

**METABOLIC ENGINEERING OF *SACCHAROMYCES*
CEREVISIAE FOR THE PRODUCTION OF
BENZYLISOQUINOLINE ALKALOIDS**

Thesis by
Kristy Hawkins

In partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
2009
(Defended October 28, 2008)

© 2009

Kristy Hawkins

All rights reserved

ACKNOWLEDGEMENTS

The completion of this thesis work was a long and arduous journey which has consumed the past five years of my life and I have many people to thank. First I would like to thank my research advisor Christina Smolke for taking me on as one of her first students. When I started in the lab, I did not even know how to use a pipette or basic principles of molecular biology. Christina gave us our first lessons in cloning but later allowed us the freedom to direct much of our own research work.

I also have to thank my colleagues in the lab that have become some of my best friends. In particular, Stephanie Culler is an amazing scientist and one of the most genuine people I have ever met. Her work ethic is unparalleled and she has kept me company during many late nights and weekends. I also have to acknowledge Dr. Kevin Hoff who has always taken the time to dwell on others' research problems, and more often than not, come up with some valuable suggestions if not solutions. Chase Beisel is another lab member who is always willing to lend an ear, and it has been a pleasure getting to know both him and his lovely wife Amy. Finally, I have to thank Michael Siddiqui for being brave enough to take on this challenging project so that my work building the BIA pathway is not lost. Michael also assisted me with some of the upstream pathway experiments presented here.

For never-ending support and unconditional love, I have to thank my parents Barbara and Terry Hawkins. They have always been incredible role models without even trying. And lastly, I have to recognize Branden Ray who has been my rock for the past

year. When I doubt myself, he reassures me; when my enthusiasm is lost, he helps me find it; and when I have a breakthrough, he is the first person I run to tell.

ABSTRACT

The engineering of synthetic metabolic pathways in microbial hosts holds much promise for the synthesis of new chemicals and materials, including a variety of natural and non-natural products. The benzyloisoquinoline alkaloids (BIAs) represent a large and structurally diverse class of plant secondary metabolites that exhibit a broad range of pharmacological activities. The reconstitution of a BIA biosynthetic pathway in an engineered microbial host offers several advantages over isolation from plants, including the targeted production of key intermediate molecules, rapid biomass accumulation, ease of purification, and the availability of genetic tools for strain engineering and pathway optimization.

Here we describe the development of a synthetic BIA pathway in an engineered yeast host which incorporates heterologous enzymes from a variety of organisms. The BIA backbone is derived from two molecules of tyrosine and is assembled through a heterologous pathway comprising enzymatic activities from plants, bacteria, and humans. Simultaneous efforts have focused on the downstream portion of this pathway to convert a commercially available substrate to the major branch point intermediate reticuline. By synthesizing both stereoisomers of reticuline from a racemic substrate, we have demonstrated production of BIA metabolites along the diversified sanguinarine/berberine and morphinan branches. Further optimization, scale-up, and a combination of bioconversions and chemical synthesis will potentially revolutionize drug discovery and manufacturing of these compounds.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT.....	v
LIST OF FIGURES AND TABLES.....	xi
CHAPTER I. INTRODUCTION	1
1.1. Yeast as a host organism	1
1.2. Tools for metabolic engineering in yeast	2
1.3. The benzyloquinoline alkaloid (BIA) pathway	4
1.4. Metabolic engineering of <i>S. cerevisiae</i> for the production of BIAs	6
CHAPTER II. THE REGULATORY ROLES OF THE GALACTOSE PERMEASE AND KINASE IN THE INDUCTION RESPONSE OF THE GAL NETWORK IN <i>SACCHAROMYCES CEREVISIAE</i>	8
Abstract.....	8
2.1. Introduction.....	9
2.2. Results	13
<i>2.2.1. Galactose permease deletion results in a linear induction response</i>	<i>13</i>
<i>2.2.2. Constitutive expression of the galactose permease results in a tunable linear response</i>	<i>14</i>
<i>2.2.3. Constitutive expression of regulatory proteins enhances the switch-like response of the network</i>	<i>16</i>
<i>2.2.4. Galactokinase deletion results in a regimed network response.....</i>	<i>20</i>

2.2.5. Population distributions in GAL2-modified strains exhibit graded responses	23
2.2.6. Deletion of the galactokinase results in multiple stable populations	25
2.3. Discussion.....	27
2.4. Materials and Methods.....	33
2.4.1. Yeast strain construction.....	33
2.4.2. Yeast expression plasmids.....	36
2.4.3. Fluorescence assays.....	38
2.4.4. Flow cytometry assays	38
 CHAPTER III. SYNTHESIS OF THE BENZYLISOQUINOLINE ALKALOID	
BACKBONE FROM TYROSINE IN SACCHAROMYCES CEREVISIAE	40
Abstract.....	40
3.1. Introduction.....	41
3.2. Results	42
3.2.1. Functional expression of tyrosine/dopa decarboxylase from <i>Papaver somniferum</i>	42
3.2.2. Production of dopamine.....	44
3.2.3. Production of 4-hydroxyphenylacetaldehyde and 3,4-dihydroxyphenylacetaldehyde.....	49
3.2.4. Production of norcoclaurine and norlaudanosoline.....	60
3.3. Discussion.....	67
3.4. Materials and Methods.....	68
3.4.1. Plasmid and yeast strain construction.....	68
3.4.2. Analysis of metabolite production	70

3.4.3. Analysis of protein levels through Western blotting	71
3.4.4. Monoamine oxidase and NCS in vitro assays.....	72

CHAPTER IV: PRODUCTION OF (R, S)-RETICULINE AND DOWNSTREAM BENZYLISOQUINOLINE ALKALOIDS IN SACCHAROMYCES CEREVISIAE 74

Abstract.....	74
4.1. Introduction.....	75
4.2. Results	76
4.2.1. Synthesis of reticuline from norlaudanoline in yeast	76
4.2.2. BIA production in the yeast strains is substrate limited	85
4.2.3. Tuning enzyme levels with a novel titration strategy.....	88
4.2.4. Synthesis of sanguinarine/berberine intermediates	93
4.2.5. N-methylation of (S)-tetrahydroberberine for the production of a noscapine pathway intermediate.....	99
4.2.6. Pathway for laudanine production	101
4.2.6. Synthesis of the morphinan intermediate salutaridine.....	102
4.3. Discussion.....	106
4.4. Materials and Methods.....	109
4.4.1. Plasmid and yeast strain construction.....	109
4.4.2. Growth conditions.....	113
4.4.3. Analysis of metabolite production	114
4.4.4. Fluorescence quantification.....	116
4.4.5. Analysis of protein levels through Western blotting	116
4.4.6. Analysis of transcript levels through qRT-PCR.....	117

4.4.7. Analysis of stereoisomer forms of BIA metabolites through capillary electrophoresis-based chiral separation.....	118
4.4.8. Preparation of cell extracts	119
CHAPTER V. PRODUCTION OF MORPHINAN ALKALOIDS IN SACCHAROMYCES CEREVISIAE	120
Abstract.....	120
5.1. Introduction.....	121
5.2. Results	123
5.2.1. Expression and activity of <i>Papaver somniferum</i> salutaridine synthase	123
5.2.2. Expression of <i>Papaver</i> salutaridine reductase variants for the production of salutaridinol.....	124
5.2.3. Expression of salutaridinol-7-O-acetyltransferase and development of a process for thebaine production	128
5.3. Discussion.....	138
5.4. Materials and Methods.....	139
5.4.1. Plasmid and yeast strain construction.....	139
5.4.2. Cell growth conditions.....	141
5.4.3. Analysis of metabolite production	141
5.4.4. Analysis of protein levels through Western blotting.....	142
5.4.5. Fluorescence quantification.....	143
CHAPTER VI. CONCLUSIONS AND FUTURE WORK.....	144
6.1. A new tool for tuning protein expression.....	144
6.2. De novo biosynthesis of BIA backbone molecules.....	145

6.3. Production of the intermediate reticuline and downstream berberine and morphinan alkaloids	146
6.4. Construction of a strain to produce downstream BIAs from tyrosine	148
REFERENCES.....	149

LIST OF FIGURES AND TABLES

Figure 1.1. General schematic of the BIA pathway	6
Figure 2.1. Diagram of the native GAL network.....	11
Figure 2.2. Population-averaged response from strains with altered Gal2p regulation...	16
Figure 2.3. Population-averaged response from strains with altered Gal3p and Gal80p regulation.....	19
Figure 2.4. Population-averaged response from strains with no Gal1p activity.....	22
Figure 2.5. Population response from strains with altered Gal2p regulation.....	24
Figure 2.6. Population response from strains with no Gal1p activity.....	26
Figure 2.7. Cassettes for chromosomal replacements.....	34
Table 2.1. List of yeast strains.....	35
Table 2.2. List of primer sequences used in the construction of plasmids and yeast strains.....	36
Figure 3.1. The native pathway for the production of the BIA precursor (<i>S</i>)-norcoclaurine synthesized from two molecules of tyrosine.	41
Figure 3.2. <i>In vivo</i> TYDC2 assays.....	43
Figure 3.3. Alternative pathways for dopamine production from tyrosine.....	44
Figure 3.4. LC-MS analysis of the growth media of CSY88 expressing AbPPO2 and TYDC2.....	46
Figure 3.5. Western blots of AbPPO2 constructs.....	47
Figure 3.6. LC-MS analysis of dopamine produced in yeast.....	49
Figure 3.7. Analysis of tyrosine derivatives.....	50
Figure 3.8. Monoamine oxidase enzyme assays.....	52

Figure 3.9. Production of norlaudanosoline from dopamine.....	54
Figure 3.10. Strains with increased norlaudanosoline accumulation.....	56
Figure 3.11. Pathway for the production of 4-HPA from tyrosine using the endogenous yeast genes <i>ARO8/9</i> and <i>ARO10</i>	57
Figure 3.12. Analysis of yeast strains overexpressing the <i>ARO</i> genes.....	59
Figure 3.13. Changes in 4-HPA, tyrosol, and 4-hydroxyphenylacetic acid production in knockout strains overexpressing Aro9p and Aro10p.....	60
Figure 3.14. Pathways for norcoclaurine and norlaudanosoline production.....	61
Figure 3.15. Expression of NCS and production of norcoclaurine.....	63
Figure 3.16. Comparison of norlaudanosoline production with NCS.....	65
Figure 3.17. Norcoclaurine production in <i>ADH</i> and <i>ALD</i> knockout strains.....	66
Table 3.1. Plasmids and yeast strains.....	70
Figure 4.1. The native BIA pathway.....	77
Figure 4.2. The engineered BIA pathway for the synthesis of (<i>R, S</i>)-reticuline from (<i>R, S</i>)-norlaudanosoline.....	78
Figure 4.3. Plasmid maps of BIA expression constructs.....	79
Figure 4.4. LC-MS/MS analysis confirms individual methyltransferase activities.....	80
Figure 4.5. Western blot analysis confirms expression of 6OMT, CNMT, and 4'OMT enzymes.....	81
Figure 4.6. Microbial production of (<i>R, S</i>)-reticuline.....	83
Figure 4.7. BIA metabolites accumulate in the growth media.....	84
Figure 4.8. Effects of enzyme levels, substrate levels, and culture time on reticuline production.....	87

Figure 4.9. A novel strategy for tuning enzyme expression levels.....	92
Figure 4.10. Microbial production of BIA metabolites along the sanguinarine and berberine branches.....	95
Figure 4.11. Comparison of BBE N-terminal truncations.....	97
Figure 4.12. Comparison of synthesized (<i>S</i>)-tetrahydroberberine to standard.....	98
Figure 4.13. Chromatograms show impurities and incomplete conversion of intermediates.....	99
Figure 4.14. Activity of TNMT on (<i>S</i>)-tetrahydroberberine.....	100
Figure 4.15. (<i>R, S</i>)-Reticuline 7OMT reaction and expression in yeast.....	101
Figure 4.16. Microbial production of morphinan alkaloids.....	105
Table 4.1. Engineered yeast strains.....	110
Table 4.2. Primer sequences used for qRT-PCR and to make stable integrations of enzyme constructs and <i>GAL2</i> knockouts.	113
Table 4.3. Summary of yields of benzyloisoquinoline alkaloids.....	115
Table 4.4. Primers used for qRT-PCR.....	118
Figure 5.1. Pathway for the production of thebaine from (<i>R</i>)-reticuline.....	122
Figure 5.2. Testing of SalR variants.....	125
Figure 5.3. Salutaridinol production.....	127
Figure 5.4. Pathways for the formation of thebaine and neodihydrothebaine.....	128
Figure 5.5. LC-MS analysis of salutaridinol-7- <i>O</i> -acetate.....	129
Figure 5.6. Codon-optimization increases SalAT expression.....	130
Figure 5.7. Comparison of CYP2D6 and SalSyn for salutaridinol-7- <i>O</i> -acetate production.....	132

Figure 5.8. Comparison of SalSyn reductase partners and signal sequences.....133

Figure 5.9. Alternative promoter and signal sequence for SalAT.....135

Table 5.1. Engineered yeast strains and plasmids.....141

CHAPTER I. INTRODUCTION

1.1. Yeast as a host organism

The extensive history of *Saccharomyces cerevisiae* in baking, brewing, and wine making industries has landed this microorganism a starring role in both traditional and emerging biotechnologies. Knowledge of yeast genetics, physiology, biochemistry, genetic engineering and fermentation strategies has accumulated over time¹. Genetic engineering in yeast is simplified by the availability of highly efficient transformation methods, expression vectors of varying copy number, and the ability to do targeted homologous recombination. In addition, *S. cerevisiae* produces no toxic metabolites and is non-pathogenic, earning it a GRAS (generally regarded as safe) classification by the U.S. Food and Drug Administration (FDA). Many physical properties such as tolerance to low pH and robust growth under high sugar and ethanol conditions contribute to yeast's popularity as a microbial cell factory.

S. cerevisiae was also the first eukaryotic organism whose complete genome was sequenced. Sequence information is compiled on multiple websites such as <http://www.yeastgenome.org> along with information on open reading frames, biochemical pathways, microarray studies, and networks of protein interactions. This greatly facilitates chromosomal manipulations and the implementation of other strategies such as deletion or upregulation of endogenous metabolic pathways.

One of the major reasons yeast was chosen for this work is the similarity of its protein machinery to that of higher eukaryotes. Enzymes from plants and humans are more likely to be properly folded and proteolytically processed in yeast versus a

prokaryotic host. Most importantly, membrane-targeted proteins such as cytochromes P540 are actively expressed in the yeast endoplasmic reticulum. When a suitable P450 reductase partner is coexpressed, these enzymes function in yeast whereas most do not in common bacterial expression hosts such as *Escherichia coli*. Membrane-bound enzymes are highly prevalent in plant secondary metabolic pathways and often catalyze complex reactions leading to valuable natural products that cannot be duplicated by synthetic chemistry approaches.

1.2. Tools for metabolic engineering in yeast

Metabolic engineering is a relatively new field, defined by Bailey in 1991 as “the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology.”² Whereas classical methods of strain improvement rely on selection, mutagenesis, and mating, metabolic engineering takes a methodical approach to the redirection of metabolic fluxes. Advances in molecular biology techniques allow the targeted modification of strains without unwanted mutations and the introduction of foreign genes into a new host organism. This means that in addition to applications aimed at increasing the production of natural yeast metabolites such as ethanol, it is now possible to instill yeast with completely novel activities and entirely heterologous pathways.

Approaches for microbial strain improvement include rational metabolic engineering, evolutionary metabolic engineering, and inverse metabolic engineering. Rational metabolic engineering is based on knowledge of enzymes, transporters, regulatory proteins, and other cellular processes while evolutionary metabolic

engineering relies on mutagenesis and selection rather than design. Inverse metabolic engineering is the identification of the genotype corresponding to a desired phenotype and the recreation of that phenotype in a different strain or organism. For the development of a heterologous BIA pathway, the methodology thusfar has consisted only of rational engineering strategies. However, future optimization efforts may turn to evolutionary tactics.

This work used recombinant DNA technology to introduce new genes into the host organism *S. cerevisiae*. Initial experiments were performed using 2 μ m-based multicopy plasmids and the strong constitutive TEF1 promoter. Once the desired strain composition was determined, we performed chromosomal integrations of the necessary genes such that coding regions are maintained in the absence of selective pressure, a requirement for the development of industrial strains. This also reduces metabolic load on the cell and creates a more homogeneous population. However, making a genetic integration also reduces the DNA copy number from 10-30 to a single copy per cell. In some cases, expression was not limiting, but in others, production suffered when DNA copy number was decreased. In such cases, genes were maintained on plasmids, but ultimately, multiple copies must be integrated into the genome. This is only one of several methods we employed to modify protein expression levels.

Besides DNA copy number, the associated promoter sequence is a common target for modifying protein expression. Chapter II describes modified yeast strain backgrounds which alter the response of the native galactose-inducible promoter³. Of particular interest is the *GAL2 Δ* strain which produced a linear response, allowing fine tuning of gene expression with galactose concentration. We used this strain background to titrate

individual enzymes involved in the BIA pathway to determine optimal expression levels⁴. The results demonstrated that substitution of a weaker promoter would not compromise production in most cases⁵. The strains resulting from these promoter replacements exhibited improved fitness and facilitated expansion of the downstream pathway(s). Since no toxic metabolites are produced and enzyme expression does not compromise growth, we overwhelmingly chose to use constitutive rather than inducible promoters throughout this work. For some applications, promoters specific to a certain growth phase such as the *HSP30* stationary phase promoter are useful as they do not require addition of an exogenous inducer.

Another strategy to increase protein expression is to optimize translation efficiency. To this end, we codon-optimized several genes in the BIA pathway for *S. cerevisiae*, taking into account both codon frequency and mRNA secondary structure. In some cases, this did nothing to improve expression (hCYP2D6) and in others, a major increase in protein was observed (SalAT). However, overwhelmingly we found that bottlenecks in our system were due to *in vivo* activity rather than expression. Many such steps will require more advanced protein engineering and evolutionary strategies beyond the scope of this work to attain a viable process. Other metabolic engineering tactics featured in this work include examination of substrate transport and modification of signal sequences for protein localization and secretion.

1.3. The benzyloquinoline alkaloid (BIA) pathway

The benzyloquinoline alkaloids (BIA) are a rich family of plant natural products with over 2,500 elucidated structures. Many compounds in this group possess potent

pharmacological activities including the analgesic morphine, the antitussive codeine, the muscle relaxant papaverine, and the antimicrobial agents sanguinarine and berberine⁶. The secondary pathways that regulate BIA production are very intricate and tightly-controlled, induced as a protective response against herbivores and pathogens. BIA-producing plants have very specific alkaloid profiles which have proved difficult to alter through rational engineering approaches.

Among the most widely studied plant hosts are *Papaver somniferum* (opium poppy), *Thalictrum flavum*, *Eschscholzia californica*, and *Coptis japonica*. Thanks to these efforts, conducted largely with the goal of microbial production, many of the enzymes involved in these complex pathways have been cloned and characterized.

The first committed step in BIA biosynthesis is the formation of the backbone molecule (*S*)-norcoclaurine via the condensation of two tyrosine derivatives, dopamine and 4-hydroxyphenylacetaldehyde (4-HPA; Fig. 1.1). (*S*)-norcoclaurine undergoes a series of methylation reactions and a 3'-hydroxylation to yield the branch point intermediate (*S*)-reticuline. Nearly all BIAs, with the exception of the dimeric bisbenzylisoquinoline alkaloids and papaverine, are derived from (*S*)-reticuline. Various branches have been well characterized to-date including the morphinan, sanguinarine, and berberine branches. Although no single branch has been cloned in its entirety, the ongoing work by plant geneticists allows us to begin reconstructing pathways for the production of BIAs in a microbial host.

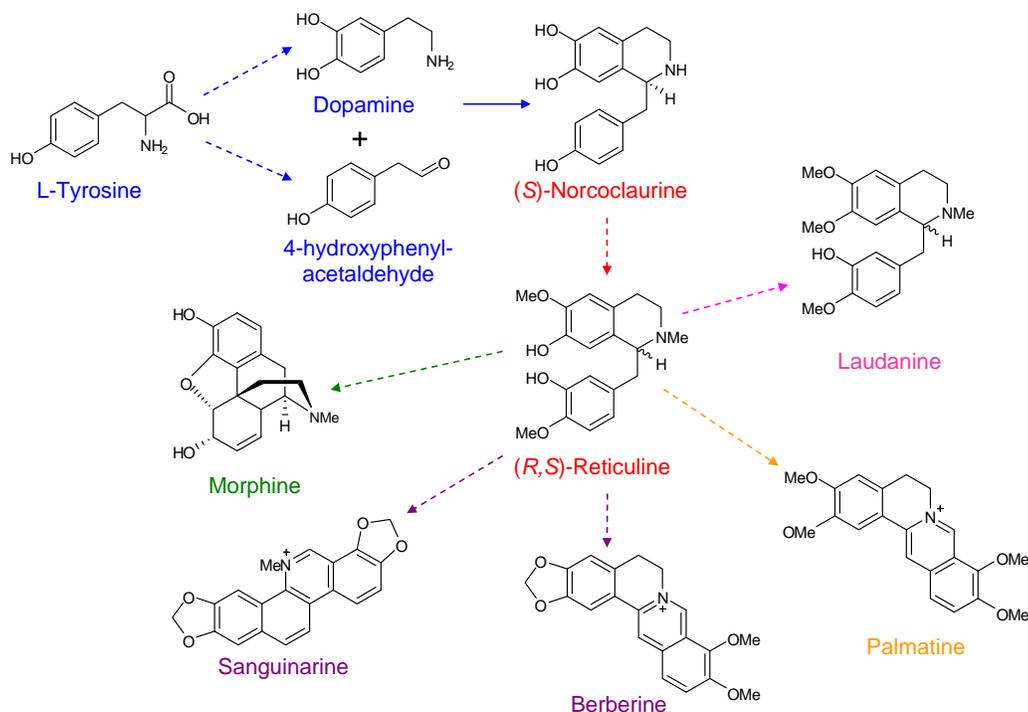


Fig. 1.1. General schematic of the BIA pathway. The upstream pathway (blue) shows the tyrosine derivatives dopamine and 4-HPA which make up the BIA backbone. The intermediate pathway (red) shows norcoclaurine, the first BIA structure, and reticuline, a major branch point intermediate. Additional branched metabolites are shown to represent the diversity of BIAs produced.

1.4. Metabolic engineering of *S. cerevisiae* for the production of BIAs

Production of BIAs in a microbial host offers several advantages to extraction from plants or chemical synthesis. Although synthesis schemes for morphine, for example, have been developed on a preparative scale, none are commercially viable. Extraction of specific BIA metabolites from plants requires rigorous purification and is somewhat limited by the natural product profiles. For instance, thebaine accounts for only 6.5% of the alkaloid content of opium poppy, but is a valuable molecule from which to derive other analgesic agents of use in the pharmaceutical industry. Reconstructing this pathway in a microbial host will allow for ‘green’ synthesis as well as accumulation of rare intermediates and even production of non-natural alkaloids.

Chapter III describes the reconstruction of the upstream portion of the BIA pathway from tyrosine to norcoclaurine. Since the enzymes catalyzing these reactions have not been elucidated in plants, a ‘bioprospecting’ approach was taken to pull activities from several diverse organisms to engineer these early steps.

Chapter IV describes the engineering of strains for the conversion of a commercially available substrate norlaudanosoline to the intermediate reticuline and beyond. Stable strains were constructed to express 6OMT, CMT, and 4’OMT activities at optimal levels. These basic strains were then used to perform additional steps to salutaridine, an intermediate along the morphinan branch using a novel activity for CYP2D6, and (*S*)-tetrahydroberberine, the penultimate metabolite along the berberine branch.

Chapter V describes an extension of the downstream morphinan branch to produce thebaine. We are able to accumulate salutaridinol-7-*O*-acetate but additional optimization and process development are needed to convert this intermediate to thebaine. Derivatives of thebaine itself are useful as analgesic pharmaceuticals and cloning of additional plant enzymes will allow production of other morphinan alkaloids.

This work covers nearly the entire BIA pathway with its many divergent branches and products. The upstream pathway lays the foundation for the *de novo* biosynthesis of the BIA backbone while simultaneous work on the later steps was conducted using a structurally similar intermediate. While significant optimization remains to be done, this work sets the stage for the construction of a single strain to convert tyrosine to potentially valuable complex BIAs.

**CHAPTER II. THE REGULATORY ROLES OF THE GALACTOSE
PERMEASE AND KINASE IN THE INDUCTION RESPONSE OF
THE GAL NETWORK IN *SACCHAROMYCES CEREVISIAE***

Adapted from Hawkins, K.M. & Smolke, C.D. *J Biol Chem* **281**, 13485-13492 (2006)

Abstract

The GAL genetic switch of *Saccharomyces cerevisiae* exhibits an ultrasensitive response to the inducer galactose as well as the ‘all-or-none’ behavior characteristic of many eukaryotic regulatory networks. We have constructed a strain which allows intermediate levels of gene expression from a tunable GAL1 promoter at both the population and single-cell level by altering the regulation of the galactose permease Gal2p. Analogous modifications to other feedback loops regulating the Gal80p repressor and the Gal3p signaling protein did not result in similarly tuned responses, indicating that the level of inducer transport is unique in its ability to control the switch response of the network. In addition, removal of the Gal1p galactokinase from the network resulted in a regimed response due to the dual role of this enzyme in galactose catabolism and transport. These two activities have competing effects on the response of the network to galactose such that transport effects of Gal1p are dominant at low galactose concentrations, whereas its catabolic effects are dominant at high galactose concentrations. In addition, flow cytometry analysis revealed the unexpected phenomenon of multiple populations in the *gal1Δ* strains which were not present in the isogenic *GAL1* background. This result indicates that Gal1p may play a previously undescribed role in the stability of the GAL network response.

2.1. Introduction

Saccharomyces cerevisiae inducible promoter systems have long been used for expression of heterologous proteins, gene function studies, and other areas of molecular genetics. Native inducible promoters such as GAL1⁷, MET25⁸, and CUP1^{9, 10}, although used successfully without modification, exhibit certain properties that are undesirable for many applications. One common feature of these systems is their autocatalytic or switch-like behavior, where addition of small amounts of inducer leads to large increases in gene expression. In prokaryotes and bacteriophages, this is generally due to cooperative interactions between transcription factors and promoter elements. In more complex eukaryotic networks, other elements such as feedback loops¹¹, zero-order sensitivity¹¹, multi-step signaling mechanisms¹¹, and nucleocytoplasmic transport of regulatory proteins¹² often contribute to nonlinear responses. In addition, native inducible promoter systems are often characterized by an all-or-none behavior, in which genes are either maximally expressed or virtually not expressed in individual cells¹³. In such cases the observed population-averaged response upon addition of inducer is due to an increase in the probability that a given cell will become fully induced. In contrast are systems that enable a homogenous cell population response and intermediate levels of gene expression in all cells proportional to the given stimulus¹³; however, examples of this are relatively rare in eukaryotic systems.

The widely used GAL promoter system is taken from an endogenous metabolic network regulating expression of a number of structural and regulatory genes required for efficient utilization of galactose as a primary carbon source (Fig. 2.1a). This complex and tightly-controlled network has served as a paradigm for gene regulatory circuits in

eukaryotic organisms. In noninducing-nonrepressing media, the Gal4p transcriptional activator binds as a dimer to recognition sites upstream of each galactose-regulated gene referred to as upstream activation sites (UASs). An inhibitory protein Gal80p dimerizes and binds to nuclear Gal4p in the absence of galactose, preventing recruitment of activator proteins by Gal4p and effectively repressing gene expression. In the presence of inducer, Gal3p becomes activated and gains affinity for Gal80p, thereby reducing the amount of Gal80p bound to Gal4p and permitting transcription from GAL promoter elements. Gal3p is an exclusively cytoplasmic protein, whereas Gal80p continuously shuttles between the nucleus and cytoplasm and becomes sequestered in the cytoplasm when bound to activated Gal3p¹⁴. In the presence of glucose, the same genes are rapidly and fully repressed by multiple mechanisms; the intracellular galactose concentration is reduced via transcriptional repression and catabolite inactivation of Gal2p¹⁵, and the Mig1p repressor inhibits both the transcription and activity of Gal4p^{16, 17}. The inducer molecule galactose is transported across the cell membrane by both a facilitated diffusion mechanism and a galactose permease protein Gal2p, which has both a high-affinity and low-affinity galactose transport mechanism¹⁸. Galactose is utilized as a sugar source by the cell through an initial conversion step catalyzed by a galactokinase Gal1p¹⁹. The levels of Gal2p, Gal3p, and Gal80p are regulated by GAL promoters, thereby forming three nested feedback control loops (Fig. 2.1b)²⁰. A number of other structural and regulatory proteins are under the control of GAL promoters with either one or two UASs.

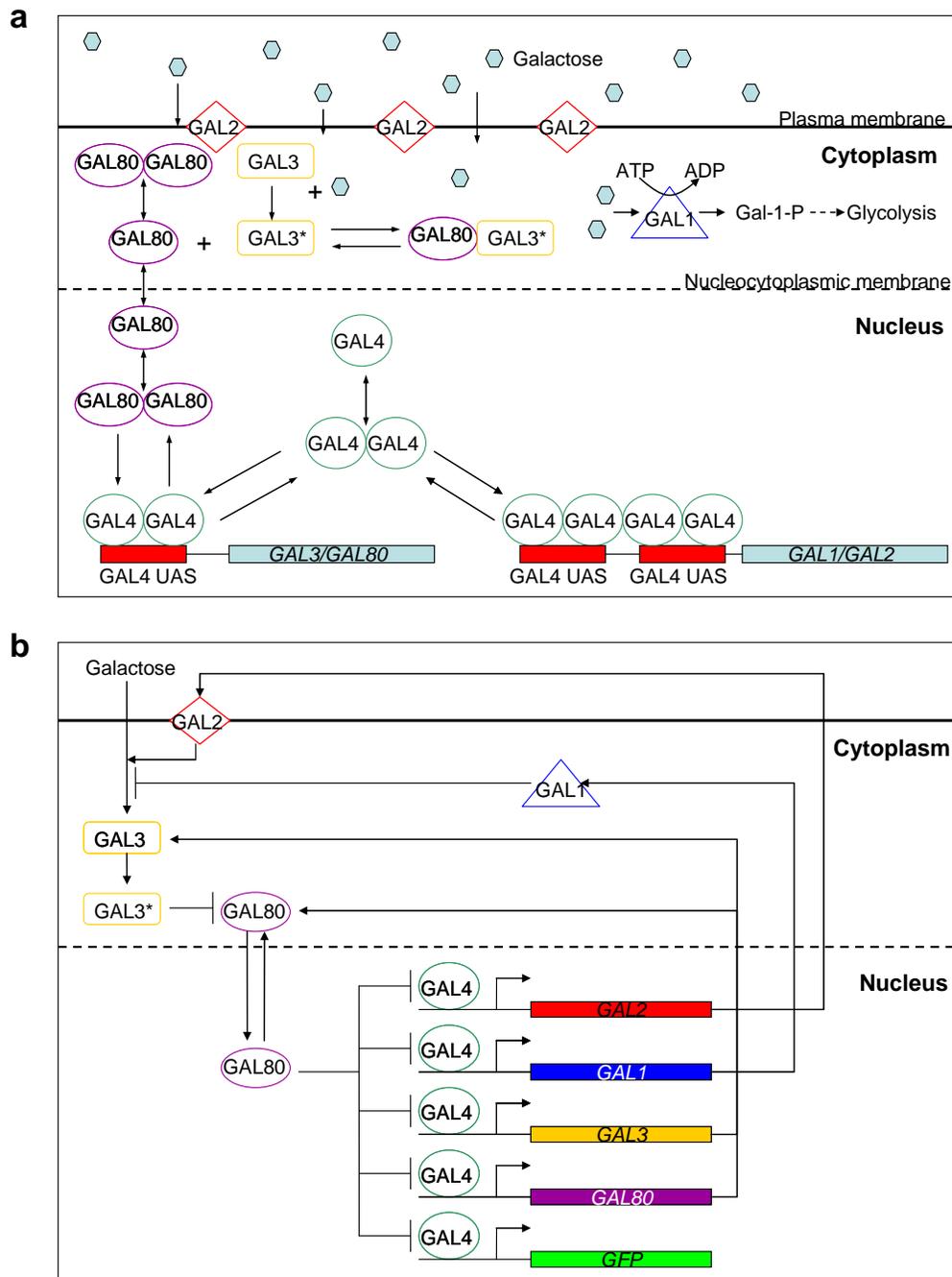


Fig. 2.1. Diagram of the native GAL network. **(a)** Schematic of the native GAL network in *S. cerevisiae*. **(b)** Schematic of the nested feedback control loops regulating the response of the GAL network in *S. cerevisiae*.

The nature of the autocatalytic response of the GAL genetic switch has been a topic of considerable research. Recent modeling work implicates the nucleocytoplasmic

shuttling of Gal80p and the feedback response of the regulatory proteins Gal3p and Gal80p as being critical to both the dynamic and steady-state performance of this system, and in particular the ultrasensitive response of the GAL induction curve^{12, 21}. Modeling has also indicated that the switch is only functional if Gal80p and Gal3p are subject to the same regulation¹². Prior work has demonstrated that the response properties of the system are highly sensitive to relative levels of Gal4p, Gal80p, and Gal3p²². The Gal2p galactose permease promoter region contains two UASs, whereas the promoter regions for Gal80p and Gal3p contain one UAS. Genes with multiple UASs are more tightly controlled by galactose, demonstrating lower basal expression and higher maximal induction; however, the effects of the permease feedback loop and transporter levels on the response of the network to varying galactose concentrations have not been examined.

Recent efforts have demonstrated that the response properties of inducible promoter systems can be altered by engineering interactions between components of the network. For example, several groups have altered the network connectivity of the arabinose metabolic network to exhibit a more tunable, homogenous response from arabinose-inducible promoter systems in *Escherichia coli* as opposed to its native, all-or-none response^{23, 24}. This tunable promoter system was designed by altering the regulation of the arabinose transporter gene from autocatalytic control to constitutive or researcher-controlled systems that resulted in a more linear induction response²³. In another recent study, the feedback loops of Gal3p and Gal80p were implicated in the memory response of the GAL network to growth history in *S. cerevisiae*²⁰. Although numerous factors certainly play critical roles in regulating the response behavior of the galactose network as described above, it is possible that the genetic switch response of the network may be

altered by removing the positive feedback control loop regulating galactose transport mediated by the Gal2p permease.

This work demonstrates that removing the positive feedback control loop regulating Gal2p expression is sufficient to alter the autocatalytic nature of this network such that the GAL promoter responds in a more linear manner to changes in galactose levels. While complete removal of the permease enables a population-averaged linear response from the GAL promoter, constitutive expression of the permease largely maintains the linear response and increases the overall magnitude of the response at a particular galactose concentration. Identical modifications to the promoter regions of the regulatory proteins Gal3p and Gal80p did not have the same effect, indicating that the feedback loop around Gal2p is unique in its ability to affect this linear versus switch-like response. The Gal2p-modified network also alters the population distribution to a more homogenous and gradual response at the single-cell level. In addition, deletion of the galactokinase Gal1p from this network has varying effects dependent on strain background and galactose concentrations due to its dual roles in catabolism and transport. At low galactose concentrations transport effects dominate such that the network response is more linear compared to the wild-type, whereas at higher galactose concentrations catabolic effects dominate such that the network response is amplified. Finally, our studies indicate that Gal1p may play a role in network stability as its removal results in the formation of multiple steady-state populations independent of strain background.

2.2. Results

2.2.1. Galactose permease deletion results in a linear induction response

Galactose is transported into the cell through both an induced high-affinity and low-affinity transport mechanism and an uninduced facilitated diffusion mechanism. The response of the GAL network was determined when the outermost positive feedback loop controlling the autocatalytic expression of the galactose permease Gal2p was removed. Initial studies examined the response of the network in the absence of the induced transport response. A *GAL2* deletion strain was constructed by inserting a kanamycin resistance marker into the *GAL2* locus of the chromosome. This system enabled the examination of the network response under conditions where the transport of galactose is limiting. Transcriptional activation, or the level of Gal4p not bound by Gal80p, in both the *gal2Δ* and the wild-type strain was determined by measuring fluorescence levels in cells harboring yEGFP under the control of a GAL1 promoter, which harbors two UASs. The data from these studies indicate that the steady-state population-averaged induction response is linear with respect to galactose in the *gal2Δ* strain across a wide range of inducer concentrations, whereas the wild-type strain demonstrates the expected autocatalytic response curve. As illustrated in Fig. 2.2a, both strains exhibit similar induction levels of approximately 25-fold over uninduced cells at the highest concentration of three percent galactose.

2.2.2. Constitutive expression of the galactose permease results in a tunable linear response

The complete removal of the induced transport mechanism eliminates the switch-like response of the GAL network. The response of the network in the presence of the inducible high- and low-affinity transport mechanisms removed from their feedback

regulation scheme was also determined. A constitutive *GAL2* strain, *tetO₂:GAL2*, was constructed to allow for tunable levels of Gal2p while removing the native positive feedback control loop. A cassette was constructed to replace the *GAL2* promoter with a tetracycline-repressible promoter²⁵. This cassette, which also contains the tTA transactivator and a kanamycin resistance gene, was inserted into the *GAL2* promoter region of the chromosome. Prior studies have indicated that in the absence of an appropriate tetracycline analog such as doxycycline, the expression levels from this promoter are approximately 10-20% of those observed from a fully induced *GAL1* promoter^{26, 27}. Similar steady-state assays of transcriptional activation in these strains were performed under varying concentrations of galactose and doxycycline. The former is expected to modulate the response of the GAL network in the presence of a constant level of galactose transporter, whereas the latter is expected to modulate the level of the galactose transporter. In the absence of doxycycline, permitting high Gal2p expression, the resulting induction curve is shifted upward compared to the *gal2Δ* strain but largely retains linearity (Fig. 2.2b). Addition of varying concentrations of doxycycline shifts the response curve to lower response levels, while at concentrations of 5 $\mu\text{g ml}^{-1}$ Gal2p expression is fully repressed and demonstrates a response identical to that of the *gal2Δ* strain. In addition, the maximum induction level observed in the *tetO₂:GAL2* strain is significantly greater than that observed in the wild-type strain. It should be noted that at high Gal2p expression levels and low galactose concentrations the response of the system is slightly nonlinear.

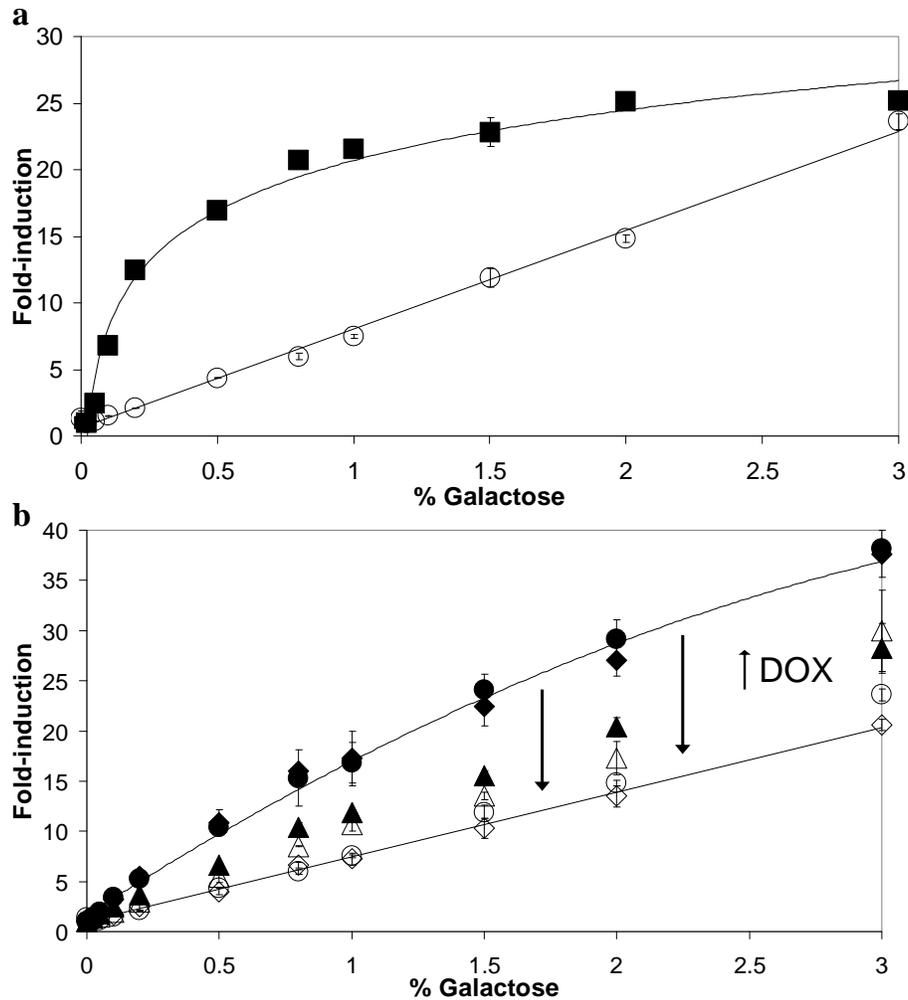


Fig. 2.2. Population-averaged response from strains with altered *Gal2p* regulation. **(a)** Population-averaged response of the *Gal2p* deletion strain (*gal2Δ*) (open circles) and the wild-type strain (filled squares). **(b)** Population-averaged response of the constitutive *Gal2p* strain (*tetO₂:GAL2*) at various concentrations of doxycycline (DOX). Levels of *Gal2p* decrease with increasing concentrations of doxycycline with full repression at concentrations over 1 mg ml⁻¹. Filled diamonds, no doxycycline; filled circles, 5 ng ml⁻¹; filled triangles, 25 ng ml⁻¹; open triangles, 50 ng ml⁻¹; open diamonds, 5 μg ml⁻¹; and open circles, *gal2Δ* strain.

2.2.3. Constitutive expression of regulatory proteins enhances the switch-like response of the network

The GAL network is regulated by three nested feedback control loops. The *Gal2p* loop is the exterior feedback loop and the presented data indicate that removal of this

loop is sufficient for modulating the sharp native response of this network to a linear response. The effects of the two interior nested loops regulating the expression of Gal80p and Gal3p on the steady-state population-averaged response of the GAL network were determined. Separate constitutive Gal80p and Gal3p strains were constructed by replacing the GAL80 and GAL3 promoters with previously described tetracycline-repressible promoter cassettes harboring the *his5*⁺ and kanamycin selection markers, respectively. In addition, a combined constitutive Gal80p/Gal3p strain was constructed by sequential insertion of these cassettes into the wild-type strain. These systems enabled the examination of the response of the GAL network under conditions where the two internally nested control loops regulating the transcriptional repressor and activator were individually and combinatorially removed. Similar steady-state population-averaged assays of transcriptional activation in these strains were performed under varying concentrations of galactose and doxycycline.

Constitutive strains for either regulatory protein Gal3p or Gal80p did not produce the same linear response observed from the constitutive Gal2p strain. The *tetO*₂:*GAL3* strain exhibited a steeper response curve under nonrepressed conditions (Fig. 2.3a). In addition, the repressed response curve demonstrated a memory response such that when doxycycline and galactose were added at the same time point the response was similar to that under the nonrepressed conditions, whereas when cells were grown in the presence of doxycycline prior to galactose addition the overall response curve was much lower. The *tetO*₂:*GAL80* strain also exhibited a steeper response curve under nonrepressed conditions (Fig. 2.3b). However, the addition of doxycycline either prior to or at the same time as the addition of galactose did not significantly alter the induction response. In

addition, the induction response from the Gal3p/Gal80p constitutive strain was much lower than any of the other strains (Fig. 2.3c). In this strain a history-dependent response was also observed in the repressed response curve such that slightly higher induction levels were observed when doxycycline and galactose were added at the same time point versus when the cells were grown in doxycycline prior to induction.

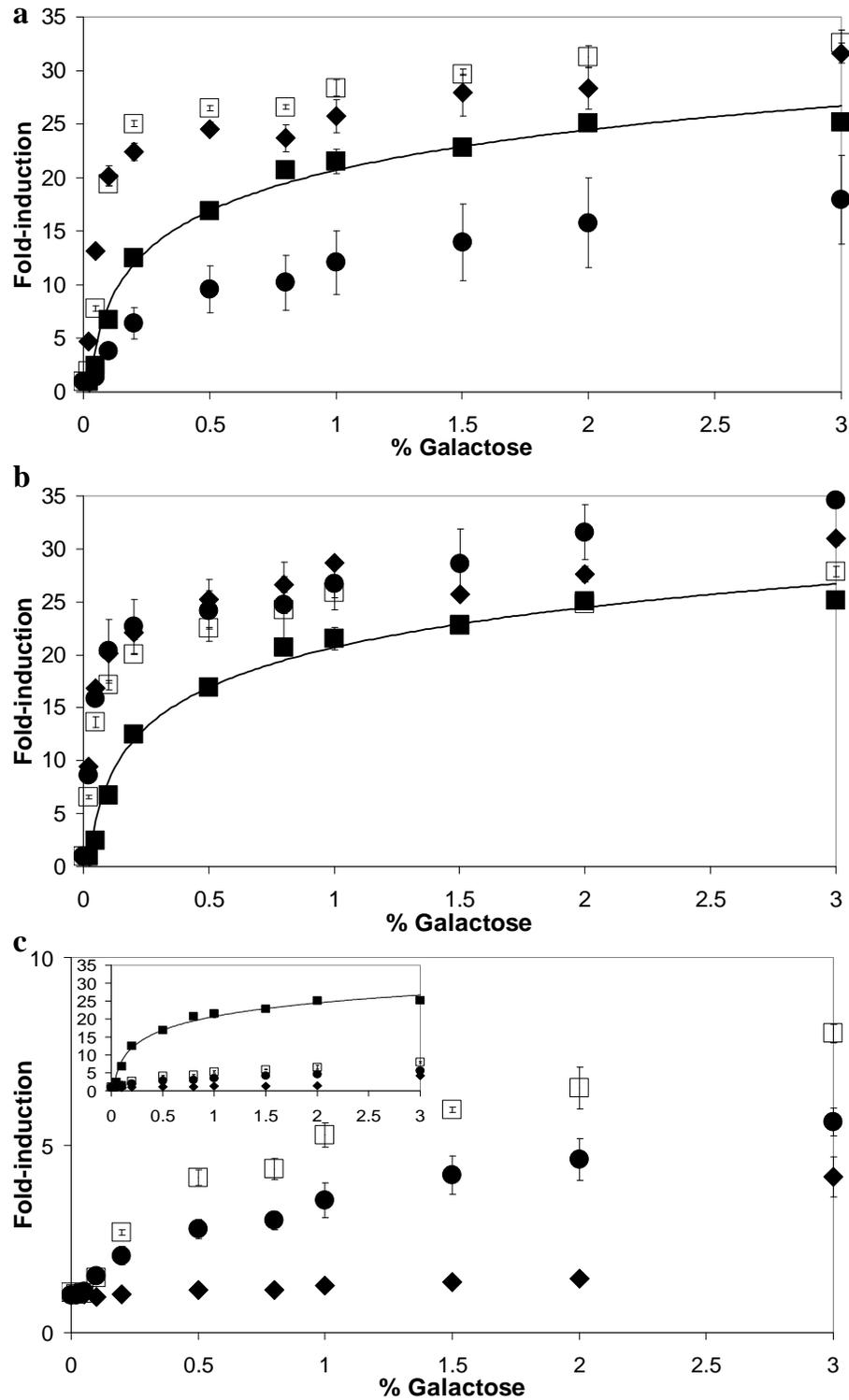


Fig. 2.3. Population-averaged response from strains with altered Gal3p and Gal80p regulation. **(a)** Population-averaged response of the wild-type strain (filled squares) and constitutive Gal3p strains (*tetO₂:GAL3*) at nonrepressed conditions (filled diamonds, 0 $\mu\text{g ml}^{-1}$ doxycycline), fully repressed conditions (open squares, 5 $\mu\text{g ml}^{-1}$ doxycycline), and

fully repressed conditions grown overnight in doxycycline (filled circles, 5 $\mu\text{g ml}^{-1}$ doxycycline). **(b)** Population-averaged response of the wild-type strain (filled squares) and constitutive Gal80p strain (*tetO₂:GAL80*) at nonrepressed conditions (filled diamonds), repressed conditions (open squares), and repressed conditions grown overnight in doxycycline (filled circles). **(c)** Population-averaged response of the constitutive Gal3p, Gal80p strain (*tetO₂:GAL3 tetO₂:GAL80*) at nonrepressed conditions (filled diamonds), repressed conditions (open squares), and repressed conditions grown overnight in doxycycline (filled circles). The inset illustrates induction levels relative to the wild-type strain (filled squares).

2.2.4. Galactokinase deletion results in a regimed network response

The data support that the nested positive and negative feedback loops in the GAL network influence the observed steady-state induction response to varying levels of galactose. The galactokinase Gal1p is also anticipated to play a key regulatory role in the response of the network as a result of its two distinct activities. The immediate role of this enzyme is in converting galactose into an energy source for the cell. Therefore, it is anticipated that removal of this activity will increase the overall response of the network at a given galactose concentration as the intracellular levels of galactose available for activating Gal3p will be effectively higher. Prior work has demonstrated higher fully induced response levels in a Gal1p knockout strain²⁸. However, the galactokinase also plays a key role in the high-affinity transport mechanisms associated with Gal2p¹⁸. To examine the role of the galactokinase on the response of the GAL network, Gal1p deletion strains were constructed in the three different Gal2p regulatory strains: wild-type, *gal2 Δ* , and *tetO₂:GAL2*. These strains were constructed by inserting a *His3MX6* selection marker into the *GALI* locus of the chromosome. These systems enable examination of the effects of the galactokinase in the response of the system under conditions where the normal Gal2p feedback control is present, Gal2p is present but the feedback control loop is removed, and in the absence of Gal2p. Similar steady-state

population-averaged assays of transcriptional activation in these strains were performed under varying concentrations of galactose.

The effects of deleting Gal1p were dependent on strain background and galactose concentration. For instance, in both the *gal1* Δ and the *gal1* Δ *tetO*₂:*GAL2* strains the induction response exhibited a more linear response in comparison to their respective *GAL1* isogenic strains (Fig. 2.4a,b). The response can be broken up into two different regimes: the low galactose regime, where the Gal1p deletion strains exhibit a lower induction response relative to their isogenic strains, and the high galactose regime, where the Gal1p deletion strains exhibit a higher induction response that increases linearly with galactose concentration relative to their isogenic strains. In the absence of Gal2p the deletion of Gal1p results in a different induction pattern (Fig. 2.4c). The response of the *gal1* Δ *gal2* Δ strain exhibits only one regime across all galactose concentrations, where the response curve maintains its linear response and is shifted upward from its isogenic strain across all galactose concentrations.

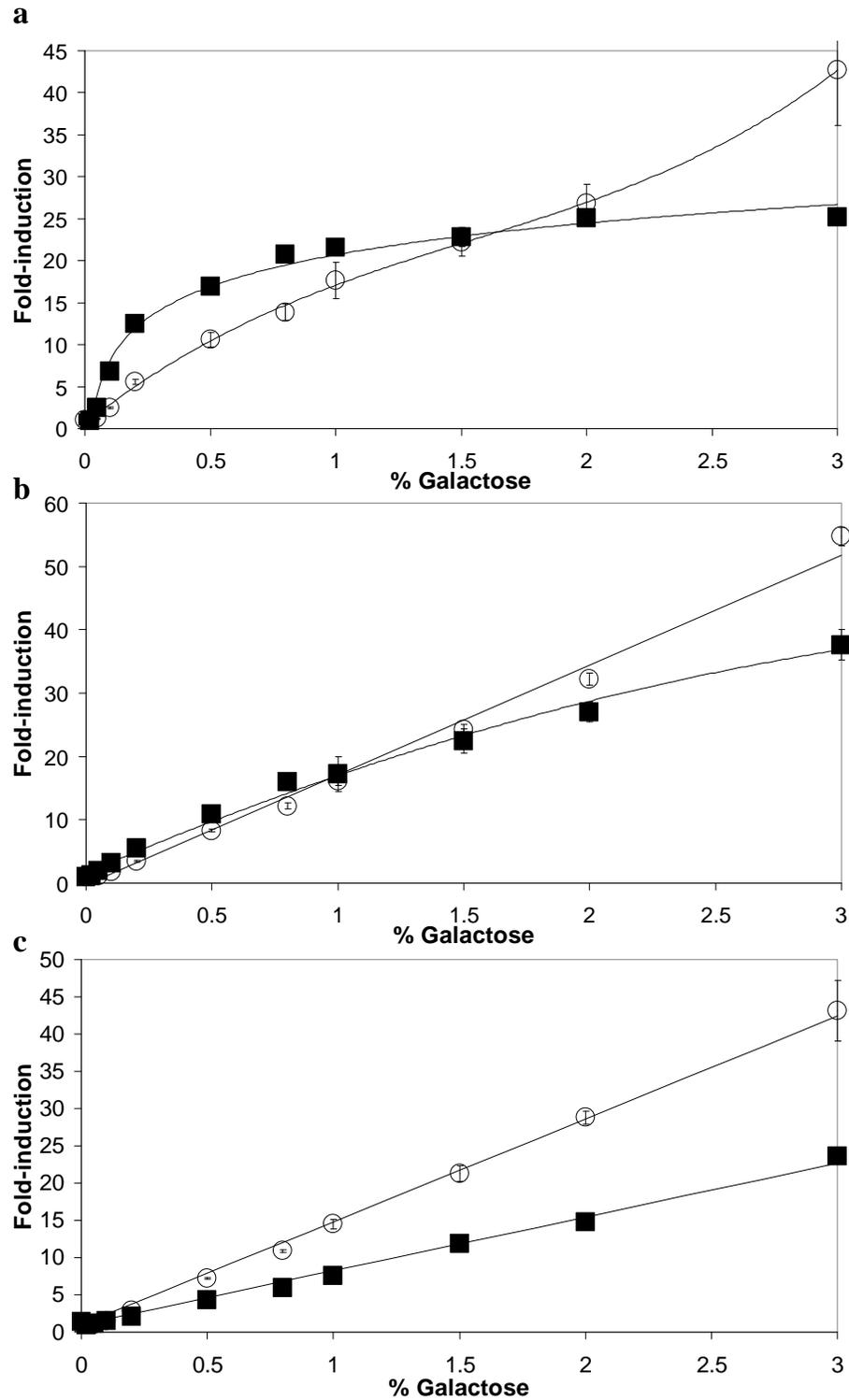


Fig. 2.4. Population-averaged response from strains with no Gal1p activity. **(a)** Population-averaged response of the Gal1p deletion strain (*gal1Δ*) (open circles) and the wild-type strain (filled squares). **(b)** Population-averaged response of the Gal1p deletion, constitutive Gal2p strain (*gal1Δ tetO₂:GAL2*) (open circles), and the corresponding

isogenic strain (*tetO₂:GAL2*) (filled squares). (c) Population-averaged response of the Gal1p, Gal2p deletion strain (*gal1Δ gal2Δ*) (open circles) and the corresponding isogenic strain (*gal2Δ*) (filled squares).

2.2.5. Population distributions in GAL2-modified strains exhibit graded responses

Alteration of the regulatory schemes at various control points modifies the steady-state population-averaged response of the GAL network. The effects of these targeted alterations on the population response of the network were determined. Flow cytometry was used to analyze the response of the cell population to alterations in Gal2p regulation. Wild-type, *gal2Δ*, and *tetO₂:GAL2* cells were cultured under the same conditions as the population-averaged studies prior to preparation for analysis. The wild-type strain exhibited two distinct populations of fully induced and uninduced cells (Fig. 2.5a). In accordance with the all-or-none effects observed in other inducible promoter systems¹³, the percentage of fully induced cells increases with increasing galactose concentrations. While both *GAL2*-modified strains, *gal2Δ* and *tetO₂:GAL2*, exhibited a significant uninduced or negative population, they did not exhibit the all-or-none effect observed with the wild-type strain. Specifically, the average level of GFP expression from the induced population and the number of cells that were induced increased with galactose concentration (Fig. 2.5b,c). This graded response was most dramatic in the *gal2Δ* strain. The *tetO₂:GAL2* strain also demonstrated a slightly graded response to galactose with a higher mean fluorescence at all concentrations compared to the *gal2Δ* strain consistent with the population-averaged data (Fig. 2.5d). The three strains exhibited similar population distributions in the fully induced state, or at high galactose concentrations, with a majority of the population expressing the maximum level of GFP. Slight

differences in the negative populations between the *tetO₂:GAL2* strain and the wild-type account for the differences in maximum induction at the population-averaged level.

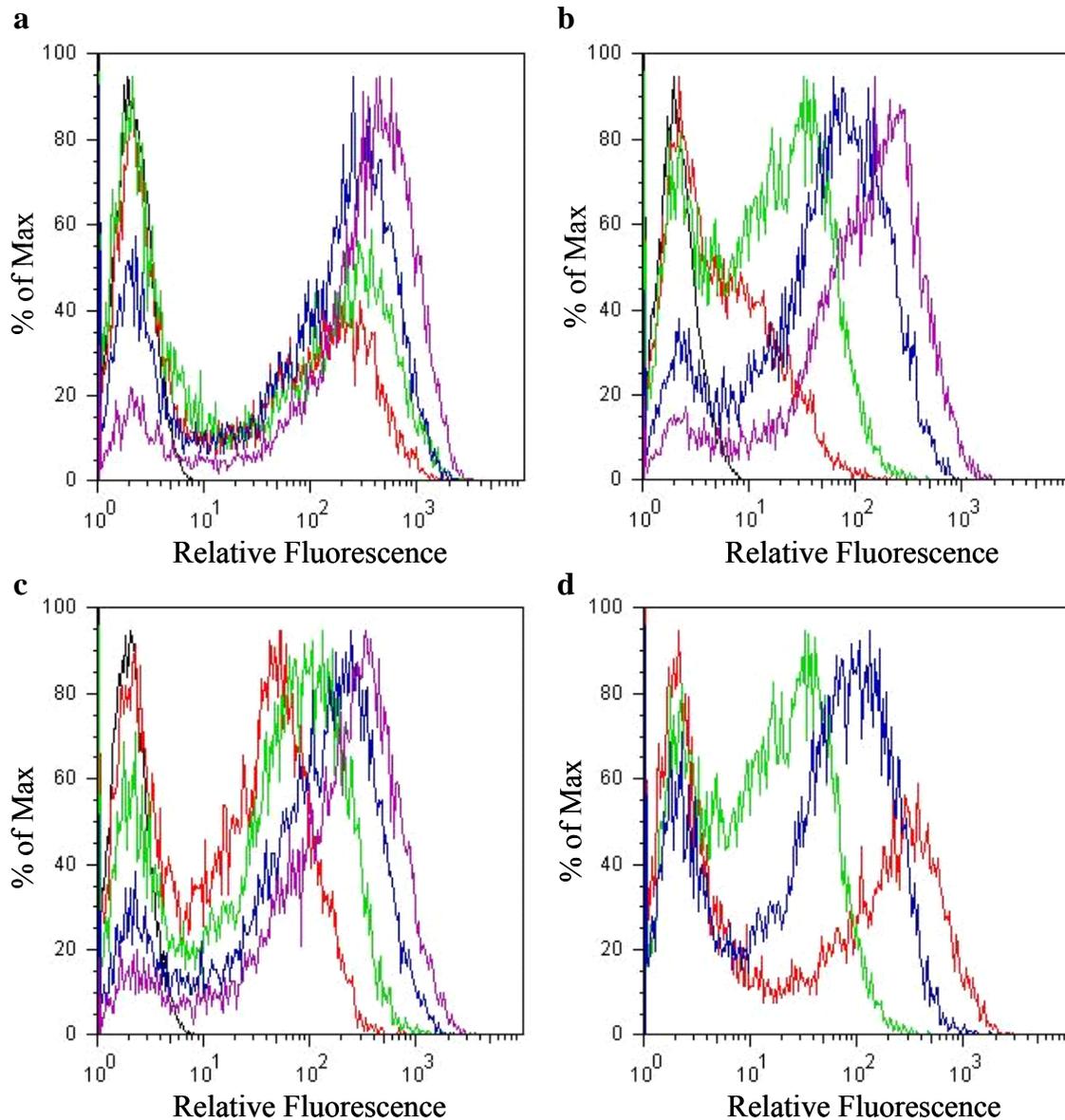


Fig. 2.5. Population response from strains with altered Gal2p regulation. For (a), (b), and (c), galactose concentration is indicated as: black (0%), red (0.2%), green (0.5%), blue (1%) and purple (3%). **(a)** Population distribution of cells with the native Gal2p positive feedback control loop (wild type) across various concentrations of galactose. **(b)** Population distribution of cells with constitutive Gal2p levels (*tetO₂:GAL2*) under nonrepressed conditions ($0 \mu\text{g ml}^{-1}$ doxycycline) across various concentrations of galactose. **(c)** Population distribution of cells with no Gal2p (*gal2Δ*) across various concentrations of galactose. **(d)** Population distributions in 0.5% galactose of wild-type cells exhibiting feedback Gal2p control (red), *tetO₂:GAL2* cells exhibiting constitutive

Gal2p expression (blue), and *gal2Δ* cells in which Gal2p is absent from the network (green).

2.2.6. Deletion of the galactokinase results in multiple stable populations

Studies support the regulated effects of the galactokinase Gal1p on the steady-state population-averaged response of the GAL network as a result of its role in the high-affinity Gal2p transport mechanism. The effects of the removal of Gal1p in a variety of Gal2p regulatory backgrounds on the population response of the network were determined. Flow cytometry analysis was conducted to determine the population response in the absence of Gal1p. *Gal1Δ*, *gal1Δ gal2Δ*, and *gal1Δ tetO₂:GAL2* cells were cultured under the same conditions as the population-averaged studies prior to preparation for analysis. The population data matches the general trends observed in the population-averaged data across different concentrations of galactose (Fig. 2.6). Interestingly, all of the Gal1p deletion strains, regardless of background, exhibited multiple, distinct cell populations across all ranges of galactose concentration measured between 0.05 and 3%. In contrast to the all-or-none response of the wild-type strain, these populations allow intermediate levels of gene expression in all *gal1Δ* strains.

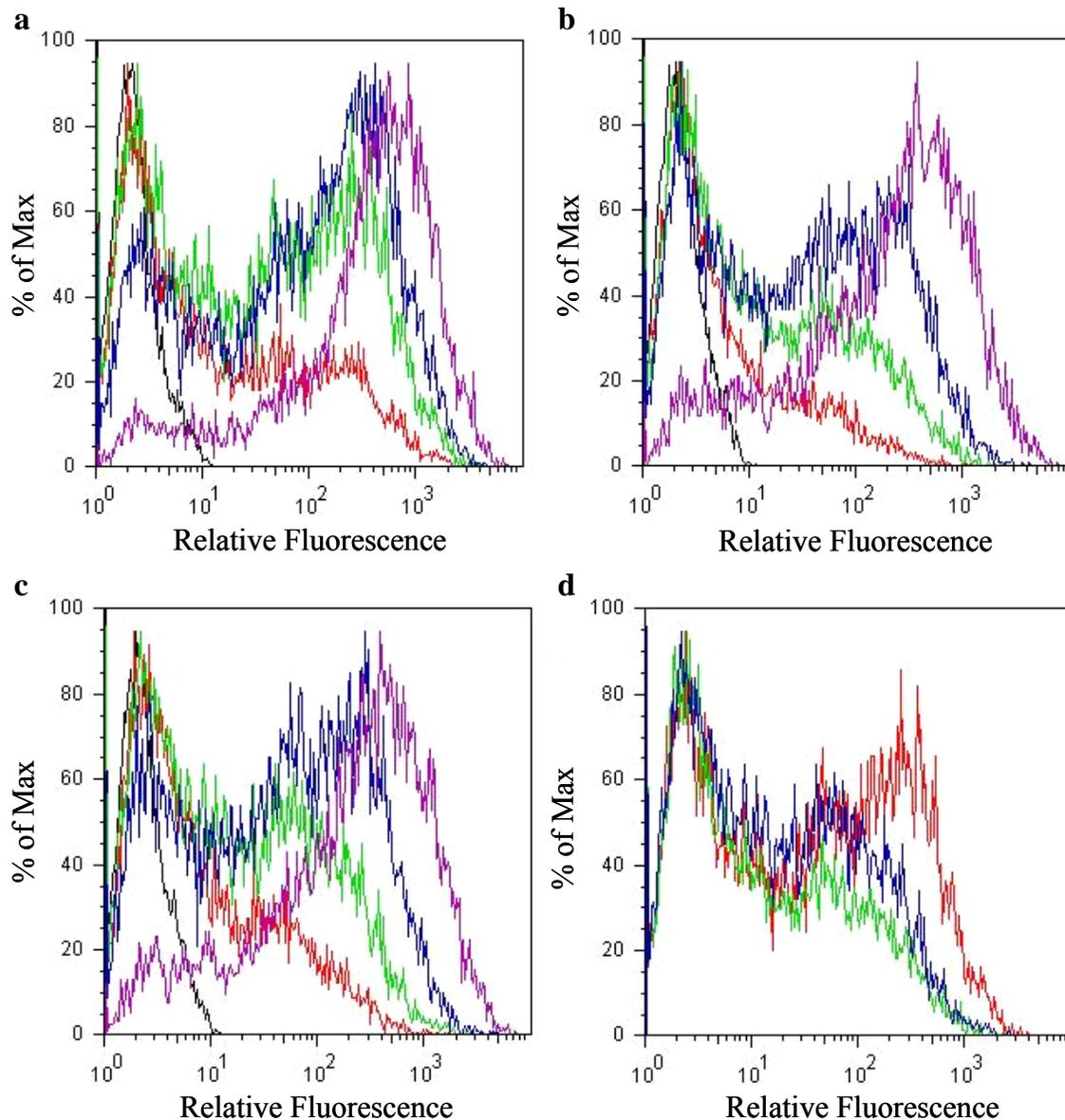


Fig. 2.6. Population response from strains with no Gal1p activity. For (a), (b), and (c), galactose concentration is indicated as: black (0%), red (0.2%), green (0.5%), blue (1%) and purple (3%). **(a)** Population distribution of cells with the native Gal2p positive feedback control loop and no Gal1p (*gal1Δ*) across various concentrations of galactose. **(b)** Population distribution of cells with constitutive Gal2p expression and no Gal1p (*gal1Δ tetO₂:GAL2*) under nonrepressed conditions ($0 \mu\text{g ml}^{-1}$ doxycycline) across various concentrations of galactose. **(c)** Population distribution of cells with no Gal2p and no Gal1p (*gal1Δ gal2Δ*) across various concentrations of galactose. **(d)** Population distributions in 0.5% galactose of *gal1Δ* cells exhibiting feedback Gal2p control and no Gal1p (red), *gal1Δ tetO₂:GAL2* cells exhibiting constitutive Gal2p expression and no Gal1p (blue), and *gal1Δ gal2Δ* cells exhibiting neither Gal2p nor Gal1p (green).

2.3. Discussion

The data from the population-averaged transcriptional activation assays demonstrate that the positive feedback loop regulating the expression of Gal2p is necessary for the sharp, autocatalytic response of the system to galactose observed in the wild-type strain. In the *gal2* Δ strain, galactose and its nonphosphorylatable analogs are transported solely by an uninduced facilitated diffusion mechanism¹⁸. We propose that under these conditions transport effects limit the intracellular galactose concentration and the ensuing network response. Specifically, there will be fewer molecules of galactose to activate Gal3p such that more Gal80p remains bound to Gal4p and therefore a decrease in the transcriptional activation response is observed. This is in contrast to the wild-type environment, where the amount of galactose in the cells increases sharply over a narrow concentration range once galactose gets into the cells as a result of the positive feedback loop regulating Gal2p.

Furthermore, we demonstrate that modulating the levels of the galactose transporter in the absence of its positive feedback control loop is an effective way of tuning the linear response of the system. In the *tetO₂:GAL2* strain, galactose is transported by a constitutive high-affinity and low-affinity transport mechanism in addition to the facilitated diffusion mechanism. While removal of the positive feedback loop eliminates much of the switch-like response of the system, under conditions of high transporter levels (low doxycycline levels) and low galactose levels the system does exhibit a slightly nonlinear response. This data indicate that at low galactose concentrations the high-affinity transport mechanism is dominant and inducer transport is not a limiting

factor in GAL promoter activation. In addition, higher maximum induction levels are observed in the *tetO₂:GAL2* strain, most likely due to the removal of the negative feedback loop on the regulation of Gal2p from increased levels of Gal80p. Furthermore, under conditions of full tetracycline repression of Gal2p, the induction curve mimics that of the *gal2Δ* strain and supports that the observed shifts in the system response with doxycycline are due solely to a corresponding change in Gal2p levels.

The data from the flow cytometry assays demonstrate that alterations in Gal2p regulation also changed the population response of the GAL network. Specifically, the positive feedback control loop regulating the expression of the galactose permease is a necessary component of the observed all-or-none response of this network. This has been demonstrated in simpler bacterial networks such as the arabinose and lactose operons²³. The results demonstrate a significant negative population irrespective of Gal2p regulation except at high galactose concentrations. The persistence of the negative population is likely due, in part, to the recently described cellular memory of this network²⁰, as cells were grown on noninducing-nonrepressing media prior to induction. Previous work has demonstrated that growing initial cell cultures in the presence of galactose will reduce, but not eliminate, this negative population²⁰.

Similar studies with the regulatory proteins indicate that the feedback loops regulating Gal3p and Gal80p levels are not necessary for the autocatalytic induction response. The *tetO₂* promoter is weaker than the native GAL promoters, yet fully induced response levels comparable to the native promoter systems are still attained when Gal3p or Gal80p are individually controlled in this manner. Reducing the levels of the repressor protein Gal80p has the anticipated effect of a higher response and a lower galactose

requirement for full induction. However, the response of the network was similar under both repressed and nonrepressed conditions in this strain (*tetO₂:GAL80*), indicating that the relative levels over which the tetO₂ promoter can regulate Gal80p expression is not sufficient for tuning the network response. Reducing the levels of the activator protein Gal3p had the unanticipated effect of also increasing the response of the network and lowering the galactose level at which full induction is observed. The sharper response curve observed under constitutive Gal3p regulation versus feedback regulation may be explained by the higher concentrations of the activator protein potentially present at lower concentrations of galactose in the constitutive strain background. In addition, unlike the *tetO₂:GAL80* strain the response of the system in the *tetO₂:GAL3* strain was highly dependent on the concentration of Gal3p at the time of induction and indicated that the relative level over which the tetO₂ promoter can regulate Gal3p expression is sufficient for tuning the network response. This difference in observed tunability may be explained by differences in the relative levels of these two regulator proteins, as Gal3p has been estimated to be at a 5-fold higher concentration than Gal80p at full induction conditions¹⁴. Furthermore, unlike the constitutive *GAL2* strain, the behavior of the complete knockouts is not replicated under conditions of full repression indicating that low levels of Gal3p and Gal80p are sufficient to maintain switch functionality. In a Gal3p knockout strain the network is not inducible with the exception of long-term adaptation occurring after several days²⁹. In a Gal80p knockout strain the Gal4p activation domain is not repressed and the population remains fully induced independent of galactose³⁰. The data indicate that basal expression from the tetO₂ promoter produces sufficient Gal3p to

activate the switch even at low inducer concentrations and enough Gal80p to fully repress Gal4p in the absence of inducer.

The response of the network under constant and equal levels of both regulatory proteins was unexpected. Previous modeling work had predicted that the switch response of the network would be unaffected if Gal80p and Gal3p were not autoregulated²². The results from these studies indicate that under equal levels of the activator and repressor proteins expressed from the tetO₂ promoter under nonrepressing conditions the galactose network is not inducible. However, when levels of these regulatory proteins are both lowered under repressed conditions the network exhibits low levels of induction that depend on the concentrations of Gal3p and Gal80p at the time of induction. The low induction levels observed in this strain may indicate the sensitivity of this network to the ratio of Gal3p/Gal80p levels, and in particular lowering this ratio to one. Finally, the observed memory response in this strain supports the sensitivity of the system to starting levels of Gal3p, attaining higher induction levels when Gal3p is present at the time of galactose addition.

These studies indicate that the feedback loops controlling the levels of these two regulatory proteins may not be essential to the native switch-like response of the GAL network. In the case of Gal3p, it has been suggested that the feedback loop is a remnant of the evolution of this regulatory protein. This signal transduction molecule is the result of paralogous evolution from Gal1p³¹ and effectively separates galactose sensing and metabolism as it has lost its galactokinase activity³². However, it is currently not clear why Gal80p evolved the same type of autoregulation mechanism if it is not necessary for

maintaining the response of the network, other than to prevent overexpression at high galactose concentrations.

The complex response properties observed in the *gal1Δ* strains are proposed to be a result of the competing roles of Gal1p in catabolism and transport in the GAL network. Removal of the Gal1p catabolic activity increases the effective concentration of galactose in the cell, which would be expected to increase the response of the network at all galactose concentrations. However, removal of the Gal1p transport activity, which effectively removes the high-affinity Gal2p transport mechanism, would be expected to decrease the response of the network particularly at low galactose concentrations, where this transport mechanism is essential to efficient galactose transport. This dual role model is supported by the data from the population-averaged transcriptional activation assays. In the absence of Gal2p the transport role of Gal1p is no longer a factor in the pathway and therefore the shifted response is solely due to the absence of galactose catabolism (*gal1Δ gal2Δ* versus *gal2Δ*). However, when Gal2p is present either at constitutive levels or under feedback regulated control both the transport and catabolic roles of Gal1p influence the response of the system. At low inducer concentrations the transport effects of Gal1p are dominant in the response of the system (observed as lower induction levels from the *gal1Δ* strains), whereas at high inducer concentrations the catabolic effects of Gal1p are dominant in the response of the system (observed as higher induction levels).

The data from the flow cytometry assays indicate that Gal1p also plays a role in effecting the population response of the GAL network. The data demonstrate the occurrence of multiple cell populations in all galactokinase deletion strains regardless of the regulation of the galactose permease or even its presence. This supports that the

occurrence of the multiple populations is due to the loss of the galactokinase function of Gal1p and not due to the loss of the high-affinity transport mechanism mediated by Gal2p. To our knowledge, the removal of a network kinase has not been demonstrated to result in the occurrence of multiple, steady-state cell populations in other networks. Multiple cell populations are often associated with different steady-state or stability regimes. It is possible that Gal1p, either through its galactokinase activity or some as yet unidentified activity, plays a role in stabilizing the population response. A recent structural study comparing Gal1p and Gal3p suggests that the galactokinase can function as a transcriptional activator³³, and the loss of this activity may potentially contribute to the emergence of multiple steady-states in the absence of Gal1p.

In summary, this work demonstrates that the removal of key regulatory loops alters the steady-state and population response of the galactose metabolic network in *S. cerevisiae*. The feedback loop regulating the Gal2p permease is critical to the observed autocatalytic induction response and all-or-none response of the system. The permease also presents a suitable control point at which titrating levels of this protein with available tools enables tuning of the network response with the *GAL2* deletion strain exhibiting a linear response between 0 and 3 % galactose. The feedback loops regulating the activator and repressor proteins are not necessary for the autocatalytic induction response and are not suitable control points for tuning the response of the system with the promoter system used in this work. In addition to further elucidating the roles of the various regulatory loops in the response of this network, this work also presents a number of engineered networks that will be useful as tunable, homogenous promoter systems in *S. cerevisiae*.

2.4. Materials and Methods

2.4.1. Yeast strain construction

The wild-type haploid yeast strain used in this study is W303 α (MAT α *his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*). All other strains were constructed by making modifications to the chromosome of this wild-type strain through standard homologous recombination procedures³⁴. For each strain an insertion cassette was constructed with the appropriate insertion sequences and regions of homology to the desired targeted sites on the chromosome. A cassette harboring an *E. coli* kanamycin resistance gene and associated promoter and terminator elements with ends homologous to regions flanking *GAL2* on the chromosome was constructed by amplifying the appropriate segment from pFA6a-ZZ-TEV-S-kanMX6³⁵. A second cassette harboring the tetO₂ response element and minimal CYC1 promoter, the tTA transactivator and associated promoter and terminator elements, and the kanamycin resistance gene and associated promoter and terminator elements with ends homologous to regions flanking the *GAL2* promoter was constructed in two steps (Fig. 2.7a). In the first step the kanamycin resistance cassette was amplified from pFA6a-ZZ-TEV-S-kanMX6 and the tetracycline-regulatable promoter cassette was amplified from pCM188²⁷ separately. In a second round PCR step, these two cassettes were combined to form one cassette by overlap extension techniques³⁴. A third cassette harboring a *Schizosaccharomyces pombe* histidine biosynthetic gene (*his5*⁺) and associated promoter and terminator elements with ends homologous to regions flanking *GAL1* on the chromosome was constructed by amplifying the appropriate segment from pFA6-S-TEV-ZZ-HIS3MX6³⁵. Analogous cassettes with regions flanking the *GAL3* and *GAL80* promoters were also constructed (Fig. 2.7b).

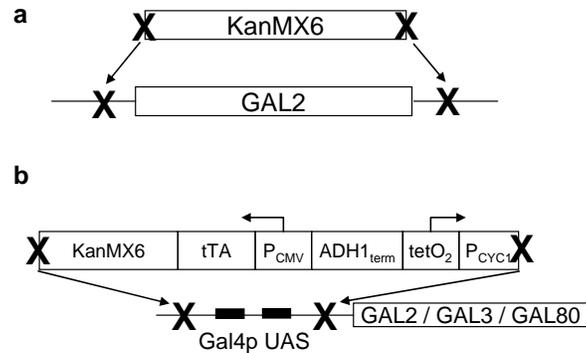


Fig. 2.7. Cassettes for chromosomal replacements. **(a)** Schematic of the *GAL2* knockout constructed by replacement of the entire coding region with the KanMX6 cassette. **(b)** General schematic of the constructs used to replace the native promoter of *GAL2*, *GAL3*, and *GAL80* with tetracycline-repressible promoters.

The individual fragments for the *GAL2*, *GAL3*, and *GAL80* promoter substitution cassettes were amplified using the TripleMaster PCR System (Eppendorf). All other cassettes were constructed with standard PCR procedures in a Dyad PCR machine (MJ Research) with Taq DNA polymerase (Roche). Oligonucleotide primers were purchased from Integrated DNA Technologies, and primer sequences are available upon request. Cassettes were transformed into the appropriate strains using a standard lithium acetate procedure³⁶. The *GAL2* knockout and *GAL2*, *GAL3*, and *GAL80* tetracycline-regulatable expression cassettes were inserted into the wild-type strain. The *GAL1* knockout cassette was inserted into the wild-type strain, the *GAL2* knockout strain, and the *GAL2* tetracycline-regulatable strain. Strains with inserted cassettes were selected by growth on synthetic complete media with the appropriate antibiotic selection and dropout media. Confirmation of cassette insertion into the correct chromosomal location was conducted by PCR amplification of the targeted region of the chromosome. A list of yeast strains and primer sequences used in this work are provided in Table 2.1 and Table 2.2.

Table 2.1. List of yeast strains.

Strain Number	Genotype	Plasmid
CSY3	<i>MATα his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1</i>	
CSY22	<i>gal2Δ</i>	
CSY13	<i>Gal2p::KanMX6-tTA-tetO₂</i>	
CSY50	wild-type	pGAL-GFP
CSY52	<i>gal2Δ</i>	pGAL-GFP
CSY40	<i>Gal2p::KanMX6-tTA-tetO₂</i>	pGAL-GFP
CSY53	<i>gal1Δ</i>	pGAL-GFP
CSY54	<i>gal1Δ gal2Δ</i>	pGAL-GFP
CSY55	<i>gal1Δ Gal2p::KanMX6-tTA-tetO₂</i>	pGAL-GFP
CSY89	<i>Gal3p::KanMX6-tTA-tetO₂</i>	pGAL-GFP
CSY90	<i>Gal80p::His3MX6-tTA-tetO₂</i>	pGAL-GFP
CSY91	<i>Gal3p::KanMX6-tTA-tetO₂ Gal80p::His3MX6-tTA-tetO₂</i>	pGAL-GFP

* All strains are derivatives of CSY3; only modifications to the wild-type background are indicated.

Table 2.2. List of primer sequences used in the construction of plasmids and yeast strains.

Name	Sequence	Description
GAL2kan.fwd	AGAATAGTAATAGTTAAGTAAACACAAGATTAACATAATAGGCAGATCCGCTAGGGATAA	Fwd primer for replacing GAL2 gene with KanMX
GAL2kan.rev	CATGAAAAATTAAGAGAGATGATGGAGCGTCTCACTTCAAGAATTCGAGCTCGTTTAAAC	Rev primer for replacing GAL2 gene with KanMX
tTA-tetO.fwd	CCTTGACTGCAATACGGCATCTACCCACCGTACTCGTCAA	Fwd primer for amplifying tTA and tetO ₂ for cassette construction
GAL2tetO.rev	TTACTATTCTTGATGATAATTGAATAAGGTGCATAATGAACCCCGAATTGATCCGGTAA	Rev primer for amplifying tTA and tetO ₂ with 40bp homology to GAL2 promoter region
KanR.fwd	ATGCCGATTGCAGTCAAGGGCCAGATCCGCTAGGGATAA	Fwd primer for KanMX portion of promoter replacement cassettes
GAL2p.kan.rev	ATATTCGAAAGGGCGGTTGCCTCAGGAAGGCACCGCGGGAATTCGAGCTCGTTTAAAC	Rev primer for KanMX portion of cassette with 40bp homology to GAL2 promoter region
GAL2p.screen.fwd	CATTAATTTTGCTTCCAAGACGACAGTAATATGTCTCCTA	Fwd primer for screening GAL2 promoter replacement
GAL2p.screen.rev	ATGGGAATCTTTACTGAGTGAAGAGATCACGTCTTC	Rev primer for screening GAL2 promoter replacement
GAL1HISko.fwd	GTGCGTCTCGTCTTACC CGTCCGTTCTCTGAAACGCAGGGCAGATCCGCTAGGGATAA	Fwd primer for amplifying HIS5 with 40bp homology to sequence upstream of GAL1
GAL1HISko.rev	CTACTCGTTATTATTGCGTATTTTGTGATGCTAAAGTTATGAATTCGAGCTCGTTTAAAC	Rev primer for amplifying HIS5 with 40bp homology to sequence downstream of GAL1
GAL1ko.screen.fwd	TGGAACCTTTCAGTAATACGCTTAACTGCTC	Fwd primer for screening of GAL1 knockout
GAL1p.fwd	TCATGAATCCCTGAATTTTCAAAAATCTTACTTTTTTTTTGG	Fwd primer for cloning of GAL1 promoter
GAL1p.rev	CATAGGATCCGTTTTTCTCCTTGAC	Rev primer for cloning of GAL1 promoter
yEGFP.fwd	GCGGATCCATGTCTAAAGGTGAAGAATTATCACTG	Fwd primer for cloning yEGFP-CLN2-PEST
yEGFP.rev	GCACGCGTTATATTACCCTGTTATCCCTAGCG	Rev primer for cloning yEGFP-CLN2-PEST
GAL3p.kan.rev	ATTAACACAGTGGTTTCTTTGCATAAACACCATCAGCCTGAATTCGAGCTCGTTTAAAC	Rev primer for KanMX portion of cassette with 40bp homology to GAL3 promoter region
GAL3tetO.rev	CTCTGACCGGAGAAGTGAATTTGGAACGTTTGTATTATCATCCCCGAATTGATCCGGTAA	Rev primer for amplifying tTA and tetO ₂ with 40bp homology to GAL3 promoter region
GAL3p.screen.fwd	CTTATTAACCGCTTTTACTATTATCTTCTACGCTGACAGT	Fwd primer for screening GAL3 promoter replacement
GAL3p.screen.rev	TGAATGTGCCACATGATAGTCCGATTTAAAGTAATTCGAC	Rev primer for screening GAL3 promoter replacement
GAL80p.his.fwd	TGACTGCCACTGGACCTGAAGACATGCAACAAAGTGAAGGAATTCGAGCTCGTTTAAAC	Rev primer for His3MX portion of cassette with 40bp homology to GAL80 promoter region
GAL80tetO.rev	GAAAGAACGGGAAACCAACTATCAGATTGTATACGCTGGCCCCGAATTGATCCGGTAA	Rev primer for amplifying tTA and tetO ₂ with 40bp homology to GAL80 promoter region
GAL80p.screen.fwd	GAACCTCCTCCAGATGGAATCCCTCCATA	Fwd primer for screening GAL80 promoter replacement
GAL80p.screen.rev	ATCTAGTGAACATGCAAGGGCCATTCTACGAAAAGATAC	Rev primer for screening GAL80 promoter replacement

2.4.2. Yeast expression plasmids

Standard molecular biology cloning techniques were used to construct the reporter plasmid used to assay Gal4p activation³⁴. The plasmid was generated by cloning into the pCM190²⁷ shuttle plasmid. This plasmid contains an *E. coli* origin of replication (f1) and selection marker for ampicillin resistance, as well as a *S. cerevisiae* 2 μ M high copy origin of replication and a selection marker for a uracil biosynthetic gene for plasmid maintenance in synthetic complete media supplemented with the appropriate amino acid

dropout solution. A yeast enhanced green fluorescent protein (yEGFP) gene with a degradation tag (CLN2-PEST)³⁷ and ADH1 terminator was inserted into the multi-cloning site of pCM190 behind the tetO₇ promoter between BamHI and MluI restriction sites. The GAL1 promoter was then cloned into this vector between EcoRI and BamHI restriction sites. The *yEGFP-CLN2-PEST* gene was amplified from pSVA15³⁷ using standard PCR procedures as previously described. The GAL1 promoter was amplified from pRS314-Gal³⁸. This promoter contains two UASs and has been used in previous studies to measure Gal4p activity levels³⁹.

The reporter plasmid was constructed using restriction endonucleases and T4 DNA ligase from New England Biolabs. Plasmids were screened by transformation into an electrocompetent *E. coli* strain, DH10B (Invitrogen; F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*dlacZ* Δ M15 Δ *lacX74 deoR recA1 endA1 araD139 Δ (*ara, leu*)7697 *galU galK* λ -*rpsL nupG*), using a Gene Pulser Xcell System (BioRAD) according to manufacturer's instructions. Subcloning was confirmed by restriction analysis. Confirmed plasmids were then transformed into the appropriate *S. cerevisiae* strains using a standard lithium acetate protocol³⁶. *E. coli* cells were grown on Luria-Bertani media (DIFCO) with 100 μ g ml⁻¹ ampicillin (EMD Chemicals) for plasmid maintenance and *S. cerevisiae* cells were grown on synthetic complete media (DIFCO) supplemented with the appropriate dropout solution (Calbiochem) for plasmid maintenance. Plasmid isolation was conducted using Perfectprep Plasmid Isolation Kits (Eppendorf) according to manufacturer's instructions.*

2.4.3. Fluorescence assays

Cell cultures were grown at 30°C in test tubes shaken at 200 rpm. Strains containing the reporter plasmid were grown in synthetic complete medium with the appropriate dropout solution (lacking uracil) and sugar source (2% raffinose, 1% sucrose). Overnight cultures were backdiluted 30-fold into fresh noninducing-nonrepressing media to an OD₆₀₀ between 0.05 and 0.1. For assaying the network response, this fresh media contained appropriate concentrations of galactose (DIFCO), doxycycline (Sigma) for tetracycline-regulatable *GAL2*, *GAL3*, and *GAL80* strains, or water (negative control). Fluorescence and OD₆₀₀ readings were measured using a Safire (TECAN) fluorescent plate reader after 8 h. Sample volumes of 200 μ L were aliquoted into 96-well flat-bottom black plates (Greiner). The excitation and emission wavelengths were set to 485 nm and 515 nm, respectively, with a bandwidth of 12 nm. Fluorescence was measured from the bottom of the plate with a gain setting of 100. Fluorescence was normalized for cell number by dividing relative fluorescence units (RFUs) by the OD₆₀₀ of the culture after subtracting the media background from each. All measurements were repeated at least in triplicate.

2.4.4. Flow cytometry assays

Yeast cells were grown according to methods detailed in fluorescence assays prior to preparation for flow cytometry analysis. After 7 h of induction, 5 ml of cells were harvested by centrifugation at 6000 rpm for 5 min, resuspended in 5 ml of phosphate buffered saline and incubated on ice for 30 min. This wash was repeated and the cell solution was subsequently filtered through a 40 μ M cell strainer (Falcon). Cells were

analyzed on a FACSCalibur instrument (Becton Dickinson; San Jose, CA) using a 15 mW Argon laser with a 400 nm excitation wavelength and a 488 nm emission wavelength. For each sample approximately 10,000 cells were analyzed and each sample was repeated in duplicate. Data from the population fluorescence was analyzed using FlowJo software (Tree Star, Inc).

**CHAPTER III. SYNTHESIS OF THE BENZYLISOQUINOLINE
ALKALOID BACKBONE FROM TYROSINE IN *SACCHAROMYCES
CEREVISIAE***

Abstract

This work focuses on the development of a recombinant pathway for the production of the benzyloquinoline alkaloid (BIA) backbone in *Saccharomyces cerevisiae*. The true intermediate in the native plant pathway is (*S*)-norcoclaurine but extensions of this engineered pathway can also use (*R*, *S*)-norlaudanoline as a precursor. These simple BIAs are derived from two molecules of tyrosine which can be synthesized by yeast in addition to being a standard media component. We engineered numerous yeast strains to express heterologous and endogenous enzymes for the production of dopamine, 4-hydroxyphenylacetaldehyde (4-HPA), and 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) intermediates. While no single strain has been optimized to the point of achieving *de novo* biosynthesis of BIAs, these experiments laid the groundwork for producing these molecules from tyrosine in yeast. This upstream pathway can later be combined with downstream steps for the total biosynthesis of a variety of pharmacologically-relevant molecules.

3.1. Introduction

The benzyloquinoline alkaloids (BIAs) are a large and diverse group of natural products with ~2,500 defined structures, including many with potent pharmacological properties⁴⁰. The complex biosynthesis of BIAs in plants begins with the condensation of the backbone structure derived from two molecules of tyrosine. Along one branch, tyrosine is converted to dopamine and a second molecule of tyrosine is converted to 4-hydroxyphenylacetaldehyde (4-HPA) along a second branch. Dopamine is the precursor for the isoquinoline moiety while 4-HPA is incorporated into the benzyl component (Fig. 3.1). Despite extensive investigations of *Papaver somniferum*, *Eschscholzia californica*, *Thalictrum flavum*, *Coptis japonica* and other BIA producing plants, only two enzymes have been isolated from these very early steps involved in BIA biosynthesis⁶.

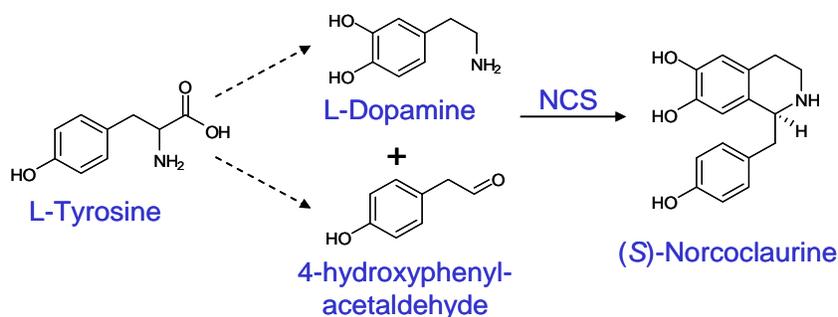


Fig. 3.1. The native pathway for the production of the BIA precursor (*S*)-norcoclaurine synthesized from two molecules of tyrosine.

The earliest enzyme cloned in this pathway is tyrosine decarboxylase (TYDC) which catalyzes the decarboxylation of tyrosine to tyramine or dihydroxyphenylalanine (L-dopa) to dopamine⁴¹. The other cloned enzyme in this upstream pathway is norcoclaurine synthase (NCS) which catalyzes the condensation of dopamine and 4-HPA, the first committed step in BIA biosynthesis. These enzymes can be incorporated into a

recombinant pathway but the remaining activities must be reconstituted using enzymes from other organisms. This work demonstrates the assembly of an engineered pathway in yeast that incorporates activities from plants, humans, bacteria, and fungi for the production of the BIA backbone molecule norcoclaurine and its analog norlaudanosoline.

3.2. Results

3.2.1. Functional expression of tyrosine/dopa decarboxylase from *Papaver somniferum*

Multiple gene sequences coding for tyrosine/dopa decarboxylase enzymes from *P. somniferum* have previously been cloned and characterized in *Escherichia coli*⁴¹. Yeast expression constructs for both TYDC1 and TYDC2 variants were made using a high-copy plasmid backbone and a strong promoter (TEF1). Initial experiments showed that TYDC2 outperformed TYDC1 *in vivo* in our yeast strains, with TYDC1 showing very little activity on tyrosine and L-dopa substrates (data not shown). This is somewhat unexpected given that previous results showed TYDC1 to have a higher specific activity for both substrates. However, kinetic data were gathered using fusion proteins with β -galactosidase which notably formed inclusion bodies in *E. coli* samples⁴¹ and may not be directly transferable to yeast. We also noticed greater conversion of the substrate tyrosine compared to L-dopa, again contradicting *in vitro* results which suggest that the specific activity for tyrosine is only 65% of that of L-dopa. However, feeding an exogenous substrate can also introduce transport limitations, and tyrosine is more likely to be actively transported across yeast cell membranes. Moreover, results have shown that catecholamines including L-dopa and dopamine activate the yeast oxidative stress response pathway due to their ability to autooxidize. This is not only a source of loss of

these intermediates, but production of these molecules may impair growth and have other unintended effects. The similar compounds tyrosine and tyramine are more stable and do not elicit this response⁴².

Results of *in vivo* assays using yeast cells constitutively expressing TYDC2 with tyrosine and L-dopa substrates added to the growth medium at a concentration of 1 mM are shown (Fig. 3.2). Tyramine production by TYDC2 in our engineered yeast strains is estimated to reach over 100 mg l⁻¹ using standard yeast media and increases with the addition of exogenous tyrosine. Endogenous tyrosine biosynthesis can also influence the yield and may be a target for future metabolic engineering efforts.

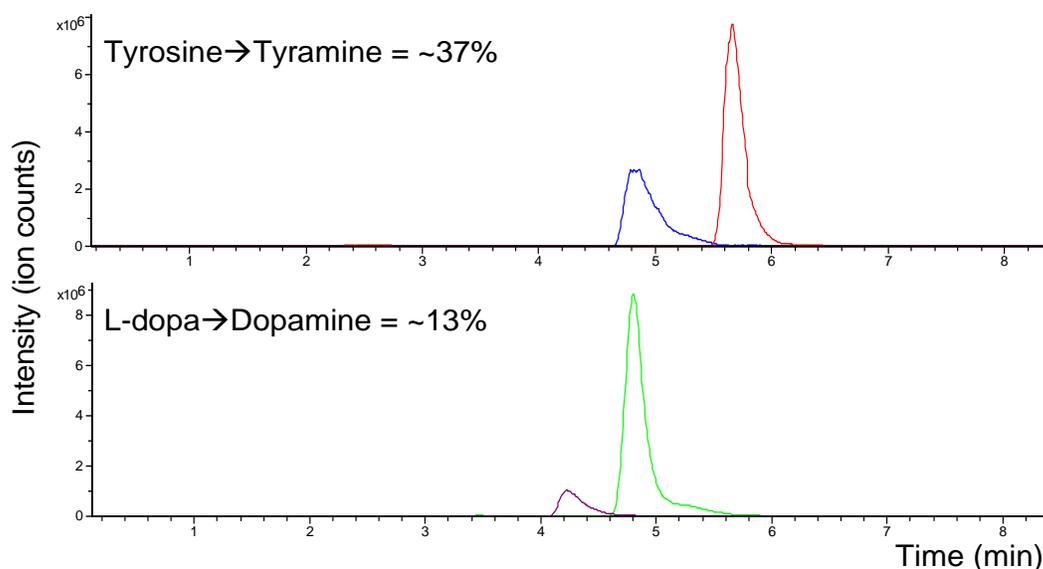


Fig. 3.2. *In vivo* TYDC2 assays. Top: LC-MS analysis of growth media of yeast strains expressing TYDC2 and supplemented with 1 mM tyrosine. Extracted ion chromatograms are shown for tyrosine ($m/z = 182$, red) and tyramine ($m/z = 138$, blue). Bottom: LC-MS analysis of growth media of yeast strains expressing TYDC2 and supplemented with 1 mM L-dopa. Extracted ion chromatograms are shown for L-dopa ($m/z = 198$, green) and dopamine ($m/z = 154$, purple). Percent conversion is calculated by the ratio of the peak areas.

3.2.2. Production of dopamine

Two different pathways for the production of dopamine that incorporate the TYDC2 activity are conceivable. One pathway requires a phenyloxidase activity to convert tyrosine to L-dopa while the other requires a phenolase activity to convert tyramine to dopamine (Fig. 3.3).

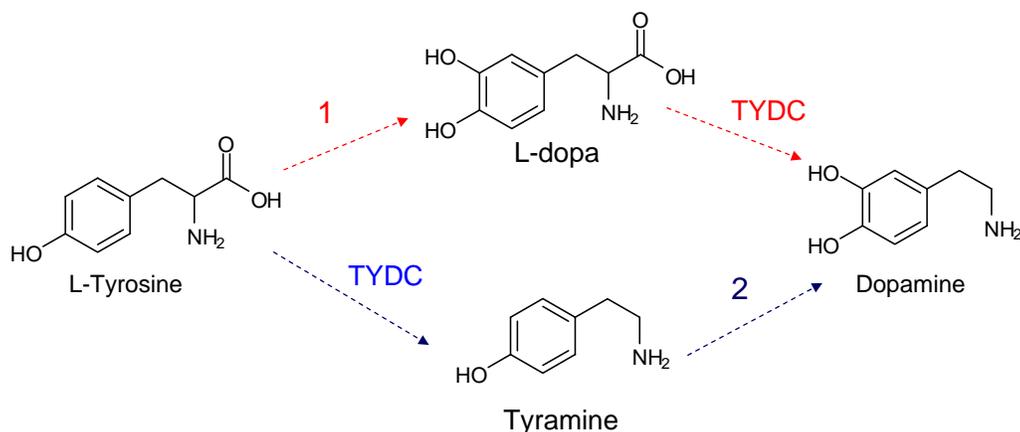


Fig. 3.3. Alternative pathways for dopamine production from tyrosine. The upper pathway (shown in red) uses an unnamed enzyme activity (1) to convert tyrosine to L-dopa and the TYDC activity to convert L-dopa to dopamine. The lower pathway (shown in blue) uses the TYDC activity to convert tyrosine to tyramine and a second enzyme activity (2) to convert tyramine to dopamine.

The oxidation of tyrosine to L-dopa can be catalyzed by three different types of enzymes: tyrosine hydroxylase, tyrosinase, and β -tyrosinase⁴³. Tyrosinase activities are widely distributed in nature as they are key enzymes involved in melanin biosynthesis. However, in addition to catalyzing the orthohydroxylation of phenolic substrates, they also catalyze the oxidation of these catechol products to quinones. In the host organism and in the presence of activating nucleophiles, quinones are spontaneously polymerized to form melanin. Reactive quinone intermediates have antibiotic properties and melanin

itself exerts many strengthening and protective effects; however, accumulation of the L-dopa intermediate is desired in this case, contrary to what happens in nature.

Yeast strains were constructed to express either the human tyrosine hydroxylase 2 (hTH2) or rat tyrosine hydroxylase (TyrH) along with the human GTP cyclohydrolase I (hGTPCHI) required to synthesize the cofactor tetrahydrobiopterin (BH₄). Preliminary results indicated no L-dopa accumulation in the growth media or cell extracts of these strains. In addition, expression levels of these recombinant proteins were assayed by Western blotting using a standard epitope tag, and all three fell below the detection threshold. Difficulties with expression coupled to the fact that hTH2 in particular has a tyrosine hydroxylase to dopa oxidase ratio of ~2:1⁴⁴ led us to pursue other enzymes for this activity. However, methods such as codon-optimization, use of more favorable TH variants (such as hTH4), and optimization of the cofactors Cu(II) and BH₄ may show more encouraging results.

We also obtained sequences coding for two tyrosinase cDNAs from *Agaricus bisporus* (AbPPO1 and AbPPO2), the common button mushroom. Results from AbPPO2 expressed in yeast were encouraging but inconsistent. One strain was able to show production of dopamine when co-transformed with a plasmid expressing TYDC2. However, the addition of 1-10 μM Cu(II)SO₄, which should facilitate the catalytic activity of the copper-containing tyrosinase enzyme, increased accumulation of L-dopa but produced no dopamine. TYDC2 remained active in the presence of Cu(II) as evidenced by tyramine accumulation but failed to convert L-dopa to dopamine under these conditions (Fig. 3.4).

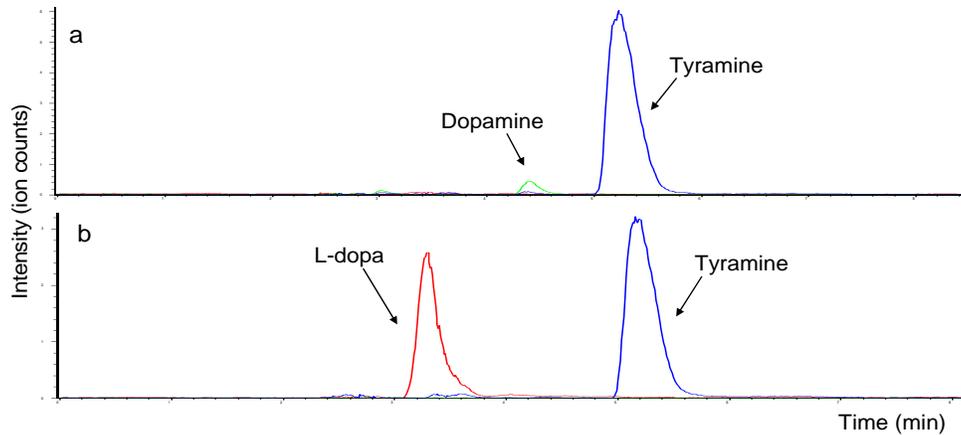


Fig. 3.4. LC-MS analysis of the growth media of CSY88 (Table 3.1) expressing AbPPO2 and TYDC2. **(a)** Production of tyramine ($m/z = 138$, blue) and dopamine ($m/z = 154$, green) are observed in synthetic complete media with no additives. **(b)** Production of tyramine ($m/z = 138$, blue) and L-dopa ($m/z = 198$) are observed when $5 \mu\text{M}$ Cu(II)SO_4 is added to the media; dopamine production is not detectable.

In an effort to increase metabolite shuttling between AbPPO2 and TYDC2 by co-localizing the enzymes *in vivo*, a fusion protein between the two enzymes was constructed and the proteins were tagged with matching C-terminal or N-terminal leucine zippers. In these strains, tyramine production remained high while L-dopa production was nearly abolished, and no dopamine was produced. As proteolytic processing of AbPPO2 in *Agaricus* is believed to take place from the C-terminal region⁴⁵, additional amino acids may interfere with processing and maturation of the protein. This is true specifically in the case of the fusion protein as it was constructed with the AbPPO2 domain on the N-terminus. Also, the C-terminal leucine zipper sequence would likely be cleaved during post-translational processing. Evidence for extensive proteolytic processing is provided by Western blotting analysis using N- and C-terminal tags (Fig. 3.5). However, the number of discrete bands suggests that the desired product does not dominate in our host. Differences in signal sequences and proteases between the two organisms may explain much of the lack of activity seen with our AbPPO2 constructs.

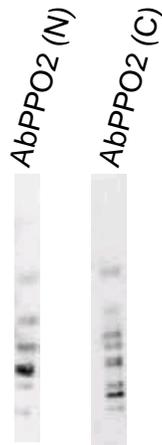


Fig. 3.5. Western blots of AbPPO2 constructs. Left: AbPPO2 with an N-terminal tag detected with an Anti-His G-HRP antibody. Right: AbPPO2 with a C-terminal tag detected with an Anti-V5 HRP antibody.

An alternative pathway for dopamine production proceeds through tyramine and requires the enzymatic addition of a 3'-hydroxyl group. The most well-characterized enzyme that performs this reaction is the human cytochrome P450 2D6 (hCYP2D6). While yeast is a suitable and often-used host for P450 expression, this class of enzymes still presents many difficulties as will be demonstrated throughout this work. Overexpression of the endogenous yeast NADPH cytochrome P450 reductase (yCPR1) or other heterologous P450 reductase is generally required to observe activity of P450 enzymes either *in vivo* or using yeast microsomal preparations. Among the most commonly used strains for expression of heterologous P450s are W(R) and WAT11 strains which contain genomic integrations of GAL-inducible yeast CPR1 and *Arabidopsis thaliana* ATR1 reductases, respectively⁴⁶. We tested both strain backgrounds with co-expression of the TYDC2 and hCYP2D6 enzymes for dopamine production. The engineered yeast cells in the WAT11 strain background showed no dopamine production when grown in the presence of galactose, indicating that the *A. thaliana* ATR1 is not a

suitable reductase partner for hCYP2D6. The W(R) strain background produced dopamine, but surprisingly, no difference was observed between induced and uninduced cultures. This suggests that leaky expression of yCPR1 from the GAL1-10 promoter is sufficient to enhance the activity of its P450 partner above that seen in the wild-type background strain. Additional expression of the yCPR1 driven by the tetO₂ promoter (which is constitutive in the absence of doxycycline) further increased dopamine production. The use of rich media (2x yeast nitrogen base and 5 g l⁻¹ casein hydrolysate instead of ammonium sulfate) also boosted dopamine production slightly but is not an economical solution for an industrial process.

We sought to improve CYP2D6 expression by optimizing codon usage of the gene for yeast. We were not able to observe either improved expression or activity but retained this sequence to reduce metabolic stress on the cells. We co-expressed this optimized sequence with the soluble yCPR Δ 33 and also constructed a fusion protein between yCYP2D6 and yCPR Δ 33. This fusion construct yielded the highest dopamine levels when expressed from a strong TEF6 promoter in a strain background with either an integrated or plasmid-based copy of TYDC2. This engineered yeast strain produced ~10 μ M dopamine while other strains fell short of this mark (Fig. 3.6). Although experimental evidence suggests that this level should be sufficient for incorporation into norcoclaurine, we have not been able to demonstrate this *in vivo* by supplementing 4-HPA. Aside from the redox balance, many other factors may be limiting, including the low affinity of CYP2D6 for the substrate tyramine and autooxidation of dopamine. Additional protein evolution, strain engineering, and optimization of media conditions are required to address these potential issues.

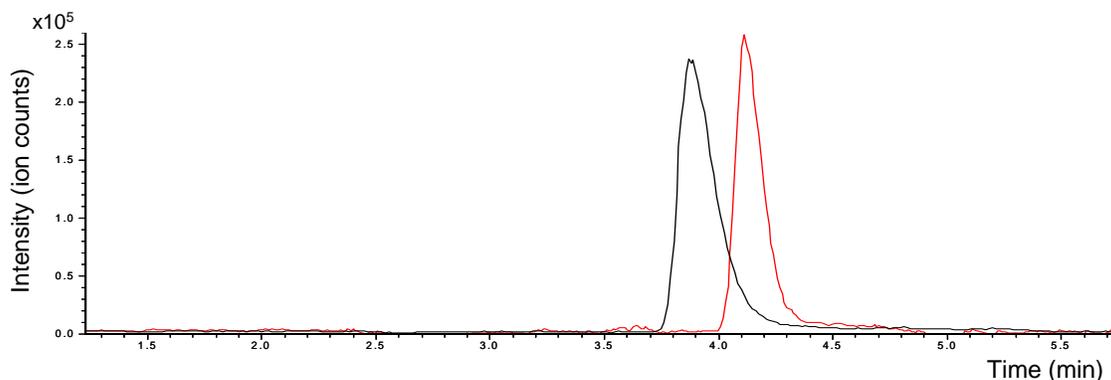


Fig. 3.6. LC-MS analysis of dopamine produced in yeast. Production of dopamine by engineered yeast strains expressing $P_{TEF1}:TYDC2$ and $P_{TEF6}:yCYP2D6-yCPR\Delta33$ from high-copy plasmids (red; pCS221 and pCS1565) compared to a 10 μ M dopamine standard (black). Extracted ion chromatograms for $m/z = 154$ are shown.

3.2.3. Production of 4-hydroxyphenylacetaldehyde and 3,4-dihydroxyphenylacetaldehyde

The other intermediate species required for norcoclaurine formation is 4-hydroxyphenylacetaldehyde (4-HPA). The characterization of this product is severely hindered by the lack of a commercially available standard. A small amount was synthesized (Bio Synthesis, Inc.) for use in downstream enzyme assays but was also very unstable in solution and difficult to detect by LC-MS due to the lack of a basic group for protonatization by electrospray ionization (ESI; Fig. 3.7). While this molecule is more amenable to UV detection, the absorbance of similar molecules in complex biological samples obstructs this type of analysis. Other detection methods such as thin-layer chromatography (TLC) and reaction with Purpald did not yield consistent and reliable results for our samples.

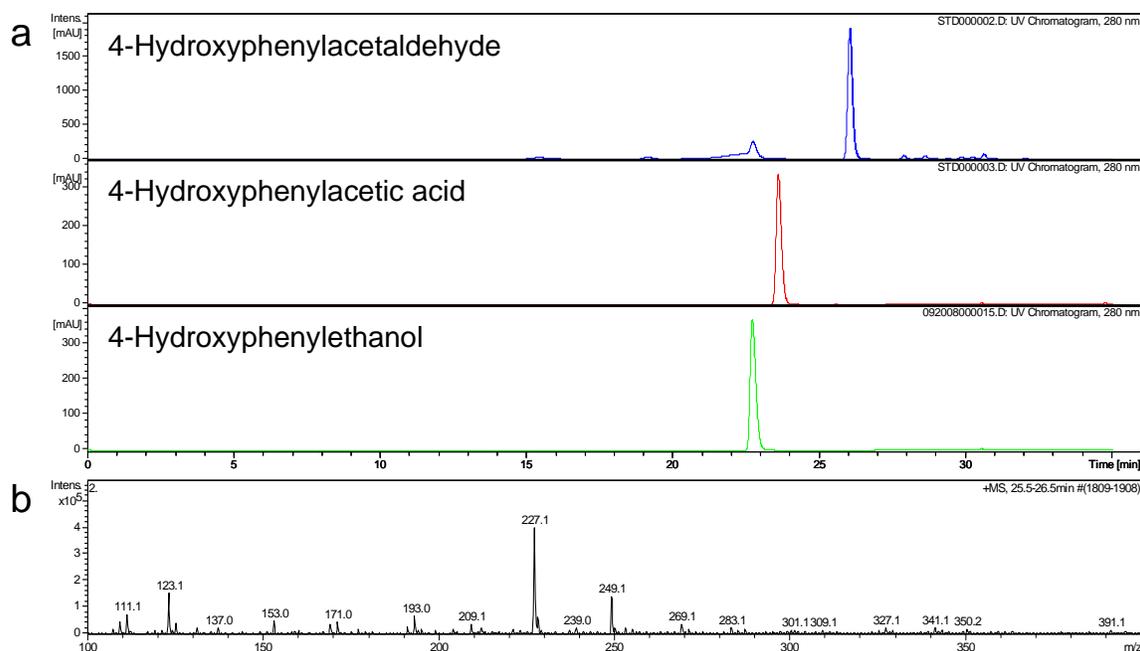


Fig. 3.7. Analysis of tyrosine derivatives. **(a)** LC-UV data of standards 4-HPA (blue; 26 min), 4-hydroxyphenylacetic acid (red; 23.5 min), and 4-hydroxyphenylethanol (green; 23.8 min). Absorbance is measured at a fixed wavelength of 280 nm. **(b)** The MS chromatogram of the 4-HPA peak at 26 min shows very low levels of the expected $m/z = 137$ ion. These small aromatic molecules lacking nitrogen are more amenable to UV detection than ESI-MS.

In addition to being chemically unstable, 4-HPA is also very enzymatically unstable *in vivo* similar to other reactive aldehyde species. We expect 4-HPA to be converted to either 4-hydroxyphenylacetic acid or 4-hydroxyphenylethanol (tyrosol) depending on the redox state of the cell. Tyrosol is one of the fusel alcohols produced via the Ehrlich pathway in yeast. Any one of five aldehyde dehydrogenase enzymes (*ALD2-6*) or seven alcohol dehydrogenase enzymes (*ADH1-7*) can potentially act on 4-HPA to divert this intermediate to central metabolism. For this reason, in many assays we sought to detect not only 4-HPA but other downstream products as well.

One possible pathway for 4-HPA production is the conversion of tyramine by a monoamine oxidase activity (Fig. 3.8a). These enzymes are ubiquitous in nature and we

constructed plasmids and yeast strains to test the *E. coli* monoamine oxidase (MaoA), the *Klebsiella aerogenes* W70 tyramine oxidase⁴⁷, the human liver monoamine oxidase (MAO-A) optimized for expression in yeast⁴⁸, and the *M. luteus* tyramine oxidase⁴⁹. Initial experiments looking at protein expression levels and activity using a standard H₂O₂ assay eliminated the *K. aerogenes* tynA while the *E. coli* MaoA and human MAO-A were shown to be actively expressed in our yeast hosts (Fig. 3.8b). The *M. luteus* tyramine oxidase was later codon-optimized and assembled using short oligonucleotides based on its use in other microbial BIA pathway work⁵⁰.

To test the *in vivo* activity of the monoamine oxidase variants, we added exogenous tyramine (or used a background strain expressing TYDC2) and performed metabolite analysis of the growth media. In samples expressing monoamine oxidase activities and supplemented with 1 mM tyramine, we used LC-UV to look for peaks corresponding to 4-HPA along with 4-hydroxyphenylacetic acid and 4-hydroxyphenylethanol (Fig 3.8c and Fig 3.7a). However, the major difference between these spectra was in the neighborhood of 6 min rather than the later elution times of the expected end products of known 4-HPA utilization pathways. The nature of these metabolites is unclear as no major ion is predominant in these peaks. The same strains grown in the absence of tyramine showed similar increases in this region, indicating that the monoamine oxidases may act on other natural metabolites as well. These *in vivo* assays also make it unclear whether the monoamine oxidase enzymes accept tyramine as a substrate. *In vitro* assays were performed using protein lysates from *E. coli* cells overexpressing the native MaoA. Although we could not detect 4-HPA from the MaoA reaction, we were able to successfully couple this product to the NCS reaction to produce

norcoclaurine *in vitro* (Fig. 3.8d). This provides evidence that the bacterial monoamine oxidase is capable of producing 4-HPA but it is currently unclear whether *in vivo* conditions in yeast are conducive to this reaction.

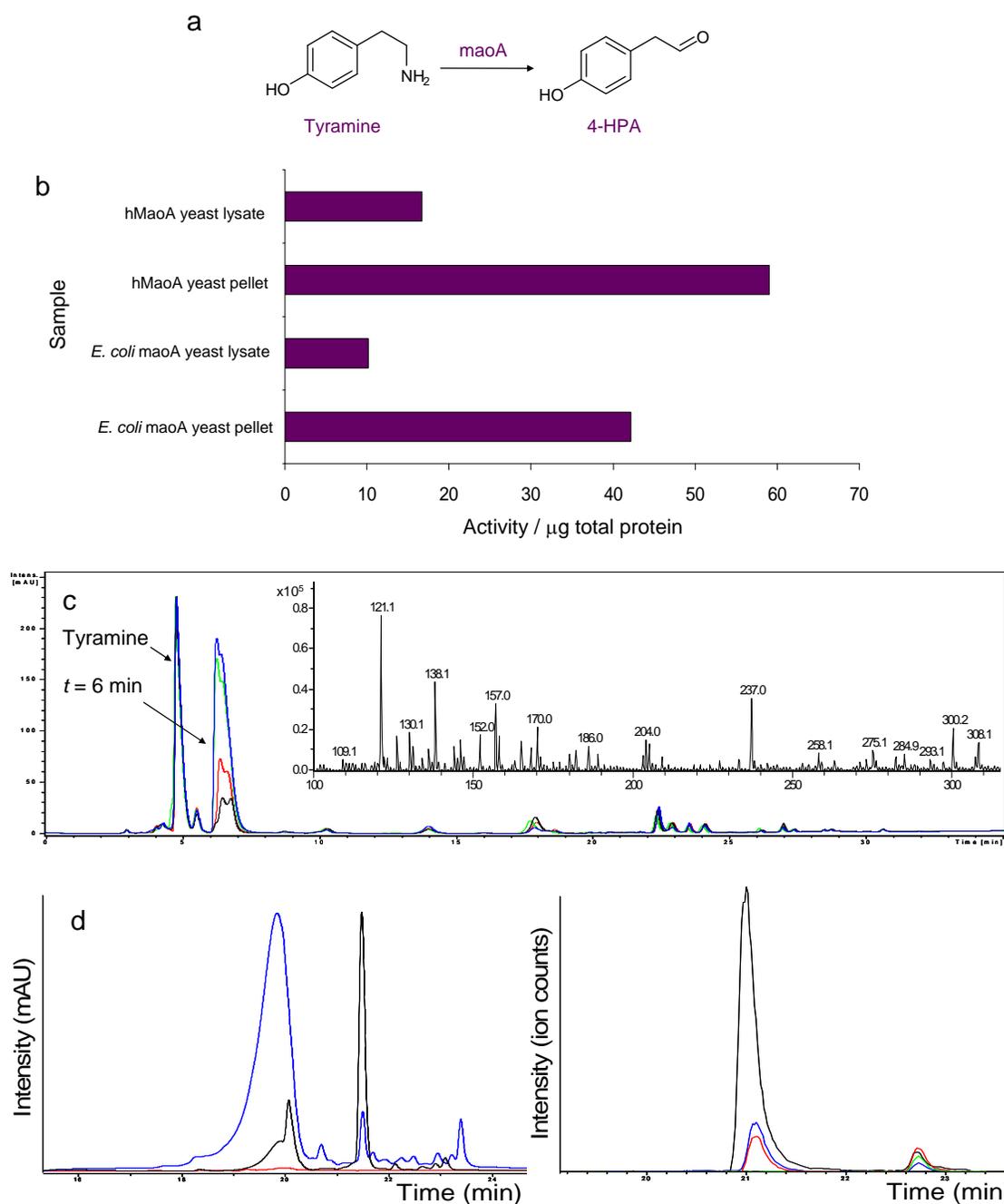


Fig. 3.8 Monoamine oxidase enzyme assays. **(a)** Pathway for the conversion of tyramine to 4-HPA using a monoamine oxidase activity. **(b)** Results of H_2O_2 assay on yeast cell lysates and lysed cell pellets. The majority of monoamine oxidase activity is found in the

cell pellet consistent with Western blotting experiments which indicate that the majority of the monoamine oxidase protein co-localizes with the cell debris upon lysis with Y-PER. **(c)** LC-UV analysis of the growth media of yeast strains supplemented with 1 mM tyramine. Negative control (no monoamine oxidase) = black; *E. coli* maoA = red; hMaoA = green; *M. luteus* maoA = blue. The inset ion chromatogram is representative of the enlarged peak at ~6 min seen in strains expressing the monoamine oxidase; no major ion is dominant. **(d)** Left: LC-UV analysis (276 nm) of *E. coli* maoA *in vitro* reaction (red), tyramine oxidase *in vitro* reaction (positive control; blue) and 4-HPA (black). Right: Extracted ion chromatograms ($m/z = 272$) of NCS *in vitro* reactions with *E. coli* maoA reaction product (red), tyramine oxidase reaction product (blue), authentic 4-HPA (black), and no 4-HPA (green).

Each of these monoamine oxidase variants from bacteria and human sources are reported to accept tyramine and dopamine along with other similar substrates. This means that these enzymes can produce 3,4-DHPA in addition to 4-HPA since tyramine and dopamine are both potentially available intermediates in our system. It has been reported previously that dopamine and its oxidized product 3,4-DHPA can spontaneously condense to form norlaudanosoline (Fig. 3.9a)⁵¹. This is also a potentially useful molecule as it differs from the natural intermediate norcoclaurine by only a 3'-hydroxyl group. Although the formation of norlaudanosoline from dopamine was recently reported to occur in *E. coli* supplemented with 5 mM dopamine⁵⁰, we found that in our strains up to 100 mM exogenous dopamine was required to detect accumulation of norlaudanosoline in the growth media. We also found that the human MAO-A is the preferred variant for this reaction as expected according to its biological function (Fig. 3.9b).

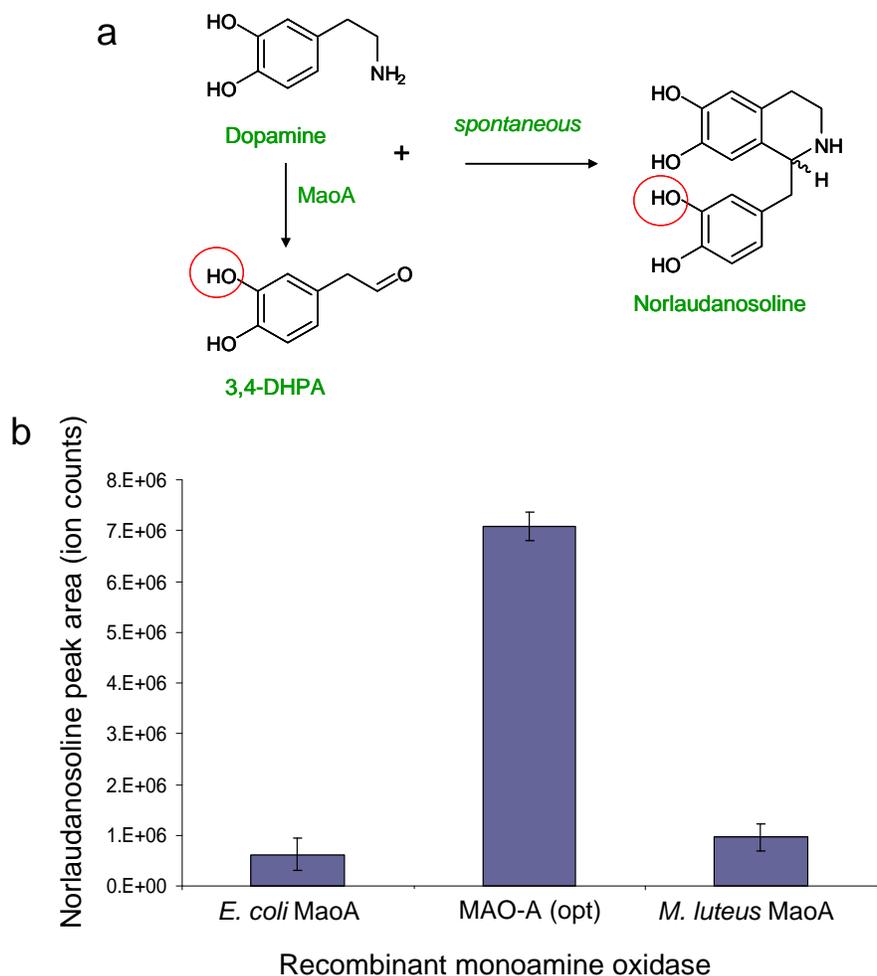


Fig. 3.9. Production of norlaudanosoline from dopamine. **(a)** Pathway for dopamine conversion to 3,4-DHPA and norlaudanosoline. The only difference between norcoclaurine and norlaudanosoline is the presence of the 3'-hydroxyl group circled in red. **(b)** Comparison of norlaudanosoline production between strains expressing the *E. coli* MaoA, the human MAO-A optimized for yeast, and the *M. luteus* MaoA supplemented with 0.1 M dopamine.

Although we were able to produce detectable levels of norlaudanosoline, it was still unclear why this reaction was so inefficient in our yeast hosts, requiring 20x more dopamine than reported for *E. coli*⁵⁰. Possible reasons include transport, spontaneous oxidation of dopamine at low pH, and degradation of the 3,4-DHPA intermediate by endogenous enzymes. In an attempt to facilitate dopamine uptake, we constructed strains to overexpress a mutant form of the general amino acid permease (Gap1^{K9K16})⁵² and

observed a 6 to 7-fold increase in norlaudanosoline production (Fig. 3.10a). We also tested single *ADH* knockouts expressing MAO-A and observed a ~30-fold increase in norlaudanosoline production from 100 mM dopamine in *adh2Δ*, *adh5Δ*, and *adh6Δ* strains (Fig. 3.10b). While we did observe increases in norlaudanosoline accumulation of ~15-fold in *ald2Δ* and *ald3Δ* strains, the results indicated that 3,4-DHPA is preferentially oxidized to the alcohol in standard yeast media. A combinatorial knockout of *ADH2*, *ADH5*, *ADH6*, *ALD2*, and *ALD3* alcohol and aldehyde dehydrogenase enzymes is expected to show additive effects. Future work includes testing of combinatorial *ADH* and *ALD* knockout strains for an increase in the flux of 4-HPA and 3,4-DHPA through our recombinant pathway.

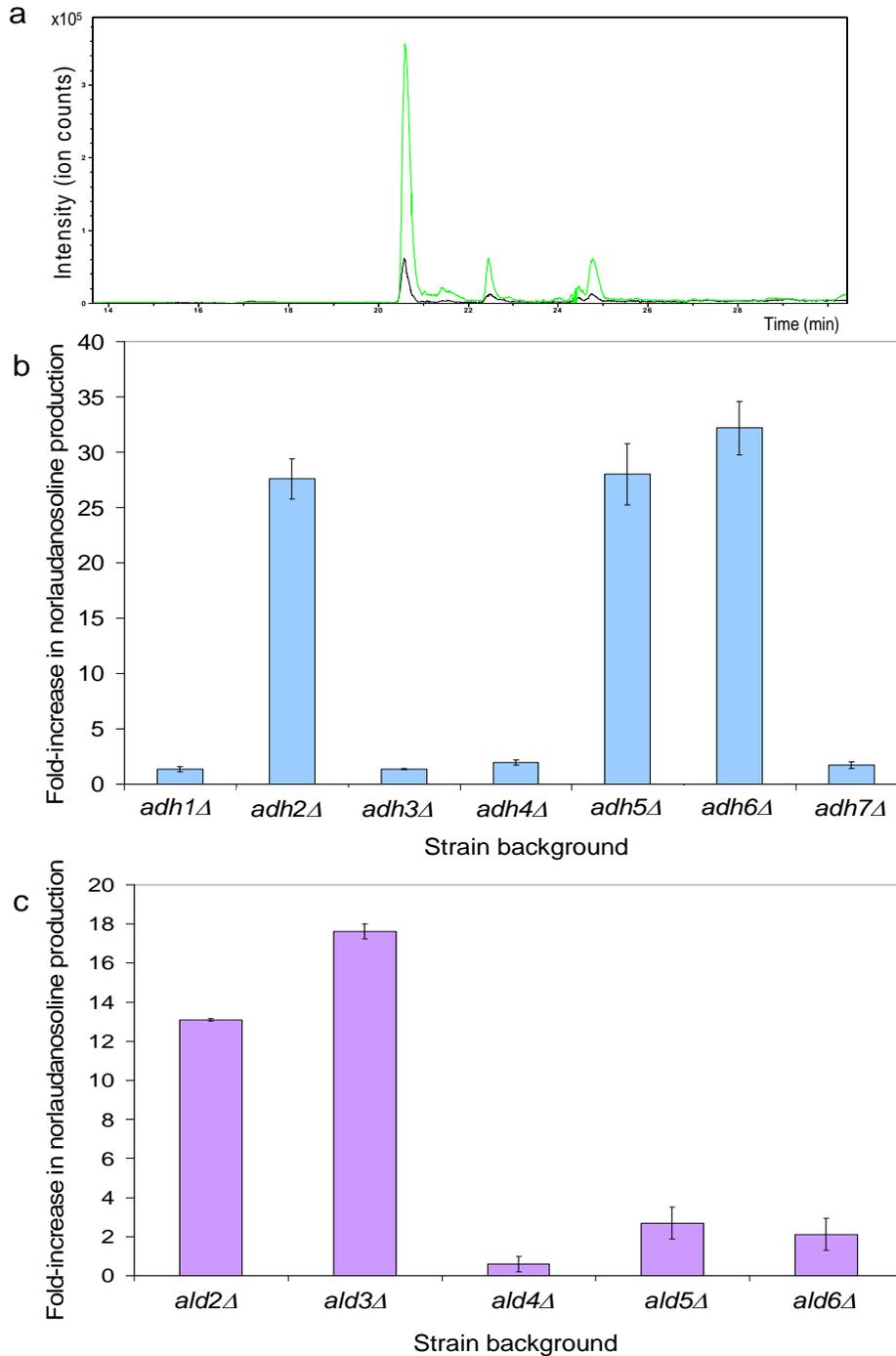


Fig. 3.10. Strains with increased norlaudanosoline accumulation. **(a)** Extracted ion chromatogram of norlaudanosoline ($m/z = 288$) produced from 1 mM dopamine in yeast strains expressing MAO-A and Gap1p^{K9K16} (green) and MAO-A only (black). The peak at 20.7 min is the correct norlaudanosoline product. **(b)** Norlaudanosoline production from 100 mM dopamine in *ADH1-7* and **(c)** *ALD2-6* single knockout strain backgrounds expressing MAO-A normalized to the same background strain expressing MAO-A with no deletions.

An alternative pathway to 4-HPA uses endogenous yeast genes *ARO8/9* and *ARO10* involved in the biosynthesis and degradation of tyrosine (Fig. 3.11). The *ARO8* and *ARO9* genes are analogous aromatic amino acid transaminase activities. Aro8p, classified as a type I aminotransferase, produces L-glutamate from 2-oxoglutarate as a by-product whereas Aro9p, a type II aminotransferase, produces L-alanine from pyruvate⁵³. The *ARO10* gene product is a decarboxylase enzyme implicated in the degradation of aromatic and branched chain amino acids for the production of fusel alcohols via the Erhlich pathway.

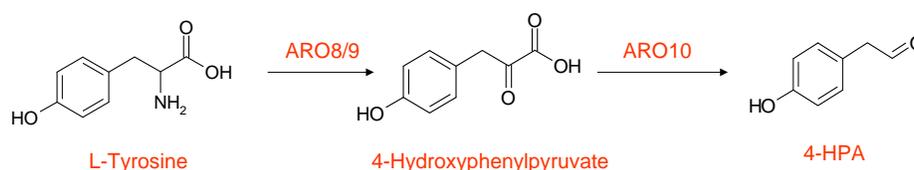


Fig. 3.11. Pathway for the production of 4-HPA from tyrosine using the endogenous yeast genes *ARO8/9* and *ARO10*.

Since the *ARO* genes are native to yeast, they are an obvious choice for this activity. However, these genes are also subject to endogenous regulation designed to prevent utilization of inferior nitrogen sources. For example, Aro9p expression levels are more than 10-fold higher when grown on urea versus ammonia, a superior nitrogen source, with similar transcriptional regulation expected for Aro10p⁵⁴. Transcriptional regulation alone is simple to overcome by expressing the gene behind a promoter that is constitutive and/or active under the desired growth conditions. However, experimental results using strains with constitutive expression of Aro10p indicate that this protein is subject to post-translational regulation and/or requires a second protein for activity⁵⁵. As the details of this regulation are still unknown, we expect Aro10p to show low activity in

cultures grown on glucose and ammonia. With this in mind, we made plasmid-based constructs to constitutively express *ARO8*, *ARO9* and *ARO10*. We transformed an empty vector control, a plasmid expressing *ARO8* or *ARO9* only, and a plasmid expressing both *ARO9* and *ARO10* and examined the growth media for the formation of UV peaks corresponding to 4-HPA and its metabolites in standard synthetic complete yeast drop-out media (Fig. 3.12). In strains expressing only *ARO8* or *ARO9*, we observed an increase in the peak identified as 4-hydroxyphenylpyruvate, and in strains expressing *ARO9* and *ARO10*, we observed this same peak plus an additional peak corresponding to tyrosol. No peaks corresponding to 4-HPA were observed. These results suggest that overexpression of *ARO8/ARO9* and *ARO10* can convert tyrosine to 4-HPA, and moreover, that the 4-HPA is efficiently converted to tyrosol under standard glucose-grown conditions.

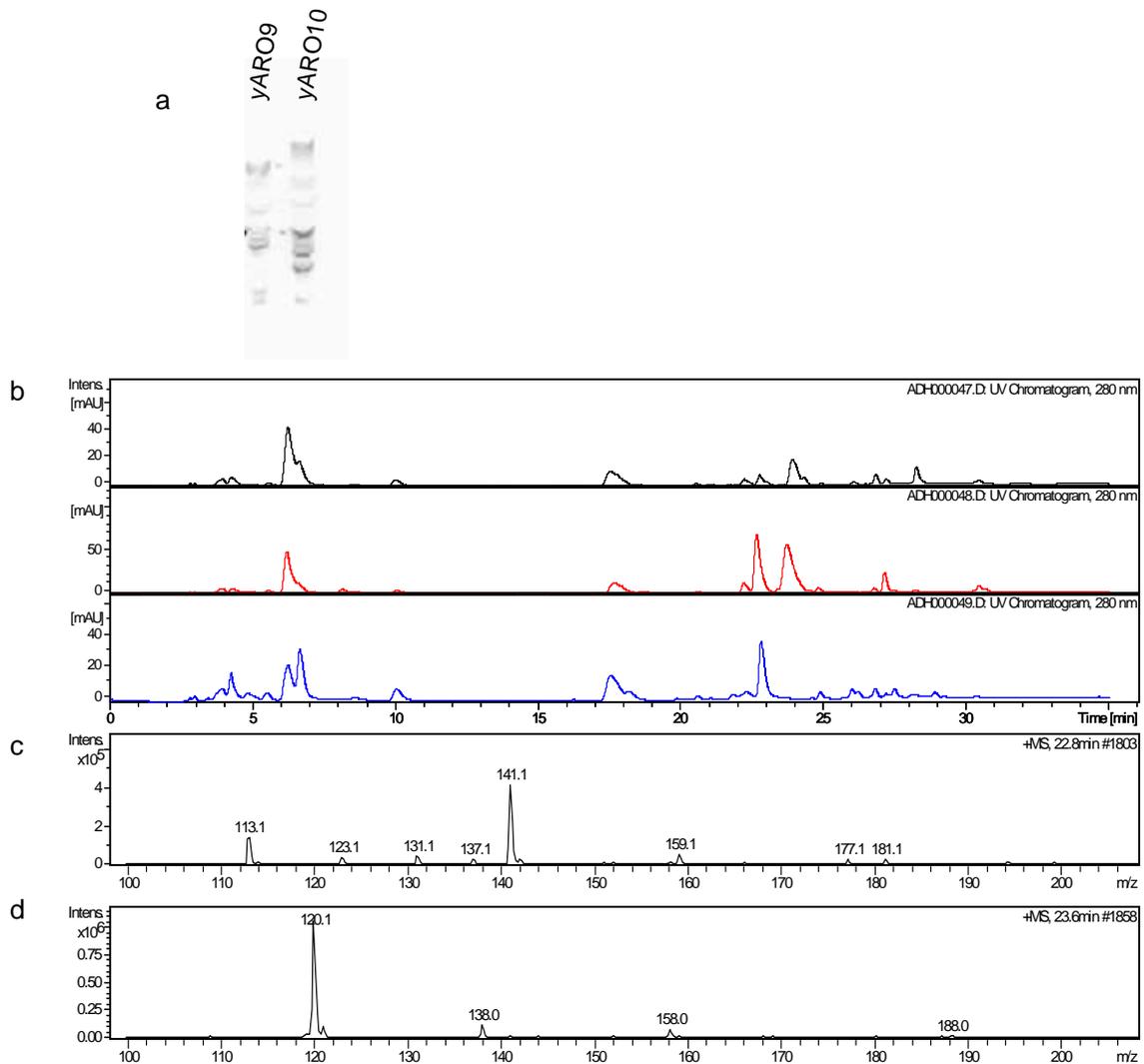


Fig. 3.12. Analysis of yeast strains overexpressing the *ARO* genes. **(a)** Western blotting using C-terminal V5 epitope tags indicate extensive proteolytic processing of Aro9p and Aro10p under glucose-grown conditions. **(b)** LC-UV chromatograms of the growth media of yeast strains constitutively expressing *ARO* genes. Negative control lacking an enzyme coding sequence (black), *ARO9* and *ARO10* (red), and *ARO9* only (blue). **(c)** Mass spectrometer data for the peak at 22.8 min identified as 4-hydroxyphenylpyruvate with the major ion $m/z = 181$. **(d)** Mass spectrometer data for the peak at 23.6 min identified as tyrosol with characteristic ions $m/z = 138$ and $m/z = 120$.

We tested a construct overexpressing Aro9p and Aro10p in the *ADH* and *ALD* knockout strains hoping to observe decreased conversion to tyrosol and the possible emergence of a peak corresponding to 4-HPA. While we did observe a slight increase in

4-HPA production in many of the knockouts, we also observed larger increases in tyrosol and 4-hydroxyphenylacetic acid. This is surprising and likely a consequence of the cells adjusting to overexpression of the ARO genes. It is unclear whether combinatorial knockouts will show more promising results.

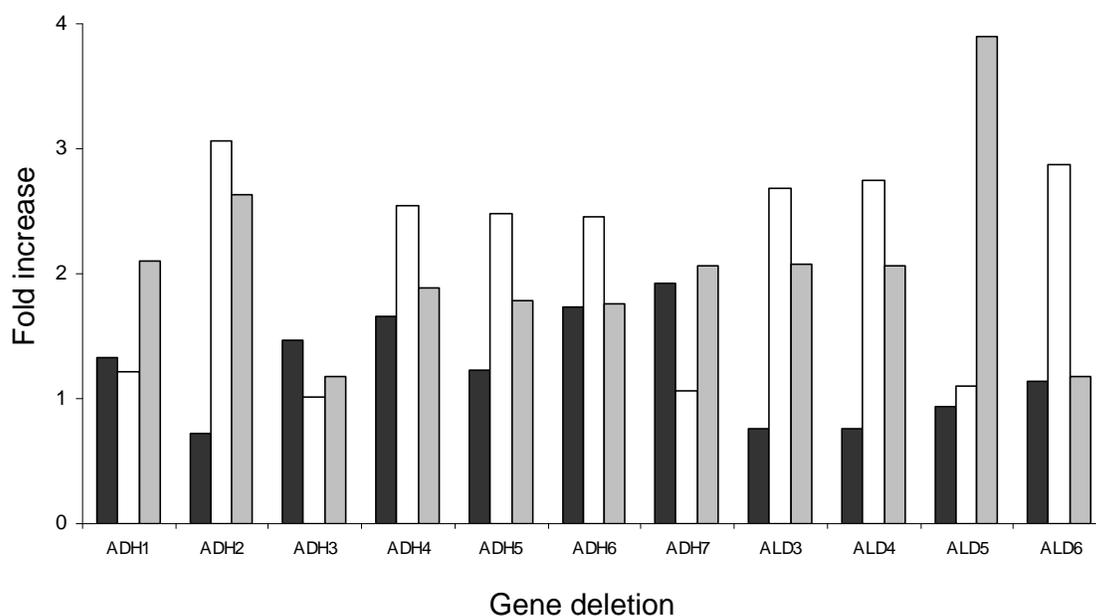


Fig. 3.13. Changes in 4-HPA, tyrosol, and 4-hydroxyphenylacetic acid production in knockout strains overexpressing Aro9p and Aro10p. Production of 4-HPA (black), tyrosol (white), and 4-hydroxyphenylacetic acid (grey) are shown for each knockout strain background; the area of each UV chromatogram peak is normalized to the control strain also expressing Aro9p and Aro10p with no gene deletions.

3.2.4. Production of norcoclaurine and norlaudanosoline

The end goal of the aforementioned engineered yeast strains is to produce a sufficient amount of dopamine and 4-HPA (or 3,4-DHPA) to be incorporated into the BIA backbone molecules norcoclaurine and/or norlaudanosoline. Both products are possible with the same combination of enzymes if a monoamine oxidase activity is used and both tyramine and dopamine substrates are present. However, assuming the ARO

genes will produce only 4-HPA from tyrosine and no other aldehydes that can react with dopamine, we should accumulate only norcoclaurine in engineered strains using this alternative pathway (Fig. 3.14)

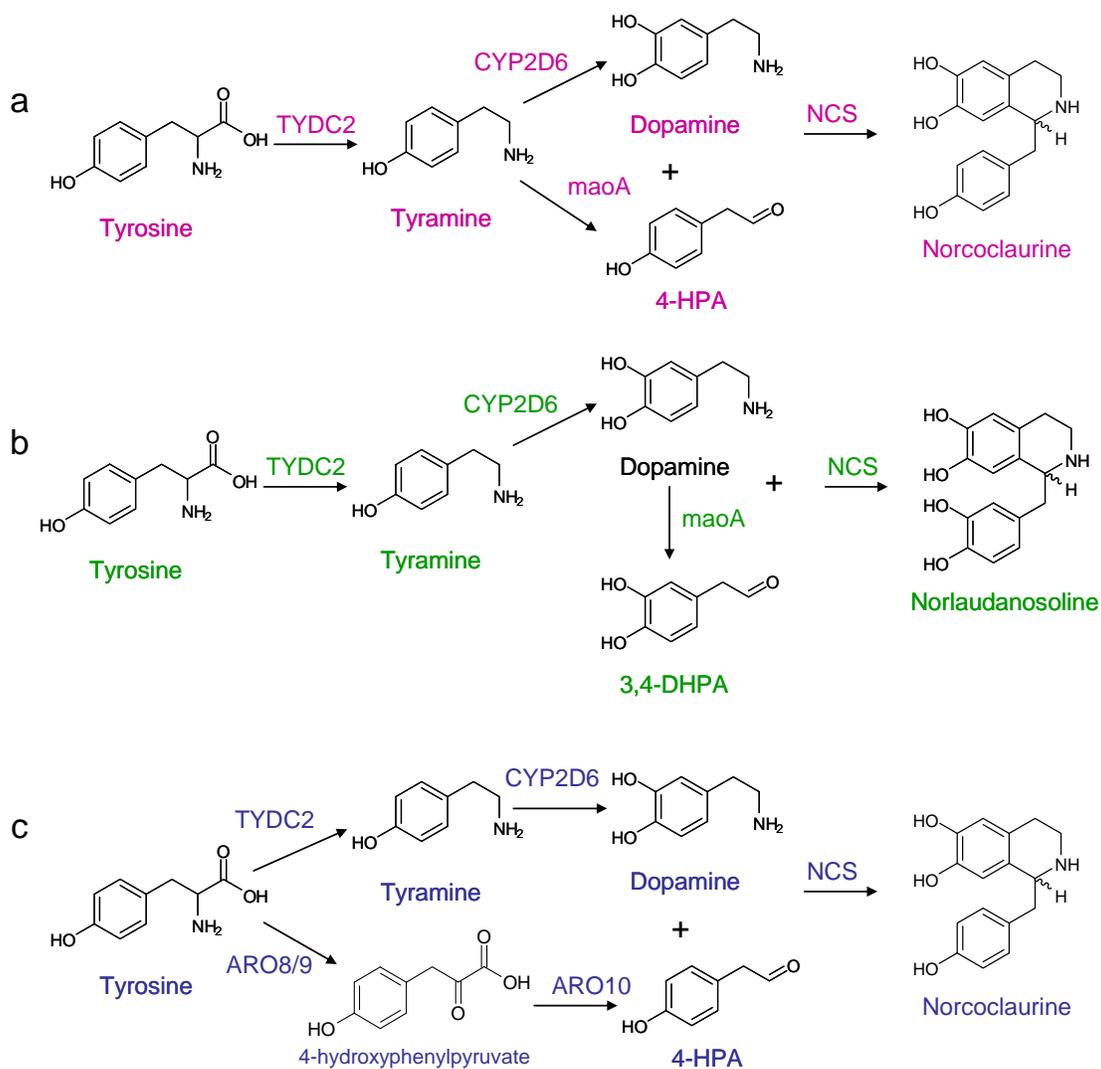


Fig. 3.14. Pathways for norcoclaurine and norlaudanosoline production. The NCS activity catalyzing the final step is optional as this condensation reaction occurs spontaneously *in vivo*. **(a)** Production of norcoclaurine using TYDC2, CYP2D6, and maoA. **(b)** Production of norlaudanosoline using the same set of enzymes as shown in (a). **(c)** Production of norcoclaurine using TYDC2 and CYP2D6 to produce dopamine and the yeast *ARO* genes to produce 4-HPA.

In plants, the enzyme norcoclaurine synthase (NCS) catalyzes the reaction between dopamine and 4-HPA. NCS has been cloned and characterized from *T. flavum*⁵⁶ and *C. japonica*⁵⁷; however, the homologous PR10 proteins from *P. somniferum* are not catalytically active. Mechanistic studies on the *T. flavum* and *C. japonica* NCS variants show that these proteins from two different families are capable of catalyzing the same reaction *in vitro* although it is unclear what role NCS plays in the native hosts.

We obtained the *T. flavum* NCS variant to test in our engineered yeast strains. Previous characterization of the *T. flavum* NCS showed that removal of the N-terminal signal sequence by making a $\Delta 10$ or $\Delta 19$ truncation facilitated soluble expression in a recombinant host⁵⁶; however, we were also able to observe expression of the full-length NCS in our strains (Fig. 3.15a). We tested each variant *in vitro* and *in vivo* with exogenous dopamine and 4-HPA. Not only did we not observe differences between the NCS variants, but we found that negative controls lacking an NCS sequence produced similar levels of norcoclaurine. We observed this same product in *in vitro* reactions containing wild-type yeast lysates but not when *E. coli* lysates were used. All *in vivo* and *in vitro* products had the same elution time and the correct mass fragmentation pattern (Fig. 3.15b)⁵⁶. Our finding that NCS is not required for the formation of the BIA backbone is in agreement with other studies using heterologous hosts⁵⁰.

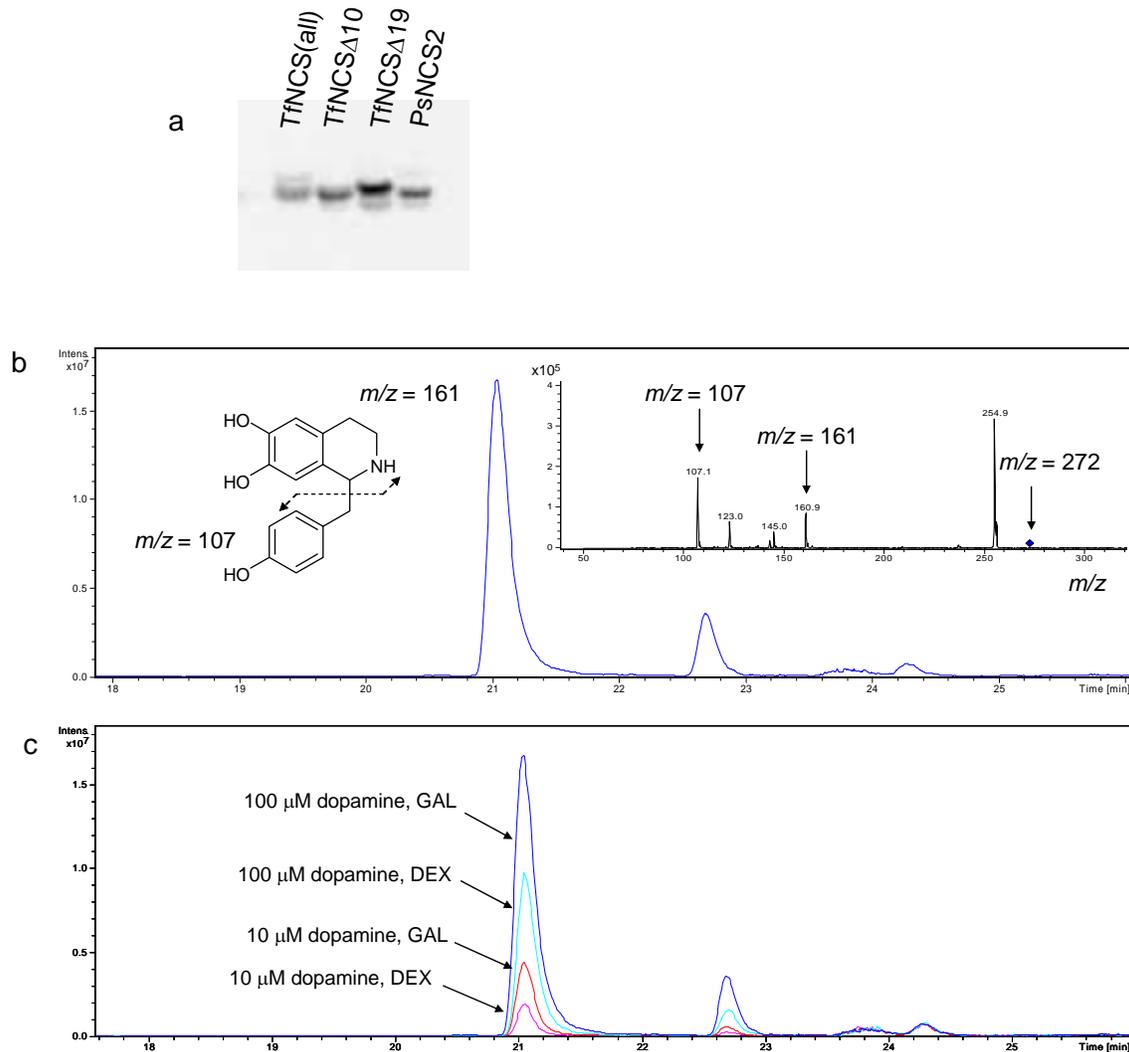


Fig. 3.15. Expression of NCS and production of norcoclaurine. **(a)** Western blot showing expression of V5-tagged proteins—*T. flavum* NCS (entire sequence), NCS Δ 10, and NCS Δ 19. The *P. somniferum* NCS2 variant is also shown but not discussed as subsequent reports indicate it is not catalytically active⁶. **(b)** Production of norcoclaurine *in vivo* by feeding dopamine and 4-HPA. The extracted ion chromatogram of $m/z = 272$ is shown in blue, with the inset showing the correct fragmentation pattern for this product eluting at 21 min; the second peak at 22.7 min also showed the correct fragmentation pattern and was consistent in all *in vivo* and *in vitro* experiments. **(c)** Norcoclaurine production by feeding 4-HPA and limiting amounts of dopamine. NCS was expressed from a GAL-inducible promoter to demonstrate that the enzyme assists this reaction when one (or both) substrates is limiting. Extracted ion chromatograms for norcoclaurine ($m/z = 272$) are shown; magenta = 10 μ M dopamine grown in dextrose; red = 10 μ M dopamine

grown in galactose; cyan = 100 μ M dopamine grown in dextrose; blue = 100 μ M dopamine grown in galactose.

We also performed experiments titrating dopamine in a strain with the TfNCS Δ 10 expressed from the GAL1-10 promoter. The results show that expression of NCS increases production of norcoclaurine when one or both substrates are limiting (Fig. 3.15c). In contrast to this result, however, we found that expression of NCS hindered norlaudanosoline production. We fed 100 mM dopamine to strains expressing MAO-A alone or with co-expression of TfNCS Δ 10 and observed much lower norlaudanosoline production in strains co-expressing TfNCS Δ 10 (Fig. 3.16). We also fused matching C-terminal leucine zippers⁵⁸ to each enzyme to facilitate co-localization of the 3,4-DHPA and NCS enzyme for a more efficient reaction and to prevent degradation of the acetaldehyde intermediate. Substantially lower norlaudanosoline production was also observed in this strain. It is not surprising that NCS does not facilitate this reaction since 3,4-DHPA is not the true substrate. However, the mechanism by which the enzyme actually interferes with norlaudanosoline production is unclear. Perhaps the substrates can bind to the protein but not react in the active site due to presence of the additional hydroxyl group.

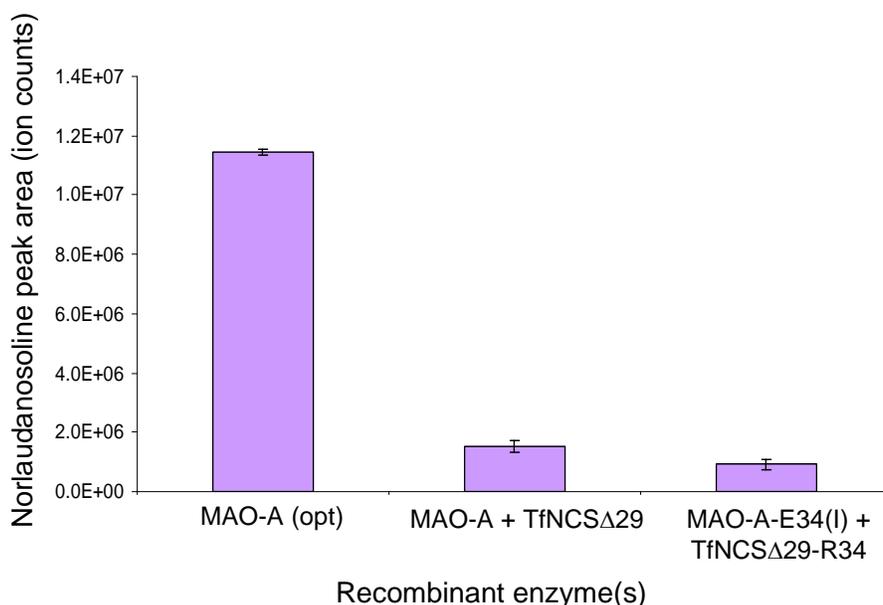


Fig. 3.16. Comparison of norlaudanosoline production with NCS. Norlaudanosoline production as measured by the extracted ion chromatogram area of strains expressing the human MAO-A optimized for yeast with or without the *T. flavum* NCSΔ29 and with C-terminal leucine zippers to co-localize the hMaoA and the *T. flavum* NCSΔ29.

We also tested production of norcoclaurine in our *ADH* and *ALD* knockout strains expressing *ARO9* and *ARO10*. For these experiments, we also had to feed a large excess of dopamine (100 mM) to see production of norcoclaurine. While we were able to observe increases in norcoclaurine production in several of the knockout strains, the results were not as dramatic as those observed for norlaudanosoline, with a maximum increase of ~3-fold compared to ~35-fold for norlaudanosoline. This is consistent with observations of 4-HPA, tyrosol, and 4-hydroxyphenylacetic acid production which suggest that the 4-HPA intermediate is otherwise diverted when only one of these enzymes is deleted. Since 4-HPA is a natural intermediate, it is also logical that multiple endogenous enzymes are capable of converting this substrate while fewer may act on 3,4-DHPA and/or have lower activity on this acetaldehyde. It is unclear from these results

whether combinatorial knockouts will have additive effects or how many deletions are required to prevent 4-HPA degradation. Once combinatorial knockouts are constructed, we can test strains co-expressing Gap1^{K9K16} to improve dopamine transport. Ultimately, however, we would like to incorporate enzymes for *in vivo* dopamine production in these strains as well. Expression of NCS is also expected to enhance norcoclaurine production when substrate concentrations are low.

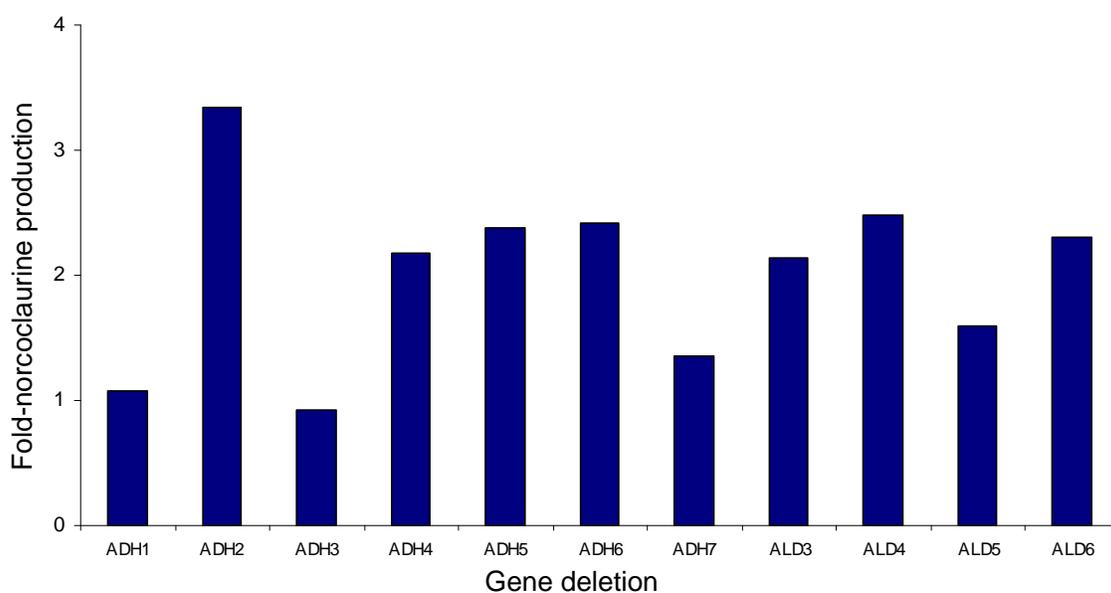


Fig. 3.17. Norcoclaurine production in *ADH* and *ALD* knockout strains. The fold increase in norlaudanosoline production from strains expressing Aro9p and Aro10p supplemented with 100 mM dopamine is normalized to the control strain with no deletions. Norcoclaurine is measured as the area of the extracted ion chromatogram peak ($m/z = 272$) with an elution time of 21.2 min.

Our results show that the production of norcoclaurine and norlaudanosoline occur readily in our yeasts hosts when sufficient precursors are available. The expression of NCS may lead to higher yields of norcoclaurine but may actually inhibit norlaudanosoline production. Expression of NCS may be one way to direct production of either norcoclaurine or norlaudanosoline as an alternative to using the *ARO* genes for

production of only the 4-HPA intermediate. It is also noteworthy that other variants such as the *Coptis japonica* NCS may have different specificities. The *C. japonica* NCS was reported to have a different mechanism⁵⁷, and although it also failed to enhance norlaudanosoline synthesis in a microbial host, it reportedly did not inhibit this reaction⁵⁰.

3.3. Discussion

We have identified several recombinant and endogenous yeast enzymes capable of performing the reactions to produce the BIA backbone molecules norcoclaurine and norlaudanosoline from tyrosine. We have constructed yeast strains to test these pathways and made several efforts to optimize production of the precursor molecules. However, we have not yet reached levels of dopamine required for the total biosynthesis of norcoclaurine and/or norlaudanosoline *in vivo*. While we have demonstrated production of tyramine and dopamine, we believe that dopamine is still limiting based on experiments showing production of BIA molecules upon exogenous addition of dopamine.

For the production of norlaudanosoline, which is not the natural intermediate but useful for synthetic pathways, we were able to demonstrate production by feeding dopamine and expressing only the human MAO-A. However, this remains an extremely inefficient process even with improvements in intracellular transport and partial elimination of a competing pathway. Again, *in vivo* dopamine production was not sufficient to be converted to 3,4-DHPA and norlaudanosoline. These negative results may be due, at least in part, to the autooxidation of dopamine and the instability of the acetaldehyde intermediate. Similarly, we were able to demonstrate production of

norcoclaurine by feeding dopamine and expressing Aro9p and Aro10p. Production was significantly enhanced in several *ADH* and *ALD* knockout backgrounds. From our best strains containing only one gene deletion and not co-expressing Gap1p^{K9K16}, we are able to produce ~ 10 mg l⁻¹ norcoclaurine or norlaudanosoline from 100 mM dopamine.

Much work remains to optimize the individual pathways for norcoclaurine and norlaudanosoline production. This may require the exploration of other enzymatic activities or protein engineering/evolution, particularly to increase the activity of CYP2D6 and/or AbPPO2. The construction of combinatorial knockouts and possibly additional strain engineering is likely required to direct the flux of 4-HPA and 3,4-DHPA towards BIA production by blocking endogenous degradation pathways. In addition, advanced engineering strategies such as building protein scaffolds may be employed to increase local concentrations of these intermediates and facilitate metabolite channeling between enzymes.

3.4. Materials and Methods

3.4.1. Plasmid and yeast strain construction

We obtained restriction enzymes, T4 DNA ligase, and other cloning enzymes from New England Biolabs. We performed polymerase chain reaction (PCR) amplifications using Expand High Fidelity PCR System (Roche). Oligonucleotide synthesis was performed by Integrated DNA Technologies. A list of selected plasmids and yeast strains is provided (Table 3.1).

We used standard molecular biology techniques to construct the BIA expression vectors⁵⁹. BIA expression constructs contained the 2 μ high-copy yeast origin of

replication along with appropriate yeast selection markers and ampicillin resistance. Recombinant enzymes were expressed from the yeast TEF1 promoter and flanked by a CYC1 terminator sequence. We constructed shuttle vectors for subcloning of 1 or 2 cDNA sequences in this fashion. The *hCYP2D6* cDNA was provided by F. Peter Guengerich (Vanderbilt University) as pCW/DB6⁶⁰ and the yeast codon-optimized version of this gene was synthesized by DNA 2.0. We PCR-amplified the endogenous yeast P450 reductase gene (*CPRI*) from W303 genomic DNA and the *Homo sapiens* *CPRI* gene from pH2E1red⁶¹.

We transformed ligation reactions into an electrocompetent *E. coli* strain, DH10B (Invitrogen; F-*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*dlacZ* Δ M15 Δ *lacX74 deoR recA1 endA1 araD139 Δ (*ara, leu*)7697 *galU galK* λ -*rpsL nupG*), using a Gene Pulser Xcell System (BioRAD) according to the manufacturer's instructions. We conducted plasmid isolation using the Wizard Plus SV Minipreps DNA purification system (Promega) according to the manufacturer's instructions. Subcloning was confirmed by restriction analysis and sequence verification (Laragen, Inc.). We transformed plasmids into the appropriate *S. cerevisiae* strains using a standard lithium acetate protocol⁶². All yeast strains used in this work were based on the haploid yeast strain W303 α (MAT α *his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*)⁶³. *E. coli* cells were grown on Luria-Bertani media (BD Diagnostics) with 100 μ g/ml ampicillin (EMD Chemicals) for plasmid maintenance, and *S. cerevisiae* cells were grown in synthetic complete media (BD Diagnostics) supplemented with the appropriate dropout solution for plasmid maintenance (Calbiochem).*

Table 3.1. Plasmids and yeast strains.

Plasmids	
pCS138	P_{TEF1} -GAP1 ^{K9K16}
pCS221	P_{TEF1} -TYDC2
pCS250	P_{TEF1} -TfNCS Δ 10
pCS251	P_{TEF1} -AbPP02
pCS283	P_{TEF1} -TYDC2, P_{TEF1} -maoA
pCS330	P_{TEF1} -TYDC2, P_{TEF1} -yCYP2D6
pCS405	$P_{GAL1-10}$ -TfNCS Δ 10(V5-HIS tag)
pCS637	P_{TEF1} -yARO10
pCS689	P_{TEF1} -yARO9, P_{TEF1} -yARO10
pCS785	P_{TEF1} -maoA(V5-HIS tag)
pCS886	P_{TEF1} -TyrH, P_{TEF1} -hGTPCHI
pCS887	P_{TEF1} -hTH2, P_{TEF1} -hGTPCHI
pCS1050	P_{TEF1} -hMaoAopt(V5-HIS tag)
pCS1227	P_{TEF1} -yARO8
pCS1480	P_{TEF6} -yCYP2D6
pCS1482	P_{TEF1} -hMaoAopt-E34(I), P_{TEF1} -TfNCS Δ 29-R34
pCS1537	P_{TEF1} -MlmaoAopt
pCS1546	P_{TEF1} -TYDC1
pCS1565	P_{TEF6} -yCYP2D6:yCPR Δ 33 fusion
pCS1591	P_{TEF1} -hMaoAopt, P_{TEF1} -TfNCS Δ 29

Yeast strains	Plasmid(s)	Integrated constructs	Plasmid-based constructs
CSY194		$P_{GAL1-10}$ -yCPR1	
CSY195		$P_{GAL1-10}$ -AtATR1	
CSY88	pCS221, pCS251		P_{TEF1} -TYDC2, P_{TEF1} -AbPP02
CSY151		wild-type	
CSY152		<i>ald2</i> Δ	
CSY153		<i>ald3</i> Δ	
CSY154		<i>ald4</i> Δ	
CSY155		<i>ald5</i> Δ	
CSY156		<i>ald6</i> Δ	
CSY417		<i>adh1</i> Δ	
CSY418		<i>adh2</i> Δ	
CSY419		<i>adh3</i> Δ	
CSY420		<i>adh4</i> Δ	
CSY421		<i>adh5</i> Δ	
CSY422		<i>adh6</i> Δ	
CSY423		<i>adh7</i> Δ	

3.4.2. Analysis of metabolite production

We evaluated BIA metabolite levels by LC-MS/MS analysis of cell extracts and growth media. At appropriate time points, aliquots of yeast cultures were centrifuged to recover cells as pellets and allow collection of the growth media. We analyzed the growth media or an appropriate dilution directly by LC-MS/MS. Samples were run on an Agilent ZORBAX SB-Aq 3 x 250 mm, 5 μ m column with 0.1% acetic acid as solvent A and methanol as solvent B. We used a gradient elution to separate the metabolites of interest as follows: 0-10 min at 100% A, 10-30 min 0-90% B, 30-35 min 90-0% B, followed by a

5 min equilibration at 100% A between samples. Following LC separation, metabolites were injected into an Agilent 6320 ion trap MSD for detection and identification. We verified chromatogram data through at least three independent experiments and from multiple strains where appropriate. Quantification of metabolites was based on the integrated area of the extracted ion chromatogram peaks⁶⁴ calculated using DataAnalysis for 6300 Series Ion Trap LC/MS Version 3.4 (Bruker Daltonik GmbH) and reported as the mean \pm s.d. When appropriate, we normalized the measured levels to a metabolite peak of known concentration in the growth media.

3.4.3. Analysis of protein levels through Western blotting

We constructed plasmids for Western blotting experiments by cloning the C-terminal epitope tag(s) from pYES-NT/A (Invitrogen) into our standard TEF1 expression vector followed by subcloning of the enzyme of interest. For the AbPPO2 construct with the N-terminal tag, we cloned this sequence into the original GAL-inducible pYES-NT/A vector and detected with the Anti-His G-HRP antibody. We transformed individual plasmids into wild-type yeast cells using a standard lithium acetate protocol. Overnight cultures were grown and backdiluted 1:100 into 100 ml cultures. Cells were grown to $OD_{600} \sim 1.5$ and pelleted by centrifugation. The media was removed and cells were washed in 1 ml PBS, pelleted, and resuspended in 0.5 ml Y-PER plus HALT protease inhibitor (Pierce). Cells were vortexed for ~ 20 min and the lysate separated by centrifugation. We estimated total protein using the Coomassie Plus Protein Assay Reagent (Pierce) and loaded $\sim 50 \mu\text{g}$ of each sample onto a protein gel. We used NuPage 4-12% Bis-Tris gels with MES running buffer and transfer buffer according to the manufacturer's instructions (Invitrogen). Proteins were blotted onto a nitrocellulose

membrane (Whatman) using a semi-dry transfer cell (Bio-Rad) for 25 min at 25 V. We incubated the membrane with the Anti-V5 HRP antibody (Invitrogen) according to the manufacturer's instructions (Invitrogen) with 5% BSA as the blocking agent. Proteins were detected with the West Pico Super Signal Detection kit (Pierce) and imaged on a ChemiDoc XRS system (Bio-Rad).

3.4.4. Monoamine oxidase and NCS *in vitro* assays

For the monoamine oxidase H₂O₂ assays, yeast cells were grown in 50-100 ml cultures and lysed with 0.5 ml Y-PER plus HALT protease inhibitor (Pierce). Cell pellets were resuspended in 0.5 ml PBS, and total protein for each sample was estimated using the Coomassie Plus Protein Assay Reagent (Pierce) to normalize activity. Standard curves were generated and plate-based assays performed using the Fluoro H₂O₂ kit (Cell Technology, Inc.) according to the manufacturer's instructions. The activity was compared to a negative control lacking a monoamine oxidase sequence (this value was subtracted from the data presented).

For the monoamine oxidase *in vitro* reactions, *E. coli* cells were used to express the recombinant proteins since *maoA* can be recovered in the soluble lysate of bacterial samples whereas most protein associates with the cell pellet in yeast. *E. coli* cells expressing *maoA* or NCS from the *P_{trc}* promoter were induced with 1 mM IPTG and grown for ~12 hr in 10 ml cultures. The cells were harvested by centrifugation and lysed with B-PER plus HALT (Pierce). Soluble lysates were desalted using NAP columns and resuspended in the appropriate reaction buffer. For *maoA* assays, the reaction consisted of 0.1 M potassium phosphate, pH 7.0, 10 mM tyramine, and 100 μ l *maoA* lysate in a

total volume of 1 ml incubated at 37°C for 1 hr. For NCS assays, the reaction consisted of 0.1 M Tris-HCL, pH 7.5, 10 mM dopamine, 10 µl maoA reaction, and 50 µl NCS lysate in a total volume of 100 µl incubated at 37°C for 1.5 hr. For the maoA reaction, purified tyramine oxidase from *Arthrobacter sp.* (Sigma) was used as a positive control in the described maoA reaction mixture and also under the recommended conditions.

**CHAPTER IV: PRODUCTION OF (*R*, *S*)-RETICULINE AND
DOWNSTREAM BENZYLISOQUINOLINE ALKALOIDS IN
*SACCHAROMYCES CEREVISIAE***

Adapted from Hawkins, K.M. & Smolke, C.D. *Nat Chem Biol* **4**, 564-573 (2008)

Abstract

The benzyloquinoline alkaloids (BIAs) are a diverse class of plant secondary metabolites that exhibit a broad range of pharmacological activities and are synthesized through biosynthetic pathways that exhibit complex enzyme activities and regulatory strategies. We have engineered yeast to produce the major branch point intermediate reticuline and downstream BIA metabolites from a commercially available substrate. An enzyme tuning strategy was implemented that identified activity differences between variants and determined optimal expression levels. By synthesizing both stereoisomer forms of reticuline and integrating enzyme activities from three plant sources and humans, we demonstrated the synthesis of metabolites in the sanguinarine/berberine and morphinan branches. We also demonstrated that a human P450 enzyme exhibits a novel activity in the conversion of (*R*)-reticuline to the morphinan alkaloid salutaridine. Our engineered microbial hosts offer access to a rich group of BIA molecules and associated activities that will be further expanded through synthetic chemistry and biology approaches.

4.1. Introduction

We have engineered yeast strains expressing combinations of enzymes from three plant sources and humans as microbial hosts for the production of a wide array of BIA metabolites. In particular, we examined the ability of different combinations of three recombinant enzymes from *Thalictrum flavum* and *Papaver somniferum* to produce the early BIA metabolite reticuline from norlaudanosoline. In addition, we described a novel enzyme tuning strategy that can be generally applied to determine optimal enzyme expression levels to conserve cellular resources and improve growth and production rates without compromising pathway flux. These studies demonstrated that reticuline production levels varied from ~10 to 150 mg l⁻¹ depending on the combination of enzyme variants used, highlighting the differences in activities and interactions among pathway variants in the engineered host. We also engineered yeast strains that produce BIA metabolites along two of the major branches from reticuline: the sanguinarine/berberine branch and the morphinan branch. Expression of three downstream enzymes from *T. flavum* and *P. somniferum* and a reductase partner from *Arabidopsis thaliana* resulted in the synthesis of (*S*)-scoulerine, (*S*)-tetrahydrocolumbamine, and (*S*)-tetrahydroberberine from (*S*)-reticuline. In addition, expression of a human P450 enzyme and its native reductase partner resulted in the synthesis of salutaridine from (*R*)-reticuline, demonstrating a novel activity for this P450 enzyme. As the enzymatic activities leading to salutaridine have not yet been cloned and characterized from plant hosts⁶⁵, our synthetic pathway highlights the diversity of BIA products that can be synthesized in microbial hosts by combining activities from diverse sources.

4.2. Results

4.2.1. Synthesis of reticuline from norlaudanosoline in yeast

In native plant hosts such as *T. flavum* and *P. somniferum*, the first committed step in BIA biosynthesis is the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPA) catalyzed by norcoclaurine synthase (NCS) to produce (*S*)-norcoclaurine (Fig. 4.1). This natural intermediate undergoes a series of methylation reactions catalyzed by norcoclaurine 6-*O*-methyltransferase (6OMT), coclaurine-*N*-methyltransferase (CNMT), 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase (4'OMT) and a hydroxylation reaction catalyzed by cytochrome P450 80B1 (CYP80B1) to produce the major branch point intermediate (*S*)-reticuline. (*S*)-Reticuline is converted to downstream metabolites along various branches, ultimately resulting in the synthesis of pharmacologically relevant molecules such as morphine and berberine.

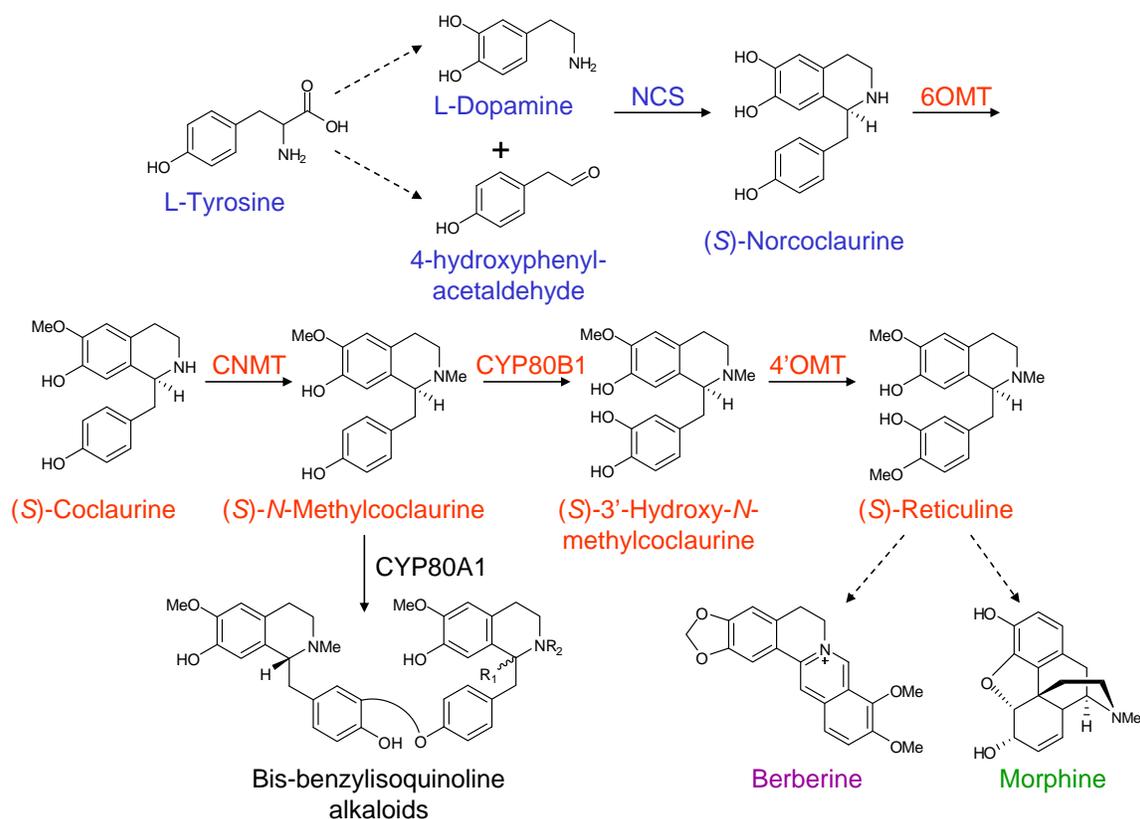


Fig. 4.1. The native BIA pathway. Color schemes for the metabolites and enzymes are as follows: upstream portion of the pathway to the first BIA backbone molecule norcoclaurine, blue; middle portion of the pathway from norcoclaurine to the branch point metabolite reticuline, red; sanguinarine/berberine branch, purple; morphine branch, green. Conversion steps for which the entire set of enzymes has not been completely elucidated or cloned are indicated by dashed arrows.

The total synthesis of the BIA backbone from tyrosine represents a significant engineering challenge as many of the plant enzymes that perform the early conversion reactions have not yet been isolated and cloned. As an alternative route, we have constructed a synthetic BIA pathway in *Saccharomyces cerevisiae* that converts the commercially available substrate (*R, S*)-norlaudanosoline to (*R, S*)-reticuline (Fig. 4.2). Norlaudanosoline differs from the natural substrate norcoclaurine by the presence of a 3'-OH group that is added to the BIA backbone by CYP80B1 in the native pathway. Therefore, our synthetic BIA pathway is comprised of three heterologous AdoMet-

dependent methyltransferase enzymes 6OMT, CNMT, and 4'OMT that convert norlaudanosoline to reticuline. Norlaudanosoline was preferred over dopamine as the starting substrate in this work as this initial conversion step proved to be extremely inefficient in yeast cells expressing either the *E. coli* or human monoamine oxidase enzyme variant with or without the *T. flavum* NCS, requiring fed dopamine concentrations of ~100 mM (data not shown).

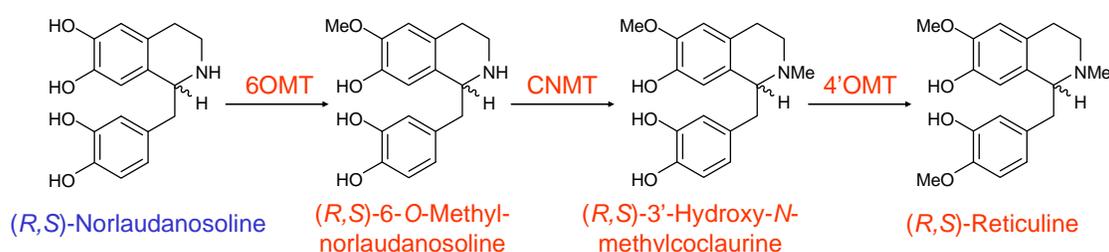


Fig. 4.2. The engineered BIA pathway for the synthesis of (*R*, *S*)-reticuline from (*R*, *S*)-norlaudanosoline. The fed substrate, (*R*, *S*)-norlaudanosoline, is indicated in blue.

Yeast strains were engineered to express one or more of the heterologous BIA pathway enzymes from *P. somniferum* and *T. flavum*. We constructed yeast BIA expression vectors, where each construct enabled the expression of one or two enzymes from constitutive TEF1 promoters (Fig. 4.3). We also constructed single-gene expression plasmids to characterize each enzyme variant individually. The resulting engineered yeast strains were grown in the presence of norlaudanosoline (or appropriate substrate) and assayed for the expected product(s).

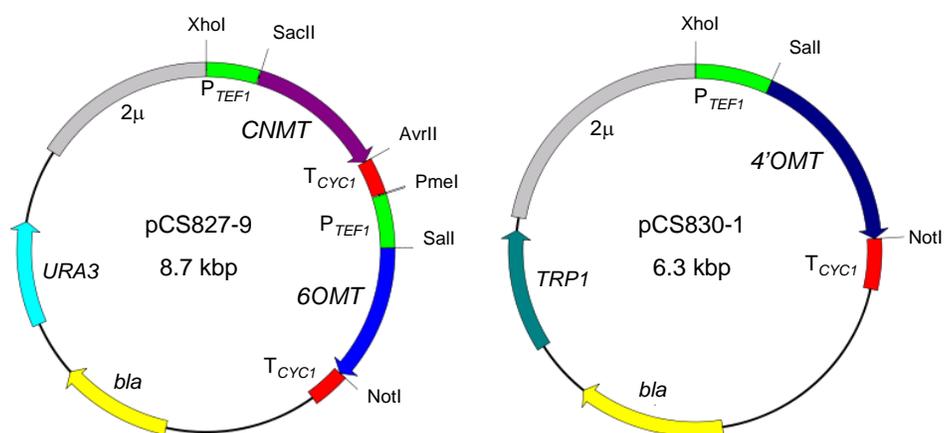


Fig. 4.3. Plasmid maps of BIA expression constructs. Heterologous cDNA sequences are expressed in yeast from constitutive TEF1 promoters on high-copy plasmids with different selection markers. The expression systems are used for the combinatorial expression of 6OMT, CNMT, and 4'OMT enzyme variants.

We tested the 6OMT and 4'OMT activities using norlaudanosoline and laudanosoline as substrates and the CNMT activities using norlaudanosoline and 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline as substrates (Fig. 4.4). As standards are not commercially available for the metabolites of interest, following separation by high performance liquid chromatography, positive identification of BIAs was confirmed using selective reaction monitoring and tandem mass spectrometry (LC-MS/MS) in which the resulting ion fragments are characteristic for a specific molecular structure. Enzymes from both plant species performed the expected methylation reactions on the provided substrate(s) when expressed individually, and yeast strains lacking the heterologous coding sequences were not able to methylate the examined substrates. No differences were observed in the relative activities of the 6OMT and the 4'OMT variants which methylated both norlaudanosoline and laudanosoline in the expected positions. However, the *P. somniferum* CNMT variant methylated significantly more of the norlaudanosoline substrate, accumulating over six times the amount of laudanosoline compared to the *T.*

flavum variant. Conversely, the *T. flavum* variant produced ~20 times the amount of *N*-methylated 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline, indicating the different substrate preferences of these orthologous proteins.

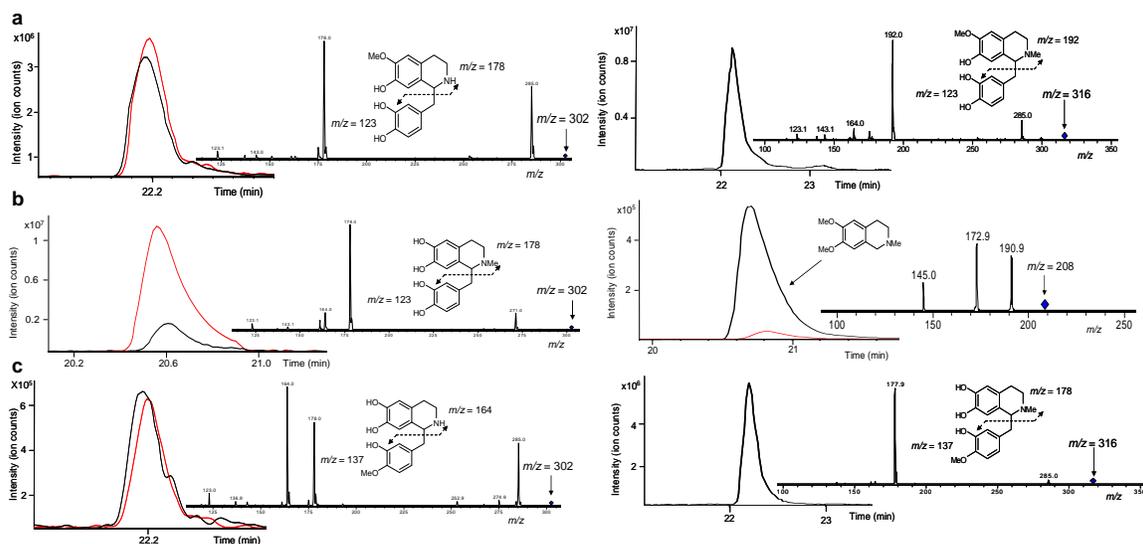


Fig. 4.4. LC-MS/MS analysis confirms individual methyltransferase activities. The growth media of engineered yeast strains supplemented with 1 mM fed substrate and grown for 48 h was analyzed for product formation. **(a)** Characterization of 6OMT fed norlaudanansoline (left) and laudanansoline (right). Left, extracted ion chromatograms for $m/z = 302$ are shown for CSY450 (red) and CSY451 (black). Right, extracted ion chromatogram for $m/z = 316$ is shown for CSY450 and is representative of CSY451. **(b)** Characterization of CNMT fed norlaudanansoline (left) and 6,7-dimethoxy-1,2,3,4-tetrahydroisoquinoline (right). Left, extracted ion chromatograms for $m/z = 302$ are shown for CSY452 (red) and CSY453 (black). Right, extracted ion chromatograms for $m/z = 208$ are shown for CSY452 (red) and CSY453 (black). **(c)** Characterization of 4'OMT fed norlaudanansoline (left) and laudanansoline (right). Left, extracted ion chromatogram for $m/z = 302$ are shown for CSY454 (red) and CSY455 (black). The 4'OMT reaction product ($m/z = 302$) appears to contain a mixture of norlaudanansoline methylated at both the 6-OH and 4'-OH positions with major fragments of $m/z = 164$ and 178 observed. The unexpected $m/z = 178$ ion is believed to be due to a co-eluting impurity in the norlaudanansoline with the parent ion $m/z = 314$. Right, extracted ion chromatogram for $m/z = 316$ is shown for CSY454 and is representative of CSY455. The fragmentation pattern of the 4'OMT reaction product does not show evidence of non-specific methylation.

We confirmed full-length expression of all methyltransferase variants by Western blotting using C-terminally tagged constructs (Fig. 4.5). Protein expression levels from

high-copy plasmids were within 2 to 3-fold of that observed from a highly-expressed yeast-enhanced GFP variant, indicating sufficient expression of the methyltransferase enzymes in the microbial host. In addition, the enzyme variants were present at similar levels such that any observed activity differences between variants cannot be attributed to differences in expression but rather some inherent property of the enzymes. The results also suggested that both CNMT variants were translated more efficiently and/or have a longer half-life in yeast compared to the 6OMT and 4'OMT enzymes.

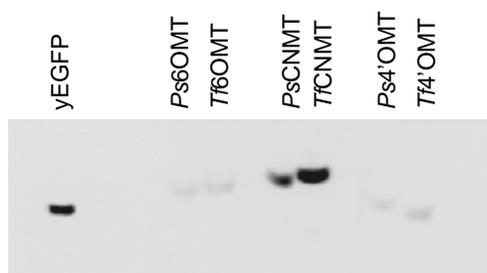


Fig. 4.5. Western blot analysis confirms protein expression of 6OMT, CNMT, and 4'OMT enzymes. Enzyme levels are comparable to levels of a yeast-enhanced GFP variant (yEGFP) expressed from the same construct (within ~2-3-fold signal). All proteins were expressed behind the TEF1 promoter on high-copy plasmids with a C-terminal V5 epitope tag. The blot was incubated with the Anti-V5 HRP antibody and a chemiluminescent assay was used for signal detection. Each lane was loaded with ~50 mg total protein. The CNMT variants are the most highly expressed and/or stable of the recombinant enzymes. At longer exposure times, degradation products were visible from the *Tf*CNMT, indicating that this protein may be subject to proteolysis in yeast, although the full-length protein still dominates.

We co-transformed yeast cells with all possible combinations of the 6OMT, CNMT, and 4'OMT enzymes from *T. flavum* and *P. somniferum* (Table 4.1). All strains harboring different combinations of the three methyltransferase enzyme variants demonstrated production of reticuline from norlaudanosoline as verified through LC-MS/MS (Fig. 4.6a). Certain enzyme combinations exhibited greater reticuline accumulation as estimated by percent conversion or the ratio of the extracted ion

chromatogram peak areas of reticuline to norlaudanosoline (Fig. 4.6b). Notably all but one strain containing the *T. flavum* CNMT variant produced ~60% less reticuline than the strains with the *P. somniferum* CNMT variant, indicating that the latter variant demonstrated higher activity in this synthetic pathway. This activity difference can be attributed to differences in substrate affinities of the CNMT variants observed in the single enzyme studies as the *P. somniferum* CNMT exhibited higher activity on the pathway substrate norlaudanosoline (Fig. 4.4). Measurements of 5-10% conversion of norlaudanosoline agreed with other estimates of reticuline concentration of ~100 μM (from 1 mM norlaudanosoline) based on comparative peak area analysis of structurally similar standards. Although BIA metabolites up to reticuline exist only in the (*S*)-conformation in plants, our studies demonstrated that the three methyltransferases accepted both stereoisomers as substrates equally *in vivo* to yield a racemic mixture of (*R*, *S*)-reticuline from (*R*, *S*)-norlaudanosoline^{66,67} (Fig. 4.6c).

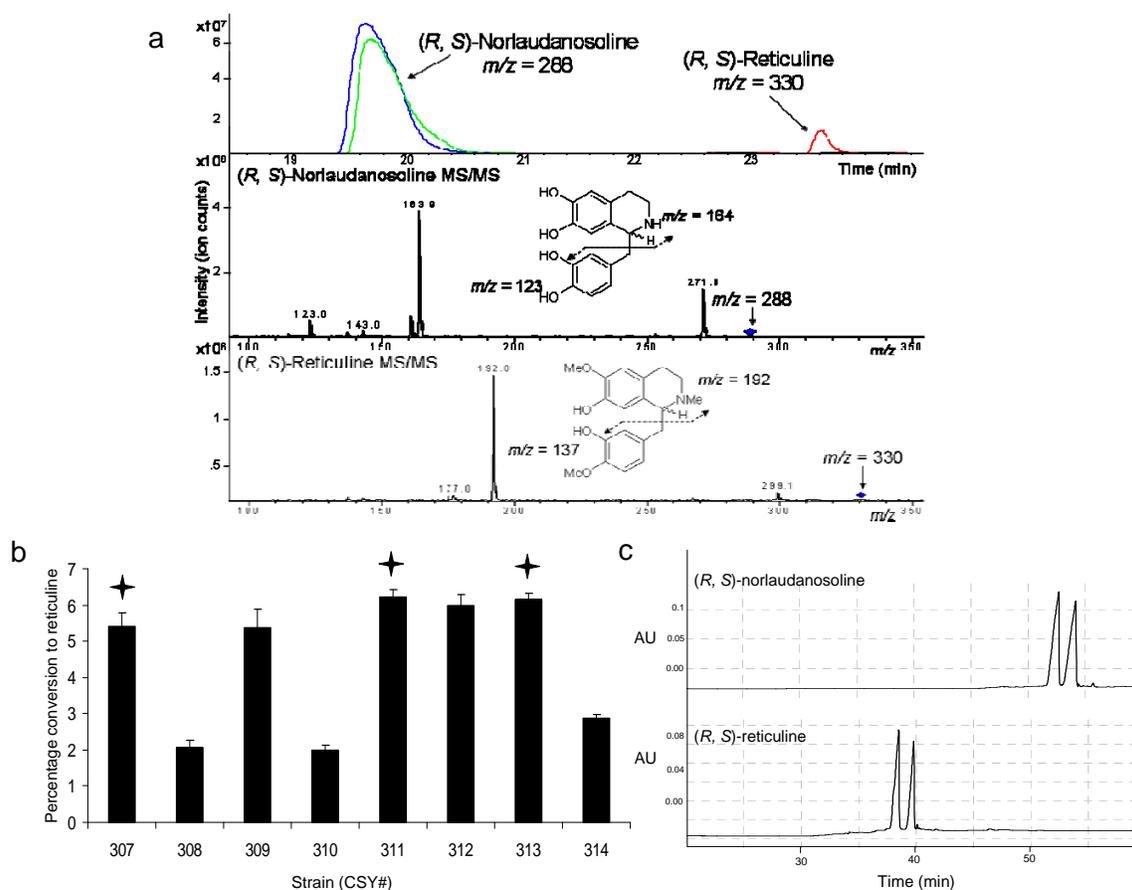


Fig. 4.6. Microbial production of (*R,S*)-reticuline. **(a)** LC-MS/MS analysis of the growth media of engineered yeast strains supplemented with 1 mM norlaudanosoline and grown for 48 h confirms reticuline production. Extracted ion chromatograms are shown for norlaudanosoline in the wild-type (blue) and CSY307 (green) strains and for reticuline in the wild-type (black, no visible peak) and CSY307 (red) strains. Reticuline is identified as the $m/z = 330$ ion eluting at 23.6 min showing the expected fragments $m/z = 192$ and 137 produced by cultures of CSY307 and similar engineered strains. Control experiments in which strains were missing any one of the three required enzymes or grown in the absence of substrate did not accumulate the metabolite peak identified as reticuline. **(b)** Reticuline production is dependent on the combination of enzyme variants expressed by the engineered yeast strains. Reticuline production is reported as percentage substrate conversion from engineered yeast strains harboring BIA expression constructs for combinatorial expression of different 6OMT, CNMT, and 4'OMT enzyme variants supplemented with 1 mM norlaudanosoline and grown for 48 h. Stars indicate enzyme combinations used in strains for stable expression. Data are reported as mean values \pm s.d. from at least three independent experiments. **(c)** Chiral analysis of (*R,S*)-norlaudanosoline and (*R,S*)-reticuline (converted by CSY288). Separation of stereoisomers was performed through capillary electrophoresis on reticuline fractions collected from the LC column.

Estimation of BIA metabolite levels in the growth media and cell extract indicated that the norlaudanosoline concentration drop across the membrane was ~10- to 30-fold and that the reticuline to norlaudanosoline ratio was slightly greater inside the cell compared to the growth media (Fig. 4.7). The results supported a passive diffusion transport mechanism in the microbial host and confirmed that the substrate was accessible to the intracellular BIA enzymes and that the synthesized metabolites accumulated in the growth media. This property greatly simplifies metabolite profiling as production levels can be estimated by direct analysis of the growth media without rigorous extraction or purification steps. However, the transport of BIA substrates across the cell membrane is somewhat limiting and highlights the importance of reconstructing this pathway in a single microbial host to avoid inefficiencies due to transport of metabolites across two cell membranes.

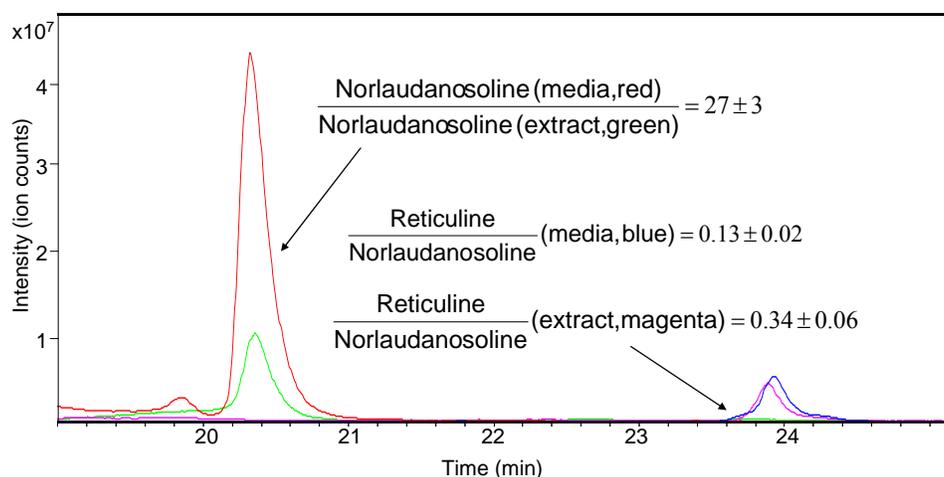


Fig. 4.7. BIA metabolites accumulate in the growth media. Analysis of norlaudanosoline (substrate) and reticuline (product) concentrations in cell extracts and growth media support a passive diffusion transport mechanism for BIA metabolites. Data shown is a 1:10 dilution of the growth media (norlaudanosoline, red; reticuline, blue) and a ~1:2 dilution of the cell extract (norlaudanosoline, green; reticuline, magenta) of CSY288 supplemented with 1 mM norlaudanosoline and grown for 48 h. Ratios of product to substrate and extracellular versus intracellular norlaudanosoline concentrations are shown as calculated by LC-MS peak areas. Chromatograms and calculations are representative of two independent experiments \pm s.d.

In an attempt to overcome transport limitations, we deleted several identified multi-drug transporters which may be actively transporting the substrate out of the cell. Using a reticuline producing background strain CSY288, total gene deletions were made for *aqr1Δ*, *flr1Δ*, *qdr1/2Δ*, *qdr3Δ* and the quintuple knockout. The *PDR1* transcription factor known to activate expression of additional proteins from this multi-drug transport family was also deleted in the CSY288 background strain. No difference in norlaudanosoline conversion was observed in any of these strain backgrounds (data not shown). However, deletion of these genes may also cause more reticuline to be retained inside the cells such that no net effect is observed in the growth media. Another approach is to express a protein which actively, and preferably selectively, transports BIA substrates into the cell but a good target candidate has not yet been identified.

4.2.2. BIA production in the yeast strains is substrate limited

We selected three combinations of the methyltransferase enzyme variants that demonstrated high conversion of norlaudanosoline to reticuline for stable expression in yeast (CSY288, CSY334, CSY456; Table 4.1). Expression of heterologous enzymes from the chromosome is anticipated to result in more consistent and stable expression over time, facilitate expansion of the synthetic pathway, and enable cultures to be grown in rich media without selective pressure. However, expression levels from chromosomal integrations are expected to be lower than those from the high-copy plasmid system due to the reduction in DNA copy number. Reticuline production levels compared favorably in all stable expression strains with only one combination of enzymes showing a decrease

in production greater than 30% upon integration of the constructs (Fig. 4.8a). The results indicated that the *P. somniferum* 6OMT variant outperformed the *T. flavum* variant at lower expression levels and exhibited higher specific activity in the synthetic pathway.

The results from the analysis of reticuline production in the stable expression strains suggested that substrate conversion was not severely hindered by enzyme expression levels. We examined the effects of substrate concentration and growth phase on reticuline production levels from cultures of CSY288. Reticuline accumulation increased roughly in proportion to the initial concentration of norlaudanosoline in the media between 0.5 to 5 mM such that percent conversion was between 7-10% in all samples, supporting a substrate limitation model (Fig. 4.8b). Substrate limitation in our synthetic BIA pathway was not unexpected since, in addition to potential transport issues, norlaudanosoline and 6-*O*-methyl norlaudanosoline are not the natural substrates for 6OMT and CNMT, respectively. For example, kinetic characterization studies on the 6OMT variant from *C. japonica* determined a K_m value of 2.23 mM for (*R*, *S*)-norlaudanosoline, with a relative activity of 76% compared to (*S*)-norcoclaurine⁶⁶. Although the 6OMT variant from *P. somniferum* used here has a reported K_m value of 10 μ M for norcoclaurine⁶⁷, the value may be significantly higher for norlaudanosoline. We also examined the time scale of production under conditions that model a batch fermentation run and observed significant reticuline production shortly after substrate addition to a shake flask culture seeded with CSY288. The production rate slowed over time, but reticuline continued to accumulate as cells entered the stationary growth phase (Fig. 4.8c).

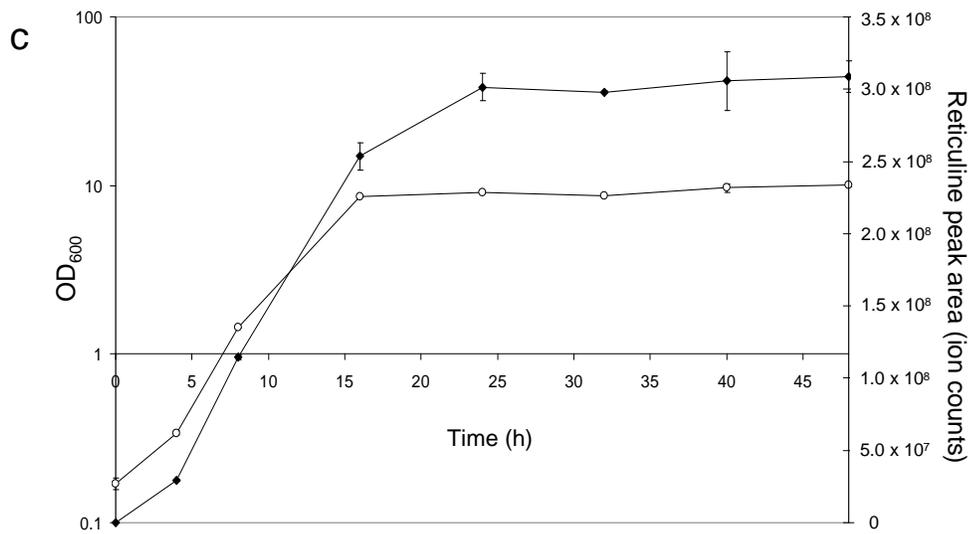
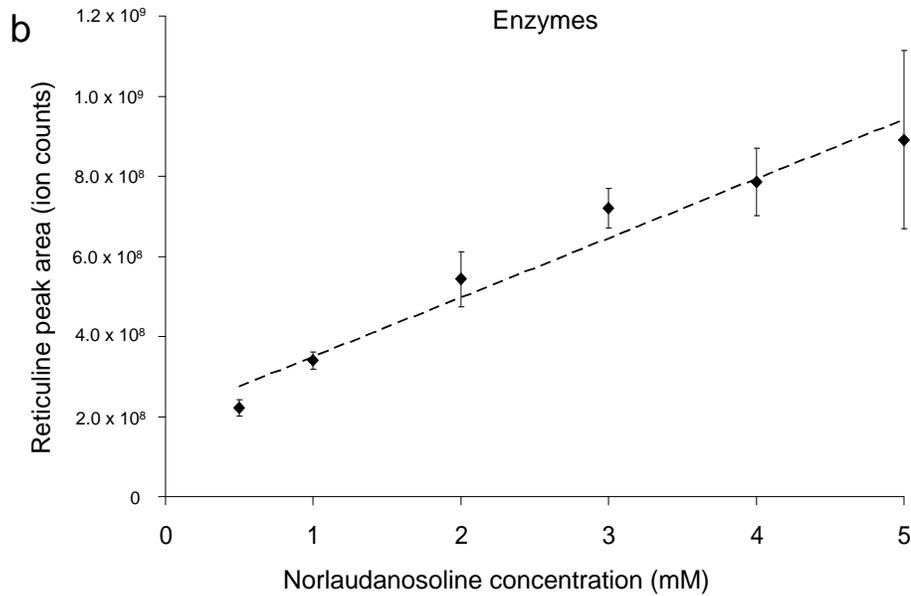
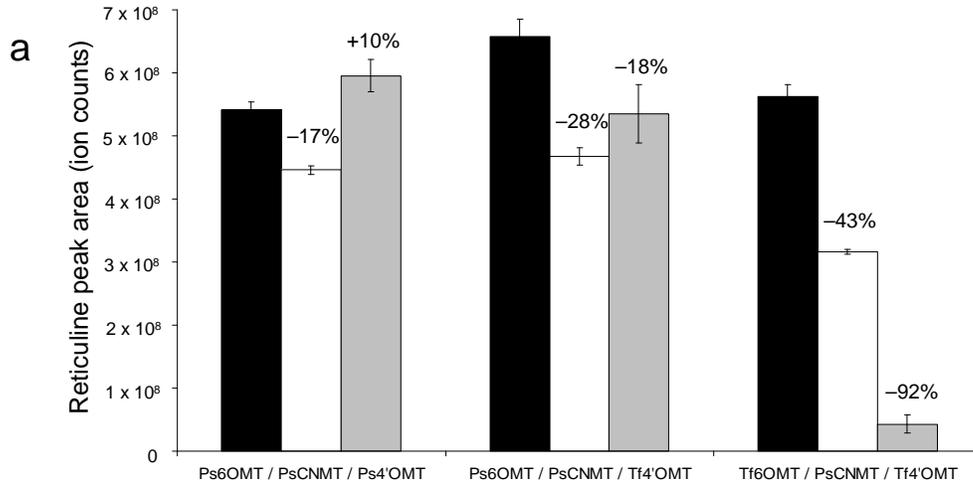


Fig. 4.8. Effects of enzyme levels, substrate levels, and culture time on reticuline production. **(a)** Reticuline production is not significantly impacted at lower enzyme expression levels. Reticuline production as measured by LC-MS peak area from three strains engineered to stably express the indicated enzyme variants from the TEF1 promoter (CSY288, CSY334, CSY456; white) or the mutant TEF7 promoter (CSY448, CSY449, and CSY458; gray) were compared to strains expressing the enzymes from plasmid-based constructs (CSY307, CSY311, CSY313; black). The percentage change in production observed from the stable strains after 48 h growth in media supplemented with 2 mM norlaudanosoline is indicated. **(b)** Reticuline production increases with substrate concentration. Reticuline production was measured in the growth media of CSY288 supplemented with a range of norlaudanosoline concentrations after 48 h growth. **(c)** Reticuline production increases with the OD₆₀₀ of the culture. Reticuline production (solid diamonds) and OD₆₀₀ (open circles) were measured at the indicated time points in the growth media of CSY288 supplemented with 2 mM norlaudanosoline added at $t = 0$. All data are reported as mean values \pm s.d. from at least three independent experiments.

4.2.3. Tuning enzyme levels with a novel titration strategy

Pathway optimization strategies often require an analysis of enzyme expression levels and their effects on metabolite accumulation and strain growth rate. Systems that enable heterologous enzymes to be expressed at a minimum level while maintaining maximum pathway flux and product accumulation are desired to avoid wasting cellular resources synthesizing excess proteins. Such strategies become more critical as an engineered pathway is extended to include additional enzymes that will further tax cellular resources. Optimization of pathway enzyme levels should result in an improved strain growth rate and a markedly reduced production time.

To optimize BIA enzyme expression levels, we developed a system that allows each enzyme to be titrated independently. We replaced the constitutive TEF1 promoter with the GAL1-10 galactose-inducible promoter for one of the three BIA enzyme coding sequences in our stable strains to allow tunable expression of one heterologous enzyme while holding the other two constant. In addition, we deleted the GAL2 permease from

each strain to enable titratable control over enzyme levels. Deletion of the *GAL2* permease has been previously shown to result in a more homogenous, linear induction response from the GAL network relative to the wild type all-or-none, switch-like response³. Our system design resulted in a series of engineered strains stably expressing the BIA enzymes in which levels of one of the three methyltransferases were precisely regulated according to galactose concentration (CSY325-329; Table 4.1).

We used the tunable strains to determine relationships between galactose concentration, enzyme expression levels, and reticuline production. The titratable yeast strains were grown in 1 mM norlaudanosoline and varying galactose concentrations, and reticuline production was analyzed at 24 h after substrate addition. The enzyme titration experiments demonstrated that in nearly all strains at ~0.5% galactose, conversion of norlaudanosoline to reticuline reached over 70% of the levels attained from fully induced conditions at 2% galactose (Fig. 4.9a). All strains demonstrated production comparable to the parent strains at maximum induction levels. Only the strain containing the GAL-inducible *T. flavum* 6OMT (CSY329; Table 4.1) did not show changes in production corresponding to galactose concentration as reticuline production was similarly low at all points, suggesting that higher expression levels were necessary to observe significant activity of this variant in our system (data not shown).

The relationship between galactose concentration and relative expression levels was obtained indirectly using an analogous fluorescent reporter system. We constructed integration cassettes containing a green fluorescent protein (yEGFP3) under the control of the constitutive *TEF1* promoter or the inducible *GAL1-10* promoter in a *GAL2Δ* background to make strains CSY428 and CSY429 (Table 4.1). We compared the relative

fluorescence levels from the GAL1-10 promoter at various galactose concentrations to fluorescence levels observed from the TEF1 promoter. The reporter protein titration experiments demonstrated that the expression from the tunable GAL promoter system at 0.5% galactose translated to ~16% expression relative to the native TEF1 promoter (Fig. 4.9b). The results indicated that the transcriptional activity of the heterologous enzyme promoters could be significantly reduced while maintaining maximal substrate conversion to reticuline.

We optimized BIA enzyme expression levels based on our titration assay results by designing integration cassettes in which the level of expression from the promoter system was minimized without compromising production. We used a mutated TEF promoter library⁵ in which TEF1 promoter variants with altered levels of gene expression had been generated and characterized. Our titration assays indicated that optimal expression levels for each of the BIA enzymes tested, with the possible exception of *P. somniferum* 4'OMT, corresponded to ~16% of the native TEF1 promoter, approximating that of the TEF7 mutant. We replaced the TEF1 promoter with the TEF7 promoter in our chromosomal integration cassettes and constructed optimized reticuline-producing strains. TEF7-substituted versions of CSY288 and CSY334 (CSY448, CSY449; Table 4.1) showed comparable production levels (Fig. 4.8a). However, CSY458 showed greatly reduced production compared to CSY456 and other TEF7-substituted strains. Since the only difference between CSY458 and CSY449 is the 6OMT variant, the data further supported that this step can become limiting at low expression levels and that the *P. somniferum* 6OMT exhibited higher specific activity than the *T. flavum* variant in the synthetic pathway. We verified trends in relative transcript levels between the TEF1 and

TEF7 integrated expression systems and the high copy plasmid-based system by qRT-PCR (Fig. 4.9c). Growth rates were not significantly and reproducibly different between the TEF1 and TEF7 strains; however, this optimization strategy, which resulted in reduced metabolic load on the cell without compromising reticuline production levels, will likely prove more important upon further extensions of the pathway. In addition, this tuning strategy can be generally applied to other recombinant pathways in yeast to determine and set the minimal expression level of heterologous enzymes for a desired product yield.

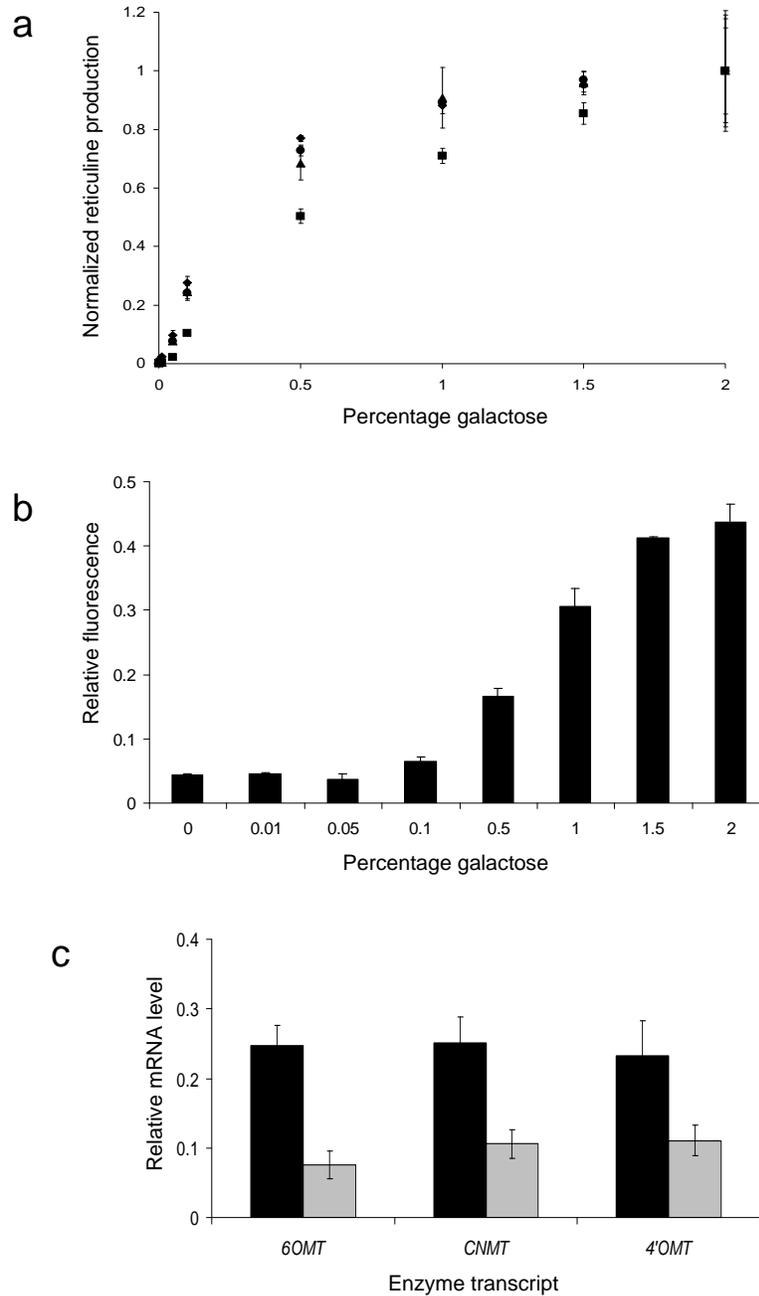


Fig. 4.9. A novel strategy for tuning enzyme expression levels. **(a)** Reticuline production as a function of galactose concentration in engineered reticuline-producing strains expressing one heterologous enzyme from the titratable GAL expression system; production is normalized to each strain in 2% galactose. CSY325, diamonds; CSY326, squares; CSY327, circles; CSY328, triangles. Cells were grown overnight in noninducing-nonrepressing media and backdiluted into media containing the indicated galactose concentration. Following 4 h of induction, 1 mM norlaundanosoline was added and supernatants were analyzed for reticuline production after 24 h. **(b)** GFP reporter strains were used to guide the determination of relative protein expression levels for the enzyme titration studies. Relative fluorescence data is reported for CSY429 normalized to

CSY428. Cells were grown overnight in noninducing-nonrepressing media and backdiluted into media containing the indicated galactose concentration. Fluorescence measurements were normalized to OD₆₀₀ for each sample and taken during the exponential growth phase. (c) qRT-PCR analysis confirms trends in relative transcript levels from the TEF1 and TEF7 integrated expression systems compared to the high-copy plasmid-based systems. Transcript levels in representative stable strains using the TEF1 promoter (CSY288; black) and the TEF7 promoter (CSY448; gray) are normalized to levels from a high-copy plasmid expression system (CSY307). All data are reported as mean values \pm s.d. from at least two independent experiments.

4.2.4. *Synthesis of sanguinarine/berberine intermediates*

Reticuline represents a major branch point intermediate in plant secondary metabolism from which a wide variety of BIA metabolites can be derived, including sanguinarine, berberine, protopine, magnoflorine, and morphinan alkaloids. The enzymes in these downstream BIA branches are of particular interest as many exhibit complex biosynthetic activities. We examined the synthesis of metabolites along a major branch, the sanguinarine/berberine branch, by expressing additional enzymes from the native plant pathways as a demonstration of the diverse natural products and activities that can be produced in our engineered yeast strains. This family of benzophenanthridine alkaloids has generated interest for their pharmacological activities, most notably as antimicrobial agents.

The first conversion step in the sanguinarine/berberine branch of the BIA pathway is performed by the berberine bridge enzyme (BBE), which catalyzes the oxidative cyclization of the N-methyl moiety of (*S*)-reticuline into the berberine bridge carbon of (*S*)-scoulerine⁶⁸ (Fig. 4.10a). This unique reaction forms the protoberberine carbon skeleton via a methylene iminium ion intermediate and cannot be replicated through synthetic chemistry approaches. We constructed a plasmid expressing the *P. somniferum* BBE cDNA for transformation into our reticuline-producing yeast strains CSY288 and

CSY334 (CSY336, CSY338; Table 4.1). LC-MS/MS analysis demonstrated that the engineered strains produce (*S*)-scoulerine (Fig. 4.10b). The relative peak areas of the appropriate ions indicated ~40% conversion of (*R, S*)-reticuline to (*S*)-scoulerine with an effective conversion of ~80% based on an equal mixture of stereoisomers. In addition, we also tested enzyme truncations of the first 25 and 41 amino acids to remove the N-terminal signal sequence⁶⁹. Neither truncated variant demonstrated increased production of (*S*)-scoulerine, and the Δ 41 truncation showed slightly compromised production (Fig. 4.11).

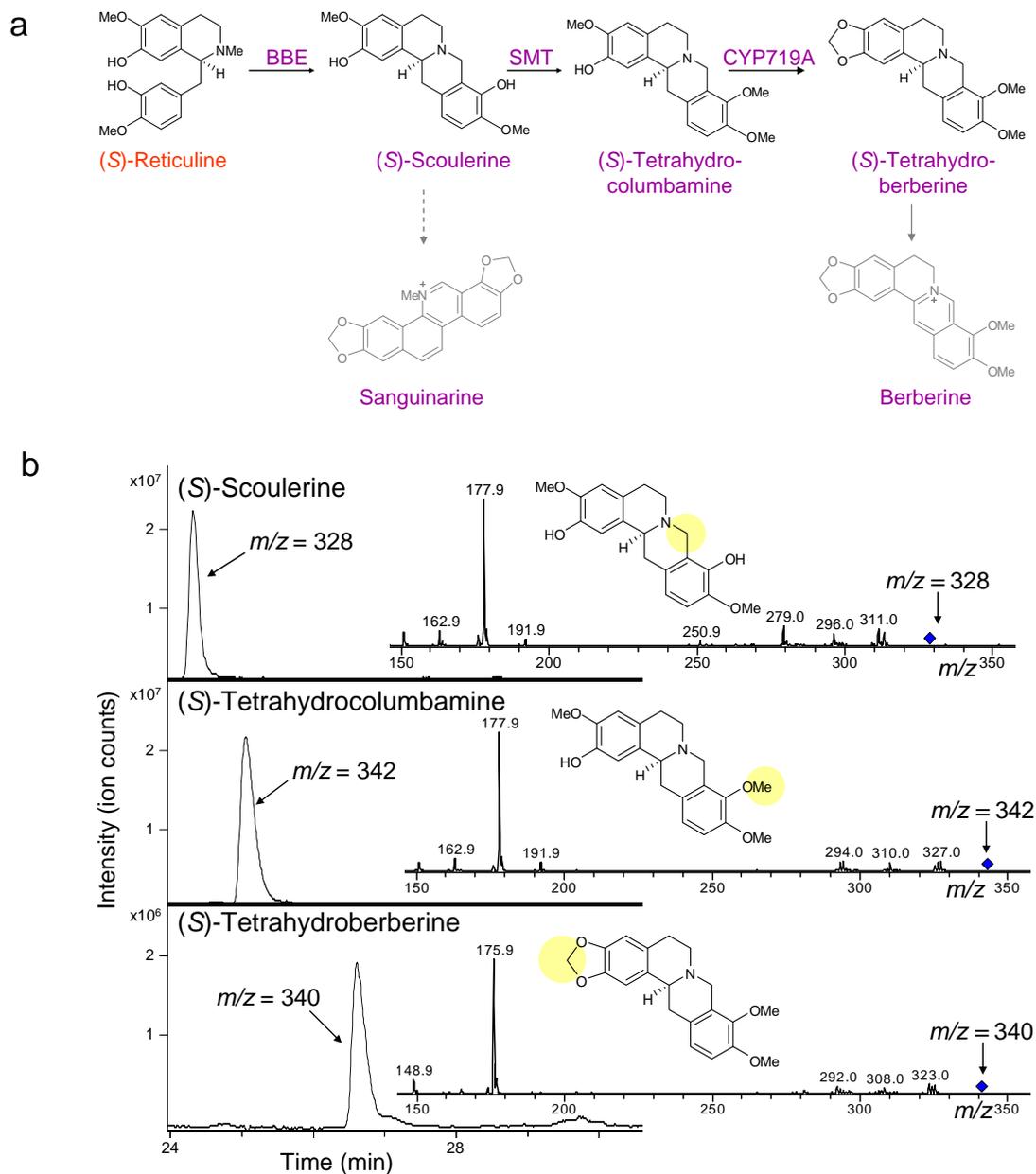


Fig. 4.10. Microbial production of BIA metabolites along the sanguinarine and berberine branches. **(a)** Pathway for the synthesis of sanguinarine/berberine metabolites from the key intermediate (*S*)-reticuline. Color schemes follow that defined in Fig. 4.1. Shaded portions of the pathway indicate steps not reconstructed in the engineered microbial host. Dashed arrows represent multiple steps and solid arrows where no enzyme is listed indicated a single step for which the enzyme has not been cloned from plant hosts. **(b)** LC-MS/MS analysis of the growth media of engineered yeast strains supplemented with 4 mM norlaudanosoline and grown for 48 h confirms (*S*)-scoulerine, (*S*)-tetrahydrocolumbamine, and (*S*)-tetrahydroberberine production. Data is shown for CSY336, CSY337, and CSY410, respectively, and is representative of analogous strains. The fragmentation pattern of the 328 ion corresponding to (*S*)-scoulerine differs

significantly from that of (*S*)-reticuline as the formation of the berberine bridge stabilizes (*S*)-scoulerine so that it does not fragment into the benzyl and isoquinoline moieties, but instead loses methyl and hydroxyl groups. The fragmentation pattern of the 342 ion identified as (*S*)-tetrahydrocolumbamine is consistent with that of the parent molecule (*S*)-scoulerine; the observed fragments exhibit an increase ($m/z = 14$) attributable to the additional methyl group. (*S*)-Tetrahydroberberine is identified by its major ion ($m/z = 340$) and exhibits a fragmentation pattern similar to the parent molecule (*S*)-tetrahydrocolumbamine; the observed fragments show a decrease ($m/z = 2$) consistent with the formation of the methylenedioxy bridge. Growth media of strains lacking the required enzyme coding sequence(s) did not display the identified metabolite peak(s).

(*S*)-Scoulerine represents a second important branch point metabolite as it can be converted to BIA metabolites along either the sanguinarine or berberine branches (Fig. 4.10a). The oxidation of (*S*)-scoulerine to (*S*)-cheilanthifoline, an intermediate metabolite along the sanguinarine branch, is catalyzed by a cytochrome P450 enzyme that has not yet been cloned from native plant hosts. Alternatively, the methylation of (*S*)-scoulerine in the 9-OH position by (*S*)-scoulerine 9-*O*-methyltransferase (SMT) results in the synthesis of (*S*)-tetrahydrocolumbamine, an intermediate metabolite along the berberine branch. We constructed a plasmid co-expressing the *P. somniferum* *BBE* and *T. flavum* *SMT* cDNAs for transformation into our reticuline-producing yeast strains (CSY337, CSY339; Table 4.1). LC-MS/MS analysis demonstrated that the heterologous SMT enzyme performed the expected methylation reaction to produce (*S*)-tetrahydrocolumbamine (Fig. 4.10b). Yeast strains producing (*S*)-tetrahydrocolumbamine from (*R, S*)-norlaudanosoline exhibited little or no accumulation of (*S*)-scoulerine, indicating efficient conversion of the substrate and yielding $\sim 60 \text{ mg l}^{-1}$ (*S*)-tetrahydrocolumbamine from 4 mM (*R, S*)-norlaudanosoline.

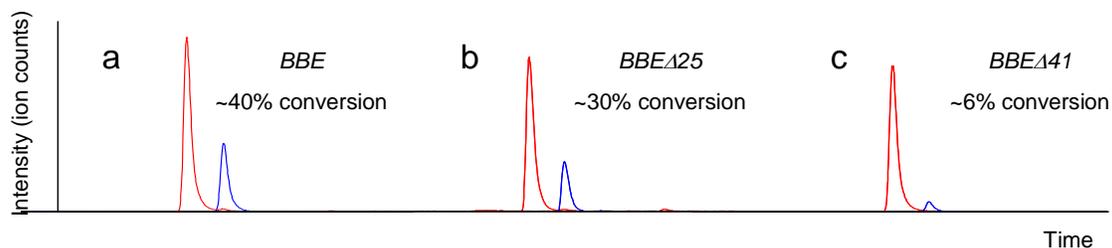


Fig. 4.11. Comparison of BBE N-terminal truncations. Extracted ion chromatograms are shown for reticuline ($m/z = 330$, red) and scoulerine ($m/z = 328$, blue). Percentage conversion is calculated as the ratio of scoulerine to reticuline measured in the growth media and values for each BBE variant are shown.

The next and penultimate metabolite along the berberine branch is (*S*)-tetrahydroberberine, or (*S*)-canadine, which is produced by a methylenedioxy bridge-forming reaction from (*S*)-tetrahydrocolumbamine (Fig. 4.10a). This reaction is catalyzed by canadine synthase, a cytochrome P450 enzyme (CYP719A1), which has been cloned from multiple berberine-producing plants⁷⁰. Previous characterization studies on the *C. japonica* CYP719A1 have been performed in a yeast strain that co-expresses the *Arabidopsis thaliana* P450 reductase ATR1⁷⁰. We therefore constructed a dual-expression plasmid containing the *T. flavum* CYP719A1 and SMT cDNAs for co-transformation with plasmids containing the *A. thaliana* ATR1 and *P. somniferum* BBE cDNAs into our reticuline-producing yeast strains (CSY399, CSY400; Table 4.1). LC-MS/MS analysis demonstrated low levels of (*S*)-tetrahydroberberine production ($\ll 5 \text{ mg l}^{-1}$) from a starting substrate concentration of 4 mM norlaudanosoline. Production of (*S*)-tetrahydroberberine was confirmed not only by its characteristic fragmentation pattern but also by co-elution experiments with authentic (*R, S*)-tetrahydroberberine (Fig. 4.12).

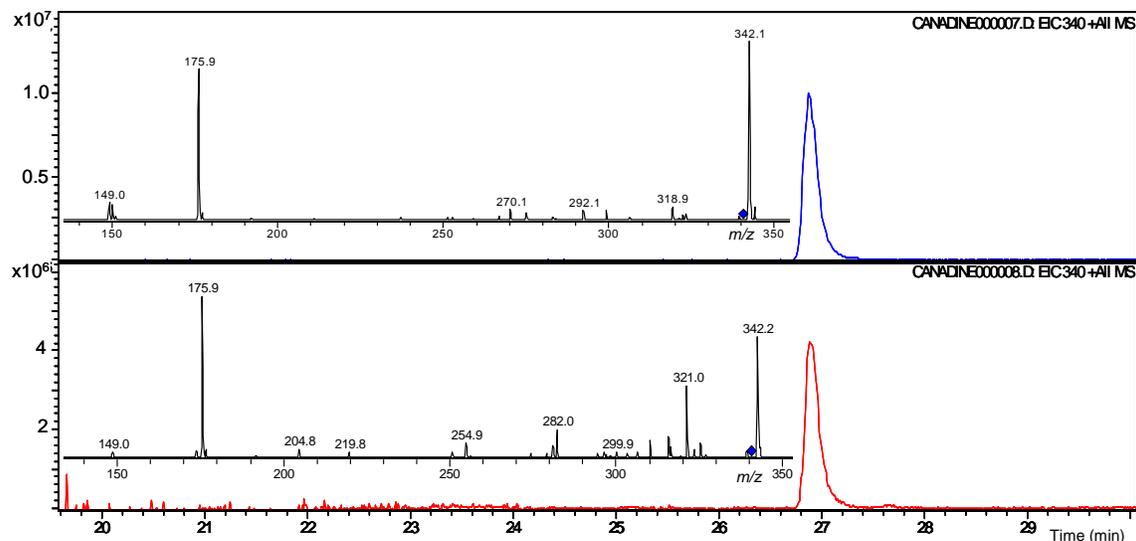


Fig. 4.12. Comparison of synthesized (*S*)-tetrahydroberberine to standard. Authentic DL-canadin or (*R*, *S*)-tetrahydroberberine (Apin Chemicals Ltd) co-elutes with the product identified as (*S*)-tetrahydroberberine synthesized by engineered yeast cells at 26.9 min using the method described. The $m/z = 340$ extracted ion chromatograms are shown for the standard (50 mM; blue) and for the yeast growth media of CSY410 supplemented with 4 mM norlaudanoline (red). MS/MS on the 340 ions shows the major ions $m/z = 149$ and $m/z = 175$ in common.

Similar analysis of engineered strains expressing endogenous levels of the yeast P450-NADPH reductase or additional copies of the yeast or human reductase did not demonstrate (*S*)-tetrahydroberberine production, suggesting that these CPR1 variants were not suitable reductase partners for CYP719A1 (data not shown). We also constructed a strain in which the *A. thaliana* *ATR1* reductase coding sequence was integrated into the chromosome of CSY288 and transformed with the plasmids expressing BBE, SMT, and CYP719A (CSY410; Table 4.1). LC-MS/MS analysis demonstrated that (*S*)-tetrahydroberberine accumulation was ~10-fold greater in CSY410 than strains with plasmid-based *ATR1* expression (Fig. 4.10b). We estimated (*S*)-tetrahydroberberine production to be ~30 mg l⁻¹ from a substrate concentration of 4 mM or ~1-2% total conversion from norlaudanoline or laudanoline prior to optimization

of this heterologous seven-enzyme pathway. The accumulation of several intermediates in the synthetic pathway highlighted remaining flux limitations in our system (Fig. 4.13).

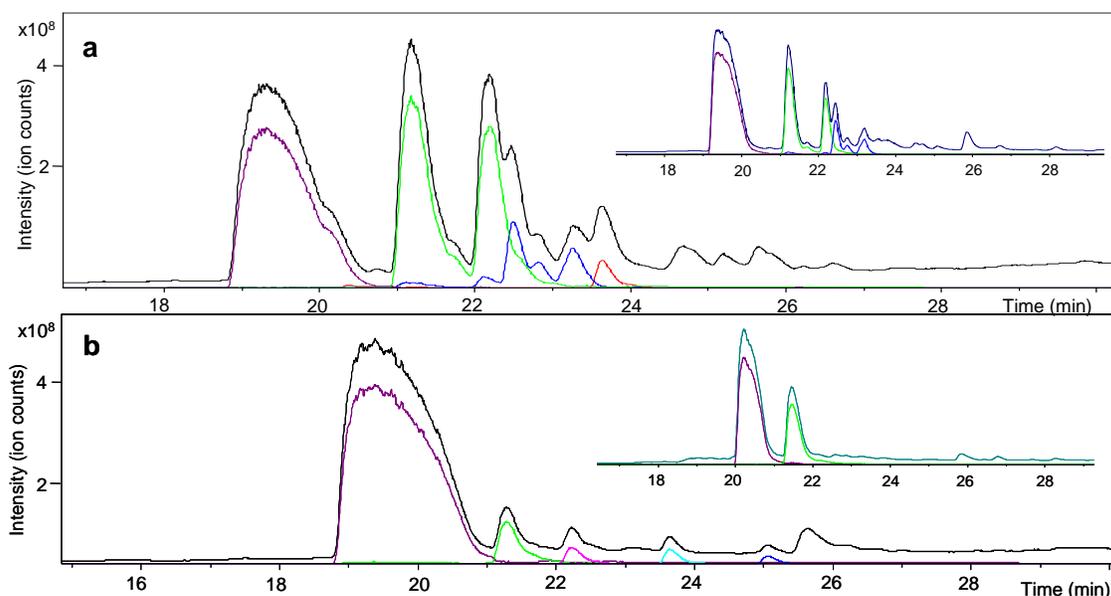


Fig. 4.13. Chromatograms show impurities and incomplete conversion of intermediates. LC-MS chromatograms of the growth media of (a) CSY288 supplemented with 2 mM norlaudanosoline and (b) CSY410 supplemented with 2 mM laudanosoline and grown for 48 h showing impurities and degradation products in the sample. (a) The total ion chromatogram is shown in black and the extracted ion chromatogram for norlaudanosoline (MS 288) is shown in purple. The major impurities are MS 314 shown in green and MS 316 shown in blue; reticuline (MS 330) is shown in red. The same 314 and 316 ions are observed in the wild-type strain supplemented with norlaudanosoline (inset). (b) The total ion chromatogram is shown in black and the extracted ion chromatogram for laudanosoline (MS 302) is shown in purple. The major impurity is MS 300 shown in green; the intermediate methyl-laudanosoline (MS 316) is shown in magenta; the intermediate reticuline (MS 330) is shown in cyan; and the intermediate tetrahydrocolumbamine (MS 342) is shown in blue. The same 300 ion is observed in the wild-type strain supplemented with laudanosoline (inset). All detectable BIA metabolites elute in this time frame with the rest of the spectra relatively flat.

4.2.5. *N*-methylation of (*S*)-tetrahydroberberine for the production of a noscapine pathway intermediate

Another rare biochemical reaction was demonstrated in our engineered yeast strains to produce a quaternary ammonium alkaloid. The tetrahydroprotoberberine *cis*-*N*-methyltransferase (TNMT) was recently cloned from opium poppy⁷¹. The native

substrate was identified as (*S*)-stylopine as this enzyme was implicated in protopine and sanguinarine biosynthesis. TNMT catalyzes the AdoMet-dependent methylation of the tertiary nitrogen of (*S*)-stylopine or similar substrate to produce a quaternary ammonium alkaloid. While exhibiting significant sequence similarity to CNMT, TNMT is the only *N*-methyltransferase in alkaloid metabolism known to produce a quaternary amine. Characterization studies of recombinant TNMT purified from *E. coli* showed a strict substrate requirement for dimethoxy or methylenedioxy functional groups at C2/3 and C9/10. The product (*S*)-tetrahydroberberine meets this requirement and TNMT was actually shown to have slightly greater activity on this substrate as opposed to stylopine⁷¹.

TNMT activity was tested in our engineered yeast strains expressing 6OMT, CNMT, and 4'OMT to produce reticuline and BBE, SMT, CYP719A/ATR1 to produce (*S*)-tetrahydroberberine. Although levels of the substrate (*S*)-tetrahydroberberine are relatively low in this extensive pathway, the correct reaction product was still detectable in the growth media. The extracted ion chromatograms and fragmentation results are shown (Fig. 4.14).

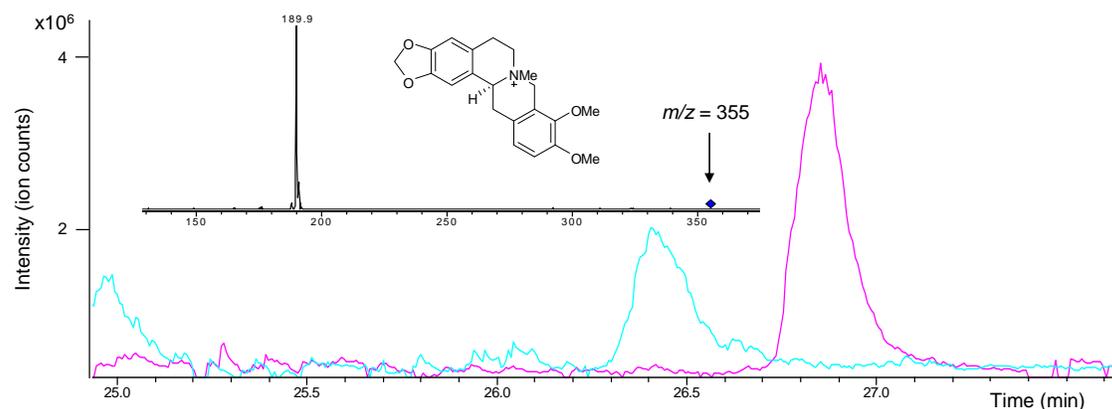


Fig. 4.14. Activity of TNMT on (*S*)-tetrahydroberberine. The product is identified as the $m/z = 355$ ion shown in cyan and elutes as 26.4 min. The major fragment $m/z = 189.9$ is 15 units higher than that from (*S*)-tetrahydroberberine (Fig. 4.10) as expected from the addition of the *N*-methyl group. The substrate (*S*)-tetrahydroberberine produced *in vivo* in yeast is shown in magenta.

4.2.6. Pathway for laudanine production

An additional branch extending from reticuline which is only known to include one additional step is the AdoMet-dependent methylation in the 7-*OH* position. The *P. somniferum* (*R, S*)-reticuline 7-*O*-methyltransferase converts reticuline to laudanine (Fig. 4.15a). The recombinant 7OMT is slightly more promiscuous than similar plant methyltransferases in that it accepts phenolic compounds; however, it does not accept *N*-demethylated substrates.

This reaction was anticipated to be straightforward in our system; however, yeast strains engineered to express 7OMT from various constructs did not show production of laudanine detectable by LC-MS/MS analysis. *In vitro* reactions set up as described previously⁶⁷ using yeast lysates containing 7OMT also did not show activity of this protein on reticuline, laudanosoline, or guaiacol. Although expression of the full-length protein was confirmed (Fig. 4.15b), no activity was observed under any conditions tested when this cDNA was expressed in yeast.

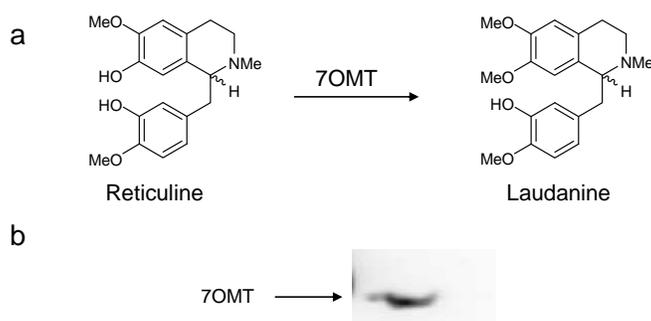


Fig. 4.15. (*R, S*)-Reticuline 7OMT reaction and expression in yeast. **(a)** The reaction performed by 7OMT to convert reticuline to laudanine. **(b)** Expression of the full-length 7OMT in reticuline-producing strain backgrounds was confirmed by Western blotting using a construct tagged with the V5 epitope.

4.2.6. Synthesis of the morphinan intermediate salutaridine

The morphinan alkaloids comprise another major family of BIA metabolites derived from the intermediate reticuline. This family of alkaloids, which includes morphine, codeine, and thebaine, has received the widest pharmacological use to-date and generated interest mainly for their analgesic properties. However, several of the enzymes that perform the early conversion steps in the morphinan branch have not been cloned and characterized, precluding engineering efforts relying on the native enzymes. To further demonstrate the diversity of small molecule products that can be synthesized and the power of reconstructing metabolic pathways in engineered hosts, we utilized an enzyme unrelated to the plant BIA pathway to synthesize an early metabolite in the morphinan branch.

Metabolites in the morphinan branch are synthesized from (*S*)-reticuline via (*R*)-reticuline following a two step-isomerization process (Fig. 4.16a). In the first step, (*S*)-reticuline is transformed to 1,2-dehydroreticuline by 1,2-dehydroreticuline synthase (DRS)⁷². The second step is a reduction of 1,2-dehydroreticuline in the presence of NADPH to (*R*)-reticuline catalyzed by 1,2-dehydroreticuline reductase (DRR)⁷³. Both enzymes have been partially purified and characterized but no sequence information is available to facilitate expression in a recombinant host⁶. With the lack of available cDNAs for enzymes performing the early isomerization steps, engineering a heterologous pathway to produce morphinan alkaloids is not possible if only (*S*)-reticuline is produced. However, using (*R, S*)-norlaudanosoline as the substrate for our engineered BIA pathway resulted in the synthesis of both forms of reticuline, allowing access to metabolites along the branch extending towards morphine production.

The first morphinan alkaloid synthesized from (*R*)-reticuline is salutaridine via a complex carbon-carbon phenol coupling reaction catalyzed by salutaridine synthase. This enzyme was recently cloned from *P. somniferum* but no characterization studies have been published to date^{6, 74}. However, a recent study demonstrating that humans are capable of synthesizing small amounts of morphine suggests the presence of analogous enzymes in other organisms. The human cytochrome P450 CYP2D6 has been implicated in this pathway, specifically catalyzing the hydroxylation of tyramine to dopamine and the demethylation of codeine to morphine⁷⁵. Furthermore, it was implied that an unidentified P450 enzyme, possibly CYP2D6, may accept reticuline as a substrate although experimental evidence was not shown. We examined the ability of CYP2D6 to convert (*R*)-reticuline to salutaridine in our engineered yeast strains (Fig. 4.16b). We constructed plasmids expressing the *CYP2D6* cDNA and various reductase partners for co-expression in our reticuline-producing yeast strain backgrounds (CSY463, CSY465-6; Table 4.1).

LC-MS/MS analysis demonstrated that CYP2D6 catalyzed the conversion of (*R*)-reticuline to salutaridine in strains producing (*R, S*)-reticuline, highlighting a previously uncharacterized activity for this enzyme (Fig. 4.16b). Surprisingly, this reaction proceeded with only endogenous levels of the yeast P450 reductase (yCPR) and observed levels of salutaridine in a CSY288 background strain were consistent with 6-8% conversion of (*R, S*)-reticuline. Codon-optimization of the CYP2D6 sequence (yCYP2D6) to improve translational efficiency in yeast did not show a significant improvement in salutaridine yield but was retained for future work to help lighten metabolic burden of this extensive pathway (Fig. 4.16c). Based on our experience with the CYP719A cofactor

expression system, a chromosomal integration of the *hCPR1* reductase gene was made in the reticuline-producing background strain. This resulted in a ~30% increase in salutaridine production which was further increased by expressing *yCYP2D6* from the TEF6 promoter with 1.2 times the transcriptional activity of the native TEF1 promoter. Additional strains constructed with the soluble *yCPRΔ33* and a fusion of the *yCYP2D6*-*yCPRΔ33* did yield additional improvements. The maximum yield of salutaridine based on ~15% conversion of reticuline is ~25 mg l⁻¹ from 5 mM norlaudanosoline. However, the activity of CYP2D6 on (*R*)-reticuline remained relatively low as cofactor and expression optimization strategies yielded only modest improvements. As human P450 enzymes are notoriously promiscuous, the affinity is likely low for (*R*)-reticuline and may require additional protein engineering or evolution to significantly improve specificity and activity for this substrate. In addition, as enzymes that exhibit this activity are cloned and characterized from plant hosts, they can be tested in our system for increased production of salutaridine⁷⁴.

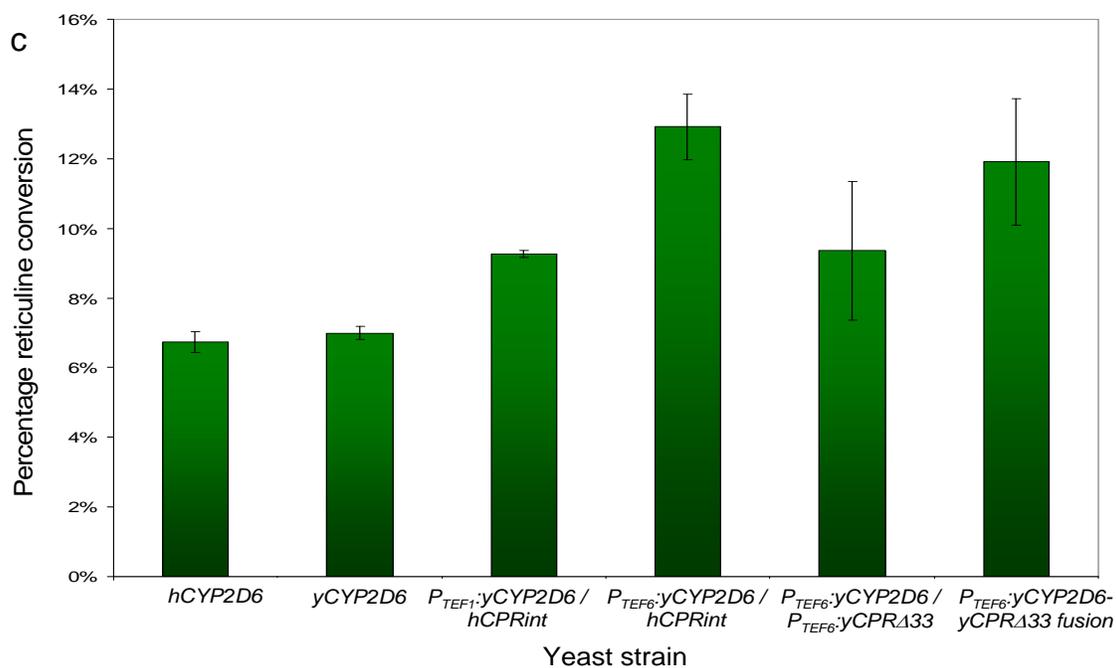
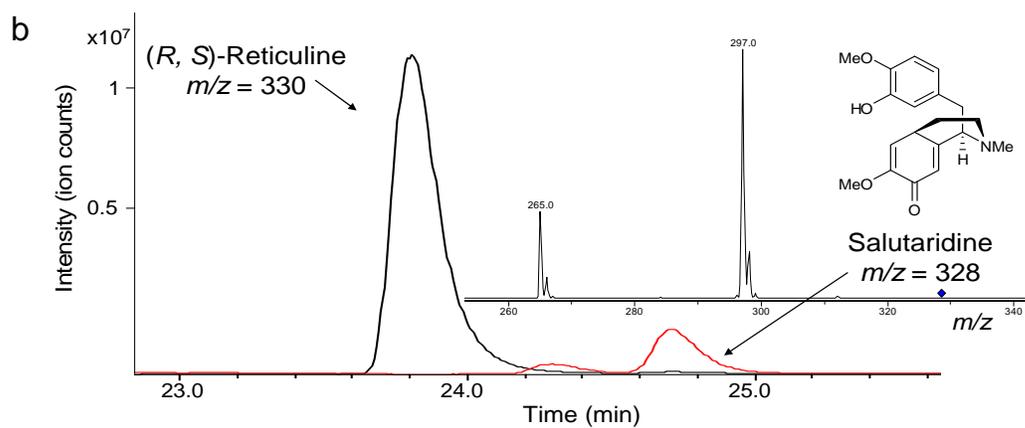
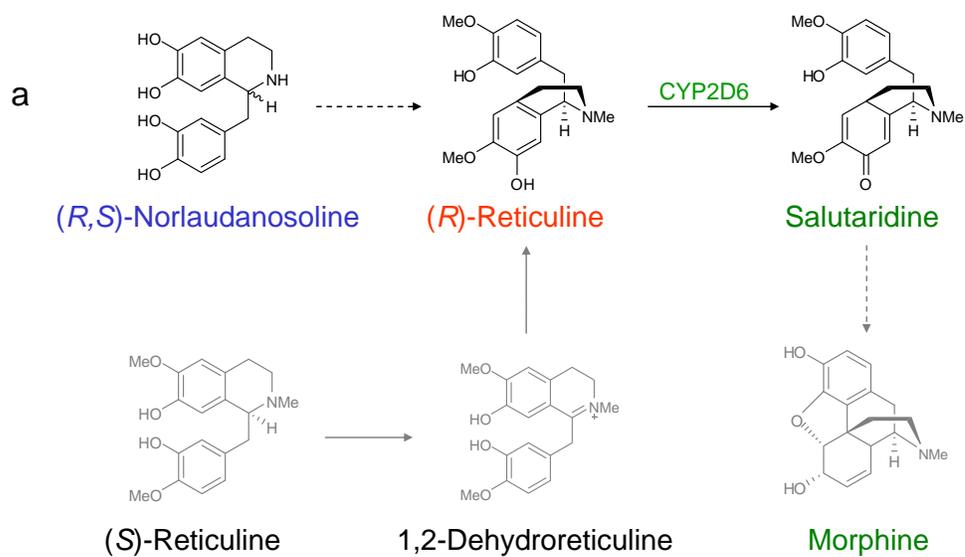


Fig. 4.16. Microbial production of morphinan alkaloids. **(a)** Native and engineered pathways for the synthesis of morphinan metabolites. Color schemes and notation follow that defined in Fig. 4.1. In the engineered pathway, (*R*)-reticuline is produced from (*R*)-norlaudanosoline. A heterologous human cytochrome P450 (*CYP2D6*) enzyme can convert (*R*)-reticuline to the first morphinan alkaloid salutaridine. In the native plant pathway, only (*S*)-reticuline is produced and two additional enzymes are required to produce (*R*)-reticuline, which is used in the synthesis of the morphinan alkaloids. **(b)** LC-MS/MS analysis of the growth media of engineered yeast strains supplemented with 4 mM norlaudanosoline and grown for 48 h confirms salutaridine production. Extracted ion chromatograms are shown for reticuline (black) and salutaridine (red) produced by CSY489 and are representative of analogous strains. The salutaridine 328 ion elutes at the same time as (*S*)-scoulerine, but exhibits a distinctly different fragmentation pattern. The major ions ($m/z = 297$ and 265) are consistent with the expected fragments reported in the literature for salutaridine⁷⁶. Strains lacking the *CYP2D6* coding sequence did not produce the metabolite peak identified as salutaridine. **(c)** Multiple salutaridine-producing strains were tested for conversion of reticuline to salutaridine. The *x*-axis shows the *CYP2D6* and *CPR* expression constructs present in the CSY288 background strain. The ‘optimal’ strain (which includes consistency) contains a chromosomal integration of *hCPR1* and increased plasmid-based expression of γ *CYP2D6*.

4.3. Discussion

The implementation of synthetic BIA pathways in a microbial host represents a significant challenge and a rich area for future research, in part due to the complex enzyme activities and regulatory strategies present in the plant hosts that may not translate directly to the engineered host by simply expressing the cloned enzyme activities. For example, a major factor that may generally limit enzymatic activity and conversion is the lack of subcellular compartmentalization in yeast. In particular, BBE and other enzymes are believed to be associated with endomembranes, forming ‘vesicles’ or ‘metabolons’ that facilitate channeling of intermediates and sequestration of toxic metabolites from the cytosol⁷⁷. Future engineering efforts may enable the assembly of synthetic enzyme complexes in heterologous hosts. A second factor that may limit conversion is the prevalence of cytochrome P450s in the native pathway, which typically exhibit low to no activity in microorganisms. Both *CYP719A1* and *CYP2D6* exhibited

significantly improved activities when coupled to the appropriate reductase partner stably integrated into the host genome as opposed to plasmid-based expression. However, both P450 enzymes did not fully consume their substrates, highlighting the potential for yield improvements at these steps through protein engineering strategies. A third factor that may limit conversion is the lack of an active transport system to facilitate passage of BIA metabolites across microbial cell membranes and allow for higher intracellular substrate concentrations. Such limitations highlight the importance for reconstruction of these pathways in a single microbial host and may represent a target for future engineering efforts to improve uptake of synthetic substrates.

As a heterologous production host for diverse BIA molecules, our engineered strains will be useful tools in furthering the characterization of a ubiquitous plant secondary metabolic pathway. First, our strains can be used as tools to characterize enzymes and probe regulatory strategies for the BIA pathway in a synthetic host. For example, we utilized our strains to probe the relative activities of enzyme variants in the engineered pathway. We further described the application of a general tool for titrating individual pathway enzymes to determine optimal expression levels. Such strategies indicated that the 6OMT and CNMT variants from *P. somniferum* exhibited superior properties in the engineered pathways relative to the *T. flavum* variants. Second, our strains can be used as functional genomics tools to further the elucidation of the BIA pathway and characterize new enzyme activities from native plant hosts or non-native sources as demonstrated here. The engineered strains can be used to screen cDNA libraries from BIA-producing plants in high-throughput assays for activities that modify

BIA metabolites. Such a tool is particularly powerful in light of the scarce availability of many BIA intermediates and challenges in plant metabolic engineering.

We have constructed unique production hosts for a diverse set of BIAs offering access to a large family of pharmacologically-relevant molecules previously unattainable from other natural sources and synthetic chemistry approaches. Our production hosts offer advantages in directing the production of specific BIA molecules, ease of purification from a GRAS organism, well-established fermentation schemes, and rapid biomass accumulation. The microbial BIA production levels reported here reached ~ 150 mg l⁻¹, similar to previously reported yields from initial engineering efforts on microbial strains to produce other natural plant products⁷⁸. The application of industrial optimization strategies is expected to significantly increase production levels. More importantly, the engineered yeast strains offer the potential to produce an even broader spectrum of BIA metabolites through extensions of this synthetic pathway. For example, as new enzyme activities are cloned from the native plant hosts, the cDNAs can be expressed in our engineered strains to produce berberine and other BIA metabolites in the sanguinarine, morphinan, and bis-benzylisoquinoline branches. As demonstrated here with a human P450 enzyme, there is also the exciting potential to express enzymes unrelated to the native BIA pathway but that accept BIA metabolites as substrates in our synthetic hosts to access a wider pool of intermediates and derivatives. Furthermore, strategies that involve recombining native and novel enzyme activities and feeding alternative substrates can be used to produce non-natural BIA molecules, expanding molecular diversity. Finally, synthetic methods can be used in conjunction with *in vivo* biosynthesis to attach various functional groups to the molecular backbones, thereby

creating a rich population of alkaloids and potentially tapping into new and enhanced pharmacological activities.

4.4. Materials and Methods

4.4.1. Plasmid and yeast strain construction

We obtained restriction enzymes, T4 DNA ligase, and other cloning enzymes from New England Biolabs. We performed polymerase chain reaction (PCR) amplifications using Expand High Fidelity PCR System (Roche). Oligonucleotide synthesis was performed by Integrated DNA Technologies. A list of strains and primer sequences for chromosomal integrations is provided (Tables 4.1 and 4.2).

Table 4.1. Engineered yeast strains.

<u>Strain</u>	<u>Integrated constructs</u>	<u>Plasmid-based constructs</u>
CSY288	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	
CSY307		<i>P_{TEF1}-Ps6OMT, P_{TEF1}-PsCNMT, P_{TEF1}-Ps4'OMT</i>
CSY308		<i>P_{TEF1}-Ps6OMT, P_{TEF1}-TfCNMT, P_{TEF1}-Ps4'OMT</i>
CSY309		<i>P_{TEF1}-Tf6OMT, P_{TEF1}-PsCNMT, P_{TEF1}-Ps4'OMT</i>
CSY310		<i>P_{TEF1}-Tf6OMT, P_{TEF1}-TfCNMT, P_{TEF1}-Ps4'OMT</i>
CSY311		<i>P_{TEF1}-Ps6OMT, P_{TEF1}-PsCNMT, P_{TEF1}-Tf4'OMT</i>
CSY312		<i>P_{TEF1}-Ps6OMT, P_{TEF1}-TfCNMT, P_{TEF1}-Tf4'OMT</i>
CSY313		<i>P_{TEF1}-Tf6OMT, P_{TEF1}-PsCNMT, P_{TEF1}-Tf4'OMT</i>
CSY314		<i>P_{TEF1}-Tf6OMT, P_{TEF1}-TfCNMT, P_{TEF1}-Tf4'OMT</i>
CSY325	<i>his3::P_{GAL1-10}-Ps6OMT-loxP-Kan^R, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, gal2::HIS3</i>	
CSY326	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{GAL1-10}-Ps4'OMT-loxP-LEU2, gal2::HIS3</i>	
CSY327	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{GAL1-10}-PsCNMT-loxP-URA3, ura3::P_{TEF1}-Ps4'OMT, gal2::HIS3</i>	
CSY328	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{GAL1-10}-Tf4'OMT-loxP-LEU2, gal2::HIS3</i>	
CSY329	<i>his3::P_{GAL1-10}-Tf6OMT-loxP-LEU2, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT, gal2::HIS3</i>	
CSY334	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	
CSY336	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-PsBBE</i>
CSY337	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-PsBBE, P_{TEF1}-TfS9OMT</i>
CSY338	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-PsBBE</i>
CSY339	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-PsBBE, P_{TEF1}-TfS9OMT</i>
CSY399	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-PsBBE, P_{TEF1}-TfS9OMT, P_{TEF1}-TfCYP719A, P_{TEF1}-AtATR1</i>
CSY400	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-PsBBE, P_{TEF1}-TfS9OMT, P_{TEF1}-TfCYP719A, P_{TEF1}-AtATR1</i>
CSY410	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-PsBBE, P_{TEF1}-TfS9OMT, P_{TEF1}-TfCYP719A</i>
CSY424	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-yCYP2D6</i>
CSY425	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-yCYP2D6</i>
CSY428	<i>ura3::P_{TEF1}-yEGFP3-loxP-LEU2</i>	
CSY429	<i>ura3::P_{GAL1-10}-yEGFP3-loxP-LEU2, gal2::Kan^R</i>	
CSY448	<i>his3::P_{TEF7mut}-Ps6OMT, leu2::P_{TEF7mut}-PsCNMT, ura3::P_{TEF7mut}-Ps4'OMT</i>	
CSY449	<i>his3::P_{TEF7mut}-Ps6OMT, leu2::P_{TEF7mut}-PsCNMT, ura3::P_{TEF7mut}-Tf4'OMT</i>	
CSY450		<i>P_{TEF1}-Ps6OMT</i>
CSY451		<i>P_{TEF1}-Tf6OMT</i>
CSY452		<i>P_{TEF1}-PsCNMT</i>
CSY453		<i>P_{TEF1}-TfCNMT</i>
CSY454		<i>P_{TEF1}-Ps4'OMT</i>
CSY455		<i>P_{TEF1}-Tf4'OMT</i>
CSY456	<i>his3::P_{TEF1}-Tf6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	
CSY458	<i>his3::P_{TEF7mut}-Tf6OMT, leu2::P_{TEF7mut}-PsCNMT, ura3::P_{TEF7mut}-Tf4'OMT</i>	
CSY463	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-hCYP2D6</i>
CSY464	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Tf4'OMT</i>	<i>P_{TEF1}-yCYP2D6:yCPR1 (fusion protein)</i>
CSY465	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-hCYP2D6, P_{TEF1}-yCPR1</i>
CSY466	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-hCYP2D6, P_{TEF1}-AtATR1</i>
CSY489	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-hCPR-loxP-LEU2</i>	<i>P_{TEF1}-hCYP2D6</i>
CSY490	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-hCYP2D6:hCPR1 (fusion protein)</i>

We used standard molecular biology techniques to construct the BIA expression vectors⁵⁹. BIA expression constructs contained the 2 μ high-copy yeast origin of replication along with appropriate yeast selection markers and ampicillin resistance. Recombinant enzymes were expressed from the yeast TEF1 promoter and flanked by a CYC1 terminator sequence. We constructed shuttle vectors for subcloning of 1 or 2 cDNA sequences in this fashion. Coding sequences for the enzymes of interest, with the exception of *hCYP2D6*, were generously donated as cDNAs from Peter Facchini

(University of Calgary) in plasmids typically suited for expression in *E. coli*. The *hCYP2D6* cDNA was provided by F. Peter Guengerich (Vanderbilt University) as pCW/DB6⁶⁰ and the yeast codon-optimized version of this gene was synthesized by DNA 2.0. We PCR-amplified the endogenous yeast P450 reductase gene (*CPR1*) from W303 genomic DNA, the *A. thaliana ATR1* gene from WAT11 genomic DNA⁷⁹, and the *Homo sapiens CPR1* gene from pH2E1red⁶¹.

We transformed ligation reactions into an electrocompetent *E. coli* strain, DH10B (Invitrogen; F-*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*dlacZ* Δ M15 Δ *lacX74 deoR recA1 endA1 araD139 Δ (*ara, leu*)7697 *galU galK λ -rpsL nupG*), using a Gene Pulser Xcell System (BioRAD) according to the manufacturer's instructions. We conducted plasmid isolation using the Wizard Plus SV Minipreps DNA purification system (Promega) according to the manufacturer's instructions. Subcloning was confirmed by restriction analysis and sequence verification (Laragen, Inc.). We transformed plasmids into the appropriate *S. cerevisiae* strains using a standard lithium acetate protocol⁶². All yeast strains used in this work were based on the haploid yeast strain W303 α (*MAT α his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*)⁶³. *E. coli* cells were grown on Luria-Bertani media (BD Diagnostics) with 100 μ g/ml ampicillin (EMD Chemicals) for plasmid maintenance, and *S. cerevisiae* cells were grown in synthetic complete media (BD Diagnostics) supplemented with the appropriate dropout solution for plasmid maintenance (Calbiochem).*

We performed chromosomal integrations of DNA fragments through homologous recombination using a standard lithium acetate transformation protocol to construct strains that stably express combinations of the BIA enzymes. We built gene insertion

cassettes harboring the appropriate BIA enzyme expression construct and associated selection marker flanked by *loxP* sites to allow removal of the selection marker following integration with a Cre-*loxP* system⁸⁰. The plasmids pUG6, pUG27, pUG72, and pUG73⁸¹ (EUROSCARF) contain geneticin resistance, *S. pombe his5⁺*, *Kluyveromyces lactis URA3*, and *K. lactis LEU2* genes, respectively, flanked by *loxP* sites and were used in the construction of the integration cassettes. We amplified and assembled the TEF1 promoter and CYC1 terminator using PCR-based methods and the primers TEF.fwd, TEF.rev, CYC1.fwd, and CYC1.rev (Table 4.2). The assembled DNA insert contained a multi-cloning site and was subcloned upstream of the selection marker in each construct. We cloned cDNAs into the multi-cloning site and amplified the entire integration cassette in two PCR steps using “A” and “B” primer sets to add ~80 nt of homology (Table 4.2). We gel purified integration cassettes using the Zymoclean Gel DNA Recovery Kit (Zymo Research) according to the manufacturer’s instructions prior to transformation into the appropriate yeast strain. Integration of the cassettes into the correct locus was confirmed by PCR analysis of the targeted region of the chromosome. To remove all selection markers from the final strains, we transformed cells with the plasmid pSH63 which expresses a GAL-inducible Cre recombinase⁸¹. Cells harboring the plasmid were induced with galactose for 24 h and then transferred to YPD media to remove the selection pressure for plasmid maintenance. We isolated and restreaked single colonies on appropriate media to verify the loss of selection markers.

We built template plasmids for chromosomal integrations for the enzyme titration studies by cloning the GAL1-10 promoter (amplified from pRS314-GAL³⁸ using primers GAL.fwd and GAL.rev) in place of the TEF1 promoter in pUG-based plasmids.

Construction of strains using the GAL1-10 promoter to control expression of single enzymes was analogous to that described for the TEF1 promoter strains with a final step to replace the *GAL2* locus with the *his5⁺* selection marker using a cassette amplified from pUG27 with primers GAL2ko.fwd and GAL2ko.rev. We confirmed integration into the targeted site with primers GAL2sc.fwd and GAL2sc.rev (Table 4.2).

Table 4.2. Primer sequences used for qRT-PCR and to make stable integrations of enzyme constructs and *GAL2* knockouts. Restriction sites are underlined.

Primer name	Sequence
TEF.fwd	ATAAAAGC <u>TTACATATGCATAGCTT</u> CAAAATGTTTCTACTCC
TEF.rev	TGCTAGC <u>GT</u> TTAAACGAATTCTGGATCCATCCTAGGAAAACCTAGATTAGATTGCTATGC
CYC1.fwd	ATGGATCCAGAA <u>TT</u> CGTTTAAACGCTAGCACCCGGGGAGGGCCGCATCATGTAATTAGTT
CYC1.rev	ATAACTGCAGAG <u>TCGAC</u> GGCCGCAAATTAAGCCTTCGA
HISint.fwdA	TATACTAAAAAATGAGCAGGTGTCGGGGCTGGCTTAACTATGCGGCATCA
HISint.revA	TATATATATCGTATGCTGCAGCTCTGGCTTATCGAAATTAATACGACTCACTA
HISint.fwdB	ATTGGCATTATCACATAATGAATTATACATTATATAAGTAATGTGATTTCTTGAAGAATATACTAAAAAATGAGCAGG
HISint.revB	AGTATCATACTGTTTCGTATACATACTTACTGACATTCATAGGTATACATATATACACATGTATATATATCGTATGCTGCAGCT
LEUint.fwdA	AGCAATATATATATATATATTGTCGGGGCTGGCTTAACTATGCGGCATCA
LEUint.revA	AGTTTATGTACAAATATCATAA <u>CTGGCTTATCGAAATTAATACGACTCACTA</u>
LEUint.fwdB	TTTTCCAATAGGTGTTAGCAATCGTCTTACTTTCTAACTTTTCTTACCTTTTACATTCAGCAATATATATATATATAT
LEUint.revB	TACCCTATGAACATATTCCATTTTGTAAATTCGTGTCGTTTCTATTATGAATTTCAATTTATAAAGTTTATGTACAAATATCATAA
URAint.fwdA	ACTGCACAGAACAACCTTGTGCGGGCTGGCTTAACTATGCGGCATCA
URAint.revA	AGTTTAGTATACATGCATTTACTGGCTTATCGAAATTAATACGACTCACTA
URAint.fwdB	GGTATATATACGCATATGTGGTGTGAAGAACATGAAATGCCCAGTATTCTTAACCAACTGCACAGAACAACCT
URAint.revB	AATCATTACGACCGAGATTCCCGGTAATAACTGATATAAATAAATGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTA
GAL.fwd	ATAAAGC <u>TTACATATG</u> TCTAGAAAATTCCTTGAATTTTCAAAAAT
GAL.rev	ATAC <u>CTAGG</u> TTTTTTCTCCTTGACGTTAAAGTATAGAG
GAL2ko.fwd	CATTAATTTTGCTTCCAAGACGACAGTAATATGTCTCCTACAATACCAGTGTGCGGGCTGGCTTAACTATGCGGCATCA
GAL2ko.rev	TATATGTACACAAATAATAGGTTTAGGTAAGGAATTTATATAATCGTAAGCTGGCTTATCGAAATTAATACGACTCACTA
GAL2sc.fwd	TGTGCATGTTATCTATATCCTTCTTTATATAGATGCTGTT
GAL2sc.rev	ATTAATTGTATGTTAGCTCAGGAATTCAACTGGAAGAAAG

4.4.2. Growth conditions

For BIA metabolite production assays with the exception of enzyme titration studies, engineered yeast strains were grown in test tube cultures in volumes ranging from 2-10 mL at 30°C and 200 rpm in the appropriate drop out media with 2% dextrose

(w/v) as a sugar source. We diluted overnight cultures 1:20 in fresh media supplemented with the appropriate pathway substrate as reported, typically norlaudanoline (CHEMOS GmbH, ACROS Organics; ~\$160 g⁻¹) or laudanoline (ACROS Organics; ~\$220 g⁻¹) at concentrations between 0.1-5 mM diluted from a 10 or 20 mM stock solution in water. With the exception of time-course experiments, we assayed cultures at 24 or 48 h following substrate addition to observe maximum accumulation as cells reached the stationary growth phase.

For the enzyme titration assays, GAL-inducible strains were grown overnight in synthetic complete media containing all amino acids with 2% raffinose (w/v), 1% sucrose (w/v) as a sugar source. Galactose was added to the media at the appropriate concentration from a 10X stock at the time of back dilution and the norlaudanoline substrate was added to a final concentration of 1 mM following an induction period of 4 h. We sampled aliquots of the growth media for LC-MS/MS analysis 24 h after substrate addition. Data is reported as reticuline production measured by LC-MS peak area.

4.4.3. Analysis of metabolite production

We evaluated BIA metabolite levels by LC-MS/MS analysis of cell extracts and growth media. At appropriate time points, aliquots of yeast cultures were centrifuged to recover cells as pellets and allow collection of the growth media. We analyzed the growth media or an appropriate dilution directly by LC-MS/MS. Samples were run on an Agilent ZORBAX SB-Aq 3 x 250 mm, 5 μm column with 0.1% acetic acid as solvent A and methanol as solvent B. We used a gradient elution to separate the metabolites of interest as follows: 0-10 min at 100% A, 10-30 min 0-90% B, 30-35 min 90-0% B, followed by a

5 min equilibration at 100% A between samples. Following LC separation, metabolites were injected into an Agilent 6320 ion trap MSD for detection and identification. We used selective reaction monitoring to isolate ions of interest for MS/MS to verify the molecular structure of each metabolite. We verified chromatogram data including fragmentation patterns through at least three independent experiments and from multiple strains where appropriate. Quantification of metabolites was based on the integrated area of the extracted ion chromatogram peaks⁶⁴ calculated using DataAnalysis for 6300 Series Ion Trap LC/MS Version 3.4 (Bruker Daltonik GmbH) and reported as the mean \pm s.d. When appropriate, we normalized the measured levels to a metabolite peak of known concentration in the growth media, typically the substrate (norlaudanosoline) peak. We also generated standard curves for norlaudanosoline and laudanosoline to relate peak area to metabolite concentration in our samples. Although slight differences in ionization efficiencies were observed between these available standards, this method further validated our estimates of metabolite concentrations based on percentage substrate conversion. A summary of the average yields for each of the synthesized BIA compounds is provided (Table 4.3).

Table 4.3. Summary of yields of benzyloquinoline alkaloids. Yields are reported from the supernatant of growth cultures of the appropriate strains fed 5 mM norlaudanosoline.

BIA compound	yield (μM)	yield (mg l⁻¹)
(R, S)-reticuline	500	164.5
(S)-scoulerine	200	65.4
(S)-tetrahydrocolumbamine	200	68.2
(S)-tetrahydroberberine	100	33.9
salutaridine	75	24.5

4.4.4. Fluorescence quantification

For fluorescence measurements of CSY428 and CSY429, overnight cultures were grown in synthetic complete media with either 2% dextrose or 2% raffinose and 1% sucrose (noninducing-nonrepressing), respectively. Cells were backdiluted into fresh media containing either dextrose (CSY428) or noninducing-nonrepressing media supplemented with the appropriate galactose concentration from a 10X stock (CSY429). We measured fluorescence (excitation 485 nm, emission 515 nm) and OD₆₀₀ values on a fluorescence plate reader (TECAN, Safire). Fluorescence values were normalized using the OD₆₀₀ reading and values from a no stain control (wild-type cells) were subtracted. We used the resulting trend between relative fluorescence and galactose concentration as a guideline to estimate relative protein expression levels in our titratable strains.

4.4.5. Analysis of protein levels through Western blotting

We constructed plasmids for Western blotting experiments by cloning the C-terminal epitope tag(s) from pYES-NT/A (Invitrogen) into a standard BIA expression vector followed by subcloning of the enzyme of interest. We transformed individual plasmids into wild-type yeast cells using a standard lithium acetate protocol. Overnight cultures were grown and backdiluted 1:100 into 100 ml cultures. Cells were grown to OD₆₀₀ ~1.2 and pelleted by centrifugation. The media was removed and cells were washed in 1 ml PBS, pelleted, and resuspended in 0.5 ml Y-PER plus HALT protease inhibitor (Pierce). Cells were vortexed for ~20 min and the lysate separated by centrifugation. We estimated total protein using the Coomassie Plus Protein Assay Reagent (Pierce) and loaded ~50 µg of each sample onto a protein gel. We used NuPage 4-12% Bis-Tris gels with MES running buffer and transfer buffer according to the

manufacturer's instructions (Invitrogen). Proteins were blotted onto a nitrocellulose membrane (Whatman) using a semi-dry transfer cell (Bio-Rad) for 20 min at 15 V. We incubated the membrane with the Anti-V5 HRP antibody (Invitrogen) according to the manufacturer's instructions (Invitrogen) with 5% nonfat milk as the blocking agent. Proteins were detected with the West Pico Super Signal Detection kit (Pierce) and imaged on a ChemiDoc XRS system (Bio-Rad). The image shown is representative of at least two independent experiments.

4.4.6. Analysis of transcript levels through qRT-PCR

We extracted total RNA from yeast cells grown in 10 ml cultures to $OD_{600} \sim 0.5$ using standard acid phenol extraction procedures⁸². Briefly, cells were pelleted, frozen in liquid nitrogen, and resuspended in a 50 mM NaOAc (pH 5.2) and 10 mM EDTA buffer. Cells were lysed by adding SDS to a final concentration of 1.6% and equal volume of acid phenol and incubating for 10 min at 65°C with intermittent vortexing. Following cooling on ice, the aqueous phase was extracted and further extraction was carried out with an equal volume of chloroform. RNA samples were ethanol precipitated and resuspended in water, and total RNA was quantified on a NanoDrop according to the manufacturer's instructions. RNA samples were treated with DNase (Invitrogen) according to the manufacturer's instructions.

We performed cDNA synthesis using gene-specific primers (Table 4.4) and Superscript III Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions with 5 µg total RNA used in each reaction. Relative transcript levels were quantified from the cDNA samples by employing an appropriate primer set and the iQ SYBR Green Supermix (Bio-Rad) according to the manufacturer's instructions on an

iCycler iQ qRT-PCR machine (Bio-Rad). We analyzed the resulting data with the iCycler iQ software according to the manufacturer's instructions. We normalized transcript levels to that of the endogenous *ACT1* gene⁸³ using *ACT1*-specific primers. Data are reported as the mean \pm s.d. from at least two independent experiments.

Table 4.4. Primers used for qRT-PCR.

<u>Primer name</u>	<u>Sequence</u>
RTGFP.fwd	GCCATGCCAGAAGGTTATGT
RTGFP.rev	ACCATTCTTTTGTTCAGCC
RT6OMT.fwd	CTCATTAGCTCCACCAGCTAAGTA
RT6OMT.rev	GGTCAAACCGTCGCCTAAAT
RTCNMT.fwd	CAAAGTGCGGGTTACTACTCTG
RTCNMT.rev	TCCACAGAACTTGTAAGTCCAGT
RT4'OMT.fwd	GGTGCCGATGATACTAGGCA
RT4'OMT.rev	CCATTCCCTTCATTAAACAAGTGGC
ACT1.fwd	GGCATCATACCTTCTACAACGAAT
ACT1.rev	GGAATCCAAAACAATACCAGTAGTTCTA

4.4.7. Analysis of stereoisomer forms of BIA metabolites through capillary electrophoresis-based chiral separation

We performed chiral analysis of norlaudanosoline, laudanosoline, and reticuline using the P/ACE MDQ Capillary Electrophoresis (CE) system (Beckman-Coulter). A bare fused silica capillary (50 μ m i.d.) with 31.0 cm effective length (41 cm total) was used. We tested buffer solutions from the chiral test kit (Beckman) with the substrates norlaudanosoline and laudanosoline and optimized conditions using the HS- β -CD separation buffer at 5% (w/v) according to the manufacturer's instructions. Injection was for 5 psi for 10 sec and voltage separation was performed at 15 kV. A diode array detector (PDA) module was used with the spectra presented taken at 280 nm.

For analysis of reticuline and other BIA metabolites, we developed HPLC methods to use water as the aqueous buffer rather than 0.1% acetic acid to avoid unnecessary noise in the CE spectra. Samples were run on a ZORBAX SB-Aq 3 x 250 mm, 5 μ m column with a gradient elution from 10% B to 90% B between 0 and 30 min where solvent B is methanol and solvent A is water. Reticuline elutes at 13 min using this method as confirmed by MS data. We collected fractions from 13.0-13.2 min and verified purity by re-running the sample(s) on LC-MS. Multiple fractions were collected and concentrated ~4 to 5-fold prior to CE analysis.

4.4.8. Preparation of cell extracts

We prepared cell extracts by washing the cell pellets three times with PBS, adding 1 ml of cold methanol at -40°C, and vortexing with glass beads to disrupt the cells. Cell debris and glass beads were pelleted by centrifugation. The methanol was transferred to a clean tube and concentrated by evaporation to a final volume approximately equal to that of the original cell mass.

CHAPTER V. PRODUCTION OF MORPHINAN ALKALOIDS IN***SACCHAROMYCES CEREVISIAE*****Abstract**

The branch of benzyloquinoline alkaloid metabolism leading to codeine and morphine includes several molecules of proven and potential pharmacological importance. Production of morphinan alkaloids in a microbial host in lieu of extraction from plants will increase the availability of these molecules for drug discovery and decrease manufacturing costs. This work builds on previously engineered strains of *Saccharomyces cerevisiae* capable of converting the substrate (*R, S*)-norlaudanosoline to salutaridine, the first benzyloquinoline alkaloid specific to the morphinan pathway. We tested the recently cloned salutaridine synthase (SalSyn) for the conversion of (*R*)-reticuline to salutaridine and compared it to CYP2D6 used in the original strains. In addition, we engineered strains to co-express salutaridine reductase (SalR) and salutaridinol-7-*O*-acetyltransferase (SalAT) for the production of salutaridinol-7-*O*-acetate. We performed preliminary work to increase accumulation of salutaridinol-7-*O*-acetate with the end goal of developing a process for thebaine production.

5.1. Introduction

The morphinan alkaloids and their derivatives are a source of many important and diverse pharmaceuticals. Isolation from plants is the sole means of obtaining the active compounds morphine, codeine, and thebaine. Additional chemical and biosynthetic derivatization further broadens the array of useful therapeutics in this category. The goal of this work is to engineer yeast for the production of morphinan alkaloids. A microbial production host avoids rigorous purification processes and removes existing limits on plant production. Additional conversion steps can also be performed in the host organism as required. Thebaine is the initial target molecule along the morphinan branch. The requisite enzymes have been cloned, and thebaine represents only 6.5% of the total alkaloid content of *Papaver somniferum* yet can be used for the production of many semisynthetic opiate drugs⁸⁴.

Previously, the conversion of norlaudanoline and the morphinan alkaloid salutaridine was demonstrated in yeast⁴. From salutaridine, additional downstream enzymes are available for the production of salutaridinol and salutaridinol-7-*O*-acetate (Fig. 5.1). We used a strain background constitutively expressing *Papaver somniferum* 6OMT, CNMT, and 4'OMT to convert (*R, S*)-norlaudanoline, a commercially available substrate with the BIA backbone structure, to (*R, S*)-reticuline. Using the (*R*)- isomer of reticuline, we are able to produce salutaridine using a cytochrome P450 activity from humans or plants. This novel activity for the human CYP2D6 was demonstrated using engineered yeast cells and optimal conversion occurred in a strain background expressing a single copy of the human cytochrome P450 reductase (hCPR1). The salutaridine synthase activity (SalSyn, CYP719B) from *P. somniferum* has recently been cloned;

however, the characterization of this enzyme expressed in a recombinant host has not yet been published. We tested this enzyme in a variety of strain backgrounds and used both P450 activities in the development and initial optimization of a process for thebaine production.

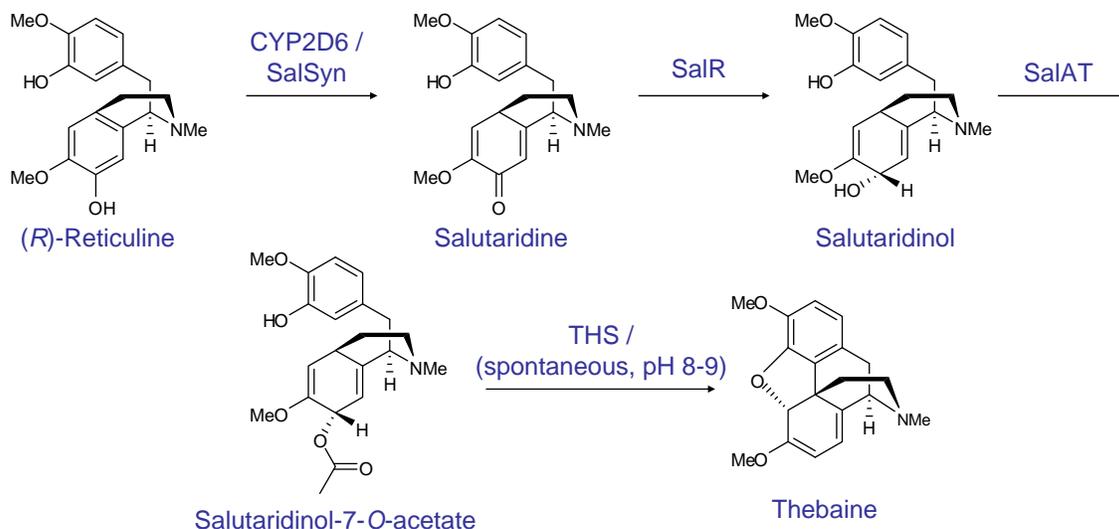


Fig. 5.1. Pathway for the production of thebaine from (*R*)-reticuline. In our engineered yeast strains, both (*R*)- and (*S*)- isomers of reticuline are produced from a racemic mixture of norlaudanosoline as the substrate for this extended branch of the pathway. Additional conversion steps by CYP2D6/SalSyn, SalR, and SalAT lead to salutaridinol-7-*O*-acetate which is spontaneously converted to thebaine at high pH and may also be facilitated by thebaine synthase (THS) in plant hosts.

The remaining steps in this pathway are catalyzed by salutaridine reductase (SalR) and salutaridinol-7-*O*-acetyltransferase (SalAT). Both enzymes have been cloned from *P. somniferum* and were tested in our yeast expression hosts. This work marks the development of a process that produces modest amounts of salutaridinol-7-*O*-acetate. However, many optimization challenges remain, primarily due to the pH dependence of thebaine formation. A two-stage process or combination of *in vivo* and *in vitro* transformations may be required.

5.2. Results

5.2.1. Expression and activity of *Papaver somniferum* salutaridine synthase

The recently cloned *P. somniferum* salutaridine synthase (SalSyn, CYP719B), the enzyme responsible for conversion of (*R*)-reticuline to salutaridine in plants, was codon-optimized for yeast and tested in our engineered strains. No accumulation of salutaridine was observed when this enzyme was tested in combination with the human, yeast, or *A. thaliana* P450 reductase. There are only two known P450 reductase sequences from poppy plants⁸⁵. Unlike humans and yeast, plants are known to have multiple P450 reductase genes such that the single reductase identified from *P. somniferum* may not be the correct partner for SalSyn. We codon-optimized this sequence for the host organism *S. cerevisiae* and assembled the gene from short oligonucleotide templates. This optimized *P. somniferum* reductase was tested in combination with SalSyn. We tested the *P. somniferum* reductase on a low-copy expression vector or integrated into the host genome and were unable to observe accumulation of salutaridine from SalSyn.

We also tested various designs of SalSyn with the N-terminal sequence replaced by the first 16-18 amino acids of bovine CYP17, which has been shown to efficiently target the yeast microsomal membrane⁸⁶. Salutaridine accumulation was not observed in these strains co-expressing the described P450 reductase partners. Although we did not observe accumulation of salutaridine attributable to SalSyn activity, we were able to show that this activity participates in the larger pathway for salutaridinol-7-*O*-acetate production and actually outperforms CYP2D6 in most strains.

5.2.2. Expression of *Papaver salutaridine reductase* variants for the production of *salutaridinol*

Additional enzymes implicated in morphine biosynthesis in plants downstream of salutaridine have been cloned and characterized. Salutaridine reductase (SalR) converts salutaridine to salutaridinol. This enzyme has been cloned from both *P. somniferum* and *P. bracteatum* species. Although the protein sequences differ by only 13 amino acids, the *P. bracteatum* variant is reportedly more stable in a bacterial host⁸⁷. We examined protein expression levels in the W303 α yeast strain and also observed higher levels of the *P. bracteatum* variant as expected (Fig. 5.2a). This variant was also shown to have more favorable catalytic properties *in vitro*^{65, 87} and was therefore the initial focus of additional studies in our yeast hosts. Optimization of codon usage did not increase expression of the *P. bracteatum* SalR as observed by Western blotting experiments (Fig. 5.2a). However, this sequence was retained as increased translational efficiency is expected to reduce metabolic burden on engineered yeast strains expressing multiple recombinant proteins. Based on previous mutagenesis studies, we also constructed V106A and R48K mutants and the combinatorial mutant, which have been shown to improve turnover rate and increase NADH affinity, respectively⁸⁷. These SalR variants were tested in combination with SalSyn and CYP2D6 (with and without SalAT) and no improvements were observed (Fig. 5.2b).

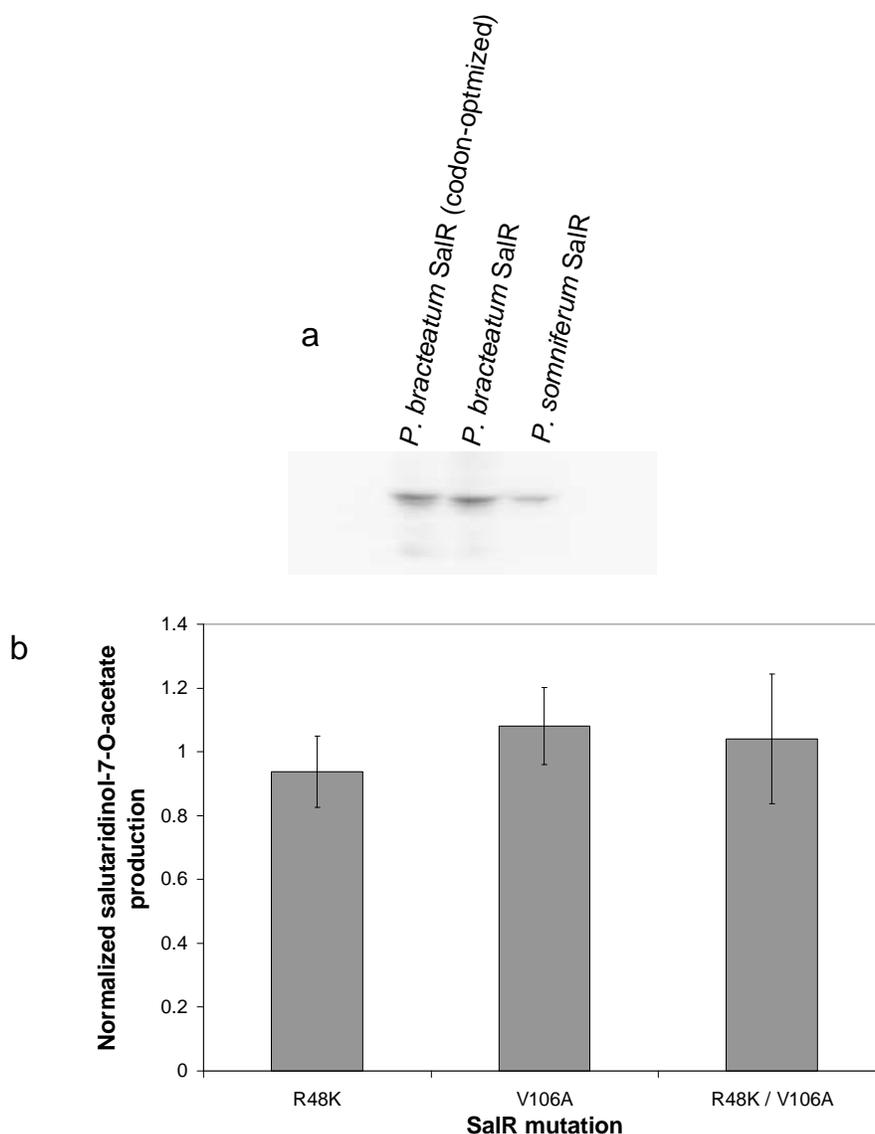


Fig. 5.2. Testing of SalR variants. **(a)** Western blotting using SalR constructs tagged with a C-terminal V5 epitope and detected with the anti-V5 HRP antibody and a chemiluminescent assay. Lanes are loaded with equal amounts of total protein from yeast lysates (50 μ g). **(b)** Comparison of salutaridinol-7-*O*-acetate production by reticuline-producing yeast strains expressing yCYP2D6, SalATopt, and a SalR variant. Values are normalized to production of salutaridinol-7-*O*-acetate from a strain expressing the codon-optimized *P. bracteatum* SalR lacking any amino acid substitutions.

Identification of the salutaridinol peak proved difficult in many of our strains as it appeared as a small tail on the reticuline peak since these metabolites have the same

molecular weight. However, one strain expressing yCYP2D6 and the codon-optimized *P. bracteatum* SalR from TEF6 promoters from a high copy plasmid (pCS1605; Table 5.1) produced a clearly discernable second peak identifiable as salutaridinol by its fragmentation pattern (Fig. 5.3)⁷⁶. We also observed a decrease but not total depletion of salutaridine. Full conversion of this substrate is unlikely since SalR is substrate inhibited at low salutaridine concentrations with a K_i of 140 μM ⁸⁷. This may also be a reason so little activity is observed in other strains even when the same *SalR* sequence is expressed. Low salutaridine production is particularly an issue when using SalSyn, but even in the case of CYP2D6 where salutaridinol is observable, it is only estimated to reach a maximum concentration of 75 μM in the growth media with a presumably similar intracellular concentration based on a passive diffusion transport mechanism for BIA molecules.

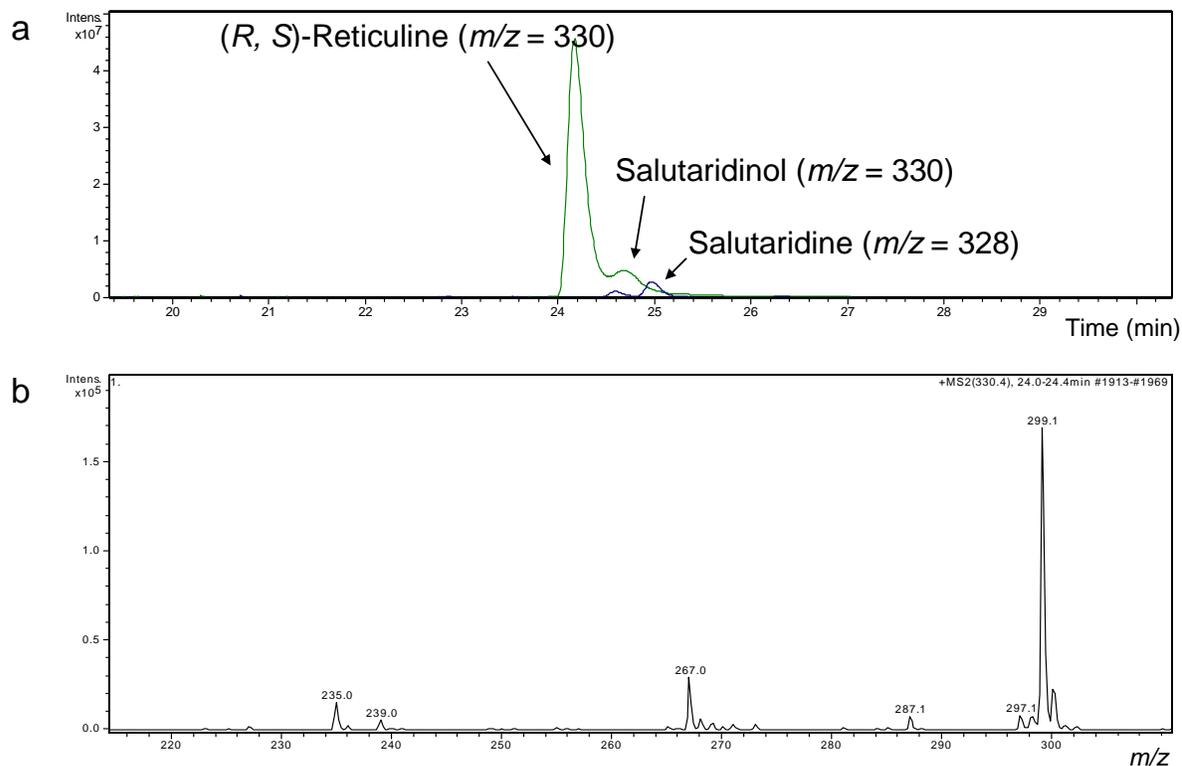


Fig. 5.3. Salutaridinol production. **(a)** Salutaridinol appears as a second peak following reticuline elution at the same molecular weight ($m/z = 330$, green). The substrate peak salutaridine is shown in blue ($m/z = 328$). **(b)** The fragmentation pattern of the peak for $m/z = 330$ at 24.4 min shows ions 299, 287, 267, and 239 that are characteristic for salutaridinol⁷⁶. Additional ions 273, 255, and 241 are visible at higher resolution. Yeast strains not expressing SalR did not show this additional peak or this fragmentation pattern for the $m/z = 330$ ion.

SalR is also an NADPH-dependent enzyme and we suspect that cofactor limitation may be an issue as NADPH is generated in only a few endogenous reactions: the two dehydrogenase reactions of the pentose-phosphate pathway, the NADP⁺-dependent isocitrate dehydrogenase reaction, the NADPH⁺-dependent acetaldehyde dehydrogenase reaction, and the reaction catalyzed by malic enzyme⁸⁸. As the flux through the pentose-phosphate pathway is tightly regulated, the shunt pathway including the mitochondrial malic enzyme was previously identified as a metabolic engineering target for creating an additional source of NADPH. Expression of a short version of

Mae1p lacking the first 90 nucleotides coding for the mitochondrial targeting sequence may be a way to increase the availability of cytosolic NADPH for the SalR reaction. We may also engineer the enzyme to have altered cofactor specificity beyond what was obtained with the R48K mutation.

5.2.3. Expression of salutaridinol-7-O-acetyltransferase and development of a process for thebaine production

The next cloned enzyme in the morphinan branch is the acetyl-CoA-dependent salutaridinol 7-O-acetyltransferase (SalAT) from opium poppy^{89, 90}. Salutaridinol-7-O-acetate undergoes a subsequent spontaneous allylic elimination to form thebaine at pH 8-9. However, at pH 6-7 the product [8,9-dihydro-5H-2,12-dimethoxy-1-hydroxy-7-methyl-dibenz[d,f]azoninium]acetate is formed which can be reduced by NaBH₄ to neodihydrothebaine⁸⁹ (Fig. 5.4).

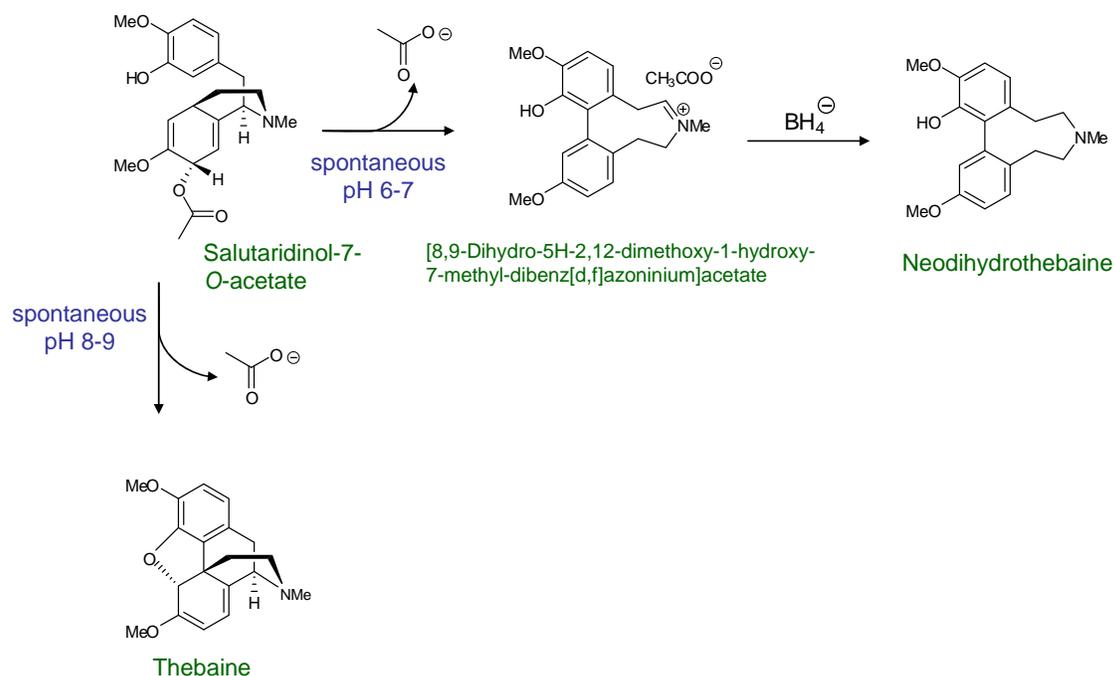


Fig. 5.4. Pathways for the formation of thebaine and neodihydrothebaine. These two very different molecules are derived from salutaridinol-7-*O*-acetate depending on pH. At pH 8-9, loss of the acetyl group leads to thebaine whereas at pH 6-7, loss of the acetyl group leads to [8,9-Dihydro-5H-2,12-dimethoxy-1-hydroxy-7-methyl-dibenz[d,f]azoninium]acetate which can be reduced to neodihydrothebaine.

A seemingly insignificant but reproducible peak was observed for salutaridinol-7-*O*-acetate when using the *P. somniferum* SalAT coding sequence in our system without additional considerations for pH optimization. This peak at $m/z = 372$ appears when using any combination of yCYP2D6 or SalSyn and *P. somniferum* or *P. bracteatum* SalR variants co-expressed with SalAT and is not present in the absence of SalAT (Fig. 5.5). There was no significant difference in salutaridinol-7-*O*-acetate accumulation observed between the various combinations of preceding enzymes tested so we sought other ways to optimize our strains.

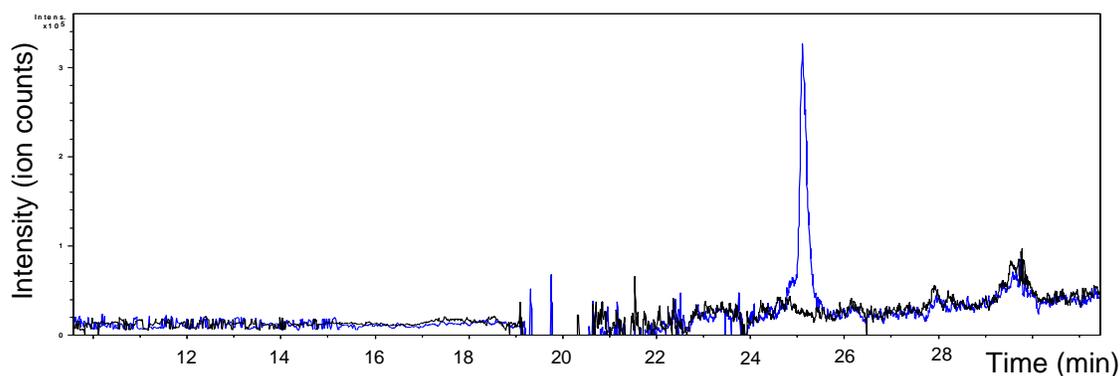


Fig. 5.5. LC-MS analysis of salutaridinol-7-*O*-acetate. Extracted ion chromatograms of $m/z = 372$ for a strain expressing yCYP2D6/hCPR1, PbSalR, and SalAT (blue) and a strain expressing yCYP2D6/hCPR1, and PbSalR only (black); both strains also express the necessary enzymes for reticuline production from norlaudanosoline. The peak identified as salutaridinol-7-*O*-acetate elutes at 25 min and co-elutes with a 313 ion consistent with the loss of the acetyl group in the mass spectrometer.

Using a construct in which the native SalAT sequence was tagged with a V5 epitope for Western blotting, no protein was observed in either the soluble fraction or aggregated with the cell pellet. Codon-optimization greatly increased expression of SalAT in our *S. cerevisiae* strains (Fig. 5.6a).

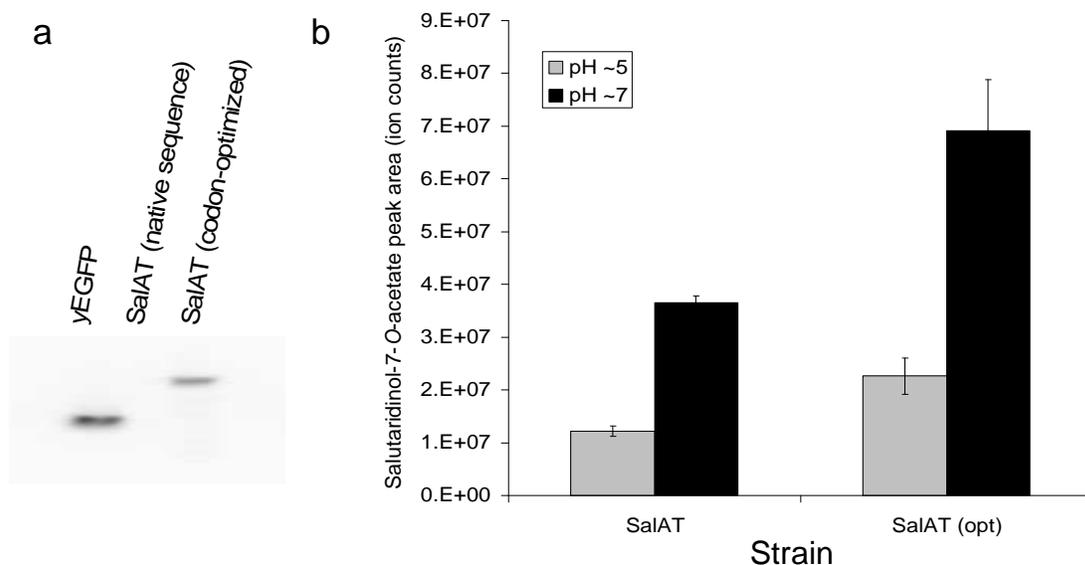


Fig. 5.6. Codon-optimization increases SalAT expression. **(a)** Western blot of SalAT constructs tagged with a C-terminal V5 epitope and incubated with the anti-V5 HRP antibody for chemiluminescent detection. Expression of SalAT using the native sequence is not observed, but the codon-optimized SalAT is expressed at levels comparable to a highly-expressed yEGFP. Lanes are equally loaded with 50 μg total protein. **(b)** Comparison of salutaridinol-7-*O*-acetate production from strains expressing SalSyn/AtATR, PsSalR, and either SalAT or the codon-optimized SalAT (CSY514 and CSY515). Cultures were grown at pH \sim 5 (standard synthetic complete media; grey) or buffered to maintain pH \sim 7 with 10 mM phosphate buffer (black).

Once expression was improved, enzymatic activity *in vivo* still remained a major hurdle with little observable increase in salutaridinol-7-*O*-acetate production. The K_m for salutaridinol is relatively low at 9 μM ⁹⁰ but still may be limiting in our system in addition to acetyl-CoA requirements. Furthermore, the pH optimum for SalAT is 7-9 with a temperature optimum of 47°C⁹⁰. These are far from yeast physiological conditions of pH

~5 and 30°C. Yeast cells are also more adaptable to acidic rather than alkaline pH, and only yeast strains evolved to be thermotolerant tend to grow at temperatures approaching or exceeding 37°C. It was quickly demonstrated that yeast cells inoculated into unbuffered media at pH 8 were able to grow but the alkaline pH was not maintained, and growth was severely retarded in media buffered to maintain the pH at 8. Other enzymatic activities, specifically CYP2D6, were also compromised at pH > 7. However, at pH ~7, cell viability is not noticeably impaired and we observed an increase in salutaridinol-7-*O*-acetate accumulation up to 3-fold in several of our strains (Fig. 5.6b and 5.7b). At this higher pH, we were also able to discern a significant difference between strains expressing SalSyn versus CYP2D6. Strains expressing SalSyn accumulated ~2-fold more salutaridinol-7-*O*-acetate compared to strains expressing CYP2D6, indicating that SalSyn has higher activity when functioning as part of the larger pathway (Fig. 5.7).

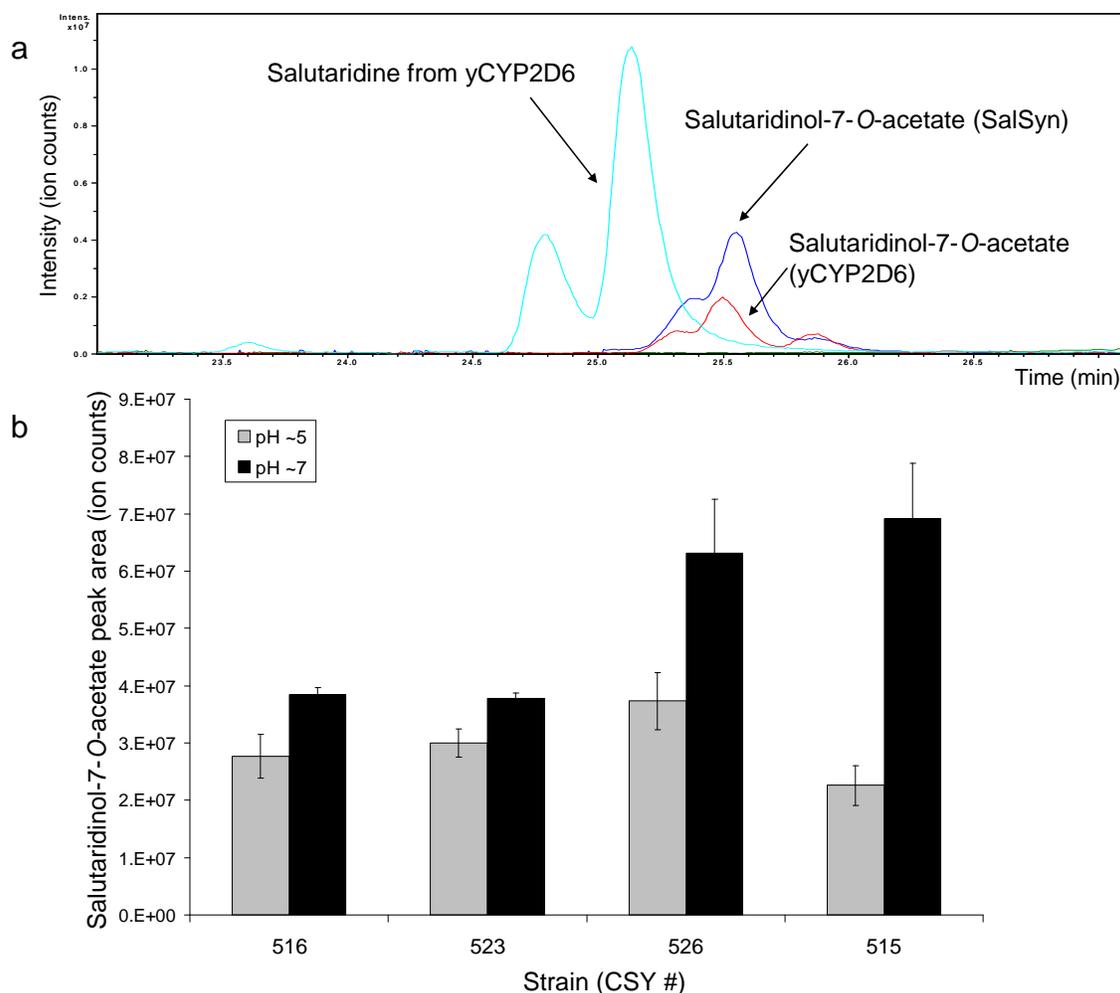


Fig. 5.7. Comparison of CYP2D6 and SalSyn for salutaridinol-7-*O*-acetate production. **(a)** Extracted ion chromatograms of salutaridine ($m/z = 328$; cyan) and salutaridinol-7-*O*-acetate ($m/z = 372$; red) from CSY523 expressing CYP2D6/hCPR1, PsSalR, and PsSalAT(opt); and salutaridine ($m/z = 328$; green, no peak) and salutaridinol-7-*O*-acetate ($m/z = 372$; blue) from CSY515 expressing SalSyn/AtATR1, PsSalR, and PsSalAT(opt). **(b)** Comparison of salutaridinol-7-*O*-acetate peak areas of strains expressing CYP2D6/hCPR1 and PbSalR or PsSalR (CSY516 and CSY523) and strains expressing SalSyn/AtATR. Data from cultures grown at pH ~5 (grey) and pH ~7 (black) are shown.

Based on this data, we chose to revisit SalSyn and optimization of the reductase partner using production of salutaridinol-7-*O*-acetate as a benchmark. The best SalR variant proved to be the original *P. somniferum* sequence (Fig. 5.7b), and this was used in combination with the codon-optimized SalAT to test various strain backgrounds with no

reductase or an integrated copy of AtATR1, hCPR1, or PsCPR1 behind the TEF1 promoter (Fig. 5.8a). We also found that one design with the N-terminal signal sequence replaced by that of bovine CYP17 outperformed the native SalSyn (Fig. 5.8b).

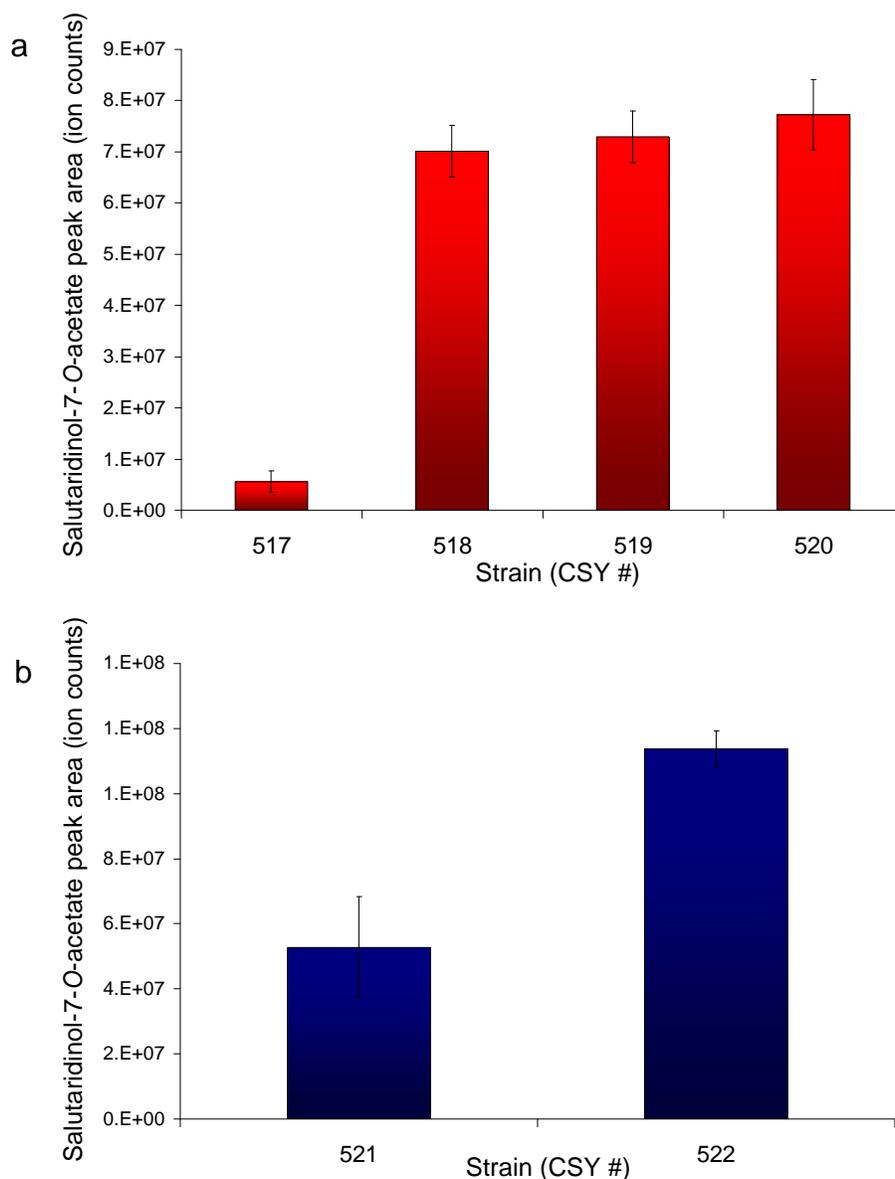


Fig. 5.8. Comparison of SalSyn reductase partners and signal sequences. **(a)** Salutaridinol-7-*O*-acetate production from strains expressing SalSyn, PbSalR, and PsSalAT (opt) along with no reductase partner (CSY517), hCPR1 (CSY518), AtATR1 (CSY519), or PsCPR (CSY520). **(b)** Salutaridinol-7-*O*-acetate production from strains expressing PsCPR, PsSalR, and either the native SalSyn sequence (CSY521) or a version of SalSyn with the first 12 amino acids replaced by the bovine CYP17 16 amino acid N-terminal signal sequence.

Although we are seeing salutaridinol-7-*O*-acetate accumulation, we would like to push this towards thebaine production which should proceed spontaneously at pH 8-9. Based on the proposed mechanism, it appears that a histidine in the active site stabilizes the leaving group and can determine the preferred location of nucleophilic attack while the molecule is still bound to SalAT such that the pH at the time of reaction is critical. Our own experimental evidence also shows that increasing the pH following the growth phase does not alter the ratio of salutaridinol-7-*O*-acetate to thebaine. We therefore sought a two-stage process in which the cells could grow and produce other intermediates up to salutaridinol shown to accumulate at physiological pH. SalAT could then be expressed along with a simultaneous increase in pH. Rather than use a promoter that requires addition of an exogenous inducer such as GAL1 or CUP1, we examined the HSP30 promoter which was shown to be induced as the cells entered stationary phase. This allows automatic induction of SalAT as the expression levels of other enzymes expressed from constitutive promoters are declining (Fig. 5.9a).

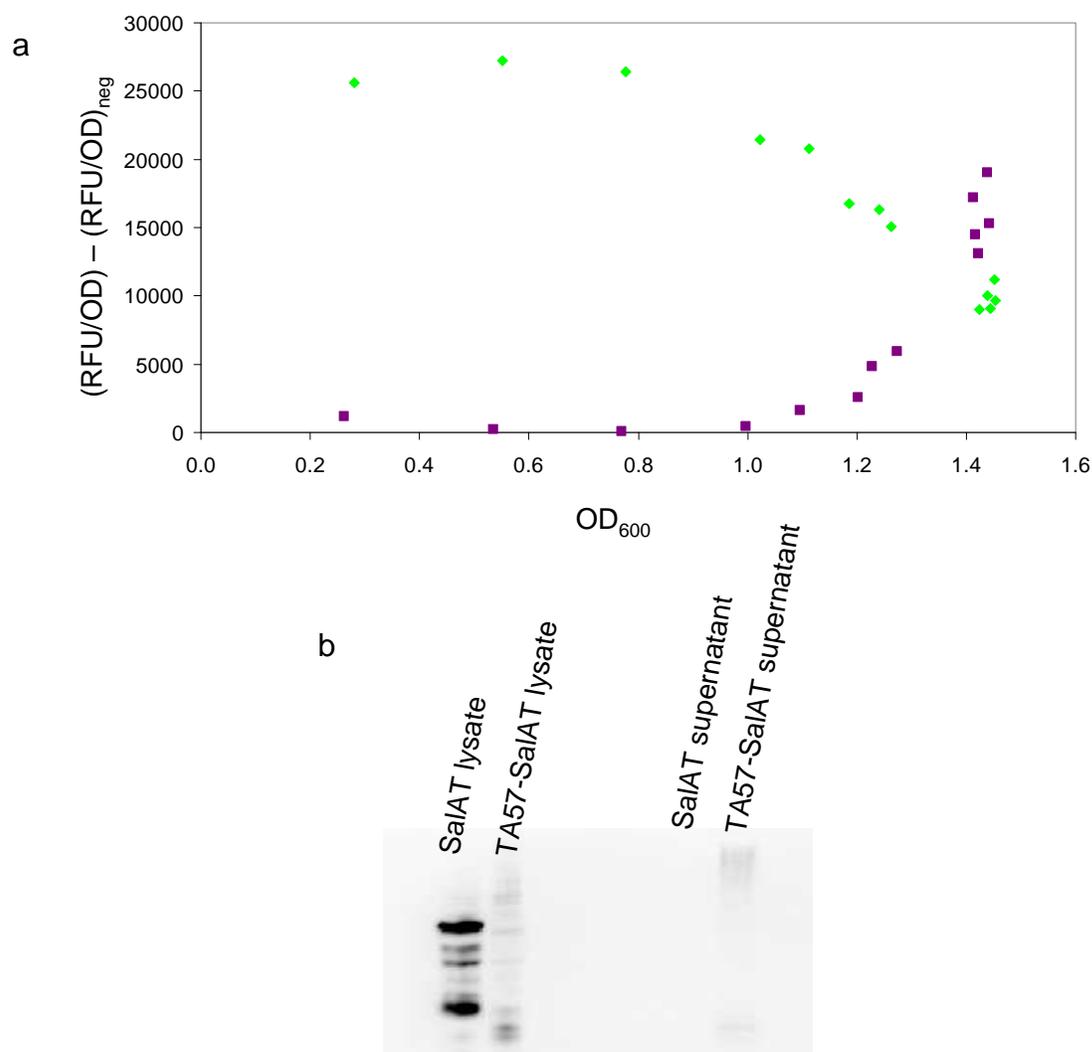


Fig. 5.9. Alternative promoter and signal sequence for SalAT. **(a)** Time course data of P_{TEF1} :yEGFP (green diamonds) and P_{HSP30} :yEGFP (purple squares). As the cells reach stationary phase, the TEF1 promoter expression decreases while the HSP30 promoter turns on. Relative fluorescence and OD_{600} were measured in a 96-well plate format and normalized to a negative control with no yEGFP expression. **(b)** Western blot of yeast lysates (50 μ g total protein) and supernatants concentrated \sim 250-fold, 20 μ l loaded. Expression and localization of the yeast codon-optimized SalAT is compared to the same coding sequence with a TA57 leader sequence for processing through the secretory pathway.

In addition, we looked at secretion of SalAT since the intracellular pH does not necessarily correspond to the extracellular pH which is much easier to measure and manipulate. The modified α -factor leader sequence in which the C-terminus was

modified from “SLDKR” so “SMAKR” was tested as well as the TA57 leader sequence lacking N-linked glycosylation sites. The spacer sequences were EEAEAEAEPK and EEGEPK, respectively⁹¹. We analyzed the growth media of a yeast strain expressing a tagged version of the secreted SalAT to examine protein levels (Fig. 5.9b). Intracellular TA57-SalAT levels were low although little protein was detected in the concentrated growth media. This could be due in large part to proteolysis either before or after the culture was harvested.

We tested each of these SalAT constructs in a reticuline-producing strain background also expressing *P_{TEF1}:hCPRI* from the chromosome, and *P_{TEF6}:yCYP2D6* and *P_{TEF6}:PbSalR* from a high-copy plasmid. For constructs expressing SalAT from the HSP30 promoter, phosphate buffer was added to increase the pH to ~7.5 as the cells entered stationary phase, typically following ~24 h growth in the presence of substrate. Constitutive SalAT constructs were tested at physiological pH or buffered to pH 7-7.5 upon backdilution in the presence of substrate.

For secreted enzymes, no salutaridinol-7-*O*-acetate production was observed under any pH conditions. This could be attributable to low or incorrect expression of the secreted enzymes or inactivity of SalAT in the growth media due to lack of acetyl-CoA or other factors. For the SalAT lacking a signal sequence, salutaridinol-7-*O*-acetate production was greater when using a strong constitutive promoter compared to the HSP30 promoter. It is unclear whether this is related to differences in expression levels between the HSP30 and TEF6 promoters or the timing of induction and pH change.

Further investigations into the mechanism of SalAT and thebaine formation are necessary for the rational engineering of improved yeast strains. While we are able to

accumulate more salutaridinol-7-*O*-acetate at pH 7 without severely compromising growth, yeast are not tolerant to more alkaline pH and higher temperatures. Even if we were to evolve such a robust strain, preceding enzymatic activities may suffer.

It is also possible that acetyl-CoA is limiting, in which case, previous studies have shown that this pool can be increased for use in a heterologous pathway⁹². Using the work engineering the pyruvate dehydrogenase bypass as a guide, we plan to construct strains to overexpress both the *S. cerevisiae* *ALD6* and the *Salmonella acs* L641P mutant. It is important to note that overexpression of *ALD6* may also increase the NADPH pool for the SalR enzyme activity.

Many of the strains we constructed for salutaridinol-7-*O*-acetate production, in the hopes of producing thebaine, accumulated similar levels of the precursor. However, one clear trend was that the combination of all *P. somniferum* enzymes was superior although CYP2D6 produced more salutaridine and *P. bracteatum* SalR produced more salutaridinol. This result, along with RNAi studies in plants, suggests the existence of protein-protein interactions^{93, 94}. Specifically, in opium poppy, depletion of SalAT results in accumulation of salutaridine rather than salutaridinol, indicating that both SalR and SalAT are required for production of salutaridinol⁹³. In addition, overexpression of SalAT in plants resulted in increased morphine, codeine, and thebaine production, indicating that SalAT is limiting in the native pathway⁹³. In our recombinant host, it is unclear whether the lack of other poppy proteins is limiting the flux to thebaine. If metabolon formation including additional uncloned genes (which may have undesired products) is an issue, it is unclear how to get around this in our host organism. One strategy is testing of variants from other organisms, particularly those known to

accumulate thebaine such as *P. bracteatum*. Other engineering and evolutionary strategies remain to be seen.

5.3. Discussion

The morphinan alkaloids represent an important branch of BIA metabolites of interest to the pharmaceutical industry. This work marks the first effort towards the development of a microbial host to perform the final conversion steps leading to thebaine production. We successfully demonstrated conversion of the commercially available substrate norlaudanoline to salutaridinol-7-*O*-acetate, the precursor to thebaine, building upon a previously characterized yeast strain background⁴. In our best strains, we are able to achieve ~7-8% conversion of reticuline which is estimated to correspond to salutaridinol-7-*O*-acetate titers of up to 15 mg l⁻¹. We also attempted to optimize the activity of the recently-cloned SalSyn (CYP719B) and were unable to observe salutaridine accumulation although the presence of downstream metabolites indicated that this enzyme is functional in our host strains. Further optimization of this first step using the native plant enzyme SalSyn may alleviate the bottleneck in this pathway. Additional optimization of SalR and SalAT activities may also be required to develop a viable process for production of these molecules. This could be through engineering yeast strains with increased cofactor pools and/or engineering the proteins themselves to have higher activity under physiological conditions. We did observe strains expressing all *P. somniferum* enzyme variants to have the greatest accumulation of salutaridinol-7-*O*-acetate. This is not terribly surprising given previous plant studies indicating protein interactions are required for activity^{93, 94}. Moreover, the lack of expression of additional

downstream enzymes may also limit the flux through this pathway if indeed they are required for complete metabolon formation and function. Finally, the sequencing of a thebaine synthase (THS) activity from one or more BIA producing plants and its incorporation into our yeast hosts may facilitate conversion of salutaridinol-7-*O*-acetate to thebaine under physiological conditions.

5.4. Materials and Methods

5.4.1. Plasmid and yeast strain construction

We obtained restriction enzymes, T4 DNA ligase, and other cloning enzymes from New England Biolabs. We performed polymerase chain reaction (PCR) amplifications using Expand High Fidelity PCR System (Roche). Oligonucleotide synthesis was performed by Integrated DNA Technologies. A list of engineered yeast strains and plasmids is provided (Table 5.1).

We used standard molecular biology techniques to construct the BIA expression vectors⁵⁹. BIA expression constructs contained the 2 μ high-copy yeast origin of replication along with appropriate yeast selection markers and ampicillin resistance. Recombinant enzymes were expressed from the yeast TEF1 promoter and flanked by a CYC1 terminator sequence. We constructed shuttle vectors for subcloning of 1 or 2 cDNA sequences in this fashion. The cDNA sequences for the native *P. somniferum* SalR, *P. somniferum* SalAT, and yeast codon-optimized SalSyn were synthesized by DNA2.0. The *P. bracteatum* SalR sequence was obtained through a series of site directed mutagenesis reactions to make the necessary 13 amino acid substitutions in the *P. somniferum* SalR sequence (QuikChange II, Stratagene). The codon-optimized sequences

for *P. bracteatum* SalR, *P. somniferum* SalAT, and *P. somniferum* CPR were designed using GeneDesigner software (DNA2.0) and assembled using primers designed by DNAWorks2.4 (Center for Information Technology, National Institutes of Health).

We transformed ligation reactions into an electrocompetent *E. coli* strain, DH10B (Invitrogen; F-*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80*dlacZ* Δ M15 Δ *lacX74 deoR recA1 endA1 araD139 Δ (*ara, leu*)7697 *galU galK λ -rpsL nupG*), using a Gene Pulser Xcell System (BioRAD) according to the manufacturer's instructions. We conducted plasmid isolation using the Wizard Plus SV Minipreps DNA purification system (Promega) according to the manufacturer's instructions. Subcloning was confirmed by restriction analysis and sequence verification (Laragen, Inc.). We transformed plasmids into the appropriate *S. cerevisiae* strains using a standard lithium acetate protocol⁶². All yeast strains used in this work were based on the haploid yeast strain W303 α (*MAT α his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*)⁶³. *E. coli* cells were grown on Luria-Bertani media (BD Diagnostics) with 100 μ g/ml ampicillin (EMD Chemicals) for plasmid maintenance, and *S. cerevisiae* cells were grown in synthetic complete media (BD Diagnostics) supplemented with the appropriate dropout solution for plasmid maintenance (Calbiochem).*

Table 5.1. Engineered yeast strains and plasmids.

<u>Strain</u>	<u>Plasmid(s)</u>	<u>Integrated constructs</u>	<u>Plasmid-based constructs</u>
CSY288		<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	
CSY409		<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	
CSY488		<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-hCPR-loxP-LEU2</i>	
CSY506		<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-PsCPR(opt)-loxP-LEU2</i>	
CSY514	pCS1357, pCS1449	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-PsSalR, P_{TEF1}-PsSalAT, P_{TEF1}-SalSyn</i>
CSY515	pCS1449, pCS1540	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-PsSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-SalSyn</i>
CSY516	pCS1480, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-hCPR-loxP-LEU2</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalAT, P_{TEF6}-yCYP2D6</i>
CSY517	pCS1558, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-CYP17-SalSynB</i>
CSY518	pCS1558, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-hCPR-loxP-LEU2</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-CYP17-SalSynB</i>
CSY519	pCS1558, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-CYP17-SalSynB</i>
CSY520	pCS1558, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-PsCPR(opt)-loxP-LEU2</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-CYP17-SalSynB</i>
CSY521	pCS1449, pCS1540	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-PsCPR(opt)-loxP-LEU2</i>	<i>P_{TEF1}-PsSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-SalSyn</i>
CSY522	pCS1558, pCS1540	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-PsCPR(opt)-loxP-LEU2</i>	<i>P_{TEF1}-PsSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-CYP17-SalSynB</i>
CSY523	pCS1480, pCS1540	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-hCPR-loxP-LEU2</i>	<i>P_{TEF1}-PsSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-yCYP2D6</i>
CSY526	pCS1449, pCS1507	<i>his3::P_{TEF1}-Ps6OMT, leu2::P_{TEF1}-PsCNMT, ura3::P_{TEF1}-Ps4'OMT, trp1::P_{TEF1}-AtATR1-loxP-Kan^R</i>	<i>P_{TEF1}-PbSalR, P_{TEF1}-PsSalATopt, P_{TEF1}-SalSyn</i>
	pCS1508		<i>P_{TEF6}-yCYP2D6, P_{TEF6}-PsSalATopt</i>
	pCS1554		<i>P_{TEF1}-PbSalRopt</i>
	pCS1605		<i>P_{TEF6}-yCYP2D6, P_{TEF6}-PbSalRopt</i>
	pCS1606		<i>P_{TEF6}-PsSalATopt</i>
	pCS1607		<i>P_{HSP30}-PsSalATopt</i>
	pCS1608		<i>P_{HSP30}-α PsSalATopt</i>
	pCS1609		<i>P_{HSP30}-TA57-SalATopt</i>

5.4.2. Cell growth conditions

For metabolite assays, overnight cultures of yeast cells were grown to stationary phase and backdiluted 1:20 into 2-3 ml fresh media plus the appropriate substrate, typically norlaudanosoline or laudanosoline added from a 20 mM stock to a final concentration of 4 mM. For buffered cultures, 0.1 M phosphate at pH 8.0 was added as a 10X stock to yield a final pH of ~7 which was maintained during cell growth. The concentration of yeast nitrogen base, ammonium sulfate, amino acids, and dextrose were adjusted so that the final concentration was consistent with standard media conditions. Cells were grown ~48 hr for most assays such that stationary phase was reached to allow for more consistent and reproducible results. No significant difference was observed if cells were grown for an additional 2-4 days.

5.4.3. Analysis of metabolite production

We evaluated BIA metabolite levels by LC-MS/MS analysis of cell extracts and growth media. At appropriate time points, aliquots of yeast cultures were centrifuged to recover cells as pellets and allow collection of the growth media. We analyzed the growth media or an appropriate dilution directly by LC-MS/MS. Samples were run on an Agilent ZORBAX SB-Aq 3 x 250 mm, 5 μ m column with 0.1% acetic acid as solvent A and methanol as solvent B. We used a gradient elution to separate the metabolites of interest as follows: 0-10 min at 100% A, 10-30 min 0-90% B, 30-35 min 90-0% B, followed by a 5 min equilibration at 100% A between samples. Following LC separation, metabolites were injected into an Agilent 6320 ion trap MSD for detection and identification. We verified chromatogram data through at least three independent experiments and from multiple strains where appropriate. Quantification of metabolites was based on the integrated area of the extracted ion chromatogram peaks⁶⁴ calculated using DataAnalysis for 6300 Series Ion Trap LC/MS Version 3.4 (Bruker Daltonik GmbH) and reported as the mean \pm s.d. When appropriate, we normalized the measured levels to a metabolite peak of known concentration in the growth media.

5.4.4. Analysis of protein levels through Western blotting

We constructed plasmids for Western blotting experiments by cloning the C-terminal epitope tag(s) from pYES-NT/A (Invitrogen) into our standard TEF1 expression vector followed by subcloning of the enzyme of interest. We transformed individual plasmids into wild-type yeast cells using a standard lithium acetate protocol. Overnight cultures were grown and backdiluted 1:100 into 100 ml cultures. Cells were grown to OD₆₀₀ ~1.5 and pelleted by centrifugation. The media was removed and cells were washed in 1 ml PBS, pelleted, and resuspended in 0.5 ml Y-PER plus HALT protease

inhibitor (Pierce). Cells were vortexed for ~20 min and the lysate separated by centrifugation. We estimated total protein using the Coomassie Plus Protein Assay Reagent (Pierce) and loaded ~50 μg of each sample onto a protein gel. We used NuPage 4-12% Bis-Tris gels with MES running buffer and transfer buffer according to the manufacturer's instructions (Invitrogen). Proteins were blotted onto a nitrocellulose membrane (Whatman) using a semi-dry transfer cell (Bio-Rad) for 25 min at 25 V. We incubated the membrane with the Anti-V5 HRP antibody (Invitrogen) according to the manufacturer's instructions (Invitrogen) with 5% BSA as the blocking agent. Proteins were detected with the West Pico Super Signal Detection kit (Pierce) and imaged on a ChemiDoc XRS system (Bio-Rad).

5.4.5. Fluorescence quantification

For measurements of HSP30 and TEF1 promoter activity using a fluorescent reporter, overnight cultures were grown in synthetic complete media with 2% dextrose and backdiluted 1:100 into fresh media to begin the time course assay. We measured fluorescence (excitation 485 nm, emission 515 nm) and OD_{600} values on a fluorescence plate reader (TECAN, Safire) at various time points between 4 and >48 hr. Fluorescence values were normalized using the OD_{600} reading and values from a no stain control (wild-type cells) were subtracted.

CHAPTER VI. CONCLUSIONS AND FUTURE WORK

6.1. A new tool for tuning protein expression

Chapter II describes a novel tool for tuning protein expression levels in yeast. By engineering strains with altered expression of the Gal2p permease, we demonstrated a modified induction response from the native GAL1-10 promoter. Specifically, the *GAL2Δ* strain resulted in a linear and homogeneous response such that protein levels could be finely tuned with galactose concentration.

We demonstrated the utility of this system in our yeast strains engineered for the production of benzyloquinoline alkaloids (BIAs). By titrating one of three enzymes in a short heterologous pathway independently, we were able to determine optimal expression levels of each enzyme. We then selected a constitutive promoter from a well-characterized promoter library⁵ that exhibited the desired expression level and made the appropriate substitutions. Constitutive promoters are preferred for this work to avoid the additional induction step and the use of costly inducers. We also observed greater conversion in the constitutively-expressing strains, likely due to the elimination of an uninduced population.

We propose this as a general strategy that can be useful for adjusting protein expression levels for pathway engineering. Although it is not applicable to all systems; for instance, the P450 activities were not tunable in this range, it can greatly simplify the promoter screening and selection process. In addition, this tuning strategy is applicable to many other areas of synthetic biology such as the construction of genetic circuits.

6.2. *De novo* biosynthesis of BIA backbone molecules

While we were able to make significant progress engineering the downstream BIA pathway starting with norlaudanosoline, we ultimately want to synthesize BIA metabolites *de novo* from two molecules of tyrosine. Since this early pathway has not been fully elucidated in plants, we took a bioprospecting approach and combined enzymatic activities from plants, bacteria, humans, and yeast to engineer a unique pathway for the production of norcoclaurine and norlaudanosoline.

We have developed two different synthesis routes for the production of dopamine and 4-hydroxyphenylacetaldehyde (4-HPA) and validated enzymatic activities for each step. However, we are only able to produce very low levels of dopamine and more optimization is required to reach levels sufficient to build the BIA backbone. This may require protein engineering or evolution of the CYP2D6 enzyme used to convert tyramine to dopamine. Similar to other human P450s, CYP2D6 exhibits low activity on a broad range of substrates as we have also used this activity to convert (*R*)-reticuline to salutaridine. Since dopamine and L-dopa activate the yeast oxidative stress response pathway, a reporter gene such as GFP fused to the FUS1 promoter can potentially be used as a high-throughput screen for increased production of either of these molecules⁴².

In addition, we have identified endogenous alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALD) activities that act on the acetaldehyde intermediates 4-HPA and 3,4-DHPA. As single knockouts allowed increased accumulation of norcoclaurine and norlaudanosoline, we are hopeful that combinatorial knockouts will show additive effects.

We have also begun work building protein scaffolds and using leucine zipper domains to co-localize enzymes and increase local metabolite concentrations. These and other strain engineering methods will be required to develop a viable process for the biosynthesis of the BIA backbone.

6.3. Production of the intermediate reticuline and downstream berberine and morphinan alkaloids

We engineered yeast strains for the production of two sets of very diverse BIAs along the berberine and morphinan branches. Currently, we are able to accumulate up to 165 mg l⁻¹ of (*R, S*)-reticuline, ~35 mg l⁻¹ of (*S*)-tetrahydroberberine, and ~15 mg l⁻¹ of salutaridinol-7-*O*-acetate. We put considerable effort into optimizing transcriptional activity of the enzymes used to convert norlaudanosoline to reticuline. However, intracellular transport is limiting in addition to the affinity of the enzymes for norlaudanosoline as it is not the natural substrate. Total biosynthesis of norcoclaurine *in vivo* will remedy both issues assuming an additional P450 activity CY80B1 can be successfully incorporated.

Along the berberine branch, we are able to accumulate (*S*)-scoulerine, (*S*)-tetrahydrocolumbamine, and (*S*)-tetrahydroberberine. Very recently, a protein with (*S*)-cheilanthifoline synthase activity was cloned from *E. californica*⁹⁵. This enzyme which accepts (*S*)-scoulerine as a substrate will allow access to other metabolites along the sanguinarine branch. In addition, the BIA palmatine can be produced from (*S*)-tetrahydrocolumbamine once the enzyme opening this branch through columbamine is cloned. And finally, (*S*)-tetrahydroberberine can be oxidized to the final product

berberine in our yeast strains once the (*S*)-tetrahydroberberine oxidase (STOX) activity is cloned from plant hosts or an alternative enzyme is identified. Although we see efficient conversion to (*S*)-tetrahydrocolumbamine, we only see ~10% conversion to (*S*)-tetrahydroberberine which makes this step a target for additional optimization.

Preliminary strains were constructed for production of morphinan alkaloids which leave room for improvement at every step. Once again, we observed low activity from both P450s tested in our system for the production of salutaridine. This is not surprising for the human P450 CYP2D6 but higher activity is expected for the plant enzyme SalSyn (CYP719B). We tested multiple reductase partners here to no avail and feel this may be the limitation in our system. Low accumulation of salutaridine also affects SalR activity as the enzyme is inhibited at low substrate concentrations. Testing of different SalR variants and mutants as well as codon-optimization did not yield significant improvements. SalR is also an NADPH-dependent enzyme so the redox balance of the cell may need to be taken into consideration. For the SalAT step, we were able to significantly improve expression through codon-optimization but are still plagued by low activity. Since this step also requires acetyl-CoA, other metabolic engineering strategies can be applied to increase this cofactor pool if necessary. In addition, the pH and temperature optima for this enzyme are not compatible with yeast physiological conditions. While we can increase the pH to ~7 in buffered media, we cannot operate at pH>8 which is the reported requirement for thebaine production. Ultimately, a 2-stage process will likely be required. Our early attempts to do this entirely *in vivo* were unsuccessful, and it is likely that the final step must be performed *in vitro*. Alternatively,

the cloning of thebaine synthase (THS) from poppy plants and its incorporation into this heterologous pathway may facilitate efficient thebaine production in our yeast hosts.

6.4. Construction of a strain to produce downstream BIAs from tyrosine

The end goal of this work is the construction of yeast strains capable of performing the total biosynthesis of complex BIAs from tyrosine. To reach this goal, much more optimization is required in each segment of the pathway, particularly the upstream portion. We have begun to merge the two pathways by producing reticuline from dopamine and coclaurine from exogenous dopamine and 4-HPA, demonstrating the feasibility of combining these segments. Also, as predicted, conversion is greater when the BIA precursor is synthesized *in vivo* rather than added to the media, avoiding transport limitations. Presuming we can construct a strain to produce norcoclaurine and norlaudanoline, there is no reason to believe we cannot begin to piece together a stable strain to produce reticuline from tyrosine. From there, plasmid-based expression of downstream enzymes should be sufficient to extend these pathways down various branches. The work described here is significant as it marks the first demonstration of reconstructing this very important pathway in a single microbial host.

REFERENCES

1. Nevoigt, E. Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **72**, 379-412 (2008).
2. Bailey, J.E. Toward a science of metabolic engineering. *Science* **252**, 1668-1675 (1991).
3. Hawkins, K.M. & Smolke, C.D. The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in *Saccharomyces cerevisiae*. *J Biol Chem* **281**, 13485-13492 (2006).
4. Hawkins, K.M. & Smolke, C.D. Production of benzyloquinoline alkaloids in *Saccharomyces cerevisiae*. *Nat Chem Biol* **4**, 564-573 (2008).
5. Nevoigt, E. et al. Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **72**, 5266-5273 (2006).
6. Ziegler, J. & Facchini, P.J. Alkaloid biosynthesis: metabolism and trafficking. *Annu Rev Plant Biol* **59**, 735-769 (2008).
7. Funk, M. et al. Vector systems for heterologous expression of proteins in *Saccharomyces cerevisiae*. *Methods Enzymol* **350**, 248-257 (2002).
8. Solow, S.P., Sengbusch, J. & Laird, M.W. Heterologous Protein Production from the Inducible MET25 Promoter in *Saccharomyces cerevisiae*. *Biotechnol Prog* **21**, 617-620 (2005).
9. Koller, A., Valesco, J. & Subramani, S. The CUP1 promoter of *Saccharomyces cerevisiae* is inducible by copper in *Pichia pastoris*. *Yeast* **16**, 651-656 (2000).
10. Mascorro-Gallardo, J.O., Covarrubias, A.A. & Gaxiola, R. Construction of a CUP1 promoter-based vector to modulate gene expression in *Saccharomyces cerevisiae*. *Gene* **172**, 169-170 (1996).
11. Koshland, D.E., Jr. The era of pathway quantification. *Science* **280**, 852-853 (1998).
12. Verma, M., Bhat, P.J. & Venkatesh, K.V. Quantitative analysis of GAL genetic switch of *Saccharomyces cerevisiae* reveals that nucleocytoplasmic shuttling of Gal80p results in a highly sensitive response to galactose. *J Biol Chem* **278**, 48764-48769 (2003).
13. Louis, M. & Becskei, A. Binary and graded responses in gene networks. *Sci STKE* **2002**, PE33 (2002).
14. Peng, G. & Hopper, J.E. Gene activation by interaction of an inhibitor with a cytoplasmic signaling protein. *Proc Natl Acad Sci U S A* **99**, 8548-8553 (2002).
15. Horak, J. & Wolf, D.H. Catabolite inactivation of the galactose transporter in the yeast *Saccharomyces cerevisiae*: ubiquitination, endocytosis, and degradation in the vacuole. *J Bacteriol* **179**, 1541-1549 (1997).
16. Lamphier, M.S. & Ptashne, M. Multiple mechanisms mediate glucose repression of the yeast GAL1 gene. *Proc Natl Acad Sci U S A* **89**, 5922-5926 (1992).
17. Nehlin, J.O., Carlberg, M. & Ronne, H. Control of yeast GAL genes by MIG1 repressor: a transcriptional cascade in the glucose response. *Embo J* **10**, 3373-3377 (1991).

18. Ramos, J., Szkutnicka, K. & Cirillo, V.P. Characteristics of galactose transport in *Saccharomyces cerevisiae* cells and reconstituted lipid vesicles. *J Bacteriol* **171**, 3539-3544 (1989).
19. Schell, M.A. & Wilson, D.B. Purification and properties of galactokinase from *Saccharomyces cerevisiae*. *J Biol Chem* **252**, 1162-1166 (1977).
20. Acar, M., Becskei, A. & van Oudenaarden, A. Enhancement of cellular memory by reducing stochastic transitions. *Nature* **435**, 228-232 (2005).
21. Ruhela, A. et al. Autoregulation of regulatory proteins is key for dynamic operation of GAL switch in *Saccharomyces cerevisiae*. *FEBS Lett* **576**, 119-126 (2004).
22. Venkatesh, K.V., Bhat, P.J., Kumar, R.A. & Doshi, P. Quantitative model for Gal4p-mediated expression of the galactose/melibiose regulon in *Saccharomyces cerevisiae*. *Biotechnol Prog* **15**, 51-57 (1999).
23. Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L. & Keasling, J.D. Homogeneous expression of the P(BAD) promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology* **147**, 3241-3247 (2001).
24. Morgan-Kiss, R.M., Wadler, C. & Cronan, J.E., Jr. Long-term and homogeneous regulation of the *Escherichia coli* araBAD promoter by use of a lactose transporter of relaxed specificity. *Proc Natl Acad Sci U S A* **99**, 7373-7377 (2002).
25. Yen, K., Gitsham, P., Wishart, J., Oliver, S.G. & Zhang, N. An improved tetO promoter replacement system for regulating the expression of yeast genes. *Yeast* **20**, 1255-1262 (2003).
26. Belli, G., Gari, E., Piedrafita, L., Aldea, M. & Herrero, E. An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res* **26**, 942-947 (1998).
27. Gari, E., Piedrafita, L., Aldea, M. & Herrero, E. A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* **13**, 837-848 (1997).
28. Stagoj, M.N., Comino, A. & Komel, R. Fluorescence based assay of GAL system in yeast *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* **244**, 105-110 (2005).
29. Bhat, P.J. & Venkatesh, K.V. Stochastic variation in the concentration of a repressor activates GAL genetic switch: implications in evolution of regulatory network. *FEBS Lett* **579**, 597-603 (2005).
30. Verma, M., Bhat, P.J. & Venkatesh, K.V. Expression of GAL genes in a mutant strain of *Saccharomyces cerevisiae* lacking GAL80: quantitative model and experimental verification. *Biotechnol Appl Biochem* **39**, 89-97 (2004).
31. Wolfe, K.H. & Shields, D.C. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708-713 (1997).
32. Bhat, P.J., Oh, D. & Hopper, J.E. Analysis of the GAL3 signal transduction pathway activating GAL4 protein-dependent transcription in *Saccharomyces cerevisiae*. *Genetics* **125**, 281-291 (1990).
33. Thoden, J.B., Sellick, C.A., Timson, D.J., Reece, R.J. & Holden, H.M. Molecular structure of *Saccharomyces cerevisiae* Gal1p, a bifunctional galactokinase and transcriptional inducer. *J Biol Chem* **280**, 36905-36911 (2005).

34. Guide to Yeast Genetics and Molecular and Cellular Biology, Vol. 194. (Elsevier Academic Press, San Francisco; 2004).
35. Longtine, M.S. et al. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953-961 (1998).
36. Gietz, R.D. & Woods, R.A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350**, 87-96 (2002).
37. Mateus, C. & Avery, S.V. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast* **16**, 1313-1323 (2000).
38. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19-27 (1989).
39. Li, J. et al. Green fluorescent protein in *Saccharomyces cerevisiae*: real-time studies of the GAL1 promoter. *Biotechnol Bioeng* **70**, 187-196 (2000).
40. Facchini, P.J. ALKALOID BIOSYNTHESIS IN PLANTS: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 29-66 (2001).
41. Facchini, P.J. & De Luca, V. Differential and tissue-specific expression of a gene family for tyrosine/dopa decarboxylase in opium poppy. *J Biol Chem* **269**, 26684-26690 (1994).
42. Staleva, L., Hall, A. & Orlow, S.J. Oxidative stress activates FUS1 and RLM1 transcription in the yeast *Saccharomyces cerevisiae* in an oxidant-dependent manner. *Mol Biol Cell* **15**, 5574-5582 (2004).
43. Ikram Ul, H. & Ali, S. Microbiological transformation of L-tyrosine to 3,4-dihydroxyphenyl L-alanine (L-dopa) by a mutant strain of *Aspergillus oryzae* UV-7. *Curr Microbiol* **45**, 88-93 (2002).
44. Haavik, J. L-DOPA is a substrate for tyrosine hydroxylase. *J Neurochem* **69**, 1720-1728 (1997).
45. Wichers, H.J. et al. Cloning, expression and characterisation of two tyrosinase cDNAs from *Agaricus bisporus*. *Appl Microbiol Biotechnol* **61**, 336-341 (2003).
46. 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).
47. Cooper, R.A. On the amine oxidases of *Klebsiella aerogenes* strain W70. *FEMS Microbiol Lett* **146**, 85-89 (1997).
48. Urban, P., Andersen, J.K., Hsu, H.P. & Pompon, D. Comparative membrane locations and activities of human monoamine oxidases expressed in yeast. *FEBS Lett* **286**, 142-146 (1991).
49. Roh, J.H. et al. Purification, cloning, and three-dimensional structure prediction of *Micrococcus luteus* FAD-containing tyramine oxidase. *Biochem Biophys Res Commun* **268**, 293-297 (2000).
50. Minami, H. et al. Microbial production of plant benzylisoquinoline alkaloids. *Proc Natl Acad Sci U S A* **105**, 7393-7398 (2008).

51. Hoover, L.K., Moo-Young, M. & Legge, R.L. Biotransformation of dopamine to norlaudanosoline by *Aspergillus niger*. *Biotechnol Bioeng* **38**, 1029-1033 (1991).
52. Soetens, O., De Craene, J.O. & Andre, B. Ubiquitin is required for sorting to the vacuole of the yeast general amino acid permease, Gap1. *J Biol Chem* **276**, 43949-43957 (2001).
53. Iraqui, I., Vissers, S., Cartiaux, M. & Urrestarazu, A. Characterisation of *Saccharomyces cerevisiae* ARO8 and ARO9 genes encoding aromatic aminotransferases I and II reveals a new aminotransferase subfamily. *Mol Gen Genet* **257**, 238-248 (1998).
54. Iraqui, I., Vissers, S., Andre, B. & Urrestarazu, A. Transcriptional induction by aromatic amino acids in *Saccharomyces cerevisiae*. *Mol Cell Biol* **19**, 3360-3371 (1999).
55. Vuralhan, Z. et al. Physiological characterization of the ARO10-dependent, broad-substrate-specificity 2-oxo acid decarboxylase activity of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **71**, 3276-3284 (2005).
56. Samanani, N., Liscombe, D.K. & Facchini, P.J. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloisoquinoline alkaloid biosynthesis. *Plant J* **40**, 302-313 (2004).
57. Minami, H., Dubouzet, E., Iwasa, K. & Sato, F. Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J Biol Chem* **282**, 6274-6282 (2007).
58. Bashor, C.J., Helman, N.C., Yan, S. & Lim, W.A. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* **319**, 1539-1543 (2008).
59. Russell, J.S.a.D.W. (ed.) *Molecular Cloning*, Edn. 3. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 2001).
60. Gillam, E.M., Guo, Z., Martin, M.V., Jenkins, C.M. & Guengerich, F.P. Expression of cytochrome P450 2D6 in *Escherichia coli*, purification, and spectral and catalytic characterization. *Arch Biochem Biophys* **319**, 540-550 (1995).
61. Mapoles, J., Berthou, F., Alexander, A., Simon, F. & Menez, J.F. Mammalian PC-12 cell genetically engineered for human cytochrome P450 2E1 expression. *Eur J Biochem* **214**, 735-745 (1993).
62. Gietz, R.D. & Woods, R.A. Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol Biol* **313**, 107-120 (2006).
63. Thomas, B.J. & Rothstein, R. Elevated recombination rates in transcriptionally active DNA. *Cell* **56**, 619-630 (1989).
64. Chen, J. et al. Analysis of major alkaloids in *Rhizoma coptidis* by capillary electrophoresis-electrospray-time of flight mass spectrometry with different background electrolytes. *Electrophoresis* **29**, 2135-2147 (2008).
65. 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).
66. Sato, F., Tsujita, T., Katagiri, Y., Yoshida, S. & Yamada, Y. Purification and characterization of S-adenosyl-L-methionine: norcoclaurine 6-O-methyltransferase from cultured *Coptis japonica* cells. *Eur J Biochem* **225**, 125-131 (1994).

67. Ounaroon, A., Decker, G., Schmidt, J., Lottspeich, F. & Kutchan, T.M. (R,S)-Reticuline 7-O-methyltransferase and (R,S)-norcoclaurine 6-O-methyltransferase of *Papaver somniferum* - cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy. *Plant J* **36**, 808-819 (2003).
68. Kutchan, T.M. & Dittrich, 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).
69. 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).
70. 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**, 38557-38565 (2003).
71. 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).
72. Hirata, K., Poeaknapo, C., Schmidt, J. & Zenk, M.H. 1,2-Dehydroreticuline synthase, the branch point enzyme opening the morphinan biosynthetic pathway. *Phytochemistry* **65**, 1039-1046 (2004).
73. De-Eknamkul W, Z.M. Purification and properties of 1,2-dehydroreticuline reductase from *Papaver somniferum* seedlings. *Phytochemistry* **31**, 813-821 (1992).
74. Liscombe, D.K. & Facchini, P.J. Evolutionary and cellular webs in benzyloquinoline alkaloid biosynthesis. *Curr Opin Biotechnol* **19**, 173-180 (2008).
75. Zhu, W., Cadet, P., Baggerman, G., Mantione, K.J. & Stefano, G.B. Human white blood cells synthesize morphine: CYP2D6 modulation. *J Immunol* **175**, 7357-7362 (2005).
76. Raith, K. et al. Electrospray tandem mass spectrometric investigations of morphinans. *J Am Soc Mass Spectrom* **14**, 1262-1269 (2003).
77. Winkel, B.S. Metabolic channeling in plants. *Annu Rev Plant Biol* **55**, 85-107 (2004).
78. Ro, D.K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943 (2006).
79. Urban, P., Mignotte, C., Kazmaier, M., Delorme, F. & Pompon, D. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J Biol Chem* **272**, 19176-19186 (1997).
80. Guldener, U., Heck, S., Fielder, T., Beinhauer, J. & Hegemann, J.H. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* **24**, 2519-2524 (1996).
81. 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**, e23 (2002).

82. Caponigro, G., Muhlrud, D. & Parker, R. A small segment of the MAT alpha 1 transcript promotes mRNA decay in *Saccharomyces cerevisiae*: a stimulatory role for rare codons. *Mol Cell Biol* **13**, 5141-5148 (1993).
83. Ng, R. & Abelson, J. Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **77**, 3912-3916 (1980).
84. Boonstra, B., Rathbone, D.A. & Bruce, N.C. Engineering novel biocatalytic routes for production of semisynthetic opiate drugs. *Biomol Eng* **18**, 41-47 (2001).
85. Rosco, A., Pauli, H.H., Priesner, W. & Kutchan, T.M. Cloning and heterologous expression of NADPH-cytochrome P450 reductases from the Papaveraceae. *Arch Biochem Biophys* **348**, 369-377 (1997).
86. Sugiura, M., Sakaki, T., Yabusaki, Y. & Ohkawa, H. Cloning and expression in *Escherichia coli* and *Saccharomyces cerevisiae* of a novel tobacco cytochrome P-450-like cDNA. *Biochim Biophys Acta* **1308**, 231-240 (1996).
87. Geissler, R., Brandt, W. & Ziegler, J. Molecular modeling and site-directed mutagenesis reveal the benzyloquinoline binding site of the short-chain dehydrogenase/reductase salutaridine reductase. *Plant Physiol* **143**, 1493-1503 (2007).
88. Moreira dos Santos, M., Raghevendran, V., Kotter, P., Olsson, L. & Nielsen, J. Manipulation of malic enzyme in *Saccharomyces cerevisiae* for increasing NADPH production capacity aerobically in different cellular compartments. *Metab Eng* **6**, 352-363 (2004).
89. Lenz, R. & Zenk, M.H. Acetyl coenzyme A:salutaridinol-7-O-acetyltransferase from papaver somniferum plant cell cultures. The enzyme catalyzing the formation of thebaine in morphine biosynthesis. *J Biol Chem* **270**, 31091-31096 (1995).
90. Grothe, T., Lenz, R. & Kutchan, T.M. Molecular characterization of the salutaridinol 7-O-acetyltransferase involved in morphine biosynthesis in opium poppy *Papaver somniferum*. *J Biol Chem* **276**, 30717-30723 (2001).
91. Kjeldsen, T. Yeast secretory expression of insulin precursors. *Appl Microbiol Biotechnol* **54**, 277-286 (2000).
92. Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.K. & Keasling, J.D. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab Eng* **9**, 160-168 (2007).
93. Allen, R.S. et al. Metabolic engineering of morphinan alkaloids by over-expression and RNAi suppression of salutaridinol 7-O-acetyltransferase in opium poppy. *Plant Biotechnol J* **6**, 22-30 (2008).
94. Allen, R.S. et al. RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat Biotechnol* **22**, 1559-1566 (2004).
95. Ikezawa, N., Iwasa, K. & Sato, F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in *Eschscholzia californica*. *Plant Cell Rep* (2008).