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 time[1]. 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 reations 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 2μm-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 promoter[3]. Of particular interest is the *GAL2Δ* 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 levels[4]. The results demonstrated that substitution of a weaker promoter would not compromise production in most cases[5]. 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 benzylisoquinoline alkaloid (BIA) pathway

The benzylisoquinoline 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 berberine[6]. 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.

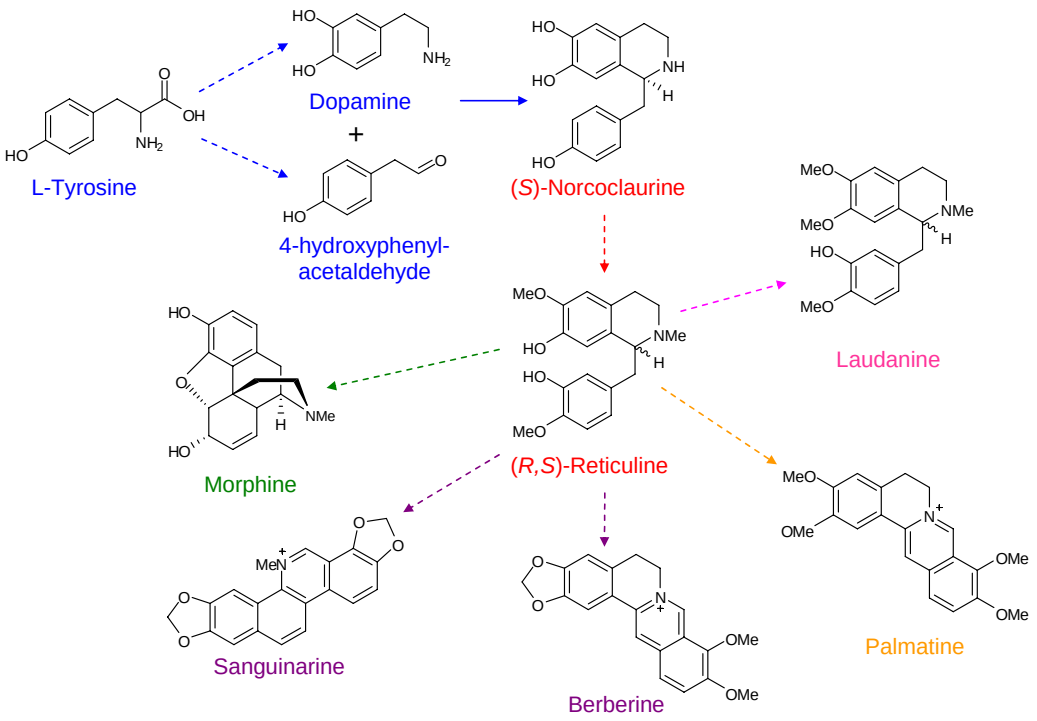

Fig. 1.1. General schematic of the BIA pathway. The upstream pathway (blue) shows the tyrosine derivatives dopamine and 4-HPA which make up the BIA backbone. The intermediate pathway (red) shows norcoclaurine, the first BIA structure, and reticuline, a major branch point intermediate. Additional branched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