4-HPAA condensation is catalyzed by NCS. Reticuline is synthesized from norcoclaurine via one hydroxylation and three methyltransferase reactions. *ARO10* overexpression also increases PAA production from PP. Dopamine and PAA condensation is catalyzed by NCS. 4'-dehydro-*N*-methylcoclaurine is synthesized from 4'-dehydronorcoclaurine via two methyltransferase reactions. **(B)** Norlaudanosoline synthesis in yeast. *Aro10* knockout reduces 4-HPAA and PAA production. *MAO* expression results in 3,4-dHPAA synthesis. Dopamine and 3,4-dHPAA condensation is catalyzed by NCS. Reticuline is synthesized from norlaudanosoline via three methyltransferase reactions. Compound abbreviations: PEP, phosphoenolpyruvate; E4P, erythrose 4-phosphate; DHAP, 2-dehydro-3-deoxy-D-arabino-heptonoate 7-phosphate; 4-HPP, 4-hydroxyphenylpyruvate; 4-HPAA, 4-hydroxyphenylacetaldehyde; PP, phenylpyruvate, PAA, phenylacetaldehyde; 3,4-dHPAA, 3,4-dihydroxyphenylacetaldehyde. Enzyme abbreviations: DODC, dopamine decarboxylase; NCS, norcoclaurine synthase; MAO, monoamine oxidase; NMCH, *N*-methylcoclaurine hydroxylase.

## 4.3 Results

### 4.3.1 *ARO10* knockout eliminates *de novo* 4'-dehydroxynorcoclaurine synthesis

In 2020, we reported *de novo* synthesis of the benzyloisoquinoline alkaloid (BIA) reticuline in yeast at gram-per-liter scale<sup>16</sup>. This was a 57,000-fold improvement over our previous reported titers and a 30-fold improvement over the highest titers reported in a microbial host to date. In addition to 4.6 g/L of reticuline, strain LP507 produced several other BIAs and alternative scaffolds (Figure 4.2A and B). By peak area, reticuline represented just 42% of all BIAs and alternative scaffolds detectible by mass spectrometry. An additional 25% were pathway intermediates from the incomplete conversion of norcoclaurine to reticuline and 33% were a variety of alternative scaffolds. Eliminating these products would further improve BIA synthesis in yeast.

The expression of second copies of *6OMT*, *CNMT*, and *NMCH* in strain LP524 reduced the relative abundance of intermediates produced as a total percentage of condensation products from 25% to 18% (Figure 4.2B, strain LP524). The aromatic AA pathway results in the synthesis of both 4-HPP and phenylpyruvate (PP) from prephenate (Figure 4.1A). Overexpression of *ARO10*, which improves synthesis of 4-HPAA and hence BIAs, also results in PAA synthesis<sup>189</sup>. The resulting PAA is then condensed with dopamine by norcoclaurine synthase (NCS) to form the undesired 4'-dehydroxynorcoclaurine<sup>39</sup>.



**Figure 4.2. Effects of *ARO10* knockout in yeast synthesizing benzyloisoquinoline alkaloids**

**(A)** Absolute abundance of condensation products in select fed-batch fermentations. Sugar feeding profiles were either a series of pulses controlled by off-gas analysis (Ps) or exponential with maintenance of a constant growth rate. End-point samples were taken and analyzed for condensation products derived from the aldehydes PAA (4'-dehydroxynorcoclaurine and derivatives), 4-HPAA, and 3,4-dHPAA (norcoclaurine/norlaudanosoline and derivatives, reticuline). **(B)** Relative abundance of condensation products in the same fed-batch fermentations. **(C)** Impact of *ARO10* knockout on dopamine, reticuline, and biomass in strain LP524. The native and heterologous copies of *ARO10* were knocked out of strain LP524 (strain LN1002) and both strains were grown in fed-batch fermentation with the same media and feeding profile. End-point samples were taken and analyzed for OD<sub>600</sub> and metabolite content. Abbreviations: MTs, methyltransferases; MAO, monoamine oxidase.

To ablate PAA synthesis, both the native and overexpressed copies of *ARO10* were knocked out of strain LP524, generating strain LN1002. Strains LP524 and LN1002 were both grown in sugar-limited fed-batch fermentations. Under these conditions, LP524 synthesized 3.0 g/L of reticuline and 0.2 g/L dopamine, while the Aro10 knockout synthesized 3.1 g/L of dopamine and 0.33 g/L reticuline (Figure 4.2C). Several 2-oxoacid decarboxylases have been identified in yeast<sup>70,190</sup>, but the 90% reduction in reticuline levels demonstrates the dominant contribution of Aro10 to 4-HPAA synthesis in the parent strain. Eliminating Aro10 activity also improved biomass accumulation in bioreactor; using the same amount of sugar and feeding profile, strain LN1002 grew to a final OD<sub>600</sub> of 78 compared to 55 for strain LP524. Mass spectrometry revealed almost complete ablation of 4'-dehydroxynorcoclaurine as well as its methylated derivatives (Figure 4.2A and Figure 4.2B).

#### **4.3.2 Expression of *HsMAO-A* for norlaudanosoline synthesis**

Strain LN1002, lacking Aro10, provides a clean starting point for exploring an alternative route to reticuline synthesis. We opted to explore the synthesis of norlaudanosoline using the aldehyde 3,4-dHPAA, which is generated from dopamine by oxidative deamination (Figure 4.1B). Monoamine oxidase from *Micrococcus luteus* (*MMAO*) has been used extensively to produce 3,4-dHPAA for norlaudanosoline synthesis in *E. coli*<sup>11,74,191</sup>. However, a bottleneck at *MMAO* was observed at just 160 mg/L reticuline<sup>74</sup>. Instead, we turned to human monoamine oxidase A (*HsMAO-A*), which has previously been suggested as a candidate for norlaudanosoline synthesis in yeast<sup>73</sup>.

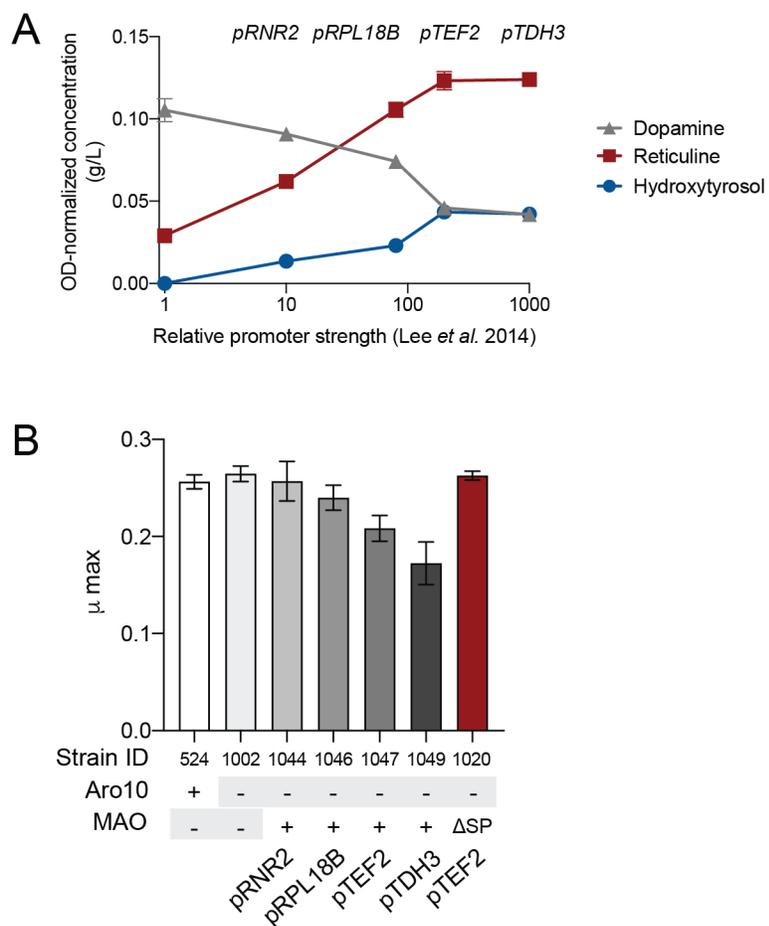
First, we explored the impact of promoter strength on *HsMAO-A*-driven BIA metabolite profile and strain fitness. *HsMAO-A* codon-optimized for expression in yeast was introduced to LN1002 under four promoters that have been well-characterized and span several orders of

magnitude of protein expression<sup>164</sup>. *HsMAO-A* activity was assessed by growing strains in deep-well plates and measuring their metabolite profile. *HsMAO-A* expression restored reticuline synthesis in LN1002 in a promoter strength-dependent manner (Figure 4.3A). Dopamine titers decreased with increasing promoter strength, which demonstrates that *HsMAO-A* is indeed functional and consuming dopamine. Hydroxytyrosol titers, which followed a similar trend as reticuline (Figure 4.3A), are an indirect readout of 3,4-dHPAA levels, because hydroxytyrosol synthesis results from the reduction of 3,4-dHPAA by yeast oxidoreductases (Figure 4.4A). Finally, successful restoration of reticuline synthesis indicated indirectly that dopamine and 3,4-dHPAA were condensed by NCS to produce norlaudanosoline, which was then methylated to form reticuline.

Growth rate assays, performed in parallel to the metabolite analysis, revealed that strains' maximum growth rate ( $\mu_{\max}$ ) decreases with increased *MAO* expression (Figure 4.3B). *HsMAO-A* localizes to the outer mitochondrial membrane in humans<sup>192</sup>. We investigated whether mitochondrial localization impacts growth rate. *HsMAO-A* harbors a 30 amino acid C-terminal signal peptide which is anchored in the mitochondrial membrane<sup>193</sup>. The signal peptide was removed from *HsMAO-A* (LN1020; *HsMAO-A* $\Delta$ SP) and this variant was expressed in yeast under control of the *TEF2* promoter. Removal of the signal peptide restored  $\mu_{\max}$  to the level observed in MAO-less strains (Figure 4.3B). However, MAO activity dropped in strain LN1020 compared to its full-length counterpart in strain LN1047, as indicated by reductions in hydroxytyrosol and reticuline titers and an increase in dopamine in deep well plate assays (Figure 4.7A, Supplemental Figures). Curiously, removal of the signal peptide did not impact mitochondrial localization of *HsMAO-A* (Figure 4.8, Supplemental Figures), leading us to conclude that the improved growth rate of LN1020 is independent of *HsMAO-A* $\Delta$ SP localization and likely results from the decrease

in MAO activity upon perturbation of its C-terminus. Oxidative deamination of dopamine also produces hydrogen peroxide, which may impose a burden to the cell. We also attempted to target *HsMAO-A* to the peroxisome using the peroxisomal targeting tag ePTS1 (LN1015; *HsMAO-A* $\Delta$ SP-p)<sup>194</sup>. However, there was no evidence that *HsMAO-A* was localized to the peroxisome, with or without removal of the signal peptide (Figure 4.8 and Figure 4.9, Supplemental Figures).

Since we observed a trade-off between growth rate and reticuline synthesis, and since reticuline synthesis plateaued between *MAO* expression from the *TEF2* and *TDH3* promoters, we chose the strain LN1047, expressing full-length *HsMAO-A* under the *TEF2* promoter, to carry forward for fed-batch fermentation.



**Figure 4.3. Promoter titration of *HsMAO-A* in an *ARO10* knockout background**

Strain LN1002, which lacks *Aro10*, was used to express no *HsMAO-A* (-) or full-length *HsMAO-A* under the control of *pRNR2*, *pRPL18B*, *pTEF2* and *pTDH3*. All strains were grown in deep well plates overnight and then used to determine **(A)** metabolite profile and **(B)** maximum growth rate. **(A)** Cultures were back-diluted into rich media in deep well plates and then analyzed for metabolite profile and final OD<sub>600</sub>. Metabolite profile of strains presented as g/L normalized to OD<sub>600</sub>. **(B)** Cultures were back-diluted into minimal media in microtiter plates in a plate reader, and maximum growth rate was obtained from the growth curves. Strains LP524 (parent of LN1002) and LN1020 (*HsMAO-A* $\Delta SP$  under the control of *pTEF2*) were also included in this growth assay.

### 4.3.3 Fed-batch fermentation with *HsMAO-A* driven by *pTEF2*

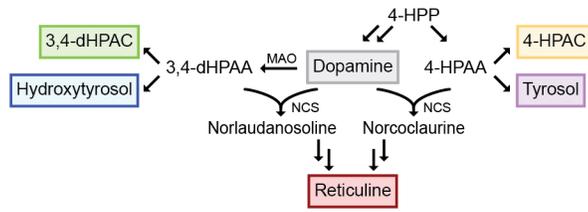
A key challenge in BIA production is yeast's propensity to transform precursor aldehydes into the corresponding fusel alcohols or acids<sup>16</sup>. Aldehydes may be oxidized or reduced by oxidoreductases depending on the redox environment of the cell<sup>70</sup>. In our prior study, significant residual 4-HPAA catabolic activity remained in strain LP507, even with the deletion of 7 oxidoreductases. We found that oxidation of 4-HPAA into 4-hydroxyphenylacetic acid (4-HPAC) was especially persistent. We circumvented this by using a fed-batch protocol that promoted periodic production and consumption of ethanol - a series of sugar pulses controlled through off-gas analysis - intended to maintain an environment where yeast has limited capacity to catabolize 4-HPAA<sup>195</sup>. Although this strategy was highly effective at preventing major 4-HPAC and tyrosol accumulation (Figure 4.4B), it is sub-optimal to intentionally ferment a fraction of carbon. Moreover, utilizing a simple, sugar-limited exponential feeding profile, which promotes an oxidative redox state, would be a superior long-term solution for BIA synthesis in industrial settings.

When a sugar-limited exponential feeding profile is used to grow strain LP524, which is similar in genotype to LP507, 4-HPAC accumulates in the fed-batch phase (Figure 4.10, Supplemental Figures). This has a negative impact on titers, yielding 3.0 g/L reticuline, 0.7 g/L tyrosol and 0.8 g/L 4-HPAC compared to 4.6 g/L reticuline, 0.3 g/L tyrosol and 0.4 g/L 4-HPAC respectively using the pulse-feeding profile (Figure 4.4B). Similarly to strain LP507, strain LP524 accumulates BIA intermediates upstream of reticuline as well as 4'-dehydronorcochlorine and methylated derivatives (Figure 4.2A and Figure 4.2B).

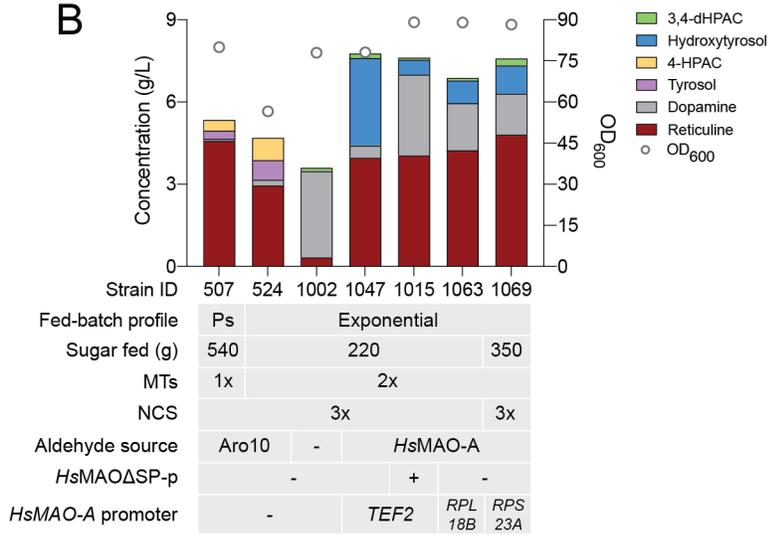
Strain LN1047, expressing *HsMAO-A* from the *TEF2* promoter, was grown in fed-batch fermentation using the same sugar-limited exponential feeding profile. The strain produced 4.0

g/L reticuline, a 33% improvement in titers over strain LP524 under the same conditions (Figure 4.4B). The strain additionally produced 3.2 g/L hydroxytyrosol, 0.44 g/L dopamine and 0.16 g/L 3,4-dHPAC, but very little 4'-dehydronorcoclaurine or methylated derivatives compared to LP524 (5% vs. 32% of total peak area as determined by LC-MS; Figure 4.2B). Pathway intermediates between norlaudanosoline/norcoclaurine and reticuline were also slightly lower (12% vs. 19% of total peak area). As a result, reticuline comprised 80% of condensation products by peak area, compared to the 42% of LP507 and the 49% of LP524. These initial results demonstrated that BIA synthesis through 3,4-dHPAA does indeed successfully reduce production of unwanted condensation products.

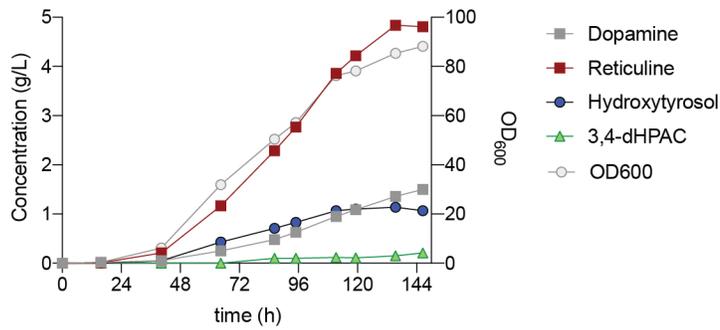
**A**



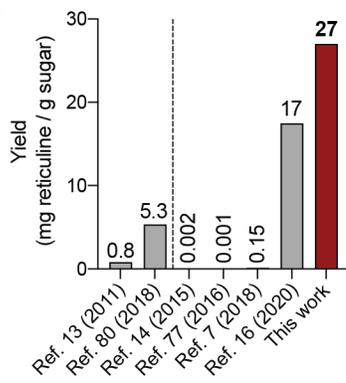
**B**



**C**



**D**



#### Figure 4.4. Optimization of reticuline production in fed-batch fermentation

(A) Pathways for fusel acid and alcohol production in strains synthesizing 4-HPAA and 3,4-dHPAA. (B) End-point analysis of quantifiable metabolites of BIA-producing strains grown in fed-batch fermentation. Strains were grown in batch phase until sugar was exhausted, then one of two sugar feeding fed-batch profiles was used - pulsed (Ps) or exponential. Pulsed fed-batch media contained 500 g/L sucrose, exponential fed-batch media contained 360 g/L sucrose. (C) Time course of fed-batch fermentation with LN1069, with *HsMAO-A* under control of *pRPS23A*. (D) *De novo* reticuline yields in selected publications expressed as mg reticuline/g sugar. On the left side of the dashed line are yields reported in *E. coli*; on the right side are those reported in yeast. Compound abbreviations: 4-HPP, 4-hydroxyphenylpyruvate; 4-HPAC, 4-hydroxyphenylacetate; 3,4-dHPAA, 3,4-dihydroxyphenylacetaldehyde; 3,4-dHPAC, 3,4-dihydroxyphenylacetate. Enzyme abbreviations: MTs, methyltransferases; NCS, norcoclaurine synthase.

#### 4.3.4 Optimization of norlaudanosoline synthesis through branch point balancing

Use of *HsMAO-A* resulted in a more targeted NCS-derived condensation metabolite profile, but the NCS substrates dopamine and 3,4-dHPAA accumulated instead. A delicate balance must be achieved in the synthesis of dopamine, 3,4-dHPAA and norlaudanosoline (Figure 4.4A). *HsMAO-A* activity for the conversion of dopamine to 3,4-dHPAA must not be too high or too low, and NCS activity must be sufficient to condense dopamine and 3,4-dHPAA before the former exits the cell or scavenging oxidoreductases reduce the latter to hydroxytyrosol. In the case of *MAO* expression from the *TEF2* promoter, 6 times more hydroxytyrosol accumulates than dopamine. This suggests that MAO activity is too strong when expression is driven by *TEF2*. Thus, two options were explored to alter the ratio of dopamine:3,4-dHPAA: using the MAO variant *HsMAO $\Delta$ SP-p*, which has reduced activity compared to wild-type *HsMAO-A*, and varying the strength of the promoter driving *HsMAO-A* expression.

In fed-batch fermentation *HsMAO $\Delta$ SP-p*, under control of the *TEF2* promoter, resulted in the synthesis of 4 g/L reticuline, this time with the additional accumulation of 3 g/L dopamine and 0.6 g/L hydroxytyrosol (Figure 4.4B). Under this scenario, MAO activity was too weak and resulted in an accumulation of 5 times more dopamine than hydroxytyrosol. The continued accumulation of both dopamine and hydroxytyrosol in fermentations with strains LN1047 and LN1015 suggested that NCS activity may not be sufficient for efficient condensation of dopamine and 3,4-dHPAA (Figure 4.4B). *NCS* expression negatively impacts strain health, which can be alleviated by peroxisomal sequestration<sup>194</sup>. Thus, an extra copy of peroxisomally-targeted NCS was introduced to strains moving forward. Next, we probed weaker promoters driving full-length *HsMAO-A* expression. With the *RPL18B* promoter driving *MAO* expression, strain LN1063 produced 4.25 g/L reticuline while continuing to produce 1.5 g/L dopamine and 0.5 g/L

hydroxytyrosol (Figure 4.4B). While representing a slight improvement in titer, dopamine was still 3 times as abundant as hydroxytyrosol, indicating that optimal MAO activity in this strain requires a promoter with strength between *pRPL18B* and *pTEF2*.

The yeast MoClo collection of characterized promoters does not include information on any promoters intermediate to *pRPL18B* and *pTEF2*<sup>164</sup>. A 2010 report from Canelas *et al.* contains RNASeq data of two common yeast strains, S288C and CEN.PK grown in both shake-flask and sugar-limited chemostat conditions<sup>196</sup>. We scanned this data for promoters whose strength was between those of *pRPL18B* and *pTEF2* in both CEN.PK and S288C (our strains are based on S288C) in both shake-flask and sugar-limited conditions. Three promoters, *pRPL39*, *pHTB1*, and *pRPS23A*, were selected. We confirmed that these promoters were indeed intermediate to *pRPL18B* and *pTEF2* in 96-well plate format, as determined by dopamine, hydroxytyrosol, and reticuline abundance (Figure 4.7B, Supplemental Figures). From this set of three new promoters, strain LN1069, expressing full length *HsMAO-A* from the *RPS23A* promoter and an additional copy of peroxisomally-targeted *NCS*, was selected for exponential fed-batch fermentation (Figure 4.4C)

In this experiment we used an additional 450 mL of fed-batch media compared to previous runs, increasing the total fed sugar from 220 g to 350 g. Under these conditions, strain LN1069 synthesized 4.8 g/L reticuline, with an additional 1.5 g/L dopamine and 1.0 g/L hydroxytyrosol (Figure 4.4B). This corresponds to an overall yield of 27 mg reticuline/g sucrose (Figure 4.4D). By mass spectrometry, reticuline comprised 83% of detectable condensation products, with another 12% attributable to pathway intermediates and 5% to 4'-dehydronorcoclaurine (Figure 4.2B). Strain LN1069, synthesizing reticuline *via* norlaudanosoline, represents an improvement of reticuline titers (4.8 vs 4.6 g/L), reticuline yield (27 vs 17 mg/g sucrose), and reticuline as a

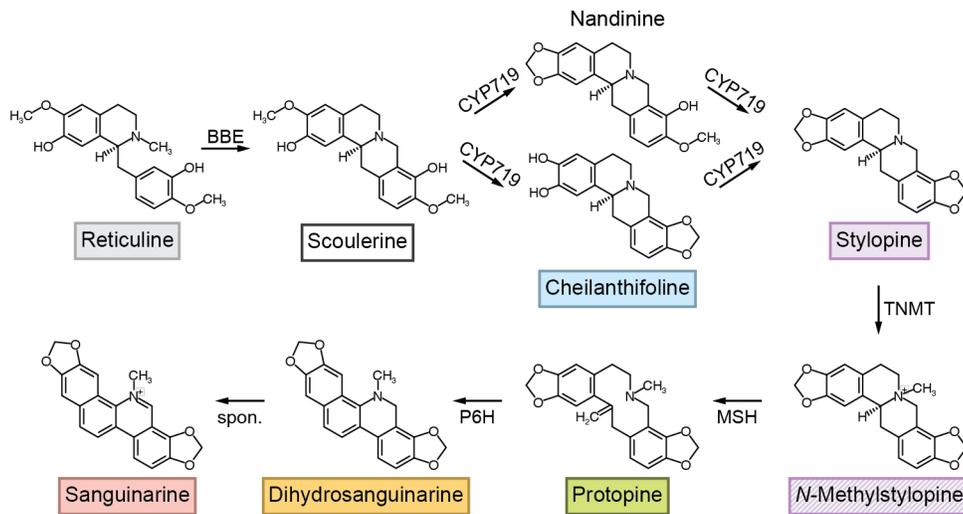
percentage of condensation products (84% vs 42%) over our previous work synthesizing reticuline *via* norcoclaurine.

#### 4.3.5 *De novo* dihydrosanguinarine synthesis *via* norlaudanosoline

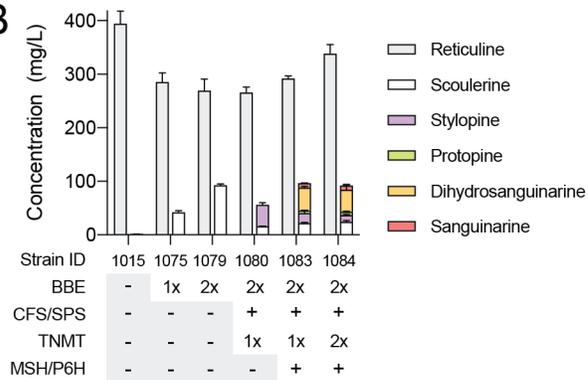
We previously published the reconstitution of a partial dihydrosanguinarine pathway in yeast<sup>23</sup>. Dihydrosanguinarine synthesis from reticuline requires six enzymes: berberine bridge enzyme (BBE) converts reticuline to scoulerine; two cytochromes P450 convert scoulerine to stylophine; tetrahydroprotoberberine *N*-methyltransferase (TNMT) expression results in *N*-methylstylophine synthesis; two more cytochromes P450 convert *N*-methylstylophine to dihydrosanguinarine (Figure 4.5A). We identified a pathway bottleneck between scoulerine and stylophine synthesis, which was resolved in a follow-up work in 2016<sup>197</sup>. Here, all six enzymes are introduced to yeast synthesizing reticuline from *de novo* norlaudanosoline (strain LN1015).

Introduction of one copy of *BBE* to strain LN1015 resulted in 15% conversion of reticuline to scoulerine in deep-well plates (Figure 4.5B, strain LN1075). A second copy of *BBE* improved conversion to 25% (LN1079). The next three pathway enzymes (cheilanthifoline synthase, CFS; stylophine synthase, SPS; tetrahydroprotoberberine *N*-methyltransferase, TNMT) were introduced into strain LN1079, generating strain LN1080 (Figure 4.5B). In our 2016 report, these three enzymes converted 100% of supplemented scoulerine to *N*-methylstylophine. Here, we observed a significant reduction in scoulerine titers (90 mg/L to 15 mg/L), but incomplete conversion of scoulerine to *N*-methylstylophine. Examination of LC-MS peak areas (Figure 4.5C) reveals some residual scoulerine and cheilanthifoline, indicating that in this strain, flux through the cytochromes P450 is somewhat insufficient. Conversion of stylophine to *N*-methylstylophine was also incomplete. Nevertheless, the reduction in scoulerine titers indicated that the strain could support the complete dihydrosanguinarine pathway.

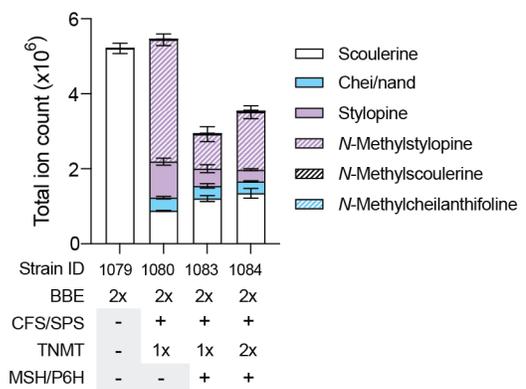
**A**



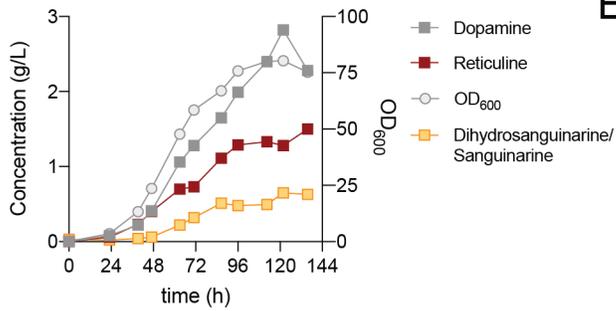
**B**



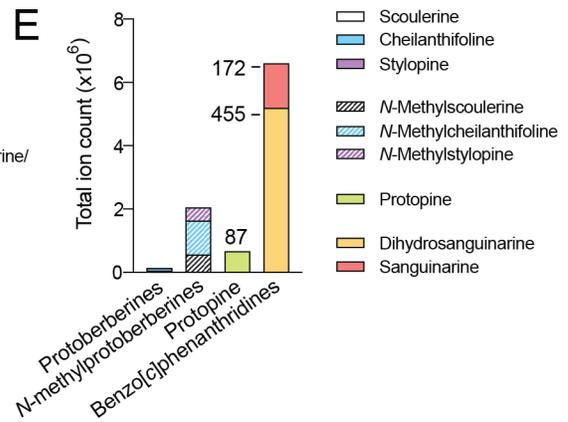
**C**



**D**



**E**



### Figure 4.5. *De novo* dihydrosanguinarine synthesis in yeast

**(A)** Dihydrosanguinarine pathway from reticuline. Conversion of reticuline to scoulerine by BBE generates the protoberberine scaffold. Scoulerine is converted to stylophine through parallel activities of two cytochromes P450 CYP719s through the intermediates nandinine and cheilanthifoline. *N*-methylation of stylophine by TNMT yields *N*-methylstylophine, which is productive for the dihydrosanguinarine pathway; TNMT can also methylate nandinine and cheilanthifoline. Hydroxylation of *N*-methylstylophine by MSH produces protopine. Further hydroxylation of protopine by P6H results in synthesis of the benzo[*c*]phenanthridine dihydrosanguinarine. Dihydrosanguinarine can be enzymatically oxidized to sanguinarine, but oxidation also occurs spontaneously. **(B)** Stepwise construction of a *de novo* dihydrosanguinarine strain. Pathway enzymes were integrated into LN1015, a strain of yeast synthesizing reticuline *via* norlaudanosoline. Metabolites were extracted and, where possible, quantified by LC-MS. **(C)** Metabolite profile of intermediates between scoulerine and *N*-methylstylophine. **(D)** Fed-batch fermentation of LN1084, a strain synthesizing dihydrosanguinarine *de novo*. Samples were regularly collected for OD<sub>600</sub> and metabolite analysis. **(E)** Metabolite profile of target and intermediate compounds at the final time point. Metabolites are organized by scaffold as described in **(A)**. Quantifiable metabolites are indicated in mg/L. Abbreviations: BBE, berberine bridge enzyme; CFS, cheilanthifoline synthase; SPS, stylophine synthase; TNMT, tetrahydroprotoberberine *N*-methyltransferase; MSH, *N*-methylstylophine hydroxylase; P6H, protopine 6-hydroxylase.

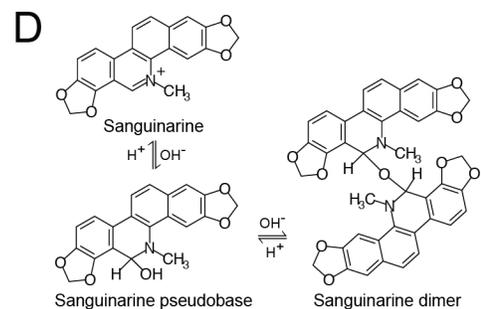
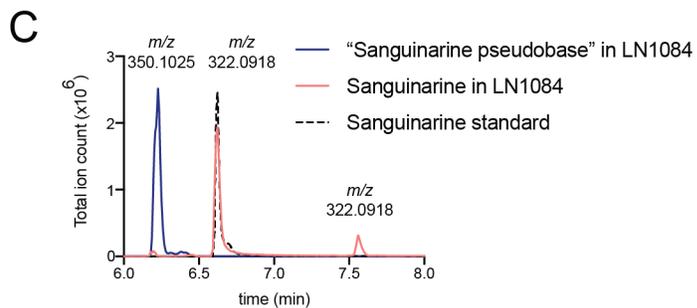
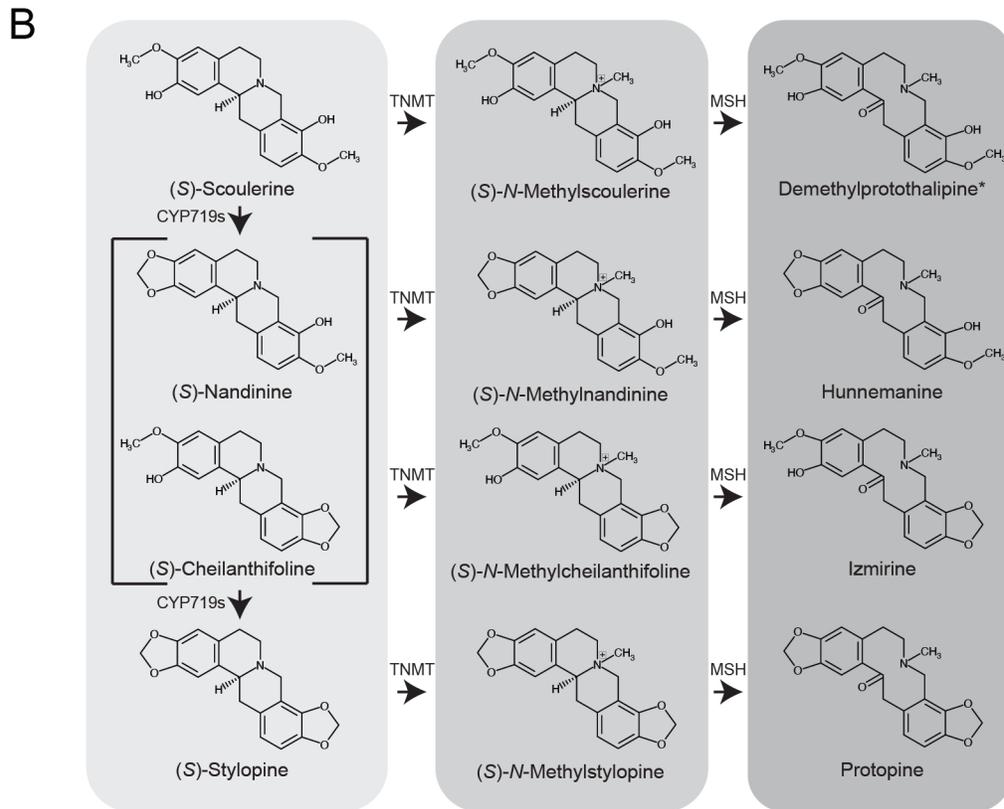
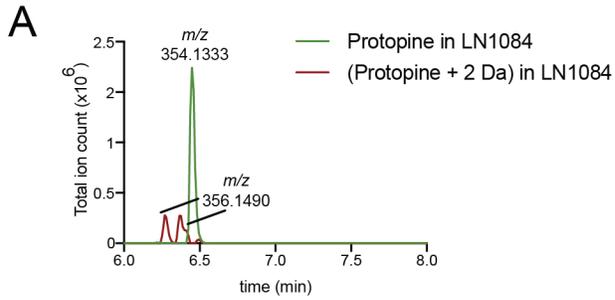
Finally, *N*-methylstylopine hydroxylase (MSH) and protopine 6-hydroxylase (P6H) were introduced to LN1080, generating a strain of yeast producing dihydrosanguinarine *de novo* from sugar (Figure 4.5B, LN1083). In deep-well plates, LN1083 synthesized 40 mg/L dihydrosanguinarine and an additional 7 mg/L sanguinarine, presumably through spontaneous oxidation. The impact of an additional copy of *TNMT* was also probed (strain LN1084). While there was no improvement in dihydrosanguinarine titers, there was an observable decrease in stylopine and an increase in *N*-methylstylopine, indicating that the extra copy of *TNMT* was effective (Figure 4.5C).

Growth of LN1084 in sugar-limited conditions in a bioreactor resulted in the synthesis of 455 mg/L dihydrosanguinarine and an additional 172 mg/L sanguinarine, for a combined output of approximately 630 mg/L (Figure 4.5D). The strain produced an additional 1.5 g/L reticuline and 87 mg/L protopine. In total, dihydrosanguinarine and sanguinarine synthesis comprised 27% of total quantifiable BIAs. The accumulation of 1.5 g/L reticuline signifies that BBE continues to be the rate-limiting step in flux through the entire dihydrosanguinarine pathway, as was previously observed in deep-well plates. Scoulerine, cheilanthifoline, and stylopine were almost undetectable, while all three *N*-methylated equivalents were observed (Figure 4.5E).

In our previous work, BIA quantification of fed-batch samples was performed using cell-free supernatant<sup>16</sup>. However, we noticed that both sanguinarine and dihydrosanguinarine were more readily detected in fermentation broth containing cells; in fact, dihydrosanguinarine was not detected in supernatant at all (Figure 4.11, Supplemental Materials). In plants, dihydrosanguinarine is localized intracellularly, while sanguinarine is extracellular but primarily associated with cell walls<sup>17</sup>. Sanguinarine salts are readily soluble in water and methanol<sup>198</sup>, but dihydrosanguinarine is considered “non-basic” and is primarily soluble in nonpolar solvents<sup>199</sup>. Accordingly, extraction

conditions were developed that balanced sample throughput, dilution accuracy, and use of C18-compatible solvents.

Several novel peaks were observed in addition to known pathway intermediates. Two distinct peaks with exact mass 356.1490 and predicted molecular formula  $C_{20}H_{21}NO_5$ , were detected (Figure 4.6A). This could correspond to protopine with two additional hydrogens. The fact that there are two peaks suggests that these compounds are derived from the activity of MSH on *N*-methylnandinine and *N*-methylcheilanthifoline to produce protopine-like compounds (Figure 4.6B). Combined, these peaks' area comprises roughly 1/3 of protopine; assuming similar ionization efficiency, these could correspond to 30 mg/L of lost carbon. An unknown compound with an exact mass of 350.1025 and predicted molecular formula  $C_{21}H_{17}NO_5$  was also observed (Figure 4.6C). This could be a hydroxylated dihydrosanguinarine. Enzymes in the CYP82 family, which includes MSH and P6H, are known to hydroxylate dihydrosanguinarine<sup>200</sup>. However, this activity has not been ascribed to MSH or P6H. Alternatively, 6-hydroxydihydrosanguinarine, also called sanguinarine pseudobase, is a naturally-occurring form of sanguinarine that exists in equilibrium with sanguinarine and sanguinarine dimer at physiological pH (Figure 4.6D)<sup>201</sup>. Sanguinarine pseudobase is not observed in authentic sanguinarine standard (Figure 4.6C), but this standard is stored in acidic conditions (0.1% formic acid in MeOH). Finally, a second peak corresponding to the same exact mass as sanguinarine but eluting 1 minute later is observed only in fermentation samples (Figure 4.6C). This peak could correspond to the in-source fragmentation of sanguinarine dimer. If these two peaks are indeed related to sanguinarine, the peak areas could correspond to an additional 150 mg/L of sanguinarine in fed-batch samples. These novel peaks should be subjected to MS<sup>2</sup> analysis in the future to provide further insight as to their identity.



**Figure 4.6. Unidentified metabolites in fermentation samples of strain LN1084**

**(A)** Broth from the fed-batch fermentation of strain LN1084, synthesizing *de novo* dihydrosanguinarine, was extracted and analyzed by HPLC-FT-MS. Two unknown peaks with exact  $m/z$   $[M+H]^+$  of 356.1490 were observed, corresponding to a molecular formula of  $C_{20}H_{21}NO_5$ . **(B)** Dihydrosanguinarine synthesis proceeds from scoulerine through protopine. Scoulerine is converted to stylophine through the action of two cytochromes P450. Stylophine is *N*-methylated by TNMT and then hydroxylated to form the protopine backbone by MSH. Accumulation of scoulerine, nandinine, or cheilanthifoline results in their *N*-methylation. Hypothetically, these compounds could also be hydroxylated by MSH; hunnemanine and izmirine have molecular formulas of  $C_{20}H_{21}NO_5$ . **(C)** Broth from the fed-batch fermentation of strain LN1084, synthesizing *de novo* dihydrosanguinarine, was extracted and analyzed by HPLC-FT-MS. A peak with exact  $m/z$   $[M+H]^+$  of 350.1025 was observed, as was an additional peak with exact  $m/z$   $[M+H]^+$  of 332.0918. **(D)** Interconversion between sanguinarine quaternary cation, sanguinarine pseudobase, and sanguinarine dimer. Sanguinarine pseudobase has a molecular formula of  $C_{21}H_{17}NO_5$ , which corresponds to exact  $m/z$   $[M+H]^+$  of 350.1025.

## 4.4 Discussion

In 2020, we reported the first gram-per-liter-scale synthesis of benzyloisoquinoline alkaloids (BIAs) in a microbial host<sup>16</sup>. Here, we demonstrate that the Ehrlich pathway enzyme Aro10 is promiscuous and contributes to side-product formation. In order to avoid this promiscuity, we re-route BIA synthesis through norlaudanosoline using the enzyme human monoamine oxidase A (*HsMAO-A*). The MAO-enabled route to reticuline synthesis results in slightly higher titers while also significantly improving yields and dramatically reducing unwanted condensation products. Finally, we highlight the potential of our reticuline platform strain by introducing a complete pathway to dihydrosanguinarine synthesis. In total, we synthesized 627 mg/L dihydrosanguinarine and sanguinarine, representing the highest titer of a BIA end-product to date.

The reticuline yield reported here is among the highest of any aromatic amino acid (AA) pathway-derived compound made in yeast<sup>65</sup>. A yield of 27 mg reticuline/g sucrose is comparable to the highest reported yields for vanillin  $\beta$ -D-glucoside (32 mg/g glucose)<sup>202</sup>, shikimate (62.5 mg/g sucrose)<sup>203</sup>, *cis, cis*-muconic acid<sup>204</sup> (66 mg/g glucose), 2-phenylethanol (76.6 mg/g glucose)<sup>205</sup> and *p*-coumarate (155 mg/g glucose)<sup>67</sup>, which is remarkable considering the reticuline pathway's length and complexity. Reticuline synthesis requires the concerted activity of 7 heterologous enzymes following the high-level production of an aromatic AA; synthesis of the aforementioned compounds requires fewer heterologous enzymes and/or they are derived far upstream of an aromatic AA.

An obvious point of improvement with LN1069, the final reticuline strain assayed in this study, is the continued accumulation of dopamine and hydroxytyrosol. In addition to 4.8 g/L reticuline, the strain also produced a combined 2.5 g/L dopamine and hydroxytyrosol. Based on the importance of oxidoreductase knockouts for norcoclaurine synthesis<sup>16</sup>, it is likely that 3,4-

dHPAA is being reduced to hydroxytyrosol before NCS has a chance to condense it with dopamine. Additional copies of NCS, elimination of 3,4-dHPAA reduction, or both, should further improve reticuline titers in this background.

It is intriguing that the dominant fusel product generated from 3,4-dHPAA in aerobic sugar-limited conditions is the fusel alcohol hydroxytyrosol (Figure 4.4B). Fusel alcohol production is typically associated with a reductive redox environment<sup>70</sup>, but off-gas analysis clearly demonstrates that the strain is not in a fermentative state (Figure 4.10B, Supplemental Figures). Cells' redox environment is best described by the ratio of NAD<sup>+</sup>/NADH, which is ~0.5 during anaerobic growth and ~130 during aerobic sugar-limited conditions, even reaching ~300 in the cytoplasm – the exact values depend on the study but the trend holds across groups<sup>206,207</sup>. The ratio of NADP<sup>+</sup>/NADPH, however, does not follow this trend. NADP<sup>+</sup>/NADPH ratios are close to 1 during aerobic, sugar-limited conditions, with NADPH outnumbering NADP<sup>+</sup> 15:1 in the cytoplasm<sup>208</sup>. It is possible that an oxidoreductase using NADPH as a co-factor is reducing 3,4-dHPAA even in sugar-limited conditions. Six NADPH-utilizing oxidoreductases have already been knocked out in the parent strain with the intention of limiting 4-HPAA reduction<sup>16</sup>. It is likely that another such enzyme has comparatively strong activity on 3,4-dHPAA compared to 4-HPAA.

The aromatic AA pathway branches at chorismate to form tryptophan, and prephenate to form tyrosine and phenylalanine (Figure 4.1). A concern when making an aromatic AA-derived product is that excess flux down the aromatic AA pathway can overflow into other branches. By LC-MS peak area, our first-generation BIA strain LP507 synthesizes ~66% norcoclaurine, derived from the aldehyde 4-HPAA from the tyrosine branch, and ~33% 4'-dehydronorcoclaurine, derived from the aldehyde PAA from the phenylalanine branch. This is in spite of the overexpression of *TYR1* to push flux towards tyrosine. In this work, we eliminated PAA synthesis by knocking out

*ARO10* (Figure 4.1B, Figure 4.2B). However, we did not address the underlying issue of excess flux towards phenylalanine. 2-Phenylethanol production in yeast is subject to a mirror image of the same issue, in which PAA synthesis is desired and 4-HPAA is an unwanted side product. Recently, a promoter-based screen was performed to identify an expression strength for *TYR1* that decreased flux into the tyrosine branch without resulting in an auxotrophy, which improved the ratio of 2-phenylethanol:tyrosol in batch fermentation<sup>205</sup>. It is possible that such a strategy would prove useful in BIA synthesis; alternatively, more copies of *TYR1* could be introduced providing that this does not also result in a phenylalanine auxotrophy.

Biomass production begins to plateau towards the end of fed-batch fermentations of BIA-synthesizing strains, whether they produce norlaudanosoline (Figure 4.4C) or norcoclaurine (Figure 4.10, Supplemental Figures). As growth plateaus so does reticuline synthesis, suggesting that BIA production is growth-coupled. It is possible that MAO activity begins to stall towards the end of the fermentation with strain LN1069, since dopamine continues to rise while hydroxytyrosol and reticuline plateau (Figure 4.4C). A broader scan of promoters for *MAO* expression may identify a candidate that is more appropriate for maintaining BIA production throughout fluctuations in growth rate. Additionally, the cause of growth inhibition should be identified and addressed if possible. All strains investigated in this work experienced a growth plateau, which means that no single heterologous compound is responsible for this effect. It is likely that the heavily engineered background has multiple sources of stress that accumulate over the course of a fermentation.

We also modified a strain synthesizing reticuline *via* norlaudanosoline (LN1015) to produce dihydrosanguinarine *de novo* from sugar. In fed-batch fermentation, we detected 630 mg/L of dihydrosanguinarine and its oxidized derivative sanguinarine from sucrose. This is the

highest production of a commercial BIA by a factor of  $>300^7$ . Dihydrosanguinarine/sanguinarine comprised about 28% of total quantifiable BIAs in fed-batch fermentation of LN1084, with reticuline comprising an additional 67%. A BBE-catalyzed bottleneck between reticuline and scoulerine during *de novo* BIA synthesis has been previously described<sup>7</sup>. Recently, BBE was shown to be poorly soluble in *E. coli*. This was ameliorated by generating an N-terminal maltose-binding protein fusion, resulting in an 80-fold improvement in BBE activity *in vivo*<sup>209</sup>. This fusion protein strategy may prove beneficial in the present system as well, enabling even higher dihydrosanguinarine titers in the future.

## 4.5 Materials and Methods

### 4.5.1 Yeast and *E. coli* growth conditions

*E. coli* was grown in liquid Luria Broth (10 g/L peptone, 5 g/L yeast extract, 10 g/L sodium chloride, LB; Fisher Bioreagents) at 37°C with shaking at 200 rpm. *E. coli* transformations were selected on solid LB with 2% agar. Antibiotics (ampicillin, 100 µg/mL; kanamycin, 50 µg/mL; chloramphenicol, 50 µg/mL; hygromycin, 200 µg/mL) were supplied as necessary for plasmid maintenance.

For genetic manipulation, yeast was grown in liquid yeast peptone dextrose (20 g/L peptone, 20 g/L dextrose, 10 g/L yeast extract, YPD; Sigma Aldrich) at 30°C with shaking at 200 rpm. Yeast transformations with Cas9-containing plasmids were selected on solid YPD with 2% agar, 200 µg/mL G418, and 200 µg/mL hygromycin. For assessment of BIA synthesis in 96-well plate format, yeast was grown in 2x synthetic complete media (2x SC: 13.6 g/L Difco Yeast Nitrogen Base (YNB), 3.84 g/L yeast synthetic drop-out medium supplements without histidine (Millipore-Sigma), 152 mg/L histidine, 40 g/L sucrose) with shaking at 400 rpm in 96-well 2 mL deep-well plates overnight, followed by a 1:50 back dilution into fresh 2x SC for 3 days. For

assessment of yeast growth in 96-well plate format, strains were grown in 2x SC overnight with shaking at 400 rpm, followed by a 1:100 back dilution into YNB with 20 g/L sucrose supplemented with 76 mg/L methionine and 76 mg/L histidine. Prior to growth in fermenter, strains were transformed with a plasmid complementing methionine and histidine auxotrophies and selected on solid 1x SC media lacking histidine (SC-His). Pre-cultures for bioreactor experiments were grown in 50 mL of SC-His in a 250 mL flask with shaking at 200 rpm.

#### **4.5.2 Strain construction**

Gene knockouts and genomic integrations were introduced to yeast *via* CRISPR-directed homologous recombination. A plasmid harboring Cas9 and an empty guide RNA transcription cassette was linearized by *NotI/BsaI* double digestion (New England Biolabs) and transformed into yeast together with a linear piece of DNA containing the guide RNA sequence flanked on either side by homology to the plasmid, which resulted in *in vivo* plasmid assembly. Linear DNA containing the guide was generated by PCR. Guide RNAs for gene knockout were selected based on a combined score from the online tools Yeast CRISPRi<sup>210</sup> and CCTOP<sup>211</sup>. Gene knockouts were generated through co-transformation of a linear fragment with 40 bp homology to either side of the gene of interest. Gene integrations were targeted to genomic regions previously identified to promote high-level *GFP* expression<sup>212,213</sup>. Gene integrations were introduced either as pre-cloned promoter-gene-terminator cassettes or as individual promoters, genes, and terminators containing 40 bp of overlap between parts. Integrations were targeted to a region of interest in *trans* through co-transformation of ~600 bp regions of homology to the genome, with 40 bp of homology to common linker sequences present at the 5' and 3' ends of gene cassettes. All transformations were performed using a standard lithium acetate/salmon sperm heat shock protocol. Yeast strains were

cured of Cas9-containing plasmids between rounds of transformations by sub-streaking on solid YPD.

The vector pGC1899, harboring expression cassettes for *SdiCFS*, *NdoSPS*, and *PsTNMT*, was constructed *via* Golden Gate assembly. Type IIS enzymes were purchased from Thermo Fisher Scientific, T7 ligase and T4 ligase buffer were purchased from New England Biolabs. Golden Gate reactions were performed as described in the Yeast Toolkit (YTK) system<sup>164</sup>. Assemblies containing *SdiCFS* were performed using “end on ligation”, due to the presence of an internal *BsaI* site.

#### **4.5.3 Growth curves and determination of maximum growth rate**

Yeast was grown in 96-well plates in YNB media supplemented with methionine and histidine, in a Tecan Sunrise plate reader. Plates were wrapped with Parafilm to prevent evaporation. OD<sub>595</sub> measurements were taken every 5 min for 48 hrs. Following background subtraction, values were normalized to starting OD, ln-transformed, and smoothed across a 20-minute interval. The slope of the curves across a 1-hour window was determined, and the maximum slope was considered  $\mu_{\max}$ .

#### **4.5.4 Fed-batch fermentation**

Fed-batch fermentations were performed in Applikon 3L BioBundle fermenters. pH was maintained at 4.5 using 4N NaOH, temperature was kept at 30 °C. Air flow was set to 1 L/min, dissolved oxygen was controlled at 35% air saturation by automatic adjustment of stirring rate. Off-gas composition (partial pressure of O<sub>2</sub> and CO<sub>2</sub>) was measured with a Tandem Multiplex gas analyzer. Precultures were grown at 30 °C for 36 hrs in 50 mL SC-His medium with shaking at 200 rpm. Cells were centrifuged for 10 min at 4000 g, washed once in 0.9% NaCl, and suspended in 50 mL 0.9% NaCl prior to inoculation in 950 mL of batch medium (initial OD<sub>600</sub>~0.2).

Following consumption of sugar as indicated by off-gas analysis, fed-batch phase was triggered with an initial feeding rate of 0.60 g/h sucrose, increased exponentially at a dilution rate of 0.025 h<sup>-1</sup>. Batch medium (per liter): 40 g sucrose, 6 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2.5 g/L KH<sub>2</sub>PO<sub>4</sub>, 1 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 mL vitamin stock, and 5 mL trace element stock. Feeding medium (per liter): 360 g sucrose, 60 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 15 g KH<sub>2</sub>PO<sub>4</sub>, 6 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 15 mL vitamin stock, and 15 mL trace element stock per liter. Vitamin stock (per liter): 2,500 mg myo-inositol, 100 mg calcium pantothenate, 100 mg thiamine hydrochloride, 100 mg pyridoxine, 100 mg nicotinic acid, 20 mg p-aminobenzoic acid, 5 mg biotin, and 5 mg folic acid. Trace element stock (per liter): 15 g Na<sub>2</sub>EDTA, 2.9 g CaCl<sub>2</sub>, 9.2 g ZnSO<sub>4</sub>·7H<sub>2</sub>O, 0.5 g CuSO<sub>4</sub>, 0.43 g MnSO<sub>4</sub>·H<sub>2</sub>O, 0.47 g CoCl<sub>2</sub>, 0.48 g Na<sub>2</sub>MoO<sub>4</sub>, and 5.1 g FeSO<sub>4</sub>·7H<sub>2</sub>O. A biomass conversion ratio of 0.59 g/L per OD<sub>600</sub> unit was determined by drying cells in pre-dried Falcon tubes in a 100°C oven overnight in triplicate.

#### **4.5.5 High pressure liquid chromatography analysis by ultraviolet absorbance (HPLC-UV) and mass spectrometry (HPLC-MS)**

Dopamine, hydroxytyrosol, tyrosol, 4-HPAC, 3,4-dHPAC, and reticuline were quantified by HPLC-UV using an Agilent 1200 HPLC system. Samples in 96-well plate format were diluted 1:2 with 100% acetonitrile (AcN) containing 0.1% trifluoroacetic acid (TFA), vortexed briefly, centrifuged for 5 min at 21,000 g, and then supernatant was analyzed. Supernatants from bioreactors were treated as above and then further diluted with 50% AcN/0.1% TFA from two- to 24-fold as appropriate to stay within the range of standard curves. Five µL of analyte was applied to an Eclipse XDB-C18 column (150 x 4.6 mm, 5 µm, Agilent Technologies) and separated using the following gradient at 1 mL/min: 0-10 min, 5-20% B; 10-15 min, 20-50% B; 15-15.1 min, 50-95% B; 15.1-25 min, 95% B; 25-28 min, 5% B where A was 0.1% TFA in water and B was 0.1% TFA in methanol. All compounds were detected at 280 nm.

Relative peak areas of BIAs and BIA-like scaffolds were assessed using an Agilent 6545 qTOF-MS. All samples were diluted 1:5 with 100% AcN containing 0.1% formic acid (FA), shaken briefly with a plate shaker, and then water containing 0.1% FA was added to bring the final AcN concentration to 15%. Samples were centrifuged for 10 min at 4,000 g, and then supernatant was analyzed. Supernatants were diluted as necessary to avoid saturation of the detector (up 100-fold for 96-well plate format and 1000-fold for bioreactor). Five  $\mu$ L of analyte was applied to a Zorbax Eclipse Plus C18 column (50 x 2.1 mm, 1.8  $\mu$ m, Agilent Technologies) and separated using the following gradient at 0.3 mL/min flow rate: 0-4 min, 2-10% B; 4-6 min, 10-85% B; 6-7 min, 85% B, 7-7.1 min, 85-2% B where A was 0.1% FA in water and B was 0.1% FA in AcN. The column was reequilibrated for 2 min in 2% B at 0.45 ml/min. Settings: column compartment, 30°C; sheath gas flow rate, 10 L/min; sheath gas temperature, 350°C; drying gas flow rate 12 L/min; drying gas temperature, 325°C; nebulizing gas, 55 psig.

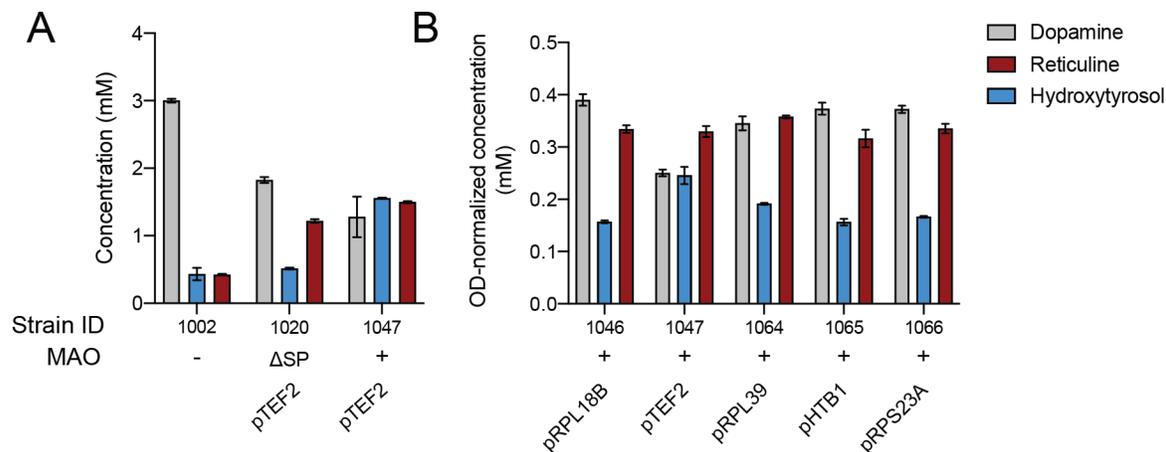
Dihydrosanguinarine pathway intermediates were assessed by HPLC-FT-MS using an Agilent 1290 Infinity II HPLC (Agilent Technologies) and a 7T-LTQ-FT-ICR (Thermo Fisher Scientific). Samples in 96-well plate format were extracted as per HPLC-qTOF-MS analysis. For determination of dihydrosanguinarine and sanguinarine concentration in bioreactors, fermentation broth was diluted 1:5 with 100% MeOH containing 0.1% FA, and then further diluted to 1:100 with 100% MeOH/0.1% FA prior to centrifugation for 5 min at 4,000 g. Samples were diluted as necessary in 100% MeOH/0.1% FA to stay within the linear range of the MS. Five  $\mu$ L of analyte was applied to a Zorbax Eclipse Plus C18 column (50 x 2.1 mm, 1.8  $\mu$ m, Agilent Technologies) and separated using the following gradient: 0.3 mL/min flow rate: 0-4 min, 2-10% B; 4-6 min, 10-85% B; 6-9 min, 85% B, 9-9.1 min, 85-2% B where A was 0.1% FA in water and B was 0.1% FA

in AcN. The column was reequilibrated for 5 min in 2% B at 0.3 ml/min. Settings: scanning range, 100-400  $m/z$ , resolution, 25,000; capillary voltage, 5 kV; source temperature, 350°C.

Sources of HPLC-UV/MS reagents: water and acetonitrile, Fisher Scientific; methanol, Sigma Aldrich; formic acid, Fluka; trifluoroacetic acid, Sigma Aldrich. Sources for authentic BIAs were: (*S*)-norcoclaurine, TRC Inc. (North York, Ontario, Canada); (*S*)-reticuline, gift from Dr. Peter Facchini; (*S*)-scoulerine, ChromaDex (Irvine, Ca, USA); (*S*)-stylophine, ChromaDex (Irvine, Ca, USA); protopine, TRC Inc.; sanguinarine, Sigma. Dihydrosanguinarine was derived from sanguinarine through sodium borohydride reduction<sup>184</sup>.

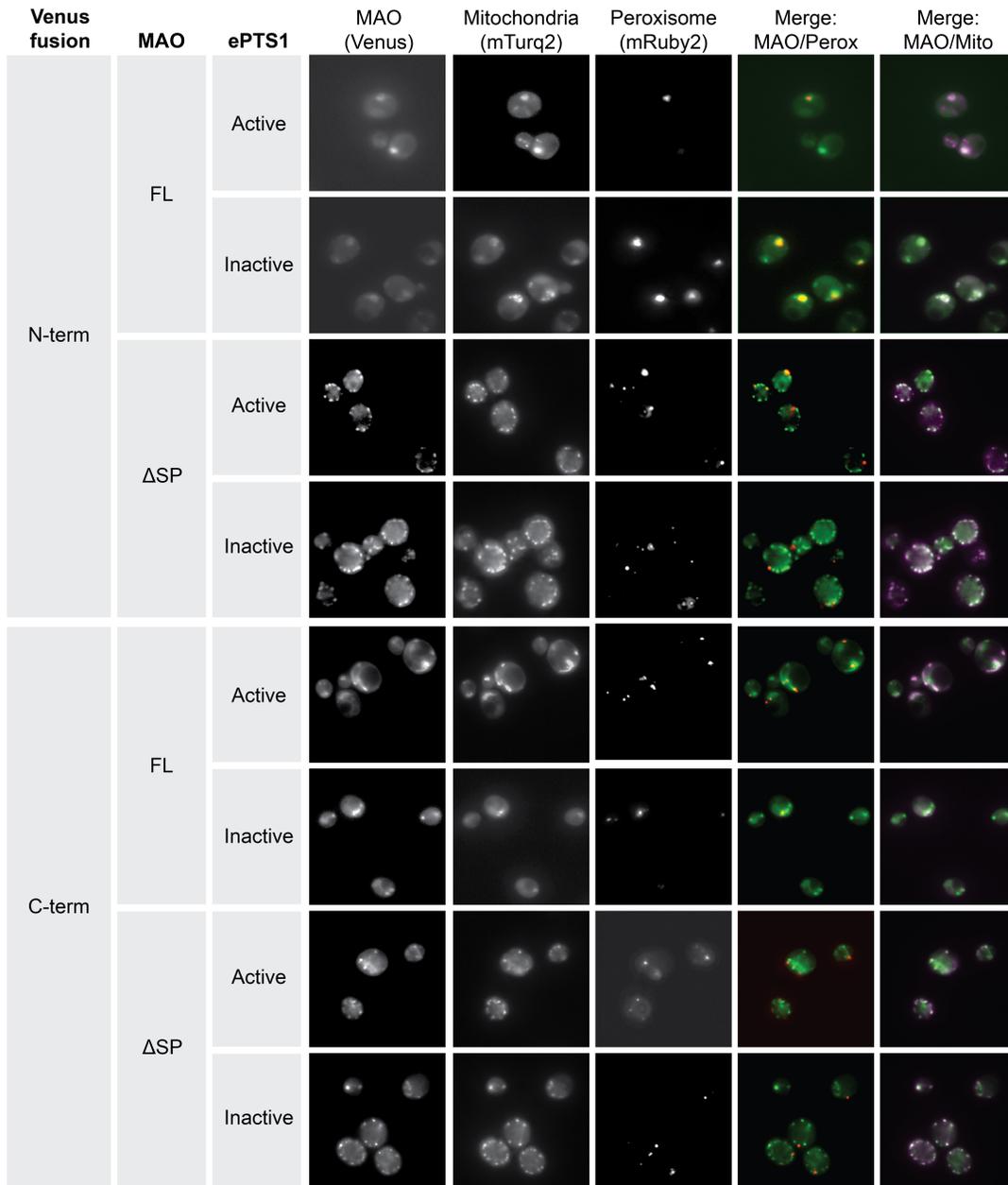
## 4.6 Supplemental materials for Chapter 4

### 4.6.1 Supplemental figures



**Figure 4.7. Metabolite profile of strains expressing *HsMAO-A* from various promoters**

*HsMAO-A*, with (+) or without its C-terminal signal peptide ( $\Delta$ SP), was integrated into strain LN1002 (*aro10* $\Delta$ ) under a variety of promoter strengths. Strains were grown in rich media in deep well plates and then analyzed for dopamine, reticuline, and hydroxytyrosol content by LC-UV. **(A)** Strain LN1020, harboring *HsMAO-A* without a C-terminal signal peptide, produces more dopamine, less hydroxytyrosol, and less reticuline than strain LN1047, harboring full-length *HsMAO-A*. **(B)** Strains LN1064, LN1065, and LN1066, expressing *HsMAO-A* from the promoters of *RPL39*, *HTB1*, and *RPS23A*, respectively, have a metabolite profile that resembles strains LN1046 and LN1047. Metabolite content is represented normalized to final OD<sub>600</sub>.



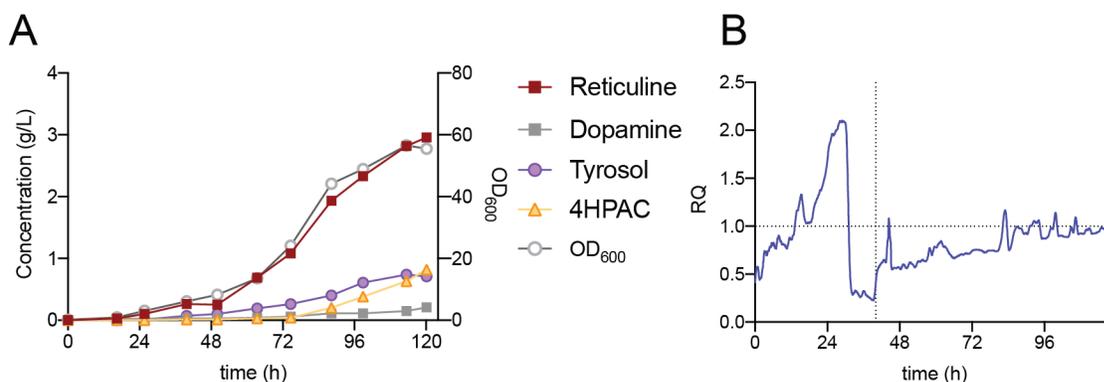
**Figure 4.8. Visualization of *HsMAO-A* localization by fluorescence microscopy**

The fluorescent protein Venus was fused to the N-terminus (N-term) or C-terminus (C-term) of full-length *HsMAO-A* (FL) or *HsMAO* $\Delta$ SP ( $\Delta$ SP). At the extreme C-terminus of each fusion protein was an active ePTS1, targeting the cargo to the peroxisome, or an inactive ePTS1, expected to maintain cytosolic localization. Each fusion protein was expressed in strain yKSS001, which harbors *Su9ss-mTurq2* to visualize mitochondria and *mRuby-ePTS1* to visualize peroxisomes. Although mitochondrial morphology differs between strains expressing *HsMAO-A* and *HsMAO* $\Delta$ SP, mitochondrial co-localization is observed throughout. Further, no differences are observed between MAO variants with active or inactive ePTS1.

VioE fusion	MAO	ePTS1	<i>PEX5</i>	<i>pex5Δ</i>
Ctrl	-	Active		
		Inactive		
N-term	FL	Active		
		Inactive		
	ΔSP	Active		
		Inactive		
C-term	FL	Active		
		Inactive		
	ΔSP	Active		
		Inactive		

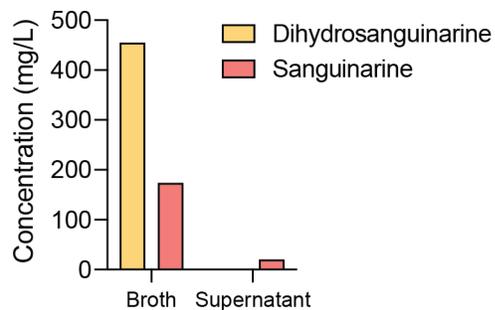
**Figure 4.9. Assessment of *Hs*MAO-A peroxisomal compartmentalization**

Agar spot assay to visualize efficiency of *Hs*MAO import to the peroxisome. Strains yPSG335 (*PEX5*) and yPSG340 (*pex5Δ*) express the violacein pathway proteins VioA and VioB in the cytosol, which convert tryptophan to IPA imine dimer (colorless). VioE, fused to *Hs*MAO-A or *Hs*MAOΔSP on the N- or C-terminus, is required to convert IPA imine dimer to prodeoxyviolacein (green). Sequestration of VioE to the peroxisome, *via* active ePTS1, prevents the formation of colour (Ctrl, active ePTS1, *PEX5*). Inactive ePTS1 reduces peroxisomal import dramatically<sup>214</sup>; such strains should be a deep green (Ctrl, inactive ePTS1, *PEX5*). *pex5Δ* strains, defective in peroxisomal import, are also deep green. A lack of difference in colour between active and inactive ePTS1 fusion proteins indicates that *Hs*MAO-A peroxisomal import is not effective. Expression of full-length *Hs*MAO-A, especially as a VioE N-terminal fusion, results in a grey colour. Abbreviations: ePTS1, enhanced peroxisomal targeting sequence 1; Ctrl, control; MAO, *Hs*MAO-A; FL, full-length; ΔSP, *Hs*MAOΔSP; N-term, N-terminal enzyme fusion; C-term, C-terminal enzyme fusion.



**Figure 4.10. Fed-batch fermentation of strain LP524**

Strain LP524 was grown in a bioreactor using an exponential, sugar-limited feeding profile. **(A)** Samples were regularly taken and assessed for biomass and for metabolite profile by LC-UV. **(B)** On-line CO<sub>2</sub> and O<sub>2</sub> off-gas traces were used to calculate the respiratory quotient (RQ) of the yeast using the calculation  $(\text{CO}_2 \text{ in exhaust} - \text{CO}_2 \text{ baseline}) / (\text{O}_2 \text{ baseline} - \text{O}_2 \text{ in exhaust}) = \text{RQ}$ . The eye is guided to RQ = 1 with a horizontal dotted line. Sugar fermentation results in RQ > 1, while aerobic carbon consumption results in RQ < 1. The beginning of fed-batch is indicated with a vertical dotted line at t = 40 hrs.



**Figure 4.11. Detection of dihydrosanguinarine and sanguinarine in fermentation broth and supernatant**

End-point samples of strain LN1084, grown in fed-batch fermentation, were extracted for metabolite analysis by LC-MS as cell-free supernatant or as broth. Dihydrosanguinarine was not detected in supernatant, vs. 455 mg/L in broth. Sanguinarine was detected at 20 mg/L in supernatant vs 174 mg/L in broth.

## 4.6.2 Supplemental materials and methods

### 4.6.2.1 Strain construction and growth media for HsMAO-A localization experiments

The *S. cerevisiae* strains for HsMAO-A localization and visualization experiments were BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) and BY4741 *pex5Δ*, ordered from Open Biosystems (GE Dharmacon). Wild-type yeast cultures were grown in YPD medium (10 g/L Bacto yeast extract, 20 g/L Bacto peptone, 20 g/L glucose). Selection of auxotrophic markers (*URA3*, *LEU2*, *HIS3*) was performed in synthetic complete (SC) medium (6.7 g/L Difco Yeast Nitrogen Base without amino acids (Spectrum Chemical); 2 g/L Drop-out Mix Synthetic minus appropriate amino acids, without Yeast Nitrogen Base (US Biological); 20 g/L glucose). All strains used are listed in Table 4.1

Golden Gate assembly reactions were transformed into chemically competent *E. coli* prepared from strain TG1 (Lucigen). Transformed cells were selected on LB containing the antibiotics chloramphenicol, ampicillin or kanamycin.

Yeast expression vectors were built using Golden Gate assembly as described in the Yeast Toolkit (YTK) system<sup>164</sup>. Integration into the yeast genome via homologous recombination at the *URA3* or *LEU2* locus was achieved by transformation of linearized plasmids (*NotI* digestion, NEB). All transformations were performed using a standard lithium acetate method, and cells were plated onto selective auxotrophic SC agar plates with 2% glucose. Individual colonies were picked as biological replicates directly from this transformation plate and grown independently for further analysis.

Chromosomal integrations of the fluorescent localization reporters mRuby2-ePTS1 (peroxisome) and Su9(1-69)-mTurquoise2<sup>215</sup> (mitochondria) were performed by co-transforming a CEN6/ARS4 CRISPR plasmid (containing Cas9, a guide RNA for targeting the YMR206W locus

and a *HIS3* auxotrophic marker) with linearized repair DNA designed for homologous recombination. Cells were plated on histidine-dropout medium, re-streaked on histidine-dropout medium and then grown in non-selective medium to remove the CRISPR plasmid. Chromosomal integration was confirmed by PCR, and removal of CRISPR plasmid was confirmed by replica plating from non-selective medium onto histidine-dropout medium (a colony will not grow on histidine-dropout medium if the CRISPR plasmid has been removed).

#### 4.6.2.2 *Fluorescence microscopy*

For *HsMAO-A* localization experiments, strains were grown to saturation in SC medium with 2% glucose and auxotrophic selection, diluted 50-fold into fresh selective medium and grown for 6–8 hrs. Cultures were then resuspended in 1x PBS before spotting onto plain glass slides for imaging on a Zeiss Axio Observer D1 microscope with an X-Cite Series 120 fluorescent lamp, a Hamamatsu Orca-Flash 4.0 digital camera and ZEN 2.6 (blue edition) software. Images were analyzed using ZEN 3.0 (blue edition) software. The fluorescent protein variants used were CFP mTurquoise2, YFP Venus and RFP mRuby2.

#### 4.6.2.3 *Spot assays for prodeoxyviolacein production and visualization*

Agar plate spots for visualization of prodeoxyviolacein production were generated by plating 5  $\mu$ l of saturated culture on SC agar plates with 2% glucose and appropriate auxotrophic selection. Plates were grown at 30 °C and imaged at 48 hrs. Three biological replicates of each strain were plated and imaged.

### 4.6.3 Supplemental tables

**Table 4.1. Strain list for Chapter 4**

Base strain for BIA synthesis			
Name	Brief description	Genotype	Ref.
LP507	Strain derived from BY4741 capable of making 4.6 g/L reticuline. See reference for full detail.	<i>BY4741</i> . Oxidoreductase knockouts: $\Delta aad3$ , $\Delta ydr541c$ , $\Delta ypr1$ , $\Delta adh6$ , $\Delta ari1$ , $\Delta gre2$ , $\Delta hfd1$ . Mitochondrial fixes: <i>MIP1</i> , <i>CAT5</i> , <i>SAL1</i> . Aromatic amino acid flux: <i>ARO4<sup>FBR</sup></i> , <i>ARO7<sup>FBR</sup></i> , <i>ARO2</i> , <i>TYRI</i> . 4-HPAA synthesis: <i>ARO10</i> . Dopamine synthesis: <i>DODC</i> , <i>CYP76AD5</i> . Norcoclaurine synthesis: <i>CjNCSΔ35</i> (2 copies). Reticuline pathway: <i>Ps6OMT</i> , <i>PsCNMT</i> , <i>Ps4'OMT</i> (2 copies), <i>AtATR2</i> .	16

Yeast strains generated for BIA synthesis			
Name	Parent	Brief Description	Genotype
LP524	LP507	Complete second copy of reticuline pathway	<b>USERXII-1</b> ::P <sub>TDH3</sub> - <i>PsCNMT</i> -T <sub>TDH1</sub> -P <sub>CCW12</sub> - <i>EcNMCH</i> -T <sub>PGII</sub>
LN1002	LP524	Knockouts of native and overexpressed Aro10	$\Delta Aro10$ , $\Delta FgF18$ ::P <sub>TDH3</sub> - <i>Aro10</i> -T <sub>TDH1</sub>
LN1044	LN1002	RNR2-MAO	<b>USERX-1</b> ::P <sub>RNR2</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1046	LN1002	RPL18B-MAO	<b>USERX-1</b> ::P <sub>RPL18B</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1047	LN1002	TEF2-MAO	<b>USERX-1</b> ::P <sub>TEF2</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1049	LN1002	TDH3-MAO	<b>USERX-1</b> ::P <sub>TDH3</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1020	LN1002	TEF2-MAOΔSP	<b>USERX-1</b> ::P <sub>TEF2</sub> - <i>HsMAO-AΔSP</i> -T <sub>TDH1</sub>
LN1015	LN1002	TEF2-MAOΔSP-p	<b>USERX-1</b> ::P <sub>TEF2</sub> - <i>HsMAO-AΔSP</i> -ePTS1-T <sub>TDH1</sub>
LN1063	LN1046	NCS-p	<b>USERX-3</b> ::P <sub>TDH3</sub> - <i>CjNCSΔ35</i> -ePTS1-T <sub>ENO2</sub>
LN1064	LN1002	RPL39-MAO	<b>USERX-1</b> ::P <sub>RPL39</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1065	LN1002	HTP1-MAO	<b>USERX-1</b> ::P <sub>HTB1</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1066	LN1002	RPS23A-MAO	<b>USERX-1</b> ::P <sub>RPS23A</sub> - <i>HsMAO-A</i> -T <sub>TDH1</sub>
LN1067	LN1064	NCS-p	<b>USERXI-2</b> ::P <sub>TDH3</sub> - <i>CjNCSΔ35</i> -ePTS1-T <sub>ENO2</sub>
LN1068	LN1065	NCS-p	<b>USERXI-2</b> ::P <sub>TDH3</sub> - <i>CjNCSΔ35</i> -ePTS1-T <sub>ENO2</sub>
LN1069	LN1066	NCS-p	<b>USERXI-2</b> ::P <sub>TDH3</sub> - <i>CjNCSΔ35</i> -ePTS1-T <sub>ENO2</sub>
LN1075	LN1015	BBE	<b>USERX-3</b> ::P <sub>TDH3</sub> <i>EcBBE</i> -T <sub>TDH1</sub>
LN1078	LN1075	CFS/SPS/TNMT	<b>USERXI-2</b> ::P <sub>PGK1</sub> - <i>SdiCFS</i> -T <sub>ADH1</sub> -P <sub>TEF1</sub> - <i>NdoSPS</i> -T <sub>PGK1</sub> -P <sub>TEF2</sub> -TNMT-T <sub>ADH1</sub>
LN1079	LN1075	BBE (2 copies)	<b>USERXI-2</b> ::P <sub>TDH3</sub> <i>EcBBE</i> -T <sub>TDH1</sub>
LN1080	LN1079	CFS/SPS/TNMT	<b>USERIV-1</b> ::P <sub>PGK1</sub> - <i>SdiCFS</i> -T <sub>ADH1</sub> -P <sub>TEF1</sub> - <i>NdoSPS</i> -T <sub>PGK1</sub> -P <sub>TEF2</sub> -TNMT-T <sub>ADH1</sub>
LN1083	LN1080	MSH/P6H	<b>USERIX-1</b> ::P <sub>PDC1</sub> -P6H-T <sub>CYC1</sub> - <b>linker</b> -P <sub>TDH3</sub> -MSH-T <sub>ADH1</sub>
LN1084	LN1080	MSH/P6H/TNMT	<b>USERIX-1</b> ::P <sub>PDC1</sub> -P6H-T <sub>CYC1</sub> - <b>linker</b> -P <sub>TDH3</sub> -MSH-T <sub>ADH1</sub> - <b>linker</b> -P <sub>FBA1</sub> -TNMT-T <sub>PGII</sub>

Yeast strains generated for assessment of <i>HsMAO-A</i> localization			
Name	Parent	Plasmid used	Description
yWCD230	BY4741	pWCD1351	Wild-type yeast ( <i>his3Δ</i> )
yWCD231	BY4741 pex5Δ::KanMX	pWCD1351	Defective peroxisomal import ( <i>his3Δ</i> )
yKSS001	yWCD230	pKS0864	YMR206WΔ::pHHF1-Su9(1-69)-mTurquoise2-tENO2-pTEF2-mRuby2-GSLGRGRR-SKL!-tPGK1
yKSS002	yKSS001	pKS0902	Microscopy: N-terminal VioE+Venus MAO, ePTS1
yKSS003	yKSS001	pKS0903	Microscopy: N-terminal VioE+Venus MAO, inactive ePTS1
yKSS004	yKSS001	pKS0904	Microscopy: N-terminal VioE+Venus MAOΔSP, ePTS1
yKSS005	yKSS001	pKS0905	Microscopy: N-terminal VioE+Venus MAOΔSP, inactive ePTS1
yKSS006	yKSS001	pKS0906	Microscopy: C-terminal VioE+Venus MAO, ePTS1
yKSS007	yKSS001	pKS0907	Microscopy: C-terminal VioE+Venus MAO, inactive ePTS1
yKSS008	yKSS001	pKS0908	Microscopy: C-terminal VioE+Venus MAOΔSP, ePTS1
yKSS009	yKSS001	pKS0909	Microscopy: C-terminal VioE+Venus MAOΔSP, inactive ePTS1
yPSG335	yWCD230	pJAS1052	pTDH3-VioA-tENO1-pTEF1-VioB-tPGK1
yPSG340	yWCD231	pJAS1052	pTDH3-VioA-tENO1-pTEF1-VioB-tPGK1
yKSS010	yPSG335	pKS0902	Violaicin: N-terminal VioE+Venus MAO, ePTS1
yKSS011	yPSG335	pKS0903	Violaicin: N-terminal VioE+Venus MAO, inactive ePTS1
yKSS012	yPSG335	pKS0904	Violaicin: N-terminal VioE+Venus MAOΔSP, ePTS1
yKSS013	yPSG335	pKS0905	N-terminal VioE+Venus MAOΔSP, inactive ePTS1
yKSS014	yPSG335	pKS0906	Violaicin: C-terminal VioE+Venus MAO, ePTS1
yKSS015	yPSG335	pKS0907	Violaicin: C-terminal VioE+Venus MAO, inactive ePTS1
yKSS016	yPSG335	pKS0908	Violaicin: C-terminal VioE+Venus MAOΔSP, ePTS1
yKSS017	yPSG335	pKS0909	Violaicin: C-terminal VioE+Venus MAOΔSP, inactive ePTS1
yKSS018	yPSG340	pKS0902	Violaicin, import defects: N-terminal VioE+Venus MAO, ePTS1
yKSS019	yPSG340	pKS0903	Violaicin, import defects: N-terminal VioE+Venus MAO, inactive ePTS1
yKSS020	yPSG340	pKS0904	Violaicin, import defects: N-terminal VioE+Venus MAOΔSP, ePTS1
yKSS021	yPSG340	pKS0905	Violaicin, import defects: N-terminal VioE+Venus MAOΔSP, inactive ePTS1
yKSS022	yPSG340	pKS0906	Violaicin, import defects: C-terminal VioE+Venus MAO, ePTS1
yKSS023	yPSG340	pKS0907	Violaicin, import defects: C-terminal VioE+Venus MAO, inactive ePTS1
yKSS024	yPSG340	pKS0908	Violaicin, import defects: C-terminal VioE+Venus MAOΔSP, ePTS1
yKSS025	yPSG340	pKS0909	Violaicin, import defects: C-terminal VioE+Venus MAOΔSP, inactive ePTS1
yKSS026	yPSG335	pKS1124	Violaicin: VioE-Venus fusion, ePTS1
yKSS027	yPSG335	pKS1125	Violaicin: VioE-Venus fusion, inactive ePTS1
yKSS028	yPSG340	pKS1124	Violaicin, import defects: VioE-Venus fusion, active ePTS1
yKSS029	yPSG340	pKS1125	Violaicin, import defects: VioE-Venus fusion, inactive ePTS1

**Table 4.2. Plasmid list for Chapter 4**

Plasmids		
Name	Brief Description	Ref.
pCAS-G418	Yeast replicative plasmid (2 $\mu$ ) with Cas9 and tRNA <sup>Tyr</sup> -driven gRNA. G418 resistance.	183
pCAS-Hyg	Yeast replicative plasmid (2 $\mu$ ) with Cas9 and tRNA <sup>Tyr</sup> -driven gRNA. Hyg resistance.	16
pBSC011	P <sub>TDH3</sub> -C <sub>j</sub> NCSD35-T <sub>ENO2</sub>	16
pGC1899	Harbors CFS, SPS, TNMT for integration into genome	Chp. 4
pGC997	Harbors MSH, P6H, TNMT for integration into genome	23
pHUM	Harbors <i>HIS3</i> , <i>URA3</i> , <i>MET17</i> for prototrophy restoration of S288C-derived strains	216
pWCD1351	<i>HIS3</i>	Chp. 4
pKS0864	Integrative plasmid harbouring pHHF1-Su9ss-mTurquoise2-tENO2-pTEF2-mRuby2-GSLGRGRR-SKL! <sup>-</sup> tPGK1 and a <i>HIS3</i> marker	Chp. 4
pKS0902	pTEF2-VioE-Venus-Hs_MAO-A-ePTS1-tADH1	Chp. 4
pKS0903	pTEF2-VioE-Venus-Hs_MAO-A-dead_PTS1-tADH1	Chp. 4
pKS0904	pTEF2-VioE-Venus-Hs_MAO-A(1-497)-ePTS1-tADH1	Chp. 4
pKS0905	pTEF2-VioE-Venus-Hs_MAO-A(1-497)-dead_PTS1-tADH1	Chp. 4
pKS0906	pTEF2-Hs_MAO-A-VioE-Venus-ePTS1-tADH1	Chp. 4
pKS0907	pTEF2-Hs_MAO-A-VioE-Venus-dead_PTS1-tADH1	Chp. 4
pKS0908	pTEF2-Hs_MAO-A(1-497)-VioE-Venus-ePTS1-tADH1	Chp. 4
pKS0909	pTEF2-Hs_MAO-A(1-497)-VioE-Venus-dead_PTS1-tADH1	Chp. 4
pKS1124	pTEF2-VioE-Venus-ePTS1-tADH1	Chp. 4
pKS1125	pTEF2-VioE-Venus-dead_PTS1-tADH1	Chp. 4

Chp.4 refers to Chapter 4 of this work

**Table 4.3. Primer list for Chapter 4**

Yeast strains			
Number	Name	Sequence (5'->3')	Description
LB33	HDV <sub>gRNA</sub> F	CACCTATATCTGCGTGTTC	When paired: amplify empty guide RNA transcription cassette from pCAS-G418/Hyg. Use with two internal primers providing guide RNA target sequence.
LB34	gRNA_scaffold_R	GTCAAGACTGTCAAGGAGG	
LB1968	X-1_UP_F	cgctcactagtagacaacacacg	Amplify X-1 upstream homology region with homology to gene cassettes
LB1969	(LV3)_X-1_UP_R	gcattttattatataagttgtttattcagagtattcc tggaagcatacagaatattcactaac	
LB1970	(LV5)_X-1_DN_F	Cctcttatattacatcaaaataagaaataattat aacaccacgattgagtgctgcactcttattc	Amplify X-1 downstream homology region with homology to gene cassettes
LB1971	X-1_DN_R	ccttttccaattcttaggctattgg	
LB1964	X-1_guide_F	gtagtacaagaacatattgGTTTTAGA GCTAGAAATAGCAAGT	Use with LB33/LB34 to introduce X-1 guide sequence to CRISPR plasmid
LB1965	X-1_guide_R	ccatattgtctgtgactacAAAGTCCCA TTCGCCACCCGAA	
LB1987	X-3_UP_F	ggctactgattttgtaagcaactc	Amplify X-3 upstream homology region with homology to gene cassettes
LB1979	(LV3)_X-3_UP_R	gcattttattatataagttgtttattcagagtattcc tgagaagaaattttggggtaatatg	
LB1980	(LV5)_X-3_DN_F	Cctcttatattacatcaaaataagaaataattat aacagggaataaggtttaaaggcactg	Amplify X-3 downstream homology region with homology to gene cassettes
LB1981	X-3_DN_R	ggtatctcaatgaacgagctcg	
LB1974	X-3_guide_F	gatcgccgaatggcacgcaGTTTTAGA GCTAGAAATAGCAAGT	

LB1975	X-3_guide_R	tgcggtgccattcgggcgtcAAAGTCCC ATTCGCCACCCGAA	Use with LB33/LB34 to introduce X-3 guide sequence to CRISPR plasmid
LB1988	XI-2_UP_F	ggtttctgaaaaagaagtagtgc	Amplify XI-2 upstream homology region with homology to gene cassettes
LB1989	(LV3)_XI-2_UP_R	gcattttattatataagttgtttattcagagtattcc tctgaaagcgctagtcgtgtgtacc	
LB1990	(LV5)_XI-2_DN_F	Cctctttatattacatcaaaataagaaaataattat aacagctttcagttttcgtggctag	
LB1991	XI-2_DN_R	ctataacatggtttacaacccgagg	Amplify XI-2 downstream homology region with homology to gene cassettes
LB1984	XI-2_guide_F	actttgtcgtttcttactttGTTTTAGAGCT AGAAATAGCAAGT	Use with LB33/LB34 to introduce XI-2 guide sequence to CRISPR plasmid
LB1985	XI-2_guide_R	aaagtaagaacgacaagtAAAGTCCC ATTCGCCACCCGAA	
LN1100	IV-1_UP_F	CGTGCGCTTGAGATTCAGT	Amplify IV-1 upstream homology region with homology to gene cassettes
LN1101	(LV3)_IV-1_UP_R	gcattttattatataagttgtttattcagagtattcc tAGAGTTCCCGTCGGAAT	
LN1102	(LV5)_IV-1_DN_F	cctctttatattacatcaaaataagaaaataattata acaCGTTACTAGCGTTGCAAGT GG	Amplify IV-1 downstream homology region with homology to gene cassettes
LN1103	IV-1_DN_R	GGATTTGGTTTAGCAGCAGTC	Introduce IV-1 guide RNA sequence to CRISPR plasmid using overlap extension
LN1106	IV-1_guide_F	CGCCGGCTGGGCAACACCTTC GGGTGGCGAATGGGACTTTCT GCAAGGAAGTTTAAGCGT	
LN1107	IV-1_guide_R	GACTAGCCTTATTTAACTTG CTATTTCTAGCTCTAAAACAC GCTTAAACTTCCTTGCAGA	
LN1092	II-1_UP_F	GCGTTCACAGTACTCTTTTA GAAC	Amplify II-1 upstream homology region with homology to gene cassettes
LN1093	(LV3)_II-1_UP_R	gcattttattatataagttgtttattcagagtattcc tAAAATAACATGTTGCGTGCA C	
LN1094	(LV5)_II-1_DN_F	cctctttatattacatcaaaataagaaaataattata acaGACAAACTTTACAAAGAA GACACCC	Amplify II-1 downstream homology region with homology to gene cassettes
LN1095	II-1_DN_R	GTATGCCGTGATATGAACAAA C	Introduce II-1 guide RNA sequence to CRISPR plasmid using overlap extension
LN1098	II-1_guide_F	CGCCGGCTGGGCAACACCTTC GGGTGGCGAATGGGACTTTAA CTGCTCAGGGCGGATAAC	
LN1099	II-1_guide_R	GACTAGCCTTATTTAACTTG CTATTTCTAGCTCTAAAACGT TATCCGCCCTGAGCAGTT	
LN1108	IX-1_UP_F	CAACTGCTAAGAACTCTGTGA TCTTC	Amplify IX-1 upstream homology region with homology to gene cassettes
LN1109	(LV3)_IX-1_UP_R	gcattttattatataagttgtttattcagagtattcc tTCGCGAGATAGAACGACATC	
LN1110	(LV5)_IX-1_DN_F	cctctttatattacatcaaaataagaaaataattata acaTTGATGACTAGCGGACT TG	Amplify IX-1 downstream homology region with homology to gene cassettes
LN1111	IX-1_DN_R	GGCAGAAAACACCCGTAGA ATAC	Introduce IX-1 guide RNA sequence to CRISPR plasmid using overlap extension
LN1114	IX-1_guide_F	CGCCGGCTGGGCAACACCTTC GGGTGGCGAATGGGACTTTAT CTTAAATGAAAGACAGAGGTT TT	

LN1115	IX-1_guide_R	GACTAGCCTTATTTAACTTG CTATTTCTAGCTCTAAAACCT CTGTCTTTCATTTAAGATAAA G	
LN14	(LV3)_pRNR2_F	aggaataactctgaataaaacaacttatataataaa aatgcAGTCGAACAAGAAGCAG G	Amplify pRNR2 with homology to gene cassettes
LN456	pRNR2_R	GGTAATTGGACAAATAAATAC GTGTATTAAG	
LN18	(LV3)_pRPL18B_F	aggaataactctgaataaaacaacttatataataaa aatgcaagaggatgtccaatatttt	Amplify pRPL18B with homology to gene cassettes
LN455	pRPL18B_R	tftgtttttgtttcttctaattgatt	
LN1080	(LV3)_pRPL39_F	aggaataactctgaataaaacaacttatataataaa aatgcCTTGGATATGTATGTTGG TCTTGTT	Amplify pRPL39A with homology to gene cassettes
LN1081	pRPL39_R	GTTGATCTATCTGTGTTTATTT GCTTG	
LN1083	(LV3)_pHTB1_F	aggaataactctgaataaaacaacttatataataaa aatgcATGATGGTTCAACAAGAC CAGA	Amplify pHTB1 with homology to gene cassettes
LN1084	pHTB1_R	TGTATGTGTGTATGGTTTATTT GTGG	
LN1086	(LV3)_pRPS23A_F	aggaataactctgaataaaacaacttatataataaa aatgcGTCGGTGCCTAGACTT TTC	Amplify pRPS23A with homology to gene cassettes
LN1087	pRPS23A_R	CTTTGTTTATTTCTGTTGTCT TAGG	
LN27	(LV3)_pTEF2_F	aggaataactctgaataaaacaacttatataataaa aatgctgataggcaagatcaatgtaaac	Amplify pTEF2 with homology to gene cassettes
LN454	pTEF2_R	gtttagtaattatagttcgttgaccg	
PP120	(LV3)_pTDH3_F	AGGAATACTCTGAATAAAAAC AACTTATATAATAAAAATGCtc gagttatcattatcaact	Amplify pTDH3 with homology to gene cassettes
LN398	pTDH3_R	TTTGTTTGTATGTGTGTTTA Ttcg	
LB1130	tTDH1_F	ATAAAGCAATCTTGATGAGGA TAATG	Amplify tTDH1 with homology to gene cassettes
LB1131	(LV5)_tTDH1_R	tggtataattatttctattttgatgtaataaagag GCCATCCTAGAACTTCAATTC ACCAC	
LN247	tPGI1_F	aacaaategctcttaaatatatacc	Amplify tPGI1 with homology to gene cassettes
PP127	(LV5)_tPGI1_R	tggtataattatttctattttgatgtaataaagag gggtatactggaggcttcat	
LN1018	Aro10g1_OE_F	CGCCGGCTGGCAACACCTTC GGGTGGCGAATGGGACTTTAT ATTCGCCTTGTGGACT	Introduce guide RNA targeting Aro10 to CRISPR plasmid using overlap extension
LN1019	Aro10g1_OE_R	GACTAGCCTTATTTAACTTG CTATTTCTAGCTCTAAAACAG TGTCCACAAGGCGAATAT	
LN1020	Aro10_repair_OE_F	ATTGCCGAGGTCATGCTGAGC ATTTGTCGTACTTTTGTGCCGT ATATTAAG	Generate repair template for native Aro10 knockout using overlap extension
LN1021	Aro10_repair_OE_R	AAAGAACTCTGTGGTAGTGGT AAAATAGCTTTAATATACGGC ACAAAAGTACGA	

LN1022	FgF18_repair_OE_F	tttcagtagatttgtaactgtgcaaccataactca tgccaatcgtc	Generate repair template for heterologous Aro10 knockout using overlap extension
LN1023	FgF18_repair_OE_R	tgtgatgaatttgagagcccacttttgggggac gattggcatgagttatg	
LN459	(pRNR2)_HsMAO-A_F	GAATCCAAACTTAATACACGT ATTTATTTGTCCAATTACCAT GGAAAATCAAGAAAAGGCAT C	Amplify <i>HsMAO-A</i> with homology to pRNR2
LN458	(pRPL18B)_HsMAO-A_F	atagaagaaaaaatcaattagaagaaaacaaa aaacaaaATGGAAAATCAAGAAA AGGCATC	Amplify <i>HsMAO-A</i> with homology to pRPL18B
LN1089	(pRPL39)_HsMAO-A-F	AATTCGAAAAAGACAAGCAA ATAAACACAGATAGATCAAC ATGGAAAATCAAGAAAAGGC ATC	Amplify <i>HsMAO-A</i> with homology to pRPL39
LN1090	(pHTB1)_HsMAO-A_F	ATAGACAAGTCAAACCACAA ATAAACCATACACACATACAA TGGAAAATCAAGAAAAGGCA TC	Amplify <i>HsMAO-A</i> with homology to pHTB1
LN1091	(pRPS23A)_HsMAO-A_F	AAATTTTACAAAAACCTAAGA CAACAGGAAATAAACAAAGA TGGAAAATCAAGAAAAGGCA TC	Amplify <i>HsMAO-A</i> with homology to pRPS23A
LN457	(pTEF2)_HsMAO-A_F	ttttagaataacggtcaacgaactataattaacta aacATGGAAAATCAAGAAAAG GCATC	Amplify <i>HsMAO-A</i> with homology to pTEF2
LN396	(pTDH3)_HsMAO-A_F	caagaacttagttcgaATAAACACAC ATAAACAAACAAAATGGAAA ATCAAGAAAAGGCATC	Amplify <i>HsMAO-A</i> with homology to pTDH3
LN397	(tTDH1)_HsMAO-A_FL_R	attcaaaaaaatcattatcctcatcaagattgct ttatTTATGATCTTGGCAACAAT TTGTAC	Amplify full-length <i>HsMAO-A</i> with homology to tTDH1
LN978	(ePTS1)_HsMAO-A_dSP_R	CAACTTAGAACGACGACCAC GACCTAATGATGGCAAGTTTC TTTCCCAAAAaGTATGAGTGA TTTCAAC	Amplify signal peptide-less <i>HsMAO-A</i> with C-terminal ePTS1 tag
LN979	(ePTS1)_TDH1t_F	TTAGGTCGTGGTCGTCGTTCT AAGTTGTAAataaagcaatcttgatgagg ataatg	Amplify tTDH1 with homology to C-terminal ePTS1 tag
LN1071	(pTEF1)_CjNCS_F	cttctgctcattagaaagaaagcatagcaatcta atctaagtttaataAACAAAtggaagaaact gtaatgttatc	Amplify <i>CjNCS</i> with homology to pTEF1
LN1072	(ePTS1)_CjNCS_R	ttaCAACTTAGAACGACGACCA CGACCTAAActcagaagattgtgcttatttt c	Amplify <i>CjNCS</i> with C- terminal ePTS1 tag
LN1070	(ePTS1)_tPGI1_F	TTAGGTCGTGGTCGTCGTTCT AAGTTGtaaacaatcgctcttaaatatata cctaaag	Amplify tPGI1 with N- terminal ePTS1 tag
LN564	(pTDH3)_EcBBE_F	caagaacttagttcgaATAAACACAC ATAAACAAACAAAATGGAAA ATAAGACACCGATTTTC	Amplify BBE with homology to pTDH3/tPGI1
LN565	(tPGI1)_EcBBE_R	ctttaatgtcttaggtatatttaagagcgatttg ttTTATATAACGACTTCTCCCCC G	

PP118	(LV3)_pPDC1_F	AGGAATACTCTGAATAAAAC AACTTATATAATAAAAATGCa catgcgactgggt	Amplify genes from pGC997 with homology to integration sites
PP124	(LV5)_tADH1_R	TGTTATAATTATTTTCTTATTT TGATGTAATATAAAGAGGgcat gccggtagag	Amplify MSH, P6H from pGC997 with homology to integration sites
PP127	(LV5)_tPGI1_R	TGTTATAATTATTTTCTTATTT TGATGTAATATAAAGAGGggtat actggaggctcat	Amplify MSH, P6H, and TNMT from pGC997 with homology to integration sites
LN1124	SPS-FP	gcatcgtctcatcgggtctcatatggaagaatcttc tggattgtctcc	Amplify stylophine synthase for Golden Gate cloning
LN1125	SPS-RP	atgccgtctcagggtctcaggatctaataaccgta atagatggagaacct	
LN1126	CFS-FP	gcatcgtctcatcgggtctcatatggaagaatcttc tggttggttacc	Amplify cheilanthifoline synthase for Golden Gate cloning
LN1127	CFS-RP	atgccgtctcagggtctcaggatctaagagttcttg gggtgatcttag	
LN1128	TNMT-FP	gcatcgtctcatcgggtctcatatgggttcaatagat gaggtcaagaag	Amplify TNMT for Golden Gate cloning
LN1129	TNMT-RP	atgccgtctcagggtctcaggatctacttcttga aaagcagctgc	

**Table 4.4. Synthetic gene list for Chapter 4**

*HsMAO-A*

atggaaaatcaagaaaaggcatctattgcagggtcacatgtttgatgttggtttataggtggaggaattcagggttgcagctgctaaattgttaa  
ctgagtacgggtgttctgttttagtattagaggctagagacagagttgggtgtagaacctatacaatcagaaacgaacatgtagactacgttga  
cgttgggtggagcttatgttggcctactcagaataggatttaagggtgctaaaggaattaggtattgaaactataaagttaatgttccagaaaga  
ttagtacagtatgttaaaggttaagacatatccatttagaggtgcctttccaccagtttggaaatcctatcgcttactggattacaataatttgggag  
aaccattgataatattgggtaaggagatccaactgacgcaccttgggaggcacagcatgcagataagtgaggacaagatgacatgaagga  
attaatcgataaaatattgttgactaaaaccgcaaggagattcgctatttattgtaaatattaacgtaacttctgaacctcatgaagttcagca  
ttgtggttcttatggatgtcaagcagtggtggtgacaacaagaatatttccagttacaacgggtggtcaagaaagaaaattcgtcggtggttct  
ggacaggtatctgaaagaattatggatttattaggtgatcaagtcaattgaatcatccagttaccatgtagaccaatcttctgataacatc  
atagaaactttaaactcatgaacattatgagtgtaaatattgcataaacgcaattctccaacctgactgcaaaaatccattttagacctgaattac  
ctgctgaaagaaaccaattaatccaagattgccaatgggtgctgtaatacaatgtatgatgtattacaagaagcttctggaagaagaag  
attactgtggttgcattatagaggatgaagatgctccaattctatcaccttagacgatacactaaacctgatggatcttacctgcataatgg  
gattcatattggccagaaaggccgatagattggcaaaagttacataaagaattagaagaagaagaatctgtgagttatgctaaagtcttagg  
ttctcaagaagccttgcctcctgctcattatgaagaaaagaactggtgtgaggaacaatattctggaggatgttactgcttatttccacctg  
gaattatgacacaatacggagagttatcaggcagccagtcggtgagaatttcttgcaggaaactgaaaccgctacaaaatggtctggttatat  
ggaaggtgcagtagaggctggtgaaagggctgctagagaagttttaaattgattaggaagaagcactgaaaaggatattgggtacaggag  
cctgagtcctaaagatgttctgctgttgaatcactcatacctttgggaaagaacttgcctcagtcctcaggttgtgaaaatcattggttctc  
tacctcagtcaccgccttgggattcgtttgtataagtagtaaaattgttccaagatca

*EcBBE (from Eschscolzia californica)*

atggaaaataagacaccgattttcttttattcaatcttctatcattgctaaattgcgcgcttggcgggaatgatctattgtcatgcctaacgtt  
caatggtgtacgtaatacactgtgttttagcggcagcagtgacttcaataggttcttcccttagtatacaaaaaccttgttccaaaa  
ctcactaatatccaagcctagtgctattatccttccggggagcaagaggatccaacactataagatgtataagaaaagggttctggacc  
ataaggctgaggtcaggaggtcattcctacgaaggtctgtcatacaccagtgacacaccttcttctatagacttaataatctaaatagggt  
aagtattgattagaatcagaacagcgtgggtgagagcggatcaactctaggagaattatattacgcgataacggaaagtcttcaaagctt  
ggatttaccgcagggtggtgccctaccgtgggtacggggggtcacatctccgggggggcttggcatgatgagtagaaagtaggggcta  
gctgctgacaatgtagtagatgcatattatcgacgcgaatggggctatcttagacaggcaagctatgggagaggacgtgttttgggccatt  
agaggcgggtggcgggtggggtatggggagccatatacgcctggaaaataaagcttctacctgtccccgaaaagtaccgtattccgtgtaa  
ctaaaaactggcgatagacgagcgcacatcactgctacacaagtggcagttctgctgctgaagaattggaagaggacttactctgtccgtt  
ctggcggggagcagatgagaagcaggtatggtgacgatgttaggcttcttcttggactgaaaacagtagcaaaagtcaacctttgattacttt  
tcctgaactaggactttagagaggattaccttgaatgtcctgggggggaatccttccgctatctagctggattgaaaccgtgagccaac  
ttaataatagattcctgaagttgatgaaagggcattcaaaaccaagttgatttaacgaaggagccgctgcttagcaaaagcgtttatggctt  
ctggagagattatctaaagagccaaatggttctatagctttgaacgggttccgggggacaaatgcaaaagatctctcagacttaccctgtcc  
ctcatcgtccggtacgaggtgatggtagaatacatcgtggcatggaatcaatcagaacaaaagaagaaaaccgagtttcttactggttgg  
aaaaagtttacgaatttatgaaaccgtttgatctaagaatccccgtttgggatacgtcaaccatatagatttagatcttgggggaaatagactgg  
gggaataagacagtagttaacaatgcgatagagatcagtcgttcatgggtgagtcataatttttccaaattacgagagattaatccgtgctaa  
aacctaatagatccaataatgtttaaaccaccgcagtcctcccccgatggccaattttgattaccttgagaagacacttgggtccgac  
gggggagaagtcgttatataa

## 5 Chapter Five: Conclusion and Perspectives

Benzylisoquinoline alkaloids (BIAs) represent a large class of plant secondary natural products. Some BIAs are essential medicines, while other BIAs have pharmaceutical potential but are too scarce to investigate further. The current source of most BIAs is plant extraction. Microbial BIA synthesis presents an opportunity to access the rest of the BIA family for research and commercial purposes. BIA synthesis in microbial hosts has been an ongoing area of research for almost 15 years. The first pathway reconstructions from supplemented precursors were reported in 2008<sup>11,12</sup>, with increasingly complex pathways introduced as the key enzymes were identified, including dihydrosanguinarine (2014)<sup>23</sup>, morphine (2015)<sup>10</sup>, and noscapine (2018)<sup>7</sup>. The first *de novo* synthesis of reticuline was demonstrated in *E. coli* in 2011<sup>191</sup> and in *S. cerevisiae* in 2015<sup>14</sup>. The most recent milestone is the gram-per-liter-scale synthesis of the BIA reticuline in yeast in 2020<sup>16</sup>. Until now, this success has not been accompanied by the introduction of a BIA pathway of commercial interest.

This thesis is a case study of dihydrosanguinarine synthesis in yeast. I was a co-first author of the first pathway reconstruction of dihydrosanguinarine synthesis from supplemented norlaudanosoline in 2014, which was one of the longest heterologous pathways introduced to yeast to date. The pathway was functional but inefficient. The first goal of the thesis was to improve upon the pathway, with the intention to introduce *de novo* BIA synthesis once it was developed. Since norlaudanosoline was being used in supplementation assays, I explored the introduction of norlaudanosoline synthesis into a BIA-production strain. Not only was norlaudanosoline synthesis achievable in yeast, but it came with clear advantages over the natural route.

This thesis demonstrates the first *de novo* dihydrosanguinarine synthesis in a microbe. *In vitro* supplementation of norlaudanosoline to strains expressing a complete dihydrosanguinarine

pathway identified a bottleneck at the first 2 cytochromes P450, resulting in a reduction in yields. First, this bottleneck was resolved by identifying superior enzyme homologs in a series of combinatorial co-expression assays (Chapter 3). Next, norlaudanosoline synthesis was introduced to yeast synthesizing BIAs *de novo* from sucrose (Chapter 4). This improved the selectivity of BIA scaffold synthesis while also reducing the complexity of the fed-batch fermentation protocol. Finally, dihydrosanguinarine synthesis was introduced to the norlaudanosoline production strain. The end result is a strain that can produce 630 mg/L of combined dihydrosanguinarine and sanguinarine using a simple exponential fed-batch fermentation protocol. Dihydrosanguinarine is the reduced form of the antimicrobial sanguinarine, which is the active ingredient in the animal feed supplement Sangrovit®; the 2L fermentation reported in this thesis produced enough dihydrosanguinarine to supply more than one ton of animal feed<sup>217</sup>.

The paradox of this work, and other long BIA pathway reconstructions in microbes, is that the primary targets are compounds that are commercially available (sanguinarine, morphine, noscapine). Although sanguinarine can be purchased as a plant extract for as low as \$3/kg on Alibaba, its analytical-grade equivalents range from \$14,000,000-\$40,000,000/kg (sites accessed: R&D Systems, Bio-Techne, Sigma) - prices that are typical for rare BIAs. Thus, the value of the research presented in this thesis is not in dihydrosanguinarine synthesis *per se*. Instead, the value is two-fold. First, we demonstrate that microbial BIA synthesis is sufficiently advanced for the production of rare BIAs for the purposes of research and potential commercialization. Second, lessons learned throughout this process can be broadly applied to future targets, either BIAs or other plant natural products.

In addition to enabling higher yields and higher selectivity of scaffold synthesis, the norlaudanosoline route to BIA synthesis might also result in a more versatile platform strain. The

condensation of dopamine and 3,4-dHPAA immediately adds a 3'-hydroxyl group to the resulting scaffold compared to norcoclaurine synthesis. The natural route to 3' hydroxylation requires an *N*-methylation event; norlaudanosoline synthesis bypasses this requirement. This opens an opportunity for the biosynthesis of any BIAs that are 3'-hydroxylated but not *N*-methylated, either natural or non-natural. One such target is the BIA tetrahydropapaverine (6,7,3'4'-tetramethylnorlaudanosoline), which is the precursor to the papaverine as well as the anesthetic atracurium. Enzymes catalyzing all four activities have now been identified<sup>218</sup> or engineered<sup>219</sup>, opening the door to high-level microbial synthesis of this BIA.

## 5.1 The necessity of multi-copy gene integration

Enzymes that performed well in *in vitro* norlaudanosoline supplementation assays in Chapter 3 demonstrated sub-optimal performance in *de novo* biosynthesis described in Chapter 4. *In vitro* supplementation assays in 96-well plates achieved full conversion of norlaudanosoline to *N*-methylstylopine. In contrast, bottlenecks at BBE, CFS, and TNMT were observed when norlaudanosoline was synthesized *de novo* in 96-well plate format. Further, bottlenecks at all cytochromes P450 were observed during fed-batch fermentation. *In vitro* supplementation assays were performed with up to 100  $\mu$ M norlaudanosoline (28 mg/L), which is ~20-fold lower than the amount of norlaudanosoline produced *de novo* in 96-well plate format. It is possible that slow turnover rates begin to hinder pathway performance as flux and titers increase. BBE from *Eschscholzia californica* has a turnover rate of 8/sec<sup>220</sup> and the TNMT from *Glaucium flavum* has a turnover of just 0.8/sec<sup>19</sup>. Fortunately, pathway flux through both BBE and TNMT improves with gene copy number. In general, the entire dihydrosanguinarine pathway downstream of reticuline should be integrated in multiple copies for optimal flux. This mirrors the upstream portion of the pathway; *NCS* and *4'OMT* were already incorporated in multiple copies in strain

LP507<sup>16</sup>, and additional copies of *NCS* and the other methyltransferases were introduced in this work.

It is very likely that the requirement for multiple copies of most pathway enzymes is a general trend for the reconstitution of plant secondary metabolite pathways in microbial hosts. A standout example of plant enzyme inefficiency is ribulose biphosphate carboxylase/oxygenase (RuBisCO), which has a  $k_{cat}$  of just 3/sec even though it is responsible for carbon fixation in plants<sup>221</sup>. Consequently, RuBisCO can be found in multiple copies across photosynthetic algae and land plants, an extreme example being 117 copies in the dinoflagellate *Prorocentrum donghaiense*<sup>222</sup>. A landmark paper detailing  $k_{cat}$  values across the BRENDA enzyme database reveals that the average  $k_{cat}$  across all characterized enzymes is just ~10/sec, with the average enzyme involved with secondary metabolism being 2.5/sec<sup>223</sup>. In this context, BBE and TNMT are perfectly average enzymes. The pertinent question for synthetic biologists may not be whether or not a second copy of a gene should be introduced, but instead, how many copies should be present?

Several CRISPR-based approaches have been developed for targeted multi-copy gene integration in yeast. One such approach is that of synthetic “landing pads” with heterologous guide RNA sites and flanking homology regions existing in predetermined copy numbers in the genome. This landing pad approach has been used to probe the importance of multi-copy *NCS* integration for norcochlorine synthesis in yeast<sup>224</sup>. Another approach relies on targeting genes to natural multi-copy “delta” elements in the yeast genome; the authors successfully integrated up to 18 copies of a 24-kb DNA cassette for (*R,R*)-2,3-butanediol synthesis from xylose in a single round of transformation<sup>225</sup>. Both systems allow researchers to rapidly ascertain the optimal copy number to balance pathway flux without unnecessarily burdening the strain. These approaches should be

routinely used to assess whether bottlenecks in heterologous secondary metabolite pathway reconstruction may be overcome with copy number increase.

It is unfortunate that so much protein must be expressed in order to support gram-scale BIA synthesis, stemming both from pathway length and from the slow turnover rates detailed above. It has been demonstrated that heterologous protein expression represents a quantifiable burden on *E. coli*, with a reduction of growth rate in a promoter strength-dependent manner<sup>226</sup>. Many of the genes used in this work are integrated in multiple places in the genome, with expression driven by the *TDH3* promoter, which is one of the highest-strength promoters identified in yeast. The strains assessed in Chapter 4 have a maximum growth rate of approximately half that of their wild-type parent (personal correspondence, Martin lab). This is undoubtedly due to multiple phenomena, including the reduced ability to catabolize aldehydes, the synthesis of proteins with inherent toxicity such as NCS<sup>194</sup>, and the potential toxicity of BIAs themselves, but the additional burden of protein production from strong promoters cannot be helping.

To some extent, the high value of many BIAs reduces the necessity to ameliorate the slow growth rate; productivity is less of a concern when the target compound costs \$100/mg vs. \$1/kg. However, the eventual growth arrest in both norcoclaurine-producing and norlaudanoline-producing BIA strains must be addressed before commercialization can occur. Many of the issues with BIA production in yeast could be addressed in future works using other synthetic biology solutions, such as compartmentalization of the entire pathway to an organelle or splitting the pathway between organisms. These issues all come with inherent challenges. Natural organelles' protein import mechanisms may be overwhelmed by heterologous protein<sup>194</sup>. BIAs and pathway intermediates such as dopamine are not well-imported by yeast in acidic media<sup>1</sup>, although a series of transporters have been identified that improve uptake of supplemented BIAs<sup>227</sup>. However, there

is potential in both of these options to reduce the double burdens of aldehyde catabolizing enzyme knockouts, necessary to support 4-HPAA/3,4-dHPAA production, and gene overexpression, necessary to support BIA condensation and diversification.

## 5.2 The impacts of enzyme promiscuity

A common theme throughout this work is side-product formation. In Chapter 3, the promiscuity of TNMT contributed to loss of carbon in the dihydrosanguinarine pathway. In Chapter 4, the concerted promiscuity of the yeast enzyme Aro10 as well as the plant enzyme NCS resulted in the synthesis of non-BIA scaffolds. Paradoxically, both resolutions to side-product formation also harnessed promiscuity.

Enzyme promiscuity is a well-documented phenomenon in plant secondary metabolism<sup>113</sup>. It is hypothesized that gene duplication in plants is one of the key factors driving enzyme promiscuity, as one copy may diverge in substrate preference while the other continues to provide the original activity<sup>228</sup>. Occasionally, new substrate preferences may provide a selective advantage, which maintains the activity and allows for additional divergence. Frequently, numerous other substrate preferences are maintained during this process<sup>229</sup>. This “generalist” approach to product synthesis is also typically associated with the poor catalytic efficiency described above<sup>230</sup>.

Plants have some control over promiscuous enzyme activity through the timing and localization of enzyme expression, both within cells and between tissues. These control mechanisms are lost when multiple promiscuous enzymes are co-expressed in a microbial host. Metabolic engineering strategies exist to re-introduce some of these control mechanisms, such as enzyme compartmentalization or scaffolding of sequential enzymes in a pathway to limit substrate diffusion (detailed in Section 2.13). In this work, we re-routed natural pathways to work with enzyme promiscuity rather than against it.

Norlaudanosoline synthesis bypasses the promiscuous 2-oxoacid decarboxylase Aro10 in favor of monoamine oxidase A, which eliminates NCS-catalyzed formation of the side-product 4'-dehydronorcoclaurine. Although in *E. coli*, spontaneous norlaudanosoline condensation occurs even in the absence of NCS, in yeast it was necessary for NCS to accept 3,4-dHPAA. Next, norlaudanosoline which is not found naturally in plants, then had to be converted into reticuline. Thus, this pathway redesign required not only the promiscuity of NCS but also that of the downstream methyltransferases.

Pathway optimization in Chapter 2 involved the intentional selection of promiscuous enzymes (Ring A- and Ring B-closing CYP719s) over those with more limited substrate specificity. The focus was for carbon to be moved through the pathway as rapidly as possible, as opposed to as specifically as possible. The complimentary promiscuities ensured that all carbon was re-captured for the purposes of dihydrosanguinarine synthesis. With the typical enzyme involved in plant secondary metabolite synthesis being both slow (Section 5.1) and promiscuous, this strategy for pathway optimization should be broadly applicable across such heterologous synthesis projects.

A native pathway is the obvious first step for proof-of-concept reconstitution, but optimization should be flexible and take non-obvious routes into consideration. In this work, it was advantageous to have a large bank of enzyme homologs to identify more promiscuous Ring A- and Ring B-closing CYP719s. A yeast-based biosensor for BIAs has recently been developed<sup>231</sup>, which could enable the use of *in vivo* mutagenesis systems like OrthoRep<sup>232</sup> or EvolvR<sup>233</sup> to achieve the same effect with a directed evolution approach.

The interplay of host enzymes and heterologous pathways has been a recurring theme throughout the introduction of BIA synthesis into yeast. The simple introduction of dopamine and

3,4-dHPAA synthesis into *E. coli* was sufficient for ~ 25 mg/L norlaudanosoline production with very little strain engineering<sup>13</sup>. The same approach resulted in just ~0.1 mg/L norcoclaurine production in yeast<sup>14</sup>. It wasn't until the discovery that yeast oxidoreductases were catabolizing 4-HPAA, and their subsequent combinatorial knockout, that high-level BIA synthesis was possible in yeast<sup>16</sup>. While tyrosol production in norcoclaurine strains is now minimal, hydroxytyrosol formation in norlaudanosoline strains is still a problem, suggesting that 3,4-dHPAA is catabolized by yet another oxidoreductase.

Yeast enzymes, especially those involved with detoxification or nutrient scavenging, often interfere with heterologous pathway reconstruction<sup>142,145,169</sup>. While we have chosen to knock out the offending enzymes, this strategy does risk the gradual accumulation of toxins that those enzymes naturally catabolize. Oxidoreductase knockouts did not affect maximum growth rate of BIA strains in microtiter plates (personal correspondence, Martin lab). However, maximum growth rate is assessed very early into such assays (typically 4-5 hours). It is possible that the deleteriousness of oxidoreductase knockouts does not become apparent until well into fed-batch fermentation, leading to the plateaus in growth demonstrated in strains in Pyne *et al.* 2020 and in this work. For that reason, a compartmentalization approach for BIA synthesis could be beneficial, because it would permit the restoration of oxidoreductases.

### 5.3 Perspectives

This thesis identifies several avenues for the immediate improvement of both reticuline and dihydrosanguinarine synthesis in yeast.

#### *Reticuline synthesis*

In addition to 4.8 g/L reticuline, we also identified 2.8 g/L of un-reacted upstream intermediates (dopamine, hydroxytyrosol). We also identified that *HsMAO-A* activity decreased with strain growth rate.

- An oxidoreductase knockout panel should be screened to identify the enzyme(s) responsible for 3,4-dHPAA reduction, which we predict might use NADPH as a cofactor.
- An additional copy of *NCS* should be integrated to promote a more efficient condensation of dopamine and 3,4-dHPAA.
- RNASeq should be performed at several points throughout fed-batch fermentation. The transcriptome profile may provide insight into the stress that BIA-producing strains experience throughout fermentation. Additionally, this will allow for the identification of promoters that are more consistent across changing growth conditions. These should be used to express *HsMAO-A* so that dopamine production doesn't dominate as fed-batch fermentation progresses.
- Once downstream pathways are optimized, additional flux may be pushed through the pentose phosphate and aromatic amino acid pathways.

#### *Dihydrosanguinarine synthesis*

Dihydrosanguinarine and sanguinarine comprised 28% of quantifiable BIAs during fed-batch fermentation of strain LN1084.

- There is a residual 1.5 g/L reticuline. A combination of more copies of *BBE* and/or MBP-*BBE* fusion protein should be tested. The fusion protein strategy may not work as well in yeast; *BBE* harbors an N-terminal vacuolar targeting tag<sup>135</sup> that may be poorly tolerated in an organism without endomembranes such as *E. coli*.

- Protoberberines did not accumulate in the fed-batch fermentation, but *N*-methylated protoberberines did, as well as protopine and potentially hunnemanine and izmirine, meaning that all of the cytochromes P450 in the dihydrosanguinarine pathway displayed some degree of insufficiency. This is likely to be further exacerbated as flux through scoulerine increases. Further pathway balancing will be required to optimize dihydrosanguinarine production. In particular, a CYP82 enzyme bank should be screened to identify enzymes with more activity on *N*-methylstylopine and protopine *in vivo*, as was recommended in Chapter 3.
- The intracellular entrapment of dihydrosanguinarine and sanguinarine presents a problem not only for efficient pathway flux but also for future downstream processing. Originally, dihydrosanguinarine was thought to be a better end product than sanguinarine due to its reduced toxicity. Further, dihydrosanguinarine oxidation results in another molar equivalent of H<sub>2</sub>O<sub>2</sub> production. However, dihydrosanguinarine is undetectable in fermentation supernatants, which is perhaps unsurprising given its intracellular localization in plants<sup>17</sup>. In plants, sanguinarine oxidation is catalyzed by dihydrobenzylphenanthridine oxidase (DBOX) in the ER lumen<sup>234</sup>, resulting in sanguinarine-rich vesicles that fuse with the plasma membrane<sup>17</sup>. Consequently, a benzo[*c*]phenanthridine-specific transporter has not been identified. However, introduction of DBOX may prove to be a net positive if more BIAs can be removed from the cell. The process should be observable by fluorescence microscopy, as dihydrosanguinarine and sanguinarine can be differentiated by their excitation and emission wavelengths<sup>17</sup>.

## 6 References

1. Narcross, L., Fossati, E., Bourgeois, L., Dueber, J. E. & Martin, V. J. J. Microbial Factories for the Production of Benzyloisoquinoline Alkaloids. *Trends Biotechnol.* **34**, 228–241 (2016).
2. Raghavendra, T. Neuromuscular blocking drugs: discovery and development. *J. R. Soc. Med.* **95**, 363–367 (2002).
3. Rida, P. C. G., LiVecche, D., Ogden, A., Zhou, J. & Aneja, R. The Noscapine Chronicle: A Pharmaco-Historic Biography of the Opiate Alkaloid Family and its Clinical Applications. *Med. Res. Rev.* **35**, 1072–1096 (2015).
4. Benej, M. *et al.* Papaverine and its derivatives radiosensitize solid tumors by inhibiting mitochondrial metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 10756–10761 (2018).
5. Stiborova, M. *et al.* Macleaya cordata extract and Sangrovit genotoxicity. Assessment in vivo. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* **152**, 35–39 (2008).
6. Gao, H. Advances of the total synthesis of morphine in the last decade. *IOP Conf. Ser. Earth Environ. Sci.* **657**, (2021).
7. Li, Y. *et al.* Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3922–E3931 (2018).
8. Kaboudin, B. & Sohrabi, M. *Chemistry and synthesis of major opium alkaloids: a comprehensive review. Journal of the Iranian Chemical Society* (Springer Berlin Heidelberg, 2021). doi:10.1007/s13738-021-02268-y.
9. Voigt, C. A. Synthetic biology 2020–2030: six commercially-available products that are changing our world. *Nat. Commun.* **11**, 10–15 (2020).
10. Galanie, S., Thodey, K., Trenchard, I. J., Interrante, M. F. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science (80-. )*. **349**, 1095–1100 (2015).
11. Minami, H. *et al.* Microbial production of plant benzyloisoquinoline alkaloids. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 7393–8 (2008).
12. Hawkins, K. M. & Smolke, C. D. Production of benzyloisoquinoline alkaloids in *Saccharomyces cerevisiae*. *Nat. Chem. Biol.* **4**, 564–573 (2008).
13. Nakagawa, A. *et al.* A bacterial platform for fermentative production of plant alkaloids. *Nat. Commun.* **2**, 1–8 (2011).
14. DeLoache, W. C. *et al.* An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat. Chem. Biol.* **11**, 465–471 (2015).
15. Nakagawa, A. *et al.* Total biosynthesis of opiates by stepwise fermentation using engineered *Escherichia coli*. *Nat. Commun.* **7**, 1–2 (2016).
16. Pyne, M. E. *et al.* A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nat. Commun.* **11**, 1–10 (2020).
17. Weiss, D., Baumert, A., Vogel, M. & Roos, W. Sanguinarine reductase, a key enzyme of benzophenanthridine detoxification. *Plant, Cell Environ.* **29**, 291–302 (2006).
18. Beaudoin, G. A. W. & Facchini, P. J. Benzyloisoquinoline alkaloid biosynthesis in opium poppy. *Planta* **240**, 19–32 (2014).
19. Lang, D. E. *et al.* Structure-function studies of tetrahydroprotoberberine N-methyltransferase reveal the molecular basis of stereoselective substrate recognition. *J. Biol. Chem.* **294**, 14482–14898 (2019).
20. Hagel, J. M. *et al.* Characterization of a flavoprotein oxidase from opium poppy catalyzing the final steps in sanguinarine and papaverine biosynthesis. *J. Biol. Chem.* **287**,

- 42972–42983 (2012).
21. Kutchan, T. M. & Dirtrich, H. Characterization and mechanism of the berberine bridge enzyme, a covalently flavinylated oxidase of benzophenanthridine alkaloid biosynthesis in plants. *J. Biol. Chem.* **270**, 24475–24481 (1995).
  22. Dastmalchi, M., Park, M. R., Morris, J. S. & Facchini, P. Family portraits: the enzymes behind benzyloquinoline alkaloid diversity. *Phytochem. Rev.* **17**, 249–277 (2018).
  23. Fossati, E. *et al.* Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in *Saccharomyces cerevisiae*. *Nat. Commun.* **5**, 3283 (2014).
  24. Li, J. W.-H. & Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **325**, 161–5 (2009).
  25. Harvey, A. L. Natural products as a screening resource. *Curr. Opin. Chem. Biol.* **11**, 480–484 (2007).
  26. David, B., Wolfender, J.-L. & Dias, D. a. The pharmaceutical industry and natural products: historical status and new trends. *Phytochem. Rev.* (2014) doi:10.1007/s11101-014-9367-z.
  27. Harvey, A. L. Natural products in drug discovery. *Drug Discov. Today* **13**, 894–901 (2008).
  28. Lahlou, M. The Success of Natural Products in Drug Discovery. *Pharmacol. & Pharm.* **04**, 17–31 (2013).
  29. Cordell, G. A. Sustainable medicines and global health care. *Planta Med.* **77**, 1129–1138 (2011).
  30. Rathbone, D. A. & Bruce, N. C. Microbial transformation of alkaloids. *Curr. Opin. Microbiol.* **5**, 274–281 (2002).
  31. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* **12**, 355–367 (2014).
  32. Pickens, L. B., Tang, Y. & Chooi, Y.-H. H. Metabolic engineering for the production of natural products. *Annu. Rev. Chem. Biomol. Eng.* **2**, 211–236 (2011).
  33. Hagel, J. M. & Facchini, P. J. Benzyloquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant Cell Physiol.* **54**, 647–72 (2013).
  34. Cordell, G. A., Quinn-Beattie, M. Lou & Farnsworth, N. R. The potential of alkaloids in drug discovery. *Phyther. Res.* **15**, 183–205 (2001).
  35. Stadler, R., Kutchan, T. M. & Zenk, M. H. (S)-norcoclaurine is the central intermediate in benzyloquinoline alkaloid biosynthesis. *Phytochemistry* **28**, 1083–1086 (1989).
  36. Winzer, T. *et al.* A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* **336**, 1704–8 (2012).
  37. Farrow, S. C. & Facchini, P. J. Dioxygenases catalyze O-demethylation and O,O-demethylation with widespread roles in benzyloquinoline alkaloid metabolism in opium poppy. *J. Bol. Chem.* **288**, 28997–29012 (2013).
  38. Takemura, T., Ikezawa, N., Iwasa, K. & Sato, F. Molecular cloning and characterization of a cytochrome P450 in sanguinarine biosynthesis from *Eschscholzia californica* cells. *Phytochemistry* **91**, 100–8 (2013).
  39. Ruff, B. M., Bräse, S. & O'Connor, S. E. Biocatalytic production of tetrahydroisoquinolines. *Tetrahedron Lett.* **53**, 1071–1074 (2012).
  40. Liscombe, D. K. & Facchini, P. J. Molecular cloning and characterization of tetrahydroprotoberberine cis-N-methyltransferase, an enzyme involved in alkaloid

- biosynthesis in opium poppy. *J. Biol. Chem.* **282**, 14741–14751 (2007).
41. Farrow, S. C., Hagel, J. M., Beaudoin, G. A. W., Burns, D. C. & Facchini, P. J. Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. *Nat. Chem. Biol.* **11**, 728–732 (2015).
  42. Singla, D., Sharma, A., Kaur, J., Panwar, B. & Raghava, G. P. S. BIADB: a curated database of benzyloisoquinoline alkaloids. *BMC Pharmacol.* **10**, 1–8 (2010).
  43. Peking Union Medical College Hospital. Pharmacokinetics and Pharmacodynamics of Higenamine in Chinese Healthy Subjects (Clinicaltrials.gov Identifier NCT01451229). (2011).
  44. Bloomer, R. J., Schriefer, J. M. & Gunnels, T. A. Clinical safety assessment of oral higenamine supplementation in healthy, young men. *Hum. Exp. Toxicol.* **34**, 935–945 (2015).
  45. Law, V. *et al.* DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **42**, 1091–1097 (2014).
  46. Wu, M.-C., Law, B., Wilkinson, B. & Micklefield, J. Bioengineering natural product biosynthetic pathways for therapeutic applications. *Curr. Opin. Biotechnol.* **23**, 931–40 (2012).
  47. Amirkia, V. & Heinrich, M. Alkaloids as drug leads - A predictive structural and biodiversity-based analysis. *Phytochem. Lett.* **10**, xlviii–53 (2014).
  48. Bowman, W. C. Neuromuscular block. *Br. J. Pharmacol.* **147**, 277–286 (2006).
  49. Food and Drug Administration. Department of Health and Human Oral Health Care Drug Products for Over-. *Fed. Regist.* **68**, 1–57 (2003).
  50. Frampton, J. E. & McTavish, D. Mivacurium: A Review of its Pharmacology and Therapeutic Potential in General Anaesthesia. *Drugs* **45**, 1066–1089 (1993).
  51. Cougar Biotechnology. A Study of Noscaphine HCl (CB3304 ) in Patients With Relapsed or Refractory Multiple Myeloma (ClinicalTrials.gov Identifier: NCT00912899). <https://clinicaltrials.gov> (2011).
  52. Kang, H., Jang, S.-W., Pak, J. H. & Shim, S. Glucocorticoids inhibit breast cancer cell migration and invasion by inhibiting MMP-9 gene expression through the suppression of NF- $\kappa$ B activation. *Mol. Cell. Biochem.* **403**, 85–94 (2015).
  53. Medicine, S. J. T. U. S. of. Efficacy and Safety of Berberine in the Treatment of Diabetes with Dyslipidemia (Clinicaltrials.gov Identifier NCT00462046). <https://clinicaltrials.gov/ct2/show/NCT00462046> (2007).
  54. Zha, W. *et al.* Berberine inhibits HIV protease inhibitor-induced inflammatory response by modulating ER stress signaling pathways in murine macrophages. *PLoS One* **5**, 2–9 (2010).
  55. Peng, L. *et al.* Antibacterial activity and mechanism of berberine against *Streptococcus agalactiae*. *Int. J. Clin. Exp. Pathol.* **8**, 5217–5223 (2015).
  56. Meade, J. a. *et al.* (-)-Stepholidine is a potent pan-dopamine receptor antagonist of both G protein- and  $\beta$ -arrestin-mediated signaling. *Psychopharmacology (Berl)*. **232**, 917–930 (2015).
  57. Rinner, U. & Hudlicky, T. Synthesis of Morphine Alkaloids and Derivatives. in *Alkaloid Synthesis* (ed. Knölker, H.-J.) 33–66 (Springer Berlin Heidelberg, 2012). doi:10.1007/128\_2011\_133.
  58. Pittard, J. & Yang, J. Biosynthesis of the Aromatic Amino Acids. *EcoSal Plus* **3**, (2008).
  59. Juminaga, D. *et al.* Modular engineering of L-tyrosine production in *Escherichia coli*.

- Appl. Environ. Microbiol.* **78**, 89–98 (2012).
60. Gold, N. D. *et al.* Metabolic engineering of a tyrosine-overproducing yeast platform using targeted metabolomics. *Microb. Cell Fact.* **14**, 73 (2015).
  61. McKenna, R., Thompson, B., Pugh, S. & Nielsen, D. R. Rational and combinatorial approaches to engineering styrene production by *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **13**, 123 (2014).
  62. Patnaik, R., Zolanz, R. R., Green, D. A. & Kraynie, D. F. Tyrosine Production by Recombinant *Escherichia coli*: Fermentation Optimization and Recovery. *Biotechnol. Bioeng.* **99**, 741–752 (2008).
  63. Rodriguez, A., Kildegaard, K. R., Li, M., Borodina, I. & Nielsen, J. Establishment of a yeast platform strain for production of p-coumaric acid through metabolic engineering of aromatic amino acid biosynthesis. *Metab. Eng.* **31**, 181–188 (2015).
  64. Westfall, P. J. *et al.* Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 111–118 (2012).
  65. Liu, Q., Liu, Y., Chen, Y. & Nielsen, J. Current state of aromatics production using yeast: achievements and challenges. *Curr. Opin. Biotechnol.* **65**, 65–74 (2020).
  66. Liu, Y. *et al.* Genetic engineering of *Escherichia coli* to improve L-phenylalanine production. *BMC Biotechnol.* **18**, 1–12 (2018).
  67. Liu, Q. *et al.* Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. *Nat. Commun.* **10**, 1–13 (2019).
  68. Guo, W. *et al.* Rewiring central carbon metabolism for tyrosol and salidroside production in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **117**, 2410–2419 (2020).
  69. Kunjapur, A. M., Tarasova, Y. & Prather, K. L. J. Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli*. *J. Am. Chem. Soc.* **136**, 11644–11654 (2014).
  70. Hazelwood, L. a, Daran, J.-M., van Maris, A. J. a, Pronk, J. T. & Dickinson, J. R. The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. *Appl. Environ. Microbiol.* **74**, 2259–66 (2008).
  71. Kunjapur, A. M. & Prather, K. L. J. Microbial engineering for aldehyde synthesis. *Appl. Environ. Microbiol.* **81**, AEM.03319-14 (2015).
  72. Rodriguez, G. M. & Atsumi, S. Toward aldehyde and alkane production by removing aldehyde reductase activity in *Escherichia coli*. *Metab. Eng.* **25**, 227–237 (2014).
  73. Hawkins, K. METABOLIC ENGINEERING OF SACCHAROMYCES CEREVISIAE FOR THE PRODUCTION OF BENZYLISOQUINOLINE ALKALOIDS Thesis by. vol. 2009 (Stanford, 2009).
  74. Matsumura, E. *et al.* Microbial production of novel sulphated alkaloids for drug discovery. *Sci. Rep.* **8**, 1–10 (2018).
  75. Trenchard, I. J., Siddiqui, M. S., Thodey, K. & Smolke, C. D. De novo production of the key branch point benzylisoquinoline alkaloid reticuline in yeast. *Metab. Eng.* (2015) doi:10.1016/j.ymben.2015.06.010.
  76. Cushing, M. L. The Oxidation of Catechol-type Substrates by Tyrosinase. *J. Am. Chem. Soc.* **70**, 1184–1187 (1948).
  77. Prieto, M. a, Perez-Aranda, a & Garcia, J. L. Characterization of an *Escherichia coli* aromatic hydroxylase with a broad substrate range. *J. Bacteriol.* **175**, 2162–7 (1993).
  78. Claus, H. & Decker, H. Bacterial tyrosinases. *Syst. Appl. Microbiol.* **29**, 3–14 (2006).

79. Hatlestad, G. J. *et al.* The beet R locus encodes a new cytochrome P450 required for red betalain production. *Nat. Genet.* **44**, 816–820 (2012).
80. Daubner, S. C., Le, T. & Wang, S. Tyrosine hydroxylase and regulation of dopamine synthesis. *Arch. Biochem. Biophys.* **508**, 1–12 (2011).
81. Nakagawa, A. *et al.* (R,S)-tetrahydropapaveroline production by stepwise fermentation using engineered *Escherichia coli*. *Sci. Rep.* **4**, 6695 (2014).
82. Ehrenworth, a. M., Sarria, S. & Peralta-Yahya, P. Pterin-Dependent Mono-oxidation for the Microbial Synthesis of a Modified Monoterpene Indole Alkaloid. *ACS Synth. Biol.* 150729131430002 (2015) doi:10.1021/acssynbio.5b00025.
83. Koopman, F. *et al.* De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **11**, 155 (2012).
84. Horwitz, A. A. *et al.* Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Syst.* 1–9 (2015) doi:10.1016/j.cels.2015.02.001.
85. Sunnadeniya, R. *et al.* Tyrosine hydroxylation in betalain pigment biosynthesis is performed by cytochrome P450 enzymes in beets (*Beta vulgaris*). *PLoS One* **11**, 1–16 (2016).
86. Stöckigt, J., Antonchick, A. P., Wu, F. & Waldmann, H. The Pictet-Spengler Reaction in Nature and in Organic Chemistry. *Angew. Chemie - Int. Ed.* **50**, 8538–8564 (2011).
87. Minami, H., Dubouzet, E., Iwasa, K. & Sato, F. Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J. Biol. Chem.* **282**, 6274–6282 (2007).
88. Kim, J.-S. *et al.* Improvement of reticuline productivity from dopamine by using engineered *Escherichia coli*. *Biosci. Biotechnol. Biochem.* **77**, 2166–8 (2013).
89. Vavricka, C. J. *et al.* Mechanism-based tuning of insect 3,4-dihydroxyphenylacetaldehyde synthase for synthetic bioproduction of benzyloquinoline alkaloids. *Nat. Commun.* **10**, 1–11 (2019).
90. Vavricka, C. *et al.* Machine learning discovery of missing links that mediate alternative branches to plant alkaloids. 1–26.
91. Lichman, B. R. *et al.* ‘Dopamine-first’ mechanism enables the rational engineering of the norcoclaurine synthase aldehyde activity profile. *FEBS J.* **282**, 1137–1151 (2015).
92. Trenchard, I. J. & Smolke, C. D. Engineering strategies for the fermentative production of plant alkaloids in yeast. *Metab. Eng.* 1–9 (2015) doi:10.1016/j.ymben.2015.05.001.
93. Ziegler, J. *et al.* Comparative transcript and alkaloid profiling in *Papaver* species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. *Plant J.* **48**, 177–192 (2006).
94. Winzer, T. *et al.* Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. *Science (80-. ).* **349**, 309–312 (2015).
95. Thodey, K., Galanie, S. & Smolke, C. D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* **10**, 837–844 (2014).
96. Unterlinner, B., Lenz, R. & Kutchan, T. M. Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy *Papaver somniferum*. *Plant J.* **18**, 465–75 (1999).
97. Hagel, J. M. & Facchini, P. J. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. *Nat. Chem. Biol.* **6**, 273–275 (2010).
98. Fossati, E., Narcross, L., Ekins, A., Falguyret, J.-P. & Martin, V. J. J. Synthesis of morphinan alkaloids in *Saccharomyces cerevisiae*. *PLoS One* **10**, 30124459 (2015).

99. Theuns, H. G. *et al.* Neodihydrothebaine and bractazonine, two dibenz[d,f]azonine alkaloids of *Papaver bracteatum*. *Phytochemistry* **23**, 1157–1166 (1984).
100. Chen, X. *et al.* A pathogenesis-related 10 protein catalyzes the final step in thebaine biosynthesis article. *Nat. Chem. Biol.* **14**, 738–743 (2018).
101. Dastmalchi, M. *et al.* Neopinone isomerase is involved in codeine and morphine biosynthesis in opium poppy. *Nat. Chem. Biol.* **15**, 384–390 (2019).
102. Beaudoin, G. A. W. Characterization of Oxidative Enzymes Involved in the Biosynthesis of Benzyloquinoline Alkaloids in Opium Poppy (*Papaver somniferum*). (University of Calgary, 2015).
103. Dang, T. T. T., Chen, X. & Facchini, P. J. Acetylation serves as a protective group in noscapine biosynthesis in opium poppy. *Nat. Chem. Biol.* **11**, 104–106 (2015).
104. Li, Y. & Smolke, C. D. Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. *Nat. Commun.* **7**, (2016).
105. Chen, X., Dang, T. T. T. & Facchini, P. J. Noscapine comes of age. *Phytochemistry* **111**, 7–13 (2015).
106. Jakočiūnas, T. *et al.* Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* 1–10 (2015) doi:10.1016/j.ymben.2015.01.008.
107. Dang, T. T. T., Onoyovwi, A., Farrow, S. C. & Facchini, P. J. *Biochemical genomics for gene discovery in benzyloquinoline alkaloid biosynthesis in opium poppy and related species. Methods in Enzymology* vol. 515 (Elsevier Inc., 2012).
108. Desgagné-Penix, I. & Facchini, P. J. Systematic silencing of benzyloquinoline alkaloid biosynthetic genes reveals the major route to papaverine in opium poppy. *Plant J.* **72**, 331–44 (2012).
109. Matasci, N. *et al.* Data access for the 1,000 Plants (1KP) project. *Gigascience* **3**, 1–10 (2014).
110. Facchini, P. J. *et al.* Synthetic biosystems for the production of high-value plant metabolites. *Trends Biotechnol.* **30**, 127–131 (2012).
111. Oye, K., Bubela, T. & Lawson, J. C. H. Regulate ‘home-brew’ opiates. *Nature* **521**, 281–283 (2015).
112. Nam, H. *et al.* Network Context and Selection in the Evolution to Enzyme Specificity. *Science (80-. )*. **337**, 1101–1104 (2012).
113. Fischbach, M. A. & Clardy, J. One pathway, many products. *Nat. Chem. Biol.* **3**, 353–355 (2007).
114. Shin, J. H., Kim, H. U., Kim, D. I. & Lee, S. Y. Production of bulk chemicals via novel metabolic pathways in microorganisms. *Biotechnol. Adv.* **31**, 925–935 (2013).
115. Aharoni, A. *et al.* The ‘evolvability’ of promiscuous protein functions. *Nat. Genet.* **37**, 73–76 (2005).
116. Yoshikuni, Y., Ferrin, T. E. & Keasling, J. D. Designed divergent evolution of enzyme function. *Nature* **440**, 1078–1082 (2006).
117. Jones, J. A., Toparlak, Ö. D. & Koffas, M. A. Metabolic pathway balancing and its role in the production of biofuels and chemicals. *Curr. Opin. Biotechnol.* **33**, 52–59 (2015).
118. Furubayashi, M. *et al.* A highly selective biosynthetic pathway to non-natural C50 carotenoids assembled from moderately selective enzymes. *Nat. Commun.* **6**, 7534 (2015).
119. Martin, V., Narcross, L., Ekins, A., Fossati, E. & Zhu, Y. Method of making a benzyloquinoline alkaloid (BIA) metabolite, enzymes therefore. (2015).
120. Smolke, C. D., Thodey, C., Trenchard, I. & Galanie, S. Benzyloquinoline alkaloids (bia)

- producing microbes, and methods of making and using the same. (2017).
121. Tang, X., Tan, Y., Zhu, H., Zhao, K. & Shen, W. Microbial conversion of glycerol to 1,3-propanediol by an engineered strain of escherichia coli. *Appl. Environ. Microbiol.* **75**, 1628–1634 (2009).
  122. NCBI. Genbank surpasses one trillion total bases of publicly available sequence data. <http://www.ncbi.nlm.nih.gov/news/01-23-2015-genbank-trillion-bases/>.
  123. Xiao, M. *et al.* Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotechnol.* **166**, 122–134 (2013).
  124. Ferrer, M. *et al.* Estimating the success of enzyme bioprospecting through metagenomics: Current status and future trends. *Microb. Biotechnol.* **9**, 22–34 (2016).
  125. Keeling, C. I. *et al.* Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *BMC Plant Biol.* **11**, 43 (2011).
  126. Liscombe, D. K., Ziegler, J., Schmidt, J., Ammer, C. & Facchini, P. J. Targeted metabolite and transcript profiling for elucidating enzyme function: isolation of novel N-methyltransferases from three benzyloquinoline alkaloid-producing species. *Plant J.* **60**, 729–43 (2009).
  127. Augustin, M. M. *et al.* Elucidating steroid alkaloid biosynthesis in *Veratrum californicum*: Production of verazine in Sf9 cells. *Plant J.* **82**, 991–1003 (2015).
  128. Marques, J. V *et al.* Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis. *J. Biol. Chem.* **288**, 466–79 (2013).
  129. Lau, W. & Sattely, E. S. Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science (80-. )*. **349**, 1224–28 (2015).
  130. Moniodis, J. *et al.* The transcriptome of sesquiterpenoid biosynthesis in heartwood xylem of Western Australian sandalwood (*Santalum spicatum*). *Phytochemistry* **113**, 79–86 (2015).
  131. Liu, Q. *et al.* Reconstitution of the costunolide biosynthetic pathway in yeast and *Nicotiana benthamiana*. *PLoS One* **6**, (2011).
  132. Qingzhu, H. *et al.* Transcriptomic analysis reveals key genes related to Betalain biosynthesis in pulp coloration of *Hylocereus polyrhizus*. *Front. Plant Sci.* **6**, 1–13 (2016).
  133. Carlson, R. Time for new DNA synthesis and sequencing cost curves. <http://www.synthesis.cc/2014/02/time-for-new-cost-curves-2014.html>.
  134. Eriksen, D. T., Hsieh, P. C. H., Lynn, P. & Zhao, H. Directed evolution of a cellobiose utilization pathway in *Saccharomyces cerevisiae* by simultaneously engineering multiple proteins. *Microb. Cell Fact.* **12**, 1–11 (2013).
  135. Bird, D. A. & Facchini, P. J. Berberine bridge enzyme, a key branch-point enzyme in benzyloquinoline alkaloid biosynthesis, contains a vacuolar sorting determinant. *Planta* **213**, 888–897 (2001).
  136. Reitman, Z. J. *et al.* Enzyme redesign guided by cancer-derived IDH1 mutations. *Nat. Chem. Biol.* **8**, 887–889 (2012).
  137. Bayer, T. S. *et al.* Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* **131**, 6508–6515 (2009).
  138. Jendresen, C. B. *et al.* Highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **81**, 4458–4476 (2015).
  139. Díaz Chávez, M. L., Rolf, M., Gesell, A. & Kutchan, T. M. Characterization of two

- methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy *Argemone mexicana*. *Arch. Biochem. Biophys.* **507**, 186–93 (2011).
140. Szczebara, F. M. *et al.* Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat. Biotechnol.* **21**, 143–149 (2003).
  141. Jiang, H., Wood, K. V. & Morgan, J. A. Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **71**, 2962–2969 (2005).
  142. Campbell, A. *et al.* Engineering of a Nepetalactol-Producing Platform Strain of *Saccharomyces cerevisiae* for the Production of Plant Seco-Iridoids. *ACS Synth. Biol.* **5**, 405–414 (2016).
  143. Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J. & Dueber, J. E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res.* **41**, 10668–10678 (2013).
  144. Leonard, E. *et al.* Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13654–13659 (2010).
  145. Avalos, J. L., Fink, G. R. & Stephanopoulos, G. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **31**, 335–41 (2013).
  146. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
  147. Choi, K. B., Morishige, T., Shitan, N., Yazaki, K. & Sato, F. Molecular cloning and characterization of coclaurine N-methyltransferase from cultured cells of *Coptis japonica*. *J. Biol. Chem.* **277**, 830–835 (2002).
  148. Choi, K. B., Morishige, T. & Sato, F. Purification and characterization of coclaurine N-methyltransferase from cultured *Coptis japonica* cells. *Phytochemistry* **56**, 649–655 (2001).
  149. Samanani, N., Park, S.-U. & Facchini, P. J. Cell type-specific localization of transcripts encoding nine consecutive enzymes involved in protoberberine alkaloid biosynthesis. *Plant Cell* **17**, 915–26 (2005).
  150. Ikezawa, N. *et al.* Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured *Coptis japonica* cells. *J. Biol. Chem.* **278**, 38557–65 (2003).
  151. Ikezawa, N., Iwasa, K. & Sato, F. Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylophine biosynthesis in *Eschscholzia californica*. *FEBS J.* **274**, 1019–35 (2007).
  152. Ikezawa, N., Iwasa, K. & Sato, F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in *Eschscholzia californica*. *Plant Cell Rep.* **28**, 123–33 (2009).
  153. Dang, T. T. T. & Facchini, P. J. Cloning and characterization of canadine synthase involved in noscapine biosynthesis in opium poppy. *FEBS Lett.* **588**, 198–204 (2014).
  154. Gesell, A. *et al.* CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy. *J. Biol. Chem.* **284**, 24432–24442 (2009).
  155. Bauer, W. & H. Zenk, M. Two methylenedioxy bridge forming cytochrome P-450 dependent enzymes are involved in (S)-stylophine biosynthesis. *Phytochemistry* **30**, 2953–2961 (1991).

156. Nelson, D. R. Cytochrome P450 Nomenclature, 2004. in *Cytochrome P450 Protocols* (eds. Phillips, I. R. & Shephard, E. A.) 1–10 (Humana Press, 2006). doi:10.1385/1-59259-998-2:1.
157. Morishige, T., Tsujita, T., Yamada, Y. & Sato, F. Molecular characterization of the S-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase involved in isoquinoline alkaloid biosynthesis in *Coptis japonica*. *J. Biol. Chem.* **275**, 23398–23405 (2000).
158. Góngora-Castillo, E. *et al.* Development of Transcriptomic Resources for Interrogating the Biosynthesis of Monoterpene Indole Alkaloids in Medicinal Plant Species. *PLoS One* **7**, (2012).
159. Li, Y. *et al.* RNA-seq based de novo transcriptome assembly and gene discovery of *Cistanche deserticola* fleshy stem. *PLoS One* **10**, 1–19 (2015).
160. Näätäsaari, L., Krainer, F. W., Schubert, M., Glieder, A. & Thallinger, G. G. Peroxidase gene discovery from the horseradish transcriptome. *BMC Genomics* **15**, 1–16 (2014).
161. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499–507 (2014).
162. Shao, Z., Zhao, H. & Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **37**, 1–10 (2009).
163. Luo, Y., Enghiad, B. & Zhao, H. New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Nat. Prod. Rep.* **33**, 174–182 (2016).
164. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
165. Dodd, D., Kiyonari, S., Mackie, R. I. & Cann, I. K. O. Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium *Prevotella bryantii* B14. *J. Bacteriol.* **192**, 2335–2345 (2010).
166. Jing, F. *et al.* Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochem.* **12**, 1–16 (2011).
167. Rueffer, M., Zumstein, G. & Zenk, M. H. Partial purification and properties of S-adenosyl-L-methionine: (S)-tetrahydroprotoberberine-cis-N-methyltransferase from suspension-cultured cells of *Eschscholtzia* and *Corydalis*. *Phytochemistry* **29**, 3727–3733 (1990).
168. O'Keefe, B. R. & Beecher, C. Isolation and Characterization of S-Adenosyl-L-Methionine:Tetrahydroberberine-cis-N-Methyltransferase from Suspension Cultures of *Sanguinaria canadensis* L. *Plant Physiol.* **105**, 395–403 (1994).
169. Brown, S., Clastre, M., Courdavault, V. & O'Connor, S. E. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3205–3210 (2015).
170. Pompon, D., Louerat, B., Bronine, A. & Urban, P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64 (1996).
171. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
172. Jennewein, S. *et al.* Coexpression in yeast of *Taxus* cytochrome P450 reductase with cytochrome p450 oxygenases involved in taxol biosynthesis. *Biotechnol. Bioeng.* **89**, 588–598 (2005).

173. Trantas, E., Panopoulos, N. & Ververidis, F. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metab. Eng.* **11**, 355–66 (2009).
174. Dai, Z. *et al.* Producing aglycons of ginsenosides in bakers' yeast. *Sci. Rep.* **4**, 1–6 (2014).
175. Bar-Even, A. & Salah Tawfik, D. Engineering specialized metabolic pathways--is there a room for enzyme improvements? *Curr. Opin. Biotechnol.* **24**, 310–9 (2013).
176. Beaudoin, G. a W. & Facchini, P. J. Isolation and characterization of a cDNA encoding (S)-cis-N-methylstylopine 14-hydroxylase from opium poppy, a key enzyme in sanguinarine biosynthesis. *Biochem. Biophys. Res. Commun.* **431**, 597–603 (2013).
177. Dang, T.-T. T. & Facchini, P. J. CYP82Y1 Is N-Methylcanadine 1-Hydroxylase, a Key Noscapine Biosynthetic Enzyme in Opium Poppy. *J. Biol. Chem.* **289**, 2013–26 (2014).
178. Min, X. J., Butler, G., Storms, R. & Tsang, A. OrfPredictor: Predicting protein-coding regions in EST-derived sequences. *Nucleic Acids Res.* **33**, 677–680 (2005).
179. Nelson, D. R. The cytochrome p450 homepage. *Hum. Genomics* **4**, 59–65 (2009).
180. Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. 545–551 (2009) doi:10.1002/yea.
181. Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D. & Hegemann, J. H. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**, 23 (2002).
182. Carter, Z. & Delneri, D. New generation of loxP-mutated deletion cassettes for the genetic manipulation of yeast natural isolates. *Yeast* **27**, 765–775 (2010).
183. Ryan, O. W. *et al.* Selection of chromosomal DNA libraries using a multiplex CRISPR system. *Elife* e03703 (2014) doi:10.7554/eLife.03703.
184. Schumacher, H. M. & Zenk, M. H. Partial purification and characterization of dihydrobenzophenanthridine oxidase from *Eschscholtzia californica* cell suspension cultures. *Plant Cell Rep.* **7**, 43–46 (1988).
185. Pyne, M. E. *et al.* *Reconstituting Plant Secondary Metabolism in Saccharomyces cerevisiae for Production of High-Value Benzylisoquinoline Alkaloids. Methods in Enzymology* vol. 575 (Elsevier Inc., 2016).
186. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
187. Nelson, D. R. & Schuler, M. a. Cytochrome P450 Genes from the Sacred Lotus Genome. *Trop. Plant Biol.* **6**, 138–151 (2013).
188. van Dijken, J. . *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb. Technol.* **26**, 706–714 (2000).
189. Kneen, M. M. *et al.* Characterization of a thiamin diphosphate-dependent phenylpyruvate decarboxylase from *Saccharomyces cerevisiae*. *FEBS J.* **278**, 1842–1853 (2011).
190. Romagnoli, G., Luttik, M. A. H., Kötter, P., Pronk, J. T. & Daran, J. M. Substrate specificity of thiamine pyrophosphate-dependent 2-oxo-acid decarboxylases in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **78**, 7538–7548 (2012).
191. Nakagawa, A. *et al.* A bacterial platform for fermentative production of plant alkaloids. *Nat. Commun.* **2**, 1–8 (2011).
192. Gaweska, H. & Fitzpatrick, P. F. Structures and mechanism of the monoamine oxidase family. *Biomol. Concepts* **2**, 365–377 (2011).

193. Son, S. Y. *et al.* Structure of human monoamine oxidase A at 2.2-Å resolution: The control of opening the entry for substrates/inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5739–5744 (2008).
194. Grewal, P. S., Samson, J. A., Baker, J. J., Choi, B. & Dueber, J. E. Peroxisome compartmentalization of a toxic enzyme improves alkaloid production. *Nat. Chem. Biol.* **17**, 96–103 (2021).
195. Davis, V. E. & Walsh, M. J. Alcohol, amines, and alkaloids: A possible biochemical basis for alcohol addiction. *Science (80-. )*. **167**, 1005–1007 (1970).
196. Canelas, A. B. *et al.* Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains. *Nat. Commun.* **1**, (2010).
197. Narcross, L., Bourgeois, L., Fossati, E., Burton, E. & Martin, V. J. J. Mining Enzyme Diversity of Transcriptome Libraries through DNA Synthesis for Benzylisoquinoline Alkaloid Pathway Optimization in Yeast. *ACS Synth. Biol.* **5**, 1505–1518 (2016).
198. Absolínová, H., Jančář, L., Jančářová, I., Vičar, J. & Kubáň, V. Acid-base behaviour of sanguinarine and dihydrosanguinarine. *Cent. Eur. J. Chem.* **7**, 876–883 (2009).
199. Dostál, J. Two Faces of Alkaloids. *J. Chem. Educ.* **77**, 993–998 (2000).
200. Hori, K. *et al.* Mining of the Uncharacterized Cytochrome P450 Genes Involved in Alkaloid Biosynthesis in California Poppy Using a Draft Genome Sequence. *Plant Cell Physiol.* **59**, 222–233 (2018).
201. Dostál, J. *et al.* Sanguinarine pseudobase: Re-examination of NMR assignments using gradient-enhanced spectroscopy. *Magn. Reson. Chem.* **36**, 869–872 (1998).
202. Brochado, A. R. *et al.* Improved vanillin production in baker's yeast through in silico design. *Microb. Cell Fact.* **9**, 84 (2010).
203. Suástegui, M. *et al.* Multilevel engineering of the upstream module of aromatic amino acid biosynthesis in *Saccharomyces cerevisiae* for high production of polymer and drug precursors. *Metab. Eng.* **42**, 134–144 (2017).
204. Wang, G. *et al.* Improvement of cis,cis-Muconic Acid Production in *Saccharomyces cerevisiae* through Biosensor-Aided Genome Engineering. *ACS Synth. Biol.* **9**, 634–646 (2020).
205. Hassing, E. J., de Groot, P. A., Marquenie, V. R., Pronk, J. T. & Daran, J. M. G. Connecting central carbon and aromatic amino acid metabolisms to improve de novo 2-phenylethanol production in *Saccharomyces cerevisiae*. *Metab. Eng.* **56**, 165–180 (2019).
206. Bekers, K. M., Heijnen, J. J. & van Gulik, W. M. Determination of the in vivo NAD:NADH ratio in *Saccharomyces cerevisiae* under anaerobic conditions, using alcohol dehydrogenase as sensor reaction. *Yeast* **32**, 541–557 (2015).
207. Canelas, A. B., Van Gulik, W. M. & Heijnen, J. J. Determination of the cytosolic free NAD/NADH ratio in *Saccharomyces cerevisiae* under steady-state and highly dynamic conditions. *Biotechnol. Bioeng.* **100**, 734–743 (2008).
208. Zhang, J. *et al.* Determination of the Cytosolic NADPH/NADP Ratio in *Saccharomyces cerevisiae* using Shikimate Dehydrogenase as Sensor Reaction. *Sci. Rep.* **5**, 1–12 (2015).
209. Zhao, W. *et al.* Biosynthesis of plant-specific alkaloids tetrahydroprotoberberines in engineered *Escherichia coli*. *Green Chem.* **12**, 810–812 (2021).
210. Smith, J. D. *et al.* Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol.* **17**, 1–16 (2016).
211. Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J. & Mateo, J. L. CCTop: An intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* **10**, 1–11

- (2015).
212. Babaei, M. *et al.* Expansion of EasyClone-MarkerFree toolkit for *Saccharomyces cerevisiae* genome with new integration sites. *FEMS Yeast Res.* **21**, 1–6 (2021).
  213. Jensen, N. B. *et al.* EasyClone: Method for iterative chromosomal integration of multiple genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **14**, 238–248 (2014).
  214. DeLoache, W. C., Russ, Z. N. & Dueber, J. E. Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. *Nat. Commun.* **7**, (2016).
  215. Westermann, B. & Neupert, W. Mitochondria-targeted green fluorescent proteins: Convenient tools for the study of organelle biogenesis in *Saccharomyces cerevisiae*. *Yeast* **16**, 1421–1427 (2000).
  216. Ralsler, M., Mülleder, M., Campbell, K., Matsarskaia, O. & Eckerstorfer, F. *Saccharomyces cerevisiae* single-copy plasmids for auxotrophy compensation, multiple marker selection, and for designing metabolically cooperating communities. *F1000Research* **5**, 1–12 (2016).
  217. Imanpoor, M. R. & Roohi, Z. Effects of Sangrovit-supplemented diet on growth performance, blood biochemical parameters, survival and stress resistance to salinity in the Caspian roach (*Rutilus rutilus*) fry. *Aquac. Res.* **47**, 2874–2880 (2016).
  218. Chang, L., Hagel, J. M. & Fachini, P. J. Isolation and characterization of O-methyltransferases involved in the biosynthesis of glaucine in *Glaucium flavum*. *Plant Physiol.* **169**, 1127–1140 (2015).
  219. D’Oelsnitz, S. *et al.* Using structurally fungible biosensors to evolve improved alkaloid biosyntheses. *bioRxiv* 1–36 (2021) doi:10.1101/2021.06.07.447399.
  220. Winkler, A. *et al.* A concerted mechanism for berberine bridge enzyme. *Nat. Chem. Biol.* **4**, 739–741 (2008).
  221. Sage, R. F. Variation in the *k<sub>cat</sub>* of Rubisco in C3 and C4 plants and some implications for photosynthetic performance at high and low temperature. *J. Exp. Bot.* **53**, 609–620 (2002).
  222. Shi, X., Zhang, H. & Lin, S. Tandem repeats, high copy number and remarkable diel expression rhythm of form II RuBisCO in *Prorocentrum donghaiense* (Dinophyceae). *PLoS One* **8**, (2013).
  223. Bar-Even, A. *et al.* The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
  224. Bourgeois, L., Pyne, M. E. & Martin, V. J. J. A Highly Characterized Synthetic Landing Pad System for Precise Multicopy Gene Integration in Yeast. *ACS Synth. Biol.* **7**, 2675–2685 (2018).
  225. Shi, S., Liang, Y., Zhang, M. M., Ang, E. L. & Zhao, H. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **33**, 19–27 (2016).
  226. Bienick, M. S. *et al.* The interrelationship between promoter strength, gene expression, and growth rate. *PLoS One* **9**, (2014).
  227. Dastmalchi, M. *et al.* Purine Permease-Type Benzylisoquinoline Alkaloid Transporters in Opium Poppy. *Plant Physiol.* **181**, 916–933 (2019).
  228. Weng, J. K., Philippe, R. N. & Noel, J. P. The rise of chemodiversity in plants. *Science (80-. )*. **336**, 1667–1670 (2012).
  229. Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505 (2010).
  230. Noda-Garcia, L. & Tawfik, D. S. Enzyme evolution in natural products biosynthesis:

- target- or diversity-oriented? *Curr. Opin. Chem. Biol.* **59**, 147–154 (2020).
231. Bean, B. D. M. *et al.* Functional expression of opioid receptors and other human GPCRs in yeast engineered to produce human sterols. 1–14 (2021).
  232. Rix, G. *et al.* Scalable continuous evolution for the generation of diverse enzyme variants encompassing promiscuous activities. *Nat. Commun.* **11**, 1–11 (2020).
  233. Halperin, S. O. *et al.* CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* **560**, 248–252 (2018).
  234. Hagel, J. M. & Facchini, P. J. Subcellular localization of sanguinarine biosynthetic enzymes in cultured opium poppy cells. *Vitr. Cell. Dev. Biol. Plant.* **48**, 233–240 (2012).